The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Assessing the relationship between DNA methylation and gene expression in germ cell tumours

By

Safiah Hamed Alhazmi, M.Sc.

Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy

December 2015

Abstract

Germ cell tumours (GCTs) are a class of tumours classified histologically into two main types: seminoma and non-seminoma. Prior studies revealed that there is a significant difference in global DNA methylation between those two types, where non-seminomas represent more differentiated cells and exhibit a high level of methylation compared with seminomas that resemble the precursor cells of GCTs. A number of studies have reported that silencing of genes by DNA methylation is a common phenomenon in many types of cancer. However, the silenced genes and the genomic targets that are methylated in GCTs have not yet been systematically identified. Furthermore, many methylation studies in GCTs do not include the level of gene expression in their investigation. We hypothesized that the methylation of genes might play an important role in gene silencing in GCTs, so the main focus of this thesis was studying the relationship between the gene methylation and gene expression in GCT cell lines representing seminoma and non-seminoma. We analysed genome methylation and gene expression of these cell lines using the Illumina infinium Human Methylome 450 bead chip system and Affymetrix Gene Chip Human Genome U133 Plus 2.0 arrays, respectively. We also compared our results with gene expression data from primary tumours in order to identify which events were shared in primary GCTs tumour. qPCR analysis was carried out after treatment of cells with the demethylation agent, 5-aza-2-deoxycytidine, to confirm that expression of identified genes was regulated by methylation.

These analyses showed that differential methylation of CpG islands between seminoma and non-seminoma cell lines correlated well with differential gene expression and revealed that hypermethylation of CpG islands near the transcriptional start site was more strongly correlated with low gene expression than was methylation of other regions.

Meanwhile, methylation analysis identified uniquely methylated genes and features for each cell line, which may imply an underlying mechanism of their development. One-hundred and forty-seven silenced genes which exhibited a difference in methylation and expression between seminoma and non-seminoma cell lines were identified, some of these genes were also differentially expressed in primary tumours. Re-expression of selected silenced genes in non-seminoma cells after treatment with 5-aza-2-deoxycytidine confirmed that methylation played a role in gene silencing. Some of the genes identified are closely associated with pluripotency and implicated in chemosensitivity (PRDM14, KLF4, TDRD12, DDX43, MNS1, and RBMXL2). Silencing of these genes could therefore account for the progression process from seminoma to non-seminoma.

PRDM14 was given special attention as it plays an important role in germ cell development and maintenance of germ cell pluripotency. The role of PRDM14 in GCT biology was studied, revealing that high expression of PRDM14 in combination with 5-aza-2-deoxycytidine treatment increased the response of cells to chemotherapy compared with those that had low levels of PRDM14. In addition, this study supports a growing body of literature on PRDM14 suggesting that this gene plays a critical role in DNA demethylation.

Publication/Poster presentation

Journal published:

Dzul Azri Mohamed Noor, Jennie N Jeyapalan, **Safiah Alhazmi**, Matthew Carr, Benjamin Squibb, Claire Wallace, Christopher Tan, Martin Cusack, Jaime Hughes, Tom Reader, Janet Shipley, Denise Sheer & Paul J Scotting (2016). *npj Genomic Medicine*.

Posters:

Safiah Alhazmi, Dzul Noor, Chris Tan, Jaime Hughes, Claire Wallace, Paul Scotting. Aberrant DNA Methylation and Silencing of Gene Expression: Possible Diagnostic and therapeutic Implications in Testicular Germ Cell Tumours presented as a poster presentation at Functional genomics and system biology conference, 21-23 November 2013. Hinxton, Cambridge, UK.

- Safiah Alhazmi, Dzul Noor, Chris Tan, Jaime Hughes, Claire Wallace, Paul Scotting. Aberrant DNA Methylation and Silencing of Gene Expression: Possible Diagnostic and therapeutic Implications in Testicular Germ Cell Tumours presented as a poster presentation at UK genome science meeting 2013. 2-4 September 2013. University of Nottingham.
- Safiah Alhazmi, Matthew Carr, Dzul Noor, Jennie N Jeyapalan, Claire Wallace, Paul J Scotting. Silencing of pluripotency-associated genes by DNA methylation in differentiated germ cell tumours presented as a poster presentation at Genome science 2014: Biology, Technology & Bioinformatics conference. 1-3 September. University of Oxford, UK.
- Safiah Alhazmi, Matthew Carr, Dzul Noor, Jennie N Jeyapalan, Claire Wallace, Paul J Scotting. Silencing of pluripotency-associated genes by DNA methylation in differentiated germ cell tumours presented as a poster presentation at the 3rd Annual HumanMethylation450 Array Analysis conference. 8-9 May 2014. University College London.
- Safiah Alhazmi, Matthew Carr, Dzul Noor, Jennie N Jeyapalan, Claire Wallace, Paul J Scotting. Silencing of pluripotency-associated genes by DNA methylation in differentiated germ cell tumours presented as a poster presentation at at the at 7th Saudi student conference. 1-2 February 2014. Edinburgh, UK.

Acknowledgements

I thank Allah (my God) who is the most gracious and the most merciful for giving me the health, strength, and ability to complete this study.

I would like to thank those who gave me academic support and assistance in laboratory techniques. First of all, I am very thankful to my supervisor Dr Paul Scotting for his patient and support to develop me on the way of being a scientist and direct my project. I have been extremely lucky to have a supervisor who is so kindly concerned about my study and taught me nearly all the basics of understanding molecular genetics.

Secondly, my sincere appreciations also go to Professor Jane Hewitt for her invaluable advice to develop my project. Meanwhile, my thanks also go to Jaime Hughes for her expert, patient assistance in laboratory techniques. I would like to thank all past and present members of D13 for their help and friendship. They provide me a positive environment for study.

A huge thank to my mother for her prayer and supporting me throughout all of my life. I would especially like to thank my lovely son (Mohammed) and all my brothers and sisters for their prayers and support.

Finally, I would like to thank King Abdulaziz University for funding my PhD study.

Table of Contents

List of Tables

List of Figures

Abbreviations

5-aza	5-azadeoxycytidine
5mC	5 methylcytosine
BLIMP-1	B-lymphocyte-induced maturation protein-1
bp	base pair
CIMP	CpG island methylator phenotype
CIS	Carcinoma in situ
CGI	CpG island
CpG	Cytosine-phosphate-Guanine
DMSO	Dimethyl sulfoxide
DNMTs	DNA methyltransferases
EC	Embryonal Carcinoma
EDTA	Ethylenediaminetetraacetic acid
ESC	Embryonic stem cell
ES	Embryonic stem
FBS	Fetal bovine serum
GCTs	Germ Cell Tumours
GCT27	Cell line from testicular Embryocarcinoma
GCT44	Cell line from testicular YST
HDAC	Histone deacetylases
ITGCNU	Intratubular germ cell neoplasia, unclassified
KLF	Krüppel-like factor
L1	LINE-1
LB	Luria broth
LINE	Long interspersed nuclear element
MBD	Methyl-CpG binding domain
MBP	Methyl-binding proteins
MGMT	O6-methylguanine–DNA methyltransferase
MLTF	Major late transcription factor

ncRNA	Non-coding RNA
NSCLC	Non-small cell lung carcinoma
NSEM	Non-seminomas
PBS	Phosphate buffered saline
PBST	PBS with Tween20
PFA	paraformaldehyde
PGCs	Primordial germ cells
RIN	RNA integrity number
RLGS	Restriction landmark genomic scanning
RMA	Robust multi-array average
RT	Reverse-Transcriptase
SEM	Seminoma
SDS	Sodium dodecyl sulfate
SDS-PAGE	SDS-Polyacrylamide gel electrophoresis
SINE	Short interspersed nuclear element
SOC	Super optimal broth with catabolite repression
TAE buffer	Buffer mixed with Tris base, acetic acid, and EDTA
TCam-2	Cell line from testicular seminoma
TERT	Teratoma
TET	Ten-eleven translocation
TF	Transcription factor
TFAP2C	Transcription factor AP-2 gamma
TGCT	Testicular germ cell tumour
tPRDM14	GCT44 cell line transfected
TSG	Tumour suppressor gene
TSS	Transcription start site
UTR	Untranslated region
YST	Yolk sac tumour

Chapter 1. Introduction

DNA methylation is a critically important epigenetic alteration that affects gene expression without altering the underlying DNA sequence. It is known that DNA methylation is involved in the normal control of cells. Aberrant DNA methylation is observed as a frequent event in cancers where it may result in transcriptional repression or silencing (Akhavan-Niaki and Samadani, 2013). Therefore, DNA methylation can be used as a way of identifying important genes that play a role in tumorigenesis.

Germ cell tumours (GCTs) are malignant or benign tumours believed to be derived from primordial germ cells (PGCs) (di Pietro *et al*., 2005, Jeyapalan *et al*., 2011). GCTs affect both children and adults, and occur in gonadal and extragonadal regions. They are classified histologically into two major groups, seminomas and non-seminomas, where the latter group exhibits a higher degree of DNA methylation than seminomas (Netto *et al*., 2008).

Recent studies have revealed that DNA methylation is associated with tumourigenesis, particularly in the silencing of critical growth regulators such as tumour suppressor genes (Baylin, 2005) or pluripotency markers, including several transcription factors (Western *et al*., 2010). However, for GCTs, the genomic targets that are methylated and which silenced genes are most likely to play a role in the tumours' biology remain to be systematically determined. This thesis makes an attempt to fill this gap to provide insight into the relationship between DNA methylation and gene expression.

1.1. DNA methylation

Over recent decades, cancer biology research has been driven by the genetic revolution that allowed the identification of structural DNA changes (mutations) that changed gene expression in hereditary and sporadic cancers. In the late 1967s, a hypothesis was proposed that there is another factor that could change gene expression without altering the sequence of DNA (Scarano *et al*., 1967), later known as an epigenetic alteration, including DNA methylation and histone modifications.

In eukaryotes, DNA methylation involves the covalent addition of a methyl group to carbon five (C5) of the cytosine base within CpG dinucleotides of DNA (Bird, 2002). The content of GC nucleotides in the human genome is approximately 42% (Jabbari and Bernardi, 2004) but the occurrence frequency of CpG dinucleotides in the DNA sequence is only 1% of expected frequency (Illingworth and Bird, 2009). This low frequency of the CpG dinucleotide could be explained by the action of spontaneous deamination. The cytosines in CpG dinucleotides often become methylated producing methylcytosines but these methylated cytosines are highly susceptible to spontaneous deamination that converts methylcytosine to thymine (Scarano *et al*., 1967). Then, during DNA replication, the mismatched TG base pair is repaired to TA resulting in a permanent alteration or mutation in the DNA sequence. Unmethylated cytosine, on the other hand, converts by spontaneous deamination to uracil, which can be repaired by the cells to cytosine. Therefore, a high level of conversion of methylated cytosine to thymine decreases the amount of CpG dinucleotides over replication times (Cooper and Krawczak, 1989). However, although the average CpG dinucleotide level within the genome is low, some regions in the genome contain a high density of CpG dinucleotides; these include repetitive

sequences and regions known as CpG islands (CGIs) (Cooper *et al*., 1983).

5-methylcytosine (5-mC), sometimes classed as the fifth nucleotide, was first discovered by Johnson and Coghill (1925) in Tubercle Bacillus. Later, in 1948, Hotchkiss proved the presence of this methylated cytosine in the DNA of calf thymus cells using paper chromatography. Kelly and Smith (1970) reported that specific methylation sensitive restriction enzymes could be used as a useful method to distinguish between methylated and unmethylated cytosines in DNA.

The first hypothesis that DNA methylation could play a role in gene expression was reported by Scarano (1973). Subsequently, many studies were carried out regarding this epigenetic phenomenon (Holliday and Pugh, 1975, Li et al., 1993, Yoder et al., 1997, Baylln et al., 1997, Payer and Lee, 2008, Udali et al., 2015). General reviews for this phenomenon are reported by Razin and Cedar (1991) as well as by Roberston (2005).

However, many researchers have argued that DNA methylation is a secondary event in gene silencing following histone modifications. This argument emerged from genetic studies in the *Neurospora crassa* model (Tamaru and Selker, 2001) showing methylation of histone methyltransferase (H3K9) led to methylate all CGI and silence certain genes. Moreover, Richards and Elgin (2002) suggested that DNA methylation might be controlled by histone methylation. Feldman *et al.* (2006) showed that methylation of histone methyltransferase (H3K9) causes local heterochromatinization followed by an increase in DNA methylation in promotor regions by the enzymes Dnmt3a/3b, thus silencing specific genes.

However, there is still more evidence required to support this argument because the human genome is more complex than *Neurospora crassa*. In addition, the integration of DNA methylation with other epigenetic modifications is a complex process and depends on multiple components. Furthermore, genes involved in H3K9 methylation could also be involved in *de novo* DNA methylation (Rose and Klose, 2014).

Whether DNA methylation is a cause or a consequence of downregulation of gene expression is still controversial. As a cause of gene repression, methylation could affect transcription factor binding sites. Alternatively, DNA methylation may stabilize chromatin modification that causes gene repression.

1.1.1 DNA methylation machinery

The DNA methylation machinery primarily consists of DNA methyltransferases, also known as DNMTases or DNMTs, a family of enzymes that are responsible for catalysing the transfer of methyl groups during the DNA methylation process. The DNA methyltransferases catalyse DNA methylation by initiating and maintaining the addition of methyl groups to the 5th carbon position of the cytosine ring within the CpG dinucleotide, thereby forming 5-methylcytosine.

There are three DNMTs responsible for the establishment and maintenance of DNA methylation, namely DNMT1, DNMT3a and DNMT3b (Gnyszka *et al*., 2013). DNMTs are classified as maintenance DNA methyltransferase (DNMT1) or *de novo* DNA methyltransferases (DNMT3a and DNMT3b). Another relevant protein, DNA cytosine-5-methyltransferase 3-like protein (DNMT3L), is recognized as a regulatory factor for the *de novo* DNA methylation process (Cheng and

Blumenthal, 2008). Additionally, DNMT2 shows sequence similarity to DNA methyltransferases including all of the conserved methyltransferase motifs but cannot methylate DNA (Robertson, 2001). Recent studies pointed out that DNMT2 methylate tRNA in some organisms such as *Arabidopsis thaliana* (Goll *et al*., 2006), *Schistosoma mansoni* and *Drosophila melanogaster* (Raddatz *et al*., 2013).

DNMT1 is the most abundant of the DNMTs, particularly in mammalian cells, and is recognised as a maintenance DNA methyltransferase. This is because it acts primarily on hemi-methylated DNA and maintains the pre-existing methylation patterns, to deliver symmetrically methylated CpG dinucleotides in the double strands of newly replicated DNA (Bestor, 1992). Robertson (2001) showed that there is no enzyme that can compensate the function of DNMT1. Furthermore, Jin *et al.* (2011) found that the deletion of both alleles of DNMT1 from mouse embryos at day E9 was lethal. DNMT3A and DNMT3B are called *de novo* methylases because they initiate methylation of unmethylated DNA (Szyf, 2009). DNMT3a methylates CpG dinucleotides at a faster rate than DNMT3b, but slightly slower than DNMT1. DNMT3L binds to the catalytic domain of *de novo* methyltransferases then accelerates their ability to attach to S-adenosyl-L-methionine (Figure 1.1) and stimulates their ability to methylate DNA (Jin *et al*., 2011).

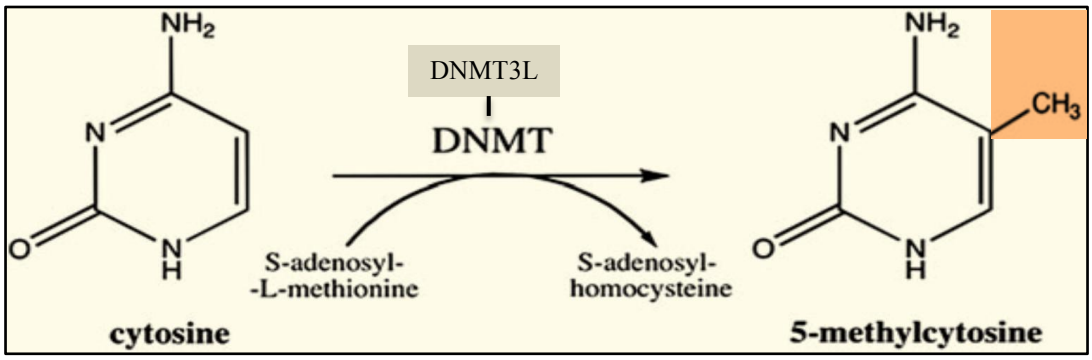


Figure 1.1: DNA methyltransferases convert cytosine to 5methylcytosine. DNMT3L binds *de novo* methyltransferases then accelerates their ability to attach to S-adenosyl-L-methionine (SAM) to stimulate their ability to methylate DNA by addition of methyl group. Figure adapted from Meng *et al.* (2015)

Other key components of the epigenetic regulation of gene expression by DNA methylation are the methylation mark readers known as methyl-CpG-binding domain (MBD) proteins (proteins with a methyl binding domain), such as MBD1-4 and MeCP2 (Figure 1.2). These bind specifically to methylated CpGs, thereby silencing transcription and modulating gene expression. Other proteins that bind to methylated CpG are classified according to their domain type. For example, the zinc finger protein family such as Kaiso related proteins, ZBTB4, ZBTB38, UHRF1 and UHRF2, use zinc finger domains to bind methylated CpGs at specific sequences to repress gene expression through interaction with histone deacetylases, to remodel chromatin into a repressive state (Parry and Clarke, 2011, Meng *et al*., 2015).

The activities of the MBD proteins are closely supported by the local chromatin structures, which are often primarily responsible for determining the repression or transcription of certain genes (Fuks *et al*., 2003). Therefore, MBDs contribute to the formation of heterochromatinc regions in the genome thus enforcing silencing of genes in these regions (Rose and Klose, 2014). For example, Fuks *et al.* (2003) reported that MeCP2 could bind to the chromatin modifying enzymes, histone lysine methyltransferases, to modify histone H3 on lysine at position 9 (H3K9me) resulting in a condensed chromatin structure and transcriptional repression of genes in that region. In this regard, the state of the chromatin structure may be critically important in the regulation of transcription and the repression of gene.

In addition, methylation can affect individual gene expression by direct blocking of transcription factor (TF) binding such as AP-2 and MLTF. Therefore, cytosine methylation at promoter sequences can prevent binding of a TF to its binding site resulting in repression of gene expression (Comb and Goodman, 1990). For example, in mouse astrocyte differentiation from neuroepithelial cells at E11.5, removal of methylation from CpG islands in the promoter region of GFAP (the glial fibrillary acidic protein gene) promotes the binding of STAT3 (signal transducer and activator of transcription 3) to its binding site in the GFAP promoter resulting in transcriptional activation of GFAP (Cheng *et al*., 2011).

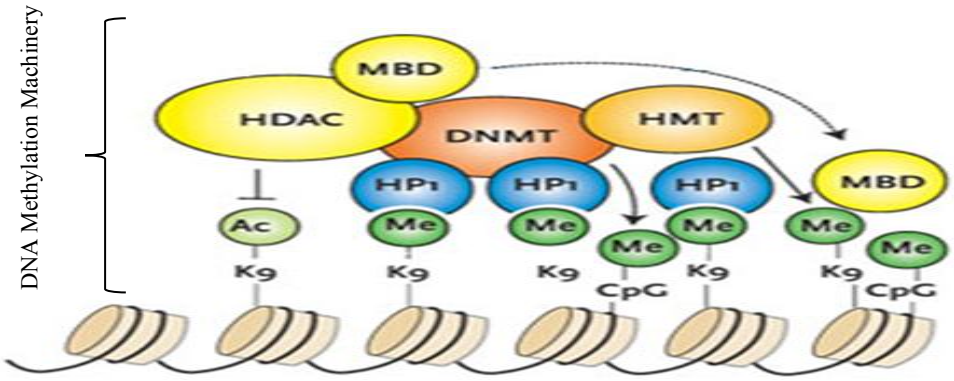


Figure1.2: DNA methylation machinery. DNA methyltransferases (DNMTs) and methyl-CpG-binding proteins (MBD) catalyse the transfer of methyl groups to histone and DNA sequence while Histone deacetylases (HDAC) remove acetyl groups from lysine residues making way for methylation. Figure adapted from Feinberg and Tycko (2004).

1.1.2. DNA methylation in the human genome

The human genome thus contains two sets of information, namely genetic and epigenetic. The genetic information, is primarily responsible for providing the needed scheme for the processing and manufacture of all important proteins and RNAs required for the survival of organisms while the epigenetic information is used to determine where, how and when the genetic information should be used, including its transcription (Esteller, 2011). DNA methylation is one of the major forms of epigenetic information in the human genome that is responsible for ensuring appropriate gene expression patterns.

Generally, the methylation of DNA in the human genome may potentially affect the binding of proteins to their target DNA sequences, thereby resulting in a number of effects on the genome. DNA methylation has diverse effects on the human genome including epigenetic inheritance, genomic stability, transcriptional repression, X-chromosome inactivation, and imprinting of specific DNA sequences. Changes in DNA methylation have in numerous cases been correlated with genetic lesions and genomic instability such as Rett syndrome that is caused by mutations in the methyl-CpG binding gene (MeCP2) which lead to suppress genes that are critical in normal brain function, and Facial anomalies syndrome which is caused by a mutation in DNMT3b that is associated with hypomethylation of pericentromic satellite regions thus genome instability (Paulsen and Ferguson-Smith, 2001).

1.1.2.1 The role of DNA methylation in human development and cell differentiation

DNA methylation is involved in the normal control of human development and cell differentiation. For example, many germline-specific genes are methylated in somatic cells but not in germ cells. DNA methylation reprogramming involves regulation of transcription factors that are important in early development (Wagner *et al*., 2014).

Normal germ cells undergo many processes, including the reprogramming of DNA methylation during embryogenesis (Mochizuki *et al*., 2012). DNA methylation is removed during zygote formation and re-established after implantation (Jin *et al*., 2011) to regulate gene expression in germ cells during embryogenesis (Messerschmidt *et al*., 2014) (Figure 1.3).

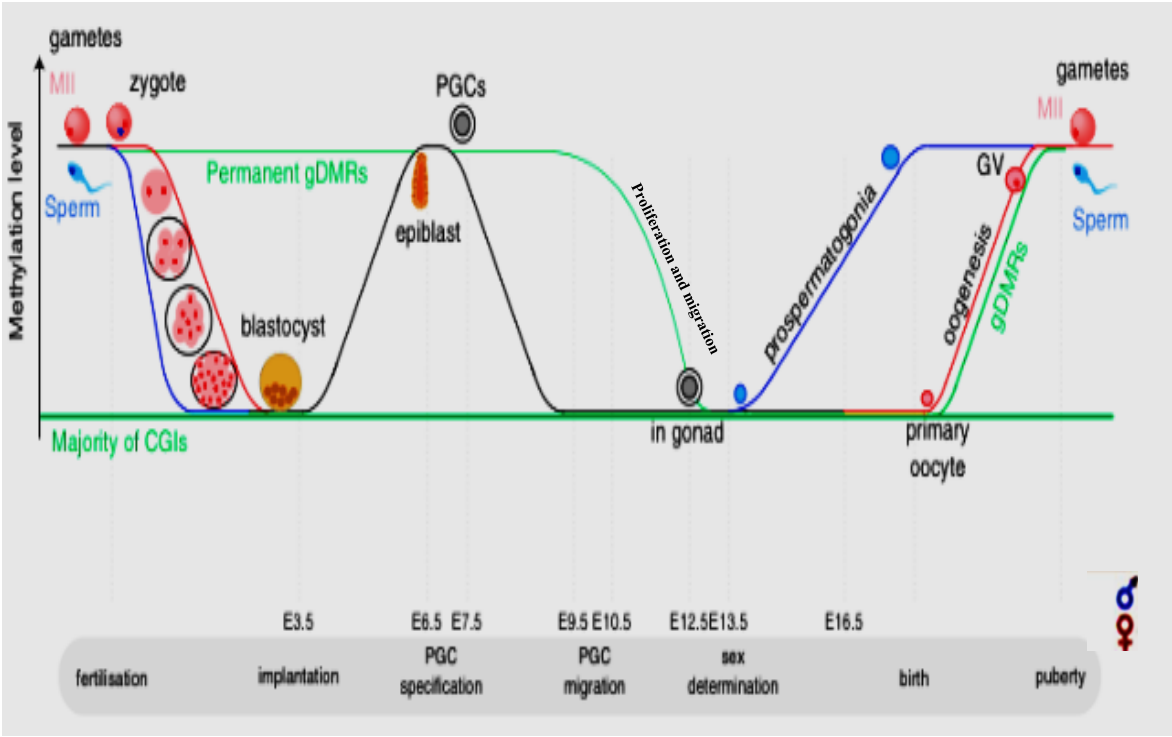


Figure 1.3: DNA methylation changes during embryogenesis: Maternal and paternal DNA in the zygote undergo reprogramming of DNA methylation. Figure adapted from Saadeh and Schulz (2014).

During embryogenesis, expression of certain genes is regulated by DNA methylation in a phenomenon called genomic imprinting resulting in parent-of-origin specific manner. Genomic imprinting is a process where a gene is silenced in one allele that inherited from one parents while other allele that inherited from other parent is expressed (Reik *et al.*, 2003). For example, IGF2 (Insulin-like growth factor 2) expressed from the paternal allele while H19 expressed from the maternal allele (Bartolomei *et al*., 1991).

1.1.2.2 Aberrant DNA methylation in cancer

Compared with the role of DNA mutation, there are relatively few studies on the role of DNA methylation in the development of cancer. The genetic hypotheses of cancer, which focus on the effect of mutations, have guided cancer research for many decades. In 1983, Feinberg and Vogelstein discovered loss of DNA methylation at CpG dinucleotides in tumour sample. This effort was the first evidence that aberrant methylation could play a role in cancer. That study triggered interest in the role of epigenetics in cancer. The widely accepted 'multiple-hits' hypothesis, was proposed by Carol O. Nordling, states that cancer is the result of accumulations of mutations (Nordling, 1953). This hypothesis formulated by Knudson who performed a statistical analysis on 48 cases of retinoblastoma and noted that a cancer is formed by two mutational hits leading to inactivation of tumour suppressor gene (TSG). Knudson (1971) pointed that the inherited form mutation of one of the two copies of a specific tumour suppressor gene is inherited in the germline and the second hit occurs in somatic cells. While the nonhereditary form, both mutations occur in somatic cells therefore, this hypothesis called the 'two-hit' hypothesis or the Knudson hypothesis. Later, in 2001, Knudson pointed out that a hit event that lead to silencing of gene expression might be caused by methylation of TSG (Knudson, 2001).

Tumorigenesis occurs as a result of both genetic and epigenetic alterations. Genetic mutations are a sequence change in the genomic DNA such as substitution, deletion or insertion of nucleotides. On the other hand, the epigenetic modifications that characterise the onset of cancer are generally attributed to the disruption of mechanisms such as DNA methylation and post-translational modifications of nucleosome positioning and histone modifications, all of which result in activation or inactivation of particular genes without changing DNA sequence.

However, tumourigenesis is widely believed to involve a complex interaction between the two sets of mechanisms: genetic or epigenetic factors (You and Jones, 2012). A number of studies revealed that DNA methylation is a crucial player in both aberrant DNA repair and genome instability in cancer by two distinct pathways, through a mechanism of DNMT1-PCNA (Proliferation cell nuclear antigen) interaction leading to hypermethylation and silencing of the mismatch repair gene (hMLH1) or by a methylation proficiency mechanism that reduces DNA methylation, thus increasing genome instability (Jones and Laird, 1999, Rizwana and Hahn, 1999, Robertson and Jones, 2000)

Single locus DNA hypermethylation

Curtin *et al.* (2011) reported that CGI hypermethylation of TSGs is a common marker of human cancers and Esteller (2007) claimed that identification of hypermethylated TSGs for each human cancer could be a target for treatment of those cancers. According to Sproul *et al.* (2012), a study that analysed the methylation profile of 1,154 cancers from seven different tissue types, more than a thousand genes are subject to CGI hypermethylation in these tissues and half of these genes show some degree of tissue specificity of gene expression.

In cancer cells, hypermethylation of CGIs located in the promoter region of TSGs can lead to transcriptional silencing or downregulation of expression that may reduce the function of that gene such silencing could contribute to dysfunction of cell signaling, DNA repair, remodeling or apoptosis for almost all types of tumour facilitating progression of tumourigenesis (Ehrlich and Jiang, 2005, Heyn and Esteller, 2012).

In other findings, it was documented that the DNA methylation pattern is similar in inherited and sporadic cancers of the same type of tumour (Esteller *et al*., 2001). DNA methylation associated with cancer varies according to tumour type (Esteller, 2011). However, there is an opinion that regulation of genes by methylation is different in each cancer and certain CpGs are more important than other in silencing of genes (Esteller, 2002).

DNA hypermethylation not only causes local silencing for specific genes but this repression can also extend to large regions of chromosomes (more than 1 Mb), leading to suppression of neighbouring unmethylated genes. This effect is referred to as long range epigenetic silencing (LRES) and might also be important in the initiation and progression of cancer (Clark, 2007, Swami, 2010, Forn *et al*., 2013).

Global genomic hypomethylation

The first epigenetic change described in human cancer was the loss of DNA methylation throughout various areas of the genome. This was demonstrated using methylation-sensitive restriction enzyme digestion followed by Southern blotting (Feinberg and Vogelstein, 1983). This study reported that adenocarcinomas of the colon had low global DNA methylation compared with ordinary colonic epithelium. Later the same year, Gama-Sosa *et al.* (1983) used methylation-sensitive restriction enzyme digestion analysis followed by high-performance liquid chromatography and showed that low content of 5mC was associated with tumour progression.

Since this discovery, many cancer studies have revealed that tumours have low 5mC compared to normal tissue. This difference is associated with low methylation in intergenic repetitive regions that can accelerate the instability of the genome. Global DNA hypomethylation in

malignancy is significantly connected with hypomethylation of certain DNA repeatitive sequence or single-duplicate DNA arrangements (Ross *et al*., 2010). Several studies reported that a low level of DNA methylation (which refers to hypomethylation) is seen in the early stage of many types of cancer (Hernandez-Blazquez *et al*., 2000, Kinney *et al*., 2008, Park *et al*., 2009).

In contrast to hypermethylation of TSGs, Okamoto (2012) noted that hypomethylation of CGIs of an oncogene may lead to increasing gene expression (Bert et al., 2013) and can contribute to an increase proliferation and tumour formation. It should be noted that during carcinogenesis the hypomethylation status is not only restricted to oncogenes but also has an effect on the whole genome. For example, when Gaudet and colleagues (2003) introduced a hypomorphic *DNMT1* allele in mice, they found increased loss of heterozygosity, aneuploidy, chromosomal instability, and the mice eventually developed aggressive T cell lymphomas. This information should be kept in mind during cancer therapy by demethylation drugs, as these might cause increased genome instability (Tang *et al*., 2009).

Hypomethylation has been proposed as a biomarker because it satisfies three requirements of a successful biomarker: the need to be an early change in a disease, showing an early discrimination of the tumour, and being detectable in the background of molecular changes in DNA of normal cells (Wilson *et al.*, 2007).

Although many studies have reported that hypermethylation is associated with the suppression of specific genes, there are a few studies revealing that hypomethylation has the same action. For example, inactivation of the synuclein γ gene (*SNCG*) in ovarian and breast cancers is associated with regional hypomethylation of a CGI of this oncogene (Gupta *et al*., 2003). However, the majority of studies

involving the role of methylation in cancer pay attention primarily to hypermethylation status in CGIs.

In conclusion, both the abnormal hypermethylation of promoter CGIs and global hypomethylation (Figure 1.4) are likely together to initiate or progress tumours (Bariol *et al*., 2003, Lujambio *et al*., 2008). In general, detecting hypermethylation of specific genes could be used as a prognostic marker while DNA hypomethylation might be helpful in distinguishing between subtypes of tumours or between normal and cancer cells where most tumours can be classified according to their degree of methylation.

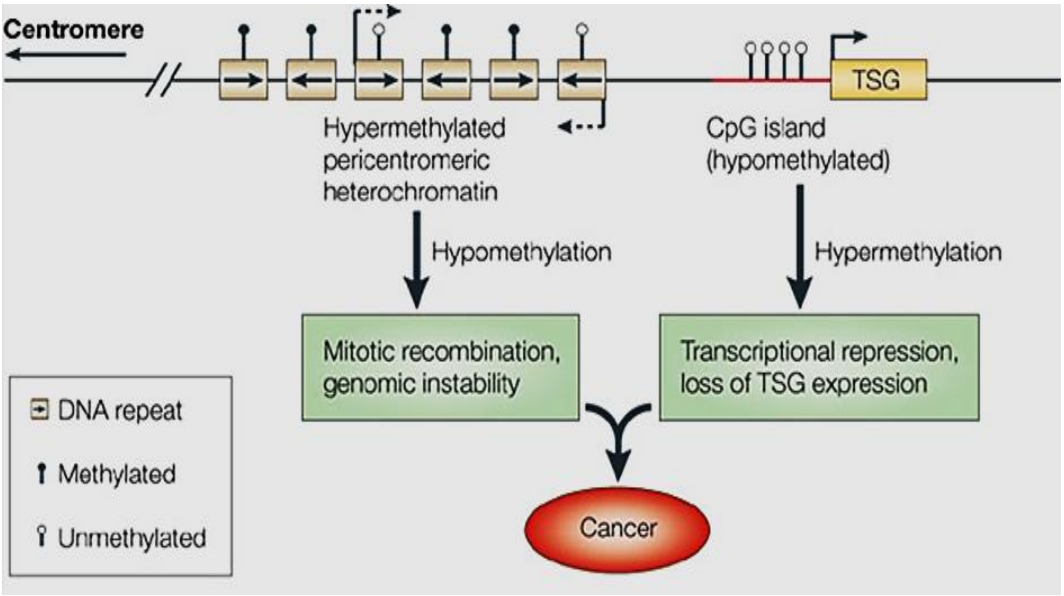


Figure 1.4: The role of DNA methylation in tumourigenesis: Hypomethylation of repetitive elements leads to genomic instability and could activate oncogenes, while hypermethylation of CGI at the promoter region leads to silencing of tumour suppressor genes. Taken from Roberston (2005).

1.1.3 CpG island methylator phenotype (CIMP)

Methylation of DNA mainly occurs in cytosine residues within dinucleotide CpG sites. CpG rich regions are called CpG islands (CGIs). The bulk of lone CpG are found in intronic and intergenic regions of the DNA sequence within the transposable elements and the repeat sequences (Wilson *et al.*, 2007).

The generally accepted definition of CGIs is a region in the human genome with high concentration of CpG dinucleotides (Juo *et al.,* 2014), where GC content is greater than 50% and an observed CpG to expected CpG ratio is greater than 60% (Gardiner-Garden and Frommer, 1987). These conspicuous islands overlap the promoter regions or are near transcription start site (TSS) of 60–70% of human genes particularly housekeeping genes and about 40% of tissue specific genes (Larsen *et al.*, 1992, Wang and Leung, 2004, Shen *et al.,* 2007).

As CGIs are frequently associated with the 5’ end of a gene, Antequera and Bird (1993) and Bird (2002) considered them as gene markers and used them to estimate the number of genes in the genome. Although the genome size and gene number of different vertebrates such as human, rat, and mouse are very similar, the distribution and number of CGIs among their genomes are varied (Gibbs *et al*., 2004, Han *et al*., 2008). The number of CGIs is approximately 27000, 19600, and 15500 within the genome of human, rat, and mouse, respectively.

About 60-90% of CpG dinucleotides are methylated in normal cells (Bogdanović and Veenstra, 2009), with the exception of CGIs (Ghoshal *et al*., 2005). Many studies reported that CGIs are differentially methylated in different tissues (tissue-specific differentially methylated regions, TDMs) and those tissues show differential gene expression

(Feltus *et al*., 2003, Bock *et al*., 2006, Previti *et al*., 2009). These findings suggest that methylation features could be related to differentiation of the cells and/or their specific function.

The CpG island methylator phenotype (CIMP), defined as widespread CpG island promoter methylation of a subset of genes, has received a lot of attention. It is predicted that the CIMP could be used as a biomarker for patients under medication (Zhang *et al*., 2011). CIMP is recognized as a common feature of human cancers (Weisenberger *et al*., 2006, Shinjo *et al.*, 2012). But why, where, and when this phenomenon occurs remains poorly understood.

The first study describing a CIMP was in colorectal cancer by Toyota and colleagues (1999), where they used bisulfite-PCR reactions and methylation-sensitive restriction enzyme digestion, followed by southern blotting. They found that colorectal cancer samples showed aberrant and extensive CpG island promoter methylation in several tumour-suppressor genes. Moreover, they showed that this CIMP was implicated in genetic instability. Silencing of the mismatch repair gene (MLH1) resulted in the silencing of hundreds of genes, including TSGs. Furthermore, they pointed out that CIMP-associated cancers had a distinct histology, epidemiology, and molecular features.

Since that time, extensive studies have supported the CIMP hypothesis and confirmed the original finding that CIMP is associated with promoter hypermethylation in many TSGs. Several studies observed that different types of cancers with varying phenotypes, such as glioblastomas, gastric cancer, pancreatic cancer, ovarian cancer and liver cancer shared common CIMP-associated genes (Issa *et al*., 2005, Teodoridis *et al*., 2008, Wu *et al*., 2010). For example, Herman *et al.* (1995) found suppression of the CDKN2/p16/MTS1 complex is

associated with aberrant DNA methylation in all common human cancers including breast cancer, prostate cancer, renal cancer and colon cancer.

In addition, a variety of TSGs such as MINT1/MINT2/ CDKN2/MLH1 were observed to be methylated at their promoter CpG islands in colorectal cancer, gastric cancer, pancreatic adenocarcinoma, and a range of human cancer cell lines (Teodoridis *et al*., 2008). Moreover, relationships between CIMP and various parameters, including microsatellite instability (MSI), rates of mutations for p53, BRAF, and KRAS, and tumour recurrence after transplantation, were identified (Samowitz *et al*., 2005, Ogino *et al*., 2009, Noushmehr *et al*., 2010, Wu *et al*., 2010).

However, a few studies did not support the CIMP. Anacleto *et al.* (2005) looked for CIMP in five genes (DAPK, MGMT, hMLH1, p16INK4a, and p14ARF) in primary colorectal cancers. That study is not compatible with the independent existence of CIMP and suggested that the association between methylation and colorectal cancers was indirect due to the correlation with MSI. Another study examined the promoter methylation of six genes (hMLH1, MGMT, $p16^{INK4A}$, $p14^{ARF}$, APC, and CDH1) in colorectal cancer and the corresponding normal tissue using methylation-specific PCR (Yamashita *et al*., 2003). They reported that their results did not support the CIMP hypothesis.

The existence of this phenotype remains controversial. Issa *et al.* (2005) pointed out that all studies that did not find the CIMP phenotype used unselected genes and non-appropriate statistical analysis methods. When Issa (2004) reanalysed the data in those papers with appropriate statistical analysis, they found a clustering of methylation of specific genes such as CDKN2A and MLH1, therefore suggesting that appropriate analysis is important to reveal a CIMP.

The fact that different cancers, such as acute lymphocytic leukaemia (Garcia-Manero *et al.*, 2002), acute myeloid leukaemia (Toyota *et al.,* 2001), gastric carcinomas (Kim *et al.,* 2003), liver cancer (Shen *et al.,* 2002), and ovarian cancer (Strathdee *et al.*, 2001), have significantly different rates of CIMP in TSGs implies that the occurrence of aberrant methylation of CGIs at promotor regions is not random.

The causes of CIMP are not well-understood and the genes that are implicated in this aberrant methylation have not generally been elucidated. However, there are other suggestions which need more studies. For example, mutations in DNA methyltransferases are considered as strong candidates for causing CIMP in cancer (Issa, 2004). Kanai *et al.* (2001) reported that overexpression of DNMT1 was significantly associated with CIMP in colorectal and stomach cancers.

1.2 Germ cell tumours

Germ cell tumours (GCTs) are malignant or benign tumours believed to be derived from primitive germ cells (Amatruda *et al*., 2013), a cell type in which extensive methylation reprogramming occurs. GCTs are the most common tumours in young men (Koul *et al*., 2002) and represent about 3% of all children's cancers (Echevarria *et al.*, 2008). The incidence of malignant testicular GCTs has increased over the past few decades (Huyghe *et al.,* 2003, Goedert *et al.,* 2007).

1.2.1 Origin of GCTs

GCTs represent a heterogeneous group of neoplasms that are believed to arise from undifferentiated primordial germ cells (PGCs) (Palmer *et al.,* 2010) that normally form gametes. During embryonic development, PGCs normally migrate to the gonads but it has been hypothesized that some may follow a midline path but instead of descending into the ovaries or testes, they may settle in extragonadal sites such as the abdomen, chest, head and pelvis. However, the origin of some types of germ cell tumours is still controversial (Scholz *et al.*, 2002, Scotting, 2006).

Germ cells normally migrate along the midline of the fetus before they finally settle into a place in reproductive organs where they eventually produce eggs in females and sperm in males. However, abnormal groupings of the germ cells can cluster together and form a tumour within and outside the testis and ovary. In adults, more than 90% of GCTs occur in the gonads, while the rest develop in extragonadal regions (Raddatz *et al*., 2013). The most common sites for GCTs outside of the reproductive tract are mediastinum, abdomen, sacrococcyx, and pelvis (Elzinga-Tinke *et al*., 2015), while some GCTs can also found in the central nervous system (Kersh *et al.,* 1988).

PGCs are pluripotent cells, which undergo three key processes during specification to oocytes or spermatozoa; repression of somatic genes such as the Hox family genes (Ohinata *et al.*, 2005), regulation of potential pluripotency genes, and genome wide epigenetic reprogramming (Yamaji *et al.*, 2008). Therefore, any disruption of these regulatory processes can result in abnormal PGC development which could lead to tumourigenesis.

In humans, PGCs arise from pluripotent epiblast cells in the embryonic yolk sac during week 3-4 after conception (Mamsen *et al.*, 2012), induced by signalling from extraembryonic tissues (Ohinata *et al.*, 2005)(Figure 1.3). The signals involved in the specification and migration of human PGCs are not fully understood. However, the specification of the PGCs in mice is initiated by expression of B-lymphocyte-induced maturation protein-1 (Blimp-1), also known as PRDM1, followed by several pluripotency markers such as Stella, Eras, c-Kit, Oct4, and Nanog (which also controls migration) (Hayashi *et al.*, 2007, Gu *et al*., 2009, Magnúsdóttir *et al.,* 2013). Expression of these pluripotency markers decreases in neighbouring somatic cells (Yabuta *et al.*, 2006).

PRDM14 also has a key role in the maintenance of pluripotency of PGCs, through repression of fibroblast growth factor (FGF) signalling, which is necessary for cellular differentiation (Grabole *et al.*, 2013). Yamaji *et al.* (2008) revealed a critical function of Prdm14 in establishment of the germ cell lineage in mice during the migration of PGCs from the dorsal mesentery of the hindgut along midline to the gonadal ridge, where PGCs undergo epigenetic reprogramming and proliferation to develop into mature germ cells.

Nettersheim *et al.* (2015) pointed out that undifferentiated GCTs share features with PGCs in the expression of pluripotency markers such as the transcription factors NANOG, OCT4 and SOX2 (West *et al.*, 2010). Therefore, studying the relationship between DNA methylation and the expression of these regulatory genes in GCTs could reveal more about the role of DNA methylation in tumourigenesis.

1.2.2 GCT classification

Human GCTs are classified into two groups according to where they occur anatomically. The majority of tumours are in a gonadal site in the testes or ovaries while the minority are in extragonadal sites such as the thorax (mediastinal tumours) and sacrococcygeal area. In adults, 95% of GCTs occur in gonads while the remaining five percent are in extragonadal tissues. While in children, 50% of GCTs is gonadal and another 50% is in extragonadal sites (Kucukoner *et al.*, 2012, Vasdev *et al.,* 2013). Regardless of location, GCTs are classified histologically into two major groups: seminomatous (also called germinomatous) and non-seminomatous (also called non-germinomatous). The latter group is subdivided into: differentiated teratomas (exhibiting a degree of 'embryonic' cellular phenotype), undifferentiated embryonal carcinomas (EC), extra-embryonic yolk sac tumours (YST) and choriocarcinomas (Oosterhuis and Looijenga, 2005). A germ cell tumour with a combination of the different histologies can be classified as a mixed-type GCT (Sesterhenn and Davis Jr, 2004).

GCTs can also be classified into five subtypes (I-V) (Figure 1.5), which incorporate multiple features (Rijlaarsdam *et al*., 2015). Each subtype has specific molecular, clinical and histopathological properties. In brief, type I (paediatric) clinically presents as a teratoma or YST which arise in children before the age of 5 years. Type I present in gonadal or extragnadal sites where extra-gonadal, sacral teratomas,

occur most frequently and are mostly benign. Type I have been believed to arise from PGCs at an early stage (Rijlaarsdam et al., 2015). Type II (adult) refers to testicular GCTs, described as seminomas or non-seminomas, and presents most frequently in the gonads where the diagnosis is most frequently in males and females between 25-35 years. The median age of clinical diagnosis of type II GCTs in males is higher than in female. The common precursor of this type is Carcinoma In Situ (CIS) (Oosterhuis and Looijenga, 2005). However, Type III, IV, and V originate from more differentiated germ cell progenitor. Type III (elderly males) is known as spermatocytic seminoma (SS), which arise after the age of 50 years and is generally benign, localised in the testis. This type is believed to originate from mature germ cells (spermatogonium or spermatocyte) and does not occur in females (Rijlaarsdam *et al*., 2015). Type IV arises only in female after the age of 50 years old and is hypothesized to originate most often from meiotic germ cells in the ovary that underwent maternal imprinting. The type V is believed to originate from the fertilization of an empty ovum by two sperm cells, resulting in completely paternally imprinted mature male germ cells. This type usually presents in the placenta or uterus (Oosterhuis and Looijenga, 2005).

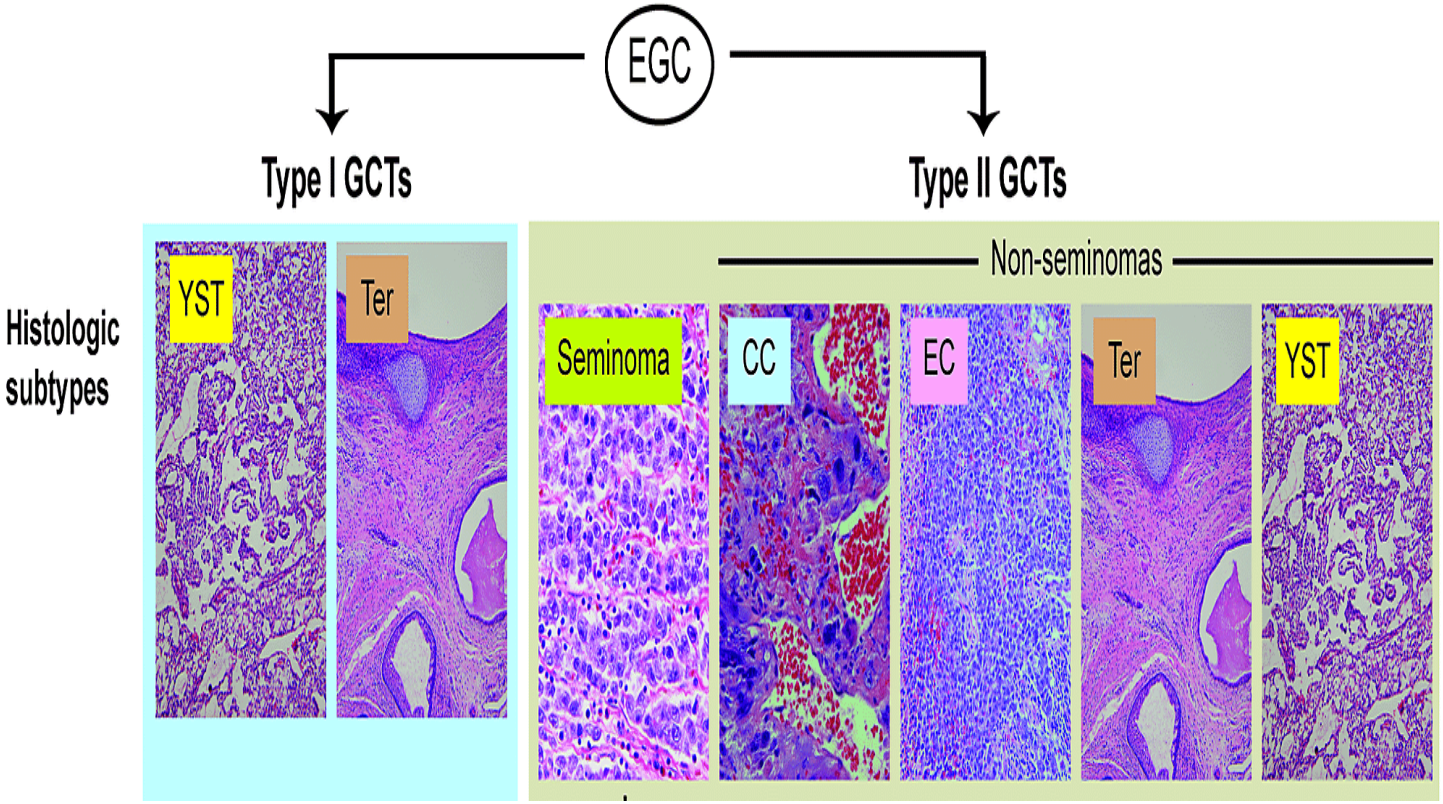


Figure 1.5. Classification of GCTs. These are divided by histological subtype: yolk sac tumour (YST), teratoma (Ter), seminoma, choriocarcinoma (CC), embryonal carcinoma (EC). Adapted from (Chen and Amatruda, 2013).

1.2.2.1 Seminomas

These are malignant tumours and can be further classified, according to location, into seminomas (testis), dysgerminomas (ovary), and extragonadal germinomas (brain). They are very sensitive to chemotherapy and radiotherapy, and are therefore highly curable (Boujelbene *et al*., 2011). The dysgerminomas and seminomas are not seen in children until they reach puberty, but they are common in young people aged between 15-44 years (Arora *et al*., 2012). A subtype of the seminomas, spermatocytic seminomas (SS), which arise exclusively in the testis, is usually found in the elderly and are not regarded as true GCTs as may do not arise from ITGCNU (Raiss *et al*., 2011).

1.2.2.2 Non-seminomas

Teratoma

These are the most common GCTs in fetuses (Frazier *et al*., 2012) and children (Curto *et al*., 2007). They may contain different types of tissues or organs such as skin or bone that are derived from the three germ layers, endoderm, mesoderm and ectoderm. Teratoma is usually a benign tumour but it can behave in a malignant fashion, especially in testes, depending on maturity and the other types of the cells involved (Ulbright, 2004). The sacrococcygeal (the tail bone of the distal end of the spinal cord) teratomas are the common GCTs found in children. Due to the fact that the sacrococcygeal tumours are sometimes visible from outside of the body, the diagnosis is usually made early and the treatment initiated early, making the prognosis of teratomas very favourable (Michael *et al*., 2000). In children, most ovarian GCTs are teratomas, followed by the YST, EC, and mixed-type tumours (Horton *et al*., 2007).

Embryonal carcinoma (EC)

The EC is an uncommon malignant GCT that most often occurs in testis, more rarely in the ovary and very unusually in neonates (Singh, 2002). It is usually mixed with other types of GCTs.

Endodermal sinus or yolk sac tumour (YST)

YST is often very aggressive and malignant when it is found in testes, ovary or sacrococcygeal regions (Yao *et al*., 2012). It can spread rapidly through the lymphatic system to any organ in the body revealing its aggressive nature (Wobbes *et al*., 1981).

Choriocarcinoma

This is a very rare malignant GCT that recapitulates the chorion layer of the placenta (namely, trophoblastic). Although rare, it is the most clinically dangerous GCT because it can grow and spread quickly to any part of the body (Worster *et al*., 2002). These cells can form a tumour in the placental cells during pregnancy, so they may spread to the infant and the mother. When the tumour develops during pregnancy, it is referred to as gestational choriocarcinoma. However, If young children develop choriocarcinoma from chorion cells that come from the placenta, this is termed non-gestational choriocarcinoma (Chen *et al*., 2003).

