Genetic-epigenetic interactions in medulloblastoma development

Dolores Mary Hamilton

Declaration

I certify that no part of the material documented in this thesis has previously been submitted for a degree or other qualification in this or any other university. I declare that this thesis represents my own unaided work, carried out by myself, except where it is acknowledged otherwise in the thesis text.

Dolores Hamilton

October 2013

Acknowledgements

Dedication

I would like to thank all my family and friends. I extend a special feeling of gratitude to my mum, Maeve, and to my sister, Roisin, whose belief in me has meant so much during the pursuit of my PhD. Their words of encouragement and push for tenacity have carried me through this highly challenging and rewarding experience.

This thesis is dedicated to the memory of my dad, Dermot Hamilton. His strength of character and unwavering hope in his fight against cancer were my inspiration to embark on my cancer research journey.

Abstract

Medulloblastoma is the most common malignant brain tumour of childhood. Transcriptomic profiling has revealed the existence of four core molecular subgroups (SHH, WNT, Group 3 and Group 4) with distinct clinical, pathologic and molecular characteristics. However, the specific molecular events associated with tumour development in these groups are poorly understood. DNA methylation plays a key role in epigenetic transcriptional regulation, and promoter hypermethylation leading to gene silencing is a common feature of medulloblastoma. DNA methylation profiling has identified distinct methylomic profiles associated with the four subgroups of medulloblastoma, and the wider role of DNA methylation in medulloblastoma now requires investigation. Using two high-throughput screening approaches, this project therefore undertook a comprehensive investigation into the potential role of specific DNA methylation events in the development of the distinct subgroups of medulloblastoma.

Using DNA methylation profiles, which were generated for 216 medulloblastomas using the GoldenGate methylation array, the first approach identified 73 CpG methylation markers (encompassing 63 genes) which significantly distinguished Group 3 and/or Group 4 medulloblastomas. Subgroup-specific differential gene expression analysis showed that, for the majority of the methylation markers identified, there was no clear inverse association between methylation and gene expression. One gene (*RHOH*) was identified which showed strong evidence of epigenetic dysregulation in medulloblastomas. *RHOH* methylation represented a potential epigenetic event in Group 4 tumours; 51% of Group 4 medulloblastomas showed aberrant hypomethylation of multiple *RHOH* promoter-associated CpG residues, which was associated with upregulated *RHOH* expression in Group 4 tumours. *RHOH* was re-expressed in 4 out of 6 methylated cell lines following treatment with the demethylating agent 5'-aza-2'-deoxycytidine (5-azaCdR). This study has thus identified a novel putative oncogenic role for *RHOH* in Group 4 medulloblastoma development.

In the second approach, a functional epigenomics screen identified 283 genes which were upregulated in 2 or more cell lines investigated (*n*=10) following 5-azaCdR treatment. Assessment of DNA methylation using the Illumina 450K methylation array identified 160 CpG residues (encompassing 21 of the 283 genes) whose methylation status was consistent with expression alterations observed after 5-azaCDR, and methylation-dependent gene regulation, in cell lines. 9/160 CpG residues (6%) showed evidence of subgroup-specific differential methylation which was concordant with differential gene expression and potential epigenetic gene regulation in medulloblastoma subgroups. These 9 sites represented 5 candidate genes (*ACTC1, ANXA2, FAM46A, PRPH* and *S100A4*). Aberrant hypermethylation of multiple gene body CpG residues was associated with *FAM46A* silencing in non-SHH tumours, while aberrant hypermethylation of multiple promoter-associated residues was associated with *ACTC1* silencing in Group 3 and Group 4 medulloblastomas. Single site

hypomethylation events associated with upregulated expression in WNT tumours were identified for *ANXA2*, *PRPH* and *S100A4*.

This study has identified six genes with putative oncogenic or tumour suppressor roles in the development of distinct subgroups of medulloblastoma through their epigenetic dysregulation. Further work is now required to validate these findings and to assess their functional significance in medulloblastoma subgroups, as well as their potential relevance in medulloblastoma sub-classification and prognostication.

List of abbreviations

3' UTR	Three prime untranslated region
5-azaCdR	5'-aza-2'-deoxycytidine
5' UTR	Five prime untranslated region
ASO	Allele-specific oligonucleotide
AT/RT	Atypical teratoid/rhabdoid tumour
ATP	Adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
C	Cytosine
cDNA	Complementary DNA
CGH	Comparative genomic hybridisation
CHARM	Comprehensive high-throughput array-based relative methylation
ChIP	Chromatin immunoprecipitation
CLL	Chronic lymphocytic leukaemia
CML	Chronic myeloid leukaemia
CNS	Central nervous system
COG	North American Children's Oncology Group
CpG	Cytosine-guanine dinucleotide
cRNA	Complementary RNA
CRUK	Cancer Research UK
CSF	Cerebrospinal fluid
CT	Computed tomography
DMARD	Denaturation of methylation differences
DMSO	Dimethylsulfoxide
DN	Desmoplastic/nodular medulloblastoma
DNA	Deoxyribose nucleic acid
DNMT	DNA methyltransferase
dsDNA	Double-stranded DNA
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
EFS	Event free survival
EGL	External granule layer
ESC	Embryonic stem cell
FCS	Fetal calf serum
FC	Fold change
FDA	Food and Drug Administration
FDR	False discovery rate
FFPE	Formalin-fixed paraffin embedded
FISH	Fluorescent in situ hybridisation
gDNA	Genomic DNA
GNP	Granule neuron precursor

GO	Gene ontology
GTP	Guanine triphosphate
Gy	Gray
HART	Hyperfractionated accelerated radiotherapy
HAT	Histone acetyl transferase
HDAC	Histone deacetylase
HFRT	Hyperfractionated radiotherapy
HMT	Histone methyltransferase
IGL	Internal granule layer
LCA	Large cell/anaplastic medulloblastoma
LFS	Li-Fraumeni syndrome
LINES	Long interspersed nuclear elements
LOH	Loss of heterozygosity
LOI	Loss of imprinting
LSO	Locus-specific oligonucleotide
MAD	Median absolute deviation
MAGIC	Medulloblastoma Advanced Genomics International Consortium
MBD	Methyl-CpG binding domain
MBEN	Medulloblastoma with extensive nodularity
MBP	Methyl binding protein
MDS	Myelodysplastic syndrome
MeCP2	Methyl-CpG binding protein 2
MeDIP	Methylated DNA immunoprecipitation
miRNA	MicroRNA
MRI	Magnetic resonance imaging
mRNA	Messenger RNA
MSI	Microsatellite instability
MSO	Methylation-specific oligonucleotides
MSP	Methylation-specific PCR
NCBI	National Center for Biotechnology Information
NDR	Nucleosome depleted region
NextGen	Next generation
NMF	Non-negative matrix factorisation
NOMe-Seq	Nucleosome occupancy methylome sequencing
NTP	Nucleoside triphosphate
OD	Optical density
OS	Overall survival
PBS	Phosphate buffered saline
PCA	Principle component analysis
PcG	Polycomb group proteins
PCR	Polymerase chain reaction

PNET	Primitive neuroectodermal tumour
PNET3	Primitive neuroectodermal tumour 3 trial
PRC2	Polycomb repressive complex 2
qPCR	Quantitative PCR
qRT-PCR	Real-time quantitative reverse transcriptase PCR
RIN	RNA Integrity number
RLGS	Restriction landmark genome scanning
RNA	Ribonucleic acid
RT-PCR	Reverse transcriptase PCR
SHH	Sonic hedgehog
SINES	Short interspersed nuclear elements
SIOP	International Society of Paediatric Oncology
siRNA	Small interfering RNA
SNP	Single nucleotide polymorphism
TE	Transposable element
TSS	Transcription start site
U	Uracil
UCSC	University of California, Santa Cruz
UKCCSG	United Kingdom Children's Cancer Study Group
UV	Ultraviolet
WGBS	Whole genome bisulfite sequencing
WGSBS	Whole genome shotgun bisulfite sequencing
WHO	World Health Organisation
WNT	Wingless

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Cancer

Cancer defines a heterogeneous group of diseases that are characterised by abnormal and uncontrolled cellular growth. Cancer, by definition, is a malignant neoplasm; cells have acquired the ability to penetrate and invade into surrounding tissue and to metastasise to more distant body sites. Disseminated tumour cells migrate from the primary site through the lymphatic system, blood vessels or across body cavities and successful metastasis depends on their ability to establish and develop into overt secondary tumours at the distant site (Hanahan and Weinberg, 2011). The presence and extent of metastases defines tumour stage and is an important prognostic indicator in most cancers (NCI, 2013).

Cancer is a genetic disease and all cancers are caused by somatic mutations, with in some cases also germline mutations that carry an inherited predisposition to cancer (Alexandrov *et al.*, 2013). While there are many widely recognised environmental causative factors in cancer, many of them, such as tobacco tar and ultra-violet (UV) radiation, exert their carcinogenic effect by targeting the genetic makeup of cells. In normal tissues the rate of cell growth, cell differentiation and cell death are tightly regulated to maintain tissue architecture and function. In cancer, a series of mutations and other genetic alterations result in the dysregulation of genes that control normal cellular growth patterns, leading to the development of large populations of cells that no longer obey the rules governing normal tissue growth (Strachan and Read, 2004b). Cancers are typically classified according to the tissue of origin; carcinomas are derived from epithelial cells, sarcomas from stromal cells and leukaemias and lymphomas from blood cell precursors. A histological examination of cells and tissue, as viewed under the microscope, is used to diagnose and also to classify tumours (Strachan and Read, 2004b).

More than 200 different cancer types have been described and cancer is a leading cause of disease and death worldwide (CRUK, 2013a). Improvements in the treatment of cancer have seen survival rates double in the UK in the last 40 years, and half of people with cancer now survive their disease for at least 5 years (CRUK, 2013a). Intensive research over the past two decades has increased understanding of the molecular pathology of cancer. This has led to the development of refined classification systems for some cancer types and to the development of novel molecularly targeted therapies,

which are starting to help deliver the promise of personalised cancer medicine with the primary aim of improving cancer survival and cure rates further. However, despite an increased understanding of the genetic and lifestyle causes of cancer, incidence rates continue to rise, and more than one in three people in the UK will be diagnosed with some form of cancer during their lifetime (CRUK, 2013a).

1.1.1 Cancer incidence and mortality

In 2010, 324,579 people were diagnosed with cancer in the UK (CRUK, 2013a). Cancer is primarily a disease of older people with ~65% of cancers diagnosed in people over the age of 65 (Figure 1.1). While incidence is significantly higher in males than in females (426 and 374 per 100,000, respectively), incidence rates have increased by 6% in females compared with just 2% in males over the last 10 years (CRUK, 2013a).

Some tissues are more prone to developing cancer than others and cancers of the breast, lung, bowel and prostate account for over half of all new cases diagnosed (Figure 1.2). Despite being rare in men, breast cancer is the most common cancer in the UK. Prostate cancer is the most common cancer in men and accounts for 25% of male cancers. Over the last 10 years, cancer types have shown varying trends and have seen large increases in the incidence of cancers such as kidney, liver, oral and malignant melanoma with the increase most likely attributable to lifestyle choices (CRUK, 2013a).

Cancer was the cause of 157,275 deaths in the UK in 2010 and accounted for 28% of all deaths (CRUK, 2013a). Deaths from cancers of the lung, bowel, breast and prostate are the most frequent and contribute 46% of all cancer deaths (Figure 1.3). Owing to an increased understanding of cancer causes and improvements in treatment, mortality has declined. The rate of decline, however, has slowed over the last 10 years (CRUK, 2013a). More insights need to be gained from the underlying molecular pathology of cancer and better targeted treatments need to be developed in order to improve the survival and cure of a disease that is increasing in incidence.
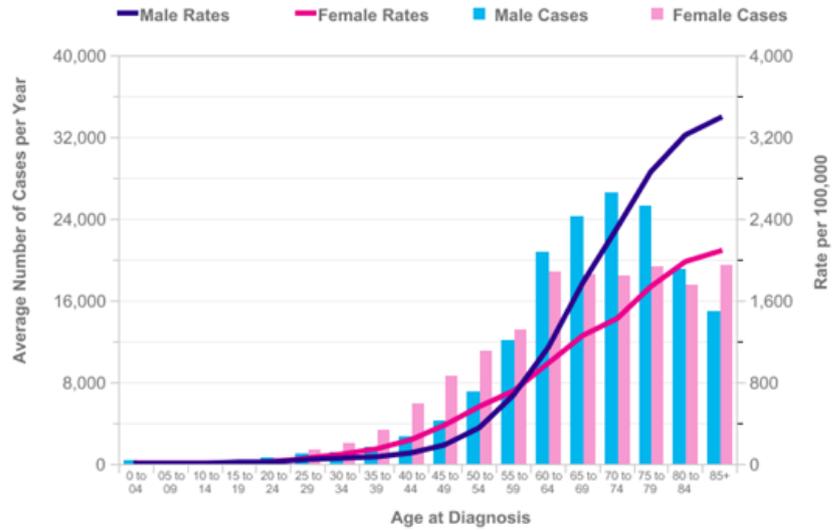
Figure 1.1. Average number of new cancers per year increases with age. Figure taken from Cancer Research UK cancer stats (CRUK, 2013a).
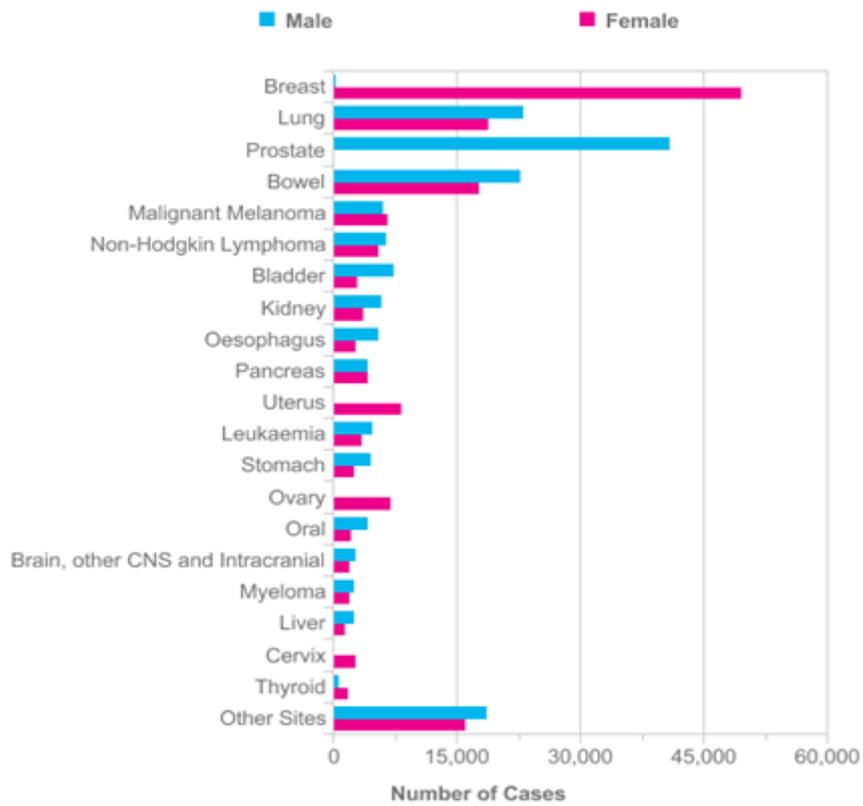
Figure 1.2. Incidence rates of common cancers. The 20 most commonly diagnosed cancers in the UK in 2010. Figure taken from Cancer Research UK cancer stats (CRUK, 2013a).
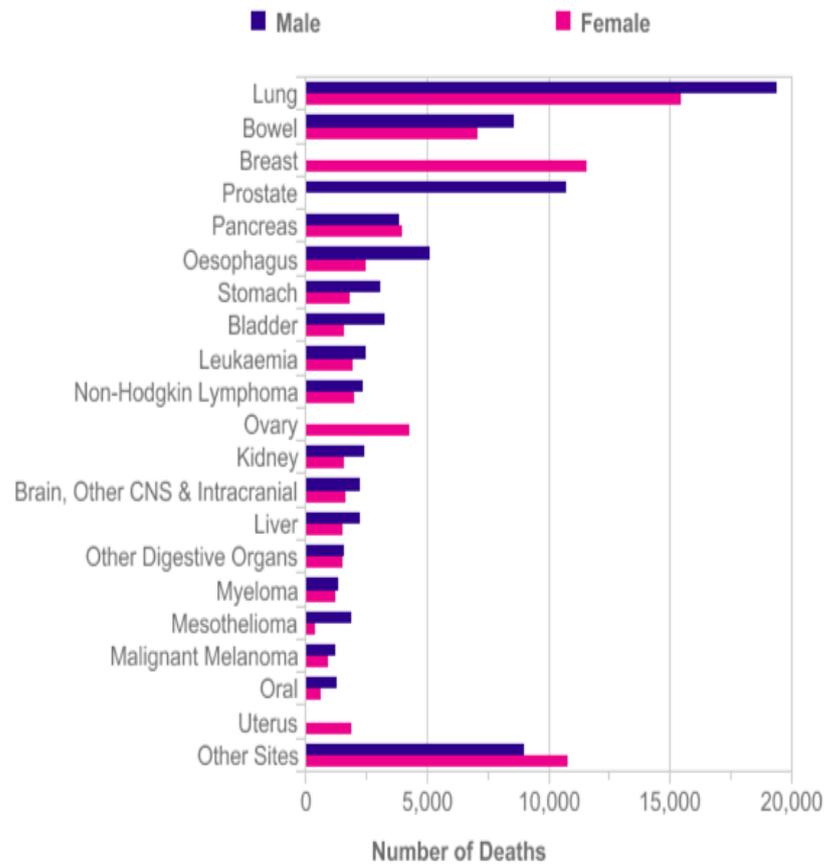
Figure 1.3. Mortality rates of common cancers in the UK. The 20 most common causes of cancer death in the UK in 2010. Figure taken from Cancer Research UK cancer stats (CRUK, 2013a).

1.2 Cancer biology

1.2.1 Cancer is a multi-step disease

Cancer is a highly complex and heterogeneous disease at the genetic level. There are many different genes that acquire mutations within and between the different tumour types, and it is the interaction between these mutational events that drives uncontrolled cellular proliferation and the development of individual cancers (Strachan and Read, 2004b).

Most cancers arise by expansion from a single cell (clonal evolution). There is no single mutation that can transform a normal cell into a malignant one and it has been estimated that any one cell would require on average 6-7 mutations to become malignant. Given that the typical mutational rate is 10^{-7} per gene per cell the chances of this occurring are very unlikely (Strachan and Read, 2004b). Cancer is the result of an accumulation of mutations, with each successive mutation providing the cell with a growth advantage and expanding the target population of cells for the next mutation (Figure 1.4). In turn, some mutations alter genomic stability and increase the overall mutational rate. As a result of this evolutionary process, cancer is a multi-step disease that develops over time, driving altered tissue growth through the progressive stages of hyperplasia (excessive cell proliferation) and dysplasia (excessive and abnormal growth of immature cells), leading ultimately to the development of malignant tumours (Strachan and Read, 2004b).

The differences in clinical behaviour and responses to treatment observed for tumours of the same type can, in part, be explained by the tumour heterogeneity that arises as a result of the clonal evolution of cancer (Croce, 2008). Furthermore, as well as the initial clone and sub-clones, tumours can also contain progenitor cancer cells that will differ in their genetic alterations and differentiation states. Tumour heterogeneity makes the identification of the initiating steps in the development of cancer critically important for the development of rational cancer therapies (Croce, 2008).
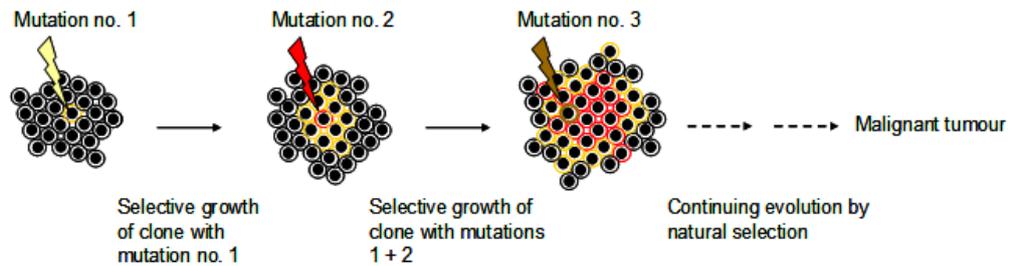
Figure 1.4. Clonal evolution in tumour cell populations. Successive mutations and clonal expansion result in malignant tumour development. Figure adapted from Strachan and Read, 2004.

Advances in whole-genome sequencing technologies have facilitated the identification of thousands of somatic mutations in a single cancer sample and have led to a recent major breakthrough in understanding the diversity and complexity of mutational processes in cancer (Alexandrov *et al.*, 2013). An analysis of over 4 million mutations from over 7,000 cancers has revealed the presence of distinct mutational signatures, some of which are shared by many cancer types while others describe a single cancer type (Alexandrov *et al.*, 2013). It is anticipated that as the number of cancer whole-genome sequencing studies increases, their findings will hold great promise to advance knowledge and understanding of cancer causes with potential important implications for treatment and prevention.

There are two broad categories of genes that are the targets of mutations in cancer: oncogenes, the normal activity of which promotes cell proliferation, growth and invasion, and tumour suppressor genes whose normal activity inhibits events that lead to cancer.

1.2.2 *Oncogenes*

Oncogenes were first identified in the 1960s when it was discovered that some animal cancers were caused by viruses, and that the transforming properties of retroviruses were entirely due to the possession of one extra gene which became known as the oncogene. Viral oncogenes are copies of normal cellular genes called 'proto-oncogenes', that have become activated following their incorporation into the viral RNA genome (Strachan and Read, 2004b). Many cellular oncogenes have subsequently been identified by transforming immortalised cells in culture with DNA extracted from tumour cells. Most human cancers are not caused by viruses and proto-oncogenes

become activated and gain oncogenic potential by genetic changes in their coding or regulatory sequences or by an increase in copy number, resulting in either qualitative or quantitative changes in gene expression. Activation of proto-oncogenes involves a dominant gain of function and their altered expression in cancer directly determines the development of the malignant phenotype (Strachan and Read, 2004b). Proto-oncogenes encode proteins that control cell proliferation and apoptosis and the functional role of oncogene products can be classified into six broad groups: transcription factors, growth factors, growth factor receptors, signal transducers, chromatin remodellers and apoptosis regulators (Croce, 2008). Examples of oncogenes, their function and the cancer types they are disrupted in are listed in Table 1.1.

Oncogene	Function	Cancer types
MYC	Transcription factor	Breast cancer/Burkitt's lymphoma
MYCN	Transcription factor	Neuroblastoma/lung cancer
ERBB2	Growth factor receptor	Breast cancer/ovarian cancer
RAS	Intracellular signalling	Colorectal cancer/lung cancer/thyroid cancer/melanoma and more
BRAF	Intracellular signalling	Melanoma/thyroid cancer/colorectal cancer
SRC	Tyrosine kinase	Colorectal carcinoma
EGFR	Growth factor receptor	Glioblastoma/non-small cell lung cancer

Table 1.1. Examples of oncogenes. Oncogenes are listed alongside their normal cellular function and examples of cancers in which they become dysregulated. Table adapted from Croce, 2008.

Oncogenes can be activated by one of three main mechanisms: amplification, mutation and chromosomal rearrangements/translocations. These changes are dominant and normally affect only one allele. Activation of proto-oncogenes by gene amplification results in cancer cells containing multiple copies of structurally normal oncogenes and the over-production of an unaltered gene product (Strachan and Read, 2004b). Amplification of the oncogene *MYCN* is a common feature in neuroblastoma (Schwab, 1990), while breast cancers often amplify *ERBB2* which encodes a cell surface growth factor receptor (Strachan and Read, 2004b).

Activation of proto-oncogenes by point mutation occurs as a result of the production of a modified gene product that confers a qualitative gain of function (Strachan and Read, 2004b). For example, specific activating point mutations in the *RAS* family genes, *HRAS*, *KRAS* and *NRAS*, are frequently found in a variety of tumours (Table 1.1). *RAS* family genes are important mediators of cell signalling and the mutant RAS protein has reduced GTPase activity, leading to a slower inactivation of GTP-RAS and an excessive cellular response to the signal from the G-protein coupled receptors (Khosravi-Far and Der, 1994).

Proto-oncogene activation can also occur as a result of chromosomal translocations that produce a novel chimeric gene product. The best known example of this is the Philadelphia (Ph) chromosome which is seen in 90% of patients with chronic myeloid leukaemia (CML). The Philadelphia chromosome results from a balanced reciprocal 9;22 translocation at well-defined breakpoints resulting in the formation of a novel fusion gene, *BCR-ABL*. This chimeric gene produces an abnormal ABL tyrosine kinase product with transforming properties (Strachan and Read, 2004b). Alternatively, translocation into a transcriptionally active region of chromatin can cause activation of proto-oncogenes. Burkitt's lymphoma is a childhood tumour that is characterised by translocations that place the *MYC* oncogene close to an immunoglobulin locus that is actively transcribed in antibody-producing-B-cells, leading to upregulated *MYC* expression (Wiseman, 2006).

1.2.2.1 Targeting oncogenic proteins

Owing to their significant transforming potential, oncogenic proteins represent highly attractive therapeutic targets. Considerable progress has been made in the development of therapies that target oncogenic tyrosine kinases which are crucial mediators in multiple signalling pathways including proliferation and migration (Baselga, 2006). Several small molecule inhibitors of tyrosine kinases, most of which act as ATP mimetics, have been developed and approved for use in multiple cancers in which activation of the targeted oncogene is a significant feature (Imai and Takaoka, 2006) (Table 1.2). Accumulating clinical trial results suggest that monotherapy with these targeted therapies may be limited, owing to the multiple pathways that can be disrupted in cancer cells, and multi-targeting therapy with these agents may be more rational (Imai and Takaoka, 2006). At present, many of these targeted therapies are licensed for use in combination regimens with conventional chemotherapies, facilitating a potential reduction in chemotherapeutic dose and reducing dose-related toxicities. Imatinib (Glivec) was one of the first successful small molecule inhibitors developed; it inactivates the kinase activity of the *BCR-ABL* fusion gene in CML and has shown remarkable efficacy inducing complete remission in most patients with Philadelphia chromosome-positive CML (Druker *et al.*, 2001; Ottmann *et al.*, 2002).

The development of monoclonal antibodies that target cell surface oncogenic proteins by eliciting an immune response has been a significant development in the treatment of cancers such as breast cancer and colorectal cancer (Table 1.2). Trastuzumab (Herceptin) is the first monoclonal antibody that targets the oncogene *ERBB2* (Carter *et al.*, 1992). It is approved for the treatment of patients with metastatic breast cancer who carry an increased *ERBB2* copy number (Krejsa *et al.*, 2006). Trastuzumab shows excellent anti-tumour activity, particularly when combined with the cytotoxic agents doxorubicin and paclitaxel (Baselga *et al.*, 1998).

The field of targeted cancer therapy is expanding rapidly and as an increasing number of molecular targets are identified, the number of novel therapies entering clinical trials is likely to increase. The potential for rational combinations of targeted therapies must also be explored with the primary aim of improving survival and limiting the emergence of drug resistance which can be seen with single agent use (Imai and Takaoka, 2006).

Targeted therapy	Molecular target	Cancer types
Small molecule inhibitors		
Imatinib (Glivec, Novartis)	ABL, PDGFR, KIT	CML/gastrointestinal stromal tumours
Gefitinib (Iressa, AstraZeneca)	EGFR	Non-small cell lung cancer
Erlotinib (Tarceva, Genentech)	EGFR	Non-small cell lung cancer
Sorafenib (Nexavar, Bayer/Onyx)	VEGFR, PDGFR, FLT3	Renal cell carcinoma
Sunitinib (Sutent, Pfizer)	VEGFR, PDGFR, FLT3	Renal cell carcinoma/gastrointestinal stromal tumours
Monoclonal antibodies		
Trastuzumab (Herceptin, Genentech)	ERBB2	Breast cancer
Cetuximab (Erbitux, ImClone)	EGFR	Colorectal cancer
Bevacizumab (Avastin, Genentech)	VEGF	Colorectal cancer/non-small cell lung carcinoma

Table 1.2. Examples of cancer therapies that target oncogenic proteins. Targeted therapies are listed by their generic drug name alongside their commercial name and the pharmaceutical company that holds the patent. Their molecular targets and the cancers they treat are listed. EGFR denotes epidermal growth factor receptor; FLT3 FMS-like tyrosine kinase 3; PDGFR platelet-derived growth factor receptor; VEGF vascular endothelial growth factor. Table adapted from Croce, 2008.

1.2.3 Tumour suppressor genes

Tumour suppressors are inactivated forms of cellular genes. They encode proteins that are negative regulators of normal growth and differentiation and have been identified by their loss of function in tumours and in studies with transgenic null mice (Strachan and Read, 2004b).

Initial insights into tumour suppressor gene inactivation came from early work carried out by Knudson in 1971 on retinoblastoma (Knudson, 1971). Retinoblastoma is an aggressive childhood cancer and ~40% of cases represent familial disease which is often bilateral compared with sporadic forms which are always unilateral. Knudson noted that the age-distribution of bilateral disease was consistent with a single mutation, while sporadic cases followed a two-hit mechanism. He reasoned that all retinoblastomas involved two hits with one being inherited in the familial cases. Knudson's hypothesis was subsequently proved when it was shown that sporadic and hereditary forms of retinoblastoma are characterised by loss of both alleles of the retinoblastoma gene (*RB1*) (Cavenee *et al.*, 1983).

The two-hit hypothesis proposed by Knudson was established as the paradigm for tumour suppressor gene inactivation; tumour suppressor gene loss of function is recessive and therefore both copies of the gene must be lost from the cell for the tumour suppressor phenotype to be evident (Strachan and Read, 2004b). In their study of retinoblastoma, Cavenee et al. observed that somatic genetic changes caused loss of heterozygosity (LOH) at markers close to the *RB1* locus (Cavenee *et al.*, 1983), and thus LOH analysis has been a major approach employed to discover the locations of tumour suppressor genes. Since the discovery of the *RB1* tumour suppressor gene, a large number of tumour suppressor genes have been identified through their association with familial and sporadic cancers. Commonly inactivated tumour suppressor genes are listed in Table 1.3.

There are many recognised mechanisms for the inactivation of tumour suppressor genes. Tumour suppressor genes may be silenced by whole chromosomal loss, chromosomal loss and duplication, gene conversion, gene deletion, mitotic recombination, point mutation or by epigenetic inactivation through promoter hypermethylation (Strachan and Read, 2004b). Methylation-dependent silencing of tumour suppressor genes is discussed in detail in section 1.6. The *BRCA1* gene which is inactivated in

approximately 10-15% of sporadic breast cancers has been shown to be silenced by DNA methylation in some of these tumours (Birgisdottir *et al.*, 2006). Genes that are inactivated by epigenetic mechanisms are attractive therapeutic targets owing to the reversible nature of epigenetic modifications (see section 1.6).

TP53 is the gene most frequently altered in human cancers. *TP53* encodes the p53 tumour suppressor protein which protects DNA from damage by blocking proliferation, stimulating DNA repair and promoting apoptosis. It is regarded as the 'guardian of the genome'. *TP53* inactivation can occur by point mutation and many thousands of inactivating mutations have been identified in human cancers (Lim *et al.*, 2007; Petitjean *et al.*, 2007).

Bi-allelic loss of tumour suppressor genes can occur via the same mechanism or by any combination of two different mechanisms. Following the identification of genes which frequently show loss of function of only one allele which is sufficient to cause an abnormal phenotype, it has been postulated that haploinsufficiency may be sufficient to cause gene inactivation and provide a growth advantage (Strachan and Read, 2004b). This phenomenon is observed in medulloblastoma with loss of one copy of the *PTCH1* gene sufficient to initiate tumour formation (Goodrich *et al.*, 1997), while loss of a single allele of the *PTEN* tumour suppressor gene in prostate cancer has been shown to promote tumour progression (Kwabi-Addo *et al.*, 2001).

Gene	Function	Cancer types
APC	Cell-cell recognition	Colorectal cancer
BRCA1	Transcriptional regulation and cell cycle control	Breast cancer/ovarian cancer
TP53	Regulation of DNA repair and apoptosis	Breast cancer/colorectal cancer/lung cancer and many more
RB1	Cell cycle control	Retinoblastoma
CDKN2A	Cell cycle control	Melanoma
VHL	Cell division and angiogenesis	Renal cell carcinoma

Table 1.3. Examples of tumour suppressor genes. Genes are listed alongside their normal cellular function and examples of cancers in which they become dysregulated. Table adapted from Strachan and Read, 2004.

1.2.4   Hallmark capabilities of cancer cells

The capabilities of a normal cell to undergo malignant transformation depend in large part on the accumulation of genetic and also epigenetic alterations that confer a selective advantage on subclones of cells, enabling their growth, invasion and eventual metastasis (Hanahan and Weinberg, 2011). In 2000, Hanahan and Weinberg described six hallmarks of cancer that were acquired by cells to enable their malignant transformation (Hanahan and Weinberg, 2000). In a recent update to their initial publication, the authors report that the six hallmarks continue to provide a solid foundation for understanding cancer biology. The hallmark capabilities are integral components of most cancer forms and they provide a framework for understanding the significant heterogeneity observed in cancers; the mechanisms by which they are acquired are heterogenous both within and across different cancer types (Hanahan and Weinberg, 2011). The six hallmarks of cancer are illustrated in Figure 1.5 and described below.
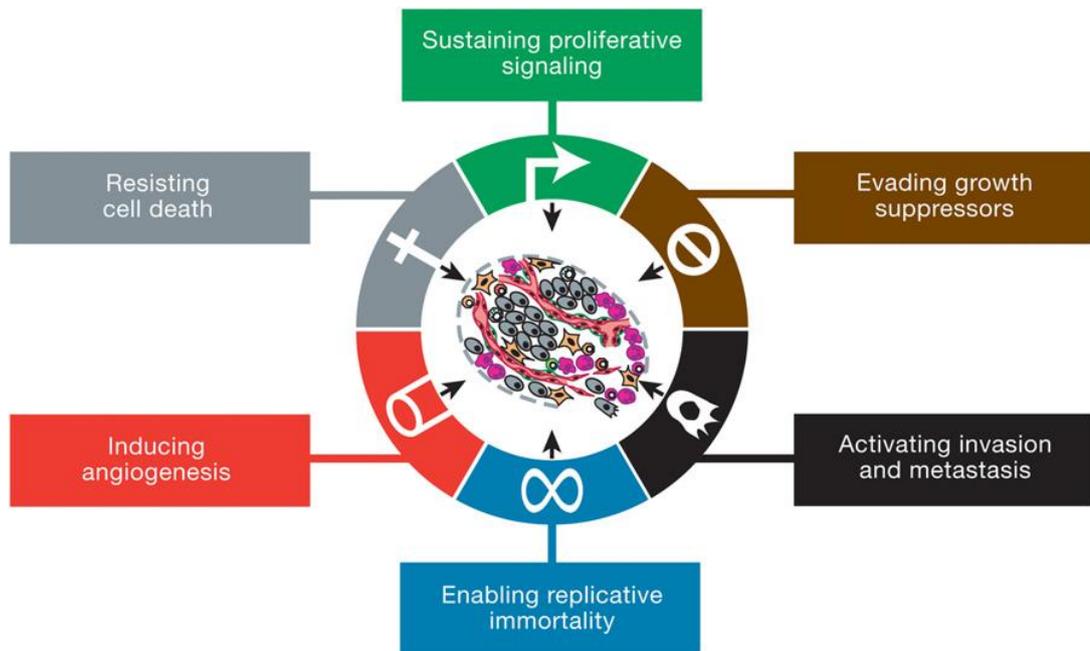
Figure 1.5. The six hallmark capabilities of cancer necessary for tumour growth. Figure taken from Hanahan and Weinberg, 2011.

1.2.4.1 Sustained proliferative signalling

Normal cells tightly regulate cell cycle processes and maintain a homeostatic state between cell growth and cell death. The most fundamental capability of cancer cells is their ability to sustain proliferation and this is mediated largely by oncogenic growth factors that bind cell surface receptors, typically the tyrosine kinases. There are a number of ways by which cancer cells can sustain proliferative signalling; they may produce growth factor ligands themselves or they may stimulate normal cells within the supporting stroma to produce growth factors. Cancer cells can also become hyper-responsive to growth factor ligand by increasing the number of receptor proteins on their surface. Aberrant and sustained signalling may also take place independent of growth factors, through the constitutive activation of signalling pathway components that are downstream of the receptors (Hanahan and Weinberg, 2011).

1.2.4.2   Evading growth suppressors

As well as inducing and sustaining growth stimulatory signals, cancer cells must also evade programmes that negatively regulate cell proliferation. This hallmark relies mainly on the actions of tumour suppressor genes. Examples include the *RB1* and *TP53* genes which, as previously described, are prototypical tumour suppressor genes that act within key cellular regulatory pathways to control cell proliferation or apoptosis, respectively (Hanahan and Weinberg, 2011).

1.2.4.3   Activating invasion and metastasis

The capability of cancer cells to invade and metastasise is not yet fully understood. The invasion-metastasis cascade depicts successive cellular changes, beginning with local invasion, dissemination of cancer cells into neighbouring lymphatic and blood vessels followed by their transportation to distant sites. Initially micrometastases form, which then colonise and form macroscropic tumours at sites distant from the primary site. Cancer cells that have acquired the ability to invade and metastasise have typically developed alterations in their shape as well as in their attachment to other cells and to the extracellular matrix (ECM). For example, E-cadherin plays a critical role in the co-ordinated assembly of epithelial cell sheets, and its reduced expression is known to promote invasion and metastasis in some human carcinomas, a feature that strongly supports its role as a key suppressor of this hallmark capability. Genes encoding other cell-cell and cell-ECM adhesion factors are altered in some highly aggressive carcinomas, while those favouring cytostasis are downregulated (Hanahan and Weinberg, 2011).

1.2.4.4   Enabling replicative immortality

Cancer cells require unlimited replicative potential in order to generate large tumour masses. In normal lineages cells can undergo only a limited number of successive cell growth and division cycles owing to the barriers of senescence and cell death. In cancer cells it is thought that telomeres protecting the ends of chromosomes play a key role in their ability to proliferate in an unlimited manner. In non-cancerous cells, telomeres shorten progressively and eventually lose their ability to protect the ends of chromosomal DNA and cells become senescent. Telomerase is the DNA polymerase

enzyme that adds telomere repeats to the ends of telomeric DNA and while it is mainly absent in normal cells it is activated in ~90% of immortalised cells, including cancer cells. The presence of telomerase activity is associated with a resistance to induction of both senescence and apoptosis (Hanahan and Weinberg, 2011).

1.2.4.5 Inducing angiogenesis

Tumour cells, like normal cells, require nutrients and oxygen and the ability to eliminate waste products. In tumours this is achieved via the newly formed tumour-associated blood vessels (angiogenesis). In normal adult cells, angiogenesis is transiently active during processes such as wound healing and in the female reproductive cycle. In contrast, in tumour cells angiogenesis is almost always active supporting the expansion of neoplastic growth. For example, the signalling proteins VEGF-A (vascular endothelial growth factor-a) and TSP1 (thrombospondin-1) induce and inhibit, respectively, the angiogenic process in some cancers by binding to cell surface receptors on vascular endothelial cells (Hanahan and Weinberg, 2011).

1.2.4.6 Resisting cell death

Programmed cell death by apoptosis is a natural barrier to the development of cancer. Apoptotic processes can be triggered in cancer cells by stresses resulting from oncogene signalling and from DNA damage. Tumour cells can evade apoptosis through a variety of mechanisms. The most common mechanism is through the loss of the *TP53* tumour suppressor gene which eliminates the critical DNA damage response from apoptotic signalling processes. Tumours may also downregulate pro-apoptotic factors such as Bax and Bim, leading to increased expression of anti-apoptotic mediators or of survival signals (Hanahan and Weinberg, 2011).

1.3 Childhood cancer

In the UK, 1,603 new cases of childhood cancer were diagnosed in 2010 (CRUK, 2013b). The three most common types of cancer diagnosed in children are leukaemia, brain and CNS tumours and lymphomas and they account for approximately two-thirds of all childhood cancers (Figure 1.6). The prospects for children with cancer have improved dramatically and almost three-quarters of children are now cured of their disease compared with a quarter in the late 1960s. Despite this improved survival, cancer is the most common cause of disease-related death in children aged 1-14 years, and 260 children under the age of 15 die from their disease each year in the UK (CRUK, 2013b). Survival rates vary according to cancer type and although tumours of the brain and CNS rank second in incidence, they are the most common cause of deaths from cancer in children (CRUK, 2013b). The majority of childhood cancers are invasive and they present significant treatment challenges, including the long-term adverse effects associated with current treatment modalities. For example, radiotherapy treatment of brain tumours in children is associated with significant long-term neuro-cognitive and neuro-endocrine defects and for these reasons is often delayed in very young children (Pizer and Clifford, 2009). There is an urgent need to further characterise the aetiology and pathogenesis of childhood cancers and to develop better treatments that will improve survival further and reduce treatment-related sequelae for long-term survivors.
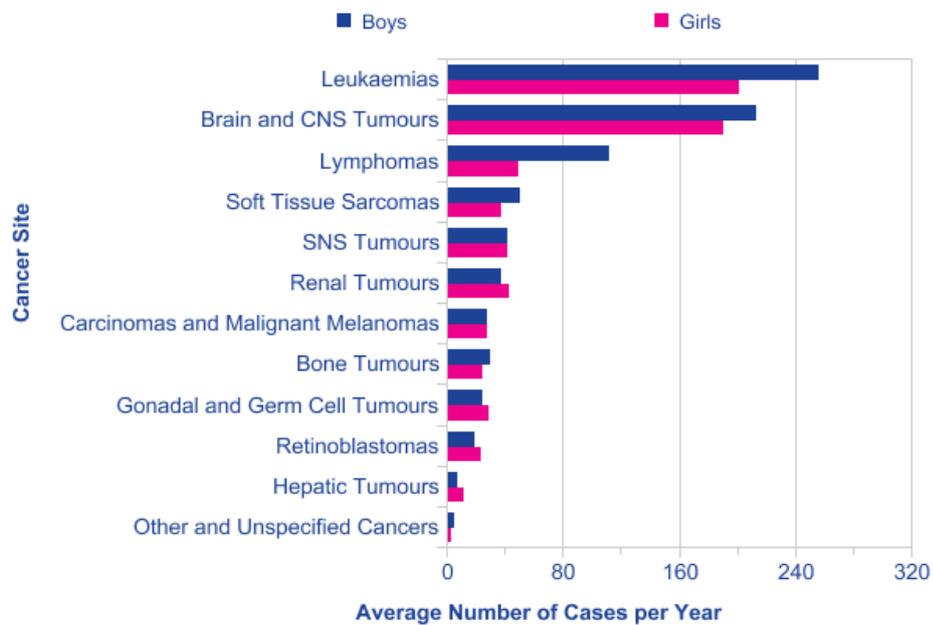
Figure 1.6. Incidence rates of the most common childhood cancers in the UK. Figure taken from Cancer Research UK cancer stats (CRUK, 2013b).

1.4 Paediatric tumours of the central nervous system

Cancers of the central nervous system (CNS) are the second most common cancer type in children after leukaemia, and are the leading cause of cancer-related mortality in childhood (Pui *et al.*, 2011; Siegel *et al.*, 2013). The latest statistics from Cancer Research UK report that astrocytomas, the most common type of glioma, are most prevalent and constitute 43% of all brain and CNS tumours in children. Approximately 76% of astrocytomas diagnosed in children are low grade and 15% are high grade. Embryonal tumours of the CNS (see section 1.4.1) are the second most frequent subgroup and medulloblastomas make up almost three-quarters of this group (CRUK, 2013b). Medulloblastoma is the most common malignant brain tumour in childhood and accounts for 15-20% of CNS tumours (Pizer and Clifford, 2009). Ependymoma and choroid plexus tumours constitute a further 10% of childhood brain and CNS tumours (CRUK, 2013b).

Paediatric tumours of the brain and CNS can present in a myriad of ways and this can lead to a delayed diagnosis. Presenting signs and symptoms will depend on the growth rate of the tumour, its location in the CNS and the child's age (Packer *et al.*, 2010). The

majority of childhood CNS tumours arise in the brain and for most their exact cause is largely unknown. Specific inherited syndromes can predispose children to these tumours in a small proportion of cases and this is discussed further in the context of medulloblastoma in section 1.5.6.1. The only environmental factor that has consistently been associated with the development of brain tumours in children has been exposure to radiation therapy (Ron *et al.*, 1988). Compared to brain tumours that arise in adults, childhood brain tumours demonstrate greater histological variation, they are more frequently metastatic at diagnosis and they are more frequently embryonal in their classification (Dolecek *et al.*, 2012).

The most recent World Health Organisation (WHO) classification of CNS tumours is based on a panel of expert neuro-pathologists and geneticists (Louis *et al.*, 2007). It is presented as the worldwide standard for the definition of brain tumours and provides clearly defined histopathological and clinical diagnostic criteria. The standardised classification and grading allows epidemiological studies and clinical trials to be conducted on an international scale. Following advances that have been made in diagnostic pathology and in tumour genetics, the diagnosis and classification of brain tumours now depends on immunohistochemical markers and genetic profiles alongside traditional histopathological features (Louis *et al.*, 2007).

1.4.1 Embryonal tumours of the CNS

Embryonal brain tumours are characterised by the proliferation of tissue that is normally seen only in the developing embryo and may remain in the brain after birth. Originally embryonal CNS tumours were grouped under a broad category of primitive neuroectodermal tumours (PNETs) based on their common undifferentiated round cell morphology (Ellison, 2002). Following progress that was made in understanding their distinctive clinical, pathological, molecular and behavioural characteristics, the 2000 WHO classification separated them into 5 distinct types; medulloblastoma, supratentorial PNET, atypical teratoid/rhabdoid tumour (AT/RT), ependymoblastoma and medulloepithelioma (Kleihues *et al.*, 2002). Since 2000 the classification of embryonal tumours has been further refined with the recognition of 5 histological variants of medulloblastoma that are associated with outcome and hence have significant clinical utility (Louis *et al.*, 2007; Gilbertson and Ellison, 2008; Gulino *et al.*, 2008).

Embryonal tumours are classified within the larger group of tumours of neuroepithelial tissue (Louis *et al.*, 2007). The current WHO classification (2007) defines three distinct types of embryonal tumour; medulloblastoma, CNS primitive neuroectodermal tumour (CNS-PNETs) and atypical teratoid/rhabdoid tumours (AT/RTs). Medulloblastoma is further subclassified into the 5 histological variants (see section 1.5.1), while CNS-PNETs encompass CNS neuroblastoma, CNS ganglioneuroblastoma, medulloepithelioma and ependymoblastoma. All embryonal CNS tumours are malignant neoplasms of WHO grade IV (Louis *et al.*, 2007) and they are associated with a comparably high mortality and with significant long-term morbidity for survivors (Pfister *et al.*, 2010). They comprise the largest group of malignant brain tumours in childhood and medulloblastomas are the most common type (Sarkar *et al.*, 2005).

1.5 Medulloblastoma

Medulloblastoma, by definition, arises in the posterior fossa which is a region of the hindbrain that contains the brainstem and the cerebellum (Figure 1.6). It was first described in 1925 by Bailey and Cushing as a tumour of primitive origin arising in the posterior fossa of young children that was distinct from gliomas (Bailey and Cushing, 1925). Since its initial description significant progress has been made in understanding the biology of medulloblastoma and this has been paralleled by refinements in its pathological classification and the recognition of medulloblastoma as a distinct disease entity (Louis *et al.*, 2007).

Medulloblastoma arises in the cerebellum, most often in the roof of the 4th ventricle, invading the cerebellar vermis and often invading through the ependyma in the floor of the ventricle to enter the brainstem (Raffel, 2004) (Figure 1.6). Less commonly the tumour arises in the cerebellar hemisphere, particularly in older children (Pizer and Clifford, 2009). Medulloblastoma is now established as a highly heterogeneous disease, both at the histopathological and the molecular level (Pizer and Clifford, 2009). The current consensus is that medulloblastoma comprises four distinct molecular subgroups that most likely arise from different cells of origin (Taylor *et al.*, 2012). Cellular origins of medulloblastoma are discussed further in section 1.5.9.

Medulloblastoma is the most common malignant brain tumour of childhood and it accounts for around 10% of all paediatric cancer deaths (Pizer and Clifford, 2009). There are approximately 1 in 200,000 children diagnosed with medulloblastoma each year (Parsons *et al.*, 2011), with approximately 90 patients diagnosed in the UK each year (Pizer and Clifford, 2009). Medulloblastomas are principally diagnosed in children less than 15 years of age, and they have a bimodal age distribution, peaking at 3 to 4 years of age and then again between 8 and 9 years of age. Around 10-15% of cases diagnosed are in children under 3 years of age (Packer *et al.*, 1999; Crawford *et al.*, 2007). For unknown reasons the incidence of medulloblastoma is 1.5-2 times higher in males than in females (Crawford *et al.*, 2007). Medulloblastoma is relatively uncommon in adults, representing 0.4-1% of adult CNS tumours (Kieran *et al.*, 2010).
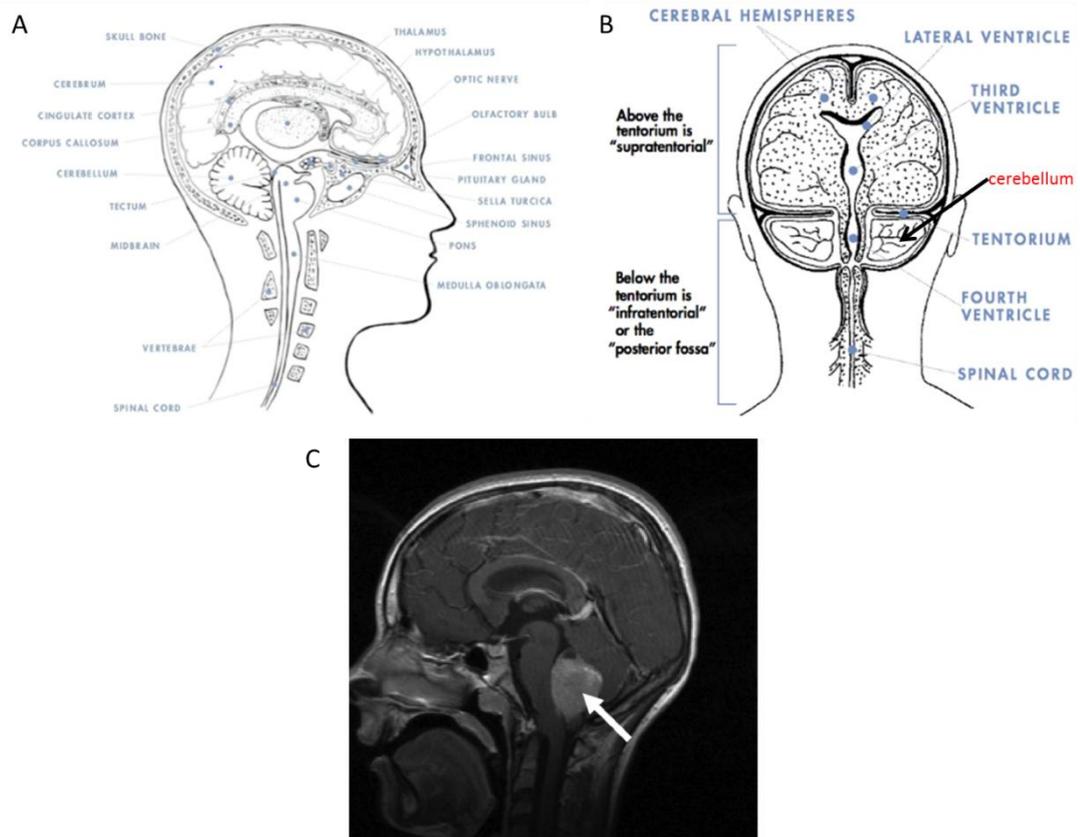
Figure 1.7. Anatomy of the brain and location of medulloblastoma. A. Sagittal section showing regions of the brain. The cerebellum is located in the hindbrain at the top of the brainstem. **B.** Schematic illustrating the location of the tentorium, fourth ventricle and the cerebellum. The tentorium separates the cerebral hemispheres from the structures of the posterior fossa. The fourth ventricle which contains cerebrospinal fluid is an expansion of the central canal of the medulla oblongata. The cerebellum is highlighted in red and is located in the posterior fossa below the tentorium. The roof of the fourth ventricle is formed by the cerebellum Schematic brain sections taken from the American Brain Tumour Association website (http://www.abta.org/understanding-brain-tumors/anatomy/physical-structure.html). **C**. T1-weighted sagittal magnetic resonance imaging (MRI) scan following gadolinium of child with medulloblastoma. Medulloblastoma is visible as grey staining mass indicated by white arrow. Image kindly provided by Professor Simon Bailey (Paediatric Brain Tumour Research Group, Newcastle University, UK).

1.5.1 Histopathological variants of medulloblastoma

The 2007 WHO classification of CNS tumours defines medulloblastoma as a grade IV malignant neoplasm and recognises the tumour heterogeneity that is reflected in its diverse histopathology (Louis *et al.*, 2007). Histopathological variants of medulloblastoma have been recognised that are associated with distinct molecular features and with distinct patterns of clinical behaviour, and are therefore of clinical utility in disease classification (Ellison *et al.*, 2003; Ellison, 2010). Based on the diverse histological phenotypes and differences in their clinical behaviour, the current WHO classification defines 5 distinct histological disease variants: classic medulloblastoma,

desmoplastic/nodular (DN) medulloblastoma, medulloblastoma with extensive nodularity (MBEN), large cell medulloblastoma and anaplastic medulloblastoma.

1.5.1.1 Classic medulloblastoma

Classic medulloblastomas are most numerous and constitute 70-80% of all medulloblastomas (McManamy *et al.*, 2007; Gilbertson and Ellison, 2008; Pizer and Clifford, 2009). The classic variant comprises sheets of small cells with a high nuclear:cytoplasmic ratio and round hyperchromatic nuclei (Ellison, 2002) (Figure 1.7A). Foci of necrosis and angiogenesis variably interrupt these sheets of undifferentiated cells and neuroblastic rosettes which consist of tumour cell nuclei arranged in a circular fashion around tangled cytoplasmic processes are observed in less than 40% of classic tumours (Ellison, 2002).

1.5.1.2 Desmoplastic/nodular medulloblastoma

The desmoplastic/nodular (DN) variant is observed in 10-15% of medulloblastomas (Ellison, 2002; McManamy *et al.*, 2007). The DN medulloblastoma is characterised by nodules of differentiated neurocytic cells and internodular desmoplasia; uniform neurocytic cells are usually round and scattered across reticulin-rich desmoplastic regions where tumour cells show more nuclear pleomorphism than intranodular cells (McManamy *et al.*, 2007) (Figure 1.7B). The extent of nodularity can be variable and an increased nodular density has been associated with an improved prognosis (McManamy *et al.*, 2007).

Desmoplasia is most often observed in very young children, representing up to 50% of tumours in infants (< 3 years at diagnosis) and only 5% of tumours in patients over 3 years of age at diagnosis (McManamy *et al.*, 2007). Several studies have shown that desmoplastic medulloblastomas presenting in infancy have a significantly better outcome than other histological phenotypes in this age group (Rutkowski *et al.*, 2005; McManamy *et al.*, 2007; Rutkowski *et al.*, 2009; Grundy *et al.*, 2010), while the significance of desmoplasia in non-infant cases appears less important (McManamy *et al.*, 2007).

1.5.1.3 Medulloblastoma with extensive nodularity

Medulloblastoma with extensive nodularity (MBEN) shares fundamental histopathological features with the D/N disease variant and it is recognised as a desmoplastic medulloblastoma (McManamy *et al.*, 2007). Compared to D/N tumours, MBENs comprise large and irregularly shaped nodules that dominate the histology with a reduced internodular desmoplastic component (Ellison, 2010) (Figure 1.7C). MBENs comprise less than 1% of medulloblastomas and most often present in infants (< 3 years of age at diagnosis) (McManamy *et al.*, 2007). They are associated with a favourable prognosis (Giangaspero *et al.*, 1999; Suresh *et al.*, 2004).

1.5.1.4 Large cell medulloblastoma

The large cell variant of medulloblastoma makes up between 2-4% of medulloblastomas (Ellison, 2002; McManamy *et al.*, 2007). It often presents with metastatic disease (Gajjar *et al.*, 2004) and is associated with a poor prognosis (Brown *et al.*, 2000; Eberhart *et al.*, 2002a; Eberhart and Burger, 2003; McManamy *et al.*, 2003). This variant comprises large round cells with a prominent single nucleolus and pleomorphic nuclei (Giangaspero *et al.*, 1992) and have a higher mitotic and apoptotic rate than other variants (McManamy *et al.*, 2003) (Figure 1.7D).

1.5.1.5 Anaplastic medulloblastoma

Anaplastic tumours comprise between 10-20% of medulloblastomas (Gilbertson and Ellison, 2008). This phenotype is characterised by marked nuclear pleomorphism accompanied by cell moulding, producing a paving-like pattern of nuclei, and cell wrapping (Brown *et al.*, 2000; Eberhart *et al.*, 2002a) (Figure 1.7E). These tumours also show increased mitotic activity and apoptosis (McManamy *et al.*, 2003). Similar to the large cell variant, anaplastic medulloblastomas are associated with a poor prognosis (Eberhart *et al.*, 2002a; McManamy *et al.*, 2003). Because the large cell and anaplastic medulloblastomas share similar histopathological features and an aggressive biological behaviour, they are frequently combined together into a single large cell/anaplastic (LCA) type in studies of medulloblastoma (Gilbertson and Ellison, 2008).
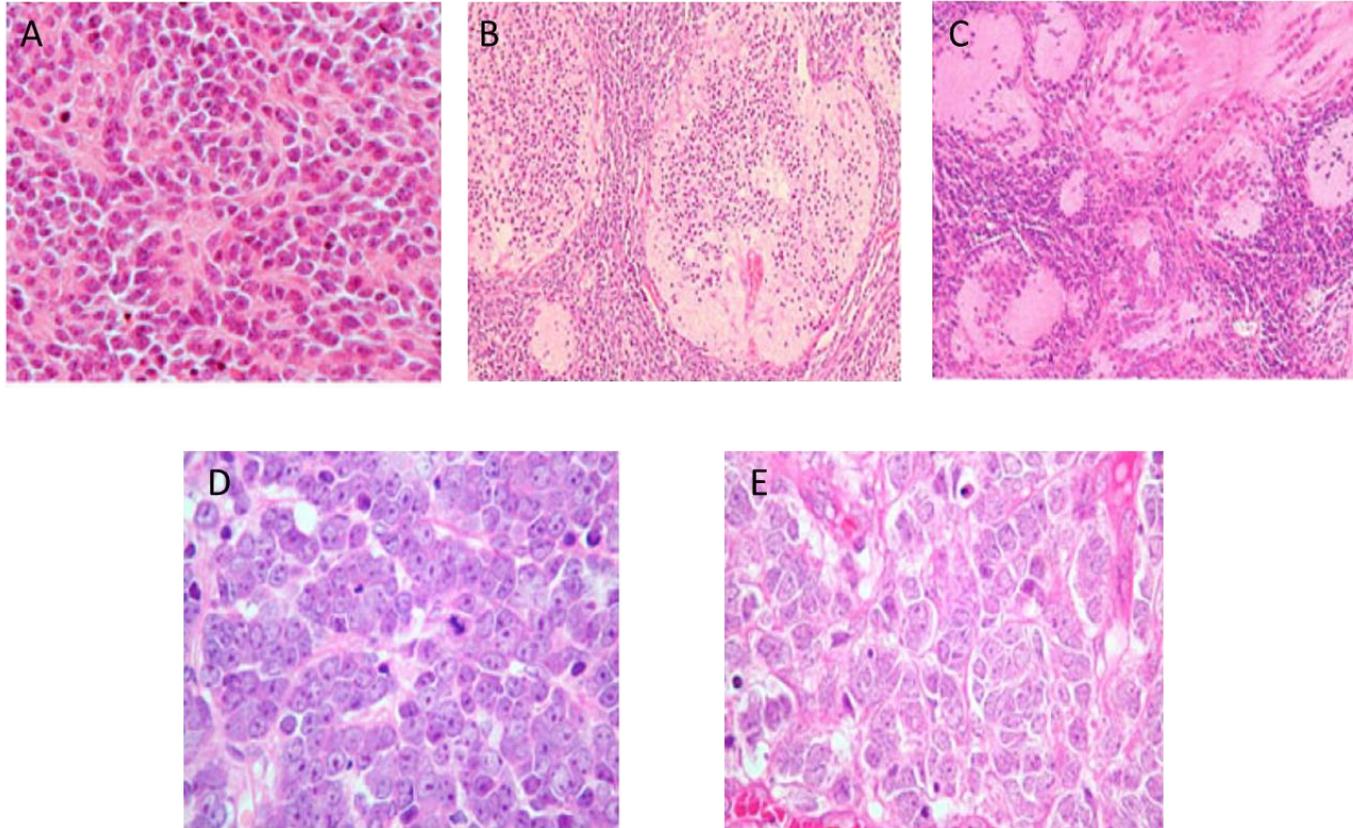
Figure 1.8. Histopathological subtypes of medulloblastoma. **A**. Classic medullloblastoma. **B**. Desmoplastic/nodular (D/N) medulloblastoma. **C**. Medulloblastoma with extensive nodularity (MBEN). **D**. Large cell medulloblastoma. **E**. Anaplastic medulloblastoma. Sections were stained with haematoxylin and eosin. Figure taken from Ellison, 2010.

1.5.2   Clinical presentation and diagnosis of medulloblastoma

Children with medulloblastoma typically present with signs and symptoms of obstruction of CSF flow and cerebellar dysfunction. Presenting symptoms include wide-based stance, ataxia and nystagmus. Behavioural changes accompanied by vomiting, headache and lethargy are due mainly to hydrocephalus and subsequent raised intracranial pressure caused by the tumour mass (Zakhary *et al.*, 2001). Infants with medulloblastoma present less characteristically with intermittent vomiting and macrocephaly (Packer *et al.*, 1999). In infants, cerebrospinal fluid obstruction with ventricular dilatation may result in downward deviation of the eyes, the so-called "sun-setting sign" (Packer *et al.*, 1999). Although a significant number of children will show evidence of metastatic spread at diagnosis, many will be asymptomatic at the time of initial presentation (Packer *et al.*, 1999).

Diagnosis of medulloblastoma is made by either computed tomography (CT) or magnetic resonance imaging (MRI). Due to the higher resolution and increased sensitivity of MRI it has become the neuroimaging technique of choice (Packer *et al.*, 1999). On MRI the tumour usually presents as a midline lesion filling the fourth ventricle and sometimes extending into the ventricular foramina (Packer *et al.*, 1999) (see Figure 1.6). The presence of metastasis at diagnosis confers a poor prognosis in medulloblastoma (Bailey *et al.*, 1995; Zeltzer *et al.*, 1999) and it is an established prognostic variable that influences patient stratification and treatment (Pizer and Clifford, 2009) (see sections 1.5.4 and 1.5.5). For these reasons MRI imaging of the entire neuraxis (brain and spinal cord) is required to determine the extent of disease dissemination before initiating treatment (Packer *et al.*, 1999).

1.5.3   Tumour staging

Approximately 30% of children show evidence of metastatic spread of medulloblastoma at diagnosis (Bailey *et al.*, 1995). Metastatic spread occurs mainly via the cerebrospinal fluid, leading to a high incidence of spinal and of diffuse leptomeningeal seeding. Solid metastases to the supratentorial region also occur. Extraneuraxial metastases are relatively uncommon and appear to occur largely as a result of surgical intervention, with preferential spread to the bone apparent (Sure *et al.*, 1995). Following MRI and possible cytological examination of the CSF, the Chang staging system (Chang *et al.*, 1969) is used to define the presence or absence of metastases. Depending on the

presence and extent of disease dissemination, medulloblastomas are staged M0-M4 as described in Table 1.4.

M stage	
M0	No gross subarachnoid or haematogenous metastasis
M1	Microscopic tumour cells found in CSF
M2	Gross nodular seeding in cerebellum, cerebral subarachnoid space, or in third or fourth ventricles
M3	Gross nodular seeding in spinal subarachnoid space
M4	Extraneuraxial metastasis

Table 1.4. Chang classification of metastasis stage in medulloblastoma. Based on MRI imaging and cytological examination of the CSF, medulloblastomas are staged M0-M4 depending on the presence and extent of metastases as detailed. Table adapted from Zeltzer et al., 1999.

1.5.4 *Current disease-risk stratification in medulloblastoma*

Since the mid-1990s, the risk stratification of medulloblastoma has been based solely on clinical disease features and stratifies patients into "standard risk" / "average risk" and "high risk", depending on their age at diagnosis, extent of surgical resection and metastatic status (Pizer and Clifford, 2009).

The presence of metastases at diagnosis, as defined by the Chang classification system (see section 1.5.3), and including spread to other areas of the brain (Chang stage M2) or to the spine (Chang stage M3) or as detected in the CSF (Chang stage M1), has consistently been found to be the strongest clinical predictor of poor patient outcome (Bailey *et al.*, 1995; Zeltzer *et al.*, 1999). The extent of surgical resection has been widely demonstrated to be an important prognostic indicator in medulloblastoma, with a gross total or near gross total surgical excision conferring a better prognosis (Albright *et al.*, 1996; Zeltzer *et al.*, 1999).

Very young children (< 3 years at diagnosis) have a much poorer prognosis than older children and are considered within their own exclusive risk category with different management strategies (Pizer and Clifford, 2009) (see section 1.5.5.1.3). The avoidance

of craniospinal radiotherapy in this age group, due to the severe neurological adverse effects of radiotherapy on the developing brain (see section 1.5.5.2), is believed to contribute to their poorer survival outcome (Pizer and Clifford, 2009).

The current risk stratification as recognised by the North American Children's Oncology Group (COG) and the International Society of Paediatric Oncology (SIOP) Europe Brain Tumour Committee recognises the following 3 risk categories (Pizer and Clifford, 2009):

(1) Standard risk / Average risk: patients $\geq$ 3 years of age without evidence of metastatic spread (M0) and having $\leq 1.5 cm^2$ of residual tumour mass after surgery.

(2) High risk: patients $\geq$ 3 years of age with evidence of disseminated disease (M1-M3) and/or a less complete surgical resection.

(3) Patients < 3 years at diagnosis.

While this clinical staging system has influenced treatment protocols and contributed to improved survival rates in medulloblastoma (see section 1.5.5), it is well-observed that patients stratified within the same group and thus receiving identical therapies can have widely disparate clinical outcomes, owing to biological differences within their tumours (Eberhart and Burger, 2003), and the aggressive therapies employed are associated with significant long-term adverse effects (see section 1.5.5.2). As a result, much of the research over the last 10 years has focused on understanding the biology of medulloblastoma and on identifying molecular biomarkers that will facilitate the development of a more accurate and refined stratification that will differentiate high risk patients from low risk patient and will inform treatment protocols accordingly.

Significant advances have been made in understanding the molecular pathology of medulloblastoma. Molecular disease biomarkers with prognostic significance have been identified and validated in clinical trial cohorts (Pizer and Clifford, 2009; Ellison *et al.*, 2011b) and have since been incorporated into a refined risk stratification model based on clinical, pathological and molecular variables that recognises low risk, standard risk and high risk disease (Ellison *et al.*, 2011b) (see section 1.5.11). This new revised stratification will form the basis of the forthcoming pan-European SIOP clinical trial (PNET5) as discussed in section 1.5.11.

1.5.5   Treatment of medulloblastoma

Based on the clinical staging system currently used (see section 1.5.4), a multimodality approach consisting of surgery, radiotherapy and chemotherapy has been designed for medulloblastomas. Advances in treatment have seen overall survival rates reach 70-80% (Northcott *et al.*, 2012a).

Surgery aims to resect as much of the primary tumour mass as is possible and is guided by the anatomy of the lesion. Most medulloblastomas are amenable to total resection if they do not extensively infiltrate the cerebellum and brainstem (Packer *et al.*, 1999). Radiotherapy remains the single most effective means of postoperative treatment for children greater than 3 years of age (Packer *et al.*, 1999) and owing to the tendency of medulloblastoma to spread via the CSF it is standard practice to irradiate the entire craniospinal axis, as this has been associated with the most significant improvement in survival (Landberg *et al.*, 1980; Bouffet *et al.*, 1992; Packer *et al.*, 2012). In children less than 3 years of age, radiotherapy is often not given initially due to its well documented adverse effects on the developing nervous system (Fouladi *et al.*, 2005; Lafay-Cousin *et al.*, 2009) (see section 1.5.5.2). The benefits of chemotherapy in the treatment of standard risk disease have been widely demonstrated. The European SIOP PNET3 study showed a significant improvement in 3-year event free survival (79% versus 64%) for patients with standard risk disease who were treated with pre-irradiation chemotherapy compared to those treated with radiotherapy alone (Taylor *et al.*, 2003a). Meanwhile in another Phase III study, chemotherapy given during and after radiotherapy for patients with standard risk medulloblastoma resulted in 5-year event free survival (EFS) and overall survival (OS) of 81% and 86% respectively (Packer *et al.*, 2006). The benefit of adjuvant chemotherapy in metastatic medulloblastoma has not been definitely proven in Phase III studies, although it is given as standard care in high-risk disease (Pizer and Clifford, 2009).

The use of molecular targeted therapies in medulloblastoma is still in its infancy and they are not currently included as standard treatments in adjuvant therapies (see section 1.5.12).

1.5.5.1 Treatment protocols

1.5.5.1.1 Current treatments for standard risk patients

In standard risk patients, post-surgical radiotherapy comprises a 'reduced dose' craniospinal dose of 23.4 Gy in addition to a local boost of 54-55.8 Gy to the posterior fossa (Pizer and Clifford, 2009). Adverse effects on neurocognitive and endocrine functions caused by craniospinal radiation (see section 1.5.5.2) have led to this lowering of the radiation dose in standard risk patients (Packer *et al.*, 2006). A current Children's Oncology Group (COG) trial is investigating the effects of lowering the radiation dose further to 18 Gy and decreasing the posterior fossa volume boosted in standard risk patients (De Braganca and Packer, 2013). Meanwhile, the European SIOP group are investigating the use of hyperfractionated radiotherapy (HFRT) using twice daily fractions that can theoretically increase the dose to the tumour without an increase in adverse effects to the normal CNS (Pizer and Clifford, 2009). The most widely used maintenance chemotherapy regimen in standard risk patients is the 'Packer' regimen which consists of 8 courses of vincristine, lomustine and cisplatin (Pizer and Clifford, 2009). Cyclophosphamide may also be used with one study reporting no significant difference in survival when lomustine was replaced with cyclophosphamide in combination with vincristine and cisplatin (Packer *et al.*, 2006).

1.5.5.1.2 Current treatments for high risk patients

Patients with high risk medulloblastoma are treated post-surgically with craniospinal radiation doses of 35-36 Gy together with a local boost of 18-20 Gy to the posterior fossa (Pizer and Clifford, 2009). The optimal chemotherapy regimen for high risk cases has yet to be defined (Pizer and Clifford, 2009). Improved survival for patients with metastatic disease has been reported with a craniospinal radiation dose of 40 Gy together with conventional chemotherapy (Pizer and Clifford, 2009), while in another trial improved survival was observed when patients received 39 Gy of craniospinal radiation followed by 4 cycles of cyclophosphamide-based, dose-intensive chemotherapy (Gajjar *et al.*, 2006). A current COG trial for patients with high risk disease is investigating the use of carboplatin as a radiation sensitiser during radiotherapy followed by chemotherapy (De Braganca and Packer, 2013). Hyperfractionated accelerated radiotherapy (HART) is employed in some protocols for

the management of high risk disease with a view to improving tumour control without increasing toxicity and has shown promising results with 5-year EFS of 70% reported (Gandola *et al.*, 2009).

1.5.5.1.3 Current treatments for patients under 3 years of age

Chemotherapy-based strategies are employed to delay or avoid the use of craniospinal radiation in very young children (Pizer and Clifford, 2009). Infants and children under the age of 3 years of age are treated with high dose chemotherapy using agents including cisplatin, cyclophosphamide and vincristine, often supported by stem cell rescue (De Braganca and Packer, 2013). Studies have also been conducted that have employed alternative treatments using intrathecal and high dose intravenous methotrexate and report that it may be equally or more effective in the treatment of infants with medulloblastoma (Chi *et al.*, 2004; Rutkowski *et al.*, 2005). The introduction of a tightly controlled conformal approach to the delivery of radiation to the posterior fossa has led to a renewed interest in using local radiotherapy in the management of very young children with medulloblastoma. In a UK study that used this approach in combination with moderately intensive chemotherapy, survival of around 50-60% was reported for non-metastatic patients (Pizer and Clifford, 2009). For children under the age of 3 who are treated with surgery and chemotherapy alone, overall survival of 60-73% is observed, while for infants with the favourable prognosis desmoplastic histology (see section 1.5.1.2) survival rates of up to 90% are seen (Rutkowski *et al.*, 2010).

1.5.5.1.4 Current treatments for relapsing patients

Despite treatment advances, approximately 40% of children diagnosed with and treated for medulloblastoma will experience tumour recurrence (Jones *et al.*, 2012). Patients who have not previously received radiotherapy and who relapse can be successfully treated with radiotherapy (Ridola *et al.*, 2007). However, for patients whose disease recurs after combined radiotherapy and chemotherapy, they have a very poor prognosis at relapse and only 2% will survive (Pizer and Clifford, 2009). High dose chemotherapy has been shown to cure a subset of patients who have an isolated local relapse and maximal resection prior to receiving the high dose chemotherapy (Pizer and Clifford, 2009). For the majority of patients, treatment will be palliative in the relapse setting.

1.5.5.2 *Treatment-associated adverse effects*

The long-term neurocognitive and endocrinologic sequelae of treatment of children with medulloblastoma have been well documented, and craniospinal irradiation has been demonstrated as a major cause of adverse effects (Packer *et al.*, 1989). Neurocognitive sequelae are seen in patients of all age groups due to whole brain radiotherapy and possibly local boost radiotherapy, with children under 8 years of age appearing more vulnerable (Packer and Vezina, 2008). Children between the ages of 3 and 7 years who are treated with 36 Gy of radiation have a 20-30 point decrease in overall intelligence (Packer and Vezina, 2008), while the reduced dose used in standard risk patients (see section 1.5.5.1.1) still results in a significant decline of 10-15 IQ points (Mulhern *et al.*, 2005). Many long-term survivors are at high risk of psychosocial difficulties including forming stable relationships and obtaining employment (Packer *et al.*, 1999). Growth hormone deficiency is the most common endocrinologic sequela of radiotherapy and hypothyroidism also occurs relatively commonly and may be caused by either hypothalamic radiation or scattered irradiation received by the thyroid gland (Packer *et al.*, 1987). Sexual hormone dysfunction may also occur, leading to delayed development (Packer *et al.*, 1999). Some long-term survivors require hearing aids due to hearing loss caused by radiotherapy and cisplatin chemotherapy (De Braganca and Packer, 2013). Secondary tumours such as gliomas and meningiomas are rare complications of therapy but can lead to significant morbidity and mortality in survivors (Packer *et al.*, 2013). Owing to these significant treatment-associated adverse effects, the follow-up care of survivors is complex and patients will attend long-term survival clinics for monitoring and support.

1.5.6 Genetic basis of medulloblastoma

Significant progress has been made in the last 10 years in understanding the genetic basis of medulloblastoma. Studies of familial cancer syndromes provided the first insights into the molecular processes that underlie the disease (see section 1.5.6.1), while in recent years, disease profiling using multiple genome-wide approaches has revealed the presence of molecular subgroups of medulloblastoma that are associated with distinct genetic abnormalities and with distinct clinicopathological features and clinical outcomes (Thompson *et al.*, 2006; Kool *et al.*, 2008; Cho *et al.*, 2011; Northcott *et al.*, 2011b; Kool *et al.*, 2012) (see section 1.5.8). Such progress has paved the way for patient stratification and treatment based on tumour biology as well as clinical and histopathological features (see section 1.5.12). It is envisaged that molecular stratification will facilitate the delivery of personalised therapies tailored to a more accurate disease risk prediction, as well as the development and delivery of novel molecularly targeted therapies, with the primary aim of improving patient outcomes and reducing the long-term sequelae associated with current treatments (Pizer and Clifford, 2009; Ellison, 2010; Ellison *et al.*, 2011b).

1.5.6.1 Familial syndromes predisposing to medulloblastoma

While less than 5% of medulloblastomas are associated with an inherited cancer syndrome, studies of the genetic defects that cause these syndromes have revealed their extensive involvement in the development of sporadic medulloblastomas. These studies have provided significant insights into the role of two key developmental signalling pathways, the Sonic hedgehog (SHH) pathway and the WNT pathway, that are now known to define distinct molecular subgroups of sporadic medulloblastoma (see section 1.5.8), as well as a role for *TP53* in medulloblastoma development.

Gorlin syndrome (Nevoid basal cell carcinoma syndrome) is an autosomal dominant disorder that predisposes to both cancer and developmental defects and is caused by mutations in the patched homologue 1 gene (*PTCH1*) (Hahn *et al.*, 1996). *PTCH1* is a key component of the SHH signalling pathway (see section 1.5.7.1) and patients with Gorlin syndrome have a 3-5% lifetime risk of developing medulloblastoma (Ramaswamy *et al.*, 2011). Recurrent somatic mutations in *PTCH1* have been identified in approximately 10% of sporadic medulloblastomas (Pietsch *et al.*, 1997; Raffel *et al.*, 1997), and overall aberrant activation of the SHH signalling pathway is observed in up

to one third of all medulloblastomas owing to recurrent mutations or copy number aberrations that target multiple pathway components (Northcott *et al.*, 2012a) (see section 1.5.7.1).

Turcot syndrome is characterised by the concurrence of multiple colorectal adenomas and a primary tumour of the CNS (Hamilton *et al.*, 1995). In a subtype of Turcot syndrome, type II, it has been shown that an increased risk of colorectal carcinoma and of medulloblastoma development is attributable to inactivating germline mutations in the adenomatous polyposis coli (*APC*) gene (Hamilton *et al.*, 1995). The *APC* gene is a tumour suppressor gene that negatively regulates β-catenin, a key transcriptional activator and effector of the WNT signalling pathway (see section 1.5.7.2). Aberrant activation of the WNT signalling pathway caused by somatic mutations of the β-catenin gene (*CTNNB1*) is recognised to occur in approximately 10% of sporadic medulloblastomas, with rarer mutations in other WNT pathway components also observed (Pizer and Clifford, 2009) (see section 1.5.7.2).

Li-Fraumeni syndrome (LFS) is an autosomal dominant hereditary disorder that is associated with a dramatically increased incidence and early onset of a number of cancers, including medulloblastoma (Li and Fraumeni, 1969). Most patients with LFS harbour germline *TP53* mutations (Malkin *et al.*, 1990). The role of the tumour suppressor gene *TP53* in driving cancers has been well documented and somatic *TP53* mutations are frequently observed in most human cancers (Lim *et al.*, 2007; Petitjean *et al.*, 2007) (see section 1.2.3). While early reports have found *TP53* mutations to be rare in medulloblastoma, recent studies have identified up to 10-15% of medulloblastomas with *TP53* mutations (Tabori *et al.*, 2010; Lindsey *et al.*, 2011; Zhukova *et al.*, 2013). Data concerning the prognostic role of *TP53* mutation in medulloblastomas has been contradictory. One study reported a significantly adverse prognosis for *TP53*-mutated medulloblastomas independent of clinical presentation (Tabori *et al.*, 2010). However, more recent studies suggest that the role of *TP53* mutation in medulloblastoma is pleiotropic and dependent on the molecular subgroup of tumours in which it occurs (Pfaff *et al.*, 2010; Lindsey *et al.*, 2011; Rausch *et al.*, 2012; Zhukova *et al.*, 2013). Enrichment of *TP53* mutations has been reported in the distinct subgroup of medulloblastomas that exhibit SHH pathway activation (Zhukova *et al.*, 2013) and in the distinct subgroup of tumours that exhibit WNT pathway activation (Pfaff *et al.*,

2010; Lindsey *et al.*, 2011; Zhukova *et al.*, 2013). It has been well established that the subgroup of medulloblastomas with WNT pathway activation are associated with a good prognosis (Ellison *et al.*, 2005; Fattet *et al.*, 2009; Ellison *et al.*, 2011b), and where somatic *TP53* mutations were observed in WNT tumours the favourable prognosis associated with this subgroup remained (Pfaff *et al.*, 2010; Lindsey *et al.*, 2011; Zhukova *et al.*, 2013). In contrast, *TP53* mutations in SHH medulloblastomas were associated with a poor outcome (Zhukova *et al.*, 2013). Furthermore, several studies have reported that patients with a *TP53* mutation associated with the SHH subtype and a poor prognosis, harboured germline *TP53* mutations and were LFS patients, indicating that the impact of *TP53* mutations is dependent on the clinical (familial *versus* sporadic disease) and the molecular contexts in which they arise (Lindsey *et al.*, 2011; Rausch *et al.*, 2012; Zhukova *et al.*, 2013).

1.5.6.2 *Recurrent chromosomal abnormalities and gene amplifications*

Using fluorescent in situ hybridisation (FISH), array CGH and more recently genome-wide SNP arrays and loss of heterozygosity (LOH) analyses, recurrent and non-random chromosomal losses and gains have been identified in medulloblastoma. Studies carried out in recent years that have recognised the distinct molecular subgroups of medulloblastoma (see section 1.5.8), have identified many of these characteristic abnormalities to be enriched in specific disease subgroup(s) (Kool *et al.*, 2012).

Isochromosome 17q (i17q) is the most frequent chromosomal abnormality in medulloblastoma, occurring in 30-40% of cases (Bigner *et al.*, 1988; Lamont *et al.*, 2004; Pfister *et al.*, 2009). An isochromosome 17 results in the loss of the short arm (17p) and duplication of the long arm (17q), leading to a single copy of 17p and three copies of 17q. The role of i17q in medulloblastoma has yet to be realised. Recent next generation sequencing studies have identified only one gene, *CTDNEP1*, located on chromosome 17p13.1 to be a recurrent target of homozygous mutation in tumours with 17p deletion, suggesting a potential tumour suppressor role for this gene on 17p (Parsons *et al.*, 2011; Jones *et al.*, 2012; Pugh *et al.*, 2012; Robinson *et al.*, 2012). The *TP53* gene, which is located at 17p13.1, has not been found to be mutated more commonly in tumours with i17q compared to those without (Tabori *et al.*, 2010). In addition, regions of high level amplification on 17q have not been identified (Pfister *et al.*, 2009). It is possible that epigenetic events such as DNA methylation (see section

1.6) may target tumour suppressor genes on 17p; epigenetic inactivation of the *HIC1* gene which is located at 17p13.3 has been associated with 17p loss in medulloblastoma (Rood *et al.*, 2002). It has also been postulated that i17q may drive clonal selection of tumours through haploinsufficiency for genes on 17p and a modest increase of expression of genes on chromosome 17q (Cunningham *et al.*, 2006). i17q is now recognised to occur most frequently in non-SHH/non-WNT medulloblastomas (Kool *et al.*, 2012) (see section 1.5.8.2). A poor prognosis associated with chromosome 17 aberrations has previously been demonstrated in several studies investigating medulloblastoma as a single disease (Lamont *et al.*, 2004; Pan *et al.*, 2005; Pfister *et al.*, 2009), and more recently it has been shown to remain an independent prognostic factor within distinct molecular subgroups of the disease (Kool *et al.*, 2012) (see section 1.5.8.2). An important role for i17q has been suggested following the observation that recurrent medulloblastomas show increased levels of 17q gain compared with the original diagnostic tumour (Korshunov *et al.*, 2008).

Other prevalent cytogenetic abnormalities that have been identified in many studies in medulloblastoma include monosomy 6, deletion of chromosomes 9q, 10q and 16q, and gain of chromosomes 1q, 7 and 18 (Northcott *et al.*, 2012a). Loss of chromosome 6 has been shown to be significantly associated with the subgroup of 'favourable outcome' WNT-activated medulloblastomas (Clifford *et al.*, 2006; Fattet *et al.*, 2009) (see section 1.5.8). Loss of chromosome 9q is most frequently observed in SHH-activated medulloblastomas (Kool *et al.*, 2012; Northcott *et al.*, 2012c) (see section 1.5.8). The *PTCH1* gene, a negative regulator of SHH signalling, is located on chromosome 9q22.32 and has been shown to be the most frequently mutated gene and the most frequent target of focal deletion in SHH medulloblastomas (Northcott *et al.*, 2012c).

Alongside gross chromosomal abnormalities, recurrent focal amplification of specific oncogenes has been identified in medulloblastomas. The oncogenes *MYC* (at 8q24) and *MYCN* (at 2p24) are the most commonly amplified loci in medulloblastoma, with each reported in approximately 4-10% of cases (Pfister *et al.*, 2009). Multiple studies have confirmed that *MYC* amplification is significantly associated with a poor prognosis and with the large cell and anaplastic histology (Brown *et al.*, 2000; Aldosari *et al.*, 2002a; Eberhart *et al.*, 2002b; Lamont *et al.*, 2004; Pfister *et al.*, 2009), and amplification of *MYC* is now an established molecular marker of poor clinical outcome in

medulloblastoma (see section 1.5.10) *MYC* amplifications are enriched in a distinct subgroup of non-SHH/non-WNT medulloblastomas that are associated with the worst clinical outcome (Cho *et al.*, 2011; Kool *et al.*, 2012) (see section 1.5.8). *MYCN* amplified medulloblastomas are clinically more heterogeneous than *MYC* amplified cases (Pfister *et al.*, 2009), although they are often found in large cell and anaplastic tumours (Aldosari *et al.*, 2002a; Eberhart *et al.*, 2002b). *MYCN* amplification, like *MYC*, is associated with a worse prognosis (Pfister *et al.*, 2009; Kool *et al.*, 2012; Ryan *et al.*, 2012), however, its significance as a prognostic marker has not yet been proven in large trial cohorts (Rutkowski *et al.*, 2007; Ellison *et al.*, 2011b). *MYC* and *MYCN* belong to the MYC family of cellular oncogenes and encode transcription factors involved in the control of cell proliferation, apoptosis and transformation (Grandori *et al.*, 2000). While the biological role of increased *MYC* expression in tumours is not yet clear, it has been shown *in vitro* to promote proliferation of cerebellar granule cells in mouse models and an anaplastic histology in cell lines (Fults *et al.*, 2002; Stearns *et al.*, 2006). Increased *MYCN* expression has been reported to drive the formation of metastatic medulloblastoma in mouse models (Swartling *et al.*, 2010).

Other oncogenes recurrently amplified in medulloblastoma include *OTX2*, which is commonly amplified in non-SHH and non-WNT tumours (de Haas *et al.*, 2006), the cell cycle progression factor *CDK6* (Mendrzyk *et al.*, 2005), and the SHH-associated *GLI2* oncogene (Northcott *et al.*, 2009; Northcott *et al.*, 2011a).

In a recent study that investigated genomic gains and losses in a cohort of 212 medulloblastomas using high-resolution SNP genotyping, focal amplications of 15 known oncogenes and focal deletions of 20 known tumour suppressor genes were identified (Northcott *et al.*, 2009). A significant finding of this study was the identification of novel and recurrent homozygous deletions and amplifications of genes targeting histone lysine methylation, particularly histone 3, lysine 9 (H3K9) (Northcott *et al.*, 2009). Post-translational modifications of histone proteins is an epigenetic mechanism that regulates gene transcription by establishing transcriptionally active (open) or silent (closed) chromatin conformations (see section 1.6). This finding is particularly interesting in light of recent genome-wide sequencing studies that have identified recurrent mutations in several chromatin-associated genes (see section

1.5.6.3), suggesting that deregulation of the epigenome is an important step in the pathogenesis of medulloblastoma.

The largest study of somatic copy number aberrations in medulloblastomas has recently been carried out by the Medulloblastoma Advanced Genomics International Consortium (MAGIC) which consists of world-wide medulloblastoma experts (Northcott *et al.*, 2012c). MAGIC investigated copy number events in 1,087 medulloblastomas by SNP arrays and concluded that somatic copy number aberrations were common in medulloblastoma and were largely enriched in distinct molecular subgroups of the disease (Northcott *et al.*, 2012c). As well as confirming well characterised copy number events, this study by Northcott et al. also identified novel genes which are recurrently affected by copy number changes in specific subgroups. These recent novel findings are discussed further in the context of their respective subgroup in section 1.5.8.

1.5.6.3 *Genetic mutations in medulloblastoma*

Until recently, there were only a small number of genes known to be recurrently mutated in medulloblastomas; these included the SHH pathway genes *PTCH1*, *SUFU* and *SMO* (see section 1.5.7.1), the WNT pathway gene *CTNNB1* (see section 1.5.7.2) and *TP53* (see section 1.5.6.1) (Northcott *et al.*, 2012a).

In recent years, whole exome and whole genome sequencing studies have significantly increased our understanding of genetic mutations in medulloblastoma, and while results suggest that paediatric medulloblastomas are characterised by fewer somatic mutations compared with adult solid tumours, these studies have identified novel recurrent mutations that may play critical roles in the development and progression of medulloblastoma (Parsons *et al.*, 2011; Jones *et al.*, 2012; Pugh *et al.*, 2012; Robinson *et al.*, 2012). The study conducted by Parsons et al. was the first whole-exome sequencing study of medulloblastoma and was carried out using Sanger sequencing technology on a cohort of 22 medulloblastomas (Parsons *et al.*, 2011). In addition to mutations in components of the SHH and WNT signalling pathways, this study identified recurrent inactivating mutations in *MLL2* and *MLL3* in 16% of patients, implying a probable tumour suppressor role in medulloblastoma (Parsons *et al.*, 2011). These genes are members of the MLL family of histone 3, lysine 4 (H3K4) trimethylases which regulate gene transcription through the methylation of H3K4 and

the propagation of active (open) transcriptional chromatin states (Vermeulen and Timmers, 2010) (see section 1.6). This study also identified mutations in *KDM6B* which encodes a histone lysine demethylase and in 2 further chromatin–associated genes, namely the SWI/SNF chromatin remodelling complex family members *SMARCA4* and *ARID1A* (Parsons *et al.*, 2011). In all cases the mutations identified were clearly distinguishable from passenger events, suggesting a critical role for histone modifiers and other chromatin-associated genes in medulloblastoma development (Parsons *et al.*, 2011).

Following the initial study by Parsons et al., a series of three independent next-generation whole exome and whole genome sequencing studies of medulloblastoma were completed (Jones *et al.*, 2012; Pugh *et al.*, 2012; Robinson *et al.*, 2012). As well as confirming the findings reported by Parsons et al., these studies have identified a common set of novel recurrent mutations that target distinct subgroups of medulloblastoma, and include additional histone modifying and other chromatin-associated genes (Jones *et al.*, 2012; Pugh *et al.*, 2012; Robinson *et al.*, 2012). Their findings are discussed further in the context of their respective disease subgroup in section 1.5.8.

Alongside copy number aberrations and mutations, epigenetic dysregulation of genes by DNA methylation has also been reported in medulloblastoma and is discussed in detail in section 1.6.

1.5.7 *Aberrant activation of developmental signalling pathways*

The Sonic hedgehog (SHH) and WNT/Wingless signalling pathways are embryonal developmental pathways that have been shown to regulate the proliferation of neural stem and/or progenitor cells in the developing cerebellum and are thus essential for normal neural and cerebellar development (Fan and Eberhart, 2008; Gilbertson and Ellison, 2008). The SHH and WNT pathways were first implicated in medulloblastoma tumourigenesis with the discovery of germline mutations in pathway components in rare inherited familial disorders Gorlin syndrome and Turcot syndrome, respectively (see section 1.5.6.1). Somatic mutations and other pathway-activating aberrations were subsequently identified in significant proportions of sporadic medulloblastomas, and SHH-activated and WNT-activated medulloblastomas are now firmly established as well defined discrete molecular subgroups of medulloblastoma which are associated

with distinct clinical, histopathological and genomic features as well as different clinical outcomes (Thompson *et al.*, 2006; Kool *et al.*, 2008; Pizer and Clifford, 2009; Cho *et al.*, 2011; Northcott *et al.*, 2011b; Kool *et al.*, 2012) (see section 1.5.8). In addition to their key roles in early development and differentiation, the SHH and WNT signalling pathways play important roles in the normal growth and patterning of fast-renewing epithelial tissues, such as the skin and intestinal lining, and therefore their aberrant regulation is also implicated in adult tumours (Taipale and Beachy, 2001).

1.5.7.1 *Sonic hedgehog signalling pathway*

The SHH signalling pathway, discovered in *Drosophila,* is a major developmental pathway in many organisms (Nusslein-Volhard and Wieschaus, 1980). Three mammalian hedgehog genes have been identified, sonic hedgehog (*SHH*), indian hedgehog (*IHH*) and desert hedgehog (*DHH*) (Huangfu and Anderson, 2005). *SHH* plays a key role in the development of the cerebellum where it regulates the proliferation of granule neurone precursor (GNP) cells (Wechsler-Reya and Scott, 1999).

SHH pathway activation occurs when the SHH ligand, which is secreted by Purkinjie neurons in the developing cerebellum (see section 1.5.7.1.1), binds to the Patched (Ptch) transmembrane receptor, abrogating its inhibitory effect on the associated G-protein coupled smoothened (Smo) receptor. (Saran, 2009). Upon its release, Smo translocates to specialised membrane structures (cilium) where it initiates a mitogenic signalling cascade, culminating in the nuclear translocation and activation of the Gli family of transcription factors and upregulated expression of several SHH target genes, including *GLI1* and *PTCH1* itself, as well as cyclin D1 and *MYCN* (Saran, 2009; Hatten and Roussel, 2011). When the SHH pathway is inactive, suppressor of fused (Sufu) binds the Gli2 and Gli3 transcription factors in the cytoplasm (Saran, 2009). Following activation Gli2 translocates to the nucleus where it is a transcriptional activator, while Gli3 acts as both an activator and a repressor of downstream signalling (Huangfu and Anderson, 2005) (Figure 1.9).

1.5.7.1.1 Sonic hedgehog signalling in cerebellar development

The cerebellum develops from granule neuron precursor (GNP) cells in the rhombic lip of the foetal brain. These cells migrate from the rhombic lip to form the external granule layer (EGL), where they undergo significant postnatal proliferation (Gilthorpe *et al.*, 2002). The SHH ligand is secreted by Purkinje neuron cells beneath the EGL (Dahmane and Ruiz i Altaba, 1999), activating the mitogenic SHH signalling pathway as described above and driving the extensive proliferation of GNPs in the EGL. Following extensive proliferation, GNPs migrate inwards past the Purkinje cells to become mature granule neurons in the internal granule layer, and the EGL no longer exists as cell proliferation stops (Saran, 2009) (Figure 1.10).

The role of SHH signalling in medulloblastoma tumourigenesis has been extensively supported by several mouse models in which the SHH pathway has been aberrantly activated and spontaneous medulloblastomas have developed that are similar to sporadic human medulloblastomas (Goodrich *et al.*, 1997; Zurawel *et al.*, 2000; Hallahan *et al.*, 2004; Lee *et al.*, 2007). The majority of published mouse models of medulloblastoma are based on SHH pathway activation and provided some of the earliest insights into the GNP cell as the most likely cell of origin for SHH activated medulloblastomas (Hatten and Roussel, 2011) (see section 1.5.9).

1.5.7.1.2 Aberrant activation of the SHH pathway in medulloblastoma

Aberrant activation of the SHH pathway in medulloblastoma results in enhanced expression of oncogenes, such as cyclin D1 and *MYCN* (see section 1.5.7.1), and leads to uncontrolled proliferation of GNPs in the EGL of the developing cerebellum (Figure 1.10). Aberrant pathway activation in sporadic tumours is associated with inactivating mutations in *PTCH1,* which have been observed in 20-30% of SHH-activated medulloblastomas (Schwalbe *et al.*, 2011; Jones *et al.*, 2012; Pugh *et al.*, 2012; Robinson *et al.*, 2012). Activating mutations in *SMO* have been reported in approximately 5% of cases, while *SUFU* mutations have also been reported in a small number of SHH tumours (Pizer and Clifford, 2009). High-level amplification of SHH downstream effectors *GLI1* and *GLI2* have also been reported in medulloblastomas (Northcott *et al.*, 2009). Recently the MAGIC consortium identified amplification of *GLI2* and focal deletion of *PTCH1* specifically enriched in SHH tumours, reporting

PTCH1 as the most frequent target of focal deletion in the distinct subgroup of SHH tumours (Northcott *et al.*, 2012c).

Overall SHH pathway activation is observed in approximately one third of all medulloblastomas and they represent a discrete molecular subgroup of tumours that are discussed in detail in section 1.5.8.
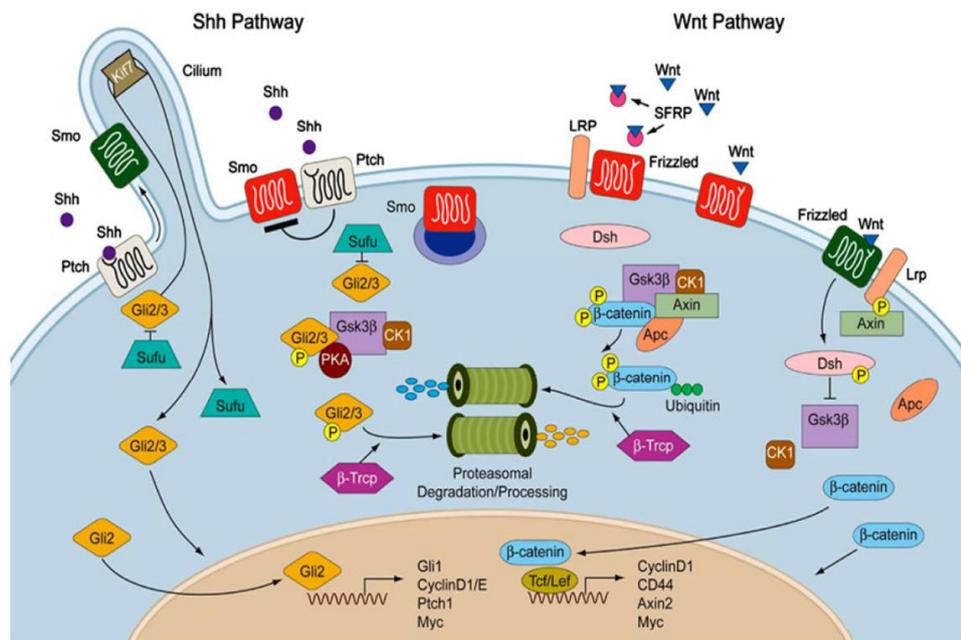
Figure 1.9. SHH and WNT signalling pathways. When the SHH pathway is inactive, Patched (Ptch) receptors inhibit Smoothened (Smo) receptors (inactive Smo receptors shown in red) and prevent access of Smoothened to the primary cilium. Suppressor of fused (Sufu) binds the downstream effectors of the SHH pathway (Gli2 and Gli3) in the cytoplasm and primary cilium. Gli2 is also targeted for degradation by the proteosomal pathway. When the SHH pathway is active, Ptch ceases to repress Smoothened (active Smo receptors-shown in green) and it moves into the primary cilium and initiates a mitogenic signalling cascade. Sufu no longer binds Gli2 and Gli3 and Gli2 translocates to the nucleus where it activates transcription of genes including the oncogenes MYC and cyclin D1. Activation of the canonical WNT pathway is initiated by binding of the secreted WNT ligand to the Frizzled receptor on the cell surface (active Frizzled receptors-shown in green). When the WNT pathway is inactive, cytosolic concentrations of downstream effector β-catenin are kept low as a result of phosphorylation and subsequent degradation via the APC/axin/Gsk3β protein complex. When the WNT pathway is active, active Frizzled receptors (shown in green) phosphorylate the dishevelled potein (Dsh) which inactivates the protein degradation complex and blocks β–catenin degradation, allowing β-catenin to translocate to the nucleus where it activates the transcription of proliferative target gene. The SFRP family of proteins negatively regulate the WNT pathway by binding WNT and preventing binding to the Frizzled receptor. Figure taken from Ellison, 2010.
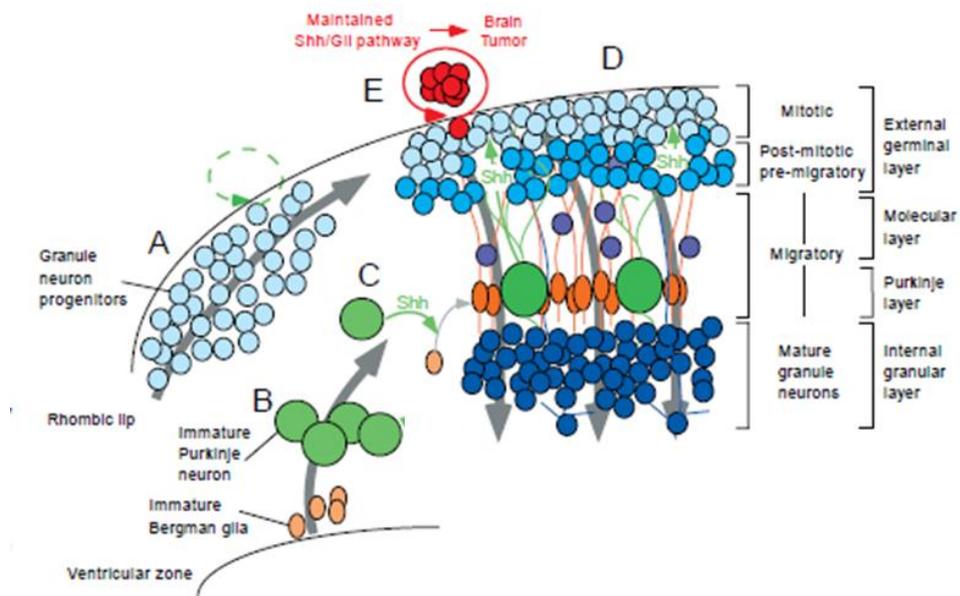
Figure 1.10. The role of SHH signalling in cerebellar development. **A**. Granule neuronal precursors (GNPs) migrate from the rhombic lip and may use the SHH pathway in a transient autocrine manner. **B.** Immature Purkinje neurons and Bergmann glia migrate from the ventricular zone towards the external granule layer (EGL). **C**. SHH secreted from Purkinje neurons induces maturation of Bergmann glia. **D**. SHH secreted by Purkinje neurons beneath the EGL drives the extensive proliferation of GNPs in the EGL. Following extensive proliferation, mature glia promote the differentiation of granule neurons and antagonise the effects of SHH. Granule cells migrate on glial fibres across the molecular and Purkinje layers to form the internal granule layer (IGL) of the cerebellum. **E**. Maintained SHH signalling in EGL cells can lead to the development of medulloblastoma. Figure adapted from Dahmane and Ruiz i Altaba, 1999.

1.5.7.2 *WNT/Wingless signalling pathway*

The WNT pathway plays a key role in the development of the central nervous system, regulating proliferation and differentiation and maintaining neural stem and precursor cell populations (Ciani and Salinas, 2005).

WNT pathway activity is directly related to the amount of free cytosolic β-catenin which is a key downstream effector of the WNT signalling pathway. Activation of the canonical WNT pathway is initiated by binding of the secreted WNT ligand to the frizzled receptor on the cell surface. In the absence of WNT, cytosolic concentrations of β-catenin are kept low as a result of complex formation with APC and axin with subsequent phosphorylation by glycogen synthase kinase 3β and casein kinase Ia (CKIa) which targets it for degradation via the ubiquitin-proteasome pathway. Upon binding the WNT ligand, frizzled phosphorylates the dishevelled protein (Dsh) which causes inactivation of the axin/Gsk3β/APC complex which targets β-catenin for phosphorylation and subsequent degradation and blocks β-catenin degradation. This

results in increased nuclear translocation of β-catenin where it binds to the gene regulatory protein Tcf/Lef and regulates the transcription of WNT target genes which include the proliferation genes cyclin D1 and *Myc* (Saran, 2009).The secreted frizzled-related protein (SFRP) comprises a family of WNT inhibitors that can bind the WNT ligand and prevent it binding to frizzled receptor (Kongkham *et al.*, 2010a) (Figure 1.9).