1.2.3 The development of GCTs

It is generally accepted that both testicular seminomas and non-seminomas GCTs originate from precursor lesions termed carcinoma in situ (CIS) (also called intratubular germ cell neoplasia unclassified (IGCNU)) (Skakkebaek *et al*., 1987, Looijenga *et al*., 2007). IGCNU is believed to arise from PGCs during embryogenesis by blocking or delaying maturation into gametes (Eckert *et al*., 2008, De Felici, 2013). The reason for this mis-maturation is not yet entirely understood. It is probable that both environmental factors and aberrant epigenetic regulation during early embryogenesis play a role in endocrinological imbalances and an excess of oestrogens, that stimulate PGCs to acquire the tumourigenic features of CIS cells (Sonne *et al*., 2008, Elzinga-Tinke *et al*., 2015). Pluripotent cells in CIS resemble PGCs in morphology and embryonal markers (both cells express OCT3/4, PLAP, c-KIT, NANOG, SOX17 and AP-2ɣ) (Dieckmann and Skakkebaek, 1999, Rajpert-De Meyts, 2006, Sonne *et al*., 2008). Moreover, several studies revealed that CIS is the precursor lesion of GCTs, based on the finding of development of CIS to invasive GCTs in patients within 5 years (von der Maase *et al*., 1986, Dieckmann and Skakkebaek, 1999) and the similarity in both CIS and GCTs in gene expression profiles (Summersgill *et al.*, 2001) , epigenetic profiles (Netto *et al*., 2008, Furukawa *et al*., 2009) and chromosomal gain of 12p (Looijenga *et al*., 2003). The arrest of PGC differentiation and aneuploidy of 12p are the first detected events in the development of CIS (Rosenberg et al., 2000, Oosterhuis et al., 2002, Ottesen et al., 2003). Therefore, it has been accepted that PGCs are the cells of origin for all GCTs regardless of their locations. The development of GCTs is summarised in figure 1.6.

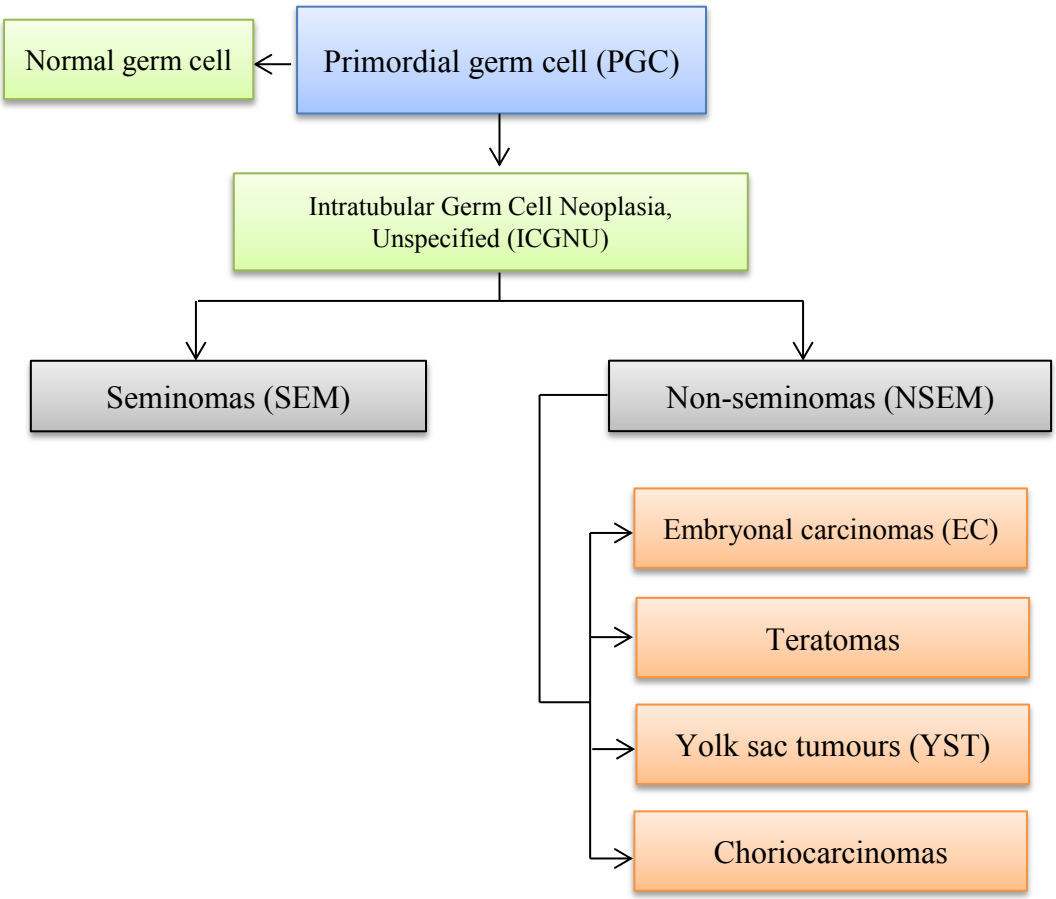


Figure 1.6: The development of testicular GCTs: Both seminomas and non-seminomas develop from intratubular germ cell neoplasia unclassified (IGCNU). Adapted from Cusack and Scotting (2013).

1.2.4 Treatment of germ cell tumours

In general, the plan for treatment of cancer depends on several factors, such as the type and location of tumour, the time of diagnosis of the tumour (newly diagnosed or recurrence after treatment), the stage of the tumour which is based on the extent of cancer cells spread and the age of the patient.

For GCTs, treatment may include standard treatments such as surgery, chemotherapy, and radiation (Gobel *et al*., 2000). Chemotherapy is used often as the first treatment if the tumour is large but not in gonadal site. Surgery is generally performed as standard for teratomas and residual tumours (Sugawara *et al*., 1999). Radiation therapy uses high-energy radiation from X-rays, gamma rays or the fast moving subatomic particles to target and destroy the cancer cells (Bamberg *et al*., 1999). In addition to the killing of the cancer cells, radiation therapy can destroy normal cells leading to physical side effects such as nausea, fatigue and hair loss.

Children with benign GCTs usually undergo surgery to remove the tumour. However, those with malignant tumours may need to receive further treatment after the staging process. The staging process is the classification system that helps doctors to determine the extent of cancer progression. These consider factors such as size of the tumour and the extent of tumour spread to other organs (Sturgeon, 2002).

1.2.5 Incidence of GCTs

The incidence of malignant GCTs is higher in males than females (Raddatz *et al*., 2013). This difference might be due to the dissimilarity of gametogenesis mechanisms between males and females. In general, non-seminomas occur more frequently in children while seminoma is

more common in younger adults and adolescences (Echevarria *et al*., 2008).

Furthermore, the incidence of the several histological subtypes of testicular GCTs have increased significantly and consistently over the past few years (Huyghe *et al*., 2003).

1.3 The role of DNA methylation in GCTs

The role of DNA methylation in GCT development and progression are not fully understood. Some investigations suggest that aberrant methylation is a key factor in the development of GCT subtypes (Smith-Sørensen *et al*., 2002, Honorio *et al*., 2003, Netto *et al*., 2008). Although the tumour subtypes (seminoma and non-seminoma) are presumed to originate from the same progenitor cells, they differ histologically, follow different progression pathways, and vary in molecular alterations. For example, seminomas are considered as hypomethylated and exhibit a relatively similar histology to the germ cell precursor (Netto *et al*., 2008), while non-seminomas are hypermethylated, tend to be more aggressive and resistant to chemotherapy and exhibit differentiation into forms resembling somatic tissues (teratomas) or extraembryonic structures such as yolk sac (yolk sac tumour) or placenta (choriocarcinoma) (Chen and Amatruda, 2013)

The progression of testicular GCTs from CIS to seminoma or non-seminomas (Hoei-Hansen *et al*., 2005, Eckert *et al*., 2008) is accompanied by profound changes in methylation levels (Almstrup *et al*., 2010). The difference in the DNA methylation levels during this progression could be an important key for understanding how these tumours occur. Identifying the methylated driver genes that play a role in this progression could lead to new powerful and specific epigenetic therapies.

Recent genome-wide studies of paediatric GCTs (Jeyapalan *et al*., 2011, Amatruda *et al*., 2013) revealed large differences in the level of methylation of many TSGs between GCT subtypes, where seminomas showed a lower methylation level than non-seminomas for those genes. The regions within these silenced genes that related to hypermethylation still needs more investigation.

Abnormal DNA methylation has been implicated in the aetiology of various classes of malignancy such as colon tumours and gliomas and can possibly be particularly applicable in GCTs because of the broad epigenetic reprogramming that happens in the germ line and early fetus during embryogenesis. Rijlaarsdam *et al.* (2015) suggested that the genome wide DNA methylation profile for GCTs subtypes provides clues to the origin and underlying developmental biology of this tumour.

Previous reports suggested a hypothesis that the methylation status of GCTs reflected the embryonic developmental phase of the PGC when the tumour emerged, with seminomas emerging from a hypomethylated PGC and non-seminomas beginning after methylation of PGCs. However, the hypomethylation status detected in IGCNU, which is accepted as the precursor of both testicular seminomas and non-seminomas, suggests that both seminomas and non-seminomas arise from a hypomethylated PGC (Amatruda *et al*., 2013).

Peltomaki *et al.* (1991) provided the first evidence of differences in methylation between seminoma and non-seminoma testicular GCTs. They inspected X-chromosomes by Southern blotting after methylation-sensitive restriction digestion. They found that specific gene promoters are hypermethylated in non-seminomas but not in seminomas. Moreover, non-seminomas generally demonstrated a hypermethylation status similar to that found in the inactive female X-chromosome and seminomas were normally hypomethylated, like the active X-

chromosome. Peltomaki *et al.* (1991) studied the methylation of the 5' region of two X-chromosome genes, hypoxanthine phosphoribosyltransferase (HPRT) and phosphoglycerate kinase (PGK), and demonstrated a contrasting pattern: seminomas were virtually unmethylated while on the contrary, non-seminomas were generally hypermethylated.

Smiraglia *et al.* (2007) utilized global restriction landmark genomic scanning (RLGS), which is a useful method used to identify aberrant CGIs hypermethylation in cancer and tissue specific methylation of CGIs, and revealed that seminomas contain much lower levels of 5mC (an average level of 0.08%) than non-seminomas (an average of 1.11%).

Others have inspected the CGIs of particular genes (non-X linked) and found that non-seminomas have a tendency to resemble numerous adult non-GCTs in that they frequently contain hypermethylation in these regions, while seminomas show hypomethylation of these genes (Koul *et al*., 2002, Okamoto and Kawakami, 2007).

In imprinting, which involves silencing of parental allele-specific methylation of particular CpG regions, many studies have demonstrated that postpubertal testicular GCTs show an absence of imprinting at both the H19 and IGF2 loci (Mishina *et al*., 1996, Looijenga *et al*., 1998, Sievers *et al*., 2005).

The study of Koul *et al.* (2002) on the two major histologies of GCT, seminomas and non-seminomas, identified distinct hypermethylation patterns in the promoter region of MGMT. They revealed that MGMT was silenced or downregulated in the majority of non-seminoma samples and in a wide variety of cancers (Esteller *et al*., 2000).

Many studies revealed that striking feature of all invasive TGCTs is gaining of the short arm of chromosome 12 (12p) and they suggested that (12p) contains relevant genes of which their overexpression is crucial for the development of this cancer (Mostert *et al.*, 1998, Rosenberg *et al.*, 2000, Looijenga *et al.*, 2003). Rodriguez *et al.* (2003) analyzed (12p) in TGCTs samples using a global approach to expression profiling targeting chromosomes for identifying particularly genes in this region that amplified and the result pointed that highly overexpression of BCAT1 was specific to non-seminomas and upregulated other genes such as CMAS, EKI1, KRAS2, SURB7, therefor this study represented BCAT1 as a candidate gene for TGCT development.

Furthermore, several methylation studies applied in testicular GCTs showed that YST had higher number of hypermethylated promoter genes than seminomatous GCTs. Nine genes that were identified (APC, RASSF1A, HOXA9, XPA1, EMX2, MSX1, NME2, SCGB3A1 and HIC1) were hypermethylated in all YST samples analysed in those studies (Koul *et al*., 2002, Lind *et al*., 2007, Furukawa *et al*., 2009). BCAT1 is a target gene for CMYC, which is involved in cell proliferation, differentiation and apoptosis, and plays an important role in many tumours.

A recent study on malignant GCTs in children showed a CpG methylator phenotype in YST including silencing of genes associated with Caspase-8-dependent apoptosis and correlated with increased expression of DNMT3B (Jeyapalan *et al.*, 2011). This finding suggested that DNMT3B overexpression could be a strong candidate to cause the methylator phenotype in YSTs.

Ushida *et al.* (2012) reported that despite non-seminomas exhibiting higher methylation than seminomas, both show significant hypomethylation of LINE-1 repeat elements compared to normal cells or other tumours of somatic tissue origin. This study supports the previous

suggestion for using serum LINE-1 hypomethylation as a potential prognostic marker for cancer (Estecio *et al*., 2007, Tangkijvanich *et al*., 2007, Jeyapalan *et al*., 2011). Tangkijvanich *et al*. (2007) revealed that a high level of LINE-1 hypomethylation was significantly associated with tumor progression using measurement of COBRA LINE-1 in the serum of hepatocellular carcinoma patients.

Recent studies revealed the important role of PRDM14, not only in demethylation activity of germ cell development (Yamaji *et al*., 2008, Okashita *et al*., 2014) but also in PGC specification (Magnúsdóttir *et al*., 2013) and maintenance of pluripotency of PGCs by suppressing DNA methylation and fibroblast growth factor (FGF) signalling which led to suppression of cellular differentiation (Grabole *et al*., 2013, Fan *et al*., 2015).

Studies of PRDM14 showed its fundamental role in repressing the expression of the methyltransferase family members, especially DNMT3B and DNMT1 (Magnúsdóttir *et al*., 2013). Moreover, PRDM14 promotes DNA demethylation through interaction with Ten-eleven translocation (TET) proteins (Okashita *et al*., 2014). Additionally, Hu *et al.* (2005) found that higher levels of PRDM14 expression correlated with decreasing methylation levels in breast cancer. PRDM14 may also regulate pluripotency factors such as NANOG and KLF4 (Grabole *et al*., 2013).

1.4 Recent global DNA methylation studies in GCTs

Two recent studies analysed global DNA methylation in GCT cell lines and provided insight into the aetiology and underlying developmental biology of GCTs. First, van der Zwan *et al.* (2014) studied the interaction between gene expression and CpG methylation but only in seminoma versus EC cell lines. Second, Rijlaarsdam *et al.* (2015) identified specific and global methylation differences between GCT subtypes, although they did not assess the relationship between methylation and gene expression.

In this study, the Illumina infinium HumanMethylome450 bead chip system and Affymetrix expression arrays were used for four GCT subtypes in order to gain a comprehensive view of the correlation between methylation and gene expression. In addition, to confirm that key genes were silenced by methylation, their expression was determined after treating the hypermethylated cells lines with the demethylating agent, 5- azacytidine (5-aza). The data were compared to gene expression in a cohort of primary GCT samples to identify how many genes differentially expressed in YST versus seminoma were common in primary tumour and cell lines. Through this approach, my study was able to determine whether silencing of methylated genes was a general phenomenon in this type of tumour.

1.5 Project Aims

The overall aim of this study is to assess the relation between DNA methylation and gene expression in GCTs, in order to determine the role of DNA methylation in GCT development through regulation of gene expression and to identify the methylated genes that most likely play a role in tumours biology.

The specific aims were:

- To investigate the global differences in methylation between seminoma and non-seminoma cell lines and detect the genomic targets that are methylated using whole genome approaches.
- To assess the correlation between methylation and gene expression.
- To identify dysregulated genes that might be silenced by methylation in GCT cell lines.
- To identify which genes were most likely to contribute towards the phenotypic differences between hypo- and hypermethylated cell lines.
- To investigate which candidate gene (s) could play a role in tumour pathology, or be a promising target for therapeutic applications.

Chapter 2. Materials and methods

2.1 Materials

2.1.1 Equipment

2100 Bioanalyzer	Agilent technologies, Uk
Autoclave	Thermo Fisher Scientific, Uk
C1000 thermal cycler (qPCR)	Bio-Rad, UK
Centrifuge 5804	Eppendorf, UK
CO2 Incubators	LabX, Canada
Dry heat block	Labnet international Inc,UK
Electroporation (ECM830)	BTX, US
Geldoc Gel Light Imaging System	Bio-Rad, UK
Gel electrophoresis equipment	Appleton woods, UK
Haemocytometer	Weber Scientific, UK
Laminar air flow cabinet	MACH-AIRE, UK
Microscope with Camera Eclipse	Nikon, Japan
Minicentrifuge	Labnet, UK
NanoDrop spectrophotometer	Thermo Scientific, UK
PCR-Thermal cycler TC-3000	TECHNE, UK
PCR-Thermal cycler GS1	G-Storm, UK

PH-Meter GP353	EDT instrument, UK
Pipette autoclaved (0.5-1000 µl)	Nichiryo, Japan
Power Pac mini-protean tetracell	Bio-Rad, UK
Refrigerated Centrifuge Hawk15/05	MSE, UK

2.1.2 Consumables

6,12,24,96 well plates	Corning incorporated, US
40 µM cell strainer	MWG, UK
96-well PCR plates (0.2 ml)	Thermo scientific, UK
Adhesive PCR sealing sheets	Thermo scientific, UK
Cell culture vented cap flasks	Corning incorporated, US
Coverslips (22mm)	VWR, US
Cryotube vials	Thermo Scientific, UK
Electroporation cuvettes	Flowgen Bioscience, UK
Eppendorf (0.5, 1.5 ml)	Star lab, UK
Hybond Extra Nitrocellulose roll	Amersham Biosciences, uk
Microscopic Superfrost Slides (1MM)	VWR, US
Parafilm M (4" X 125')	Pechiney, Chicago
Pipette tips DNase /RNase free	Star lab, UK
Serological Pipette (1, 10, 25 ml)	Greiner bio-one, UK

2.1.3 Reagents and Standards

2-Mercaptoethanol	Sigma, UK
Ammonium persulfate (APS)	Sigma, UK
Chloroform	Fisher Scientific, UK
Crystal violet dye	Sigma, UK
Dimethyl sulfoxide (DMSO)	Sigma, UK
Ethanol	Sigma, UK
Hi-Res standard agarose	Geneflow, UK
Isopropanol	Fisher Scientific, UK
Methanol	Sigma, UK
Mounting medium with DAPI	Vector, UK
Oligo(dT)18 Primer	Thermo Scientific, UK
Paraformaldehyde (PFA)	Sigma, UK
Phosphate buffered saline (PBS)	Sigma, UK
Proteinase K	Sigma, UK
Random primer	Promega, UK
RedSafe DNA stain	Chembio, UK
Sequencing grade solution dNTP's	Amersham, Biotech Inc, USA
Tri reagent	Sigma, UK

Tris base	Invitrogen, UK
Triton x-100	Sigma, UK
Trypsin-EDTA	Sigma, UK
Trypan blue solution (0.4%)	Sigma, UK
Tween 20	Sigma, UK

2.1.4 Chemicals

Cisplatin (cis-diamminodichloroplatinumII)	Accord, UK
Decitabine (5-aza-2'deoxycytidine)	Sigma-Aldrich, UK

2.1.5 Marker and ladder

DNA ladder (1Kb, 100bp)	NEB, UK
Protein ladder (SeeBluePlus2Prestained)	Life Technologies,UK

2.1.6 Enzymes and Kits

DNasa1, Amplification grade	Invitrogen, UK
GenElute PCR clean-up kit	Sigma, UK
GenElute plasmid miniprep kit	Sigma, Uk
Phusion High-Fidelity PCR Master mix	BioLab, UK
QIA quick Gel extraction kit	Qiagen, UK
Restriction endonucleases	NEB, UK
RNeasy mini kit	Qiagen, UK
QIA Shredder	Qiagen, UK

SensiFAST SYBR No-Rox kit	Bioline, Uk
SuperScript III Reverse Transcriptase	Invitrogen, UK
T4 ligase	Invitrogen, UK

2.1.7 Antibiotics

Ampicillin	Sigma-Aldrich, UK
G418 (Geneticin)	Sigma-Aldrich, UK
Kanamycin	Sigma-Aldrich, UK
Penicillin/streptomycin	Sigma-Aldrich, UK

2.1.8 Antibodies

-Primary antibodies

Mouse Anti 5-methylcytidine	AbD serotec, Bio Rad, UK
Myc-Tag (9B11) Mouse mAb	Cell Signaling Technology,UK
Anti-PRDM14 antibody	Abcam, UK

-Secondary antibodies

Alexa Fluor®488 Goat anti-Rabbit IgG	Thermo Scientific, UK
Donkey anti-Mouse IgG, IRDye®800CW	LI-COR, USA
TEXAS RED® anti-Mouse IgG	VECTOR, USA

2.1.9 Plasmids

pcDNA 3.1	Invitrogen, UK
pEGFP GFP-C1	Clontech Laboratories, UK

2.1.10 *E. coli* strain

Competent cells (E.coli Laboratory strains, alpha selected silver efficiency Bioline, UK).

2.1.11 Cell lines

Human testicular germ cell tumour (TGCT) cell lines analysed in this study were purchased from the Institute for Cancer Research, Sutton, London.

TCAM2	Seminoma (SEM)
GCT44	Non-seminoma yolk sac tumour (YST)
GCT27	Non-seminoma embryonal carcinoma (EC)
NT2D1	Non-seminoma teratocarcinoma (TERT)

2.1.12 Database and software

The following data bases were used to annotate or identify gene and protein features as pointed:

Promoter prediction	Eukaryotic Promoter Database	http://epd.vital-it.ch
CpG sites and islands	MethPrimer	http://www.urogene.org/methprimer/
Primer design	Primer3web version 4.0.0	http://primer3.ut.ee/
BLAT Search Genome	UCSC Genome Browser NCBI Build 36.1 Feb 2009(GRCh37/hg19)	http://www.genome.ucsc.edu/cgi-bin/hgBlat
Biological process & molecular function	Atlas of genetics and cytogenetics in oncology and haematology	www.atlasgeneticsoncology.org/index.html www.genecards.org www.geneontology.org
Venn diagram	GeneVenn software developed at the University of Kent	www.cs.kent.ac.uk/people/staff/pjr/EulerVennCircles/EulerVennApplet.html
Gene Sequence for designing primer	CDS to get cDNA sequence	http://www.ncbi.nlm.nih.gov/gene/
Nucleotide BLAST	Standard Nucleotide BLAST	http://blast.ncbi.nlm.nih.gov/blast.cgi
Check Restriction Enzymes cutting	NEBcutter V2.0	http://nc2.neb.com/NEBcutter2/
Convert DNA sequence into its reverse	Reverse Complement	www.bioinformatics.org/sms/rev_comp.html

2.1.13 Primers

Gene	Forward (5 → 3)	Reverse (5 → 3)	Product size (bp)
PCR Primers			
Beta-actin	TCTACAATGACCTGCGTGTG	ATCTCCTTCTGCATCCTGTC	684
CTHRC1	AAACTGGAAATGAACGGCCC	TGTGAAATACCAACGCTGACA	381
EPCAM	CTGGCCGTAAACTGCTTTGT	TCCAGATCCAGTTGTTCCCC	629
OCT4	TCCCTTCGCAAGCCCTCAT	TGATGTCCTGGGACTCCTC	474
Sox17	GCGCCCTTCACGTGTACTA	TAATATACCGCGGAGCTGGC	351
RT-qPCR Primers			
BST1	GCTGGTGTCGACAGAAAAAT	TTTTCCAAAAGGAATCCACA	96
CTHRC1	ACACCCAACTACAAGCAGTG	GCATTTTAGCCGAAGTGAGC	141
CYCLO-PHILIN-B	AACAGCAAATTCCATCGTGT	TCACCGTAGATGCTCTTTCC	96
DDX43	CCAGACTGGAACAGGAAAGA	AATTCCCGAGTGGGAGTTAG	126
EPCAM	CAAAACTTGGGAGAAGAGCA	CACTGCTTGGCCTTAAAGAG	103
GGCT	AGGGATAGCCACCATTTTTC	AAGGGTGATTTCCCTGATTC	140
HIST1H4C	TCCAGGGCATTACAAAACCG	CCTTAAGCACACCTCGAGTC	102
HPRT	AATTCTTTGCTGACCTGCTG	TATGTCCCCTGTTGACTGGT	136
KLF4	ACGGCTGTGGATGGAAATTC	TTCATGTGTAAGGCGAGGTG	136
L1TD1	TGGGGACTCTGGGAAGAAAA	GTGCCAGTTCCTTGACTTGT	99
Ly75	TGATCAGGCATTGCACAGTT	TTCCCTCATCAGTCTGCTCA	124
MGMT	AGGTGTTATGGAAGCTGCTG	GACAGGATTGCCTCTCATTG	128
PON3	TTACCAACTCCCTCCTGTCA	CCATTGGCACTACAAAATCC	118
PRDM14	TTTATCGCCAAAGGAGTC	GTACCTCCTTTTCCATCT	146
RBMXL2	CTCCTGATGAAAGACCGAGA	GACTTGCCGTTCATGTCTCT	107
RPRM	TAGCCTGTACATAATGCGCG	ACGAGGAAGTTGATCATGCC	129
SOX15	TTCCCACTGCAAACTGGAAG	TAGTGGGTATAGGTGGGCAG	93
SOX17	ATGACTCCGGTGTGAATCTC	TCACACGTCAGGATAGTTGC	101
TDRD12	CATTGACTTCTGCCGAGACA	TCCACAGCACATAATCTGGC	126
TRIL	TGGAGAGTCTAGTCAAGCTG	ACTCCAGATGTAGGTAGAGCA	102
TRIM95	CAAGGCGATAAGGAAGCTGT	CAACATCACAGAGAGCCGTT	97
Cloning primers			
PRDM14	AGTAGAATTCATGGCTCTACCCCGGCCA	TGATCTCGAGGTGAGTCATTGTGCCTGGC	1862

2.2 Methods

2.2.1 Cell culture

2.2.1.1 Preparation of growth media

Seminoma cells were cultured in RPMI-1640 medium while non-seminoma cells were grown in DMEM (Dulbecco's Modified Eagle Medium). Both media were routinely supplemented with 10% FBS (Fetal Bovine Serum) and 1% P/S (Penicillin/streptomycin). Cells were incubated in a humidified 5% CO_2 incubator at 37 °C and passaged 1:2 or 1:4 in fresh medium when they reached around 80% confluence. All media, supplements and antibiotics were purchased from Sigma-Aldrich, UK.

2.2.1.2 Recovering cells from frozen cell stocks

Cells were thawed from frozen cryotube stocks by defrosting them in a 37 °C water bath for about 10s then transferring them quickly to a centrifuge tube containing 10 ml of medium. After that, cells were subjected to centrifugation at 100 x *g* for 5 min, supernatant was removed and the cell pellet was resuspended in 1ml of fresh medium then transferred to a culture flask containing pre-warmed complete medium.

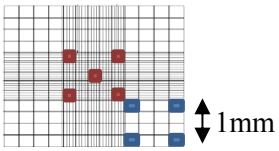
2.2.1.3 Maintenance and passage of cells

Media were changed every 2-3 days. Once the cells were 80% confluent, they were passaged in fresh media in a ratio of 1:2 or 1:4 depending on the growth rate. To split cells, media were removed and cells were washed with 2 ml of 1xPBS (phosphate buffered saline) followed by 2-4 ml of Trypsin-EDTA to detach cells from the flask surface, cells were left for 4-5 minutes or until they were dissociated. Medium was added and subjected to centrifugation at 100×*g* for 5 min.

Supernatant was removed and cell pellets were resuspended in 1 ml of medium. For continuation of cell line, cells were transferred to new culture flasks with pre-warmed media. For biomolecular experiments the resuspended cells were counted and harvested for further analysis or stored at -80°C.

2.2.1.4 Cell counting

To determine the number of viable cells, after resuspending the cells in 1 ml of medium, 10 µl of cell suspension was mixed with 10µl of Trypan blue stain, then transferred to a haemocytometer. Under the microscope, cells were counted in the 5 red squares (as in the figure below). Cells on the bottom line and left hand line of the square should be included- the average was taken then multiplied by 25 to give the number of cells present in the 1mm² area. In addition, cells in the 4 blue squares were counted and the average was taken and multiplied by 16 to give the number of cells in the 1mm² area. The average of these two numbers (in two areas) was taken and multiplied by 10^4 to give the number of cells in 1ml of culture.



2.2.1.5 Storage of cells

To store cells, cells were detached using Trypsin-EDTA then subjected to centrifugation at 100× g for 5 min. Pellets of approximately 1×10^6 cells were resuspended in a cold mixture of 6ml FBS and 10% DMSO (dimethyl sulfoxide) then stored in cryotubes at -80 °C then after 24 hours the cryotubes could be transferred to liquid nitrogen for longer storage.

2.2.1.6 Cell staining and fixation

In a 6-well plate, medium was removed then cells were fixed by using 1ml/well of 4% PFA (paraformaladehyde) then incubating at room

temperature for 20 minutes. After that, PFA was removed and 0.5 ml/well of 0.1% crystal violet was used for 5 minutes to stain the cells. Wells were washed with distilled water then left to dry at room temperature.

2.2.2 Cell treatments

2.2.2.1 Decitabine (5-aza-2'deoxycytidine (5-aza))

For the cytotoxicity assay, cells were seeded at a density of $2x10^4$ cells/well in 6-well plates. Next day, cells were exposed to different concentrations of 5-aza ranging from 0.1 to 20 μM. After 24h, the viable cells in each well were counted using a haemocytometer.

For DNA demethylation and re-expression studies, cells were cultured in 6-well plates, and when they reached approximately 30-50% confluence, they were treated with a range of 0.1-20 μM 5-aza. Media supplemented with 5-aza were changed for 1-3 days. Media were replaced subsequently with a drug-free media then cells were harvested in a time course assay and stored for later nucleic acid extraction. Untreated control cells were also prepared.

2.2.2.2 Cisplatin (cis-diamminedichloroplatinumII)

Cells which had been transfected with a gene of interest were subsequently exposed to cisplatin at a range of 5 to 30 μm for 2 h then the medium was refreshed with drug-free medium. For each treatment, three independent biological and technical replicate experiments were performed.

2.2.2.3 Generating an antibiotic (G418) kill curve

To determine the optimal concentration of antibiotic for selecting stable transfected cell lines, cells were plated in 6-well plates at the

density of 2 × 10^4 per well in complete growth medium. When cells reached 60-70% confluence, media was removed and replaced with a medium containing antibiotic (G418) with different concentrations (from 0 to 100µg/ml) while one well was left without treatment as a control. The medium containing antibiotic was replaced every three days for one week. The culture was examined every day for signs of visual toxicity using light microscope to assess the appearance of cell dead comparing to untreated cell. The`low´dose was considered when the minimal visual toxicity was apparent even after seven days of treatment while a dose was considered `high´ when the visual toxicity was apparent within 2-3 days of treatment and all cells were dead before the seventh days. The optimal dose was selected when all cells were dead after one week of treatment.

2.2.2.4 Clonogenic Survival Assay

A clonogenic assay was used to assess the response of cells to treatment with chemotherapy, by determining the ability of a single cell to grow and form a colony after treatment (the definition of a colony is one consisting of approximately at least 50 cells). Once cells had become confluent in a T75 flask, they were harvested using trypsin-EDTA then resuspended with 3 ml medium. Cells were passed through a 40 µM cell strainer to ensure the formation of a single cell suspension. To plate 100 cells/well in 6-well plate, 10 µl of cells were taken then mixed with 10µl of Trypan blue then the number of cells in 1ml was determined using a haemocytometer (as in section 2.2.1.4) then calculation was done using the formula:

$$C1V1 = C2V2$$

$$A^* \times V1 = (300 \text{ cells}/ 3\text{ml}) \times 60\text{ml}$$

$$V1 = 33.3 \text{ cells/ml} \times 60\text{ml}$$

$$= (0.1998/ A) \text{ ml}$$

= (199.8/A) µl

A*: the number of cells in 1 ml

The cells were transferred to a T75 flask containing 60 ml medium. The cells were mixed thoroughly before transferring 3 ml into each well of a 6-well plate. Plates were incubated at 37 °C overnight to allow cells to attach. Next day, cisplatin was added at a range of 2 to 30 µM for 2h, medium was changed with drug-free completed medium. Cells were incubated for 1-2 weeks until colonies containing approximately 50 cells were formed.

When colonies were visible, fixation and staining was carried out (as in section 2.2.1.6). Under the microscope, the number of colonies that survived after the treatment were counted (where a colony was defined as with at least 50 cells) then the following equation was applied to investigate any improvement in response towards chemotherapy.

$$\text{Surviving fraction (SF)} = \frac{\text{Number of colonies formed after treatment}}{\text{Number of cells seeded x PE}}$$

Where

$$\text{Plating efficiency (PE)} = \frac{\text{Number of colonies formed}}{\text{Number of cells seeded}} \text{ X } 100$$

2.2.3 General molecular biological methods

2.2.3.1 Genomic DNA Extraction

300µl lysis buffer (Tris-EDTA buffer) and 3.3µl Proteinase K (10µg/µl) were added to cell pellets and mixed. Samples were incubated at 37 °C for 1 hour. Samples were subjected to centrifugation at 13,000 x *g* for 5 min and supernatants were transferred to clean Eppendorf tubes. 300µl isopropanol was added to samples and mixed. DNA precipitates were transferred to new tubes and washed with 500µl of 70% ethanol then subjected to centrifugation at 13,000 x *g* for 5 min. Ethanol was removed and left to dry at room temperature. Finally, 20-40µl of distilled water was added to resuspend the DNA which was stored at -20 °C.

2.2.3.2 RNA Extraction

Cell pellets ($2X10^6$ cells) were resuspended in 1ml TRI reagent then incubated at room temperature (RT) for 5 minutes. 200µl chloroform was added to this mixture, the tube was inverted at least three times till the mixture become cloudy then incubated at RT for 10 min. To separate the layers, the mixture was subjected to centrifugation at 13,000 x *g* for 15 min at 4 °C. The top aqueous layer was transferred into a new ml tube containing 300µl isopropanol. The tube was inverted five times then incubated at RT for 5 min. To pellet RNA, the solution was subjected to centrifugation at 13,000 x *g* for 15 min at 4 °C. After discarding the supernatant, 300µl of 70% ethanol was added to the pellet without mixing, then subjected to centrifugation at 13,000 x *g* for 5 min at 4 °C. Ethanol was removed and the pellet was left to dry for five minutes at room temperature. The pellet was resuspended in 20-40µl sterile deionized water then stored at -80°C or treated with DNAase to remove DNA.

For RT-qPCR experiments, total RNA was isolated using the RNeasy mini kit and QIA Shredder according to the manufacturer's instructions. RNA concentration was determined with the NanoDrop spectrophotometer. The quality of RNA was assessed by using the Agilent Bioanalyzer according to the manufacturer's instructions.

2.2.3.3 DNAase treatment

Contaminating DNA was removed from RNA by adding 2µl of 10× DNAase reaction buffer and 2.3µl of DNAaseI to 20µl of RNA sample then incubating at RT for 10 min. 1µl EDTA was then added and the tube was incubated at 65 °C for 10 min then on ice for one min. RNA concentration was measured using Nano drop then stored immediately in -80 °C or converted to cDNA.

2.2.3.4 cDNA synthesis

To convert RNA to cDNA, a mixture of 0.5µl of each of Random primer, oligo$(dT)_{18}$ primer, dNTP (10µM), and ddH2O were added to 1-2µg of RNA then sterile water was added up to 14µl. The mixture was incubated at 65 °C for 5 min then on ice for 1 min. 4µl of 5×First strand buffer and 1µl of 0.1M DTT were added to the mixture. 1µl of superscriptIII was added to all samples except No-RT samples. Samples were incubated in a Techne thermal cycler, at 25 °C for 5min, 50 °C for 45min and 70 °C for 15min. cDNA was stored at -20 °C.

2.2.3.5 Measuring DNA and RNA concentration

DNA and RNA concentrations were measured using a Nanodrop where the 260/280 nm reading should be 1.8-2.2 for pure DNA and 1.7-2.0 for pure RNA.

The quality and concentration of RNA for microarray experiments was further analysed using an Agilent 2100 Bioanalyzer which gives measurements for the RNA integrity number (RIN) which should be between 8-10.

2.2.4 Gene expression analysis

2.2.4.1 Gene expression microarray

After extracting RNA and determining the concentration and quality of RNA samples, approximately 100 ng/µl of RNA sample was sent to Nottingham Arabidopsis Stock Centre at Sutton Bonington campus at the University of Nottingham to perform a gene expression assay using Affymetrix HumanGeneChip U133 Plus 2.0 arrays. Raw data reflecting the fluorescence intensities were statistically analysed to determine fold changes as compared to control samples.

2.2.4.2 Primer design

Primers were designed using the NCBI Genome Browser to get the sequence of the interested gene. Then, the sequence was inputted into Primer3 Software to set length and sequence of primer, PCR product size, and relative annealing temperature (Tm). The specificity of choosing primer pairs was checked by using BLAST function (blast nucleotide) of the NCBI Genome Browser Software. Optimising the annealing temperature of primers was achieved by using gradient function of PCR G-Storm machine testing across a range of Tm to

determine the optimum annealing temperature where non-specific bands and insufficient primer-template hybridization were not seen.

2.2.4.3 Polymerase chain reaction (PCR)

PCR was generally carried out in 0.2ml PCR tubes with a reaction volume of 10-50µl in a PCR-Thermal cycler TC-3000. A quick spin was applied to the mixture in the PCR reaction tube. Phusion High-Fidelity PCR master mix with HF buffer was used with the following reaction conditions:

Component	10µl Reaction	25µl Reaction	Final concentration
10µM Forward Primer	0.25 µl	0.5 µl	0.5 µM
10µM Reverse Primer	0.25 µl	0.5 µl	0.5 µM
Phusion master mix	5 µl	12.5 µl	1x
Template DNA	100ng	100ng	<250ng
Nuclease-free water	4.5 µl	11.5 µl	

General thermocycling conditions for settings a PCR are summarised as:

Step		**Temp**	**Time**
Initial denaturation		98 °C	30 sec
25-40 Cycles	Melting	95 °C	1 min
	Annealing	56-72 °C	1 min
	Extension	72 °C	30 sec/kb
Final extension		72 °C	5 min
Hold		4 °C	

2.2.4.4 Gradient annealing temperature optimization

A gradient PCR reaction was performed on all primer pairs to test the best annealing temperature. In each tube, 10 µl mixture was made as follow: 4.5 µl of sterile water, 0.25 µl of forward primer (10 µM),

0.25 µl of reverse primer (10 µM) and 5 µl of Phusion High-Fidelity PCR master mix with HF buffer. Then 1 µl of cDNA template was added to each tube. A gradient PCR reaction was designed as follows: initial denaturation at 95 °C for 1 minute, and then 40 cycles at 95 °C for 1 min, annealing (a gradient) between the lowest and highest temperature that would be appropriate for the primers (for example between 54-72 °C) for 1min, and extension at 72 °C for 30 seconds and final extension at 72 °C for 5 minutes.

2.2.4.5 Reverse transcriptase qPCR (RT-qPCR)

RT-qPCR was used to quantify cDNA after reverse transcription. The relative amount of amplified cDNA is measured through the emission of SYBR green dye fluorescence when it is intercalated into double-stranded DNA. qPCR was carried out using a C1000 thermal cycler with SensiFast SYBR No-ROX kit. Primers were designed using the Primer3 web site and choosing optimal primer pairs according to the following rules:

-Sequence: complementary sequences within and between primers, mismatches, and T base at the 3'end of the primer should be avoided.

-% GC content: should be between 40-50%

-Length of PCR product: should be less than 150 bp

-Length of primers: should be between 18-30 bp

-Melting temperature: should be around 56-60 °C

A master mixes for each primer pair was prepared by mixing 11µl of nuclease-free water, 0.25µl of each primer (10µM), and 12.5µl of 2x SensiFast SYBR No-ROX Mix. 24µL of master mix was placed into each well of 96-well plate then 100 ng of cDNA, no-RT sample, and NTC (non template control) were added to assigned wells. After that, the plate was placed in C1000 thermal cycler machine and the program was set to PCR cycling condition as follows:

cycles	Temp.	Time	Notes
1	95 °C	2 min	Polymerase activation
40	95 °C	5 sec	Denaturation
	58 °C	10 sec	Annealing
	72 °C	20 sec	Extension

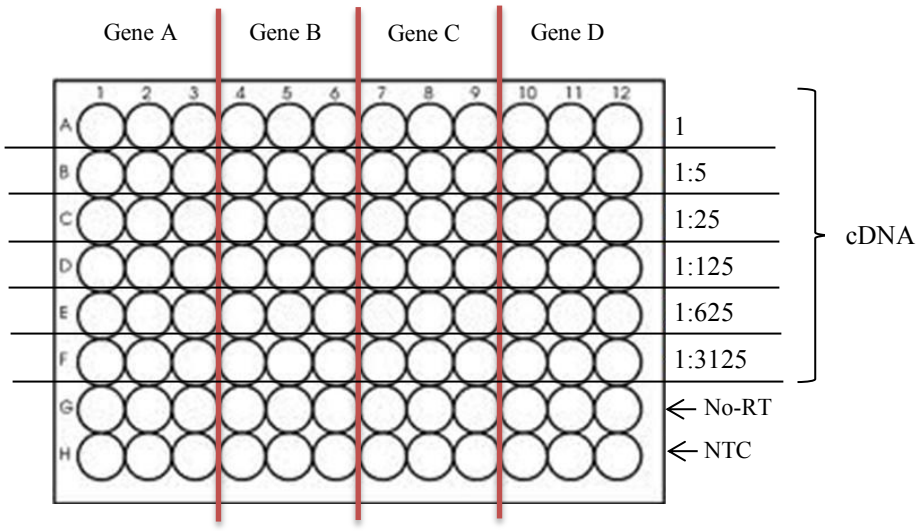


Figure 2.1: The design of standard curve plate to assess primer efficiency: Serial dilutions of cDNA were carried out to assess the efficiency of each primer. No-RT (No reverse transcriptase) and NTC (No template control) were used as negative controls. The sample was applied in three triplicate for each gene.

The data was analysed using C1000 thermal cycler software and nonspecific amplification was detected by melting-curve analysis. All qPCR results were imported into Microsoft Excel 2010. To determine the standard threshold and amplification efficiency of each set of primers, a standard curve was generated for each primer pair by amplifying six serial dilutions of template cDNA (1, 1:5, 1:25, 1:125, 1:625, 1:3125). To confirm the accuracy and reproducibility of qPCR results, three independent experimental runs were performed for each gene. To accept the qPCR results, the efficiency of the primer was required to be 90-110%, the R2 of the standard curve was closed to 1, and the slope of the curve was between -3.2 and -3.6. The single melt curve peak of each primer sets was important for assay specificity validation.

The threshold cycle (C_t) represented the cycle number at which the fluorescent signal crosses the threshold of the background level. C_t level was inversely proportional and correlated with the amount of the initial template (cDNA) concentration. The threshold cycle (C_t) value of each sample was calculated compared to the housekeeping gene (ß-actin) using the Pfaffl equation (Pfaffl, 2001). The Pffafl equation was used to determine the relative amount of RNA of the target gene in different samples.

$$\text{Ratio} = \frac{(E_{\text{target}})^{\Delta CP_{\text{taregt}}\ (control\text{-}sample)}}{(E_{\text{ref}})^{\Delta CP_{\text{ref}}\ (control\text{-}sample)}}$$

In the above equation, ratio refers to the expression ratio between the sample and calibrator. ΔCP_{target} (calibrator-test) = (C_t of the target gene in the calibrator - C_t of the target gene in the test sample. ΔCP_{ref} (calibrator-test) = (C_t of the reference gene in the calibrator - C_t of the reference gene in the test sample).

E_{target} and E_{ref} are equal to 1+E where E is the efficiency value in a standard curve of exponential amplification of target and reference genes, respectively. Target means gene of interest while reference means control gene (housekeeping gene).

2.2.5 Molecular cloning and transformation

2.2.5.1 Restriction enzyme digestion

To clone insert cDNA into plasmid, both were digested with the appropriate restriction enzymes. The restriction sites were introduced flanking a gene of interest sequence by adding the appropriate recognition sequences to the primer sequences. The digestion was carried out in a total reaction of 50µl by incubating the mixture in the water bath at 37 °C for 1 hour as follows:

	50µl Reaction
10XBuffer	5µl
DNA	1µg
Enzyme	2µl/enzyme
ddH2O	Up to 50µl

After digestion, DNA and vector fragments were purified from the digestion mixture by using GenElute PCR clean-up kit and GenElute plasmid miniprep kit respectively following the manufacturer's instructions.

2.2.5.2 Gel extraction

To obtain DNA/plasmid, agarose gel electrophoresis was carried out then both desired bands were cut out viewing with an UV illumination box. A QIA quick Gel extraction kit was used according to the manufacturer's instructions to extract cDNA or plasmid.

2.2.5.3 Ligation reaction

DNA of interest was ligated into vector at a molecular ratio of 1:3 vector to insert using T4 ligase as followed:

Component	20 µl Reaction	Control
10x T4 DNA ligase buffer	2µl	2µl
Vector	50ng	50ng
Insert DNA	150ng	-
Nuclease-free water	Up to 20µl	Up to 20µl
T4 DNA ligase	1µl	1µl

The mixture was mixed briefly and subjected to centrifugation at 13,000 x *g* for 30 sec then incubated at RT for 2 hours then chilled on ice.

2.2.5.4 Transformation

Competent *E. coli* (25µl) were defrosted on ice then ligation reaction were transformed into competent cells as follows: 5µl of cDNA/plasmid-ligation was added once competent cells had just defrosted, mixed gently, then placed on ice for 30 min. Afterwards, the tube was placed on heat block at 42 °C for 40 sec then immediately cooled on ice for 5 min. 1ml of SOC medium was added to the mixture and incubated at 37 °C with shaking for 1 hour. 1ml of mixture was transferred to a new tube and subjected to centrifugation at 13,000 x *g* for 1 min. The supernatant was removed and the pellet was resuspended with 100µl of Mu medium. Finally, the mixture was spread onto LB agar plates containing the appropriate antibiotic. The plate was left inverted at 37 °C for 12-18 hours for selection of transformed cells.

The next day, single colonies were picked and cultured in tubes containing 4ml of Mu media and 4µl of appropriate antibiotic, then incubated overnight at 37 °C with vigorous shaking. To isolate and purify DNA from bacteria culture, plasmid Mini prep kit was used according to manufacturer's instructions.

2.2.5.5 DNA Sequencing

Sanger DNA sequencing was carried out by sending 5µl of purified plasmid (100ng/µl) or PCR product (1ng/100bp) and 5µl of appropriate primers to Source Bioscience Technology. Sequence files were analysed by using FinchTV 1.4.0 program and aligned with the Blast function of the NCBI genome browser.

2.2.5.6 Transfection by electroporation

Electroporation was used for both transient and stable transfection of mammalian cells. When cells in T75 flasks were at high confluence (70-90%), cells were harvested and subjected to centrifugation. Pellets of approximately 1×10^7 cells were resuspended in 500 µl of medium (without serum) then 10 µg of plasmid (containing gene of interest and antibiotic selection) and 8 µg of GFP plasmid were added and mixed then placed into an electroporation cuvette with 4 mm gap. The cuvette was placed into the electroporation apparatus and shocked once at high-voltage electrical pulse of defined magnitude and length. The cells were then placed in normal complete cell growth medium (set in section 2.2.1.1) in T75 flask and incubated at 37 °C and 5% CO_2. The number of shocks, the voltage and capacitance settings were adjusted to 310 volt with one pulse at length 10msec. Transfected cells could be visualized by observation of GFP expression (green) under fluorescence microscope over 24-72 hours to determine the efficiency of transfection.

2.2.5.7 Generation of stable cell lines

48-72 hours post transfection, selection antibiotic was added at low, optimal and high dose (set in section 2.2.2.3) to three transected flasks and to 6-well plate as follows: one well was left without antibiotic as control, one with low dose, two with optimal dose, and two with high dose of antibiotic. Every 2-3 days, medium with the appropriate concentration of antibiotic was replaced and cells were examined visually for toxicity. By 9 days post transfection, surviving cells were frozen with medium lacking selection antibiotic as a polyclonal cell line.

2.2.6 Bacteria culture

2.2.6.1 Preparing bacteria growth medium

- LB-Agar plate: 5g NaCl, 5g tryptone, 2.5g Yeast Extract and 7.5 Agar were mixed together then ddH2O was added to make up to 500ml then autoclaved for 121 °C for 20 min. Medium was cooled to 50 °C then appropriate antibiotic was added. The solution was poured in 10 cm petri dishes then stored inverted at 4 °C.

- 1Liter of LB-media (Luria broth): 0.5 g of NaCl, 10 g of Bacto Tryptone, and 5 g of Yeast extract were dissolved in 1L of ddH2O then 2 ml of 1M NaOH was added then the mixture was autoclaved at 121 °C for 15min.

-1Liter of Mu medium: as same as LB media but 10 g of NaCl instead of 5g.

-1Liter of SOC medium: 20 g (2%) BactoTrypton, 5 g (0.5%) BactoYeastExtract, 0.6 g (10 mM) NaCl, 0.2 g (3 mM) KCl were dissolved in 1 L of ddH2O then autoclave at 121 °C for 15min then added to the cooled sterile filtered solution: 10 ml of each of (10 Mm) MgCl2 , (10 mM MgSO4) , and (20mM) Glucose.

2.2.6.2 Making glycerol stocks

500 µl of an overnight culture of bacteria transformed with plasmid containing gene of interest was added to 500 µl of 90% glycerol, mixed and stored at -80°C for long term storage.

2.2.7 Protein analysis

2.2.7.1 Coverslip preparation

Coverslips were soaked in 38% 1M HCl for 30 minutes then rinsed with sterile distilled water followed by 95% ethanol. They were left to dry in the hood and were stored at room temperature in a sterile container. The coverslips were placed in a 6-well plate and covered with 2ml of 100µg/ml Poly-D-Lysine then incubated overnight at 37 °C. Next day, the coverslips were washed with sterile 1x PBS before plating cells.

2.2.7.2 Immunofluorescent staining

Cells were plated at the density of 2×10^5 per well in complete growth medium in a 6-well plate containing prepared coverslips. Next day, culture medium was removed and cells were fixed with 4% PFA for 10 minutes. PFA was removed and wells were washed with 1x PBS. After removing 1x PBS, cells were permeabilised by incubating them with 0.2% Triton X-100 in 1x PBS for 10 minutes at room temperature. The solution was removed and cells were incubated with blocking solution (5% marvel milk and 0.1% triton X-100 in 1x PBS) for 30 minutes at room temperature then washed twice with 1x PBS. All subsequent steps were carried out on parafilm. Coverslips were removed from the plate and placed cell-side down on a sheet of parafilm containing drops of about 20µl of primary antibody for each coverslip and incubating for 1 hour at room temperature or overnight at 4 °C in humidified chamber. Coverslips were washed 3 times in 1x PBS. Then, coverslips were incubated with fluorescently labeled secondary antibody for 1 hour at room temperature followed by 3 washes in 1x PBS. Coverslips were mounted by placing cell-side down onto microscopic slides containing one drop of mounting medium with DAPI (4’,6-Diamidino-2-phenylindole) then sealing around coverslips perimeter with nail polish.

2.2.7.3 Western blot

- Sample preparation

Cell pellets were dissolved and mixed well by pipetting in 30-50 µl cold 1x PBS then 50µl of boiling 2x Laemli buffer and mixed quickly. Samples were boiled at 100 °C for 10 min then incubated on ice for 10 min.

2.2.7.4 SDS-PAGE

Base resolving gel

Reagent	**12.5%**	**2x (12.5%)**
Acryl-Bis 29:1	2.14ml	4.28ml
Solution B*	1.43ml	2.86ml
Water (SDW)	1.44ml	2.88ml
APS 10%	75µl	150µl
TEMED (add last)	7.5µl	15µl

* Solution B:
-1.57M Tris HCl pH8.8 (19g in 100ml)
-0.4% SDS (Sodium Dodecyl Sulfate)
- *APS (Ammonium persulfate) freshly prepared*

Stacking gel

Reagent	**12.5%**	**2x (12.5%)**
Acryl-Bis 29:1	0.65ml	1.3ml
Solution C*	1.25ml	2.5ml
Water (SDW)	1.73ml	3.46ml
APS 10%	75µl	150µl
TEMED (add last)	7.5µl	15µl

* Solution C:
-0.5M Tris HCl Ph6.8 (6.06g in 100ml)
-0.4% SDS (Sodium Dodecyl Sulfate)

First, base resolving gel was poured between two glass plates in a gel caster then quickly overlaid with 70% ethanol until set to flatten the top of the base gel then left for around 20 minutes till the gel solidified. Ethanol was removed and stacking gel was poured. The comb was quickly inserted before the stacking gel polymerized.