1.5.7.2.1 WNT signalling in cerebellar development

The significance of the WNT signalling pathway in cerebellar development is not fully understood, however, it has been shown that deletion of *Wnt-1* from mice completely blocks cerebellar development (McMahon and Bradley, 1990; Thomas and Capecchi, 1990). A recently developed mouse model which targeted expression of *CTNNB1* to progenitor cells of the lower rhombic lip, a germinal matrix situated between the neural tube and roofplate of the fourth ventricle of the developing cerebellum and gives rise to diverse neuronal brainstem lineages, developed tumours that transcriptionally resembled human WNT medulloblastomas (Gibson *et al.*, 2010). These tumours arose from the dorsal brainstem, the hindmost region of the brainstem, and provides very strong evidence that the WNT subgroup of tumours are anatomically distinct from the EGL-derived SHH medulloblastomas and arise outside the cerebellum from progenitor cells of the dorsal brainstem (Gibson *et al.*, 2010) (see section 1.5.9).

1.5.7.2.2 Aberrant activation of the WNT pathway in medulloblastoma

Activating mutations in the *CTNNB1* gene that encodes β-catenin are observed in approximately 10% of sporadic medulloblastomas (Kool *et al.*, 2012) and in up to 90% of WNT-activated medulloblastomas (Schwalbe *et al.*, 2011; Jones *et al.*, 2012; Pugh *et al.*, 2012; Robinson *et al.*, 2012). Activating mutations in *CTNNB1* leads to upregulated transcription of oncogenic WNT target genes. Other less common mutations have been found in the genes encoding APC and axin, each affecting approximately 2.5% of cases (Pizer and Clifford, 2009). Aberrant pathway activation has also been identified as a result of epigenetic silencing of the SFRP family of WNT inhibitors by DNA methylation, suggesting a potential tumour suppressor role for this gene family in medulloblastoma development and a possible mechanism by which some medulloblastomas may aberrantly activate the WNT signalling pathway (Kongkham *et al.*, 2010a).

WNT-activated medulloblastomas are significantly associated with loss of chromosome 6 (Clifford *et al.*, 2006; Fattet *et al.*, 2009). A nucleopositive β-catenin phenotype has been shown to correlate strongly with β-catenin mutations and to identify a subset of tumours with a favourable clinical outcome (Ellison *et al.*, 2005; Fattet *et al.*, 2009), and thus represents a marker of good prognosis that can be readily applied in the clinical setting to identify the favourable risk WNT-activated medulloblastomas (see section 1.5.10). WNT-activated medulloblastomas represent a discrete subgroup of tumours that are discussed in detail in section 1.5.8.

1.5.8   High-throughput analysis of the medulloblastoma genome and transcriptome

Significant advances have been made in the field of medulloblastoma genomics in recent years. Tumour profiling using high density expression and SNP arrays, and more recently next generation sequencing technologies, have provided significant insights into the molecular heterogeneity that underpins medulloblastoma. This has led to the firm recognition that medulloblastoma comprises distinct molecular subgroups, which behave very differently and will therefore require accurate diagnosis and different therapeutic approaches in the clinic.

1.5.8.1   Transcriptional profiling of medulloblastoma

Several groups have carried out transcriptomic profiling of moderate to large independent cohorts of primary medulloblastomas and revealed the existence of distinct subgroups of the disease based on their gene expression profiles (Thompson *et al.*, 2006; Kool *et al.*, 2008; Cho *et al.*, 2011; Northcott *et al.*, 2011b). All groups unequivocally identified the presence of mutually exclusive SHH-activated and WNT-activated subgroups, but they differed in the number of non-SHH/non-WNT subgroups that were found; Thompson et al, and Northcott et al. identified two further transcriptomic subgroups while Kool et al. identified three and Cho et al. identified four additional groups. Across the studies, clinico-pathological and genetic features were found to significantly differ between the molecular subgroups identified. All groups reported significant enrichment of monosomy 6 in WNT tumours as previously reported (Clifford *et al.*, 2006; Fattet *et al.*, 2009). The WNT subgroup was associated with a classic histology, while SHH tumours were enriched for chromosome 9q loss, the desmoplastic nodular phenotype and represented the majority of infant (< 3 years) cases

(Thompson *et al.*, 2006; Kool *et al.*, 2008; Cho *et al.*, 2011; Northcott *et al.*, 2011b). While a variable number of non-SHH/non-WNT subgroups were identified, they were associated with similar features across the different studies. The non-SHH/non-WNT tumour subgroups represented approximately 60% of all medulloblastomas, and while they were not found to be associated with any specific signalling pathways, they did show upregulated expression of neuronal and photoreceptor genes. These tumour subgroups were associated with the chromosomal abnormality i17q and with LCA histology (Kool *et al.*, 2008; Cho *et al.*, 2011; Northcott *et al.*, 2011b). The majority of metastatic cases fell into the non-SHH/non-WNT tumour subgroups (Kool *et al.*, 2008; Cho *et al.*, 2011; Northcott *et al.*, 2011b), with metastasis an established clinical predictor of poor patient outcome (Bailey *et al.*, 1995; Zeltzer *et al.*, 1999). In their study, Cho et al., identified a specific non-SHH/non-WNT subgroup that was enriched for *MYC* amplification and displayed a particularly aggressive behaviour (Cho *et al.*, 2011).

Several groups have designed assays that could be applied in the clinical setting to diagnose individual subgroups of medulloblastoma. In one study, immunohistochemistry, using validated subgroup-specific immunohistochemical markers, successfully distinguished SHH, WNT and non-SHH/WNT tumours (Ellison *et al.*, 2011a). Northcott et al. designed a novel four antibody approach that reliably assigned 98% of formalin-fixed paraffin embedded (FFPE) tumour samples to their four previously defined transcriptomic subgroups (Northcott *et al.*, 2011b). Minimal mRNA expression signatures have been designed and validated that reliably distinguish distinct medulloblastoma subgroups, and can routinely and rapidly be applied in the clinical setting requiring only small amounts of RNA (Schwalbe *et al.*, 2011; Northcott *et al.*, 2012d). The ability to accurately and rapidly diagnose specific medulloblastoma subgroups in the clinic will be vital for future clinical trials that will tailor therapies and employ novel molecularly targeted therapies according to tumour biology.

1.5.8.2 Four core molecular subgroups of medulloblastoma

As a result of the variable number of molecular subgroups that were identified from the series of transcriptomic profiling studies undertaken (see section 1.5.8.1), a consensus conference was held in the autumn of 2010 and after reviewing all the evidence, medulloblastoma experts from around the world reached a consensus that there are four core molecular subgroups of medulloblastoma, namely SHH, WNT, Group 3 and Group 4 (Taylor *et al.*, 2012). The WNT and SHH subgroups were named after the key signalling pathways thought to play significant roles in their pathogenesis; less is known about the underlying biology of the remaining two subgroups and so they were given generic names until they can be further refined molecularly (Taylor *et al.*, 2012). The evidence strongly suggests that there are subtypes within the core subgroups but as yet their exact number, nature and composition are unknown (Taylor *et al.*, 2012).

Based on the consensus four subgroups (SHH, WNT, Group 3 and Group 4), Kool et al. carried out a meta-analysis of all molecular and clinical data that was available from independent studies for a total of 550 primary medulloblastomas, and from this they defined the key clinical, histological and genomic features of the four subgroups. They also performed a comprehensive survival analysis, assessing the prognostic significance of subgroup status and the prognostic relevance of molecular markers significantly enriched within distinct subgroups (Kool *et al.*, 2012). This meta-analysis represents the largest series of biology data on medulloblastoma reported to date.

Northcott et al. conducted a meta-analysis of a series of three medulloblastoma next-generation sequencing studies that have recently been completed (Jones *et al.*, 2012; Pugh *et al.*, 2012; Robinson *et al.*, 2012), and defined a common set of novel and known recurrent mutations targeting distinct medulloblastoma subgroups (Northcott *et al.*, 2012a). They also included in their analysis the large-scale copy number study recently completed by the MAGIC consortium (Northcott *et al.*, 2012c) and defined the most prevalent copy number aberrations targeting distinct subgroups (Northcott *et al.*, 2012a). The key features that define the four core subgroups of medulloblastoma, as described by the aforementioned extensive meta-analysis studies, are described below and are summarised in Table 1.5.

1.5.8.2.1 SHH subgroup of medulloblastoma

The SHH subgroup represents approximately 28% of all medulloblastomas (Kool *et al.*, 2012). They are an intermediate prognosis subgroup, with overall survival rates ranging from ~60% to 80% (Cho *et al.*, 2011; Northcott *et al.*, 2011a; Northcott *et al.*, 2011b; Kool *et al.*, 2012). SHH medulloblastomas occur almost equally in males and females and they exhibit a bimodal age distribution, accounting for the majority of both infant and adult medulloblastomas but only for a minority of childhood cases. There is a strong association between infant SHH tumours and desmoplastic/nodular histology, and SHH tumours have a better outcome in infants compared to children and adults (Kool *et al.*, 2012) (Table 1.5). These findings are consistent with previous studies reporting the association between desmoplastric/nodular histology, infant medulloblastoma and favourable prognosis (see section 1.5.1.2).

There are frequent deletions of chromosomes 9q and 10q in SHH tumours (Kool *et al.*, 2012). Copy number aberrations affecting the SHH target genes, *MYCN* and *GLI2* and mutations of *PTCH1* are typical of this subgroup (Northcott *et al.*, 2012a) (Table 1.5). Mutations affecting the histone modifier *MLL2* (Parsons *et al.*, 2011) were enriched in both SHH and WNT medulloblastomas compared with Group 3 and Group 4 (Northcott *et al.*, 2012a). Mutations in *BCOR* and *LDB1*, components of the nuclear receptor co-repressor (N-CoR) complex which is associated with histone deacetylase (HDAC) activity, were mutually exclusive in SHH tumours (Northcott *et al.*, 2012a). While gene mutations were generally low frequency (Table 1.5), these findings suggest that disruption of the histone code and chromatin organisation may be a significant feature of SHH medulloblastoma (Table 1.5).

1.5.8.2.2 WNT subgroup of medulloblastoma

WNT tumours represent the smallest subgroup (11%). They occur almost equally in males and females and are rare in infants, typically occurring in children over the age of 3 years. WNT medulloblastomas have the best survival outcome of all the subgroups and this has been firmly established in several studies (Ellison *et al.*, 2005; Gajjar *et al.*, 2006; Fattet *et al.*, 2009; Kool *et al.*, 2012), with over 95% of patients surviving their disease (Kool *et al.*, 2012). WNT tumours are mainly classic histology and are infrequently metastatic (Kool *et al.*, 2012) (Table 1.5).

WNT medulloblastomas are characterised by loss of chromosome 6 (Clifford *et al.*, 2006; Kool *et al.*, 2012). Apart from this, they exhibit very few other chromosomal abnormalities or copy number events (Kool *et al.*, 2012) (Table 1.5). Mutations in *CTNNB1* are almost exclusively found in WNT tumours; *CTNNB1* is the most frequently mutated gene in medulloblastoma and is mutated in 91% of WNT tumours (Northcott *et al.*, 2012a). *DDX3X* is the second most frequently mutated gene and mutations are found in approximately 50% of WNT tumours as well as in SHH medulloblastomas (Northcott *et al.*, 2012a). *DDX3X* is involved in multiple cellular processes including cell cycle regulation, and it is thought that it is required to maintain the lineage of progenitors of the lower rhombic lip (Robinson *et al.*, 2012), which are the proposed cell of origin of WNT medulloblastomas (see section 1.5.9). As well as enrichment of mutations in *MLL2*, mutations in the chromatin modifier *SMARCA4* are frequently found in WNT tumours (Northcott *et al.*, 2012a) (Table 1.5).

1.5.8.2.3 Group 3 medulloblastomas

Group 3 medulloblastomas currently have the worst outcome of the four subgroups (Cho *et al.*, 2011; Northcott *et al.*, 2011b; Kool *et al.*, 2012). They are more common in males, are rarely found in adult populations and are frequently of the LCA histological subtype. Group 3 tumours are frequently metastatic at presentation (Kool *et al.*, 2012) (Table 1.5).

MYC amplification is highly enriched in Group 3 tumours and is observed in approximately 17% of Group 3 medulloblastomas (Cho *et al.*, 2011; Kool *et al.*, 2012). i17q is frequently found in these tumours (Table 1.5). Multiple studies have confirmed that *MYC* amplification is significantly associated with a poor prognosis and with the large cell and anaplastic histology (Brown *et al.*, 2000; Aldosari *et al.*, 2002a; Eberhart *et al.*, 2002b; Lamont *et al.*, 2004; Pfister *et al.*, 2009). In their meta-analysis, Kool et al. reported that all patients with *MYC* or *MYCN* amplification had a worse prognosis compared to cases without amplification. However, in the context of Group 3 tumours, they found that patients without *MYC* do equally poor as those carrying the amplification (Kool *et al.*, 2012).

In recent studies, an aberrant H3K27 (histone 3, lysine 27) methylation state was reported to occur specifically in Group 3 and Group 4 medulloblastomas, further

implicating disruption of the histone code as a significant oncogenic feature in medulloblastomas (Robinson *et al.*, 2012; Dubuc *et al.*, 2013). Recurrent mutations in the histone modifiers *MLL2* and *MLL3* and in the chromatin modifier *SMARCA4* are found in Group 3 tumours (Northcott *et al.*, 2012a). Gene mutations are, however, low frequency, and many tumours are not defined by mutations, leaving many driver events still undetected (Table 1.5).

1.5.8.2.4 Group 4 medulloblastomas

Group 4 are the most common medulloblastoma subgroup and they have an intermediate prognosis similar to SHH tumours. They are more common in males and are mainly of classic histology, although they can be LCA. Metastases are found in approximately one third of patients (Kool *et al.*, 2012) (Table 1.5). Group 4 medulloblastomas that present with metastases have been shown to have a poorer clinical outcome compared with those without (Kool *et al.*, 2012).

i17q is frequently found in Group 4 tumours (Kool *et al.*, 2012) and loss of one copy of chromosome X in female Group 4 medulloblastomas has also been reported (Northcott *et al.*, 2012a). The proto-oncogenes *MYCN* and *CDK6* are recurrently amplified in Group 4 tumours (Northcott *et al.*, 2012a) (Table 1.5), and *MYCN* amplification was associated with a worse outcome (Kool *et al.*, 2012). Additionally, chromosome 17 aberrations in Group 4 tumours were associated with a worse outcome (Kool *et al.*, 2012).

Recurrent mutations in the histone modifying genes, *MLL2* and *MLL3*, are present in Group 4 medulloblastomas (Northcott *et al.*, 2012a) (Table 1.5). While gene mutations are generally low frequency, of particular interest was the finding that *KDM6A*, a H3K27 demethylase, is the most frequently mutated gene in Group 4. An additional chromatin remodelling gene, *ZMYM3*, is also recurrently mutated in Group 4 (Northcott *et al.*, 2012a) (Table 1.5). As previously mentioned, aberrant H3K27 methylation has been found to be prevalent in both Group 3 and Group 4 medulloblastomas (Robinson *et al.*, 2012; Dubuc *et al.*, 2013), suggesting that dysregulation of epigenetic processes may be a significant feature of Group 3 and Group 4 pathology.

	SHH	WNT	Group 3	Group 4
Clinical features				
Prevalence	28%	11%	27%	34%
Gender ratio (male/female)	~1/1	~1/1	~2/1	~2/1
Age	adults>infants>children	children>adults	children>infants; rarely adults	children>adults; rarely infants
Histology	classic>D/N>LCA	classic; rarely LCA	classic>LCA; rarely D/N	classic; rarely LCA ; rarely D/N
Metastasis	~15%	~10%	~30%	~30%
Overall survival (5 years)	~60% (Infant SHH ~75%)	~95%	~50%	~60%
Genomic features				
Expression signature	SHH *MYCN*/minimal *MYC*	WNT	Photoreceptor *MYC*/minimal *MYCN*	Neuronal *MYCN*/minimal *MYC*
Cytogenetics	-9q; -10q; -11;-16, -17p; +3q	-6	-8; -10q; -16q;-17p; +1q; +7; +17q; i17q	-8; -11; -16q;-17p; +7; +17q; i17q
Driver genes **Recurrent mutations**	*PTCH1* (24%) *MLL2* (12%) *TP53* (11%) *DDX3X* (11%) *BCOR* (7%) *LDB1* (7%) *TCF4* (5%)	*CTNNB1* (91%) *DDX3X* (50%) *SMARCA4* (25%) *MLL2* (13%) *TP53* (13%)	*SMARCA4* (11%) *SPTB* (6%) *MLL2* (5%) *CTDNEP1* (5%) *LRP1B* (5%) *TNXB* (5%)	*KDM6A* (12%) *MLL3* (5%) *ZMYM3* (5%) *CBFA2T2* (3%)
Driver genes **Focal SCNAs**	*MYCN* (8%) *GLI2* (5%) *PTCH1* (4%)		*MYC* (17%)	*MYCN* (6%) *CDK6* (5%)

Table 1.5. Summary of the clinical, histopathological and genetic features associated with the four molecular subgroups of medulloblastoma. Clinical features, expression signatures and cytogenetic features were described by Kool et al. following their meta-analysis of clinical and molecular data for 550 primary medulloblastomas (Kool *et al.*, 2012). Subgroup-specific mutational events and copy number events are summarised along with their prevalence within each subgroup, adapted from (Northcott *et al.*, 2012a). Medulloblastoma is characterised by relatively few somatic mutational events which are generally low frequency and many driver events have yet to be detected, particularly in the Group 3 and Group 4 subgroups. D/N: desmoplastic nodular; LCA: large cell anaplastic.

1.5.9 Cellular origins of medulloblastoma

Current evidence in mouse models of medulloblastoma suggests that the four molecular subgroups of medulloblastoma most likely have different cellular origins. The vast majority of mouse models available recapitulate the SHH subgroup of tumours (Hatten and Roussel, 2011) and have been available for more than a decade. Inactivation of *Ptch1* (*Ptch1*$^{+/-}$) (Goodrich *et al.*, 1997; Wetmore *et al.*, 2000) and activation of smoothened (*Smoa1*) (Hallahan *et al.*, 2004; Hatton *et al.*, 2008) are the most commonly used initiating events to generate tumours replicating human SHH medulloblastomas in mice. Using models such as these, SHH medulloblastomas have been shown to arise in cerebellar granule neuron precursors (GNPs) of the cerebellar external granule cell layer (EGL) (Schuller *et al.*, 2008; Yang *et al.*, 2008) (see Figure 1.10).

Mice that express an activated mutant β-catenin allele alongside *TP53* deletion develop tumours that recapitulate human WNT medulloblastomas and have identified lower rhombic lip progenitors of the dorsal brainstem as the cell of origin for WNT tumours (Gibson *et al.*, 2010).

Against a background of *TP53* deletion, *MYC*-driven mouse models of Group 3 medulloblastomas have recently been developed, that have aggressive phenotypes similar to the human Group 3 tumours (Kawauchi *et al.*, 2012; Pei *et al.*, 2012). These models have identified 3 potential cells of origin for Group 3 tumours: atonal homologue 1 (ATOH1)-positive GNPs from the EGL (Kawauchi *et al.*, 2012); ATOH1-negative GNPs from the EGL (Kawauchi *et al.*, 2012) and prominin 1(PROM1)-positive, lineage negative neural stem cells (NSCs) (Pei *et al.*, 2012). Meanwhile the cellular origin of Group 4 medulloblastomas is still unknown.

These findings that distinct subgroups arise from different cellular origins will have important implications for the study of abnormal tumour-specific events in medulloblastoma. Normal control cerebellar tissue comprises a heterogeneous population of cells and signals generated will represent an average of these cell populations, and may not be reflective of the precise cell of origin of the tumour. In order to make rational and accurate assessments of primary tumour *versus* normal it will be necessary to account for subgroup-specific differences. This presents a significant challenge in the field of medulloblastoma research as the cellular origins of Group 4 and

the precise cell of origin of Group 3 tumours are not yet known and also acquiring such cell-type specific controls will not be easy.

1.5.10 Medulloblastoma molecular markers as prognostic indicators

As previously discussed, positive nuclear staining for the WNT pathway marker β-catenin (Ellison *et al.*, 2005; Fattet *et al.*, 2009) and amplification of the *MYC* oncogene (Brown *et al.*, 2000; Aldosari *et al.*, 2002a; Eberhart *et al.*, 2002b; Lamont *et al.*, 2004; Pfister *et al.*, 2009) are firmly established molecular prognostic markers in medulloblastoma. Both markers have been validated in multiple clinical trials cohorts and have subsequently been incorporated into a new risk stratification model that recognises low-risk (favourable risk), standard-risk and high-risk categories of medulloblastoma, and will form the basis of the forthcoming PNET5 clinical trial (Pizer and Clifford, 2009) (see section 1.5.11).

While no other molecular markers have been validated in trials cohorts, progress is being made and it is anticipated that as more biological insights are gained and molecular subclassification further refined, markers will be identified that will predict survival both between and within molecular subgroups and more accurately inform treatment choices for patients in the clinical setting.

FSTL5 expression has recently been identified as a marker of poor prognosis both within and across medulloblastoma subgroups (Remke *et al.*, 2011). Remke et al. reported subgroup-specific expression patterns for *FSTL5*, with lowest expression in the favourable-risk WNT tumours and over-expression in a subset of non-SHH/non-WNT tumours which are associated with a poorer clinical outcome. Immunostaining for *FSTL5* confirmed that its expression was prognostic across the subgroups in univariate and multivariate survival models, and also within the non-SHH/non-WNT tumours (Remke *et al.*, 2011). *FSTL5* expression represents a promising marker for identifying a subset of high-risk and lower-risk non-SHH/non-WNT tumours for more effective therapeutic stratification.

In another study, upregulated expression of the histone deacetylases, *HDAC5* and *HDAC9*, has been shown to correlate with poor prognostic medulloblastoma subgroups (defined as tumours with balanced chromosome 6 or gain of 6q or 17q) and to be associated with poor overall survival (Milde *et al.*, 2010). Over-expression of both

HDAC5 and *HDAC9* was identified as independent high-risk indicators and their knockdown in medulloblastoma cells reduced growth and induced apoptosis, suggesting a potential prognostic and functional role for *HDAC5* and *HDAC9* expression in medulloblastoma (Milde *et al.*, 2010). Recently, independent high-risk DNA methylation biomarkers have been identified in medulloblastoma (Schwalbe *et al.*, 2013) (see section1.7.2.2).

1.5.11 Improved risk stratification models and future clinical trials in medulloblastoma

It is clear that the identification of prognostically important histological and molecular features offers significant potential for improved stratification of medulloblastoma. Molecular subclassification is refining the discovery of prognostic biomarkers as well as novel therapeutic targets (see section 1.5.12) and it is beginning to inform the therapeutic management of medulloblastoma.

Before prognostic markers can be incorporated into a stratification model for future clinical trials they must be robustly validated in independent patient cohorts. Table 1.6 details the molecular and histopathological markers that have shown significant associations with prognosis in ≥ 2 clinical trials and have been incorporated into an improved stratification model alongside current clinical risk variables (see section 1.5.4). The new refined stratification model identifies three distinct medulloblastoma risk groups: high-risk, standard risk and low-risk (favourable-risk) (Pizer and Clifford, 2009). The utility of the molecular markers alongside clinicopathological disease features in disease outcome prediction was validated in a primary medulloblastoma cohort of 207 patients (3-16 years) from the SIOP PNET3 trial (Ellison *et al.*, 2011b).

This new stratification model will form the basis of the forthcoming PNET5 clinical trial which will test whether children (3-16 years of age) with favourable-risk disease may benefit from de-escalated therapy, with the aim of maintaining survival rates while reducing the late effects of treatment. This trial will also aim to improve survival rates in the standard-risk group following removal of the favourable-risk cases (Pizer and Clifford, 2009).

	Disease marker
Low-risk	β-catenin nucleopositive **and** No MYC amplification **and** Total surgical resection **and** No evidence of metastatic disease (M0) **and** Classic or D/N histology
Standard-risk	β- catenin nucleonegative **and** No MYC amplification **and** Total surgical resection **and** No evidence of metastatic disease **and** Classic or D/N histology
High-risk	β-catenin nucleopositive or nucleonegative **and** MYC amplification **or** Sub-total surgical resection **or** Evidence of metastatic disease (M1-M3) **or** LCA histology

Table 1.6. Refined risk stratification model for children (3-16 years) with medulloblastoma. Proposed new stratification model for non-infant children with medulloblastoma based on validated clinical, histopathological and molecular disease biomarkers. This model will be used to inform treatment strategies in the forthcoming PNET5 clinical trial. Table adapted from Pizer and Clifford, 2009.

1.5.12 Molecular targeted therapy in medulloblastoma

The use of molecularly targeted therapies in medulloblastoma is still in its infancy. The greatest advances have been made in targeting the SHH signalling pathway and small molecule inhibitors of the SHH pathway have been developed, most of which act on smoothened (SMO) (Low and de Sauvage, 2010; Ng and Curran, 2011). SMO-inhibitors have recently entered clinical trials for SHH-driven medulloblastomas. Despite showing a pronounced initial response, studies have shown that both humans and mice with PTCH-and SMO-driven tumours quickly develop resistance, suggesting these agents may be ineffective as monotherapy (Rudin *et al.*, 2009; Metcalfe and de Sauvage, 2011). Furthermore, SMO-inhibitors are unlikely to be active against tumours driven by amplifications and mutations in downstream pathway components such as SUFU and GLI (Ramaswamy *et al.*, 2011). A mutation in SMO has also been identified that has no effect on SHH signalling but disrupts drug binding, reducing its effect and leading to disease recurrence (Yauch *et al.*, 2009). While SMO-inhibitors may be ideally suited to young infants with germline *PTCH1* mutations, such as those observed in Gorlin syndrome medulloblastoma (see section 1.5.6.1), there are concerns of developmental complications with their use in young infants, with SHH pathway inhibition found to disrupt bone growth in young mice (Kimura *et al.*, 2008).

A targeted WNT pathway inhibitor (XAV939) has also been developed and is thought to inhibit WNT signalling through the stabilisation of axin, which is a key component of the β-catenin destruction complex (Huang *et al.*, 2009) (see section1.5.7.2). The application of this therapy to medulloblastoma is unclear as the majority of WNT medulloblastomas are characterised by activating mutations in β-catenin, which acts downstream of axin. Furthermore consideration must be given to the importance of WNT signalling in maintaining stem cell populations (Ciani and Salinas, 2005) and the wide ranging adverse effects that would be associated with its inhibition.

Owing to the identification of recurrent mutations in histone modifiers and other chromatin-modifying genes (see section 1.5.6.3 and 1.5.8.2), an exciting avenue that is currently being explored is the use of targeted epigenetic therapy in medulloblastoma. Epigenetic modifications, which include the addition of methyl groups to DNA and the post-translational addition of among others acetyl and methyl groups to core histone proteins, are discussed in section 1.6. Briefly, the interaction between these different

epigenetic mechanisms plays a key role in establishing chromatin architecture and regulating gene expression (see section 1.6). Epigenetic modifications are reversible making them an attractive therapeutic target (Dawson and Kouzarides, 2012) and this is currently an area of active interest in the pharmaceutical industry, particularly in the development of drugs targeting histone lysine post-translational modifications. Such targeted agents offer the potential for altering transcriptional programmes in cancer cells. Targeted epigenetic therapy has shown considerable promise for the indirect targeting of the *MYC* oncogene, the inhibition of which has proved very challenging to scientists over the years (Batora *et al.*, 2013). A recently developed small-molecule bromodomain inhibitor, JQ1, has been shown to downregulate *MYC* expression (Mertz *et al.*, 2011; Loven *et al.*, 2013). Bromodomains are commonly found in chromatin-associated proteins that recognise and selectively bind to acetyl-lysine residues, mediating their transcriptional effects (Filippakopoulos and Knapp, 2012). JQ1 interferes with the capacity of bromodomain-containing proteins to bind acetylated lysines and through this mechanism downregulates the expression of *MYC* (Mertz *et al.*, 2011; Loven *et al.*, 2013). JQ1 represents an exciting and promising treatment for the poor prognostic Group 3 medulloblastomas and *in vitro* testing in medulloblastoma is currently underway (Batora *et al.*, 2013). DZNep (3-Deazaneplanocin A) is another targeted epigenetic therapy that is soon to be tested in upcoming Phase II clinical trials in America in subsets of medulloblastoma (Gajjar *et al.*, 2013). DZNep is an inhibitor of EZH2 which encodes a histone lysine methyltransferase and functions within the Polycomb repressive complex 2 (PRC2) to promote histone 3, lysine 27 trimethylation (H3K27me3), a mark of transcriptionally silent chromatin (Batora *et al.*, 2013) (see section 1.6). As previously discussed in section 1.5.8.2, an aberrant H3K27 methylation state has been reported to occur specifically in Group 3 and Group 4 meduloblastomas and DZNep, therefore, represents an exciting novel targeted therapy for Group 3 and Group 4 patients harbouring these defects.

1.5.13 Summary

Significant advances in the field of medulloblastoma genomics have led to the recognition of medulloblastoma as a disease comprising four very distinct molecular subgroups, which will require accurate diagnosis in the clinic and different therapeutic strategies. While these genomic advances will translate to improved risk stratification and to improved treatment options for patients with medulloblastoma, it is becoming increasingly clear that the molecular pathogenesis of the disease cannot be completely attributed to genetic events. Recent whole exome and whole genome sequencing studies have shown that medulloblastoma is characterised by relatively few somatic mutational events, and events identified are generally low frequency (Parsons *et al.*, 2011; Jones *et al.*, 2012; Pugh *et al.*, 2012; Robinson *et al.*, 2012). As a result, there are many currently undetected driver events in medulloblastoma pathogenesis, particularly in the Group 3 and Group 4 tumour subgroups (see section 1.5.8.2).

There is increasing recognition that epigenomic deregulation may play a critical role in the development of medulloblastoma (Jones *et al.*, 2013). Tumour suppressor genes that are silenced by DNA methylation in medulloblastoma have been known for quite some time (see section 1.6.3.1) and, as discussed, several chromatin modifying genes which are targeted by mutation in medulloblastoma have recently been identified (see sections 1.5.6.3 and 1.5.8.2). Comprehensive characterisation of the medulloblastoma epigenome offers considerable potential for the identification of further driver events involved in tumour initiation and progression. Studies of epigenetic deregulation in medulloblastoma have gained considerable interest and pace in recent years and have been paralleled with advances in the technology for studying genome-wide DNA methylation. DNA methylation is a critical epigenetic mechanism of gene regulation and its importance in cancer has been widely established (Jones and Laird, 1999; Jones and Baylin, 2002; Baylin and Jones, 2011). The role of DNA methylation in regulating gene expression during normal development and during tumourigenesis is discussed in detail in the sections that follow.

1.6 Epigenetics

Epigenetics is defined as heritable changes in gene activity and expression that occur without a change in the primary DNA sequence (Bird, 2007; Goldberg *et al.*, 2007). Epigenetic mechanisms, which include DNA methylation (Feinberg and Tycko, 2004; Jones and Baylin, 2007), histone modifications (Kouzarides, 2007; Li *et al.*, 2007), nucleosome remodelling (Li *et al.*, 2007) and the impact of non-coding RNAs (Mattick and Makunin, 2006), play major roles in regulating gene activity and expression during development and differentiation as well as in disease (Jones and Baylin, 2007; Jaenisch and Young, 2008). Gene transcription is a remarkably intricate process that is tightly regulated at many genetic and epigenetic levels (Lemon and Tjian, 2000). It is becoming increasingly clear that epigenetic control of gene regulation is itself highly complex, with significant crosstalk between different epigenetic processes evident (Goldberg *et al.*, 2007). Much of the research in the field of epigenetics has focused on two processes, namely DNA methylation and histone modifications, and how the interplay between them establishes a transcriptionally active euchromatic state or transcriptionally silent heterochromatic state (Jones and Baylin, 2002; Jaenisch and Bird, 2003; Jones and Baylin, 2007; Blackledge and Klose, 2011; Poetsch and Plass, 2011).

1.6.1 DNA methylation and CpG islands

DNA methylation has been described as a "powerful mechanism for the suppression of gene activity" (Jones and Laird, 1999). In the human genome, the predominant form of DNA methylation occurs symmetrically at carbon 5 position of cytosine residues on both strands of CpG dinucleotides (Bird, 2002). DNA methylation at CpG dinucleotides acts as a stable and heritable epigenetic mark that is generally associated with repressed chromatin states and inhibition of transcriptional activation (Klose and Bird, 2006) (see section 1.6.2). While CpG methylation affects approximately 70-80% of CpG sites, both the CpG sites and their degrees of methylation are unevenly distributed in the genome (Bird *et al.*, 1987). Of great significance is the presence of CpG islands which are unmethylated GC-rich regions that are predominantly found at the promoter region of many genes (Bird *et al.*, 1987).

A CpG island is formally defined as a region of at least 200 base-pairs with an average CG content greater than 50% and an observed-to-expected CpG ratio greater than 0.6

(Gardiner-Garden and Frommer, 1987). CpG islands are a feature of housekeeping genes as well as many genes with a tissue-restricted expression pattern (Gardiner-Garden and Frommer, 1987). Approximately 70% of genes are associated with CpG islands, the great majority of which are unmethylated at all stages during development and in all tissue types (Bird, 2002). During normal development a small number of CpG islands become methylated and this methylated state is associated with long-term gene silencing (Bird, 2002). Developmentally programmed CpG island methylation is involved in fundamental biological processes such as genomic imprinting and X chromosome inactivation (Bird, 2002).

Genomic imprinting, which describes the non-equivalence in allele expression for certain gene loci in a parent-of-origin specific manner (Reik *et al.*, 2001) and X chromosome inactivation in females involve the silencing of one allele only (Bird, 2002). The importance of DNA methylation in mediating this silencing has been demonstrated in mouse models in which DNA methyltransferase enzymes (DNMTs) are disrupted leading to reactivation of X-linked genes (Sado *et al.*, 2000) and the disruption of the monoallelic expression of several imprinted genes (Li *et al.*, 1993). In normal somatic cells, CpG island methylation is also a feature of some tissue-specific genes and is thought to be critical in regulating their cell-type specific expression patterns (Strathdee *et al.*, 2004b). Examples include the *maspin* gene which is methylated and repressed in mesenchymal and haematopoietic cells (Futscher *et al.*, 2002) and *MCJ* which is methylated and repressed in epithelial cells (Strathdee *et al.*, 2004a). De novo methylation of CpG islands also occurs in certain tissues during aging and is a significant feature of cancer cells (Bird, 2002) (see section 1.6.4).

Transcriptionally active genes are associated with a methylation-free CpG island, however, the exact mechanisms by which CpG islands remain methylation-free and the precise mechanisms by which they contribute to gene regulation are not yet fully understood (Blackledge and Klose, 2011). As some tissue-specific genes have CpG islands that remain unmethylated in all tissues regardless of expression, an unmethylated CpG island is not indicative of active gene transcription but rather establishes a 'permissive' state of gene transcription (Bird, 2002). Early evidence suggests that a methylation-free CpG island may be created by gene promoters that are active in early development and that this may be as a result of a transcriptionally active

chromatin state excluding the DNA methylation machinery (Bird, 2002). These findings suggest that unmethylated CpG islands represent footprints of embryonic promoter activity (Bird, 2002). The unmethylated state of promoter CpG islands has been shown to be strongly influenced by transcription factor binding with CpG islands acquiring methylation if they are depleted of transcription factor binding sites (Brandeis *et al.*, 1994; Macleod *et al.*, 1994). Recent studies have identified a new family of DNA binding proteins (ZF-CxxC proteins) that are specifically recruited to unmethylated CpG dinucleotides and participate in establishing a transcriptionally active chromatin state and in gene activation (Blackledge *et al.*, 2010; Thomson *et al.*, 2010). These studies demonstrated that the ZF-CxxC domain of this protein family specifically targets unmethylated CpG islands, providing the first evidence that CpG islands are directly interpreted through recognition of non-methylated DNA (Blackledge *et al.*, 2010; Thomson *et al.*, 2010; Blackledge and Klose, 2011).

Outside of the predominantly unmethylated CpG islands, the remaining CpG sites in the genome are largely methylated (Bird, 2002). Activation of transposable element-derived promoters is a well-documented consequence of loss of this DNA methylation (Bird, 2002). Transposable element (TE)-derived repetitive DNA sequences, such as LINES (long interspersed nuclear elements) and SINES (short interspersed nuclear elements) retrotransposons, comprise almost half of the human genome and are mobile DNA sequences that can migrate to different regions of the genome (Smit, 1996). TE sequences are heavily methylated and it is thought that this methylation contributes to the suppression of transposition and TE-mediated changes in gene expression and the overall maintenance of genomic stability (Morgan *et al.*, 1999; Slotkin and Martienssen, 2007).

1.6.2 DNA methylation, histone modifications and chromatin structure

Genomic DNA is packaged with histone proteins to form protein/DNA complexes known as chromatin. The basic unit of chromatin is the nucleosome which is composed of ~146bp of DNA wrapped around an octamer of the four core histone proteins (H2A, H2B, H3 and H4) (Luger *et al.*, 1997). The amino terminal tails of the core histone proteins are subject to several post-translational modifications by histone modifying enzymes. These modifications include acetylation, methylation, phosphorylation, ubiquitination and sumoylation and are critical for regulating chromatin structure and

function, and influence many DNA processes including gene transcription (Koch *et al.*, 2007).

Acetylation and methylation of histone lysine residues are the two most common post-translational modifications that interact with DNA methylation to establish transcriptionally active euchromatic states and transcriptionally inactive heterochromatic states (Jones and Baylin, 2002; Jones and Baylin, 2007; Blackledge and Klose, 2011; Poetsch and Plass, 2011). Euchromatin is a region where the DNA is accessible to the transcription machinery and other DNA binding proteins, representing an open conformation due to the relaxed state of well-spaced nucleosomes. In contrast, heterochromatic regions are areas where DNA is packaged into tightly compacted nucleosomes rendering it inaccessible to transcription factors and other chromatin-associated proteins (Jones and Baylin, 2002) (Figure 1.11). The balance between euchromatin and heterochromatin in a given cell type ensures that the gene expression pattern is stably inherited in daughter cells (Bird, 2002).

Unmethylated CpG islands, especially those that are associated with active gene promoters, reside in regions of euchromatin composed of widely and irregularly spaced nucleosomes (Jones and Baylin, 2002). Active gene promoters are characterised by nucleosome depleted regions (NDRs) immediately upstream from their transcription start site, while flanking nucleosomes carry the histone modification H3 trimethylated on lysine 4 (H3K4me3) as well as highly acetylated histone H3 and H4 lysine residues (Jenuwein and Allis, 2001; Baylin and Jones, 2011) (Figure 1.11). These flanking nucleosomes are also associated with the histone variant H2A.Z which dynamically regulates chromatin conformation by creating domains that are poised for transcriptional activation through its ability to destabilise nucleosomes to facilitate access to DNA and gene transcription (Fan *et al.*, 2002; Baylin and Jones, 2011).

In contrast, methylated DNA resides in regions of transcriptionally silent heterochromatin composed of tightly packed nucleosomes (Lin *et al.*, 2007; Baylin and Jones, 2011) and where it occurs at gene promoters, they lack the H2A.Z histone variant and are accompanied by repressive histone modifications, such as H3K27me3 and H3K9me3 (Jones and Baylin, 2002; Lin *et al.*, 2007) (Figure 1.11). Polycomb group (PcG) proteins are transcriptional repressors that play a central role in gene regulation during differentiation and development through their histone modification activities, in

particular applying the repressive H3K27me3 inhibitory mark and conferring repressive chromatin structure at gene promoters (Bracken and Helin, 2009) (Figure 1.11). The PcG proteins form multiprotein repressive complexes, known as Polycomb repressive complexes (PRCs), and they have been shown to co-occur with the active histone mark, H3K4me3, around the TSS of key developmental and lineage-specific genes in embryonic stem cells (ESCs) (Mikkelsen *et al.*, 2007). This dynamic bivalent chromatin state is thought to confer regulatory flexibility by maintaining ESC pluripotency through their quiescence while allowing for their rapid transcriptional activation during differentiation (Mikkelsen *et al.*, 2007). Bivalent gene promoters are prone to *de novo* DNA methylation in cancer and pre-cancerous cells (Ohm *et al.*, 2007; Widschwendter *et al.*, 2007).
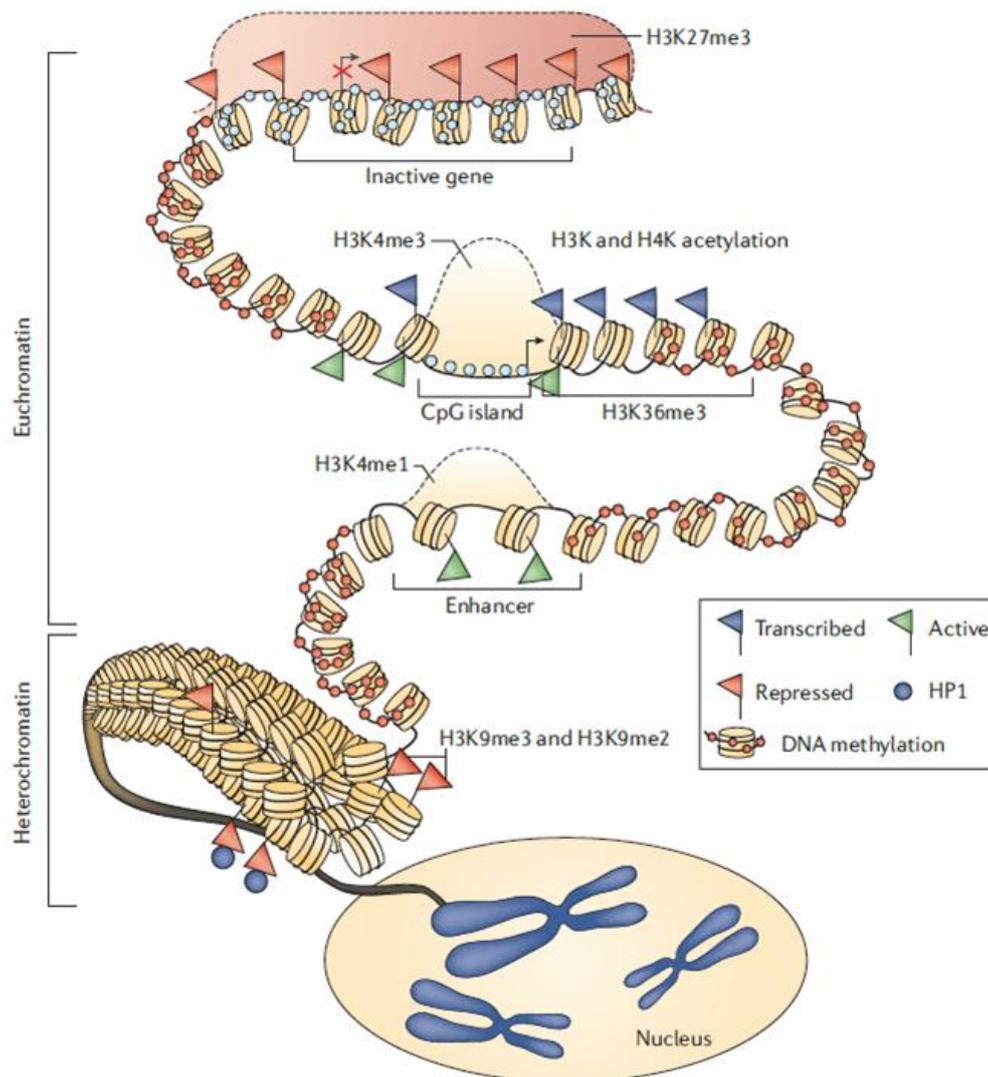
Figure 1.11. Model of the overall structure of the epigenome in normal human cells. Diagram shows the balanced state of chromatin in the genome, which maintains the packaging state of DNA; DNA methylation, histone modifications and nucleosome positioning associated with the different states of euchromatin and heterochromatin are shown. The top of the figure shows a silenced gene (indicated by a red X over the transcription start site); its promoter CpG island is occupied by a repressive Polycomb group (PcG) complex (indicated by the red shaded area) that mediates chromatin changes, including the repressive histone modification trimethylation of lysine 27 on histone 3 (H3K27me3). CpG dinucleotides within this region are unmethylated (blue circles) (unmethylated DNA is not indicative of gene transcription), and nucleosomes are postioned over the transcriptional start sites. Sites upstream from the gene promoter are heavily methylated (red circles). The gene promoter beneath the silenced gene has adopted a fully active transcriptional state, associated with the active H3K4me3 marks and acetylation of key H3 and H4 lysines. Nucleosomes are absent from the transcription start site. Towards the bottom of the figure the packaging of the majority of DNA into a transcriptionally repressed heterochromatin state is shown; nucleosomes are densely packed, repressive H3K9me2 and H3K9me3 marks are present and DNA is heavily methylated. Figure taken from Baylin and Jones, 2011.

1.6.2.1 *DNA methylation and transcriptional repression*

The exact mechanisms by which DNA methylation mediates transcriptional silencing are not yet completely understood however insights have been gained from an increased understanding of the DNA methylation machinery.

There are three active DNA methyltransferase enzymes (DNMT1, DNMT3A and DNMT3B) that catalyse the transfer of a methyl group from *S*-adenosyl-L-methionine to cytosine residues; DNMT1 is the major maintenance methyltransferase that has a preference for hemimethylated CpG sites and is responsible for copying pre-existing methylation patterns onto the new strand during DNA replication. DNMT3A and DNMT3B are *de novo* methyltransferases that are active on unmethylated DNA and are responsible for establishing methylation patterns during early development (Klose and Bird, 2006). The importance of DNMT3A and DNMT3B enzymes during development is evident by the fact that knock-out mice lacking either of them show different defects and die at different developmental stages (Okano *et al.*, 1999). De novo methylation is active in germ cells and in the early embryo; *de novo* methylation in developing germ cells gives rise to substantially methylated genomes which then undergo a wave of genome-wide demethylation in the early embryo pre-implantation stage, followed shortly afterwards by large-scale de novo methylation in somatic cell lineages (Strachan and Read, 2004a).

Studies have suggested that de novo methylation by DNMT3 may be targeted to specific genomic regions either through recognition of specific histone modifications such as unmethylated H3K4, or through their direct recruitment by histone methyltransferases such as those responsible for writing the repressive histone mark H3K9me3 (Chen and Riggs, 2011). Further strengthening their role as transcriptional repressors, all three DNMTs have been shown to interact with histone deacetylase enzymes which are central to many gene silencing protein complexes, to establish a transcriptionally silent chromatin state (Fuks *et al.*, 2000; Robertson *et al.*, 2000; Rountree *et al.*, 2000).

It is thought that DNA methylation can repress gene transcription through two broad mechanisms. The first mechanism involves direct repression by inhibiting the binding

of sequence-specific transcription factors whose binding sites contain CpG sites such as MYC and MYCN (Tate and Bird, 1993).

The second mechanism of methylation-dependent gene repression is mediated by methyl-CpG binding proteins (MBPs) that specifically bind methylated DNA. The first MBP identified was the methyl-CpG binding protein 2 (MeCP2) (Lewis *et al.*, 1992), and this was followed by methyl-CpG binding domain proteins (MBD1, MBD2 and MBD4) which all bind methylated DNA with a higher affinity than unmethylated DNA via their conserved MBD homology domain (Hendrich and Bird, 1998). The MBPs may compete with transcription factors for their DNA binding sites but a major breakthrough in understanding methylation-mediated repression came with the finding that MeCP2 interacts with a co-repressor complex containing histone deacetylase enzymes (HDACs) (Jones *et al.*, 1998; Nan *et al.*, 1998). Further studies reported similar findings with other MBPs (Ng *et al.*, 1999; Feng and Zhang, 2001), strongly supporting their role as transcriptional repressors through the recruitment of histone deacetylation to methylated DNA in regions of transcriptional silencing (Jones and Baylin, 2002) (Figure 1.12). Additionally MBPs have been shown to participate in chromatin remodelling complexes, which are large multi-protein complexes that regulate gene expression through nucleosome restructuring (Saha *et al.*, 2006); MBPs have been identified in complexes that contain transcriptional co-repressors, such as Sin3A as well as histone deacetylases, and are recruited to methylated genomic regions where they establish a transcriptionally inactive chromatin state (Wade *et al.*, 1999; Zhang *et al.*, 1999; Feng and Zhang, 2001).

For many years the causal role of DNA methylation in initiating gene silencing versus maintaining a silent state induced by another mechanism has been the topic of intense debate. Early studies investigating de novo methylation during embryonic development support DNA methylation as a secondary event that affects genes that have already been silenced by other mechanisms, effectively stabilising their inactivation (Bird, 2002). Recent studies that have been carried out using a novel Nucleosome occupancy methylome sequencing (NOMe-Seq) approach have strongly supported this notion that promoter DNA methylation occurs secondary to other inactivating events and acts to effectively stabilise or "lock" the transcriptionally inactive state (Han *et al.*, 2011; Kelly *et al.*, 2012).

NOMe-Seq is a single molecule, high resolution nucleosome positioning assay based on enzyme accessibility to GpC sites that also retains endogenous DNA methylation information from CpG sites. This novel approach allowed the investigators to study both DNA methylation and nucleosome positioning on the same individual DNA molecule (Han *et al.*, 2011; Kelly *et al.*, 2012). Using NOMe-Seq to look at individual DNA strands they found that promoter DNA methylation occurred secondary to the DNA becoming inaccessible by nucleosome occupancy, with DNA methylation requiring the nucleosomes to be present (Han *et al.*, 2011; Kelly *et al.*, 2012). Of further interest the investigators also reported that the chromatin architecture of non-CpG island promoters was almost identical to that of CpG island promoters and reported gene silencing directly associated with methylated non-CpG island gene promoters (Han *et al.*, 2011; Kelly *et al.*, 2012). Their findings suggest that epigenetic mechanisms of gene regulation are shared between CpG island promoters and non-CpG island promoters. These findings will have important implications in the study of cancer epigenetics, and studies of aberrant DNA methylation in cancer should include non-CpG island promoter methylation as this may also contribute to tumourigenesis in a similar manner to established CpG island promoter hypermethylation (Han *et al.*, 2011) (see section 1.6.4).

While DNA methylation interacts with other epigenetic events, including histone modifications and nucleosome positioning (relationships are summarised in Table 1.7), experimental evidence suggests that DNA methylation is the dominant event and acts to seal in transcriptional repression (Jones and Baylin, 2002). This evidence comes from the use of HDAC inhibitors such as trichostatin that alone do not reactivate aberrantly methylated and silenced genes in tumour cells. Reactivation of many aberrantly methylated and silenced genes in tumour cells can only be achieved by using agents that block DNMT activity, such as 5-aza-2'-deoxycytidine, either alone or in combination with HDAC inhibitors (Cameron *et al.*, 1999).
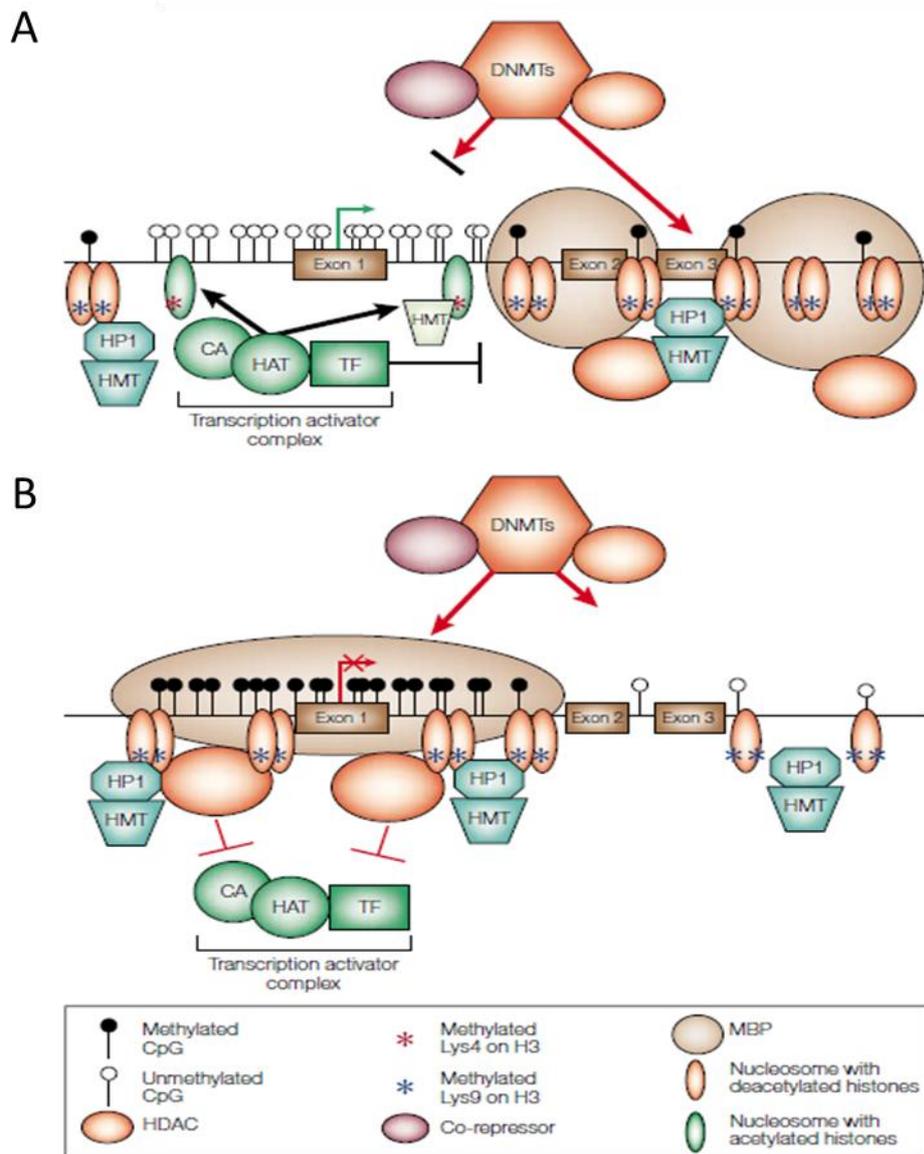
Figure 1.12. A CpG-rich promoter in transcriptionally active and transcriptionally repressed states. A. A typical actively transcribed promoter that contains a CpG island; CpG island is unmethylated; nucleosomes are widely spaced and contain the active acetylated histone marks, mediated by histone acetyltransferase (HAT) and the active H3K4me3 mark, mediated by lysine 4 methyltransferase (HMT). The open chromatin structure allows entry of the transcription activator complex which consists of a co-activator protein (CA), transcription factor and HAT. DNMTs are excluded from the promoter and directed toward flanking repressed chromatin, associated with deacetylated histones that contain H3K9 methylation marks, dense DNA methylation, methyl binding protein (MBP)- and histone deacetylase (HDAC)-complexes and tightly packed nucleosomes. **B.** The transcriptionally repressed promoter is associated with dense CpG island methylation. MBPs bind to the DNA and recruit complexes containing HMT and HDAC complexes, which deacetylate the nucleosomes alongside methylation of H3K9 to generate a tightly compacted chromatin structure that is inaccessible to the transcriptional machinery. Figure taken from Jones and Baylin, 2002.

	Active chromatin	**Repressive chromatin**
Structure	Less condensed, open, accessible; wide and irregular spaced nucleosomes	Condensed, closed, inaccessible; densely packed nucleosomes
Activity	Active, expressed	Silent, repressed
DNA methylation	Hypomethylation	Hypermethylation
Histone acetylation	Hyperacetylation of histone H3 and H4	Hypoacetylation of histone H3 and H4
Histone methylation	H3K4me2, H3K4me3	H3K27me2, H3K27me3, H3K9me2, H3K9me3

Table 1.7. Summary of DNA methylation and histone modifications associated with transcriptionally active and transcriptionally repressed chromatin states.

1.6.3 *DNA methylation in cancer*

The disruption of epigenetic mechanisms is now established as a hallmark of cancer (Wong *et al.*, 2007), and genetic and epigenetic processes are believed to co-operate at all stages of cancer development (Jones and Baylin, 2007; Baylin and Jones, 2011). It has been recognised for many years that DNA methylation patterns are widely disrupted in tumour cells; global loss of methylation across the genome is accompanied by gene-specific promoter hypermethylation (Feinberg and Tycko, 2004; Esteller, 2005). It is clear that genetic aberrations alone cannot account for all the molecular events that drive tumourigenesis, and genetic and epigenetic processes are inextricably linked in the regulation of gene transcription (Lemon and Tjian, 2000). Understanding the functional role of aberrant DNA methylation events in cancer is fundamental to our understanding of the molecular pathogenesis of cancer types. The significant technological advances in DNA methylation analysis that have been made over the years mean that entire cancer methylomes can now be interrogated at single base-pair resolution (Laird, 2010). Such advances have facilitated the molecular characterisation of many tumour types as well

as the identification of methylation markers that predict patient outcome and represent novel therapeutic targets (Laird, 2003) (see section 1.6.3.4).

1.6.3.1 *Promoter hypermethylation and transcriptional silencing*

The most widely studied alteration of DNA methylation in tumours is the silencing of tumour suppressor genes by CpG island hypermethylation (Jones and Baylin, 2002; Jones and Baylin, 2007; Esteller, 2008) and this plays a key tumourigenic role contributing to all the typical hallmarks of a cancer cell that result from tumour suppressor inactivation (Hanahan and Weinberg, 2011) (see section 1.2.4). Promoter methylation is now established, alongside genetic mutations and loss of heterozygosity (LOH), as a potential contributor of one hit, according to Knudson's two-hit hypothesis, for loss of function of a tumour suppressor gene (Jones and Laird, 1999).

It is thought that between 5-10% of all promoter CpG islands that are normally unmethylated become aberrantly methylated in cancer, affecting hundreds of genes across almost all cancer types (Jones and Baylin, 2002; Jones and Baylin, 2007). The functional significance of CpG island promoter hypermethylation is appreciated by the many genes affected that are involved in processes such as cell cycle regulation, tumour cell invasion, chromatin remodelling, transcription and apoptosis, emphasising the importance of methylation-mediated gene silencing as a major driver for cancer initiation and progression (Robertson, 2005) (Table 1.8). For many of these genes promoter hypermethylation is the only mechanism for their loss of function (Jones and Baylin, 2002).

Evidence from familial cancer syndromes suggests that aberrant promoter hypermethylation can be an early event in tumourigenesis and provide a selective advantage to tumour cells (Jones and Baylin, 2002). Germ-line mutations in *VHL* (von Hippel-Lindau), *BRCA1* (breast cancer 1, early onset) and *STK11* (serine/threonine kinase 11) genes cause familial forms of renal, breast and colon cancer, respectively; these genes are often silenced by promoter DNA methylation in sporadic cases of these cancers (Jones and Baylin, 2002). Further evidence of DNA hypermethylation being an early driver event comes from the identification of genes which are involved in key developmental pathways and are hypermethylated and silenced in different tumour types (Baylin and Ohm, 2006). In colon cancer (Suzuki *et al.*, 2004) and also in

medulloblastoma (Kongkham *et al.*, 2010a), the *SFRP* (secreted frizzled related protein) family of WNT pathway inhibitors are known to be silenced by DNA methylation. This shows how epigenetic inactivation can complement single driver mutations, such as those in the WNT pathway genes, APC and β-catenin, and aberrantly activate pathway signalling (Baylin and Jones, 2011).

Tumour suppressor genes that are frequently mutated or epigenetically inactivated in cancers often reside in genomic regions characterised by LOH (loss of heterozygosity) caused by chromosomal deletions (Jones and Baylin, 2002). The tumour suppressor genes, *RASSF1A* (located at 3p21) and *HIC1* (located at 17p13.3) are in chromosomal regions that are characterised by frequent LOH in several tumour types, and both genes are frequently silenced by promoter hypermethylation in many cancer types. Interestingly, gene mutations have not been consistently found in these regions, and *RASSF1A* and *HIC1* constitute the principal tumour suppressor genes in these regions and are defined solely by epigenetic events (Jones and Baylin, 2002).

The importance of methylation-mediated gene silencing in cancer is also recognised by the occurrence of changes that predispose to mutational events during tumour progression. This was first shown for the mismatch-repair gene MLH1 (mutL homologue 1, colon cancer, non-polyposis type 2) which is hypermethylated and silenced in sporadic tumours that have microsatellite instability (MSI) (Jones and Baylin, 2002). *MGMT* (O-6-methylguanine-DNA methyltransferase) is another DNA repair gene that is silenced by promoter hypermethylation in several different tumour types, leading to a failure to repair DNA and a predisposition to mutations in key genes such as *TP53* and *K-RAS* (Esteller *et al.*, 2000b; Esteller *et al.*, 2001). These findings show how CpG methylation can directly cause genetic changes in cancer and emphasise the interaction between genetics and epigenetics in tumourigenesis (Baylin and Jones, 2011). Mutational events have also been identified that are associated with epigenomic events in cancers, for example, mutations of the *IDH1* (isocitrate dehydrogenase 1) gene in gliomas have been found to correlate significantly with a DNA CpG island hypermethylated phenotype (Noushmehr *et al.*, 2010). The TET family proteins have recently been identified as enzyme systems that actively remove DNA methylation, forming 5-hydroxylmethyl cytosine from methylated cytosine. These proteins have been discovered to be crucially important during development and tumourigenesis

(Kriaucionis and Heintz, 2009; Sharma *et al.*, 2010). Mutations in the *TET2* gene have been identified in gliomas and leukaemias (Figueroa *et al.*, 2010; Noushmehr *et al.*, 2010). In leukaemias, *TET2* mutations were associated with increased numbers of hypermethylated genes (Figueroa *et al.*, 2010). Furthermore, recent whole-exome sequencing studies have identified frequent mutations in genes that are known to play roles in epigenome organisation (Ernst *et al.*, 2010; Jiao *et al.*, 2011; Yap *et al.*, 2011), suggesting that many abnormal epigenetic events may lie downstream of genetic events (Baylin and Jones, 2011). An emerging important role for promoter hypermethylation is the transcriptional repression of multiple microRNAs (miRNAs) (Baylin and Jones, 2011). MicroRNAs are small non-coding RNAs that target mRNA transcripts for cleavage or translational repression and can, therefore, influence the output of many protein coding genes (Bartel, 2004). Their repression can lead to pathway disruption and upregulation of their oncogenic targets (Saito *et al.*, 2006) and has been shown to produce a metastatic phenotype (Lujambio *et al.*, 2008).

Important roles for DNA methylation outside of gene promoters and CpG islands are emerging. From their genome-wide study of DNA methylation patterns in colon cancer, Irizarry et al. reported that most DNA methylation alterations in colon cancer did not occur in promoters or in CpG islands but occurred in regions up to 2kb distant from the promoter CpG island, which they termed 'CpG island shores'(Irizarry *et al.*, 2009). CpG island shore methylation was strongly associated with gene expression and with tissue of origin, consistent with the notion that epigenetic changes cause cancer largely through the disruption of differentiation processes (Irizarry *et al.*, 2009). In their findings, Irizarry et al. reported hypermethylation events enriched closer to associated CpG islands, while hypomethylation events were enriched further from the CpG island. This is consistent with the DNA methylation 'spreading theory' which describes that de novo DNA methylation may begin at flanking CpG sites and progressively invade into the core of the CpG island (Stirzaker *et al.*, 2004). The exact mechanism of how 'shore' methylation regulates gene transcription is not yet fully understood and one suggestion is that it may do so through the control of enhancer elements (Tsai and Baylin, 2011).

Gene function	Gene	Cancer type
Cell-cycle regulation	*RB1*	Retinoblastoma
	CDKN2A	Colon, lung and many more
Tumour-cell invasion	*CDH1*	Breast, gastric, thyroid, leukaemia, liver
	CDH13	Lung, ovarian, pancreatic
	TIMP3	Brain ,kidney
	VHL	Kidney
DNA repair/detoxification	*MLH1*	Colon, endometrial, gastric
	MGMT	Colon, brain, lung, breast
	BRCA1	Breast, ovarian
	GSTP1	Prostate, liver, colon, breast, kidney
Chromatin remodelling	*SMARCA3*	Colon, gastric
Cell signalling	*RASSF1A*	Lung, liver, brain (incl. medulloblastoma)
	SOCS1	Colon, liver, multiple myeloma
	SFRP	Colon, medulloblastoma
Transcription	*ESR1*	Colon, breast, lung, leukaemia
Apoptosis	*DAPK1*	Lymphoma
	CASP8	Medulloblastoma

Table 1.8. Examples of genes silenced by aberrant promoter DNA methylation in cancer. Gene function and cancers in which the methylation-dependent gene silencing is observed are listed for each gene. Table adapted from Robertson, 2005.

1.6.3.2 DNA hypomethylation in cancer

In 1983, DNA methylation was first linked to cancer when it was shown that the genomes of cancer cells were largely hypomethylated compared to their normal counterparts (Feinberg and Vogelstein, 1983a). This global DNA hypomethylation occurs primarily in satellite and pericentromeric repetitive DNA sequences, regions that are heavily methylated and associated with a heterochromatic state in normal cells (De Smet and Loriot, 2010). Genomic instability is a hallmark of cancer cells, and it is thought that global DNA hypomethylation contributes to tumourigenesis through the reactivation of transposable elements (see section 1.6.1), causing genomic instability and the disruption of gene structure (Robertson, 2005; De Smet and Loriot, 2010). Reactivation of transposon promoters following demethylation may also lead to aberrant gene regulation by interfering with transcription or by generating antisense transcripts (Robertson and Wolffe, 2000). Loss of methylation has been identified as an early event in cancer and has been associated with a severe and metastatic phenotype in many cancers (Widschwendter *et al.*, 2004).

Gene-specific DNA hypomethylation can lead to gene activation, and although not as frequently reported as tumour suppressor gene silencing by promoter hypermethylation, a small number of genes have been identified that are reactivated following loss of methylation from their promoter region. The *HRAS* oncogene has been shown to be reactivated in colorectal and small cell lung carcinomas following hypomethylation of its normally methylated CpG island promoter (Feinberg and Vogelstein, 1983b). The *MAGE* (melanoma antigen) family of cancer-testis genes (expressed normally in the testis and aberrantly in tumours) encode tumour antigens and are frequently demethylated and overexpressed in cancers (De Smet *et al.*, 1999). Other genes reported to be affected by promoter CpG demethylation accompanied by overexpression include the *S100A4* (S100 calcium binding protein A4) gene in colon cancer (Nakamura and Takenaga, 1998), *cyclin D2* (Oshimo *et al.*, 2003) and *maspin* (Akiyama *et al.*, 2003) in gastric tumours and the *HPV16* (human papillomavirus 16) gene in cervical cancer (Badal *et al.*, 2003). DNA hypomethylation can also lead to loss of imprinting (LOI) and this can drive tumourigenesis. This is best characterised by the abnormal activation of the maternally inherited allele of the *IGF2* (insulin-like growth factor II) gene which is observed in approximately 40% of colorectal cancers (Wong *et al.*, 2007).

1.6.3.3 Genome-wide analysis of DNA methylation patterns in tumours

Over the last decade, significant advances have been made in the technologies used to analyse DNA methylation. As a result of these advances, cancer epigenetic studies have progressed from assessing the methylation status of specific regions in a small number of genes to an unbiased, genome-wide interrogation of DNA methylation.

DNA methylation information is erased by standard molecular biology techniques, including cloning in bacteria and PCR. As a result, almost all sequence-specific DNA methylation analysis techniques rely on a methylation-dependent pre-treatment of the DNA, which can be carried out using one of three main approaches: endonuclease digestion, affinity enrichment and bisulfite conversion (Laird, 2010). Following methylation-dependent treatment of genomic DNA, array-based hybridisation and sequencing technologies are commonly employed to investigate methylation at CpG dinucleotides on a genome-wide scale (Laird, 2010). The development of DNA methylation microarray hybridisation techniques made a significant impact on the study of cancer epigenomics, allowing the study of DNA methylation on a genome-wide scale for a large number of samples (Laird, 2010). As a result of the advances that have been made in sequencing technologies in recent years, it is likely that array-based analyses will shift to analysis by next-generation sequencing technologies such as whole genome shotgun bisulfite sequencing (WGSBS) (Lister *et al.*, 2009), which is considered the ultimate comprehensive single base-pair resolution DNA methylation analysis technique (Laird, 2010).

The first locus-specific DNA methylation analyses used methylation-sensitive restriction endonucleases to fragment DNA in a methylation-dependent manner which could then be identified using gel electrophoresis and Southern blotting (Kaput and Sneider, 1979). Methylation-sensitive restriction digestion followed by PCR across the restriction site is a very sensitive technique that is still currently in use, however, these methods are prone to false positive results due to incomplete digestion for reasons other than DNA methylation (Laird, 2010).

Restriction landmark genome scanning (RLGS) was the first genome-wide technique for studying DNA methylation patterns (Hayashizaki *et al.*, 1993). This endonuclease-dependent method uses two-dimensional gel electrophoresis to investigate differences in

methylation which are detected as differences in the pattern of restriction fragments generated, however, the process is labour-intensive and has largely been replaced by methods that do not rely on gel electrophoresis (Laird, 2010). Comprehensive high-throughput array-based relative methylation (CHARM) analysis is a genome-wide array-based method that uses the methylation-dependent endonuclease *Mcr*BC, which provides greater sensitivity to densely methylated regions compared with a methylation-sensitive enzyme (Irizarry *et al.*, 2008). It was using this method, to assess genome-wide DNA methylation in colon cancers, that led Irizarry *et al*. to report the presence of CpG hypermethylation in 'CpG island shores' (Irizarry *et al.*, 2009) (see section 1.6.3.1).

MeDIP (methylated DNA immunoprecipitation) is an array-based analysis that utilises affinity enrichment of methylated regions by immunoprecipitation of denatured genomic DNA with an antibody specific for methylated cytosine, followed by hybridisation to an array (Weber *et al.*, 2005; Weber *et al.*, 2007). Affinity-based methods, such as MeDIP, allow for rapid and efficient genome-wide assessment of DNA methylation, however, they do not provide information on individual CpG residues and require adjustment for varying CpG density across the genome (Laird, 2010). As with other array-based methods, affinity enrichment techniques are being adapted for analysis by next generation sequencing (Down *et al.*, 2008).

Bisulfite conversion of DNA followed by sequencing is considered to be the gold-standard approach for characterising DNA methylation patterns at base-pair resolution (Clark *et al.*, 1994). The method of sequencing bisulfite-converted DNA to map 5-methylcytosine was first developed by Frommer et al. (Frommer *et al.*, 1992). Briefly, sodium bisulfite treatment converts non-methylated cytosines to uracil while methylated cytosines remain unchanged. Converted uracil residues are subsequently read as thymine following PCR amplification and sequencing and the only cytosines that remain in the sequence are methylated. Bisulfite-converted genomic DNA can be assayed both on microarrays and by next generation sequencing to investigate genome-wide DNA methylation patterns. The methylation array platforms utilise amplified bisulfite-converted genomic DNA, hybridised to microarrays containing methylation-specific oligonucleotides (MSO) to determine a quantitative measure of methylation at CpG sites (Gitan *et al.*, 2002). Illumina have developed three generations of methylation

microarrays with increasing genomic coverage; the first generation GoldenGate Cancer Panel I microarray targets 1,505 CpG sites across a targeted panel of 807 known cancer genes (Bibikova *et al.*, 2006). Subsequently the Infinium (27,000 CpG sites) (Bibikova *et al.*, 2009) and the Infinium 450K (450,000CpG sites) (Bibikova *et al.*, 2011) were released that achieve unprecedented coverage to allow a truly genome-wide characterisation of DNA methylation. Both the GoldenGate Cancer Panel I and the Infinuim 450K microarrays were used in this project and are discussed further in section 2.7.

Unlike other methylation analysis methods, whole-genome shotgun bisulfite sequencing (WGSBS) generates quantitative genome-wide methylation profiles at single nucleotide resolution (Laird, 2010). The first human methylomes published using this technique were in 2009 (Lister *et al.*, 2009) and it is likely that this gold standard approach will supersede all other techniques in the near future. While bisulfite-based methods are accurate and reproducible (Laird, 2010), the major sources of bias and error are incomplete bisulfite conversion and differential PCR efficiencies for methylated compared with unmethylated versions of the same sequence (Warnecke *et al.*, 1997). Nanopore sequencing is a novel sequencing approach that is currently in development. It offers the potential for direct sequencing of methylated cytosines without the need for prior bisulfite treatment of DNA (Branton *et al.*, 2008; Clarke *et al.*, 2009), and may be the next major advance in high-throughput DNA methylation analysis.

As the number of epigenomic studies carried out using unbiased, genome-wide approaches increases, new and interesting findings are being made about the functional significance of DNA methylation outside of gene promoters and CpG islands, which have been the focus of many conventional approaches. For example, the discovery by Irizarry *et al*. that functionally relevant hypermethylation occurs outside of gene promoters and CpG islands, at CpG sites located in regions up to 2kb distant from CpG islands, termed 'CpG island shores'(Irizarry *et al.*, 2009). Also, potential roles for intragenic DNA methylation in controlling alternative promoter usage (Maunakea *et al.*, 2010) and in alternative splicing have been reported (Shukla *et al.*, 2011; Brown *et al.*, 2012). The ability to assess DNA methylation genome-wide, at single nucleotide resolution, presents exciting new opportunities to more fully explore the role of DNA methylation in gene regulation both in normal cells and in cancer.

Another genome-wide approach used in DNA methylation analysis involves expression profiling of cells treated with DNA methyltransferase inhibitors, such as 5'-aza-2'-deoxycytidine, to discover genes that are epigenetically dysregulated by DNA methylation in cancer cells. This approach was used in this project for investigations reported in Chapter 4, where it is discussed in further detail.

1.6.3.4 The clinical and therapeutic role of DNA methylation biomarkers in cancer

Aberrant DNA methylation in cancer has long held promise as a potential biomarker strategy (Laird, 2003), and as epigenome-wide investigations gather pace, methylation patterns are being recognised that are biomarkers for tumour type, early detection, assessment of prognosis and predicted response to therapy (Baylin and Jones, 2011). Commercial assays are currently awaiting approval for the detection of hypermethylated genes in stool and blood DNA as markers for colon cancer risk and detection (Lofton-Day *et al.*, 2008; Glockner *et al.*, 2009), as well as the detection of GSTP1 (glutathione *S*-transferase PI) hypermethylation in tumour biopsy and urine samples for prostate cancer (Cairns *et al.*, 2001; Rosenbaum *et al.*, 2005).

Prognostic DNA methylation biomarkers have been identified that can identify patients with Stage I non-small-cell lung cancer who are considered high-risk and would therefore benefit from intensified adjuvant therapeutic strategies (Brock *et al.*, 2008). In patients with glioma, hypermethylation and subsequent silencing of the DNA repair gene MGMT can predict response to the chemotherapeutic agent, temozolamide (Esteller *et al.*, 2000a).

An increasing number of studies are recognising that DNA methylation patterns may reflect cellular origins; Fernandez et al. profiled 1054 cancers on the GoldenGate Caner Panel I array and reported that patterns of methylation could identify tumour type origin of cancers of unknown primary site (Fernandez *et al.*, 2012). In another study conducted across 7 different cancer types, the authors reported that differential methylation profiles reflected the developmental history and transcriptional state of the different cells of origin (Sproul *et al.*, 2012). Thus characterisation of cancer methylomes may provide clues into the precise cell of origin of the cancer type and subtypes.

The DNA methylation mark is reversible and therefore presents an exciting and attractive therapeutic target (Issa and Kantarjian, 2005).The azanucleosides, 5'-azacytidine and 5'-aza-2'-deoxycytidine (decitabine) were developed over 40 years ago as cytostatic agents (Sorm *et al.*, 1964). They have long been known to cause global demethylation in human cell line model systems (Jones and Taylor, 1980) and do so through the inhibition of all DNA methyltransferases (DNMTs) (Stresemann and Lyko, 2008). This finding led to their development as epigenetic drugs and following substantial refinements in their clinical dosing schedules to minimise drug-related toxicities, they have showed significant clinical benefits in the treatment of myelodysplastic syndrome (MDS), a proleukaemic bone marrow disorder (Silverman *et al.*, 2002; Kantarjian *et al.*, 2006). Subsequently, 5'-azacytidine (Vidaza) and decitabine (Dacogen) received approval by the FDA for the treatment of MDS and it is hoped that they will show similar therapeutic benefit in other haematological malignancies and in solid tumours.

While it is clear that the azanucleosides have cytotoxic effects and it has been shown that they cause DNA demethylation in patients (Mund *et al.*, 2005; Soriano *et al.*, 2007), their precise mode of action in inducing the clinical responses observed are still not fully understood. The rationale behind their use is in their ability to re-activate genes that have been aberrantly silenced by hypermethylation; many aberrantly methylated and silenced genes in tumour cells are reactivated following demethylation using DNMT inhibitors (Cameron *et al.*, 1999). The identification of robust associations between critical epigenetic reactivation events and patient responses will be an important area of research in azanucleoside-based therapies. While treatment with azanucleosides is effective and improves overall survival rates (Fenaux *et al.*, 2009), most patients eventually develop resistance (Tuma, 2009). Furthermore, treatment with azanucleosides results in non-specific overall DNA demethylation and multiple regulatory pathways may be affected (Komashko and Farnham, 2010). Significant efforts are therefore being made to develop more selective DNMT inhibitors (Chik and Szyf, 2011).

1.7 The role of DNA methylation in medulloblastoma

The vast majority of DNA methylation studies in medulloblastoma to-date have focused on hypermethylation of gene promoter regions and subsequent gene silencing. These studies have led to the identification of a number of genes that are epigenetically silenced by promoter hypermethylation in medulloblastoma (Table 1.9), many of them emerging from classical candidate gene approaches. Utilising DNA methylation microarray technologies, two recent independent studies have revealed the presence of discrete methylomic subgroups of medulloblastoma, and demonstrated that genome-wide DNA methylation profiles can reliably and reproducibly classify tumours into the four core molecular subgroups (SHH, WNT, Group 3 and Group 4). Alongside epigenetic classification, the value of DNA methylation biomarkers in disease prognostication is being recognised as the number of epigenome-wide studies in medulloblastoma increases.

1.7.1 Gene silencing by promoter hypermethylation in medulloblastoma

The disruption of DNA methylation at the whole-genome level was first alluded to in medulloblastoma by an early study that used restriction landmark genomic scanning (RLGS) and reported up to 1% of all CpG islands in primary medulloblastomas and 6% in medulloblastoma cell lines to be aberrantly methylated (Fruhwald *et al.*, 2001). This was followed by a series of candidate gene-specific investigations to look for evidence of methylation-dependent expression in medulloblastoma; individual genes were selected for analysis either because they were known to be inactivated by genetic mechanisms in medulloblastoma or because they were known to be epigenetically inactivated in other cancers (Lindsey *et al.*, 2005).

Initial methods for interrogating promoter methylation status relied on bisulfite conversion of genomic DNA followed by PCR amplification using primers specific for either the methylated or unmethylated template (methylation-specific PCR). These methods have proved useful to confirm the presence of promoter hypermethylation in a number of genes in medulloblastoma, of which *HIC1* (Rood *et al.*, 2002), *RASSF1A* (Lusher *et al.*, 2002) and *CASP8* (Zuzak *et al.*, 2002) were consistently found to be hypermethylated and silenced in primary tumours, and represent the most frequent epigenetically inactivated genes in medulloblastoma (Lindsey *et al.*, 2005). While the functional significance of their silencing has not been fully investigated in

medulloblastoma, studies in other tumour types have provided insights into the tumourigenic effect of their loss and *HIC1*, *RASSF1A* and *CASP8* represent strong candidates for tumour suppressor genes in medulloblastoma pathogenesis (Lindsey *et al.*, 2005).

A genome-wide approach based on expression profiling of cells that have been treated with the DNA methyltransferase inhibitor, 5'-aza-2'-deoxycytidine (5-azaCdR), has been useful in identifying further candidate tumour suppressor genes in medulloblastoma. Coupled with 5-azaCdR-induced expression changes, methylation-specific PCR has identified aberrant promoter methylation and subsequent epigenetic inactivation of *COL1A2* (Anderton *et al.*, 2008), *SPINT2* (Kongkham *et al.*, 2008) and the *SFRP* gene family (*SFRP1*, *SFRP2*, *SFRP3*) (Kongkham *et al.*, 2010a) in primary tumours.

Epigenetic inactivation of *COL1A2* has previously been reported in other cancers (Sengupta *et al.*, 2003; Chiba *et al.*, 2005). *COL1A2* encodes collagen type 1 which is a major component of the extracellular matrix, the disruption of which is a frequent event in promoting tumour growth and progression (Kalluri and Zeisberg, 2006). Anderton et al. identified CpG island promoter hypermethylation associated with *COL1A2* silencing in the majority (77%) of primary tumours. Interestingly, they identified a subset of tumours that were unmethylated at the *COL1A2* promoter; these tumours represented infant desmoplastic/nodular cases and suggest a disease subgroup-specific role for *COL1A2* epigenetic regulation in medulloblastoma (Anderton *et al.*, 2008).

SPINT2 promoter methylation was detected in ~35% of primary medulloblastomas and was associated with transcriptional silencing (Kongkham *et al.*, 2008). *SPINT2* is a negative regulator of HGF/MET signalling, which is known to be dysregulated in medulloblastomas (Binning *et al.*, 2008). HGF/MET signalling is disrupted in many different cancer types, contributing to tumourigenesis and metastasis through its programme of invasive growth signalling (Comoglio *et al.*, 2008). Kongkham et al. showed that re-expression of *SPINT2* in medulloblastoma cell lines reduced proliferation and motility and demonstrated improved survival times *in vivo* following its re-rexpression (Kongkham *et al.*, 2008). Their findings strongly support *SPINT2* as a putative tumour suppressor gene in medulloblastoma and further implicate dysregulation of the HGF/MET signalling pathway in medulloblastoma.

The *SFRP* gene family (*SFRP1*, *SFRP2* and *SFRP3*) are inhibitors of WNT signalling and have been identified as putative tumour suppressor genes silenced by DNA methylation (Kongkham *et al.*, 2010a). These findings suggest that epigenetic dysregulation of the WNT pathway may complement driver mutations in key pathway genes, such as β-catenin, to aberrantly activate WNT signalling. Additional evidence for epigenetic dysregulation of key developmental pathways in medulloblastoma comes from the *PTCH1* gene which is a negative regulator of the SHH pathway. Based on 5-azaCdR-induced expression changes and assessment of CpG island promoter methylation using a denaturation of methylation differences (DAMD) assay (genome-wide assay based on increased melting temperature of methylated DNA), epigenetic silencing of *PTCH1* by promoter DNA methylation was demonstrated in medulloblastoma (Diede *et al.*, 2010).

Epigenetic regulation of the *S100* gene family has also been investigated in medulloblastoma. The S100 family of calcium binding proteins are multifunctional proteins involved in a variety of cellular processes including cell growth and cell cycle regulation (Lindsey *et al.*, 2007). By assessing 5-azaCdR-induced expression changes, Lindsey et al. screened 16 *S100* genes for evidence of epigenetic regulation in medulloblastoma. Two members, *S100A6* and *S100A10*, demonstrated hypermethylation of their CpG island promoter associated with their transcriptional silencing in medulloblastoma. For one member, *S100A4*, a tumour-specific hypomethylated pattern was observed which was associated with elevated expression. *S100A4* does not contain a promoter-associated CpG island, and epigenetic regulation of *S100A4* in medulloblastoma was associated with hypomethylation of 2 critical intragenic CpG sites located within the 1st intron. *S100A4* hypomethylation has been detected in other cancers (Nakamura and Takenaga, 1998; Rosty *et al.*, 2002), and high levels of *S100A4* expression have been associated with metastatic progression (Hernan *et al.*, 2003). Alongside tumour suppressor gene silencing by promoter hypermethylation, aberrant hypomethylation events leading to oncogene activation may play a significant role in medulloblastoma pathogenesis.

Gene	Gene function	References
CDKN2A	Cell cycle regulator	(Fruhwald *et al.*, 2001)
RASSF1A	Interacts with DNA repair protein XPA, inhibits accumulation of cyclin D1 and induces cell cycle arrest	(Lusher *et al.*, 2002)
CASP8	Central role in execution of apoptosis	(Zuzak *et al.*, 2002)
HIC1	Growth regulator	(Rood *et al.*, 2002)
MCJ (DNAJD1)	Hypermethylated in many types of cancer, loss of expression correlates with increased resistance to antineoplastic drugs	(Lindsey *et al.*, 2006)
ZIC2	Transcriptional repressor	(Pfister *et al.*, 2007)
S100 family (S100A6, S100A10)	Encodes multifunctional calcium binding protein family, involved in variety of cellular processes including cell growth and cell cycle regulation	(Lindsey *et al.*, 2007)
SPINT2	Inhibitor of HGF/MET signalling	(Kongkham *et al.*, 2008)
COL1A2	Encodes pro-alpha2 chain of type I collagen	(Anderton *et al.*, 2008)
SFRP family (SFRP 1,2,3)	Inhibitors of WNT pathway	(Kongkham *et al.*, 2010a)
KLF4	Transcription factor, key regulator in embryonic development	(Nakahara *et al.*, 2010)
PTCH1	Negative regulator of SHH pathway	(Diede *et al.*, 2010)

Table 1.9. Examples of genes silenced by promoter hypermethylation in medulloblastoma. Gene function and reference to the studies in which they were first identified are listed for each gene. Table adapted from Batora et al., 2013.

1.7.2 Genome-wide characterisation of the medulloblastoma methylome

1.7.2.1 Medulloblastoma comprises four discrete methylomic subgroups

The current gold standard assay for establishing medulloblastoma subgroup status is expression profiling by microarray (Taylor *et al.*, 2012) (see section 1.5.8).

Recent studies have demonstrated independently that DNA methylation assayed on DNA methylation microarrays can accurately and reliably subgroup medulloblastomas into the four core molecular subgroups (SHH, WNT, Group 3, Group 4) (Schwalbe *et al.*, 2013) (Hovestadt *et al.*, 2013).

In the first study, a large series of 230 medulloblastomas were profiled on the Illumina GoldenGate Cancer Panel I array (see section 1.6.3.3) (Schwalbe *et al.*, 2013). The cohort was demographically representative of the disease and all histopathological variants were included. Using a non-negative matrix factorization (NMF) (Brunet *et al.*, 2004) consensus clustering approach, Schwalbe et al. demonstrated the presence of four robust DNA methylation subgroups; findings were cross-validated in a training set comprised of fresh-frozen tumour extracts, and a test set comprised of DNA extracted from formalin-fixed paraffin embedded (FFPE) tissue (Schwalbe *et al.*, 2013). Overall 216/230 (94%) tumours could be reliably and reproducibly assigned to a subgroup. The DNA methylation subgroups showed significant associations with established expression subgroup markers, demonstrating that the DNA methylation subgroups and gene expression subgroups were closely related (Schwalbe *et al.*, 2013). The DNA methylation subgroups were also significantly associated with distinct clinical, pathological and genomic features, consistent with those observed for their transcriptomic counterparts (Schwalbe *et al.*, 2013).

In the second study, Hovestadt et al. generated genome-wide methylation profiles of a large series of 276 medulloblastomas, comprising both frozen and FFPE samples, on the Illumina Infinium 450K array (see section 1.6.3.3) (Hovestadt *et al.*, 2013). Similar to the study by Schwalbe et al., this study showed that DNA methylation profiling allows robust molecular subgrouping of medulloblastomas from both frozen and FFPE samples; four methylomic subgroups were identified that correlated strongly with previously defined transcriptomic subgroups of medulloblastoma (Hovestadt *et al.*, 2013).

It is firmly recognised that the four distinct molecular subgroups of medulloblastoma will require different therapeutic approaches (Kool *et al.*, 2012; Northcott *et al.*, 2012a). Accurate tumour classification in the clinic will therefore be paramount to implement future risk stratification models based on molecular features, and for the design and implementation of prospective clinical trials assessing personalised and molecularly targeted therapies. The DNA methylation profiling studies described here demonstrate the clinical utility of DNA methylation to accurately subclassify medulloblastomas (Hovestadt *et al.*, 2013; Schwalbe *et al.*, 2013). A DNA-based platform has clear advantages over an RNA based platform due to the superior stability of DNA, and also performs well on either frozen or FFPE tumour material and small amounts of starting material (Hovestadt *et al.*, 2013; Schwalbe *et al.*, 2013). An additional advantage of the Infinium 450K array is the ability to generate genome-wide copy-number profiles using the intensity measures of the methylation probes, facilitating a more refined molecular classification in the clinical setting (Hovestadt *et al.*, 2013).

1.7.2.2 DNA methylation biomarkers in medulloblastoma

It is anticipated that, with increasing DNA methylation profiling of well-defined clinical trial cohorts of medulloblastomas, an increasing number of methylation biomarkers will be identified that will significantly improve and refine disease risk stratification both between and within the molecular subgroups.

In their DNA methylation profiling study, Schwalbe et al. identified two novel independent high-risk methylation biomarkers in medulloblastoma (Schwalbe *et al.*, 2013). Survival analysis was carried out on a cohort of 191 patients that were included in the PNET3 clinical trial (see section 1.5.5). Owing to the very good prognosis of the WNT subgroup of tumours (see section 1.5.8.2), the prognostic potential of methylation events was tested across the non-WNT tumours. *MXI1* and *IL8* promoter methylation were both significantly associated with a worse prognosis in the non-WNT tumours, and the addition of these two methylation markers to current stratification models based on known clinical and molecular risk variables significantly improved survival prediction of favourable-risk and high-risk medulloblastomas (Schwalbe *et al.*, 2013). This improved risk stratification model incorporating methylation biomarkers offers the potential to more accurately identify favourable-risk patients in the clinic who may

benefit from de-escalated treatment regions, while targeting intensified regimens to those with high-risk disease.

Promoter methylation of the *hTERT* (human telomerase) gene has been identified as a clinically useful biomarker for leptomeningeal metastasis detection in cerebrospinal fluid (Bougel *et al.*, 2013). Telomerase, the enzyme that maintains telomeres, is active in 90% of advanced cancers but not in normal tissues, and promoter hypermethylation associated with increased *hTERT* expression in cancers is well described (Bougel *et al.*, 2013; Castelo-Branco *et al.*, 2013). Approximately 30% of children show evidence of metastatic spread of medulloblastoma at diagnosis (Bailey *et al.*, 1995), with spread occurring mainly via the cerebrospinal fluid and leading to a high incidence of spinal and of diffuse leptomeningeal seeding (Sure *et al.*, 1995). Early detection of metastastic spread is critical to accurate therapeutic stratification (see section 1.5.4). *hTERT* promoter hypermethylation has been detected in medulloblastomas (Castelo-Branco *et al.*, 2013). Its detection in the CSF provides an attractive method, as an adjunct to routine cytology, for the non-invasive diagnosis of leptomeningeal metastasis at an early stage of diagnosis (Bougel *et al.*, 2013). Furthermore, *hTERT* promoter hypermethylation has been shown to correlate with high-grade paediatric brain tumours and with tumour progression and poor prognosis (Castelo-Branco *et al.*, 2013).

As well as the identification of prognostic methylation markers, genome-wide DNA methylation profiling of large series of primary medulloblastomas holds significant promise for the discovery of novel epigenetically regulated genes that may play key roles in driving tumour initiation and progression. The identification of critical functional DNA methylation events could help further understanding of tumour pathogenesis as well as identify potential epigenetic therapeutic targets. Methylomic subclassification of medulloblastoma could enable a refined discovery of critical methylation events that are specific to the development of distinct subgroups of the disease.

1.8 Summary and aims

Medulloblastoma is the most common malignant brain tumour of childhood and accounts for approximately 10% of all childhood cancer deaths. While 5-year survival rates of ~80% have been achieved using current clinical risk stratification models, there remains a proportion of high-risk cases for which prognosis is very poor. Furthermore, survivors of medulloblastoma suffer considerable long-term adverse effects associated with their treatment. Significant advances have been made in understanding medulloblastoma genomics, and have led to the recognition of medulloblastoma as a disease that comprises four core transcriptomic subgroups (SHH, WNT, Group 3 and Group 4). The subgroups of medulloblastoma are associated with distinct clinical behaviours and genetic features and will, therefore, require different therapeutic strategies in the clinic.

While molecular subclassification is beginning to inform treatment of medulloblastoma, more insights need to be gained from the underlying molecular pathology of the different disease subgroups, to identify critical driver events that could potentially be targeted therapeutically. Group 3 and Group 4 medulloblastomas are less well-defined molecularly compared with the SHH and WNT subgroups. Group 3 and Group 4 tumours have not been found to be associated with specific signalling pathways and whole exome and whole genome sequencing studies have not identified high frequency driver mutational events. Group 3 tumours carry a dismal prognosis and Group 4, the largest subgroup, comprises subsets of tumours with poor clinical outcome. There is an urgent need to understand the molecular defects that are driving these particularly aggressive phenotypes and to develop better therapies to improve survival.

The role of the epigenome in medulloblastoma is increasingly being recognised, and the diversity of genetic/epigenetic interactions is beginning to be more fully understood. Recent whole exome and whole genome sequencing studies have estimated that at least one third of all medulloblastomas are affected by mutations in genes that regulate chromatin structure and function. Such events have the potential to widely alter the transcriptome of medulloblastomas. Furthermore, dysregulation of chromatin modifiers was found to be subgroup-specific and was a significant feature of Group 3 and Group 4 tumours, suggesting that epigenetic processes may be crucial in their development.

DNA methylation plays a key role in epigenetic transcriptional regulation. Tumour suppressor genes that are silenced by promoter DNA methylation in medulloblastoma have been known for quite some time. Many of these genes were identified using classic candidate gene approaches and did not recognise the distinct subgroups of medulloblastoma.

Significant advances have been made in the technologies used to study DNA methylation, and it is now possible to interrogate this epigenetic mark at single base-pair resolution and on a genome-wide scale. Profiling primary medulloblastomas on DNA methylation microarrays has revealed the existence of four methylomic subgroups of the disease that are strongly associated with their transcriptomic counterparts (SHH, WNT, Group 3 and Group 4). The methylomic subgroups of medulloblastoma hold considerable promise for the refined discovery of DNA methylation events that are associated with gene expression, and that may play a critical role in the development of distinct disease subgroups and may potentially be targeted therapeutically.

The genome-wide role of DNA methylation in medulloblastoma subgroups has not been widely investigated. Furthermore, the majority of studies to date have focused on promoter-associated CpG island methylation. As an increasing number of cancer methylomes are being characterised, roles for DNA methylation outside of the gene promoter and outside of CpG islands are being recognised. An unbiased assessment of CpG methylation in distinct medulloblastoma subgroups must be undertaken to fully comprehend the gene-wide role of DNA methylation in medulloblastoma development.

1.8.1 Aims

This project aimed to investigate the role of DNA methylation in distinct subgroups of medulloblastoma, using two high-throughput screening approaches:

1. In the first approach, detailed in Chapter 3, the specific aims were:
   - to identify Group 3 and/or Group 4-specific DNA methylation events using methylation profiles generated for a large cohort of primary medulloblastomas on the Illumina GoldenGate Cancer Panel I methylation microarray, and
   - to investigate their association with gene expression in Group 3 and/or Group 4 tumours and in medulloblastoma cell lines treated

with the demethylating agent 5'-aza-2'-deoxycytidine (5-azaCdR).

2. In the second approach, detailed in Chapter 4, the specific aims were:

- to identify candidate epigenetically regulated genes using a functional epigenomics analysis of gene expression alterations in cell lines following 5-azaCdR treatment, combined with a genome-wide assessment of DNA methylation, and

- to investigate them further for evidence of subgroup-specific differential methylation concordant with differential gene expression and epigenetic gene regulation in primary tumours.

Chapter 2   Materials and methods

2.1 Primary tumour cohorts

2.1.1 *Primary medulloblastoma cohort 1*

Cohort 1, a representative cohort of 216 primary medulloblastomas (Table 2.1), was used for investigations reported in Chapter 3. The cohort comprised 75 (35%) medulloblastomas collected in Newcastle, UK (designated with the prefix NMB); 12 (5%) medulloblastomas kindly provided by Dr. Richard Gilbertson (Department of Developmental Neurobiology, St. Jude Children's Research Hospital, Memphis, TN, USA) (prefixed RJG) and 129 (60%) medulloblastomas from the PNET3 clinical trials cohort (prefixed PNET3). The SIOP/UKCCSG PNET3 (Primitive neuro-ectodermal tumour 3) clinical trial took place from March 1992 to January 2000 and recruited patients from across Europe aged from 3 to 16 years and defined as standard-risk. Patients were placed into one of two treatment arms (radiotherapy alone *versus* application of chemotherapy prior to radiotherapy). The study reported no statistically significant difference in overall survival for standard-risk patients treated with pre-radiation chemotherapy compared with those treated with radiotherapy alone (Taylor *et al.*, 2003b).

The histological subtypes of cohort 1, according to the current WHO classification (Louis *et al.*, 2007) (see section 1.5.1), comprised 169 (78%) classic, 21 (10%) LCA, 1 (0.5%) MBEN and 25 (12%) DN tumours. Histological variants were confirmed on review by Professor David Ellison (Chair of Pathology, St. Jude Children's Research Hospital, Memphis, TN, USA). There were 14 (6.5%) infant cases (< 3 years at diagnosis), 195 (90.5%) children (3-15 years) and 7 (3%) adults (≥16 years). The median age at diagnosis was 8.4 years. The cohort comprised 131 (61%) male and 85 (39%) female cases, with a male: female ratio of 1.5:1. M stage classification, as defined by the Chang staging system (Chang *et al.*, 1969) (see section 1.5.3) was available for most cases; M0 and M1 cases were assigned M negative status, while M2 and M3 cases were assigned M positive status. There were no M4 cases identified in the cohort. The majority of tumours were M-negative (167/207 (81%)), while 40 cases (19%) were M-positive. Patient age, sex and M stage data were kindly collated and provided by Professor Simon Bailey (Newcastle University Paediatric Brain Tumour Research Group).

DNA methylation profiles of the primary tumours in cohort 1 were assessed on the high-throughput Illumina GoldenGate Cancer panel I methylation microarray platform (see section 2.7.2). The 216 primary medulloblastomas were reliably subclassified according to their methylation profiles by Dr. Ed Schwalbe (Newcastle University Paediatric Brain Tumour Research Group). Dr. Schwalbe demonstrated by cross-validation in training and test cohorts that medulloblastoma comprises four robust DNA methylation subgroups (SHH, WNT, Group 3 and Group 4) that are strongly associated with the four principal transcriptomic subgroups of the same nomenclature and which have distinct molecular, clinical and pathological disease features (Schwalbe *et al.*, 2013)**.** The proportion of tumours classified in each methylomic subgroup was comparable to those described for their transcriptomic counterparts (Kool *et al.*, 2012). Group 4 tumours formed the largest group (44% [94 tumours]), followed by SHH (23% [50 tumours]) and Group 3 (20% [44 tumours]). WNT tumours represented the smallest group (13% [28 tumours]) (Schwalbe *et al.*, 2013) (Table 2.1).