The gel and plate were transferred to a gel tank filled with SDS running buffer. The comb was removed and wells were rinsed with running buffer to remove excess acrylamide. 5µl of protein ladder and 5-20µl of samples were loaded per well. The samples were run at 40 volt until the marker reached the base gel, then the voltage was increased to 80V for 2-4 hours.

Running Buffer (10X)

30.3g Tris

188g glycine

10ml 10% SDS

Make up to 1L with sterile water.

-Protein transfer

When markers reach the bottom of the gel, the gel was removed from the tank then rinsed with a transfer buffer. The nitrocellulose membrane was cut to gel size and wet and soaked with filter paper and sponge in transfer buffer. The gel was transferred to transfer tank and run at 30V (~400A) overnight or 100V for 2 hours.

Transfer Buffer (10x)

30.3g Tris

144g glycine

Make up to 1L with sterile water

Transfer buffer (1X)

100ml of transfer buffer (10X)

200ml methanol

Make up to 1L with sterile water.

-Protein detection

The apparatus was dismantled and the membrane was removed from the tank and placed in a container containing 20 ml of blocking buffer for 1 hour at room temperature with rocking. The membrane was washed once with PBST (PBS with Tween20) for 5 minutes. Primary antibody was diluted to appropriate concentration in 20 ml of 2% Marvel milk/PBST and membrane was incubated in this solution at room temperature in rotator for 2h. The membrane was then washed with PBST four times for 5 minutes each on a rocker. After diluting of fluorescently-labelled secondary antibody to the required concentration in 20 ml of PBST, the membrane was incubated in this solution for 2h on the rotator at room temperature. Four washing steps of the membrane with PBST in a covered plate for 5 minutes on a rocker were preceded the detection of target protein.

Blocking buffer

1g of Marvel milk dissolved in 20 ml PBST*

*PBST: 1.25 ml of 20% Tween20 mixed in 500 ml of 1x PBS.

-Detection of protein signal

Using LiCor secondary antibody, membrane was left in wash buffer and taken to the LiCor scanner. The stained membrane was scanned using software (Odyssey infrared image) to visualise protein bands.

2.2.8 Preparing standard reagents

2.2.8.1 Agarose gel

- Preparation of 1L of 0.5M EDTA stock solution: 186.12g of EDTA was dissolved in 800ml of distilled water then NaOH was added to adjust pH to 8 then deionized water was added to reach 1000ml.

- Preparation of 1L of 50× TAE buffer: 242g of Tris base was dissolved completely in 750ml deionized distilled water then 57.1ml of glacial acetic acid and 100ml of EDTA stock solution (pH 8) were added to solution to make up 1 liter.

- Preparation of 1L of working solution 1×TAE buffer: 20ml of 50× TAE buffer was mixed with distilled water to bring the final volume up to 1 liter.

- Preparation of agarose gel: To prepare 1% agarose gel, 0.6g of agarose was added to 60ml of 1×TAE buffer. The solution was dissolved by boiling in the microwave at high temperature for 5 minutes. The solution was left to cool at room temperature to around 50 °C then 1µl of RedSafe was added before pouring the solution into a gel casting tray

with a comb and was left to solidify at room temperature. The gel was then submerged into a gel tank containing 1X TAE buffer.

2.2.8.2 Chemicals

-Decitabine (5-aza-2'deoxycytidine)

A stock solution of 50mM of 5-aza was made by diluting 5mg of 5-aza stock with 438µl 1x PBS. Then, we prepared a working solution of 5mM by mixing 100µl of stock solution with 900µl of DMEM.

- Cisplatin (cis-diamminedichloroplatinumII)

A working solution of 0.33mM was made by adding 150µl of cisplatin (stock was 1mg/ml) to 1,350µl of 0.9% NaCl.

2.2.9 Statistical analysis

Affymetrix expression array data was received as CEL. file format. This data was first preprocessed using the using R programme with 'Affy' package in the bioconducter package provided by www.bioconductor.org (Bioconductor version 2.11 -BiocInstaller 1.8.3) (this analysis was performed in collaborated with MSci student, Claire Wallace). Data was normalised using the RMA (Robust multi-array average) method and filtered to exclude the probes that their expression were below control background levels. Fold changes in expression between each probe of each cell lines relative to seminoma were calculated. The data was exported as a .txt file in order to be read and analysed in Excel (full details of the workflow in Appendix II). The Excel tool PivotTable was used to assign average expression intensity values to each gene. This data was then processed in Excel and a right-tailed Welch's T-test was performed (by MSci student,Matthew Carr) for each gene comparing seminoma samples to non-seminoma samples.

Illumina infinium methylome 450k array data was obtained in Excel file. Pivot table and V-LOOKUP function were used to calculate β value.

To analyse the significance of the findings and compare the difference before and after treatment, the data were statistically calculated using two tailed paired t-test.

Chapter 3. Genome-wide methylation profile of GCT cell lines

3.1 Introduction

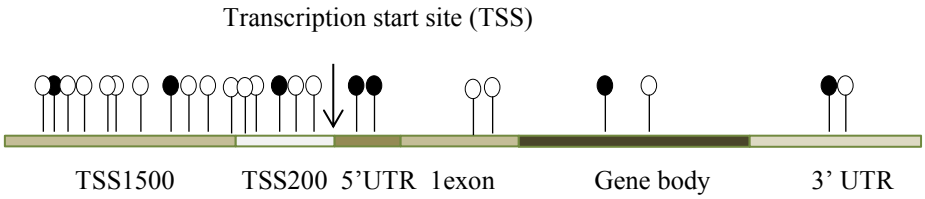
Aberrant DNA methylation has been shown in many tumours including GCTs (Vega *et al*., 2012, Brait *et al*., 2012). Two recent studies of the global methylation of paediatric GCTs demonstrated the hypermethylation of many candidate TSGs (Jeyapalan *et al*., 2011, Amatruda *et al*., 2013). However, both studies were limited by the use of an array platform, Illumina GoldenGate arrays, which analysed only 1-3 CpGs of 807 genes. The former study, which was applied in our lab, demonstrated that there was a significant difference in methylation between seminomas and YST, where the seminomas exhibited a lower level of DNA methylation than non-seminomas. Thus, in this chapter I sought to confirm and extend this result using an array system, the Infinium Human Methylation450 BeadChip platform, which covers 99% of the Reference Sequence (RefSeq) genes with an average of 17 CpG sites per gene to gain a comprehensive view of the methylation differences between subtypes of GCT cell lines.

3.2 Results

3.2.1 DNA methylation level across the genome

In order to investigate the methylation features in the four main histologic subtypes of adult testicular GCT cell lines, the Infinium Human Methylation450 BeadChip platform was used. This Infinium array platform measures DNA methylation levels using >485,000 probes which cover 99% of the Reference Sequence (RefSeq) genes (National Center for Biotechnology Information (NCBI), Bethesda, MD, USA). There is an average of 17 CpG sites per gene region distributed across gene regulatory regions, such as the promoter, 5'UTR (untranslated region), first exon, gene body, and 3'UTR. Sites near the transcription start site (TSS) labelled as TSS 1500 (1500bp upstream of the TSS) and TSS200 (200bp upstream of the TSS) were covered. This platform covers 96% of CGIs where the CGI refers to a region of at least 500bp, with greater than 55% GC content and an observed-to-expected CpG ratio >0.65. The other gene regions were: shores which are regions 2kb either side of an island, shelves which indicate regions 2kb outside of the shores and finally other (open sea) indicate the remaining sequences (Figure 3.1).

A)



B)

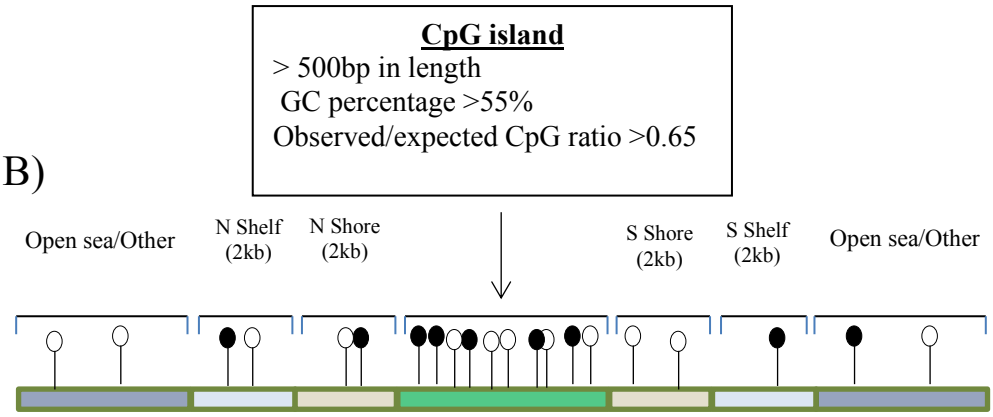


Figure 3.1: Diagrams illustrate gene regions and Illumina-annotated probes: A) Based on gene regions: sites in the promoter region, 5'UTR, gene body, and 3'UTR. B) Based on their sites relative to CpG island: shores were 2kb up and down-stream of islands while shelves were 2kb outside of the shore. Open sea probes were not annotated into any of the other Illumina classes. S refers to south, N refers to north.

The Infinium Human Methylation 450 BeadChip combined two Infinium assays, I and II, to cover most islands in the whole genome (Figure 3.2). The Infinium I assay uses two probes, one for methylated DNA (M) and other for the unmethylated (U), for each CpG site. The attachment process of each probe to the DNA strand depends on the bisulfite converted genomic DNA where the 3’ terminal end of each probe is designed to attach to either a cytosine base when it is methylated or to a thymine base when the cytosine in the DNA is unmethylated and not protected from bisulfite conversion. The Infinium II has one probe for both alleles. The 3’ terminus of the probe is designed to be complementary to DNA bisulfite conversion strand by incorporation of an A or T base (unmethylated site) in red (U) and the incorporation of C or G signals (methylated site) in green (M). Another feature of Infinium II is that it can measure the methylation level of up to three underlying CpG sites. Using the two bead types in the Infinium Human Methylation 450 BeadChip covers more of the CpG sites in the whole genome.

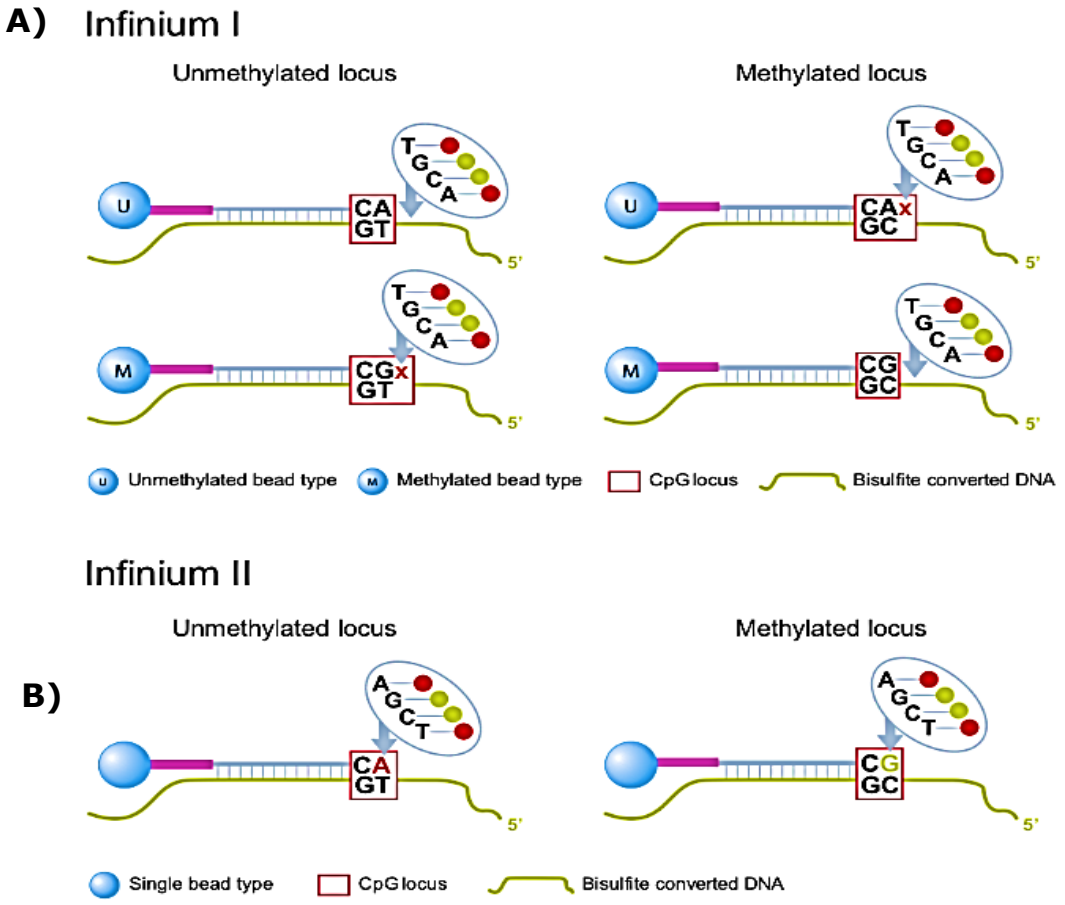


Figure 3.2: Infinium Human Methylation450 array scheme: It combines two technically distinct assays in one platform, the Infinium I assay and the Infinium II assay. A) The Infinium I assay uses two probes. One for methylated and the other for the unmethylated locus. B) The Infinium II assay uses only one probe with a two-labelled base signal: a green signal for a methylated and a red signal for an unmethylated locus (www.illumina.com).

This assay was carried out (by our lab colleague, Dr. D. Noor) by extracting genomic DNA (500ng) from four subtypes of TGCT cell lines (representing seminoma, YST, EC and teratoma) as mentioned in section 2.2.3.1. The genomic DNA was first subjected to bisulfite modification using a Zymo EZ DNA methylation kit according to the manufacturer's protocol. Subsequently, the samples were sent to Queen Mary's University of London to carry out the Infinium methylation 450 array. The raw data was received as an Infimum dataset with values representing a level of methylation at individual CpG sites as a beta-value (β) ranging from 0 (unmethylated) to 1 (fully methylated). Each probe is defined according to the Illumina website (www.illumina.com) as a measure of location relative to UCSC-reference-group. This indicates the location in relation to the closest gene as within 1500 or 200 base pairs from the transcription start site (represented as TSS 1500 or TSS 200), 5` UTR, gene body and 3` UTR) and as a measure of location relative to UCSC- CpG-Island (island, north/south shore (within 2kb up/downstream of an island respectively), north/south shelf (2-4kb up/downstream of an island, respectively), and blank indicating away from an island).

Analyses were carried out using the Excel programme. The Excel tool `Pivot Table` was used to determine average methylation (β-value) and differential methylation (delta-β-value) to each gene in relation to locations with respect to CpG islands and gene regions. Before analysis, probes that could lead to biological and technical bias such as a gender-specific bias were excluded from the analysis and also probes that were labelled for multiple gene names. Therefore, data for X and Y chromosomes and probes that represent multiple gene names or that did not represent any gene symbol were excluded. The analysis utilised two criteria which have previously been used in many investigational studies (Bibikova *et al*., 2011, Shen *et al*., 2013).

First, a beta value (β), obtained as a quantitative reading that varies between 0 (fully unmethylated) and 1 (fully methylated), was used to estimate the methylation level of the CpG locus using the ratio of intensities between the methylated and unmethylated probe. According to the Illumina website (www.illumina.com) that describes the methylation level in Infinium Methylation450 data, a methylation score of 0.3 or less represents unmethylated sites, scores of 0.6 and above represents methylated sites and values between 0.29 and 0.59 represents partially methylated sites.

Second, delta-beta values were calculated which in our study indicate the difference in methylation between seminoma and non-seminoma samples by subtracting the average beta value of seminoma from that of non-seminoma samples.

Initially, the distributions of differentially methylated regions across the genome for four testicular GCT cell lines were determined to investigate the regions that could be affected by methylation. Some previous studies reported that specific regions of some genes have methylated features such as methylation of promoter CpGs which may suppress gene expression (Irizarry *et al*., 2009). As the Infinium Human Methylation450 BeadChip array covers different regions of genes (such as island, shore, shelf) we calculated the average methylation level in the four GCT cell lines in different regions of genes by using the 'pivot table' function of Excel for the data of this array. As shown in Figure 3.3, the CGI regions in all cell lines had the lowest level of methylation while the methylation level increased in regions outside the CGI. However, when comparing non-seminoma to seminoma cell lines, non-seminoma had higher numbers of methylated CpGs in islands (where β-value $\geqslant$0.6) than the seminoma cell line. The YST cells showed lower number of methylated CpGs in shelves and others `open sea` than the seminoma

cells. For the EC and teratoma cell lines, the number of methylated CpGs was also higher compared with seminoma in all regions.

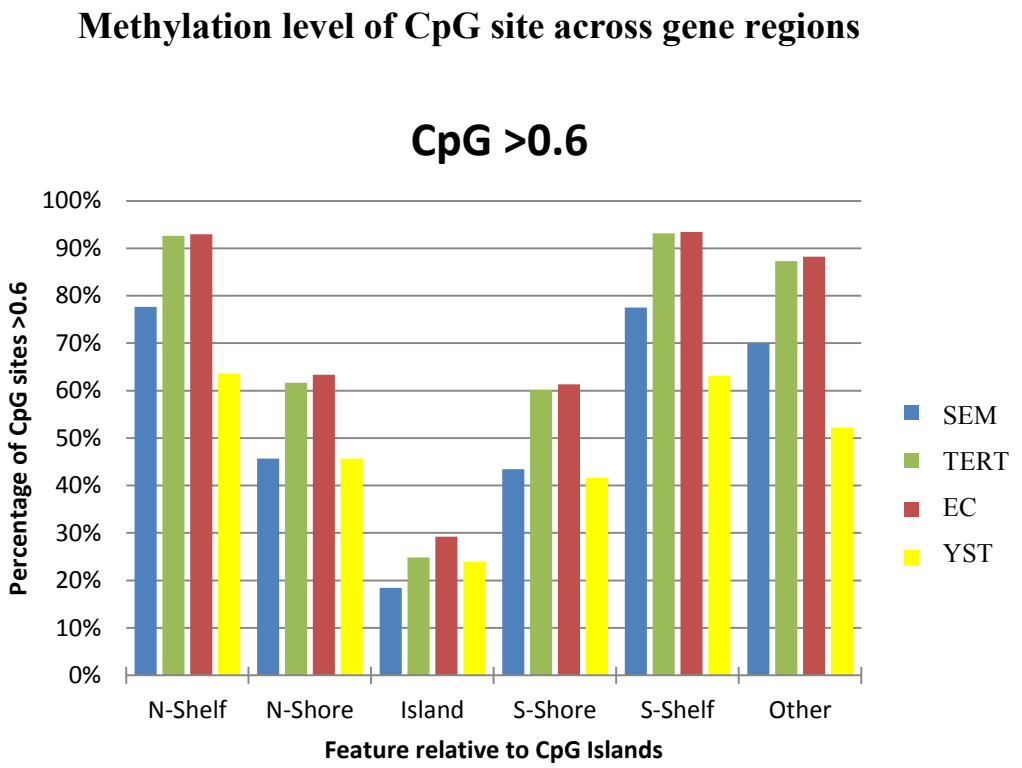


Figure 3.3: Percentage of CpG sites methylated (β-value ≥0.6) relative to CGIs. Methylation level of CpG sites across gene regions shows that island region had the lowest level of methylation in all cell lines. TERT and EC had a similar degree of methylation level at all regions and had higher level of methylation compared to seminoma. The YST cells had lowest level of methylation in shelves and open sea while their methylation was higher than seminoma in island region.

Several studies have reported that methylation of CpGs at promoter regions is associated with transcriptional repression (Fouse *et al*., 2008, Bird, 2011, Deaton and Bird, 2011, Vavouri and Lehner, 2012). Therefore, the data was re-analysed to assess the methylation level of CpGs at regions relative to the transcription start sites (TSS). The average β value for each cell line was plotted (Figure 3.4). There was a clear reduction of methylation level near the TSS (TSS155, TSS200, 5'UTR, and 1st exon) and the level increased in regions more distal from TSS (3'UTR and gene body). In general, methylation of CpG sites in SEM showed low in all regions (Figure 3.4) compared to the non-seminomas in general which indicated that seminomas have lower levels of methylation compared with non-seminomas. This finding for the whole genome is consistent with the previous study by Jeyapalan (2011) that used arrays in which fewer genes were analysed.

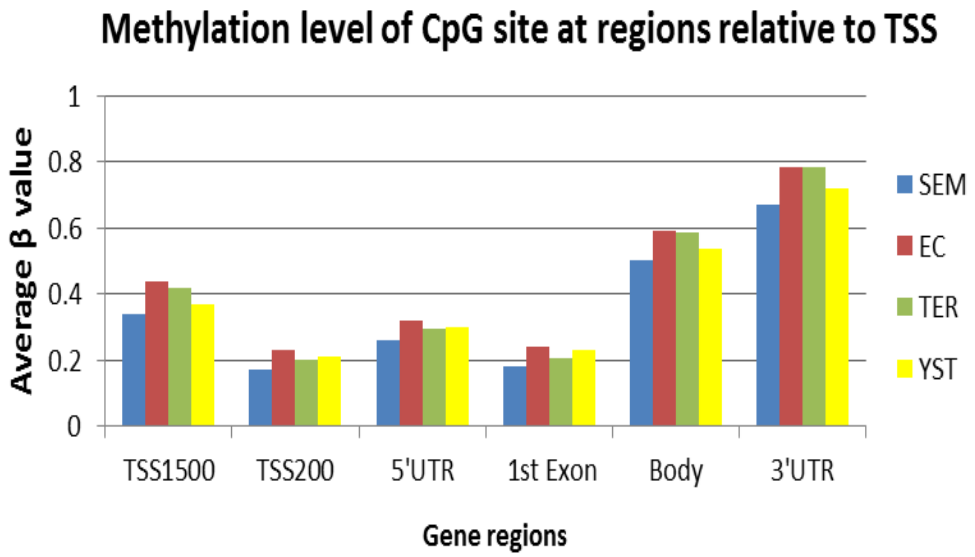


Figure 3.4: Methylation levels at CpG site at regions relative to TSS: Average methylation levels in four TGCT cell lines shows low level of methylation of CpGs near the TSS (TSS155, TSS200, 5'UTR, and 1st exon) comparing to regions away from TSS (3'UTR and gene body).

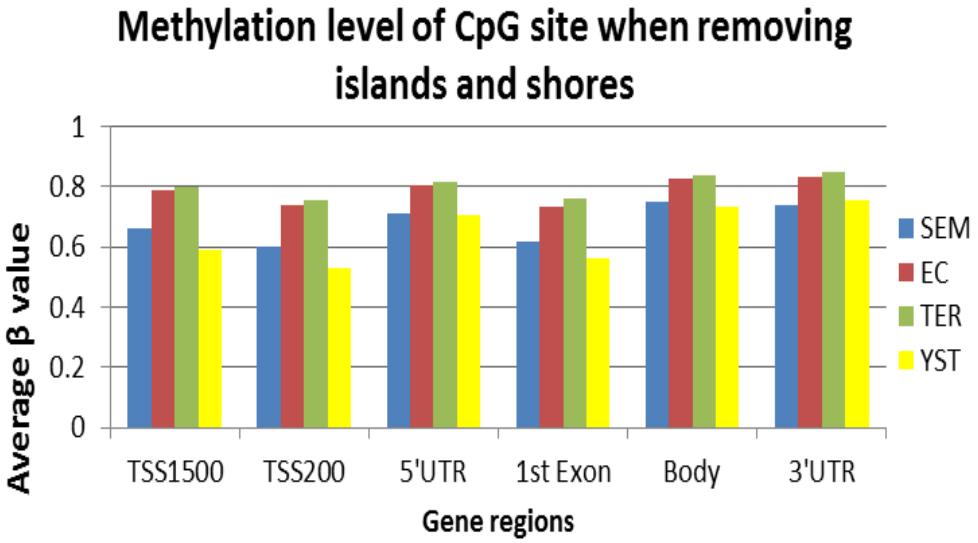


Figure 3.5: Methylation levels when removing islands and shores: Average methylation level after removing island and shores were similar either near or away from TSS

The results also showed that islands and shores had the lowest level of methylation compared with other regions (Figure 3.3), therefore these were removed from the analysis and the graph was re-plotted (Figure 3.5). The result showed that the methylation levels were similar in all regions. This result illustrated that the methylation level for YST were lower than the SEM cell line after removing CpG site at islands and shores. Also, this finding demonstrated that the low level of methylation in regions near TSS was due to low methylation of CpG site in islands and shores.

3.2.2 Methylome profile across four cell lines

I next compared the methylation levels between the four cell lines to establish similarities and differences between them, based on average methylation of CGIs near the TSSs. From the previous quantitative standpoint, there were considerable differences in the methylation between seminoma and non-seminoma and between the non-seminoma cell lines themselves.

To determine the association between tumour types and normal cells in terms of their methylation, the CpG methylation level in 20 normal cells samples obtained from The Cancer Genome Atlas project TCG (TCGA-07-0227-20A-01D-XXXX-05), where XXXX represents the ID for each normal cell as (1 Brain (1481) , 1 Central Nervous System tissue (1467), 2 colon (1651 and 1407), 2 Head/Neck (1433 and 1511), 7 Kidney (1500, 1670, 1275, 1418, 1536, 1590, and 1424), 4 lung (1626, 1633, 1551, and 1440), 1 Prostate (1578), 1 Rectum (1658), 1 Stomach (1601)) (http://cancergenome.nih.gov/), and the four GCT cell lines included in this study was compared. This is represented by Dr. Jeyapalan as a heat map with Ward's hierarchical clustering that consists of columns and rows representing the samples and the CpG sites respectively (Figure 3.6). It is clear that YST cells had a high

proportion of methylation compared to SEM and control samples. Also, the seminoma sample was the most similar in methylation to the control samples. This result demonstrated that the YST cell line was associated with a high level of global methylation (hypermethylation) whereas the SEM was hypomethylated and there was much less difference in methylation between SEM and normal cells (Figure 3.6). Interestingly, the distribution of DNA methylation showed a similar pattern in all normal cells from different patients and different tissues which gave more evidence that the methylation patterns that were seen in tumour cell lines was aberrant and could be associated with development of cancer. With respect to the methylation level in normal cells compared with GCT cell lines, the high degree of similarity in methylation of normal cells and undifferentiated SEM cells and the presence of a difference in both EC and TERT, in addition to a substantial difference in differentiated YST cells suggests that aberrant methylation could be a key factor in the differentiation of GCTs.

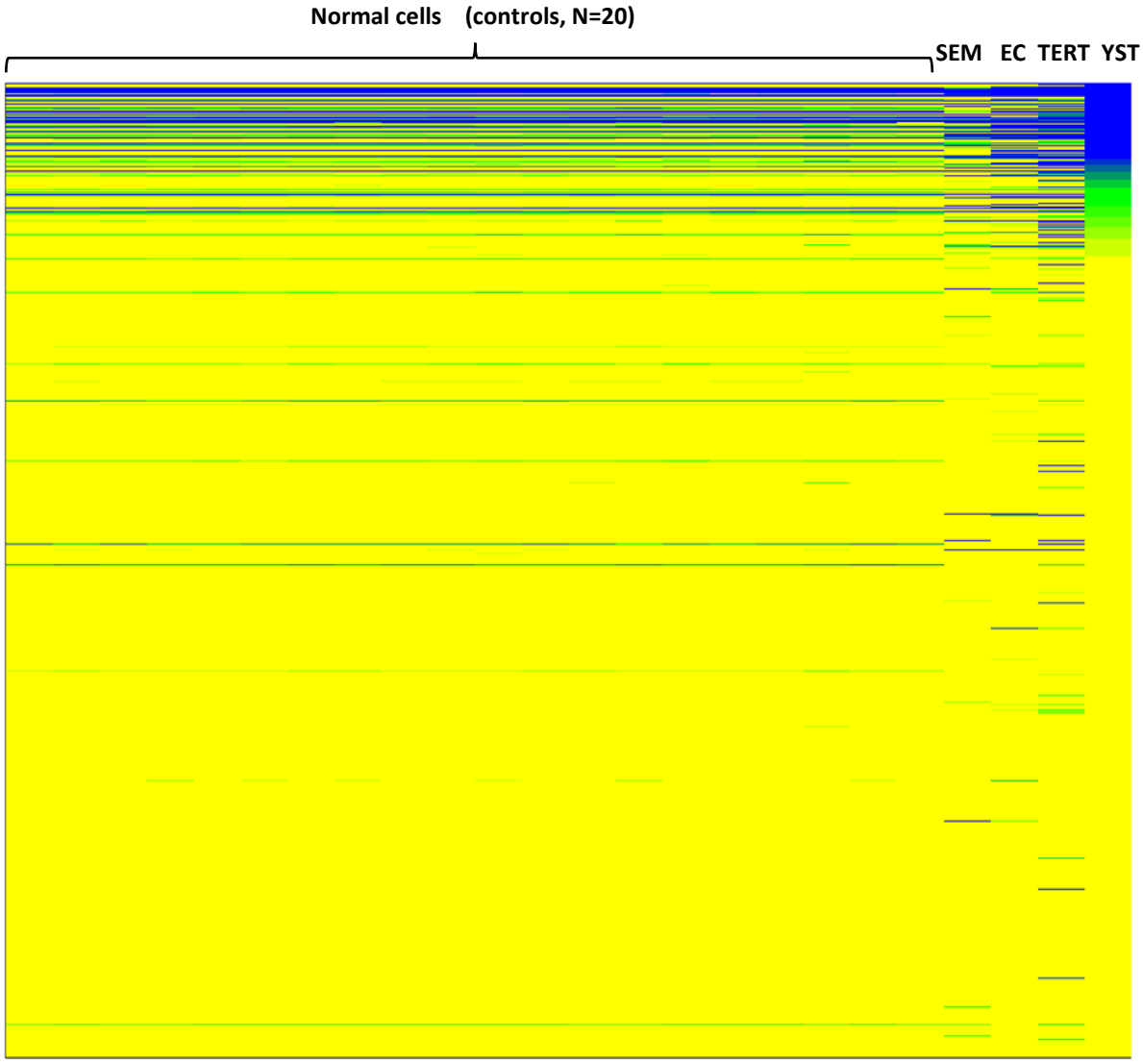


Figure 3.6: Heat map represents methylation of CpGs in normal cells and the four cell lines in this study. The highest methylation is shown in blue, the lowest in yellow, and partially methylated in green. The heat map shows that YST has a highest density of methylation compared to SEM and controls (Image kindly provided by Jennie Jeyapalan).

3.2.3 Anti-5-methylcytidine immunostaining showing differences in total methylation between SEM and YST

To verify the difference of global methylation between SEM and YST cell lines, anti-5-methyl-C immunostaining was performed (as section 2.2.7.2). To detect nuclear 5-methyl-C, a primary antibody (Mouse Anti-5-methylcytidine, Bio Rad, UK) was used. Fluorescently labelled secondary antibody (TEXAS RED® anti-mouse, Vector, USA) was then applied for visual analysis using a fluorescent microscope. Stronger nuclear staining intensity for anti 5-methylcytidine was done in YST cells compared with SEM. To obtain further evidence for the presence of methylation difference between cells, YST cells were treated with 5μM 5-aza (set in section 2.2.2.1) and compared with untreated YST cells. The result revealed a strong reduction in the anti-5-methylcytidine staining in YST cells after using a demethylation agent (5-aza) (Figure 3.7). These results provide clear confirmation that YST cells have a hypermethylation feature and thus this methylation can be substantially reduced by treatment of the cells with 5-aza.

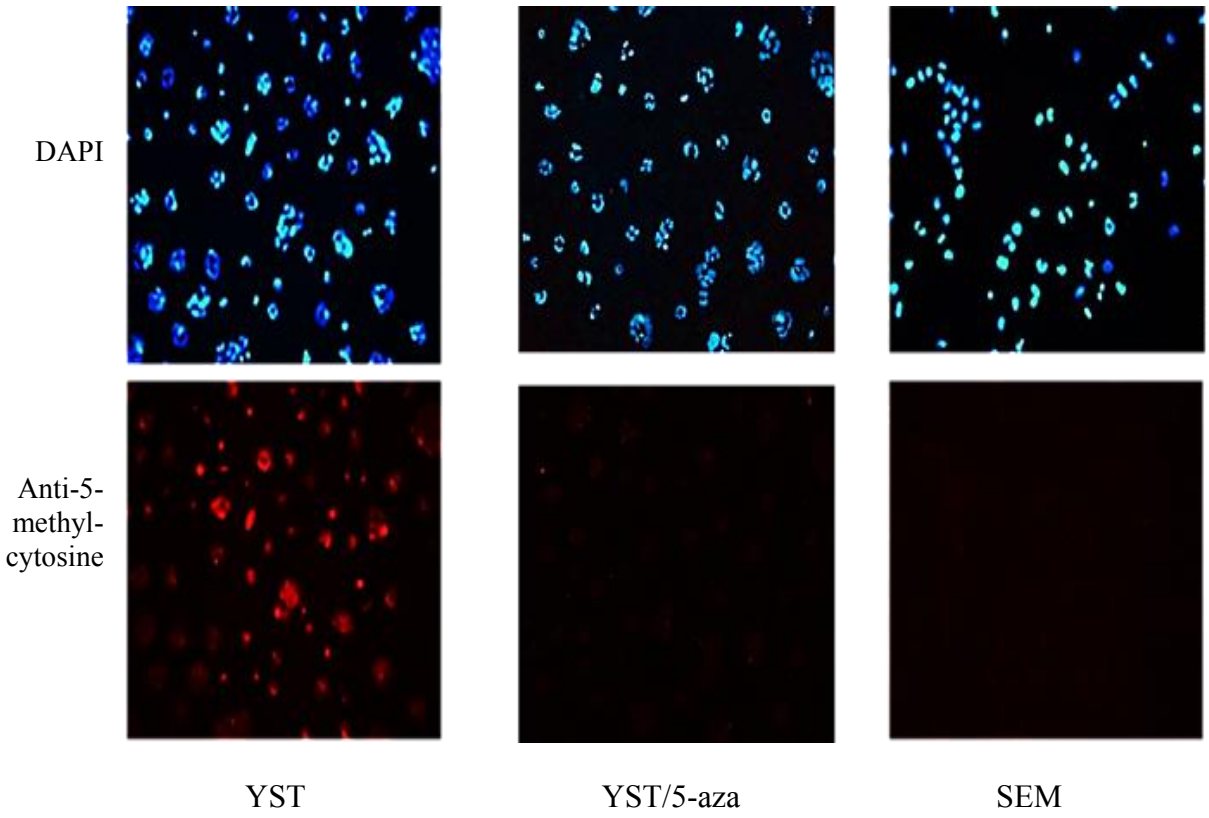


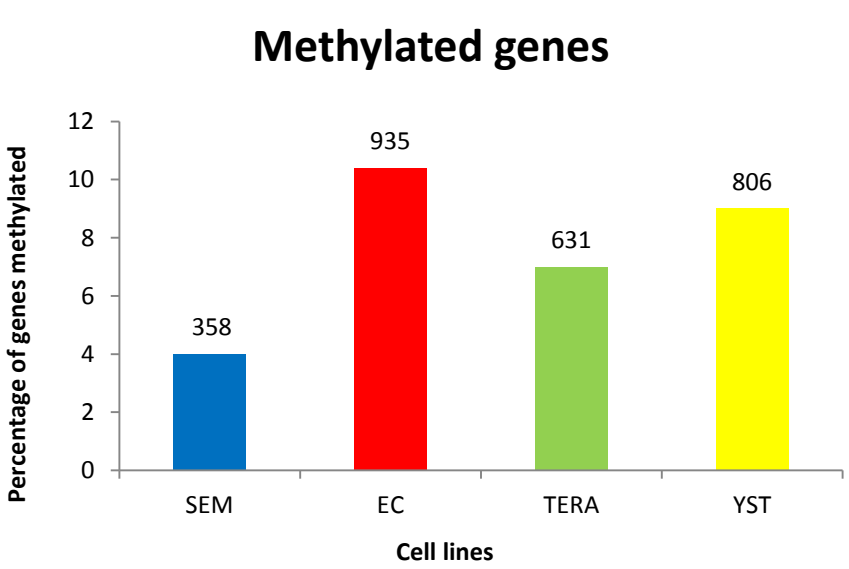
Figure 3.7: Immunostaining using anti-5-methylcytidine: There was a strong staining intensity of anti 5-methylcytidine in YST cells compared to SEM and 5-aza-treated YST. Red fluorescence is Anti-5-methylcytidine staining. DAPI is blue staining.

3.2.4 Identification of methylated genes in GCT cell lines

According to the results presented above, which showed that the most differentially methylated regions in the GCT cell lines were near the TSSs, the number and percentage of genes hypermethylated at CpGs near TSS were calculated (in collaboration with a MSci student, Matthew Carr) for each cell line. A gene was considered to be hypermethylated when the average of β-value across all CpGs near the TSS ≥0.6. This analysis showed that the highest number of hypermethylated genes was in the EC cell lines (935), then YST (806), followed by TERT (631) and SEM (358) (Figure 3.8-A).

The similarities and differences between the cell lines were summariSed in a Venn diagram (Figure 3.8 B). This shows the numbers of methylated genes which were common among four cell lines and the numbers which were uniquely methylated for each cell line. A gene was considered to be uniquely methylated when the average of β-values across all CpGs near TSS ≥0.6 in one cell line and < 0.6 in the other cell lines. It was apparent from the Venn diagram that 94% of genes (337/358) methylated in the seminoma cell line were also methylated in EC and TERA cells and 62% of these genes (222/358) were methylated in all subtypes. For YST, 66% of genes (536/805) methylated in the YST cell line were methylated in EC and TERA cells, with 27% (222/805) of them methylated in all cell lines (Figure 3.8-B). The YST cell line had the highest number (270 genes) of uniquely methylated genes relative to others, whereas only 16 genes were uniquely methylated in the SEM cell line.

A



B

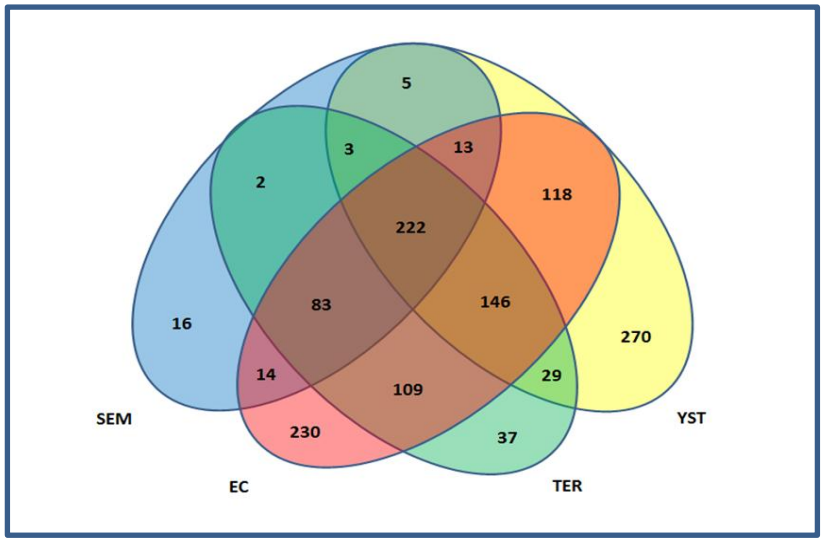


Figure 3.8: The population of methylated genes in GCT cell lines. Gene methylation based on the average of β-values across CpGs located near the TSS region. (A) Represents a percentage of methylated genes in each cell line, where values above bars indicate the number of methylated genes. (B) Venn diagram represents the overlap between genes methylated in each cell line.

3.3 Discussion

Global hypomethylation and promoter hypermethylation are core characteristics of many tumours and there is expanding evidence that these epigenetic changes could play role in cancer development (Vega *et al*., 2012, Brait *et al*., 2012). More recent studies for germ cell tumours revealed that hypermethylation of particular areas of DNA in addition to global hypomethylation could play a critical role in tumorigenesis (Jeyapalan *et al*., 2011, Amatruda *et al*., 2013). In this study, measurements of global methylation level of adult testicular GCT cell lines using high-throughput approaches reveals that seminomas have a low level of global methylation (hypomethylation) while non-seminomas, represented by three types YST, EC, and TERT, showed higher level of global methylation (hypermethylation). This finding was consistent with the general pattern described by Jeyapalan *et al.* (2011) and Amatruda (2013) studies that used paediatric germ cell tumours samples.

Furthermore, the methylation levels in normal control tissues were relatively similar to those of SEM (Figure 3.6). The low level of methylation in SEM is consistent with studies revealing that seminomas resemble PGCs and CIS (Skakkebaek *et al*., 1987, Eckert *et al*., 2008, De Felici, 2013, Elzinga-Tinke *et al*., 2015).

Analysis for different gene regions of four cell lines showed that the CGI region in all cell lines had the lowest level of methylation while the methylation level increased in regions outside the CGI. However, comparison of non-seminoma to seminoma cell lines showed the three non-seminomas had higher levels of CGI methylation than the seminoma cell line. This finding supported other studies suggesting that CGI methylation could have a central role in tumourigenesis, in particular, the methylation status of CGIs may be associated with the expression of critical genes that play roles in the development of cancer,

or differentiation of tumour cells (Brown and Strathdee, 2002, Feltus *et al*., 2003, Bock *et al*., 2006, Previti *et al*., 2009).

The regulatory importance of CGI methylation in gene expression may reflect their location in the promoter region, where the promoter has been proven for its role in regulation of gene expression. Many studies have reported that genes with a CGI promoter have transcriptional regulatory features where the methylation of promoter CpGs may suppress gene expression (Bird, 2011, Deaton and Bird, 2011, Vavouri and Lehner, 2012).

Both the EC and TERT cell lines, which are clinically more differentiated than seminomas (Cusack and Scotting, 2013), showed higher methylation than the SEM cell line and had a similar degree of methylation at all regions that were analysed as above in Figure 3.3. It is striking to see the high similarity in methylation level at all regions in these two cell lines where in fact the teratoma cell line (NT2D1) used in this analysis was derived from an EC component of a teratoma (Yao *et al*., 2007, Boucher and Bennett, 2003).

Regarding YST, it was clear that CpG sites near promoter regions had more methylation than seminomas. In addition, when comparing the methylation level at different regions, the islands and shores in YST were more methylated than those in SEM. This finding pointed to island and shores playing a critical role in global methylation difference between SEM and YST. Also, the result showed that the YST cell line had the highest number of uniquely methylated genes relative to the other cell lines which suggested that those uniquely methylated genes could contribute to the differences in phenotype seen between SEM and YST (Feltus *et al*., 2003).

Chapter 4. Relation between methylation and gene expression in GCT cell lines

4.1 Introduction

Many studies have described changes in the DNA methylation pattern in varies types of cancer cells compared with normal cells from the same tissues and identified the effect of methylation on gene expression (Goel *et al*., 2004, Zöchbauer-Müller *et al*., 2001, Herman *et al*., 1995). Furthermore, it was noted that the reduced methylation level of a whole genome (global hypomethylation) is related to proto-oncogene activation and chromosomal instability, whereas increased methylation levels of the CGIs associated with specific genes (regional hypermethylation) is associated with the silencing of tumour suppressor gene expression (Subramanian and Govindan, 2008).

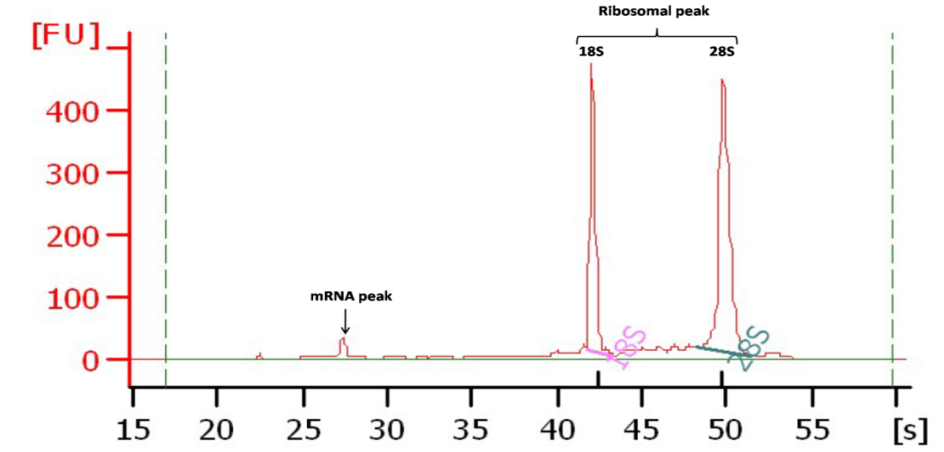
Recently developed methylation and expression arrays have revealed a correlation of methylation and expression of many genes in many classes of tumours analysed (Martin-Subero *et al*., 2009). Therefore, in this chapter, I sought to determine the correlation between methylation and gene expression in GCT cell lines using microarray analysis.

4.2 Results

4.2.1 The correlation between differential expression and methylation in different gene regions between seminoma and non-seminoma cell lines

Gene expression analysis was performed in collaboration with my colleague (Dr. D.Noor) on the same four GCT cell lines using Affymetrix Human GeneChip U133 Plus 2.0 arrays. RNA was isolated from the four cell lines using the RNeasy Mini Kit and QIAshredder, according to manufacturer instructions. The quality of RNA was assessed using the Agilent Bioanalyzer. Samples with an RNA integrity number (RIN) of more than 9.0 were selected for the Affymetrix gene expression array (Figure 4.1).

Arrays were performed at the Nottingham Arabidopsis Stock Centre, University of Nottingham Sutton Bonington Campus, using Affymetrix GeneChip Human Genome U133 Plus 2.0 arrays. The high quality RNAs were labelled and hybridised to probes designed to attach to more than 30,000 genes for an analysis of over 47,000 transcripts with respect to GenBank®, dbEST, and RefSeq.



Overall Results for sample 5 : yst1

RNA Area:	1,605.8
RNA Concentration:	1,203 ng/µl
rRNA Ratio [28s / 18s]:	1.7
RNA Integrity Number (RIN):	9.5 (B.02.07)

Figure 4.1: The RNA integrity number (RIN) for one TGCT cell line sample: The RNA integrity measurement was performed by Agilent 2100 Bioanalyzer software, which showed a high integrity of RNA (9.5)

The expression data sheet was obtained as a CEL file and was then analysed in collaboration with MSci student in our lab (Claire Wallace) using the statistical software R, with packages provided by www.bioconductor.org. Data was preprocessed using the RMA method to normalize the data (AppendixII). The data was exported as a .txt file then analysed using Excel. Probes that had expression outputs below control background were excluded. The Excel tool `pivot table` was used to assess the average expression intensity values of each gene. The fold change in expression for each cell line was calculated relative to seminoma.

The methylation data that was obtained from the Illumina Infinium HumanMethylome450 array were analysed using Excel spreadsheets (the analysis of data was explained previously in chapter 3, page 69-70). In brief, quantitative measurements of methylation across CpGs sites were assigned as beta values (0<beta<1, 0 represents an unmethylated site, and 1 represents an methylated site). Delta-beta values of differential methylation were calculated as the difference in methylation between seminoma and non-seminoma by subtracting the average beta value of seminoma from that of non-seminoma.

Expression and methylation data were merged in a single spreadsheet in Excel using the Excel function `VLOOKUP`. Further statistical analysis was carried out using functions in Excel. According to previous results and investigations that showed that methylation of CGI promoters results in gene silencing, I focused on average beta-value methylation of islands and their flanking regions (shores and shelves) relative to regions near the TSS (TSS1500, TSS200 and 5`UTR).

To determine whether there was a statistically significant correlation of differential gene expression with each level of differential methylation between seminomas and non-seminomas under the null

hypothesis that lower gene expression does not correlate with methylation, calculations of Pearson's Chi-squared test and student's t-tests with p-values (a p-values of 0.05, 0.01, and 0.001, represented as *, **, and *** respectively, were chosen as significance markers) were applied for association of methylation with gene expression between the two samples over a range of delta-β-value categories (Δβ intervals ranging from no methylation difference (0-0.05) between the two cell lines to fully methylated in non-seminoma cell line and unmethylated in seminoma (0.9-0.95)). Contingency tables were created for the observed and expected number of genes differentially expressed between seminoma and non-seminoma cell lines. For each category, the expected frequency of genes showing more than two fold differential expression were calculated under the null hypothesis that lower gene expression does not correlate with methylation. Correlation was measured by counting the number of genes in each category of differential methylation that were also differentially expressed, then the percentage correlation was calculated for each delta-beta category. The data and degree of correlation between the level of differential methylation of the various gene elements (islands, shores and shelves) and differential gene expression was presented in (Table 4.1) then plotted graphically using Excel graph tools showing the percentage of genes with more than two fold differential expression for each delta-β-value category (Figure 4.2) (See Appendix I for other Tables and Figures).

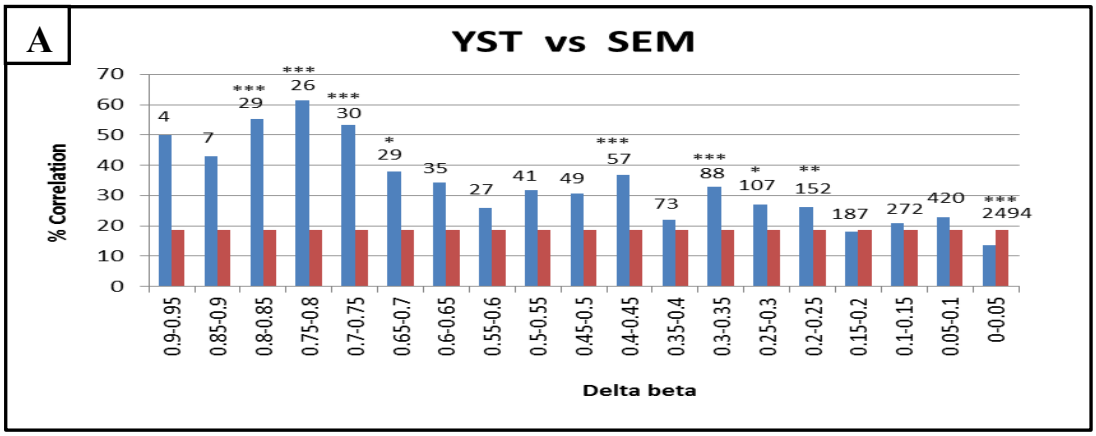
The most striking result to emerge from these data is that a delta-β-value of above 0.7 in the EC and TERT, and 0.65 in the YST consistently correlated significantly with a difference in expression greater than two-fold, which is clear when comparing YST with seminoma cells (Figure 4.2-A), where the correlation percentage was over 40%. For all non-seminoma cell lines, the correlation between

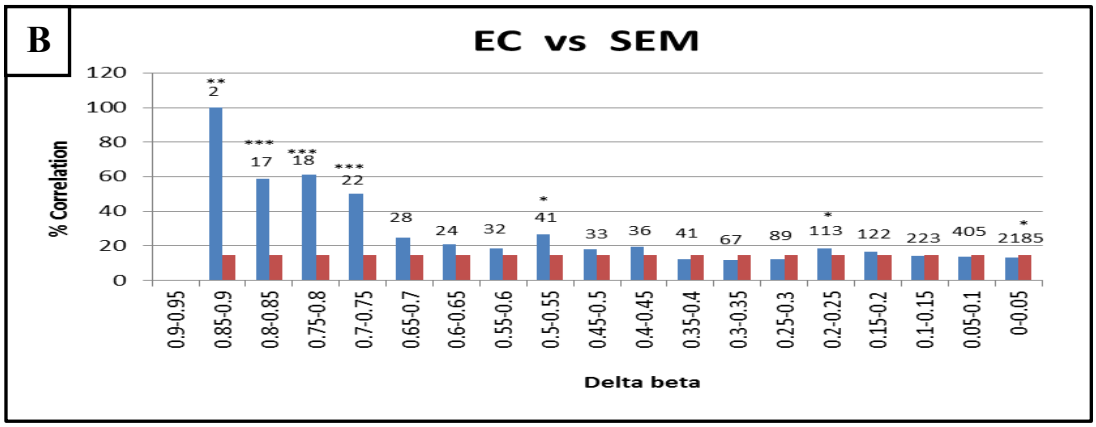
differential methylation and differential expression was stronger in islands near the TSS than other regions which revealed that methylation of islands is likely to be of biological significance with respect to gene silencing.

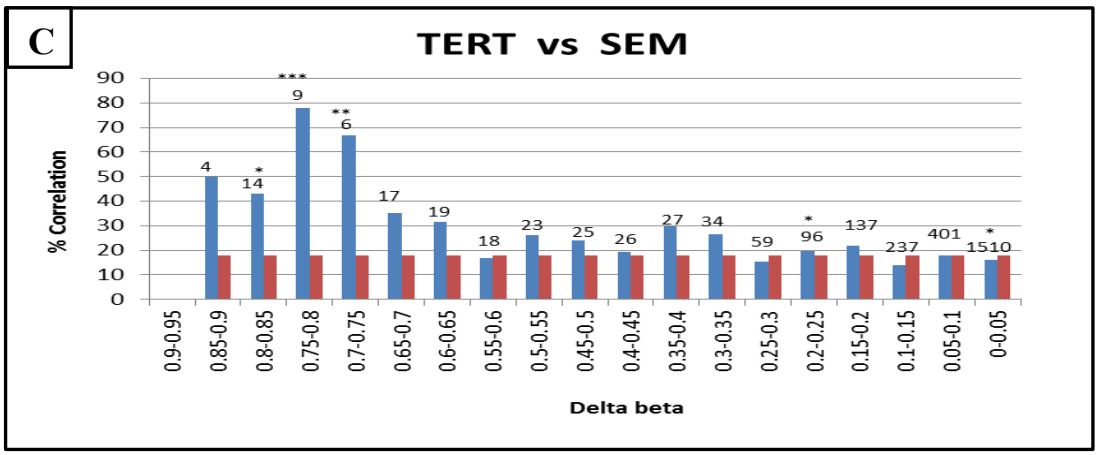
Table 4.1: The contingency table shows observed and expected numbers of genes differentially expressed for ranges of differential methylation.

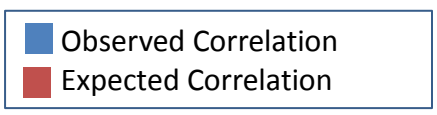
Delta beta	SEM/YST		No difference	No difference	YST/SEM		Total	P-value	Sig.
	Genes >2 fold	Genes >2 fold			Genes > 2 fold	Genes > 2 fold			
	Observed	Expected	Observed	Expected	Observed	Expected		Chi squared (2df)	
0.9-0.95	2	0.75	2	2.86	0	0.39	4	0.2537	
0.85-0.9	3	1.31	4	5.01	0	0.68	7	0.2157	
0.8-0.85	16	5.42	10	20.75	3	2.82	29	0	***
0.75-0.8	16	4.86	8	18.6	2	2.53	26	0	***
0.7-0.75	16	5.61	9	21.47	5	2.92	30	0	***
0.65-0.7	11	5.42	16	20.75	2	2.82	29	2.93E-02	*
0.6-0.65	12	6.55	21	25.04	2	3.41	35	5.57E-02	
0.55-0.6	7	5.05	17	19.32	3	2.63	27	0.5819	
0.5-0.55	13	7.67	26	29.34	2	3.99	41	0.0789	
0.45-0.5	15	9.17	30	35.06	4	4.77	49	0.1018	
0.4-0.45	21	10.66	29	40.79	7	5.55	57	0.001	***
0.35-0.4	16	13.66	49	52.23	8	7.11	73	0.6998	
0.3-0.35	29	16.46	47	62.97	12	8.57	88	0.0006	***
0.25-0.3	29	20.02	64	76.56	14	10.42	107	0.0257	*
0.2-0.25	40	28.43	90	108.76	22	14.81	152	0.0033	**
0.15-0.2	34	34.98	128	133.8	25	18.22	187	0.2458	
0.1-0.15	57	50.88	185	194.62	30	26.49	272	0.4326	
0.05-0.1	96	78.57	279	300.52	45	40.91	420	0.0545	
0-0.05	339	466.53	1939	1784.54	216	242.93	2494	0.00E+00	***
Total	772	772	2953	2953	402	402	4127		

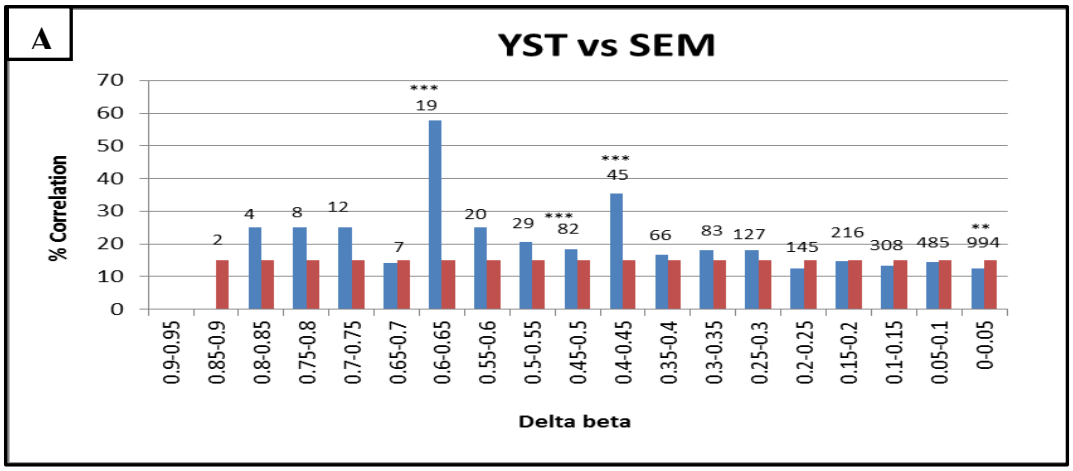
Islands

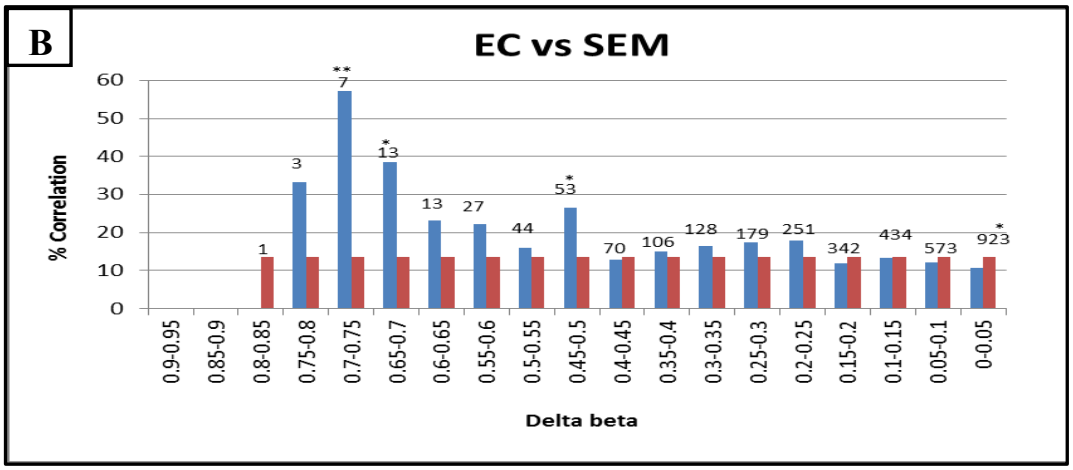

A
YST vs SEM
% Correlation
Delta beta
0.9-0.95
0.85-0.9
0.8-0.85
0.75-0.8
0.7-0.75
0.65-0.7
0.6-0.65
0.55-0.6
0.5-0.55
0.45-0.5
0.4-0.45
0.35-0.4
0.3-0.35
0.25-0.3
0.2-0.25
0.15-0.2
0.1-0.15
0.05-0.1
0-0.05

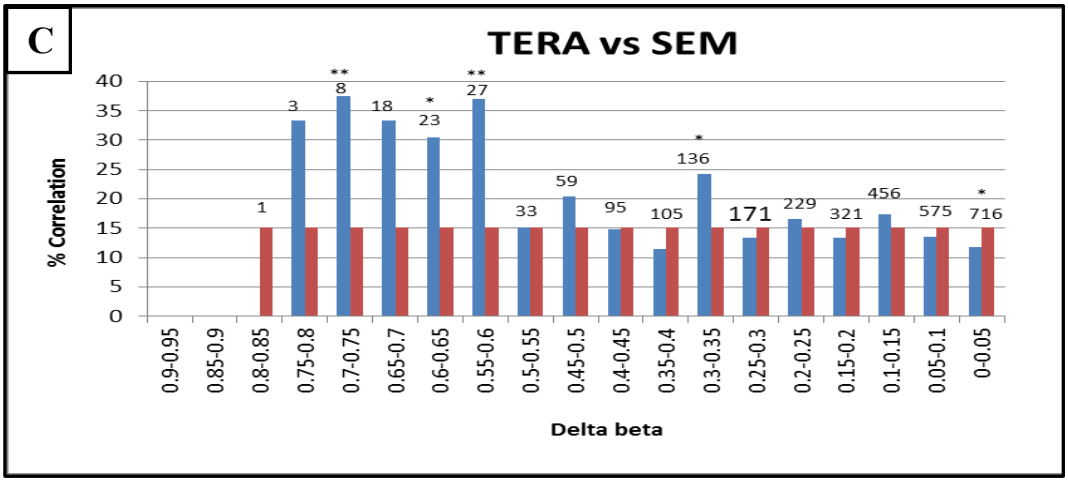

B
EC vs SEM
% Correlation
Delta beta
0.9-0.95
0.85-0.9
0.8-0.85
0.75-0.8
0.7-0.75
0.65-0.7
0.6-0.65
0.55-0.6
0.5-0.55
0.45-0.5
0.4-0.45
0.35-0.4
0.3-0.35
0.25-0.3
0.2-0.25
0.15-0.2
0.1-0.15
0.05-0.1
0-0.05


C
TERT vs SEM
% Correlation
Delta beta
0.9-0.95
0.85-0.9
0.8-0.85
0.75-0.8
0.7-0.75
0.65-0.7
0.6-0.65
0.55-0.6
0.5-0.55
0.45-0.5
0.4-0.45
0.35-0.4
0.3-0.35
0.25-0.3
0.2-0.25
0.15-0.2
0.1-0.15
0.05-0.1
0-0.05

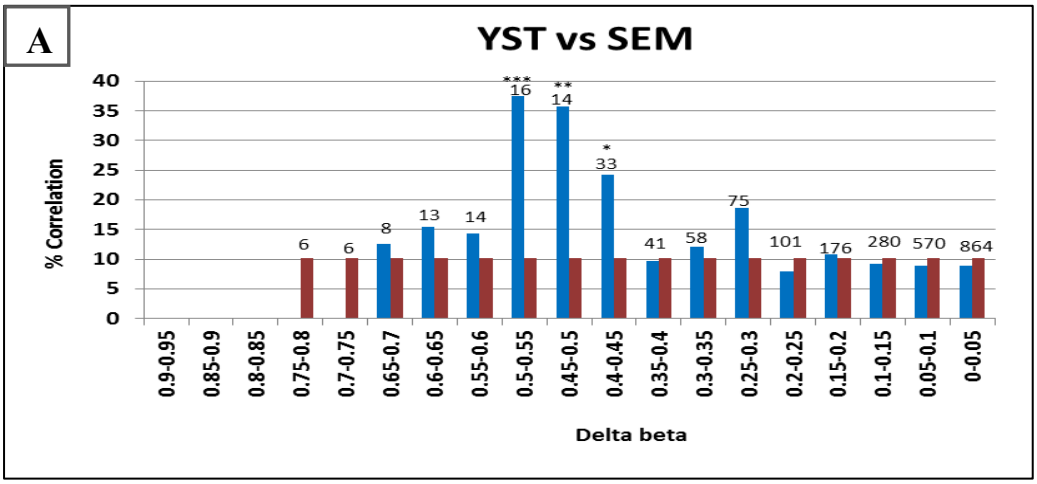

Observed Correlation
Expected Correlation

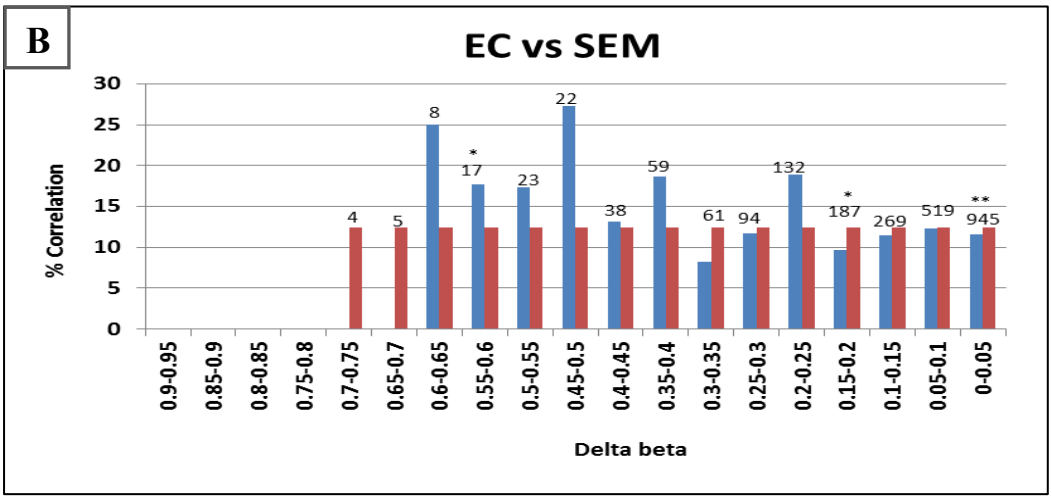

A
YST vs SEM
% Correlation
Delta beta
0.9-0.95
0.85-0.9
0.8-0.85
0.75-0.8
0.7-0.75
0.65-0.7
0.6-0.65
0.55-0.6
0.5-0.55
0.45-0.5
0.4-0.45
0.35-0.4
0.3-0.35
0.25-0.3
0.2-0.25
0.15-0.2
0.1-0.15
0.05-0.1
0-0.05


B
EC vs SEM
% Correlation
Delta beta
0.9-0.95
0.85-0.9
0.8-0.85
0.75-0.8
0.7-0.75
0.65-0.7
0.6-0.65
0.55-0.6
0.5-0.55
0.45-0.5
0.4-0.45
0.35-0.4
0.3-0.35
0.25-0.3
0.2-0.25
0.15-0.2
0.1-0.15
0.05-0.1
0-0.05


C
TERA vs SEM
% Correlation
Delta beta
0.9-0.95
0.85-0.9
0.8-0.85
0.75-0.8
0.7-0.75
0.65-0.7
0.6-0.65
0.55-0.6
0.5-0.55
0.45-0.5
0.4-0.45
0.35-0.4
0.3-0.35
0.25-0.3
0.2-0.25
0.15-0.2
0.1-0.15
0.05-0.1
0-0.05

North-Shelves





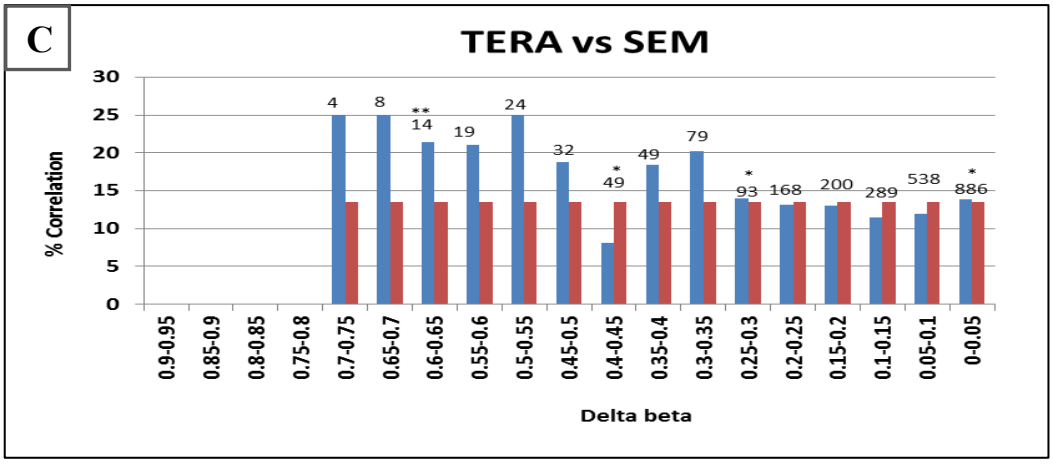


Figure 4.2. The histograms represent the correlation between differential gene expression of over two-fold and differential methylation between seminoma and non-seminoma at islands, shores and shelves: the observed (blue) and expected (red), percentage correlation for each delta-β-value was calculated as (Number of correlating genes/total number of genes *100) for seminoma versus YST (A), EC (B) and TERT (teratoma) (C). Significance of the chi-squared tests of association are shown (p=0.05*, 0.01**, 0.001***), and the total number of genes in each category is displayed above the bar.

These analyses revealed that the observed correlation differs greatly from the expected level at high delta-beta values (>0.6) and become more similar to expected or no difference at low delta-beta values Figure 4.2. Some lower delta-β-value categories show a statistical significant association due to a large number of genes in these categories thus in fact they did not reveal the real correlation where the percentage level of association is much smaller means that a correlation reflects a random association. For example, when comparing the correlation percentage with statistical significance in YST at different regions (CGIs, shores, shelves) (Table 4.2), there were 152 genes exhibiting delta-β-values between 0.2-0.25, with only 26% (40/152) showing a correlation with decreased expression, but this still has a p-value less than 0.05. On the other hand, some categories with the highest differential methylation values correlated with decreased expression and the percentage levels of association is high but these categories did not show a statistical significant association due to the small numbers of genes in these categories. For example, comparing YST with seminoma, seven genes exhibit delta-β-values between 0.85 and 0.9, of which three (43%) show a correlation with decreased expression. In addition, in comparing the YST and SEM cell lines, there was greater than two times the value expected at random. Therefore, regarding those outputs, a threshold of differential methylation was established to >0.65 and genes whose differential methylation <0.65 was excluded for further analysis.

Table 4.2: Comparing correlation percentage and significance in YST

Delta beta	No. of genes	CGIs		No. of genes	Shores		No. of genes	Shelves	
		%	p-value		%	p-value		%	p-value
0.9-0.95	4	50	0.254	-	-	-	-	-	-
0.85-0.9	7	43	0.216	2	0	0.696	-	-	-
0.8-0.85	29	55	0 ***	4	25	0.555	-	-	-
0.75-0.8	26	62	0 ***	8	25	0.711	6	0	0.687
0.7-0.75	30	53	0 ***	12	25	0.171	6	0	0.432
0.65-0.7	29	38	0.029 **	7	14	0.617	8	13	0.585
0.6-0.65	35	34	0.056	19	58	0 ***	13	15	0.727
0.55-0.6	27	26	0.582	20	25	0.176	14	14	0.781
0.5-0.55	41	32	0.079	29	21	0.678	16	38	0.001 *
0.45-0.5	49	31	0.102	82	18	0 ***	14	36	0.004 **
0.4-0.45	57	37	0.001 ***	45	36	0 ***	33	24	0.026 *
0.35-0.4	73	22	0.699	66	17	0.769	41	10	0.125
0.3-0.35	88	33	0 ***	83	18	0.323	58	12	0.180
0.25-0.3	107	27	0.026 *	127	18	0.584	75	19	0.017
0.2-0.25	152	26	0.003 **	145	12	0.696	101	8	0.372
0.15-0.2	187	18	0.246	216	15	0.405	176	11	0.411
0.1-0.15	272	21	0.433	308	13	0.617	280	9	0.367
0.05-0.1	420	23	0.055	485	14	0.582	570	9	0.627
0-0.05	2494	14	0 ***	994	13	0.008 **	864	9	0.147
Total	4127			2652			2275		

4.2.2: Identifying genes that are both differentially methylated and differentially expressed between seminoma and non-seminoma cell lines

Based on the results indicating that the methylation of CGIs near the TSS could contribute to the expression of genes, genes were identified and listed in Figure 4.3 that were both differentially methylated and reciprocally differentially expressed between non-seminoma and seminoma cell lines (these data presented in collaboration with MSci student, Matthew Carr). Genes were selected only when the difference of expression between seminoma and non-seminoma was two-fold greater in SEM according to the microarray array data and was significantly differentially methylated with a delta-β-value ≥0.65 at CGI near the TSS.

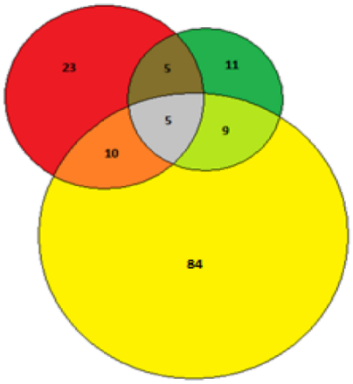
The YST cell line had the most genes that were uniquely methylated and which exhibited a reduced expression compared with seminoma (84 genes) as shown in Venn diagram (Figure 4.3-A). EC and teratoma had 23 and 11 methylated and less strongly expressed genes, respectively, whereas three genes were methylated with reduced expression showed in both EC and teratoma.

A

Genes unique to GCT27		
C11orf70	CHODL	CNTLN
CRIP1	CTSO	CXCR7
CYP2R1	DMRT1	GLDN*
H19	HSPA12B	IQCA1
KDELR3	KRT7	LRRC34
MGLL	MGMT	NUDT15
NUDT19	RPL39L	SLC7A4
UCP2	ZC3HAV1L	

Genes in GCT27 and NT2D1	
ABCA13*	C5orf58
GREM2*	PRKCH*
SDR42E1	

Genes unique to NT2D1		
CAT	CHMP4C	CR1L
GALNT3	LPHN3*	ME1
PRMT6	QPCT	RBM12B
SULT1A1	VAV1*	



Genes in GCT44 and NT2D1		
CYP1B1*	ECHDC2	EDNRB
ELMO3	EPS8L1	PKP3
RNLS	SOX15*	TEAD3*

Genes in GCT27 and GCT44			
ACOT4	BST1	CYB5R2	DUSP23
HSPA2	LTBP2	MAPK13	PLD6
TDRD12	TUSC1		

Genes in all three cell lines	
DDX43	PON3
RBMXL2	ST14
TACSTD2	

Genes unique to GCT44			
ABHD14A	ACADL*	ALDH1A3	ALX1
AMPD3	ARPC1B*	B3GNT7	BASP1
C17orf104	C1orf51	C22orf32	C6orf132
C8orf42*	CD274	CLDN7	CST6
CTHRC1	CXCL14	CYB5R1*	DSCR6*
ECHDC3	EPCAM*	EPS8L2	ESRP2
F2RL1*	FABP5	FAM109B	FAM184A
FAM84B*	FUT1	FUT11	GGCT
GIPC2	GPX7	GRAMD2*	GSTO2
GUCA1A*	HAS3	HCG11*	HIST1H4C
HSBP1L1	HYLS1	IER2*	JMJD8
KLF4*	LGALS3	LGALS8	LMO7*
LRRFIP1	LY75	MEIS1*	MNS1
MTL5*	MYD88*	NAAA	OVOL1
OXCT1	PARP12	PLEKHG4	PLXND1*
PPP1R14A	PRDM14	RASD1*	RASGEF1A*
RCSD1	RNLS*	RPRM	SDK1
SH2D3A	SLC10A4	SLC25A38*	SOX17
SUSD3*	TMEM168	TMEM179B	TNFRSF10C
TNFRSF10D*	TRIL	TRIM25	TRIM59
TTYH2	TUBB6	ZIC1	ZMIZ1*

B

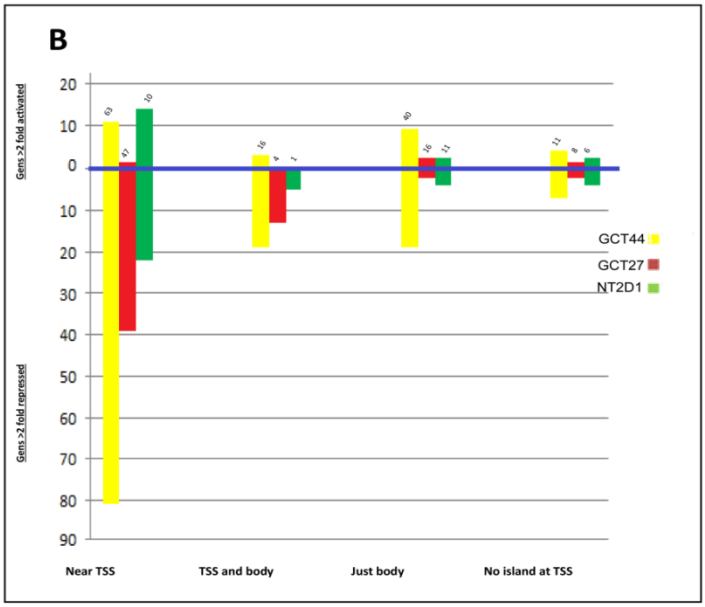


Figure 4.3: Venn diagram represents significantly correlating genes which were both differentially methylated and expressed in all three nonseminomas, relative to the seminoma cell line : A) Tables contain genes that are expressed (≥2-fold difference in expression) and also differentially methylated (delta-β ≥0.7 for EC and teratoma, ≥0.65 for YST). Colours in the tables match those of the Venn diagram. B) The numbers of genes which showed differential expression and were differentially methylated.

4.2.3 Correlation between methylation of CpG islands in gene bodies and gene silencing

Some researchers have argued that the methylation of CGIs in gene bodies is more correlated to activation of genes than to depression (Aran *et al*., 2013, Yang *et al*., 2014). However, using the same criteria to analyse the correlation between methylation of CGIs and gene expression as applied to the promoter regions where delta-β- value >0.65 for differential methylation and a two-fold difference in gene expression, the results revealed that increased methylation of CGIs in the body was more strongly associated with gene silencing than activation (Figure 4.3-B). The result of this numerical assessment indicated that 45 out of 128 genes in the YST cells line exhibiting increased methylation of CGI in the body compared with the seminoma cell line showed a decrease in gene expression by two-fold or greater and a similar relationship was seen in the EC and teratoma cell lines (Figure 4.3-B). There was a strong correlation between high methylation of CGIs in the gene body and low levels of expression in the identified genes, regardless of whether this association was seen also in CGI promoter region near TSS.

4.2.4 Optimizing conditions for RT-qPCR experiment

In order to validate the difference in expression levels of genes that were analysed in the microarray assay (section 4.2.2), RT-qPCR was used. First, RT-qPCR experiment required optimization steps including designing optimal primer pairs, determining primer efficiency and the standard threshold of each set of primers.

The primer pair used in SYBR RT-qPCR was designed using the primer3 web site http://primer3.ut.ee/ according to the specific rules (following the protocol in section 2.2.4.2). The standard threshold of each set of primers was determined according to the qPCR results of five fold serial template cDNA dilution from the seminoma cell line sample following the protocol in section 2.2.4.5 (Figure 4.4).

The standard curve plots showed the PCR cycle number (C_T) versus log cDNA quantity, which should produce a straight line with a slope of -3.3 with acceptable limits between -3.1 to -3.6 (Fraga *et al*., 2008). In addition to the efficiency value which is between 90 and 110, R^2 value (a statistical term that indicates how good one value is at predicting another) should be >0.99 (Fraga *et al*., 2008) . Moreover, the `no reverse transcriptase control` (-RT) and the `no template control` (NTC) samples should not show any fluorescent signal. The best primer efficiency is determined when the melting curve of each primer shows only one peak which means it is specific and producing no primer dimers (Figure 4.5).

Each RT-qPCR experiment was performed in three triplicates for each selected gene (Figure 4.6). All the RT-qPCR expression results of target genes were normalized against the expression of the housekeeping gene β-actin because the expression of this gene was relatively stable for all cell lines and treated cells used in this study

while others examined housekeeping genes (GAPDH, Cyclophilin B, and HPRT1) were not (data not shown). The Pffafl equation was used to determine the relative expression of the target gene (Pfaffl, 2001).

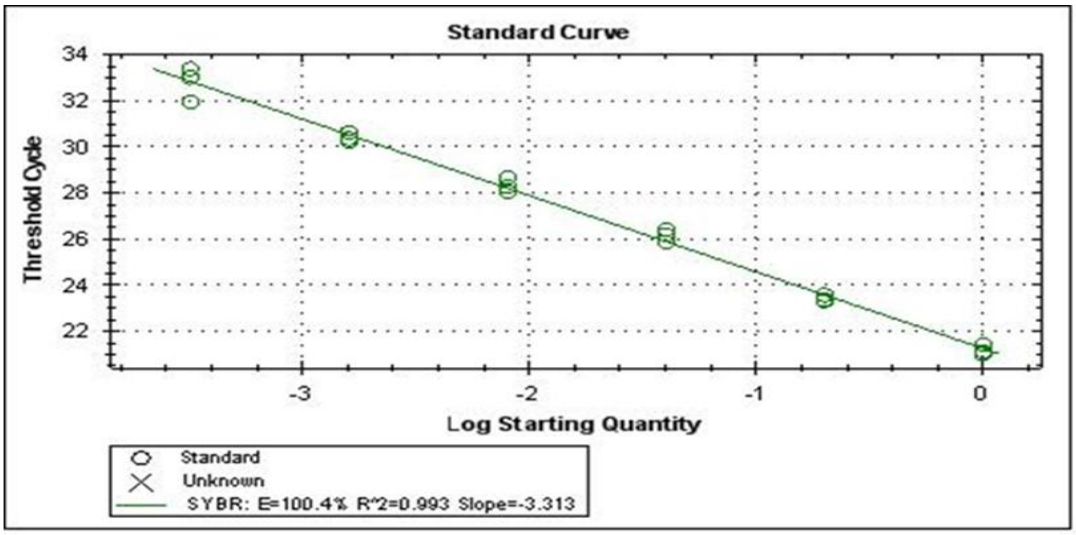


Figure 4.4: Example of optimization of primer efficiency for EPCAM primer. Optimization was performed by using five fold serial dilutions of cDNA from seminoma cell line as template.

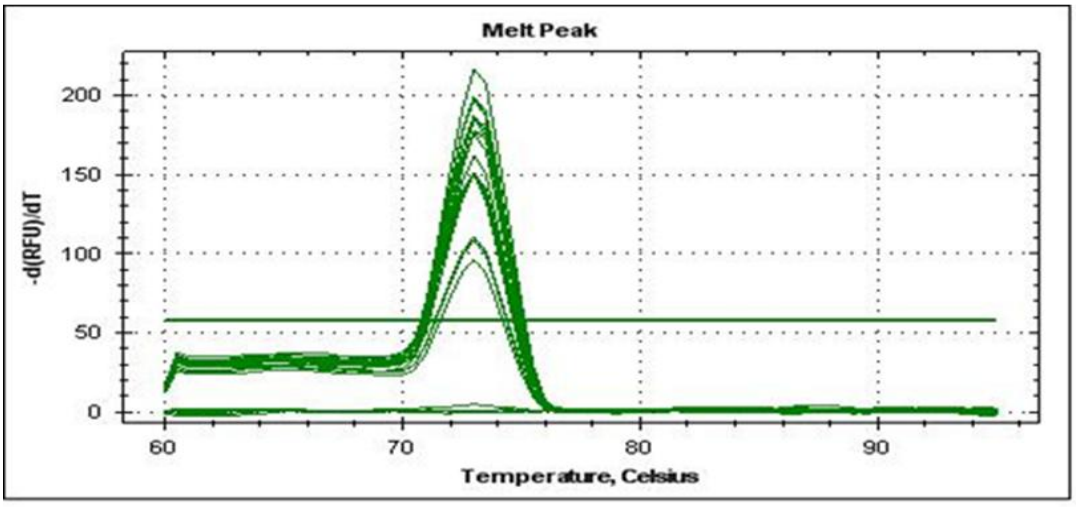


Figure 4.5: Example of melting curve for β-actin: A symmetrical curve of a single peak represents the ideal melting curve.

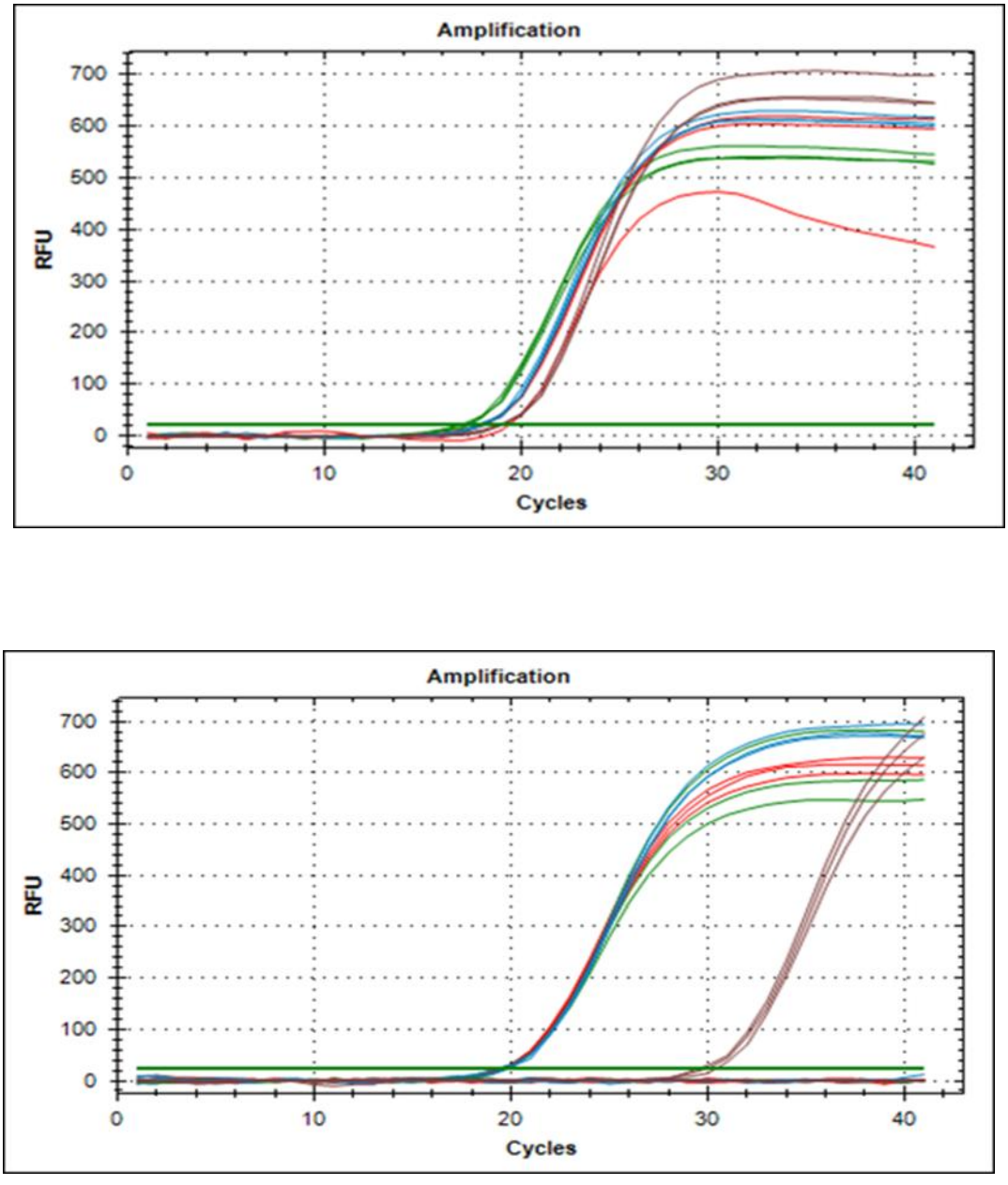


Figure 4.6: Example of graphs represent the amplification plots for β-Actin and EPCAM, respectively. Relative fluorescence (RFU) is plotted against PCR cycles number for treated and untreated samples.

4.2.5 Identifying genes that could implicated in a phenotypic difference between seminoma and YST cell lines

Some genes that were methylated and repressed in YST relative to SEM could play a potential role in the phenotypic difference between those cell lines. To validate whether those genes are indeed implicated in a phenotypic difference between seminoma and YST, the result in this study was compared with the Affymetrix expression data of Palmer *et al*. (2008) study which included primary tumour samples of a cohort of paediatric seminomas and YSTs from many different anatomical locations using the same array that I used in this study, the Affymetrix HumanGeneChip U133 Plus 2.0 arrays (Figure 4.7).

Table 4.2 shows 17 out of 147 genes which revealed a significant correlation between differential methylation and differential expression in YST versus SEM cell lines. Gene ontology (www.geneontology.org) along with the possible biology importance of these genes showed their importance in development and differentiation. Notably, several of these genes are involved in pluripotency and male gamete production: KLF4, PRDM14, DDX43 and TDRD12. Twenty one genes were consistently and significantly expressed at higher levels in primary seminomatous tumours than in YST (Table 4.3). 339 genes were differentially expressed in both studies. Eleven of those genes also were highly expressed in the seminoma cell line compared with YST in this study. The overlap of genes which showed differential expression in the cell line study and the primary tumour study indicated that the cell lines could be used as a good model for studying gene expression in this tumour type.

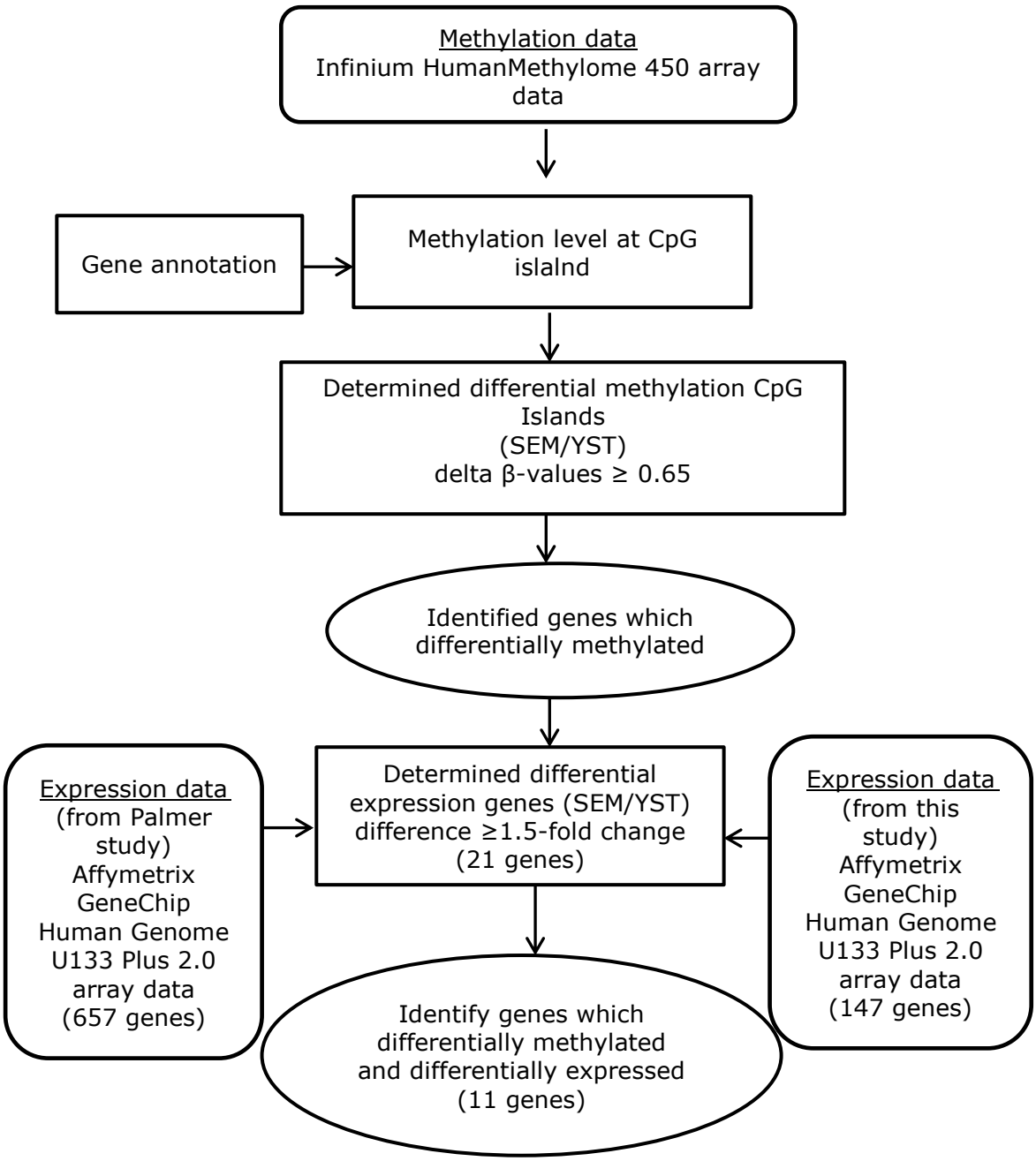


Figure 4.7: Methods overview to identify the list of genes (Table 4.3) that were methylated and repressed in YST relative to SEM

Table 4.3: Genes methylated and repressed in YST relative to SEM

Gene	Infinium Methylation450K Array			Affymetrix gene expression	Function/Importance
	β-value		Delta β-value	Fold change (SEM/YST)	
	YST	SEM			
EPCAM	0.907	0.125	0.782	9.13	Induces cell proliferation (Münz *et al*., 2004) – Oncogenic signalling molecule (Baeuerle and Gires, 2007)
SOX17	0.949	0.108	0.841	5.36	Regulation of embryonic development (Kamachi *et al*., 2000)
HIST1H4C	0.942	0.123	0.819	4.09	Core component of nucleosome (Churikov *et al*., 2004)
SOX15	0.949	0.108	0.841	1.80	Protein binding regulation of transcription (Maruyama *et al*., 2005)
GGCT	0.859	0.064	0.795	6.91	Cell proliferation (Oakley *et al*., 2008)
CTHRC1	0.944	0.064	0.88	8.35	A novel secreted protein in injured arteries and promotes cell migration (Pyagay *et al*., 2005)
TRIM59	0.895	0.134	0.761	4.72	Ubiquitin-protein ligase activity- Play role in innate immunity (Ozato *et al*., 2008)
LY75	0.841	0.172	0.669	2.57	Immune response (Kurup *et al*., 2007)
PRDM14	0.898	0.203	0.695	1.67	Transcription factor- Germ cell development- pluripotency- epigenetics reprogramming (Yamaji *et al*., 2008)
KLF4	0.733	0.082	0.651	5.09	Transcription factor- maintaining embryonic stem cells (Okita *et al*., 2007)
DDX43	0.872	0.098	0.774	5.22	RNA helicase is associated with RNA processes (Abdelhaleem, 2004)
BST1	0.923	0.152	0.771	2.12	Promotes pre-B-cell growth (Yokoyama *et al*., 2015)
RPRM	0.838	0.052	0.786	1.59	Induces cell cycle arrest (Xu *et al*., 2012)
PON3	0.949	0.229	0.72	1.61	Anti-oxidative mechanism- protects against apoptosis (Schweikert *et al*., 2012)
TRIL	0.853	0.103	0.75	4.31	Innate recognition of microbial products (Carpenter *et al*., 2009)
TDRD12	0.953	0.329	0.624	6.39	ATP-binding RNA helicase- Unique piRNA biogenesis factor (Pandey *et al*., 2013)
MGMT	0.179	0.061	0.118	1.14	Involved in DNA repair (Sharma *et al*., 2009)

Table 4.4: List of genes that were differentially methylated and expressed in my study and significantly differentially expressed in the Palmer *et al* (2008) study

Gene	Fold change	Palmer fold change	Palmer P-value
GGCT	120.36	1.78	0.01483
TDRD12	84.38	7.90	0.00002
RBMXL2	43.48	1.78	0.00359
DDX43	37.32	4.15	0.00562
KLF4	19.41	6.07	0.00004
ECHDC3	14.53	1.59	0.00173
TRIL	14.14	1.43	0.01529
OXCT1	17.24	1.68	0.00007
TMEM168	9.87	1.67	0.00324
GUCA1A	7.14	1.63	0.00121
LY75	5.95	3.86	0.00308
PARP12	4.35	2.97	0.00086
MNS1	3.61	1.10	0.00613
SOX15	3.49	2.37	0.00017
PRDM 14	3.18	1.40	0.00329
PON3	3.04	2.62	0.00246
RPRM	3.01	2.54	0.00011
LRRFIP1	2.83	1.24	0.00600
ARPC1B	2.60	2.41	0.00015
NAAA	2.42	1.20	0.01633
AMPD3	2.06	1.52	0.03360

4.2.6 Validation of the relationship between methylation and gene silencing

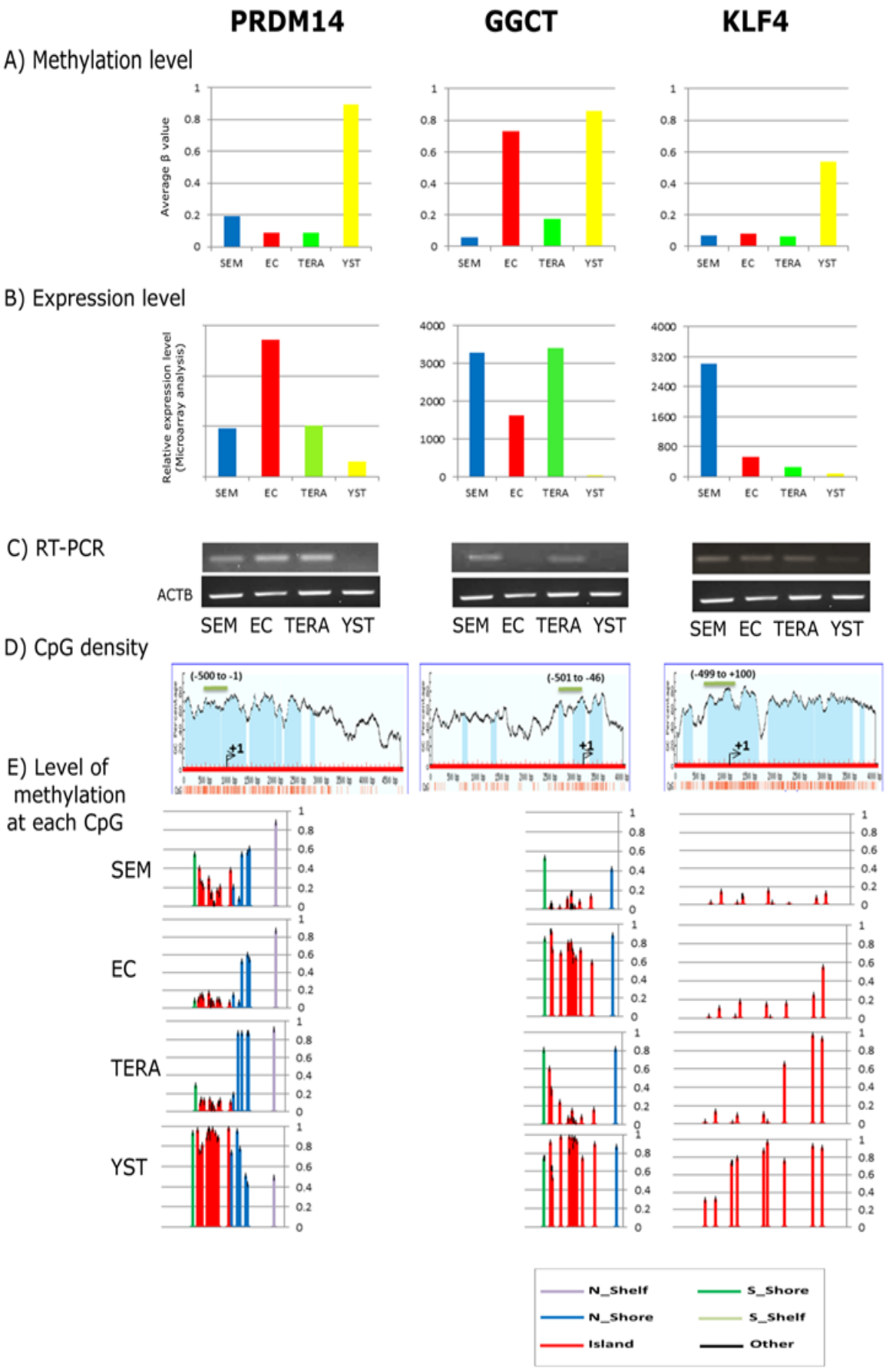
Seventeen of those genes that showed significant correlation of differential methylation and differential expression between the non-seminoma and seminoma cell lines were analysed for further validation. The methylation level of CGI near TSS according to methylation data of Infinium array for these genes is presented in Figure 4.8-A. This shows that the average methylation level (β-value) at the CGIs was high in non-seminoma cell lines, particularly in YST, versus SEM. Conversely, the expression level as analysed by Affymetrix expression arrays is presented in figure 4.8-B showing that those genes are highly expressed in SEM and downregulated in non-seminoma or silenced in YST.

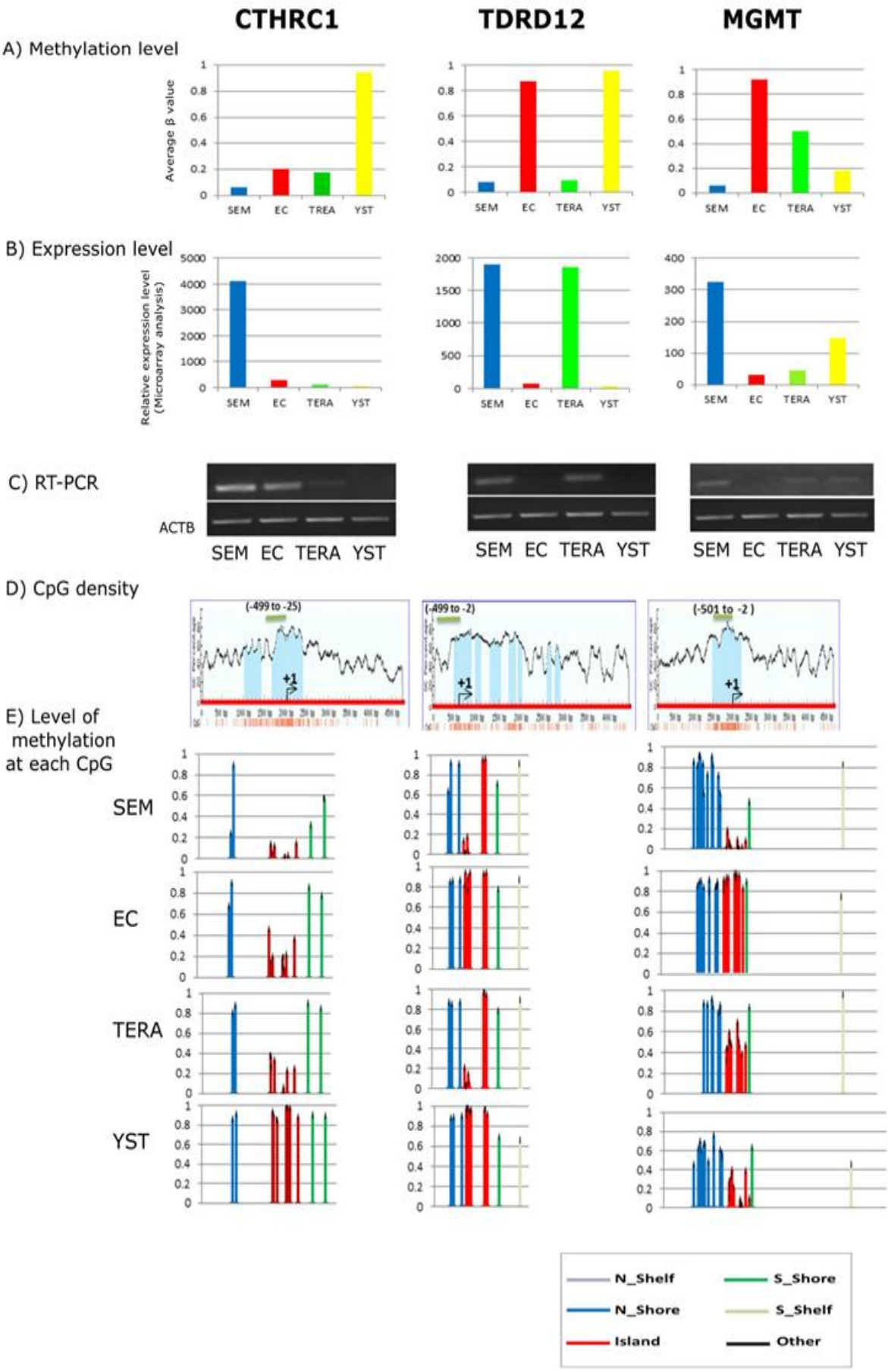
The results verified a significant negative correlation between CGI methylation near TSS and expression for all of those genes. In cases where hypermethylation was detected, transcription was often downregulated in tumour cell lines (as in YST) compared with high expression when the gene was hypomethylated (as in SEM).