Sample ID	Age at diagnosis (years)	Sex	Histological subtype	M Stage	Methylomic subgroup
NMB112	43.0	F	CLAS	0	SHH
NMB138	3.3	M	CLAS	0	SHH
NMB141	11.7	F	CLAS	0	SHH
NMB143	2.8	M	LCA	0	SHH
NMB148	1.6	M	CLAS	0	SHH
NMB154	4.7	M	CLAS	0	SHH
NMB156	7.3	M	DN	0	SHH
NMB159	22.8	M	DN	NA	SHH
NMB181	8.4	M	DN	0	SHH
NMB200	1.3	F	DN	0	SHH
NMB202	40.0	F	DN	0	SHH
NMB253	0.4	M	DN	0	SHH
NMB33	1.5	F	DN	0	SHH
NMB64	1.5	F	MBEN	0	SHH
NMB79	3.5	F	CLAS	1	SHH
NMB81	14.2	F	DN	0	SHH
RJG112	0.7	F	DN	0	SHH
RJG116	19.0	M	DN	0	SHH
RJG126	2.6	M	DN	1	SHH
RJG127	4.8	F	DN	0	SHH
RJG142	1.0	F	DN	0	SHH
PNET30113	4.1	F	CLAS	0	SHH
PNET30124	6.2	F	LCA	0	SHH
PNET30160	8.8	F	DN	0	SHH
PNET30165	12.4	M	CLAS	0	SHH
PNET30175	5.4	F	CLAS	0	SHH
PNET30179	3.5	F	CLAS	0	SHH
PNET30185	10.3	M	CLAS	0	SHH
PNET30019	15.4	M	CLAS	0	SHH
PNET30193	7.3	F	LCA	0	SHH
PNET30195	6	F	DN	0	SHH
PNET30199	15.5	F	DN	0	SHH

Sample ID	Age at diagnosis (years)	Sex	Histological subtype	M Stage	Methylomic subgroup
PNET30201	6.6	M	LCA	0	SHH
PNET30030	14.6	M	DN	0	SHH
PNET30035	11	M	CLAS	0	SHH
PNET30038	10.8	F	LCA	0	SHH
PNET30044	9.2	M	DN	0	SHH
PNET30047	7.3	M	DN	0	SHH
PNET350011	13.4	M	CLAS	0	SHH
PNET350019	15.4	M	CLAS	0	SHH
PNET350021	14	M	CLAS	1	SHH
PNET350063	5.3	F	CLAS	0	SHH
PNET350091	4.4	M	CLAS	1	SHH
PNET350104	5.4	M	LCA	1	SHH
PNET350116	3.8	F	CLAS	0	SHH
PNET350165	14	F	CLAS	0	SHH
PNET350176	7.8	F	DN	0	SHH
PNET350218	3.4	M	DN	0	SHH
PNET350290	9.7	F	CLAS	0	SHH
PNET350075	8	F	CLAS	0	SHH
NMB131	10.3	M	CLAS	0	WNT
NMB135	11.2	F	CLAS	0	WNT
NMB139	12.7	M	CLAS	0	WNT
NMB93	10	M	CLAS	0	WNT
NMB94	9.0	F	CLAS	0	WNT
PNET30119	15.8	M	CLAS	0	WNT
PNET30131	5.4	F	LCA	0	WNT
PNET30139	7.6	F	CLAS	0	WNT
PNET30147	10.8	F	CLAS	0	WNT
PNET350129	9.9	M	CLAS	0	WNT
PNET30112	7.1	F	CLAS	0	WNT
PNET30137	13.4	M	LCA	0	WNT
PNET30152	10.2	F	CLAS	0	WNT
PNET30172	8.4	M	CLAS	0	WNT

Sample ID	Age at diagnosis (years)	Sex	Histological subtype	M Stage	Methylomic subgroup
PNET30180	9.9	M	CLAS	0	WNT
PNET30002	12.8	F	CLAS	0	WNT
PNET30202	9.7	F	CLAS	0	WNT
PNET30028	7.2	F	CLAS	1	WNT
PNET30039	8.5	F	CLAS	0	WNT
PNET350045	6.8	M	CLAS	0	WNT
PNET350056	10.8	M	CLAS	0	WNT
PNET350075	8	F	CLAS	0	WNT
PNET350080	10.3	M	CLAS	0	WNT
PNET350090	10.3	F	CLAS	0	WNT
PNET350170	10.9	F	CLAS	1	WNT
PNET30051	7.5	F	CLAS	0	WNT
PNET30052	9.4	M	CLAS	0	WNT
PNET30009	6.3	F	CLAS	0	WNT
NMB136	10.5	M	CLAS	1	Group 3
NMB147	19.6	M	CLAS	0	Group 3
NMB153	3.3	M	CLAS	NA	Group 3
NMB157	5.4	M	DN	1	Group 3
NMB164	7.1	F	LCA	0	Group 3
NMB168	9.8	F	LCA	0	Group 3
NMB169	8.9	M	LCA	0	Group 3
NMB17	4.0	M	CLAS	0	Group 3
NMB171	7.7	M	CLAS	NA	Group 3
NMB188	8.6	F	CLAS	NA	Group 3
NMB227	4.0	M	CLAS	1	Group 3
NMB60	5.0	M	CLAS	0	Group 3
NMB65	9.3	M	CLAS	0	Group 3
NMB89	4.6	F	CLAS	0	Group 3
RJG113	2.5	M	LCA	1	Group 3
RJG121	4.1	M	LCA	0	Group 3
RJG122	2.5	M	LCA	0	Group 3
RJG131	5.7	F	DN	1	Group 3

Sample ID	Age at diagnosis (years)	Sex	Histological subtype	M Stage	Methylomic subgroup
RJG135	3.6	F	CLAS	0	Group 3
PNET30105	8.3	M	CLAS	0	Group 3
PNET30106	6.4	M	CLAS	0	Group 3
PNET30116	9.5	M	CLAS	0	Group 3
PNET30012	4.7	M	CLAS	0	Group 3
PNET30120	6.3	M	CLAS	0	Group 3
PNET30129	7.9	M	LCA	1	Group 3
PNET30018	3.4	M	CLAS	0	Group 3
PNET30210	4.6	M	CLAS	0	Group 3
PNET30031	7.2	M	CLAS	0	Group 3
PNET30048	14.4	M	CLAS	0	Group 3
PNET350010	5.9	M	CLAS	0	Group 3
PNET350034	14.3	M	CLAS	0	Group 3
PNET350049	13.1	F	LCA	1	Group 3
PNET350106	7.9	F	CLAS	1	Group 3
PNET350124	7.1	M	CLAS	1	Group 3
PNET350137	6	F	CLAS	0	Group 3
PNET350150	5.1	M	CLAS	1	Group 3
PNET350166	10.6	M	CLAS	0	Group 3
PNET350169	3.1	M	LCA	1	Group 3
PNET350172	5.4	M	CLAS	0	Group 3
PNET350174	12.5	M	CLAS	1	Group 3
PNET350184	11.8	M	CLAS	0	Group 3
PNET350248	3.9	M	CLAS	0	Group 3
PNET350256	4.2	M	CLAS	1	Group 3
PNET350259	9.1	M	CLAS	1	Group 3
NMB109	8.0	M	CLAS	0	Group 4
NMB110	10.0	F	CLAS	0	Group 4
NMB111	10.0	M	CLAS	0	Group 4
NMB125	6.0	F	CLAS	NA	Group 4
NMB126	8.4	F	CLAS	NA	Group 4
NMB129	5.7	M	LCA	NA	Group 4

Sample ID	Age at diagnosis (years)	Sex	Histological subtype	M Stage	Methylomic subgroup
NMB134	4.9	M	CLAS	1	Group 4
NMB140	9.4	M	CLAS	0	Group 4
NMB142	12.1	M	CLAS	1	Group 4
NMB144	5.1	M	CLAS	1	Group 4
NMB149	6.3	M	LCA	NA	Group 4
NMB151	9.8	M	CLAS	1	Group 4
NMB152	0.5	M	CLAS	0	Group 4
NMB155	19.8	M	CLAS	NA	Group 4
NMB16	0.1	M	CLAS	0	Group 4
NMB162	5.1	M	CLAS	0	Group 4
NMB165	12.7	M	CLAS	0	Group 4
NMB166	9.7	F	CLAS	0	Group 4
NMB167	6.5	F	CLAS	0	Group 4
NMB180	10.7	M	CLAS	0	Group 4
NMB184	8.6	M	CLAS	0	Group 4
NMB185	14.0	M	CLAS	1	Group 4
NMB186	8.0	M	CLAS	0	Group 4
NMB187	3.7	M	CLAS	0	Group 4
NMB189	8.6	M	CLAS	0	Group 4
NMB190	11.7	M	CLAS	1	Group 4
NMB199	9.8	M	CLAS	1	Group 4
NMB203	6.2	M	CLAS	0	Group 4
NMB252	13.1	F	CLAS	0	Group 4
NMB254	1.5	F	DN	0	Group 4
NMB43	9.0	M	CLAS	0	Group 4
NMB45	12.6	M	CLAS	0	Group 4
NMB51	6.8	M	CLAS	0	Group 4
NMB52	8.6	F	CLAS	0	Group 4
NMB76	7.5	M	CLAS	0	Group 4
NMB77	8.5	F	CLAS	0	Group 4
NMB78	5.5	M	CLAS	1	Group 4
NMB80	10.2	F	CLAS	0	Group 4

Sample ID	Age at diagnosis (years)	Sex	Histological subtype	M Stage	Methylomic subgroup
NMB82	5.4	M	CLAS	0	Group 4
NMB88	17.0	F	CLAS	0	Group 4
PNET30146	9.1	F	CLAS	0	Group 4
RJG124	6.3	M	DN	1	Group 4
RJG141	9.8	F	CLAS	0	Group 4
PNET30107	5.9	F	CLAS	0	Group 4
PNET30121	13.2	M	CLAS	0	Group 4
PNET30126	8.4	M	CLAS	0	Group 4
PNET30013	9.9	M	CLAS	0	Group 4
PNET30132	5.1	F	CLAS	0	Group 4
PNET30134	8.4	M	LCA	0	Group 4
PNET30145	5.0	M	CLAS	0	Group 4
PNET30015	12.9	M	CLAS	0	Group 4
PNET30161	8.1	M	CLAS	0	Group 4
PNET30164	8.3	F	CLAS	0	Group 4
PNET30166	8.6	M	CLAS	0	Group 4
PNET30186	11.6	F	CLAS	0	Group 4
PNET30191	7.8	M	LCA	0	Group 4
PNET30033	8.8	F	CLAS	0	Group 4
PNET350015	4.6	F	CLAS	1	Group 4
PNET350035	9.9	F	CLAS	0	Group 4
PNET350040	15.6	M	CLAS	0	Group 4
PNET350044	8.1	M	CLAS	0	Group 4
PNET350057	14.0	M	CLAS	0	Group 4
PNET350058	11.6	M	CLAS	1	Group 4
PNET350068	6.3	F	CLAS	1	Group 4
PNET350088	15.2	F	CLAS	0	Group 4
PNET350099	13.0	F	CLAS	0	Group 4
PNET350132	14.1	M	CLAS	0	Group 4
PNET350133	9.5	F	CLAS	0	Group 4
PNET350136	12.1	M	CLAS	0	Group 4
PNET350142	9.4	F	CLAS	1	Group 4

Sample ID	Age at diagnosis (years)	Sex	Histological subtype	M Stage	Methylomic subgroup
PNET350147	5.3	M	CLAS	0	Group 4
PNET350161	6.8	F	CLAS	0	Group 4
PNET350163	6.8	F	CLAS	0	Group 4
PNET350167	6.9	F	CLAS	0	Group 4
PNET350189	14.3	M	CLAS	0	Group 4
PNET350193	14.1	F	CLAS	0	Group 4
PNET350198	9.5	M	CLAS	1	Group 4
PNET350208	9.7	F	CLAS	1	Group 4
PNET350209	9.1	M	CLAS	0	Group 4
PNET350212	10.8	M	CLAS	1	Group 4
PNET350217	9.5	M	CLAS	1	Group 4
PNET350224	14.9	M	CLAS	0	Group 4
PNET350241	13.1	M	CLAS	1	Group 4
PNET350244	14.1	M	CLAS	0	Group 4
PNET350253	9.6	M	CLAS	0	Group 4
PNET350254	5.6	F	CLAS	0	Group 4
PNET350284	11.2	M	CLAS	0	Group 4
PNET350292	3.3	M	CLAS	1	Group 4
PNET30054	6.2	M	CLAS	0	Group 4
PNET30062	6.6	M	CLAS	0	Group 4
PNET30065	5.7	F	CLAS	0	Group 4
PNET30066	5.9	M	CLAS	0	Group 4
PNET30072	7.6	F	CLAS	0	Group 4
PNET30083	4.5	F	CLAS	0	Group4

Table 2.1. Primary medulloblastoma cohort 1 clinical data. Patient ID, age in years at diagnosis and sex are shown. Histological subtype is coded as follows: CLAS- classic; LCA- large cell/anaplastic; DN- desmoplastic/nodular; MBEN- MB with extensive nodularity. M stage is coded as follows: (0)– M-negative status (Chang stage M0/M1); (1) – M- positive status (Chang stage M2/M3). Tumours were subclassified into 4 principal methylomic subgroups (SHH, WNT, Group 3, Group 4) according to their Illumina GoldenGate methylation profiles.

2.1.2   *Primary medulloblastoma cohort 2*

Primary medulloblastoma cohort 2, detailed in Table 2.2, was used for investigations reported in Chapter 4. It was a representative cohort comprising 109 primary medulloblastomas. With the exception of one PNET3 case (PNET30200), the cohort comprised NMB tumours collected in Newcastle, UK. The histological subtypes of cohort 2, according to the current WHO classification (Louis *et al.*, 2007) (see section 1.5.1), comprised 76 (70%) classic, 17 (15.5%) LCA, 3 (3%) MBEN and 9 (8%) DN tumours. There were 4 (3.5%) medulloblastoma tumours whose histology was ambiguous and they could not be classified according to the WHO criteria. These tumours were classified MB-NOS (medulloblastoma- not otherwise specified). The histological variant for 88 tumours (81%) was centrally reviewed and confirmed by a panel of 3 expert neuropathologists. The remaining cases (19%) had their histology determined locally at their treatment centre. There were 19 (17.5%) infant cases (< 3 years at diagnosis), 82 (75%) children (3-15 years) and 8 (7.5%) adults (≥16 years). The median age at diagnosis was 6 years. The cohort comprised 67(61.5%) male and 42 (38.5%) female cases, with a male: female ratio of 1.6:1. M stage classification, as defined by the Chang staging system (Chang *et al.*, 1969) (see section 1.5.3) was available for most cases; M-negative status was assigned to 69/104 tumours (66%) , while 35 cases (34%) were M-positive. Patient age, sex and M stage data were again kindly collated and provided by Professor Simon Bailey (Newcastle University Paediatric Brain Tumour Research Group).

DNA methylation profiles of the primary tumours in cohort 2 were assessed on the genome-wide Illumina 450K methylation microarray platform (see section 2.7.3). Using the same methodologies previously described for the GoldenGate methylation microarray (Schwalbe *et al.*, 2013)**,** the 109 primary tumours were reliably subclassified into the 4 principal molecular subgroups (SHH, WNT, Group 3, Group 4) according to their 450K methylation profiles by Dr. Ed Schwalbe (Newcastle University Paediatric Brain Tumour Research Group). The proportion of tumours classified in each methylomic subgroup was comparable to those described for their transcriptomic counterparts (Kool *et al.*, 2012). Group 4 formed the largest group (34% [37 tumours]), followed by SHH (27% [29 tumours]) and Group 3 (27% [30 tumours]). WNT tumours represented the smallest group (12% [13 tumours]) (Table 2.2).

Sample ID	Age at diagnosis (years)	Sex	Histological subtype	M Stage	Methylomic subgroup
NMB138	3.3	M	CLAS	0	SHH
NMB168	9.8	F	LCA	0	SHH
NMB254	1.4	F	DN	0	SHH
NMB324	0.7	M	CLAS	NA	SHH
NMB460	3.9	M	DN	0	SHH
NMB465	1.5	F	MBEN	1	SHH
NMB471	0.9	F	MBEN	1	SHH
NMB483	1	M	MBEN	1	SHH
NMB555	0.3	F	MB-NOS	1	SHH
NMB612	2.5	F	DN	0	SHH
NMB625	0.3	F	CLAS	0	SHH
NMB670	1	M	CLAS	1	SHH
NMB674	2.6	F	DN	0	SHH
NMB690	2.6	M	DN	1	SHH
NMB108	35	F	CLAS	1	SHH
NMB112	43	F	CLAS	0	SHH
NMB147	19.6	M	CLAS	0	SHH
NMB156	7.3	M	DN	0	SHH
NMB18	20.3	M	DN	NA	SHH
NMB347	9.7	M	CLA	0	SHH
NMB370	6.4	F	CLA	0	SHH
NMB408	17.6	M	LCA	1	SHH
NMB435	14.8	M	DN	1	SHH
NMB437	14.2	F	CLA	1	SHH
NMB548	14.5	M	LCA	0	SHH
NMB549	5.6	F	LCA	0	SHH
NMB645	12.3	M	MB-NOS	0	SHH
NMB719	7.4	M	LCA	1	SHH
NMB738	9.82	F	CLA	1	SHH
NMB131	10.4	M	CLAS	0	WNT
NMB139	12.7	M	CLAS	0	WNT
NMB191	14.4	M	CLAS	0	WNT

Sample ID	Age at diagnosis (years)	Sex	Histological subtype	M Stage	Methylomic subgroup
NMB264	5.1	F	CLAS	1	WNT
NMB380	7.8	M	CLAS	1	WNT
NMB389	16.6	F	CLAS	0	WNT
NMB390	16.6	M	CLAS	0	WNT
NMB409	8.8	F	LCA	0	WNT
NMB417	8.6	M	CLAS	0	WNT
NMB436	4.9	M	CLAS	0	WNT
NMB458	16	F	LCA	0	WNT
NMB649	6	F	CLAS	0	WNT
NMB94	9.4	F	CLAS	0	WNT
NMB127	13.6	M	LCA	0	Group 3
NMB169	8.9	M	LCA	0	Group 3
NMB17	4.7	M	CLAS	0	Group 3
NMB176	4.3	F	CLAS	0	Group 3
NMB184	8.5	M	CLAS	0	Group 3
NMB188	8.6	F	CLAS	NA	Group 3
NMB277	15.1	F	CLAS	1	Group 3
NMB329	3.3	M	CLAS	1	Group 3
NMB330	4.7	M	CLAS	0	Group 3
NMB335	4.2	M	LCA	0	Group 3
NMB344	5	M	CLAS	1	Group 3
NMB374	0.6	F	CLAS	0	Group 3
NMB381	4.6	M	CLAS	1	Group 3
NMB384	4.5	M	MB-NOS	0	Group 3
NMB411	15.8	M	CLAS	1	Group 3
NMB440	5.6	M	LCA	1	Group 3
NMB459	4.4	M	CLAS	1	Group 3
NMB484	1.2	M	CLAS	0	Group 3
NMB490	1.9	F	LCA	1	Group 3
NMB519	4.2	M	LCA	0	Group 3
NMB535	2	M	MB-NOS	1	Group 3
NMB536	8.4	M	CLAS	0	Group 3

Sample ID	Age at diagnosis (years)	Sex	Histological subtype	M Stage	Methylomic subgroup
NMB60	5	M	CLAS	0	Group 3
NMB610	8.8	M	CLAS	0	Group 3
NMB633	4.5	F	CLAS	0	Group 3
NMB666	2.2	M	CLAS	1	Group 3
NMB684	2.6	F	CLAS	1	Group 3
NMB762	0.8	M	CLAS	1	Group 3
NMB783	3.4	F	CLAS	NA	Group 3
PNET30200	5.2	M	LCA	0	Group 3
NMB109	8	M	CLAS	0	Group 4
NMB110	10	F	CLAS	0	Group 4
NMB125	6	F	CLAS	0	Group 4
NMB126	8.4	F	CLAS	NA	Group 4
NMB132	9.3	M	CLAS	0	Group 4
NMB134	4.9	M	CLAS	1	Group 4
NMB142	12.1	M	CLAS	1	Group 4
NMB151	9.8	M	CLAS	1	Group 4
NMB166	9.7	F	CLAS	0	Group 4
NMB173	9.1	M	CLAS	0	Group 4
NMB178	5	M	CLAS	1	Group 4
NMB189	11.4	M	CLAS	0	Group 4
NMB203	6.2	M	CLAS	0	Group 4
NMB250	4.7	M	CLAS	0	Group 4
NMB260	11.3	F	CLAS	0	Group 4
NMB262	13.3	F	CLAS	0	Group 4
NMB266	7.9	M	LCA	0	Group 4
NMB283	7.5	M	LCA	0	Group 4
NMB360	11	M	CLAS	1	Group 4
NMB365	4.6	M	CLAS	0	Group 4
NMB368	11.8	M	CLAS	0	Group 4
NMB373	5.2	M	CLAS	1	Group 4
NMB401	5.7	M	CLAS	1	Group 4
NMB403	5.9	M	CLAS	1	Group 4

Sample ID	Age at diagnosis (years)	Sex	Histological subtype	M Stage	Methylomic subgroup
NMB410	15.6	M	CLAS	0	Group 4
NMB416	4.6	F	CLAS	0	Group 4
NMB438	5	M	CLAS	0	Group 4
NMB445	7.8	M	DN	0	Group 4
NMB457	4.7	F	CLAS	0	Group 4
NMB529	8	F	LCA	0	Group 4
NMB531	2.7	F	CLAS	0	Group 4
NMB632	4.8	M	CLAS	0	Group 4
NMB713	4.9	F	CLAS	0	Group 4
NMB715	4.6	F	CLAS	0	Group 4
NMB716	3	F	CLAS	0	Group 4
NMB77	8.5	F	CLAS	1	Group 4
NMB82	5.4	M	CLAS	0	Group 4

Table 2.2. Primary medulloblastoma cohort 2 clinical data. Patient ID, age in years at diagnosis and sex are shown. Histological subtype is coded as follows: CLAS- classic; LCA- large cell/anaplastic; DN-desmoplastic/nodular; MBEN- MB with extensive nodularity; MB-NOS-MB-not otherwise specified. M stage is coded as follows: (0)– M-negative status (Chang stage M0/M1); (1) – M- positive status (Chang stage M2/M3). Tumours were subclassified into 4 principal methylomic subgroups (SHH, WNT, Group 3, Group 4) according to their Illumina 450K methylation profiles.

2.2 Medulloblastoma cell lines

The origins and karyotypes of the medulloblastoma cell lines used in this project are provided in Table 2.3.

Cell line	Reference	Cell line derived from	Karyotypes	Probable molecular subgroup
Med1 (MHH-Med-1)	(Pietsch *et al.*, 1994)	Male, 10 years; establishd from cells recovered from CSF of patient with relapsed cerebellar PNET	Near diploid stem line karyotypes with clonal structural abnormalities. Partial trisomy of 1q; 46, XY; +del(1)(p22), add(4)(p160, -9, t(10;13)(q24-26;q14)	Group3/SHH
Med8A	No reference	Male; no other data	Incomplete karyotype only. Near triploid chromosome number with multiple structural abnormalities; i17q, dmin	Group 3
DAOY	(Jacobsen *et al.*, 1985)	Male, 4 years; established from a biopsy of a solid mass tumour removed from the posterior fossa. Diagnosed as desmoplastic medulloblastoma.	89 to 196 chromosomes. 95% of cells were tetraploid, remainder near octoploid. Del(1)(q32) with inv(1)(q23q32); del(1)(p13); del(3)(p21); t(1q5q); del(6)(q15q25); del(7)(q32); t(1;7)(q32;q36), t(5p8q), t(3;9)(p21;q34), i(13q); 15p+; del(20)(q12q13); i(21q). Absence of Y (90%), monosomy of 22, 2 copies of 4, trisomy of 10, duplicate X, small number of dmin	Group 3/SHH
D283 (D283 Med)	(Friedman *et al.*, 1985)	Male, 6 years; derived from the peritoneal implant and malignant ascitic fluid of patient with metastatic medulloblastoma.	47 chromosomes, including 3 marker chromosomes, 8q+, 17p+, 20q+	Group 3
D341	(Friedman *et al.*, 1988)	3.5 years; derived from large cerebellar midline mass at craniectomy,interpreted histologically as medullblastoma. Patient subsequently developed metastatic lesions (4 months post diagnosis)	47, XY; +del(1)(p13), +8, i(17)(q10), -22, dmin	Group 3
D384 (D384 Med)	(He *et al.*, 1991)	Male, 18 months; taken from the posterior fossa and massive leptomeningeal metastases. Large cell, anaplastic tumour.	45, X; -Y, -8, +der(8)t(1;8)(p11;q24), dmin	Group 3
D556	(Aldosari *et al.*, 2002b)	Female, 7 years; Large cell, anaplastic tumour.	46, XX; +1, dic(1;13)(p11.1;q34), i(17)(q10), dmin	Group 3
D425 (D425 Med)	(He *et al.*, 1991)	Male, 6 years; taken from posterior fossa	46, XY; del(10)(q22), i(17)(q10), dmin, cells contained 76-78 chromosomes	Group 3
D458 (D458 Med)	(He *et al.*, 1991)	Male, 6 years; obtained from tumour cells that had spread to the CSF of the D425 Med patient, 6 months postdiagnosis, after radiation and chemotherapy had failed.	46, XY; del(3)(p11), del(6)(q21q25), del(10)(q22), i(17)(q10), dmin, cells contained 76-78 chromosomes	Group 3
UW228-2	(Keles *et al.*, 1995)	Female, 9 years; midline posterior fossa mass involving cerebellar vermis. Histological examination revealed hallmarks of medulloblastoma.	59, XX; +1,+7, +9, +15, +17, +19, +20, +21, +4mar	Group 3/SHH
UW228-3	(Keles *et al.*, 1995)	Female, 9 years; midline posterior fossa mass involving cerebellar vermis. Histological examination revealed hallmarks of medulloblastoma.	58, XX; +1, +7, +9, +15, +19, +20, +21, +4mar	Not profiled on 450K methylation array

Table 2.3. Medulloblastoma cell line origins and karyotypes. Abbreviations: CSF, cerebrospinal fluid; del, deletion; add, additional; -, chromosome absent; I, isochromosome; t, translocation; amp, amplification; mar, marker chromosome; dmin, double minute. MHH-Med-1 and MEB-Med8A were kind gifts from Professor T. Pietsch, University of Bonn Medical Centre, Germany. D458 Med, D384 Med and D556 Med were kindly provided by Dr. D. Bigner, Duke University, USA. UW228-2 and UW228-3 were kind gifts from Dr. J. R. Silber, University of Washington, USA. DAOY, D283 Med, D341 Med were obtained from the American Type Culture Collection. Medulloblastoma cell line origins and karyotypes adapted from (Langdon *et al.*, 2006). DNA methylation profiles for 10 cell lines were assessed on the genome-wide Illumina 450K methylation microarray. Probable molecular subgroup based on DNA methylation patterns is shown. For 7 cell lines their DNA methylation patterns clustered with medulloblastoma Group 3 primary tumours. DNA methylation patterns for Med1, DAOY and UW228-2 clustered between Group 3 and SHH tumour subgroups (Ed Schwalbe, Newcastle Paediatric Brain Tumour Group; unpublished data).

2.3 Non-neoplastic cerebellar samples

The methylation profiles of 21 non-neoplastic cerebellar samples were assessed on the GoldenGate methylation microarray platform (see section 2.7.2) and used for investigations reported in Chapter 3. The cohort comprised 4 foetal, 12 infant (newborn–25 months) and 4 adult (43 years-68 years) samples, with the age of 1 patient sample unknown. The methylation profiles of 17 non-neoplastic cerebellar samples were assessed on the 450K methylation microarray platform (see section 2.7.3) and used for investigations reported in Chapter 4. The cohort comprised 3 foetal, 10 infant (newborn–25 months) and 3 adult (43 years–67 years) samples, with the age of 1 sample unknown. All samples consisted of post-mortem material from patients who died of non-neoplastic conditions.

2.4 Primary medulloblastoma transcriptomic datasets

Gene expression microarray data was available for 4 independent primary medulloblastoma cohorts:

- **Kool dataset**; comprised 62 primary tumours assayed on the Affymetrix U133+2 microarray platform (Kool *et al.*, 2008).
- **Fattet dataset**; comprised 40 primary tumours assayed on the Affymetrix U133+2 microarray platform (Fattet *et al.*, 2009).
- **Cho dataset**; comprised 204 primary tumours assayed on the Affymetrix U133A platform (Cho *et al.*, 2011).
- **Northcott** dataset; comprised 103 primary tumours assayed on the Affymetrix Exon 1.0ST array (Northcott *et al.*, 2011b).

The Kool, Fattet and Northcott datasets were available from Gene Expression Omnibus (GEO) (www.ncbi.nlm.nih.gov/geo; GSE10327, GSE12992 and GSE21140, respectively) and were downloaded into the statistical computing program R (v 2.12) (R foundation) using the *GEOquery* package available from the Bioconductor website (www.bioconductor.org). The Cho dataset was kindly provided by Dr. Dan Williamson (Newcastle University Paediatric Brain Tumour Group). Raw data (.CEL files) for the Kool and Fattet datasets, both analysed on the U133+2 platform, were downloaded from GEO and normalised together (hereafter referred to as the Kool/Fattet dataset) using the

Robust Multi-array Average (RMA) normalisation approach for raw Affymetrix data (Irizarry *et al.*, 2003). Supporting annotation packages for the Affymetrix U133+2 (hgu133plus2.db) and the U133A (hgu133a.db) platforms were downloaded into R (v 2.12) from the Bioconductor website. The annotation data file for the gene-level transcript cluster identifiers in the Affymetrix exon array was downloaded from the Affymetrix website (www.affymetrix.com/). The annotation files contained functional annotation data assigned by Affymetrix and facilitated the mapping of Affymetrix probe identifiers to important information such as chromosomal location and gene identification. In each dataset probes that mapped to an Entrez Gene ID were selected for analysis and all data was $\log_2$ transformed.

2.4.1 Medulloblastoma subgroup designation using a four metagene classifier

Original transcriptomic profiling results for 3 of the available datasets reported a variable number of molecular disease subgroups; Cho *et al*. reported 6 medulloblastoma subgroups, while Kool *et al*. reported 5 and Northcott *et al*. reported 4 discrete subgroups. All studies identified a WNT-pathway activated and SHH-pathway activated subgroup with a variable number of pathway- independent groups (see section 1.5.8.1). Following an international meta-analysis of available data there is now a consensus that medulloblastoma comprises 4 principal subgroups (SHH, WNT, Group 3 and Group 4) (see section 1.5.8.2). Through the application of NMF (non-negative matrix factorisation) to training and test cohorts, Dr. Dan Williamson (Newcastle University Paediatric Brain Tumour Group) identified and validated a 4 expression metagene classifier that reliably subclassifies medulloblastomas into the 4 principal molecular subgroups. Briefly, NMF is a method for reducing the complexity in high-dimensional data and is widely used for class discovery in biological research (Brunet *et al.*, 2004). Through the identification of expression patterns, NMF describes the variation within a dataset by reducing it to a small number of relevant metagenes (aggregate patterns of gene expression) that provide a robust clustering of samples to reveal class memberships.

Dr. Williamson applied the 4 metagene classifier to the Kool/Fattet, Cho and Northcott datasets and reliably subclassified ~90% of the tumours in each dataset into the 4 subgroups (Table 2.4). These 3 transcriptomic datasets, subgrouped according to the

universal metagene classifier, were used for investigations reported in Chapter 3 and Chapter 4.

	Transcriptomic dataset		
Consensus Subgroup	Kool/Fattet	Northcott	Cho
SHH	21	31	48
WNT	14	8	16
Group3	17	20	46
Group4	42	35	59
Not classified	8	9	25

Table 2.4. Subclassification of independent transcriptomic datasets using a universal four metagene classifier. The number of tumours reliably subclassified into each consensus medulloblastoma subgroup are shown for each dataset. Approximately 90% of tumours in each dataset were reliably subclassified using the universal 4 metagene classifier.

2.5 Culture of medulloblastoma cell lines

Ten medulloblastoma cell lines (DAOY, D283 Med, D425 Med, D458 Med, D341 Med, D384 Med, D556 Med, MHH-MED-1, MEB-MED-8A, UW228-2) (see section 2.2) were cultured in the presence and absence of the demethylating agent 5'-aza-2'-deoxycytidine (5-azaCdR) and used for investigations reported in Chapter 4.

2.5.1 Reviving cell lines stored in liquid nitrogen

All work with cell line cultures was carried out in a Class II tissue culture hood using sterilised pipettes and sterile techniques. Cryotube vials (Nunc®, Thermo Scientific, Loughborough, UK) containing 1ml of frozen cells (5×10^6 cells) (see section 2.5.4) were defrosted rapidly by placing in a water bath pre-warmed to 37^0C. The thawed cells were added drop-wise and slowly to 5ml of complete growth medium (see section 2.5.2) in a 15ml centrifuge tube (VWR International Ltd, Lutterworth, UK). Complete growth media had been pre-warmed to 37^0C in a water bath. The cells and medium were gently mixed by pipetting up and down and then centrifuged at 250 x g (relative centrifugal force, RCF) for 5 minutes in a Thermo Scientific Multifuge 3S+ centrifuge (Thermo Scientific, Loughborough, UK) to remove the cryoprotectant agent dimethylsulfoxide (DMSO) (Sigma-Aldrich Ltd, Gillingham, UK) (see section 2.5.4). The supernatant was aspirated and discarded and cells were resuspended in 5mls of fresh complete growth

medium and transferred to a sterile 25cm^2 cell culture flask (Corning®, Sigma Aldrich Ltd, Gillingham, UK). Cultures were incubated at 37^{0}C in a humidified incubator with 5% CO_2 (Sanyo Ltd, Loughborough, UK).

2.5.2 Maintaining cell line cultures

In order for cells to grow and proliferate in culture, they require a source of nutrients, hormones and growth factors. To prepare complete growth media, each 500ml of appropriate medium (Table 2.5) was supplemented with 10% (v/v) fetal calf serum (FCS) (Sigma-Aldrich Ltd, Gillingham, UK), 1% (v/v) L-glutamine solution 200mM (Sigma-Aldrich Ltd, Gillingham, UK) and 1% (v/v) sodium pyruvate solution 100mM (Sigma-Aldrich Ltd, Gillingham, UK). To control bacterial contamination 1% (v/v) penicillin/streptomycin mixture (10,000I.U/10,000µg/ml) (Sigma-Aldrich Ltd, Gillingham, UK) was added to media.

Cell cultures were regularly examined by microscopy to assess growth characteristics (Table 2.5), to observe the morphology and viability of cells and to check for microbial contamination. The contamination of cell cultures by mycoplasmas is a major problem. Mycoplasmas are resistant to the antibiotics used, they are difficult to detect and contamination can mediate a wide variety of negative effects on cultures (Sung *et al.*, 2006). Actively growing cell cultures were routinely tested for mycoplasma infection every 2 months and prior to freezing stocks (see section 2.5.4) and were uniformly negative.

The cell lines have infinite replicative capacity and following an initial recovery period immediately after seeding, will grow exponentially until they reach confluency in the flask or exhaust the nutrient supply. The media in which the cells were cultured was changed every 2-3 doublings (Table 2.5) to replenish nutrients and remove waste as the cell concentration increased. For cells growing adherently (Table 2.5), media was aspirated and replaced with fresh media that had been pre-warmed to 37^{0}C in a water bath. For non-adherent cell lines (Table 2.5), the cell suspension was transferred to a centrifuge tube and centrifuged at 250 x g (RCF) for 5 minutes in a Thermo Scientific Multifuge 3S+ centrifuge (Thermo Scientific, Loughborough, UK), the supernatant was aspirated and the cell pellet was resuspended in the appropriate volume of pre-warmed media.

Cell Line	Growth characteristics	Approximate doubling time (hrs)	Media
DAOY	Adherent	24	DMEM + 10% FCS
D283 Med	40% adherent	36	DMEM + 10% FCS
D425 Med	Non-adherent	24-36	DMEM + 10% FCS
D458 Med	60-70% adherent	36+	DMEM + 10% FCS
D341 Med	Non-adherent	36-48	DMEM + 10% FCS
D384 Med	80% non-adherent	48	DMEM + 10% FCS
D556 Med	95% non-adherent	48+	DMEM + 10% FCS
MHH-MED-1	40% adherent	36-48	DMEM + 10% FCS
MHH-MED-8A	Adherent	36-48	DMEM + 10% FCS
UW228-2	Adherent	24	DMEM/Ham's F12 +10% FCS

Table 2.5. Growth characteristics, approximate doubling times and media for medulloblastoma cell line cultures. Media was either Dulbecco's Modified Eagle's Medium (DMEM) or a 1:1 mixture of DMEM and Ham's Nutrient Mixture F12 (DMEM/Ham's F12). Media was supplied by Sigma Aldrich LTD, Gillingham, UK.

Cell lines must be subcultured on a regular basis to maintain them in the exponential phase of growth and ensure viability. Confluency of cells was examined microscopically and once cells reached approximately 70-80% confluence they were subcultured to reduce cell density. Cells were subcultured usually once or twice weekly and were generally subcultured at a 1:2 to 1:4 ratio depending on the growth characteristics of the individual cell line. Adherent cells were trypsinised first to form a cell suspension for subculture. Trypsin is a proteolytic enzyme that detaches adherent cells from the surface of the culture flask. A 1x trypsin/EDTA solution was prepared from 10 x trypsin/EDTA (Sigma-Aldrich Ltd, Gillingham, UK) using sterile phosphate buffered saline (PBS) (Sigma-Aldrich Ltd, Gillingham, UK). After aspirating the media cells were washed twice with sterile PBS (pre-warmed to 37°C) to remove traces of media as trypsin is inactivated by FCS. Enough 1 x trypsin/EDTA solution (pre-warmed to 37°C) was added to the culture flask to cover the layer of cells and the flask was incubated at 37^{0}C for approximately 2 minutes. The flask was gently tapped to aid detachment and suspension of cells which was confirmed microscopically. Complete medium with serum was added to inhibit further trypsin activity and the trypsinised cells were transferred to a centrifuge tube and centrifuged at 250 x g (RCF) for 5 minutes. For non-adherent cells the cell suspension was transferred directly to a

centrifuge tube and centrifuged with no trypsinisation step. Following centrifugation the supernatant was aspirated and the cell pellet was resuspended in fresh complete media pre-warmed to 37^0C and transferred to sterile culture flasks.

2.5.3 Cell counting

Cell counts were performed prior to freezing stocks of cells and at regular intervals during subculture to monitor growth rates and to check viability. After resuspending pellets in fresh media as described above, a 1:1 mixture of cell suspension and 1 x trypan blue stain (Sigma-Aldrich Ltd, Gillingham, UK) was prepared in a 1.5ml microcentrifuge tube (Starlab UK LTD, Milton Keynes, UK). Trypan blue stain is not absorbed by healthy viable cells but enters dead or damaged cells, allowing them to be counted. 10µl of the cell/trypan blue mixture was loaded onto a dual chamber counting slide (Bio-Rad Laboratories Ltd, Hemel Hempstead, UK) and cell counts were performed using a Bio-Rad TC10 Automated cell counter (Bio-Rad). The automated counter calculated a total cell count and a live cell count.

2.5.4 Cryopreservation of cell lines

Cell cultures are suitable for long-term storage in a liquid nitrogen freezer at temperatures below -130^0C. Prior to freezing, cell pellets were produced in the same way as during subculture (see section 2.5.2) and counted as described in section 2.5.3. Frozen stocks of cells were prepared at a concentration of 5×10^6 cells / ml. Cell pellets were resuspended in an appropriate volume of chilled FCS and placed on ice for 10 minutes. Freeze medium was prepared containing 20% DMSO and 80% FCS chilled on ice. An equal volume of freeze medium was added drop-wise to the chilled cell suspension and the suspension was left on ice for 10 minutes to allow the DMSO to permeate the cells. Addition of the cryoprotectant agent DMSO prevents the formation of intracellular ice crystals and prevents cell lysis during thawing. 1ml aliquots of cell suspension were transferred to cryotube vials (Nunc®, Thermo Scientific, Loughborough, UK) and stored overnight at -80^0C in a Nalgene® Mr Frosty freezing container (Sigma-Aldrich Ltd, Gillingham, UK). The container facilitates a slow cooling rate, generally 1^0C per minute and minimises intracellular ice formation during freezing. Frozen cells were transferred to liquid nitrogen freezer tanks the following day for long term storage.

2.5.5 5'-aza-2'-deoxycytidine treatment

5'-aza-2'-deoxycytidine (5-azaCdR) is a potent DNA demethylating agent. It is an azanucleoside that is incorporated into replicating DNA, replacing cytosine residues. Once incorporated it forms a covalent bond with DNA methyltransferase enzymes, inhibiting methyltransferase function (Stresemann and Lyko, 2008). Treatment of medulloblastoma cells *in vitro* with 5-azaCdR has been found to reactivate epigenetically silenced tumour suppressor genes (see section 1.7.1).

A 500µM stock solution of 5-azaCdR was prepared by dissolving 5-azaCdR powder (Sigma-Aldrich Ltd, Gillingham, UK) in DMEM. 0.5ml aliquots of stock solution were stored in sterile universal containers (Fisher Scientific UK Ltd, Loughborough, UK) at -20^0C. Cell cultures were treated with 5µM 5-azaCdR for 72 hours. Media and 5-azaCdR were renewed every 24 hours as 5-azaCdR is unstable at 37^0C. Media replenishment was carried out as described in section 2.5.2. 5-azaCdR is incorporated into newly synthesised DNA over several cell doublings and cells were initially seeded at an appropriate density that avoided the need to subculture over the 72 hours and ensured cells were approximately 80% confluent at the end of the treatment period. Following treatment with 5-azaCdR, cells were harvested for RNA and DNA extraction as described below. At high doses 5-azaCdR is toxic to cells (Stresemann and Lyko, 2008). The optimum concentration of 5-azaCdR and the optimum time for cell line DNA demethylation and gene reactivation had previously been assessed and determined by Dr. Jennifer Anderton (Anderton J.A. PhD Thesis, Newcastle University).

2.5.6 Harvesting cells for RNA extraction

Cells need to be harvested for RNA extraction when they are still actively proliferating in the exponential phase and pellets need to be frozen rapidly to prevent degradation of RNA. Cells were harvested when they were approximately 80% confluent. Immediately prior to harvesting the culture flasks were placed on ice. For adherent cell cultures the media was discarded and the surface of cells washed with sufficient chilled sterile PBS for 1-2 minutes. The PBS was aspirated and replaced with 5mls of fresh chilled PBS. Cell scrapers (VWR International Ltd, Lutterworth, UK) were used to lift cells adhering to the flask surface. The surface was washed with chilled PBS and the cell suspension transferred to a centrifuge tube and centrifuged at 750 x g (RCF) for 5 minutes at 4^0C in a Thermo Scientific Multifuge 3SR+ centrifuge (Thermo Scientific, Loughborough,

UK). A small volume of suspension was retained to perform cell counts as described in section 2.5.3. Following centrifugation the supernatant was discarded, the pellet re-suspended in an appropriate volume of chilled PBS and 1ml aliquots containing 5×10^6 cells transferred to 1.5ml microcentrifuge tubes. The tubes were centrifuged at 450 x g (RCF) for 1 minute at 4^0C in an Eppendorf centrifuge 5415R (Eppendorf Inc., New York, USA) to remove PBS and the pellets were immediately snap frozen in dry ice and stored at -80^0C. For non-adherent cells the media was transferred to a centrifuge tube and centrifuged at 450 x g (RCF) for 5 minutes at 4^0C. The cell pellet was washed in chilled PBS and centrifuged as before. The washed pellet was re-suspended in PBS, aliquoted and frozen as described for adherent cells.

2.5.7 Harvesting cells for DNA extraction

The procedure used for harvesting cells for DNA extraction was the same as that for RNA described above in section 2.5.6, with the exception that cells were not kept on ice as DNA is less susceptible to degradation by exogenous exonucleases.

2.6 Extraction and assessment of nucleic acids

2.6.1 RNA extraction from cells

Total RNA was extracted from harvested cells (see section 2.5.6) using phenol-chloroform extraction and alcohol precipitation methods. TRIzol® Reagent (Invitrogen Ltd, Paisley, UK) is a monophasic solution of phenol and guanidine isothiocyanate that lyses cells and dissolves cellular components producing a homogenate. Addition of chloroform to the homogenate causes the separation of a colourless upper aqueous layer, an interphase layer and a lower red phenol-chloroform organic layer. RNA is precipitated from the upper aqueous phase with isopropanol. After washing with ethanol the RNA pellet was re-suspended in RNase free water. The procedure was carried out according to manufacturer's instructions and purified RNA was stored at -80^0C.

2.6.2 DNA extraction from cells

Total DNA was extracted from harvested cells (see section 2.5.7) using the DNeasy® Blood and Tissue kit (Qiagen Ltd, Manchester, UK) according to the manufacturer's instructions. The procedure uses proteinase K to lyse cells and the lysate is passed through a silica-based membrane column that selectively binds DNA. The DNA is washed to remove contaminants and eluted from the membrane in elution buffer (AE).

Purified DNA was stored at -20^0C. The DNeasy® procedure will produce DNA yields from samples as small as 100 cells up to samples as large as 5×10^6 cells. The procedure will recover DNA fragments between 100bp and 50kb in size.

DNA was extracted from frozen primary tumour samples using TRIzol® Reagent (Invitrogen Ltd, Paisley, UK) (see section 2.6.1). DNA was extracted from the interphase and phenol-chloroform organic layer with ethanol according to manufacturer's instructions. After washing, DNA was resuspended in 10mM Tris-Cl, pH8.5 and stored at -20^0C.

2.6.3   Qualitative and quantitative assessment of nucleic acids

2.6.3.1   Spectrophotometry

UV spectrophotometry can be used to determine the concentration of RNA or DNA in a sample. Both RNA and DNA absorb UV light very efficiently and have a maximum absorption at 260nm. In contrast to the nucleic acids, proteins have an absorption maximum of 280nm. The NanoDrop 1000 spectrophotometer (Thermo Scientific, Loughborough, UK) was used to measure the concentration and assess the purity of DNA and RNA samples. The machine measures the optical density (OD) of a 1µl sample at wavelengths of 260 and 280nm and calculates the concentration of nucleic acid as well as the OD_{260}/OD_{280} ratio. Pure DNA has an OD_{260}/OD_{280} ratio ~ 1.7-1.9 and pure RNA has a ratio ~2.0. Obtaining lower ratio values can be indicative of protein contamination. NanoDrop spectrophotometry measurements are prone to overestimation as the presence of DNA, RNA and phenol contaminants in a sample will also be detected. The NanoDrop spectrophotometer produces an absorbance scan with a symmetric peak at 260nm confirming high nucleic acid purity.

2.6.3.2   Fluorometry

Fluorescence-based assays for the quantitation of nucleic acids are more sensitive and selective then spectrophotometric analysis. The Qubit® dsDNA BR Assay kit (Invitrogen Ltd, Paisley, UK) along with the Qubit® 2.0 fluorometer (Invitrogen) were used to determine the concentration of DNA in a sample. The assay uses advanced molecular probe fluorophores that fluoresce upon binding to double-stranded DNA (dsDNA). In contrast to UV-absorbance analysis, the Qubit® assay selectively measures the concentration of dsDNA and is tolerant to the presence of common contaminants.

The assay provides no measurement of the purity of DNA in a sample and can accurately calculate concentrations from 100pg/µl to 1000ng/µl. The Qubit® dsDNA BR Assay was performed according to manufacturer's instructions.

2.6.3.3 Bioanalyzer

Gene expression microarrays require high quality RNA. Prior to biotin RNA labelling (see section 2.6.4) for subsequent microarray analysis (see section 2.8), a more sensitive and accurate method for determining the quality and quantity of RNA was employed. The Agilent 2100 Bioanalyzer (Agilent Technologies UK Ltd, Wokingham, UK) is a chip-based platform that utilises capillary electrophoresis to determine quality control metrics for DNA, RNA and protein. The Agilent RNA 6000 Nano Kit (Agilent Technologies UK Ltd) was used to prepare the chip and samples according to the manufacturer's instructions. The Bioanalyzer can assess a total RNA concentration from 5-500ng/µl from a 1µl sample. The Agilent 2100 software (Agilent Technologies UK Ltd) calculates a total RNA concentration, the 28S:18S rRNA ratio and the RNA Integrity number (RIN) which assesses total RNA quality in terms of extent of degradation. High quality RNA will produce a Bioanalyzer electropherogram and gel-like image output showing 2 distinct ribosomal peaks and bands, respectively (Figure 2.1(A)). High quality RNA will have a RIN greater than 9 and a 28S:18S ratio of 2 or more. For microarray analysis a RIN higher than 8 is generally required for cultured cells. Degraded RNA will produce a RIN closer to 1 and a 28S:18S ratio less than 2, with degradation products appearing between the ribosomal peaks and bands of the electropherogram and gel-like image respectively (Figure 2.1(B)).
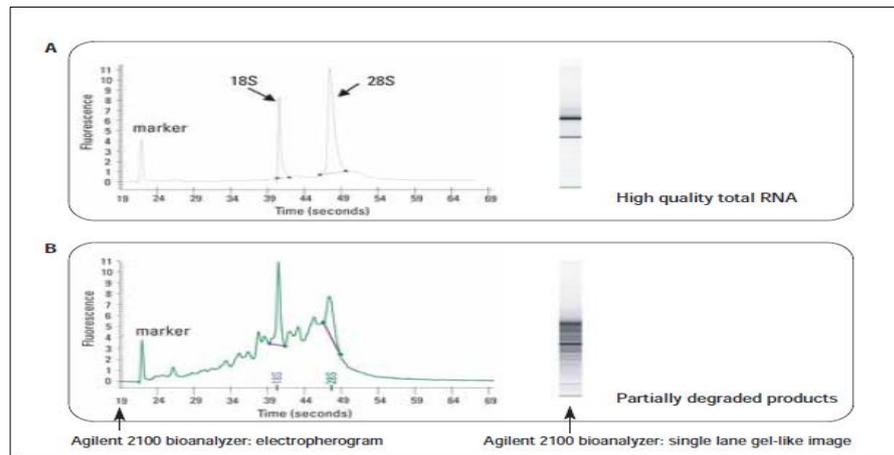
Figure 2.1. Analysis of total RNA integrity using the Agilent 2100 Bioanalyzer. (A) Upper electropherogram and gel-like image show high quality total RNA with the 18S and 28S subunit as 2 distinct peaks and bands. (B) Degradation products from partially degraded total RNA will produce peaks between the 2 ribosomal peaks in the electropherogram and between the ribosomal bands on the gel-like image. Figure taken from Agilent 2100 Bioanalyzer Application Compendium (Agilent Technologies UK Ltd, Wokingham, UK).

2.6.4   RNA amplification and biotinylation

Total RNA extracted from the 10 medulloblastoma cell lines cultured in the presence and absence of 5-azaCdR (see section 2.5.5) was amplified and biotinylated prior to hybridisation with the Illumina HumanHT-12 v4.0 BeadChip (see section 2.8.1). Direct biotin labelling of target RNA transcripts is commonly employed in microarray analysis and Cy3 fluorescent conjugates are used to generate a fluorescent signal and detect transcript levels following hybridisation to complementary probe sequences. For each cell line sample, 200ng of total RNA, determined by Bioanalyzer assessment (see section 2.6.3.3), was amplified and biotin-labelled using the Illumina® TotalPrep RNA Amplification Kit (Illumina, San Diego, CA,USA), according to the manufacturer's instructions. The procedure consisted of an initial reverse transcription step using an oligo (dT) primer and ArrayScript reverse transcriptase to generate first strand complementary DNA (cDNA). This was followed by second strand cDNA synthesis using DNA polymerase and RNase H to generate double-stranded DNA (dsDNA). Following purification the dsDNA template underwent in vitro transcription using T7 RNA polymerase and biotin-NTP mix during a 14 hour incubation step to generate multiple copies of biotinylated complementary RNA (cRNA). Purified cRNA was eluted in 200µl of nuclease-free water and the yield was assessed by NanoDrop (see section 2.6.3.1). cRNA samples were stored at -80^{0}C prior to being sent for expression microarray analysis (see section 2.8).

2.7 DNA methylation profiling

2.7.1 Introduction to high-throughput array-based DNA methylation profiling

DNA methylation microarray technologies allow for the interrogation of cytosine methylation at CpG dinucleotides on a genome–scale and at base-pair resolution. Advances in the technology in recent years have led to the development of high density arrays providing coverage of more than 450,000 CpG sites representing in excess of 20,000 genes.

The assays used in this project are based on genotyping bisulfite-converted genomic DNA (gDNA). The method of sequencing bisulfite-converted DNA to map 5-methylcytosine was first developed by Frommer et al. (Frommer *et al.*, 1992). Briefly, sodium bisulfite treatment converts non-methylated cytosines to uracil while methylated cytosines remain unchanged. Converted uracil residues are subsequently read as thymine following PCR amplification and sequencing and the only cytosines that remain in the sequence are methylated. DNA methylation profiling methodologies based on bisulfite conversion pre-treatment cannot discriminate between 5-methylcytosine and 5-hydroxymethylcytosine, which is a key intermediate in active demethylation pathways mediated by the TET family of proteins (Baylin and Jones, 2011). The methylation array platforms described below utilise amplified bisulfite-converted gDNA, hybridised to microarrays containing methylation-specific oligonucleotides (MSO) to determine a quantitative measure of methylation at CpG sites (Gitan *et al.*, 2002).

2.7.2 GoldenGate Methylation Cancer Panel I platform

The GoldenGate Methylation Cancer Panel I microarray interrogates the methylation status of 1,505 CpG sites selected from a target panel of 807 genes. The selected genes represent tumour suppressor genes, oncogenes, imprinted genes, genes involved in DNA repair, cell cycle control, differentiation and apoptosis, as well as previously reported differentially methylated genes (Bibikova *et al.*, 2006). Assay probes have been designed around the promoter and first exon of genes and are annotated according to the number of base pairs upstream or downstream from the transcription start site (TSS) they occur. Probes located upstream of the TSS are labelled with P (promoter) while those located downstream of the TSS are labelled with E (exon). Therefore, the

probe ASCL2_E76_R, is measuring methylation at the *ASCL2* gene, 76 base pairs downstream of the TSS, with the reverse strand being assayed. Probes are situated both within (n=1044) and outside (n=461) of CpG islands (CpG islands were defined by Illumina). The number of sites per gene assayed on the array ranges from one CpG site (28.6% genes) to three or more sites (14.1% genes). The assumption is made that CpG residues adjacent to the target site have the same methylation status. This assumption has been validated using bisulfite sequencing (Bibikova *et al.*, 2006; Schwalbe *et al.*, 2013).

2.7.2.1 GoldenGate methylation microarray assay

Microarray analysis was performed on the Illumina GoldenGate Cancer Panel I methylation array at the Wellcome Trust Centre for Human Genetics, Oxford, UK according to the manufacturer's protocols (Illumina, San Diego, CA,USA).

The GoldenGate methylation microarray assay is a bead array assay based on Illumina's BeadArray Technology. The assay uses 3-micron silica beads randomly assembled in microwells. Each bead is covered with hundreds of thousands of copies of a specific oligonucleotide that captures its complementary target sequence and produces a fluorescent signal, the intensity of which is read by the Bead Array reader. There are approximately 30 beads per target sequence on the array platform. A decoding algorithm is applied to determine the location and identity of each bead due to their random assembly, allowing for high density packing on a miniaturised platform (Gunderson *et al.*, 2004).

Two pairs of probes have been designed for each CpG site being interrogated on the GoldenGate methylation platform; one pair of allele-specific (ASO) and locus-specific (LSO) oligonucleotides for the methylated state and a corresponding pair for the unmethylated state. Pooled query oligos anneal to bisulfite converted gDNA with subsequent extension and ligation occurring from the ASO to the matched LSO creating PCR templates. The methylated and unmethylated templates are selectively amplified by PCR using fluorescently labelled universal PCR primers and hybridised to the bead array bearing the complementary address sequences (Figure 2.2). Ninety-six well arrays were scanned in the Bead Array reader and the intensity of fluorescent signals of both the methylated and unmethylated alleles measured.
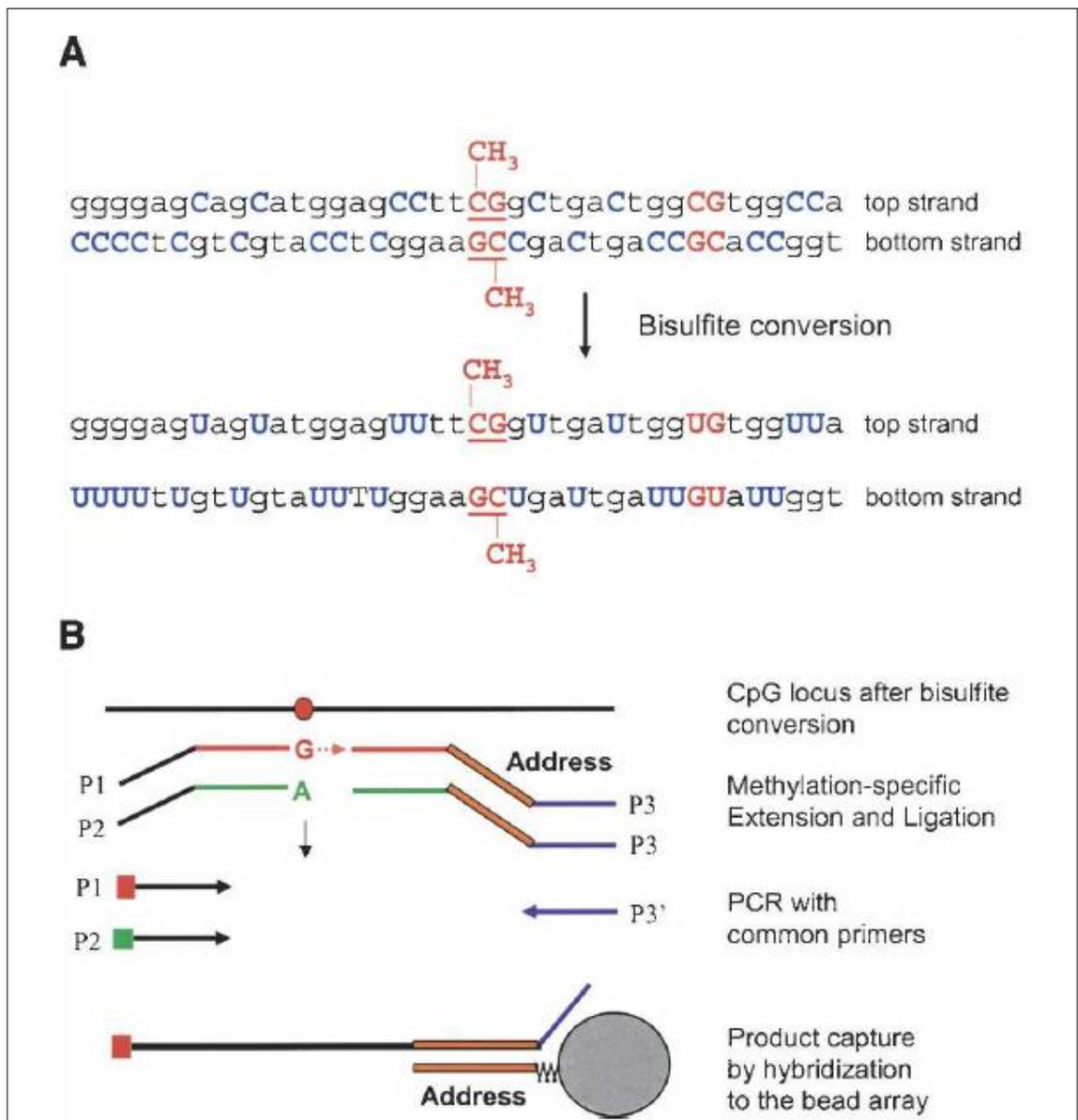
Figure 2.2. Principles of the Illumina GoldenGate methylation array technology. (A) Bisulfite treatment of gDNA converts unmethylated cytosine (C) residues to uracil (U), while methylated cytosines remain unchanged. (B) Two pairs of probes per CpG site are hybridized to target bisulfite-converted gDNA, with one probe pair specific to a methylated state and the other specific to an unmethylated state (recognised as a C/T polymorphism). Probe pairs comprise an allele-specific oligonucleotide (ASO) that discriminates the C/T polymorphism and a locus-specific oligonucleotide (LSO) that discriminates the CpG site. ASO/LSO probe pairs bind in this discriminatory manner to the converted uracil and methylated cytosine residues, with an initial ASO extension phase followed by ligation to corresponding LSO creating a template which is PCR-amplified using fluorescently labelled universal primers that are associated with the "T" (unmethylated) allele or the "C" (methylated) allele respectively. The fluorescently labelled PCR products are hybridized to the BeadArray platform based on complementary address sequences. The methylation status of each CpG site is determined by measuring the ratio of the fluorescent signal from the methylated allele to the sum of the signals of both methylated and unmethylated alleles. This is defined as the β-value and provides a continuous measure of DNA methylation levels ranging from 0 (completely unmethylated) to 1 (completely methylated). Figure taken from Bibikova et al., 2006.

BeadStudio is a general purpose program for the analysis of all Illumina array data. Raw data was exported and processed by Dr. Ed Schwalbe (Newcastle University Paediatric Brain Tumour Research Group) in BeadStudio v3.2 (Illumina, San Diego, CA, USA). The methylation status of each interrogated CpG locus was recorded as a β-value, calculated according to the following equation:

$$\beta = \frac{\max(y_{meth}, 0)}{\max(y_{unmeth}, 0) + \max(y_{meth}, 0) + 100}$$

where y_{meth} and y_{unmeth} are the intensities measured by the methylated and unmethylated probes, respectively. A constant bias of 100 is added to the denominator to regularise β when both methylated and unmethylated probe intensities are low (Bibikova *et al.*, 2006). The β-value provides a continuous measure of DNA methylation levels ranging from 0 (completely unmethylated) to 1 (completely methylated). Array quality control measures and normalisation processes were carried out by Dr. Ed Schwalbe in BeadStudio v3.2 (Schwalbe *et al.*, 2013). The GoldenGate methylation array was used for investigations reported in Chapter 3.

Initial evaluation of the GoldenGate methylation array technology by Bibikova *et al.* found it to be specific and highly sensitive, capable of detecting a change in β-value ≥ 0.2 with 95% confidence. The high reproducibility of the assay was determined using technical replicates (average $R^2 \geq 0.98$). The accuracy of the array was validated by methylation-specific PCR (MSP) which is able to sensitively determine methylation levels of individual genes (Bibikova *et al.*, 2006). Bisulfite sequencing studies have also validated the accuracy of the GoldenGate methylation array (Bibikova *et al.*, 2006; Schwalbe *et al.*, 2013).

2.7.3 Infinium Methylation 450K platform

The 450K DNA Methylation microarray is the most recent methylation array developed by Illumina and follows the success of the GoldenGate platform described above and its successor the Infinium 27K platform.

The 450K array interrogates the methylation status of approximately 475,000 CpG sites, providing coverage of a total of 21,231 UCSC RefGenes. The content of the 450K array was selected with the guidance of a consortium of DNA methylation experts and

provides comprehensive coverage across the complete gene and CpG island regions (Bibikova *et al.*, 2011). Probes have been designed covering multiple gene sub-regions as defined by Illumina; probes target promoter regions at 200bp and 1500bp blocks upstream of the transcription start site (annotated TSS200 and TSS1500, respectively) as well as the 5' and 3' UTR, first exon and gene body regions. Ninety-six per cent of CpG islands are represented on the array as well as CpG island shores (regions 0-2kb from CpG islands) and CpG island shelves (regions 2-4kb from CpG islands). Shore and shelf regions located upstream of their associated CpG island are annotated "north" (N_Shore and N_shelf), while regions located downstream of their associated island are annotated "south" (S_Shore and S_shelf). CpG islands, island shores and island shelves were defined by Illumina and they also include an "Open Sea" annotation, defining isolated CpGs in the genome. In addition to coding RNA transcripts probes have been designed that target non-coding RNAs and intergenic regions previously shown to be biologically significant or informative (Bibikova *et al.*, 2011).

2.7.3.1 *Infinium 450K methylation microarray assay*

Microarray analysis was performed on the Illumina Infinium HumanMethylation450 BeadChip at the Wellcome Trust Clinical Research Facility, Edinburgh, UK according to the manufacturer's protocols (Illumina, San Diego, CA,USA).

As for the GoldenGate platform, the 450K methylation array is a bead array assay (see section 2.7.2.1). It differs from the GoldenGate assay in that it utilises Illumina's Infinium technology, employing both Infinium I and Infinium II assays to achieve high density coverage (Figure 2.3).

The Infinium beads have long target-specific probes designed for individual CpG sites. The Infinium I methylation-specific assay uses two bead types per CpG locus, one designed against the unmethylated site and one against the methylated site. The 3'terminus of the probe is designed to match either the protected cytosine (methylated) or the thymine resulting from bisulfite conversion (unmethylated) and incorporates a fluorescent label for detection (Figure 2.3(A)). The Infinium I assay has been designed based on the assumption that CpG sites adjacent to the target locus have the same methylation status (Bibikova *et al.*, 2011).The Infinium II assay uses one bead type per locus for CpG sites located in regions of low CpG density. The underlying CpG sites

are represented by a degenerate R-base, allowing the methylation status of the target locus to be assessed independently of assumptions on the status of neighbouring sites. The 3' terminus of the probe matches the base directly upstream of the target site and a single base extension adds either a labelled G or A base depending on the methylation status of the cytosine residue in the targeted CpG site (Figure 2.3(B)).
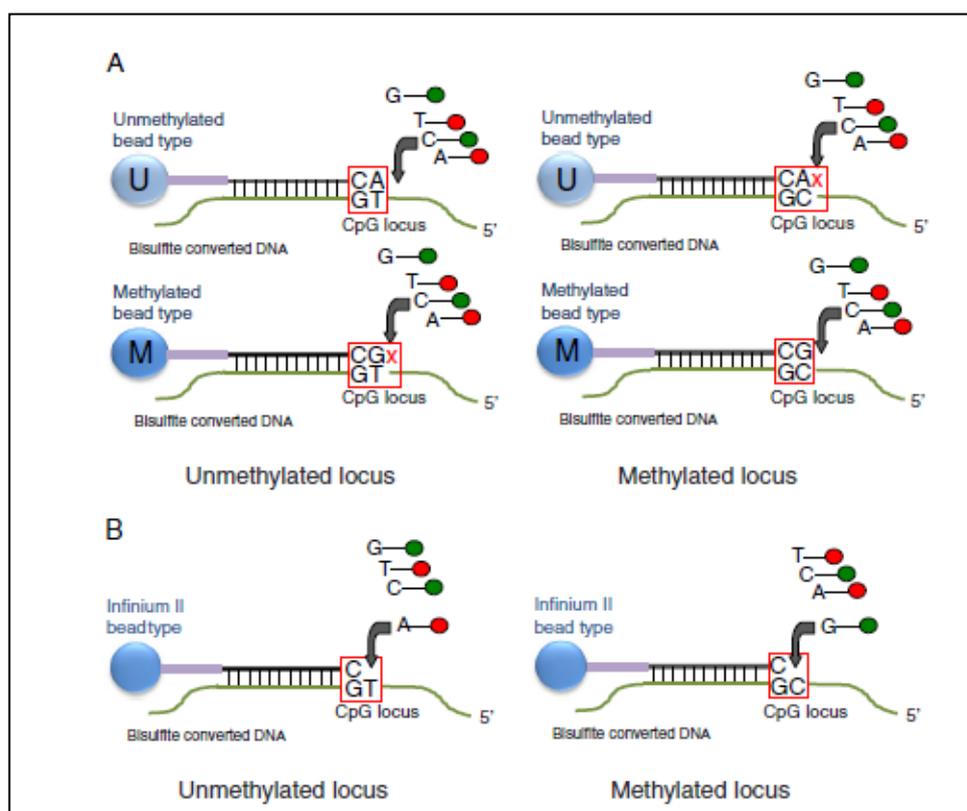
Figure 2.3. Principles of the Illumina Infinium 450K methylation array technology. **(A)** Infinium I assay: two bead types correspond to each CpG locus on the target bisulfite-converted gDNA- one to the methylated state(C) and the other to the unmethylated state (recognised as a C/T polymorphism). Following complementary recognition of the methylated or unmethylated state the probe incorporates a fluorescently labelled nucleotide. Probe design assumes the same methylation status for adjacent sites. **(B)** Infinium II assay: one bead type corresponds to each target CpG site and the probe can contain up to 3 underlying CpG sites, represented by a degenerate R base corresponding to "C" in the CpG site. Methylation state is detected by single base extension with labelled "A" base incorporated at an unmethylated site and labelled "G" at a methylated site. Infinium II assay allows target loci to be assessed independently of neighbouring CpG site status and is used in regions of low CpG density. Methylation status is measured using two methods: the β-value method provides a continuous measure of methylation ranging from 0 (completely unmethylated) to 1 (completely methylated); the M-value method calculates the log2 ratio of intensities of methylated probe *versus* unmethylated probe and also provides a continuous measure with positive M-values reflecting increasing methylation and negative M-values representing decreasing methylation. The M-value method is advantageous in microarray analysis because, unlike the β-value, it is statistically valid for many of the existing microarray statistical frameworks. Figure taken from Bibikova et al., 2011.

The performance of the 450K methylation array has been validated by whole genome bisulfite sequencing (WGBS) and has been found to be highly reproducible for technical replicates (average R^2=0.99). The array can reliably detect a change in beta value of 0.2 with 99% confidence (Bibikova *et al.*, 2011). Within the Newcastle University Paediatric Brain Tumour Group, Dr. Ed Schwalbe has carried out extensive array validation experiments, including assessments of intra- and inter-replicate reproducibility, alongside quality control measures. Independent bisulfite sequencing studies carried out within the group have validated the accuracy of the array.

There are currently two methods for measuring and reporting on the methylation levels determined by the Infinium 450K array. The first one is the β-value which is the ratio of the fluorescent signal from the methylated allele to the sum of the fluorescent signals of both methylated and unmethylated alleles and provides a continuous measure of methylation levels ranging from 0 (completely unmethylated) to 1 (completely methylated) (see section 2.7.2.1). The second method is the M-value which is calculated as the log2 ratio of the intensities of methylated probe versus unmethylated probe as in the following equation:

$$M = \log2 \left[\frac{\max(y_{meth}, 0) + 1}{\max(y_{unmeth}, 0) + 1} \right]$$

A constant bias of 1 is added to the intensity values to prevent large changes in M-value that may occur due to small intensity estimation errors. M-values close to zero indicate a hemi-methylated state, while positive M-values reflect increasing methylation and negative values reflect decreasing methylation. The M-value is widely used in mRNA expression microarray analysis and, therefore, the benefits of using this method in methylation analysis are that many of the existing microarray statistical frameworks can be applied to methylation data analysis (Du *et al.*, 2010).

Twelve samples were processed in parallel on Infinium HumanMethylation450 BeadChips. GenomeStudio is the successor to the BeadStudio software for the analysis of Illumina array data. Raw data was exported and processed by Dr. Ed Schwalbe (Newcastle University Paediatric Brain Tumour Research Group) in GenomeStudio

(Illumina, San Diego, CA, USA). Array quality control measures and normalisation processes were carried out by Dr. Ed Schwalbe in GenomeStudio. The 450K methylation array was used for investigations reported in Chapter 4.

2.8 Gene expression microarray analysis

2.8.1 HumanHT-12 v4.0 Expression BeadChip

Expression alterations in medulloblastoma cell lines following treatment with the demethylating agent 5-azaCdR (see section 2.5.5) were assessed by gene expression microarray analysis using the Illumina HumanHT-12 v4.0 BeadChip and used for investigations reported in Chapter 4. The HumanHT-12 Expression BeadChip (Illumina, San Diego, CA, USA) provides whole genome gene expression profiling. The chip contains a total of 47,231 probes covering 31,335 genes. Probes were designed to cover NCBI RefSeq content, as well as supplemental UniGene content. The platform utilises Illumina's bead array direct hybridisation assay. Similar to the methylation microarray platforms described above, the expression BeadChip comprises arrays of randomly positioned 3-micron diameter silica beads. A specific 50-mer oligonucleotide probe sequence is assigned to each bead type and with approximately 700,000 copies of the same probe sequence attached to each bead and an average of 30 beads per bead type this high level of redundancy supports high quality reproducible data. Every bead on the array is identified by a 29-mer address sequence using a series of decoding hybridisations (Gunderson *et al.*, 2004).

2.8.2 Expression microarray assay

Microarray analysis was performed on the Illumina HumanHT-12 BeadChip at the Wellcome Trust Clinical Research Facility, Edinburgh, UK according to the manufacturer's protocols (Illumina, San Diego, CA, USA). Amplified biotinylated RNA (cRNA) (see section 2.6.4) was sent for array analysis at a concentration of 150ng/µl as measured by NanoDrop (see section 2.6.3.1). Prior to microarray analysis, the quality and quantity of cRNA were evaluated at the Wellcome Trust Facility using the Agilent 2100 Bioanalyzer (Agilent Technologies UK Ltd, Wokingham, UK). Cell line pairs (5-azaCdR treated and untreated) were processed on the same BeadChip (12 samples/chip) and all samples were processed in a single batch using 2 chips to minimise batch effect. Reproducibility of the assay had previously been demonstrated by Illumina with high

concordance between technical replicates. The assay can detect a statistically significant fold change ≤ 1.35, depending on the number of replicates included, and this is detectable over a dynamic range ≥ 3logs, making it highly sensitive and precise (Illumina, 2011).

Arrays were scanned in the Bead Array reader and the intensity of fluorescent signal for target and control probes measured. Quality assessment and processing of the raw bead level data was carried out by the Wellcome Trust Facility in Edinburgh using the Illumina GenomeStudio Gene Expression Module (Illumina, San Diego, CA, USA). There were 7 control categories built into the assay controlling for sample, labelling, hybridisation and signal generation. The performance of the controls was evaluated by the Wellcome Trust Facility alongside the assessment of systematic differences between arrays and spatial artefacts to identify problematic or outlier arrays.

The Wellcome Trust Facility exported the GenomeStudio output for all target and control probes and provided the non-normalised, non-transformed bead summary data in text files. Bead summary data was created using Illumina's default method of removing outliers greater than 3 median absolute deviations (MADs) from the median and calculating the mean and variance of the remaining beads for each probe type. Alongside bead summary probe profiles, summarised gene profiles were also provided which were calculated by averaging the intensity values from all the probe types in a gene probe set. For the investigations reported in Chapter 4 the individual summarised probe intensities were used. The detection *p*-value was included in the summarised probe profiles characterising the chance that the target probe sequence signal was distinguishable from the negative controls. Bead summary data was analysed using the R package *beadarray* (Dunning *et al.*, 2007) as described in Chapter 4, section 4.3.3.2.

2.9 Reverse transcriptase polymerase chain reaction

2.9.1 Introduction to PCR and reverse trancriptase PCR

Polymerase chain reaction (PCR) is a powerful DNA amplification technique that is fundamental to many molecular biology laboratory protocols. The modern PCR technique was initially developed by Dr. Kary Mullis in 1983 (Bartlett and Stirling, 2003). The technique is based on amplification of specific target DNA sequences using single-stranded oligonucleotides (primers) which are complementary to opposite strands of the target DNA. As PCR progresses under specific thermal cycling conditions, the multiple copies of double-stranded DNA generated are used as template for further replication in a chain reaction in which the DNA template is exponentially amplified. PCR-based assays are used for a variety of biological applications, including DNA cloning for sequencing and the quantification of expression of target genes

There are 3 amplification phases during a PCR reaction; an initial exponential phase doubles the amount of PCR product at every cycle assuming the reaction is 100% efficient, a linear phase when the reaction is slowing due to consumption of reagents and a final plateau or end-point phase when the reaction has stopped due to exhaustion of reagents. Endpoint or traditional PCR (see section 2.9.3) uses agarose gel to detect PCR amplification products from the end point of the reaction and results are semi-quantitative and based on nucleotide size discrimination (see section 2.9.3.1). Real-Time quantitative PCR (qPCR) (see section 2.9.4) detects amplification during the entirety of the reaction as it progresses and measures product accumulation in the exponential phase, making it a more sensitive and precise method for quantifying DNA and gene expression.

Reverse transcriptase PCR (RT-PCR) amplifies complementary DNA (cDNA) that has been reverse transcribed from messenger RNA (mRNA), permitting the determination of target gene expression levels. One-step RT-PCR includes the reverse transcriptase step in the same tube as the PCR reaction. For all RT-PCR experiments reported in this project a two-step RT-PCR method was employed which involves an initial reverse transcriptase step to synthesise cDNA (see section 2.9.2) followed by PCR of cDNA amplicons in a separate reaction.

2.9.2   cDNA synthesis

Prior to performing PCR, cell line RNA was reverse transcribed using the iScript cDNA synthesis kit (Bio-Rad Laboratories Ltd, Hemel Hempstead, UK), according to the manufacturer's instructions. The reverse transcriptase enzyme RNase H+ in the iScript kit can synthesise cDNA from RNA input ranging from 100fg to 1μg of total RNA. A blend of oligo dT and random hexamer primers which anneal to the poly(A) tail of mRNA and random complementary sites on mRNA respectively, optimise cDNA synthesis. For all RT-PCR experiments described in this project, 1μg of cell line total RNA was initially reverse transcribed using the iScript kit and the reaction was incubated in a GeneAmp$^®$ 9700 Thermocycler (Applied Biosystems, Foster City, USA) as per protocol.

2.9.3   RT-PCR to assess re-expression of RASSF1A in medulloblastoma cell lines following 5-azaCdR treatment

Expression of the *RASSF1A* transcript in medulloblastoma cell lines before and after 5-azaCdR treatment (see section 2.5.5) was examined by endpoint RT-PCR for investigations reported in Chapter 4. *RASSF1A* was used as a positive control for the demethylating treatment of cell lines due to the robust evidence for its methylation-mediated transcriptional silencing and reactivation following 5-azaCdR treatment in multiple medulloblastoma cell lines (Lusher *et al.*, 2002).

The PCR process involves multiple thermal cycles, with each cycle comprising 3 successive reactions of DNA denaturation at 93-95^{0}C, primer annealing at temperatures usually ranging from 50-70^{0}C and DNA synthesis/extension typically at 72^{0}C. Correct primer design is essential for selective target amplification. The desired characteristics of PCR primers include a length of 18-25 base pairs, a GC content of 40-60%, a similar melting temperature between primer pairs, and no stretches of sequence complementary within or between primers at the 3' end. The selectivity of target amplification is dependent on using an optimised annealing temperature which is typically 3-5 ^{0}C below the melting temperature of the primers. The DNA synthesis/extension step employs a thermostable taq DNA polymerase (Applied Biosystems, Foster City, USA), which is isolated and purified from the bacterium *Thermophilus aquaticus,* and incorporates deoxynucleoside triphosphates (dNTPs) into a newly synthesised DNA strand complementary to the template in the 5' to 3' direction.

For this project, a previously published *RASSF1A* primer pair was used (Dammann *et al.*, 2000). Primer sequences for *RASSF1A* were as follows:

- 5' CAGATTGCAAGTTCACCTGCCACTA 3' (forward)
- 5' GATGAAGCCTGTGTAAAGAACCGTCCT 3' (reverse)

RASSF1A PCR product size was 242bp. The β-actin housekeeping gene was used to control for any differences in starting cDNA quantites between samples. β-actin is a commonly used endogenous control gene in PCR applications as it is constitutively expressed in all cells. A previously published β-actin primer pair was used (Horikoshi *et al.*, 1992). Primer sequences for β-actin were as follows:

- 5'GAGCGGGAAATCGTGCGTGACATT 3' (forward)
- 5' GATGGAGTTGAAGGTAGTTTCGTG 3' (reverse)

β-actin PCR product size was 232bp. Primers were custom synthesised by Sigma Aldrich (Sigma Aldrich Ltd, Gillingham, UK) and supplied as lyophilised pellets. Primers were reconstituted in nuclease-free water (Qiagen Ltd, Manchester, UK) at a concentration of 100μM, and stored at -20^0C alongside 10μM working stock aliquots.

20μl PCR reactions were set up using the reagents listed in Table 2.6. Reagents were kept on ice and added to 0.2ml PCR tubes (Starlab UK LTD, Milton Keynes, UK) in the order listed in Table 2.6. Reactions were set up in a LabCaire PCR clean air cabinet (Scientific Laboratory Supplies Ltd, Nottingham, UK), and cDNA template was added last at the bench. Cell line RNA was initially reverse transcribed to cDNA as described in section 2.9.2, and 3μl of cDNA template were added to the PCR reagents.

Reagent	Volume (µl) for 20µl reaction	Final concentration
Nuclease-free water	8.8	-
10X PCR Buffer II	2	1X PCR Buffer II
Forward primer (10µM)	2	1µM
Reverse primer (10µM)	2	1µM
Magnesium chloride	1.2	1.5mM
dNTPs(5mM)	0.8	0.2mM
AmpliTaq Gold (5U/µl)	0.2	0.05U/µl

Table 2.6. PCR reaction reagents and concentrations for amplification of *RASSF1A* and β-actin cDNA. Reagents were obtained from Applied Biosystems (Applied Biosystems Foster City, USA). Nuclease-free water was obtained from Qiagen (Qiagen Ltd, Manchester, UK). Magnesium chloride is added to the reaction becauseTaq polymerase is a magnesium-dependent enzyme.

PCR reactions were performed in GeneAmp® 9700 Thermocyclers (Applied Biosystems, Foster City, USA), according to the thermal cycling conditions listed in Table 2.7.

Stage	Step	Process	Temperature (^{0}C)	Time	No. of cycles
Stage 1	Step1	Incubation	95^{0}C	10 minutes	1 cycle
Stage 2	Step 1	Denaturation	94^{0}C	50 seconds	
	Step 2	Primer annealing	60^{0}C (*RASSF1A*) 56^{0}C (*β-actin*)	50 seconds	35cycles(*RASSF1A*) 25cycles(*β-actin*)
	Step 3	Synthesis/extension	72^{0}C	50 seconds	
Stage 3	Step 1	Extension	72^{0}C	7 minutes	1 cycle

Table 2.7. PCR thermal cycling conditions for amplification of *RASSF1A* and β-actin cDNA. An initial incubation step (Hot Start) ensures that active polymerase enzyme is generated at temperatures where the DNA is fully denatured. Optimum primer annealing temperature and cycling conditions for *RASSF1A* had previously been determined by Lusher et al. (Lusher *et al.*, 2002).

2.9.3.1 Agarose gel electrophoresis

Agarose gel electrophoresis is commonly used to separate DNA molecules on the basis of size. By applying an electric field, the negatively charged DNA molecules migrate through the agarose matrix towards the positively charged cathode. Smaller molecules migrate faster than larger molecules. The concentration of agarose influences product separation, and larger molecules are resolved better using a low concentration gel while smaller molecules separate better at high concentration gel. The *RASSF1A* and β-actin PCR amplification products were small fragment sizes (242bp and 232bp, respectively), and 2% agarose gels were used to detect them from the end point of the PCR reaction. Agarose gel results are semi-quantitative. GelRed Nucleic Acid Gel Stain (Biotium, Hayward, California, USA) was added to the agarose gel matrix in order to visualise the PCR products under UV light. GelRed is an intercalating nucleic acid stain and when exposed to UV light it fluoresces in a manner which intensifies after binding to DNA. A 100bp DNA ladder (Promega, Madison, WI, USA) was used to allow estimation of product size for the bands detected.

To prepare the gels, agarose powder (Invitrogen Ltd, Paisley, UK) was dissolved in 1 x TBE (Tris-Borate-EDTA) by heating in a microwave on medium power for 2-3 minutes. TBE is an electrophoresis buffer and when mixed with agarose it produces a gelatinous conductive medium. The agarose/TBE solution was allowed to cool slightly to ~50^0C before adding the GelRed stain; 1µl of a 10,000X stock solution of GelRed was added per 10ml agarose/TBE solution. After allowing the gel to set at room temperature in a gel cast, it was placed in a gel rig (Life Technologies Ltd, Carlsbad, CA, USA) and immersed in 1 x TBE buffer solution. PCR samples and 100bp DNA ladder were prepared for loading on the gel by diluting 5:1 in 6X Blue/Orange Loading Dye (Promega, Madison, WI, USA). Gels were run at 100 volts using a Powerpac (Bio-Rad Laboratories Ltd, Hemel Hempstead, UK) and PCR cDNA amplification products were visualised under UV light using a transilluminator (Syngene, Cambridge, UK).

2.9.4   Real-time quantitative reverse transcriptase PCR

Real-time RT-PCR (qRT-PCR) measures amplicon accumulation during the highly reproducible and precise exponential phase of the PCR reaction allowing for accurate and sensitive quantification of gene expression. qRT-PCR was used to quantify target gene expression levels in medulloblastoma cell lines for investigations reported in Chapter 3.

SYBR Green I Dye (Invitrogen Ltd, Paisley, UK) was used to detect and quantify cDNA amplification for all qRT-PCR reactions described in Chapter 3. SYBR Green dye fluoresces upon binding the minor groove of double stranded DNA (dsDNA) and the emitted fluorescence is proportional to the amount of amplified dsDNA detected in every cycle

A real-time PCR reaction is characterised by the cycle number when target amplification is first detected, known as the C_T (threshold cycle). SYBR Green fluorescence is normalised to the internal passive reference dye ROX and produces a normalised reporter signal (R_n). The magnitude of target signal is determined by subtracting the R_n produced during the early cycles of PCR and is defined as the ΔR_n. The qPCR instrument produces an amplification plot of ΔR_n *versus* cycle number and sets a threshold level of fluorescence that is within the early stage of the exponential phase. The C_T is thus defined as the cycle number at which fluorescence passes this threshold and is a relative measure of the concentration of target sequence in the PCR reaction.

qRT-PCR reactions were carried out using Platinum® SYBR® Green qPCR SuperMix-UDG with ROX (Invitrogen Ltd, Paisley, UK). This is a ready-to-use mix containing all the reagents required for qPCR, including taq polymerase, dNTPs and magnesium chloride along with SYBR Green I dye and the passive ROX reference dye. Cell line RNA was first reverse transcribed to cDNA as described in section 2.9.2. Full details describing how qRT-PCR reactions for target genes were set up are provided in Chapter 3, section 3.3.10. Reactions were set up in 384-well plates (Applied Biosystems, Foster City, USA) and run on the ABI PRISM 7900 HT Detection System (Applied Biosystems, Foster City, USA).

Unlike traditional PCR (see section 2.9.3), qPCR instruments employ a 2 step thermal cycling process. The initial denaturation step occurs at 95^0C with primer annealing and subsequent extension both occurring at 60^0C. Amplicon size is smaller in qRT-PCR reactions (~150bp) compared with traditional RT-PCR (~300bp), permitting higher reaction efficiencies in the exponential phase. Primer Express$^®$ software (Applied Biosystems, Foster City, USA) was used to design primers compatible with the thermal cycling conditions described above. Full details of primer design using Primer Express 2.0 software are provided in Chapter 3, section 3.3.10.

SYBR Green I dye binds to any double-stranded DNA and, therefore, a well optimised qRT-PCR reaction is essential to minimise non-specific binding and the generation of false positive signals. Full details on reaction optimisation for target genes, including optimisation of primer concentration and validation of primer specificity using melting curve analysis are provided in Chapter 3, section 3.3.10.

2.9.4.1 *Relative quantification using the delta-delta C_T method*

There are two methods of quantification of cDNA from qRT-PCR assays; absolute quantification interpolates unknown quantities from a standard curve produced under the same reaction conditions, while relative quantification, also known as the delta delta C_T ($\Delta\Delta C_T$) method, measures changes in gene expression relative to a reference sample (calibrator).

For all qRT-PCR experiments performed in Chapter 3, the $\Delta\Delta C_T$ method was used to calculate relative expression levels of target genes. The $\Delta\Delta C_T$ method measures the difference in threshold cycles for the target sample and the designated reference sample, which have both been normalised to the same endogenous control, and calculates the amount of target according to the following formula:

$$2^{-\Delta\Delta C_T}$$

The $\Delta\Delta C_T$ calculation is only valid if the efficiencies of the target and endogenous control amplifications are approximately equal. This requires a validation experiment to assess how the ΔC_T (C_T target- C_T endogenous control) varies with template dilution. For amplification efficiencies that are approximately equal the ΔC_T will not vary considerably over a range of concentrations. Full details of validation experiments for

each target gene and the calculation of relative target gene expression levels using the $\Delta\Delta C_T$ calculation are provided in Chapter 3, section 3.3.11.

2.10 Statistical methods

Statistical analyses were performed using the R Statistical Computing Program, versions 2.12 and 2.15 (R foundation; www.R-project.org). R packages required for particular analyses were downloaded from the Bioconductor website (www.bioconductor.org). Statistical significance was tested using Mann-Whitney U tests and *t*-tests for investigations reported in Chapter 3. For investigations reported in Chapter 4 the R package *limma* (Smyth, 2004) was used. *Limma* uses a linear models approach to the analysis of microarray experiments and uses a different method of variance estimation compared to conventional *t*-tests, computing a moderated *t*-statistic. *P*-values were used to assess significance. Where necessary raw *p*-values were corrected for multiple hypothesis testing using the Benjamini-Hochberg false discovery rate (FDR) correction (Benjamini and Hochberg, 1995) and a FDR-adjusted *p*-value < 0.05 was considered statistically significant. Pearson's product-moment correlation coefficient (*r*) was calculated to assess the direction and strength of the linear association between CpG site methylation and the level of gene expression. A correlation coefficient *r* value $< (-0.7)$ was indicative of a strong inverse relationship between methylation and gene expression. Further details of specific analyses are shown where appropriate.

Chapter 3

Global investigation into DNA methylation-dependent gene regulation in Group 3 and Group 4 medulloblastomas

3.1 Introduction

Aberrant DNA methylation has an established role in driving tumourigenesis (Jones and Laird, 1999). In particular, CpG island promoter hypermethylation of tumour suppressor genes has been recognised as the most frequent mechanism for gene inactivation in cancers (Esteller, 2002; Baylin and Ohm, 2006; Tsai and Baylin, 2011), and, although not as frequently reported, an association between hypomethylation and increased gene expression has been found for several oncogenes (Feinberg and Vogelstein, 1983a; Hanada *et al.*, 1993; Feinberg and Tycko, 2004). An investigation into DNA methylation-dependent gene regulation in medulloblastoma offers the potential for the identification of disease biomarkers and novel therapeutic targets.

There are approximately 30 genes known to be aberrantly silenced by DNA methylation in medulloblastoma (Dubuc *et al.*, 2012). Many of these, including the most frequently reported *RASSF1A*, *CASP8* and *HIC1* (Lindsey *et al.*, 2005) have been identified from candidate gene small scale approaches. More recently, genome-wide methods, based on microarray gene expression profiling of cell lines before and after treatment with a demethylating agent, have identified *COL1A2* (Anderton *et al.*, 2008), *SPINT2* (Kongkham *et al.*, 2008) and *SFRP* (Kongkham *et al.*, 2010a) as novel tumour suppressor genes in medulloblastoma.

The investigations reported in this chapter are based on results generated from a newer DNA methylation microarray technology (see section 2.7). Tumour profiling on the GoldenGate methylation platform has provided the opportunity to assess methylation patterns in medulloblastoma more widely, focusing on the promoter region of hundreds of cancer related genes (see section 2.7.2). Methylation profiling of a large cohort of primary medulloblastomas using the Illumina GoldenGate array has revealed the existence of four discrete methylomic subgroups of the disease that are strongly associated with their previously defined transcriptomic counterparts (SHH, WNT, Group 3 and Group 4) (Schwalbe *et al.*, 2013) (see section 1.7.2.1).

Medulloblastoma is thus a highly heterogenous tumour and is no longer considered a single disease. These 4 principal subgroups of medulloblastoma are associated with distinct molecular, clinical and pathological features and survival outcome, and will likely require different therapeutic strategies in the clinic (Kool *et al.*, 2012; Taylor *et*

al., 2012). The WNT and SHH subgroups of medulloblastoma are molecularly well-defined by key aberrations in their respective signalling pathways (see section 1.5.7), and the WNT subgroup is associated with a favourable prognosis compared to the other tumour groups (Ellison *et al.*, 2005; Clifford *et al.*, 2006; Kool *et al.*, 2012). Different cellular origins have been identified for the WNT and SHH subgroups (Schuller *et al.*, 2008; Gibson *et al.*, 2010). In contrast, very little is known about the underlying molecular pathology of the generically named Group 3 and Group 4 tumours. These two tumour subgroups make up approximately 60% of medulloblastomas. Group 3 medulloblastomas currently have the worst outcome of the 4 subgroups; *MYC* amplification is enriched in Group 3 tumours, they are associated with LCA histology and are frequently metastatic at diagnosis (Kool *et al.*, 2012) (see section 1.5.8.2). Group 4 is the largest tumour subgroup and carries an intermediate prognosis similar to SHH tumours. They are mainly classic tumours but can be of the LCA histological type and metastases are found in this subgroup. It is the Group 4 tumours whose molecular pathology remains the most elusive of all the subgroups (Kool *et al.*, 2012) (see section 1.5.8.2). While advances in the treatment of medulloblastoma have improved survival, the aggressive therapeutic regimens currently employed are associated with significant long-term adverse effects in the majority of survivors and for high-risk patients the prognosis remains poor (Pizer and Clifford, 2009).

Previously, epigenetic regulation of *COL1A2* was found to be significantly associated with distinct tumour-types of medulloblastoma, discriminating infant DN tumours from non-infant DN cases/non-desmoplastic tumours (Anderton *et al.*, 2008). The presence of distinct methylomic subgroups offers the potential to further our understanding of the underlying molecular pathogenesis of the distinct subgroups of medulloblastoma. Paediatric medulloblastoma is characterised by relatively few mutational events (Parsons *et al.*, 2011; Jones *et al.*, 2012) and, therefore, the potential role of subgroup-specific DNA methylation events must be investigated if we are to identify additional critical genes and novel pathways that may be driving tumourigenesis and that may be targeted therapeutically.

To date, there have been no detailed studies into the role of DNA methylation in the pathogenesis of Group 3 and Group 4 medulloblastomas. Using high-throughput DNA methylation microarrays, this chapter reports a global investigation into DNA

methylation-dependent gene regulation and its potential role in Group 3 and Group 4 medulloblastoma tumourigenesis. Unravelling the molecular biology of Group 3 and Group 4 medulloblastomas is vital if we are to deliver better and more personalised therapies to these patients, with the primary goal of improving survival and minimising the long-term adverse outcomes associated with aggressive treatment.

3.2 Aims

The work reported in this chapter aimed to:

- Identify significantly differentially methylated CpG sites that distinguish the Group 3 and the Group 4 medulloblastomas from other tumour subgroups.

- Identify those differentially methylated genes that also exhibit significant differential expression in Group 3 and Group 4 tumours compared to other tumour subgroups across multiple independent primary tumour cohorts.

- Evaluate the relationship between the patterns of differential methylation and gene expression observed across tumour subgroups, and identify putative candidate Group 3- and Group 4-specific epigenetically regulated genes.

- For the selected list of candidate genes, investigate the relationship between DNA methylation and gene expression in a panel of medulloblastoma cell lines treated with the demethylating agent 5-azaCdR.

3.3 Materials and Methods

3.3.1 GoldenGate Cancer Panel I Methylation Microarray

The Illumina GoldenGate Cancer Panel I methylation microarray (see section 2.7.2) was used to assess patterns of methylation in a defined cohort of primary medulloblastomas, medulloblastoma cell lines and non-neoplastic cerebellar samples. The array interrogates the methylation status of 1,505 CpG sites located in the promoter region of over 800 genes. The methylation microarray analysis was performed at the Wellcome Trust Centre for Human Genetics, Oxford, UK, according to manufacturer's protocols (Illumina, San Diego, CA, USA). All raw data processing and normalisation were carried out by Dr. Ed Schwalbe (Schwalbe *et al.*, 2013), and final methylation β-value scores were kindly provided by Dr. Schwalbe. The β-value measures methylation status on a continuous scale, ranging from 0 (completely unmethylated) to 1 (completely methylated) (see section 2.7.2.1). The Illumina annotation data file for the GoldenGate methylation array was kindly provided by Dr. Ed Schwalbe and facilitated the mapping of probes to their official gene symbol, chromosome location and CpG island status.

3.3.1.1 Primary medulloblastoma cohort

A representative cohort of 216 primary medulloblastomas (see section 2.1.1) was analysed on the GoldenGate methylation array and used for investigations reported in this chapter. The cohort included all known medulloblastoma histopathological subtypes and patient ages ranged from 1.2 months to 43 years old. The 4 principal molecular subgroups of medulloblastoma (SHH, WNT, Group 3 and Group 4) were represented in the cohort. Newcastle Research Ethics Committee approval had been obtained for the collection, storage and biological study of all material.

3.3.1.2 Medulloblastoma cell lines

Eleven medulloblastoma cell lines (Med1, Med8A, DAOY, D283, D341, D384, D556, D425, D458, UW228-2, UW228-3) (see section 2.2) were analysed on the GoldenGate methylation array.

3.3.1.3 Non-neoplastic cerebellar samples

The methylation profiles of 13 non-neoplastic cerebellar samples assayed on the GoldenGate methylation array were kindly provided by Dr. Christine Ladd-Acosta (Johns Hopkins University School of Medicine, Baltimore, USA). The cohort comprised 10 infant and 2 adult samples with the patient age of 1 sample unknown. Eight further non-neoplastic cerebellar samples were collected in Newcastle, UK and comprised 4 foetal, 2 newborn and 2 adult samples. Samples consisted of post-mortem material from patients who died from non-neoplastic conditions.

3.3.1.4 Extraction and preparation of DNA

DNA extractions for samples analysed on the GoldenGate methylation array were carried out by Dr. Meryl Lusher, Dr. Janet Lindsey and Mr. Kieran O'Toole (Newcastle University Paediatric Brain Tumour Research Group). The primary cohort comprised samples extracted from both frozen tumour tissue and formalin-fixed paraffin embedded (FFPE) tissue. DNA was extracted from FFPE samples using a Qiamp DNA FFPE tissue kit (Qiagen Ltd, Manchester, UK) and from cell lines using a Qiagen DNeasy kit (Qiagen Ltd, Manchester, UK), according to manufacturer's instructions (see section 2.6.2). DNA was extracted from frozen tumour samples using TRIzol® Reagent (Invitrogen Ltd, Paisley, UK), according to manufacturer's instructions (see section 2.6.2). Aliquots of 1µg of DNA at 100ng/µl, as measured by NanoDrop (see section 2.6.3.1), were sent for methylation array profiling.

3.3.2 Medulloblastoma methylomic subgroups

The 216 primary medulloblastomas were reliably subclassified according to their GoldenGate methylation profiles by Dr. Ed Schwalbe (Newcastle University Paediatric Brain Tumour Research Group). Dr. Schwalbe demonstrated by cross-validation in training and test cohorts that medulloblastoma comprises four robust DNA methylation subgroups (SHH, WNT, Group 3 and Group 4) that are strongly associated with the four principal transcriptomic subgroups of the same nomenclature and that have distinct molecular, clinical and pathological disease features (Schwalbe *et al.*, 2013).

The proportion of tumours classified in each methylomic subgroup was comparable to those described for their transcriptomic counterparts (Kool *et al.*, 2012). Group 4 tumours formed the largest group (44% [94 tumours]), followed by SHH (23% [50

tumours]) and Group 3 (20% [44 tumours]). WNT tumours represented the smallest group (13% [28 tumours]) (Table 2.1).

3.3.3 Medulloblastoma gene expression profiles

3.3.3.1 Primary tumour transcriptomic datasets

Three independent *in silico* transcriptomic datasets (Kool/Fattet, Cho and Northcott; see section 2.4) were used for investigations reported in this chapter. Expression profiles and supporting annotation data files were accessed as described in section 2.4. The primary tumours in each dataset were reliably subclassified into the four principal subgroups of medulloblastoma (SHH, WNT, Group 3 and Group 4) by Dr. Dan Williamson (Newcastle University Paediatric Brain Tumour Research Group) using a 4 metagene classifier as described in section 2.4.1.

3.3.3.2 Cell line transcriptomic dataset

Microarray gene expression profiling of 3 medulloblastoma cell lines, cultured in the presence and absence of the demethylating agent 5-azaCdR, had previously been carried out within the Newcastle University Paediatric Brain Tumour Research Group (Anderton *et al.*, 2008)**.** The Affymetrix U133A gene expression microarray platform was used to analyse expression changes following demethylation of the MED8A, D283 and D425 cell lines (see section 2.2). This *in silico* cell line transcriptomic dataset was used to facilitate selection of candidate epigenetically regulated genes for further investigation.

3.3.4 Differential methylation analysis

3.3.4.1 Identification of Group 3 and Group 4 DNA methylation markers

The Goldengate methylation β-value (see section 2.7.2.1) has a bounded range from 0 (completely unmethylated) to 1 (completely methylated). The distribution of β-values is bimodal with the majority of samples being either fully methylated or fully unmethylated (Figure 3.1). Due to this non-normal distribution the non-parametric Mann-Whitney U test was used to identify CpG sites with significantly different methylation profiles in the Group 3 and the Group 4 medulloblastomas compared to the other tumour subgroups.
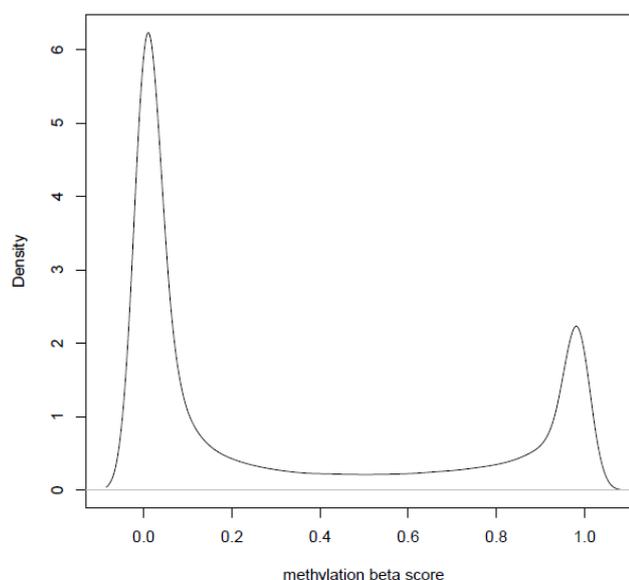
Figure 3.1. Density distribution of methylation β-values in primary medulloblastoma cohort. Kernel density plot illustrates the bimodal distribution of methylation β-values for the 216 primary tumours analysed on the GoldenGate methylation microarray.

Applying the Mann-Whitney U test, pair-wise comparisons of the methylation status of the 1,505 CpG sites on the array were tested as follows:

- Group 3 tumours were tested against the combined SHH, WNT and Group 4 tumours [G3 *vs* Other]

- Group 4 tumours were tested against the combined SHH, WNT and Group 3 tumours [G4 *vs* Other]

- The combined Group 3 and Group 4 tumours were tested against the combined SHH and WNT tumours [(G3+G4) *vs* (SHH+WNT)]

The latter comparison was tested because, although Group 3 and Group 4 are distinct medulloblastoma subgroups, studies suggest they are likely more similar to each other than to SHH or WNT tumours (Taylor *et al.*, 2012). Investigating the role of the Group (3+4) methylation markers in gene regulation offers the potential to identify underlying molecular pathology common to both Group 3 and Group 4 tumours.

The Mann-Whitney U significance tests were conducted using the R Statistical Computing Program (v. 2.12). Raw *p*-values were corrected for multiple hypothesis testing using the Benjamini-Hochberg false discovery rate (FDR) correction (Benjamini and Hochberg, 1995). A FDR-adjusted *p*-value < 0.05 was considered statistically significant. To increase the sensitivity of the significance test and ensure that only

biologically relevant probes were selected an additional filter of an absolute difference in mean β-value (Δβ) greater than 0.34 was applied. This value was chosen as it is double the limit of detection of the GoldenGate assay and is detectable by bisulfite sequencing (Illumina, 2007).

Heatmaps, representing the methylation β-values as colours, were produced to visualise subgroup-specific methylation profiles. Principle component analysis (PCA), a statistical procedure that reveals the covariance structure of a complex dataset by reducing it to a smaller number of dimensions (principal components) was a useful tool for the visualisation and interpretation of differential methylation observed between medulloblastoma subgroups. The principal components of a complex dataset account for as much of the variation in the data as possible. Combined boxplots and strip-plots, depicting the distribution of β-values through quartiles, provided a visual representation of the distribution of methylation values within and between subgroups for individual CpG sites and facilitated the selection of candidate epigenetically regulated genes for further investigation (see section 3.3.6). Heatmaps, PCA plots and combined boxplots and strip-plots were constructed using R (v.2.12).

3.3.4.2 Tumour specificity of Group 3 and Group 4 methylation markers

The Group 3 and Group 4 methylation markers identified in section 3.3.4.1 were tested in their respective subgroup against the non-neoplastic cerebellar samples. The Mann-Whitney U test was used to assess the tumour specificity of the subgroup-specific methylation markers. As described above, a FDR-adjusted p-value < 0.05 was considered statistically significant and an additional filter of an absolute difference in mean β value greater than 0.34 was applied. Heatmaps were constructed to visualise the results.

3.3.5 Differential gene expression analysis

Genes, corresponding to the significant differentially methylated CpG sites identified in section 3.3.4.1, were investigated for evidence of differential gene expression in 3 independent primary tumour transcriptomic datasets (see section 3.3.3.1).

Methylation and expression probes were paired on the basis of their official gene symbol identifiers as annotated by Illumina and Affymetrix respectively. For the Affymetrix U133A (Cho dataset) and U133plus 2.0 (Kool/Fattet dataset) expression

array platforms, probes annotated with the "_at" suffix were preferentially selected over those suffixed with "_a_at", "_s_at" or "_x_at", as they belong to unique probe sets that do not cross-hybridise to any other sequences (Leong *et al.*, 2005). One probe per gene was selected for differential expression analysis. For the Affymetrix Exon 1.0ST array platform (Northcott dataset) the corresponding transcript cluster identifier was selected. Transcript clusters comprised all probes targeting individual exons of a given gene and provided gene-level expression estimates.

Assuming normality, the parametric unpaired *t*-test was used to identify differentially expressed genes. The same pair-wise comparisons were tested as described in section 3.3.4.1; genes corresponding to the Group 3-specific methylation markers were tested in a [G3 *vs* Other] comparison, Group 4-specific markers in a [G4 *vs* Other] and the Group (3+4) markers in a [(G3+G4) *vs* (SHH+WNT)] comparison.

Unpaired *t*-tests were conducted using the R Statistical Computing Program (v. 2.12). Raw *p*-values were corrected for multiple hypothesis testing using a Benjamini-Hochberg false discovery rate (FDR) correction (Benjamini and Hochberg, 1995)**.** A FDR-adjusted *p*-value < 0.05 was considered statistically significant. To increase the sensitivity of the significance test and ensure the biological relevance of differential gene expression observed an absolute difference in mean $\log_2$ signal intensity (mean log ratio) threshold was applied. For each dataset, the average absolute difference in mean $\log_2$ intensities for all probes in all pair-wise comparisons was calculated and applied as the difference threshold. A mean log ratio threshold of 0.3 was applied to the Kool/Fattet and Northcott datasets and a threshold of 0.4 was applied to the Cho dataset (see Appendix B). Genes with a FDR-adjusted *p*-value < 0.05 and a mean log ratio > threshold in at least 2 out of 3 transcriptomic datasets were considered significantly differentially expressed.

Combined boxplots and strip-plots illustrating the patterns of differential gene expression across the tumour subgroups were constructed in R (v.2.12). Assessment of the boxplots was performed to facilitate the selection of candidate epigenetically regulated genes for further investigation (see section 3.3.6).

3.3.6 Identification of methylation events with a putative effect on gene expression

The patterns of significant differential methylation and gene expression observed across the tumour subgroups were assessed using the combined boxplots and strip-plots. Methylation of the corresponding CpG site was considered to have a putative effect on gene expression if a significantly higher subgroup-specific mean methylation status was associated with a significantly lower equivalent subgroup-specific mean gene expression level, and conversely if a significantly lower mean methylation status was associated with a significantly higher mean gene expression level (inverse relationship). Genes were considered to show evidence of potential epigenetic gene regulation by DNA methylation if an inverse subgroup-specific methylation-expression relationship was observed in 2 or more transcriptomic datasets.

3.3.7 Correlating DNA methylation to gene expression using Pearson correlation

Tumour-matched CpG methylation and gene expression data was only available for 9 primary tumour samples. The primary tumours were Newcastle medulloblastoma samples (NMB111, NMB125, NMB131, NMB134, NMB135, NMB139, NMB143, NMB82, NMB93) and had been expression profiled by Kool *et al*. (Kool *et al.*, 2008). The 9 samples comprised four WNT subgroup tumours, one SHH and four Group 4 tumours (Table 2.1).

The association between CpG methylation and the level of gene expression was assessed by calculating Pearson's product-moment correlation coefficient (r) The level of gene expression was considered to be inversely correlated with CpG methylation if the coefficient r value obtained was less than (–0.7).