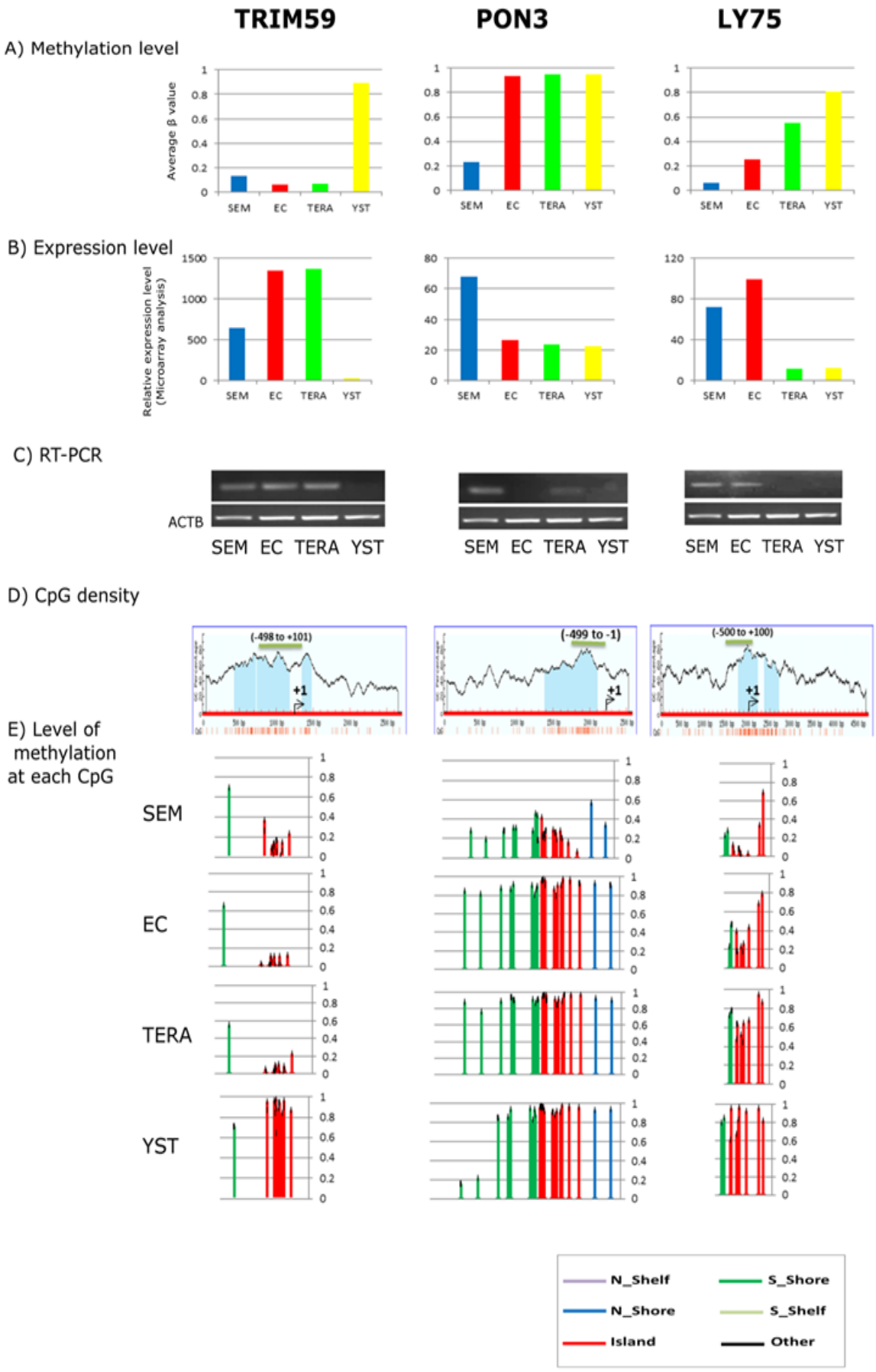
In order to validate the expression of those candidate genes that were identified by the Affymetrix array, reverse transcriptase PCR experiment was performed. RNA was extracted from the same four cell lines used in the microarray analysis using the TRI reagent and chloroform protocol as described in section 2.2.3.2. Any contaminating DNA was removed by applying DNAase treatment then cDNA synthesis was carried out using Superscript III. PCR products were analysed by agarose gel electrophoresis (Figure 4.8-C). The results indicated that Affymetrix expression array data truly reflected the expression changes of those genes.

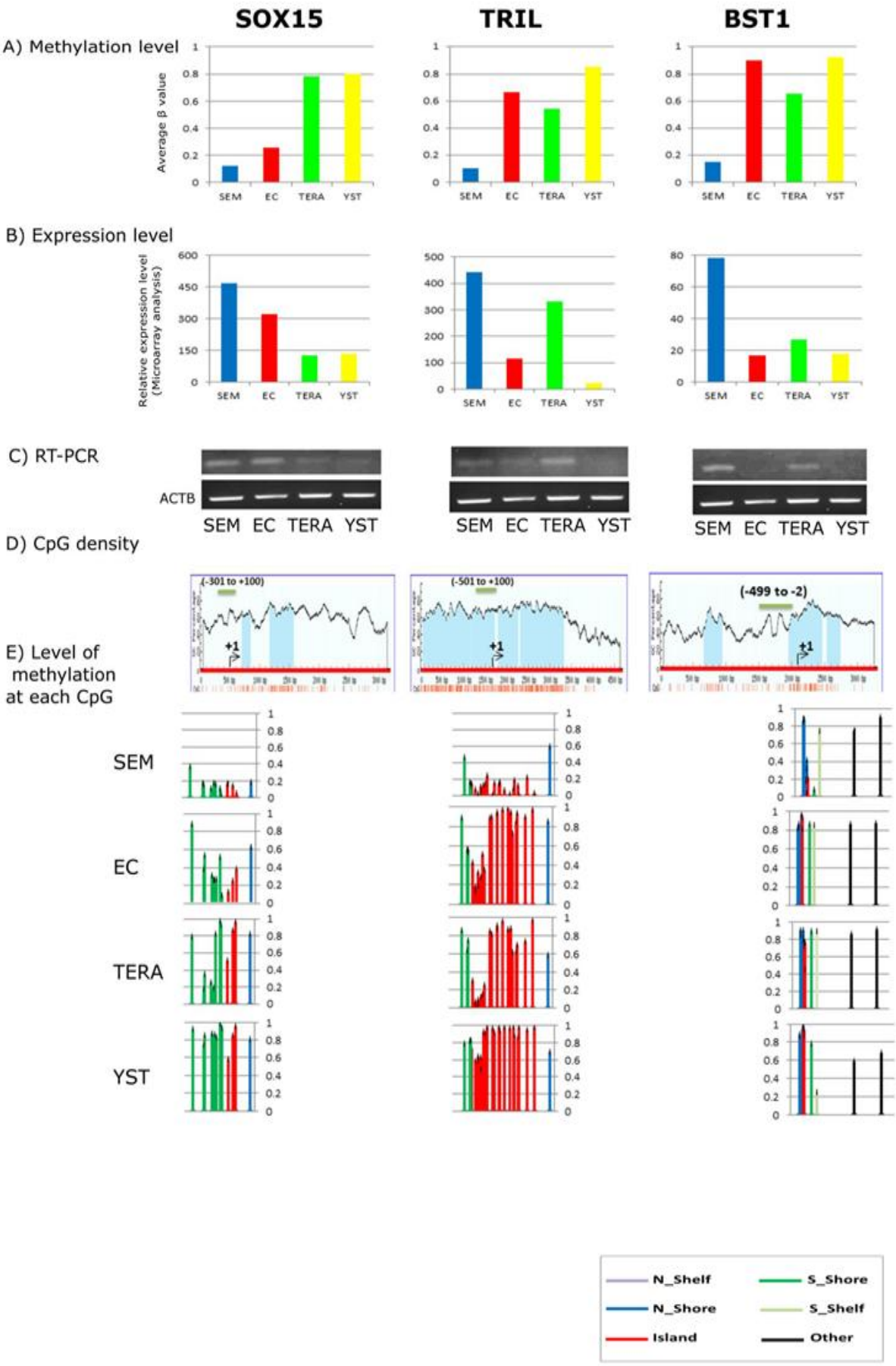
The positions of all the CpG sites and islands within each candidate gene were identified with reference to the gene sequence (regarding to NCBI website) by submitting that sequence to the MethPrimer programme (http://www.urogene.org/methprimer/). The locations of the transcription start site (TSS) and the promoter region were predicted using the Eukaryotic Promoter Database (http://epd.vital-it.ch). From Figure 4.8-D, it is clear that the methylated CGIs were concentrated near TSS. Moreover, a high level of methylation at CGIs near the TSS was strongly correlated with gene silencing as shown in Figure 4.8-E meaning that methylation of CGI around the promoter regions could play an important role in silencing of the methylated genes.

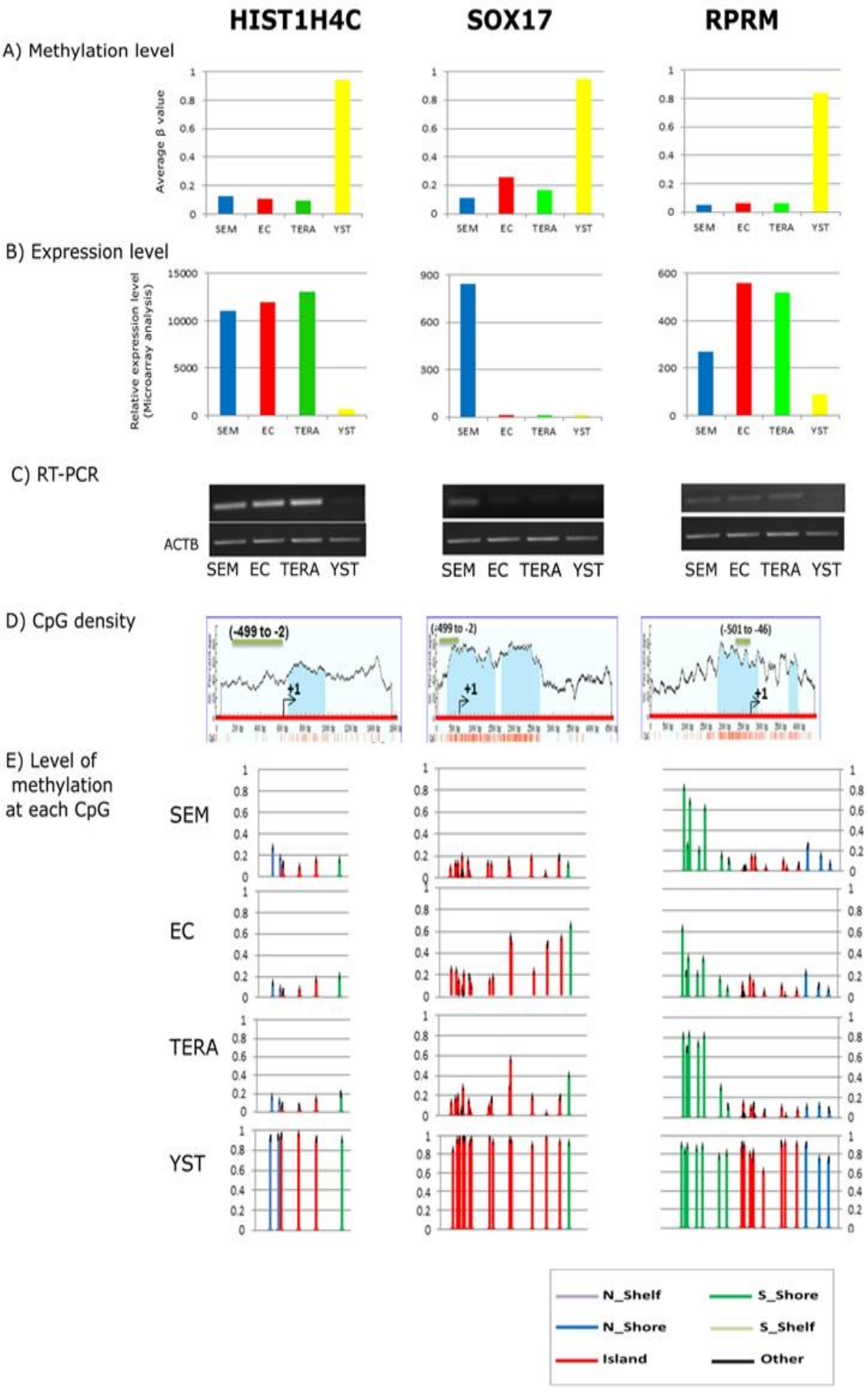
Furthermore, the methylation level of the identified CGIs located near to TSSs was analysed (Figure 4.8) to determine to which extent this level related to gene expression, most of the genes which showed hypermethylation of CGIs in non-seminomas correlated with silencing of the gene (such as GCCT, KLF4, and PON3). However, for some genes, such as TRIL, only some of CpGs-associated CGI in EC and TERT were methylated while others were not and the gene was downregulated in EC and expressed in TERA. On the other hand, when the CGIs of TDRD12 were analysed, approximately half of the CpGs were hypomethylated in SEM and TERT and the other half were hypermethylated but the gene was expressed in both cell lines. We conclude from this result that some CpGs could be more important than others in relation to gene expression.


PRDM14
GGCT
KLF4
A) Methylation level
Average β value
SEM EC TERA YST
B) Expression level
Relative expression level (Microarray analysis)
C) RT-PCR
ACTB
SEM EC TERA YST
D) CpG density
(-500 to -1)
(-501 to -46)
(-499 to +100)
+1
E) Level of methylation at each CpG
SEM
EC
TERA
YST
N_Shelf
N_Shore
Island
S_Shore
S_Shelf
Other


CTHRC1
TDRD12
MGMT
A) Methylation level
Average β value
SEM EC TREA YST
B) Expression level
Relative expression level (Microarray analysis)
SEM EC TERA YST
C) RT-PCR
ACTB
SEM EC TERA YST
D) CpG density
(-499 to -25)
(-499 to -2)
(-501 to -2)
+1
E) Level of methylation at each CpG
SEM
EC
TERA
YST
N_Shelf
N_Shore
Island
S_Shore
S_Shelf
Other


TRIM59
PON3
LY75
A) Methylation level
Average β value
SEM EC TERA YST
B) Expression level
Relative expression level (Microarray analysis)
C) RT-PCR
ACTB
D) CpG density
(-498 to +101)
(-499 to -1)
(-500 to +100)
+1
E) Level of methylation at each CpG
SEM
EC
TERA
YST
N_Shelf
N_Shore
Island
S_Shore
S_Shelf
Other


SOX15
TRIL
BST1
A) Methylation level
Average β value
SEM EC TERA YST
B) Expression level
Relative expression level (Microarray analysis)
C) RT-PCR
ACTB
D) CpG density
E) Level of methylation at each CpG
(-301 to +100)
(-501 to +100)
(-499 to -2)
+1
SEM
EC
TERA
YST
N_Shelf
N_Shore
Island
S_Shore
S_Shelf
Other


HIST1H4C
SOX17
RPRM
A) Methylation level
Average β value
SEM EC TERA YST
B) Expression level
Relative expression level (Microarray analysis)
C) RT-PCR
ACTB
D) CpG density
(-499 to -2)
(-499 to -2)
(-501 to -46)
+1
E) Level of methylation at each CpG
SEM
EC
TERA
YST
N_Shelf
N_Shore
Island
S_Shore
S_Shelf
Other

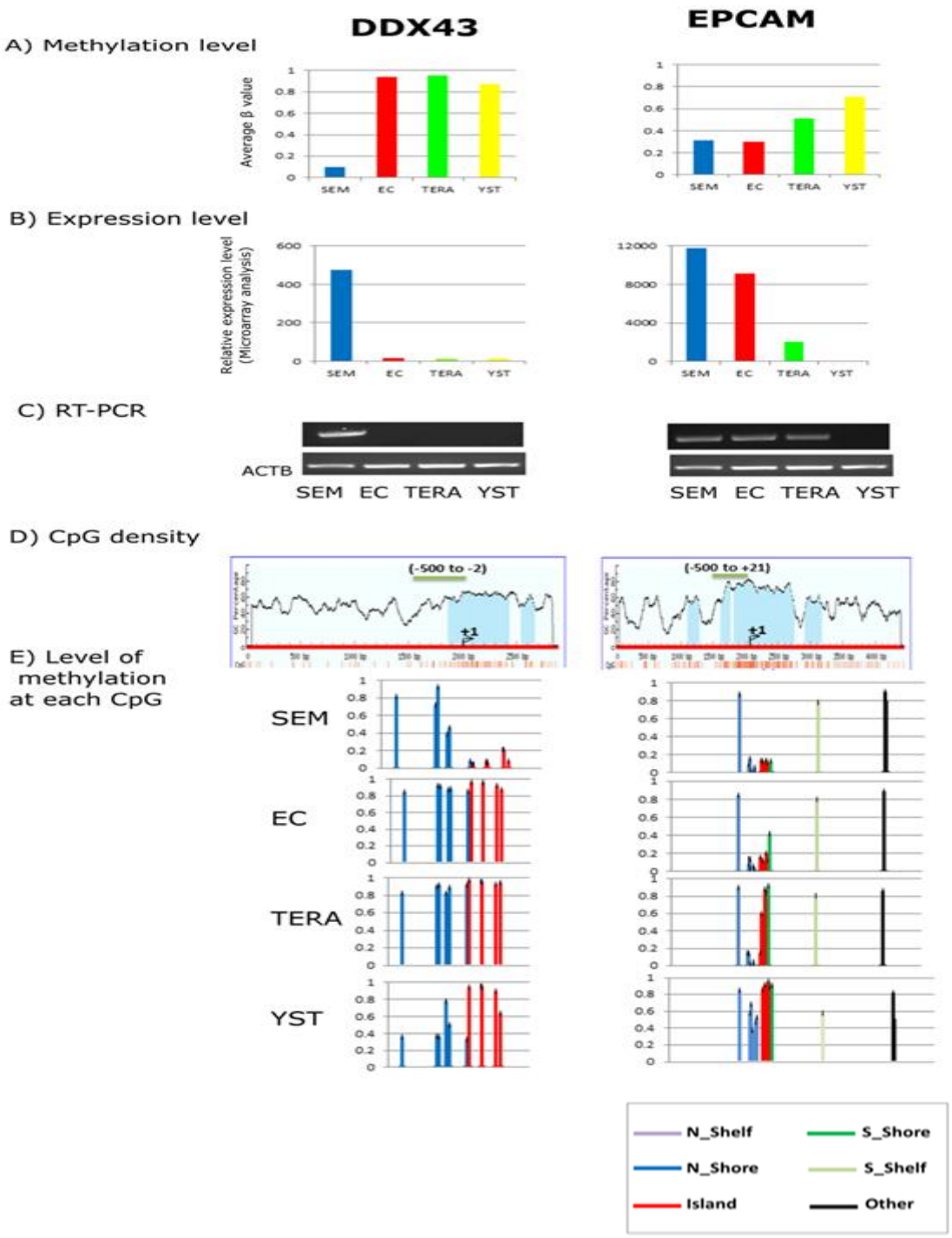


Figure 4.8: Figure represents a side by side comparison of level of CGI methylation to gene expression: a significant negative correlation between CGI methylation and expression were detected for each candidate gene. (A) methylation level across CGI in the four cell lines (B) expression level from microarray analysis (C) RT-PCR analysis (D) graph of CpG density across the gene structure where green bar and +1 indicate position of the promoter and TSS, respectively. (E) Methylation level at each CpG included in 450K chips relative to CpG site features in panel (D).

4.2.7 Validation of genes silenced by methylation

This study revealed a negative correlation between gene methylation and expression, indicating DNA methylation could play a role in silencing of those genes. To confirm that those genes were silenced by methylation, the expression of five selected genes was examined after treating the YST cells with a demethylating agent, 5-aza-2-deoxycytidine (5-aza), that has been widely used as a DNA methylation inhibitor to induce gene expression (Christman, 2002). The YST cell line was chosen for this analysis because it was highly methylated (Figure 4.8). YST cells were treated with 5μM of 5-aza for 2 days. Following treatment, expression of five out of the 17 genes (in Table 4.3), which their expression was highly in seminoma comparing to YST (fold change >4), was re-examined by RT-PCR and RT-qPCR. The primer efficiencies for candidate genes that were used in the Pfaffl equation are shown in Table 4.4. The result showed that all five genes were re-expressed following 5-aza treatment which verified that methylation of these genes could be associated with their silencing. The greatest effect of demethylation was seen for HISTH4C1, SOX17 and TDRD12, which were induced more than 200-fold in treated cells with a significant difference compared with untreated cells with p-value <0.001 using a paired t-test (Figure 4.9).

Table 4.5: Primer efficiency values obtained from standard curves for candidate genes

Gene name	Primer efficiency value %	R2	Slope of standard curve
ACTNB	98.9%	0.994	-3.349
EPCAM	100.4%	0.993	-3.313
TDRD12	100.5%	0.997	-3.309
HIST4H1C	99%	0.998	-3.384
SOX17	98.3%	0.986	-3.363
KLF4	99%	0.993	-3.443

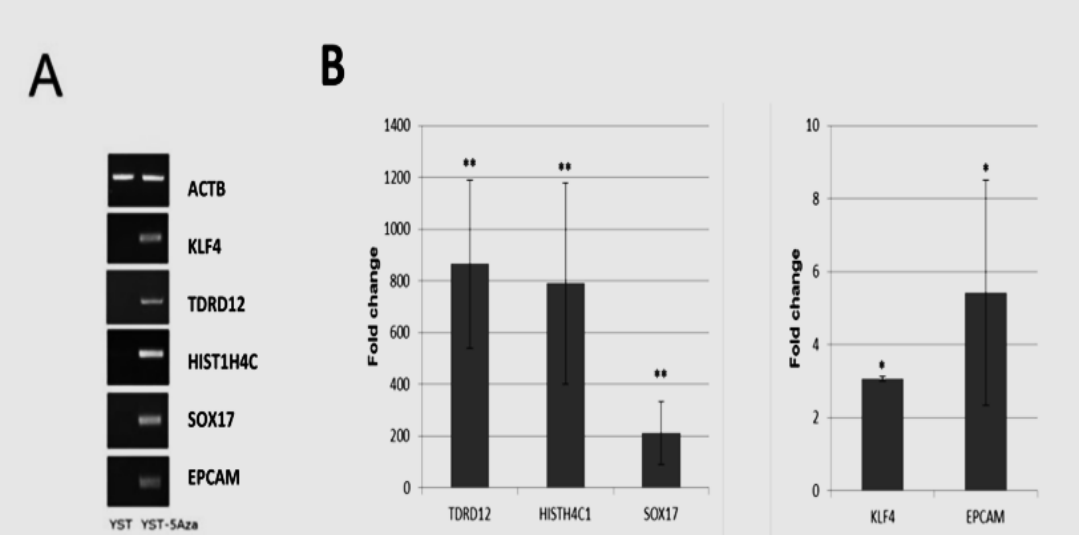


Figure 4.9: Gene expression analysis for selected genes in YSTs following treatment with 5-aza: (A) Gel image for PCR showing re-expression of 5 genes following treatment with 5-aza (B) RT-qPCR results for the same experiments. *denotes p-value <0.01 and **denotes p-value <0.001 using paired t-test. Error bars represent standard deviation.

4.2.8 Potential relationship between methylation and cancer

To detect which of the genes that show correlation between hypermethylation and reduced expression might be implicated in cancer-related events, the methylation levels of TSS-associated CGI were compared between GCT cell lines and a series of normal tissue samples for those genes. Infinium HumanMethylome450 array methylation data, including 21 normal tissues samples which included 4 prostate samples, 5 whole blood samples and 12 blood samples from 4 new born, 5 adults, and 3 adults >80 years old, was provided by Dr. Jeyapalan. High similarity in methylation patterns was found between all the normal samples and the seminoma cell line. Only three genes (DDX43, RBMXL2, and TDRD12) that showed a difference in methylation were hypomethylated in SEM versus normal tissue. Those genes known as testis-specific genes and were also highly methylated in all control samples and non-seminoma cell lines compared with seminoma (Figure 4.10).

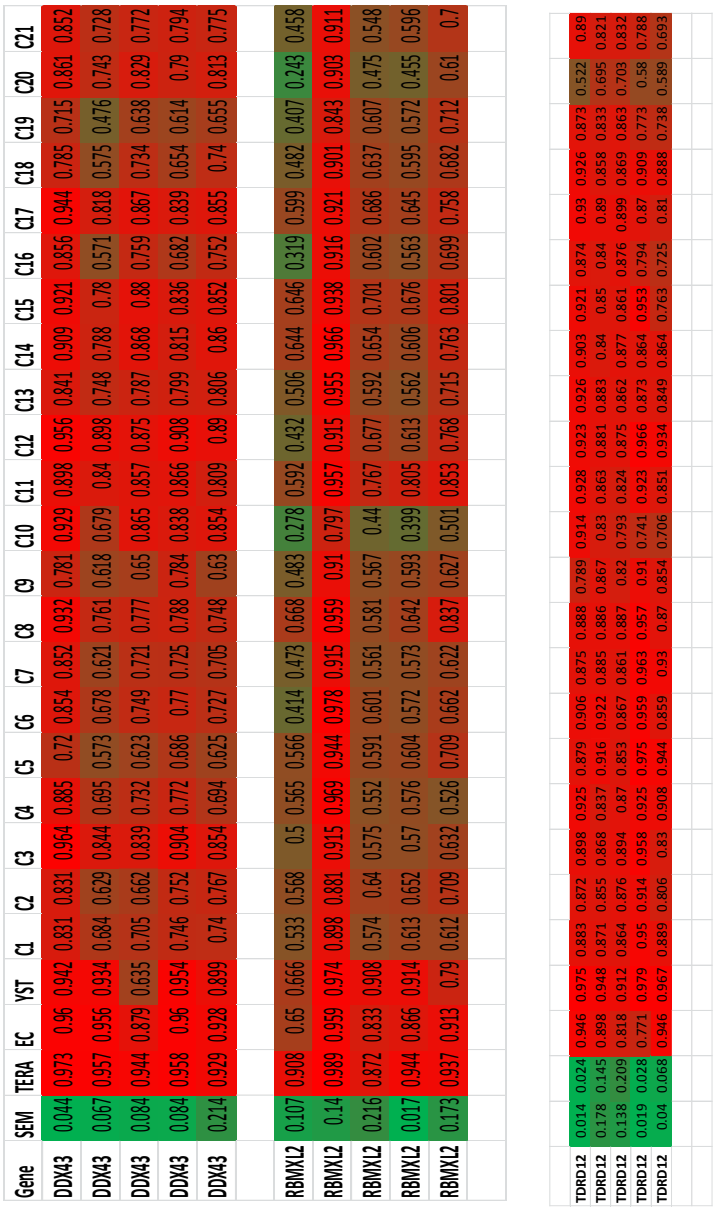


Figure 4.10: Comparison of the methylation level (β-value) between GCT cell lines and normal tissues. Red colour represents high methylation while green represents low methylation. C1-C4: normal prostate tissue, C5-C9: Whole blood samples, C10-C12: Blood samples from > 80 year olds, C13-C15 and C20-C21: Adult blood samples, and C16-C19: new-born blood samples.

4.3 Discussion

In this chapter, one of the most significant findings of this study that emerged from combining methylation and expression array data is a correlation of CGI hypermethylation with reduced gene expression in the non-seminomas compared with the seminoma cell lines, which agrees with the hypothesis that methylation of CGIs correlates with gene silencing (Robertson and Jones, 2000, Baylin, 2005, Esteller, 2007). The methylation of islands showed a stronger correlation with reduced expression than the methylation of shores or shelves.

A list of genes that were both differentially methylated and differentially expressed in non-seminoma cell lines relative to the SEM cell line was identified. The YST cell line had the most genes that were uniquely methylated and exhibited a reduced expression compared to seminoma (59 genes). Among genes differentially methylated at a TSS-associated CGI between non-seminoma and seminoma cell lines, about half showed a correlating decreased expression in the non-seminoma cell lines. Thus, it seems that differential methylation could play a substantial role in the differential gene expression seen between seminoma and non-seminoma cells.

Furthermore, when comparing the relationship between hypermethylation of CGIs in gene bodies to gene expression, the correlation was more often associated with silencing of genes rather than activation.

147 genes, out of 7242 that produced reliable signals in the expression array data, are both differentially methylated and expressed in non-seminoma relative to seminoma. 21 genes in our data set were also differentially expressed in cohorts of primary tumour. Nine of those genes have known importance in development and differentiation

processes. Moreover, some of these genes are involved in pluripotency and male gamete production such as KLF4, PRDM14, DDX43 and TDRD1. These genes could play a potential role in gene silencing that lead to the phenotypic difference between those cell lines.

Based on previous findings that explored a negative correlation between methylation and expression, we investigated whether methylation played a critical role in silencing of genes using the DNA-demethylating agent 5-aza-2-deoxycytidine. The expression of candidate genes was reactivated after treatment with 5-aza which verified that methylation played a role in regulation of gene expression of those genes that were tested and showed a negative correlation between methylation and expression. As genes were re-expressed after treatment using just a DNA demethylating agent (5-aza), DNA methylation in GCTs may regulate gene expression by a mechanism which does not involve modifying chromatin.

Chapter 5. The role of PRDM14 in GCTs

5.1 Introduction

From the analysis described in chapter 4, PRDM14 was found to be both differentially methylated and expressed between the seminoma and YST cell lines, where it was methylated and down regulated in YST. PRDM14 is a key regulator of PGC specification. Interestingly, recent studies found that PRDM14 has a significant role in the conversion of embryonic stem cells (ES) to PGCs by many mechanisms such as activating PGC genes, suppression of somatic genes, and promotion of DNA demethylation (Magnúsdóttir and Surani, 2014, Okashita *et al*., 2014, Nakaki and Saitou, 2014).

It is also noteworthy that PRDM14 demethylates DNA directly through two mechanisms, repression of *de novo* DNA methytransferases (DNMT3a and DNMT3b), DNMT3L, and DNMT1 cofactor (UHRF1) (Grabole *et al*., 2013) and activation of the TET enzymes (Okashita *et al*., 2014).

The cure rate for patients with GCTs is about 95% at early tumour stages but this is reduced for late-stage tumour and those with non-seminoma. Seminomas are considered very sensitive to chemotherapy such as cisplatin (di Pietro *et al*., 2005, Duale *et al*., 2007) while non-seminoma cells (especially YST) are less sensitive to therapy (Houldsworth *et al*., 1998).

O'Byrne *et al.* (2011) demonstrated that epigenetic changes, including DNA methylation, are associated with resistance to standard chemotherapies such as cisplatin. Moreover, Zeller *et al.* (2012) identified some DNA methylation drivers that play critical roles in chemoresistance. However, the actual mechanism of resistance of

cancer cells to cisplatin is still unclear but it is known that cisplatin induces cell death by trigging apoptosis.

Cancer cells can evade apoptosis by utilising a variety of mechanisms; some are specific to particular tumour types. One of these mechanisms, related to this study, is methylation-induced silencing of tumour suppressor pathways that promote evasion of cell death (Anglim *et al*., 2008). Interestingly, recent evidence suggested that overexpression of PRDM14 facilitates apoptosis in HPV-positive cancers through upregulating the apoptosis regulators NOXA and PUMA (Snellenberg *et al*., 2014).

Collectively, these studies suggest that PRDM14 could play a key role in GCT progression and demethylation of DNA in addition to sensitive to chemotherapy. Therefore, in this chapter, the potential function of PRM14 in germ cell carcinogenesis and chemosensitivity has been studied by generating YST cells stably overexpressing PRDM14.

5.2. Results

5.2.1 Production of YST cells stably overexpressing PRDM14

To study the role of PRDM14 in DNA methylation and GCTs biology, a PRDM14 expression construct was generated and transfected into the YST cell line (GCT44). The full-length PRDM14 coding region was amplified from the SEM cell line (TCAM-2) by reverse transcription-PCR (RT-PCR). The PCR product was cloned into myc-pcDNA3.1 vector (Invitrogen, UK) to create the myc-pcDNA3.1-PRDM14 expression construct. The cloning site for PRDM14 was introduced using EcoR1 and Xba1 restriction enzymes (Figure 5.1). The sequence of myc-pcDNA3.1-PRDM14 was verified by sequencing (Appendix II).

To determine the transfection efficiency, GFP plasmid was transiently transfected alone and with myc-pcDNA3.1-PRDM14 expression construct into YST cells using electroporation following (section 2.2.5.6). The percentage transfection efficiency was determined visually under the fluorescent microscope 24 hour after transfection, by calculating proportion of cells that expressed GFP. The transfection efficiency was more than 80% (Figure 5.2).

Stable transfection was achieved using G418 antibiotic after 48 hour post transfection followed section 2.2.5.7. As shown in Figure 5.3, a kill curve for G418 was generated. The lowest concentration of G418, that minimal visual toxicity was apparent even after seven days of treatment, was 20 μM. While the highest concentration was 100 μM, when all cells were dead within 2-3 days of treatment. The optimal concentration of G418 was 80μM, chosen as all cells were dead after one week of treatment. Three independent polyclonal stably-PRDM14 expressing cell lines were generated for YST.

Clone Name: PRDM14 (Full-length)
Vector: pcDNA3.1/myc tag- size 5.5 Kb
Resistance: Ampicillin-Neomycine
Source: cDNA from seminoma (TCam2) cells
Gene: PRDM14: PR Domain-Containing Protein 14 (Homo sapiens)
Cloning sites: EcoR1 and Xba1
Insert length: 1737bp
Total: 1737 + 5.5 =7.237
Sequence:

GAC TTG GGA TCC AGT GTG GTG GAA TTC GCC CGC ATG GCT.....TACTAG TCT AGA

Myc Tag **BamH1** **EcoR1** **PRDM14 Sequence** **Xba1**

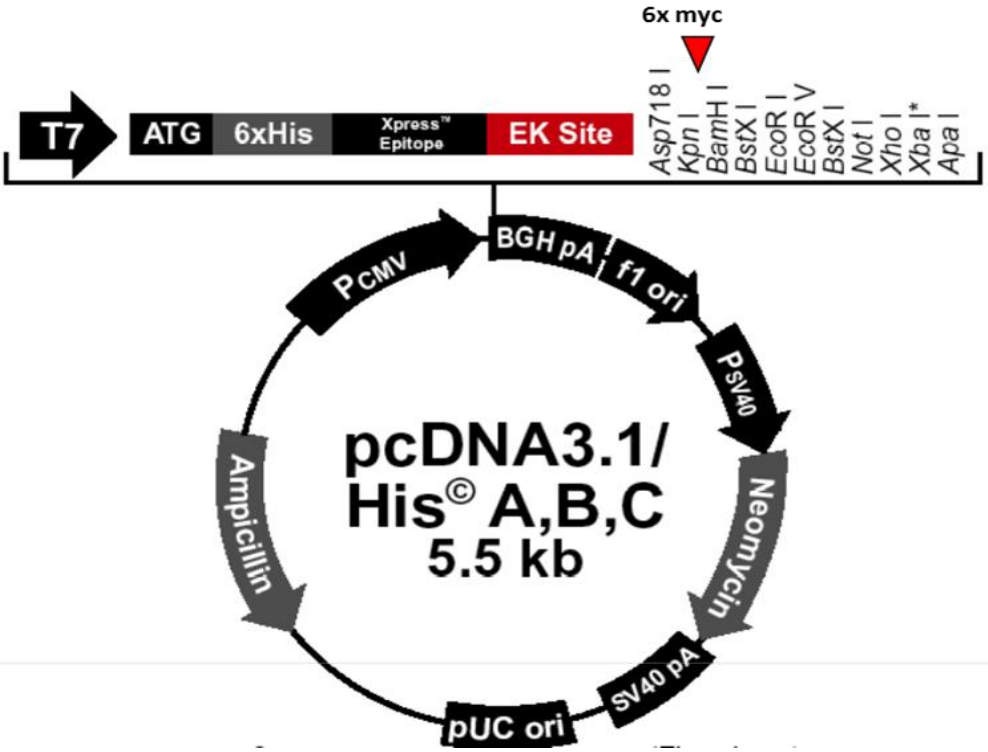


Figure 5.1: pcDNA3.1/Myc-tag expression vector map: Human PRDM14 was cloned from seminoma cell line and inserted into the pcDNA3.1 expression vector with Myc-tag at the 5' end.

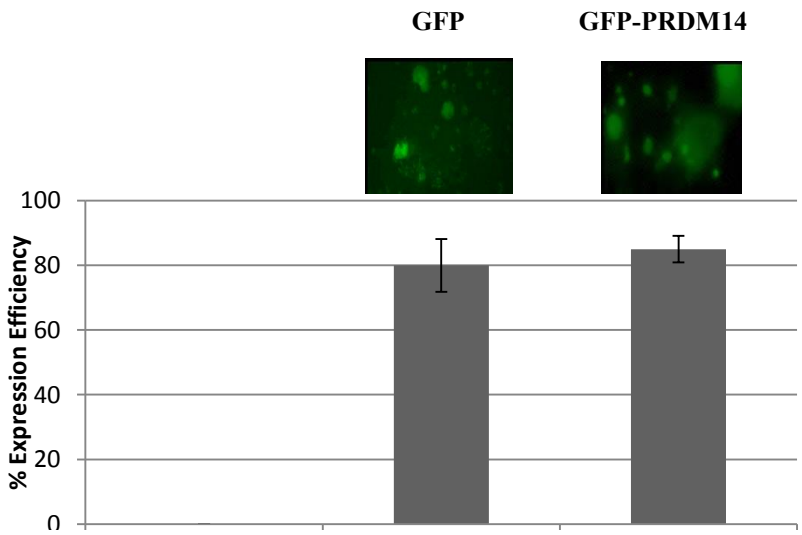


Figure 5.2: Transfection efficiency of PRDM14: Assessing the transfection efficiency by counting the cells that expressed GFP versus the total cells in the same population by fluorescent microscope. Images showed cells that expressed GFP. The graph showed high expression efficiency (≥ 80%). Error bars represent standard deviations of three microscopic fields were counted.

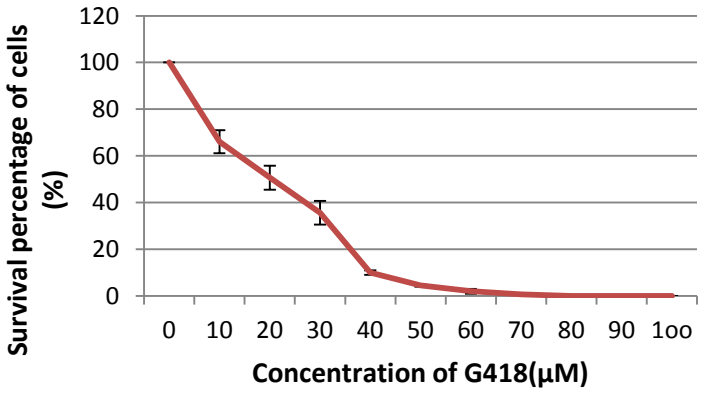


Figure 5.3: Generation of a kill curve for (G418): YST cells were exposed to increasing concentration of G418 for one week and percentage survival was calculated. Surviving cells were counted using trypan blue staining assay. The values were expressed as mean. Error bars represent standard deviations of three microscopic fields were counted.

Polyclonal stably-PRDM14 expressing cells in addition to transient transfection were selected for PRDM14 expression analysis by RT-PCR (Figure 5.4) and RT-qPCR (Figure 5.5) compared to control (untransfected YST) and SEM cell line. The data showed that there is expression for PRDM14 in transient and stable transfected YST cells comparing to the control YST cells which means that these samples could be sufficient for further analysis.

The purpose of using three independent polyclonal populations is to investigate the consistency of expression of PRDM14 in stably transfected YST cells comparing to untransfected YST. The PRDM14 expression was observed in all three populations in addition to transient transfected cell and SEM cell lines (Figure 5.4 and 5.5). However, variability was observed in signal intensity with three polyclonal stable transfected cell populations. This may be due to survival and outgrowth of the nontransgenic cells within the polyclonal cell population or due to differences in heterogeneity of these populations rather than transfection efficiency.

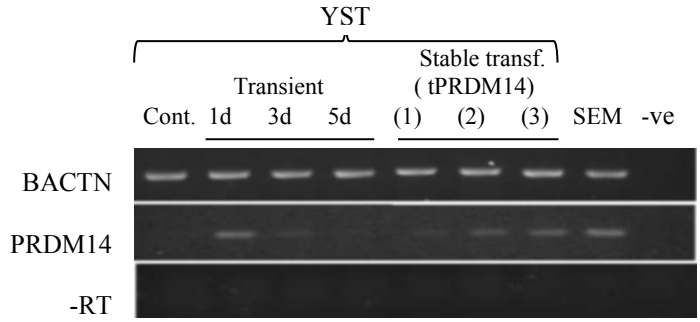


Figure 5.4: RT-PCR for PRDM14 expression after transfection: PRDM14 expression was shown in both transient and stable transfected YST cell lines. At first day (1d) of transfection, PRDM14 expression was clear but in third day (3d) the expression decreased and no expression in fifth day (5d). While in stable transfection, there was a clear expression in three independent samples compared to control (Cont.), untransfected YST. (SEM) represents seminoma cell line where PRDM14 is expressed.(-ve) negative control for PCR experiment. B-actin (BACTN) was used as loading control. (–RT) no reverse transcriptase samples.

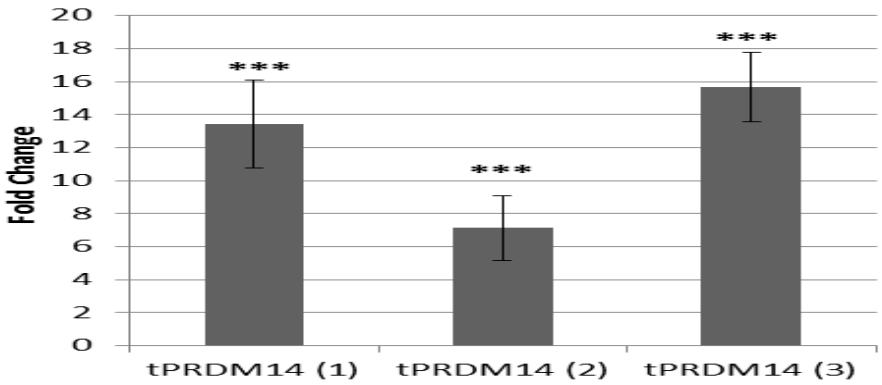


Figure 5.5: RT-qPCR for assessing YST cells stably overexpressing PRDM14: Bar graphs showing PRDM14 expression level after stable transfection comparing to untransfected YST. Experiment was performed for three independent samples. All three independent transfections revealed significantly increased in PRDM14 expression. *** denotes p-value <0.001 using paired t-test. Error bars represent standard deviation.

5.2.2. Validation of expression of PRDM14 protein using western blot

Expression of PRDM14 protein in the transfected YST cell line was evaluated by western blot analysis. After transient and stable transfection with myc-pcDNA3.1 tagged PRDM14, cells were harvested then cell lysates were used to investigate protein expression of PRDM14. The samples were run on SDS-PAGE to separate the proteins according to their molecular weight followed by western blot analysis as described in section 2.2.7.4. Primary mouse polyclonal antibody (Myc-Tag Mouse mAb) was used to bind to the Myc tagged proteins. Secondary antibody (Donkey anti-Mouse IgG, IRDye®800CW, LI-COR, USA) was used to visualise the primary antibody that bound to a target protein.

Western blot analysis showed the presence of myc-tagged PRDM14 in the transfected YST cell lines. In Figure 5.6, bands for PRDM14 protein were detected in lanes 6-9 that presented myc-tagged PRDM14 protein expression at approximate 77kDa which proved that the PRDM14 protein was expressed in the transfected YST cell. In addition, protein signals detected for stable transfection samples showed greater intensities compared to the protein signals of transient transfected samples at 7 days where protein was assumed to have degraded after that time. This observation provided evidence that polyclonal cultures stably expressed PRDM14 protein.

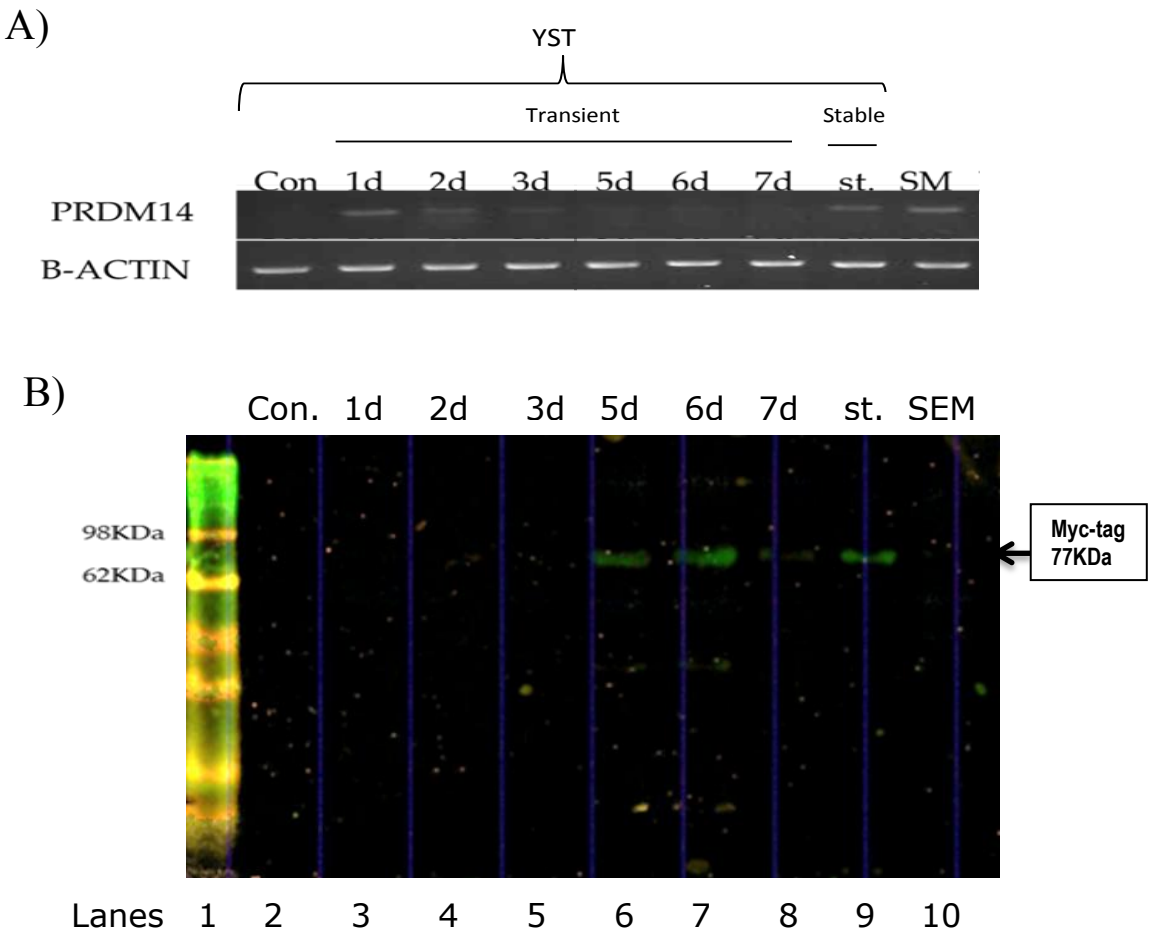


Figure 5.6: RT-PCR and western blot analysis of PRDM14 expression in transfected YST cell lines.
A) RT-PCR showed expression of PRDM14 in transient transfection for 1-7 days (1d-7d) which showed decreasing of PRDM14 expression after 3 day (3d) while stable transfection (st.) showed increased of expression. Untransfected YST cells (Con.), seminoma cells (SEM)
B) Western blot: PRDM14 protein expression in YST cells was detected after transfection using anti-Myc. Both transient and stable transfected YST cell lines were tested. Transient transfection at 5 and 6 days (lanes 6 and 7, respectively) showed expression of PRDM14 while in 7 day of transient transfection (lane 8), the expression of PRDM14 protein decreased. Stable transfection (lane 9) showed increased expression of PRDM14 protein. Ladder (lane 1), Untransfected YST cells (Con.) (lane 2), transient transfection for 1-3 days (lanes 3-5, respectively), and seminoma cells (SEM) (lane 10).

5.2.3. Overexpression of PRDM14 is associated with changes in the expression of other genes

To investigate the effect of PRDM14 on regulation of other genes, RT-qPCR analysis was performed to examine the effect of overexpressing PRDM14 on the expression of some candidate genes; these were found to be methylated and silenced in YST compared with SEM and are known to have an important role in germ cell progenitors and/or pluripotency. In this study, expression levels of these genes in YST cells stably overexpressing PRDM14 were compared with those in untransfected YST cells. All quantification data were normalized to β-actin which acts as an internal control. All three independent stable cell lines that significant overexpressed PRDM14 showed an increase in the expression of KLF2, OCT4, RASSF2, TDRD12, and KLF4 with a 2-4x fold change for KLF2 ($p < 0.01$), OCT4 ($p < 0.001$) and TDRD12 ($p < 0.05$), and a 7-17x fold change for RASSF2 ($p < 0.01$) and KLF4 ($p < 0.001$) (Figure 5.7). All of these genes in addition to PRDM14 were methylated and downregulated in YST compared with SEM cell lines, but after overexpression of PRDM14, those genes were up-regulated in all transfected YST samples. This revealed that their expression levels were positively correlated with PRDM14 expression levels. Therefore, we suggest that PRDM14 could be involved in regulation of these genes.

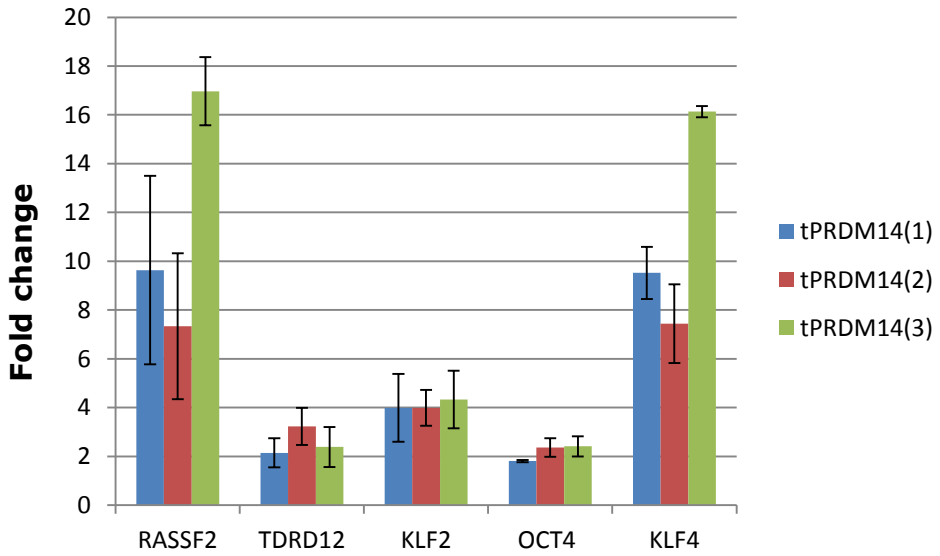


Figure 5.7: Gene expression analysis in YST cells overexpressing PRDM14. Bar graphs showing the average fold change of gene expression for selected genes after overexpression of PRDM14 in all three independent transfected YST samples compared with untransfected YST cell lines. Error bars represent standard deviation of three technical replicates.

5.2.4. Assessing the global DNA methylation level after overexpression of PRDM14 in YST

To further assess the effect of PRDM14 on GCT biology, the DNA methylation level in YST cells stably overexpressing PRDM14 was investigated. Stably transfected and untransfected YST cells were fixed and subjected to immunostaining with primary antibody (mouse Anti 5-methylcytidine) and secondary antibody (TEXAS RED® anti-mouse IgG) for methylation detection in addition to primary antibody (Anti-PRDM14 antibody) and secondary antibody (Alexa Fluor®488 Goat anti-Rabbit IgG) for PRDM14 detection (as mentioned in section 2.2.7.2). The methylation level of overexpressing PRDM14 YST cells to SEM and untransfected YST was compared. Immunohistochemical examination of 5mC can be used to assess global DNA methylation, especially when the number of samples available for assessment is small or when the methylation status of cancer cells cannot be normalized to normal tissues (Piyathilake *et al*., 2000, Wermann *et al*., 2010).

The results indicated that global methylation of overexpressed PRDM14 YST cells was decreased compared with untransfected YST (Figure 5.8), which revealed that PRDM14 could play a key role in demethylation machinery.