3.3.8 In silico analysis of gene expression changes following demethylating treatment in medulloblastoma cell lines

The *in silico* cell line transcriptomic dataset, described in section 3.3.3.2, was used to assess microarray gene expression changes following 5-azaCdR treatment of 3 medulloblastoma cell lines.

Based on the magnitude of mean log ratios observed for significant differentially expressed genes in primary tumours (see section 3.3.5 and Appendix B), probes

detecting an expression increase >1.5 fold following 5-azaCdR were considered upregulated by 5-azaCdR. To determine whether transcriptional silencing and upregulation following 5-azaCdR treatment was associated with DNA methylation, the GoldenGate methylation β-values of the corresponding CpG sites were considered; sites with a β value ≥0.7 were classed as methylated (M), those with a β value <0.3 were classed as unmethylated (U) and those with values between 0.3 and 0.7 were classed as part-methylated (PM). Genes which were upregulated by 5-azaCdR treatment consistent with their CpG methylation were considered to show evidence of methylation-dependent gene regulation in cell lines.

3.3.9 Identification of subgroup-specific candidate epigenetically regulated genes

Putative subgroup-specific candidate epigenetically regulated genes were identified as those that:

- were significantly differentially methylated at one or more CpG sites in Group 3 and/or Group 4 tumours, and patterns of differential methylation showed an inverse association with equivalent subgroup-specific differential gene expression in at least 2 transcriptomic datasets (see section 3.3.6)

And either

- showed a strong inverse correlation between methylation and gene expression, as measured by Pearson's correlation coefficient *r* (*r*< -0.7) (see section 3.3.7)

Or

- showed >1.5-fold increase in expression following 5-azaCdR treatment in at least one methylated cell line, as determined by *in silico* analysis of 5-azaCdR-induced gene expression changes (see section 3.3.8)

It was decided not to require the Group 3 and Group 4 methylation markers to also be tumour-specific (see section 3.3.4.2) in the candidate gene selection criteria. It was recognised that this could lead to some candidate genes having an aberrant methylation status that would potentially be functionally significant in SHH or WNT tumours. However, as it is likely that Group 3 and Group 4 tumours have different cellular origins to each other and to SHH and WNT tumours (Kool *et al.*, 2012) (see section

1.5.9), and in the absence of appropriate cell-type specific controls to allow an accurate comparison of tumour *versus* normal, tumour specificity was not a requirement. The tumour-specific nature of Group 3- and/or Group 4-specific methylation markers was, however, used as a reference to appraise the potential functional significance of subgroup-specific DNA methylation events. Candidate genes were selected for further investigation of a direct relationship between CpG methylation and gene expression by qRT-PCR assessment of expression changes in medulloblastoma cell lines treated with 5-azaCdR (see section 3.3.10).

3.3.10 Real-time quantitative RT-PCR

Real-time quantitative reverse transcriptase PCR (qRT-PCR) (see section 2.9.4) was performed in a panel of medulloblastoma cell lines to assess candidate gene expression levels before and after treatment with the demethylating agent 5-azaCdR. Relative gene expression levels before and after 5-azaCdR treatment were determined using the delta delta C_T method (see section 2.9.4.1).

3.3.10.1 Medulloblastoma cell lines

A panel of 6 medulloblastoma cell lines (D283, D384, D425, D556, DAOY and MED1) (see section 2.2) had previously been cultured in the presence of 5µM 5-azaCdR for 72 hours. RNA extracted from the untreated and 5-azaCdR-treated cell line pairs was kindly provided by Dr. Janet Lindsey (Newcastle University Paediatric Brain Tumour Group).

3.3.10.2 cDNA synthesis

The concentration of RNA in the cell line samples was determined by NanoDrop (see section 2.6.3.1) and 1µg of cell line total RNA was reverse transcribed to cDNA using the iScript cDNA synthesis kit (Bio-Rad Laboratories Ltd, Hemel Hempstead, UK) (see section 2.9.2), according to the manufacturer's instructions.

3.3.10.3 Primer design

qRT-PCR primers were designed using transcript information obtained from the EMBL-EBI Ensembl Genome Browser (www.ensembl.org). To prevent genomic DNA detection and amplification during the PCR reaction primers were designed to span an exon-exon junction. In circumstances where genes had more than one transcript the

UCSC Genome Browser (http://genome.ucsc.edu/) was used to identify exons common to multiple isoforms.

The sequences for gene-specific forward and reverse primers (Table 3.1) were designed using the Primer Express 2.0 software (Applied Biosystems, Foster City, USA). The software automates primer design and was designed specifically for Applied Biosystems real-time PCR instruments. Primer Express facilitates the design of primers for applications using SYBR Green I dye detection (see section 2.9.4); the designed primers required minimal optimisation (see section 3.3.10.4) and were compatible with the thermal cycling conditions (see section 3.3.10.5). All primers were designed to anneal at 60^0C and to generate amplicons with a maximum length of 150bp to facilitate optimum reaction efficiencies. To ensure that primers spanned an exon-exon junction and that no SNPs were present in the sequence, the primer pairs were assessed using the *in-silico* PCR function on the UCSC Genome Bioinformatics website (http://genome.ucsc.edu/). The NCBI BLAST tool was used to ensure target specificity of primer sets (http://blast.ncbi.nlm.nih.giv/). BLAST aligns and compares query DNA sequences against gene sequence databases and facilitates the identification of non-specific amplification.

The isoform A transcript of the *RASSF1* gene was used as a positive control for the demethylating treatment of the cell lines with 5-azaCdR. *RASSF1A* was used due to the robust evidence for its methylation-dependent transcriptional silencing and re-activation following 5-azaCdR treatment in multiple medulloblastoma cell lines (Lusher *et al.*, 2002)**.** A previously published *RASSF1A* primer set was used (Palakurthy *et al.*, 2009) (Table 3.1). Primer sequences for the endogenous control genes, *TBP* and *GAPDH*, were kindly provided by Dr. Sarra Ryan and Dr. Lisa Russell, respectively (Newcastle University Leukaemia Research Cytogenetics Group) (Table 3.1). *TBP* and *GAPDH* are common endogenous control genes in PCR applications. They are housekeeping genes essential for the maintenance of basic cellular functions and are constitutively expressed in all cells. Endogenous controls were necessary to normalise any differences in starting cDNA quantities between samples.

Primers were custom synthesised by Sigma Aldrich (Sigma Aldrich Ltd, Gillingham, UK) and supplied as lyophilised pellets. Upon delivery primers were reconstituted in

nuclease-free water (Qiagen Ltd, Manchester, UK) at a concentration of 100µM and stored at -20^0C alongside 10µM working stock aliquots.

Gene Name	F primer (5'-3')	R primer (5'-3')	Amplicon size (bp)	Exons spanned
DDR2	CAGCTTCCAGTCAGTGGTCAGA	CCAGCCTTCCATATTTGGCA	51	5&6
HDAC1	GCCATCCTGGAACTGCTAAAGT	AATGTCAATGTACAGCACCCTCTG	51	5&6
MET	GCTACACACTGGTTATCACTGGGA	CATTCAATGGGATCTTCGTGATC	51	4&5
RHOH	CCAGTTGAAGACTAGGCTTTGGAG	CATCCAAGCACCGTCTGCTT	51	2&3
RASSF1A	ACCTCTGTGGCGACTTCATC	CGGTAGTGGCAGGTGAACTT	51	1&2
GAPDH	CAAGGTCATCCATGACAACTTTG	GTCCACCACCCTGTTGCTGTAG	51	7&8
TBP	TGTATTAACAGGTGCTAAAGTCAG	TTTTCAAATGCTTCATAAATTTCTGC	51	7&8

Table 3.1. Primer sequences for qRT-PCR. Forward and reverse primer sequences, amplicon size and exons spanned during primer design are shown for target and control genes.

3.3.10.4 Primer validation and optimisation

A melt curve analysis was performed immediately after each real-time PCR run to confirm target specificity of primers. The Primer Express software sets a default maximum melting temperature (T_m) of 85^0C for amplicons generated during the PCR reaction, with most falling between 75^0C and 85^0C. The melt curve analysis was applied between $60-95^0$C at 10^0C intervals and produced a plot of change in fluorescence against temperature. For each primer set the melt curve should have one peak occurring at the appropriate T_m for the amplicons generated and this should be consistent across all samples tested. A melt curve will identify any non-specific amplification, including primer-dimer formation (see Figure 3.2). Primer-dimers occur when primers anneal to each other creating short sequence by-products of the PCR reaction that can adversely affect the quality of PCR results.

The formation of primer-dimers can be minimised by optimising the concentration of primers being added to the reaction. The minimum primer concentration was chosen that produced sufficient early detection (lowest C_T/ highest fluorescent signal) while minimising primer-dimer formation (see Table 3.2). All primer sets were validated using melt curves and, at the primer concentrations detailed in Table 3.2, non-specific amplification was not evident.

PCR reaction efficiencies for the target and endogenous control genes were validated as described in section 3.3.11.1. Inspection of C_T values and signal intensities produced from these validation experiments facilitated the optimisation of input cDNA concentration. The optimum C_T values for gene quantification are generally between 15 and 30. For genes whose initial C_T values fell outside this range (highly and lowly expressed) the concentration of input cDNA was optimised (Table 3.2).
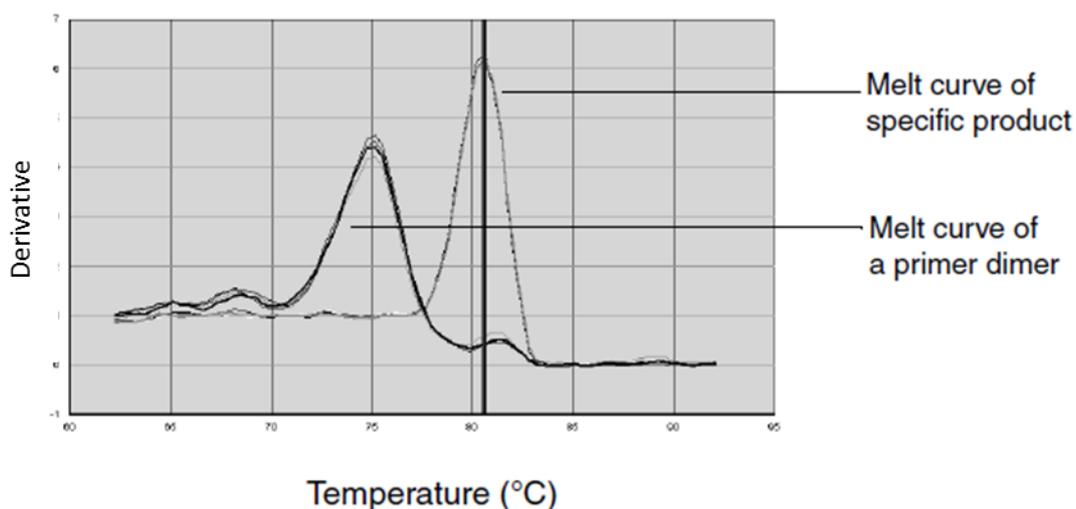
Figure 3.2. Melt curve analysis to detect non-specific PCR amplification. The melt curve represents a plot of the first derivative of rate of change in fluorescence as a function of temperature and shows typical primer-dimer formation. Primer-dimers are short sequence PCR by-products caused by primers annealing to each other. Primer-dimers have a characteristically lower T_m compared to that of the larger specific target PCR product. Figure taken from the Applied Biosystems SYBR® GreenPCR Master Mix User Guide (Applied Biosystems, Foster City, USA).

3.3.10.5 Protocol for qRT-PCR

qRT-PCR reactions were carried out using Platinum SYBR® Green qPCR SuperMix-UDG with ROX (Invitrogen Ltd, Paisley, UK). 10µl reactions were set up as follows: 5.1µl SYBR Green superMix, 0.5µl nuclease-free water, 0.2µl forward primer and 0.2µl reverse primer. 6µl of reaction mix and 4µl of cDNA template were added to wells on a 384 well-plate. Each reaction was performed in triplicate with no-template controls. Reactions were set up in a LabCaire PCR clean air cabinet (Scientific Laboratory Supplies Ltd, Nottingham, UK) and the cDNA template was added at the bench. Table 3.2 details the primer concentration and the input cDNA concentration used for each target gene reaction.

Real-time PCR amplification was carried out using the ABI PRISM 7900 HT Detection System (Applied Biosystems, Foster City, USA) according to the following protocol: 10

minutes at 95^0C (incubation), followed by 40 cycles of 15 seconds at 95^0C (denaturation) and 60 seconds at 60^0C (annealing/extension). The qRT-PCR results presented in section 3.4.9 are representative of triplicate PCR runs carried out on different days.

Amplicon	Primer conc. in 10µl reaction	cDNA conc. in 10µl reaction
DDR2	0.2µM	0.2ng/µl
HDAC1	0.2µM	0.2ng/µl
MET	0.05µM	2ng/µl
RHOH	0.05µM	2ng/µl
RASSF1A	0.2µM	7ng/µl
GAPDH	0.2µM	0.2ng/µl
TBP	0.2µM	0.2ng/µl

Table 3.2. Optimised primer concentration and cDNA template concentration for target and control genes. Concentration of primers and cDNA template added to each 10µl PCR reaction are shown for target and control genes tested by qRT-PCR.

3.3.11 Relative quantification using the delta delta C_T method

The delta delta C_T ($\Delta\Delta C_T$) method (see section 2.9.4.1) was used to determine the relative gene expression levels in medulloblastoma cell lines before and after treatment with 5-azaCdR. A validation experiment was first performed to determine if the reaction efficiencies of the target and endogenous control genes were approximately equal, ensuring the validity of the $\Delta\Delta C_T$ calculation.

3.3.11.1 Validation of PCR reaction efficiencies

PCR reaction efficiencies for the target and endogenous control genes were assessed by carrying out reactions using two- and three-fold serial dilutions of a pooled medulloblastoma cell line cDNA sample. The concentration of cDNA in the pooled sample was determined by NanoDrop (see section 2.6.3.1) and dilutions were prepared spanning 5-6 input cDNA concentrations. Reactions were performed in triplicate as described in section 3.3.10.5. C_T values were exported from the instrument software to Microsoft Excel 2007 and the ΔC_T (C_T target – C_T endogenous control) was calculated for each serial dilution of cDNA template. A plot of log input cDNA amount *versus* ΔC_T was produced in Microsoft Excel. The efficiencies of target and endogenous control genes were considered equal if the absolute value of the slope was ≤0.2.

3.3.11.2 Relative quantification of candidate gene expression before and after 5-azaCdR treatment

The $\Delta\Delta C_T$ method measures changes in gene expression relative to a reference sample (calibrator sample). For investigations reported in this chapter the sample with the lowest expressed levels (highest C_T values) was designated the calibrator sample. Following each PCR run C_T values were exported to Microsoft Excel 2007 and the mean C_T and standard deviation values of the replicate sample results were calculated. The mean C_T was used to determine relative target gene expression as follows:

- ΔC_T was calculated by: C_T target $-$ C_T endogenous control.
- $\Delta\Delta CT$ was calculated by: $\Delta C_{T \text{ test sample}}$ - $\Delta C_{T \text{ calibrator sample.}}$
- Relative gene expression in the treated and untreated cell line pairs was calculated using the formula $2^{-\Delta\Delta CT}$.

For each candidate gene, the fold-change in expression following 5-azaCdR treatment was calculated by normalising the relative mean expression in the treated cell line to the relative mean expression in the untreated cell line in Excel 2007. Graphs illustrating relative mean expression and mean fold-change in expression from triplicate PCR runs were produced in GraphPad Prism (v.4.03).

Genes that showed an expression increase > 2-fold following 5-azaCdR treatment were classed as being significantly upregulated. To determine whether the transcriptional silencing and gene upregulation following 5-azaCdR treatment observed was associated with DNA methylation, the methylation status of the corresponding GoldenGate CpG probe (see section 3.3.6) was considered. Using the GoldenGate methylation β-value scores, loci with a β value ≥0.7 were classed as methylated (M), those with a β value <0.3 were classed as unmethylated (U) and those with values between 0.3 and 0.7 were classed as part-methylated (PM). Genes that were silenced and upregulated by 5-azaCdR in methylated cell lines but not in unmethylated cell lines had a methylation status consistent with their epigenetic regulation and showed strong evidence of methylation-dependent transcriptional regulation in medulloblastoma cell lines. Candidate gene CpG methylation was considered to be functionally significant in medulloblastoma cell lines if evidence of epigenetic gene regulation was seen in 4 or more cell lines.

3.4 Results

This chapter aimed to investigate the role of gene promoter DNA methylation in the regulation of gene expression in the Group 3 and Group 4 medulloblastoma subgroups. Using the high-throughput GoldenGate methylation microarray a panel of 807 genes were assessed for potential epigenetic regulation in a representative cohort of 216 primary medulloblastomas.

3.4.1   Novel methylation markers of Group 3 and Group 4 medulloblastomas

Using the Mann-Whitney U test, novel methylation markers of Group 3, Group 4 and of Group (3+4) medulloblastomas were identified as described in section 3.3.4.1.

There were 23 significant differentially methylated CpG sites in Group 3 tumours (Table 3.3 and Figure 3.3(A)); 31 CpG sites were Group 4-specific (Table 3.4 and Figure 3.3(B)) and 48 CpG sites significantly discriminated Group (3+4) from both SHH and WNT tumours (Table 3.5 and Figure 3.4)

Three Group 3-specific events (13%) were also specific to Group (3+4) (Table 3.3). In the larger Group 4 there were 26 events (84%) that were also significant for Group (3+4) (Table 3.4). Significance testing was carried out by combining distinct subgroups of medulloblastoma together to facilitate a 2 group comparison (see section 3.3.4.1). The high proportion of Group 4 events that are also Group (3+4) events highlights the potential limitations of the experimental design and significance test applied in accommodating the unequal variances and sample sizes across the 4 distinct subgroups of medulloblastoma. The results do, however, strongly support the current view that while Group 3 and Group 4 are distinct molecular entities in medulloblastoma they are more similar to each other than to SHH and WNT tumours (Taylor *et al.*, 2012).

A total of 73 CpG sites (encompassing 63 genes) were identified that significantly distinguish the Group 3 and/or Group 4 medulloblastomas and represent novel methylation markers of these poorly characterised subgroups (Figure 3.5).

Group 3 *vs* Other					
Methylation CpG probe	**CpG Island**	**Mean β value (Group 3)**	**Mean β value (Other)**	**Absolute Difference in mean β value (Δβ)**	**Corrected *p*-value**
BCAP31_P1072_F	Y	0.39	0.86	0.47	5.47E-10
BCR_P422_F	Y	0.64	0.28	0.36	5.06E-09
BLK_P668_R *	N	0.26	0.82	0.56	1.28E-13
CCL3_E53_R	N	0.50	0.87	0.36	8.93E-11
CHI3L2_P226_F *	N	0.27	0.66	0.39	2.05E-10
FES_E34_R	Y	0.58	0.10	0.48	7.79E-13
FES_P223_R	Y	0.60	0.13	0.47	2.14E-11
FGF1_P357_R	N	0.61	0.22	0.39	2.12E-10
FRZB_E186_R	Y	0.52	0.08	0.43	1.08E-08
HFE_E273_R *	Y	0.88	0.41	0.48	6.35E-10
HLA-DOB_P357_R	N	0.42	0.79	0.37	9.25E-11
IL1RN_P93_R	Y	0.58	0.13	0.45	1.42E-13
KRT1_P798_R	N	0.54	0.93	0.39	5.31E-14
MMP14_P13_F	Y	0.53	0.09	0.46	1.04E-12
PLA2G2A_E268_F	N	0.47	0.88	0.41	7.22E-12
RARRES1_P426_R	Y	0.63	0.21	0.41	8.93E-11
SERPINA5_E69_F	N	0.54	0.88	0.34	5.54E-12
TGFB2_E226_R	Y	0.50	0.09	0.41	7.31E-06
THBS2_P605_R	N	0.46	0.88	0.43	1.04E-12
TP73_P945_F	Y	0.56	0.19	0.37	6.79E-07
TRIP6_E33_F	Y	0.59	0.18	0.41	2.64E-11
WRN_P969_F	Y	0.65	0.173	0.49	1.37E-12
ZNFN1A1_E102_F	Y	0.41	0.90	0.49	5.31E-14

Table 3.3. Medulloblastoma Group 3-specific DNA methylation events. Results of the Mann-Whitney U tests assessing methylation in Group 3 tumours against methylation in all other tumour subgroups are detailed. Methylation probe ID and presence of a CpG island, as defined by Illumina, are shown. 23 CpG sites were significantly differentially methylated with a FDR-adjusted p-value <0.05 and a Δβ value >0.34. 61% (14/23) of sites were located in a CpG island. One gene (*FES*) had 2 CpG sites both located in a CpG island that had significantly different methylation levels in Group 3 tumours compared to all other subgroups. * Probes were also significantly differentially methylated in Group (3+4) tumours compared to (SHH+WNT) tumours.

Group 4 *vs* Other					
Methylation CpG probe	**CpG Island**	**Mean β value (Group 4)**	**Mean β value (Other)**	**Absolute Difference in mean β value (Δβ)**	**Corrected *p*-value**
ASCL2_P609_R **	Y	0.12	0.49	0.37	2.34E-10
CCKAR_E79_F **	N	0.08	0.58	0.50	2.30E-15
CCKAR_P270_F **	N	0.10	0.59	0.49	3.86E-14
CSF3R_P472_F	N	0.36	0.71	0.36	4.63E-17
DDR2_E331_F **	N	1.0	0.64	0.35	2.66E-10
FZD9_E458_F **	Y	0.13	0.47	0.34	6.71E-09
HDAC1_P414_R **	Y	0.85	0.44	0.41	9.42E-17
HIC_seq_48_S103_R	Y	0.33	0.83	0.50	4.25E-16
HLA-DPA1_P28_R	N	0.85	0.50	0.35	1.59E-15
IFNGR2_P377_R **	Y	0.09	0.52	0.43	1.05E-17
IGF1_E394_F	N	0.78	0.40	0.38	1.35E-13
IL16_P93_R **	N	0.21	0.56	0.35	1.11E-10
L1CAM_P19_F **	Y	0.89	0.36	0.53	4.10E-16
LEFTY2_P719_F **	N	0.86	0.49	0.37	1.15E-17
MEST_E150_F **	Y	0.29	0.64	0.35	1.39E-09
MEST_P4_F **	Y	0.39	0.73	0.34	4.65E-10
MEST_P62_R **	Y	0.32	0.70	0.38	1.88E-11
MET_E333_F **	Y	0.66	0.28	0.38	3.52E-10
MMP10_E136_R **	N	0.82	0.42	0.40	9.40E-14
NOTCH4_P938_F	N	0.73	0.38	0.35	3.02E-12
PADI4_P1158_R **	N	0.75	0.41	0.35	5.08E-12
PIK3R1_P307_F **	N	0.08	0.60	0.52	6.39E-22
PLG_E406_F **	N	0.29	0.68	0.39	1.95E-15
RAN_P581_R	Y	0.21	0.64	0.43	4.69E-16
RHOH_P121_F **	N	0.41	0.82	0.41	9.63E-14
SPP1_E140_R **	N	0.75	0.37	0.38	3.92E-14

		Group 4 *vs* Other			
Methylation CpG probe	CpG Island	Mean β value (Group 4)	Mean β value (Other)	Absolute Difference in mean β value (Δβ)	Corrected *p* value
TM7SF3_P1068_R	N	0.53	0.88	0.35	1.59E-15
TRIM29_E189_F **	Y	0.74	0.33	0.41	2.71E-16
WNT10B_P993_F **	Y	0.76	0.33	0.42	8.03E-17
ZIM3_P718_R **	N	0.42	0.77	0.35	1.46E-12
ZNF264_P397_F **	Y	0.09	0.46	0.36	1.53E-09

Table 3.4. Medulloblastoma Group 4-specific DNA methylation events. Results of the Mann-Whitney U tests assessing methylation in Group 4 tumours against methylation in all other tumour subgroups are detailed. Methylation probe ID and presence of a CpG island, as defined by Illumina, are shown. 31 CpG sites were significantly differentially methylated with a FDR-adjusted p-value <0.05 and a Δβ value >0.34. 48% (15/31) of sites were located in a CpG island. Two genes, *CCKAR* and *MEST*, had 2 and 3 CpG sites, respectively, that had significantly different methylation levels in Group 4 tumours compared to all other subgroups. *MEST* CpG sites were located in a CpG island while *CCKAR* sites were not. ** Probes were also significantly differentially methylated in Group (3+4) tumours compared to (SHH+WNT) tumour

Group (3+4) *vs* (SHH+WNT)					
Methylation CpG probe	**CpG Island**	**Mean β value (Group(3+4))**	**Mean β value (SHH+WNT)**	**Absolute Difference in mean β value (Δβ)**	**Corrected *p*-value**
ASCL2_E76_R	Y	0.03	0.46	0.43	3.67E-09
ASCL2_P360_F	Y	0.04	0.56	0.52	5.82E-14
ASCL2_P609_R	Y	0.14	0.67	0.53	1.81E-18
AXIN1_P995_R	Y	0.20	0.57	0.37	2.25E-08
BGN_P333_R	N	0.63	0.26	0.37	3.64E-16
BLK_P668_R	N	0.58	0.93	0.35	8.15E-13
CAPG_E228_F	N	0.95	0.57	0.38	4.89E-22
CCKAR_E79_F	N	0.12	0.80	0.68	9.83E-26
CCKAR_P270_F	N	0.11	0.85	0.74	3.33E-27
CEACAM1_E57_R	N	0.13	0.51	0.37	3.35E-16
CHI3L2_P226_F	N	0.46	0.80	0.34	3.95E-14
CSF3R_P472_F	N	0.44	0.77	0.34	4.74E-15
CYP2E1_E53_R	N	0.66	0.29	0.37	1.05E-12
DDR2_E331_F	N	0.98	0.45	0.53	2.92E-18
FGF1_E5_F	N	0.55	0.19	0.36	6.55E-16
FZD9_E458_F	Y	0.14	0.63	0.49	2.59E-17
HDAC1_P414_R	Y	0.74	0.40	0.34	1.00E-10
HFE_E273_R	N	0.75	0.08	0.67	1.13E-23
HIC1_seq_48_S103_R	Y	0.46	0.88	0.42	5.23E-12
IFNGR2_P377_R	Y	0.16	0.64	0.48	4.91E-20
IL16_P226_F	N	0.11	0.46	0.34	6.77E-11
IL16_P93_R	N	0.22	0.75	0.53	4.89E-22
IL1RN_E42_F	N	0.65	0.32	0.34	2.84E-15
IL8_P83_F	N	0.92	0.57	0.35	9.45E-14
L1CAM_P19_F	Y	0.78	0.25	0.54	1.66E-15

Group (3+4) *vs* (SHH+WNT)					
Methylation CpG probe	CpG Island	Mean β value (Group(3+4))	Mean β value (SHH+WNT)	Absolute Difference in mean β value (Δβ)	Corrected *p*-value
LCN2_P86_R	N	0.86	0.52	0.34	5.07E-16
LEFTY2_P719_F	N	0.80	0.38	0.43	9.14E-20
MEST_E150_F	Y	0.33	0.78	0.45	4.98E-15
MEST_P4_F	Y	0.42	0.87	0.46	6.87E-16
MEST_P62_R	Y	0.36	0.84	0.48	2.68E-17
MET_E333_F	Y	0.66	0.06	0.59	3.16E-21
MMP10_E136_R	N	0.75	0.33	0.42	8.87E-15
MSH2_P1008_F	Y	0.13	0.64	0.51	3.64E-16
PADI4_P1158_R	N	0.68	0.33	0.35	5.45E-11
PIK3R1_P307_F	N	0.15	0.78	0.64	1.69E-25
PLG_E406_F	N	0.31	0.88	0.58	2.58E-29
RHOH_P121_F	N	0.50	0.90	0.40	1.05E-16
SPARC_P195_F	N	0.63	0.15	0.47	4.89E-22
SPP1_E140_R	N	0.73	0.21	0.52	5.84E-23
STAT5A_E42_F	N	0.90	0.47	0.42	4.17E-20
TAL1_P594_F	Y	0.13	0.54	0.42	8.10E-15
TFPI2_P152_R	Y	0.49	0.13	0.36	2.58E-18
TRIM29_E189_F	Y	0.69	0.20	0.50	1.06E-21
VAV1_E9_F	Y	0.96	0.51	0.45	1.31E-18
VAV1_P317_F	N	0.90	0.40	0.49	3.16E-21
WNT10B_P993_F	Y	0.68	0.22	0.46	2.23E-18
ZIM3_P718_R	N	0.48	0.86	0.38	8.49E-16
ZNF264_P397_F	Y	0.15	0.56	0.41	7.70E-11

Table 3.5. **Medulloblastoma Group (3+4)-specific DNA methylation events**. Results of the Mann-Whitney U tests assessing methylation in Group (3+4) tumours against methylation in the SHH and WNT subgroups are detailed. Methylation probe ID and presence of a CpG island, as defined by Illumina, are shown. 48 CpG sites were significantly differentially methylated with a FDR-adjusted p-value <0.05 and a $\Delta\beta$ value >0.34. 44% (21/48) of sites were located in a CpG island. Five genes, *ASCL2,CCKAR, IL16,MEST* and *VAV1*, had 2 or more CpG sites that had significantly different methylation levels in Group (3+4) tumours compared to SHH and WNT tumours. CpG sites for *ASCL2* and *MEST* were located within a CpG island. *VAV1* had one CpG site (VAV1_E9_F) located within an island while the other (VAV1_P317_F) was outside an island. Sites for *CCKAR* and *IL16* were not located within CpG islands.
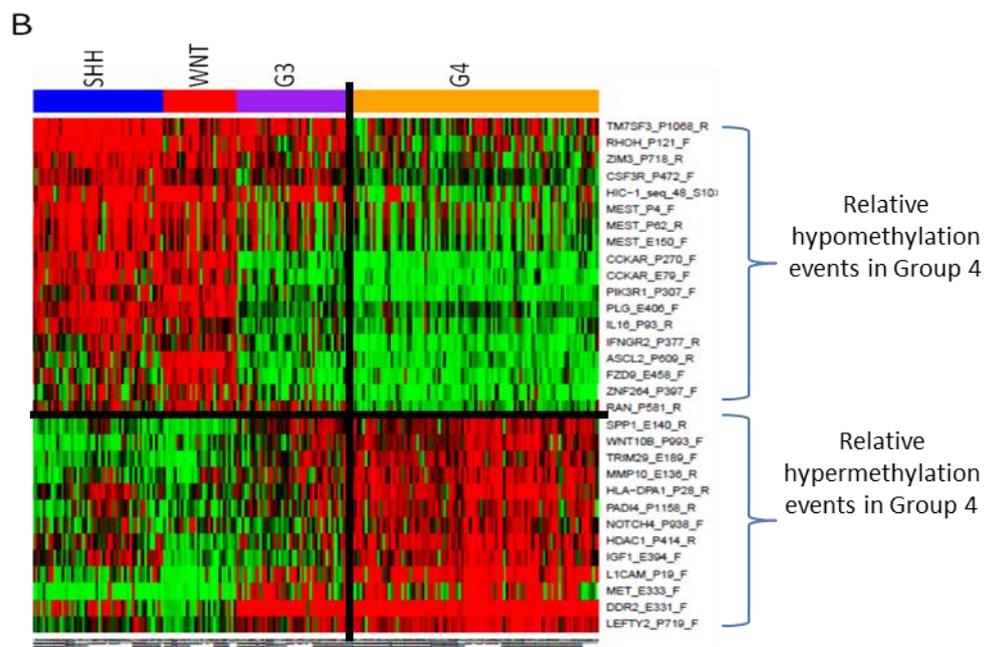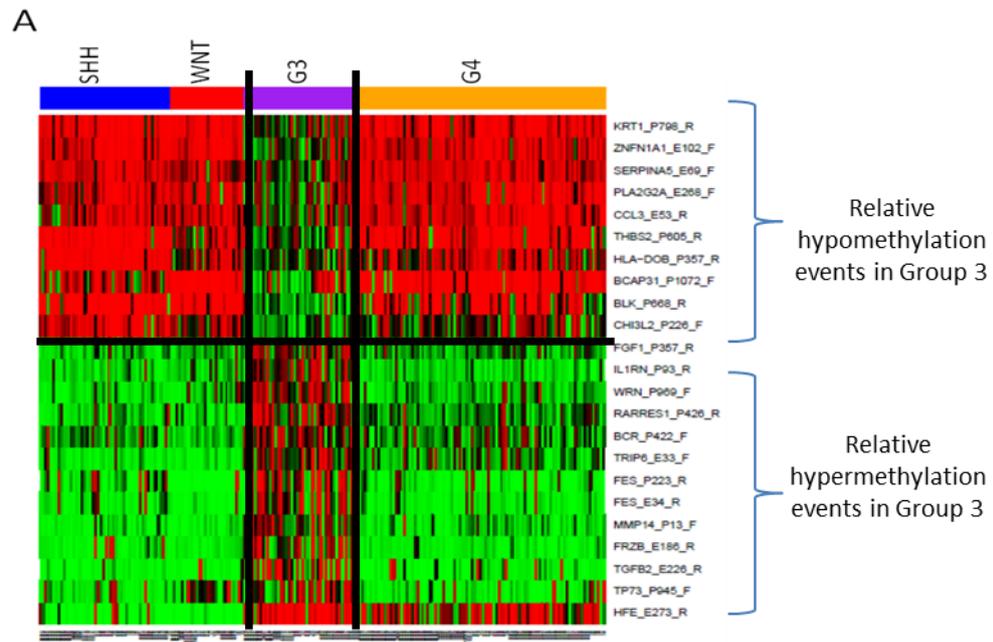
Figure 3.3. Heatmaps illustrating differential methylation patterns for Group 3-specific and Group 4-specific methylation markers. Colour gradients represent the continuous β-values between 0: completely unmethylated (GREEN) and 1: completely methylated (RED). **(A)** Group 3-specific methylation events; the differential pattern of methylation for the 23 CpG sites clearly distinguishes Group3 tumours from all other subgroups. **(B)** Group 4-specific methylation events; 26 of the 31 Group 4-specific CpG sites also significantly distinguish Group (3+4) from SHH and WNT tumours. The differential pattern of methylation for the Group 4 sites clearly illustrates the similarity to Group 3 tumours. Subgroups are represented on heatmaps by different colours: SHH, blue; WNT, red; Group 3(G3), purple; Group 4 (G4), orange.
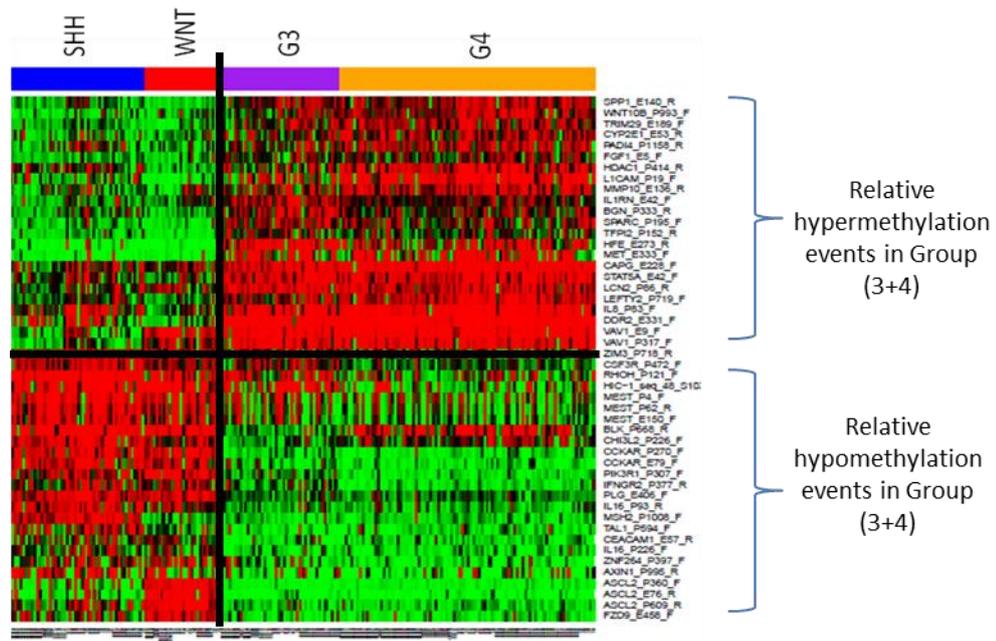
Figure 3.4 Heatmap illustrating differential methylation pattern for Group (3+4)-specific methylation markers. Colour gradients represent the continuous β-values between 0: completely unmethylated (GREEN) and 1: completely methylated (RED).The differential pattern of methylation for the 48 CpG sites clearly distinguishes Group (3+4) tumours from SHH and WNT tumours. Subgroups are represented on heatmaps by different colours: SHH, blue; WNT, red; Group 3(G3), purple; Group 4 (G4), orange.
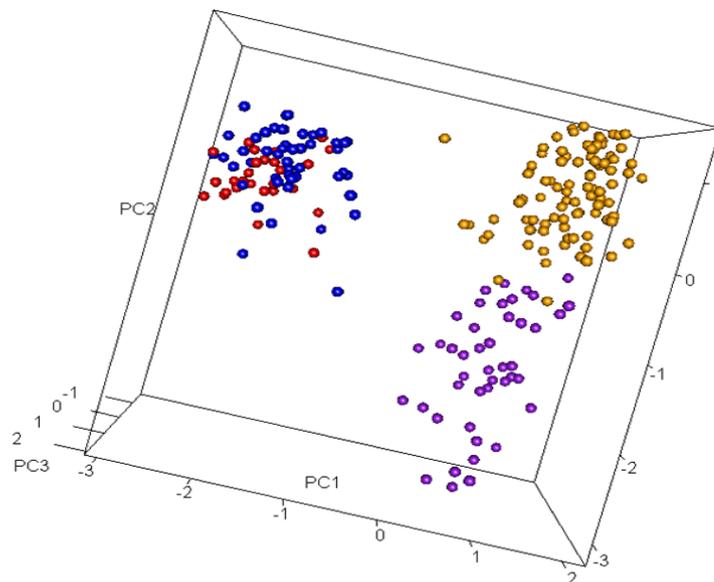
Figure 3.5. 3-D PCA plot of the Group 3 and Group 4 methylation markers. Plot illustrates how the 73 Group 3- and/or Group 4- specific CpG sites effectively distinguish Group 3 (purple) and Group 4 (orange) tumours from the SHH (blue) and WNT (red) tumours. The methylation markers identify Group 3 and Group 4 as distinct subgroups that are more similar to each other than to SHH and WNT.

3.4.2 Tumour specificity of Group 3 and Group 4 methylation markers

The Mann-Whitney U test was next used to assess the tumour specificity of the Group 3-, Group 4- and the Group (3+4)-specific methylation markers as described in section 3.3.4.2.

The majority of Group 3 methylation markers (21/23; 91%) were significantly differentially methylated between Group 3 tumours and non-neoplastic cerebella (Appendix A); 12 CpG sites were hypermethylated and 9 were hypomethylated in Group 3 (Figure 3.6(A)). Less than half of Group 4 methylation markers (13/31; 42%) were significantly differentially methylated between Group 4 and non-neoplastic cerebella (Appendix A); 5 CpG sites were hypermethylated and 8 were hypomethylated in Group 4 (Figure 3.6(B)). Similarly, 42% (20/48) of Group (3+4) methylation markers had significantly different methylation levels in the Group (3+4) tumours compared to the non-neoplastic cerebella (Appendix A); 9 CpG sites were hypermethylated and 11 were hypomethylated in Group (3+4) tumours (Figure 3.6(C)). The tumour-specific markers showed more variable patterns of methylation in the Group 3 and Group 4 primary tumour samples compared with the non-neoplastic cerebellar samples, which showed greater consistency in their methylated and unmethylated states (Figure 3.6).
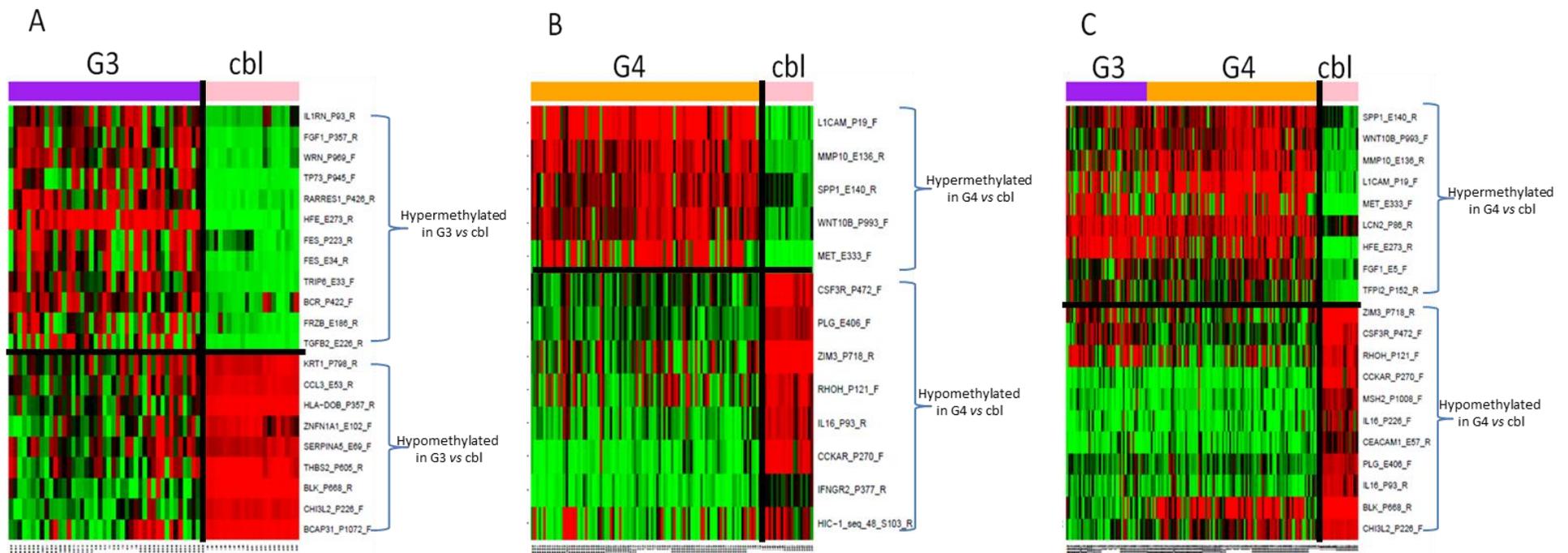
Figure 3.6. Heatmaps illustrating differential methylation patterns for the tumour-specific Group 3 and Group 4 methylation markers. Colour gradients represent the continuous β-values between 0: completely unmethylated (GREEN) and 1: completely methylated (RED). Heatmaps demonstrate the significant differential patterns of methylation observed between primary tumour subgroup and non-neoplastic cerebella (cbl). **(A)** 21/23 Group 3-specific CpG sites are also tumour-specific, encompassing 12 hypermethylated and 9 hypomethylated events. **(B)** 13/31 Group 4-specific CpG sites are also tumour-specific, encompassing 5 hypermethylated and 8 hypomethylated events. **(C)** 20/48 Group (3+4)-specific CpG sites are also tumour-specific, encompassing 9 hypermethylated and 11 hypomethylated events. Tumour-specific markers show less variable methylation patterns in non-neoplastic cerebella (cbl) compared with Group 3 and Group 4 primary tumours. Primary tumour subgroups and non-neoplastic cerebella are represented on heatmaps by different colours: Group 3(G3), purple; Group 4 (G4), orange; non-neoplastic cerebella (cbl), pink.

3.4.3 Differential gene expression across multiple primary cohorts

The 63 genes containing Group 3-, Group 4- and Group (3+4)-specific CpG sites (see section 3.4.1) were next investigated for evidence of equivalent subgroup-specific differential gene expression in 3 independent primary medulloblastoma cohorts, using unpaired *t*-tests as described in section 3.3.5.

In general, across the 3 transcriptomic datasets, the absolute difference in mean $\log_2$ expression levels (mean log ratio) observed for all pair-wise comparisons were small (see Appendix B). A total of 15 out of the 63 genes were significantly differentially expressed in a subgroup-specific manner (Table 3.6). The results can be summarised as follows:

- Group 3-specific events:- of the 22 genes differentially methylated at 1 or more CpG sites, 3 (14%) were differentially expressed in Group 3 tumours.

- Group 4-specific events:- of the 28 genes differentially methylated at 1 or more CpG sites, 8 (29%) were differentially expressed in Group 4 tumours.

- Group (3+4)-specific events:- of the 41 genes differentially methylated at 1 or more CpG sites, 7 (17%) were differentially expressed in Group (3+4) tumours.

3.4.4 Methylation events with a putative effect on gene expression

The patterns of differential methylation and gene expression across the tumour subgroups were assessed using combined boxplots and strip-plots for the 15 genes identified in section 3.4.3. Subgroup-specific methylation-expression relationships were evaluated as described in section 3.3.6 to evaluate potential epigenetic gene regulation by DNA methylation.

There was no evidence of an inverse methylation-expression relationship for the three Group 3-specific genes (Table 3.6). Three Group 4-specific genes, *HDAC1*, *IGF1* and *RHOH*, showed evidence of potential methylation-dependent gene regulation in primary tumours by demonstrating an inverse methylation-expression relationship in 2 or more transcriptomic datasets (Table 3.6). Similarly, three Group (3+4)-specific genes (*BGN*, *DDR2*, *MET*) showed evidence of a direct relationship between CpG methylation and

gene expression (Table 3.6). The patterns of differential methylation and gene expression observed were diverse and gene-specific. The strongest evidence of putative subgroup-specific epigenetic regulation was seen for *HDAC1*, *DDR2* and *MET*. These genes showed >2-fold changes in expression across the 3 transcriptomic datasets consistent with their methylation profiles as described below.

Differential methylation analysis identified the *HDAC1* CpG probe, HDAC1_P414_R, as a Group 4- and a Group (3+4)-specific marker. The magnitude of methylation change (Δβ) and corresponding expression change (mean log ratio) observed were greater for Group 4 *versus* all other groups (Tables 3.4, 3.5 and 3.6). From this point forward *HDAC1* was considered a Group 4-specific gene. The significantly higher methylation of the HDAC1_P414_R probe in the Group 4 tumours correlated with significantly lower gene expression levels in Group 4 (Figure 3.7(A)).

The *DDR2* CpG probe, DDR2_E331_F, was identified as a Group 4-and Group (3+4)-specific marker. The magnitude of methylation change (Δβ) and corresponding expression change (mean log ratio) observed were greater for Group (3+4) *versus* (SHH+WNT) (Tables 3.4, 3.5 and 3.6). From this point forward *DDR2* was considered a Group (3+4)-specific gene. This was reinforced by the strong differential methylation patterns observed across the 4 tumour subgroups (Figure 3.7(B)). The significantly higher methylation status of the DDR2_E331_F probe in the Group (3+4) tumours correlated with significantly lower gene expression in Group (3+4) compared to the (SHH + WNT) tumours (Figure 3.7(B)). The differential pattern of methylation seen across the 4 tumour subgroups was consistent with the differential pattern of gene expression and epigenetic gene regulation, with the greatest corresponding differences observed between the Group (3+4) tumours and the WNT subgroup, while the SHH subgroup of tumours had an intermediate expression and methylation profile (Figure 3.7(B)).

The *MET* CpG probe, MET_E333_F, was a Group 4- and a Group (3+4)-specific methylation marker. The magnitude of methylation change (Δβ) observed was greater for Group (3+4) *versus* (SHH+WNT) (Tables 3.4 and 3.5). Methylation of MET_E333_F showed a strong differential pattern across the 4 tumour subgroups, clearly distinguishing Group (3+4) tumours from (SHH+WNT) tumours (Figure 3.7(C)). Differential expression analysis identified *MET* as significantly differentially

expressed between Group (3+4) and (SHH+WNT) tumours and between Group 4 and all other subgroups (Table 3.6). Assessment of the boxplots, however, showed a strong SHH-specific differential expression profile (Figure 3.7(C)). This result highlights the potential limitations of combining discrete disease subgroups of unequal size and variance together to facilitate a 2 group comparison for differential expression analysis using *t*-tests (see section 3.3.5). The subgroup-specific methylation profile of MET_E333_F was consistent with potential epigenetic transcriptional regulation in the SHH, Group 3 and Group 4 medulloblastomas (Figure3.7 (C)); upregulated expression was observed in SHH tumours consistent with an unmethylated state while Group 3 and Group 4 tumours had significantly higher methylation levels consistent with their transcriptional silencing. The methylation-expression relationship did not extend to WNT tumours. They were unmethylated in a manner similar to SHH tumours but had a silenced expression profile similar to Group 3 and Group 4 tumours, suggesting that methylation-independent mechanisms are responsible for repressing *MET* expression in this distinct subgroup of tumours.

RHOH (Group 4-specific) and *BGN* (Group (3+4)-specific) were differentially expressed in 2 of the 3 transcriptomic datasets and both genes showed evidence of an inverse methylation-expression relationship in the 2 datasets (Table 3.6). The lower methylation of the *RHOH* CpG probe, RHOH_P121_F, in Group 4 tumours correlated with a higher gene expression. The differential gene expression observed for *RHOH* was most significant between Group 4 and (SHH+WNT) tumours consistent with the differential patterns of methylation (Figure 3.8(A)). Higher methylation of the *BGN* CpG probe, BGN_P333_R, in Group (3+4) tumours correlated with lower gene expression which was most significant between Group (3+4) and WNT tumours consistent with its methylation profile (Figure 3.8(B)). *IGF1* (Group 4-specific) was differentially expressed in 3 transcriptomic datasets and showed an inverse methylation-expression relationship in 2 (Table 3.6). Increased methylation of the *IGF1* CpG probe, IGF1_E394_F, in Group 4 correlated with lower *IGF1* expression (Figure 3.8(C)).The magnitude of subgroup-specific expression changes observed for *RHOH*, *BGN* and *IGF1* were not as large as those observed for *HDAC1*, *DDR2* and *MET* (Figures 3.7 and 3.8).

The remaining genes that were differentially methylated and differentially expressed in Group 3 and/or Group 4 tumours did not show evidence of an association between methylation and expression indicative of methylation-dependent gene regulation (Table 3.6). Assessment of the differential patterns of CpG methylation and gene expression across the tumour subgroups thus identified 6 genes (*HDAC1*, *DDR2*, *MET*, *RHOH*, *BGN* and *IGF1*), that showed evidence of potential subgroup-specific epigenetic regulation by DNA methylation, for further analysis.

| Group | Significant differentially methylated CpG probe | CpG island | Chromosome | Gene symbol | Significant differential expression | | | | | | | | | | | |
| | | | | | Kool/Fattet | | | | Northcott | | | | Cho | | | |
					Expression probe	Mean log ratio	Corrected *p*-value	Inverse reln exp	Expression probe	Mean log ratio	Corrected *p*-value	Inverse reln	Expression probe	Mean log ratio	Corrected *p*-value	Inverse reln
Group 3 specific	BCAP31_P1072_F	Y	X	BCAP31	200837_at	0.59	0.002072	N	4026669	0.47	0.001676	N	200837_at	0.76	5.59E-09	N
	CHI3L2_P226_F	N	1	CHI3L2	213060_s	0.37	0.022322	N			NS	-	213060_s_at	0.64	0.001476	N
	FGF1_P357_R	N	5	FGF1	205117_at	0.79	0.022322	N			NS	-	205117_at	0.71	0.003815	N
Group 4 specific	DDR2_E331_F	N	1	DDR2	205168_at	1.08	8.80E-06	Y	2364231	1.03	1.80E-06	Y	205168_at	0.92	2.55E-06	Y
	HDAC1_P414_R	Y	1	HDAC1	201209_at	1.99	2.14E-18	Y	2328868	1.45	9.73E-14	Y	201209_at	2.75	8.22E-21	Y
	IGF1_E394_F	N	12	IGF1	209540_at	1.04	0.000115	Y	3468345	0.30	0.001287	Y	209540_at	0.98	0.003876	N
	LEFTY2_P719_F	N	1	LEFTY2	206012_at	0.58	0.001686	N	2458629	0.81	6.89E-07	Y	206012_at	0.79	4.47E-06	N
	MET_E333_F	Y	7	MET	203510_at	1.19	0.000205	Y	3020343	1.29	2.63E-06	Y	203510_at	2.18	2.45E-09	Y
	RAN_P581_R	Y	12	RAN	200750_s	0.69	3.70E-13	N	3438027	0.33	0.001247	N	200750_s_at	0.74	4.61E-11	N
	RHOH_P121_F	N	4	RHOH	204951_at	0.40	0.035632	Y	2724671	0.27	0.00088	N	204951_at	0.40	0.01436	Y
	TM7SF3_P1068_R	N	12	TM7SF3	217974_at	0.42	0.006573	N	3448481	0.66	9.14E-07	N	217974_at	0.77	1.39E-07	N
Group (3+4) specific	BGN_P333_R	N	X	BGN	201262_s	0.57	0.001578	Y	3995633	0.58	0.001107	Y			NS	-
	DDR2_E331_F	N	1	DDR2	205168_at	1.36	8.80E-06	Y	2364231	1.34	7.33E-06	Y	205168_at	1.70	3.63E-10	Y
	HDAC1_P414_R	Y	1	HDAC1	201209_at	1.28	3.25E-06	Y	2328868	1.0	3.39E-07	Y	201209_at	1.68	9.75E-11	Y
	MEST_E150_F *	Y	7	MEST	202016_at	1.32	8.01E-05	N			NS	-	202016_at	1.50	1.30E-09	N
	MET_E333_F	Y	7	MET	203510_at	1.12	0.009999	Y*	3020343	1.75	2.66E-06	Y*	203510_at	3.13	9.75E-11	Y*
	MSH2_P1008_F	Y	2	MSH2	209421_at	0.40	0.001876	N	2480992	0.34	0.006251	N	209421_at	0.59	1.30E-06	N
	STAT5A_E42_	N	17	STAT5A	203010_at	0.47	2.78E-05	N	3721658	0.51	1.09E-05	N	203010_at	0.67	7.48E-06	N

Table 3.6. Subgroup-specific differential CpG methylation and gene expression. Genes that contained 1 or more significant differentially methylated CpG sites and were also significantly differentially expressed are shown. Subgroup-specificity, significant GoldenGate CpG probe and presence of a CpG island, along with results of the differential expression analysis are detailed; genes with a FDR-adjusted *p*-value <0.05 and a mean log ratio > threshold in at least 2 out of 3 transcriptomic datasets were considered to be significantly differentially expressed. A mean log ratio threshold of 0.3 was applied to the Kool/Fattet and Northcott datasets and a threshold of 0.4 was applied to the Cho dataset. *DDR2*, *HDAC1* and *MET* showed methylation and expression patterns that were specific to Group 4 tumours and to Group (3+4). Genes highlighted in yellow showed a direct inverse relationship (inverse reln) between CpG methylation and gene expression, and this relationship was observed in at least 2 transcriptomic datasets. * only one of three differentially methylated sites is shown. NS=not significant.

CpG methylation   Gene expression

A   *HDAC1*

HDAC1_P414_R   201209_at

B   *DDR2*

DDR2_E331_F   205168_at
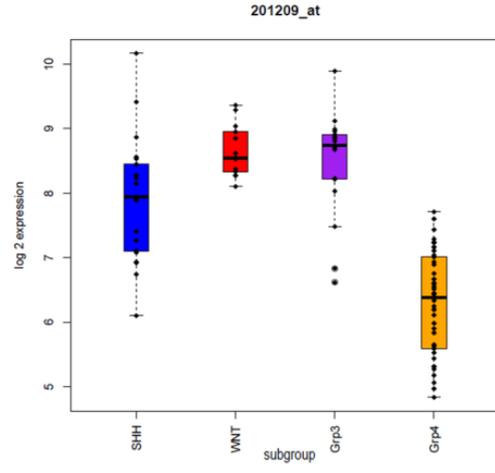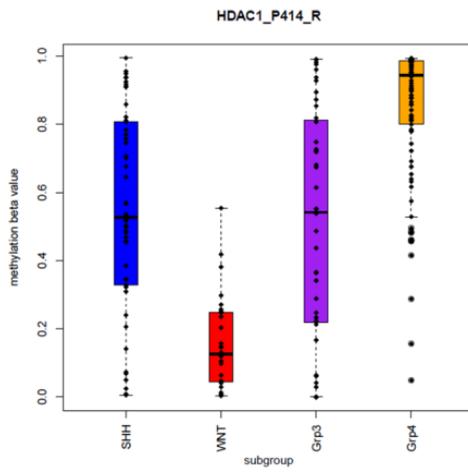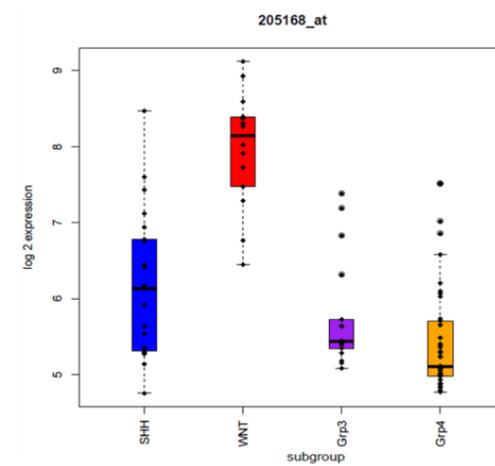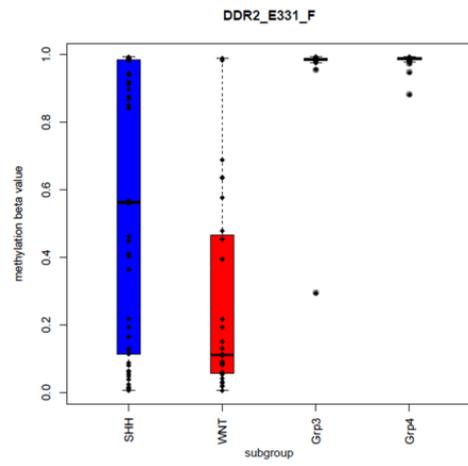
C   *MET*

MET_E333_F   203510_at
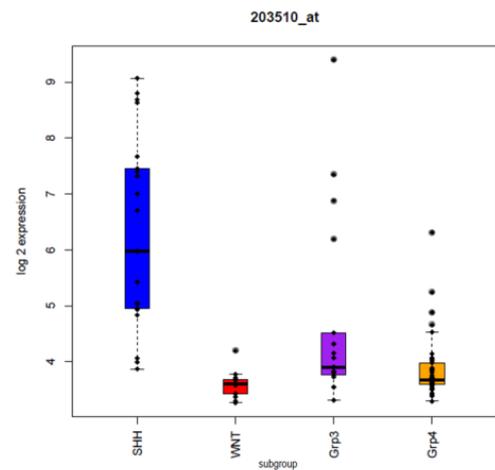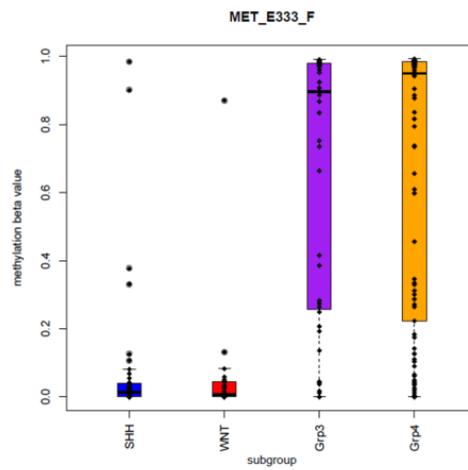
Figure 3.7. Combined boxplots and strip-plots showing differential patterns of CpG methylation and gene expression in primary tumours for *HDAC1*, *DDR2* and *MET*. (**A**) *HDAC1*: significantly higher methylation of HDAC1_P414_R in Group 4 *vs* other (p=9.4E-17) correlates with significantly lower gene expression in Group 4 (p=2.1E-18). (**B**) *DDR2*: significantly higher methylation of DDR2_E331_F in Group (3+4) *vs* other (p=2.9E-18) correlates with significantly lower gene expression in Group (3+4) (p=8.8E-06). (**C**) *MET*: significantly higher methylation of MET_E333_F in Group (3+4) vs other (p=3.2E-21) correlates with significantly lower gene expression in Group (3+4) vs other (p=0.01). Boxplots show that *MET* expression is lower in Group (3+4) compared to the SHH subgroup but not compared to the WNT subgroup. Expression profiles for all genes are representative of the Kool/Fattet dataset. Subgroups are represented on boxplots by different colours: SHH, blue (n=50); WNT, red (n=28); Group 3(Grp3), purple (n=44); Group 4 (Grp4), orange (n=94).

CpG methylation   **Gene expression**

A   *RhoH*
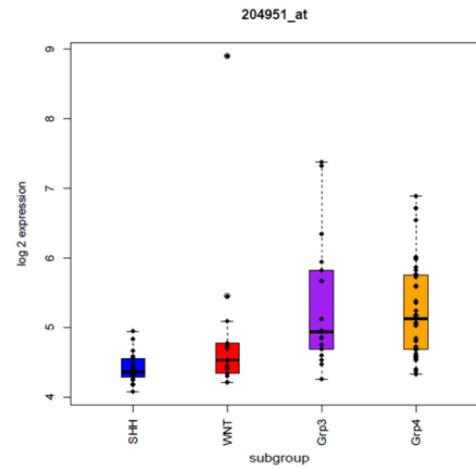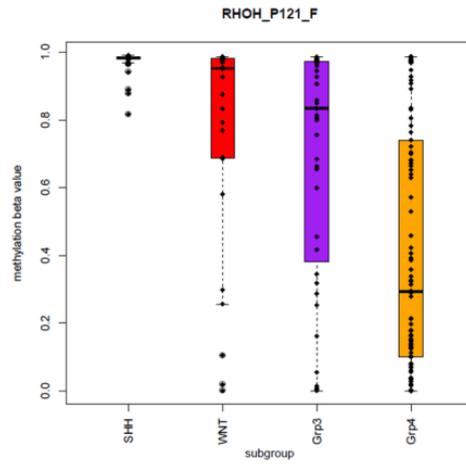
B   *BGN*
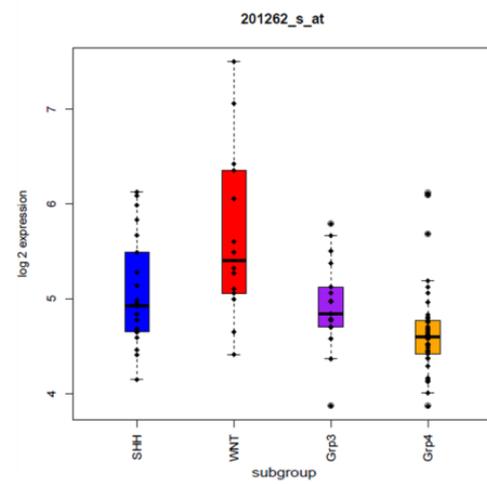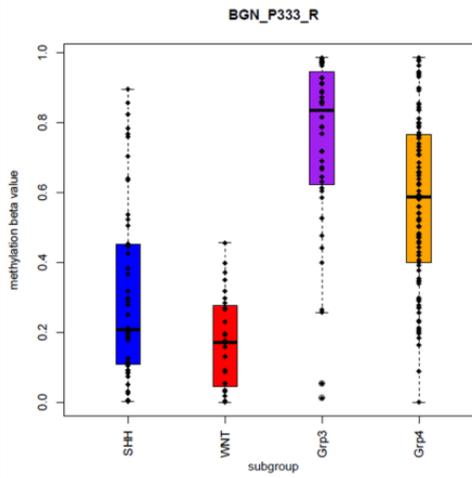
C   *IGF1*
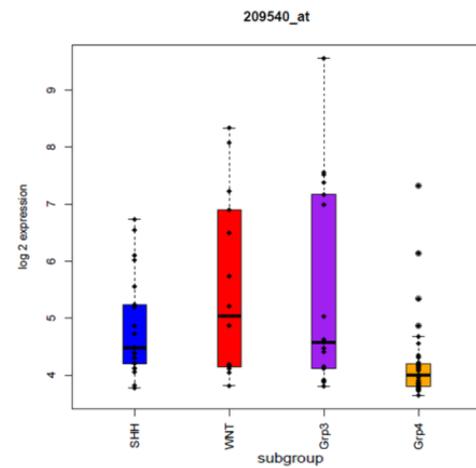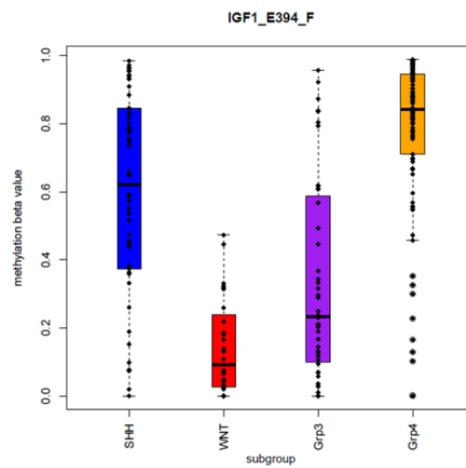
Figure 3.8. Combined boxplots and strip-plots showing differential patterns of CpG methylation and gene expression in primary tumours for *RHOH*, *BGN* and *IGF1*. **(A)** *RHOH* significantly lower methylation of RHOH_P121_F in Group 4 *vs* other (p=9.6E-14) correlates with significantly higher gene expression in Group 4 (p=0.04). Consistent with the methylation profile, gene expression difference is most distinguishable between Group 4 and (SHH+WNT) tumours, with patterns closely related in Group 3 and Group 4. **(B)** *BGN*: significantly higher methylation of BGN_P333_R in Group (3+4) *vs* other (p=3.6E-16) correlates with significantly lower gene expression in Group (3+4) (p=0.002). Consistent with the methylation profile, gene expression difference is most distinguishable between Group (3+4) and WNT tumours. **(C)** *IGF1*: significantly higher methylation of IGF1_E394_F in Group 4 *vs* other (p=1.4E-13) correlates with significantly lower gene expression in Group 4 (p=1.2E-04). Expression profiles for all genes are representative of the Kool/Fattet dataset. Subgroups are represented on boxplots by different colours: SHH, blue (n=50); WNT, red (n=28); Group 3(Grp3), purple (n=44); Group 4 (Grp4), orange (n=94).

3.4.5 Correlation between CpG methylation and gene expression in primary tumours

Methylation and gene expression data were available for 9 matched primary tumour samples. To support the subgroup-specific inverse methylation-expression relationships observed for the 6 genes identified in section 3.4.4, and to determine the strength of the association between CpG methylation and gene expression, Pearson's correlation coefficient (r) was calculated.

HDAC1 and *DDR2* showed strong inverse correlations ($r < (-0.7)$) between the level of gene expression and CpG methylation (Figure 3.9 (A) and (B)) providing further supportive evidence for their methylation-dependent regulation in primary medulloblastomas. *IGF1*, *BGN* and *RHOH* did not show the same strong inverse correlation between CpG methylation and gene expression (data not shown). Pearson correlation r values were (-0.4), (-0.5) and (-0.5) for *IGF1*, *BGN* and *RHOH*, respectively.

The weakest correlation between CpG methylation and gene expression was seen for *MET* ($r= (-0.2)$) (data not shown). Based on the subgroup-specific methylation-expression relationships observed for *MET* (Figure 3.7(C)), a strong inverse correlation would not be expected. *MET* showed evidence of potential epigenetic regulation in SHH, Group 3 and Group 4 tumours but not in WNT tumours. In the WNT subgroup, an unmethylated state did not correlate with upregulated expression as it did in the SHH group (Figure 3.7(C)), suggesting that other methylation-independent gene regulatory mechanisms are repressing transcription in the WNT subgroup of tumours.

A
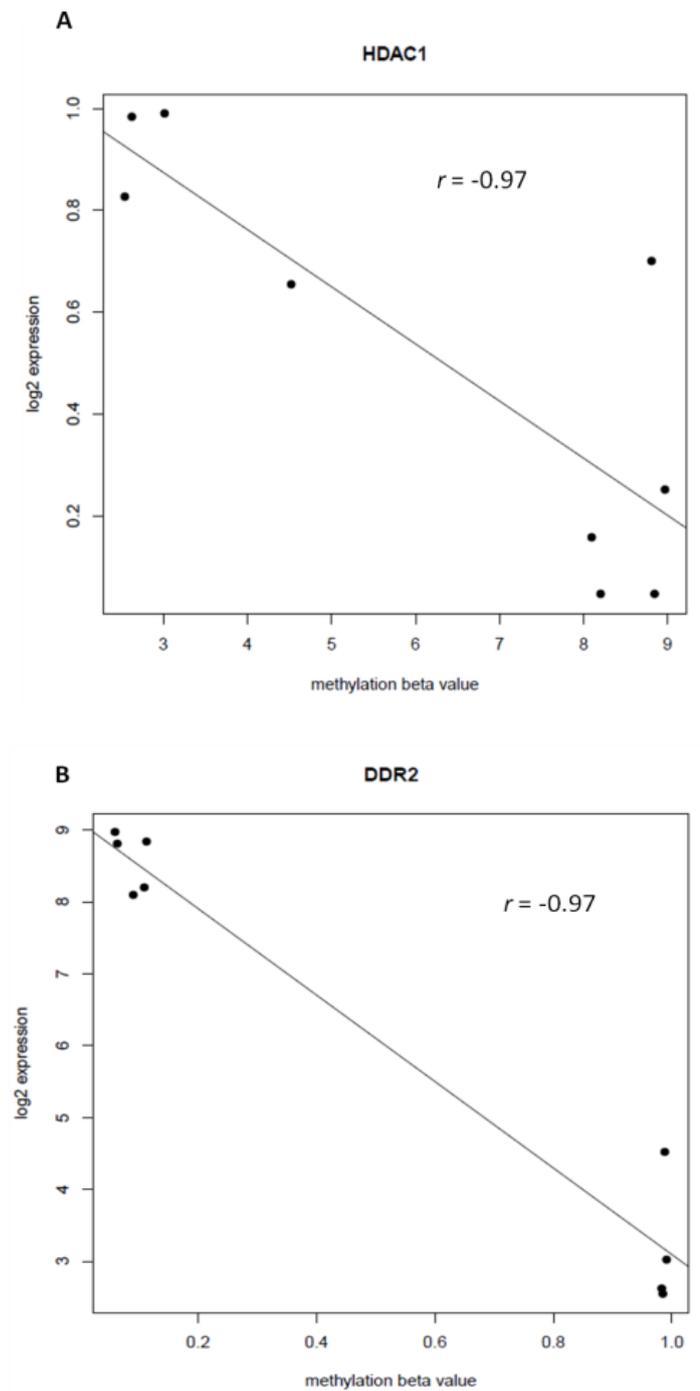
HDAC1

B   ### DDR2

Figure 3.9. Linear correlation between CpG methylation and gene expression for *HDAC1* and *DDR2*. (A) *HDAC1* and (B) *DDR2* show strong inverse correlation between CpG methylation β-values and level of gene expression. Pearson correlation *r* values measuring the strength of the inverse relationship are shown.

3.4.6 In silico gene expression analysis in cell lines treated with 5-azaCdR

The six genes identified in section 3.4.4 were next investigated for evidence of methylation-dependent transcriptional silencing and upregulation by 5-azaCdR in 3 medulloblastoma cell lines (see section 3.3.8). Affymetrix U133A microarray probe signal intensities before and after 5-azaCdR (see section 3.3.3.2) are detailed in Table 3.7.

Genes which showed an increase in expression >1.5-fold following 5-azaCdR treatment were considered upregulated. *DDR2*, *MET* and *RHOH* showed evidence of upregulated gene expression following treatment with 5-azaCdR consistent with a methylated status in at least one cell line. *DDR2* and *MET* were part-methylated/methylated in the 3 cell lines and both genes showed > 1.5-fold increases in expression following 5-azaCdR treatment of D283 (Table 3.7). *RHOH* was methylated in the 3 cell lines and expressed at low levels. RHOH showed increased expression in 2 out of 3 methylated cell lines following treatment, and was considered upregulated in D425, where it showed a 5.2-fold increase in expression after 5-azaCdR treatment (Table 3.7).

HDAC1 showed reduced endogenous expression in the methylated D283 cell line compared to the 2 unmethylated cell lines, but this was not accompanied by upregulated *HDAC1* expression following 5-azaCdR treatment of D283. *IGF1* was unmethylated in D283 and D425 and showed 1.6-fold and 7.1-fold increases in expression following 5-azaCdR treatment, respectively, suggesting that methylation-independent mechanisms were responsible for the expression changes observed. There was no evidence of methylation-dependent re-expression of *BGN* in the cell lines following 5-azaCdR treatment (Table 3.7).

Gene Symbol	Affy U133A probe ID	MED8A					D283					D425			
		5-azaCdR treated signal	Untreated signal	Fold-change in expression after treatment	Methylation status in MED8A		5-azaCdR treated signal	Untreated signal	Fold-change in expression after treatment	Methylation status in D283		5-azaCdR treated signal	Untreated signal	Fold-change in expression after treatment	Methylation status in D425
HDAC1	201209_at	642	602	1.1	U		443	544	0.8	M		927	946	1	U
DDR2	205168_at	602	545	1.1	PM		43	20	2.2	M		92	98	1	M
MET	203510-	16	13	1.2	M		105	66	1.6	M		177	142	1.2	M
RHOH	204951_at	18	14	1.3	M		8	29	0.3	M		31	6	5.2	M
BGN	201262_s_at	82	91	0.9	U		9	16	0.6	PM		17	20	0.9	U
IGF1	209540_at	6	42	0.1	U		47	29	1.6	U		50	7	7.1	U

Table 3.7. Gene expression changes from microarray data following 5-azaCdR treatment of 3 medulloblastoma cell lines. Gene symbol and corresponding Affymetrix U133A microarray probe identifier are shown. For each of the 3 cell lines (MED8A, D283 and D425) the 5-azaCdR treated signal intensity, untreated signal intensity, fold-change in expression signal and methylation status are shown for the 6 genes. Fold-change in expression after 5-azaCdR treatment was calculated by normalising the treated signal to the corresponding untreated signal. Genes with fold-change > 1.5 were considered upregulated by 5-azaCdR. Methylation status describes the GoldenGate β-value scores for the CpG sites identified in section 3.4.4 that have a putative effect on gene expression; U (unmethylated): $\beta<0.3$, PM (part-methylated): $0.3\geq\beta<0.7$, M (methylated): $\beta\geq0.7$.

3.4.7 Subgroup-specific candidate epigenetically regulated genes

Subgroup-specific candidate epigenetically regulated genes were identified according to the criteria defined in section 3.3.9. Based on the strength of evidence for their potential epigenetic regulation, four candidate genes *(HDAC1, DDR2, MET and RHOH)* were identified (Table 3.8). Candidate genes were investigated for evidence of a direct relationship between CpG methylation and gene expression by qRT-PCR assessment of expression changes in medulloblastoma cell lines treated with 5-azaCdR (see section 3.4.8).

BGN and *IGF1* both contained a CpG residue that showed evidence of a subgroup-specific methylation-expression relationship following visual inspection of the boxplots (Figure 3.8). However, neither gene showed a strong inverse linear correlation between CpG methylation and gene expression (see section 3.4.5), or *in silico* evidence of methylation-dependent re-expression by 5-azaCdR treatment in medulloblastoma cell lines (see section 3.4.6). These genes were not considered further (Table 3.8).

Gene	Subgroup-specific (Group 3 and/or Group 4) methylation pattern	Subgroup-specific inverse methylation-expression relationship	Inverse correlation between CpG methylation and gene expression for 9 matched primary tumour samples Pearson correlation $r < (-0.7)$	Methylation-dependent re-expression in ≥ 1 cell line after 5-azaCdR (*in silico* microarray data)	Subgroup-specific candidate epigenetically regulated gene
BGN	Group (3+4)	Y	N	N	N
DDR2	Group (3+4)	Y	Y	N	Y
HDAC1	Group 4	Y	Y	N	Y
IGF1	Group 4	Y	N	N	N
MET	Group (3+4)	Y	N	Y	Y
RHOH	Group 4	Y	N	Y	Y

Table 3.8. Identification of subgroup-specific candidate epigenetically regulated genes. To be selected as candidates for further investigation, genes were required to demonstrate a subgroup-specific inverse methylation-expression relationship and either a strong inverse correlation (r<(-0.7)) between CpG methylation and gene expression for 9 primary tumour samples for which methylation and expression data were both available, or evidence of methylation-dependent re-expression in at least 1 cell line following 5-azaCdR, from *in silico* analysis of microarray data. Subgroup-specificity of differential methylation event(s) is shown for each gene in table.

3.4.8 qRT-PCR assessment of candidate gene expression changes in medulloblastoma cell lines following 5-azaCdR treatment

qRT-PCR was performed for the candidate genes to investigate methylation-dependent expression in medulloblastoma cell lines. Relative gene expression levels were measured before and after treatment with the demethylating agent 5-azaCdR using the delta delta C_T method as described in section 3.3.11.2.

3.4.8.1 Validation of PCR reaction efficiencies

To ensure the validity of the delta delta C_T method, PCR reaction efficiencies for the target and endogenous control genes were assessed as described in section 3.3.11.1. Target and control gene amplification efficiencies were successfully validated. Absolute slope values for the ΔC_T versus log cDNA input plots ranged from 0.02-0.2, demonstrating approximately equal amplification efficiencies for the targets and the endogenous control genes. Figure 3.10 shows the results of the validation experiment for the target gene *MET* and the endogenous control gene *GAPDH*.

The performance of the endogenous control genes was equivalent and consistent in the untreated and 5-azaCdR treated cell lines in all experimental runs. Figure 3.11 shows the equivalent performance of the control genes in the 12 cell line samples for one experimental run. All results presented in section 3.4.9 have been normalised to the *GAPDH* endogenous control gene.
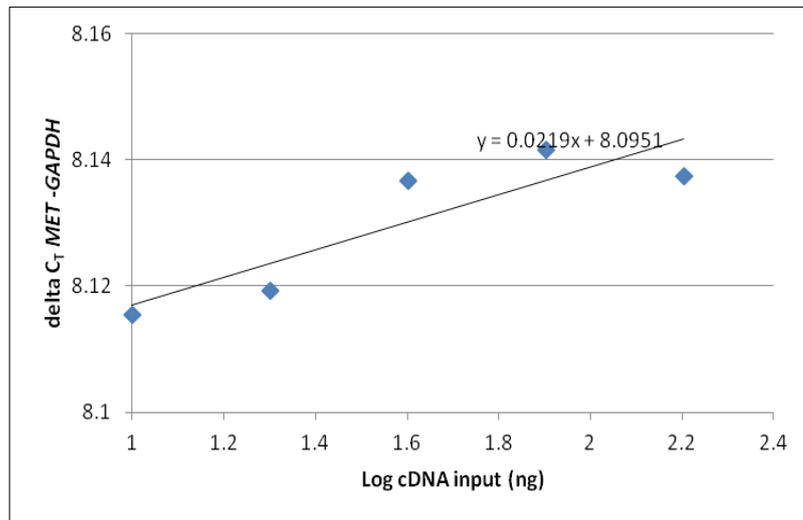
Figure 3.10. Validation of the delta delta C_T method of relative quantification. Difference in C_T value (delta C_T) between target gene MET and endogenous control gene GAPDH over a range of pooled medulloblastoma cell line cDNA concentrations are shown. Plot shows a consistent delta C_T across cDNA input range producing an absolute slope value of 0.02, demonstrating equal reaction efficiencies of target and control gene and validating delta delta CT method of gene quantification. All target genes were validated with an absolute slope value ≤0.2.
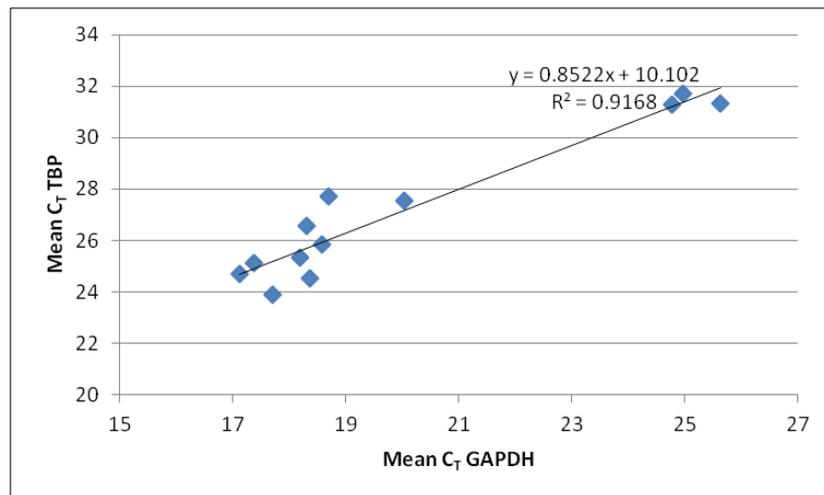
Figure 3.11. Equivalent performance of endogenous control genes in qRT-PCR reactions. Mean C_T values for *TBP* show strong correlation with mean C_T values for *GAPDH* in 12 cell line samples, demonstrating high equivalence in performance of the 2 endogenous control genes in the cell lines. This result was consistently produced in all experimental runs.

3.4.9 qRT-PCR results

Genes were considered upregulated by 5-azaCdR if they showed an expression increase > 2-fold following treatment. Genes that demonstrated evidence of methylation-dependent transcriptional regulation in 4 or more cell lines ($\geq 67\%$) were considered to show strong evidence of epigenetic regulation by DNA methylation.

3.4.9.1 RASSF1A

The isoform A transcript of the *RASSF1* gene was used as a positive control for the demethylating treatment of the cell lines with 5-azaCdr (see section 3.3.10.3).

Increased *RASSF1A* expression $\geq$ 100-fold was observed across the 6 cell lines following 5-azaCdR treatment (Figure 3.12). Hypermethylation of the *RASSF1A* CpG island-associated promoter has previously been established in all medulloblastoma cell lines tested (Lusher *et al.*, 2002). The methylation of *RASSF1A* was consistent with the epigenetic transcriptional regulation seen in cell lines. This result demonstrated the success of the 5-azaCdR treatment in demethylating the cell lines and proved the technical validity and robustness of this method to assess methylation-dependent gene regulation in medulloblastoma cell lines for the selected candidate genes.
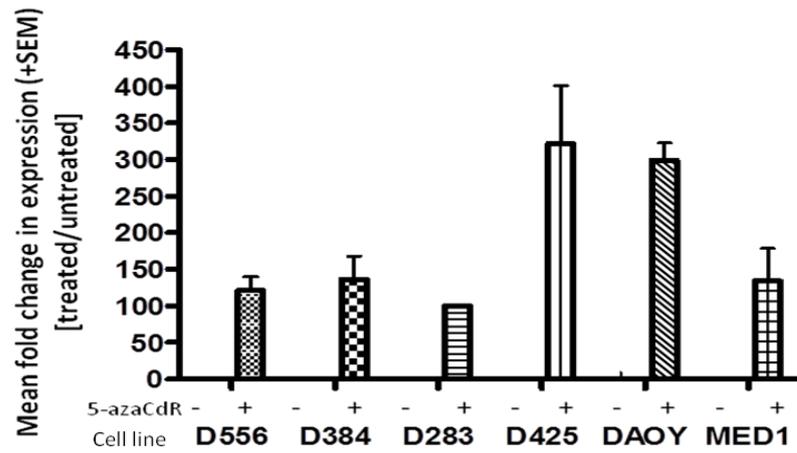
Figure 3.12. Methylation dependent re-expression of *RASSF1A* in medulloblastoma cell lines. *RASSF1A* was re-expressed by 5-azaCdR in the 6 cell lines tested. Greater than 100-fold increase in expression was observed in all cell lines following treatment. *RASSF1A* promoter-associated CpG island is hypermethylated in medulloblastoma cell lines (Lusher *et al*.,2002), consistent with its epigenetic transcriptional regulation observed. Presence or absence of 5-azaCdR is shown as follows: (-) untreated cell line; (+) 5-azaCdR treated cell line. Mean fold change in expression following treatment was calculated by normalising relative mean expression in the treated cell line (+) to relative mean expression in its untreated pair (-).

3.4.9.2 DDR2

DDR2 did not show strong evidence of methylation-dependent re-expression by 5-azaCdR in medulloblastoma cell lines (Figure 3.13). The DDR2_E331_F CpG probe was constitutively methylated ($\beta \geq 0.7$) in the cell lines. With the exception of D283, which showed a 2.2-fold increase in expression following 5-azaCdR treatment, expression remained relatively unchanged after treatment. D425 showed a small decrease in expression following treatment.
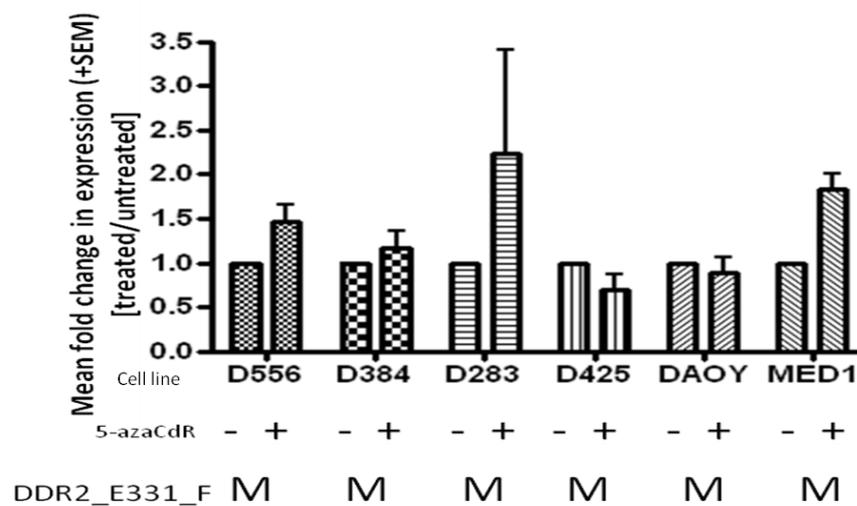
Figure 3.13. *DDR2* **expression is relatively unchanged following demethylation of cell lines. The DDR2_E331_F probe was constitutively methylated in cell lines.** Upregulated *DDR2* expression was observed in one cell line (D283) consistent with a methylated status, while expression remained relatively unchanged in the remaining cell lines. Presence or absence of 5-azaCdR is shown as follows: (-) untreated cell line; (+) 5-azaCdR treated cell line. Mean fold change in expression following treatment was calculated by normalising relative mean expression in the treated cell line (+) to relative mean expression in its untreated pair (-). Methylation status describes the GoldenGate β-value scores for the DDR2_E331_F CpG probe; M (methylated): $\beta \geq 0.7$.

3.4.9.3 HDAC1

No evidence of methylation-dependent transcriptional regulation was found for *HDAC1*. Expression levels remained relatively constant before and after 5-azaCdR irrespective of the methylation status of the HDAC1_P414_R CpG probe (Figure 3.14(A)). The methylated cell line D283 did show an increased expression following treatment but it was not significantly upregulated (>2-fold increase) by 5-azaCdR. There was no strong evidence of transcriptional silencing in methylated cell lines compared with unmethylated cell lines (Figure 3.14(B)). *HDAC1* expression in the methylated D283 and DAOY cell lines was considerably reduced compared with the unmethylated D425 but not compared with other unmethylated cell lines (D556, D384 and MED1) (Figure 3.14(B)).

(A)
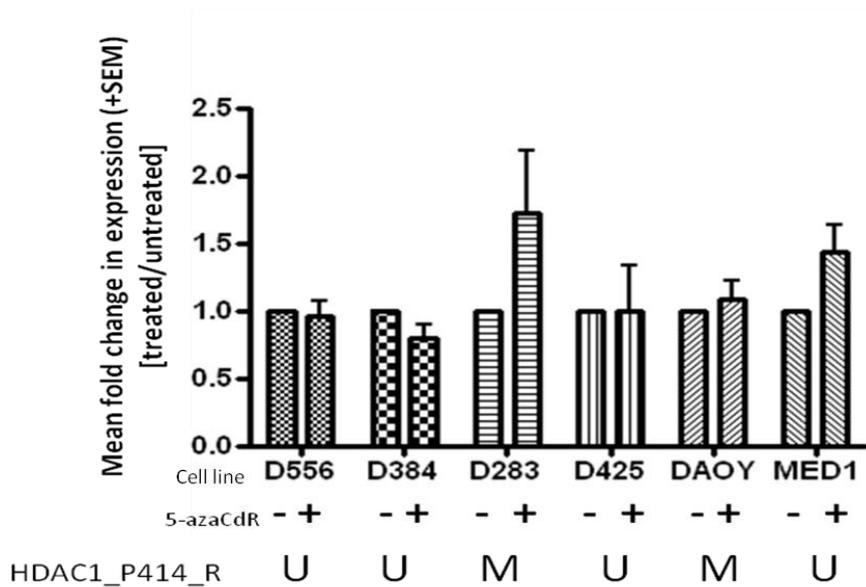
(B)
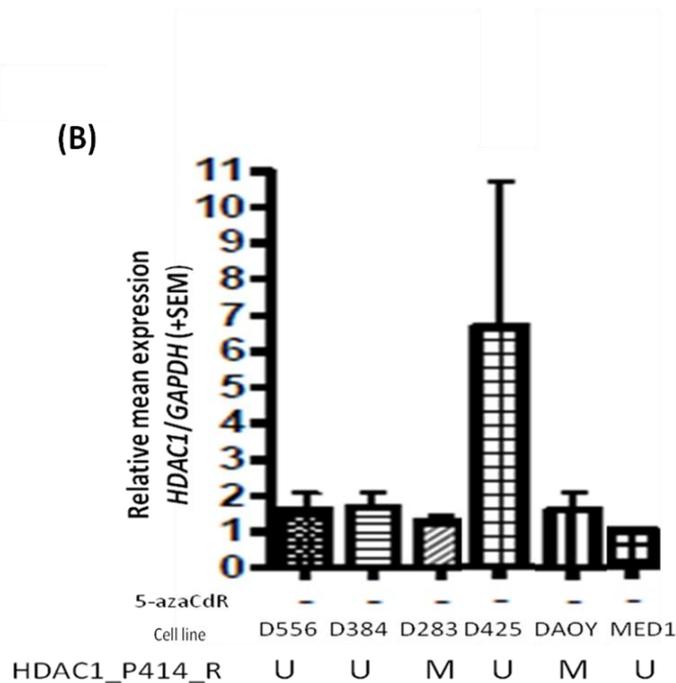
Figure 3.14. *HDAC1* **does not show evidence of methylation-dependent transcriptional regulation in medulloblastoma cell lines**. **(A)** *HDAC1* expression levels were largely unchanged following 5-azaCdR treatment irrespective of methylation status of the HDAC1_P414_R CpG probe in cell lines. Presence or absence of 5-azaCdR is shown as follows: (-) untreated cell line; (+) 5-azaCdR treated cell line. Mean fold change in expression following treatment was calculated by normalising relative mean expression in the treated cell line (+) to relative mean expression in its untreated pair (-). **(B)** Relative mean *HDAC1* expression levels in the untreated cell lines do not show evidence of transcriptional silencing associated with *HDAC1* methylation. Expression was reduced in *HDAC1* methylated cell lines compared with unmethylated D425 but not compared to the 3 other unmethylated lines. Methylation status describes the GoldenGate β-value scores for the HDAC1_P414_R CpG probe; U (unmethylated): β<0.3, M (methylated): β≥0.7.

3.4.9.4 MET

MET showed strong evidence of methylation-dependent transcriptional regulation in medulloblastoma cell lines. The MET_E333_F CpG probe was methylated/part-methylated in 5 out of 6 cell lines. Three out of five *MET* methylated cell lines (D556, D384 and MED1) were transcriptionally silenced compared to the unmethylated DAOY cell line (Figure 3.15(B)) and were upregulated following 5-azaCdR treatment (Figure 3.15 (A)). The unmethylated DAOY cell line showed significantly higher endogenous expression levels compared to methylated lines (Figure 3.15(B)), and did not show upregulation following 5-azaCdR (Figure 3.15(A)). The methylated D283 and D425 cell lines did not show re-expression of *MET* following 5-azaCdR treatment (Figure 3.15(A)), suggesting that methylation-independent mechanisms are involved in *MET* regulation in these cell lines.

The methylation status of the *MET* CpG probe (MET_E333_F) was consistent with its epigenetic transcriptional regulation in 4 out of 6 medulloblastoma cell lines (D556, D384, MED1 and DAOY), providing strong evidence of methylation-dependent expression of the *MET* oncogene in medulloblastoma.
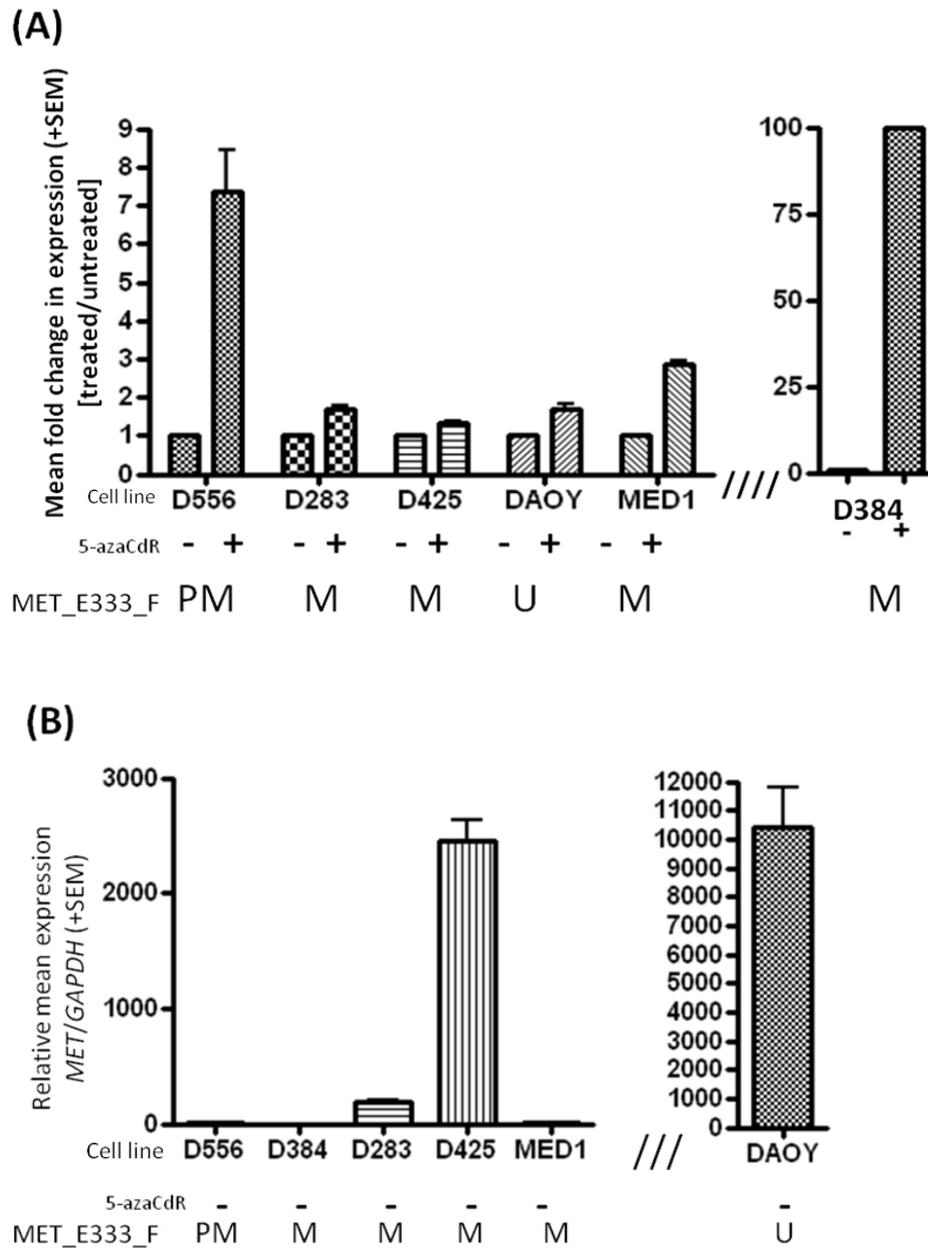
Figure 3.15. *MET* **shows evidence of epigenetic regulation by DNA methylation in medulloblastoma cell lines**. **(A)** Three out of five *MET* methylated/part-methylated cell lines (D556, D384 and MED1) show upregulated *MET* expression following 5-azaCdR treatment. DAOY did not show upregulated expression consistent with its unmethylated status. Presence or absence of 5-azaCdR is shown as follows: (-) untreated cell line; (+) 5-azaCdR treated cell line. Mean fold change in expression following treatment was calculated by normalising relative mean expression in the treated cell line (+) to relative mean expression in its untreated pair (-).**(B)** *MET* is transcriptionally silenced in the methylated cell lines, D556, D384 and MED1, compared to the unmethylated DAOY cell line, which shows significantly higher endogenous *MET* expression. Endogenous expression of *MET* is reduced in the methylated D283 cell line but does not undergo significant upregulation by 5-azaCdR. Methylation-dependent re-expression is not observed in the methylated D425 cell line. Methylation status describes the GoldenGate β-value scores for the MET_E333_F CpG probe; PM (part-methylated): $0.3 \geq \beta < 0.7$, M (methylated): $\beta \geq 0.7$.

3.4.9.5 RHOH

The RHOH_P121_F CpG probe was constitutively methylated ($\beta \geq 0.7$) in the cell lines. Methylation-dependent re-expression following 5-azaCdR was observed in 4 out of the 6 cell lines (Figure 3.16). MED1 showed a 1.9-fold increase in expression after 5-azaCdR, falling just below the threshold (2-fold) to be considered upregulated. *RHOH* expression was reduced in D384 after 5-azaCdR, suggesting that gene expression is not regulated by DNA methylation in this cell line.

Four out of six medulloblastoma cell lines (D556, D283, D425 and DAOY) showed upregulation of *RHOH* following 5-azaCdR treatment consistent with methylation of the RHOH_P121_F CpG probe, and thus providing strong evidence of epigenetic regulation of *RHOH* by DNA methylation in medulloblastoma.
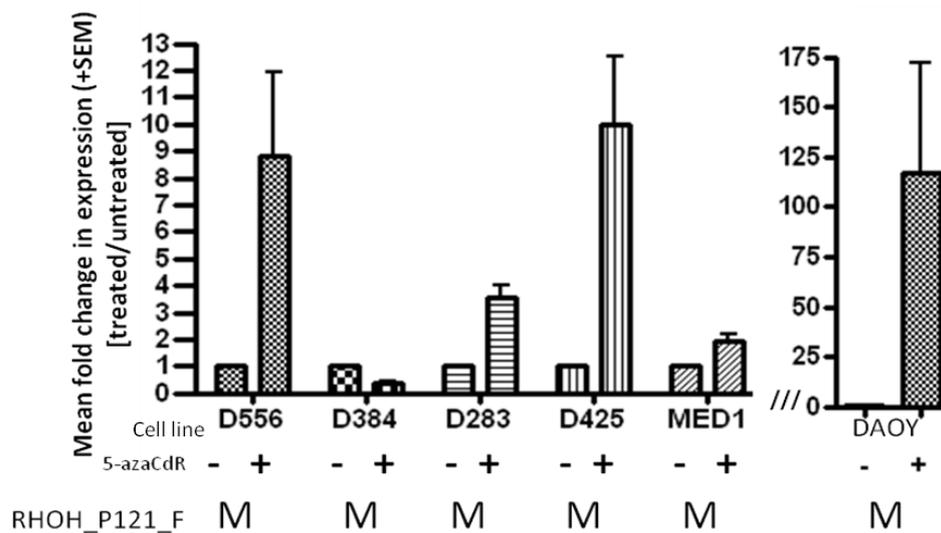
Figure 3.16. *RHOH shows evidence of epigenetic regulation by DNA methylation in medulloblastoma cell lines.* *RHOH* is constitutively methylated in the cell lines. Four cell lines (D556, D283, D425and DAOY) show significant re-expression of *RHOH* by 5-azaCdR, providing strong evidence of methylation-dependent transcriptional regulation. Increased expression following 5-azaCdR treatment is observed in MED1, while D384 shows downregulation of *RHOH* following 5-azaCdR treatment. Presence or absence of 5-azaCdR is shown as follows: (-) untreated cell line; (+) 5-azaCdR treated cell line. Mean fold change in expression following treatment was calculated by normalising relative mean expression in the treated cell line (+) to relative mean expression in its untreated pair (-). Methylation status describes the GoldenGate β-value scores for the RHOH_P121_F CpG probe; M (methylated): $\beta \geq 0.7$.

3.5 Discussion

3.5.1 Novel methylation markers of Group 3 and Group 4 medulloblastomas

The four principal subgroups of medulloblastoma are associated with distinct DNA methylomic profiles (Schwalbe *et al.*, 2013). This study has identified novel CpG methylation biomarkers of the poorly characterised Group 3 and Group 4 tumours. There were 23 Group 3, 31 Group 4 and 48 Group (3+4) methylation biomarkers identified. The results reported in this study strongly support the current opinion that Group 3 and Group 4 tumours are molecularly more similar to each other than they are to SHH and WNT tumours (Taylor *et al.*, 2012) with 13% of Group 3 and 84% of Group 4 markers also discriminating Group (3+4) tumours.

The number of medulloblastomas within each molecular subgroup is not equal (see section 3.3.2) and assessment of boxplots illustrating the subgroup-specific patterns of methylation for the markers identified showed unequal variances between groups. The non-parametric Mann-Whitney U test was applied to identify subgroup-specific methylation markers, and while this test is assumed to be relatively insensitive to unequal sample sizes, the large proportion of Group 4 events that were also Group (3+4)-specific suggest that it was potentially limited in accommodating the unequal sizes and variances that exist between the subgroups, particularly after combining distinct groups together to facilitate a two group comparison. A Kruskal-Wallis test comparing the 4 subgroups, followed by *post hoc* Mann-Whitney U tests comparing individual pair-wise comparisons may have improved sensitivity to identify more Group 4 events that do not also discriminate Group (3+4). Applying a difference threshold to the significance test, in the form of an absolute difference in mean methylation scores, increased the sensitivity of the test to identify true positive subgroup-specific events and the methylation markers identified clearly distinguished the Group 3 and Group 4 tumours from each other and from the SHH and WNT tumours, while also recognising Group 3 and Group 4 as being more similar to each other than to SHH and WNT tumours (Figure 3.5).

The methylation markers were variably located within and outside of CpG islands within gene promoter regions, and for a small number of genes (*n*=6) two or three CpG sites were concordantly differentially methylated in a subgroup-specific manner. The majority of Group 3-markers (91%) were also tumour-specific and, therefore, any

subsequent functional gene regulatory effect identified would signify a significantly aberrant event in this subgroup of tumours. In contrast, less than 50% of Group 4- and Group (3+4)-markers were tumour-specific. The tumour-specific events identified were hypermethylated and hypomethylated in equal numbers compared to normal cerebella. It was important to exercise caution interpreting the results of tumour-specificity because the cellular origin of Group 3 and Group 4 tumours is still unknown and it is likely they have different cellular origins to each other and to SHH and WNT tumours (Kool *et al.*, 2012). It has already been shown that the SHH and WNT subgroups have different cellular origins, with SHH tumours most likely arising from GNP cells of the developing cerebellum (Schuller *et al.*, 2008), while WNT tumours are believed to arise outside the cerebellum from progenitor cells of the dorsal brainstem (Gibson *et al.*, 2010). Until cell of origin can be identified we must consider with caution the utility of non-matched cerebellar samples as controls for a disease that comprises four distinct subgroups with different cellular origins. The control cerebellar samples will include a mix of different cell types and the methylation β value will reflect the average methylation of these mixed cell populations. In this study non-neoplastic cerebellar samples were used as a useful control reference to evaluate potentially aberrant methylation in specific subgroups.

3.5.2 *Differential methylation and gene expression in Group 3 and Group 4 medulloblastomas*

Sixty-three genes were differentially methylated at one or more CpG sites in Group 3 and Group 4 tumours. Fifteen genes (24%) were also differentially expressed in Group 3 and Group 4. Upon assessment of the differential patterns of methylation and expression, three Group 4-specific genes showed an inverse methylation-expression relationship consistent with epigenetic regulation, three Group (3+4)-specific genes showed a relationship consistent with epigenetic regulation while there were no Group 3 genes that showed patterns consistent with epigenetic regulation. These results suggest that the majority of discriminatory methylation markers do not have an effect on gene expression in Group 3 and Group 4 and that the association between DNA methylation and gene expression is weak.

In general, the differential patterns of gene expression were not as strong as the differential patterns of CpG methylation and the differences in gene expression

observed between subgroups were small for many genes. It is important to remember that these genes are not representative of the most differentially expressed genes in medulloblastoma subgroups, but rather they were pre-selected for differential analysis based on them having one or more differentially methylated CpG sites. The small magnitude of expression change does not negate its biological importance, particularly if it is associated with dysregulation of a normal regulatory mechanism, such as DNA methylation, that is known to drive tumourigenesis in its aberrant form.

The patterns of differential methylation and gene expression observed across the tumour subgroups were diverse and gene-specific. The inverse methylation-expression relationships ranged from strong subgroup-specific correlations for *HDAC1* and *DDR2* (Figure 3.7(A) and (B)) to more subtle expression changes inversely correlating with CpG methylation of *RHOH*, *BGN* and *IGF1* (Figure 3.8(A), (B) and (C)). Contributing to the complexity of methylation and gene expression relationships is an unmethylated promoter state that correlates with transcriptional competence but is not indicative of it (Bird, 2002). *MET* was transcriptionally silenced in Group 3 and Group 4 tumours correlating with a methylated state and while an unmethylated state correlated with upregulated expression in SHH tumours this was not true for WNT tumours which were unmethylated and transcriptionally repressed (Figure 3.7(C)). The patterns of methylation and expression observed were consistent with epigenetic regulation in the SHH, Group 3 and Group 4 tumours. Interestingly the lack of a relationship in WNT tumours suggests that different gene regulatory mechanisms may be employed by subgroups of a disease with distinct molecular biology.

BGN is located on the X chromosome and showed an inverse methylation-expression relationship consistent with epigenetic regulation in medulloblastoma subgroups. It is important to consider the sex chromosomes because females who are diploid X carry a methylated silenced X chromosome (Wutz and Gribnau, 2007) while the single X chromosome of males is largely unmethylated. Differential methylation of X-chromosome probes may reflect gender-specific differences rather than a true biological effect and their potential to confound a differential analysis should not be ignored. It is important to note also the loss of one copy of the X chromosome that is found in 80% of females with Group 4 tumours (Taylor *et al.*, 2012) resulting in the loss of methylation in female tumours which can further confound results. Aberrant methylation of X-

linked genes has previously been reported in medulloblastoma and may contribute to tumourigenesis, particularly in male patients (Anderton *et al.*, 2008).

For those genes that were differentially methylated and differentially expressed in Group 3 and Group 4 tumours but showed no evidence of a methylation-expression relationship consistent with epigenetic gene regulation, the patterns of methylation and expression were varied; some genes showed higher levels of methylation correlating with higher levels of expression while for others the patterns were more complex and not consistent across multiple tumour subgroups.

3.5.3 Majority of DNA methylation markers do not show a clear association with gene expression

The results of this study show that the majority of DNA methylation markers specific to Group 3 and Group 4 medulloblastomas do not show a clear association with gene expression. There are a number of possible reasons for this observed weak association between methylation and expression. In model systems DNA methylation has been found to repress transcription in a manner that depends on the location and density of CpGs relative to the gene promoter (Bird, 2002). It is possible that differentially methylated sites identified in this study are too distant from the transcription start site to influence expression or that their methylation status is not representative of the neighbouring region Although bisulfite sequencing has confirmed methylation status consistent with adjacent CpG sites for probes on the GoldenGate array (Bibikova et al., 2006), this may not hold true for all probes. It is also possible that the changes in methylation observed across the tumour subgroups are secondary or 'passenger' events that are not directly involved in driving tumourigenesis but rather accompany it. As previously discussed it is highly probable that the four distinct subgroups of medulloblastoma arise from different cells of origin. Another explanation for the lack of functional relationship observed between DNA methylation and gene expression is that the distinct methylation patterns associated with the disease subgroups may reflect their different cell of origin and developmental stage. A study in breast cancer found that aberrant promoter hypermethylation targeted genes that were already repressed in the normal tissue in a cell lineage-specific manner and concluded that the hypermethylation event did not cause gene silencing with which it was associated but rather marked the cell lineage from which the tumour developed (Sproul *et al.*, 2011). The authors

extended this study further across 7 different cancer types and concluded the same finding that differential methylation profiles reflect the developmental history and transcriptional state of the different cells of origin (Sproul *et al.*, 2012).

Gene transcription is a remarkably intricate process that is tightly regulated at many genetic and epigenetic levels (Lemon and Tjian, 2000). Epigenetic transcriptional regulation relies on the interplay between different mechanisms that control chromatin organisation, including DNA methylation, histone modifications and nucleosome remodelling (Jones and Baylin, 2007; Blackledge and Klose, 2011; Tsai and Baylin, 2011). While there are many reports in the literature of a direct association between DNA methylation and gene expression this does not appear to be true for the majority of DNA methylation markers of Group 3 and Group 4 medulloblastomas. Given the established complexity of epigenetic regulation it would be prudent to investigate epigenetic regulation within the wider context of nucleosome structure and chromatin organisation. Several chromatin modifiers have recently been identified as frequently mutated in medulloblastomas (Parsons *et al.*, 2011; Robinson *et al.*, 2012). Whole genome sequencing of primary tumours identified mutations that were enriched in specific subgroups and found mutations in genes controlling histone methylation targeting Group 3 and Group 4 tumours exclusively (Robinson *et al.*, 2012). Investigations into the deregulation of wider epigenetic mechanisms may yield a greater understanding of any functional role of aberrant DNA methylation in gene regulation and in driving medulloblastoma tumourigenesis.

3.5.4 *Identification of subgroup-specific candidate epigenetically regulated genes in medulloblastoma*

Four genes (*HDAC1*, *DDR2*, *MET* and *RHOH*) were identified as candidate epigenetically regulated genes in primary medulloblastomas. Genes were selected based on the strength of evidence for epigenetic regulation from the data that was available (see Table 3.8). This study was limited by not having gene expression data for all the samples for which there was GoldenGate methylation data. Associations between subgroup-specific patterns of differential methylation and gene expression had to be determined, in large part, by a comparison of means based on the results of the differential methylation and expression analyses of microarray data. Methylation-expression relationships were evaluated from visual inspection of boxplots illustrating

the differential patterns across the tumour subgroups and this was supported by Pearson's correlation analysis of a small sample of 9 tumours for which methylation and expression data were available. Based on the limited tumour-matched data available, *in silico* analysis of microarray gene expression changes by 5-azaCdR in medulloblastoma cell lines was carried out, to further support evidence of a potential functional relationship between methylation and gene expression in medulloblastoma subgroups. Tumour-specificity of the Group 3 and Group 4 methylation markers was considered to assess potential aberrant methylation and the potential oncogenic or tumour suppressor role of candidate genes in medulloblastoma development. Gene expression data for normal cerebellar samples was not available to confirm aberrant gene expression patterns.

HDAC1 showed strong evidence of methylation-dependent transcriptional regulation in primary tumours. Due to the strength of the differential patterns of methylation and expression observed across the tumour subgroups and the strong inverse correlation between methylation and expression, consistent with epigenetic gene regulation, *HDAC1* was selected as a candidate gene. The *HDAC1* CpG site (HDAC1_P414_R) is located in a CpG island-associated promoter and is methylated in Group 4 tumours, correlating with a lower gene expression in this subgroup (Figure 3.7(A)). The methylation status of HDAC1_P414_R in Group 4 tumours was not significantly different to non-neoplastic cerebella (p=0.19, $\Delta\beta$<0.34; Appendix A) and was therefore not tumour-specific for Group 4 tumours, suggesting that hypomethylation of *HDAC1* in SHH, WNT and Group 3 tumours effects upregulated gene expression in these tumour subgroups (Figure 3.7(A)). A literature search did not identify any previously published reports of epigenetic *HDAC1* regulation in cancers.

DDR2 showed a strong methylation-expression relationship in primary tumours and showed evidence of methylation-dependent re-expression following 5-azaCdR in the D283 cell line. The *DDR2* CpG site (DDR2_E331_F) is located in a non-CpG island promoter region and had a significantly higher methylation status in Group (3+4) tumours, correlating with a significantly lower gene expression (Figure 3.7(B)).The methylation status of the CpG site in Group (3+4) tumours was not tumour-specific (p=0.01, $\Delta\beta$<0.34; Appendix A), suggesting that hypomethylation of *DDR2* in SHH and WNT tumours effects upregulated gene expression in these tumour subgroups compared

to the methylated and silenced Group 3 and Group 4 tumours (Figure 3.7(B)). In order to be significantly tumour-specific, methylation markers had to satisfy $p < 0.05$ and $\Delta\beta > 0.34$. *DDR2* has previously been reported to be hypomethylated in a subset of hepatocellular carcinomas and is thought to contribute to tumour progression (Hernandez-Vargas *et al.*, 2010).

MET showed evidence of potential epigenetic gene regulation in SHH, Group 3 and Group 4 tumours. It also showed evidence of methylation-dependent re-expression following 5-azaCdR in the D283 cell line. The MET_E333_F probe is located in a CpG island-associated promoter and was differentially methylated in Group (3+4) tumours compared with (SHH+WNT) tumours (Figure 3.7(C)). The methylation status of MET_E333_F in Group (3+4) tumours was significantly different to non-neoplastic cerebella (p=2.3E-10, $\Delta\beta$>0.34; Appendix A), suggesting that it is aberrantly methylated in Group 3 and Group 4 tumours. Expression of *MET* in SHH tumours was significantly higher compared to Group 3 and Group 4 and this was consistent with a significantly lower methylation in SHH.(Figure 3.7(C)). The methylation-expression relationship did not extend to WNT tumours which were unmethylated and silenced for *MET*. It is important to remember that while an unmethylated state permits gene transcription it is not indicative of it (Bird, 2002) and these results would suggest that methylation-independent mechanisms are responsible for *MET* silencing in WNT tumours.

The proto-oncogene *MET* encodes the tyrosine kinase receptor for the hepatocyte growth factor (HGF) ligand and HGF/MET signalling has been implicated in a variety of cancer types, contributing to tumourigenesis, disease progression and metastasis through its programme of invasive growth signalling (Comoglio *et al.*, 2008). Aberrant MET signalling in medulloblastoma is well documented and it was not surprising to find upregulated *MET* expression specifically in SHH tumours. The role of MET signalling in normal cerebellar development directing the proliferation, migration and invasion of GNP cells (cell of origin of SHH medulloblastomas) has been well established (Ieraci *et al.*, 2002). In mouse models with activated SHH signalling, overexpression of the *MET* receptor ligand HGF in GNP cells increased new medulloblastoma formation compared with SHH activation alone (Binning *et al.*, 2008) and this apparent co-operation between SHH and MET signalling suggests a selective advantage for *MET* overexpression in SHH medulloblastoma formation and

progression. MET dysregulation in tumours is most frequently caused by overexpression of *MET* or its ligand *HGF* (Comoglio *et al.*, 2008) and there are a number of mechanisms reported to cause overexpression of *MET* in medulloblastomas. *MET* amplifications have been identified in a significant proportion of primary medulloblastomas (Kongkham *et al.*, 2010b). Silencing of MET pathway inhibitors has also been implicated in medulloblastoma pathogenesis (Kongkham *et al.*, 2008). While mutations in MET pathway genes have been identified in several cancer types (Comoglio *et al.*, 2008), a recent study found that mutations played a minor role in pathway dysregulation in medulloblastomas (Onvani *et al.*, 2012).

The strong patterns of differential methylation and expression, consistent with epigenetic regulation of *MET* in SHH, Group 3 and Group 4 tumours suggest that the *MET* promoter may potentially be controlled by DNA methylation in these tumour subgroups and that absence of promoter DNA methylation in SHH tumours may cause upregulated *MET* expression independent of amplification and mutation. WNT medulloblastomas arise from a different cell of origin to SHH tumours (Gibson *et al.*, 2010) and it is most likely that Group 3 and Group 4 tumours have different cellular origins (Kool *et al.*, 2012). There may therefore not be the same selective advantage for *MET* overexpression in these tumour subgroups as for SHH tumours. If this is the case the differential pattern of methylation between Group (3+4) tumours and WNT tumours would suggest that biologically distinct subgroups may use different mechanisms to regulate key signalling pathways and gene regulation may be highly subgroup-specific. *MET* was selected as a candidate gene due to its established oncogenic role in medulloblastoma and the strength of evidence for methylation-dependent transcriptional regulation in SHH, Group 3 and Group 4 medulloblastomas.

The *RHOH* CpG site (RhoH_P121_F) is located in a non-CpG island promoter. The significantly lower methylation of RHOH_P121_F in Group 4 tumours correlated with significantly higher gene expression (Figure 3.8(A)). *RHOH* expression profiles for Group 3 and Group 4 tumours were more similar to each other than to SHH and WNT and this was consistent with their methylation profiles and epigenetic regulation. The difference in gene expression observed between the tumour subgroups was more subtle than that observed for *HDAC1*, *DDR2* and *MET*, however, small changes in expression can be biologically important, particularly for low expressed genes. *In silico* analysis of

gene expression microarray data showed > 5-fold increase in *RHOH* expression in the methylated D425 cell line following 5-azaCdR treatment, providing further evidence for epigenetic regulation. The methylation status of the RHOH_P414_R probe in Group 4 tumours was significantly different to non-neoplastic cerebella and the marker was, therefore, tumour-specific (*p*=7.6E-06, Δβ>0.34; Appendix A). The *RHOH* marker was hypomethylated in Group 4 tumours, suggesting a potential oncogenic role for this gene in this subgroup of medulloblastomas. *RHOH* belongs to the family of Rho GTPases that play key roles in intracellular signalling and regulate cellular morphology, motility and proliferation (Williams *et al.*, 2008). *RHOH* is a hematopoietic-specific member of the protein family and overexpression of *RHOH* has been reported in lymphomas, myelomas and leukaemias (Sanchez-Aguilera *et al.*, 2010). Upregulated expression of *RHOH* has been reported to contribute to the initiation and progression of B-cell chronic lymphocytic leukaemia (CLL) and to be associated with an unfavourable prognosis (Sanchez-Aguilera *et al.*, 2010; Troeger *et al.*, 2012)**.**

Most studies have reported on tumour suppressor gene silencing by DNA hypermethylation of CpG island-associated promoters and so it was interesting to note that for 3 of the candidate genes the potential functional relationship was associated with hypomethylation of their respective CpG sites, suggesting a potential oncogenic role in tumour development. One candidate gene, *RHOH*, showed aberrant hypomethylation with potential functional significance in Group 4 tumours. Hypomethylation leading to activation of oncogenes has previously been reported (Feinberg and Vogelstein, 1983a; Hanada *et al.*, 1993). Recently *OCT4*, a transcription factor involved in maintaining self-renewal of embryonic stem cells, has been identified to be upregulated by DNA hypomethylation in gliomas (Shi *et al.*, 2013). The differentially methylated CpG sites identified in this study for *HDAC1* and *MET* were located in CpG island-associated promoters, while the *DDR2* and *RHOH* CpG sites were located in non-CpG island promoters. Evidence suggests that non-CpG island promoters behave in a manner similar to CpG island-associated promoters (Marx *et al.*, 2013; Oster *et al.*, 2013).

3.5.5 MET and RHOH show strong evidence of methylation-dependent expression which may be functionally relevant in subgroups of medulloblastoma

qRT-PCR assessment of expression changes following 5-azaCdR treatment in medulloblastoma cell lines identified a direct relationship between promoter CpG methylation and gene expression for 2 out of the 4 candidate genes analysed (*MET* and *RHOH)*, providing strong evidence of epigenetic regulation by DNA methylation for *MET* and *RHOH* in medulloblastoma (Table 3.9). There was no evidence of methylation-dependent expression of *HDAC1* and *DDR2* in medulloblastoma cell lines, suggesting that the subgroup-specific patterns of methylation and expression observed for these genes are not functionally related.

The result for *RHOH* strongly supports its epigenetic regulation by hypomethylation in Group 4 medulloblastomas and identifies *RHOH* as a novel candidate oncogene in Group 4 medulloblastoma pathogenesis (see section 3.5.6).

The result for *MET* strongly supports DNA methylation-dependent gene regulation as a candidate mechanism of MET pathway regulation/dysregulation in distinct subgroups of medulloblastoma. The functional significance of upregulated *MET* expression in driving medulloblastoma tumourigenesis , progression and metastasis has been demonstrated *in vitro* and *in vivo* (Li *et al.*, 2005; Binning *et al.*, 2008; Kongkham *et al.*, 2008; Provencal *et al.*, 2009). *MET* is an attractive drug target in many cancers and there are a number of mechanisms by which the MET signalling pathway and its downstream effectors can be targeted therapeutically using small molecule inhibitors and blocking antibodies (Gherardi *et al.*, 2012). Medulloblastoma cells have previously shown sensitivity to small molecule inhibitors of MET (Kongkham *et al.*, 2010b). The method of targeting MET signalling will depend on the mechanism of pathway activation. The findings reported here suggest an alternative mechanism of MET pathway activation in SHH medulloblastomas that could act independent of gene amplification and mutation and could help guide choice of targeted therapy. Perhaps the most interesting observation is that the Group 3 and Group 4 tumours appear to require promoter DNA methylation to repress *MET* expression while this is not the case for WNT tumours which are unmethylated and repressed, suggesting that distinct biological disease subgroups potentially use different gene regulatory mechanisms to control gene expression programmes. This may be explained in part by different cells of origin and

different selective pressures acting in concert with different tumour biology to drive tumourigenesis, as observed with MET and SHH signalling in SHH medulloblastoma formation (Binning *et al.*, 2008).

Gene	Subgroup-specific (Group 3 and/or Group 4) methylation pattern	Subgroup-specific inverse methylation-expression relationship	Tumour-specific methylation event	Aberrant methylation state	Methylation-dependent expression in ≥ 4/6 cell lines (qRT-PCR; -/+ 5azaCdR)
DDR2	Group (3+4)	Y	N	Hypomethylatd in SHH + WNT	N
HDAC1	Group 4	Y	N	Hypomethylated in SHH, WNT and Group 3	N
MET	Group (3+4)	In SHH, Group 3 and Group 4 but not WNT	Y	Hypermethylated in Group (3+4)	Y
RHOH	Group 4	Y	Y	Hypomethylated in Group 4	Y

Table 3.9. Summary table of Group 3 and/or Group 4–specific candidate epigenetically regulated genes. Candidate genes showed evidence of an inverse methylation/expression relationship in Group 3 and/or Group 4 tumours. For two candidate genes, *MET* and *RHOH*, their subgroup-specific methylation was also tumour-specific, indicating an aberrant event in Group 3and/or Group 4. Candidate genes were investigated further for evidence of methylation-dependent gene expression in medulloblastoma cell lines by qRT-PCR. Expression was assessed in cell lines before and after treatment with the demethylating agent 5-azaCdR. *MET* and *RHOH* show strong evidence of a direct relationship between CpG methylation and gene expression in cell lines, which may be functionally relevant in the development of distinct subgroups of medulloblastoma.

3.5.6 RHOH – a novel candidate oncogene in Group 4 medulloblastomas

This work has identified a potential functional relationship between DNA methylation and gene expression for the *RHOH* gene in Group 4 medulloblastomas. *RHOH* was found to be hypomethylated in Group 4 tumours consistent with upregulated expression, and suggestive of a potential oncogenic role in Group 4 tumourigenesis (Figure 3.17). The promoter-associated *RHOH* CpG site was unmethylated ($\beta \leq 0.3$) in 51% (48/94) of Group 4 tumours. There were no unmethylated control cerebellar samples (0/21). Aberrant hypomethylation ($\beta \leq 0.3$) was also observed in 21% (9/44) of Group 3 tumours and in 18% (5/28) of WNT tumours, while 100% (50/50) of SHH tumours were methylated ($\beta \geq 0.7$) at this residue (Figure 3.17).

RHOH is located on chromosome 4p13 and is a member of the Rho GTPase family of proteins that function in intracellular signalling and regulate cell proliferation and motility (Williams *et al.*, 2008). Unlike other Rho GTPases, *RHOH* is GTPase-deficient and is dependent on interactions with other GTPases, post-translational modifications and expression levels for cellular activity (Li *et al.*, 2002). *RHOH* is expressed in haematopoietic cells and has been identified to be dysregulated in lymphomas and leukaemias (Sanchez-Aguilera *et al.*, 2010). Upregulated *RHOH* expression has been reported to promote migration of B cells in CLL suggesting a critical role in disease progression (Sanchez-Aguilera *et al.*, 2010; Troeger *et al.*, 2012).
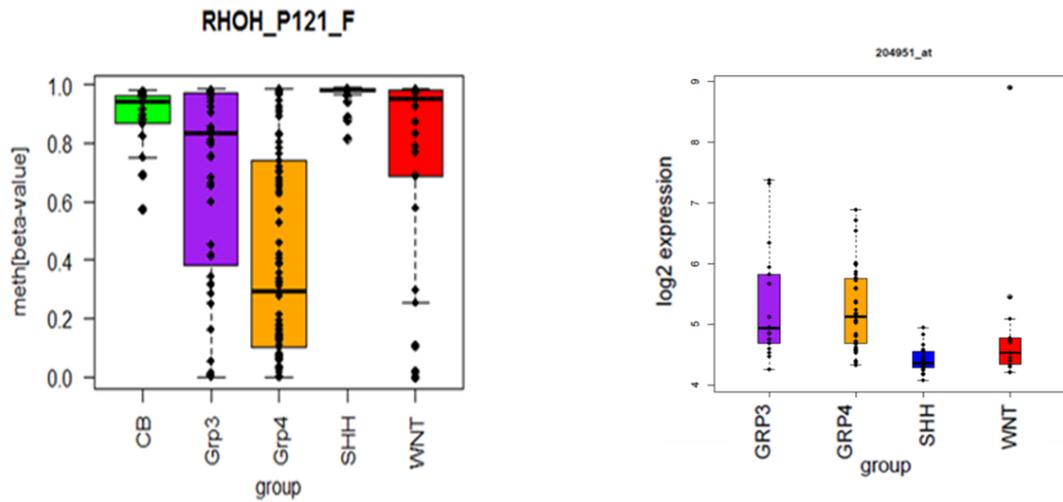
Figure 3.17. **Patterns of CpG methylation and gene expression for *RHOH*.** Combined boxplots and strip-plots illustrate the differential patterns of CpG methylation and gene expression for the *RHOH* promoter-associated CpG residue RHOH_P121_F. Methylation of RHOH_P121_F in control cerebellar samples (CB) is also shown. The CpG residue is hypomethylated in Group 4 medulloblastomas and hypomethylation is associated with significantly higher expression in Group 4 compared with the other tumour subgroups. Primary tumour subgroups and control cerebella (CB) are represented on boxplots by different colours: SHH, blue; WNT, red; Group 3(Grp3), purple; Group 4 (Grp4), orange; control cerebella (CB), green. Expression profiles are representative of the Kool/Fattet transcriptomic dataset. Expression profiles for control cerebella were not available for this study.

3.6 Further work

Investigations reported in this chapter have identified novel DNA methylation markers of the poorly characterised Group 3 and Group 4 medulloblastomas. While the majority of these markers did not show a clear association with gene expression, this work has identified a potential novel candidate oncogene in Group 4 tumours, which may be upregulated by DNA hypomethylation in this tumour subgroup. It will now be essential to validate these findings *in vitro* and *in vivo* and to elucidate any functional role for *RHOH* in Group 4 medulloblastoma pathogenesis. The results reported in this study for *RHOH* are based on the methylation status of a single CpG residue located within the non-CpG island gene promoter. It will be necessary to validate the methylation status of the wider promoter region in primary tumours and in normal cerebella by bisulfite sequencing (see section 1.6.3.3). qRT-PCR will need to be performed in primary tumours and normal cerebella to assess the subgroup-specific expression changes and to validate a functional methylation-expression relationship in primary medulloblastomas, and the upregulated expression consistent with hypomethylation in Group 4 tumours. It will be important to determine whether the increased *RHOH* mRNA levels translate to

elevated protein levels in Group 4 tumours by western blot analysis. To determine any functional role for *RHOH* and perhaps for wider Rho signalling pathways in Group 4 medulloblastomas, proliferation, apoptosis and migration assays will need to be conducted *in vitro* and *in vivo* following modulation of *RHOH* expression. Results from functional assays will reveal any involvement of *RHOH* in the initiation and progression of Group 4 medulloblastomas and identify its potential as a therapeutic drug target within these tumours. There are currently no mouse models of Group 4 medulloblastomas reported. While *RHOH* represents a Group 4-specific gene, the patterns of differential methylation and expression observed are more similar in Group 3 and Group 4 tumours compared to SHH and WNT tumours. Aberrant hypomethylation is also observed in 21% (9/44) of Group 3 tumours (Figure 3.17). Recently mouse models of Group 3 medulloblastomas that express *c-MYC* have been generated (Kawauchi *et al.*, 2012; Pei *et al.*, 2012). A potential functional role for *RHOH* in medulloblastoma tumourigenesis could be investigated using these models of Group 3 tumours.

Further work will be required to validate the methylation status of the wider promoter CpG island of *MET* in primary tumours and in normal cerebella by bisulfite sequencing, and qRT-PCR will be required to validate a functional relationship between methylation and expression in tumour subgroups. It may also be interesting to investigate expression of other MET pathway components in WNT, Group 3 and Group 4 tumours to possibly gain an understanding of why methylation may be required for *MET* silencing in Group 3 and Group 4 tumours but not in the WNT subgroup.

The work in this study did not extend to an investigation into the prognostic relevance of the Group 3 and Group 4 methylation markers identified. Prognostically significant methylation biomarkers have recently been identified in medulloblastoma, which offer the potential for improved disease risk stratification owing to their high clinical utility (Castelo-Branco *et al.*, 2013; Schwalbe *et al.*, 2013) (see section 1.7.2.2). It will be important to assess the independent prognostic relevance of the Group 3 and Group 4 methylation markers across all primary tumours and tumour subgroups. For those markers whose methylation is variable within the subgroup(s), their intra-group prognostic relevance will need to be investigated. Prognostication within a subgroup may identify subsets of tumours which behave differently, and which may benefit from

either less intensive or more aggressive therapies depending on the prognosis conferred. It will be particularly important to investigate the prognostic value of the Group 4-specific *RHOH* methylation marker, as it could potentially represent a prognostic marker with functional relevance.

Alongside an investigation into the potential role of Group 3 and Group 4 methylation events as independent prognostic biomarkers, it will be important to assess their ability to improve the accuracy of survival prediction alongside established prognostic features. Group 3 medulloblastomas are enriched for *MYC* amplification and LCA histology and are frequently metastatic. Metastases and LCA histology are also reported in Group 4 tumours (Kool *et al.*, 2012). These are established prognostic features in medulloblastoma (Ellison *et al.*, 2011b). It will be important to assess the ability of the methylation markers to improve the accuracy and significance of current survival prediction by adding them to a multivariate Cox regression model (Cox, 1972) containing the known risk variables. Developing a more accurate and refined disease risk stratification is essential to improve survival and minimise late adverse treatment effects through the delivery of more personalised therapies (Pizer and Clifford, 2009).

3.7 Summary

This study has identified 2 genes, *MET* and *RHOH*, which show strong evidence for a direct role of DNA methylation in transcriptional regulation within distinct subgroups of medulloblastoma. *RHOH* has been identified as a potential novel oncogene epigenetically regulated by promoter DNA hypomethylation in Group 4 medulloblastomas. Further work must now be carried out to validate epigenetic regulation of *RHOH* in medulloblastoma and to determine any functional role in Group 4 tumourigenesis. This study has also identified a potential candidate mechanism of *MET* regulation and dysregulation by DNA methylation in distinct medulloblastoma subgroups that requires further validation.

This study has been limited by the lack of tumour-matched methylation and expression data available, and methylation-expression associations have had to be determined largely by a comparison of means based on subgroup-specific differential analysis. Tumour-matched data would permit a more accurate statistical correlation on a tumour by tumour basis. Using the methods described in this chapter, the results suggest that DNA methylation may directly regulate relatively few genes in Group 3 and Group 4 medulloblastomas. While a simplistic relationship between DNA methylation and gene expression may not exist for the majority of methylation events identified, transcriptional regulation is highly complex and DNA methylation status alone may not be sufficient to explain gene expression differences.

Investigations in this chapter have focused on a role for promoter DNA methylation of a targeted panel of cancer-related genes. Epigenome-wide studies in cancer are beginning to reveal roles for DNA methylation outside of the gene promoter and outside of CpG islands (Irizarry *et al.*, 2009; Kulis *et al.*, 2013). Using genome-wide approaches, an unbiased assessment of gene-wide DNA methylation offers the potential to gain greater insights into the role of DNA methylation in medulloblastoma development and to identify further epigenetically regulated genes within the subgroups. The genome-wide role of DNA methylation in medulloblastoma has not been widely investigated. Using a functional epigenomics approach, investigations reported in Chapter 4 sought to identify genes whose expression was upregulated in cell lines following demethylating treatment, and to undertake a comprehensive characterisation of DNA methylation

events that may be associated with expression changes, and that may play a key role in the development of distinct subgroups of medulloblastoma.

Chapter 4

Functional epigenomics identifies novel candidate epigenetically regulated genes in distinct subgroups of medulloblastoma

4.1 Introduction

Most studies of cancer DNA methylation have focused on defining aberrant gene promoter and, in particular, CpG island-associated events and their contribution to tumour development. Using the first generation Illumina GoldenGate methylation microarray (Bibikova *et al.*, 2006), investigations reported in Chapter 3 provided a detailed study into the role of promoter DNA methylation in gene regulation for 807 cancer-related genes in the distinct Group 3 and Group 4 subgroups of medulloblastoma. The coverage offered by DNA methylation analysis methods has increased substantially in recent years and it is now possible to study DNA methylation at significantly higher resolution and on a genome-wide scale. Using the latest generation Illumina Infinium 450K methylation array (Bibikova *et al.*, 2011), the investigations reported in this chapter sought to undertake a more comprehensive characterisation of DNA methylation events that may regulate gene expression and that may play a role in the development of distinct subgroups of medulloblastoma.

The Illumina Infinium 450K methylation array provides genome-wide coverage of over 450,000 CpG sites (see section 2.7.3). This high density array interrogates sites located across the complete gene and CpG island regions (Bibikova *et al.*, 2011), permitting an unbiased study into the role of gene-wide DNA methylation in gene regulation. The 450K methylation array targets CpG sites in gene promoters in 200 base pair and 1500 base pair blocks upstream of the transcription start site (designated TSS200 and TSS1500, respectively). The 5'UTR and 3'UTR, first exon and gene body regions are also targeted independently (Bibikova *et al.*, 2011). As the number of epigenome-wide studies increases potential roles for intragenic gene body DNA methylation in gene regulation processes are being recognised. It has been confirmed in human brain samples that intragenic DNA methylation plays a role in controlling alternative promoter usage (Maunakea *et al.*, 2010) and roles in alternative splicing have also been reported (Shukla *et al.*, 2011; Brown *et al.*, 2012). Studies have also demonstrated that a large number of enhancer regions show an inverse correlation between DNA methylation and gene expression levels or enhancer activity (Schmidl *et al.*, 2009; Hoivik *et al.*, 2011; Aran *et al.*, 2013), and of particular interest was the finding that enhancer methylation changes were associated with gene expression alterations while promoter methylation remained unaffected (Aran *et al.*, 2013). In one study in a large

cohort of CLL patients, the strongest correlation between gene expression and DNA methylation was found for intragenic CpGs rather than for those located at 5' regions, and the negative correlation was frequently related to enhancer regions, suggesting that in some cases the methylation status of distal regulatory regions might be more predictive of expression levels than promoter methylation (Kulis *et al.*, 2012).

The 450K methylation array provides enhanced coverage of CpG island regions; 96% of CpG islands are represented on the array as well as CpG island shores and CpG island shelves (Bibikova *et al.*, 2011). It has recently been suggested that many CpG islands are flanked by CpG island shores, which have been defined as regions immediately outside of CpG islands and located within 2kb of the island, where the density of CpG sites has fallen to approximately one tenth the density in the island itself (Irizarry *et al.*, 2009). CpG island shelves are 2kb regions immediately outside of CpG island shores with further reduced CpG density (Bibikova *et al.*, 2011). The importance of DNA methylation in these regions adjacent to CpG islands was first recognised in the colon cancer methylome (Irizarry *et al.*, 2009). Using an epigenome-wide approach, Irizarry et al. reported that most methylation changes in colon cancer occurred not in promoters and not in CpG islands but rather in the CpG island shore regions, and were strongly related to gene expression (see section 1.6.3.1).

The azanucleoside 5'-aza-2'-deoxycytidine (5-azaCdR) is a demethylating drug that incorporates into the genome of actively replicating cells, where it establishes an irreversible covalent bond with DNA methyltransferase enzymes (DNMTs), blocking their methyltransferase function and promoting their subsequent degradation (Stresemann and Lyko, 2008). As a result, DNA methylation marks are lost during DNA replication (Stresemann and Lyko, 2008). Studies have shown that 5-azaCdR induces significant demethylation in human cell lines (Jones and Taylor, 1980). Investigations reported in this chapter utilised a genome-wide microarray-based analysis of gene expression changes induced by the demethylating agent 5-azaCdR to discover candidate epigenetically silenced genes in cancer cell lines that were further investigated for evidence of methylation-dependent regulation in primary tumours. This pharmacological-based approach relies on the functional criterion of transcriptional activation following demethylation and has proved an effective screening method to discover epigenetically regulated genes in cancer cells (Suzuki *et al.*, 2002). Using this

genome-wide functional screening approach several genes have been identified that are silenced by promoter hypermethylation in primary medulloblastomas, including *COL1A2* (Anderton *et al.*, 2008), *SPINT2* (Kongkham *et al.*, 2008) and the *SFRP* family of WNT inhibitors (Kongkham *et al.*, 2010a).

Medulloblastoma comprises four core methylomic subgroups (SHH, WNT, Group 3 and Group 4) that are strongly associated with their transcriptomic counterparts (Schwalbe *et al.*, 2013). It is recognised that these subgroups are distinct molecular entities and thus future clinical trials and treatments should take subgroup status into account (Northcott *et al.*, 2012b; Taylor *et al.*, 2012). It is believed that most human cancers are associated with aberrant gene expression that can be attributed in part to epigenetic alterations (Sadikovic *et al.*, 2008). There are currently approximately 30 genes known to be aberrantly silenced by DNA methylation in the medulloblastoma epigenome (Dubuc *et al.*, 2012). DNA methylation studies in medulloblastoma have previously been restricted to specific promoter-associated loci in a limited number of genes and the functional significance of gene-wide DNA methylation in the pathogenesis of distinct subgroups of medulloblastoma has not been widely investigated. The identification of subgroup-specific functionally important DNA methylation events offers the potential for the development of novel therapeutic targets and disease biomarkers that will facilitate the delivery of more personalised and targeted therapies, with the primary aim of improving survival and reducing the late adverse effects associated with current aggressive radiotherapy and chemotherapy regimens.

The genome-wide role of DNA methylation in the distinct subgroups of medulloblastoma has not been widely investigated. Using a functional epigenomics analysis of methylation-dependent gene expression alterations by 5-azaCdR, combined with a genome-wide analysis of DNA methylation, investigations reported in the chapter sought to undertake a comprehensive characterisation of CpG methylation events that are associated with expression alterations and that may play a role in the pathogenesis of distinct subgroups of medulloblastoma.

4.2 Aims

Using genome-wide expression and methylation microarray technologies, the work reported in this chapter aimed to:

- Identify genes that undergo methylation-dependent transcriptional silencing and re-expression following 5-azaCdR treatment in medulloblastoma cell lines.

- Identify critical CpG residues whose methylation status is consistent with methylation-dependent gene regulation in cell lines

- Assess differential CpG methylation and gene expression in primary medulloblastomas, in a subgroup-specific context.

- Identify subgroup-specific candidate epigenetically regulated genes, defined by those genes which show consistent methylation-dependent gene regulation in cell lines, alongside differential CpG methylation inversely correlated with gene expression in distinct subgroups of primary tumours.

- For candidate genes assess patterns of CpG methylation across genomic regions and evaluate the potential effect of region-wide DNA methylation on gene expression.

4.3 Materials and methods

4.3.1 5'-Aza-2'-deoxycytidine treatment of medulloblastoma cell lines and RNA extraction.

Ten medulloblastoma cell lines (DAOY, D283 Med, D425 Med, D458 Med, D341 Med, D384 Med, D556 Med, MHH-MED-1, MEB-MED-8A and UW228-2) (see section 2.2) were cultured in the presence and absence of the demethylating agent 5'-aza-2'-deoxycytidine (5-azaCdR) as described in section 2.5.5, followed by RNA extraction as described in section 2.6.1. The quality and quantity of RNA were assessed by Bioanalyzer (see section 2.6.3.3).

4.3.2 Re-expression of RASSF1A in medulloblastoma cell lines following 5-azaCdR treatment

To test the effectiveness of 5-azaCdR treatment in cell lines, and prior to sending samples for expression microarray analysis, *RASSF1A* expression was assessed by RT-PCR, using cDNA synthesised from extracted RNA (see section 2.9.3). *RASSF1A* was used as a positive control for the demethylating treatment of the 10 cell lines due to the robust evidence for its methylation-dependent transcriptional silencing and re-expression following 5-azaCdR treatment in multiple cell lines (Lusher *et al.*, 2002). Agarose gel electrophoresis was used to determine the presence and size of PCR products (see section 2.9.3.1).

4.3.3 Whole genome gene expression profiling

4.3.3.1 HumanHT-12 microarray assay

Whole genome gene expression profiles for the 10 untreated and 5-azaCdR-treated cell line pairs were generated on the Human HT12-v4.0 microarray platform (see section 2.8). Extracted cell line RNA was amplified and biotinylated as described in section 2.6.4, and aliquots of 750ng of cRNA at 150ng/µl were sent for microarray gene expression profiling. Microarray analysis was performed at the Wellcome Trust Clinical Research Facility, Edinburgh, UK, according to manufacturer's protocols (Illumina, San Diego, CA, USA). Cell line pairs were processed on the same BeadChip and all samples were processed in a single batch using 2 chips to minimise batch effect. Quality assessment and processing of raw bead level data was carried out by the Wellcome Trust Facility in Edinburgh using the Illumina GenomeStudio Gene Expression Module

(Illumina, San Diego, CA, USA). All categories of Illumina controls performed as expected and no samples failed internal quality control measures. The number of detected genes ($p<0.05$) ranged from 10,000 to 14,000 across the cell line samples. The bead summary data output from GenomeStudio (see section 2.8.2) for target and control probes was kindly provided by the Wellcome Trust Facility. The bead summary data provided the average signal intensity and the detection p-value for each probe. The average standard error associated with bead to bead variability and the average number of beads per probe on each array was also provided.

4.3.3.2 Quality control, array normalisation and data filtering

Bead summary data was analysed in R (v 2.15) using the *beadarray* package (Dunning *et al.*, 2007). The supporting annotation package for the HT12-v4.0 platform (illuminaHumanv4.db) was downloaded into R (v 2.15) from the Bioconductor website (www.bioconductor.org), and facilitated the mapping of Illumina probe sequences to important functional information including chromosome location and gene identification.

Prior to normalisation the data was $\log_2$ transformed to stabilise variance. Boxplots of the average probe signal intensity levels and the average number of beads per probe type were constructed to assess any obvious differences in sample performance within and between BeadChips. Normalisation of high-throughput microarray data is important to correct for variation of non-biological origin across a chip and between chips so that all samples in an experiment are comparable, and it is important to select an appropriate normalisation method to avoid removing important biological effects. The method of quantile normalisation has previously been found to demonstrate favourable performance for high density microarray data (Bolstad *et al.*, 2003). The goal of the quantile method is to make the distribution of intensities for each array in a set the same and given that genes do not change between arrays and intensity distribution between arrays should be the same this method is deemed appropriate (Bolstad *et al.*, 2003). The log2 summarised data were quantile normalised in R (v 2.15).

Transcript probe filtering measures were carried out to improve the power to detect differential gene expression. There were four probe quality categories ('Perfect', 'Good', 'Bad' and 'No match') defined within the Illumina annotation package that

have been shown to correlate with expression level (Dunning *et al.*, 2007). 'Bad' and 'No match' probes were considered non-responding probes and were filtered from the data. This filtering step is reported to be equivalent to removing lowly-expressed probes as well as removing those with high expression caused by non-specific hybridisation (Dunning *et al.*, 2007). It is also common practice in microarray data analysis to remove probes not detected in a proportion of samples. In this study, to account for genes being transcriptionally silenced by DNA methylation and subsequently re-activated following demethylation, probes not detected with a *p*-value <0.05 in at least 1 cell line were filtered out. It was important to consider probes that mapped to the sex chromosomes because females who are diploid X carry a methylated silenced X chromosome (Wutz and Gribnau, 2007), while the single X chromosome of males is largely unmethylated. Probes that mapped to the sex chromosomes were removed to eliminate the potential for gender-specific differences to confound differential analysis. Finally, probes that did not map to an official gene symbol or chromosome location according to Illumina's annotation were removed as they were considered uninformative for downstream analysis.

4.3.4   Genome-wide DNA methylation profiling

4.3.4.1   Infinium 450K methylation microarray

For investigations reported in this chapter the Infinium 450K methylation platform (Bibikova *et al.*, 2011) (see section 2.7.3) was used to assess patterns of CpG methylation in a defined cohort of primary medulloblastomas, in cell lines and in non-neoplastic cerebellar samples. The 450K microarray interrogates the methylation status of over 450,000 CpG sites corresponding to 21,231 genes. The array provides comprehensive CpG coverage across the complete genome with probes distributed across the gene promoter, 5' UTR, 1st exon, gene body and the 3'UTR. The array provides coverage of CpG islands, CpG island shores and island shelves as well as isolated genomic sites. The 450K methylation microarray analysis was performed at the Wellcome Trust Clinical Research facility, Edinburgh, UK, according to manufacturer's protocols (Illumina, San Diego, CA, USA). Raw data processing and normalisation measures were carried out by Dr. Ed. Schwalbe (Newcastle University Paediatric Brain Tumour Group) and final methylation M-value scores were kindly provided by Dr. Schwalbe. The M-value provides a continuous measure of DNA methylation and is

calculated as the $\log_2$ ratio of the intensities of methylated probe versus unmethylated probe (see section 2.7.3.1); M-values close to zero indicate a hemi-methylated state, while positive M-values reflect increasing methylation and negative values reflect decreasing methylation. M-values can be transformed to their corresponding β-values using the "m2beta()" function within the *lumi* package in R. Compared with methylation β-values, M-values have greater homoscedasticity and this allows DNA methylation data to be appropriately analysed by robust statistical methods that are commonly employed for expression microarray analysis (Du *et al.*, 2010).

4.3.4.2 Primary medulloblastoma cohort 2

A representative cohort of 109 primary medulloblastomas (see section 2.1.2) was analysed on the Infinium 450K methylation array and used for investigations reported in this chapter. The cohort included all known medulloblastoma histopathological subtypes and patient ages ranged from 3.6 months to 43 years old. The 4 principal molecular subgroups of medulloblastoma (SHH, WNT, Group 3 and Group 4) were represented in the cohort. Newcastle Research Ethics Committee approval had been obtained for the collection, storage and biological study of all material.

4.3.4.3 Medulloblastoma cell lines

The methylation profiles of 10 medulloblastoma cell lines (DAOY, D283 Med, D425 Med, D458 Med, D341 Med, D384 Med, D556 Med, MHH-MED-1, MEB-MED-8A and UW228-2) (see section 2.2) were assessed on the 450K methylation array. Cell lines were cultured as described in section 2.5.

4.3.4.4 Non-neoplastic cerebellar samples

The methylation profiles of 17 non-neoplastic cerebellar samples (see section 2.3) were assessed on the 450K methylation array. Samples consisted of post-mortem material from patients who died from non-neoplastic conditions and included foetal, infant and adult patient samples.

4.3.4.5 DNA extraction

DNA was extracted from cell lines using a Qiagen DNeasy kit (Qiagen Ltd, Manchester, UK) (section 2.6.2), according to manufacturer's instructions. DNA

extractions from primary tumours and non-neoplastic cerebellar samples were carried out by Dr. Janet Lindsey, Dr. Rebecca Hill and Dr. Chris Howell (Newcastle University Paediatric Brain Tumour Research Group) using TRIzol Reagent (Invitrogen Ltd, Paisley, UK) (see section 2.6.2), according to manufacturer's instructions. Aliquots of 1µg of DNA at 50ng/µl, as measured by Qubit fluorometry (see section 2.6.3.2) were sent for methylation array profiling.

4.3.5   Medulloblastoma methylomic subgroups

The 109 primary tumours were reliably subclassified into the four principal molecular subgroups (SHH, WNT, Group 3 and Group 4) according to their 450K methylation profiles by Dr. Ed Schwalbe (Newcastle University Paediatric Brain Tumour Research Group). Using the same methodologies described for the GoldenGate methylation microarray (Schwalbe *et al.*, 2013), Dr. Schwalbe identified four robust methylomic subgroups that were strongly associated with their transcriptomic counterparts.

The proportion of tumours classified in each methylomic subgroup was comparable to those described for their transcriptomic counterparts (Kool *et al.*, 2012); Group 4 formed the largest group (34% [37 tumours]), followed by SHH (27% [29 tumours]) and Group 3 (27% [30 tumours]). WNT tumours represented the smallest group (12% [13 tumours]) (Table 2.2).

4.3.6   Primary medulloblastoma transcriptomic datasets

The three transcriptomic datasets (Kool/Fattet, Cho and Northcott; see section 2.4) that were used for investigations reported in Chapter 3 were also used for investigations reported in this chapter. Expression profiles and supporting annotation data files were accessed as described in section 2.4. The primary tumours in each dataset were reliably subclassified into the four principal subgroups of medulloblastoma (SHH, WNT, Group 3 and Group 4) by Dr. Dan Williamson (Newcastle University Paediatric Brain Tumour Research Group) using a 4 metagene classifier as described in section 2.4.1.

4.3.7   Gene upregulation following 5-azaCdR treatment of medulloblastoma cell lines

A fold change (FC) estimation approach was used to identify genes whose expression was upregulated following 5-azaCdR treatment of cell lines. Cell lines were treated independently. For each cell line pair the average probe signal intensity (non-log scale)

for the treated cell line was normalised to its untreated counterpart to determine the FC in expression after 5-azaCdR. The mean FC and the standard deviation of the mean were calculated for each cell line and a FC threshold of [meanFC + 3SD] was applied. Genes were considered upregulated by 5-azaCdR treatment and thus showed evidence of potential methylation-dependent transcriptional regulation in cell lines if the FC in expression exceeded the threshold in 2 or more cell lines.

Gene ontology (GO) analysis of the upregulated genes was performed using the open access DAVID Bioinformatics Resource (http://david.abcc.ncifcrf.gov/). The DAVID Functional Annotation Clustering algorithm was applied to the list of Illumina probe identifiers. This algorithm measures the relationships between GO annotation terms and groups them into GO annotation groups, facilitating a more focused biological interpretation of the data (Huang da *et al.*, 2007).

4.3.8 *Identification of critical CpG sites associated with methylation-dependent expression alterations in cell lines*

Genes which were upregulated by 5-azaCdR treatment in medulloblastoma cell lines (see section 4.3.7) were selected for assessment of their methylation status in order to identify CpG sites whose methylation was consistent with methylation-dependent expression alterations observed following 5-azaCdR treatment. Methylation status was assessed using the Infinium 450K methylation array (see section 4.3.4.1). The 450K annotation package (IlluminaHumanMethylation450k.db) was downloaded from the Bioconductor website (www.bioconductor.org) into R (v 2.15). The package contained functional annotation data assigned by Illumina for all probes on the 450K array and facilitated the mapping of probes to important information including chromosome location and gene identification as well as genomic location (promoter, body, UTRs) and CpG density of neighbouring region (island, shore, isolated site).

Methylation and expression probes were paired on the basis of their official gene symbol identifiers as annotated by Illumina in the relevant annotation packages. Methylation M-values were converted to their corresponding β-values using the *lumi* package which was downloaded from the Bioconductor website into R. All CpG methylation probes for each upregulated gene were initially extracted from the cell line 450K methylation data. Probes with β-values ≤ 0.3 were considered unmethylated and those with β-values >0.3 were considered methylated. Probes that were unmethylated in

all 10 cell lines (constitutively unmethylated) were removed from the analysis as these were deemed uninformative and would not be associated with any upregulated expression observed in the cell lines following 5-azaCdR treatment. The remaining CpG probes were investigated further for evidence of a critical role in epigenetic gene regulation.

4.3.8.1 *Methylation-dependent re-expression in cell lines*

Following the removal of constitutively unmethylated probes the remaining CpG probes were considered constitutively methylated or variably methylated as follows:

- Constitutively methylated: β-value >0.3 in all 10 cell lines.
- Variably methylated: variable β-values ranging from 0 (completely unmethylated) to 1 (completely methylated) across the 10 cell lines.

To determine whether methylation status was consistent with methylation-dependent re-expression in cell lines following 5-azaCdR treatment, the fold changes in expression observed across the 10 cell lines for the upregulated genes were converted into a matrix of 0's and 1's in R. Fold changes greater than the cell line threshold were assigned the nominal value 1 (upregulated), while those not exceeding the threshold were assigned the nominal value 0 (not upregulated). Similarly, the methylation β-values for the corresponding CpG probes in the 10 cell lines were converted into a matrix of 0's (unmethylated; β ≤0.3) and 1's (methylated, β>0.3). The corresponding gene-specific rows and cell line-specific columns of the 2 matrices were added together to give a results matrix containing the values 0, 1 and 2 which were interpreted as a "consistent relationship" or "inconsistent relationship" as follows:

- Results value 0: "consistent relationship"; CpG site was unmethylated and gene not upregulated by 5azaCdR.
- Results value 2: "consistent relationship"; CpG site was methylated and gene upregulated by 5azaCdR.
- Results value 1: "inconsistent relationship"; CpG site was unmethylated and gene was upregulated OR CpG site was methylated and gene was not upregulated by 5azaCdR.

CpG sites were considered to have a methylation status consistent with methylation-dependent gene re-expression following 5-azaCdR treatment if the following criteria were met:

- Constitutively methylated probes: a "consistent relationship" (results value 2) was observed in ≥ 60% cell lines
- Variably methylated probes: a "consistent relationship" (results value 0 or 2) was observed in ≥ 60% cell lines AND ≥ 60% methylated cell lines were upregulated (results value 2).

Bar graphs were constructed in R (v 2.15) to illustrate methylation-dependent gene re-expression following 5-azaCdR treatment in cell lines.

4.3.8.2 Correlation between CpG methylation and gene expression in cell lines

The variably methylated CpG sites whose methylation status was consistent with methylation-dependent re-expression observed in cell lines (see section 4.3.8.1) were further investigated for evidence of a significant inverse methylation-expression relationship. The association between CpG methylation and endogenous gene expression in the 10 untreated cell lines was assessed by calculating Pearson's product-moment correlation coefficient r and associated p-value. In order to control for the large number of correlation tests being performed raw p-values were corrected using the Benjamini-Hochberg false discovery rate (FDR) correction (Benjamini and Hochberg, 1995). The level of endogenous gene expression was considered to be significantly inversely correlated with CpG methylation if the coefficient r value was $< (-0.7)$ and the adjusted p-value was <0.05. Linear regression plots were constructed in R (v 2.15) to illustrate significant inverse methylation-expression relationships in cell lines.

4.3.9 Identification of candidate epigenetically regulated genes in medulloblastoma cell lines

Genes that corresponded to the constitutively methylated probes whose methylation was consistent with methylation-dependent re-expression in at least 60% of cell lines following 5-azaCdR treatment (see section 4.3.8.1) were selected as candidate epigenetically regulated genes. Similarly, genes that corresponded to the variably methylated probes whose methylation status was consistent with expression alterations as described in section 4.3.8.1 and whose methylation showed a significant inverse

correlation with endogenous gene expression in cell lines (see section 4.3.8.2) were selected as candidate epigenetically regulated genes.

4.3.10 Linear models for microarray data: *limma*

The candidate epigenetically regulated genes which were identified in cell lines (see section 4.3.9) were investigated for evidence of significant differential CpG methylation and associated differential gene expression in primary medulloblastoma subgroups using the R package *limma* (Smyth, 2004). The output statistics for significance analysis using *limma* are the moderated *t*-statistic and the moderated *F*-statistic, which have the same interpretation as the ordinary *t*-statistic and *F*-statistic in conventional *t*-tests and ANOVA, except that *limma* uses a different method of variance estimation (Smyth, 2004). *Limma* uses a linear models approach to the analysis of microarray experiments and after fitting a linear model to each probe in the data it uses an empirical Bayes method to shrink probe-wise sample variances towards a common value and increase degrees of freedom for individual variances (Smyth, 2004). The effectiveness of *limma* in differential expression analysis has previously been demonstrated (Irizarry, 2005), and it has been found to be superior to the *t*-test and ANOVA, providing an increase in power and improvement in the false-positive rate (Jeanmougin *et al.*, 2010).

A further advantage of *limma* is its capability to assess differences in a multifactorial designed experiment where multiple contrasts between targets can be analysed simultaneously. For investigations reported in this chapter all pair-wise comparisons (SHH*v*WNT, SHH*v*Grp3, SHH*v*Grp4, WNT*v*Grp3, WNT*v*Grp4, Grp3*v*Grp4) could be tested simultaneously and subsequent subgroup-specificity determined (see sections 4.3.11 and 4.3.12). *Limma* required 2 matrices to be specified; a design matrix that identified the subgroup of each sample in the data (SHH, WNT, Grp3 or Grp4) and a contrast matrix that specified the pair-wise comparisons to be made. *Limma* provided two functions that performed the significance tests and adjusted *p*-values for multiple testing. The "topTable()" function computed the moderated *t*-statistic for each probe and contrast, testing each individual contrast to 0, alongside a moderated *F*-statistic testing whether all the contrasts were 0. *P*-values were adjusted for multiple testing using the Benjamini-Hochberg false discovery rate correction (Benjamini and Hochberg, 1995). *P*-values were adjusted for the number of probes tested as well as the number of contrasts and all tests were treated as equivalent regardless of which probe or contrast

they related to, facilitating a comparison of the number of differential events between contrasts to be made. The "topTable()" output also included a $\log_2$ fold change (logFC) value. The logFC represented the value of the contrast. LogFC thresholds (difference thresholds) were applied to the differential methylation analysis (see section 4.3.11) and to the differential expression analysis (see section 4.3.12). The "decideTests()" function produced a numeric matrix with the elements -1, 0 or 1 depending on whether each *t*-statistic was classified as significantly negative, not significant or significantly positive, respectively, according to the adjusted *p*-value and logFC thresholds imposed.

4.3.10.1 Differential methylation in primary tumours using linear models

Limma is a parametric statistical model that was designed for gene expression microarray data analysis, where the assumption of homoscedasticity used by this model is appropriately made (Jeanmougin *et al.*, 2010). The M-value method of measuring methylation with Infinium arrays (see section 4.3.4.1) has been recommended for differential methylation analysis due to its approximate homoscedasticity, making it statistically valid to use methods such as *limma* (Du *et al.*, 2010). In comparison, the β-value measure of methylation (see section 2.7.3.1) has a finite range from 0 to 1 and shows severe heteroscedasticity at high and low levels of methylation due to the compressed nature of the data, making it inappropriate to use *limma* for differential analysis using β-values (Du *et al.*, 2010).

4.3.11 Identification of subgroup-specific differentially methylated CpG sites

Candidate genes with strong evidence of methylation-dependent transcriptional regulation in medulloblastoma cell lines (see section 4.3.9) were selected for differential methylation analysis in primary tumours and non-neoplastic cerebellar samples. The critical CpG sites corresponding to the candidate genes were investigated for evidence of subgroup-specific and tumour-specific differential methylation using *limma*.

SHH, WNT, Group 3 and Group 4 tumours and non-neoplastic cerebellar samples were identified in the design matrix. The contrast matrix included all pair-wise contrasts. An example of the pair-wise contrasts tested to determine SHH-specific events were as follows:

- SHH *v* WNT; SHH *v* Group 3; SHH *v* Group 4; SHH *v* cerebella

Alongside single group contrasts, paired group contrasts were tested as follows:

- (SHH+WNT) *v* (G3+G4); (SHH+G3) *v* (WNT+G4); (WNT+G3) *v* (SHH+G4); (G3+G4) *v* cerebella

These contrasts were tested to identify events specific to Group (3+4) tumours. As previously discussed in Chapter 3, although Group 3 and Group 4 are distinct subgroups they are likely more similar to each other than to SHH or WNT tumours (Taylor *et al.*, 2012).

Contrasts were required to satisfy a FDR-adjusted *p*-value <0.05 and a logFC (absolute difference in mean M-value scores) > 0.9 to be judged significant. A mean M-value difference threshold within the range 0.4 and 1.4 is recommended for differential methylation analysis (Du *et al.*, 2010). Du et al. demonstrated that thresholds within this range are optimum as they can increase the true positive rate without deteriorating the detection rate of the significance test. A threshold of 0.9 was selected as it is midway in the recommended range.

The "decideTests()" function of *limma* (see section 4.3.10) was applied to classify results as significantly negative, not significant or significantly positive using the elements -1, 0 and 1 respectively. Individual subgroup-specificity was determined by overlapping the results for the appropriate individual contrasts and selecting only those events that were significant (-1 or 1) in all selected contrasts. This was carried out in R. As an example, SHH-specific events were identified as those events that were significant (-1 or 1) in the contrasts SHH *v* WNT, SHH *v* G3 and SHH *v* G4. Group (3+4)-specific events were identified as those events that were significant (-1 or 1) in the (SHH+WNT) *v* (G3+G4) contrast but were not significant (0) in the (SHH+G3) *v* (WNT+G4) and the (WNT+G3) *v* (SHH+G4) contrasts.

To further increase the sensitivity of the significance test and to ensure that only biologically relevant probes were selected, an additional filter of an absolute difference in mean β-value (Δβ) greater than 0.2 was applied. This value was chosen as it is the limit of detection of the Infinium 450K assay (Bibikova *et al.*, 2011). This final filter was applied after subgroup specificity was determined using limma.

As an example, SHH-specific events were required to satisfy the following Δβ filter:

- |mean β-value (SHH) – mean β-value (WNT+G3+G4)| > 0.2

Contrasts assessing tumour subgroup methylation against non-neoplastic cerebella methylation were tested in the contrast matrix. Similar to the investigations reported in Chapter 3, the difference in methylation between the subgroup-specific event and the non-neoplastic cerebella (tumour-specificity) was used as a reference to appraise the potential functional significance of DNA methylation of candidate genes in medulloblastoma subgroups.

4.3.12 Identification of subgroup-specific differentially expressed genes

Candidate genes with strong evidence of methylation-dependent transcriptional regulation in medulloblastoma cell lines (see section 4.3.9) were selected for differential expression analysis in three independent primary tumour datasets (see section 4.3.6) using *limma*. Transcript probe identifiers were selected as previously described in section 3.3.5.

SHH, WNT, Group 3 and Group 4 tumours were identified in the design matrix. The contrast matrix was identical to the one used for differential methylation analysis (see section 4.3.11) with the exception that non-neoplastic cerebella were not included. Expression profiles for non-neoplastic cerebella were not available for this study.

An individual logFC (absolute difference in mean log2 expression values) threshold was applied to each transcriptomic dataset. The threshold applied was the mean logFC for all the contrasts and probes tested in that dataset, and was calculated in R from the limma "topTable()" output (see section 4.3.10). The following logFC thresholds were applied: Kool/Fattet dataset, 0.5; Cho dataset, 0.7; Northcott dataset, 0.4. Contrasts were required to satisfy a FDR-adjusted *p*-value <0.05 and a logFC > threshold to be judged significant. The "decideTests()" function (see section 4.3.10) was applied to classify results as significantly negative, not significant or significantly positive using the elements -1, 0 and 1 respectively. Subgroup specificity was determined as described in section 4.3.11.

4.3.13 Identification of subgroup-specific candidate epigenetically regulated genes

The significance results for genes that were differentially methylated and differentially expressed in the same primary tumour subgroup were compared for evidence of an inverse methylation-expression relationship. Genes were considered to show evidence of a relationship consistent with potential methylation-dependent regulation if a significantly negative (-1) subgroup-specific methylation profile correlated with a significantly positive (1) subgroup-specific expression profile or vice versa. The inverse methylation-expression relationship had to be observed in at least 2 transcriptomic datasets. Combined boxplots and strip-plots were constructed in R and provided a visual representation of the differential patterns of CpG methylation and gene expression for candidate genes.

Subgroup-specific candidate epigenetically regulated genes were defined by those genes which showed consistent methylation-dependent gene regulation in cell lines, alongside differential CpG methylation inversely correlated with gene expression in distinct subgroups of primary tumours. Figure 4.1 summarises the steps involved in the selection of subgroup-specific candidate epigenetically regulated genes.

Genome-wide analysis of gene expression changes following 5-azaCdR treatment of medulloblastoma cell lines

10 untreated and 5-azaCdR-treated cell line pairs

↓

Gene expression profiled on Illumina HT12v4.0 microarray

↓

Per cell line pair, fold change in expression following 5-azaCdRtreatment (FC) calculated along with mean FC and standard deviation (SD) of the mean

↓

FC > [(mean FC +3SD)] in ≥ 2 cell lines

↓

Genes whose expression is upregulated following 5-azaCdR

↓

Genome-wide assessment of CpG methylation associated with gene expression changes following 5-azaCdR treatment of medulloblastoma cell lines

Illumina 450K cell line methylation data (10 cell lines)

↓

CpG sites corresponding to upregulated genes

Constitutively methylated CpG sites (β>0.3 in 10 cell lines)

Methylation consistent with gene upregulation in ≥ 60% cell lines

Variably methylated CpG sites (0≤β≤1 across 10 cell lines)

- Methylation status consistent with expression changes observed in ≥ 60% cell lines **AND**
- Methylation consistent with gene upregulation in ≥ 60% methylated cell lines **AND**
- Significant inverse linear correlation between CpG methylation and endogenous gene expression ($r<(-0.7)$; $p<0.05$)

Critical CpG sites associated with methylation-dependent gene expression in cell lines

↓

Candidate epigenetically regulated genes in medulloblastoma cell lines

↓

Differential methylation and gene expression analysis in primary tumours

Differential methylation and gene expression analysis in primary tumours

Differential CpG methylation

↓

Critical CpG sites identified in cell lines

↓

Illumina 450K methylation array; 109 primary tumours: 4 subgroups (SHH, WNT, Grp3, Grp4)

↓

R package *limma*

↓

Subgroup-specific significant differential CpG methylation: $p<0.05$, logFC>0.9, $\Delta\beta>0.2$

Differential gene expression

↓

Candidate genes identified in cell lines

↓

3 independent transcriptomic datasets; 357 primary tumours: 4 subgroups (SHH, WNT, Grp3, Grp4)

↓

R package *limma*

↓

Subgroup-specific significant differential gene expression: $p<0.05$, logFC > (0.4-0.7), ≥ 2 independent datasets

Subgroup-specific differential CpG methylation inversely associated with differential gene expression

↓

Subgroup-specific candidate epigenetically regulated genes in primary tumours

Figure 4.1. Selection of subgroup-specific candidate epigenetically regulated genes in primary medulloblastomas. Flow diagram illustrates the steps followed to select candidate epigenetically regulated genes in distinct subgroups of medulloblastoma. Genes were selected from an initial functional epigenomics screen in cell lines. Candidate genes that showed methylation-dependent gene regulation in cell lines were investigated for evidence of differential methylation and gene expression in distinct subgroups of primary tumours. Full details of differential methylation and gene expression analyses are provided in sections 4.3.11 and 4.3.12 respectively. For differential methylation analysis, a logFC threshold (absolute difference in mean M-values) and a $\Delta\beta$ threshold (absolute difference in mean β-values) were applied as described in section 4.3.11 to increase the sensitivity of the significance test. For differential expression analysis, logFC thresholds (absolute difference in mean $\log_2$ expression) ranging from 0.4 to 0.7, were applied to the 3 independent transcriptomic datasets as described in section 4.3.12. Significant subgroup-specific differential gene expression was required in at least 2 out of the 3 datasets. Candidate subgroup-specific epigenetically regulated genes were defined by those genes which showed consistent methylation-dependent gene regulation in cell lines, alongside differential CpG methylation inversely correlated with gene expression in distinct subgroups of primary tumours.

4.3.13.1 Assessment of region-wide DNA methylation for candidate epigenetically regulated genes

For subgroup-specific candidate genes identified (see section 4.3.13), the methylation status of CpG residues adjacent to the critical significantly differentially methylated sites on the Illumina 450K array, were analysed in primary medulloblastomas and in the normal cerebellum. Residues with 450K array β-value scores ≤0.3 were considered unmethylated (U); β-value scores >0.7 were considered methylated (M) and scores >0.3 but ≤0.7 were considered part-methylated (PM). Data was represented as clock diagrams produced in Excel 2007 to visually determine whether the methylation status of the differentially methylated sites, which were associated with gene expression in primary tumours, was reflective of the wider genomic region in which they were located.

4.3.13.2 RNA-Seq versus microarray expression profiling of candidate gene expression changes following 5-azaCdR treatment in cell lines

RNA-Seq is a high-throughput deep sequencing technology that provides resolution at the single base level. It is a more sensitive method of quantification of transcripts and their isoforms and has a better dynamic range compared with microarray technologies (Wang *et al.*, 2009). It can therefore provide a useful validation of microarray profiling results. Towards the end of this project RNA sequencing reads were available for 6 untreated and 5-azaCdR-treated medulloblastoma cell line pairs (D283, D341, D425, D556, D458, MED1), and were kindly provided by Dr. Richard Birnie (Newcastle University Paediatric Brain Tumour Group) for the candidate genes of interest. The correlation between RNA-Seq reads and log_2 expression levels generated on the Illumina HT12v4.0 microarray for candidate genes identified in section 4.3.13 was assessed by calculating Pearson's product moment correlation coefficient (*r*). RNA-Seq transcript reads were considered to validate methylation-dependent expression alterations observed by microarray analysis if the coefficient *r* value obtained was greater than 0.7.

4.4 Results

This chapter aimed to investigate the genome-wide role of DNA methylation in the regulation of gene expression in distinct subgroups of medulloblastoma. Using a functional epigenomics analysis of methylation-dependent gene expression alterations, combined with a genome-wide analysis of DNA methylation, the investigations reported in this chapter provide a comprehensive characterisation of methylation events that are associated with expression alterations and that may play a role in the development of distinct subgroups of medulloblastoma.

4.4.1 Re-expression of RASSF1A in medulloblastoma cell lines following treatment with 5-azaCdR

The isoform A transcript of the *RASSF1* gene was used as a positive control for the demethylating treatment of cell lines with 5-azaCdR (see section 4.3.2). Re-expression of *RASSF1A* was observed in all 10 cell lines following 5-azaCdR treatment (Figure 4.2(A)). Hypermethylation of the *RASSF1A* CpG island-associated promoter has previously been established in multiple medulloblastoma cell lines (Lusher *et al.*, 2002), indicating that *RASSF1A* promoter hypermethylation is a common event in medulloblastoma cell lines, consistent with its observed epigenetic inactivation and re-expression by 5-azaCdR. Equivalent expression levels were observed for the β-actin control transcript in the treated and untreated cell line pairs (Figure 4.2(B)). Subsequent studies which have assessed genome-wide patterns of DNA methylation in a subset of cell lines after 5-azaCdR have demonstrated global demethylation following treatment (Ed Schwalbe, Newcastle University Paediatric Brain Tumour Group; unpublished data).
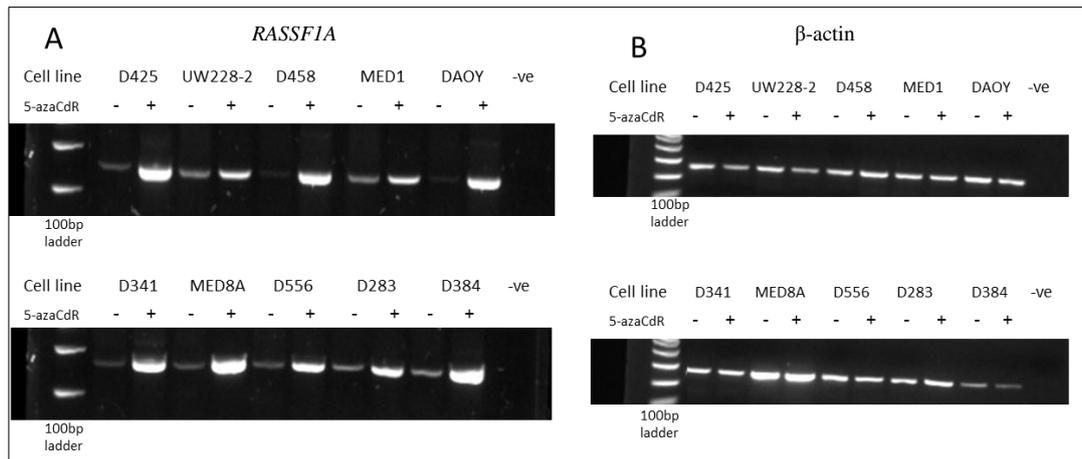
Figure 4.2. Re-expression of *RASSF1A* following 5-azaCdR treatment of medulloblastoma cell lines. (A) Reverse transcription PCR analysis of *RASSF1A* gene expression levels in the 10 medulloblastoma cell lines with (+) and without (-) treatment with 5-azaCdR. Methylation-dependent re-expression of *RASSF1A* was observed in all 10 cell lines following 5-azaCdR treatment. **(B)** Reverse transcription PCR analysis of the β-actin endogenous control gene in the 10 medulloblastoma cell lines with (+) and without (-) treatment with 5-azaCdR. β-actin was expressed at equivalent levels in the treated and untreated cell line pairs. PCR was assessed against a negative control reaction (-ve), which was set up identically to the experimental PCR but without template DNA. Results of negative control reaction show no PCR product demonstrating that there was no contamination of PCR reactions from non-sample DNA.

4.4.2 Whole genome gene expression profiling

The Illumina HT12-v4.0 microarray profiles the expression of 47,231 mRNA transcripts (encompassing over 30,000 genes) and was used to assess expression alterations in medulloblastoma cell lines following treatment with 5-azaCdR.

4.4.2.1 Quality control and array normalisation

The 10 untreated and 5-azaCdR-treated cell line pairs were processed in a single batch on two BeadChips. There were no obvious differences in the distribution of average probe signal intensities for samples within and between chips (Figure 4.3(A)), and the average number of beads per probe type represented on each array was approximately equal for all samples (Figure 4.3(B)). The distribution of average probe signal intensities was identical for all samples following quantile normalisation of data (Figure 4.3 (C)).

4.4.2.2 Data filtering

A total of 12,881 probes were assigned a 'Bad' or 'No match' quality score as defined by Illumina after normalisation. By their definition these probes are considered to have undesirable properties and have been shown to correlate with lowly expressed probes and with probes expressed at high levels due to non-specific hybridisation (Dunning *et al.*, 2007). They were classed as non-responding probes and removed from the data.

All samples showed high numbers of probes detected with $p<0.05$, with similar numbers within and between chips (mean=17,497 and SD=1,299). These levels were similar to other published reports using Illumina chips (Powell *et al.*, 2012) and were indicative of a high overall sample expression quality. There were 13,111 (28%) probes that were not detected ($p<0.05$) in any cell line sample and were therefore filtered from the data, while 8,286 (18%) probes were detected in all samples. An additional 740 probes were removed from the data. They represented probes that mapped to the sex chromosomes or that did not map to an official gene symbol or chromosome location according to Illumina's annotation.

Overall, data filtering measures removed 26,732 (57%) probes that were considered unreliable or uninformative. The remaining 20,499 (43%) probes were investigated for evidence of methylation-dependent gene expression alterations in medulloblastoma cell lines following 5-azaCdR treatment.
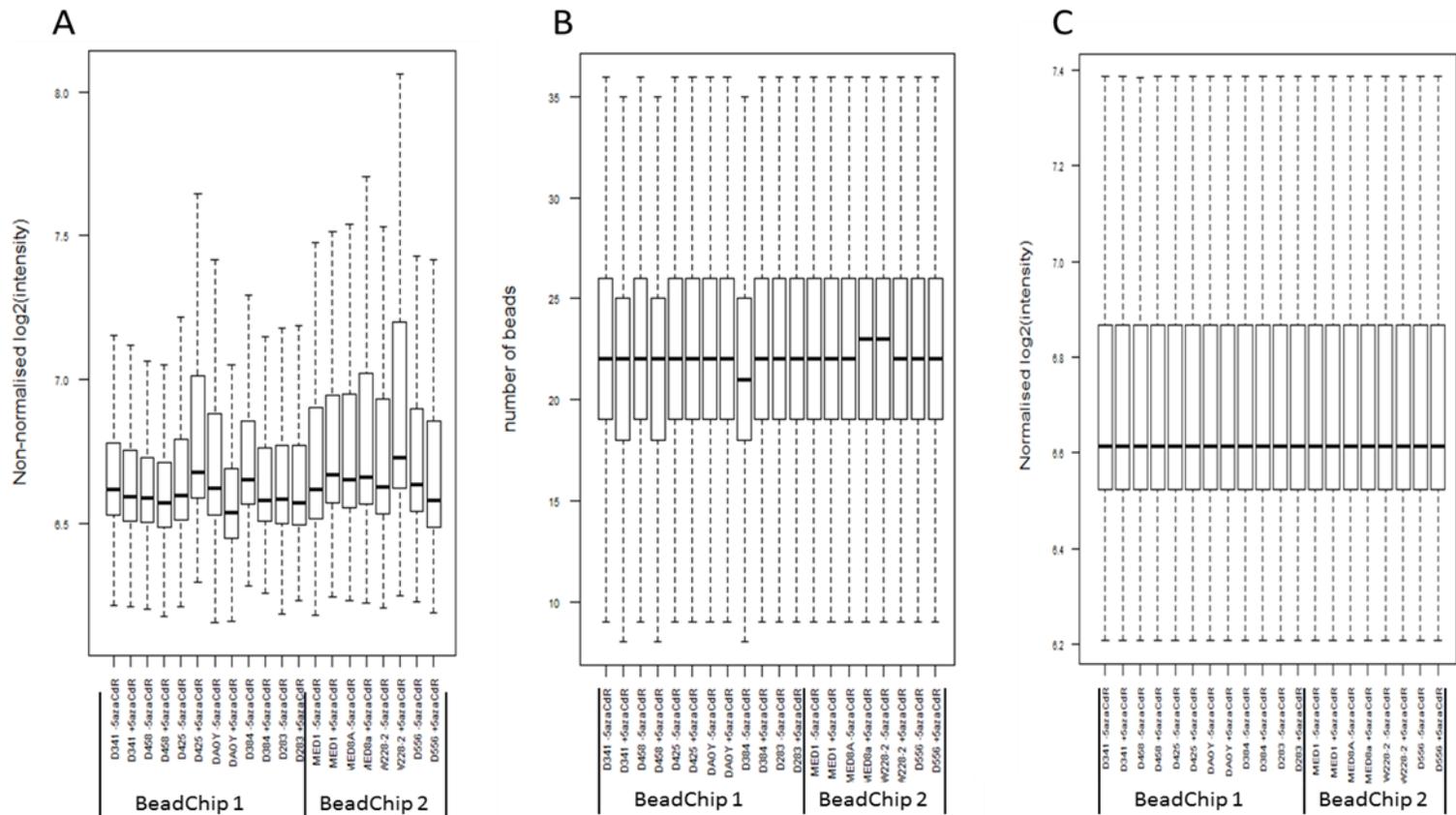
Figure 4.3. Quality control and normalisation of Illumina gene expression bead summary data. **(A).** Boxplots illustrating the distribution of average probe signal intensities before data normalisation for all samples processed across the 2 BeadChips. Distributions are similar for all samples and no arrays were identified as problematic or outliers. **(B).** Boxplots illustrating the similar average number of beads per probe type represented on each array across the 2 BeadChips. **(C).** Boxplots illustrating the identical distributions of average probe signal intensities following quantile normalisation of bead summary data for all samples. Twelve samples (6 cell line pairs) were processed on BeadChip1 and eight samples (4 cell line pairs) were processed on BeadChip2. BeadChips were processed in a single batch to minimise batch effects.

4.4.3 Upregulated gene expression following 5-azaCdR treatment of medulloblastoma cell lines

FC thresholds were defined as described in section 4.3.7. The FC thresholds applied to the 10 cell lines ranged from 1.3 to 2.9. Analysis of individual cell lines showed that between 0.2% and 1.7% of probes detected an expression increase greater than the threshold following 5-azaCdR treatment (Table 4.1).

Cell line	meanFC+3SD	FC threshold	FC>threshold (no. probes)	FC> threshold (% probes)
D341	1.50	1.5	165	0.8
D458	1.57	1.6	257	1.3
D425	1.78	1.8	252	1.2
DAOY	1.62	1.6	264	1.3
D384	1.57	1.6	200	1.0
D283	1.45	1.5	244	1.2
MED1	1.34	1.3	357	1.7
MED8A	1.81	1.8	148	0.7
UW228-2	2.93	2.9	46	0.2
D556	1.44	1.4	106	0.5

Table 4.1. **Fold change thresholds for upregulated gene expression in medulloblastoma cell lines following 5-azaCdR treatment**. The FC threshold applied to each cell line was the mean FC (fold change in expression after 5-azaCdR) + 3SD of the mean. The number and percentage of transcript probes that exceeded the threshold in each cell line are shown. Probes were considered upregulated by 5-azaCdR if they detected an expression increase greater than the threshold in 2 or more cell lines following 5-azaCdR treatment.

A total of 307 probes (1.5%), encompassing 283 genes, detected an expression increase greater than the threshold in 2 or more cell lines and were considered upregulated by 5-azaCdR treatment. 58% (179/307) of upregulated probes were observed in only 2 cell lines and 1% (3/307) were observed in 7 cell lines. Microarray analysis did not detect any upregulated probes following 5-azaCdR treatment in more than 7 out of the 10 cell lines (Table 4.2).

Number of cell lines	Number of probes upregulated (%)
2	179 (58%)
3	70 (23%)
4	24 (8%)
5	21 (7%)
6	10 (3%)
7	3 (1%)
8	0 (0%)
9	0 (0%)
10	0 (0%)

Table 4.2. Microarray analysis of gene upregulation in medulloblastoma cell lines. The number of probes upregulated by 5-azaCdR treatment and the number of cell lines in which upregulation was observed are shown.

GO analysis of upregulated probes using the DAVID Bioinformatics Resource revealed enrichment for genes involved in the maintenance, structure and function of the intracellular cytoskeleton as well as genes involved in the regulation of apoptosis and stimulus response genes. There was no specific pathway enrichment identified among the upregulated genes (Figure 4.4).

Annotation Cluster 1		Enrichment Score: 6.38				Count	P_Value	Benjamini
GOTERM_CC_FAT	contractile fiber part		RT			19	4.7E-13	1.5E-10
GOTERM_CC_FAT	sarcomere		RT			18	5.0E-13	7.8E-11
GOTERM_CC_FAT	contractile fiber		RT			19	1.6E-12	1.6E-10
SP_PIR_KEYWORDS	muscle protein		RT			14	2.1E-12	8.2E-10
GOTERM_CC_FAT	myofibril		RT			18	4.1E-12	3.2E-10
GOTERM_CC_FAT	actin cytoskeleton		RT			20	1.4E-7	8.5E-6
GOTERM_CC_FAT	I band		RT			10	3.0E-7	1.5E-5
GOTERM_CC_FAT	Z disc		RT			9	8.5E-7	3.8E-5
GOTERM_CC_FAT	cytoskeleton		RT			46	9.2E-6	3.2E-4
SP_PIR_KEYWORDS	actin binding		RT			7	2.0E-5	1.1E-3
GOTERM_BP_FAT	muscle contraction		RT			11	2.5E-4	1.9E-2
GOTERM_BP_FAT	muscle system process		RT			11	5.2E-4	2.6E-2
GOTERM_MF_FAT	cytoskeletal protein binding		RT			20	7.3E-4	5.9E-2
GOTERM_BP_FAT	actin filament-based process		RT			13	7.5E-4	3.3E-2
GOTERM_MF_FAT	actin binding		RT			15	1.1E-3	7.1E-2
SP_PIR_KEYWORDS	actin-binding		RT			10	9.1E-3	1.0E-1
SP_PIR_KEYWORDS	cytoskeleton		RT			17	2.1E-2	2.0E-1
Annotation Cluster 2		Enrichment Score: 5.6				Count	P_Value	Benjamini
GOTERM_BP_FAT	response to organic substance		RT			36	1.1E-8	1.8E-5
GOTERM_BP_FAT	response to abiotic stimulus		RT			21	3.7E-6	9.9E-4
GOTERM_BP_FAT	response to mechanical stimulus		RT			9	3.8E-6	8.7E-4
GOTERM_BP_FAT	response to hormone stimulus		RT			20	1.3E-5	2.1E-3
GOTERM_BP_FAT	response to endogenous stimulus		RT			20	5.0E-5	6.2E-3
Annotation Cluster 3		Enrichment Score: 4.77				Count	P_Value	Benjamini
GOTERM_BP_FAT	regulation of apoptosis		RT			34	1.5E-6	7.8E-4
GOTERM_BP_FAT	regulation of programmed cell death		RT			34	1.8E-6	7.3E-4
GOTERM_BP_FAT	regulation of cell death		RT			34	2.0E-6	6.3E-4
GOTERM_BP_FAT	programmed cell death		RT			26	3.2E-5	4.3E-3
GOTERM_BP_FAT	apoptosis		RT			25	7.1E-5	7.5E-3
GOTERM_BP_FAT	cell death		RT			27	1.7E-4	1.5E-2
GOTERM_BP_FAT	death		RT			27	1.9E-4	1.6E-2
Annotation Cluster 4		Enrichment Score: 4.74				Count	P_Value	Benjamini
GOTERM_BP_FAT	response to inorganic substance		RT			17	3.4E-7	2.7E-4
GOTERM_BP_FAT	response to metal ion		RT			11	6.0E-5	6.9E-3
GOTERM_BP_FAT	response to calcium ion		RT			7	3.0E-4	2.1E-2
Annotation Cluster 5		Enrichment Score: 4.26				Count	P_Value	Benjamini
GOTERM_BP_FAT	regulation of apoptosis		RT			34	1.5E-6	7.8E-4
GOTERM_BP_FAT	regulation of programmed cell death		RT			34	1.8E-6	7.3E-4
GOTERM_BP_FAT	regulation of cell death		RT			34	2.0E-6	6.3E-4
GOTERM_BP_FAT	negative regulation of apoptosis		RT			17	3.0E-4	2.1E-2
GOTERM_BP_FAT	negative regulation of programmed cell death		RT			17	3.5E-4	2.1E-2
GOTERM_BP_FAT	negative regulation of cell death		RT			17	3.7E-4	2.1E-2
GOTERM_BP_FAT	anti-apoptosis		RT			10	7.8E-3	1.9E-1
Annotation Cluster 6		Enrichment Score: 4.04				Count	P_Value	Benjamini
GOTERM_BP_FAT	response to extracellular stimulus		RT			16	4.3E-6	8.6E-4
GOTERM_BP_FAT	response to nutrient levels		RT			15	5.6E-6	1.0E-3
GOTERM_BP_FAT	response to nutrient		RT			10	5.7E-4	2.7E-2
GOTERM_BP_FAT	response to vitamin		RT			6	4.9E-3	1.3E-1

238

Figure 4.4. Gene ontology of probes upregulated by 5-azaCdR The top 6 annotation clusters computed using the DAVID Functional Annotation Clustering Algorithmn are shown. For each cluster the group of enriched terms having similar biological meaning due to sharing similar gene members are detailed, along with the number of genes involved in each term, and the modified Fisher Exact p-value for gene enrichment.The overall Enrichment Score for the group was computed using the p-values of each term member; the higher the enrichment score, the more enriched the group of terms. 131 clusters were identified from the 307 upregulated Illumina probe IDs that were entered into the DAVID Bioinformatics Annotation Clustering Tool. The top 6 clusters had enrichment scores greater than 4. Functional Annotation Clustering Report down;loaded from the DAVID Bioinformatics website (http://david.abcc.ncifcrf.gov/).

4.4.4 Critical CpG sites associated with methylation-dependent gene regulation in cell lines

The 283 genes upregulated by 5-azaCdR treatment in cell lines encompassed 5,251 CpG probes that were extracted from the cell line 450K methylation array data. Of these, 1,013 (19%) were constitutively unmethylated in the cell lines and were removed from the data as they would not be associated with any methylation-dependent expression changes observed for their respective genes. There were 1,579 (30%) constitutively methylated probes (see section 4.3.8.1) and 2,659 (51%) variably methylated probes (see section 4.3.8.1) that were investigated for evidence of a critical role in epigenetic gene regulation.

4.4.4.1 Constitutively methylated CpG sites

Constitutively methylated CpG sites were considered to show strong evidence of a critical role in methylation-dependent gene regulation if re-expression was observed in $\geq$ 60% of cell lines following 5-azaCdR treatment. A "consistent relationship" (see section 4.3.8.1) between CpG methylation and gene re-expression following 5-azaCdR was observed in $\geq$ 60% of cell lines for 71 (4%) constitutively methylated CpG sites. The 71 sites (encompassing 12 genes) were considered to show strong evidence of a critical role in epigenetic gene regulation. As an example of the constitutively methylated critical CpG sites identified, Figure 4.5 illustrates methylation-dependent re-expression of the gene *SYT11* in cell lines.
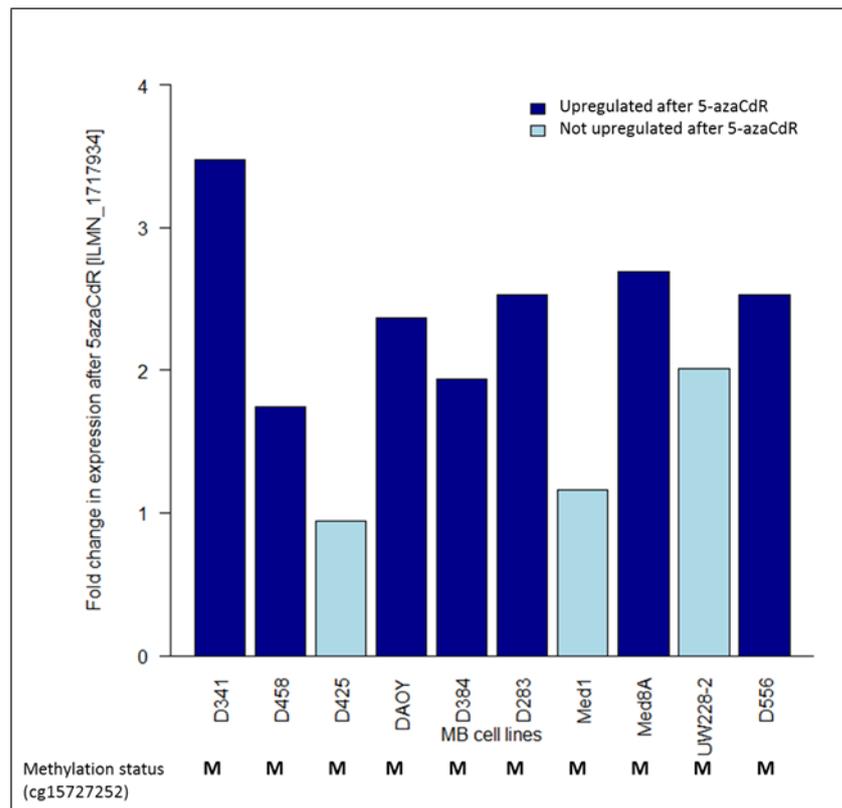
Figure 4.5. Methylation-dependent re-expression of *SYT11* in medulloblastoma cell lines. Bar-graph illustrates the fold change (FC) in expression after 5-azaCdR treatment in 10 cell lines for the *SYT11* transcript probe ILMN_171793. Dark blue bars represent a FC > threshold for that cell line, indicative of upregulated gene expression by 5-azaCdR; light blue bars represent a FC < threshold for that cell line, indicative of no upregulated expression by 5-azaCdR. Methylation status M = methylated (β-values >0.3). Methylation of the *SYT11* CpG site cg15727252 is consistent with methylation-dependent re-expression observed in 70% of cell lines.

4.4.4.2 Variably methylated CpG sites

There were 291 variably methylated CpG sites (11%) that showed a "consistent relationship" (see section 4.3.8.1) between CpG methylation status and gene expression alterations observed following 5-azaCdR treatment in ≥ 60% of cell lines. The 291 CpG sites corresponded to 66 genes.

Of the 291 sites, 89 also showed a significant inverse correlation ($r < -0.7$; FDR adjusted *p*-value<0.05) between methylation and endogenous gene expression levels in medulloblastoma cell lines. Methylation of these 89 CpG sites (encompassing 17 genes) was considered to show strong evidence of an association with transcriptional silencing and re-expression by 5-azaCdR and thus a critical role in epigenetic gene regulation. As an example of the variably methylated critical CpG sites identified, Figure 4.6 illustrates the "consistent relationship" that was observed between the methylation status of the

TUBA1C CpG probe shown (cg27632435) and expression alterations in 80% of cell lines following 5-azaCdR treatment, with 67% of methylated cell lines showing upregulated expression after treatment. Figure 4.7 illustrates the significant inverse correlation between methylation of the same *TUBA1C* CpG site (cg27632435) and endogenous gene expression in cell lines. Together these results demonstrate that methylation of this CpG site was consistent with *TUBA1C* transcriptional silencing and upregulation by 5-azaCdR in cell lines.
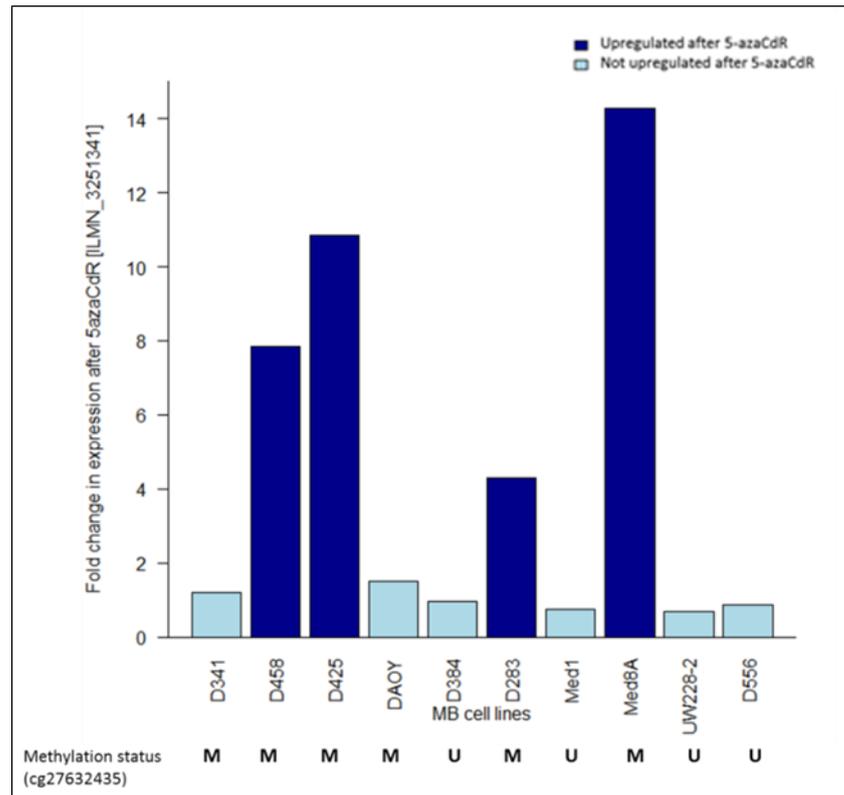
Figure 4.6. Methylation-dependent regulation of *TUBA1C* in medulloblastoma cell lines. Bar-graph illustrates the fold change (FC) in expression after 5-azaCdR treatment of 10 cell lines for the *TUBA1C* transcript probe ILMN_3251341. Dark blue bars represent a FC > threshold for that cell line, indicative of upregulated gene expression by 5-azaCdR; light blue bars represent a FC < threshold for that cell line, indicative of no upregulated expression by 5-azaCdR. Methylation status M = methylated (β-values >0.3). Methylation status U= unmethylated ($\beta \leq 0.3$). Methylation of the *TUBA1C* CpG probe cg27632435 is consistent with methylation-dependent re-expression observed in 67% (4/6) of methylated cell lines. Consistent with methylation-dependent expression, 100% (4/4) of unmethylated cell lines do not show re-expression after 5-azaCdR. Overall methylation status is consistent with epigenetic gene regulation in 80% of cell lines.
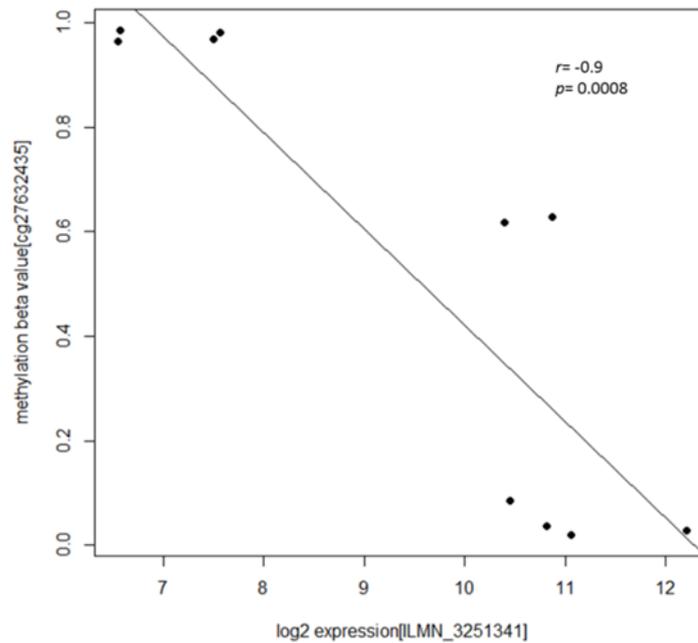
Figure 4.7. Inverse linear correlation between CpG methylation and gene expression for *TUBA1C*. Methylation status of the *TUBA1C* CpG probe cg27632435 is significantly inversely correlated with gene expression in 10 cell lines. Pearson correlation *r* value measuring the strength of the inverse relationship and the FDR-adjusted *p*-value of the correlation test are shown. Results show that methylation of the CpG site is associated with *TUBA1C* transcriptional silencing in cell lines.

4.4.5 Candidate epigenetically regulated genes in medulloblastoma cell lines

A total of 160 CpG sites (71 constitutively methylated and 89 variably methylated) were identified that showed strong evidence of a critical role in epigenetic gene regulation in medulloblastoma cell lines. The 160 CpG sites encompassed 21 genes. All genes were associated with DNA methylation events occurring at multiple residues, ranging from 18 critical CpG sites in the gene *SQSTM1* to 2 critical CpG sites in *FABP5* and *S100A10* (Table 4.3).

Gene	Chromosome	Number of critical CpG sites identified
ACTC1	15	7
ANXA2	15	6
DAZL	3	16
FABP5	8	2
FAM46A	6	4
HSPB1	7	7
IER3	6	4
MAP1LC3A	20	17
MFGE8	15	5
PRPH	12	4
S100A10	1	2
S100A4	1	5
S100A6	1	8
SAP25	7	6
SQSTM1	5	18
SYT11	1	3
TAGLN	11	5
TEAD2	19	4
TGFBI	5	14
TUBA1C	12	13
UBB	17	10

Table 4.3. Candidate epigenetically regulated genes in medulloblastoma cell lines. Twenty-one genes contained multiple CpG residues which had a methylation status consistent with their epigenetic transcriptional regulation in cell lines. Chromosome location and number of critical CpG sites associated with methylation-dependent gene regulation are shown for each gene.

50% (80/160) of the CpG sites identified were located in CpG islands and the majority of these were island-associated promoters (Figure 4.8). 32% (52/160) of sites were located in CpG island shores and island shelves and these were distributed equally between promoter regions and gene body/3'UTR regions. A further 18% (29/160) of residues were isolated sites distributed equally between promoter and gene body/3'UTR regions (Figure 4.8).
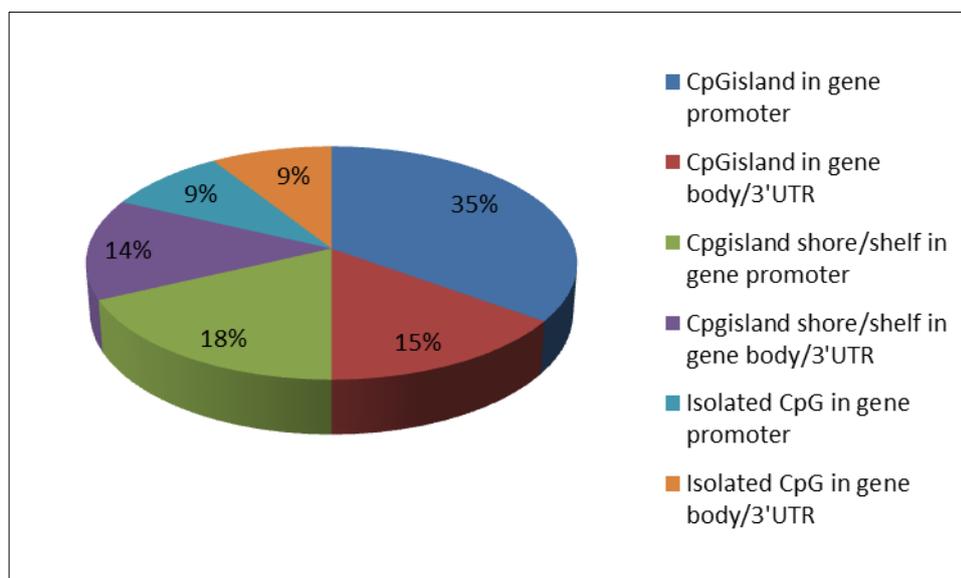
Figure 4.8. Genomic distribution of CpG sites associated with methylation-dependent gene regulation in cell lines. Pie chart illustrates the percentage of critical CpG sites in the different genomic regions (as defined by Illumina) and in different regions of CpG density (as defined by Illumina). CpG sites were located in gene promoter, gene body and 3'UTR functional regions and were within CpG islands and island shores/shelves as well as isolated non-CpG island sites within the gene. 50% of sites were located in a CpG island while a further 32% were located in a CpG island shore/shelf. 62% of sites were located in gene promoter regions. Pie chart was produced in Excel 2007.

4.4.6 Subgroup-specific differential CpG methylation

Of the 160 critical CpG sites whose methylation was consistent with epigenetic gene regulation in cell lines, 32 (20%) were significantly differentially methylated in specific subgroups of primary medulloblastomas (see section 4.3.11). The 32 differentially methylated sites encompassed 10 of the 21 (48%) candidate genes (Table 4.4). Four genes (*ANXA2*, *MAP1LC3A*, *MFGE8* and *TUBA1C*) were differentially methylated in more than 1 subgroup; differential methylation of *MFGE8* occurred at different CpG sites in the different subgroups while for *ANXA2*, *MAP1LC3A* and *TUBA1C* differential methylation occurred at the same residues in different subgroups and could be explained by the variable methylation observed for these sites in some tumour subgroups

compared to others. Figure 4.9 illustrates the variable methylation observed for *TUBA1C* in the SHH and Group 3 tumours that led to the same CpG site being significantly differentially methylated in WNT tumours and in Group 4 tumours. There were 4 CpG sites (encompassing 2 genes) that were significantly differentially methylated between Group (3+4) and (SHH+WNT) tumours (Table 4.4). Interestingly, there were no sites that were differentially methylated between (SHH+Group 3) and (WNT+Group 4) or between (SHH+Group 4) and (WNT+Group 3), supporting the current opinion that Group 3 and Group 4 tumours are more similar to each other than they are to SHH or WNT tumours (Taylor *et al.*, 2012).

Subgroup	Gene	Number of differentially methylated CpG loci
SHH	*FAM46A*	3
	MFGE8	1
WNT	*ANXA2*	1
	MAP1LC3A	2
	MFGE8	1
	PRPH	1
	S100A4	2
	TUBA1C	13
Group 3	*TAGLN*	2
	MAP1LC3A	1
Group 4	*ANXA2*	1
	MAP1LC3A	1
	TUBA1C	13
Group (3+4)	*SOSTM1*	2
	ACTC1	2

Table 4.4. Subgroup-specific differential CpG methylation in primary medulloblastomas. Ten candidate genes were differentially methylated at 1 or more CpG sites in specific subgroups of medulloblastoma. Subgroup specificity and number of differentially methylated CpG sites are shown for each gene. Four genes were differentially methylated in more than 1 subgroup; differential methylation of *MFGE8* occurred at different CpG sites in different subgroups while for *ANXA2*, *MAP1LC3A* and *TUBA1C* differential methylation occurred at the same site(s) in the different subgroups. A total of 32 critical CpG sites were differentially methylated in a subgroup-specific manner in primary medulloblastomas.
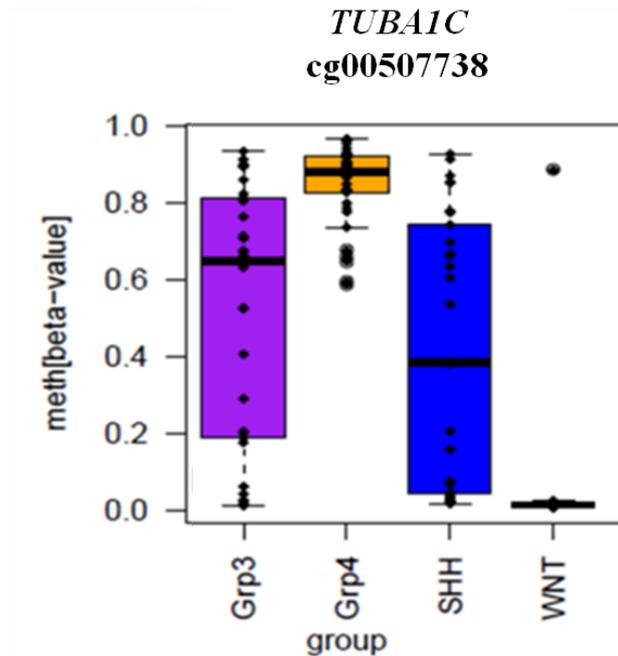
Figure 4.9. Differential methylation of *TUBA1C* in WNT and in Group 4 tumours. Boxplots illustrate the differential patterns of methylation β-values of the *TUBA1C* CpG probe cg00507738 across the tumour subgroups. The same site is significantly differentially methylated in WNT tumours compared to (SHH+Grp3+Grp4) tumours and in Grp4 tumours compared to (SHH+WNT+Grp3) tumours. The methylation β-values are variable in the SHH and Group 3 tumours compared to the uniform WNT and Group 4 tumours explaining specificity in 2 distinct subgroups. The same differential patterns of methylation were observed for the 13 WNT-specific and Group 4-specific *TUBA1C* CpG sites (see Table 4.4).

4.4.7 Subgroup-specific differential gene expression

Candidate genes were next investigated for evidence of subgroup-specific differential gene expression (see section 4.3.12). Of the 21 candidate genes (encompassing the 160 critical CpG sites) identified in cell lines (see section 4.4.5), 15 were significantly differentially expressed in specific subgroups of primary medulloblastomas in at least 1 transcriptomic dataset and 11 genes (52%) were differentially expressed in the same subgroup in 2 or more datasets (Table 4.5). The following subgroup-specific differentially expressed genes were identified in 2 or more datasets:

- SHH-specific: *FAM46A*, *PRPH*
- WNT-specific: *ANXA2*, *HSPB1*, *PRPH*, *S100A4*, *S100A6*, *TGFBI*
- Group 3-specific: *HSPB1*, *SYT11*
- Group (3+4)-specific: *ACTC1*, *MFG8A*, *SYT11*, *TEAD2*

No genes were identified that were differentially expressed in Group 4 tumours in 2 or more datasets.