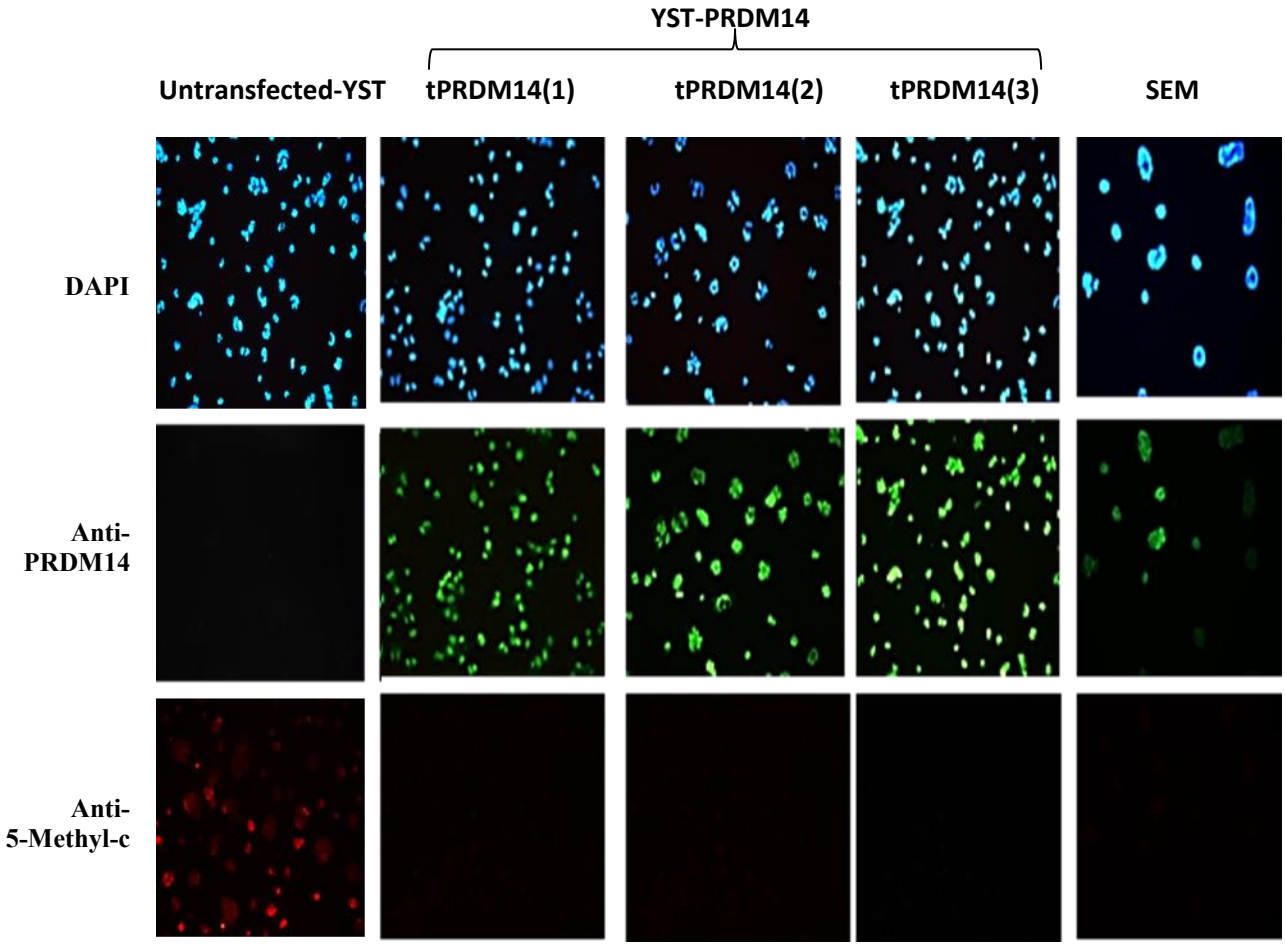


Figure 5.8: Immunostaining analysis to assess the methylation level after PRDM14 transfection: It showed that methylation levels decreased in all three independent YST-PRDM14 samples compared with untransfected YST. All samples recorded in the same camera setting.

5.2.5. Clonogenic survival assay for YST cells stably expressing PRDM14 following treatment with cisplatin

The clonogenic cell survival assay (or a colony formation assay) is an *in vitro* assay based on evaluating the ability of a single cell to form a colony after treatment with ionizing radiation or chemotherapy agents such as etoposide and cisplatin. This assay was described initially in the 1950s for studying the effectiveness of radiation on cells (Franken *et al*., 2006) but can also be used for other cytotoxic agents. This assay was used to assess the effect of PRDM14 overexpression on YST sensitivity to cisplatin.

To determine whether PRDM14 overexpression improved the response of the YST cell line towards cisplatin, untransfected and stably transfected cells were grown until confluent then subjected to clonogenic assay (section 2.2.2.4). In brief, 100 cells/well were plated in each well of a 6-well plate and when cells had adhered to the surface of the well, they were treated with cisplatin at a range of 2 to 30 μM for 2h. The medium was then replaced with drug-free medium. Cells were then incubated for 1-2 weeks. Cells were fixed with 4% paraformaldehyde then stained with 0.5% crystal violet to identify the surviving cells.

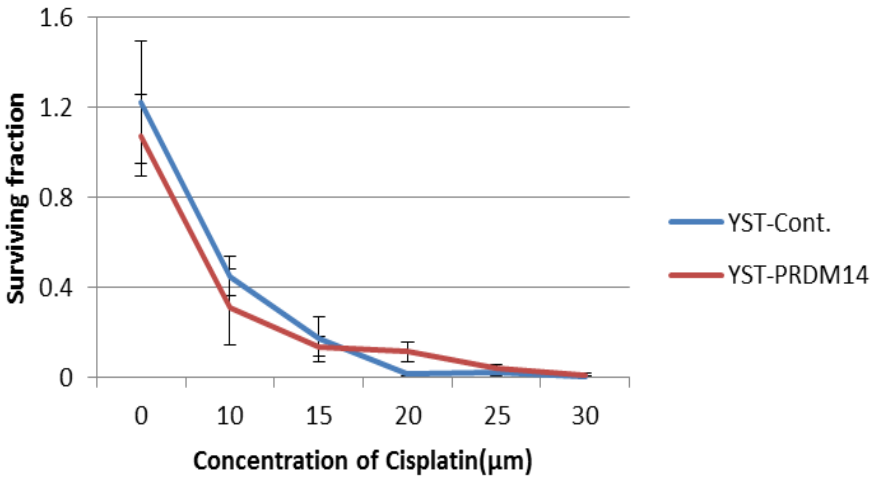
The number of colonies that survived after the treatment was counted under microscope. The following equation was applied to measure the surviving fraction which determines the significant improvement in response towards chemotherapy.

$$\text{Surviving fraction (SF)} = \frac{\text{Number of colonies formed after treatment}}{\text{Number of cells seeded x PE}}$$

$$\text{Plating efficiency (PE)} = \frac{\text{Number of colonies formed}}{\text{Number of cells seeded}} \text{ X } 100$$

Using this assay, YST-PRDM14 cell lines showed no consistent response to cisplatin treatment with different concentrations compared to untransfected YST. YST-PRDM14 cells showed more sensitivity to cisplatin with low concentrations compared to untransfected YST cell (10 and 15μM) then showed resistance to cisplatin with high concentrations (20-30μM) (Figure 5.9).

A)



B)

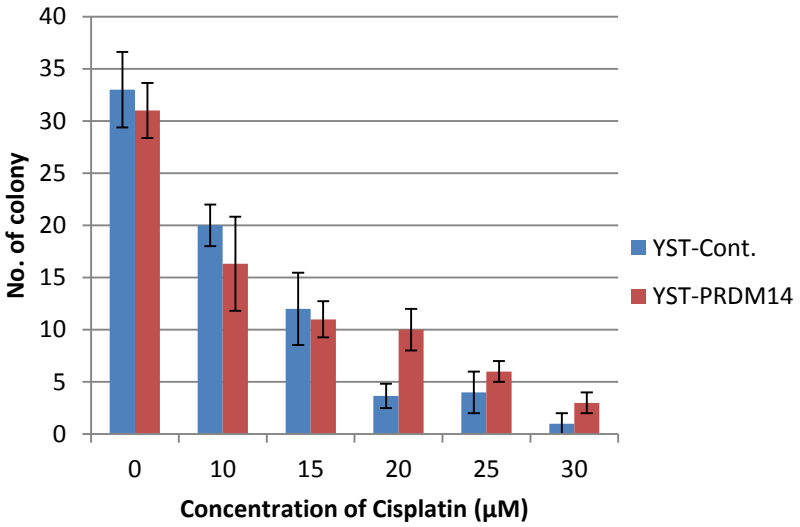


Figure 5.9: Colongenic assay: A) Surviving fraction after cisplatin treatment showed that there is no significant change in response to cisplatin after overexpression of PRDM14. B) Average number of colonies after cisplatin treatment showed that PRDM14 did not improve response to cisplatin treatment. Error bars represent standard error of three biological triplicates.

5.2.6. Sensitivity towards cisplatin after inhibition of DNA methylation in PRDM14 positive YST cells

It has been suggested that hypermethylation is related to resistance to chemotherapy in many cell lines such as ovarian cancer and NSCLC (Yu *et al*., 2011, Zeller *et al*., 2012, Zhang *et al*., 2014). Also, demethylating agents (5-aza) enhance sensitivity of several tumour cell lines towards cisplatin (Appleton *et al*., 2007).

In this study, it was confirmed that PRDM14 is highly expressed in hypomethylated seminomas and silenced in hypermethylated YST and re-expression of PRDM14 in YST cells led to demethylation of DNA. Therefore, we expect from reviewing previous literature and from our results that overexpression of PRDM14 may reduce methylation in YST and so promote sensitivity of YST cells towards cisplatin (as summarised in Figure 5.10). Untransfected YST cells and stably transfected PRDM14 cells were treated with (5μM) of 5-aza for 24 hours then the medium was changed with drug-free medium for one day after that the cells were subsequently exposed to cisplatin at a range of 10 to 30 μM for 2 h then the medium was refreshed with drug-free medium for 24h. Viable cells were counted by a haematocytometer.

Overexpression of PRDM14 showed a higher response rate to cisplatin after global demethylation in PRDM14 positive YST cells compared with their counterparts. The result revealed that PRDM14 in combination with 5-aza can significantly ($p= 0.02$) induce cisplatin-induced cytotoxicity (Figure 5.11).

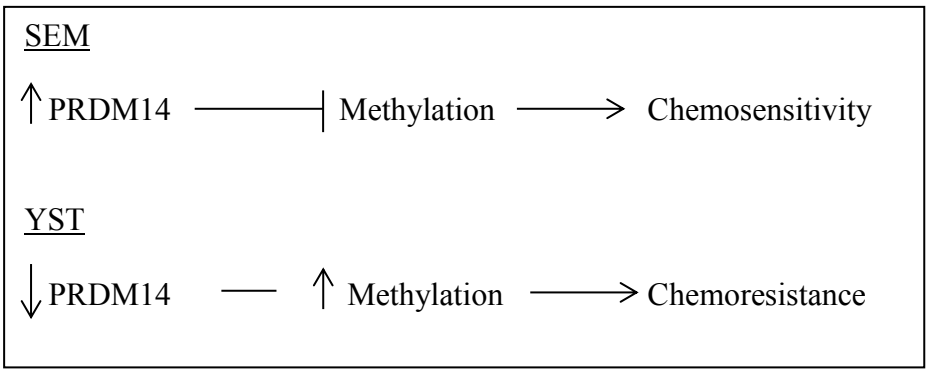


Figure 5.10: Summary of proposed action of PRDM14

Cytotoxicity-Cisplatin

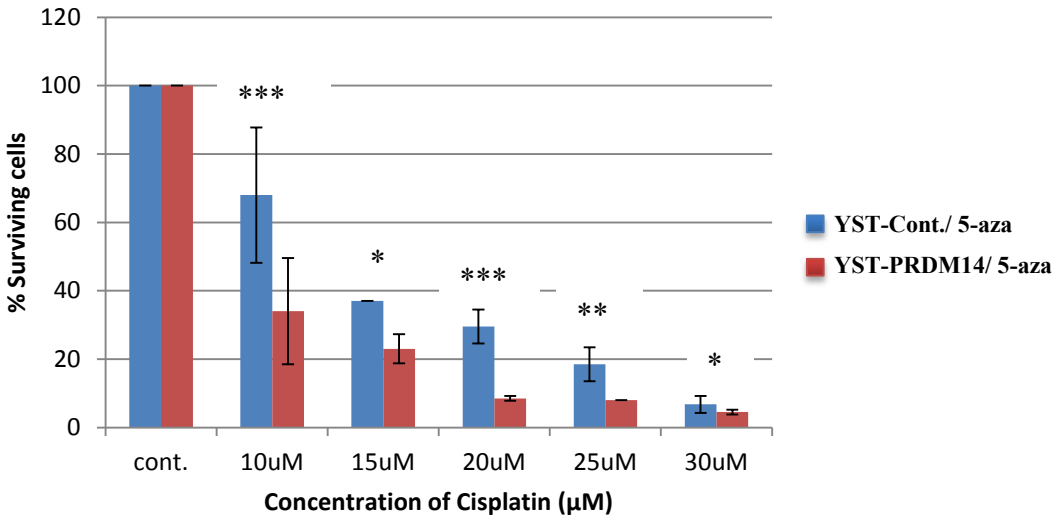


Figure 5.11: Sensitivity towards cisplatin concentrations: Response of cells to cisplatin after using 5µM 5-aza showed that there was increased in response to cisplatin after overexpression of PRDM14. Percentage of surviving cells showed that PRDM14 with 5-aza significantly improve response of cells to cisplatin (10-30µM) (*denotes p-value<0.05, **denotes p-value <0.01, and ***denotes p-value < 0.001 using a paired t-test). Error bars represent standard deviation of three biological triplicates.

5.3 Discussion

Studies in normal embryonic stem cells (ESCs) revealed that PRDM14 is expressed at much lower levels than in PGCs (Yamaji *et al*., 2013, Grabole *et al*., 2013) which suggested that PRDM14 could play role in the development or progression of GCTs. Moreover, it was pointed that knock-down of PRDM14 led to differentiation of PGCs to extraembryonic endoderm (Ma *et al*., 2011) or embryonic cell (Yamaji *et al*., 2013). This study showed that PRDM14 is differentially methylated and expressed in YST compared to SEM cell lines which supports a role of PRDM14 in the biological differences between those two types of GCTs.

Recent attention has focused on the importance of PRDM14 associated demethylation in the differentiation of PGCs (Okashita *et al*., 2014). Several studies have demonstrated that overexpression of PRDM14 contributes to maintenance of pluripotency in ESCs by repressing the methylation machinery (Grabole *et al*., 2013, Yamaji *et al*., 2013).

In this chapter, the role of PRDM14, which is methylated and silenced in YST, in the biology of GCTs was tested by assessing the effects of increasing expression of PRDM14 in three aspects; the regulation of the expression of selected genes, the global methylation levels, and the sensitivity towards cisplatin.

The results showed that overexpression of PRDM14 promotes upregulation of some pluripotency associated genes (KLF2, KLF4, OCT-4) as well as RAS associated gene (RASSF2), and TDRD12 (also known as ECAT8). This result is consistent with studies that revealed that PRDM14 regulates the pluripotency process by regulating pluripotency associated genes (Ma *et al*., 2011, Chan *et al*., 2013). A recent study

reported that PRDM14 is considered a unique regulator for pluripotency genes (Nakaki and Saitou, 2014), thus increased expression of PRDM14 may maintain seminomas in an undifferentiated state. The effect of PRDM14 expression in regulating TDRD12 and RASSF2 has not yet been examined in other studies. This work has demonstrated that RASSF2 and TDRD12 are differentially methylated and expressed in YST compared to seminoma (hypomethylated and expressed in seminoma). Therefore, PRDM14 may regulate TDRD12 and RASSF2 indirectly by its hypomethylation action. Guerrero-Setas *et al.* (2013) pointed out that RASSF2 is silenced by hypermethylation and involved in the progression of many cancers so it is hypothesized that there is an underlying pathway for progression of GCTs, which could involve upregulated RASSF2 by a demethylation action through overexpression of PRDM14. With regard to TDRD12 , Almatrafi *et al.* (2014) revealed that TDRD12 is a human cancer germline gene that has expression restricted to the germ cells of the gonads, therefore TDRD12 might also be involved in GCT biology.

The second aspect of the role of PRDM14 in the biology of GCTs was assessing the methylation level in YST cells when PRDM14 was overexpressed. There is a rapidly growing literature supporting a role of PRDM14 in epigenetic reprogramming during germ cell development via demethylation of DNA in PGCs. Some have suggested that this process occurs in combination with expression of other genes such as Blimp-1 and Prmt5 (Nagamatsu *et al*., 2011). Our result showed that overexpression of PRDM14 alone can decrease the level of methylation so our study supports other research that considered PRDM14 alone is a key trigger for the genome-wide DNA demethylation in PGCs directly by recruiting the chromatin regulator polycomb repressive complex 2 (PRC2) and PRC2 and repressing Dnmt3b and Dnmt3l resulting in

demethylating DNA (Nakaki and Saitou, 2014, Burton *et al*., 2013, Chan *et al*., 2013).

Generally, seminoma cells (where PRDM14 is highly expressed compared with YST cells) are extremely sensitive to cisplatin-based chemotherapy drugs, while YST cells are known as more aggressive and chemoresistant. Seminoma is classified as the low/or intermediate risk group of GCT types. It responds well to traditional oncology treatments based on the characteristic of its cells (Looijenga *et al*., 2011).

Some studies have identified a number of genes capable of driving chemoresistance or chemosensitivity of cancer cells such as a study by Whitehurst and colleagues (2007), which identified a number of cancer testis genes that might be involved in chemotherapeutic resistance. Others revealed that DNA demethylation increases sensitivity of neuroblastoma cancer cells to chemotherapy by controlling the expression of specific genes such as CASP8 and RASSF1A, which play a critical role in apoptosis and mitotic arrest, resulting in impaired resistance to cisplatin treatment thus enhancing cell apoptosis (Charlet *et al*., 2012). Therefore, manipulation of epigenetic modifications may provide a novel epigenetic therapy for cancer.

The effect of overexpressed PRDM14 on sensitivity to chemotherapy was investigated in this study after inhibition of DNA methylation by 5aza in PRDM14 positive YST cells. The results showed that PRDM14 supported the cytotoxic effects and apoptosis of cisplatin after using 5-aza on YST cells in culture. But when the sensitivity of overexpressed PRDM14 YST cells to cisplatin was assessed without using demethylation agent 5-aza, the result showed no effect. In conclusion, overexpression of PRDM14 could promote sensitivity of cells towards cisplatin in combination with 5-aza. The findings agree with other studies showed that using 5-aza in combination with chemotherapeutic

drugs could be useful to enhance more sensitivity to chemotherapy (Beyrouthy *et al*., 2009, Charlet *et al*., 2012, Liu *et al*., 2015).

Chapter 6. Genomic screening for genes upregulated by demethylation revealed novel targets of epigenetic silencing in GCTs

6.1 Introduction

Recently, there is an increased interest in the effect of epigenetic alterations on certain diseases where some studies demonstrated that epigenetic abnormalities are implicated in the development of many different types of cancer and could be more destructive than genetic mutations (Rothstein *et al.*, 2009, Charles *et al.*, 2012). Moreover, several studies revealed that epigenetic mechanisms such as DNA methylation and histone modifications could play a key role in the development of chemoresistance to existing drugs (Kelly *et al*., 2010, Crea *et al*., 2011, Zhang *et al*., 2014).

Therefore, epigenetic cancer therapy studies that focus on creating new anti-cancer treatments that target the epigenome and identification of candidate methylated genes that are implicated in the tumourigenesis or drug resistance hold a great promise for overcoming such aggressive diseases in the future. Additionally, there is a suggestion that inhibition of hypermethylation events that occur early in development of tumours could be a good strategy to identify genes that are silenced by methylation. Several studies may have a potential diagnostic and prognostic significance or could reveal a potential mechanistic explanation for why some cells are more resistant to chemotherapy than others.

Therefore, the identification of methylation profiles in many cancers (Esteller, 2011) and methylation marker genes of each tumour (Hartmann *et al.,* 2009) could be used to further understand and diagnose the diseases or treatment plans where each type of cancer has

a unique DNA methylation pattern and CGI hypermethylation profile (methylotype).

According to my previous results, that showed a clear difference in methylation between YST (as YST showed uniquely distribution of methylated region and genes compared with two other types of non-seminomas) and SEM cell lines, the experiment was carried out focusing on comparing the methylation status between SEM and YST cell lines. I hypothesised that such methylation differences may reflect functional differences between the two common types of GCT, seminomas and non-seminomas.

The aim of my work described in this chapter was to identify dysregulated genes which might be silenced by methylation in GCT cell lines by treating the hypermethylated YST cell line with the demethylating drug (5-aza) then performing a genome wide screen using microarray analysis to assess the expression profile of treated and untreated cells.

6.2 Results

6.2.1 Cytotoxic assay

Previous studies have demonstrated that 5-aza treatment leads to re-expression of specific genes in different cancer types using a wide range of concentrations and treatment regimens. Previous studies using 5-aza in GCT cell lines also used several regimens, so there was no agreed protocol to treat YST cell line with 5-aza.

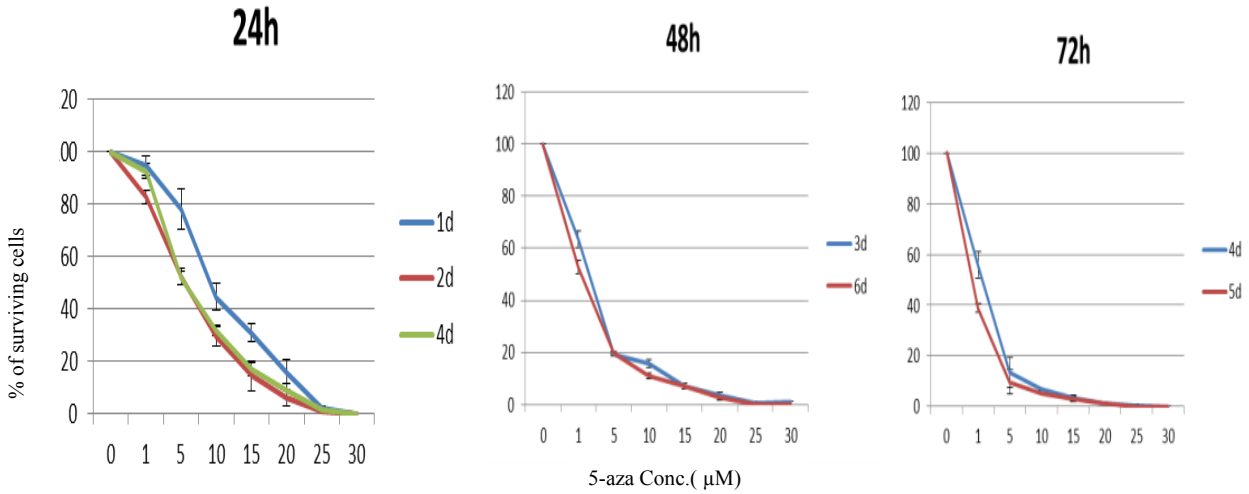
Importantly, the inhibition of methylation by 5-aza should be used at a non-toxic concentration to reactivate silenced genes. This drug inhibits DNA methylation by incorporating into the DNA during DNA synthesis (Chik and Szyf, 2011) and by depletion of DNA methyltransferases (Ghoshal *et al*., 2005). However, 5-aza is highly toxic, each type of human cancer cell having a different sensitivity to 5-aza treatment (Zhu *et al*., 2004). In addition to concentration of dose, Momparler and Goodman (1977) found in their *in vitro* studies on 5-aza that exposure time plays an important role in its activity and cytotoxic effect.

Considering these issues, a cytotoxicity assay was performed to determine the half maximal inhibitory concentration (IC50) of this drug on the highly methylated cell line (YST) (leading to 50% cell death in treated cells compared with untreated cells). Three independent experiments were performed, with each experiment performed in three triplicates following the protocol described in section 2.2.2.1. Cells were plated at $2x10^4$ cells per well in a 6-well plate. After 24/48/72 hours (Figure 6.1 -A), cells were treated with different concentrations of 5-aza, ranging from 1 to 30 μM. Media were replaced subsequently with a drug-free media then cells were harvested in a time course assay. The viable cells in each well were counted using a haemocytometer. The

results showed that the optimal concentration and time for treatment with 5-aza was 10μM for 24 hours then the cells were harvested after 2 days of treatment

IC50 tool kit program (www.ic50.tk) was used for dose-response experiments to calculate the inhibitory concentrations IC50 value and to get sigmoidal curve-fitting. For treated YST compared with untreated control, IC50 was approximately 10μM (IC50=8.98169 +/- 1.964) (Figure 6.1 B).

A)



B) Results: 24h-2d

IC50	8.98169 +/- 1.964

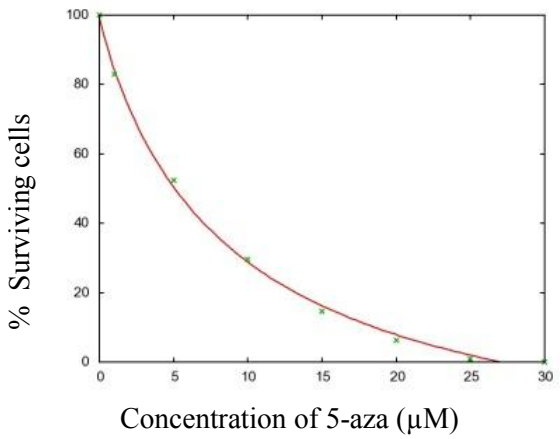


Figure 6.1: Cytotoxicity assay on the YST cell line.
A) Cells were treated with different concentrations of 5-aza, then harvested at different times. The optimal concentration and time for treatment with 5-aza were 10µM for 24 hours, with cells harvested after two days.
B) The IC_{50} for treated YST compared with untreated controls was 10µM (IC_{50}=8.98169 +/- 1.964).

6.2.2 Correlation between aberrant DNA methylation and silencing of gene expression

Many studies in recent years have revealed that treatment of cells with 5-aza-2-deoxycytidine verifies that 5-methylcytosine plays an important role in gene regulation (Luo *et al*., 2009, Yuan *et al*., 2004, Baylin, 2005, Zheng *et al*., 2012, Moore *et al*., 2013).

To test the hypothesis that aberrant DNA methylation affects GCT cells by altering gene expression, the approach used is shown in (Figure 6.2). Through this approach, YST cells were treated with 10μM 5-aza to identify which genes were silenced by methylation. After treatment, RNA was extracted and the quality of total RNA in each sample was determined by Agilent 2100 Bioanalyzer (Figure 6.3, Figure 6.4, and Table 6.1).

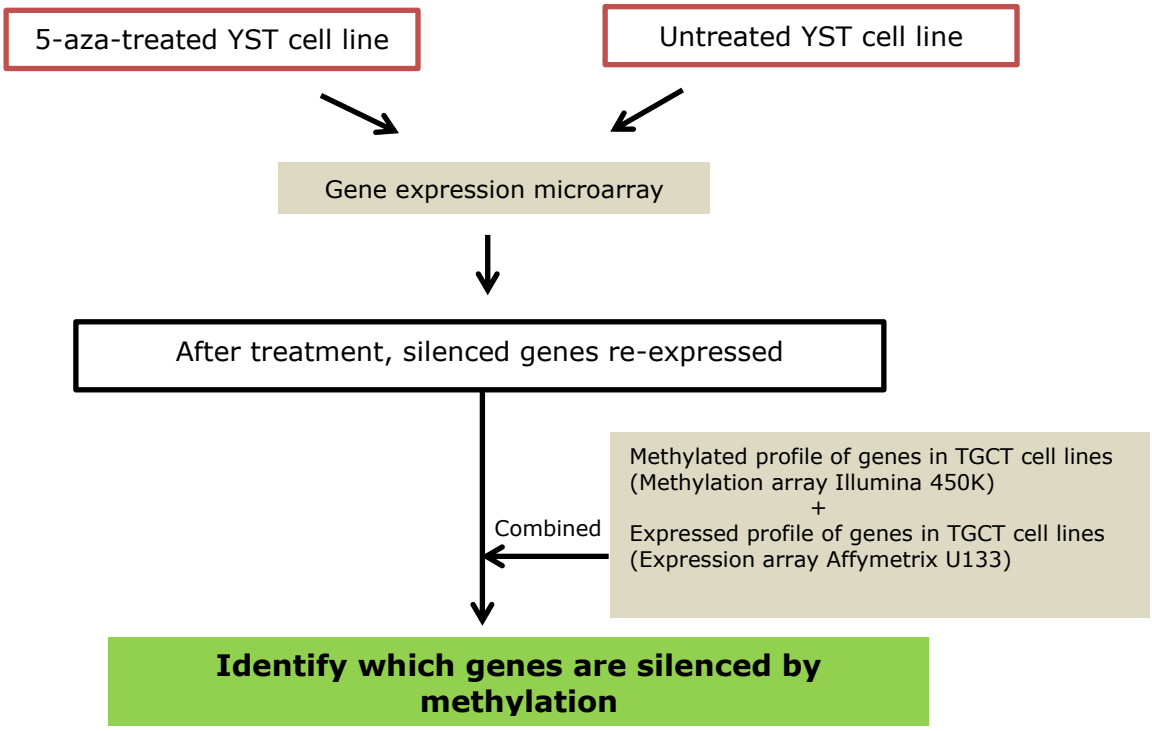


Figure 6.2 Schematic of the experimental approaches: YST cell line was treated with 5-aza then gene expression microarray carried out for treated and untreated cells to detect upregulated genes. Combining the data with that previously generated, Methylation array Illumina 450K and Affymetrix expression array U133 for TGCT cell lines, the genes that silenced by methylation were identified.

Sample 1b

rRNA ratio [28s/18s] = 2.0 RIN = 10 RNA conc = 92ng/ul

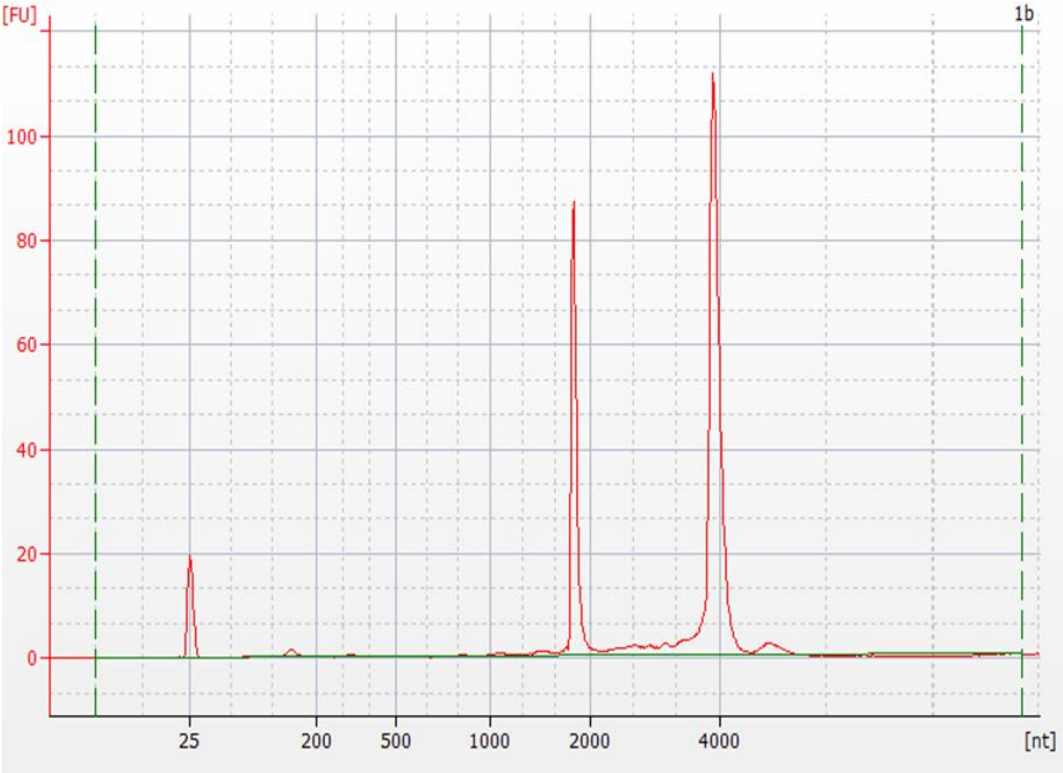


Figure 6.3: The RNA integrity number (RIN) for one YST cell line sample (1b): The RNA integrity measurement was performed by Agilent 2100 Bioanalyzer software, which showed a high integrity of RNA (10).

Table 6.1: Concentration and integrity number of RNA samples.

Sample	Concentration (ng/µl)	RNA integrity number (RIN)	rRNA ratio [28s/18s]
Control- untreated YST (1a)	71	10	2.1
Control- untreated YST (1b)	92	10	2
Treated YST-2days (2a)	168	10	2
Treated YST-2days (2b)	89	10	2
Treated YST-8days (8a)	129	10	2.1
Treated YST-8days (8b)	120	10	2.2

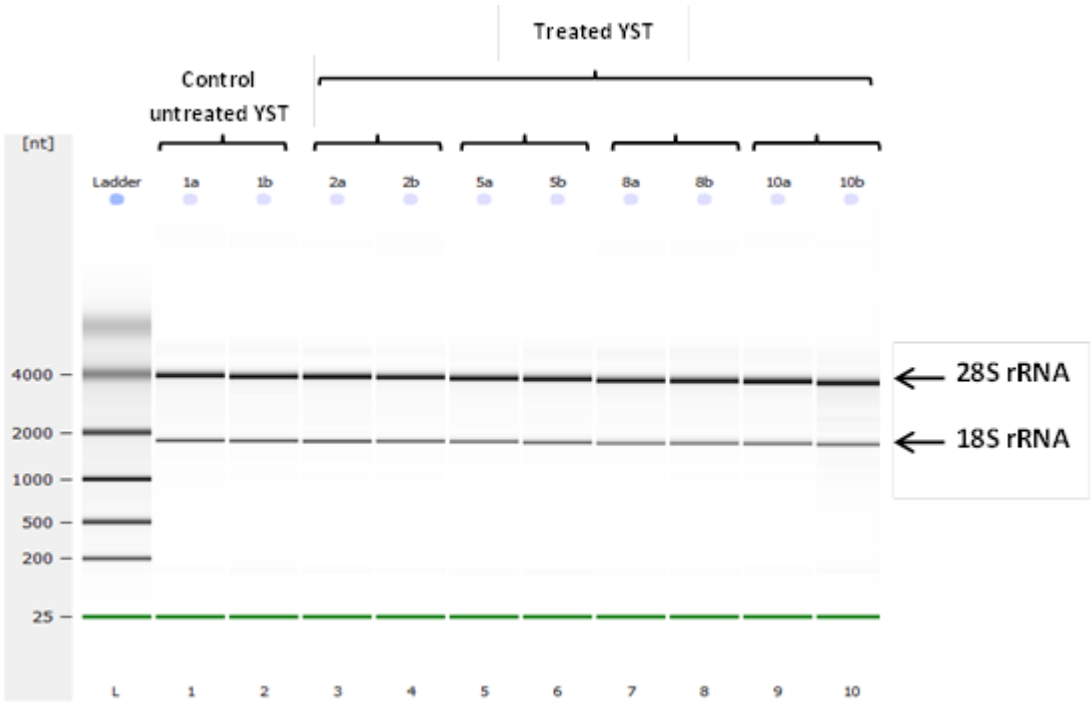


Figure 6.4: Gel-like image of RNA samples on Agilent 2100 Bioanalyzer: RNA samples were separated and detected by laser fluorescence beam. Bioanalyzer software generated a gel-like image that estimated the intactness of RNA based on the size distribution of ribosomal RNA subunits (18S and 28S ratios).

After determining the concentration and the quality of RNA samples, samples (1b, 2b, and 8b) were sent to Nottingham Arabidopsis Stock Centre at Sutton Bonington campus at the University of Nottingham for further analysis using Affymetrix HumanGeneChip U133 Plus 2.0 arrays. The raw data of microarray assay was stored as a CEL files then pre-processed using the statistical software R (Appendix II). Data was filtered such that probes which showed expression below control background were excluded. Fold changes in expression between each probe of treated cells relative to untreated cells were calculated. The data was exported as a txt file then analysed in Excel.

Analysis of the comparative gene expression before and after treatment (for two time-points) identified a set of genes that were re-expressed following demethylation where the significant differences in their expression were more than 2 fold changes at two or eight days (Figure 6.5). Twenty-three genes were upregulated after two days of treatment and twenty-two genes at eight days, while five genes were upregulated at both days (Table 6.2).

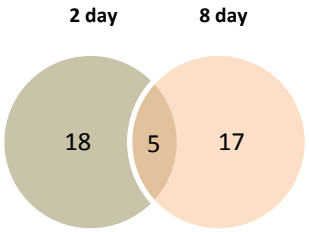


Figure 6.5: number of genes upregulated after treatment at two time points.

Table 6.2: List showed some genes re-expressed at one day of 5-aza treatment and others re-expressed at two days.

Gene	Fold change		Expression level				Methylation level (Beta)	
	2d	8d	Untreated YST	treated		SEM	YST	SEM
				2day	8day			
EPCAM	0.44	2.16	20.81	28.32	93.23	10631	0.9	0.12
SOX17	2	1.6	21.5	80.79	69.35	917.1	0.9	0.1
HIST1H4C	3.65	3.39	37.82	474	396	1605	0.9	0.12
APOB	3.45	4.56	36.46	397	858	31.43	0.51	0.66
FIGNL1	2.12	2.27	14.55	63.59	70.19	1071	0.81	0.24
CTHRC1	1.01	2.04	8.83	17.81	36.32	3138	0.94	0.06
PTPRO	2.22	1.56	597	2776	1758	334	0.6	0.14
PPFIA4	2.07	0.72	33.01	138.6	54.3	33.9	0.9	0.9

6.2.3 Identification of novel dysregulated genes which might be silenced by methylation in TGCT cell lines

To indicate which methylated genes were upregulated after treatment, I combined the results of expression microarrays that were performed for testicular GCT cell lines and treated YST cell line to compare the expression of genes before and after treatment with 5-aza. Then I used the data from published microarray experiments for normal testicular cells as a control (GEO accession: GSM380048) for comparing the gene expression at normal condition.

Five methylated genes (APOB, FIGNL1, HIST1H4C, MAGEB2, and TCEAL7) were re-expressed after treatment at two time points (2d and 8d). Three of those genes (FIGNL1, HIST1H4C, and TCEAL7) were expressed in normal testicular cells while others were silent. When comparing the expression and methylation levels of these genes in YST with seminoma, there was a significant difference in expression with fold change > 2 in these three genes. FIGNL1 and HIST1H4C were also differentially methylated in promoter CGIs between YST and seminoma.

6.2.4 Validation of microarray data

To validate microarray data, some genes were selected, according to three criteria; they showed differential expression and differential methylation between SEM and YST, their expression were very high in SEM, and re-expressed after 5-aza treatment. RT-qPCR was carried out for the same RNA samples that were sent for Affymetrix expression array analysis. The threshold cycle (C_t) values of each sample were defined by automated machine thresholds and the relative expression level of each gene was calculated according to the Pfaffl equation comparison with the housekeeping gene (β-actin).

The expression of the first two genes in table 6.2 (EPCAM and SOX17) which showed the above three criteria, were examined by PCR and qPCR (Figure 6.6 and 6.7) to validate the microarray results. As illustrated in figure 6.7 for qPCR analysis, the expression of those two genes upregulated after treatment of YST cell lines with (10μM) 5-aza comparing to untreated YST but still less than their expression in the seminoma cell line. The results of PCR and qPCR expression analysis for these candidate genes indicated that Affymetrix expression array data truly reflected the gene expression changes in response to 5-aza treatment.

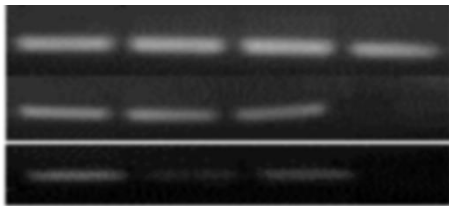


Figure 6.6: Reverse-Transcriptase PCR was carried out for the selected methylated genes in four cell lines.

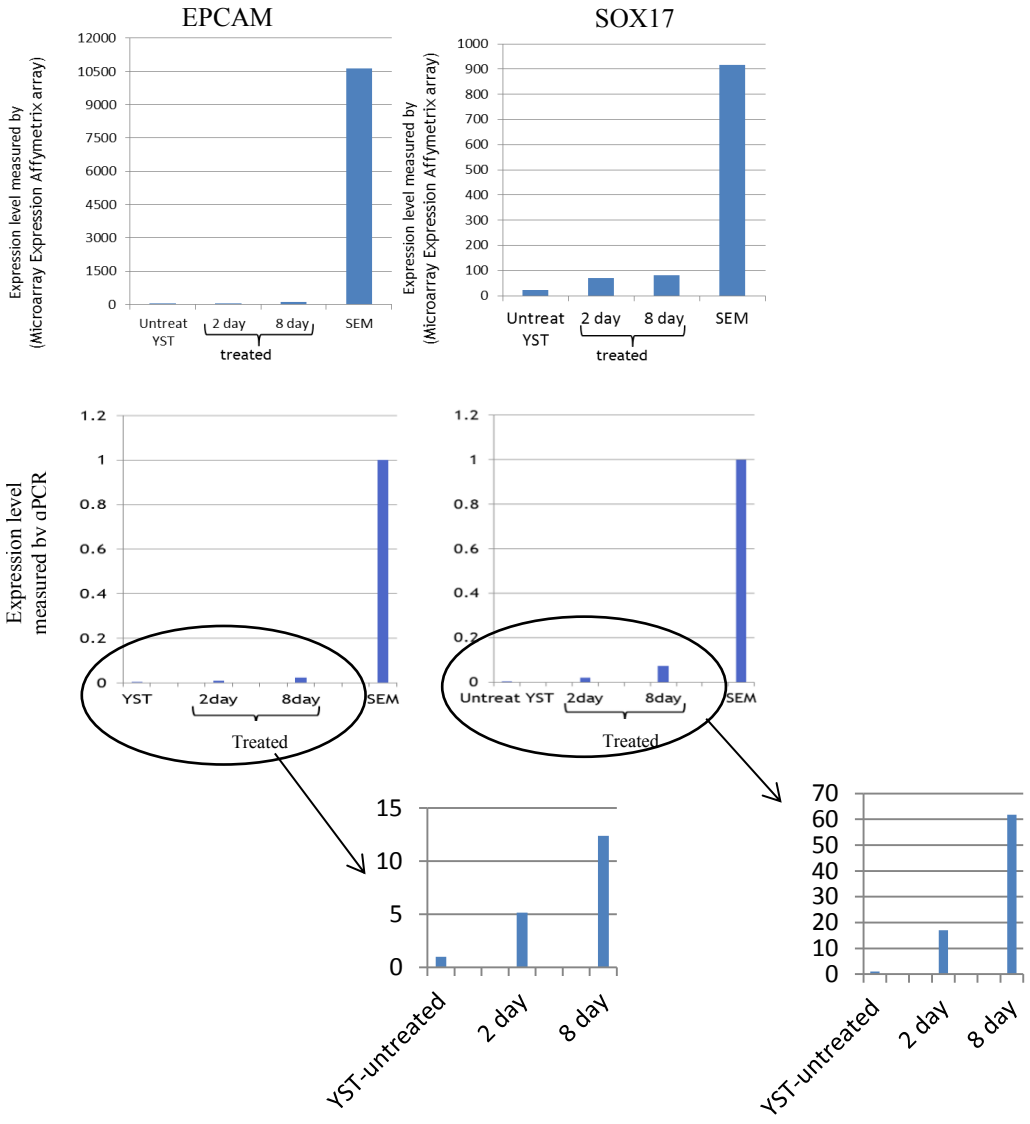


Figure 6.7: RT-qPCR results indicate that Affymetrix expression array data truly reflected the expression changes in response to 5-aza treatment.

6.3 Discussion

The role of aberrant DNA methylation in reprogramming gene expression and development of cancer has been studied using the hypomethylating agent, 5-aza (Beyrouthy *et al*., 2009) which reduces the level of DNA methylation and causes changes in treated cells such as activation of silent genes, inhibition of DNA methyltransferases, and decondensation of chromatin (Jüttermann *et al*., 1994). Jones and Taylor (1980) suggested that 5-azacytidine experiments could provide early clues to the impact of DNA methylation on gene expression. Treatment of the hypermethylated YST cell line with 5-aza and analysis of the gene expression in these cells before and after treatment using microarray assay, showed five methylated genes (APOB, FIGNL1, HIST1H4C, MAGEB2, and TCEAL7) re-expressed after treatment and three of those genes (FIGNL1, HIST1H4C, and TCEAL7) were expressed in normal testicular cells while others were silent. When comparing the expression and methylation levels of these genes in YST with seminoma, there was a significant difference in expression. FIGNL1 and HIST1H4C were also differentially methylated at CGI between YST and seminoma. This finding suggested that CGI methylation could be causative of reduced gene expression in those genes. Future research should therefore be concentrated on those methylated silenced genes that showed stable re-expression after DNA demethylation to provide a better understanding of the correlation of methylation and expression in GCTs.

Chapter 7. General discussion

7.1 Methylation profile of GCT cell lines

The first conclusion derived from the results in this thesis is that non-seminoma cells have more methylated CpGs than seminoma cells, which is consistent with previous studies (Smiraglia *et al*., 2002, Jeyapalan *et al*., 2011). The similarities and differences between the cell lines revealed that the YST cell line had the highest number (270 genes) of uniquely methylated genes relative to others, whereas only 16 genes were uniquely methylated in the SEM cell line. The difference in methylation status in two common GCT subtypes (seminomas and non-seminomas) could be due to the time point of their developmental arrest (Okamoto and Kawakami, 2007), where the histological and phenotypic characteristics of each type of GCT are dependent on the degree of differentiation (Wermann *et al*., 2010).

7.2 Progression of GCTs

The difference of methylation level between SEM and non-seminoma cell lines support the suggestion that methylation may play a critical role in the progression of this tumour and could help to determine the degree of differentiation of the tumour cells (Wermann *et al*., 2010) as both subtypes develop and differentiate from the same hypomethylated progenitor cells (PGCs) (Almstrup *et al*., 2010). PGCs undergo methylation erasure and become completely unmethylated at early phases of migration (Lind *et al*., 2007). There is evidence that mismaturation of PGCs during embryogenesis leads to the formation of CIS, then testicular seminomas subsequently arise from CIS (Netto *et al*., 2008). The finding that seminoma cells are similar to normal testicular tissue in methylation status and the difference in methylation between seminomas and non-seminomas both support the hypothesis

that seminomas could be progressed to non-seminomas by gaining methylation (Netto *et al*., 2008) (Figure 7.1).

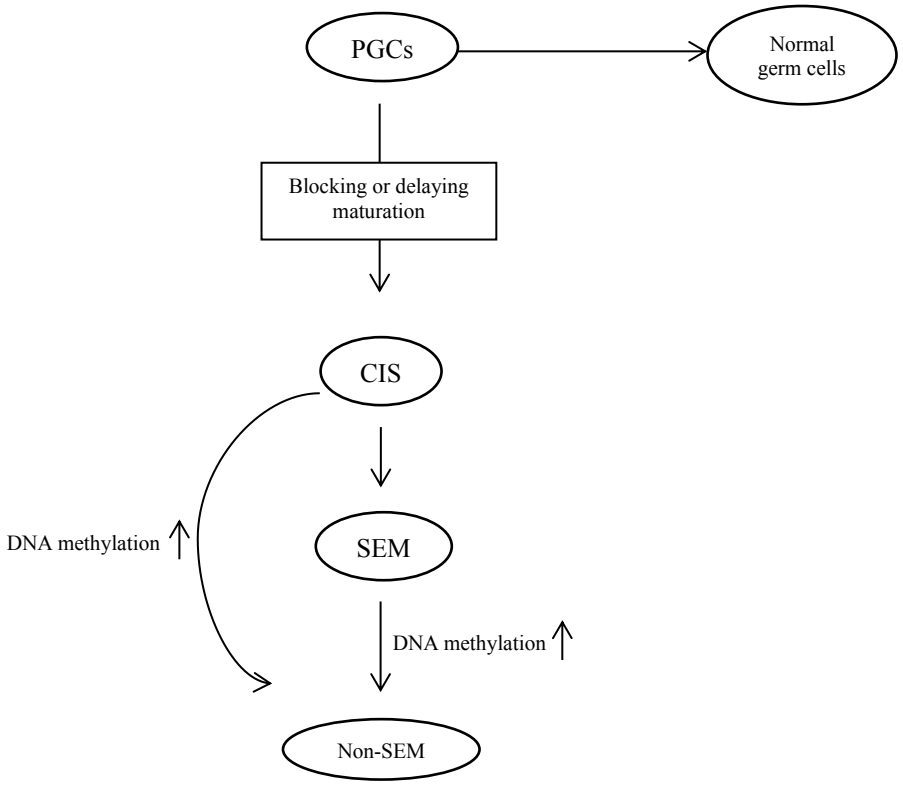


Figure 7.1: Model representing the progression of testicular GCTs. In this current model, mismaturation of PGCs leads to formation of CIS. By gaining methylation, it is suggested that non-SEM differentiate from SEM or directly from CIS.

The nonrandom pattern of methylation that is observed during development, where DNA methylation is erased during zygote formation and re-established after implantation (Jin *et al*., 2011), demonstrates that there is an underlying biological mechanism leading to regulated expression of specific genes by methylation. DNA methylation is established in early development by *de novo* methyltransferases DNMT3A and DNMT3B and maintenance by DNA methyltransferase DNMT1 (Bird, 2002) to regulate gene expression in germ cells during embryogenesis (Messerschmidt *et al*., 2014) (summarised in Figure 1.3). But the mechanism of gains of methylation in non-seminomas types of GCTs remains under investigation. There are some possible mechanisms that lead to hypermethylation status of non-seminomas.

One of them is through overexpression of DNA methyltransferases (DNMTs) which has been reported in many cancer studies (Etoh *et al*., 2004, Teodoridis *et al*., 2008). Several studies show that DNMT3L expression is very low in many normal cells such as testis, ovary and thymus (Aapola *et al*., 2004) but high in GCTs. DNMT3L is essential for normal development during embryogenesis and it is expressed in testis to form prospermatogonia during the perinatal period (Aapola *et al*., 2004). Its expression then decreases sharply after birth when prospermatogonia differentiate to spermatogonia. Okamato (2012) found that the presence of embryonal carcinoma is associated with high expression of DNMT3L in male mice after birth and suggested that DNMT3L could be used as specific marker for the diagnosis of human embryonal carcinoma (EC) (where EC in that study was considered as a malignant counterpart of human embryonic stem cells that differentiate into non-seminoma). DNMT3L is recognised as a regulatory factor for the *de novo* DNA methylation process (Chédin *et al*., 2002) and therefore it could contribute to an increase methylation in the non-seminomas.

Another possible mechanism for hypermethylation is through reduced expression of the TET proteins (Williams *et al*., 2012), leading to gain of DNA methylation (Huang and Rao, 2014). Therefore, it is suggested that down regulation of TET protein could potentially contribute in the hypermethylation status of the non-seminomas. However, mechanisms related to increased DNMT or decreased TET enzyme expression could also explain the differences in methylation level in CGI of specific genes in YST.

Several studies have reported that aberrant promoter hypermethylation leads to suppression of specific tumour suppressor genes (TSG) and thus development or progression of different human cancers (Astuti *et al*., 2001, Burbee *et al*., 2001, Dammann *et al.*, 2001, Honorio *et al*., 2003). Among those genes are MLH1 and MTS1 in testicular GCTs (Chaubert *et al*., 1997, OLASZ *et al*., 2005). Chaubert *et al.* (1997) suggested that silencing of hypermethylated MTS1 plays a role in the development of GCTs, while Olasz *et al.* (2005) found that hMLH1 hypermethylation status correlates with the loss of its protein expression leading to microsatellite instability.