Subgroup	Gene	Significant differential gene expression		
		Kool/Fattet	Cho	Northcott
SHH	*ACTC1*			√
	FAM46A	√	√	√
	PRPH	√	√	
	S100A6		√	
	TEAD2	√		
WNT	*ACTC1*	√		
	ANXA2	√	√	√
	HSPB1	√	√	√
	PRPH	√	√	
	S100A4	√	√	
	S100A6	√	√	
	TEAD2	√		
	TGFBI	√	√	√
Group 3	*ANXA2*	√		
	HSPB1	√	√	√
	IER3	√		
	SYT11	√	√	√
	TEAD2			√
Group 4	*ANXA2*		√	
	FAM46A		√	
	SYT11		√	
	TAGLN			√
	TEAD2			√
Group (3+4)	*ACTC1*		√	√
	FABP5	√		
	MFGE8		√	√
	S100A6			√
	SYT11		√	√
	TEAD2	√		√

Table 4.5. Subgroup-specific differential gene expression in primary medulloblastomas. Significant subgroup-specific differential gene expression (√) is shown for each transcriptomic dataset. A total of 15 genes were significantly differentially expressed in 1 or more subgroups in at least 1 dataset and 11 genes were significantly differentially expressed in the same subgroup in 2 or more datasets. Significant subgroup-specific differential expression was determined as described in section 4.3.12.

4.4.8 Subgroup-specific candidate epigenetically regulated genes in primary tumours

Subgroup-specific candidate epigenetically regulated genes were defined by those genes which showed methylation-dependent gene regulation in cell lines, alongside differential CpG methylation inversely correlated with gene expression in distinct subgroups of primary tumours (Figure 4.1).

The following 5 genes showed evidence of methylation-dependent regulation in cell lines, alongside differential CpG methylation inversely correlated with gene expression in distinct subgroups of primary tumours, consistent with their epigenetic regulation:

- SHH-specific: *FAM46A* (Figure 4.10)
- WNT-specific: *ANXA2* (Figure 4.11), *PRPH* (Figure 4.12), *S100A4* (Figure 4.13)
- Group (3+4)-specific: *ACTC1* (Figure 4.14)

To illustrate the identification of subgroup-specific candidate genes, Table 4.6 details the significant differentially methylated and significant differentially expressed events in SHH tumours compared to all other subgroups, and explains the identification of *FAM46A* as a candidate epigenetically regulated gene in SHH tumours.

The 5 candidate genes encompassed 9 critical CpG sites that showed an association with gene expression in primary tumours. These represented 6% of the 160 critical sites that showed an association with expression in cell lines. Figure 4.15 summarises the number of genes and critical CpG sites identified at each step of the analysis, which led to the identification of the 5 candidate epigenetically regulated genes in primary tumours.

	450K methylation	Gene	WNT*v*SHH	G3*v*SHH	G4*v*SHH	SHH*v*Cerebella	SHH*v*(WNT+G3+G4)	SHH*v*Cerebella
			Contrasts				**Absolute difference in mean β-values(Δβ)**	
Significant differential CpG methylation	cg04399083	*FAM46A*	1	1	1	0	0.323823	0.089449
	cg11075693	*FAM46A*	1	1	1	0	0.347132	0.077789
	cg22388948	*FAM46A*	1	1	1	0	0.595152	0.038517
	cg08896420	*MFGE8*	1	1	1	-1	0.466019	0.216803
Significant differential gene expression	**KoolFattet dataset**							
	Transcript Probe ID	*Gene*	**WNTvSHH**	**G3vSHH**	**G4vSHH**			
	224973_at	*FAM46A*	-1	-1	-1			
	213847_at	*PRPH*	1	-1	-1			
	238323_at	*TEAD2*	1	-1	-1			
	Cho dataset							
	Transcript Probe ID	*Gene*	**WNTvSHH**	**G3vSHH**	**G4vSHH**			
	221766_s_at	*FAM46A*	-1	-1	-1			
	213847_at	*PRPH*	1	-1	-1			
	217728_at	*S100A6*	1	-1	-1			
	Northcott dataset							
	Transcript Probe ID	*Gene*	**WNTvSHH**	**G3vSHH**	**G4vSHH**			
	3588303	*ACTC1*	1	1	1			
	2962383	*FAM46A*	-1	-1	-1			

Table 4.6. Significant differential CpG methylation and gene expression in SHH medulloblastomas. *Limma* significance results are shown for significantly differentially methylated and significantly differentially expressed events specific to SHH tumours; 1= significantly positive event; 0= not significant; -1= significantly negative event. 450K CpG methylation probes, transcript probes from the 3 datasets (Kool/Fattet, Cho and Northcott) and the contrasts tested using limma are shown. The CpG methylation probes and the transcript probes shown are significantly different (1 or -1) in the 3 contrasts: WNTvSHH, Grp3vSHH and Grp4vSHH, making them SHH-specific events. CpG methylation was judged significantly different (1 or -1) only if the FDR-adjusted *p*-value was < 0.05 and the absolute difference in mean M-value methylation scores was >0.9. An additional filter of absolute difference in mean β-values>0.2 also had to be satisfied (see section 4.3.12). Gene expression was judged significantly different (1 or -1) only if the FDR-adjusted *p*-value was <0.05 and the absolute difference in log2 expression levels was >0.5 for the Kool/Fattet dataset; >0.7 for the Cho datasset and >0.4 for the Northcott dataset (see section 4.3.13). *FAM46A* is the only gene differentially methylated and differentially expressed in SHH tumours; the significantly lower methylation status (1) in SHH tumours correlates with significantly higher expression (1) in the 3 transcriptomic datasets, identifying *FAM46A* as a strong candidate epigenetically regulated gene in SHH tumours. The 3 SHH-specific *FAM46A* CpG loci do not have a significantly different methylation profile in SHH tumours compared to non-neoplastic cerebella (limma result 0 and Δβ<0.34), suggesting that a hypermethylated status in WNT,Group3 and Group4 tumours is associated with transcriptional silencing in these tumour subgroups (see Figure 4.9). Expression data was not available for non-neoplastic cerebellar samples to assess aberrant gene expression. G3=Group 3 and G4=Group 4.
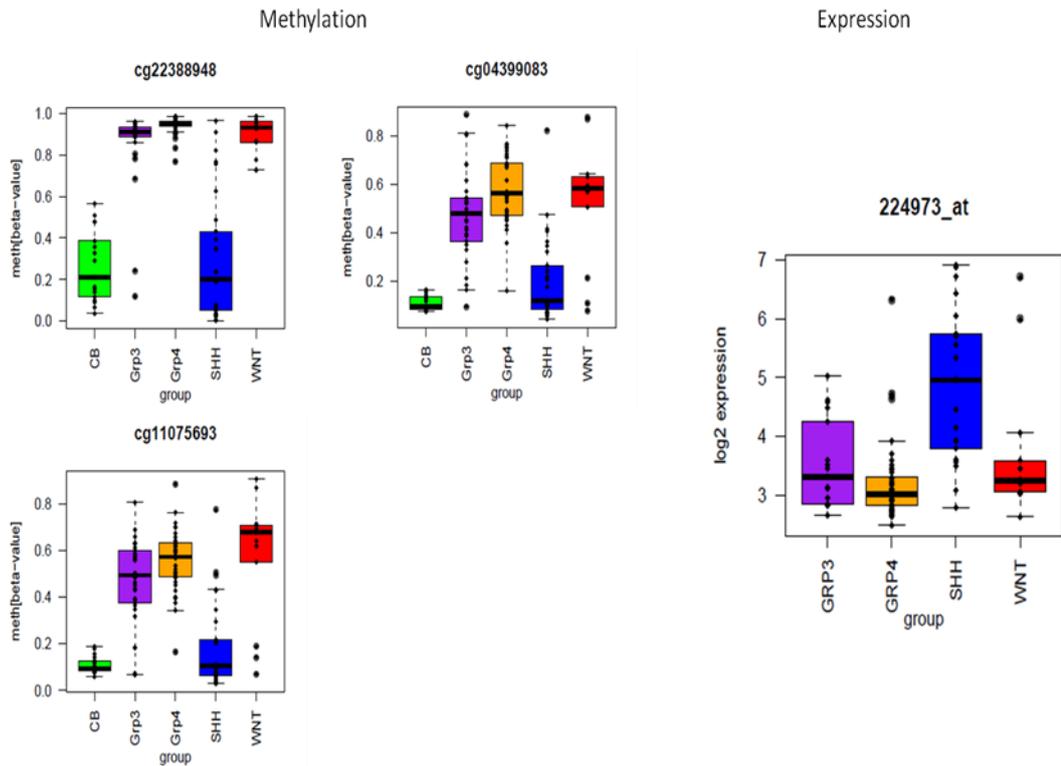
Figure 4.10. Epigenetic regulation of *FAM46A* in primary medulloblastomas. Boxplots illustrate the differential patterns of CpG methylation for the 3 critical *FAM46A* CpG probes (cg22388948, cg04399083 and cg11075693) and differential patterns of gene expression for the Kool/Fattet *FAM46A* transcript probe (224973_at). Significantly lower *FAM46A* methylation correlated with significantly higher gene expression in SHH tumours. The same pattern of significant differential gene expression was observed in Cho and Northcott datasets (not shown). Methylation profiles of critical *FAM46A* sites in SHH tumours were not significantly different to non-neoplastic cerebella (CB) (cg22388948, p=0.7, $\Delta\beta$=0.04; cg11075693, p=0.5, $\Delta\beta$=0.08; cg04399083, p=0.2, $\Delta\beta$=0.09), and were therefore not tumour-specific. These results suggest that *FAM46A* is potentially silenced by hypermethylation of critical CpG sites in WNT, Grp 3 and Grp 4 medulloblastomas.

ANXA2

Methylation

cg02072495

Expression

213503_x_at
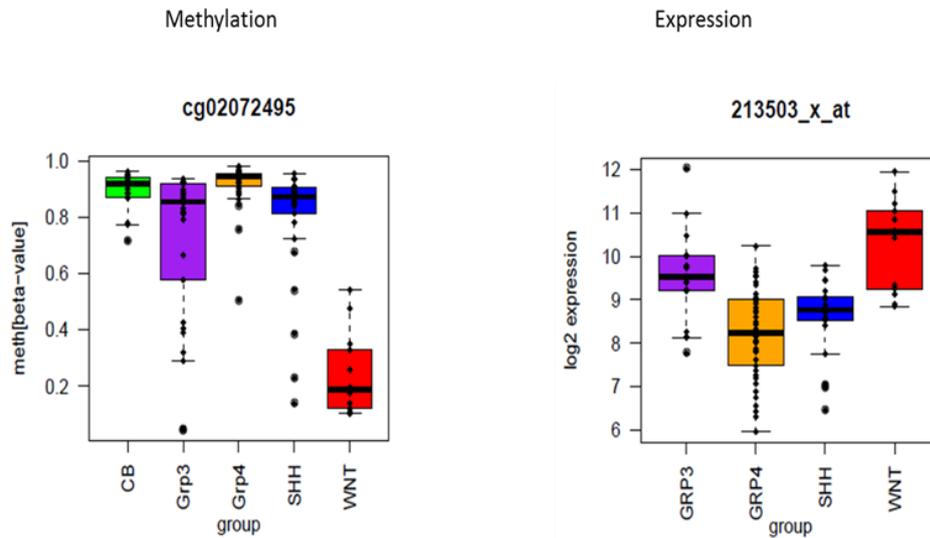
Figure 4.11. Epigenetic regulation of *ANXA2* in primary medulloblastomas. Boxplots illustrate the differential patterns of CpG methylation for critical *ANXA2* CpG probe (cg02072495) and differential patterns of gene expression for the Kool/Fattet *ANXA2* transcript probe (213503_x_at). Significantly lower *ANXA2* methylation correlated with significantly higher gene expression in WNT tumours. The same pattern of significant differential gene expression was observed in Cho and Northcott datasets (not shown). Methylation profile of critical *ANXA2* CpG site in WNT tumours was significantly different to non-neoplastic cerebella (CB) ($p<1 \times 10^{-6}$, $\Delta\beta=0.7$) and was therefore tumour-specific. These results suggest that *ANXA2* is potentially upregulated by hypomethylation of critical CpG site(s) in WNT medulloblastomas.

PRPH

Methylation

cg25763520

Expression

213847_at
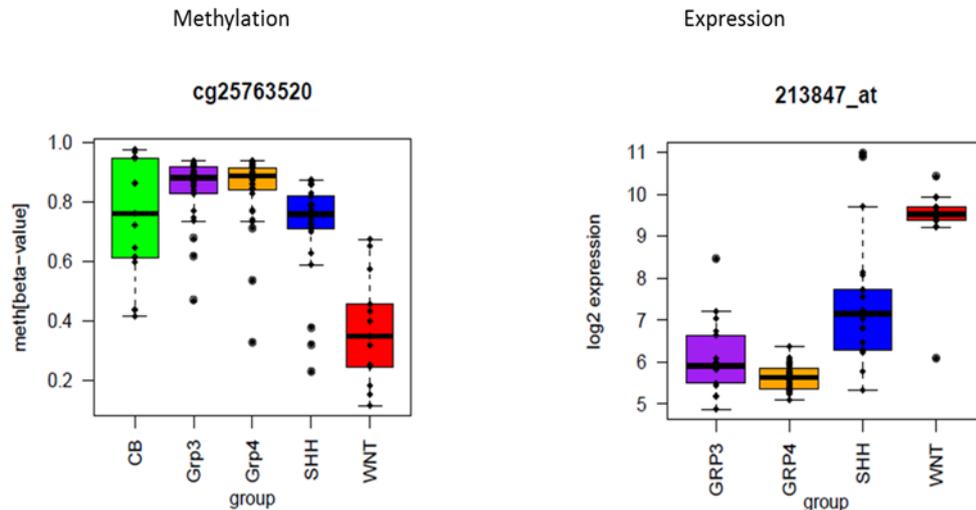
Figure 4.12. Epigenetic regulation of *PRPH* in primary medulloblastomas. Boxplots illustrate the differential patterns of CpG methylation for critical *PRPH* CpG probe (cg25763520) and differential patterns of gene expression for the Kool/Fattet *PRPH* transcript probe (213847_at). Significantly lower *PRPH* methylation correlated with significantly higher gene expression in WNT tumours. The same pattern of significant differential gene expression was observed in Cho dataset (not shown). Methylation profile of critical *PRPH* CpG site in WNT tumours was significantly different to non-neoplastic cerebella (CB) ($p<1 \times 10^{-6}$, $\Delta\beta=0.4$) and was therefore tumour-specific. These results suggest that PRPH is potentially upregulated by hypomethylation of critical CpG site(s) in WNT medulloblastomas.
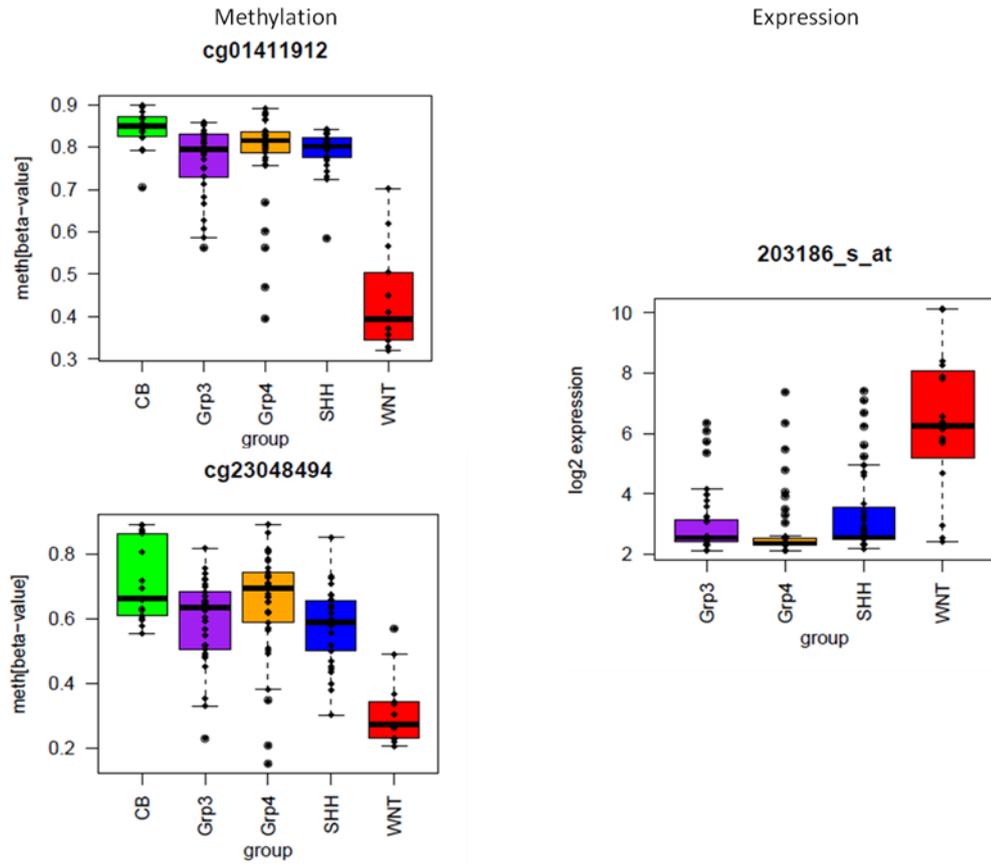
Figure 4.13. Epigenetic regulation of *S100A4* in primary medulloblastomas. Boxplots illustrate the differential patterns of CpG methylation for the 2 critical *S100A4* CpG probes (cg01411912 and cg23048494) and differential patterns of gene expression for the Cho *S100A4* transcript probe (203186_s_at). Significantly lower *S100A4* methylation correlated with significantly higher gene expression in WNT tumours. The same pattern of significant differential gene expression was observed in Kool/Fattet dataset (not shown). Methylation profiles of critical *S100A4* CpG sites in WNT tumours were significantly different to non-neoplastic cerebella (CB)(cg01411912: $p < 1 \times 10^{-6}$, $\Delta\beta = 0.4$; cg23048494: $p < 1 \times 10^{-6}$, $\Delta\beta = 0.4$), and were therefore tumour-specific. These results suggest that *S100A4* is potentially upregulated by hypomethylation of critical CpG sites in WNT medulloblastomas.
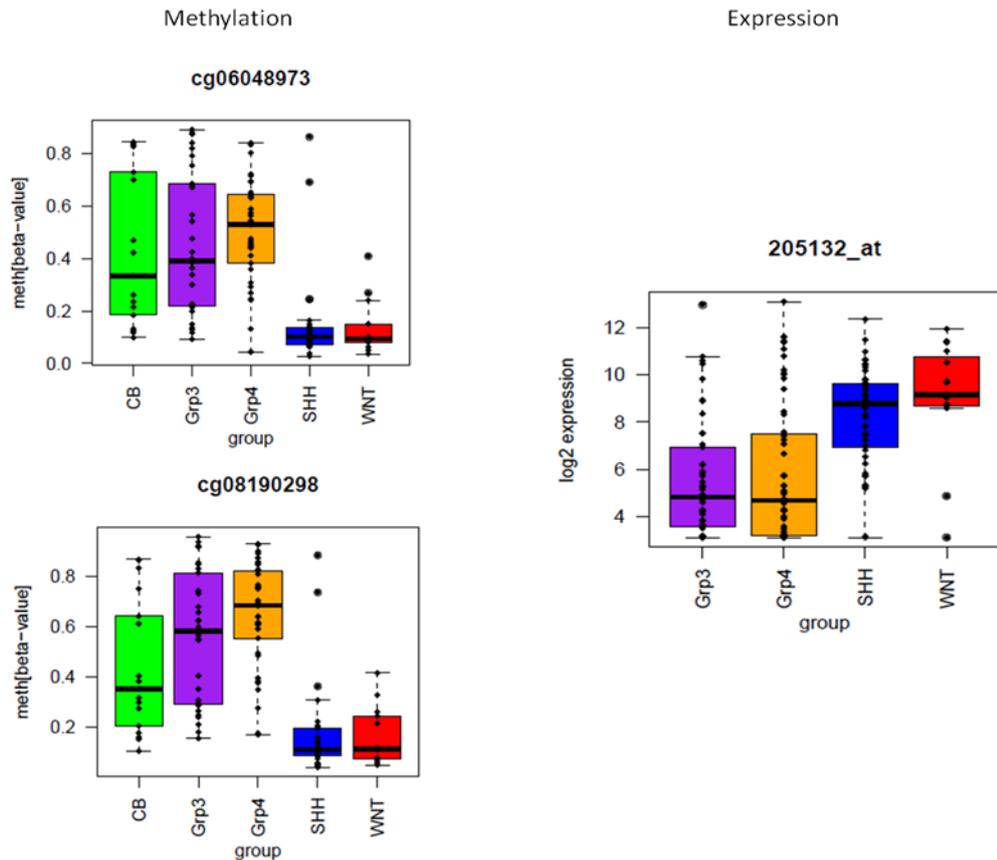
Figure 4.14. Epigenetic regulation of *ACTC1* in primary medulloblastomas. Boxplots illustrate the differential patterns of CpG methylation for the 2 critical *ACTC1* CpG probes (cg06048973 and cg08190298) and differential patterns of gene expression for the Cho *ACTC1* transcript probe (205132_at). Significantly higher *ACTC1* methylation correlated with significantly lower gene expression in Group (3+4) tumours. The same pattern of significant differential gene expression was observed in Northcott dataset (not shown). Methylation profile of critical *ACTC1* CpG probe cg08190298 in Group (3+4) tumours was significantly different to non-neoplastic cerebella (CB) (p=0.009, $\Delta\beta$=0.2) and was therefore tumour-specific. Methylation profile of critical *ACTC1* site cg06048973 in Group (3+4) tumours was not significantly different to non-neoplastic cerebella (CB) (p=0.5, $\Delta\beta$=0.04) and was therefore not tumour-specific. Methylation profiles of both sites in (SHH+WNT) tumours were significantly different to non-neoplastic cerebella (CB) (cg06048973: $p<1\times10^{-6}$, $\Delta\beta$=0.3; cg08190298: p=1$\times10^{-5}$, $\Delta\beta$=0.3), and were therefore tumour-specific in (SHH+WNT) tumours. These results suggest that *ACTC1* is potentially upregulated by hypomethylation of critical CpG sites in (SHH+WNT) medulloblastomas Non-neoplastic cerebellar samples have an intermediate methylation profile between Group (3+4) and (SHH+WNT) tumours and so potential silencing by hypermethylation of critical sites in Group (3+4) tumours cannot be excluded.

Genome-wide analysis of gene expression changes following 5-azaCdR treatment of medulloblastoma cell lines

10 untreated and 5-azaCdR- treated cell line pairs / Gene expression profiled on Illumina HT12v4.0 microarray

↓

307 transcript probes upregulated with FC > [(mean FC +3SD)] in ≥ 2 cell lines following 5-azaCdR treatment

↓

283 genes upregulated following 5-azaCdR treatment of cell lines

↓

Genome-wide assessment of CpG methylation associated with gene expression changes following 5-azaCdR treatment of medulloblastoma cell lines

Illumina 450K cell line methylation data (10 cell lines)

↓

4238 constitutively methylated and variably methylated CpG sites corresponding to upregulated genes

↓

160 critical CpG sites associated with methylation-dependent gene expression in cell lines

↓

21 candidate epigenetically regulated genes in medulloblastoma cell lines

↓

Differential methylation and gene expression analysis in primary tumours

Differential CpG methylation	Differential gene expression
32 out of 160 critical CpG sites (10 out of 21 genes) significantly differentially methylated in distinct subgroups of medulloblastoma	15 out of 21 genes significantly differentially expressed in distinct subgroups of medulloblastoma

↓

9 critical CpG sites were significantly differentially methylated in subgroups of primary medulloblastomas and showed an inverse relationship between subgroup-specific methylation and gene expression

↓

5 subgroup-specific candidate epigenetically regulated genes in primary medulloblastomas

- *FAM46A*, *ANXA2*, *PRPH*, *S100A4*, *ACTC1*

Figure 4.15. Identification of subgroup-specific candidate epigenetically regulated genes in primary tumours. Flow diagram summarises the number of genes and critical CpG sites identified at each step of the analysis, which led to the identification of 9 critical CpG sites, encompassing 5 genes, which showed strong evidence of methylation-dependent gene expression in distinct subgroups of medulloblastoma.

4.4.9 Region-wide CpG methylation and candidate gene expression

For the 5 candidate genes, patterns of CpG methylation across the gene were assessed using the 450K array. For each gene, adjacent CpG residues on the 450K array were selected encompassing the critical differentially methylated CpG site(s) that were associated with subgroup-specific gene expression.

4.4.9.1 FAM46A

The 3 critical CpG sites identified for *FAM46A* (Figure 4.10) were adjacent to each other on the 450K array and were located in a CpG island and CpG island shore within the gene body (Figure 4.16). In total, there were 4 gene body sites represented on the array and their methylation status was consistent with each other in primary tumours and in non-neoplastic cerebella (Figure 4.16), suggesting that their methylation may reflect wider gene body island and island shore methylation. This pattern was supported by the strength of the linear relationship (r=0.9) between the methylation β-value of the critical gene body site (cg04399083) and the mean methylation β-value of the 4 gene body CpG sites in primary tumours and non-neoplastic cerebella (Figure 4.17). The critical CpG site (cg04399083) was aberrantly methylated (β>0.3) in 72% of primary tumours and in 100% of these methylated tumours the wider gene body, as represented on the array, was also aberrantly methylated (β>0.3 in ≥75% gene body sites); the site was unmethylated (β≤0.3) in 28% of primary tumours and in 100% of normal cerebella and in 77% of unmethylated tumours and 88% of normal cerebella the wider gene body, as represented on the array, was also unmethylated (β≤0.3 in ≥75% gene body sites). Aberrant methylation of the critical *FAM46A* CpG sites was seen in WNT, Group 3 and Group 4 tumours (Figure 4.10). These results suggest that hypermethylation of the gene body may potentially effect *FAM46A* silencing in WNT, Group 3 and Group 4 tumours. The CpG island-associated promoter of *FAM46A* was consistently unmethylated in primary tumours and normal cerebella (Figure 4.16). *FAM46A* represents a strong candidate epigenetically regulated gene in primary medulloblastomas (Table 4.7).
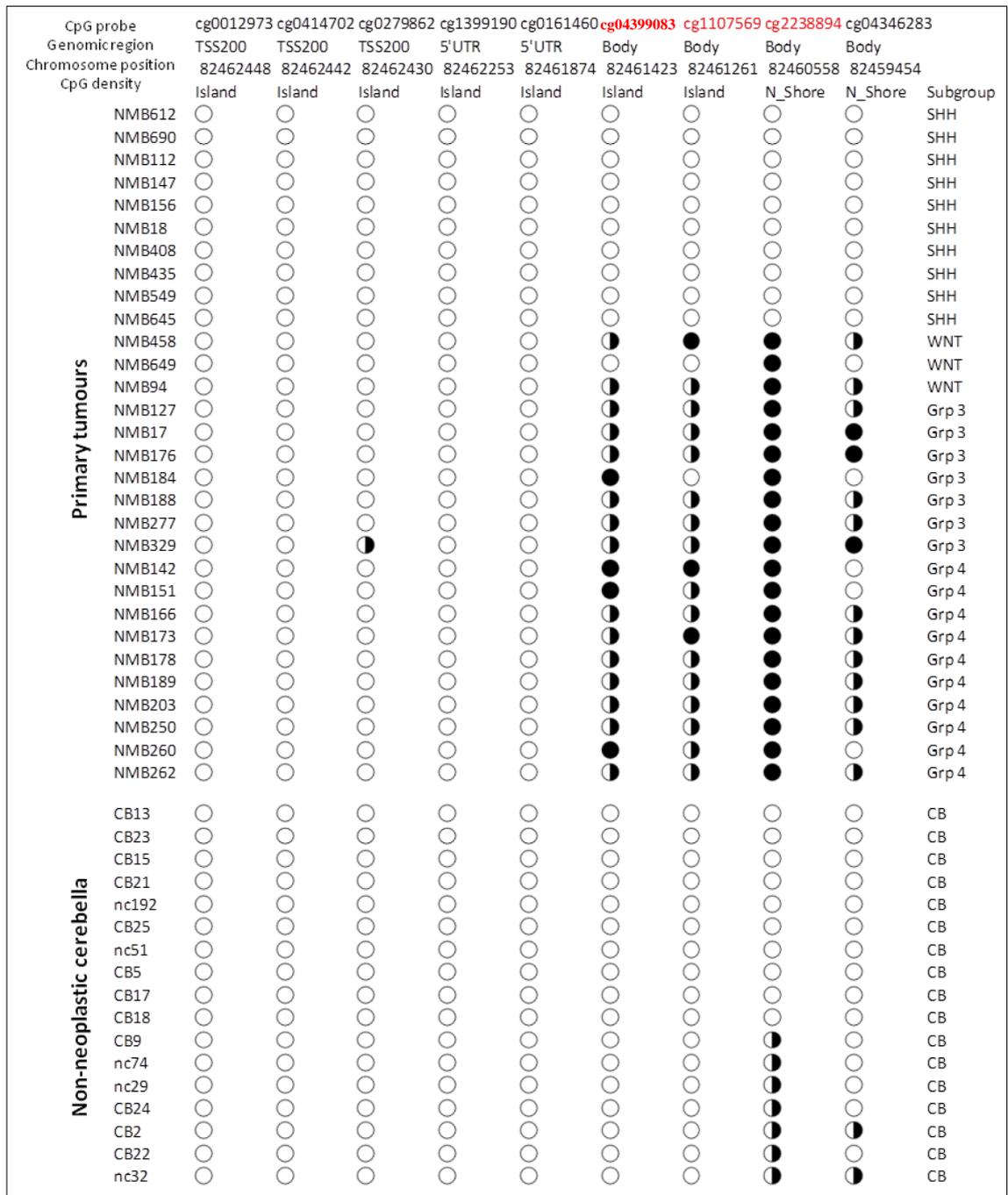
Figure 4.16. Methylation status of *FAM46A* in primary tumours and non-neoplastic cerebella. The methylation status of nine adjacent CpG probes on the 450K methylation array , encompassing the 3 critical CpG sites identified for *FAM46A* (annotated in red) are shown in a representative sample of 30 primary tumours and in normal cerebellar samples. Genomic region and CpG density were defined by Illumina: TSS200: 200base pairs upstream of transcription start site; 5'UTR: 5' untranslated region; Body: gene body; island: CpG island; N_shore: less dense CpG region located within 2kb upstream of the associated CpG island. Filled circles: β≥0.7; half-filled circles: 0.3<β<0.7; empty circles: β≤0.3. Aberrant hypermethylation of the 4 gene body sites on the array can be seen in a subset of primary tumours. *FAM46A* CpG island-associated promoter was consistently unmethylated in primary tumours and non-neoplastic cerebella. Subgroup status of primary tumours is shown. An unmethylated state (β≤0.3) is observed exclusively in SHH tumours, while hypermethylation of the gene body region (β>0.3 in ≥75% gene body sites) is observed in WNT, Group 3 (GRP3) and Group 4 (GRP4) tumours, consistent with the subgroup-specific patterns of methylation observed for the 3 individual critical CpG sites (see Figure 4.10).
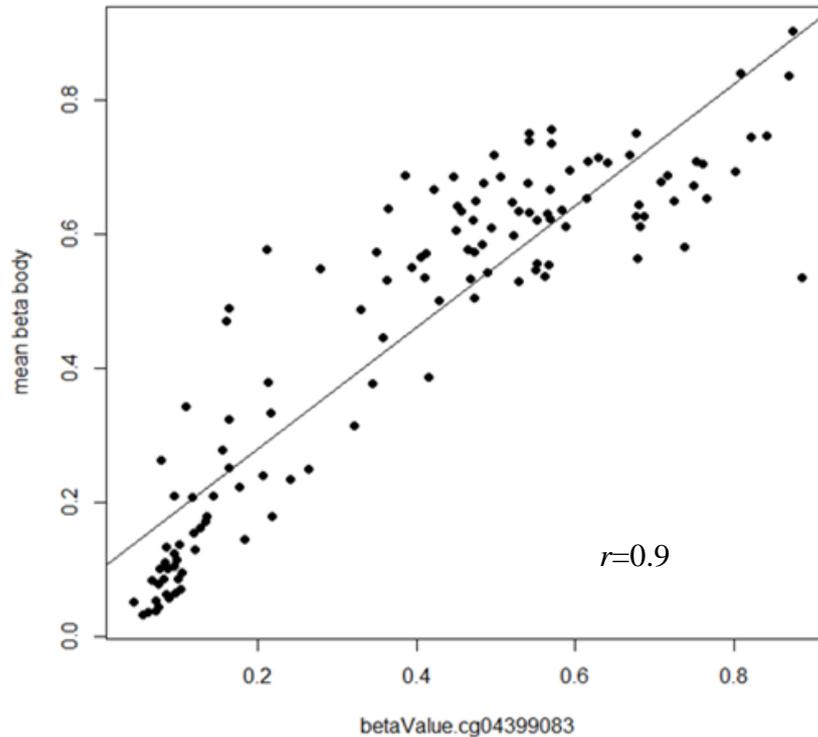
Figure 4.17 *FAM46A* gene body methylation. Linear association between the methylation β-value of the *FAM46A* gene body CpG site (cg04399083) and the mean methylation β-value of the 4 gene body sites represented on the 450K array is shown for the 109 primary tumours and 17 non-neoplastic cerebellar samples assayed on the Illumina Infinium 450K methylation array. Pearson correlation *r* value measuring the strength of the linear association is shown.

4.4.9.2 ANXA2

The critical CpG site identified for *ANXA2* (cg02072495) (Figure 4.11) was located in a CpG island shore in the 5'UTR (Figure 4.18). The site was hypomethylated in the WNT subgroup of tumours and hypomethylation was associated with significantly increased expression in WNT tumours compared with the other tumour subgroups (Figure 4.11). A hypomethylated state (β≤0.3) was observed in 69% of WNT tumours in contrast to a methylated state (β>0.3) observed in 100% of non-neoplastic cerebella and in 93% of SHH, 90% of Group 3 and 100% of Group 4 tumours. The aberrant hypomethylated state of this site in the subset of primary tumours was not reflected across the wider island and island shore-associated 5'UTR as it was represented on the array (Figure 4.18) (Table 4.7).
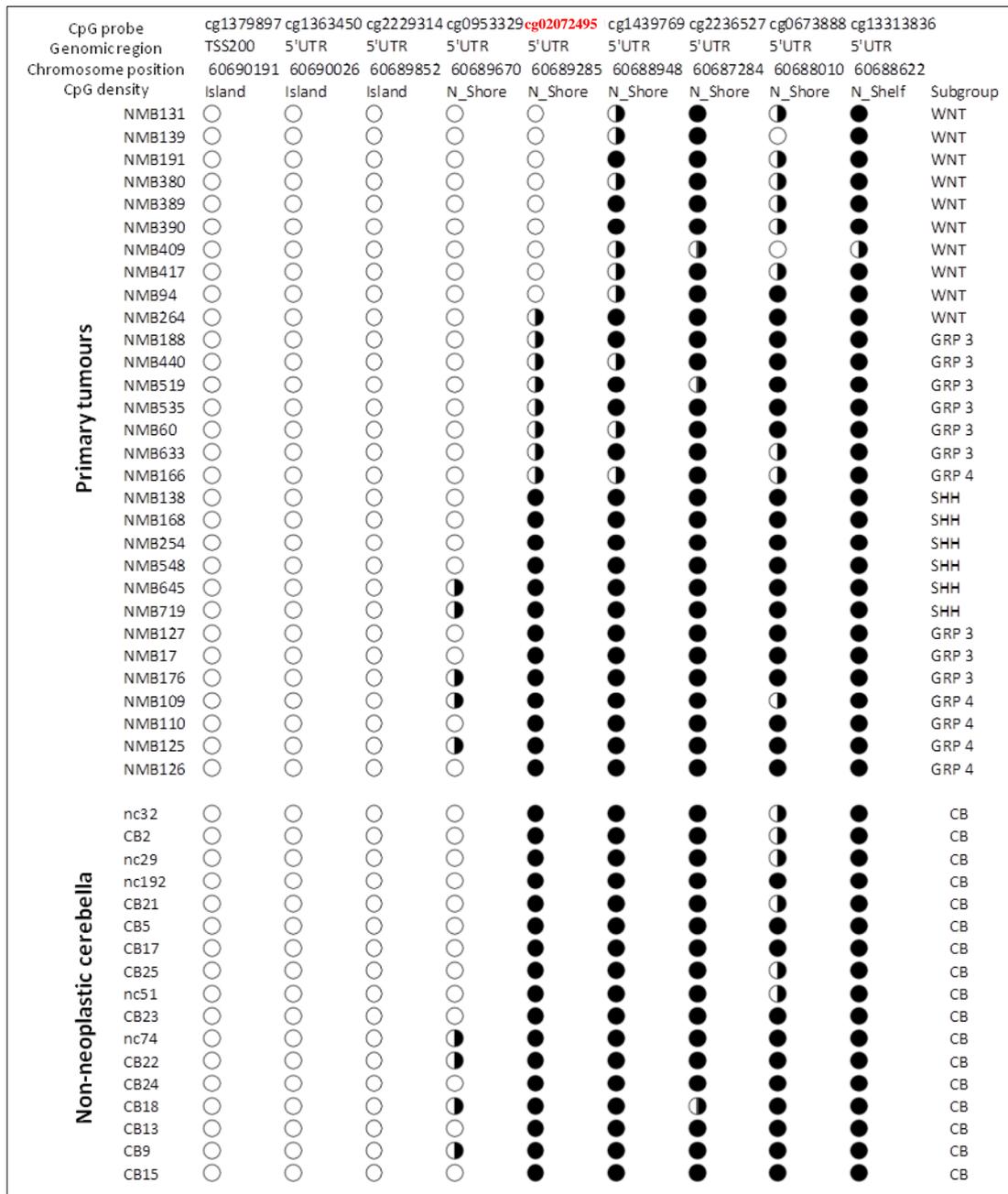
Figure 4.18. Methylation status of *ANXA2* in primary tumours and non-neoplastic cerebella. The methylation status of nine adjacent CpG probes on the 450K methylation array , encompassing the critical CpG site identified for *ANXA2*(annotated in red) are shown in a representative sample of 30 primary tumours and in normal cerebellar samples. Genomic region and CpG density were defined by Illumina: TSS200: 200base pairs upstream of transcription start site; 5'UTR: 5' untranslated region; island: CpG island; N_shore: less dense CpG region located within 2kb upstream of the associated CpG island; N_shelf: 2kb region upstream of the associated CpG island shore. Filled circles: $\beta \geq 0.7$; half-filled circles: $0.3 < \beta < 0.7$; empty circles: $\beta \leq 0.3$. Aberrant hypomethylation of the critical site can be seen in a subset of primary tumours. The methylation status of the wider 5'UTR on the array was consistent in all primary tumours and non-neoplastic cerebella. Subgroup status of primary tumours is shown. The critical CpG site is hypomethylated ($\beta \leq 0.3$) in 69% of WNT tumours. A hypomethylated state is also observed in 10% of Group 3 (GRP 3) tumours. Methylation ($\beta > 0.3$) consistent with non-neoplastic cerebella is observed in the majority of SHH, Group 3 (GRP 3) and Group 4 (GRP 4) tumours (see Figure 4.11)

4.4.9.3 PRPH

The critical CpG site identified for *PRPH* (cg25763520) (Figure 4.12) was located in a CpG island shore in a 1500 base pair block upstream of the TSS in the gene promoter (Figure 4.19). The site was hypomethylated in the WNT subgroup of tumours and hypomethylation was associated with significantly increased expression in WNT tumours compared with the other tumour subgroups (Figure 4.12). A hypomethylated state ($\beta \leq 0.3$) was observed in 38% of WNT tumours in contrast to a methylated state ($\beta > 0.3$) in 100% of non-neoplastic cerebella and in 97% of SHH, 100% of Group 3 and 100% of Group 4 tumours. In the subset of tumours where this site was hypomethylated, aberrant hypomethylation of adjacent island shore-associated promoter sites was not observed on the array (Figure 4.19) (Table 4.7). The methylation status of adjacent promoter sites on the array was variable in the non-neoplastic cerebellar samples (Figure 4.19). Interestingly, in tumours that were hypomethylated at the critical promoter site there was dense aberrant methylation of the CpG island-associated gene body region on the array (Figure 4.19).
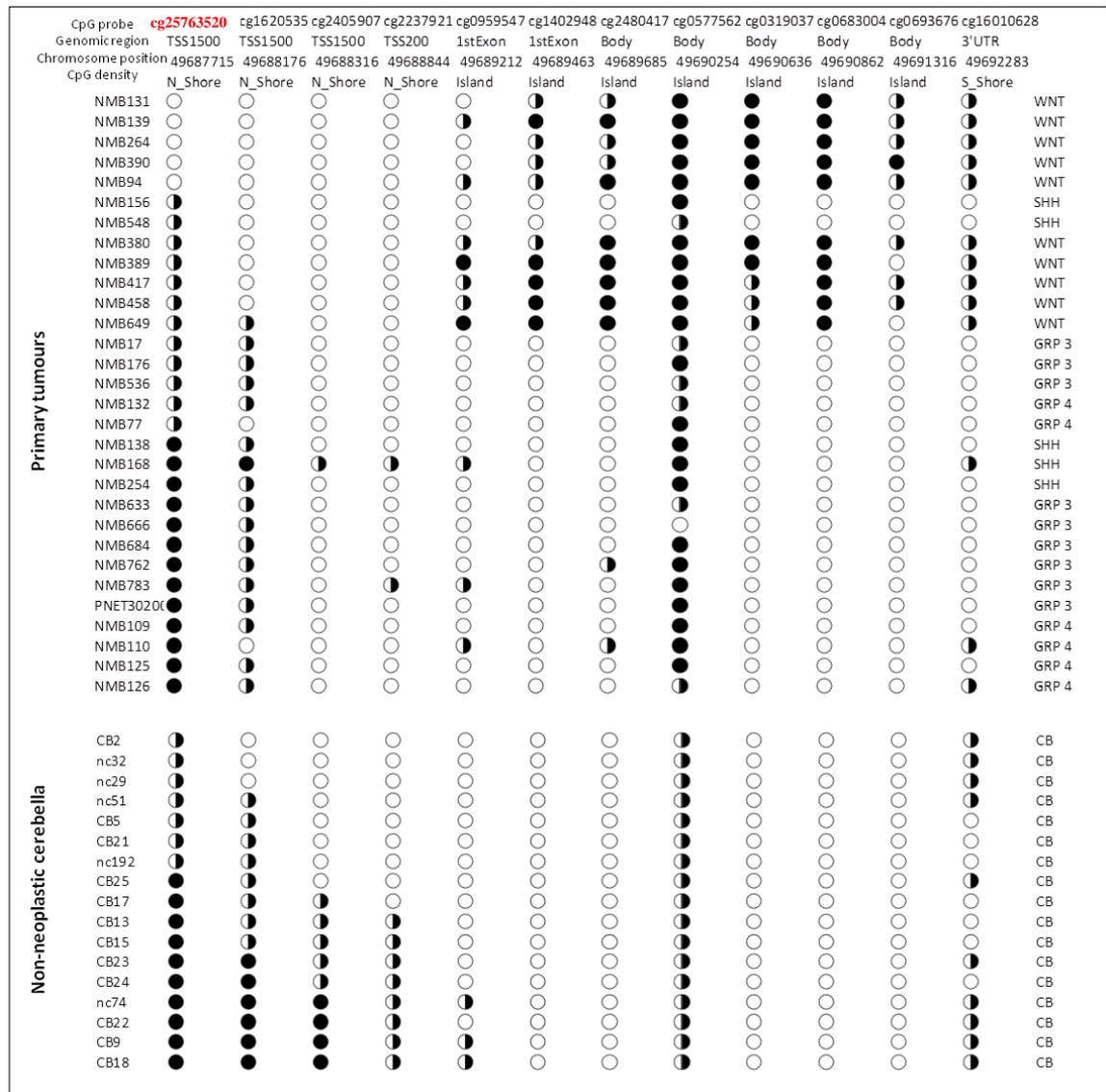
Figure 4.19. Methylation status of *PRPH* in primary tumours and non-neoplastic cerebella. The methylation status of 14 adjacent CpG probes on the 450K methylation array , encompassing the critical CpG site identified for *PRPH* (annotated in red) are shown in a representative sample of 30 primary tumours and in normal cerebella samples. Genomic region and CpG density were defined by Illumina: TSS1500: 1500base pairs upstream of transcription start site; TSS200: 200base pairs upstream of transcription start site; 5'UTR: 5' untranslated region; body: gene body; 3'UTR: 3' untranslated region; island: CpG island; N_shore: less dense CpG region located within 2kb upstream of the associated CpG island; S-shore: less dense CpG region located within 2kb downstream of the associated CpG island;. Filled circles: β≥0.7; half-filled circles: 0.3<β<0.7; empty circles: β≤0.3. Aberrant hypomethylation of the critical site can be seen in a subset of primary tumours. The methylation status of the wider gene promoter region represented on the array was variable in non-neoplastic cerebella. Tumours that were hypomethylated at the critical site demonstrated dense aberrant methylation of the CpG island-associated gene body region represented on the array. Subgroup status of primary tumours is shown. The critical CpG site is hypomethylated (β≤0.3) in 38% of WNT tumours. A hypomethylated state is also observed in 3% of SHH tumours. Methylation (β>0.3) consistent with non-neoplastic cerebella is observed in 97% of SHH and 100% of Group 3 (GRP3) and Group 4 (GRP4) tumours (see Figure 4.12).

4.4.9.4 S100A4

There were 2 differentially methylated critical CpG probes identified for *S100A4* (Figure 4.13); 1 probe (cg23048494) was located in the 5'UTR and the other (cg01411912) was located in the gene body. *S100A4* has a non CpG island-associated promoter and there were a total of 8 CpG sites represented on the 450K array and covering mainly the gene promoter (Figure 4.20).

The 2 critical sites were hypomethylated in the WNT subgroup of tumours and hypomethylation was associated with significantly increased expression in WNT tumours compared with the other tumour subgroups (Figure 4.13). For the 5'UTR site, a hypomethylated state ($\beta \leq 0.3$) was observed in 56% of WNT tumours in contrast to a methylated state ($\beta > 0.3$) in 100% of non-neoplastic cerebella and in 100% of SHH, 97% of Group 3 and 93% of Group 4 tumours. In tumours where this site was hypomethylated, aberrant hypomethylation of the wider 5'UTR and gene promoter region represented on the array was not observed (Figure 4.20) (Table 4.7). Methylation of the wider region was consistent in primary tumours and non-neoplastic cerebella (Figure 4.20). The aberrant hypomethylation of the critical gene body CpG site (cg0141191) was more subtle; this site was fully methylated ($\beta \geq 0.7$) in 100% of non-neoplastic cerebella and in 97% of SHH, 80% of Group 3 and 87% of Group 4 tumours. In contrast, the site was fully methylated in only 8% of WNT tumours and was part-methylated ($0.3 < \beta < 0.7$) in the remaining 92% of WNT tumours. This site was not unmethylated ($\beta \leq 0.3$) in any primary tumour or normal cerebellar samples. There were no other gene body CpG sites represented on the array. Aberrant hypomethylation of the 5'UTR site (cg2304849) was observed in all WNT tumours that were aberrantly hypomethylated at the gene body site (cg0141191).
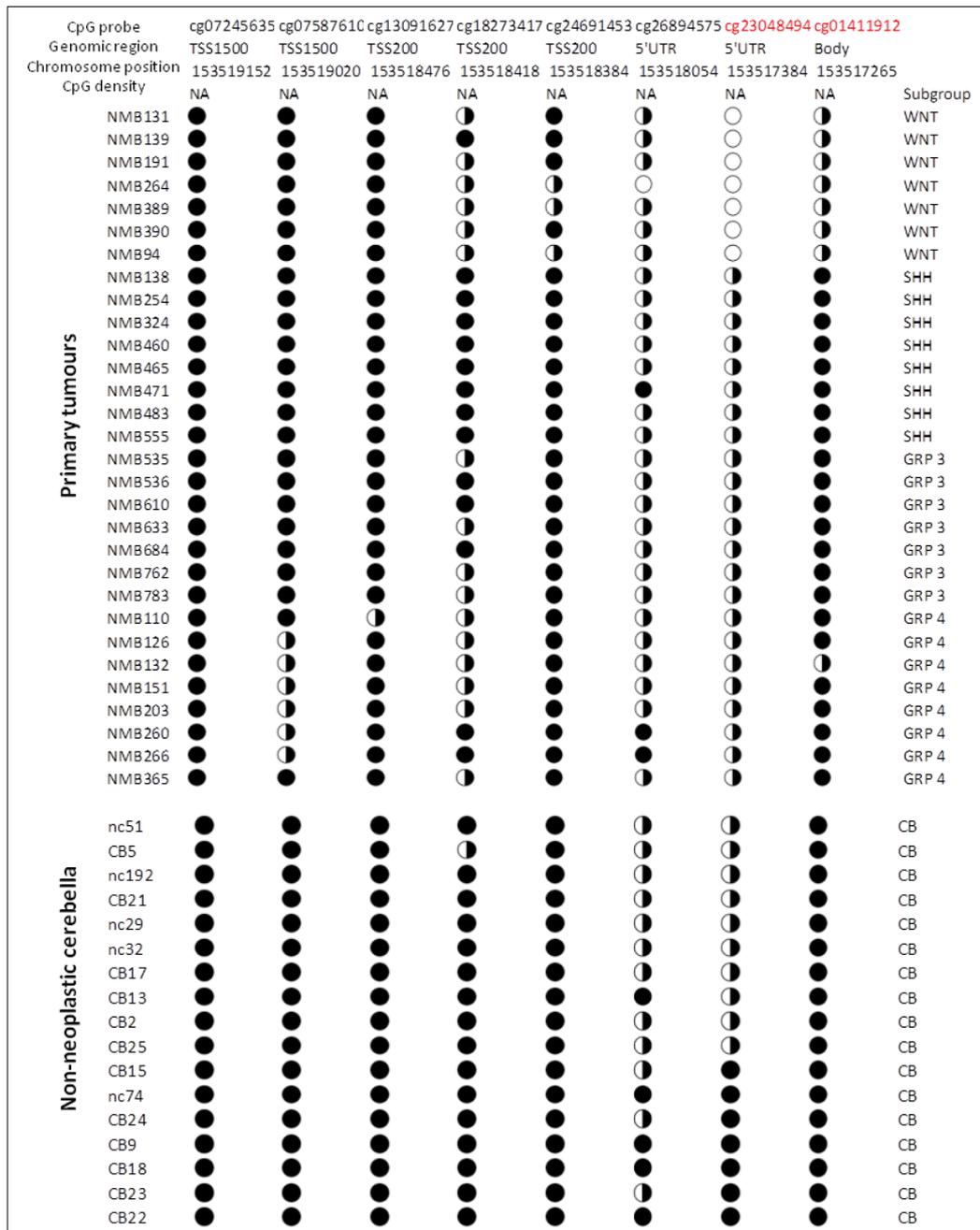
Figure 4.20. Methylation status of *S100A4* in primary tumours and non-neoplastic cerebella. The methylation status of the 8 CpG probes on the 450K methylation array are shown in a representative sample of 30 primary tumours and in normal cerebellar samples. The critical CpG sites identified for *S100A4* are annotated in red. Genomic region and CpG density were defined by Illumina: TSS1500: 1500base pairs upstream of transcription start site; TSS200: 200base pairs upstream of transcription start site; 5'UTR: 5' untranslated region; body: gene body; NA: non-CpG island site. Filled circles: β≥0.7; half-filled circles: 0.3<β<0.7; empty circles: β≤0.3. Aberrant hypomethylation of the critical sites can be seen in a subset of primary tumours. All tumours that were hypomethylated at the cg014119 site were also hypomethylated at the cg2304849 site. The wider gene promoter region represented on the array was consistently methylated in primary tumours and in non-neoplastic cerebella. Subgroup status of primary tumours is shown. Hypomethylation of the 2 critical CpG sites was observed mainly in WNT tumours, while the majority of SHH, Group 3 (GRP3) and Group 4 (GRP4) tumours were methylated consistent with the non-neoplastic cerebella (see Figure 4.13).

4.4.9.5 *ACTC1*

The 2 differentially methylated critical CpG sites identified for *ACTC1* (Figure 4.14) were adjacent to each other on the 450K array and were located in the non-CpG island promoter region (Figure 4.21). The sites showed significantly higher methylation in the Group (3+4) tumours compared with (SHH+WNT) tumours and this was associated with significantly lower gene expression in Group (3+4) compared with (SHH+WNT) (Figure 4.14). Both sites showed variable methylation in the non-neoplastic cerebella (Figure 4.14 and Figure 4.21). For the cg06048973 probe, methylation (β>0.3) was observed in 51/109 primary tumours and in 80% of these methylated tumours the wider promoter region, as represented on the array (8 CpG sites), was also methylated (β>0.3 in ≥75% promoter sites); the site was unmethylated (β≤0.3) in 58/109 primary tumours and in 79% of these unmethylated tumours the wider promoter region was similarly unmethylated (β≤0.3 in ≥75% promoter sites). Methylation (β>0.3) of the site was observed in 9/17 normal cerebellar samples but only 44% of these showed methylation of the wider region and for those that did, methylation was localised to the first 3 promoter sites on the array and was not reflective of all promoter sites (Figure 4.21). These results suggest that potential methylation-dependent silencing of *ACTC1* in the Group 3 and Group 4 tumours requires dense aberrant methylation of the gene promoter region. This is supported by the strong linear relationship (r=0.8) observed in the primary tumours and normal cerebella between the methylation β-value of the critical gene promoter site (cg06048973) and the mean methylation β-value of the 8 promoter sites represented on the array (Figure 4.22). *ACTC1* represents a strong candidate epigenetically regulated gene in primary medulloblastomas (Tabl e 4.7).
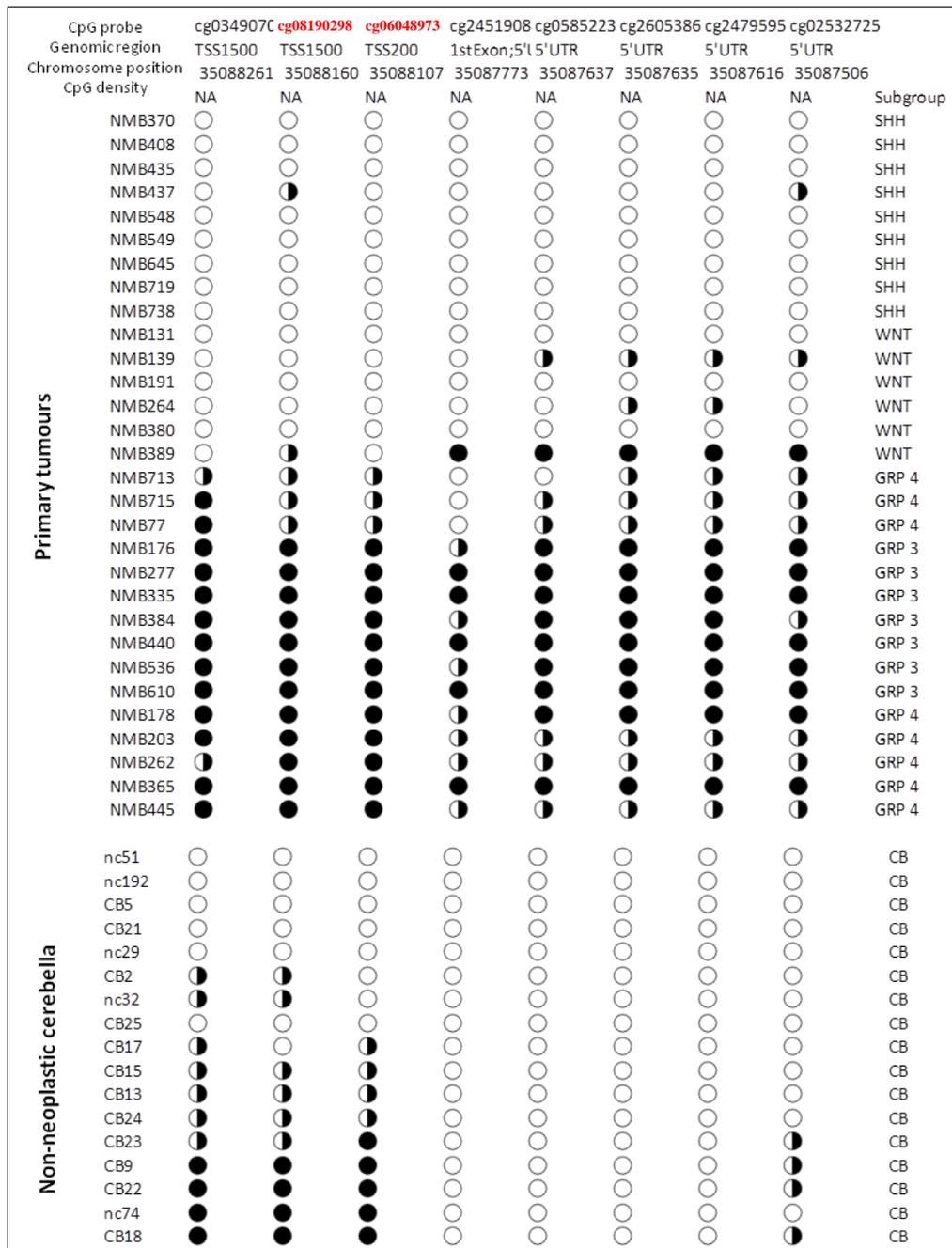
Figure 4.21. Methylation status of *ACTC1* in primary tumours and non-neoplastic cerebella. The methylation status of the 8 promoter-associated CpG sites represented on the array are shown in a representative sample of 30 primary tumours and in normal cerebellar samples.. The 2 critical CpG sites identified for *ACTC1* are annotated in red. Genomic region and CpG density were defined by Illumina: TSS1500: 1500base pairs upstream of transcription start site; TSS200: 200base pairs upstream of transcription start site; 5'UTR: 5' untranslated region ; NA: non-CpG island site Filled circles: β≥0.7; half-filled circles: 0.3<β<0.7; empty circles: β≤0.3. Methylation of the 2 critical sites in primary tumours is strongly associated with aberrant hypermethylation of the wider gene promoter region as represented on the array. Subgroup status of primary tumours is shown. Dense methylation of the gene promoter (β>0.3 in ≥75% promoter sites) was observed mainly in Group 3 (GRP3) and Group 4 (GRP4) tumours consistent with the subgroup-specific patterns of methylation observed for the 2 critical sites (see Figure 4.13).
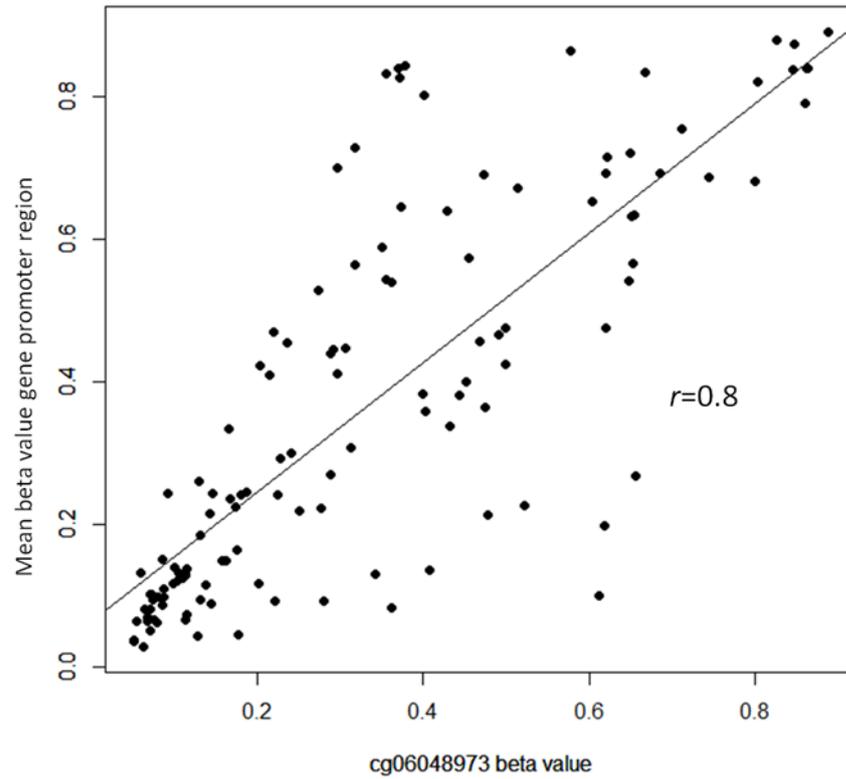
Figure 4.22. *ACTC1* **promoter methylation**. Linear association between the methylation β-value of the *ACTC1* gene promoter CpG site (cg06048973) and the mean methylation β-value of the 8 gene promoter CpG sites represented on the 450K array is shown for the 109 primary tumours and 17 non-neoplastic cerebellar samples assayed on the Illumina Infinium 450K methylation array. Pearson correlation *r* value measuring the strength of the linear association is shown.

Gene	Methylation-dependent transcriptional silencing and upregulation by 5azaCdR treatment in cell lines	Subgroup-specific methylation pattern of critical CpG sites associated with expression alterations in cell lines	Subgroup-specific inverse methylation-expression relationship	Tumour-specific methylation event(s)	Aberrant methylation state	Location of subgroup-specific critical CpG site(s)	Aberrant methylation patterns observed across multiple adjacent CpG sites
FAM46A	Y	SHH	Y	N	Hypermethylated in WNT, Group 3 and Group 4	Gene Body	Y
ACTC1	Y	Group (3+4)	Y	Y	Hypermethylated in Goup (3+4)	Gene promoter	Y
PRPH	Y	WNT	Y	Y	Hypomethylated in WNT subgroup	Gene promoter	N
ANXA2	Y	WNT	Y	Y	Hypomethylated in WNT subgroup	Gene promoter	N
S100A4	Y	WNT	Y	Y	Hypomethylated in WNT subgroup	Gene promoter and Gene body	N

Table 4.7. Summary table of subgroup-specific candidate epigenetically regulated genes. *FAM46A* and *ACTC1* are the strongest candidate genes identified in this study as aberrant methylation of multiple adjacent CpG residues may regulate their transcription in primary medulloblastomas. CpG residues adjacent to the critical site(s) identified in *PRPH*, *ANXA2* and *S100A4* do not show similar aberrant methylation patterns.

4.4.10 RNA-Seq validates microarray expression alterations following 5-azaCdR treatment in medulloblastoma cell lines

Results from the more sensitive RNA-Seq method of gene quantification validated the methylation-dependent microarray expression alterations observed in 6 medulloblastoma cell lines following 5-azaCdR treatment, for the 5 candidate epigenetically regulated genes identified in this chapter (Figure 4.23). Strong linear relationships (r>0.7) were observed between gene expression detected in the untreated and 5-azaCdR-treated cell lines by microarray profiling on the Illumina HT12 array and by RNA-Seq for the 5 candidate genes.
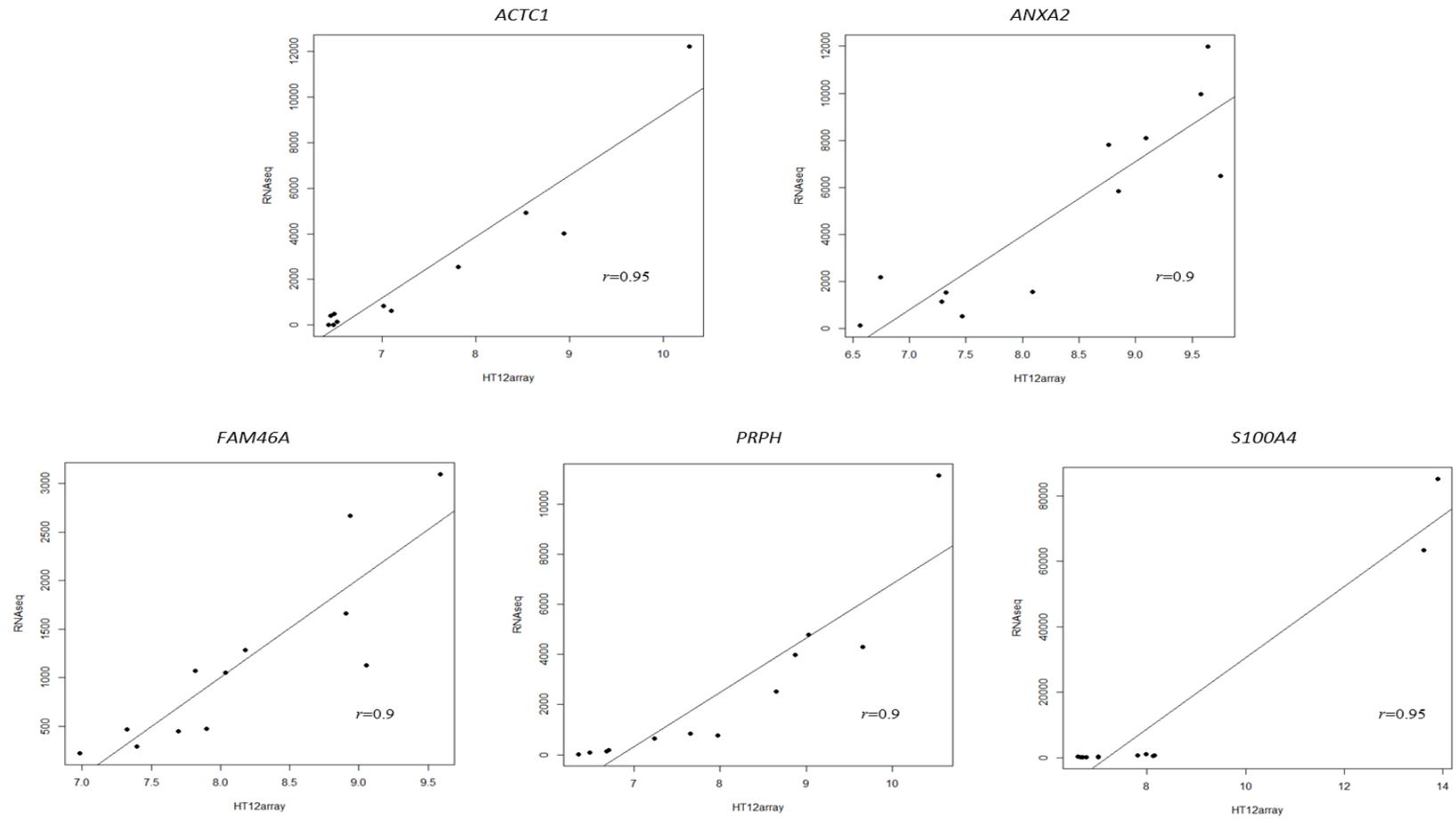
Figure 4.23. RNA-Seq versus microarray profiling of gene expression in medulloblastoma cell lines for candidate genes. Plots show the strong linear association between RNA-Seq and microarray profiling of gene expression in 6 medulloblastoma cell lines before and after treatment with 5-azaCdR. Results are shown for the 5 candidate genes. RNA-Seq transcript reads validate the methylation-dependent expression changes observed by microarray profiling for candidate genes. Pearson correlation r values measuring the strength of the linear association are shown.

4.5 Discussion

4.5.1 Upregulated gene expression following 5-azaCdR treatment in cell lines

A fold change (FC) estimation approach (see section 4.3.7) was used to identify genes whose expression was upregulated following 5-azaCdR treatment. A FC threshold (meanFC+3SD) was applied to each cell line and probes detecting an expression increase greater than the threshold following 5-azaCdR treatment in 2 or more cell lines were considered upregulated. This approach allowed cell lines to be treated independently. By doing so it broadened the scope of the analysis to facilitate the identification of both constitutive and variable methylation in cell lines that was associated with gene expression changes by 5-azaCdR consistent with potential epigenetic regulation (see section 4.4.4).

The FC thresholds applied to the cell lines ranged from 1.3 to 2.9 (Table 4.1). The minimum detection threshold of 1.3 has previously been reported in the literature for differential gene expression with the Illumina HT12v4.0 microarray (Kibriya *et al.*, 2011). Microarrays tend to have a low dynamic range compared with the more sensitive methods of qRT-PCR and RNA-Seq and they can lead to small yet biologically significant fold changes in gene expression (Mutch *et al.*, 2002). Expression microarray methods have demonstrated high accuracy to detect these small biologically important changes and results have been found to underestimate the fold change, implicating the importance of detecting small signal changes (Yao *et al.*, 2004). Signal quenching is known to occur in expression microarray analysis and is a source of non-linearity between signal intensity and gene expression (Ramdas *et al.*, 2001) and it was therefore considered important to select fold change thresholds that best described the data.

There were 283 genes upregulated by 5-azaCdR in medulloblastoma cell lines. The upregulated genes were enriched for genes encoding components of the intracellular cytoskeleton and involved in processes relating to the intracellular cytoskeleton and muscle contraction (Figure 4.4). The cytoskeleton plays important roles in establishing cell shape and in cell division, motility and cell-matrix interactions. It is now understood that in response to the cellular microenvironment mechanical signalling through the intracellular cytoskeleton can regulate cancer cell proliferation and drive

cancer progression (Provenzano and Keely, 2011), illustrating potential key roles for their epigenetic dysregulation in medulloblastoma tumourigenesis.

The upregulated genes were also enriched for stress response genes and genes involved in the regulation of apoptosis (Figure 4.4). The role of 5-azaCdR in demethylating cell lines and patient samples has been well established (Bender *et al.*, 1998; Christman, 2002; Stresemann and Lyko, 2008). It is also clear that 5-azaCdR has cytotoxic effects and as such it is a licensed epigenetic therapy for the treatment of myelodysplastic syndrome (MDS) (Stresemann and Lyko, 2008). At high doses 5-azaCdR can induce pronounced toxicities. The optimum concentration of 5-azaCdR (5µM) and the optimum time for cell line DNA demethylation and gene reactivation (72 hours) had previously been assessed and determined using the *RASSF1A* transcript by Dr. Jennifer Anderton (Anderton J.A. PhD Thesis, Newcastle University). For investigations reported in this chapter, *RASSF1A* expression in the untreated and treated cell line pairs was assessed by RT-PCR prior to microarray profiling. *RASSF1A* silencing and re-expression by 5-azaCdR was evident in all 10 cell lines tested (Figure 4.2) serving as a positive control for the demethylation of cell lines by 5-azaCdR. At these doses adapted to optimise epigenetic effects it would still be expected to observe cellular stress responses and apoptosis owing to the cytotoxic effects of 5-azaCdR. With a pharmacological-based approach using 5-azaCdR it is important to be aware that some genes may be upregulated as an indirect consequence of exposure to the drug and not directly related to their methylation. Studies have shown that some changes in gene expression following 5-azaCdR are not the direct result of demethylation but rather the result of secondary effects (Komashko and Farnham, 2010). It is therefore necessary to analyse methylation patterns in cell lines and to assess upregulated expression associated with a methylated state and consistent with a direct association between DNA methylation and gene expression.

4.5.2 Critical CpG sites associated with methylation-dependent gene regulation in cell lines

Assessment of DNA methylation in accordance with transcriptional silencing and re-expression by 5-azaCdR treatment in cell lines identified critical CpG sites that showed an association between their methylation and expression alterations seen for the upregulated genes, consistent with their potential epigenetic regulation (see section

4.4.4). The genome-wide 450K methylation array was used to assess the DNA methylation status of upregulated genes in cell lines and permitted an unbiased investigation into the potential role of gene-wide DNA methylation in gene regulation. There were over 5,251 CpG probes on the 450K array that corresponded to the 283 upregulated genes and 4,238 (81%) of these were constitutively or variably methylated in the cell lines and, therefore, potentially associated with the observed upregulated expression by 5-azaCdR.

In order for CpG sites to be considered to play a potential critical role in epigenetic gene regulation, an association between their DNA methylation status and expression alterations by 5-azaCdR treatment was required in at least 6 out of the 10 cell lines (see section 4.3.8.1). It was considered that by meeting this criterion, genes would demonstrate strong evidence of potential methylation-dependent regulation in cell lines while also recognising that different cell lines may utilise different mechanisms of gene regulation dependent on the varying selective pressures that may be acting to maintain their unlimited proliferation. A total of 160 (4%) constitutively and variably methylated CpG sites were found to have a strong association between their methylation and gene re-expression by 5azaCdR treatment in cell lines. These sites encompassed 21 of the 283 upregulated genes (7%), suggesting that for the majority of upregulated genes the changes in expression observed due to 5-azaCdR treatment were not due to the relief of repression caused by methylation, as has previously been reported (Komashko and Farnham, 2010). These results reinforce the importance of assessing CpG methylation alongside re-expression by 5-azaCdR treatment and with the identification of 21 candidate epigenetically regulated genes in medulloblastoma cell lines they highlight the utility of this functional epigenomics approach to screen for evidence of epigenetic gene regulation on a genome-wide scale. Of the 21 candidate genes identified in cell lines, 2 genes (*S100A4* and *DAZL*) had previously been reported to undergo methylation-dependent re-expression by 5-azaCdR treatment in medulloblastoma cell lines (Anderton *et al.*, 2008).

The high resolution of the 450K methylation array facilitated the identification of multiple critical CpG sites for each candidate gene (Table 4.3). Overall the sites were located across all functional gene regions (promoter, gene body and 3'UTR), with equal numbers in promoter and non-promoter regions, and 82% were situated in CpG islands

and CpG island shores/shelves (Figure 4.8). While the role of CpG island-associated promoter methylation in gene regulation has been well established, these results suggest a strong potential role for CpG island and CpG island shore methylation in regions outside of the promoter, in the epigenetic regulation of genes in medulloblastoma.

4.5.3 Differential CpG methylation associated with differential gene expression in distinct subgroups of medulloblastoma

Candidate genes which showed strong evidence of methylation-dependent transcriptional regulation in medulloblastoma cell lines were investigated for evidence of a methylation-expression relationship in primary medulloblastoma subgroups. The R package *limma* was used to assess differential methylation of the critical CpG sites identified in cell lines and differential expression of candidate genes in primary medulloblastomas (see section 4.3.10). A major practical advantage of *limma* was its capability to assess multiple contrasts simultaneously and to select the overlapping significant contrasts that differentiated one subgroup from all others. This approach and the experimental design made the differential analyses more resistant to the effects of unequal group sizes and unequal variances that were observed in the differential analyses described in Chapter 3, where distinct groups were combined together to facilitate a single pair-wise comparison using Mann-Whitney U tests (see section 3.3.4) and *t*-tests (see section 3.3.5). Another practical advantage of limma was the ease of interpretation of significance results; the 'decideTests()' function reported the results of the significance test, incorporating both the FDR-adjusted *p*-value threshold and the difference threshold into the final result and indicated whether the contrast tested was significantly negative or significantly positive. This facilitated a straightforward determination of a significant inverse methylation-expression relationship in distinct subgroups of medulloblastoma (Table 4.6).

Similar to the investigations reported in Chapter 3, a technical limitation of this study was that tumour-matched methylation and expression data was not available. As a result, direct correlation of methylation and expression could not be assessed on a sample by sample basis to support the subgroup-specific patterns of methylation and expression observed, and findings were based on a comparison of means only.

For the majority of critical CpG sites that showed an association with gene expression in cell lines, this relationship was not reflected in the differential patterns of methylation

and gene expression observed in the primary tumour subgroups; 9 of the 160 CpG sites identified in cell lines (6%) showed differential methylation associated with differential gene expression and consistent with their potential epigenetic regulation in distinct subgroups of medulloblastoma (see section 4.4.8). The sites encompassed 5 of the 21 genes (24%) identified in cell lines. These results suggest that patterns of differential methylation in medulloblastoma subgroups are not strongly associated with the patterns of differential gene expression, consistent with the findings in Chapter 3 which investigated Group 3- and Group 4-specific methylation events (see section 3.5.3).

4.5.4 Functional epigenomics identifies subgroup-specific candidate epigenetically regulated genes in medulloblastoma

By carrying out a functional epigenomics screen in medulloblastoma cell lines, followed by a differential methylation/expression analysis of critical events in primary medulloblastomas, the investigations reported in this chapter have identified 5 candidate genes that showed strong evidence of methylation-dependent expression in distinct molecular subgroups of the disease; *FAM46A* hypermethylation was specific to non-SHH tumours; *ANXA2* hypomethylation, *PRPH* hypomethylation and *S100A4* hypomethylation were specific to WNT tumours and *ACTC1* hypermethylation was specific to Group (3+4) tumours (Table 4.7). Epigenetic deregulation of *S100A4* has previously been reported in medulloblastoma (Lindsey *et al.*, 2007) while the other 4 genes represent novel genes potentially involved in the pathogenesis of medulloblastoma. *FAM46A* and *ACTC1* were the strongest candidate epigenetically regulated genes identified in this study as their aberrant methylation patterns extended across multiple adjacent CpG residues (Table 4.7).

Since the 4 distinct molecular subgroups of medulloblastoma potentially arise from distinct stem and progenitor cell populations (Schuller *et al.*, 2008; Yang *et al.*, 2008; Gibson *et al.*, 2010; Kawauchi *et al.*, 2012; Pei *et al.*, 2012) an accurate assessment of tumour versus normal tissue should account for these subgroup-specific differences and comparisons should be made to the correct cell type from which the tumour originates. Control cerebella samples will contain a heterogeneous population of cells that may mask the signal from the actual cell of origin. Until the cell of origin can be identified for all the subgroups of medulloblastoma and in the absence of appropriate cell type-matched controls, the tumour-specificity of the candidate gene methylation events was

considered only to appraise the potential functional relevance of methylation-dependent gene expression in the tumour subgroups.

4.5.4.1 *FAM46A*

There were 4 critical CpG sites belonging to the *FAM46A* gene whose methylation was associated with methylation-dependent expression changes in medulloblastoma cell lines treated with 5-azaCdR (Table 4.3). RNA-Seq results for 6 of the untreated and 5-azaCdR-treated cell line pairs showed strong correlation with the expression microarray results (r=0.9) (Figure 4.23), validating the methylation-dependent expression changes observed and providing further supporting evidence of epigenetic regulation. Three of the four sites identified in cell lines were differentially methylated in SHH tumours compared to the other subgroups and their methylation was in accordance with differential expression in the SHH subgroup and potential epigenetic regulation in primary tumours (Figure 4.10). The methylation status of the 3 CpG sites in SHH tumours was not significantly different to the normal cerebella and they represented aberrant hypermethylation events in the WNT, Group 3 and Group 4 tumours (Figure 4.10). The 3 sites were located in the gene body of *FAM46A* in a CpG island and island shore and assessment of adjacent sites on the 450K array suggests that aberrant methylation in non-SHH tumours may be reflective of aberrant methylation across the wider CpG island and island shore-associated gene body region (Figure 4.16). These findings are based on only 4 gene body CpG probes on the 450K array and would require validation by bisulfite sequencing of the CpG island-associated gene body.

Based on the findings of this investigation, *FAM46A* represents a potential novel tumour suppressor gene in WNT, Group 3 and Group 4 medulloblastomas, which may be silenced by aberrant methylation of a CpG island and island shore located in the gene body. *FAM46A* (family with sequence similarity 46, member A) is located on chromosome 6q14 and the normal functions of the encoded protein are unknown. There is very little reported in the literature on any role for this gene in tumourigenesis. One study has reported differential *FAM46A* expression in breast cancer patients who respond and who don't respond to 5-fluorouracil (5-FU) treatment and suggest that it may represent a predictive marker of response to therapy in breast cancer (Tsao *et al.*, 2010).

4.5.4.2 ANXA2, PRPH and S100A4

ANXA2, *PRPH* and *S100A4* had similar patterns of differential methylation and gene expression specific to the WNT subgroup of tumours (Figure 4.11, Figure 4.12 and Figure 4.13). There were 6 critical *ANXA2* CpG sites, 4 *PRPH* sites and 5 S100A4 sites whose methylation was associated with methylation-dependent expression changes in medulloblastoma cell lines treated with 5-azaCdR (Table 4.3). RNA-Seq results for 6 of the untreated and 5-azaCdR-treated cell line pairs showed strong correlation with the expression microarray results (Figure 4.23), validating the methylation-dependent expression changes observed and providing further supporting evidence of epigenetic regulation.

Of the CpG sites that were identified in cell lines, there was 1 critical site in the *ANXA2* gene, 1 in the *PRPH* gene and 2 in the *S100A4* gene that were differentially methylated in WNT tumours compared to the other tumour subgroups, and their methylation was in accordance with differential expression in the WNT subgroup and potential epigenetic regulation in primary tumours (Figure 4.11, Figure 4.12 and Figure 4.13). The methylation status of the sites in WNT tumours was significantly different to the normal cerebella for the 3 genes and they represented hypomethylation events in the WNT tumours (Figure 4.11, Figure 4.12 and Figure 4.13).

For all 3 genes, assessment of adjacent probes on the 450K array did not support strong evidence of the site-specific hypomethylation being reflected across the wider gene region. In the case of *S100A4* there was only 1 gene body site represented on the array and so methylation patterns across the wider gene body region could not be investigated by microarray analysis. The findings reported here are based on only those sites that are present on the 450K array and they would require validation by bisulfite sequencing of the surrounding region for each candidate gene. While aberrant methylation states affecting multiple adjacent CpG residues provides more convincing evidence of potential epigenetic regulation, methylation at a single CpG residue has been shown to be sufficient to negatively affect binding of transcription factors to the DNA (Tate and Bird, 1993). In other studies, epigenetic regulation of *S100A4* has been shown to occur at 2 critical intragenic CpG sites located within its first intron (Rosty *et al.*, 2002; Lindsey *et al.*, 2007).

Based on the results reported here, *ANXA2*, *PRPH* and *S100A4* represent potential oncogenes in WNT medulloblastomas, which may be upregulated by hypomethylation of critical CpG sites. *S100A4*, located on chromosome 1q21, is known to be regulated epigenetically by DNA methylation of key intragenic CpG sites (Rosty *et al.*, 2002) and hypomethylation of *S100A4* has been found in medulloblastoma (Lindsey *et al.*, 2007) as well as in other cancers, including pancreatic and colon adenocarcinoma (Nakamura and Takenaga, 1998; Rosty *et al.*, 2002). High levels of *S100A4* have been shown to be associated with metastatic progression of many cancers including medulloblastoma (Hernan *et al.*, 2003). This study has found *S100A4* hypomethylation to occur almost exclusively in the WNT subgroup of tumours (see section 4.4.9.4) and has identified a potential critical site located within the gene promoter that may play a role in gene regulation.

PRPH is located on chromosome 12q12-q13 and encodes a cytoskeletal protein found in neurons of the peripheral nervous system. A recent study in neuroblastoma has identified promoter methylation of *PRPH* as a prognostic biomarker, reporting that reduced expression levels associated with a methylated state contributed to a more aggressive phenotype of the disease (Decock *et al.*, 2012). The authors report that high levels of *PRPH* are associated with a more differentiated state and suggest that it may also serve as a differentiation marker for tumours derived from the neural crest (Decock *et al.*, 2012).

ANXA2, located on chromosome 15q22.2, encodes a member of the calcium-dependent phospholipid binding proteins that play a role in the regulation of cellular growth and signal transduction pathways. Increased expression of *ANXA2* correlates with malignant changes in multiple cancers and in breast cancer elevated expression has been shown to drive proliferation, migration and invasion of cells through regulation of downstream targets including c-myc and cyclin D1 (Wu *et al.*, 2012).

4.5.4.3 ACTC1

There were 7 critical CpG sites belonging to the *ACTC1* gene whose methylation was associated with methylation-dependent expression changes in medulloblastoma cell lines treated with 5-azaCdR (Table 4.3). RNA-Seq results for 6 of the untreated and 5-azaCdR-treated cell line pairs showed strong correlation with the expression microarray results (r=0.95) (Figure 4.23), validating the methylation-dependent expression changes observed and providing further supporting evidence of epigenetic regulation. Two of the seven sites identified in cell lines were differentially methylated between Group (3+4) and (SHH+WNT) tumours and their methylation was in accordance with differential expression between the same groups and potential epigenetic regulation in primary tumours (Figure 4.14). Methylation of the 2 CpG sites in control cerebella was variable and for 1 probe (cg06048973) it was significantly hypomethylated in (SHH+WNT) tumours, while the other probe (cg08190298) was significantly hypermethylated in Group (3+4) and significantly hypomethylated in (SHH+WNT) (Figure 4.14). The 2 sites were located in the non CpG island-associated promoter region of *ACTC1*, 200bp and 1500bp upstream from the TSS (Figure 4.21). Assessment of adjacent promoter sites on the 450K array suggests that, while methylation of the critical sites was variable in normal cerebella, aberrant dense methylation of the wider promoter region occurred in tandem with methylation of these sites in primary tumours (Figure 4.21). These findings are based on probes represented on the 450K array and would require validation by bisulfite sequencing of the wider *ACTC1* promoter.

Based on the findings of this investigation, *ACTC1* represents a potential novel tumour suppressor gene in Group 3 and Group 4 medulloblastomas, which may be silenced by aberrant dense methylation of its gene promoter. *ACTC1* is located on chromosome 15q14 and encodes for a member of the actin family of proteins, which are involved in various types of cell motility. Studies in cancer have found that ACTC1 is commonly differentially expressed among subclasses of bladder cancer (Zaravinos *et al.*, 2011), and based on its expression profile it has been identified as a cancer-specific biomarker in prostate cancer (Huang *et al.*, 2010).

4.5.5 Genes previously identified to be regulated by DNA methylation in medulloblastoma

The candidate epigenetically regulated genes identified in this chapter include one gene, *S100A4*, which has previously been found to be regulated by DNA methylation in primary medulloblastomas (Lindsey *et al.*, 2007). The 4 remaining genes (*ACTC1*, *ANXA2*, *FAM46A* and *PRPH*) represent novel medulloblastoma genes. *COL1A2* (Anderton *et al.*, 2008), *SPINT2* (Kongkham *et al.*, 2008) and the *SFRP* family of WNT inhibitors (Kongkham *et al.*, 2010a) have previously been identified to be epigenetically silenced in primary medulloblastomas using the same pharmacological-based genome-wide functional screening approach in cell lines that was employed for investigations reported here, yet it must be noted that they were not identified as candidate genes in this study.

In the studies reported by Kongkham et al. that identified *SPINT2* (Kongkham *et al.*, 2008) and the *SFRP* family (Kongkham *et al.*, 2010a), the Affymetrix HGU133A+2 was used to assess gene expression changes following demethylating treatment of cell lines with 5-azaCdR. These genes were not found to be upregulated following 5-azaCdR treatment of cell lines on the Illumina HT12 platform used in this study, suggesting that this microarray-based approach is sensitive to the type of platform used to assess gene expression alterations. Discrepancies among microarray studies using different platforms are frequently reported and this can often be attributed to the different analytical frameworks employed for data analysis (Konishi *et al.*, 2008), which may explain why these genes were not identified on the different platform used in this investigation.

The Affymetrix HG-U133A platform was used in the study that identified epigenetic regulation of *COL1A2* in primary medulloblastomas (Anderton *et al.*, 2008). This study investigated methylation-dependent expression alterations by microarray analysis of 3 medulloblastoma cell lines and validated the findings in 2 of the cell lines by qRT-PCR. Using the Illumina HT12 microarray and a larger panel of 10 medulloblastoma cell lines, *COL1A2* expression was upregulated following 5-azaCdR treatment in cell lines in the investigations reported in this chapter, however, following analysis of its methylation status in the 10 cell lines used in this study, it did not meet the criterion required to be considered a candidate epigenetically regulated gene in cell lines; an

association between methylation status and expression alterations was not observed in at least 6 out of the 10 cell lines. Based on their findings of methylation-dependent gene expression in 3 medulloblastoma cell lines, Anderton et al went on to show methylation-dependent expression of *COL1A2* in primary tumours (Anderton *et al.*, 2008). This discrepancy in findings by using a larger panel of cell lines may be caused by some cell lines acquiring artifactual gains of methylation that are not associated with expression alterations (Bender *et al.*, 1998; Smiraglia *et al.*, 2001), reminding us of the need to exercise caution with cell line models and of the importance of carrying out validation studies in primary tumours. The requirement to show an association between CpG methylation and expression alterations in 6 out of 10 cell lines may have been too strict for the *COL1A2* gene and perhaps this may be explained by the different cell lines utilising different mechanisms of *COL1A2* gene regulation depending on selective pressures operating to maintain their unlimited growth. If this were the case, it highlights the importance of analysing the cell lines independently and not assuming they will behave collectively.

In this study, assessment of *RASSF1A* expression by RT-PCR in the untreated and 5-azaCdR treated cell line pairs found methylation-dependent transcriptional silencing and re-activation of *RASSF1A* by 5-azaCdR in all 10 cell lines (see section 4.4.1). These results were consistent with the robust evidence for epigenetic regulation of *RASSF1A* by DNA methylation in medulloblastoma cell lines (Lusher *et al.*, 2002). However, *RASSF1A* was not identified as upregulated following 5-azaCdR by microarray analysis of the same panel of 10 untreated and 5-azaCdR treated cell lines. RT-PCR primers were designed specifically against the isoform A transcript of the *RASSF1* gene (see section 2.9.3). The Illumina HT12 array does not measure expression of individual exons and, therefore, does not distinguish between transcript variants of a particular gene and so it is likely that the expression of *RASSF1A* is being masked by other *RASSF1* transcript variants. This discrepancy in *RASSF1A* detection highlights a limitation of this array to detect expression alterations that may affect particular transcript variants of a gene.

RHOH was identified in Chapter 3 as a potential oncogene, upregulated in Group 4 medulloblastomas by hypomethylation of a critical promoter-associated CpG site (see section 3.5.6). qRT-PCR assessment of expression changes following 5-azaCdR

treatment in medulloblastoma cell lines found *RHOH* to be re-expressed in 4 cell lines following 5-azaCdR treatment and re-expression was associated with methylation in these cell lines, consistent with epigenetic regulation (see section 3.4.9.5). In the investigations reported in this chapter, microarray analysis of expression changes in the same cell lines did not identify upregulated *RHOH* expression following 5-azaCdR treatment. This discrepancy in RHOH detection using different methods suggests that high-throughput microarray expression profiling may be limited in its sensitivity to detect expression alterations of some genes following demethylating treatment of cell lines with 5-azaCdR.

RNA-Seq is a more sensitive method of quantification of transcripts and their isoforms and it has a better dynamic range compared with microarray technologies (Wang *et al.*, 2009). RNA sequencing results that became available towards the end of this project validated the expression alterations detected by microarray in 6 cell lines for the candidate genes identified (Figure 4.23). Many of the discrepancies in gene detection that have been recognised here as a result of different microarray platforms and different methods of detection may be overcome by using the more sensitive and high-throughput RNA-Seq technology to perform a genome-wide screen of expression alterations by demethylating treatment in cell lines. A universal analytic pipeline should be developed to facilitate consistency between studies. The use of RNA-Seq technology to identify epigenetically regulated genes in medulloblastoma is discussed further in Chapter 6.

4.5.6 *Assessment of RHOH methylation using the Illumina 450K methylation array*

Using the Ilumina GoldenGate Cancer Panel I methylation microarray, investigations reported in Chapter 3 identified *RHOH* as a potential candidate oncogene in Group 4 medulloblastomas; upregulated *RHOH* expression in Group 4 tumours was associated with hypomethylation of a critical promoter-associated CpG site (see section 3.5.6). Following investigations in this chapter which used the latest generation Illumina 450K methylation microarray, it was decided to assess methylation of the wider *RHOH* promoter region using this higher density methylation array.

The critical *RHOH* CpG site, which was associated with gene expression and was identified from the GoldenGate methylation array in Chapter 3 (RHOH_P121_F), was

located 121 base pairs upstream from the transcription start site in the non-CpG island-associated gene promoter. In the cohort of 216 primary medulloblastomas (see section 2.1.1) and 21 non-neoplastic cerebellar samples (see section 2.3) that were analysed on the GoldenGate methylation array, this site was methylated ($\beta > 0.3$) in 100% of non-neoplastic cerebella, 100% of SHH tumours, 82% of WNT tumours and 79% of Group 3 tumours. In contrast, aberrant hypomethylation ($\beta \leq 0.3$) was observed in 51% of Group 4 tumours (Figure 4.24). This same site was represented on the 450K methylation array (cg26296101), and showed similar patterns of subgroup-specific and tumour-specific methylation in the smaller, and largely independent, cohort of 109 tumours (see section 2.1.2) and 17 non-neoplastic cerebellar samples (see section 2.3) that were analysed on the 450K array. In the 450K cohort, methylation ($\beta > 0.3$) was observed in 100% of non-neoplastic cerebella, 100% of SHH tumours, 100% of WNT tumours and 80% of Group 3 tumours. Aberrant hypomethylation ($\beta \leq 0.3$) was observed in 30% of Group 4 tumours (Figure 4.24).
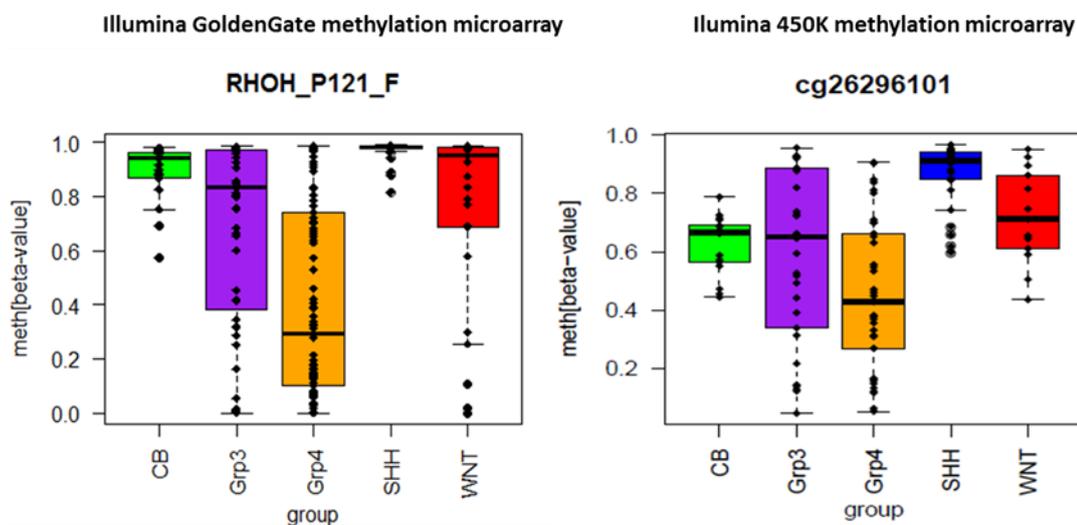
Figure 4.24 Methylation patterns of the critical *RHOH* CpG site in primary tumours and non-neoplastic cerebella assessed using the GoldenGate methylation microarray and the Infinium 450K methylation microarray. Combined boxplots and strip-plots show differential patterns of CpG methylation for the same critical RHOH CpG site assessed on the GoldenGate methylation microarray (RHOH_P121_F) and assessed on the Infinium 450K methylation microarray (cg26296101)**,** in different cohorts of primary tumours and non-neoplastic cerebellar samples. A cohort of 216 primary tumours and 21 non-neoplastic cerebella were analysed on the GoldenGate methylation microarray. A largely independent cohort of 109 tumours and 17 non-neoplastic cerebellar samples were analysed on the 450K methylation microarray. The same differential patterns were observed in different cohorts assessed on different microarray platforms.

There were 7 adjacent promoter-associated sites in the 450K methylation data for *RHOH* (Figure 4.25). Patterns of aberrant hypomethylation equivalent to the critical CpG site (cg2629610), which was observed predominantly in Group 4 tumours, were not observed across all promoter-associated sites, but were observed at sites immediately adjacent to the critical site on the array (Figure 4.25). The residue immediately upstream of this site on the array (cg0565810) showed a consistent methylation status; the critical site was aberrantly hypomethylated ($\beta \leq 0.3$) in 17 out of 109 tumours and in 13 of these hypomethylated tumours (76%) the adjacent upstream residue was also hypomethylated. In 99% of methylated tumours ($\beta > 0.3$) the adjacent upstream residue was also methylated. The residue immediately downstream of the critical site on the array (cg0080439) was unmethylated ($\beta \leq 0.3$) in 66 out of 109 (61%) primary tumours compared with only 2 out of 17 (12%) normal cerebella. Interestingly this aberrant hypomethylated state was observed in 100% of Group 4 tumours, which represented 56% of the unmethylated tumours. The remaining promoter sites were consistently methylated in all primary tumours and non-neoplastic cerebella (Figure 4.25). Validation studies investigating methylation-dependent upregulation of *RHOH* in Group 4 tumours should focus on the critical CpG site identified (cg2629610) and sites which are immediately upstream and downstream of it in the gene promoter.
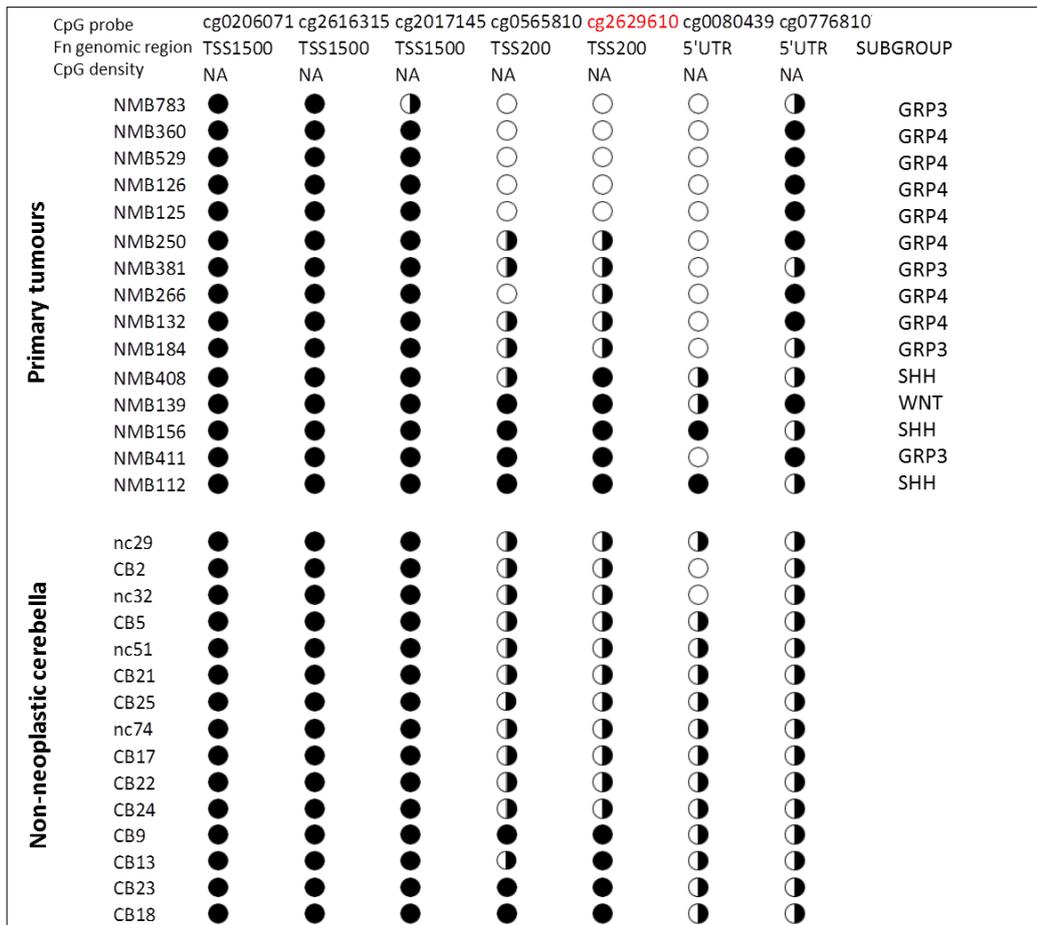
Figure 4.25. Methylation status of *RHOH* gene promoter in primary tumours and non-neoplastic cerebella. The methylation status of the 7 promoter-associated CpG sites represented on the 450K array are shown in a representative sample of 15 primary tumours and 15 normal cerebellar samples.. The critical CpG residue identified for *RHOH* is annotated in red. Functional (F^n) genomic region and CpG density were defined by Illumina: TSS1500: 1500 base pairs upstream of transcription start site; TSS200: 200base pairs upstream of transcription start site; 5' UTR: 5' untranslated region ; NA: non-CpG island site Filled circles: $\beta \geq 0.7$; half-filled circles: $0.3 < \beta < 0.7$; empty circles: $\beta \leq 0.3$. Consistent aberrant hypomethylation was observed at the CpG site immediately upstream (cg0565810) of the critical site on the array. Aberrant hypomethylation was observed predominantly in Group 4 tumours. The site immediately downstream (cg0080439) of the critical site was aberrantly hypomethylated in 100% of Group 4 tumours, representing 56% of all tumours that were hypomethylated at this site.

4.6 Further work

Investigations reported in this chapter have identified 5 candidate genes in medulloblastoma (*ACTC1*, *ANXA2*, *FAM46A*, *PRPH* and *S100A4*) which show strong evidence of DNA methylation-dependent gene regulation and which may play important roles in the development of distinct subgroups of the disease. The strongest evidence of potential epigenetic regulation was found for *FAM46A* and *ACTC1*, with aberrant methylation patterns extending across multiple adjacent CpG residues in the gene body and in the gene promoter regions, respectively (Table 4.7). Genes were identified using high-throughput genome-wide expression and methylation microarrays and it will now be essential to validate findings using more sensitive methods.

RNA-Seq data for the 5 genes provided a useful validation of methylation-dependent expression alterations observed by microarray analysis in cell lines following 5-azaCdR treatment (Figure 4.23). The methylation status in cell lines before and after treatment with 5-azaCdR will need to be validated by bisulfite sequencing and qRT-PCR will need to be performed to confirm evidence of a direct association between DNA methylation and gene expression. It is known that immortalised cell lines frequently contain artifactual gains of methylation at many loci (Bender *et al.*, 1998; Smiraglia *et al.*, 2001) and it will be necessary to validate these *in vitro* methylation findings in the 4 subgroups of primary medulloblastomas and in the normal cerebellum by bisulfite sequencing and to validate the associated gene expression changes in primary tumours by qRT-PCR. Expression will also need to be assessed in the normal cerebellum to confirm the potential tumour suppressor or oncogenic role of genes identified.

The high-resolution 450K methylation array permitted an assessment of gene-wide DNA methylation. Based on the array results for the candidate genes identified, it will be important to assess hypermethylation of the CpG island-associated gene body of *FAM46A* in non-SHH tumours (Figure 4.16) and the dense methylation of the non-CpG island-associated promoter of *ACTC1* in Group 3 and Group 4 tumours (Figure 4.21) by bisulfite sequencing. *S100A4* has previously been reported to be epigenetically regulated in medulloblastoma by methylation of critical intragenic sites (Rosty *et al.*, 2002; Lindsey *et al.*, 2007). Investigations in this chapter have identified a site in the non-CpG island-associated promoter of *S100A4* (Figure 4.20) that may also play a critical role in gene regulation and methylation of the wider promoter region will need

to be investigated by bisulfite sequencing, as well as any associated expression changes by qRT-PCR in primary tumours. For *ANXA2* validation studies will need to focus on the CpG island and island shore-associated promoter to determine if methylation at a single CpG site is sufficient to affect transcription (Figure 4.18). For *PRPH,* both promoter and gene body methylation will need to be investigated to determine whether hypermethylation of the CpG island-associated gene body is required alongside hypomethylation of the gene promoter to cause upregulated expression (Figure 4.19).

In the case of aberrant gene body methylation that was identified in *FAM46A* and *PRPH*, it would be interesting to investigate potential alternate promoter usage or alternate splicing events as intragenic DNA methylation has been reported in recent studies to play a role in these processes (Kulis *et al.*, 2013). PCR primers could be designed against alternate transcript variants to assess any differential expression and ChIP-seq would allow characterisation of transcription factor binding sites, nucleosome positioning and histone modifications that are indicative of active promoters such as H3K4me3.

Following validation of the methylation status and associated gene expression changes in primary tumours and the normal cerebella, it will be necessary to determine whether the differences in mRNA levels observed are reflected at the protein level by western blot analysis or immunohistochemistry. To determine any functional role for the epigenetic dysregulation of these genes in medulloblastoma, functional studies including proliferation, apoptosis and migration assays will need to be conducted *in vitro* and *in vivo* following appropriate modulation of gene expression. There are currently mouse models for SHH (Gilbertson and Ellison, 2008; Hatten and Roussel, 2011), WNT (Gibson *et al.*, 2010) and Group 3 (Kawauchi *et al.*, 2012; Pei *et al.*, 2012) medulloblastomas. Results from functional assays will reveal any involvement in the initiation and progression of the distinct subgroup(s) of medulloblastoma, in which epigenetic dysregulation was observed, and identify their potential as therapeutic drug targets. High levels of *S100A4* expression have previously been reported to be associated with metastatic progression of medulloblastoma (Hernan *et al.*, 2003). *S100A4* was identified to be potentially dysregulated by DNA hypomethylation in the discrete subgroup of WNT tumours and, therefore, a more focused study into any functional role in the context of WNT tumourigensis should be carried out. Elevated

expression of a second WNT-specific gene identified in this study (*ANXA2*) has also been found to correlate with metastasis in cancers (Wu *et al.*, 2012). It is interesting that elevated expression of both these metastasis genes has been found in the WNT subgroup of tumours which confer the best prognosis of all tumour subgroups (Ellison *et al.*, 2005; Clifford *et al.*, 2006; Kool *et al.*, 2012), and any potential role in WNT tumourigenesis will need to be investigated further.

Potential epigenetic dysregulation for 3 of the candidate genes (*ANXA2*, *PRPH* and *S100A4*) was specific to the WNT subgroup of tumours. WNT tumour status confers a very good long-term prognosis in comparison to the other tumour subgroups (Ellison *et al.*, 2005; Clifford *et al.*, 2006; Kool *et al.*, 2012), and long-term survival rates likely exceed 90% (Taylor *et al.*, 2012). As a result, the role of WNT-specific methylation biomarkers in disease prognostication is limited. For the non-WNT tumours, identification of prognostic disease biomarkers which may also have functional relevance offers the potential for the development of a more accurate and refined risk stratification model.

FAM46A was hypermethylated in Group 3, Group 4 and WNT tumours. At 1 critical CpG site (cg11075693) complete methylation ($\beta \geq 0.7$) was observed in 40% of WNT tumours but only in 3% and 9% of Group 3 and Group 4 tumours, respectively (Figure 4.10). It needs to be investigated whether these methylated Group 3 and Group 4 tumours share a similar good prognosis as WNT tumours and whether this methylation marker has utility as an independent prognostic marker in Group 3 and Group 4 tumours to identify a subset of good prognosis tumours which may benefit from less aggressive therapy. The *ACTC1* critical CpG site (cg08190298) was aberrantly methylated in 40% of Group (3+4) tumours compared with only 5% of (SHH+WNT) tumours (Figure 4.14). It will be important to assess the independent prognostic relevance of this methylation marker within Group 3 and Group 4 tumours to determine whether it identifies a subset of tumours within these groups that behave differently. It will also be important to assess the ability of this marker to improve the accuracy of survival prediction alongside established prognostic features. *MYC* amplification, which is enriched in Group 3 tumours, and LCA histology and metastases which are frequently present in Group 3 and are also found in Group 4 (Kool *et al.*, 2012) are established prognostic features in medulloblastoma (Ellison *et al.*, 2011b). The identification of

prognostic disease biomarkers, leading to the development of a more accurate and refined risk stratification model, is essential to improve survival and to minimise late adverse treatment effects through the delivery of more personalised therapies (Pizer and Clifford, 2009).

4.7 Summary

This chapter reports a robust and effective genome-wide screen to identify candidate epigenetically regulated genes in distinct subgroups of medulloblastoma. Five candidate genes were identified (*ACTC1*, *ANXA2*, *FAM46A*, *PRPH*, *S100A4*), representing both potential candidate tumour suppressor genes and oncogenes, and for four of them they represent novel medulloblastoma genes. Aberrant methylation states, which may affect gene transcription and play a role in medulloblastoma tumourigenesis, were observed at multiple adjacent CpG residues in the *FAM46A* and *ACTC1* genes and provided the strongest evidence of epigenetic regulation for these genes. *S100A4* has previously been reported to be hypomethylated in medulloblastoma and to have a potential oncogenic role (Lindsey *et al.*, 2007). Investigations reported here have identified *S100A4* hypomethylation to be specific to the discrete WNT subgroup of tumours and, therefore, future studies on this gene should focus on this tumour subgroup. It will now be necessary to validate the methylation-expression relationships observed for these genes in the distinct tumour subgroups, and to determine any functional role for their epigenetic dysregulation in driving tumourigenesis and progression within medulloblastoma subgroups.

The relatively small number of candidate genes that have been identified from this genome-wide screen suggests that any direct association between DNA methylation and gene expression is highly gene-specific, and while it does exist for select genes, it does not hold true for the broader differential patterns of methylation and gene expression observed in the subgroups of medulloblastoma. In the investigations reported here, only 6% of methylation events that were associated with gene expression in medulloblastoma cell lines showed a similar association reflected in their differential subgroup-specific patterns of methylation and expression in primary tumours. A technical limitation of this study is recognised with the lack of matched-tumour methylation and expression data available, which would allow for a direct correlation analysis between methylation and expression rather than a comparison of means. Using the methods described, the

results reported in this Chapter suggest that any functional association between patterns of methylation and expression in medulloblastoma subgroups is weak. This is consistent with the findings in Chapter 3 which focused on Group 3 and Group 4 tumours, and also with the findings of other groups investigating the role of DNA methylation in different cancers (Sproul *et al.*, 2011; Jung *et al.*, 2012; Kulis *et al.*, 2012; Sproul *et al.*, 2012).

As an increasing number of cancer methylomes are being characterised it is becoming increasingly recognised that methylation patterns may, to some extent, reflect patterns in the distinct cell of origin from which the tumour subgroup arises (Kulis *et al.*, 2012; Batora *et al.*, 2013). Additionally, aberrant patterns of methylation have been found to represent markers of cell lineage rather than tumour progression, reflecting the developmental history and transcriptional state of the different cells of origin (Sproul *et al.*, 2011; Sproul *et al.*, 2012).

Epigenetic gene regulation is complex involving multiple and inter-related chromatin components, and it is likely that the roles of DNA methylation within this are complex too. It has long been acknowledged that DNA methylation may not always directly cause the silencing with which it is associated, but rather may target genes that are already silenced by another mechanism, and effectively "lock" or stabilise the transcriptionally silent state (Bird, 2002). Epigenetic dysregulation in cancers involves all components of chromatin (Jones and Baylin, 2007). Recently, whole exome and whole genome sequencing have identified recurrent mutations affecting several chromatin-modifying genes and implicating deregulation of the epigenome as an important step in medulloblastoma pathogenesis (Jones *et al.*, 2013). It is important to consider that the interaction between different epigenetic mechanisms may play a major role in transcription regulation in medulloblastoma, rather than DNA methylation alone. In future studies it will be important to study aberrant DNA methylation alongside histone modifications and nucleosome structure if we are to gain a fuller understanding of its role in gene regulation in medulloblastoma.

Advances in technologies, which have led to the high resolution 450K methylation array and now whole genome shotgun bisulfite sequencing (WGSBS) (Lister *et al.*, 2009), means that DNA methylation can be studied across the complete gene, to include promoter and intragenic sites. Studies to date suggest that a comprehensive assessment of gene body methylation has the potential to reveal further candidate epigenetically

regulated genes. It has recently been observed in multiple genome-wide epigenomic studies that intragenic DNA methylation can be positively correlated with gene expression both in normal development and in cancer cells (Kulis *et al.*, 2013). This relationship has also been reported in a context-dependent manner relating to whether sites were located inside or outside of CpG islands in the gene body (Varley *et al.*, 2013). These findings will have major implications for future studies into the genome-wide role of DNA methylation in gene regulation, whereby both positive and negative correlations should be investigated in the gene body. Further epigenetically regulated genes may be identified if studies are focused on gene regulatory elements, such as enhancer sites and transcription factor binding sites which are located outside of typical gene promoters and within the gene body. An increasing number of epigenome-wide studies are identifying significant correlative relationships between DNA methylation and gene expression for these important regulatory regions, revealing potential new roles for DNA methylation in epigenetic gene regulation (Kulis *et al.*, 2013).

Chapter 5   Discussion

5.1 Background and project summary

Medulloblastoma, the most common malignant brain tumour of childhood, is an invasive embryonal tumour of the cerebellum and accounts for approximately 10% of all childhood cancer deaths (Pizer and Clifford, 2009). 5-year survival rates of up to80% are currently being achieved using risk-adapted combination therapies (Northcott *et al.*, 2012a). Despite improved survival, there remains a proportion of high-risk patients for whom prognosis is very poor, and survivors of medulloblastoma suffer considerable long-term neurocognitive and endocrinologic adverse effects associated with their treatment (see section 1.5.5.2). At present, risk stratification is based on clinical features alone (see section 1.5.4). In the last 10 years, significant advances have been made in understanding medulloblastoma genomics, and have led to the recognition of medulloblastoma as a disease that comprises four core transcriptomic subgroups (SHH, WNT, Group 3 and Group 4). These subgroups are associated with distinct clinical behaviours and genetic features and will likely require different therapeutic strategies in the clinic (Kool *et al.*, 2012; Taylor *et al.*, 2012).

Molecular subclassification is beginning to inform treatment of medulloblastoma. Refined risk stratification models have been developed that incorporate molecular disease biomarkers and that can identify patients with a favourable prognosis, for whom less aggressive treatments could potentially be administered (Pizer and Clifford, 2009; Ellison *et al.*, 2011b). Extensive research is underway to individualise treatments according to the biology of a patient's tumour. Despite these advances, more insights need to be gained from the underlying molecular pathology of the different disease subgroups, and in particular, the less well-defined Group 3 and Group 4 tumours, to identify critical driver events in their pathogenesis that could potentially be targeted therapeutically. Group 3 medulloblastomas carry the worst prognosis of all the subgroups, while the largest subgroup, Group 4, comprises subsets of tumours with poor clinical outcome (see section 1.5.8.2). There is an urgent need to understand the molecular defects that are driving these particularly aggressive phenotypes and to develop better therapies to improve survival.

The role of the epigenome in medulloblastoma development is increasingly being recognised. DNA methylation plays a key role in epigenetic transcriptional regulation (see section 1.6.2), and the hypermethylation of promoter-associated GpG islands

leading to epigenetic gene silencing is a common feature of medulloblastoma. There are approximately 30 genes known to be aberrantly silenced by promoter hypermethylation in medulloblastoma (Dubuc *et al.*, 2012). Many of these, including *RASSF1A*, *HIC1* and *CASP8*, which are the most frequent epigenetically inactivated genes in medulloblastoma, have been identified using classical candidate gene approaches (Lindsey *et al.*, 2005). Furthermore, many of these previous studies have focused on hypermethylation of CpG island promoters and did not recognise the distinct molecular subgroups of medulloblastoma.

It is now possible to study DNA methylation on a genome-wide scale and at base-pair resolution (see section 1.6.3.3). Profiling primary medulloblastomas on DNA methylation microarrays has revealed the existence of four methylomic subgroups of the disease that are strongly associated with their transcriptomic counterparts (SHH, WNT, Group 3 and Group 4) (Schwalbe *et al.*, 2013). The methylomic subgroups can be used to assess the relationship between DNA methylation and gene expression patterns in medulloblastoma, and can facilitate a more refined discovery of DNA methylation events that are associated with gene expression, and that may play a critical role in the development of distinct disease subgroups. Epigenome-wide studies of DNA methylation in cancer are also starting to reveal roles for DNA methylation outside of the gene promoter and outside of CpG islands (Irizarry *et al.*, 2009; Kulis *et al.*, 2013). Using genome-wide approaches, an unbiased assessment of gene-wide DNA methylation in distinct medulloblastoma subgroups can now be undertaken to gain greater insights into the role of DNA methylation in medulloblastoma development.

The genome-wide role of DNA methylation in the distinct molecular subgroups of medulloblastoma has not been widely investigated. This project aimed to identify additional epigenetically regulated genes in medulloblastoma with a focus on subgroup-specificity, using two high-throughput screening approaches: i). The GoldenGate Cancer Panel I methylation microarray (see section 2.7.2) was used to identify Group 3- and Group 4-specific promoter-associated DNA methylation events and, combined with assessment of *in silico* transcriptomic datasets, their association with gene expression was evaluated for evidence of potential methylation-dependent gene regulation in Group 3 and Group 4 tumours. Candidate epigenetically regulated genes identified in primary tumours were then taken forward for further assessment of a direct relationship between

DNA methylation and gene expression in medulloblastoma cell lines treated with the demethylating agent, 5'-aza-2'-deoxycytidine (5-azaCdR) (Chapter 3); ii). The Illumina HT12v4.0 expression microarray (see section 2.8.1) was used to carry out a genome-wide screen to identify genes transcriptionally upregulated by demethylation (5-azaCdR) treatment of medulloblastoma cell lines and, combined with a genome-wide analysis of DNA methylation using the Illumina Infinium 450K methylation array (see section 2.7.3), methylation events that were associated with 5-azaCdR-induced expression alterations in cell lines were identified. Candidate epigenetically regulated genes identified in cell lines were then taken forward for further assessment of potential methylation-dependent gene regulation in primary medulloblastoma subgroups (Chapter 4).

Approach (i) (Chapter 3) identified 73 promoter-associated CpG residues (encompassing 63 genes) that significantly distinguished Group 3 and/or Group 4 medulloblastomas from all other subgroups (Figure 3.5). Of the 63 genes, four (*DDR2*, *HDAC1*, *MET* and *RHOH*) showed strong evidence of an inverse methylation/expression relationship in primary tumours and were selected as subgroup-specific candidate epigenetically regulated genes (Table 5.1). They were investigated further for evidence of a direct relationship between methylation and gene expression in cell lines treated with 5-azaCdR by qRT-PCR. For two of the candidate genes (*MET* and *RHOH*), their subgroup-specific methylation event was also tumour-specific (Table 5.1).

qRT-PCR assessment of expression changes in cell lines (*n*=6) following 5-azaCdR treatment identified a direct relationship between methylation and gene expression for two of the four candidate genes tested (*MET* and *RHOH*) (Table 5.1). In primary tumours, the inverse methylation/expression relationship observed for *MET* was seen in SHH, Group 3 and Group 4 tumours but did not extend to the WNT subgroup (Figure 3.7C). As a result, unlike its methylation pattern, the pattern of *MET* expression was not specific to Group (3+4), and the methylation event represented a potential candidate mechanism of activation of the known *MET* oncogene in SHH tumours.

In primary tumours, the promoter-associated *RHOH* residue (RHOH_P121_F) was hypomethylated in Group 4 tumours and this hypomethylated state was associated with significantly higher expression in Group 4 compared with the other tumour subgroups

(Figure 5.1). Subsequent assessment of *RHOH* promoter methylation using the genome-wide Infinium 450K methylation array suggests that multiple promoter-associated CpG residues may play a role in regulating gene transcription (Figure 4.25). Using approach (i), a potential novel oncogenic role for *RHOH* in Group 4 medulloblastoma pathogenesis has been identified.

Gene	Subgroup-specific (Group 3 and/or Group 4) methylation pattern	Subgroup-specific inverse methylation-expression relationship	Tumour-specific methylation event	Aberrant methylation state	Methylation-dependent expression in ≥ 4/6 cell lines (qRT-PCR; -/+ 5azaCdR)
DDR2	Group (3+4)	Y	N	Hypomethylated in SHH + WNT	N
HDAC1	Group 4	Y	N	Hypomethylated in SHH, WNT and Group 3	N
MET	Group (3+4)	In SHH, Group 3 and Group 4 but not WNT	Y	Hypermethylated in Group (3+4)	Y
RHOH	Group 4	Y	Y	Hypomethylated in Group 4	Y

Table 5.1. Summary table of Group 3 and/or Group 4–specific candidate epigenetically regulated genes identified in Chapter 3. Candidate genes showed evidence of an inverse methylation/expression relationship in Group 3 and/or Group 4 tumours. For two candidate genes, *MET* and *RHOH*, their subgroup-specific methylation was also tumour-specific, indicating an aberrant event in Group 3and/or Group 4. Candidate genes were investigated further for evidence of methylation-dependent gene expression in medulloblastoma cell lines using qRT-PCR. Expression was assessed in cell lines before and after treatment with the demethylating agent 5-azaCdR.
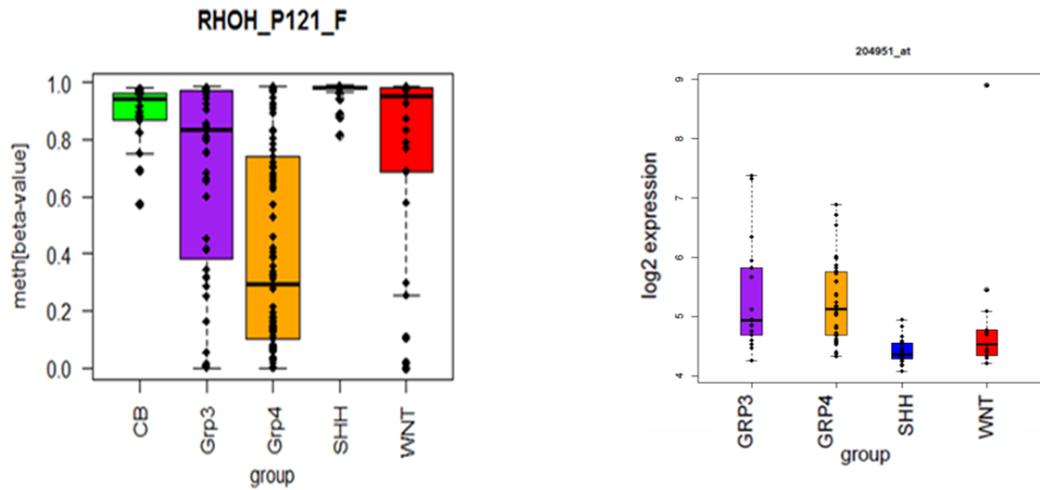
Figure 5.1. Patterns of CpG methylation and gene expression in primary medulloblastomas for RHOH. Combined boxplots and strip-plots illustrate the differential patterns of CpG methylation and gene expression for the RHOH promoter-associated CpG residue RHOH_P121_F. Methylation of RHOH_P121_F in control cerebellar samples (CB) is also show. The CpG residue is hypomethylated in Group 4 medulloblastomas and hypomethylation is associated with significantly higher expression in Group 4 compared with the other tumour subgroups.

The second approach taken by this project was a genome-wide screen that identified 283 genes that were upregulated in 2 or more medulloblastoma cell lines (*n*=10) following 5-azaCdR treatment (Chapter 4). Using the Illumina 450K methylation array, an unbiased assessment of DNA methylation identified 160 CpG residues (encompassing 21 of the 283 genes) whose methylation was consistent with gene expression alterations observed after 5-azaCdR treatment. These 21 genes were selected as candidate epigenetically regulated genes in cell lines and were taken forward for further investigation of epigenetic gene regulation in primary medulloblastoma subgroups.

Subgroup-specific differential methylation and gene expression analyses identified 5 of the 21 genes that showed evidence of methylation-dependent gene expression in primary medulloblastoma subgroups and that may play a role in medulloblastoma development (Table 5.2 and Figure 5.2). The genes identified varied with regard to their subgroup-specificity, location and number of CpG residues potentially involved, and their potential role in medulloblastoma pathogenesis. One gene identified, *S100A4*, has previously been reported to be epigenetically upregulated in medulloblastomas and this study refined its epigenetic upregulation to the WNT subgroup of tumours (Figure 5.2). The remaining four genes represent novel medulloblastoma genes.

Gene	Methylation-dependent transcriptional silencing and upregulation by 5azaCdR treatment in cell lines	Subgroup-specific methylation pattern of critical CpG sites associated with expression alterations in cell lines	Subgroup-specific inverse methylation-expression relationship	Tumour-specific methylation event(s)	Aberrant methylation state	Location of subgroup-specific critical CpG site(s)	Aberrant methylation patterns observed across multiple adjacent CpG sites
FAM46A	Y	SHH	Y	N	Hypermethylated in WNT, Group 3 and Group 4	Gene Body	Y
ACTC1	Y	Group (3+4)	Y	Y	Hypermethylated in Goup (3+4)	Gene promoter	Y
PRPH	Y	WNT	Y	Y	Hypomethylated in WNT subgroup	Gene promoter	N
ANXA2	Y	WNT	Y	Y	Hypomethylated in WNT subgroup	Gene promoter	N
S100A4	Y	WNT	Y	Y	Hypomethylated in WNT subgroup	Gene promoter and Gene body	N

Table 5.2. Summary table of subgroup-specific candidate epigenetically regulated genes identified in Chapter 4. Candidate genes showed evidence of methylation-dependent gene regulation in cell lines, alongside differential CpG methylation inversely correlated with gene expression in distinct subgroups of primary tumours. Aberrant methylation patterns extended across multiple adjacent CpG residues in FAM46A and ACTC1, making them the strongest candidates identified in Chapter 4. CpG residues adjacent to the critical site(s) identified in *PRPH*, *ANXA2* and *S100A4* did not show similar aberrant methylation patterns.
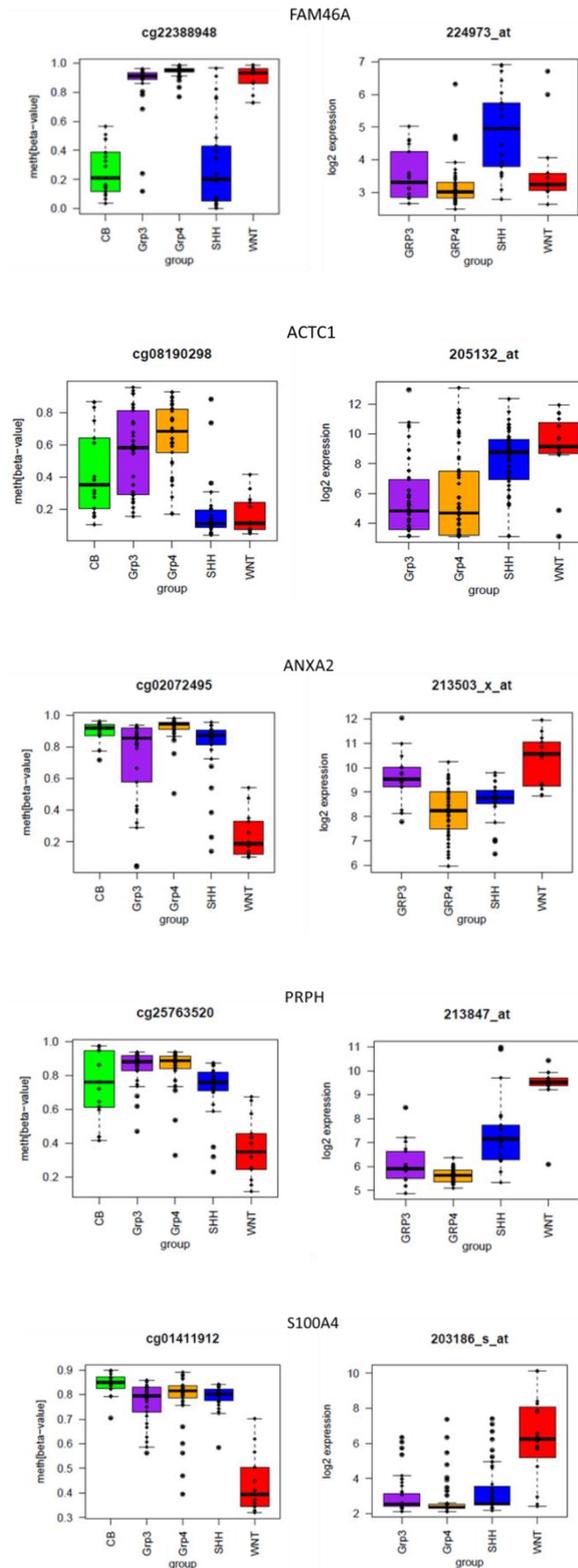
Figure 5.2. Patterns of subgroup-specific CpG methylation and gene expression for the 5 candidate epigenetically regulated genes identified in Chapter 4. Boxplots depicting methylation patterns are on the left and boxplots depicting gene expression patterns are on the right. Patterns of CpG methylation in control cerebellar samples (CB) are also shown for each gene.

5.2 High-throughput screening approaches to the identification of epigenetically regulated genes in medulloblastoma subgroups

5.2.1 *DNA methylation microarrays and the identification of subgroup-specific methylation markers*

DNA methylation microarrays provide the opportunity to study DNA methylation at base-pair resolution and on a genome-wide scale. Tumour profiling on methylation microarrays has recognised DNA methylation patterns as cancer-specific markers with a high potential for clinical applications (Baylin and Jones, 2011). The four principal subgroups of medulloblastoma are associated with distinct methylomic profiles (Schwalbe *et al.*, 2013). Investigations in Chapter 3 were focused on the determination of DNA methylation events that significantly discriminated the Group 3 and/or Group 4 medulloblastomas and that were associated with gene expression in these subgroups. Methylation profiles generated on the Illumina GoldenGate Cancer Panel I methylation microarray were used to identify methylation markers of Group 3 and Group 4 medulloblastomas.

The intuitive β-value measure of methylation reported by the GoldenGate array follows a non-parametric, bi-modal distribution and, therefore, statistical models that have been developed for gene expression microarray analysis, and that largely follow the assumptions of normality cannot be applied (Du *et al.*, 2010). By applying non-parametric Mann-Whitney U tests, novel methylation markers were identified that clearly distinguished the Group 3 and/or Group 4 tumours (Figure 3.5). While Group 3 and Group 4 tumours are considered to be more similar to each other than to SHH or WNT tumours (Taylor *et al.*, 2012), the high proportion of Group 4-specific events (84%) that also discriminated Group (3+4) tumours suggests that the experimental design (combining distinct subgroups together into one test group) and statistical test applied may have been limited in accommodating the differences in size and variance that exist between the tumour subgroups. This was also apparent in the differential expression analysis where the *t*-test was applied using a similar experimental design; as an example, *MET* was significantly differentially expressed when (SHH+WNT) were tested against Group (3+4) but expression profiles revealed similar patterns of expression in Group 3, Group 4 and in the smaller WNT subgroup compared with the SHH subgroup which had markedly different expression (Figure 3.7C). It is important,

therefore, to exercise caution when interpreting significance test results if distinct disease subgroups of unequal size and variance are combined together into one test group for the purpose of a two group comparison. While the Mann-Whitney U test is assumed to be relatively insensitive to unequal sample size, a Kruskal-Wallis test accompanied by *post hoc* Mann-Whitney U tests may have improved test sensitivity to detect Group 4 events that do not also discriminate Group (3+4), although these methods would be labour-intensive in high-throughput screening methods.

In Chapter 4, investigations were not initially directed by subgroup-specific methylation patterns, but rather, subgroup-specificity was later determined for a subset of CpG sites whose methylation was consistent with epigenetic gene regulation in cell lines. The limitations associated with subgroup-specific differential methylation analysis in Chapter 3 (see above) were not encountered in Chapter 4, because the M-value method of measuring methylation could be used with the Infinium 450K methylation array that was used for the investigations reported in Chapter 4. M-value measures of methylation follow a more normal distribution (Du *et al.*, 2010) and enabled the use of the statistical model *limma* (Smyth, 2004), which is commonly employed for differential expression analysis in microarray experiments. Using *limma*, contrast matrices could be designed that tested all possible pair-wise comparisons simultaneously in an automated way, and events that were specific to one subgroup but not to any other group could easily be determined from the *limma* output. As well as facilitating a more direct interpretation of subgroup-specificity, *limma* also facilitated a more direct comparison of differential methylation and expression results which is discussed further in section 5.2.2.

5.2.2 Association between differential patterns of DNA methylation and gene expression in medulloblastoma subgroups

5.2.2.1 Identification of subgroup-specific inverse methylation-expression relationships

The lack of primary tumour samples for which matched methylation and gene expression data were available has been one of the main limitations of this project. Associations between subgroup-specific patterns of differential methylation and gene expression have had to be determined, in large part, by a comparison of means based on the results of the differential methylation and expression analyses of microarray data.

For investigations reported in Chapter 3, genes that were both significantly differentially methylated and significantly differentially expressed in Group 3 and/or Group 4 tumours were subject to visual assessment of boxplots that illustrated the differential patterns and subgroup-specific inverse methylation/expression relationships were evaluated. This presents an obvious weakness in that the associations described are broad subgroup associations that don't take into account the expression/methylation relationship in individual tumour samples.

Matched methylation and expression data for the primary tumour cohort would permit statistical correlation on a tumour by tumour basis and provide a statistical measure of the strength of the inverse methylation/expression relationship observed, both between subgroups and within subgroups where methylation patterns were variable. In the primary tumour cohort that was used for investigations reported in Chapter 3 (Table 2.1), there were 9 primary tumour samples for which gene expression data was also available. While this permitted a sample by sample statistical correlation analysis, it was limited by the small number of tumour samples which did not represent all tumour subgroups. Statistical correlation based on the 9 matched samples was calculated alongside an *in silico* analysis of methylation-dependent gene expression changes in cell lines (n=3) following 5-azaCdR treatment. Results from these were used to support the broad methylation-expression associations observed from inspection of the boxplots and the selection of candidate epigenetically regulated genes that would be investigated further.

In the primary tumour cohort used for investigations reported in Chapter 4 (Table 2.2), there were no primary tumours for which gene expression data was also available. No statistical correlation on a tumour by tumour basis could be performed and inverse subgroup-specific associations were inferred from a comparison of means that were tested using *limma*. Using *limma* to conduct both differential methylation and expression analyses facilitated a more direct and automated comparison of significant subgroup-specific differential methylation and gene expression; inverse subgroup-specific relationships could be determined directly from the *limma* output which described whether the contrast was significantly positive or significantly negative using the nominal values 1 and -1, respectively (Table 4.6). Overall the application of a bioinformatics statistical model to differential methylation and expression analysis

provided a more robust experimental design where multiple groups were being considered together, and had considerable practical advantages for high-throughput analysis in terms of speed and ease of interpretation of results.

5.2.2.2 *Patterns of DNA methylation do not show clear associations with gene expression*

Investigations in Chapter 3 identified 73 promoter-associated CpG sites that were significantly differentially methylated in Group 3 and/or Group 4 tumours. These CpG sites encompassed 63 genes. Only 15 (24%) of these 63 genes were significantly differentially expressed in an equivalent subgroup-specific manner and only 4 showed strong inverse associations between the subgroup-specific patterns of methylation and gene expression observed. Following further analyses, one gene (*RHOH*) was identified that showed the strongest evidence of methylation-dependent gene expression in the subgroups that were targeted in this study; *RHOH* was hypomethylated and upregulated in Group 4 medulloblastomas (Figure 5.1). Furthermore, there were no Group 3-specific methylation events that showed any evidence of an inverse methylation-expression relationship for genes that were also differentially expressed in Group 3. For investigations reported in Chapter 4, 160 CpG sites, that represented the 21 genes which were upregulated in cell lines following 5-azaCdR treatment consistent with a methylated status, were examined. Of the 160 CpG sites tested, 9 (6%) showed evidence of subgroup-specific differential methylation concordant with differential gene expression, and potential epigenetic gene regulation in medulloblastoma subgroups. These 9 sites encompassed 5 candidate epigenetically regulated genes (Figure 5.2).

Overall this project finds that the patterns of methylation observed in Group 3 and Group 4 medulloblastomas, and the patterns observed for the 21 candidate genes tested in primary tumours in Chapter 4 do not show strong associations with gene expression in medulloblastoma subgroups. While a direct relationship between DNA methylation and gene expression clearly exists for some genes (see section 1.6.3) and 6 candidate genes have been identified in this study that show strong evidence of methylation-dependent gene expression, this effect would appear to be gene-specific. A general weak association between DNA methylation patterns and gene expression in primary tumours has been reported by several studies (Etcheverry *et al.*, 2010; Sproul *et al.*, 2011; Jung *et al.*, 2012; Kulis *et al.*, 2012; Sproul *et al.*, 2012). In their study, Jung et al.

used both a linear and non-linear approach to examine methylation-expression correlations in multiple myeloma. Using both approaches they identified between 2.1% and 25.3% of promoter-associated CpG probes on the GoldenGate methylation array that were correlated, either positively or negatively, with gene expression, suggestive of a weak association between DNA methylation and gene expression in multiple myeloma for the ~800 genes assayed on the GoldenGate array (Jung *et al.*, 2012).

DNA methylation interacts with other epigenetic events, including histone modifications and nucleosome positioning, to establish chromatin architecture and regulate gene expression (see section 1.6.2). The removal of DNA methylation marks is required to establish a permissive state for subsequent gene expression, but the exact role of DNA methylation in transcriptional regulation has long been a source of debate, and it is not always clear whether gene transcription serves to block DNA methylation or whether DNA methylation acts to stabilise genes that are already transcriptionally silent (Jones, 2012). Emerging studies suggest that DNA methylation may not be the predominant mechanism for epigenetic gene silencing (Jones, 2012). Furthermore, recent studies have shown that the promoters of active genes are located in nucleosome-depleted regions of the genome and that DNA methylation occurs secondary to the DNA becoming inaccessible by nucleosome occupancy (Han *et al.*, 2011; Jones, 2012; Kelly *et al.*, 2012). These studies support the long-held notion that promoter DNA methylation acts secondary to other inactivating mechanisms to stabilise transcriptionally silent states and strongly suggests that methylation requires the nucleosomes to be present. It is possible that the general lack of correlation between DNA methylation and gene expression observed may be because, for the majority of genes, methylation is necessary but not sufficient for regulating gene expression and that other mechanisms, such as histone modifications and nucleosome occupancy are more dominant.

While it has been widely established that gene promoter methylation is associated with transcriptional repression, genome-wide studies are beginning to yield interesting insights regarding the potential roles of DNA methylation outside of the gene promoter. In one study in chronic lymphocytic leukaemia, the authors reported that while DNA methylation and gene expression were poorly correlated, the majority of significantly correlative relationships were found for CpG sites located in the gene body, and both

negative and positive correlations were identified (Kulis *et al.*, 2012). Multiple studies have identified a positive correlation between intragenic DNA methylation and gene expression, and it is thought that it may impact alternative splicing mechanisms (Ball *et al.*, 2009; Irizarry *et al.*, 2009; Shukla *et al.*, 2011; Brown *et al.*, 2012; Jones, 2012; Kulis *et al.*, 2012). In their study, Kulis et al. identified 13 genes that had cancer-specific methylation changes close to alternative splice sites. They found a relationship between DNA methylation and alternative splicing for exon 5 of *PTPRC* (encodes the cell surface marker CD45) in CLL patients; DNA methylation was correlated with exon skipping, while lack of methylation was correlated with exon inclusion. They also identified extensive hypomethylation that targeted mainly enhancer regions within the gene body, where methylation was negatively correlated with gene expression and enhancer activity (Kulis *et al.*, 2012). In another study, the relationship between gene body methylation and gene expression has been reported in a context-dependent manner relating to whether CpG sites were located inside or outside of CpG islands in the gene body (Varley *et al.*, 2013).

It would appear from the studies described above, that the function of DNA methylation varies with context and that the relationship between DNA methylation and gene expression is more complex than what is currently understood. These findings will have major implications for future studies into the genome-wide role of DNA methylation both in development and in cancer, and it is clear that positive and negative methylation-expression correlations will need to be comprehensively studied if the role of DNA methylation in gene regulation is to be more fully understood and appreciated.

A further explanation for the lack of a direct association between DNA methylation patterns and gene expression may lie in the increasing recognition that methylation patterns may, to some extent, reflect patterns in the distinct cell of origin from which the tumour subgroup arises (Kulis *et al.*, 2012; Batora *et al.*, 2013). DNA methylation patterns in medulloblastoma could, therefore, hold considerable promise for the identification of the cells of origin of medulloblastoma subgroups (cell of origin of Group 4 tumours is currently unknown) as well as the potential subtypes that exist within the core subgroups (Batora *et al.*, 2013). Furthermore, aberrant patterns of DNA methylation have been identified in multiple cancer types that represent markers of cell lineage; cancer-specific hypermethylation patterns have been identified that don't play a

direct role in the transcriptional silencing with which they are associated, but rather reflect the developmental history and transcriptional state of the different cells of origin (Sproul *et al.*, 2011; Sproul *et al.*, 2012). These findings suggest that, for the large part, DNA methylation may not play a functional tumour-promoting role but may be a passenger event that holds considerable value for the identification of cellular origins in cancers.

5.2.3 *Future approaches for the identification of epigenetically regulated genes in medulloblastoma*

The development of DNA methylation microarrays was a major step forward in the era of epigenomics allowing whole cancer methylomes to be characterised at increasingly high resolution. Advances in the technology have seen the coverage offered by Illumina microarray platforms increase substantially from 1,505 promoter-associated CpG sites in ~800 genes (Bibikova *et al.*, 2006) to more than 450,000 CpG sites spanning complete gene regions of more than 20,000 genes (Bibikova *et al.*, 2011). The development of the 450K methylation array permits the genome-wide study of DNA methylation in an unbiased way; CpG sites outside of the gene promoter and outside of CpG islands can be assessed for functional roles in normal development and in cancer (Bibikova *et al.*, 2011). Furthermore the high density resolution of the 450K array permits the assessment of differentially methylated regions as well as the study of single CpG sites.

The investigations reported in Chapter 4, using the 450K methylation array, assessed methylation-expression relationships at the resolution of the single CpG site. There were 4,238 CpG sites that encompassed the 283 genes which were upregulated in cell lines following 5-azaCdR treatment, and that were assessed for evidence of methylation consistent with the upregulated expression observed. By considering the methylation status of each site individually, only 160 CpG sites (~4% and encompassing only 21 genes) were taken forward for further analysis of subgroup-specific methylation-expression relationships in primary tumours. This approach may have been limited and more sites, and genes, may have been identified initially in cell lines if regions of methylation were considered as an average measure, rather than single sites. Indeed for 2 of the candidate genes identified in Chapter 4 (*FAM46A* and *ACTC1*), subsequent assessment of their gene-wide DNA methylation patterns using probes on the 450K

array, identified methylation of multiple CpG residues in the gene body of *FAM46A* (see Figure 4.16) and multiple CpG residues in the non-CpG island promoter of *ACTC1* (see Figure 4.21) to play a potential role in their gene regulation.

In model systems, DNA methylation has been found to repress transcription in a manner that depends on the location and density of CpG sites relative to the gene promoter (Bird, 2002). While dense promoter hypermethylation is known to silence gene transcription (see section 1.6.2.1), genome-wide studies are identifying roles for DNA methylation in other functional genomic regions, such as 'CpG island shores' (Irizarry *et al.*, 2009), enhancer elements (Kulis *et al.*, 2012) and intragenic alternate promoters (Maunakea *et al.*, 2010). As a result, an increasing number of bioinformatics tools are being developed that facilitate the identification of functionally important genomic regions and that permit methylation to be assessed as an average measure over a defined number of CpG residues.

In a key study, Irizarry et al. used CHARM analysis (see section 1.6.3.3) to interrogate the colon cancer methylome alongside matched normal mucosa tissue (Irizarry *et al.*, 2009). They took the approach of averaging methylation M-values for each probe across the cancer samples and across the matched normal samples. Differential tumour *versus* normal methylation was quantified by the difference in averaged M-values and contiguous statistically significant values were grouped into regions. Using this method, they reported for the first time the presence of CpG island shores (see section 1.6.3.1). They found that most differentially methylated regions (DMRs) in colon cancer occurred in CpG island shores and that methylation of shores was strongly related to gene expression. Their study provides strong support for the study of region-wide DNA methylation in non-promoter and non-CpG island regions in medulloblastoma, where similar approaches could be taken to the identification of differentially methylated regions that characterise individual subgroups and their potential role in gene expression regulation.

Additional epigenetically regulated genes may also be identified by targeting specific functional genomic regions, such as enhancer elements, alternate intragenic promoters and common alternative splice sites. In their study of CLL, Kulis et al., showed that gene body hypomethylation was enriched for enhancer regions in B cells and for binding sites of transcription factors such as IRF and OCT1 (Kulis *et al.*, 2012). They

identified differentially methylated gene body CpG sites between subtypes of CLL that showed both positive and negative correlations with gene expression. A particularly interesting finding was that gene body CpG sites whose methylation status was negatively correlated with expression were related to enhancer elements, suggesting that the methylation status of distal regulatory regions may be more predictive for expression levels than promoter methylation (Kulis *et al.*, 2012). A further interesting finding from this study was that the *PTK2* gene (encodes cytoplasmic protein tyrosine kinase 2) exhibited simultaneous negative correlation between gene expression and promoter methylation and positive correlation in the gene body, highlighting again the importance of searching not only for negative methylation-expression correlations but also positive ones. These findings need to be further validated and the mechanisms by which DNA methylation at enhancers may regulate transcription are not known, however, they provide precedent for the investigation of DNA methylation in regulatory genomic regions in other cancers including medulloblastoma. The genome of a cell can be characterised into functional chromatin states by carrying out ChIP-Seq followed by bioinformatics modelling of a common set of chromatin states to characterise regulatory elements and their functional interaction which enable the linking of enhancers to putative target genes (Ernst *et al.*, 2011).

While DNA methylation microarrays have significantly advanced knowledge and understanding of cancer methylomes, bisulfite sequencing is considered to be the gold standard for validating DNA methylation patterns (see section 1.6.3.3). The availability of NextGen sequencing which is a high throughput sequencing technique, offers much higher coverage and a relatively lower cost for genome-wide sequencing than previous technologies (Zuo *et al.*, 2009). Whole genome shotgun bisulfite sequencing (WGSBS) is a NextGen sequencing approach that, unlike other methylation analysis methods, generates quantitative genome-wide methylation profiles at the ultimate comprehensive single-base pair resolution (Laird, 2010). WGSBS is more sensitive than microarray technologies, which are not well suited for the detection of low-frequency DNA methylation states in a sample. WGSBS can generate thousands of sequencing reads per locus which provides the potential for the detection of rare DNA methylation variants that may play a role in the development and progression of the disease phenotype (Korshunova *et al.*, 2008). It is likely that WGSBS will supersede all other techniques in the near future for the investigation of cancer methylomes.

RNA-Seq is another NextGen sequencing application that offers several advantages over microarray technologies for the detection of gene expression changes in cell lines following 5-azaCdR treatment. RNA-Seq is a powerful method for discovering, profiling and quantifying RNA transcripts across the entire transcriptome; it provides discreet, digital sequencing read counts that can be aligned to a particular sequence, supporting increased dynamic range and fold-change estimates for a more sensitive detection of genes, transcripts and differential expression compared with microarray methods. It also has the capacity to capture subtle gene expression changes that may be masked by noise levels often experienced with microarray studies (Ozsolak and Milos, 2011). The investigations in Chapter 4 used a microarray-based assessment of gene expression changes following 5-azaCdR treatment. Microarray-based approaches are sensitive to the type of platform used and discrepancies in the genes detected between different studies are observed (see section 4.5.5). The use of RNA-Seq, accompanied by a universal analytical pipeline, offers a more sensitive and reproducible method of detection of gene expression changes following demethylation treatment and would facilitate consistency between studies. In addition to gene expression analysis, RNA-Seq also enables the identification of alternative splicing events, allele-specific expression, the detection of rare and novel transcripts and non-coding RNAs (Ozsolak and Milos, 2011). The use of RNA-Seq to determine drug-induced expression changes coupled with DNA methylation analysis using WGSBS would provide the most comprehensive and refined assessment of the genome-wide role of DNA methylation in regulating the expression of genes and their transcript variants.

Gene transcription is a remarkably complex process and epigenetic regulation relies on the interplay between different mechanisms that control chromatin organisation, including DNA methylation, histone modifications and nucleosome remodelling (see section 1.6.2). As previously discussed in section 3.5.3 and section 4.6, it will be important to investigate aberrant DNA methylation alongside histone modifications and nucleosome structure if we are to gain a fuller understanding of its role in epigenetic gene regulation. Chromatin immunoprecipitation coupled with ultra-high-throughput parallel DNA sequencing (ChIP-Seq) allows for the investigation of genome-wide protein-DNA interactions (Ma and Wong, 2011). Using antibodies specific for various histone modifications, ChIP-Seq can be used to characterise nucleosomal changes that accompany DNA methylation changes in cancer cells, providing a greater insight into

their interactions in mediating transcriptional responses. A further important aspect of epigenetic gene regulation to consider is the role of micro-RNAs (miRNAs). Micro-RNAs are small non-coding RNAs that target mRNA transcripts for cleavage or transcriptional repression and can, therefore, influence the expression of many protein coding genes (Bartel, 2004). An important emerging new role for aberrant promoter DNA hypermethylation is the transcriptional repression of multiple microRNAs (Baylin and Jones, 2011). This repression has been found to be associated with pathway disruption and upregulation of oncogenic targets of the microRNAs (Saito *et al.*, 2006), as well as the constitutive activation of signalling that promotes metastasis (Lujambio *et al.*, 2008). Interestingly, downregulation of the miRNA-29 family has been found to be associated with overexpression of the DNA methyltransferases, which may then permit aberrant methylation in cancers (Fabbri *et al.*, 2007). A study of DNA methylation which extends to non-coding RNAs could reveal further epigenetic events as well as key pathways and genes that may play a critical role in medulloblastoma development.

5.3 Identification of novel putative oncogenes and tumour suppressor genes

Using the two high-throughput approaches described in section 5.1, this project has identified six genes that show strong evidence of methylation-dependent gene expression in distinct subgroups of medulloblastoma. For five of the genes identified (*RHOH*, *FAM46A*, *ANXA2*, *ACTC1*, *PRPH*), they represent novel medulloblastoma genes. *S100A4* has previously been shown to be epigenetically regulated in medulloblastoma (Lindsey *et al.*, 2007) and this project has identified its epigenetic regulation to be specific to the WNT subgroup of tumours (Figure 5.2). Interestingly, for 3 of the 6 genes identified (*S100A4*, *PRPH* and *ANXA2*), the methylation-expression relationships were specific to the WNT subgroup, where they all represented putative oncogenes in WNT medulloblastoma pathogenesis (Figure 5.2).

The determination of a putative oncogenic or tumour suppressor role for the candidate genes identified was based on their methylation profiles in normal cerebellar tissue. The lack of gene expression data for non-neoplastic samples presented a limitation to the complete evaluation of the effect of aberrant methylation on gene expression in primary tumour subgroups. Control cerebellar tissue contains a heterogeneous population of cells and the methylation signals generated will represent an average of these mixed cell

types and may mask the true signal from the actual cell of origin. Since the 4 molecular subgroups of medulloblastoma potentially arise from distinct stem and progenitor cell populations (see section 1.5.9), an accurate assessment of tumour *versus* normal tissue should account for these subgroup-specific differences and comparisons should be made to the correct cell type from which the tumour originates. Acquiring accurate cell type-matched control tissue for the distinct subgroups of medulloblastoma presents a significant challenge to researchers, and until the cell of origin can be identified for all subgroups, and in the absence of appropriate cell type-matched controls, we must consider with caution the utility of non-matched cerebellar samples as controls for the evaluation of subgroup-specific events.

For each of the candidate genes identified, DNA methylation of multiple CpG residues could be assessed using the genome-wide 450K methylation array, and the potential role of region-wide methylation on gene expression could be evaluated. Three genes (*FAM46A*, *ACTC1* and *RHOH*) demonstrated aberrant hypo- or hypermethylation of multiple adjacent CpG residues that may play a potential role in regulating gene expression (Figures 4.16, 4.21 and 4.25, respectively). These 3 genes were considered the strongest candidate epigenetically regulated genes. Nevertheless, methylation at a single CpG site has been shown to be sufficient to negatively affect binding of transcription factors to the DNA (Tate and Bird, 1993), and *ANXA2*, *PRPH* and *S100A4* still warrant further investigation.

The candidate genes identified in this project have been shown to be involved in key tumourigenic processes, including proliferation, migration and metastasis, in other cancer types (see section 3.5.6 and section 4.5.4). Their potential deregulation by aberrant DNA methylation may represent driver events and potential therapeutic targets in the distinct subgroups of medulloblastoma. Further *in vitro* and *in vivo* studies are required to validate the subgroup-specific methylation-expression relationships identified for each of the candidate genes, and any functional role they may play in driving medulloblastoma development (see section 5.3.1). Furthermore, the prognostic potential of the aberrant DNA methylation events must be investigated in well-defined clinical trial cohorts of medulloblastoma (see section 5.3.2).

5.3.1 Further work to investigate their functional significance in medulloblastoma

For the candidate genes identified in this project, bisulfite sequencing (see section 1.6.3.3) will need to be carried out in medulloblastoma cell lines before and after 5-azaCdR treatment, to validate their methylation status and the subsequent demethylation induced by 5-azaCdR. For the 3 strongest candidates, the gene region of focus will be the non-CpG island promoter regions of *RHOH* (Figure 4.25) and *ACTC1* (Figure 4.21) and the gene body region of *FAM46A* (Figure 4.19). For the remaining genes, sequencing approaches will need to focus on the critical CpG site identified in cell lines and in primary tumours and any immediately adjacent CpG residues in the surrounding region.

To accompany bisulfite sequencing, microarray gene expression changes detected in cell lines following 5-azaCdR will need to be validated by a more sensitive method, such as qRT-PCR, to confirm a direct relationship between DNA methylation and gene expression. Methylation-dependent expression of *RHOH* in cell lines has been validated by qRT-PCR for investigations reported in Chapter 3 of this project (see section 3.4.9.5). For the genes identified in Chapter 4, methylation-dependent expression of *S100A4* has previously been validated in medulloblastoma cell lines by RT-PCR (Lindsey *et al.*, 2007). This validation has not yet been carried out for *FAM46A*, *ACTC1*, *PRPH* or *ANXA2*.

Cell lines provide a resource of unlimited amounts of high quality tumour-derived DNA and RNA. However, studies have shown that medulloblastoma cell lines may not represent all medulloblastoma subtypes. They are known to contain molecular defects such as *MYC* amplification and 17p loss which are associated with more aggressive tumours (Langdon *et al.*, 2006). More recently,using a non-negative matrix factorisation (NMF) (Brunet *et al.*, 2004) consensus clustering approach, Ed Schwalbe (Newcastle University Paediatric Brain Tumour Group) demonstrated that multiple medulloblastoma cell line methylation patterns cluster with the worst prognosis *MYC*-amplified Group 3 medulloblastomas (Table 2.3; data not yet published). Furthermore, cell lines are known to undergo *de novo* methylation during *in vitro* cell culture and acquire artifactual methylation marks that are not associated with primary tumours (Bender *et al.*, 1998; Smiraglia *et al.*, 2001). For these reasons cell lines should be used

with caution in methylation studies and careful consideration should be given to the subgroup-specific nature of the DNA methylation event being investigated. For the candidate genes identified, bisulfite sequencing and qRT-PCR assessment of gene expression will be necessary in primary tumour samples to validate the subgroup-specific methylation-expression relationships observed in this project. Methylation status and gene expression will also need to be assessed in control cerebellar tissue samples to confirm the potential tumour suppressor or oncogenic role of the genes.

Medulloblastoma cell lines provide highly valuable models, particularly for the Group 3 and Group 4 tumour subgroups (see above), for functional assays to investigate the role of epigenetic deregulation in tumour development and progression. To determine any functional role, gene expression will need to be modulated in cell line model systems. This could be achieved through the use of lentiviral vector constructs that would deliver siRNA sequences into cells to silence target transcripts that are unmethylated and expressed in cell lines. Similarly, lentiviral vector systems could deliver and constitutively express genes that are methylated and silenced in cell lines. The effect of gene knockdown or overexpression on the phenotype of cells would then need to be investigated using assays that assess cell proliferation, migration and apoptosis, alongside investigations of any signalling pathways which the genes may function in or impact on.

The use of cancer cell lines in functional studies may be limited by their high homogeneity compared with tumours, which contain a heterogeneous population of cells. Moreover, cell lines growing in culture lack the complex *in vivo* environment with which tumour cells interact to sustain growth and migration capabilities (Vargo-Gogola and Rosen, 2007; Gazdar *et al.*, 2010). It is therefore important to carry out *in vivo* validation studies for those findings that are validated initially *in vitro*. Mouse models are currently available for SHH, WNT and Group 3 medulloblastomas (see section 1.5.9), and using lentiviral vector systems, gene expression may be similarly modulated in animal models as in cell line models. The effects of gene modulation may be monitored through the constitutive expression of an optical reporter gene and imaging systems. The candidate gene *RHOH* represents a Group 4-specific event and there are currently no mouse models that have been developed that represent this subgroup of medulloblastomas. The critical promoter-associated *RHOH* CpG residue

(RHOH_P121_F) was hypomethylated (β≤0.3) in 51% of Group 4 medullloblastomas, but was also hypomethylated in a considerable proportion (21%) of Group 3 tumours. It may therefore be considered rational to study any functional role of *RHOH* hypomethylation *in vivo* in Group 3 mouse models. The results from *in vitro* and subsequent *in vivo* functional studies will reveal any role for epigenetic deregulation of the candidate genes in the initiation and/or progression of the distinct subgroup(s) in which the aberrant epigenetic event is found and will identify their potential as therapeutic targets within these tumours.

5.3.2   Further work to investigate their clinical relevance in medulloblastoma

Aberrant DNA hypermethylation events that are associated with aberrant gene silencing in tumours, and that demonstrate a functional role in the initiation and/or progression of tumours may represent therapeutic targets that could potentially be targeted with demethylating agents such as 5'-aza-2'-deoxycytidine (Dacogen), which is currently approved for the treatment of MDS (see section 1.6.3.4). For those potential oncogenic genes whose expression is upregulated by an aberrant hypomethylation event, it may be possible to target their over-expressed protein product or other upstream or downstream mediators using either immunotherapy or targeted small molecule inhibitors (see section 1.2.2.1). The design and development of such targeted therapies would necessitate the accurate identification of patients in the clinical setting who carry the aberration and who would potentially benefit from the targeted therapies.

As well as representing potential key driver events in tumourigenesis, aberrant DNA methylation markers are beginning to show considerable promise as prognostic biomarkers in several cancer types, including in medulloblastoma (see section 1.6.3.4 and section 1.7.2.2). The identification of subgroup-specific methylation biomarkers offers the potential to identify functionally relevant markers that may also be prognostically relevant not only between the subgroups but, depending on their intra-subgroup variability, also within their respective subgroups.

It will be of great interest to investigate the potential prognostic relevance of the aberrant subgroup-specific methylation events associated with the candidate genes identified. For 3 of the candidate genes that were identified in Chapter 4 (*ANXA2*, *PRPH* and *S100A4*), their potential epigenetic dysregulation was specific to the WNT subgroup of tumours. WNT tumour status confers a very good long-term prognosis in

comparison to the other tumour subgroups (Ellison *et al.*, 2005; Clifford *et al.*, 2006; Kool *et al.*, 2012), with long-term survival rates likely exceeding 90% (Taylor *et al.*, 2012). As a result, the role of WNT-specific methylation biomarkers in disease prognostication is limited. The remaining candidate genes identified (*RHOH*, *FAM46A* and *ACTC1*) had aberrant methylation events specific to non-WNT tumour subgroups, and it will be important to investigate their prognostic relevance across all medulloblastomas in well-defined clinical trial cohorts for which mature event-free (EFS) and overall (OS) survival data is available. It will also be important to assess their prognostic relevance within their respective subgroups, to determine whether they can identify subsets of tumours within the core subgroup that behave differently, and that ultimately may benefit from less aggressive or more intensive therapeutic strategies depending on their risk. It will be important to assess the ability of the methylation markers to improve the accuracy and significance of current survival prediction by adding them to survival models, such as a Cox proportional hazards model (see below), which contain established medulloblastoma prognostic variables (see section 3.6 and section 4.6). The identification of prognostic disease biomarkers offers the potential for the development of a more accurate and refined risk stratification that will be essential to the delivery of personalised therapies, with the primary goal of improving survival and minimising the long-term adverse outcomes of aggressive treatment (Pizer and Clifford, 2009).

Alongside survival analysis, it will be important to investigate any significant correlations between the candidate gene methylation biomarkers and selected clinico-pathological features, such as age, gender, histological subtype and metastatic status. Significant associations would need to be considered further in survival analysis studies, as it would be important to assess any independent contribution to survival outcome prediction that the methylation biomarker confers, particularly if it is associated with an established high risk variable such as LCA histology or a positive metastatic status. These studies will be of great interest and the clinical data is now available to carry them out.

EFS and OS are defined as the time periods between diagnosis of disease to relapse or death, respectively. In medulloblastoma, EFS is considered a more accurate reflection of survival risk because patient survival following relapse is rare (Pizer and Clifford,

2009), but time to death may be influenced by variable therapeutic strategies in the relapse setting or by factors that are unrelated to their disease. Kaplan-Meier survival curves and log-rank tests are commonly employed in disease survival analysis. The Kaplan-Meier estimator (Kaplan and Meier, 1958) measures the fraction of patients living for a certain amount of time with or without an event occurring. The log-rank test (Mantel, 1966; Bland and Altman, 2004) is a hypothesis test that can be used to compare the survival distributions between different Kaplan-Meier curves for different groups tested. It tests the null hypothesis that there is no difference between survival curves. The Cox proportional hazards model (Cox, 1972) can also be used to compare different survival distributions. Compared to log-rank tests which require DNA methylation to be considered as a binary variable (methylated *versus* not methylated), the Cox model tests methylation as a continuous variable as measured by its β-value. Cox models can also assess the effects of one or more covariates that may be associated with the time that passes before an event occurs. Cox models are widely used in multivariate survival analysis studies.

5.4 Summary

This project presents a comprehensive investigation into the genome-wide role of DNA methylation in distinct molecular subgroups of medulloblastoma. Two high-throughput screening approaches have successfully identified six genes that show strong evidence of methylation-dependent gene regulation in medulloblastoma subgroups and that may play key roles in their distinct pathogenesis. These genes represent both putative oncogenes and tumour suppressor genes that are deregulated by aberrant methylation states at one or more CpG sites which are located in CpG island and non-island promoters and in the gene body. While for the large part, patterns of DNA methylation did not show clear associations with gene expression, this project focused only on negative methylation-expression correlations and assessment of methylation-expression relationships were carried out at the level of single CpG residues. It is likely that further epigenetically regulated genes may be identified in future genome-wide studies that investigate both positive and negative methylation-expression correlations, and that specifically target regions of the genome that contain important regulatory elements such as enhancers and transcription factor binding sites. Furthermore, the complete characterisation of subgroup-specific epigenetic events, to include DNA methylation

patterns alongside accompanying histone modifications and nucleosome structure, may provide greater insights into the complexity of epigenetic gene dysregulation in medulloblastoma and the complex role that DNA methylation may play in mediating any transcriptional responses.

Alongside the identification of additional epigenetic events that may be critical in the development of distinct subgroups of medulloblastoma, the six genes identified in this project now merit further investigation. The aberrant DNA methylation events and associated gene transcriptional effects may represent potential driver events in medulloblastoma subgroup pathogenesis that could be targeted therapeutically. The subgroup-specific aberrant DNA methylation events also have the potential to represent prognostic markers that could refine and improve current stratification models and allow appropriate therapies to be more accurately tailored to a patient's risk.

Appendix A

Group 3 v cerebella			Group 4 v cerebella			Group (3+4) v cerebella		
CpG probe	Δβ	p value	CpG probe	Δβ	p value	CpG probe	Δβ	p value
BCAP31_P1072_F	0.59	1.94E-07	ASCL2_P609_R	0.10	0.0002	ASCL2_E76_R	0.05	4.72E-06
BCR_P422_F	0.40	2.19E-05	CCKAR_E79_F	0.18	8.76E-08	ASCL2_P360_F	0.07	0.0001
BLK_P668_R	0.73	3.68E-10	CCKAR_P270_F	0.78	1.93E-11	ASCL2_P609_R	0.08	0.0008
CCL3_E53_R	0.44	1.31E-08	CSF3R_P472_F	0.59	1.93E-11	AXIN1_P995_R	0.09	0.006
CHI3L2_P226_F	0.66	7.57E-16	DDR2_E331_F	0.02	0.002	BGN_P333_R	0.04	0.31
FES_E34_R	0.54	2.18E-08	FZD9_E458_F	0.06	0.97	BLK_P668_R	0.41	2.08E-08
FES_P223_R	0.42	0.000135	HDAC1_P414_R	0.03	0.19	CAPG_E228_F	0.15	3.51E-08
FGF1_P357_R	0.54	9.77E-11	HIC-	0.35	0.0004	CCKAR_E79_F	0.14	2.79E-06
FRZB_E186_R	0.42	0.0009	HLA-DPA1_P28_R	0.14	9.67E-06	CCKAR_P270_F	0.77	1.06E-11
HFE_E273_R	0.87	3.68E-10	IFNGR2_P377_R	0.36	1.03E-10	CEACAM1_E57_R	0.50	8.21E-11
HLA-DOB_P357_R	0.56	3.68E-10	IGF1_E394_F	0.15	2.99E-05	CHI3L2_P226_F	0.47	3.07E-11
IL1RN_P93_R	0.36	1.09E-05	IL16_P93_R	0.64	4.85E-11	CSF3R_P472_F	0.51	5.88E-11
KRT1_P798_R	0.38	1.94E-07	L1CAM_P19_F	0.80	1.43E-10	CYP2E1_E53_R	0.06	0.973921
MMP14_P13_F	0.05	0.53	LEFTY2_P719_F	0.16	2.60E-05	DDR2_E331_F	0.01	0.006
PLA2G2A_E268_F	0.23	0.03	MEST_E150_F	0.13	0.97	FGF1_E5_F	0.40	6.87E-09
RARRES1_P426_R	0.58	7.20E-09	MEST_P4_F	0.14	0.79	FZD9_E458_F	0.08	0.993912
SERPINA5_E69_F	0.34	1.17E-06	MEST_P62_R	0.09	0.56	HDAC1_P414_R	0.14	0.73
TGFB2_E226_R	0.48	0.0007	MET_E333_F	0.65	6.53E-10	HFE_E273_R	0.73	2.53E-11
THBS2_P605_R	0.52	6.96E-12	MMP10_E136_R	0.71	1.93E-11	HIC-	0.22	0.11
TP73_P945_F	0.54	1.34E-07	NOTCH4_P938_F	0.13	0.11	IFNGR2_P377_R	0.29	9.36E-08
TRIP6_E33_F	0.54	7.18E-12	PADI4_P1158_R	0.02	0.18	IL16_P226_F	0.61	2.53E-11
WRN_P969_F	0.57	3.07E-09	PIK3R1_P307_F	0.19	2.31E-07	IL16_P93_R	0.63	2.42E-11
ZNFN1A1_E102_F	0.44	2.45E-07	PLG_E406_F	0.56	1.93E-11	IL1RN_E42_F	0.06	0.20
			RAN_P581_R	0.02	0.54	IL8_P83_F	0.06	0.681
			RHOH_P121_F	0.47	7.55E-06	L1CAM_P19_F	0.69	4.12E-09
			SPP1_E140_R	0.42	2.09E-09	LCN2_P86_R	0.35	1.09E-09
			TM7SF3_P1068_R	0.28	0.001	LEFTY2_P719_F	0.10	0.003
			TRIM29_E189_F	0.06	0.05	MEST_E150_F	0.16	0.79
			WNT10B_P993_F	0.59	1.43E-10	MEST_P4_F	0.17	0.99
			ZIM3_P718_R	0.57	2.18E-11	MEST_P62_R	0.13	0.99
			ZNF264_P397_F	0.11	1.06E-05	MET_E333_F	0.65	2.25E-10
						MMP10_E136_R	0.64	2.89E-11
						MSH2_P1008_F	0.66	1.06E-11
						PADI4_P1158_R	0.08	0.99
						PIK3R1_P307_F	0.13	0.0002
						PLG_E406_F	0.55	1.06E-11
						RHOH_P121_F	0.39	0.0004
						SPARC_P195_F	0.08	0.23

Group 3 *v* cerebella				Group 4 *v* cerebella				Group (3+4) *v* cerebella		
CpG probe	Δβ	p value		CpG probe	Δβ	p value		CpG probe	Δβ	p value
								SPP1_E140_R	0.39	2.88E-09
								STAT5A_E42_F	0.00	0.21
								TAL1_P594_F	0.08	0.79
								TFPI2_P152_R	0.43	8.21E-11
								TRIM29_E189_F	0.01	0.31
								VAV1_E9_F	0.17	5.88E-11
								VAV1_P317_F	0.28	7.06E-10
								WNT10B_P993_F	0.51	7.30E-10
								ZIM3_P718_R	0.51	2.53E-11
								ZNF264_P397_F	0.05	0.0008

Results of Mann-Whitney U tests, assessing methylation of Group 3-, Group 4- and Group (3+4)-specific events against methylation of non-neoplastic cerebellar samples. Δβ is the absolute difference in mean β values between groups tested and *p*-values were corrected using the Benjamini-Hochberg false discovery rate correction. Probe methylation was considered significantly different between tumours and control cerebella, and thus classed tumour-specific, if the FDR-adjusted *p*-value was <0.05 and the Δβ value was >0.34.

Appendix B

	Kool/Fattet			Cho			Northcott		
	Probe ID	Mean log ratio	*p*-value	Probe ID	Mean log ratio	*p*-value	Probe ID	Mean log ratio	*p*-value
Group 3 v Other	200837_at	0.59	2.00E-03	200837_at	0.76	5.59E-09	3528864	1.20	8.44E-06
	205117_at	0.79	0.02	206391_at	0.34	1.00E-04	2702752	0.31	3.32E-05
	205667_at	0.60	0.02	213060_s_at	0.64	1.40E-03	2380590	1.00	3.10E-04
	213060_s_at	0.37	0.02	203697_at	0.73	3.10E-03	4026669	0.47	1.60E-03
	206391_at	0.28	0.04	209129_at	0.82	3.10E-03	3001479	0.15	0.03
	220406_at	0.08	0.36	205117_at	0.71	3.80E-03	2351687	0.12	0.08
	203649_s_at	0.22	0.49	205418_at	0.15	3.80E-03	3015911	0.14	0.09
	205900_at	0.07	0.55	160020_at	0.14	0.01	2899110	0.14	0.10
	206255_at	0.10	0.55	205038_at	0.11	0.05	2879166	0.18	0.14
	209443_at	0.20	0.55	206255_at	0.13	0.09	3608427	0.12	0.15
	203697_at	0.20	0.59	205671_s_at	0.11	0.10	2590715	0.39	0.22
	205038_at	0.04	0.59	211330_s_at	0.08	0.33	2950145	0.08	0.38
	205114_s_at	0.31	0.59	216244_at	0.09	0.33	2317317	0.07	0.41
	209129_at	0.26	0.59	217223_s_at	0.12	0.59	3085990	0.05	0.58
	217223_s_at	0.15	0.59	205114_s_at	0.13	0.67	3549740	0.07	0.59
	232546_at	0.11	0.59	220406_at	0.03	0.69	3092663	0.15	0.62
	235754_at	0.06	0.59	203083_at	0.16	0.80	2985781	0.02	0.86
	203083_at	0.23	0.62	209443_at	0.06	0.80	2400027	0.01	0.88
	160020_at	0.03	0.72	205667_at	0.07	0.95	2501204	0.01	0.88
	205671_s_at	0.03	0.72	203649_s_at	0.01	0.96	3455728	0.01	0.88
	216244_at	0.01	0.74	205900_at	0.01	0.96	3754009	0.03	0.88
	205418_at	0.02	0.86	220804_s_at	0.00	0.96	3939183	0.02	0.88
Group 4 v Other	201209_at	1.99	2.14E-18	201209_at	2.75	8.22E-21	2328868	1.45	9.73E-14
	200750_s_at	0.69	3.70E-13	200750_s_at	0.74	4.61E-11	2458629	0.81	6.89E-07
	205168_at	1.08	8.80E-06	203510_at	2.18	2.45E-09	3448481	0.66	9.14E-07
	209540_at	1.04	1.10E-04	217974_at	0.77	1.39E-07	2364231	1.02	1.80E-06
	203510_at	1.19	2.10E-04	205168_at	0.92	2.55E-06	3020343	1.29	2.63E-06
	206012_at	0.58	1.60E-03	206012_at	0.79	4.47E-06	2724671	0.27	8.80E-04
	202504_at	0.18	2.10E-03	202016_at	0.84	5.80E-04	2950329	0.71	1.20E-03
	207639_at	0.21	6.60E-03	209540_at	0.98	3.90E-04	3438027	0.33	1.20E-03
	217974_at	0.42	6.60E-03	204951_at	0.40	0.01	3872138	0.28	1.20E-03
	205680_at	0.15	0.02	201642_at	0.17	0.04	3468345	0.30	1.30E-03
	205917_at	0.29	0.04	209875_s_at	0.68	0.04	3843058	0.18	0.01
	209977_at	0.09	0.04	211990_at	0.52	0.10	3453513	0.14	0.02
	204951_at	0.40	0.04	204584_at	0.43	0.13	3359224	0.18	0.02

	Kool/Fattet			Cho			Northcott		
	Probe ID	**Mean log ratio**	***p*-value**	**Probe ID**	**Mean log ratio**	***p*-value**	**Probe ID**	**Mean log ratio**	***p*-value**
Group 4 v Other	202016_at	0.51	0.10	206213_at	0.19	0.13	2813060	0.31	0.09
	211173_at	0.10	0.10	205247_at	0.20	0.16	2934682	0.07	0.13
	211412_at	0.05	0.10	211412_at	0.10	0.23	2406783	0.08	0.16
	211990_at	0.39	0.19	211173_at	0.09	0.39	3007829	0.08	0.16
	208461_at	0.06	0.24	207607_at	0.00	0.41	3394660	0.05	0.25
	201642_at	0.07	0.29	209827_s_at	0.05	0.59	3604287	0.12	0.25
	203591_s_at	0.07	0.54	212239_at	0.15	0.59	2322848	0.06	0.33
	212239_at	0.17	0.54	205917_at	0.12	0.62	3024025	0.10	0.60
	204584_at	0.17	0.60	205680_at	0.05	0.64	2735027	0.22	0.60
	205247_at	0.05	0.60	209977_at	0.03	0.68	2949901	0.02	0.68
	206213_at	0.03	0.77	208461_at	0.01	0.73	2764478	0.02	0.83
	207607_at	0.01	0.81	203591_s_at	0.01	0.79	3388785	0.02	0.83
	209875_s_at	0.03	0.98	202504_at	0.00	0.87	3706113	0.01	0.83
	1553022_at	0.00	0.99	207639_at	0.01	0.87	3918635	0.03	0.83
	1555016_at	0.00	0.99				4026798	0.04	0.83
Group (3+4) v (SHH+WNT)	201209_at	1.28	3.25E-06	201209_at	1.68	9.75E-11	2328868	1.00	3.39E-07
	205168_at	1.36	8.80E-06	203510_at	3.13	9.75E-11	3020343	1.75	2.66E-06
	203010_at	0.47	2.78E-05	205168_at	1.70	3.63E-10	2364231	1.34	7.33E-06
	202016_at	1.32	8.01E-05	202016_at	1.50	1.30E-09	3721658	0.51	1.09E-05
	212849_at	0.34	8.01E-05	209421_at	0.59	1.30E-06	2724671	0.26	8.32E-05
	204951_at	0.66	1.40E-03	203010_at	0.67	7.48E-06	3675047	0.17	1.80E-04
	201262_s_at	0.57	1.60E-03	204951_at	0.38	1.40E-04	2899110	0.28	3.60E-04
	209421_at	0.40	1.90E-03	212849_at	0.26	0.01	3995633	0.58	1.10E-03
	201642_at	0.16	7.90E-03	212239_at	0.54	0.01	2480992	0.34	6.30E-03
	203510_at	1.12	0.01	204584_at	0.82	0.02	2764478	0.15	6.30E-03
	207639_at	0.23	0.02	212667_at	0.50	0.02	3863669	0.14	6.50E-03
	202504_at	0.16	0.03	201642_at	0.21	0.02	3453513	0.14	0.02
	206219_s_at	0.12	0.09	211330_s_at	0.16	0.05	3872138	0.20	0.02
	205917_at	0.25	0.16	209875_s_at	0.70	0.09	2411198	0.14	0.04
	206255_at	0.30	0.16	206213_at	0.16	0.15	3394660	0.08	0.05
	213060_s_at	0.28	0.16	206255_at	0.23	0.15	3818547	0.10	0.05
	211412_at	0.04	0.18	1431_at	0.17	0.16	3272981	0.13	0.11
	205117_at	0.41	0.23	211412_at	0.06	0.21	2882098	0.23	0.13
	205680_at	0.09	0.27	211173_at	0.11	0.28	2934682	0.07	0.14
	209498_at	0.13	0.27	206012_at	0.28	0.29	3190190	0.09	0.14

	Kool/Fattet			Cho			Northcott		
	Probe ID	**Mean log ratio**	***p*-value**	**Probe ID**	**Mean log ratio**	***p*-value**	**Probe ID**	**Mean log ratio**	***p*-value**
Group (3+4) v (SHH+WNT)	209875_s_at	0.54	0.29	206219_s_at	0.05	0.29	2458629	0.28	0.17
	206213_at	0.09	0.32	209277_at	0.39	0.29	2879166	0.14	0.17
	212531_at	0.10	0.36	211506_s_at	0.13	0.32	2735027	0.55	0.17
	203591_s_at	0.11	0.37	202504_at	0.03	0.38	2322848	0.07	0.19
	201850_at	0.18	0.45	207639_at	0.09	0.38	3061621	0.44	0.27
	235754_at	0.04	0.45	209827_s_at	0.08	0.38	3024025	0.17	0.28
	209977_at	0.04	0.48	207607_at	0.00	0.41	3843058	0.09	0.28
	200665_s_at	0.18	0.48	205917_at	0.17	0.41	3085990	0.06	0.33
	209277_at	0.08	0.49	209977_at	0.05	0.41	2351687	0.05	0.35
	216244_at	0.02	0.63	209498_at	0.05	0.45	2813060	0.11	0.60
	206012_at	0.09	0.70	216244_at	0.03	0.45	2406783	0.03	0.61
	1555016_at	0.02	0.84	203591_s_at	0.03	0.46	3007829	0.03	0.65
	1431_at	0.02	0.84	212531_at	0.05	0.46	3706113	0.02	0.68
	204584_at	0.07	0.93	205680_at	0.07	0.55	3359224	0.04	0.72
	208461_at	0.01	0.93	213060_s_at	0.17	0.55	3388785	0.04	0.78
	211173_at	0.01	0.93	201850_at	0.14	0.57	3918635	0.03	0.83
	211506_s_at	0.02	0.93	205117_at	0.14	0.63	2731332	0.03	0.84
	212239_at	0.04	0.93	216928_at	0.02	0.64	2501204	0.01	0.84
	207607_at	0.01	0.93	208461_at	0.01	0.66	3604287	0.02	0.85
	1553022_at	0.00	0.98	201262_s_at	0.01	0.82	4026798	0.01	0.98
	216928_at	0.00	0.98				2562271	0.00	0.98

Results of unpaired *t*-tests, assessing differential gene expression between medulloblastoma subgroups across 3 independent transcriptomic datasets. Genes that contained subgroup-specific differentially methylated CpG sites were tested for evidence of differential expression in the equivalent subgroup. Mean log ratio is the absolute difference in mean $\log_2$ signal intensity between groups tested and *p*-values were corrected using the Benjamini-Hochberg false discovery rate correction. The average absolute difference in mean $\log_2$ intensities for all tests carried out was 0.3 in the Kool/Fattet dataset, 0.4 in the Cho dataset and 0.3 in the Northcott dataset. These values were applied as the mean log ratio thresholds. Transcripts with a FDR-adjusted *p*-value <0.05 and a mean log ratio > threshold in at least 2 out of 3 datasets were considered significantly differentially expressed.

References

Akiyama, Y., Maesawa, C., Ogasawara, S., Terashima, M. and Masuda, T. (2003) 'Cell-type-specific repression of the maspin gene is disrupted frequently by demethylation at the promoter region in gastric intestinal metaplasia and cancer cells', *Am J Pathol*, 163(5), pp. 1911-9.

Albright, A.L., Wisoff, J.H., Zeltzer, P.M., Boyett, J.M., Rorke, L.B. and Stanley, P. (1996) 'Effects of medulloblastoma resections on outcome in children: a report from the Children's Cancer Group', *Neurosurgery*, 38(2), pp. 265-71.

Aldosari, N., Bigner, S.H., Burger, P.C., Becker, L., Kepner, J.L., Friedman, H.S. and McLendon, R.E. (2002a) 'MYCC and MYCN oncogene amplification in medulloblastoma. A fluorescence in situ hybridization study on paraffin sections from the Children's Oncology Group', *Arch Pathol Lab Med*, 126(5), pp. 540-4.

Aldosari, N., Wiltshire, R.N., Dutra, A., Schrock, E., McLendon, R.E., Friedman, H.S., Bigner, D.D. and Bigner, S.H. (2002b) 'Comprehensive molecular cytogenetic investigation of chromosomal abnormalities in human medulloblastoma cell lines and xenograft', *Neuro Oncol*, 4(2), pp. 75-85.

Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Aparicio, S.A., Behjati, S., Biankin, A.V., Bignell, G.R., Bolli, N., Borg, A., Borresen-Dale, A.L., Boyault, S., Burkhardt, B., Butler, A.P., Caldas, C., Davies, H.R., Desmedt, C., Eils, R., Eyfjord, J.E., Foekens, J.A., Greaves, M., Hosoda, F., Hutter, B., Ilicic, T., Imbeaud, S., Imielinsk, M., Jager, N., Jones, D.T., Jones, D., Knappskog, S., Kool, M., Lakhani, S.R., Lopez-Otin, C., Martin, S., Munshi, N.C., Nakamura, H., Northcott, P.A., Pajic, M., Papaemmanuil, E., Paradiso, A., Pearson, J.V., Puente, X.S., Raine, K., Ramakrishna, M., Richardson, A.L., Richter, J., Rosenstiel, P., Schlesner, M., Schumacher, T.N., Span, P.N., Teague, J.W., Totoki, Y., Tutt, A.N., Valdes-Mas, R., van Buuren, M.M., van 't Veer, L., Vincent-Salomon, A., Waddell, N., Yates, L.R., Zucman-Rossi, J., Futreal, P.A., McDermott, U., Lichter, P., Meyerson, M., Grimmond, S.M., Siebert, R., Campo, E., Shibata, T., Pfister, S.M., Campbell, P.J. and Stratton, M.R. (2013) 'Signatures of mutational processes in human cancer', *Nature*, 500(7463), pp. 415-21.

Anderton, J.A., Lindsey, J.C., Lusher, M.E., Gilbertson, R.J., Bailey, S., Ellison, D.W. and Clifford, S.C. (2008) 'Global analysis of the medulloblastoma epigenome identifies disease-subgroup-specific inactivation of COL1A2', *Neuro Oncol*, 10(6), pp. 981-94.

Aran, D., Sabato, S. and Hellman, A. (2013) 'DNA methylation of distal regulatory sites characterizes dysregulation of cancer genes', *Genome Biol*, 14(3), p. R21.

Badal, V., Chuang, L.S., Tan, E.H., Badal, S., Villa, L.L., Wheeler, C.M., Li, B.F. and Bernard, H.U. (2003) 'CpG methylation of human papillomavirus type 16 DNA in cervical cancer cell lines and in clinical specimens: genomic hypomethylation correlates with carcinogenic progression', *J Virol*, 77(11), pp. 6227-34.

Bailey, C.C., Gnekow, A., Wellek, S., Jones, M., Round, C., Brown, J., Phillips, A. and Neidhardt, M.K. (1995) 'Prospective randomised trial of chemotherapy given before radiotherapy in childhood medulloblastoma. International Society of Paediatric Oncology (SIOP) and the (German) Society of Paediatric Oncology (GPO): SIOP II', *Med Pediatr Oncol*, 25(3), pp. 166-78.

Bailey, P. and Cushing, H. (1925) 'Medulloblastoma Cerebelli: a Common Type of Midcerebellar Glioma of Childhood', *Archives of Neurology and Psychiatry*, 14(2), pp. 192-224.

Ball, M.P., Li, J.B., Gao, Y., Lee, J.H., LeProust, E.M., Park, I.H., Xie, B., Daley, G.Q. and Church, G.M. (2009) 'Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells', *Nat Biotechnol*, 27(4), pp. 361-8.

Bartel, D.P. (2004) 'MicroRNAs: genomics, biogenesis, mechanism, and function', *Cell*, 116(2), pp. 281-97.

Bartlett, J.M. and Stirling, D. (2003) 'A short history of the polymerase chain reaction', *Methods Mol Biol*, 226, pp. 3-6.

Baselga, J. (2006) 'Targeting tyrosine kinases in cancer: the second wave', *Science*, 312(5777), pp. 1175-8.

Baselga, J., Norton, L., Albanell, J., Kim, Y.M. and Mendelsohn, J. (1998) 'Recombinant humanized anti-HER2 antibody (Herceptin) enhances the antitumor activity of paclitaxel and doxorubicin against HER2/neu overexpressing human breast cancer xenografts', *Cancer Res*, 58(13), pp. 2825-31.

Batora, N.V., Sturm, D., Jones, D.T., Kool, M., Pfister, S.M. and Northcott, P.A. (2013) 'Transitioning from genotypes to epigenotypes: Why the time has come for medulloblastoma epigenomics', *Neuroscience*.

Baylin, S.B. and Jones, P.A. (2011) 'A decade of exploring the cancer epigenome - biological and translational implications', *Nat Rev Cancer*, 11(10), pp. 726-34.

Baylin, S.B. and Ohm, J.E. (2006) 'Epigenetic gene silencing in cancer - a mechanism for early oncogenic pathway addiction?', *Nat Rev Cancer*, 6(2), pp. 107-16.

Bender, C.M., Pao, M.M. and Jones, P.A. (1998) 'Inhibition of DNA methylation by 5-aza-2'-deoxycytidine suppresses the growth of human tumor cell lines', *Cancer Res*, 58(1), pp. 95-101.

Benjamini, Y. and Hochberg, Y. (1995) 'Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing.', *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1), pp. 289-300.

Bibikova, M., Barnes, B., Tsan, C., Ho, V., Klotzle, B., Le, J.M., Delano, D., Zhang, L., Schroth, G.P., Gunderson, K.L., Fan, J.B. and Shen, R. (2011) 'High density DNA methylation array with single CpG site resolution', *Genomics*, 98(4), pp. 288-95.

Bibikova, M., Le, J., Barnes, B., Saedinia-Melnyk, S., Zhou, L., Shen, R. and Gunderson, K.L. (2009) 'Genome-wide DNA methylation profiling using Infinium(R) assay', *Epigenomics*, 1(1), pp. 177-200.

Bibikova, M., Lin, Z., Zhou, L., Chudin, E., Garcia, E.W., Wu, B., Doucet, D., Thomas, N.J., Wang, Y., Vollmer, E., Goldmann, T., Seifart, C., Jiang, W., Barker, D.L., Chee, M.S., Floros, J. and Fan, J.B. (2006) 'High-throughput DNA methylation profiling using universal bead arrays', *Genome Res*, 16(3), pp. 383-93.

Bigner, S.H., Mark, J., Friedman, H.S., Biegel, J.A. and Bigner, D.D. (1988) 'Structural chromosomal abnormalities in human medulloblastoma', *Cancer Genet Cytogenet*, 30(1), pp. 91-101.

Binning, M.J., Niazi, T., Pedone, C.A., Lal, B., Eberhart, C.G., Kim, K.J., Laterra, J. and Fults, D.W. (2008) 'Hepatocyte growth factor and sonic Hedgehog expression in cerebellar neural progenitor cells costimulate medulloblastoma initiation and growth', *Cancer Res*, 68(19), pp. 7838-45.

Bird, A. (2002) 'DNA methylation patterns and epigenetic memory', *Genes Dev*, 16(1), pp. 6-21.

Bird, A. (2007) 'Perceptions of epigenetics', *Nature*, 447(7143), pp. 396-8.

Bird, A.P., Taggart, M.H., Nicholls, R.D. and Higgs, D.R. (1987) 'Non-methylated CpG-rich islands at the human alpha-globin locus: implications for evolution of the alpha-globin pseudogene', *Embo j*, 6(4), pp. 999-1004.

Birgisdottir, V., Stefansson, O.A., Bodvarsdottir, S.K., Hilmarsdottir, H., Jonasson, J.G. and Eyfjord, J.E. (2006) 'Epigenetic silencing and deletion of the BRCA1 gene in sporadic breast cancer', *Breast Cancer Res*, 8(4), p. R38.

Blackledge, N.P. and Klose, R. (2011) 'CpG island chromatin: a platform for gene regulation', *Epigenetics*, 6(2), pp. 147-52.

Blackledge, N.P., Zhou, J.C., Tolstorukov, M.Y., Farcas, A.M., Park, P.J. and Klose, R.J. (2010) 'CpG islands recruit a histone H3 lysine 36 demethylase', *Mol Cell*, 38(2), pp. 179-90.

Bland, J.M. and Altman, D.G. (2004) 'The logrank test', *BMJ*, 328(7447), p. 1073.

Bolstad, B.M., Irizarry, R.A., Astrand, M. and Speed, T.P. (2003) 'A comparison of normalization methods for high density oligonucleotide array data based on variance and bias', *Bioinformatics*, 19(2), pp. 185-93.

Bouffet, E., Bernard, J.L., Frappaz, D., Gentet, J.C., Roche, H., Tron, P., Carrie, C., Raybaud, C., Joannard, A., Lapras, C. and et al. (1992) 'M4 protocol for cerebellar medulloblastoma: supratentorial radiotherapy may not be avoided', *Int J Radiat Oncol Biol Phys*, 24(1), pp. 79-85.

Bougel, S., Lhermitte, B., Gallagher, G., de Flaugergues, J.C., Janzer, R.C. and Benhattar, J. (2013) 'Methylation of the hTERT promoter: a novel cancer biomarker for leptomeningeal metastasis detection in cerebrospinal fluids', *Clin Cancer Res*, 19(8), pp. 2216-23.

Bracken, A.P. and Helin, K. (2009) 'Polycomb group proteins: navigators of lineage pathways led astray in cancer', *Nat Rev Cancer*, 9(11), pp. 773-84.

Brandeis, M., Frank, D., Keshet, I., Siegfried, Z., Mendelsohn, M., Nemes, A., Temper, V., Razin, A. and Cedar, H. (1994) 'Sp1 elements protect a CpG island from de novo methylation', *Nature*, 371(6496), pp. 435-8.

Branton, D., Deamer, D.W., Marziali, A., Bayley, H., Benner, S.A., Butler, T., Di Ventra, M., Garaj, S., Hibbs, A., Huang, X., Jovanovich, S.B., Krstic, P.S., Lindsay, S.,

Ling, X.S., Mastrangelo, C.H., Meller, A., Oliver, J.S., Pershin, Y.V., Ramsey, J.M., Riehn, R., Soni, G.V., Tabard-Cossa, V., Wanunu, M., Wiggin, M. and Schloss, J.A. (2008) 'The potential and challenges of nanopore sequencing', *Nat Biotechnol*, 26(10), pp. 1146-53.

Brock, M.V., Hooker, C.M., Ota-Machida, E., Han, Y., Guo, M., Ames, S., Glockner, S., Piantadosi, S., Gabrielson, E., Pridham, G., Pelosky, K., Belinsky, S.A., Yang, S.C., Baylin, S.B. and Herman, J.G. (2008) 'DNA methylation markers and early recurrence in stage I lung cancer', *N Engl J Med*, 358(11), pp. 1118-28.

Brown, H.G., Kepner, J.L., Perlman, E.J., Friedman, H.S., Strother, D.R., Duffner, P.K., Kun, L.E., Goldthwaite, P.T. and Burger, P.C. (2000) '"Large cell/anaplastic" medulloblastomas: a Pediatric Oncology Group Study', *J Neuropathol Exp Neurol*, 59(10), pp. 857-65.

Brown, S.J., Stoilov, P. and Xing, Y. (2012) 'Chromatin and epigenetic regulation of pre-mRNA processing', *Hum Mol Genet*, 21(R1), pp. R90-6.

Brunet, J.P., Tamayo, P., Golub, T.R. and Mesirov, J.P. (2004) 'Metagenes and molecular pattern discovery using matrix factorization', *Proc Natl Acad Sci U S A*, 101(12), pp. 4164-9.

Cairns, P., Esteller, M., Herman, J.G., Schoenberg, M., Jeronimo, C., Sanchez-Cespedes, M., Chow, N.H., Grasso, M., Wu, L., Westra, W.B. and Sidransky, D. (2001) 'Molecular detection of prostate cancer in urine by GSTP1 hypermethylation', *Clin Cancer Res*, 7(9), pp. 2727-30.

Cameron, E.E., Bachman, K.E., Myohanen, S., Herman, J.G. and Baylin, S.B. (1999) 'Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer', *Nat Genet*, 21(1), pp. 103-7.

Carter, P., Presta, L., Gorman, C.M., Ridgway, J.B., Henner, D., Wong, W.L., Rowland, A.M., Kotts, C., Carver, M.E. and Shepard, H.M. (1992) 'Humanization of an anti-p185HER2 antibody for human cancer therapy', *Proc Natl Acad Sci U S A*, 89(10), pp. 4285-9.

Castelo-Branco, P., Choufani, S., Mack, S., Gallagher, D., Zhang, C., Lipman, T., Zhukova, N., Walker, E.J., Martin, D., Merino, D., Wasserman, J.D., Elizabeth, C., Alon, N., Zhang, L., Hovestadt, V., Kool, M., Jones, D.T., Zadeh, G., Croul, S., Hawkins, C., Hitzler, J., Wang, J.C., Baruchel, S., Dirks, P.B., Malkin, D., Pfister, S., Taylor, M.D., Weksberg, R. and Tabori, U. (2013) 'Methylation of the TERT promoter and risk stratification of childhood brain tumours: an integrative genomic and molecular study', *Lancet Oncol*, 14(6), pp. 534-42.

Cavenee, W.K., Dryja, T.P., Phillips, R.A., Benedict, W.F., Godbout, R., Gallie, B.L., Murphree, A.L., Strong, L.C. and White, R.L. (1983) 'Expression of recessive alleles by chromosomal mechanisms in retinoblastoma', *Nature*, 305(5937), pp. 779-84.

Chang, C.H., Housepian, E.M. and Herbert, C., Jr. (1969) 'An operative staging system and a megavoltage radiotherapeutic technic for cerebellar medulloblastomas', *Radiology*, 93(6), pp. 1351-9.

Chen, Z.X. and Riggs, A.D. (2011) 'DNA methylation and demethylation in mammals', *J Biol Chem*, 286(21), pp. 18347-53.

Chi, S.N., Gardner, S.L., Levy, A.S., Knopp, E.A., Miller, D.C., Wisoff, J.H., Weiner, H.L. and Finlay, J.L. (2004) 'Feasibility and response to induction chemotherapy intensified with high-dose methotrexate for young children with newly diagnosed high-risk disseminated medulloblastoma', *J Clin Oncol*, 22(24), pp. 4881-7.

Chiba, T., Yokosuka, O., Fukai, K., Hirasawa, Y., Tada, M., Mikata, R., Imazeki, F., Taniguchi, H., Iwama, A., Miyazaki, M., Ochiai, T. and Saisho, H. (2005) 'Identification and investigation of methylated genes in hepatoma', *Eur J Cancer*, 41(8), pp. 1185-94.

Chik, F. and Szyf, M. (2011) 'Effects of specific DNMT gene depletion on cancer cell transformation and breast cancer cell invasion; toward selective DNMT inhibitors', *Carcinogenesis*, 32(2), pp. 224-32.

Cho, Y.J., Tsherniak, A., Tamayo, P., Santagata, S., Ligon, A., Greulich, H., Berhoukim, R., Amani, V., Goumnerova, L., Eberhart, C.G., Lau, C.C., Olson, J.M., Gilbertson, R.J., Gajjar, A., Delattre, O., Kool, M., Ligon, K., Meyerson, M., Mesirov, J.P. and Pomeroy, S.L. (2011) 'Integrative genomic analysis of medulloblastoma identifies a molecular subgroup that drives poor clinical outcome', *J Clin Oncol*, 29(11), pp. 1424-30.

Christman, J.K. (2002) '5-Azacytidine and 5-aza-2'-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy', *Oncogene*, 21(35), pp. 5483-95.

Ciani, L. and Salinas, P.C. (2005) 'WNTs in the vertebrate nervous system: from patterning to neuronal connectivity', *Nat Rev Neurosci*, 6(5), pp. 351-62.

Clark, S.J., Harrison, J., Paul, C.L. and Frommer, M. (1994) 'High sensitivity mapping of methylated cytosines', *Nucleic Acids Res*, 22(15), pp. 2990-7.

Clarke, J., Wu, H.C., Jayasinghe, L., Patel, A., Reid, S. and Bayley, H. (2009) 'Continuous base identification for single-molecule nanopore DNA sequencing', *Nat Nanotechnol*, 4(4), pp. 265-70.

Clifford, S.C., Lusher, M.E., Lindsey, J.C., Langdon, J.A., Gilbertson, R.J., Straughton, D. and Ellison, D.W. (2006) 'Wnt/Wingless pathway activation and chromosome 6 loss characterize a distinct molecular sub-group of medulloblastomas associated with a favorable prognosis', *Cell Cycle*, 5(22), pp. 2666-70.

Comoglio, P.M., Giordano, S. and Trusolino, L. (2008) 'Drug development of MET inhibitors: targeting oncogene addiction and expedience', *Nat Rev Drug Discov*, 7(6), pp. 504-16.

Cox, D.R. (1972) 'Regression Models and Life-tables', *Journal of the Royal Statistical Society. Series B (Methodological)*, 34(2), pp. 187-220.

Crawford, J.R., MacDonald, T.J. and Packer, R.J. (2007) 'Medulloblastoma in childhood: new biological advances', *Lancet Neurol*, 6(12), pp. 1073-85.

Croce, C.M. (2008) 'Oncogenes and cancer', *N Engl J Med*, 358(5), pp. 502-11.

CRUK (2013a) *Cancer statistics in the UK 2010*. Available at: www.cancerresearchuk.org/cancer-info/cancerstats (Accessed: September 20, 2013).

CRUK (2013b) *Cancer Statistics in the UK 2010: Childhood cancer statistics*. Available at: http://www.cancerresearchuk.org/cancer-info/cancerstats/childhoodcancer/incidence/#brain (Accessed: September 20, 2013).

Cunningham, S.C., Gallmeier, E., Hucl, T., Dezentje, D.A., Abdelmohsen, K., Gorospe, M. and Kern, S.E. (2006) 'Theoretical proposal: allele dosage of MAP2K4/MKK4 could rationalize frequent 17p loss in diverse human cancers', *Cell Cycle*, 5(10), pp. 1090-3.

Dahmane, N. and Ruiz i Altaba, A. (1999) 'Sonic hedgehog regulates the growth and patterning of the cerebellum', *Development*, 126(14), pp. 3089-100.

Dammann, R., Li, C., Yoon, J.H., Chin, P.L., Bates, S. and Pfeifer, G.P. (2000) 'Epigenetic inactivation of a RAS association domain family protein from the lung tumour suppressor locus 3p21.3', *Nat Genet*, 25(3), pp. 315-9.

Dawson, M.A. and Kouzarides, T. (2012) 'Cancer epigenetics: from mechanism to therapy', *Cell*, 150(1), pp. 12-27.

De Braganca, K.C. and Packer, R.J. (2013) 'Treatment Options for Medulloblastoma and CNS Primitive Neuroectodermal Tumor (PNET)', *Curr Treat Options Neurol*.

de Haas, T., Oussoren, E., Grajkowska, W., Perek-Polnik, M., Popovic, M., Zadravec-Zaletel, L., Perera, M., Corte, G., Wirths, O., van Sluis, P., Pietsch, T., Troost, D., Baas, F., Versteeg, R. and Kool, M. (2006) 'OTX1 and OTX2 expression correlates with the clinicopathologic classification of medulloblastomas', *J Neuropathol Exp Neurol*, 65(2), pp. 176-86.

De Smet, C. and Loriot, A. (2010) 'DNA hypomethylation in cancer: Epigenetic scars of a neoplastic journey', *Epigenetics*, 5(3).

De Smet, C., Lurquin, C., Lethe, B., Martelange, V. and Boon, T. (1999) 'DNA methylation is the primary silencing mechanism for a set of germ line- and tumor-specific genes with a CpG-rich promoter', *Mol Cell Biol*, 19(11), pp. 7327-35.

Decock, A., Ongenaert, M., Hoebeeck, J., De Preter, K., Van Peer, G., Van Criekinge, W., Ladenstein, R., Schulte, J.H., Noguera, R., Stallings, R.L., Van Damme, A., Laureys, G., Vermeulen, J., Van Maerken, T., Speleman, F. and Vandesompele, J. (2012) 'Genome-wide promoter methylation analysis in neuroblastoma identifies prognostic methylation biomarkers', *Genome Biol*, 13(10), p. R95.

Diede, S.J., Guenthoer, J., Geng, L.N., Mahoney, S.E., Marotta, M., Olson, J.M., Tanaka, H. and Tapscott, S.J. (2010) 'DNA methylation of developmental genes in pediatric medulloblastomas identified by denaturation analysis of methylation differences', *Proc Natl Acad Sci U S A*, 107(1), pp. 234-9.

Dolecek, T.A., Propp, J.M., Stroup, N.E. and Kruchko, C. (2012) 'CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2005-2009', *Neuro Oncol*, 14 Suppl 5, pp. v1-49.

Down, T.A., Rakyan, V.K., Turner, D.J., Flicek, P., Li, H., Kulesha, E., Graf, S., Johnson, N., Herrero, J., Tomazou, E.M., Thorne, N.P., Backdahl, L., Herberth, M., Howe, K.L., Jackson, D.K., Miretti, M.M., Marioni, J.C., Birney, E., Hubbard, T.J., Durbin, R., Tavare, S. and Beck, S. (2008) 'A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis', *Nat Biotechnol*, 26(7), pp. 779-85.

Druker, B.J., Talpaz, M., Resta, D.J., Peng, B., Buchdunger, E., Ford, J.M., Lydon, N.B., Kantarjian, H., Capdeville, R., Ohno-Jones, S. and Sawyers, C.L. (2001) 'Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia', *N Engl J Med*, 344(14), pp. 1031-7.

Du, P., Zhang, X., Huang, C.C., Jafari, N., Kibbe, W.A., Hou, L. and Lin, S.M. (2010) 'Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis', *BMC Bioinformatics*, 11, p. 587.

Dubuc, A.M., Mack, S., Unterberger, A., Northcott, P.A. and Taylor, M.D. (2012) 'The epigenetics of brain tumors', *Methods Mol Biol*, 863, pp. 139-53.

Dubuc, A.M., Remke, M., Korshunov, A., Northcott, P.A., Zhan, S.H., Mendez-Lago, M., Kool, M., Jones, D.T., Unterberger, A., Morrissy, A.S., Shih, D., Peacock, J., Ramaswamy, V., Rolider, A., Wang, X., Witt, H., Hielscher, T., Hawkins, C., Vibhakar, R., Croul, S., Rutka, J.T., Weiss, W.A., Jones, S.J., Eberhart, C.G., Marra, M.A., Pfister, S.M. and Taylor, M.D. (2013) 'Aberrant patterns of H3K4 and H3K27 histone lysine methylation occur across subgroups in medulloblastoma', *Acta Neuropathol*, 125(3), pp. 373-84.

Dunning, M.J., Smith, M.L., Ritchie, M.E. and Tavare, S. (2007) 'beadarray: R classes and methods for Illumina bead-based data', *Bioinformatics*, 23(16), pp. 2183-4.

Eberhart, C.G. and Burger, P.C. (2003) 'Anaplasia and grading in medulloblastomas', *Brain Pathol*, 13(3), pp. 376-85.

Eberhart, C.G., Kepner, J.L., Goldthwaite, P.T., Kun, L.E., Duffner, P.K., Friedman, H.S., Strother, D.R. and Burger, P.C. (2002a) 'Histopathologic grading of medulloblastomas: a Pediatric Oncology Group study', *Cancer*, 94(2), pp. 552-60.

Eberhart, C.G., Kratz, J.E., Schuster, A., Goldthwaite, P., Cohen, K.J., Perlman, E.J. and Burger, P.C. (2002b) 'Comparative genomic hybridization detects an increased number of chromosomal alterations in large cell/anaplastic medulloblastomas', *Brain Pathol*, 12(1), pp. 36-44.

Ellison, D. (2002) 'Classifying the medulloblastoma: insights from morphology and molecular genetics', *Neuropathol Appl Neurobiol*, 28(4), pp. 257-82.

Ellison, D.W. (2010) 'Childhood medulloblastoma: novel approaches to the classification of a heterogeneous disease', *Acta Neuropathol*, 120(3), pp. 305-16.

Ellison, D.W., Clifford, S.C., Gajjar, A. and Gilbertson, R.J. (2003) 'What's new in neuro-oncology? Recent advances in medulloblastoma', *Eur J Paediatr Neurol*, 7(2), pp. 53-66.

Ellison, D.W., Dalton, J., Kocak, M., Nicholson, S.L., Fraga, C., Neale, G., Kenney, A.M., Brat, D.J., Perry, A., Yong, W.H., Taylor, R.E., Bailey, S., Clifford, S.C. and Gilbertson, R.J. (2011a) 'Medulloblastoma: clinicopathological correlates of SHH, WNT, and non-SHH/WNT molecular subgroups', *Acta Neuropathol*, 121(3), pp. 381-96.

Ellison, D.W., Kocak, M., Dalton, J., Megahed, H., Lusher, M.E., Ryan, S.L., Zhao, W., Nicholson, S.L., Taylor, R.E., Bailey, S. and Clifford, S.C. (2011b) 'Definition of disease-risk stratification groups in childhood medulloblastoma using combined clinical, pathologic, and molecular variables', *J Clin Oncol*, 29(11), pp. 1400-7.

Ellison, D.W., Onilude, O.E., Lindsey, J.C., Lusher, M.E., Weston, C.L., Taylor, R.E., Pearson, A.D. and Clifford, S.C. (2005) '{beta}-Catenin Status Predicts a Favorable Outcome in Childhood Medulloblastoma: The United Kingdom Children's Cancer Study Group Brain Tumour Committee', *J Clin Oncol*, 23(31), pp. 7951-7957.

Ernst, J., Kheradpour, P., Mikkelsen, T.S., Shoresh, N., Ward, L.D., Epstein, C.B., Zhang, X., Wang, L., Issner, R., Coyne, M., Ku, M., Durham, T., Kellis, M. and Bernstein, B.E. (2011) 'Mapping and analysis of chromatin state dynamics in nine human cell types', *Nature*, 473(7345), pp. 43-9.

Ernst, T., Chase, A.J., Score, J., Hidalgo-Curtis, C.E., Bryant, C., Jones, A.V., Waghorn, K., Zoi, K., Ross, F.M., Reiter, A., Hochhaus, A., Drexler, H.G., Duncombe, A., Cervantes, F., Oscier, D., Boultwood, J., Grand, F.H. and Cross, N.C. (2010) 'Inactivating mutations of the histone methyltransferase gene EZH2 in myeloid disorders', *Nat Genet*, 42(8), pp. 722-6.

Esteller, M. (2002) 'CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future', *Oncogene*, 21(35), pp. 5427-40.

Esteller, M. (2005) 'Aberrant DNA methylation as a cancer-inducing mechanism', *Annu Rev Pharmacol Toxicol*, 45, pp. 629-56.

Esteller, M. (2008) 'Epigenetics in cancer', *N Engl J Med*, 358(11), pp. 1148-59.

Esteller, M., Garcia-Foncillas, J., Andion, E., Goodman, S.N., Hidalgo, O.F., Vanaclocha, V., Baylin, S.B. and Herman, J.G. (2000a) 'Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents', *N Engl J Med*, 343(19), pp. 1350-4.

Esteller, M., Risques, R.A., Toyota, M., Capella, G., Moreno, V., Peinado, M.A., Baylin, S.B. and Herman, J.G. (2001) 'Promoter hypermethylation of the DNA repair gene O(6)-methylguanine-DNA methyltransferase is associated with the presence of G:C to A:T transition mutations in p53 in human colorectal tumorigenesis', *Cancer Res*, 61(12), pp. 4689-92.

Esteller, M., Toyota, M., Sanchez-Cespedes, M., Capella, G., Peinado, M.A., Watkins, D.N., Issa, J.P., Sidransky, D., Baylin, S.B. and Herman, J.G. (2000b) 'Inactivation of the DNA repair gene O6-methylguanine-DNA methyltransferase by promoter hypermethylation is associated with G to A mutations in K-ras in colorectal tumorigenesis', *Cancer Res*, 60(9), pp. 2368-71.