Moreover, Honorio *et al.* (2003) pointed out that although mutations of RASSF1A are rare, promoter hypermethylation and transcriptional silencing of RASSF1A were detected in neuroblastoma, breast, lung, and kidney cancers (Dammann *et al*., 2001, Burbee *et al*., 2001, Astuti *et al*., 2001, Morrissey *et al*., 2001). Four independent studies revealed that RASSF1A is hypermethylated in non-seminomas versus seminoma (Koul *et al*., 2002, Lind *et al.*, 2007, Jeyapalan *et al.*, 2011) and other studies concerned with the role of this gene in GCTs found that promotor methylation of RASSF1A occurs early in the development of GCTs and it was concluded that the inactivation of RASSF1A could play a role in progression of seminoma to non-seminoma (Christoph *et al.*, 2007, Tian *et al*., 2011). Thus, more studies are

needed to give evidence about the role of these candidate genes in aberrant methylation and progression of GCTs and other cancers.

7.3 CpG Island Methylator Phenotype (CIMP)

This study revealed high levels of methylation in or around CGIs of a subset of genes in YST correlated well with gene silencing, which supports the CpG island methylator phenotype (CIMP) seen in other cancers such as colorectal cancers (Toyota *et al*., 1999) and gliomas (Noushmehr *et al*., 2010).

Notably, analysis of EC and teratoma cells showed that the methylation of most CpGs sites in these cells was higher than seminoma, not just in island regions but also across CpGs in all regions which means that the hypermethylation status of these cells, EC and teratoma did not represent CIMP.

The cause of CIMP in specific genes and the mechanism for this phenomenon, where methylation in CGIs result in gene silencing, remains unclear. However, there are some suggestions that need more investigation. One acceptable possibility was proposed by Turker (2006) who suggested a model to explain the cause and mechanism of aberrant DNA methylation at CGIs, leading to silencing of specific gene (Figure 7.2). The author proposed that genes are more susceptible to CIMP when they have a distinct short sequence close to the promoter, called a 'methylation center', which attracts *de novo* DNA methyltransferases based on specific features of this center such as the presence of retrotransposons. Aberrant DNA methylation starts in the methylation center then spreads to reach the TSS, leading to inactivating of the gene if there is no barrier system protecting against such methylation spreading.

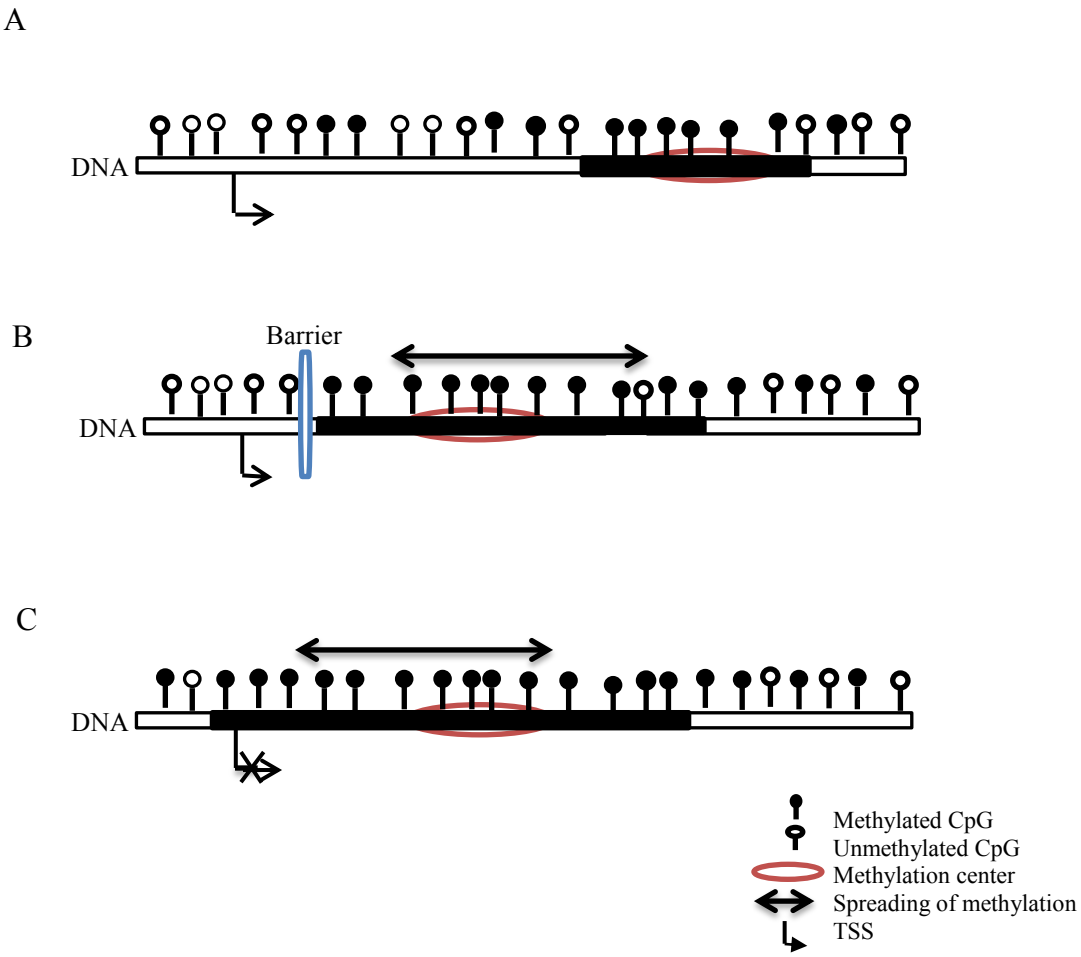


Figure 7.2: A model of proposed CIMP mechanism in cancer: A) A gene is not susceptible to CIMP when the methylation center is away from the TSS. B) A gene is susceptible to CIMP when the methylation center is close to the promoter but there is a specific barrier (could be transcription factor or co-activator) that prevents spreading methylation to TSS. C) CIMP and transcriptional suppression occur when spreading of methylation blocks the promoter.

7.4 CGI methylation is negatively correlated with gene expression

Methylation of CGIs in or near gene promoters was more strongly correlated with reduced gene expression than other regions. However, the results showed that increasing methylation of CGIs in the gene body was also more strongly associated with gene silencing than activation. There was a strong correlation between high methylation of CGIs in the gene body with suppression of gene expression, in regardless of whether this association was seen also in the CGI promoter region. However, this finding is contrary to the results of some other genome-wide sequencing studies, which revealed that DNA methylation of gene bodies correlates with gene activation rather than silencing (Lister and Ecker, 2009, Ball *et al*., 2009, Yang *et al*., 2014). The expression of genes in those studies could be related to the presence of ncRNAs (non-coding RNA) that regulate gene expression (Mercer *et al*., 2009, Deaton and Bird, 2011) rather than CGI in the body region or may be the CGIs in the gene body work functionally as alternative promoters for some genes (Carninci *et al*., 2006, Maunakea *et al*., 2010)

Recent studies showed that differential methylation status of regions other than CGI to be more closely related to regulation of gene expression. Irizarry *et al.* (2009) found that methylation status of CpG shores was associated strongly with gene expression in colon cancer. While Rao *et al.* (2013) showed hypermethylation of CGI shores to be negatively correlated with Cav1 expression in breast cancer. More recently, a study in medulloblastomas using next-generation-sequencing data analysis and whole-genome bisulphite sequencing showed that CGI shore methylation is most closely related to gene expression (Hovestadt *et al*., 2014). However, it seems that methylation of specific regions in the genome, such as island and shores, is significantly associated with control of gene expression but more research is needed to detect if

these regions are specific for the same genes in each cancer or varies between different cancers.

7.5 Identification of regulatory genes that most likely contribute to the phenotypic differences observed between seminomas and non-seminomas

Combining both methylation and expression data identified genes whose expression could be altered by methylation and appeared likely to be causative of phenotypic differences between seminomas and non-seminomas. In depth study of protein pathways and the interaction of those genes, that were methylated and silenced in YST relative to SEM could reveal potential mechanisms that cause differences in phenotypes and sensitivity to chemotherapy between these two types of tumours.

This study identified a list of genes that were both differentially methylated and expressed between GCT types by analysis of Illumina 450k array and Affymetrix expression array data. Twenty-one out of 108 genes silenced in the YST cell line were differentially methylated and expressed in non-seminoams was relative to seminomas in cell lines and in primary tumour samples. Six of them (KLF4, TDRD12, DDX43, MNS1, RBMXL2 and PRDM14) are associated with PGCs and/or pluripotency.

KLF4 is a transcription factor that plays a role in the regulation of differentiation and proliferation. KLF4 is highly expressed in PGCs (Behr *et al*., 2007), seminomas, and CIS (Godmann *et al*., 2009). In our expression data, KLF4 expression was high in seminoma and silenced in YST, which suggests that repression of KLF4 by methylation in YST may represent the differentiation status of this tumour from the pluripotent cells in CIS towards differentiated YST cells (Li *et al*., 2005, Maruyama *et al*., 2005)

The other identified genes (TDRD12, DDX43, MNS1, RBMXL2) have particular roles in reproductive processes and male gamete production and are therefore known as testes specific genes (Mathieu *et al*., 2010, Pandey *et al*., 2013, Greenbaum *et al*., 2011, Zhou *et al*., 2012). TDRD12 (tudor domain-containing protein) plays a role in the biogenesis of piRNAs which are also testis specific (Pandey *et al*., 2013). Expression of DDX43 and MNS1 is restricted to the normal testes tissue. DDX43 (also called HAGE), encodes a 'cancer testis antigen', and is an RNA dependent helicase which is also expressed in many cancers (Mathieu *et al*., 2010). MNS1 (Meiosis-specific nuclear structural1) is involved in spermiogenesis (Zhou *et al*., 2012). RBMXL2 (also called hnRNP G-T) is a heterogeneous nuclear ribonucleoprotein and known as a germ-cell specific splicing regulator (Greenbaum *et al*., 2011). However, the silencing of these genes by methylation in non-seminomas could contribute to phenotypic differences between seminoma and YST or may play a role in the differentiation. In order to validate that those genes are causative of a phenotypic difference between seminoma and YST, overexpression/loss of function studies should be applied.

PRDM14 is an important gene that needs more analysis because it has been demonstrated that PRDM14 plays an important role in germ cell development and specification (Magnúsdóttir and Surani, 2014), maintenance of germ cell pluripotency (Grabole *et al*., 2013) and recently it was investigated as a demethylation factor in germ cells (Okashita *et al*., 2014).

7.6 Role of PRDM14 in demethylation

PRDM14 is a zinc finger transcription factor that regulates pluripotency factors to maintain the pluripotent state in PGCs by suppressing the differentiation genes (Grabole *et al*., 2013). Ma *et al.* (2011) found that knockdown of PRDM14 in mouse ESCs led to differentiation to extraembryonic endoderm fates which are similar to that observed in YST tumour cells (Oosterhuis and Looijenga, 2005). In our data, PRDM14 was expressed in seminoma and silenced in YST, which suggests that silencing of PRDM14 by methylation could result in progression of seminoma to non-seminomas.

PRDM14 function is also related to DNA demethylation (Grabole *et al*., 2013) and may inhibit the transcription of DNA methyltransferases, in particular DNMT3B and DNMT1. PRDM14 could trigger passive DNA demethylation by inhibition of the maintenance activity of DNA machinery during *de novo* DNA synthesis in DNA replication (Magnúsdóttir *et al*., 2013).

Moreover, PRDM14 may cause active DNA demethylation through the TET-BER cycle, but the precise mechanism associated with TET-BER cycle remains unclear (Okashita *et al*., 2014). TET proteins (TET1 and 2) enzymatically oxidise 5'methyl-cytosine (5mCpG) to 5 hydroxymethyl-cytosine (5hmCpG) then 5hmC is further oxidized by TET proteins to produce formyl-cytosine and 5 carboxyl-cytosineare, where the later forms are subsequently repaired by thymine DNA glycosylase (TDG) and the base excision repair (BER) pathway, TDG/BER repair pathway, to form 5CpG (Okashita *et al*., 2014) (Figure 7.3). It has been suggested that methylated cytosine is converted to cytosine in an active demethylated pathway, therefore PRDM14 could play a critical role in this pathway. In addition, differential expression of PRDM14 in seminoma relative to non-seminoma could be causative of differential

methylation seen in seminoma compared with non-seminoma, through accelerating TET protein function to active demethylation.

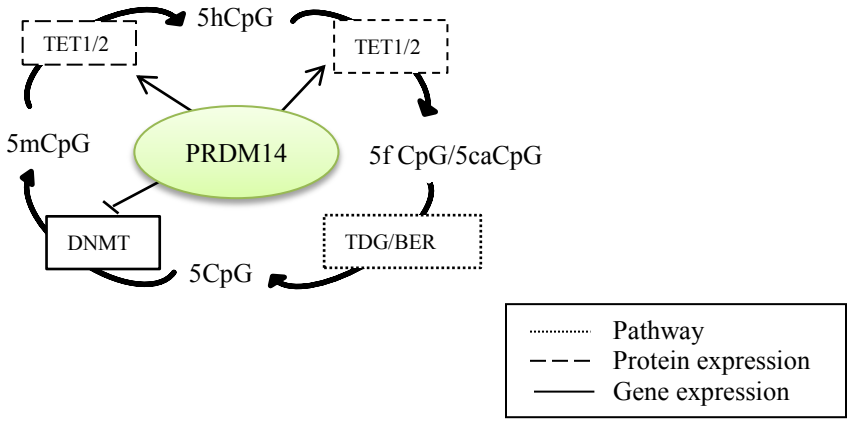


Figure 7.3: A scheme for active and passive DNA demethylation by PRDM14: Active demethylation action of PRDM14 promotes the TET/BER pathway to convert 5mCpG to unmodified 5CpG, while passive demethylation based on overexpression of PRDM14 may suppress transcriptional expression of DNA methyltransferases.

7.7 Cell lines as a useful model for methylation studies

Primary tumour methylation data was analysed and compared with the methylation data of the corresponding cell lines. The results showed differential methylation between seminoma and non-seminoma in both cell lines and primary tumour, supporting the hypothesis that the methylation level in the cell lines reflects that seen in primary tumours (Barretina *et al*., 2012). A significant correlation between CGI methylation and gene silencing is clear in this study. To determine whether DNA methylation is the cause of gene silencing, The cell lines was used as a model to verify that removing methylation, by using the DNA-demethylating agent 5-aza-2-deoxycytidine, led to re-expression of the selected genes. Cell lines provide an advantage for testing drugs or compounds because repeating the experiment many times on genetically identical cells. Selected genes were found to re-expressed after treatment with demethylating agent (5-aza), indicating that methylation could be a key factor in silencing of those genes, which were differentially expressed in non-seminomas versus seminomas in cell lines and were also differentially expressed in primary tumour study (Palmer *et al*., 2008). Comparison of expression data of cell lines with expression data from a paediatric GCT samples and primary tumour samples, the findings support the hypothesis that expression status in the cell lines between seminoma and non-seminoma is similar to that seen in primary tumours. This result is consistent with the finding of Ueki *et al.* (2002), who found that the methylation status of most tumour suppressor genes in pancreatic cell lines is similar to that in corresponding primary tumours. Despite the wide spread use of cell lines in cancer research for investigation of genetics and epigenetics studies where the results are usually extrapolated to the primary tumours, Smiraglia *et al.* (2001) reported that the hypermethylation seen in some cancer cell lines was not reflected in primary tumours.

However, an alternative interpretation of the observations in their study is that they did not compare cell lines to the actual tumour of origin from which those cell lines were derived. For example, they compared the NCCIT cell line (non-seminoma cell line) derived from human embryo placenta with primary testicular GCTs. Therefore, they found differences between cell lines and primary tissue in methylation analysis. Another explanation for their observations is that the cell lines represent a minority population of tumour cells from which they were derived. Moreover, the result of comparing tumour cell lines with normal control tissues showed that levels of methylation in seminoma closely resemble that of normal tissues, which supported the studies demonstrating clustering of seminomas with normal testicular tissue based on methylation levels (Jeyapalan *et al*., 2011). Despite the comments of Smiraglia *et al.* (2001), the use of tumour cancer cell lines as a tool, in genetic and epigenetics studies, is widely accepted (Ueki *et al*., 2002, Kao *et al*., 2009, Gazdar *et al*., 2010).

7.8 Potential therapeutic applications targeting DNA methylation

Many of the results in this study highlighted the potential role of DNA methylation in many biological processes and provide further insights into the possibility of targeting DNA methylation in the future for cancer therapy in several ways.

First, distinguishing between two subtypes of the same class of tumour, seminoma and non-seminomas, depending on the global methylation level and identifying uniquely methylated genes for each subtype might provide useful therapeutic markers for assisting other diagnosis of different types of GCTs.

Second, the progression state of GCTs related to methylation level could be used for cancer detection or even to identify those at risk of developing cancer later in life and detecting treatment purposes.

Interestingly, in this study only five pluripotency-associated genes were identified as a functional group that were silenced by methylation. They were differentially methylated and differentially expressed in non-seminomas relative to seminomas. This suggests that these could be considered as biomarkers in GCT development. These genes need further validation in a larger number of patient samples to assess their reliability before being considered as potential biomarkers in GCTs.

Furthermore, using DNA methylation and methylated genes as a potential biomarkers is a useful tool because a small amount of DNA is enough for technical analysis and DNA can be obtained from any biological tissue samples or bodily fluid such as blood and urine, where tumour DNA can be released. In addition, DNA is highly stable and can be stored for a long time, to analyse the samples before and after treatment to investigate the response of a tumour to treatment or progression of a tumour. For example, it was confirmed that MGMT methylation is a useful predictive marker for expected the response of glioblastoma patients to treatment with the alkylating agent temozolomide, where several clinical trials showed an increase in survival rate for those patients who have a methylated CGI promoter of MGMT when treated with temozolomide (Schaefer *et al*., 2010).

There are still a long way to elucidate the full understanding the role of methylation in cancer. This study revealed some insights in the correlation between methylation and expression in many genes in GCT cell lines. In particular, five of these genes that are closely associated with pluripotency and implicated in chemosensitivity might be promising targets for potential therapies in germ cell tumours which need further

analysis. In addition, these genes could help in understanding the reasons for the difference between seminoma and non-seminomas in many aspects such as aggressive behavior of non-seminoma or the sensitivity feature of seminomas to chemotherapy or the methylation difference between these types of GCTs. Moreover, PRDM14 might be used as the demethylation therapy to alter YST cells to have a more seminoma feature and thus become more sensitive to chemotherapy.

Summary:

The present study was designed to determine the correlation between DNA methylation and gene expression in two common classes of GCTs, seminomas and non-seminomas, using lab experiments and both genome expression and methylation data analysis for cell lines and primary tumours.

The most striking finding to emerge from this study was that non-seminoma cell lines revealed a very different methylator phenotype compared with seminomas. New potentially biologically important genes associated with the germ cell state and/or pluripotency process were identified; PRDM14, KLF4, TDRD12, DDX43, MNS1, and RBMXL2. The silenced and methylated genes could play a role in progression of GCTs and might provide new potential therapeutic targets for the treatment of GCTs. PRDM14 was suggested as a demethylating factor in this tumour and may cause an increase in expression of some pluripotency and/or testis genes as well as increase chemosensitivity to cisplatin in seminomas.

For future work, in order to understand more about the function of PRDM14 in GCTs, genome expression array analysis following overexpression PRDM14 in YST cells in addition to loss of function in SEM cells is necessary. In addition, it is useful to study the methylation

status of different regions of the identified genes in this study that upregulated in overexpressing PRDM14 YST cells and compared them with those in untransfected YST cells to assess the role of PRDM14 in the regional methylation of these gene.

References

AAPOLA, U., MAENPAA, K., KAIPIA, A. & PETERSON, P. 2004. Epigenetic modifications affect Dnmt3L expression. *Biochemical Society Journal,* 380**,** 705-713.

ABDELHALEEM, M. 2004. Do human RNA helicases have a role in cancer? *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer,* 1704**,** 37-46.

ALMATRAFI, A., FEICHTINGER, J., VERNON, E. G., ESCOBAR, N. G., WAKEMAN, J. A., LARCOMBE, L. D. & MCFARLANE, R. J. 2014. Identification of a class of human cancer germline genes with transcriptional silencing refractory to the hypomethylating drug 5-aza-2′-deoxycytidine. *Oncoscience,* 1**,** 745.

ALMSTRUP, K., NIELSEN, J. E., MLYNARSKA, O., JANSEN, M. T., JØRGENSEN, A., SKAKKEBÆK, N. E. & RAJPERT-DE MEYTS, E. 2010. Carcinoma in situ testis displays permissive chromatin modifications similar to immature foetal germ cells. *British Journal of Cancer,* 103**,** 1269-1276.

AMATRUDA, J. F., ROSS, J. A., CHRISTENSEN, B., FUSTINO, N. J., CHEN, K. S., HOOTEN, A. J., NELSON, H., KURIGER, J. K., RAKHEJA, D. & FRAZIER, A. L. 2013. DNA methylation analysis reveals distinct methylation signatures in pediatric germ cell tumors. *BMC Cancer,* 13**,** 313.

ANACLETO, C., LEOPOLDINO, A. M., ROSSI, B., SOARES, F. A., LOPES, A., ROCHA, J. C. C., CABALLERO, O., CAMARGO, A. A., SIMPSON, A. J. G. & PENA, S. D. J. 2005. Colorectal Cancer "Methylator Phenotype": Fact or Artifact? *Neoplasia,* 7**,** 331-335.

ANGLIM, P. P., GALLER, J. S., KOSS, M. N., HAGEN, J. A., TURLA, S., CAMPAN, M., WEISENBERGER, D. J., LAIRD, P. W., SIEGMUND, K. D. & LAIRD-OFFRINGA, I. A. 2008. Identification of a panel of sensitive and specific DNA methylation markers for squamous cell lung cancer. *Molecular Cancer,* 7**,** 62.

ANTEQUERA, F. & BIRD, A. 1993. Number of CpG islands and genes in human and mouse. *Proceedings of the National Academy of Sciences of the United States of America,* 90**,** 11995-11999.

APPLETON, K., MACKAY, H. J., JUDSON, I., PLUMB, J. A., MCCORMICK, C., STRATHDEE, G., LEE, C., BARRETT, S., READE, S. & JADAYEL, D. 2007. Phase I and pharmacodynamic trial of the DNA methyltransferase inhibitor decitabine and carboplatin in solid tumors. *Journal of Clinical Oncology,* 25**,** 4603-4609.

ARAN, D., SABATO, S. & HELLMAN, A. 2013. DNA methylation of distal regulatory sites characterizes dysregulation of cancer genes. *Genome Biol,* 14**,** R21.

ARORA, R. S., ALSTON, R. D., EDEN, T. O. B., GERACI, M. & BIRCH, J. M. 2012. Comparative incidence patterns and trends of gonadal and extragonadal germ cell tumors in England, 1979 to 2003. *Cancer,* 118**,** 4290-4297.

ASTUTI, D., AGATHANGGELOU, A., HONORIO, S., DALLOL, A., MARTINSSON, T., KOGNER, P., CUMMINS, C., NEUMANN, H., VOUTILAINEN, R. & DAHIA, P. 2001. RASSF1A promoter region CpG island hypermethylation in phaeochromocytomas and neuroblastoma tumours. *Oncogene,* 20**,** 7573-7577.

BAEUERLE, P. A. & GIRES, O. 2007. EpCAM (CD326) finding its role in cancer. *British Journal of Cancer,* 96**,** 417-423.

BALL, M. P., LI, J. B., GAO, Y., LEE, J.-H., LEPROUST, E. M., PARK, I.-H., XIE, B., DALEY, G. Q. & CHURCH, G. M. 2009. Targeted and genome-scale strategies

reveal gene-body methylation signatures in human cells. *Nature Biotechnology,* 27**,** 361-368.

BAMBERG, M., KORTMANN, R.-D., CALAMINUS, G., BECKER, G., MEISNER, C., HARMS, D. & GÖBEL, U. 1999. Radiation therapy for intracranial germinoma: results of the German cooperative prospective trials MAKEI 83/86/89. *Journal of Clinical Oncology,* 17**,** 2585-2585.

BARIOL, C., SUTER, C., CHEONG, K., KU, S.-L., MEAGHER, A., HAWKINS, N. & WARD, R. 2003. The relationship between hypomethylation and CpG island methylation in colorectal neoplasia. *The American Journal of Pathology,* 162**,** 1361-1371.

BARTOLOMEI, M. S., ZEMEL, S. & TILGHMAN, S. M. 1991. Parental imprinting of the mouse H19 gene. *Nature,* 351**,** 153-155.

BAYLIN, S. B. 2005. DNA methylation and gene silencing in cancer. *Nature Clinical Practice Oncology,* 2**,** S4-S11.

BAYLLN, S. B., HERMAN, J. G., GRAFF, J. R., VERTINO, P. M. & ISSA, J.-P. 1997. Alterations in DNA Methylation: A Fundamental Aspect of Neoplasia. *In:* GEORGE, F. V. W. & GEORGE, K. (eds.) *Advances in Cancer Research.* Academic Press.

BEHR, R., DELLER, C., GODMANN, M., MÜLLER, T., BERGMANN, M., IVELL, R. & STEGER, K. 2007. Krüppel-like factor 4 expression in normal and pathological human testes. *Molecular Human Reproduction,* 13**,** 815-820.

BERT, SAUL A., ROBINSON, MARK D., STRBENAC, D., STATHAM, AARON L., SONG, JENNY Z., HULF, T., SUTHERLAND, ROBERT L., COOLEN, MARCEL W., STIRZAKER, C. & CLARK, SUSAN J. 2013. Regional Activation of the Cancer Genome by Long-Range Epigenetic Remodeling. *Cancer Cell,* 23**,** 9-22.

BESTOR, T. H. 1992. Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain. *The EMBO Journal,* 11**,** 2611-2617.

BEYROUTHY, M. J., GARNER, K. M., HEVER, M. P., FREEMANTLE, S. J., EASTMAN, A., DMITROVSKY, E. & SPINELLA, M. J. 2009. High DNA methyltransferase 3B expression mediates 5-aza-deoxycytidine hypersensitivity in testicular germ cell tumors. *Cancer Research,* 69**,** 9360-9366.

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. & SHEN, R. 2011. High density DNA methylation array with single CpG site resolution. *Genomics,* 98**,** 288-295.

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

BIRD, A. 2011. The dinucleotide CG as a genomic signalling module. *Journal of Molecular Biology,* 409**,** 47-53.

BOCK, C., PAULSEN, M., TIERLING, S., MIKESKA, T., LENGAUER, T. & WALTER, J. 2006. CpG island methylation in human lymphocytes is highly correlated with DNA sequence, repeats, and predicted DNA structure. *PLoS Genet,* 2**,** e26.

BOUCHER, S. & BENNETT, S. A. L. 2003. Differential connexin expression, gap junction intercellular coupling, and hemichannel formation in NT2/D1 human neural progenitors and terminally differentiated hNT neurons. *Journal of Neuroscience Research,* 72**,** 393-404.

BOUJELBENE, N., COSINSCHI, A., BOUJELBENE, N., KHANFIR, K., BHAGWATI, S., HERRMANN, E., MIRIMANOFF, R.-O., OZSAHIN, M. & ZOUHAIR, A. 2011. Pure seminoma: A review and update. *Radiat Oncol,* 6**,** 90.

BRAIT, M., MALDONADO, L., BEGUM, S., LOYO, M., WEHLE, D., TAVORA, F. F., LOOIJENGA, L. H. J., KOWALSKI, J., ZHANG, Z. & ROSENBAUM, E. 2012. DNA methylation profiles delineate epigenetic heterogeneity in seminoma and non-seminoma. *British Journal of Cancer,* 106**,** 414-423.

BROWN, R. & STRATHDEE, G. 2002. Epigenomics and epigenetic therapy of cancer. *Trends in Molecular Medicine,* 8**,** S43-S48.

BURBEE, D. G., FORGACS, E., ZÖCHBAUER-MÜLLER, S., SHIVAKUMAR, L., FONG, K., GAO, B., RANDLE, D., KONDO, M., VIRMANI, A. & BADER, S. 2001. Epigenetic inactivation of RASSF1A in lung and breast cancers and malignant phenotype suppression. *Journal of the National Cancer Institute,* 93**,** 691-699.

BURTON, A., MULLER, J., TU, S., PADILLA-LONGORIA, P., GUCCIONE, E. & TORRES-PADILLA, M.-E. 2013. Single-cell profiling of epigenetic modifiers identifies PRDM14 as an inducer of cell fate in the mammalian embryo. *Cell Reports,* 5**,** 687-701.

CARNINCI, P., SANDELIN, A., LENHARD, B., KATAYAMA, S., SHIMOKAWA, K., PONJAVIC, J., SEMPLE, C. A., TAYLOR, M. S., ENGSTRÖM, P. G. & FRITH, M. C. 2006. Genome-wide analysis of mammalian promoter architecture and evolution. *Nature Genetics,* 38**,** 626-635.

CARPENTER, S., CARLSON, T., DELLACASAGRANDE, J., GARCIA, A., GIBBONS, S., HERTZOG, P., LYONS, A., LIN, L.-L., LYNCH, M. & MONIE, T. 2009. TRIL, a functional component of the TLR4 signaling complex, highly expressed in brain. *The Journal of Immunology,* 183**,** 3989-3995.

CHAN, Y. S., GÖKE, J., LU, X., VENKATESAN, N., FENG, B., SU, I. & NG, H. H. 2013. A PRC2-Dependent Repressive Role of PRDM14 in Human Embryonic Stem Cells and Induced Pluripotent Stem Cell Reprogramming. *Stem Cells,* 31**,** 682-692.

CHARLES, M. A., JOHNSON, I. T. & BELSHAW, N. J. 2012. Supra-physiological folic acid concentrations induce aberrant DNA methylation in normal human cells in vitro. *Epigenetics,* 7**,** 689-694.

CHARLET, J., SCHNEKENBURGER, M., BROWN, K. W. & DIEDERICH, M. 2012. DNA demethylation increases sensitivity of neuroblastoma cells to chemotherapeutic drugs. *Biochemical Pharmacology,* 83**,** 858-865.

CHAUBERT, P., GUILLOU, L., KURT, A.-M., BERTHOLET, M.-M., METTHEZ, G., LEISINGER, H.-J., BOSMAN, F. & SHAW, P. 1997. Frequent p16INK4 (MTS1) gene inactivation in testicular germ cell tumors. *The American Journal of Pathology,* 151**,** 859.

CHÉDIN, F., LIEBER, M. R. & HSIEH, C.-L. 2002. The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. *Proceedings of the National Academy of Sciences of the United States of America,* 99**,** 16916-16921.

CHEN, K. S. & AMATRUDA, J. F. 2013. A big catch for germ cell tumour research. *PLoS Genet,* 9**,** e1003481.

CHEN, V. W., RUIZ, B., KILLEEN, J. L., COTÉ, T. R., WU, X. C., CORREA, C. N. & HOWE, H. L. 2003. Pathology and classification of ovarian tumors. *Cancer,* 97**,** 2631-2642.

CHENG, P.-Y., LIN, Y.-P., CHEN, Y.-L., LEE, Y.-C., TAI, C.-C., WANG, Y.-T., CHEN, Y.-J., KAO, C.-F. & YU, J. 2011. Interplay between SIN3A and STAT3 mediates

chromatin conformational changes and GFAP expression during cellular differentiation. *PLoS ONE,* 6**,** e22018.

CHENG, X. & BLUMENTHAL, R. M. 2008. Mammalian DNA methyltransferases: a Structural Perspective. *Structure,* 16**,** 341-350.

CHIK, F. & SZYF, M. 2011. Effects of specific DNMT gene depletion on cancer cell transformation and breast cancer cell invasion; toward selective DNMT inhibitors. *Carcinogenesis,* 32**,** 224-232.

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**,** 5483-5495.

CHRISTOPH, F., KEMPKENSTEFFEN, C., WEIKERT, S., KRAUSE, H., SCHOSTAK, M., MILLER, K. & SCHRADER, M. 2007. Frequent epigenetic inactivation of p53 target genes in seminomatous and nonseminomatous germ cell tumors. *Cancer Letters,* 247**,** 137-142.

CHURIKOV, D., ZALENSKAYA, I. A. & ZALENSKY, A. O. 2004. Male germline-specific histones in mouse and man. *Cytogenetic and Genome Research,* 105**,** 203-214.

CLARK, S. J. 2007. Action at a distance: epigenetic silencing of large chromosomal regions in carcinogenesis. *Human Molecular Genetics,* 16**,** R88-R95.

COMB, M. & GOODMAN, H. M. 1990. CpG methylation inhibits proenkephalin gene expression and binding of the transcription factor AP-2. *Nucleic Acids Research,* 18**,** 3975-3982.

COOPER, D. N. & KRAWCZAK, M. 1989. Cytosine methylation and the fate of CpG dinucleotides in vertebrate genomes. *Human genetics,* 83**,** 181-188.

COOPER, D. N., TAGGART, M. H. & BIRD, A. P. 1983. Unmethylated domains in vertebrate DNA. *Nucleic Acids Research,* 11**,** 647-658.

CREA, F., NOBILI, S., PAOLICCHI, E., PERRONE, G., NAPOLI, C., LANDINI, I., DANESI, R. & MINI, E. 2011. Epigenetics and chemoresistance in colorectal cancer: an opportunity for treatment tailoring and novel therapeutic strategies. *Drug Resistance Updates,* 14**,** 280-296.

CURTIN, K., SLATTERY, M. L. & SAMOWITZ, W. S. 2011. CpG island methylation in colorectal cancer: past, present and future. *Pathology Research International,* 2011.

CURTO, M. L., D'ANGELO, P., CECCHETTO, G., KLERSY, C., DALL'IGNA, P., FEDERICO, A., SIRACUSA, F., ALAGGIO, R., BERNINI, G. & CONTE, M. 2007. Mature and immature teratomas: results of the first paediatric Italian study. *Pediatric Surgery International,* 23**,** 315-322.

CUSACK, M. & SCOTTING, P. 2013. DNA methylation in germ cell tumour aetiology: current understanding and outstanding questions. *Reproduction,* 146**,** R49-R60.

DAMMANN, R., YANG, G. & PFEIFER, G. P. 2001. Hypermethylation of the cpG island of Ras association domain family 1A (RASSF1A), a putative tumor suppressor gene from the 3p21. 3 locus, occurs in a large percentage of human breast cancers. *Cancer Research,* 61**,** 3105-3109.

DE FELICI, M. 2013. Origin, migration, and proliferation of human primordial germ cells. *Oogenesis.* Springer.

DEATON, A. M. & BIRD, A. 2011. CpG islands and the regulation of transcription. *Genes & Development,* 25**,** 1010-1022.

DI PIETRO, A., DE VRIES, E. G., GIETEMA, J. A., SPIERINGS, D. C. & DE JONG, S. 2005. Testicular germ cell tumours: the paradigm of chemo-sensitive solid tumours. *The International Journal of Biochemistry & Cell Biology,* 37**,** 2437-2456.

DIECKMANN, K. P. & SKAKKEBAEK, N. E. 1999. Carcinoma in situ of the testis: review of biological and clinical features. *International Journal of Cancer,* 83**,** 815-822.

DUALE, N., LINDEMAN, B., KOMADA, M., OLSEN, A.-K., ANDREASSEN, A., SODERLUND, E. J. & BRUNBORG, G. 2007. Molecular portrait of cisplatin induced response in human testis cancer cell lines based on gene expression profiles. *Molecular Cancer,* 6**,** 53.

ECHEVARRIA, M. E., FANGUSARO, J. & GOLDMAN, S. 2008. Pediatric central nervous system germ cell tumors: a review. *The Oncologist,* 13**,** 690-699.

ECKERT, D., BIERMANN, K., NETTERSHEIM, D., GILLIS, A. J. M., STEGER, K., JÄCK, H.-M., MÜLLER, A. M., LOOIJENGA, L. H. J. & SCHORLE, H. 2008. Expression of BLIMP1/PRMT5 and concurrent histone H2A/H4 arginine 3 dimethylation in fetal germ cells, CIS/IGCNU and germ cell tumors. *BMC Developmental Biology,* 8**,** 106.

EHRLICH, M. & JIANG, G. 2005. DNA Hypo-vs. Hypermethylation in Cancer. *DNA Methylation and Cancer Therapy.* Springer.

ELZINGA-TINKE, J. E., DOHLE, G. R. & LOOIJENGA, L. H. 2015. Etiology and early pathogenesis of malignant testicular germ cell tumors: towards possibilities for preinvasive diagnosis. *Asian Journal of Andrology,* 17**,** 381.

ESTECIO, M. R., GHARIBYAN, V., SHEN, L., IBRAHIM, A. E., DOSHI, K., HE, R., JELINEK, J., YANG, A. S., YAN, P. S. & HUANG, T. H. M. 2007. LINE-1 hypomethylation in cancer is highly variable and inversely correlated with microsatellite instability. *PLoS ONE,* 2**,** e399.

ESTELLER, M. 2002. CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. *Oncogene,* 21**,** 5427-5440.

ESTELLER, M. 2007. Epigenetic gene silencing in cancer: the DNA hypermethylome. *Human Molecular Genetics,* 16**,** R50-R59.

ESTELLER, M. 2011. Cancer Epigenetics for the 21st Century What's Next? *Genes & Cancer,* 2**,** 604-606.

ESTELLER, M., FRAGA, M. F., GUO, M., GARCIA-FONCILLAS, J., HEDENFALK, I., GODWIN, A. K., TROJAN, J., VAURS-BARRIÈRE, C., BIGNON, Y.-J. & RAMUS, S. 2001. DNA methylation patterns in hereditary human cancers mimic sporadic tumorigenesis. *Human Molecular Genetics,* 10**,** 3001-3007.

ESTELLER, M., GARCIA-FONCILLAS, J., ANDION, E., GOODMAN, S. N., HIDALGO, O. F., VANACLOCHA, V., BAYLIN, S. B. & HERMAN, J. G. 2000. Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. *New England Journal of Medicine,* 343**,** 1350-1354.

ETOH, T., KANAI, Y., USHIJIMA, S., NAKAGAWA, T., NAKANISHI, Y., SASAKO, M., KITANO, S. & HIROHASHI, S. 2004. Increased DNA methyltransferase 1 (DNMT1) protein expression correlates significantly with poorer tumor differentiation and frequent DNA hypermethylation of multiple CpG islands in gastric cancers. *The American Journal of Pathology,* 164**,** 689-699.

FAN, L., JIANG, J., GAO, J., SONG, H., LIU, J., YANG, L., LI, Z., CHEN, Y., ZHANG, Q. & WANG, X. 2015. Identification and Characterization of a PRDM14 Homolog in Japanese Flounder (Paralichthys olivaceus). *International Journal of Molecular Sciences,* 16**,** 9097-9118.

FEINBERG, A. P. & TYCKO, B. 2004. The history of cancer epigenetics. *Nature Reviews Cancer,* 4**,** 143-153.

FEINBERG, A. P. & VOGELSTEIN, B. 1983. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature,* 301**,** 89-92.

FELDMAN, N., GERSON, A., FANG, J., LI, E., ZHANG, Y., SHINKAI, Y., CEDAR, H. & BERGMAN, Y. 2006. G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nature Cell Biology,* 8**,** 188-194.

FELTUS, F., LEE, E., COSTELLO, J., PLASS, C. & VERTINO, P. 2003. Predicting aberrant CpG island methylation. *Proceedings of the National Academy of Sciences of the United States of America,* 100**,** 12253-12258.

FORN, M., MUÑOZ, M., TAURIELLO, D. V. F., MERLOS-SUÁREZ, A., RODILLA, V., BIGAS, A., BATLLE, E., JORDÀ, M. & PEINADO, M. A. 2013. Long range epigenetic silencing is a trans-species mechanism that results in cancer specific deregulation by overriding the chromatin domains of normal cells. *Molecular Oncology,* 7**,** 1129-1141.

FOUSE, S. D., SHEN, Y., PELLEGRINI, M., COLE, S., MEISSNER, A., VAN NESTE, L., JAENISCH, R. & FAN, G. 2008. Promoter CpG Methylation Contributes to ES Cell Gene Regulation in Parallel with Oct4/Nanog, PcG Complex, and Histone H3 K4/K27 Trimethylation. *Cell Stem Cell,* 2**,** 160-169.

FRAGA, D., MEULIA, T. & FENSTER, S. 2008. Real-Time PCR. *Current Protocols Essential Laboratory Techniques***,** 1-10.

FRANKEN, N. A., RODERMOND, H. M., STAP, J., HAVEMAN, J. & VAN BREE, C. 2006. Clonogenic assay of cells in vitro. *Nature Protocols,* 1**,** 2315-2319.

FRAZIER, A. L., WELDON, C. & AMATRUDA, J. Fetal and neonatal germ cell tumors. 2012. Elsevier, 17**,** 222-230.

FUKS, F., HURD, P. J., WOLF, D., NAN, X., BIRD, A. P. & KOUZARIDES, T. 2003. The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. *Journal of Biological Chemistry,* 278**,** 4035-4040.

FURUKAWA, S., HARUTA, M., ARAI, Y., HONDA, S., OHSHIMA, J., SUGAWARA, W., KAGEYAMA, Y., HIGASHI, Y., NISHIDA, K. & TSUNEMATSU, Y. 2009. Yolk sac tumor but not seminoma or teratoma is associated with abnormal epigenetic reprogramming pathway and shows frequent hypermethylation of various tumor suppressor genes. *Cancer Science,* 100**,** 698-708.

GAMA-SOSA, M. A., SLAGEL, V. A., TREWYN, R. W., OXENHANDLER, R., KUO, K. C., GEHRKE, C. W. & EHRLICH, M. 1983. The 5-methylcytosine content of DNA from human tumors. *Nucleic Acids Research,* 11**,** 6883-6894.

GARCIA-MANERO, G., DANIEL, J., SMITH, T. L., KORNBLAU, S. M., LEE, M.-S., KANTARJIAN, H. M. & ISSA, J.-P. J. 2002. DNA methylation of multiple promoter-associated CpG islands in adult acute lymphocytic leukemia. *Clinical Cancer Research,* 8**,** 2217-2224.

GARDINER-GARDEN, M. & FROMMER, M. 1987. CpG islands in vertebrate genomes. *Journal of Molecular Biology,* 196**,** 261-282.

GAUDET, F., HODGSON, J. G., EDEN, A., JACKSON-GRUSBY, L., DAUSMAN, J., GRAY, J. W., LEONHARDT, H. & JAENISCH, R. 2003. Induction of tumors in mice by genomic hypomethylation. *Science,* 300**,** 489-492.

GAZDAR, A. F., GIRARD, L., LOCKWOOD, W. W., LAM, W. L. & MINNA, J. D. 2010. Lung cancer cell lines as tools for biomedical discovery and research. *Journal of the National Cancer Institute,* 102**,** 1310-1321.

GHOSHAL, K., DATTA, J., MAJUMDER, S., BAI, S., KUTAY, H., MOTIWALA, T. & JACOB, S. T. 2005. 5-Aza-deoxycytidine induces selective degradation of DNA methyltransferase 1 by a proteasomal pathway that requires the KEN box, bromo-

adjacent homology domain, and nuclear localization signal. *Molecular and Cellular Biology,* 25**,** 4727-4741.

GIBBS, R. A., WEINSTOCK, G. M., METZKER, M. L., MUZNY, D. M., SODERGREN, E. J., SCHERER, S., SCOTT, G., STEFFEN, D., WORLEY, K. C. & BURCH, P. E. 2004. Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature,* 428**,** 493-521.

GNYSZKA, A., JASTRZĘBSKI, Z. & FLIS, S. 2013. DNA methyltransferase inhibitors and their emerging role in epigenetic therapy of cancer. *Anticancer Research,* 33**,** 2989-2996.

GOBEL, U., SCHNEIDER, D. T., CALAMINUS, G., HAAS, R. J., SCHMIDT, P. & HARMS, D. 2000. Germ-cell tumors in childhood and adolescence. *Annals of Oncology,* 11**,** 263-272.

GODMANN, M., GASHAW, I., EILDERMANN, K., SCHWEYER, S., BERGMANN, M., SKOTHEIM, R. & BEHR, R. 2009. The pluripotency transcription factor Krüppel-like factor 4 is strongly expressed in intratubular germ cell neoplasia unclassified and seminoma. *Molecular Human Reproduction,* 15**,** 479-488.

GOEDERT, J. J., PURDUE, M. P., MCNEEL, T. S., MCGLYNN, K. A. & ENGELS, E. A. 2007. Risk of germ cell tumors among men with HIV/acquired immunodeficiency syndrome. *Cancer Epidemiology Biomarkers & Prevention,* 16**,** 1266-1269.

GOEL, A., ARNOLD, C. N., NIEDZWIECKI, D., CARETHERS, J. M., DOWELL, J. M., WASSERMAN, L., COMPTON, C., MAYER, R. J., BERTAGNOLLI, M. M. & BOLAND, C. R. 2004. Frequent inactivation of PTEN by promoter hypermethylation in microsatellite instability-high sporadic colorectal cancers. *Cancer Research,* 64**,** 3014-3021.

GOLL, M. G., KIRPEKAR, F., MAGGERT, K. A., YODER, J. A., HSIEH, C.-L., ZHANG, X., GOLIC, K. G., JACOBSEN, S. E. & BESTOR, T. H. 2006. Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. *Science,* 311**,** 395-398.

GRABOLE, N., TISCHLER, J., HACKETT, J. A., KIM, S., TANG, F., LEITCH, H. G., MAGNÚSDÓTTIR, E. & SURANI, M. A. 2013. Prdm14 promotes germline fate and naive pluripotency by repressing FGF signalling and DNA methylation. *EMBO Reports,* 14**,** 629-637.

GREENBAUM, M. P., IWAMORI, T., BUCHOLD, G. M. & MATZUK, M. M. 2011. Germ cell intercellular bridges. *Cold Spring Harbor Perspectives in Biology,* 3**,** a005850.

GU, Y., RUNYAN, C., SHOEMAKER, A., SURANI, A. & WYLIE, C. 2009. Steel factor controls primordial germ cell survival and motility from the time of their specification in the allantois, and provides a continuous niche throughout their migration. *Development,* 136**,** 1295-1303.

GUERRERO-SETAS, D., PÉREZ-JANICES, N., BLANCO-FERNANDEZ, L., OJER, A., CAMBRA, K., BERDASCO, M., ESTELLER, M., MARIA-RUIZ, S., TORREA, N. & GUARCH, R. 2013. RASSF2 hypermethylation is present and related to shorter survival in squamous cervical cancer. *Modern Pathology,* 26**,** 1111-1122.

GUPTA, A., GODWIN, A. K., VANDERVEER, L., LU, A. & LIU, J. 2003. Hypomethylation of the synuclein γ gene CpG island promotes its aberrant expression in breast carcinoma and ovarian carcinoma. *Cancer Research,* 63**,** 664-673.

HAN, L., SU, B., LI, W.-H. & ZHAO, Z. 2008. CpG island density and its correlations with genomic features in mammalian genomes. *Genome Biololgy,* 9**,** R79.

HARTMANN, O., SPYRATOS, F., HARBECK, N., DIETRICH, D., FASSBENDER, A., SCHMITT, M., EPPENBERGER-CASTORI, S., VUAROQUEAUX, V., LEREBOURS, F. & WELZEL, K. 2009. DNA methylation markers predict outcome in node-positive, estrogen receptor-positive breast cancer with adjuvant anthracycline-based chemotherapy. *Clinical Cancer Research,* 15**,** 315-323.

HAYASHI, K., DE SOUSA LOPES, S. M. C. & SURANI, M. A. 2007. Germ cell specification in mice. *Science,* 316**,** 394-396.

HERMAN, J. G., MERLO, A., MAO, L. I., LAPIDUS, R. G., ISSA, J.-P. J., DAVIDSON, N. E., SIDRANSKY, D. & BAYLIN, S. B. 1995. Inactivation of the CDKN2/p16/MTS1 gene is frequently associated with aberrant DNA methylation in all common human cancers. *Cancer Research,* 55**,** 4525-4530.

HERNANDEZ-BLAZQUEZ, F., HABIB, M., DUMOLLARD, J., BARTHELEMY, C., BENCHAIB, M., DE CAPOA, A. & NIVELEAU, A. 2000. Evaluation of global DNA hypomethylation in human colon cancer tissues by immunohistochemistry and image analysis. *Gut,* 47**,** 689-693.

HEYN, H. & ESTELLER, M. 2012. DNA methylation profiling in the clinic: applications and challenges. *Nature Reviews Genetics,* 13**,** 679-692.

HOEI-HANSEN, C., ALMSTRUP, K., NIELSEN, J., BRASK SONNE, S., GRAEM, N., SKAKKEBAEK, N., LEFFERS, H. & MEYTS, R. D. 2005. Stem cell pluripotency factor NANOG is expressed in human fetal gonocytes, testicular carcinoma in situ and germ cell tumours. *Histopathology,* 47**,** 48-56.

HOLLIDAY, R. & PUGH, J. 1975. DNA modification mechanisms and gene activity during development. *Science,* 187**,** 226-232.

HONORIO, S., AGATHANGGELOU, A., WERNERT, N., ROTHE, M., MAHER, E. R. & LATIF, F. 2003. Frequent epigenetic inactivation of the RASSF1A tumour suppressor gene in testicular tumours and distinct methylation profiles of seminoma and nonseminoma testicular germ cell tumours. *Oncogene,* 22**,** 461-466.

HORTON, Z., SCHLATTER, M. & SCHULTZ, S. 2007. Pediatric germ cell tumors. *Surgical Oncology,* 16**,** 205-213.

HOTCHKISS, R. D. 1948. The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography. *Journal of Biological Chemistry,* 175**,** 315-332.

HOULDSWORTH, J., XIAO, H., MURTY, V., CHEN, W., RAY, B., REUTER, V. E., BOSL, G. J. & CHAGANTI, R. 1998. Human male germ cell tumor resistance to cisplatin is linked to TP53 gene mutation. *Oncogene,* 16**,** 2345-2349.

HOVESTADT, V., JONES, D. T., PICELLI, S., WANG, W., KOOL, M., NORTHCOTT, P. A., SULTAN, M., STACHURSKI, K., RYZHOVA, M. & WARNATZ, H.-J. 2014. Decoding the regulatory landscape of medulloblastoma using DNA methylation sequencing. *Nature,* 510**,** 537-541.

HU, M., YAO, J., CAI, L., BACHMAN, K. E., VAN DEN BRÛLE, F., VELCULESCU, V. & POLYAK, K. 2005. Distinct epigenetic changes in the stromal cells of breast cancers. *Nature Genetics,* 37**,** 899-905.

HUANG, Y. & RAO, A. 2014. Connections between TET proteins and aberrant DNA modification in cancer. *Trends in Genetics,* 30**,** 464-474.

HUYGHE, E., MATSUDA, T. & THONNEAU, P. 2003. Increasing incidence of testicular cancer worldwide: a review. *The Journal of Urology,* 170**,** 5-11.

ILLINGWORTH, R. S. & BIRD, A. P. 2009. CpG islands – 'A rough guide'. *FEBS Letters,* 583**,** 1713-1720.

IRIZARRY, R. A., LADD-ACOSTA, C., WEN, B., WU, Z., MONTANO, C., ONYANGO, P., CUI, H., GABO, K., RONGIONE, M. & WEBSTER, M. 2009. The human colon cancer methylome shows similar hypo-and hypermethylation at conserved tissue-specific CpG island shores. *Nature Genetics,* 41**,** 178-186.

ISSA, J.-P. 2004. CpG island methylator phenotype in cancer. *Nature Reviews Cancer,* 4**,** 988-993.

ISSA, J.-P. J., SHEN, L. & TOYOTA, M. 2005. CIMP, at last. *Gastroenterology,* 129**,** 1121-1124.

JABBARI, K. & BERNARDI, G. 2004. Cytosine methylation and CpG, TpG (CpA) and TpA frequencies. *Gene,* 333**,** 143-149.

JEYAPALAN, J. N., NOOR, D. A. M., LEE, S. H., TAN, C. L., APPLEBY, V. A., KILDAY, J. P., PALMER, R. D., SCHWALBE, E. C., CLIFFORD, S. C. & WALKER, D. A. 2011. Methylator phenotype of malignant germ cell tumours in children identifies strong candidates for chemotherapy resistance. *British Journal of Cancer,* 105**,** 575-585.

JIN, B., LI, Y. & ROBERTSON, K. D. 2011. DNA Methylation Superior or Subordinate in the Epigenetic Hierarchy? *Genes & Cancer,* 2**,** 607-617.

JOHNSON, T. B. & COGHILL, R. D. 1925. Researches on pyrimidines. The discovery of 5-methyl-cytosine in tuberculunic acid, the nucleic acid of the *Tubercle bacillus*. *Journal of the American Chemical Society,* 47**,** 2838-2844.

JONES, P. A. & LAIRD, P. W. 1999. Cancer-epigenetics comes of age. *Nature Genetics,* 21**,** 163-167.

JONES, P. A. & TAYLOR, S. M. 1980. Cellular differentiation, cytidine analogs and DNA methylation. *Cell,* 20**,** 85-93.