Etcheverry, A., Aubry, M., de Tayrac, M., Vauleon, E., Boniface, R., Guenot, F., Saikali, S., Hamlat, A., Riffaud, L., Menei, P., Quillien, V. and Mosser, J. (2010) 'DNA methylation in glioblastoma: impact on gene expression and clinical outcome', *BMC Genomics*, 11, p. 701.

Fabbri, M., Garzon, R., Cimmino, A., Liu, Z., Zanesi, N., Callegari, E., Liu, S., Alder, H., Costinean, S., Fernandez-Cymering, C., Volinia, S., Guler, G., Morrison, C.D., Chan, K.K., Marcucci, G., Calin, G.A., Huebner, K. and Croce, C.M. (2007) 'MicroRNA-29 family reverts aberrant methylation in lung cancer by targeting DNA methyltransferases 3A and 3B', *Proc Natl Acad Sci U S A*, 104(40), pp. 15805-10.

Fan, J.Y., Gordon, F., Luger, K., Hansen, J.C. and Tremethick, D.J. (2002) 'The essential histone variant H2A.Z regulates the equilibrium between different chromatin conformational states', *Nat Struct Biol*, 9(3), pp. 172-6.

Fan, X. and Eberhart, C.G. (2008) 'Medulloblastoma stem cells', *J Clin Oncol*, 26(17), pp. 2821-7.

Fattet, S., Haberler, C., Legoix, P., Varlet, P., Lellouch-Tubiana, A., Lair, S., Manie, E., Raquin, M.A., Bours, D., Carpentier, S., Barillot, E., Grill, J., Doz, F., Puget, S., Janoueix-Lerosey, I. and Delattre, O. (2009) 'Beta-catenin status in paediatric medulloblastomas: correlation of immunohistochemical expression with mutational status, genetic profiles, and clinical characteristics', *J Pathol*, 218(1), pp. 86-94.

Feinberg, A.P. and Tycko, B. (2004) 'The history of cancer epigenetics', *Nat Rev Cancer*, 4(2), pp. 143-53.

Feinberg, A.P. and Vogelstein, B. (1983a) 'Hypomethylation distinguishes genes of some human cancers from their normal counterparts', *Nature*, 301(5895), pp. 89-92.

Feinberg, A.P. and Vogelstein, B. (1983b) 'Hypomethylation of ras oncogenes in primary human cancers', *Biochem Biophys Res Commun*, 111(1), pp. 47-54.

Fenaux, P., Mufti, G.J., Hellstrom-Lindberg, E., Santini, V., Finelli, C., Giagounidis, A., Schoch, R., Gattermann, N., Sanz, G., List, A., Gore, S.D., Seymour, J.F., Bennett, J.M., Byrd, J., Backstrom, J., Zimmerman, L., McKenzie, D., Beach, C. and Silverman, L.R. (2009) 'Efficacy of azacitidine compared with that of conventional care regimens in the treatment of higher-risk myelodysplastic syndromes: a randomised, open-label, phase III study', *Lancet Oncol*, 10(3), pp. 223-32.

Feng, Q. and Zhang, Y. (2001) 'The MeCP1 complex represses transcription through preferential binding, remodeling, and deacetylating methylated nucleosomes', *Genes Dev*, 15(7), pp. 827-32.

Fernandez, A.F., Assenov, Y., Martin-Subero, J.I., Balint, B., Siebert, R., Taniguchi, H., Yamamoto, H., Hidalgo, M., Tan, A.C., Galm, O., Ferrer, I., Sanchez-Cespedes, M., Villanueva, A., Carmona, J., Sanchez-Mut, J.V., Berdasco, M., Moreno, V., Capella, G., Monk, D., Ballestar, E., Ropero, S., Martinez, R., Sanchez-Carbayo, M., Prosper, F., Agirre, X., Fraga, M.F., Grana, O., Perez-Jurado, L., Mora, J., Puig, S., Prat, J., Badimon, L., Puca, A.A., Meltzer, S.J., Lengauer, T., Bridgewater, J., Bock, C. and Esteller, M. (2012) 'A DNA methylation fingerprint of 1628 human samples', *Genome Res*, 22(2), pp. 407-19.

Figueroa, M.E., Abdel-Wahab, O., Lu, C., Ward, P.S., Patel, J., Shih, A., Li, Y., Bhagwat, N., Vasanthakumar, A., Fernandez, H.F., Tallman, M.S., Sun, Z., Wolniak, K., Peeters, J.K., Liu, W., Choe, S.E., Fantin, V.R., Paietta, E., Lowenberg, B., Licht, J.D., Godley, L.A., Delwel, R., Valk, P.J., Thompson, C.B., Levine, R.L. and Melnick, A. (2010) 'Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation', *Cancer Cell*, 18(6), pp. 553-67.

Filippakopoulos, P. and Knapp, S. (2012) 'The bromodomain interaction module', *FEBS Lett*, 586(17), pp. 2692-704.

Fouladi, M., Gilger, E., Kocak, M., Wallace, D., Buchanan, G., Reeves, C., Robbins, N., Merchant, T., Kun, L.E., Khan, R., Gajjar, A. and Mulhern, R. (2005) 'Intellectual and functional outcome of children 3 years old or younger who have CNS malignancies', *J Clin Oncol*, 23(28), pp. 7152-60.

Friedman, H.S., Burger, P.C., Bigner, S.H., Trojanowski, J.Q., Brodeur, G.M., He, X.M., Wikstrand, C.J., Kurtzberg, J., Berens, M.E., Halperin, E.C. and et al. (1988) 'Phenotypic and genotypic analysis of a human medulloblastoma cell line and transplantable xenograft (D341 Med) demonstrating amplification of c-myc', *Am J Pathol*, 130(3), pp. 472-84.

Friedman, H.S., Burger, P.C., Bigner, S.H., Trojanowski, J.Q., Wikstrand, C.J., Halperin, E.C. and Bigner, D.D. (1985) 'Establishment and characterization of the human medulloblastoma cell line and transplantable xenograft D283 Med', *J Neuropathol Exp Neurol*, 44(6), pp. 592-605.

Frommer, M., McDonald, L.E., Millar, D.S., Collis, C.M., Watt, F., Grigg, G.W., Molloy, P.L. and Paul, C.L. (1992) 'A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands', *Proc Natl Acad Sci U S A*, 89(5), pp. 1827-31.

Fruhwald, M.C., O'Dorisio, M.S., Dai, Z., Tanner, S.M., Balster, D.A., Gao, X., Wright, F.A. and Plass, C. (2001) 'Aberrant promoter methylation of previously unidentified target genes is a common abnormality in medulloblastomas--implications for tumor biology and potential clinical utility', *Oncogene*, 20(36), pp. 5033-42.

Fuks, F., Burgers, W.A., Brehm, A., Hughes-Davies, L. and Kouzarides, T. (2000) 'DNA methyltransferase Dnmt1 associates with histone deacetylase activity', *Nat Genet*, 24(1), pp. 88-91.

Fults, D., Pedone, C., Dai, C. and Holland, E.C. (2002) 'MYC expression promotes the proliferation of neural progenitor cells in culture and in vivo', *Neoplasia*, 4(1), pp. 32-9.

Futscher, B.W., Oshiro, M.M., Wozniak, R.J., Holtan, N., Hanigan, C.L., Duan, H. and Domann, F.E. (2002) 'Role for DNA methylation in the control of cell type specific maspin expression', *Nat Genet*, 31(2), pp. 175-9.

Gajjar, A., Chintagumpala, M., Ashley, D., Kellie, S., Kun, L.E., Merchant, T.E., Woo, S., Wheeler, G., Ahern, V., Krasin, M.J., Fouladi, M., Broniscer, A., Krance, R., Hale, G.A., Stewart, C.F., Dauser, R., Sanford, R.A., Fuller, C., Lau, C., Boyett, J.M., Wallace, D. and Gilbertson, R.J. (2006) 'Risk-adapted craniospinal radiotherapy followed by high-dose chemotherapy and stem-cell rescue in children with newly

diagnosed medulloblastoma (St Jude Medulloblastoma-96): long-term results from a prospective, multicentre trial', *Lancet Oncol*, 7(10), pp. 813-20.

Gajjar, A., Hernan, R., Kocak, M., Fuller, C., Lee, Y., McKinnon, P.J., Wallace, D., Lau, C., Chintagumpala, M., Ashley, D.M., Kellie, S.J., Kun, L. and Gilbertson, R.J. (2004) 'Clinical, histopathologic, and molecular markers of prognosis: toward a new disease risk stratification system for medulloblastoma', *J Clin Oncol*, 22(6), pp. 984-93.

Gajjar, A., Packer, R.J., Foreman, N.K., Cohen, K., Haas-Kogan, D. and Merchant, T.E. (2013) 'Children's Oncology Group's 2013 blueprint for research: central nervous system tumors', *Pediatr Blood Cancer*, 60(6), pp. 1022-6.

Gandola, L., Massimino, M., Cefalo, G., Solero, C., Spreafico, F., Pecori, E., Riva, D., Collini, P., Pignoli, E., Giangaspero, F., Luksch, R., Berretta, S., Poggi, G., Biassoni, V., Ferrari, A., Pollo, B., Favre, C., Sardi, I., Terenziani, M. and Fossati-Bellani, F. (2009) 'Hyperfractionated accelerated radiotherapy in the Milan strategy for metastatic medulloblastoma', *J Clin Oncol*, 27(4), pp. 566-71.

Gardiner-Garden, M. and Frommer, M. (1987) 'CpG islands in vertebrate genomes', *J Mol Biol*, 196(2), pp. 261-82.

Gazdar, A.F., Girard, L., Lockwood, W.W., Lam, W.L. and Minna, J.D. (2010) 'Lung cancer cell lines as tools for biomedical discovery and research', *J Natl Cancer Inst*, 102(17), pp. 1310-21.

Gherardi, E., Birchmeier, W., Birchmeier, C. and Vande Woude, G. (2012) 'Targeting MET in cancer: rationale and progress', *Nat Rev Cancer*, 12(2), pp. 89-103.

Giangaspero, F., Perilongo, G., Fondelli, M.P., Brisigotti, M., Carollo, C., Burnelli, R., Burger, P.C. and Garre, M.L. (1999) 'Medulloblastoma with extensive nodularity: a variant with favorable prognosis', *J Neurosurg*, 91(6), pp. 971-7.

Giangaspero, F., Rigobello, L., Badiali, M., Loda, M., Andreini, L., Basso, G., Zorzi, F. and Montaldi, A. (1992) 'Large-cell medulloblastomas. A distinct variant with highly aggressive behavior', *Am J Surg Pathol*, 16(7), pp. 687-93.

Gibson, P., Tong, Y., Robinson, G., Thompson, M.C., Currle, D.S., Eden, C., Kranenburg, T.A., Hogg, T., Poppleton, H., Martin, J., Finkelstein, D., Pounds, S., Weiss, A., Patay, Z., Scoggins, M., Ogg, R., Pei, Y., Yang, Z.J., Brun, S., Lee, Y., Zindy, F., Lindsey, J.C., Taketo, M.M., Boop, F.A., Sanford, R.A., Gajjar, A., Clifford, S.C., Roussel, M.F., McKinnon, P.J., Gutmann, D.H., Ellison, D.W., Wechsler-Reya, R. and Gilbertson, R.J. (2010) 'Subtypes of medulloblastoma have distinct developmental origins', *Nature*, 468(7327), pp. 1095-9.

Gilbertson, R.J. and Ellison, D.W. (2008) 'The origins of medulloblastoma subtypes', *Annu Rev Pathol*, 3, pp. 341-65.

Gilthorpe, J.D., Papantoniou, E.K., Chedotal, A., Lumsden, A. and Wingate, R.J. (2002) 'The migration of cerebellar rhombic lip derivatives', *Development*, 129(20), pp. 4719-28.

Gitan, R.S., Shi, H., Chen, C.M., Yan, P.S. and Huang, T.H. (2002) 'Methylation-specific oligonucleotide microarray: a new potential for high-throughput methylation analysis', *Genome Res*, 12(1), pp. 158-64.

Glockner, S.C., Dhir, M., Yi, J.M., McGarvey, K.E., Van Neste, L., Louwagie, J., Chan, T.A., Kleeberger, W., de Bruine, A.P., Smits, K.M., Khalid-de Bakker, C.A., Jonkers, D.M., Stockbrugger, R.W., Meijer, G.A., Oort, F.A., Iacobuzio-Donahue, C., Bierau, K., Herman, J.G., Baylin, S.B., Van Engeland, M., Schuebel, K.E. and Ahuja, N. (2009) 'Methylation of TFPI2 in stool DNA: a potential novel biomarker for the detection of colorectal cancer', *Cancer Res*, 69(11), pp. 4691-9.

Goldberg, A.D., Allis, C.D. and Bernstein, E. (2007) 'Epigenetics: a landscape takes shape', *Cell*, 128(4), pp. 635-8.

Goodrich, L.V., Milenkovic, L., Higgins, K.M. and Scott, M.P. (1997) 'Altered neural cell fates and medulloblastoma in mouse patched mutants', *Science*, 277(5329), pp. 1109-13.

Grandori, C., Cowley, S.M., James, L.P. and Eisenman, R.N. (2000) 'The Myc/Max/Mad network and the transcriptional control of cell behavior', *Annu Rev Cell Dev Biol*, 16, pp. 653-99.

Grundy, R.G., Wilne, S.H., Robinson, K.J., Ironside, J.W., Cox, T., Chong, W.K., Michalski, A., Campbell, R.H., Bailey, C.C., Thorp, N., Pizer, B., Punt, J., Walker, D.A., Ellison, D.W. and Machin, D. (2010) 'Primary postoperative chemotherapy without radiotherapy for treatment of brain tumours other than ependymoma in children under 3 years: results of the first UKCCSG/SIOP CNS 9204 trial', *Eur J Cancer*, 46(1), pp. 120-33.

Gulino, A., Arcella, A. and Giangaspero, F. (2008) 'Pathological and molecular heterogeneity of medulloblastoma', *Curr Opin Oncol*, 20(6), pp. 668-75.

Gunderson, K.L., Kruglyak, S., Graige, M.S., Garcia, F., Kermani, B.G., Zhao, C., Che, D., Dickinson, T., Wickham, E., Bierle, J., Doucet, D., Milewski, M., Yang, R., Siegmund, C., Haas, J., Zhou, L., Oliphant, A., Fan, J.B., Barnard, S. and Chee, M.S. (2004) 'Decoding randomly ordered DNA arrays', *Genome Res*, 14(5), pp. 870-7.

Hahn, H., Wicking, C., Zaphiropoulous, P.G., Gailani, M.R., Shanley, S., Chidambaram, A., Vorechovsky, I., Holmberg, E., Unden, A.B., Gillies, S., Negus, K., Smyth, I., Pressman, C., Leffell, D.J., Gerrard, B., Goldstein, A.M., Dean, M., Toftgard, R., Chenevix-Trench, G., Wainwright, B. and Bale, A.E. (1996) 'Mutations of the human homolog of Drosophila patched in the nevoid basal cell carcinoma syndrome', *Cell*, 85(6), pp. 841-51.

Hallahan, A.R., Pritchard, J.I., Hansen, S., Benson, M., Stoeck, J., Hatton, B.A., Russell, T.L., Ellenbogen, R.G., Bernstein, I.D., Beachy, P.A. and Olson, J.M. (2004) 'The SmoA1 mouse model reveals that notch signaling is critical for the growth and survival of sonic hedgehog-induced medulloblastomas', *Cancer Res*, 64(21), pp. 7794-800.

Hamilton, S.R., Liu, B., Parsons, R.E., Papadopoulos, N., Jen, J., Powell, S.M., Krush, A.J., Berk, T., Cohen, Z., Tetu, B. and et al. (1995) 'The molecular basis of Turcot's syndrome', *N Engl J Med*, 332(13), pp. 839-47.

Han, H., Cortez, C.C., Yang, X., Nichols, P.W., Jones, P.A. and Liang, G. (2011) 'DNA methylation directly silences genes with non-CpG island promoters and establishes a nucleosome occupied promoter', *Hum Mol Genet*, 20(22), pp. 4299-310.

Hanada, M., Delia, D., Aiello, A., Stadtmauer, E. and Reed, J.C. (1993) 'bcl-2 gene hypomethylation and high-level expression in B-cell chronic lymphocytic leukemia', *Blood*, 82(6), pp. 1820-8.

Hanahan, D. and Weinberg, R.A. (2000) 'The hallmarks of cancer', *Cell*, 100(1), pp. 57-70.

Hanahan, D. and Weinberg, R.A. (2011) 'Hallmarks of cancer: the next generation', *Cell*, 144(5), pp. 646-74.

Hatten, M.E. and Roussel, M.F. (2011) 'Development and cancer of the cerebellum', *Trends Neurosci*, 34(3), pp. 134-42.

Hatton, B.A., Villavicencio, E.H., Tsuchiya, K.D., Pritchard, J.I., Ditzler, S., Pullar, B., Hansen, S., Knoblaugh, S.E., Lee, D., Eberhart, C.G., Hallahan, A.R. and Olson, J.M. (2008) 'The Smo/Smo model: hedgehog-induced medulloblastoma with 90% incidence and leptomeningeal spread', *Cancer Res*, 68(6), pp. 1768-76.

Hayashizaki, Y., Hirotsune, S., Okazaki, Y., Hatada, I., Shibata, H., Kawai, J., Hirose, K., Watanabe, S., Fushiki, S., Wada, S. and et al. (1993) 'Restriction landmark genomic scanning method and its various applications', *Electrophoresis*, 14(4), pp. 251-8.

He, X.M., Wikstrand, C.J., Friedman, H.S., Bigner, S.H., Pleasure, S., Trojanowski, J.Q. and Bigner, D.D. (1991) 'Differentiation characteristics of newly established medulloblastoma cell lines (D384 Med, D425 Med, and D458 Med) and their transplantable xenografts', *Lab Invest*, 64(6), pp. 833-43.

Hendrich, B. and Bird, A. (1998) 'Identification and characterization of a family of mammalian methyl-CpG binding proteins', *Mol Cell Biol*, 18(11), pp. 6538-47.

Hernan, R., Fasheh, R., Calabrese, C., Frank, A.J., Maclean, K.H., Allard, D., Barraclough, R. and Gilbertson, R.J. (2003) 'ERBB2 up-regulates S100A4 and several other prometastatic genes in medulloblastoma', *Cancer Res*, 63(1), pp. 140-8.

Hernandez-Vargas, H., Lambert, M.P., Le Calvez-Kelm, F., Gouysse, G., McKay-Chopin, S., Tavtigian, S.V., Scoazec, J.Y. and Herceg, Z. (2010) 'Hepatocellular carcinoma displays distinct DNA methylation signatures with potential as clinical predictors', *PLoS One*, 5(3), p. e9749.

Hoivik, E.A., Bjanesoy, T.E., Mai, O., Okamoto, S., Minokoshi, Y., Shima, Y., Morohashi, K., Boehm, U. and Bakke, M. (2011) 'DNA methylation of intronic enhancers directs tissue-specific expression of steroidogenic factor 1/adrenal 4 binding protein (SF-1/Ad4BP)', *Endocrinology*, 152(5), pp. 2100-12.

Horikoshi, T., Danenberg, K.D., Stadlbauer, T.H., Volkenandt, M., Shea, L.C., Aigner, K., Gustavsson, B., Leichman, L., Frosing, R., Ray, M. and et al. (1992) 'Quantitation of thymidylate synthase, dihydrofolate reductase, and DT-diaphorase gene expression in human tumors using the polymerase chain reaction', *Cancer Res*, 52(1), pp. 108-16.

Hovestadt, V., Remke, M., Kool, M., Pietsch, T., Northcott, P.A., Fischer, R., Cavalli, F.M., Ramaswamy, V., Zapatka, M., Reifenberger, G., Rutkowski, S., Schick, M., Bewerunge-Hudler, M., Korshunov, A., Lichter, P., Taylor, M.D., Pfister, S.M. and Jones, D.T. (2013) 'Robust molecular subgrouping and copy-number profiling of medulloblastoma from small amounts of archival tumour material using high-density DNA methylation arrays', *Acta Neuropathol*, 125(6), pp. 913-6.

Huang da, W., Sherman, B.T., Tan, Q., Kir, J., Liu, D., Bryant, D., Guo, Y., Stephens, R., Baseler, M.W., Lane, H.C. and Lempicki, R.A. (2007) 'DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists', *Nucleic Acids Res*, 35(Web Server issue), pp. W169-75.

Huang, H.C., Zheng, S., VanBuren, V. and Zhao, Z. (2010) 'Discovering disease-specific biomarker genes for cancer diagnosis and prognosis', *Technol Cancer Res Treat*, 9(3), pp. 219-30.

Huang, S.M., Mishina, Y.M., Liu, S., Cheung, A., Stegmeier, F., Michaud, G.A., Charlat, O., Wiellette, E., Zhang, Y., Wiessner, S., Hild, M., Shi, X., Wilson, C.J., Mickanin, C., Myer, V., Fazal, A., Tomlinson, R., Serluca, F., Shao, W., Cheng, H., Shultz, M., Rau, C., Schirle, M., Schlegl, J., Ghidelli, S., Fawell, S., Lu, C., Curtis, D., Kirschner, M.W., Lengauer, C., Finan, P.M., Tallarico, J.A., Bouwmeester, T., Porter, J.A., Bauer, A. and Cong, F. (2009) 'Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling', *Nature*, 461(7264), pp. 614-20.

Huangfu, D. and Anderson, K.V. (2005) 'Cilia and Hedgehog responsiveness in the mouse', *Proc Natl Acad Sci U S A*, 102(32), pp. 11325-30.

Ieraci, A., Forni, P.E. and Ponzetto, C. (2002) 'Viable hypomorphic signaling mutant of the Met receptor reveals a role for hepatocyte growth factor in postnatal cerebellar development', *Proc Natl Acad Sci U S A*, 99(23), pp. 15200-5.

Illumina (2011) *Array-based gene expression analysis datasheet*. Available at: http://res.illumina.com/documents/products/datasheets/datasheet_gene_exp_analysis.pdf (Accessed: August 10th, 2013).

Illumina, I. (2007) *GoldenGate Assay for Methylation and BeadArray Tecnology*. Available at: www.illumina.com/Documents/products/technotes/technote_goldengate_assay_methylation.pdf (Accessed: Retrieved June,18,2010).

Imai, K. and Takaoka, A. (2006) 'Comparing antibody and small-molecule therapies for cancer', *Nat Rev Cancer*, 6(9), pp. 714-27.

Irizarry, R. (2005) 'From CEL files to annotated lists of interesting genes.', in R.Gentlemen, V.C., S. Dudoit, R. Irizarry, W. Huber (ed.) *Bioinformatics and Computational Biology Solutions using R and Bioconductor.* New York: Springer, pp. 431-442.

Irizarry, R.A., Bolstad, B.M., Collin, F., Cope, L.M., Hobbs, B. and Speed, T.P. (2003) 'Summaries of Affymetrix GeneChip probe level data', *Nucleic Acids Res*, 31(4), p. e15.

Irizarry, R.A., Ladd-Acosta, C., Carvalho, B., Wu, H., Brandenburg, S.A., Jeddeloh, J.A., Wen, B. and Feinberg, A.P. (2008) 'Comprehensive high-throughput arrays for relative methylation (CHARM)', *Genome Res*, 18(5), pp. 780-90.

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. and Feinberg, A.P. (2009) 'The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores', *Nat Genet*, 41(2), pp. 178-86.

Issa, J.P. and Kantarjian, H.M. (2005) 'Introduction: emerging role of epigenetic therapy: focus on decitabine', *Semin Hematol*, 42(3 Suppl 2), pp. S1-2.

Jacobsen, P.F., Jenkyn, D.J. and Papadimitriou, J.M. (1985) 'Establishment of a human medulloblastoma cell line and its heterotransplantation into nude mice', *J Neuropathol Exp Neurol*, 44(5), pp. 472-85.

Jaenisch, R. and Bird, A. (2003) 'Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals', *Nat Genet*, 33 Suppl, pp. 245-54.

Jaenisch, R. and Young, R. (2008) 'Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming', *Cell*, 132(4), pp. 567-82.

Jeanmougin, M., de Reynies, A., Marisa, L., Paccard, C., Nuel, G. and Guedj, M. (2010) 'Should we abandon the t-test in the analysis of gene expression microarray data: a comparison of variance modeling strategies', *PLoS One*, 5(9), p. e12336.

Jenuwein, T. and Allis, C.D. (2001) 'Translating the histone code', *Science*, 293(5532), pp. 1074-80.

Jiao, Y., Shi, C., Edil, B.H., de Wilde, R.F., Klimstra, D.S., Maitra, A., Schulick, R.D., Tang, L.H., Wolfgang, C.L., Choti, M.A., Velculescu, V.E., Diaz, L.A., Jr., Vogelstein, B., Kinzler, K.W., Hruban, R.H. and Papadopoulos, N. (2011) 'DAXX/ATRX, MEN1, and mTOR pathway genes are frequently altered in pancreatic neuroendocrine tumors', *Science*, 331(6021), pp. 1199-203.

Jones, D.T., Jager, N., Kool, M., Zichner, T., Hutter, B., Sultan, M., Cho, Y.J., Pugh, T.J., Hovestadt, V., Stutz, A.M., Rausch, T., Warnatz, H.J., Ryzhova, M., Bender, S., Sturm, D., Pleier, S., Cin, H., Pfaff, E., Sieber, L., Wittmann, A., Remke, M., Witt, H., Hutter, S., Tzaridis, T., Weischenfeldt, J., Raeder, B., Avci, M., Amstislavskiy, V., Zapatka, M., Weber, U.D., Wang, Q., Lasitschka, B., Bartholomae, C.C., Schmidt, M., von Kalle, C., Ast, V., Lawerenz, C., Eils, J., Kabbe, R., Benes, V., van Sluis, P., Koster, J., Volckmann, R., Shih, D., Betts, M.J., Russell, R.B., Coco, S., Tonini, G.P., Schuller, U., Hans, V., Graf, N., Kim, Y.J., Monoranu, C., Roggendorf, W., Unterberg, A., Herold-Mende, C., Milde, T., Kulozik, A.E., von Deimling, A., Witt, O., Maass, E., Rossler, J., Ebinger, M., Schuhmann, M.U., Fruhwald, M.C., Hasselblatt, M., Jabado, N., Rutkowski, S., von Bueren, A.O., Williamson, D., Clifford, S.C., McCabe, M.G., Collins, V.P., Wolf, S., Wiemann, S., Lehrach, H., Brors, B., Scheurlen, W., Felsberg, J., Reifenberger, G., Northcott, P.A., Taylor, M.D., Meyerson, M., Pomeroy, S.L., Yaspo, M.L., Korbel, J.O., Korshunov, A., Eils, R., Pfister, S.M. and Lichter, P. (2012)

'Dissecting the genomic complexity underlying medulloblastoma', *Nature*, 488(7409), pp. 100-5.

Jones, D.T., Northcott, P.A., Kool, M. and Pfister, S.M. (2013) 'The role of chromatin remodeling in medulloblastoma', *Brain Pathol*, 23(2), pp. 193-9.

Jones, P.A. (2012) 'Functions of DNA methylation: islands, start sites, gene bodies and beyond', *Nat Rev Genet*, 13(7), pp. 484-92.

Jones, P.A. and Baylin, S.B. (2002) 'The fundamental role of epigenetic events in cancer', *Nat Rev Genet*, 3(6), pp. 415-28.

Jones, P.A. and Baylin, S.B. (2007) 'The epigenomics of cancer', *Cell*, 128(4), pp. 683-92.

Jones, P.A. and Laird, P.W. (1999) 'Cancer epigenetics comes of age', *Nat Genet*, 21(2), pp. 163-7.

Jones, P.A. and Taylor, S.M. (1980) 'Cellular differentiation, cytidine analogs and DNA methylation', *Cell*, 20(1), pp. 85-93.

Jones, P.L., Veenstra, G.J., Wade, P.A., Vermaak, D., Kass, S.U., Landsberger, N., Strouboulis, J. and Wolffe, A.P. (1998) 'Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription', *Nat Genet*, 19(2), pp. 187-91.

Jung, S., Kim, S., Gale, M., Cherni, I., Fonseca, R., Carpten, J. and Salhia, B. (2012) 'DNA methylation in multiple myeloma is weakly associated with gene transcription', *PLoS One*, 7(12), p. e52626.

Kalluri, R. and Zeisberg, M. (2006) 'Fibroblasts in cancer', *Nat Rev Cancer*, 6(5), pp. 392-401.

Kantarjian, H., Issa, J.P., Rosenfeld, C.S., Bennett, J.M., Albitar, M., DiPersio, J., Klimek, V., Slack, J., de Castro, C., Ravandi, F., Helmer, R., 3rd, Shen, L., Nimer, S.D., Leavitt, R., Raza, A. and Saba, H. (2006) 'Decitabine improves patient outcomes in myelodysplastic syndromes: results of a phase III randomized study', *Cancer*, 106(8), pp. 1794-803.

Kaplan, E.L. and Meier, P. (1958) 'Nonparametric estimation from incomplete observations', *Journal of the American Statistical Association*, 53(282), pp. 457-481.

Kaput, J. and Sneider, T.W. (1979) 'Methylation of somatic vs germ cell DNAs analyzed by restriction endonuclease digestions', *Nucleic Acids Res*, 7(8), pp. 2303-22.

Kawauchi, D., Robinson, G., Uziel, T., Gibson, P., Rehg, J., Gao, C., Finkelstein, D., Qu, C., Pounds, S., Ellison, D.W., Gilbertson, R.J. and Roussel, M.F. (2012) 'A mouse model of the most aggressive subgroup of human medulloblastoma', *Cancer Cell*, 21(2), pp. 168-80.

Keles, G.E., Berger, M.S., Srinivasan, J., Kolstoe, D.D., Bobola, M.S. and Silber, J.R. (1995) 'Establishment and characterization of four human medulloblastoma-derived cell lines', *Oncol Res*, 7(10-11), pp. 493-503.

Kelly, T.K., Liu, Y., Lay, F.D., Liang, G., Berman, B.P. and Jones, P.A. (2012) 'Genome-wide mapping of nucleosome positioning and DNA methylation within individual DNA molecules', *Genome Res*, 22(12), pp. 2497-506.

Khosravi-Far, R. and Der, C.J. (1994) 'The Ras signal transduction pathway', *Cancer Metastasis Rev*, 13(1), pp. 67-89.

Kibriya, M.G., Raza, M., Jasmine, F., Roy, S., Paul-Brutus, R., Rahaman, R., Dodsworth, C., Rakibuz-Zaman, M., Kamal, M. and Ahsan, H. (2011) 'A genome-wide DNA methylation study in colorectal carcinoma', *BMC Med Genomics*, 4, p. 50.

Kieran, M.W., Walker, D., Frappaz, D. and Prados, M. (2010) 'Brain tumors: from childhood through adolescence into adulthood', *J Clin Oncol*, 28(32), pp. 4783-9.

Kimura, H., Ng, J.M. and Curran, T. (2008) 'Transient inhibition of the Hedgehog pathway in young mice causes permanent defects in bone structure', *Cancer Cell*, 13(3), pp. 249-60.

Kleihues, P., Louis, D.N., Scheithauer, B.W., Rorke, L.B., Reifenberger, G., Burger, P.C. and Cavenee, W.K. (2002) 'The WHO classification of tumors of the nervous system', *J Neuropathol Exp Neurol*, 61(3), pp. 215-25; discussion 226-9.

Klose, R.J. and Bird, A.P. (2006) 'Genomic DNA methylation: the mark and its mediators', *Trends Biochem Sci*, 31(2), pp. 89-97.

Knudson, A.G., Jr. (1971) 'Mutation and cancer: statistical study of retinoblastoma', *Proc Natl Acad Sci U S A*, 68(4), pp. 820-3.

Koch, C.M., Andrews, R.M., Flicek, P., Dillon, S.C., Karaoz, U., Clelland, G.K., Wilcox, S., Beare, D.M., Fowler, J.C., Couttet, P., James, K.D., Lefebvre, G.C., Bruce, A.W., Dovey, O.M., Ellis, P.D., Dhami, P., Langford, C.F., Weng, Z., Birney, E., Carter, N.P., Vetrie, D. and Dunham, I. (2007) 'The landscape of histone modifications across 1% of the human genome in five human cell lines', *Genome Res*, 17(6), pp. 691-707.

Komashko, V.M. and Farnham, P.J. (2010) '5-azacytidine treatment reorganizes genomic histone modification patterns', *Epigenetics*, 5(3).

Kongkham, P.N., Northcott, P.A., Croul, S.E., Smith, C.A., Taylor, M.D. and Rutka, J.T. (2010a) 'The SFRP family of WNT inhibitors function as novel tumor suppressor genes epigenetically silenced in medulloblastoma', *Oncogene*, 29(20), pp. 3017-24.

Kongkham, P.N., Northcott, P.A., Ra, Y.S., Nakahara, Y., Mainprize, T.G., Croul, S.E., Smith, C.A., Taylor, M.D. and Rutka, J.T. (2008) 'An epigenetic genome-wide screen identifies SPINT2 as a novel tumor suppressor gene in pediatric medulloblastoma', *Cancer Res*, 68(23), pp. 9945-53.

Kongkham, P.N., Onvani, S., Smith, C.A. and Rutka, J.T. (2010b) 'Inhibition of the MET Receptor Tyrosine Kinase as a Novel Therapeutic Strategy in Medulloblastoma', *Transl Oncol*, 3(6), pp. 336-43.

Konishi, T., Konishi, F., Takasaki, S., Inoue, K., Nakayama, K. and Konagaya, A. (2008) 'Coincidence between transcriptome analyses on different microarray platforms using a parametric framework', *PLoS One*, 3(10), p. e3555.

Kool, M., Korshunov, A., Remke, M., Jones, D.T., Schlanstein, M., Northcott, P.A., Cho, Y.J., Koster, J., Schouten-van Meeteren, A., van Vuurden, D., Clifford, S.C., Pietsch, T., von Bueren, A.O., Rutkowski, S., McCabe, M., Collins, V.P., Backlund, M.L., Haberler, C., Bourdeaut, F., Delattre, O., Doz, F., Ellison, D.W., Gilbertson, R.J., Pomeroy, S.L., Taylor, M.D., Lichter, P. and Pfister, S.M. (2012) 'Molecular subgroups of medulloblastoma: an international meta-analysis of transcriptome, genetic aberrations, and clinical data of WNT, SHH, Group 3, and Group 4 medulloblastomas', *Acta Neuropathol*, 123(4), pp. 473-84.

Kool, M., Koster, J., Bunt, J., Hasselt, N.E., Lakeman, A., van Sluis, P., Troost, D., Meeteren, N.S., Caron, H.N., Cloos, J., Mrsic, A., Ylstra, B., Grajkowska, W., Hartmann, W., Pietsch, T., Ellison, D., Clifford, S.C. and Versteeg, R. (2008) 'Integrated genomics identifies five medulloblastoma subtypes with distinct genetic profiles, pathway signatures and clinicopathological features', *PLoS One*, 3(8), p. e3088.

Korshunov, A., Benner, A., Remke, M., Lichter, P., von Deimling, A. and Pfister, S. (2008) 'Accumulation of genomic aberrations during clinical progression of medulloblastoma', *Acta Neuropathol*, 116(4), pp. 383-90.

Korshunova, Y., Maloney, R.K., Lakey, N., Citek, R.W., Bacher, B., Budiman, A., Ordway, J.M., McCombie, W.R., Leon, J., Jeddeloh, J.A. and McPherson, J.D. (2008) 'Massively parallel bisulphite pyrosequencing reveals the molecular complexity of breast cancer-associated cytosine-methylation patterns obtained from tissue and serum DNA', *Genome Res*, 18(1), pp. 19-29.

Kouzarides, T. (2007) 'Chromatin modifications and their function', *Cell*, 128(4), pp. 693-705.

Krejsa, C., Rogge, M. and Sadee, W. (2006) 'Protein therapeutics: new applications for pharmacogenetics', *Nat Rev Drug Discov*, 5(6), pp. 507-21.

Kriaucionis, S. and Heintz, N. (2009) 'The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain', *Science*, 324(5929), pp. 929-30.

Kulis, M., Heath, S., Bibikova, M., Queiros, A.C., Navarro, A., Clot, G., Martinez-Trillos, A., Castellano, G., Brun-Heath, I., Pinyol, M., Barberan-Soler, S., Papasaikas, P., Jares, P., Bea, S., Rico, D., Ecker, S., Rubio, M., Royo, R., Ho, V., Klotzle, B., Hernandez, L., Conde, L., Lopez-Guerra, M., Colomer, D., Villamor, N., Aymerich, M., Rozman, M., Bayes, M., Gut, M., Gelpi, J.L., Orozco, M., Fan, J.B., Quesada, V., Puente, X.S., Pisano, D.G., Valencia, A., Lopez-Guillermo, A., Gut, I., Lopez-Otin, C., Campo, E. and Martin-Subero, J.I. (2012) 'Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia', *Nat Genet*, 44(11), pp. 1236-42.

Kulis, M., Queiros, A.C., Beekman, R. and Martin-Subero, J.I. (2013) 'Intragenic DNA methylation in transcriptional regulation, normal differentiation and cancer', *Biochim Biophys Acta*.

Kwabi-Addo, B., Giri, D., Schmidt, K., Podsypanina, K., Parsons, R., Greenberg, N. and Ittmann, M. (2001) 'Haploinsufficiency of the Pten tumor suppressor gene promotes prostate cancer progression', *Proc Natl Acad Sci U S A*, 98(20), pp. 11563-8.

Lafay-Cousin, L., Bouffet, E., Hawkins, C., Amid, A., Huang, A. and Mabbott, D.J. (2009) 'Impact of radiation avoidance on survival and neurocognitive outcome in infant medulloblastoma', *Curr Oncol*, 16(6), pp. 21-8.

Laird, P.W. (2003) 'The power and the promise of DNA methylation markers', *Nat Rev Cancer*, 3(4), pp. 253-66.

Laird, P.W. (2010) 'Principles and challenges of genomewide DNA methylation analysis', *Nat Rev Genet*, 11(3), pp. 191-203.

Lamont, J.M., McManamy, C.S., Pearson, A.D., Clifford, S.C. and Ellison, D.W. (2004) 'Combined histopathological and molecular cytogenetic stratification of medulloblastoma patients', *Clin Cancer Res*, 10(16), pp. 5482-93.

Landberg, T.G., Lindgren, M.L., Cavallin-Stahl, E.K., Svahn-Tapper, G.O., Sundbarg, G., Garwicz, S., Lagergren, J.A., Gunnesson, V.L., Brun, A.E. and Cronqvist, S.E. (1980) 'Improvements in the radiotherapy of medulloblastoma, 1946-1975', *Cancer*, 45(4), pp. 670-8.

Langdon, J.A., Lamont, J.M., Scott, D.K., Dyer, S., Prebble, E., Bown, N., Grundy, R.G., Ellison, D.W. and Clifford, S.C. (2006) 'Combined genome-wide allelotyping and copy number analysis identify frequent genetic losses without copy number reduction in medulloblastoma', *Genes Chromosomes Cancer*, 45(1), pp. 47-60.

Lee, Y., Kawagoe, R., Sasai, K., Li, Y., Russell, H.R., Curran, T. and McKinnon, P.J. (2007) 'Loss of suppressor-of-fused function promotes tumorigenesis', *Oncogene*, 26(44), pp. 6442-7.

Lemon, B. and Tjian, R. (2000) 'Orchestrated response: a symphony of transcription factors for gene control', *Genes Dev*, 14(20), pp. 2551-69.

Leong, H.S., Yates, T., Wilson, C. and Miller, C.J. (2005) 'ADAPT: a database of affymetrix probesets and transcripts', *Bioinformatics*, 21(10), pp. 2552-3.

Lewis, J.D., Meehan, R.R., Henzel, W.J., Maurer-Fogy, I., Jeppesen, P., Klein, F. and Bird, A. (1992) 'Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA', *Cell*, 69(6), pp. 905-14.

Li, B., Carey, M. and Workman, J.L. (2007) 'The role of chromatin during transcription', *Cell*, 128(4), pp. 707-19.

Li, E., Beard, C. and Jaenisch, R. (1993) 'Role for DNA methylation in genomic imprinting', *Nature*, 366(6453), pp. 362-5.

Li, F.P. and Fraumeni, J.F., Jr. (1969) 'Soft-tissue sarcomas, breast cancer, and other neoplasms. A familial syndrome?', *Ann Intern Med*, 71(4), pp. 747-52.

Li, X., Bu, X., Lu, B., Avraham, H., Flavell, R.A. and Lim, B. (2002) 'The hematopoiesis-specific GTP-binding protein RhoH is GTPase deficient and modulates

activities of other Rho GTPases by an inhibitory function', *Mol Cell Biol*, 22(4), pp. 1158-71.

Li, Y., Lal, B., Kwon, S., Fan, X., Saldanha, U., Reznik, T.E., Kuchner, E.B., Eberhart, C., Laterra, J. and Abounader, R. (2005) 'The scatter factor/hepatocyte growth factor: c-met pathway in human embryonal central nervous system tumor malignancy', *Cancer Res*, 65(20), pp. 9355-62.

Lim, Y.P., Lim, T.T., Chan, Y.L., Song, A.C., Yeo, B.H., Vojtesek, B., Coomber, D., Rajagopal, G. and Lane, D. (2007) 'The p53 knowledgebase: an integrated information resource for p53 research', *Oncogene*, 26(11), pp. 1517-21.

Lin, J.C., Jeong, S., Liang, G., Takai, D., Fatemi, M., Tsai, Y.C., Egger, G., Gal-Yam, E.N. and Jones, P.A. (2007) 'Role of nucleosomal occupancy in the epigenetic silencing of the MLH1 CpG island', *Cancer Cell*, 12(5), pp. 432-44.

Lindsey, J.C., Anderton, J.A., Lusher, M.E. and Clifford, S.C. (2005) 'Epigenetic events in medulloblastoma development', *Neurosurg Focus*, 19(5), p. E10.

Lindsey, J.C., Hill, R.M., Megahed, H., Lusher, M.E., Schwalbe, E.C., Cole, M., Hogg, T.L., Gilbertson, R.J., Ellison, D.W., Bailey, S. and Clifford, S.C. (2011) 'TP53 mutations in favorable-risk Wnt/Wingless-subtype medulloblastomas', *J Clin Oncol*, 29(12), pp. e344-6; author reply e347-8.

Lindsey, J.C., Lusher, M.E., Anderton, J.A., Gilbertson, R.J., Ellison, D.W. and Clifford, S.C. (2007) 'Epigenetic deregulation of multiple S100 gene family members by differential hypomethylation and hypermethylation events in medulloblastoma', *Br J Cancer*, 97(2), pp. 267-74.

Lindsey, J.C., Lusher, M.E., Strathdee, G., Brown, R., Gilbertson, R.J., Bailey, S., Ellison, D.W. and Clifford, S.C. (2006) 'Epigenetic inactivation of MCJ (DNAJD1) in malignant paediatric brain tumours', *Int J Cancer*, 118(2), pp. 346-52.

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. and Ecker, J.R. (2009) 'Human DNA methylomes at base resolution show widespread epigenomic differences', *Nature*, 462(7271), pp. 315-22.

Lofton-Day, C., Model, F., Devos, T., Tetzner, R., Distler, J., Schuster, M., Song, X., Lesche, R., Liebenberg, V., Ebert, M., Molnar, B., Grutzmann, R., Pilarsky, C. and Sledziewski, A. (2008) 'DNA methylation biomarkers for blood-based colorectal cancer screening', *Clin Chem*, 54(2), pp. 414-23.

Louis, D.N., Ohgaki, H., Wiestler, O.D., Cavenee, W.K., Burger, P.C., Jouvet, A., Scheithauer, B.W. and Kleihues, P. (2007) 'The 2007 WHO classification of tumours of the central nervous system', *Acta Neuropathol*, 114(2), pp. 97-109.

Loven, J., Hoke, H.A., Lin, C.Y., Lau, A., Orlando, D.A., Vakoc, C.R., Bradner, J.E., Lee, T.I. and Young, R.A. (2013) 'Selective inhibition of tumor oncogenes by disruption of super-enhancers', *Cell*, 153(2), pp. 320-34.

Low, J.A. and de Sauvage, F.J. (2010) 'Clinical experience with Hedgehog pathway inhibitors', *J Clin Oncol*, 28(36), pp. 5321-6.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F. and Richmond, T.J. (1997) 'Crystal structure of the nucleosome core particle at 2.8 A resolution', *Nature*, 389(6648), pp. 251-60.

Lujambio, A., Calin, G.A., Villanueva, A., Ropero, S., Sanchez-Cespedes, M., Blanco, D., Montuenga, L.M., Rossi, S., Nicoloso, M.S., Faller, W.J., Gallagher, W.M., Eccles, S.A., Croce, C.M. and Esteller, M. (2008) 'A microRNA DNA methylation signature for human cancer metastasis', *Proc Natl Acad Sci U S A*, 105(36), pp. 13556-61.

Lusher, M.E., Lindsey, J.C., Latif, F., Pearson, A.D., Ellison, D.W. and Clifford, S.C. (2002) 'Biallelic epigenetic inactivation of the RASSF1A tumor suppressor gene in medulloblastoma development', *Cancer Res*, 62(20), pp. 5906-11.

Ma, W. and Wong, W.H. (2011) 'The analysis of ChIP-Seq data', *Methods Enzymol*, 497, pp. 51-73.

Macleod, D., Charlton, J., Mullins, J. and Bird, A.P. (1994) 'Sp1 sites in the mouse aprt gene promoter are required to prevent methylation of the CpG island', *Genes Dev*, 8(19), pp. 2282-92.

Malkin, D., Li, F.P., Strong, L.C., Fraumeni, J.F., Jr., Nelson, C.E., Kim, D.H., Kassel, J., Gryka, M.A., Bischoff, F.Z., Tainsky, M.A. and et al. (1990) 'Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms', *Science*, 250(4985), pp. 1233-8.

Mantel, N. (1966) 'Evaluation of survival data and two new rank order statistics arising in its consideration', *Cancer Chemother Rep*, 50(3), pp. 163-70.

Marx, A., Kahan, T. and Simon, I. (2013) 'Integrative Analysis of Methylome and Transcriptome Reveals the Importance of Unmethylated CpGs in Non-CpG Island Gene Activation', *Biomed Res Int*, 2013, p. 785731.

Mattick, J.S. and Makunin, I.V. (2006) 'Non-coding RNA', *Hum Mol Genet*, 15 Spec No 1, pp. R17-29.

Maunakea, A.K., Nagarajan, R.P., Bilenky, M., Ballinger, T.J., D'Souza, C., Fouse, S.D., Johnson, B.E., Hong, C., Nielsen, C., Zhao, Y., Turecki, G., Delaney, A., Varhol, R., Thiessen, N., Shchors, K., Heine, V.M., Rowitch, D.H., Xing, X., Fiore, C., Schillebeeckx, M., Jones, S.J., Haussler, D., Marra, M.A., Hirst, M., Wang, T. and Costello, J.F. (2010) 'Conserved role of intragenic DNA methylation in regulating alternative promoters', *Nature*, 466(7303), pp. 253-7.

McMahon, A.P. and Bradley, A. (1990) 'The Wnt-1 (int-1) proto-oncogene is required for development of a large region of the mouse brain', *Cell*, 62(6), pp. 1073-85.

McManamy, C.S., Lamont, J.M., Taylor, R.E., Cole, M., Pearson, A.D., Clifford, S.C. and Ellison, D.W. (2003) 'Morphophenotypic variation predicts clinical behavior in childhood non-desmoplastic medulloblastomas', *J Neuropathol Exp Neurol*, 62(6), pp. 627-32.

McManamy, C.S., Pears, J., Weston, C.L., Hanzely, Z., Ironside, J.W., Taylor, R.E., Grundy, R.G., Clifford, S.C. and Ellison, D.W. (2007) 'Nodule formation and desmoplasia in medulloblastomas-defining the nodular/desmoplastic variant and its biological behavior', *Brain Pathol*, 17(2), pp. 151-64.

Mendrzyk, F., Radlwimmer, B., Joos, S., Kokocinski, F., Benner, A., Stange, D.E., Neben, K., Fiegler, H., Carter, N.P., Reifenberger, G., Korshunov, A. and Lichter, P. (2005) 'Genomic and protein expression profiling identifies CDK6 as novel independent prognostic marker in medulloblastoma', *J Clin Oncol*, 23(34), pp. 8853-62.

Mertz, J.A., Conery, A.R., Bryant, B.M., Sandy, P., Balasubramanian, S., Mele, D.A., Bergeron, L. and Sims, R.J., 3rd (2011) 'Targeting MYC dependence in cancer by inhibiting BET bromodomains', *Proc Natl Acad Sci U S A*, 108(40), pp. 16669-74.

Metcalfe, C. and de Sauvage, F.J. (2011) 'Hedgehog fights back: mechanisms of acquired resistance against Smoothened antagonists', *Cancer Res*, 71(15), pp. 5057-61.

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P., Lee, W., Mendenhall, E., O'Donovan, A., Presser, A., Russ, C., Xie, X., Meissner, A., Wernig, M., Jaenisch, R., Nusbaum, C., Lander, E.S. and Bernstein, B.E. (2007) 'Genome-wide maps of chromatin state in pluripotent and lineage-committed cells', *Nature*, 448(7153), pp. 553-60.

Milde, T., Oehme, I., Korshunov, A., Kopp-Schneider, A., Remke, M., Northcott, P., Deubzer, H.E., Lodrini, M., Taylor, M.D., von Deimling, A., Pfister, S. and Witt, O. (2010) 'HDAC5 and HDAC9 in medulloblastoma: novel markers for risk stratification and role in tumor cell growth', *Clin Cancer Res*, 16(12), pp. 3240-52.

Morgan, H.D., Sutherland, H.G., Martin, D.I. and Whitelaw, E. (1999) 'Epigenetic inheritance at the agouti locus in the mouse', *Nat Genet*, 23(3), pp. 314-8.

Mulhern, R.K., Palmer, S.L., Merchant, T.E., Wallace, D., Kocak, M., Brouwers, P., Krull, K., Chintagumpala, M., Stargatt, R., Ashley, D.M., Tyc, V.L., Kun, L., Boyett, J. and Gajjar, A. (2005) 'Neurocognitive consequences of risk-adapted therapy for childhood medulloblastoma', *J Clin Oncol*, 23(24), pp. 5511-9.

Mund, C., Hackanson, B., Stresemann, C., Lubbert, M. and Lyko, F. (2005) 'Characterization of DNA demethylation effects induced by 5-Aza-2'-deoxycytidine in patients with myelodysplastic syndrome', *Cancer Res*, 65(16), pp. 7086-90.

Mutch, D.M., Berger, A., Mansourian, R., Rytz, A. and Roberts, M.A. (2002) 'The limit fold change model: a practical approach for selecting differentially expressed genes from microarray data', *BMC Bioinformatics*, 3, p. 17.

Nakahara, Y., Northcott, P.A., Li, M., Kongkham, P.N., Smith, C., Yan, H., Croul, S., Ra, Y.S., Eberhart, C., Huang, A., Bigner, D., Grajkowska, W., Van Meter, T., Rutka, J.T. and Taylor, M.D. (2010) 'Genetic and epigenetic inactivation of Kruppel-like factor 4 in medulloblastoma', *Neoplasia*, 12(1), pp. 20-7.

Nakamura, N. and Takenaga, K. (1998) 'Hypomethylation of the metastasis-associated S100A4 gene correlates with gene activation in human colon adenocarcinoma cell lines', *Clin Exp Metastasis*, 16(5), pp. 471-9.

Nan, X., Ng, H.H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N. and Bird, A. (1998) 'Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex', *Nature*, 393(6683), pp. 386-9.

NCI (2013) *National Cancer Institute Factsheet - Cancer Staging*. Available at: www.cancer.gov/cancertopics/factsheet/detection/staging (Accessed: October 6th, 2013).

Ng, H.H., Zhang, Y., Hendrich, B., Johnson, C.A., Turner, B.M., Erdjument-Bromage, H., Tempst, P., Reinberg, D. and Bird, A. (1999) 'MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex', *Nat Genet*, 23(1), pp. 58-61.

Ng, J.M. and Curran, T. (2011) 'The Hedgehog's tale: developing strategies for targeting cancer', *Nat Rev Cancer*, 11(7), pp. 493-501.

Northcott, P.A., Hielscher, T., Dubuc, A., Mack, S., Shih, D., Remke, M., Al-Halabi, H., Albrecht, S., Jabado, N., Eberhart, C.G., Grajkowska, W., Weiss, W.A., Clifford, S.C., Bouffet, E., Rutka, J.T., Korshunov, A., Pfister, S. and Taylor, M.D. (2011a) 'Pediatric and adult sonic hedgehog medulloblastomas are clinically and molecularly distinct', *Acta Neuropathol*, 122(2), pp. 231-40.

Northcott, P.A., Jones, D.T., Kool, M., Robinson, G.W., Gilbertson, R.J., Cho, Y.J., Pomeroy, S.L., Korshunov, A., Lichter, P., Taylor, M.D. and Pfister, S.M. (2012a) 'Medulloblastomics: the end of the beginning', *Nat Rev Cancer*, 12(12), pp. 818-34.

Northcott, P.A., Korshunov, A., Pfister, S.M. and Taylor, M.D. (2012b) 'The clinical implications of medulloblastoma subgroups', *Nat Rev Neurol*, 8(6), pp. 340-51.

Northcott, P.A., Korshunov, A., Witt, H., Hielscher, T., Eberhart, C.G., Mack, S., Bouffet, E., Clifford, S.C., Hawkins, C.E., French, P., Rutka, J.T., Pfister, S. and Taylor, M.D. (2011b) 'Medulloblastoma comprises four distinct molecular variants', *J Clin Oncol*, 29(11), pp. 1408-14.

Northcott, P.A., Nakahara, Y., Wu, X., Feuk, L., Ellison, D.W., Croul, S., Mack, S., Kongkham, P.N., Peacock, J., Dubuc, A., Ra, Y.S., Zilberberg, K., McLeod, J., Scherer, S.W., Sunil Rao, J., Eberhart, C.G., Grajkowska, W., Gillespie, Y., Lach, B., Grundy, R., Pollack, I.F., Hamilton, R.L., Van Meter, T., Carlotti, C.G., Boop, F., Bigner, D., Gilbertson, R.J., Rutka, J.T. and Taylor, M.D. (2009) 'Multiple recurrent genetic events converge on control of histone lysine methylation in medulloblastoma', *Nat Genet*, 41(4), pp. 465-72.

Northcott, P.A., Shih, D.J., Peacock, J., Garzia, L., Morrissy, A.S., Zichner, T., Stutz, A.M., Korshunov, A., Reimand, J., Schumacher, S.E., Beroukhim, R., Ellison, D.W., Marshall, C.R., Lionel, A.C., Mack, S., Dubuc, A., Yao, Y., Ramaswamy, V., Luu, B., Rolider, A., Cavalli, F.M., Wang, X., Remke, M., Wu, X., Chiu, R.Y., Chu, A., Chuah, E., Corbett, R.D., Hoad, G.R., Jackman, S.D., Li, Y., Lo, A., Mungall, K.L., Nip, K.M., Qian, J.Q., Raymond, A.G., Thiessen, N.T., Varhol, R.J., Birol, I., Moore, R.A., Mungall, A.J., Holt, R., Kawauchi, D., Roussel, M.F., Kool, M., Jones, D.T., Witt, H., Fernandez, L.A., Kenney, A.M., Wechsler-Reya, R.J., Dirks, P., Aviv, T., Grajkowska, W.A., Perek-Polnik, M., Haberler, C.C., Delattre, O., Reynaud, S.S., Doz, F.F., Pernet-Fattet, S.S., Cho, B.K., Kim, S.K., Wang, K.C., Scheurlen, W., Eberhart, C.G., Fevre-Montange, M., Jouvet, A., Pollack, I.F., Fan, X., Muraszko, K.M., Gillespie, G.Y., Di

Rocco, C., Massimi, L., Michiels, E.M., Kloosterhof, N.K., French, P.J., Kros, J.M., Olson, J.M., Ellenbogen, R.G., Zitterbart, K., Kren, L., Thompson, R.C., Cooper, M.K., Lach, B., McLendon, R.E., Bigner, D.D., Fontebasso, A., Albrecht, S., Jabado, N., Lindsey, J.C., Bailey, S., Gupta, N., Weiss, W.A., Bognar, L., Klekner, A., Van Meter, T.E., Kumabe, T., Tominaga, T., Elbabaa, S.K., Leonard, J.R., Rubin, J.B., et al. (2012c) 'Subgroup-specific structural variation across 1,000 medulloblastoma genomes', *Nature*, 488(7409), pp. 49-56.

Northcott, P.A., Shih, D.J., Remke, M., Cho, Y.J., Kool, M., Hawkins, C., Eberhart, C.G., Dubuc, A., Guettouche, T., Cardentey, Y., Bouffet, E., Pomeroy, S.L., Marra, M., Malkin, D., Rutka, J.T., Korshunov, A., Pfister, S. and Taylor, M.D. (2012d) 'Rapid, reliable, and reproducible molecular sub-grouping of clinical medulloblastoma samples', *Acta Neuropathol*, 123(4), pp. 615-26.

Noushmehr, H., Weisenberger, D.J., Diefes, K., Phillips, H.S., Pujara, K., Berman, B.P., Pan, F., Pelloski, C.E., Sulman, E.P., Bhat, K.P., Verhaak, R.G., Hoadley, K.A., Hayes, D.N., Perou, C.M., Schmidt, H.K., Ding, L., Wilson, R.K., Van Den Berg, D., Shen, H., Bengtsson, H., Neuvial, P., Cope, L.M., Buckley, J., Herman, J.G., Baylin, S.B., Laird, P.W. and Aldape, K. (2010) 'Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma', *Cancer Cell*, 17(5), pp. 510-22.

Nusslein-Volhard, C. and Wieschaus, E. (1980) 'Mutations affecting segment number and polarity in Drosophila', *Nature*, 287(5785), pp. 795-801.

Ohm, J.E., McGarvey, K.M., Yu, X., Cheng, L., Schuebel, K.E., Cope, L., Mohammad, H.P., Chen, W., Daniel, V.C., Yu, W., Berman, D.M., Jenuwein, T., Pruitt, K., Sharkis, S.J., Watkins, D.N., Herman, J.G. and Baylin, S.B. (2007) 'A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing', *Nat Genet*, 39(2), pp. 237-42.

Okano, M., Bell, D.W., Haber, D.A. and Li, E. (1999) 'DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development', *Cell*, 99(3), pp. 247-57.

Onvani, S., Terakawa, Y., Smith, C., Northcott, P., Taylor, M. and Rutka, J. (2012) 'Molecular genetic analysis of the hepatocyte growth factor/MET signaling pathway in pediatric medulloblastoma', *Genes Chromosomes Cancer*, 51(7), pp. 675-88.

Oshimo, Y., Nakayama, H., Ito, R., Kitadai, Y., Yoshida, K., Chayama, K. and Yasui, W. (2003) 'Promoter methylation of cyclin D2 gene in gastric carcinoma', *Int J Oncol*, 23(6), pp. 1663-70.

Oster, B., Linnet, L., Christensen, L.L., Thorsen, K., Ongen, H., Dermitzakis, E.T., Sandoval, J., Moran, S., Esteller, M., Hansen, T.F., Lamy, P., Laurberg, S., Orntoft, T.F. and Andersen, C.L. (2013) 'Non-CpG island promoter hypomethylation and miR-149 regulate the expression of SRPX2 in colorectal cancer', *Int J Cancer*, 132(10), pp. 2303-15.

Ottmann, O.G., Druker, B.J., Sawyers, C.L., Goldman, J.M., Reiffers, J., Silver, R.T., Tura, S., Fischer, T., Deininger, M.W., Schiffer, C.A., Baccarani, M., Gratwohl, A., Hochhaus, A., Hoelzer, D., Fernandes-Reese, S., Gathmann, I., Capdeville, R. and O'Brien, S.G. (2002) 'A phase 2 study of imatinib in patients with relapsed or refractory

Philadelphia chromosome-positive acute lymphoid leukemias', *Blood*, 100(6), pp. 1965-71.

Ozsolak, F. and Milos, P.M. (2011) 'RNA sequencing: advances, challenges and opportunities', *Nat Rev Genet*, 12(2), pp. 87-98.

Packer, R.J., Cogen, P., Vezina, G. and Rorke, L.B. (1999) 'Medulloblastoma: clinical and biologic aspects', *Neuro Oncol*, 1(3), pp. 232-50.

Packer, R.J., Gajjar, A., Vezina, G., Rorke-Adams, L., Burger, P.C., Robertson, P.L., Bayer, L., LaFond, D., Donahue, B.R., Marymont, M.H., Muraszko, K., Langston, J. and Sposto, R. (2006) 'Phase III study of craniospinal radiation therapy followed by adjuvant chemotherapy for newly diagnosed average-risk medulloblastoma', *J Clin Oncol*, 24(25), pp. 4202-8.

Packer, R.J., Macdonald, T. and Vezina, G. (2010) 'Central nervous system tumors', *Hematol Oncol Clin North Am*, 24(1), pp. 87-108.

Packer, R.J., Macdonald, T., Vezina, G., Keating, R. and Santi, M. (2012) 'Medulloblastoma and primitive neuroectodermal tumors', *Handb Clin Neurol*, 105, pp. 529-48.

Packer, R.J., Meadows, A.T., Rorke, L.B., Goldwein, J.L. and D'Angio, G. (1987) 'Long-term sequelae of cancer treatment on the central nervous system in childhood', *Med Pediatr Oncol*, 15(5), pp. 241-53.

Packer, R.J., Sutton, L.N., Atkins, T.E., Radcliffe, J., Bunin, G.R., D'Angio, G., Siegel, K.R. and Schut, L. (1989) 'A prospective study of cognitive function in children receiving whole-brain radiotherapy and chemotherapy: 2-year results', *J Neurosurg*, 70(5), pp. 707-13.

Packer, R.J. and Vezina, G. (2008) 'Management of and prognosis with medulloblastoma: therapy at a crossroads', *Arch Neurol*, 65(11), pp. 1419-24.

Packer, R.J., Zhou, T., Holmes, E., Vezina, G. and Gajjar, A. (2013) 'Survival and secondary tumors in children with medulloblastoma receiving radiotherapy and adjuvant chemotherapy: results of Children's Oncology Group trial A9961', *Neuro Oncol*, 15(1), pp. 97-103.

Palakurthy, R.K., Wajapeyee, N., Santra, M.K., Gazin, C., Lin, L., Gobeil, S. and Green, M.R. (2009) 'Epigenetic silencing of the RASSF1A tumor suppressor gene through HOXB3-mediated induction of DNMT3B expression', *Mol Cell*, 36(2), pp. 219-30.

Pan, E., Pellarin, M., Holmes, E., Smirnov, I., Misra, A., Eberhart, C.G., Burger, P.C., Biegel, J.A. and Feuerstein, B.G. (2005) 'Isochromosome 17q is a negative prognostic factor in poor-risk childhood medulloblastoma patients', *Clin Cancer Res*, 11(13), pp. 4733-40.

Parsons, D.W., Li, M., Zhang, X., Jones, S., Leary, R.J., Lin, J.C., Boca, S.M., Carter, H., Samayoa, J., Bettegowda, C., Gallia, G.L., Jallo, G.I., Binder, Z.A., Nikolsky, Y., Hartigan, J., Smith, D.R., Gerhard, D.S., Fults, D.W., VandenBerg, S., Berger, M.S., Marie, S.K., Shinjo, S.M., Clara, C., Phillips, P.C., Minturn, J.E., Biegel, J.A., Judkins,

A.R., Resnick, A.C., Storm, P.B., Curran, T., He, Y., Rasheed, B.A., Friedman, H.S., Keir, S.T., McLendon, R., Northcott, P.A., Taylor, M.D., Burger, P.C., Riggins, G.J., Karchin, R., Parmigiani, G., Bigner, D.D., Yan, H., Papadopoulos, N., Vogelstein, B., Kinzler, K.W. and Velculescu, V.E. (2011) 'The genetic landscape of the childhood cancer medulloblastoma', *Science*, 331(6016), pp. 435-9.

Pei, Y., Moore, C.E., Wang, J., Tewari, A.K., Eroshkin, A., Cho, Y.J., Witt, H., Korshunov, A., Read, T.A., Sun, J.L., Schmitt, E.M., Miller, C.R., Buckley, A.F., McLendon, R.E., Westbrook, T.F., Northcott, P.A., Taylor, M.D., Pfister, S.M., Febbo, P.G. and Wechsler-Reya, R.J. (2012) 'An animal model of MYC-driven medulloblastoma', *Cancer Cell*, 21(2), pp. 155-67.

Petitjean, A., Mathe, E., Kato, S., Ishioka, C., Tavtigian, S.V., Hainaut, P. and Olivier, M. (2007) 'Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: lessons from recent developments in the IARC TP53 database', *Hum Mutat*, 28(6), pp. 622-9.

Pfaff, E., Remke, M., Sturm, D., Benner, A., Witt, H., Milde, T., von Bueren, A.O., Wittmann, A., Schottler, A., Jorch, N., Graf, N., Kulozik, A.E., Witt, O., Scheurlen, W., von Deimling, A., Rutkowski, S., Taylor, M.D., Tabori, U., Lichter, P., Korshunov, A. and Pfister, S.M. (2010) 'TP53 mutation is frequently associated with CTNNB1 mutation or MYCN amplification and is compatible with long-term survival in medulloblastoma', *J Clin Oncol*, 28(35), pp. 5188-96.

Pfister, S., Remke, M., Benner, A., Mendrzyk, F., Toedt, G., Felsberg, J., Wittmann, A., Devens, F., Gerber, N.U., Joos, S., Kulozik, A., Reifenberger, G., Rutkowski, S., Wiestler, O.D., Radlwimmer, B., Scheurlen, W., Lichter, P. and Korshunov, A. (2009) 'Outcome prediction in pediatric medulloblastoma based on DNA copy-number aberrations of chromosomes 6q and 17q and the MYC and MYCN loci', *J Clin Oncol*, 27(10), pp. 1627-36.

Pfister, S., Schlaeger, C., Mendrzyk, F., Wittmann, A., Benner, A., Kulozik, A., Scheurlen, W., Radlwimmer, B. and Lichter, P. (2007) 'Array-based profiling of reference-independent methylation status (aPRIMES) identifies frequent promoter methylation and consecutive downregulation of ZIC2 in pediatric medulloblastoma', *Nucleic Acids Res*, 35(7), p. e51.

Pfister, S.M., Korshunov, A., Kool, M., Hasselblatt, M., Eberhart, C. and Taylor, M.D. (2010) 'Molecular diagnostics of CNS embryonal tumors', *Acta Neuropathol*, 120(5), pp. 553-66.

Pietsch, T., Scharmann, T., Fonatsch, C., Schmidt, D., Ockler, R., Freihoff, D., Albrecht, S., Wiestler, O.D., Zeltzer, P. and Riehm, H. (1994) 'Characterization of five new cell lines derived from human primitive neuroectodermal tumors of the central nervous system', *Cancer Res*, 54(12), pp. 3278-87.

Pietsch, T., Waha, A., Koch, A., Kraus, J., Albrecht, S., Tonn, J., Sorensen, N., Berthold, F., Henk, B., Schmandt, N., Wolf, H.K., von Deimling, A., Wainwright, B., Chenevix-Trench, G., Wiestler, O.D. and Wicking, C. (1997) 'Medulloblastomas of the desmoplastic variant carry mutations of the human homologue of Drosophila patched', *Cancer Res*, 57(11), pp. 2085-8.

Pizer, B.L. and Clifford, S.C. (2009) 'The potential impact of tumour biology on improved clinical practice for medulloblastoma: progress towards biologically driven clinical trials', *Br J Neurosurg*, 23(4), pp. 364-75.

Poetsch, A.R. and Plass, C. (2011) 'Transcriptional regulation by DNA methylation', *Cancer Treat Rev*, 37 Suppl 1, pp. S8-12.

Powell, J.E., Henders, A.K., McRae, A.F., Caracella, A., Smith, S., Wright, M.J., Whitfield, J.B., Dermitzakis, E.T., Martin, N.G., Visscher, P.M. and Montgomery, G.W. (2012) 'The Brisbane Systems Genetics Study: genetical genomics meets complex trait genetics', *PLoS One*, 7(4), p. e35430.

Provencal, M., Labbe, D., Veitch, R., Boivin, D., Rivard, G.E., Sartelet, H., Robitaille, Y., Gingras, D. and Beliveau, R. (2009) 'c-Met activation in medulloblastoma induces tissue factor expression and activity: effects on cell migration', *Carcinogenesis*, 30(7), pp. 1089-96.

Provenzano, P.P. and Keely, P.J. (2011) 'Mechanical signaling through the cytoskeleton regulates cell proliferation by coordinated focal adhesion and Rho GTPase signaling', *J Cell Sci*, 124(Pt 8), pp. 1195-205.

Pugh, T.J., Weeraratne, S.D., Archer, T.C., Pomeranz Krummel, D.A., Auclair, D., Bochicchio, J., Carneiro, M.O., Carter, S.L., Cibulskis, K., Erlich, R.L., Greulich, H., Lawrence, M.S., Lennon, N.J., McKenna, A., Meldrim, J., Ramos, A.H., Ross, M.G., Russ, C., Shefler, E., Sivachenko, A., Sogoloff, B., Stojanov, P., Tamayo, P., Mesirov, J.P., Amani, V., Teider, N., Sengupta, S., Francois, J.P., Northcott, P.A., Taylor, M.D., Yu, F., Crabtree, G.R., Kautzman, A.G., Gabriel, S.B., Getz, G., Jager, N., Jones, D.T., Lichter, P., Pfister, S.M., Roberts, T.M., Meyerson, M., Pomeroy, S.L. and Cho, Y.J. (2012) 'Medulloblastoma exome sequencing uncovers subtype-specific somatic mutations', *Nature*, 488(7409), pp. 106-10.

Pui, C.H., Gajjar, A.J., Kane, J.R., Qaddoumi, I.A. and Pappo, A.S. (2011) 'Challenging issues in pediatric oncology', *Nat Rev Clin Oncol*, 8(9), pp. 540-9.

Raffel, C. (2004) 'Medulloblastoma: molecular genetics and animal models', *Neoplasia*, 6(4), pp. 310-22.

Raffel, C., Jenkins, R.B., Frederick, L., Hebrink, D., Alderete, B., Fults, D.W. and James, C.D. (1997) 'Sporadic medulloblastomas contain PTCH mutations', *Cancer Res*, 57(5), pp. 842-5.

Ramaswamy, V., Northcott, P.A. and Taylor, M.D. (2011) 'FISH and chips: the recipe for improved prognostication and outcomes for children with medulloblastoma', *Cancer Genet*, 204(11), pp. 577-88.

Ramdas, L., Coombes, K.R., Baggerly, K., Abruzzo, L., Highsmith, W.E., Krogmann, T., Hamilton, S.R. and Zhang, W. (2001) 'Sources of nonlinearity in cDNA microarray expression measurements', *Genome Biol*, 2(11), p. RESEARCH0047.

Rausch, T., Jones, D.T., Zapatka, M., Stutz, A.M., Zichner, T., Weischenfeldt, J., Jager, N., Remke, M., Shih, D., Northcott, P.A., Pfaff, E., Tica, J., Wang, Q., Massimi, L., Witt, H., Bender, S., Pleier, S., Cin, H., Hawkins, C., Beck, C., von Deimling, A., Hans, V., Brors, B., Eils, R., Scheurlen, W., Blake, J., Benes, V., Kulozik, A.E., Witt, O.,

Martin, D., Zhang, C., Porat, R., Merino, D.M., Wasserman, J., Jabado, N., Fontebasso, A., Bullinger, L., Rucker, F.G., Dohner, K., Dohner, H., Koster, J., Molenaar, J.J., Versteeg, R., Kool, M., Tabori, U., Malkin, D., Korshunov, A., Taylor, M.D., Lichter, P., Pfister, S.M. and Korbel, J.O. (2012) 'Genome sequencing of pediatric medulloblastoma links catastrophic DNA rearrangements with TP53 mutations', *Cell*, 148(1-2), pp. 59-71.

Reik, W., Dean, W. and Walter, J. (2001) 'Epigenetic reprogramming in mammalian development', *Science*, 293(5532), pp. 1089-93.

Remke, M., Hielscher, T., Korshunov, A., Northcott, P.A., Bender, S., Kool, M., Westermann, F., Benner, A., Cin, H., Ryzhova, M., Sturm, D., Witt, H., Haag, D., Toedt, G., Wittmann, A., Schottler, A., von Bueren, A.O., von Deimling, A., Rutkowski, S., Scheurlen, W., Kulozik, A.E., Taylor, M.D., Lichter, P. and Pfister, S.M. (2011) 'FSTL5 is a marker of poor prognosis in non-WNT/non-SHH medulloblastoma', *J Clin Oncol*, 29(29), pp. 3852-61.

Ridola, V., Grill, J., Doz, F., Gentet, J.C., Frappaz, D., Raquin, M.A., Habrand, J.L., Sainte-Rose, C., Valteau-Couanet, D. and Kalifa, C. (2007) 'High-dose chemotherapy with autologous stem cell rescue followed by posterior fossa irradiation for local medulloblastoma recurrence or progression after conventional chemotherapy', *Cancer*, 110(1), pp. 156-63.

Robertson, K.D. (2005) 'DNA methylation and human disease', *Nat Rev Genet*, 6(8), pp. 597-610.

Robertson, K.D., Ait-Si-Ali, S., Yokochi, T., Wade, P.A., Jones, P.L. and Wolffe, A.P. (2000) 'DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters', *Nat Genet*, 25(3), pp. 338-42.

Robertson, K.D. and Wolffe, A.P. (2000) 'DNA methylation in health and disease', *Nat Rev Genet*, 1(1), pp. 11-9.

Robinson, G., Parker, M., Kranenburg, T.A., Lu, C., Chen, X., Ding, L., Phoenix, T.N., Hedlund, E., Wei, L., Zhu, X., Chalhoub, N., Baker, S.J., Huether, R., Kriwacki, R., Curley, N., Thiruvenkatam, R., Wang, J., Wu, G., Rusch, M., Hong, X., Becksfort, J., Gupta, P., Ma, J., Easton, J., Vadodaria, B., Onar-Thomas, A., Lin, T., Li, S., Pounds, S., Paugh, S., Zhao, D., Kawauchi, D., Roussel, M.F., Finkelstein, D., Ellison, D.W., Lau, C.C., Bouffet, E., Hassall, T., Gururangan, S., Cohn, R., Fulton, R.S., Fulton, L.L., Dooling, D.J., Ochoa, K., Gajjar, A., Mardis, E.R., Wilson, R.K., Downing, J.R., Zhang, J. and Gilbertson, R.J. (2012) 'Novel mutations target distinct subgroups of medulloblastoma', *Nature*, 488(7409), pp. 43-8.

Ron, E., Modan, B., Boice, J.D., Jr., Alfandary, E., Stovall, M., Chetrit, A. and Katz, L. (1988) 'Tumors of the brain and nervous system after radiotherapy in childhood', *N Engl J Med*, 319(16), pp. 1033-9.

Rood, B.R., Zhang, H., Weitman, D.M. and Cogen, P.H. (2002) 'Hypermethylation of HIC-1 and 17p allelic loss in medulloblastoma', *Cancer Res*, 62(13), pp. 3794-7.

Rosenbaum, E., Hoque, M.O., Cohen, Y., Zahurak, M., Eisenberger, M.A., Epstein, J.I., Partin, A.W. and Sidransky, D. (2005) 'Promoter hypermethylation as an independent

prognostic factor for relapse in patients with prostate cancer following radical prostatectomy', *Clin Cancer Res*, 11(23), pp. 8321-5.

Rosty, C., Ueki, T., Argani, P., Jansen, M., Yeo, C.J., Cameron, J.L., Hruban, R.H. and Goggins, M. (2002) 'Overexpression of S100A4 in pancreatic ductal adenocarcinomas is associated with poor differentiation and DNA hypomethylation', *Am J Pathol*, 160(1), pp. 45-50.

Rountree, M.R., Bachman, K.E. and Baylin, S.B. (2000) 'DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci', *Nat Genet*, 25(3), pp. 269-77.

Rudin, C.M., Hann, C.L., Laterra, J., Yauch, R.L., Callahan, C.A., Fu, L., Holcomb, T., Stinson, J., Gould, S.E., Coleman, B., LoRusso, P.M., Von Hoff, D.D., de Sauvage, F.J. and Low, J.A. (2009) 'Treatment of medulloblastoma with hedgehog pathway inhibitor GDC-0449', *N Engl J Med*, 361(12), pp. 1173-8.

Rutkowski, S., Bode, U., Deinlein, F., Ottensmeier, H., Warmuth-Metz, M., Soerensen, N., Graf, N., Emser, A., Pietsch, T., Wolff, J.E., Kortmann, R.D. and Kuehl, J. (2005) 'Treatment of early childhood medulloblastoma by postoperative chemotherapy alone', *N Engl J Med*, 352(10), pp. 978-86.

Rutkowski, S., Cohen, B., Finlay, J., Luksch, R., Ridola, V., Valteau-Couanet, D., Hara, J., Garre, M.L. and Grill, J. (2010) 'Medulloblastoma in young children', *Pediatr Blood Cancer*, 54(4), pp. 635-7.

Rutkowski, S., Gerber, N.U., von Hoff, K., Gnekow, A., Bode, U., Graf, N., Berthold, F., Henze, G., Wolff, J.E., Warmuth-Metz, M., Soerensen, N., Emser, A., Ottensmeier, H., Deinlein, F., Schlegel, P.G., Kortmann, R.D., Pietsch, T. and Kuehl, J. (2009) 'Treatment of early childhood medulloblastoma by postoperative chemotherapy and deferred radiotherapy', *Neuro Oncol*, 11(2), pp. 201-10.

Rutkowski, S., von Bueren, A., von Hoff, K., Hartmann, W., Shalaby, T., Deinlein, F., Warmuth-Metz, M., Soerensen, N., Emser, A., Bode, U., Mittler, U., Urban, C., Benesch, M., Kortmann, R.D., Schlegel, P.G., Kuehl, J., Pietsch, T. and Grotzer, M. (2007) 'Prognostic relevance of clinical and biological risk factors in childhood medulloblastoma: results of patients treated in the prospective multicenter trial HIT'91', *Clin Cancer Res*, 13(9), pp. 2651-7.

Ryan, S.L., Schwalbe, E.C., Cole, M., Lu, Y., Lusher, M.E., Megahed, H., O'Toole, K., Nicholson, S.L., Bognar, L., Garami, M., Hauser, P., Korshunov, A., Pfister, S.M., Williamson, D., Taylor, R.E., Ellison, D.W., Bailey, S. and Clifford, S.C. (2012) 'MYC family amplification and clinical risk-factors interact to predict an extremely poor prognosis in childhood medulloblastoma', *Acta Neuropathol*, 123(4), pp. 501-13.

Sadikovic, B., Al-Romaih, K., Squire, J.A. and Zielenska, M. (2008) 'Cause and consequences of genetic and epigenetic alterations in human cancer', *Curr Genomics*, 9(6), pp. 394-408.

Sado, T., Fenner, M.H., Tan, S.S., Tam, P., Shioda, T. and Li, E. (2000) 'X inactivation in the mouse embryo deficient for Dnmt1: distinct effect of hypomethylation on imprinted and random X inactivation', *Dev Biol*, 225(2), pp. 294-303.

Saha, A., Wittmeyer, J. and Cairns, B.R. (2006) 'Chromatin remodelling: the industrial revolution of DNA around histones', *Nat Rev Mol Cell Biol*, 7(6), pp. 437-47.

Saito, Y., Liang, G., Egger, G., Friedman, J.M., Chuang, J.C., Coetzee, G.A. and Jones, P.A. (2006) 'Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells', *Cancer Cell*, 9(6), pp. 435-43.

Sanchez-Aguilera, A., Rattmann, I., Drew, D.Z., Muller, L.U., Summey, V., Lucas, D.M., Byrd, J.C., Croce, C.M., Gu, Y., Cancelas, J.A., Johnston, P., Moritz, T. and Williams, D.A. (2010) 'Involvement of RhoH GTPase in the development of B-cell chronic lymphocytic leukemia', *Leukemia*, 24(1), pp. 97-104.

Saran, A. (2009) 'Medulloblastoma: role of developmental pathways, DNA repair signaling, and other players', *Curr Mol Med*, 9(9), pp. 1046-57.

Sarkar, C., Deb, P. and Sharma, M.C. (2005) 'Recent advances in embryonal tumours of the central nervous system', *Childs Nerv Syst*, 21(4), pp. 272-93.

Schmidl, C., Klug, M., Boeld, T.J., Andreesen, R., Hoffmann, P., Edinger, M. and Rehli, M. (2009) 'Lineage-specific DNA methylation in T cells correlates with histone methylation and enhancer activity', *Genome Res*, 19(7), pp. 1165-74.

Schuller, U., Heine, V.M., Mao, J., Kho, A.T., Dillon, A.K., Han, Y.G., Huillard, E., Sun, T., Ligon, A.H., Qian, Y., Ma, Q., Alvarez-Buylla, A., McMahon, A.P., Rowitch, D.H. and Ligon, K.L. (2008) 'Acquisition of granule neuron precursor identity is a critical determinant of progenitor cell competence to form Shh-induced medulloblastoma', *Cancer Cell*, 14(2), pp. 123-34.

Schwab, M. (1990) 'Amplification of the MYCN oncogene and deletion of putative tumour suppressor gene in human neuroblastomas', *Brain Pathol*, 1(1), pp. 41-6.

Schwalbe, E.C., Lindsey, J.C., Straughton, D., Hogg, T.L., Cole, M., Megahed, H., Ryan, S.L., Lusher, M.E., Taylor, M.D., Gilbertson, R.J., Ellison, D.W., Bailey, S. and Clifford, S.C. (2011) 'Rapid diagnosis of medulloblastoma molecular subgroups', *Clin Cancer Res*, 17(7), pp. 1883-94.

Schwalbe, E.C., Williamson, D., Lindsey, J.C., Hamilton, D., Ryan, S.L., Megahed, H., Garami, M., Hauser, P., Dembowska-Baginska, B., Perek, D., Northcott, P.A., Taylor, M.D., Taylor, R.E., Ellison, D.W., Bailey, S. and Clifford, S.C. (2013) 'DNA methylation profiling of medulloblastoma allows robust subclassification and improved outcome prediction using formalin-fixed biopsies', *Acta Neuropathol*, 125(3), pp. 359-71.

Sengupta, P.K., Smith, E.M., Kim, K., Murnane, M.J. and Smith, B.D. (2003) 'DNA hypermethylation near the transcription start site of collagen alpha2(I) gene occurs in both cancer cell lines and primary colorectal cancers', *Cancer Res*, 63(8), pp. 1789-97.

Sharma, S.V., Lee, D.Y., Li, B., Quinlan, M.P., Takahashi, F., Maheswaran, S., McDermott, U., Azizian, N., Zou, L., Fischbach, M.A., Wong, K.K., Brandstetter, K., Wittner, B., Ramaswamy, S., Classon, M. and Settleman, J. (2010) 'A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations', *Cell*, 141(1), pp. 69-80.

Shi, J., Shi, W., Ni, L., Xu, X., Su, X., Xia, L., Xu, F., Chen, J. and Zhu, J. (2013) 'OCT4 is epigenetically regulated by DNA hypomethylation of promoter and exon in primary gliomas', *Oncol Rep*, 30(1), pp. 201-6.

Shukla, S., Kavak, E., Gregory, M., Imashimizu, M., Shutinoski, B., Kashlev, M., Oberdoerffer, P., Sandberg, R. and Oberdoerffer, S. (2011) 'CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing', *Nature*, 479(7371), pp. 74-9.

Siegel, R., Naishadham, D. and Jemal, A. (2013) 'Cancer statistics, 2013', *CA Cancer J Clin*, 63(1), pp. 11-30.

Silverman, L.R., Demakos, E.P., Peterson, B.L., Kornblith, A.B., Holland, J.C., Odchimar-Reissig, R., Stone, R.M., Nelson, D., Powell, B.L., DeCastro, C.M., Ellerton, J., Larson, R.A., Schiffer, C.A. and Holland, J.F. (2002) 'Randomized controlled trial of azacitidine in patients with the myelodysplastic syndrome: a study of the cancer and leukemia group B', *J Clin Oncol*, 20(10), pp. 2429-40.

Slotkin, R.K. and Martienssen, R. (2007) 'Transposable elements and the epigenetic regulation of the genome', *Nat Rev Genet*, 8(4), pp. 272-85.

Smiraglia, D.J., Rush, L.J., Fruhwald, M.C., Dai, Z., Held, W.A., Costello, J.F., Lang, J.C., Eng, C., Li, B., Wright, F.A., Caligiuri, M.A. and Plass, C. (2001) 'Excessive CpG island hypermethylation in cancer cell lines versus primary human malignancies', *Hum Mol Genet*, 10(13), pp. 1413-9.

Smit, A.F. (1996) 'The origin of interspersed repeats in the human genome', *Curr Opin Genet Dev*, 6(6), pp. 743-8.

Smyth, G.K. (2004) 'Linear models and empirical bayes methods for assessing differential expression in microarray experiments', *Stat Appl Genet Mol Biol*, 3, p. Article3.

Soriano, A.O., Yang, H., Faderl, S., Estrov, Z., Giles, F., Ravandi, F., Cortes, J., Wierda, W.G., Ouzounian, S., Quezada, A., Pierce, S., Estey, E.H., Issa, J.P., Kantarjian, H.M. and Garcia-Manero, G. (2007) 'Safety and clinical activity of the combination of 5-azacytidine, valproic acid, and all-trans retinoic acid in acute myeloid leukemia and myelodysplastic syndrome', *Blood*, 110(7), pp. 2302-8.

Sorm, F., Piskala, A., Cihak, A. and Vesely, J. (1964) '5-Azacytidine, a new, highly effective cancerostatic', *Experientia*, 20(4), pp. 202-3.

Sproul, D., Kitchen, R.R., Nestor, C.E., Dixon, J.M., Sims, A.H., Harrison, D.J., Ramsahoye, B.H. and Meehan, R.R. (2012) 'Tissue of origin determines cancer-associated CpG island promoter hypermethylation patterns', *Genome Biol*, 13(10), p. R84.

Sproul, D., Nestor, C., Culley, J., Dickson, J.H., Dixon, J.M., Harrison, D.J., Meehan, R.R., Sims, A.H. and Ramsahoye, B.H. (2011) 'Transcriptionally repressed genes become aberrantly methylated and distinguish tumors of different lineages in breast cancer', *Proc Natl Acad Sci U S A*, 108(11), pp. 4364-9.

Stearns, D., Chaudhry, A., Abel, T.W., Burger, P.C., Dang, C.V. and Eberhart, C.G. (2006) 'c-myc overexpression causes anaplasia in medulloblastoma', *Cancer Res*, 66(2), pp. 673-81.

Stirzaker, C., Song, J.Z., Davidson, B. and Clark, S.J. (2004) 'Transcriptional gene silencing promotes DNA hypermethylation through a sequential change in chromatin modifications in cancer cells', *Cancer Res*, 64(11), pp. 3871-7.

Strachan, T. and Read, A.P. (2004a) 'Chapter 10: Human Gene Expression', in *Human Molecular Genetics 3*. New York: Garland Publishing, pp. 296-297.

Strachan, T. and Read, A.P. (2004b) 'Chapter 17: Cancer genetics', in *Human Molecular Genetics 3*. New York: Garland Publishing, pp. 488-507.

Strathdee, G., Davies, B.R., Vass, J.K., Siddiqui, N. and Brown, R. (2004a) 'Cell type-specific methylation of an intronic CpG island controls expression of the MCJ gene', *Carcinogenesis*, 25(5), pp. 693-701.

Strathdee, G., Sim, A. and Brown, R. (2004b) 'Control of gene expression by CpG island methylation in normal cells', *Biochem Soc Trans*, 32(Pt 6), pp. 913-5.

Stresemann, C. and Lyko, F. (2008) 'Modes of action of the DNA methyltransferase inhibitors azacytidine and decitabine', *Int J Cancer*, 123(1), pp. 8-13.

Sung, H., Kang, S.H., Bae, Y.J., Hong, J.T., Chung, Y.B., Lee, C.K. and Song, S. (2006) 'PCR-based detection of Mycoplasma species', *J Microbiol*, 44(1), pp. 42-9.

Sure, U., Bertalanffy, H., Isenmann, S., Brandner, S., Berghorn, W.J., Seeger, W. and Aguzzi, A. (1995) 'Secondary manifestation of medulloblastoma: metastases and local recurrences in 66 patients', *Acta Neurochir (Wien)*, 136(3-4), pp. 117-26.

Suresh, T.N., Santosh, V., Yasha, T.C., Anandh, B., Mohanty, A., Indiradevi, B., Sampath, S. and Shankar, S.K. (2004) 'Medulloblastoma with extensive nodularity: a variant occurring in the very young-clinicopathological and immunohistochemical study of four cases', *Childs Nerv Syst*, 20(1), pp. 55-60.

Suzuki, H., Gabrielson, E., Chen, W., Anbazhagan, R., van Engeland, M., Weijenberg, M.P., Herman, J.G. and Baylin, S.B. (2002) 'A genomic screen for genes upregulated by demethylation and histone deacetylase inhibition in human colorectal cancer', *Nat Genet*, 31(2), pp. 141-9.

Suzuki, H., Watkins, D.N., Jair, K.W., Schuebel, K.E., Markowitz, S.D., Chen, W.D., Pretlow, T.P., Yang, B., Akiyama, Y., Van Engeland, M., Toyota, M., Tokino, T., Hinoda, Y., Imai, K., Herman, J.G. and Baylin, S.B. (2004) 'Epigenetic inactivation of SFRP genes allows constitutive WNT signaling in colorectal cancer', *Nat Genet*, 36(4), pp. 417-22.

Swartling, F.J., Grimmer, M.R., Hackett, C.S., Northcott, P.A., Fan, Q.W., Goldenberg, D.D., Lau, J., Masic, S., Nguyen, K., Yakovenko, S., Zhe, X.N., Gilmer, H.C., Collins, R., Nagaoka, M., Phillips, J.J., Jenkins, R.B., Tihan, T., Vandenberg, S.R., James, C.D., Tanaka, K., Taylor, M.D., Weiss, W.A. and Chesler, L. (2010) 'Pleiotropic role for MYCN in medulloblastoma', *Genes Dev*, 24(10), pp. 1059-72.

Tabori, U., Baskin, B., Shago, M., Alon, N., Taylor, M.D., Ray, P.N., Bouffet, E., Malkin, D. and Hawkins, C. (2010) 'Universal poor survival in children with medulloblastoma harboring somatic TP53 mutations', *J Clin Oncol*, 28(8), pp. 1345-50.

Taipale, J. and Beachy, P.A. (2001) 'The Hedgehog and Wnt signalling pathways in cancer', *Nature*, 411(6835), pp. 349-54.

Tate, P.H. and Bird, A.P. (1993) 'Effects of DNA methylation on DNA-binding proteins and gene expression', *Curr Opin Genet Dev*, 3(2), pp. 226-31.

Taylor, M.D., Northcott, P.A., Korshunov, A., Remke, M., Cho, Y.J., Clifford, S.C., Eberhart, C.G., Parsons, D.W., Rutkowski, S., Gajjar, A., Ellison, D.W., Lichter, P., Gilbertson, R.J., Pomeroy, S.L., Kool, M. and Pfister, S.M. (2012) 'Molecular subgroups of medulloblastoma: the current consensus', *Acta Neuropathol*, 123(4), pp. 465-72.

Taylor, R.E., Bailey, C.C., Robinson, K., Weston, C.L., Ellison, D., Ironside, J., Lucraft, H., Gilbertson, R., Tait, D.M., Walker, D.A., Pizer, B.L., Imeson, J. and Lashford, L.S. (2003a) 'Results of a randomized study of preradiation chemotherapy versus radiotherapy alone for nonmetastatic medulloblastoma: The International Society of Paediatric Oncology/United Kingdom Children's Cancer Study Group PNET-3 Study', *J Clin Oncol*, 21(8), pp. 1581-91.

Taylor, R.E., Bailey, C.C., Robinson, K., Weston, C.L., Ellison, D., Ironside, J., Lucraft, H., Gilbertson, R., Tait, D.M., Walker, D.A., Pizer, B.L., Imeson, J., Lashford, L.S., International Society of Paediatric, O. and United Kingdom Children's Cancer Study, G. (2003b) 'Results of a randomized study of preradiation chemotherapy versus radiotherapy alone for nonmetastatic medulloblastoma: The International Society of Paediatric Oncology/United Kingdom Children's Cancer Study Group PNET-3 Study', *J Clin Oncol*, 21(8), pp. 1581-91.

Thomas, K.R. and Capecchi, M.R. (1990) 'Targeted disruption of the murine int-1 proto-oncogene resulting in severe abnormalities in midbrain and cerebellar development', *Nature*, 346(6287), pp. 847-50.

Thompson, M.C., Fuller, C., Hogg, T.L., Dalton, J., Finkelstein, D., Lau, C.C., Chintagumpala, M., Adesina, A., Ashley, D.M., Kellie, S.J., Taylor, M.D., Curran, T., Gajjar, A. and Gilbertson, R.J. (2006) 'Genomics identifies medulloblastoma subgroups that are enriched for specific genetic alterations', *J Clin Oncol*, 24(12), pp. 1924-31.

Thomson, J.P., Skene, P.J., Selfridge, J., Clouaire, T., Guy, J., Webb, S., Kerr, A.R., Deaton, A., Andrews, R., James, K.D., Turner, D.J., Illingworth, R. and Bird, A. (2010) 'CpG islands influence chromatin structure via the CpG-binding protein Cfp1', *Nature*, 464(7291), pp. 1082-6.

Troeger, A., Johnson, A.J., Wood, J., Blum, W.G., Andritsos, L.A., Byrd, J.C. and Williams, D.A. (2012) 'RhoH is critical for cell-microenvironment interactions in chronic lymphocytic leukemia in mice and humans', *Blood*, 119(20), pp. 4708-18.

Tsai, H.C. and Baylin, S.B. (2011) 'Cancer epigenetics: linking basic biology to clinical medicine', *Cell Res*, 21(3), pp. 502-17.

Tsao, D.A., Chang, H.J., Lin, C.Y., Hsiung, S.K., Huang, S.E., Ho, S.Y., Chang, M.S., Chiu, H.H., Chen, Y.F., Cheng, T.L. and Shiu-Ru, L. (2010) 'Gene expression profiles for predicting the efficacy of the anticancer drug 5-fluorouracil in breast cancer', *DNA Cell Biol*, 29(6), pp. 285-93.

Tuma, R.S. (2009) 'Epigenetic therapies move into new territory, but how exactly do they work?', *J Natl Cancer Inst*, 101(19), pp. 1300-1.

Vargo-Gogola, T. and Rosen, J.M. (2007) 'Modelling breast cancer: one size does not fit all', *Nat Rev Cancer*, 7(9), pp. 659-72.

Varley, K.E., Gertz, J., Bowling, K.M., Parker, S.L., Reddy, T.E., Pauli-Behn, F., Cross, M.K., Williams, B.A., Stamatoyannopoulos, J.A., Crawford, G.E., Absher, D.M., Wold, B.J. and Myers, R.M. (2013) 'Dynamic DNA methylation across diverse human cell lines and tissues', *Genome Res*, 23(3), pp. 555-67.

Vermeulen, M. and Timmers, H.T. (2010) 'Grasping trimethylation of histone H3 at lysine 4', *Epigenomics*, 2(3), pp. 395-406.

Wade, P.A., Gegonne, A., Jones, P.L., Ballestar, E., Aubry, F. and Wolffe, A.P. (1999) 'Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation', *Nat Genet*, 23(1), pp. 62-6.

Wang, Z., Gerstein, M. and Snyder, M. (2009) 'RNA-Seq: a revolutionary tool for transcriptomics', *Nat Rev Genet*, 10(1), pp. 57-63.

Warnecke, P.M., Stirzaker, C., Melki, J.R., Millar, D.S., Paul, C.L. and Clark, S.J. (1997) 'Detection and measurement of PCR bias in quantitative methylation analysis of bisulphite-treated DNA', *Nucleic Acids Res*, 25(21), pp. 4422-6.

Weber, M., Davies, J.J., Wittig, D., Oakeley, E.J., Haase, M., Lam, W.L. and Schubeler, D. (2005) 'Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells', *Nat Genet*, 37(8), pp. 853-62.

Weber, M., Hellmann, I., Stadler, M.B., Ramos, L., Paabo, S., Rebhan, M. and Schubeler, D. (2007) 'Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome', *Nat Genet*, 39(4), pp. 457-66.

Wechsler-Reya, R.J. and Scott, M.P. (1999) 'Control of neuronal precursor proliferation in the cerebellum by Sonic Hedgehog', *Neuron*, 22(1), pp. 103-14.

Wetmore, C., Eberhart, D.E. and Curran, T. (2000) 'The normal patched allele is expressed in medulloblastomas from mice with heterozygous germ-line mutation of patched', *Cancer Res*, 60(8), pp. 2239-46.

Widschwendter, M., Fiegl, H., Egle, D., Mueller-Holzner, E., Spizzo, G., Marth, C., Weisenberger, D.J., Campan, M., Young, J., Jacobs, I. and Laird, P.W. (2007) 'Epigenetic stem cell signature in cancer', *Nat Genet*, 39(2), pp. 157-8.

Widschwendter, M., Jiang, G., Woods, C., Muller, H.M., Fiegl, H., Goebel, G., Marth, C., Muller-Holzner, E., Zeimet, A.G., Laird, P.W. and Ehrlich, M. (2004) 'DNA hypomethylation and ovarian cancer biology', *Cancer Res*, 64(13), pp. 4472-80.

Williams, D.A., Zheng, Y. and Cancelas, J.A. (2008) 'Rho GTPases and regulation of hematopoietic stem cell localization', *Methods Enzymol*, 439, pp. 365-93.

Wiseman, A. (2006) 'Gene-hypermutated isoforms of MYC protein: Tumour-suppressor P-53 protective-initiation of apoptosis that normally prevents Burkitt's lymphoma tumorigenesis can be enhanced', *Med Hypotheses*, 66(5), pp. 1046-7.

Wong, J.J., Hawkins, N.J. and Ward, R.L. (2007) 'Colorectal cancer: a model for epigenetic tumorigenesis', *Gut*, 56(1), pp. 140-8.

Wu, B., Zhang, F., Yu, M., Zhao, P., Ji, W., Zhang, H., Han, J. and Niu, R. (2012) 'Up-regulation of Anxa2 gene promotes proliferation and invasion of breast cancer MCF-7 cells', *Cell Prolif*, 45(3), pp. 189-98.

Wutz, A. and Gribnau, J. (2007) 'X inactivation Xplained', *Curr Opin Genet Dev*, 17(5), pp. 387-93.

Yang, Z.J., Ellis, T., Markant, S.L., Read, T.A., Kessler, J.D., Bourboulas, M., Schuller, U., Machold, R., Fishell, G., Rowitch, D.H., Wainwright, B.J. and Wechsler-Reya, R.J. (2008) 'Medulloblastoma can be initiated by deletion of Patched in lineage-restricted progenitors or stem cells', *Cancer Cell*, 14(2), pp. 135-45.

Yao, B., Rakhade, S.N., Li, Q., Ahmed, S., Krauss, R., Draghici, S. and Loeb, J.A. (2004) 'Accuracy of cDNA microarray methods to detect small gene expression changes induced by neuregulin on breast epithelial cells', *BMC Bioinformatics*, 5, p. 99.

Yap, D.B., Chu, J., Berg, T., Schapira, M., Cheng, S.W., Moradian, A., Morin, R.D., Mungall, A.J., Meissner, B., Boyle, M., Marquez, V.E., Marra, M.A., Gascoyne, R.D., Humphries, R.K., Arrowsmith, C.H., Morin, G.B. and Aparicio, S.A. (2011) 'Somatic mutations at EZH2 Y641 act dominantly through a mechanism of selectively altered PRC2 catalytic activity, to increase H3K27 trimethylation', *Blood*, 117(8), pp. 2451-9.

Yauch, R.L., Dijkgraaf, G.J., Alicke, B., Januario, T., Ahn, C.P., Holcomb, T., Pujara, K., Stinson, J., Callahan, C.A., Tang, T., Bazan, J.F., Kan, Z., Seshagiri, S., Hann, C.L., Gould, S.E., Low, J.A., Rudin, C.M. and de Sauvage, F.J. (2009) 'Smoothened mutation confers resistance to a Hedgehog pathway inhibitor in medulloblastoma', *Science*, 326(5952), pp. 572-4.

Zakhary, R., Evren Keles, G., Aldape, K. and Berger, M.S. (2001) 'Medulloblastoma and primitive neuroectodermal tumours', in Kaye, A.H. and Laws Jr, E.R. (eds.) *Brain Tumours- An encyclopaedic approach*. 2nd edition edn. Churchill Livingstone, pp. 605-615.

Zaravinos, A., Lambrou, G.I., Boulalas, I., Delakas, D. and Spandidos, D.A. (2011) 'Identification of common differentially expressed genes in urinary bladder cancer', *PLoS One*, 6(4), p. e18135.

Zeltzer, P.M., Boyett, J.M., Finlay, J.L., Albright, A.L., Rorke, L.B., Milstein, J.M., Allen, J.C., Stevens, K.R., Stanley, P., Li, H., Wisoff, J.H., Geyer, J.R., McGuire-Cullen, P., Stehbens, J.A., Shurin, S.B. and Packer, R.J. (1999) 'Metastasis stage, adjuvant treatment, and residual tumor are prognostic factors for medulloblastoma in children: conclusions from the Children's Cancer Group 921 randomized phase III study', *J Clin Oncol*, 17(3), pp. 832-45.

Zhang, Y., Ng, H.H., Erdjument-Bromage, H., Tempst, P., Bird, A. and Reinberg, D. (1999) 'Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation', *Genes Dev*, 13(15), pp. 1924-35.

Zhukova, N., Ramaswamy, V., Remke, M., Pfaff, E., Shih, D.J., Martin, D.C., Castelo-Branco, P., Baskin, B., Ray, P.N., Bouffet, E., von Bueren, A.O., Jones, D.T., Northcott, P.A., Kool, M., Sturm, D., Pugh, T.J., Pomeroy, S.L., Cho, Y.J., Pietsch, T., Gessi, M., Rutkowski, S., Bognar, L., Klekner, A., Cho, B.K., Kim, S.K., Wang, K.C., Eberhart, C.G., Fevre-Montange, M., Fouladi, M., French, P.J., Kros, M., Grajkowska, W.A., Gupta, N., Weiss, W.A., Hauser, P., Jabado, N., Jouvet, A., Jung, S., Kumabe, T., Lach, B., Leonard, J.R., Rubin, J.B., Liau, L.M., Massimi, L., Pollack, I.F., Shin Ra, Y., Van Meir, E.G., Zitterbart, K., Schuller, U., Hill, R.M., Lindsey, J.C., Schwalbe, E.C., Bailey, S., Ellison, D.W., Hawkins, C., Malkin, D., Clifford, S.C., Korshunov, A., Pfister, S., Taylor, M.D. and Tabori, U. (2013) 'Subgroup-Specific Prognostic Implications of TP53 Mutation in Medulloblastoma', *J Clin Oncol*, 31(23), pp. 2927-35.

Zuo, T., Tycko, B., Liu, T.M., Lin, J.J. and Huang, T.H. (2009) 'Methods in DNA methylation profiling', *Epigenomics*, 1(2), pp. 331-45.

Zurawel, R.H., Allen, C., Wechsler-Reya, R., Scott, M.P. and Raffel, C. (2000) 'Evidence that haploinsufficiency of Ptch leads to medulloblastoma in mice', *Genes Chromosomes Cancer*, 28(1), pp. 77-81.

Zuzak, T.J., Steinhoff, D.F., Sutton, L.N., Phillips, P.C., Eggert, A. and Grotzer, M.A. (2002) 'Loss of caspase-8 mRNA expression is common in childhood primitive neuroectodermal brain tumour/medulloblastoma', *Eur J Cancer*, 38(1), pp. 83-91.