JUO, Y., JOHNSTON, F., ZHANG, D., JUO, H., WANG, H., PAPPOU, E., YU, T., EASWARAN, H., BAYLIN, S. & ENGELAND, M. 2014. Prognostic value of CpG island methylator phenotype among colorectal cancer patients: a systematic review and meta-analysis. *Annals of Oncology***,** 25, 2314-2327.

JÜTTERMANN, R., LI, E. & JAENISCH, R. 1994. Toxicity of 5-aza-2'-deoxycytidine to mammalian cells is mediated primarily by covalent trapping of DNA methyltransferase rather than DNA demethylation. *Proceedings of the National Academy of Sciences of the United States of America,* 91**,** 11797-11801.

KAMACHI, Y., UCHIKAWA, M. & KONDOH, H. 2000. Pairing SOX off: with partners in the regulation of embryonic development. *Trends in Genetics,* 16**,** 182-187.

KANAI, Y., USHIJIMA, S., KONDO, Y., NAKANISHI, Y. & HIROHASHI, S. 2001. DNA methyltransferase expression and DNA methylation of CPG islands and peri-centromeric satellite regions in human colorectal and stomach cancers. *International Journal of Cancer,* 91**,** 205-212.

KAO, J., SALARI, K., BOCANEGRA, M., CHOI, Y.-L., GIRARD, L., GANDHI, J., KWEI, K. A., HERNANDEZ-BOUSSARD, T., WANG, P. & GAZDAR, A. F. 2009. Molecular profiling of breast cancer cell lines defines relevant tumor models and provides a resource for cancer gene discovery. *PLoS ONE,* 4**,** e6146.

KELLY JR, T. J. & SMITH, H. O. 1970. A restriction enzyme from Hemophilus influenzae: II. Base sequence of the recognition site. *Journal of Molecular Biology,* 51**,** 393-409.

KELLY, T. K., DE CARVALHO, D. D. & JONES, P. A. 2010. Epigenetic modifications as therapeutic targets. *Nature Biotechnology,* 28**,** 1069-1078.

KERSH, C. R., CONSTABLE, W. C., EISERT, D. R., SPAULDING, C. A., HAHN, S. S., JENRETTE, J. M. & MARKS, R. D. 1988. Primary central nervous system germ cell tumors. Effect of histologic confirmation on radiotherapy. *Cancer,* 61**,** 2148-2152.

KIM, H., KIM, Y. H., KIM, S. E., KIM, N. G., NOH, S. H. & KIM, H. 2003. Concerted promoter hypermethylation of hMLH1, p16INK4A, and E-cadherin in gastric carcinomas with microsatellite instability. *The Journal of Pathology,* 200**,** 23-31.

KINNEY, S. R. M., SMIRAGLIA, D. J., JAMES, S. R., MOSER, M. T., FOSTER, B. A. & KARPF, A. R. 2008. Stage-specific alterations of DNA methyltransferase expression, DNA hypermethylation, and DNA hypomethylation during prostate cancer progression in the transgenic adenocarcinoma of mouse prostate model. *Molecular Cancer Research,* 6**,** 1365-1374.

KNUDSON, A. G. 1971. Mutation and cancer: statistical study of retinoblastoma. *Proceedings of the National Academy of Sciences of the United States of America,* 68**,** 820-823.

KNUDSON, A. G. 2001. Two genetic hits (more or less) to cancer. *Nature Reviews Cancer,* 1**,** 157-162.

KÖBERLE, B., TOMICIC, M. T., USANOVA, S. & KAINA, B. 2010. Cisplatin resistance: preclinical findings and clinical implications. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer,* 1806**,** 172-182.

KOUL, S., HOULDSWORTH, J., MANSUKHANI, M. M., DONADIO, A., MCKIERNAN, J. M., REUTER, V. E., BOSL, G. J., CHAGANTI, R. S. & MURTY, V. V. 2002. Characteristic promoter hypermethylation signatures in male germ cell tumors. *Molecular Cancer,* 1**,** 8.

KUCUKONER, M., INAL, A., KAPLAN, M. A., URAKCI, Z., FIRAT, U., UCMAK, F., ISIKDOGAN, A. & TEKBAS, G. 2012. Germ Cell Tumor Located in Gastrointestinal System: A Report of Two Cases. *World Journal of Oncology,* 3**,** 134-137.

KURUP, V. P., BARRIOS, C. S., RAJU, R., JOHNSON, B. D., LEVY, M. B. & FINK, J. N. 2007. Immune response modulation by curcumin in a latex allergy model. *Clinical and Molecular Allergy,* 5**,** 1.

LARSEN, F., GUNDERSEN, G., LOPEZ, R. & PRYDZ, H. 1992. CpG islands as gene markers in the human genome. *Genomics,* 13**,** 1095-1107.

LI, E., BEARD, C. & JAENISCH, R. 1993. Role for DNA methylation in genomic imprinting. *Nature,* 366**,** 362-365.

LI, Y., MCCLINTICK, J., ZHONG, L., EDENBERG, H. J., YODER, M. C. & CHAN, R. J. 2005. Murine embryonic stem cell differentiation is promoted by SOCS-3 and inhibited by the zinc finger transcription factor Klf4. *Blood,* 105**,** 635-637.

LIND, G. E., SKOTHEIM, R. I. & LOTHE, R. A. 2007. The epigenome of testicular germ cell tumors. *Acta Pathologica, Microbiologica, et Immunologica Scandinavica,* 115**,** 1147-1160.

LISTER, R. & ECKER, J. R. 2009. Finding the fifth base: genome-wide sequencing of cytosine methylation. *Genome Research,* 19**,** 959-966.

LIU, K., HUANG, W., GAO, W. & HE, W. 2015. Effect of combined 5-aza-2'deoxycytidine and cisplatin treatment on the P15 lung adenocarcinoma cell line. *Oncology Letters,* 9**,** 2007-2012.

LOOIJENGA, L., GILLIS, A., STOOP, H., BIERMANN, K. & OOSTERHUIS, J. 2011. Dissecting the molecular pathways of (testicular) germ cell tumour pathogenesis;

from initiation to treatment-resistance. *International Journal of Andrology,* 34**,** e234-e251.

LOOIJENGA, L. H., VERKERK, A. J., DEKKER, M. C., GURP, R. J., GILLIS, A. J. M. & OOSTERHUIS, J. W. 1998. Genomic imprinting in testicular germ cell tumours. *Acta Pathologica, Microbiologica, et Immunologica Scandinavica,*106**,** 187-197.

LOOIJENGA, L. H. J., GILLIS, A. D., STOOP, H. J., HERSMUS, R. & OOSTERHUIS, J. W. 2007. Chromosomes and Expression in Human Testicular Germ-Cell Tumors. *Annals of the New York Academy of Sciences,* 1120**,** 187-214.

LOOIJENGA, L. H. J., ZAFARANA, G., GRYGALEWICZ, B., SUMMERSGILL, B., DEBIEC-RYCHTER, M., VELTMAN, J., SCHOENMAKERS, E. F. P. M., RODRIGUEZ, S., JAFER, O. & CLARK, J. 2003. Role of gain of 12p in germ cell tumour development. *Acta Pathologica, Microbiologica, et Immunologica Scandinavica,* 111**,** 161-170.

LUJAMBIO, A., CALIN, G. A., VILLANUEVA, A., ROPERO, S., SÁNCHEZ-CÉSPEDES, M., BLANCO, D., MONTUENGA, L. M., ROSSI, S., NICOLOSO, M. S. & FALLER, W. J. 2008. A microRNA DNA methylation signature for human cancer metastasis. *Proceedings of the National Academy of Sciences of the United States of America,* 105**,** 13556-13561.

LUO, J., LI, Y., YANG, L., LIU, X., WEN, S., WANG, F., ZHANG, J., ZHANG, X. & DONG, Z. 2009. Effect of a methylation inhibitor 5-aza-2'-deoxycytidine on SHP-1 gene expression, proliferation and apoptosis in K562 cells. *Journal of Experimental Hematology/Chinese Association of Pathophysiology,* 17**,** 309-314.

MA, Z., SWIGUT, T., VALOUEV, A., RADA-IGLESIAS, A. & WYSOCKA, J. 2011. Sequence-specific regulator Prdm14 safeguards mouse ESCs from entering extraembryonic endoderm fates. *Nature Structural & Molecular Biology,* 18**,** 120-127.

MAGNÚSDÓTTIR, E., DIETMANN, S., MURAKAMI, K., GÜNESDOGAN, U., TANG, F., BAO, S., DIAMANTI, E., LAO, K., GOTTGENS, B. & SURANI, M. A. 2013. A tripartite transcription factor network regulates primordial germ cell specification in mice. *Nature Cell Biology,* 15**,** 905-915.

MAGNÚSDÓTTIR, E. & SURANI, M. A. 2014. How to make a primordial germ cell. *Development,* 141**,** 245-252.

MAMSEN, L. S., BRØCHNER, C. B., BYSKOV, A. G. & MØLLGARD, K. 2012. The migration and loss of human primordial germ stem cells from the hind gut epithelium towards the gonadal ridge. *International Journal of Developmental Biology,* 56**,** 771-8.

MARTIN-SUBERO, J. I., AMMERPOHL, O., BIBIKOVA, M., WICKHAM-GARCIA, E., AGIRRE, X., ALVAREZ, S., BRUGGEMANN, M., BUG, S., CALASANZ, M. J. & DECKERT, M. 2009. A comprehensive microarray-based DNA methylation study of 367 hematological neoplasms. *PLoS ONE,* 4**,** e6986.

MARUYAMA, M., ICHISAKA, T., NAKAGAWA, M. & YAMANAKA, S. 2005. Differential roles for Sox15 and Sox2 in transcriptional control in mouse embryonic stem cells. *Journal of Biological Chemistry,* 280**,** 24371-24379.

MATHIEU, M. G., LINLEY, A. J., REEDER, S. P., BADOUAL, C., TARTOUR, E., REES, R. C. & MCARDLE, S. E. 2010. HAGE, a cancer/testis antigen expressed at the protein level in a variety of cancers. *Cancer Immunity Archive,* 10**,** 2.

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. 2010. Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature,* 466**,** 253-257.

MENG, H., CAO, Y., QIN, J., SONG, X., ZHANG, Q., SHI, Y. & CAO, L. 2015. DNA Methylation, Its Mediators and Genome Integrity. *International Journal of Biological Sciences,* 11**,** 604-617.

MERCER, T. R., DINGER, M. E. & MATTICK, J. S. 2009. Long non-coding RNAs: insights into functions. *Nature Reviews Genetics,* 10**,** 155-159.

MESSERSCHMIDT, D. M., KNOWLES, B. B. & SOLTER, D. 2014. DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. *Genes & Development,* 28**,** 812-828.

MICHAEL, H., LUCIA, J., FOSTER, R. S. & ULBRIGHT, T. M. 2000. The pathology of late recurrence of testicular germ cell tumors. *The American Journal of Surgical Pathology,* 24**,** 257-273.

MISHINA, M., OGAWA, O., KINOSHITA, H., OKA, H., OKUMURA, K., MITSUMORI, K., KAKEHI, Y., REEVE, A. E. & YOSHIDA, O. 1996. Equivalent parental distribution of frequently lost alleles and biallelic expression of the H19 gene in human testicular germ cell tumors. *Cancer Science,* 87**,** 816-823.

MOCHIZUKI, K., TACHIBANA, M., SAITOU, M., TOKITAKE, Y. & MATSUI, Y. 2012. Implication of DNA Demethylation and Bivalent Histone Modification for Selective Gene Regulation in Mouse Primordial Germ Cells. *PLoS ONE,* 7**,** e46036.

MOMPARLER, R. L. & GOODMAN, J. 1977. In vitro cytotoxic and biochemical effects of 5-aza-2′-deoxycytidine. *Cancer Research,* 37**,** 1636-1639.

MOORE, L. D., LE, T. & FAN, G. 2013. DNA methylation and its basic function. *Neuropsychopharmacology,* 38**,** 23-38.

MORRISSEY, C., MARTINEZ, A., ZATYKA, M., AGATHANGGELOU, A., HONORIO, S., ASTUTI, D., MORGAN, N. V., MOCH, H., RICHARDS, F. M. & KISHIDA, T. 2001. Epigenetic inactivation of the RASSF1A 3p21. 3 tumor suppressor gene in both clear cell and papillary renal cell carcinoma. *Cancer Research,* 61**,** 7277-7281.

MOSTERT, M. C., VERKERK, A. J., VAN DE POL, M., HEIGHWAY, J., MARYNEN, P., ROSENBERG, C., GEURTS VAN KESSEL, A., ECHTEN, J. V., DE JONG, B. & OOSTERHUIS, J. W. 1998. Identification of the critical region of 12p over-representation in testicular germ cell tumors of adolescents and adults. *Oncogene,* 16.

MÜNZ, M., KIEU, C., MACK, B., SCHMITT, B., ZEIDLER, R. & GIRES, O. 2004. The carcinoma-associated antigen EpCAM upregulates c-myc and induces cell proliferation. *Oncogene,* 23**,** 5748-5758.

NAGAMATSU, G., KOSAKA, T., KAWASUMI, M., KINOSHITA, T., TAKUBO, K., AKIYAMA, H., SUDO, T., KOBAYASHI, T., OYA, M. & SUDA, T. 2011. A germ cell-specific gene, Prmt5, works in somatic cell reprogramming. *Journal of Biological Chemistry,* 286**,** 10641-10648.

NAKAKI, F. & SAITOU, M. 2014. PRDM14: a unique regulator for pluripotency and epigenetic reprogramming. *Trends in Biochemical Sciences,* 39**,** 289-298.

NETTERSHEIM, D., JOSTES, S., SHARMA, R., SCHNEIDER, S., HOFMANN, A., FERREIRA, H. J., HOFFMANN, P., KRISTIANSEN, G., ESTELLER, M. B. & SCHORLE, H. 2015. BMP Inhibition in Seminomas Initiates Acquisition of Pluripotency via NODAL Signaling Resulting in Reprogramming to an Embryonal Carcinoma. *PLoS Genet,* 11**,** e1005415.

NETTO, G. J., NAKAI, Y., NAKAYAMA, M., JADALLAH, S., TOUBAJI, A., NONOMURA, N., ALBADINE, R., HICKS, J. L., EPSTEIN, J. I. & YEGNASUBRAMANIAN, S. 2008. Global DNA hypomethylation in intratubular germ cell neoplasia and seminoma, but not in nonseminomatous male germ cell tumors. *Modern Pathology,* 21**,** 1337-1344.

NORDLING, C. 1953. A new theory on the cancer-inducing mechanism. *British Journal of Cancer,* 7**,** 68.

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. W., 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. & ALDAPE, K. 2010. Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. *Cancer Cell,* 17**,** 510-522.

O'BYRNE, K. J., BARR, M. P. & GRAY, S. G. 2011. The role of epigenetics in resistance to cisplatin chemotherapy in lung cancer. *Cancers,* 3**,** 1426-1453.

OAKLEY, A. J., YAMADA, T., LIU, D., COGGAN, M., CLARK, A. G. & BOARD, P. G. 2008. The identification and structural characterization of C7orf24 as γ-glutamyl cyclotransferase an essential enzyme in the γ-glutamyl cycle. *Journal of Biological Chemistry,* 283**,** 22031-22042.

OGINO, S., NOSHO, K., KIRKNER, G. J., KAWASAKI, T., MEYERHARDT, J. A., LODA, M., GIOVANNUCCI, E. L. & FUCHS, C. S. 2009. CpG island methylator phenotype, microsatellite instability, BRAF mutation and clinical outcome in colon cancer. *Gut,* 58**,** 90-96.

OHINATA, Y., PAYER, B., O'CARROLL, D., ANCELIN, K., ONO, Y., SANO, M., BARTON, S. C., OBUKHANYCH, T., NUSSENZWEIG, M. & TARAKHOVSKY, A. 2005. Blimp1 is a critical determinant of the germ cell lineage in mice. *Nature,* 436**,** 207-213.

OKAMOTO, K. 2012. Epigenetics: a way to understand the origin and biology of testicular germ cell tumors. *International Journal of Urology,* 19**,** 504-511.

OKAMOTO, K. & KAWAKAMI, T. 2007. Epigenetic profile of testicular germ cell tumours. *International Journal of Andrology,* 30**,** 385-392.

OKASHITA, N., KUMAKI, Y., EBI, K., NISHI, M., OKAMOTO, Y., NAKAYAMA, M., HASHIMOTO, S., NAKAMURA, T., SUGASAWA, K. & KOJIMA, N. 2014. PRDM14 promotes active DNA demethylation through the Ten-eleven translocation (TET)-mediated base excision repair pathway in embryonic stem cells. *Development,* 141**,** 269-280.

OKITA, K., ICHISAKA, T. & YAMANAKA, S. 2007. Generation of germline-competent induced pluripotent stem cells. *Nature,* 448**,** 313-317.

OLASZ, J., MANDOKY, L., GÉCZI, L., BODROGI, I., CSUKA, O. & BAK, M. 2005. Influence of hMLH1 methylation, mismatch repair deficiency and microsatellite instability on chemoresistance of testicular germ-cell tumors. *Anticancer Research,* 25**,** 4319-4324.

OOSTERHUIS, J. W. & LOOIJENGA, L. H. J. 2005. Testicular germ-cell tumours in a broader perspective. *Nature Reviews Cancer,* 5**,** 210-222.

OOSTERHUIS, J. W., VAN GURP, R., ROSENBERG, C. & LOOIJENGA, L. H. J. 2002. Overrepresentation of the Short Arm of Chromosome 12 Is Related to Invasive

Growth of Testicular Seminomas and Non-seminomas. *Germ Cell Tumours V*. Springer.

OTTESEN, A. M., SKAKKEBÆK, N. E., LUNDSTEEN, C., LEFFERS, H., LARSEN, J. & MEYTS, R. D. 2003. High-resolution comparative genomic hybridization detects extra chromosome arm 12p material in most cases of carcinoma in situ adjacent to overt germ cell tumors, but not before the invasive tumor development. *Genes, Chromosomes and Cancer,* 38**,** 117-125.

OZATO, K., SHIN, D.-M., CHANG, T.-H. & MORSE, H. C. 2008. TRIM family proteins and their emerging roles in innate immunity. *Nature Reviews Immunology,* 8**,** 849-860.

PALMER, R. D., BARBOSA-MORAIS, N. L., GOODING, E. L., MURALIDHAR, B., THORNTON, C. M., PETT, M. R., ROBERTS, I., SCHNEIDER, D. T., THORNE, N. & TAVARÉ, S. 2008. Pediatric malignant germ cell tumors show characteristic transcriptome profiles. *Cancer Research,* 68**,** 4239-4247.

PALMER, R. D., MURRAY, M. J., SAINI, H. K., VAN DONGEN, S., ABREU-GOODGER, C., MURALIDHAR, B., PETT, M. R., THORNTON, C. M., NICHOLSON, J. C. & ENRIGHT, A. J. 2010. Malignant germ cell tumors display common microRNA profiles resulting in global changes in expression of messenger RNA targets. *Cancer Research,* 70**,** 2911-2923.

PANDEY, R. R., TOKUZAWA, Y., YANG, Z., HAYASHI, E., ICHISAKA, T., KAJITA, S., ASANO, Y., KUNIEDA, T., SACHIDANANDAM, R. & CHUMA, S. 2013. Tudor domain containing 12 (TDRD12) is essential for secondary PIWI interacting RNA biogenesis in mice. *Proceedings of the National Academy of Sciences of the United States of America,* 110**,** 16492-16497.

PARK, S. Y., YOO, E. J., CHO, N. Y., KIM, N. & KANG, G. H. 2009. Comparison of CpG island hypermethylation and repetitive DNA hypomethylation in premalignant stages of gastric cancer, stratified for Helicobacter pylori infection. *The Journal of Pathology,* 219**,** 410-416.

PARRY, L. & CLARKE, A. R. 2011. The roles of the methyl-CpG binding proteins in cancer. *Genes & Cancer,* 2**,** 618-630.

PAULSEN, M. & FERGUSON-SMITH, A. C. 2001. DNA methylation in genomic imprinting, development, and disease. *The Journal of Pathology,* 195**,** 97-110.

PAYER, B. & LEE, J. T. 2008. X chromosome dosage compensation: how mammals keep the balance. *Annual Review of Genetics,* 42**,** 733-772.

PELTOMÄKI, P. 1991. DNA methylation changes in human testicular cancer. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease,* 1096**,** 187-196.

PFAFFL, M. W. 2001. A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Research,* 29**,** e45-e45.

PIYATHILAKE, C. J., JOHANNING, G. L., FROST, A. R., WHITESIDE, M. A., MARINE, U., GRIZZLE, W. E., HEIMBURGER, D. C. & NIVELEAU, A. 2000. Immunohistochemical evaluation of global DNA methylation: comparison with in vitro radiolabeled methyl incorporation assay. *Biotechnic & Histochemistry,* 75**,** 251-258.

PREVITI, C., HARARI, O., ZWIR, I. & DEL VAL, C. 2009. Profile analysis and prediction of tissue-specific CpG island methylation classes. *BMC Bioinformatics,* 10**,** 116.

PYAGAY, P., HEROULT, M., WANG, Q., LEHNERT, W., BELDEN, J., LIAW, L., FRIESEL, R. E. & LINDNER, V. 2005. Collagen triple helix repeat containing 1, a

novel secreted protein in injured and diseased arteries, inhibits collagen expression and promotes cell migration. *Circulation Research,* 96**,** 261-268.

RADDATZ, G., GUZZARDO, P. M., OLOVA, N., FANTAPPIÉ, M. R., RAMPP, M., SCHAEFER, M., REIK, W., HANNON, G. J. & LYKO, F. 2013. Dnmt2-dependent methylomes lack defined DNA methylation patterns. *Proceedings of the National Academy of Sciences of the United States of America,* 110**,** 8627-8631.

RAISS, G. G., ANDALOUSSI, M. M. B., RAISSOUNI, S. S., MRABTI, H. H. & ERRIHANI, H. H. 2011. Spermatocytic seminoma at the national institute of oncology in Morocco. *BMC Research Notes,* 4**,** 218.

RAJPERT-DE MEYTS, E. 2006. Developmental model for the pathogenesis of testicular carcinoma in situ: genetic and environmental aspects. *Human Reproduction Update,* 12**,** 303-323.

RAO, X., EVANS, J., CHAE, H., PILROSE, J., KIM, S., YAN, P., HUANG, R., LAI, H., LIN, H. & LIU, Y. 2013. CpG island shore methylation regulates caveolin-1 expression in breast cancer. *Oncogene,* 32**,** 4519-4528.

RAZIN, A. & CEDAR, H. 1991. DNA methylation and gene expression. *Microbiological Reviews,* 55**,** 451-458.

REIK, W., SANTOS, F. & DEAN, W. 2003. Mammalian epigenomics: reprogramming the genome for development and therapy. *Theriogenology,* 59**,** 21-32.

RICHARDS, E. J. & ELGIN, S. C. 2002. Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. *Cell,* 108**,** 489-500.

RIJLAARSDAM, M. A., TAX, D. M. J., GILLIS, A. J. M., DORSSERS, L. C. J., KOESTLER, D. C., DE RIDDER, J. & LOOIJENGA, L. H. J. 2015. Genome wide DNA methylation profiles provide clues to the origin and pathogenesis of germ cell tumors. *PLoS ONE,* 10,7.

RIZWANA, R. & HAHN, P. J. 1999. CpG methylation reduces genomic instability. *Journal of Cell Science,* 112**,** 4513-4519.

ROBERTSON, K. D. 2001. DNA methylation, methyltransferases, and cancer. *Oncogene,* 20**,** 3139-3155.

ROBERTSON, K. D. 2005. DNA methylation and human disease. *Nature Reviews Genetics,* 6**,** 597-610.

ROBERTSON, K. D. & JONES, P. A. 2000. DNA methylation: past, present and future directions. *Carcinogenesis,* 21**,** 461-467.

RODRIGUEZ, S., JAFER, O., GOKER, H., SUMMERSGILL, B. M., ZAFARANA, G., GILLIS, A. J. M., VAN GURP, R., OOSTERHUIS, J. W., LU, Y. J. & HUDDART, R. 2003. Expression profile of genes from 12p in testicular germ cell tumors of adolescents and adults associated with i (12p) and amplification at 12p11. 2–p12. 1. *Oncogene,* 22**,** 1880-1891.

ROSE, N. R. & KLOSE, R. J. 2014. Understanding the relationship between DNA methylation and histone lysine methylation. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms,* 1839**,** 1362-1372.

ROSENBERG, C., VAN GURP, R. J., GEELEN, E., OOSTERHUIS, J. W. & LOOIJENGA, L. H. 2000. Overrepresentation of the short arm of chromosome 12 is related to invasive growth of human testicular seminomas and nonseminomas. *Oncogene,* 19**,** 5858-5862.

ROTHSTEIN, M. A., CAI, Y. & MARCHANT, G. E. 2009. The ghost in our genes: legal and ethical implications of epigenetics. *Health Matrix,* 19**,** 1.

SAADEH, H. & SCHULZ, R. 2014. Protection of CpG islands against de novo DNA methylation during oogenesis is associated with the recognition site of E2f1 and E2f2. *Epigenetics & Chromatin,* 7**,** 26.

SAMOWITZ, W. S., ALBERTSEN, H., HERRICK, J., LEVIN, T. R., SWEENEY, C., MURTAUGH, M. A., WOLFF, R. K. & SLATTERY, M. L. 2005. Evaluation of a Large, Population-Based Sample Supports a CpG Island Methylator Phenotype in Colon Cancer. *Gastroenterology,* 129**,** 837-845.

SCARANO, E. 1973. DNA Methylation. *Nature,* 246**,** 539-539.

SCARANO, E., IACCARINO, M., GRIPPO, P. & PARISI, E. 1967. The heterogeneity of thymine methyl group origin in DNA pyrimidine isostichs of developing sea urchin embryos. *Proceedings of the National Academy of Sciences of the United States of America,* 57**,** 1394-1400.

SCHAEFER, M., POLLEX, T., HANNA, K., TUORTO, F., MEUSBURGER, M., HELM, M. & LYKO, F. 2010. RNA methylation by Dnmt2 protects transfer RNAs against stress-induced cleavage. *Genes & Development,* 24**,** 1590-1595.

SCHOLZ, M., ZEHENDER, M., THALMANN, G., BORNER, M., THÖNI, H. & STUDER, U. 2002. Extragonadal retroperitoneal germ cell tumor: evidence of origin in the testis. *Annals of Oncology,* 13**,** 121-124.

SCHWEIKERT, E. M., DEVARAJAN, A., WITTE, I., WILGENBUS, P., AMORT, J., FÖRSTERMANN, U., SHABAZIAN, A., GRIJALVA, V., SHIH, D. M. & FARIAS-EISNER, R. 2012. PON3 is upregulated in cancer tissues and protects against mitochondrial superoxide-mediated cell death. *Cell Death & Differentiation,* 19**,** 1549-1560.

SCOTTING, P. 2006. Are cranial germ cell tumours really tumours of germ cells? *Neuropathology and Applied Neurobiology,* 32**,** 569-574.

SESTERHENN, I. A. & DAVIS JR, C. J. 2004. Pathology of germ cell tumors of the testis. *Cancer Control,* 11**,** 374-387.

SHARMA, S., SALEHI, F., SCHEITHAUER, B. W., ROTONDO, F., SYRO, L. V. & KOVACS, K. 2009. Role of MGMT in tumor development, progression, diagnosis, treatment and prognosis. *Anticancer Research,* 29**,** 3759-3768.

SHEN, J., WANG, S., ZHANG, Y.-J., WU, H.-C., KIBRIYA, M. G., JASMINE, F., AHSAN, H., WU, D. P. H., SIEGEL, A. B. & REMOTTI, H. 2013. Exploring genome-wide DNA methylation profiles altered in hepatocellular carcinoma using Infinium HumanMethylation 450 BeadChips. *Epigenetics,* 8**,** 34-43.

SHEN, L., AHUJA, N., SHEN, Y., HABIB, N. A., TOYOTA, M., RASHID, A. & ISSA, J.-P. J. 2002. DNA methylation and environmental exposures in human hepatocellular carcinoma. *Journal of the National Cancer Institute,* 94**,** 755-761.

SHEN, L., KONDO, Y., GUO, Y., ZHANG, J., ZHANG, L., AHMED, S., SHU, J., CHEN, X., WATERLAND, R. A. & ISSA, J. 2007. Genome-wide profiling of DNA methylation reveals a class of normally methylated CpG island promoters. *PLoS Genet,* 3**,** 2023-2036.

SHINJO, K., OKAMOTO, Y., AN, B., YOKOYAMA, T., TAKEUCHI, I., FUJII, M., OSADA, H., USAMI, N., HASEGAWA, Y. & ITO, H. 2012. Integrated analysis of genetic and epigenetic alterations reveals CpG island methylator phenotype associated with distinct clinical characters of lung adenocarcinoma. *Carcinogenesis***,** bgs154.

SIEVERS, S., ALEMAZKOUR, K., ZAHN, S., PERLMAN, E. J., GILLIS, A. J. M., LOOIJENGA, L. H. J., GÖBEL, U. & SCHNEIDER, D. T. 2005. IGF2/H19 imprinting analysis of human germ cell tumors (GCTs) using the methylation-

sensitive single-nucleotide primer extension method reflects the origin of GCTs in different stages of primordial germ cell development. *Genes, Chromosomes and Cancer,* 44**,** 256-264.

SINGH, N. 2002. The pathology of ovarian embryonal carcinoma. *CME Journal of Gynecologic Oncology,* 7**,** 215-218.

SKAKKEBAEK, N. E., BERTHELSEN, J. G., GIWERCMAN, A. & MÜLLER, J. 1987. Carcinoma-in-situ of the testis: possible origin from gonocytes and precursor of all types of germ cell tumours except spermatocytoma. *International Journal of Andrology,* 10**,** 19-28.

SMIRAGLIA, D. J., KAZHIYUR-MANNAR, R., OAKES, C. C., WU, Y.-Z., LIANG, P., ANSARI, T., SU, J., RUSH, L. J., SMITH, L. T. & YU, L. 2007. Restriction landmark genomic scanning (RLGS) spot identification by second generation virtual RLGS in multiple genomes with multiple enzyme combinations. *BMC Genomics,* 8**,** 446.

SMIRAGLIA, D. J., RUSH, L. J., FRÜHWALD, M. C., DAI, Z., HELD, W. A., COSTELLO, J. F., LANG, J. C., ENG, C., LI, B. & WRIGHT, F. A. 2001. Excessive CpG island hypermethylation in cancer cell lines versus primary human malignancies. *Human Molecular Genetics,* 10**,** 1413-1419.

SMIRAGLIA, D. J., SZYMANSKA, J., KRAGGERUD, S. M., LOTHE, R. A., PELTOMAKI, P. & PLASS, C. 2002. Distinct epigenetic phenotypes in seminomatous and nonseminomatous testicular germ cell tumors. *Oncogene,* 21**,** 3909-3916.

SMITH-SØRENSEN, B., LIND, G. E., SKOTHEIM, R. I., FOSSÅ, S. D., FODSTAD, Ø., STENWIG, A. E., JAKOBSEN, K. S. & LOTHE, R. A. 2002. Oncogenomics. *Oncogene,* 21**,** 8878-8884.

SNELLENBERG, S., CILLESSEN, S. A., VAN CRIEKINGE, W., BOSCH, L., MEIJER, C. J., SNIJDERS, P. J. & STEENBERGEN, R. D. 2014. Methylation-mediated repression of PRDM14 contributes to apoptosis evasion in HPV-positive cancers. *Carcinogenesis***,** 35, 11.

SONNE, S. B., KRISTENSEN, D. M., NOVOTNY, G. W., OLESEN, I. A., NIELSEN, J. E., SKAKKEBÆK, N. E., MEYTS, R. D. & LEFFERS, H. 2008. Testicular dysgenesis syndrome and the origin of carcinoma in situ testis. *International Journal of Andrology,* 31**,** 275-287.

SPROUL, D., KITCHEN, R. R., NESTOR, C. E., DIXON, J. M., SIMS, A. H., HARRISON, D. J., RAMSAHOYE, B. H. & MEEHAN, R. R. 2012. Tissue of origin determines cancer-associated CpG island promoter hypermethylation patterns. *Genome Biology,* 13**,** R84.

STRATHDEE, G., APPLETON, K., ILLAND, M., MILLAN, D. W., SARGENT, J., PAUL, J. & BROWN, R. 2001. Primary ovarian carcinomas display multiple methylator phenotypes involving known tumor suppressor genes. *The American Journal of Pathology,* 158**,** 1121-1127.

STURGEON, C. 2002. Practice guidelines for tumor marker use in the clinic. *Clinical Chemistry,* 48**,** 1151-1159.

SUBRAMANIAN, J. & GOVINDAN, R. 2008. Molecular genetics of lung cancer in people who have never smoked. *The Lancet Oncology,* 9**,** 676-682.

SUGAWARA, Y., ZASADNY, K. R., GROSSMAN, H. B., FRANCIS, I. R., CLARKE, M. F. & WAHL, R. L. 1999. Germ cell tumor: differentiation of viable tumor, mature teratoma, and necrotic tissue with FDG PET and kinetic modeling 1. *Radiology,* 211**,** 249-256.

SUMMERSGILL, B., OSIN, P., LU, Y. J., HUDDART, R. & SHIPLEY, J. 2001. Chromosomal imbalances associated with carcinoma in situ and associated testicular germ cell tumours of adolescents and adults. *British Journal of Cancer,* 85, 213.

SWAMI, M. 2010. Cancer epigenetics: long-range silencing. *Nature Reviews Cancer,* 10**,** 238-239.

SZYF, M. 2009. Epigenetics, DNA methylation, and chromatin modifying drugs. *Annual Review of Pharmacology and Toxicology,* 49**,** 243-263.

TAMARU, H. & SELKER, E. U. 2001. A histone H3 methyltransferase controls DNA methylation in Neurospora crassa. *Nature,* 414**,** 277-283.

TANG, M., XU, W., WANG, Q., XIAO, W. & XU, R. 2009. Potential of DNMT and its epigenetic regulation for lung cancer therapy. *Current Genomics,* 10**,** 336.

TANGKIJVANICH, P., HOURPAI, N., RATTANATANYONG, P., WISEDOPAS, N., MAHACHAI, V. & MUTIRANGURA, A. 2007. Serum LINE-1 hypomethylation as a potential prognostic marker for hepatocellular carcinoma. *Clinica Chimica Acta,* 379**,** 127-133.

TEODORIDIS, J. M., HARDIE, C. & BROWN, R. 2008. CpG island methylator phenotype (CIMP) in cancer: causes and implications. *Cancer Letters,* 268**,** 177-186.

TIAN, Y., HOU, Y., ZHOU, X., CHENG, H. & ZHOU, R. 2011. Tumor suppressor RASSF1A promoter: p53 binding and methylation. *PLoS ONE,* 6.

TOYOTA, M., AHUJA, N., OHE-TOYOTA, M., HERMAN, J. G., BAYLIN, S. B. & ISSA, J.-P. J. 1999. CpG island methylator phenotype in colorectal cancer. *Proceedings of the National Academy of Sciences of the United States of America,* 96**,** 8681-8686.

TOYOTA, M., KOPECKY, K. J., TOYOTA, M.-O., JAIR, K.-W., WILLMAN, C. L. & ISSA, J.-P. J. 2001. Methylation profiling in acute myeloid leukemia. *Blood,* 97**,** 2823-2829.

UDALI, S., GUARINI, P., MORUZZI, S., RUZZENENTE, A., TAMMEN, S. A., GUGLIELMI, A., CONCI, S., PATTINI, P., OLIVIERI, O., CORROCHER, R., CHOI, S.-W. & FRISO, S. 2015. Global DNA methylation and hydroxymethylation differ in hepatocellular carcinoma and cholangiocarcinoma and relate to survival rate. *Hepatology,* 62**,** 496-504.

UEKI, T., WALTER, K. M., SKINNER, H., JAFFEE, E., HRUBAN, R. H. & GOGGINS, M. 2002. Aberrant CpG island methylation in cancer cell lines arises in the primary cancers from which they were derived. *Oncogene,* 21**,** 2114-2117.

ULBRIGHT, T. M. 2004. Gonadal teratomas: a review and speculation. *Advances in Anatomic Pathology,* 11**,** 10-23.

USHIDA, H., KAWAKAMI, T., MINAMI, K., CHANO, T., OKABE, H., OKADA, Y. & OKAMOTO, K. 2012. Methylation profile of DNA repetitive elements in human testicular germ cell tumor. *Molecular Carcinogenesis,* 51**,** 711-722.

VAN DER ZWAN, Y. G., RIJLAARSDAM, M. A., ROSSELLO, F. J., NOTINI, A. J., DE BOER, S., WATKINS, D. N., GILLIS, A. J. M., DORSSERS, L. C. J., WHITE, S. J. & LOOIJENGA, L. H. J. 2014. Seminoma and embryonal carcinoma footprints identified by analysis of integrated genome-wide epigenetic and expression profiles of germ cell cancer cell lines. *PLoS ONE,* 9, 6.

VASDEV, N., MOON, A. & THORPE, A. C. 2013. Classification, epidemiology and therapies for testicular germ cell tumours. *International Journal of Developmental Biology,* 57**,** 133-139.

VAVOURI, T. & LEHNER, B. 2012. Human genes with CpG island promoters have a distinct transcription-associated chromatin organization. *Genome Biology,* 13**,** R110.

VEGA, A., BAPTISSART, M., CAIRA, F., BRUGNON, F., LOBACCARO, J.-M. A. & VOLLE, D. H. 2012. Epigenetic: a molecular link between testicular cancer and environmental exposures. *Frontiers in Endocrinology,* 3,150.

VON DER MAASE, H., RØRTH, M., WALBOM-JØRGENSEN, S., SØRENSEN, B. L., CHRISTOPHERSEN, I. S., HALD, T., JACOBSEN, G. K., BERTHELSEN, J. G. & SKAKKEBÆK, N. E. 1986. Carcinoma in situ of contralateral testis in patients with testicular germ cell cancer: study of 27 cases in 500 patients. *British Medical Journal,* 293**,** 1398-1401.

WAGNER, J. R., BUSCHE, S., GE, B., KWAN, T., PASTINEN, T. & BLANCHETTE, M. 2014. The relationship between DNA methylation, genetic and expression inter-individual variation in untransformed human fibroblasts. *Genome Biology,* 15**,** R37.

WANG, Y. & LEUNG, F. C. 2004. An evaluation of new criteria for CpG islands in the human genome as gene markers. *Bioinformatics,* 20**,** 1170-1177.

WEISENBERGER, D. J., SIEGMUND, K. D., CAMPAN, M., YOUNG, J., LONG, T. I., FAASSE, M. A., KANG, G. H., WIDSCHWENDTER, M., WEENER, D. & BUCHANAN, D. 2006. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. *Nature Genetics,* 38**,** 787-793.

WERMANN, H., STOOP, H., GILLIS, A. J., HONECKER, F., VAN GURP, R. J., AMMERPOHL, O., RICHTER, J., OOSTERHUIS, J. W., BOKEMEYER, C. & LOOIJENGA, L. H. 2010. Global DNA methylation in fetal human germ cells and germ cell tumours: association with differentiation and cisplatin resistance. *The Journal of Pathology,* 221**,** 433-442.

WEST, F. D., TERLOUW, S. L., KWON, D. J., MUMAW, J. L., DHARA, S. K., HASNEEN, K., DOBRINSKY, J. R. & STICE, S. L. 2010. Porcine induced pluripotent stem cells produce chimeric offspring. *Stem Cells and Development,* 19**,** 1211-1220.

WESTERN, P. S., VAN DEN BERGEN, J. A., MILES, D. C. & SINCLAIR, A. H. 2010. Male fetal germ cell differentiation involves complex repression of the regulatory network controlling pluripotency. *The FASEB Journal,* 24**,** 3026-3035.

WHITEHURST, A. W., BODEMANN, B. O., CARDENAS, J., FERGUSON, D., GIRARD, L., PEYTON, M., MINNA, J. D., MICHNOFF, C., HAO, W. & ROTH, M. G. 2007. Synthetic lethal screen identification of chemosensitizer loci in cancer cells. *Nature,* 446**,** 815-819.

WILLIAMS, K., CHRISTENSEN, J. & HELIN, K. 2012. DNA methylation: TET proteins—guardians of CpG islands? *EMBO Reports,* 13**,** 28-35.

WILSON, A. S., POWER, B. E. & MOLLOY, P. L. 2007. DNA hypomethylation and human diseases. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer,* 1775**,** 138-162.

WOBBES, T., OLDHOFF, J., SCHRAFFORDT, K. H. & DE VRIES, J. A. 1981. The treatment of infantile embryonal carcinoma (Yolk-Sac Tumour) of the testis in children. *Surgery in Infancy and Childhood,* 33**,** 349-354.

WORSTER, A., SHARMA, S., MOOKADAM, F. & OPIE, J. 2002. Acute presentation of choriocarcinoma: a case study and review of the literature. *Canadian Journal of Emergency Medicine,* 4**,** 111-114.

WU, L.-M., ZHANG, F., ZHOU, L., YANG, Z., XIE, H.-Y. & ZHENG, S.-S. 2010. Predictive value of CpG island methylator phenotype for tumor recurrence in hepatitis B virus-associated hepatocellular carcinoma following liver transplantation. *BMC Cancer,* 10**,** 399-399.

XU, M., KNOX, A. J., MICHAELIS, K. A., KISELJAK-VASSILIADES, K., KLEINSCHMIDT-DEMASTERS, B. K., LILLEHEI, K. O. & WIERMAN, M. E. 2012. Reprimo (RPRM) is a novel tumor suppressor in pituitary tumors and regulates survival, proliferation, and tumorigenicity. *Endocrinology,* 153**,** 2963-2973.

YABUTA, Y., KURIMOTO, K., OHINATA, Y., SEKI, Y. & SAITOU, M. 2006. Gene expression dynamics during germline specification in mice identified by quantitative single-cell gene expression profiling. *Biology of Reproduction,* 75**,** 705-716.

YAMAJI, M., SEKI, Y., KURIMOTO, K., YABUTA, Y., YUASA, M., SHIGETA, M., YAMANAKA, K., OHINATA, Y. & SAITOU, M. 2008. Critical function of Prdm14 for the establishment of the germ cell lineage in mice. *Nature Genetics,* 40**,** 1016-1022.

YAMAJI, M., UEDA, J., HAYASHI, K., OHTA, H., YABUTA, Y., KURIMOTO, K., NAKATO, R., YAMADA, Y., SHIRAHIGE, K. & SAITOU, M. 2013. PRDM14 ensures naive pluripotency through dual regulation of signaling and epigenetic pathways in mouse embryonic stem cells. *Cell Stem Cell,* 12**,** 368-382.

YAMASHITA, K., DAI, T., DAI, Y., YAMAMOTO, F. & PERUCHO, M. 2003. Genetics supersedes epigenetics in colon cancer phenotype. *Cancer Cell,* 4**,** 121-131.

YANG, X., HAN, H., DE CARVALHO, D. D., LAY, F. D., JONES, P. A. & LIANG, G. 2014. Gene body methylation can alter gene expression and is a therapeutic target in cancer. *Cancer Cell,* 26**,** 577-590.

YAO, X.-D., HONG, Y.-P., YE, D.-W. & WANG, C.-F. 2012. Primary yolk sac tumor of seminal vesicle: a case report and literature review. *World Journal of Surgical Oncology,* 10**,** 189.

YAO, Z. X., HAN, Z., XU, J., GREESON, J., LECANU, L. & PAPADOPOULOS, V. 2007. 22R-Hydroxycholesterol induces differentiation of human NT2 precursor (Ntera2/D1 teratocarcinoma) cells. *Neuroscience,* 148**,** 441-453.

YODER, J. A., WALSH, C. P. & BESTOR, T. H. 1997. Cytosine methylation and the ecology of intragenomic parasites. *Trends in Genetics,* 13**,** 335-340.

YOKOYAMA, S., AL MAHMUDA, N., MUNESUE, T., HAYASHI, K., YAGI, K., YAMAGISHI, M. & HIGASHIDA, H. 2015. Association Study between the CD157/BST1 Gene and Autism Spectrum Disorders in a Japanese Population. *Brain Sciences,* 5**,** 188-200.

YOU, JUENG S. & JONES, PETER A. 2012. Cancer Genetics and Epigenetics: Two Sides of the Same Coin? *Cancer Cell,* 22**,** 9-20.

YU, W., JIN, C., LOU, X., HAN, X., LI, L., HE, Y., ZHANG, H., MA, K., ZHU, J. & CHENG, L. 2011. Global analysis of DNA methylation by Methyl-Capture sequencing reveals epigenetic control of cisplatin resistance in ovarian cancer cell. *PLoS ONE,* 6**,** e29450.

YUAN, B.-Z., JEFFERSON, A. M., POPESCU, N. C. & REYNOLDS, S. H. 2004. Aberrant gene expression in human non small cell lung carcinoma cells exposed to demethylating agent 5-aza-2'-deoxycytidine. *Neoplasia,* 6**,** 412-419.

ZELLER, C., DAI, W., STEELE, N. L., SIDDIQ, A., WALLEY, A. J., WILHELM-BENARTZI, C., RIZZO, S., VAN DER ZEE, A., PLUMB, J. & BROWN, R. 2012.

Candidate DNA methylation drivers of acquired cisplatin resistance in ovarian cancer identified by methylome and expression profiling. *Oncogene,* 31**,** 4567-4576.

ZHANG, Y.-W., ZHENG, Y., WANG, J.-Z., LU, X.-X., WANG, Z., CHEN, L.-B., GUAN, X.-X. & TONG, J.-D. 2014. Integrated analysis of DNA methylation and mRNA expression profiling reveals candidate genes associated with cisplatin resistance in non-small cell lung cancer. *Epigenetics,* 9**,** 896-909.

ZHENG, Z., LI, L., LIU, X., WANG, D., TU, B., WANG, L., WANG, H. & ZHU, W.-G. 2012. 5-Aza-2′-deoxycytidine reactivates gene expression via degradation of pRb pocket proteins. *The FASEB Journal,* 26**,** 449-459.

ZHOU, J., YANG, F., LEU, N. A. & WANG, P. J. 2012. MNS1 is essential for spermiogenesis and motile ciliary functions in mice. *PLoS Genetics,* 8,3.

ZHU, W.-G., HILEMAN, T., KE, Y., WANG, P., LU, S., DUAN, W., DAI, Z., TONG, T., VILLALONA-CALERO, M. A. & PLASS, C. 2004. 5-aza-2′-deoxycytidine activates the p53/p21Waf1/Cip1 pathway to inhibit cell proliferation. *Journal of Biological Chemistry,* 279**,** 15161-15166.

ZÖCHBAUER-MÜLLER, S., FONG, K. M., MAITRA, A., LAM, S., GERADTS, J., ASHFAQ, R., VIRMANI, A. K., MILCHGRUB, S., GAZDAR, A. F. & MINNA, J. D. 2001. 5′ CpG island methylation of the FHIT gene is correlated with loss of gene expression in lung and breast cancer. *Cancer Research,* 61**,** 3581-3585.

Appendix

Appendix I. Tables and histograms represent the correlation between differential expression and differential methylation between seminoma and non-seminoma at islands, shores and shelves: Contingency tables and histograms of the observed and expected number of genes differentially expressed the ranges of differential methylation between YST and seminoma. P-values of the chi-squared test of association between methylation and expression are given as well as the observed (blue) and expected (red) percentage correlation for each delta-β value. The total number of genes in each category is displayed above the bar.

Islands

EC

Delta beta	Genes >2 fold Observed	Genes >2 fold Expected	No difference Observed	No difference Expected	Genes >2 fold Observed	Genes >2 fold Expected	Total	Pvalue Chi squared(2df)	Significance
0.85-0.9	2	0.295597484	0	1.50485992	0	0.199542596	2	0.003132395	**
0.8-0.85	10	2.512578616	7	12.79130932	0	1.696112064	17	1.64917E-06	***
0.75-0.8	11	2.660377358	7	13.54373928	0	1.795883362	18	1.76475E-07	***
0.7-0.75	11	3.251572327	10	16.55345912	1	2.194968553	22	1.93137E-05	***
0.65-0.7	7	4.13836478	17	21.06803888	4	2.793596341	28	0.193469499	
0.6-0.65	5	3.547169811	17	18.05831904	2	2.394511149	24	6.97E-01	
0.55-0.6	6	4.729559748	23	24.07775872	3	3.192681532	32	0.818267241	
0.5-0.55	11	6.059748428	29	30.84962836	1	4.090623213	41	0.039290292	*
0.45-0.5	6	4.877358491	20	24.83018868	7	3.29245283	33	0.068119332	
0.4-0.45	7	5.320754717	26	27.08747856	3	3.591766724	36	0.714933726	
0.35-0.4	5	6.059748428	33	30.84962836	3	4.090623213	41	0.731249644	
0.3-0.35	8	9.902515723	51	50.41280732	8	6.684676958	67	0.729361598	
0.25-0.3	11	13.15408805	68	66.96626644	10	8.879645512	89	0.774892088	
0.2-0.25	21	16.70125786	74	85.02458548	18	11.27415666	113	0.037847341	*
0.15-0.2	20	18.03144654	85	91.79645512	17	12.17209834	122	0.268067228	
0.1-0.15	32	32.9591195	169	167.7918811	22	22.24899943	223	0.980494788	
0.05-0.1	56	59.85849057	298	304.7341338	51	40.40737564	405	0.204503054	
0-0.05	288	322.9402516	1698	1644.059463	199	218.0002859	2185	0.027240555	*
Total	517	517	2632	2632	349	349	3498		

TERT

Delta beta	Genes >2 fold Observed	Genes >2 fold Expected	No difference Observed	No difference Expected	Genes >2 fold Observed	Genes >2 fold Expected	Total	Pvalue Chi squared(2df)	Significance
0.85-0.9	2	0.71525169	2	2.83095417	0	0.45379414	4	0.222529841	
0.8-0.85	6	2.503380917	8	9.908339594	0	1.588279489	14	0.032716401	*
0.75-0.8	7	1.609316304	2	6.369646882	0	1.021036814	9	1.60814E-05	***
0.7-0.75	4	1.072877536	2	4.246431255	0	0.68069121	6	0.007244111	**
0.65-0.7	6	3.039819684	9	12.03155522	2	1.928625094	17	1.61E-01	
0.6-0.65	6	3.39744553	11	13.44703231	2	2.155522164	19	2.94E-01	
0.55-0.6	3	3.218632607	12	12.73929376	3	2.042073629	18	0.776036054	
0.5-0.55	6	4.11269722	14	16.27798648	3	2.609316304	23	0.537041382	
0.45-0.5	6	4.470323065	17	17.69346356	2	2.836213373	25	0.671273963	
0.4-0.45	5	4.649135988	19	18.4012021	2	2.949661908	26	0.838737752	
0.35-0.4	8	4.827948911	14	19.10894065	5	3.063110443	27	0.096580144	
0.3-0.35	9	6.079639369	18	24.06311044	7	3.857250188	34	0.064212898	
0.25-0.3	9	10.54996243	40	41.756574	10	6.693463561	59	0.380028012	
0.2-0.25	19	17.16604057	58	67.94290008	19	10.89105935	96	0.021402327	*
0.15-0.2	30	24.4973704	90	96.96018032	17	15.54244929	137	0.392130026	
0.1-0.15	33	42.37866266	172	167.7340346	32	26.88730278	237	0.206362026	
0.05-0.1	72	71.70398197	279	283.8031555	50	45.49286251	401	0.76757054	
0-0.05	245	270.0075131	1117	1068.685199	148	171.3072878	1510	0.021584965	*
Total	476	476	1884	1884	302	302	2662		

North-Shores

YST

Delta beta	Genes >2 fold Observed	Genes >2 fold Expected	No difference Observed	No difference Expected	Genes >2 fold Observed	Genes >2 fold Expected	Total	Pvalue Chi squared(2df)	Significance
0.85-0.9	0	0.297888386	2	1.468325792	0	0.233785822	2	0.696215854	
0.8-0.85	1	0.595776772	2	2.936651584	1	0.467571644	4	0.554525153	
0.75-0.8	2	1.191553544	5	5.873303167	1	0.935143288	8	0.710751299	
0.7-0.75	3	1.787330317	6	8.809954751	3	1.402714932	12	0.170515456	
0.65-0.7	1	1.042609351	6	5.139140271	0	0.818250377	7	0.617486934	
0.6-0.65	11	2.829939668	6	13.94909502	2	2.220965309	19	7.76E-07	***
0.55-0.6	5	2.978883861	11	14.68325792	4	2.33785822	20	0.175787922	
0.5-0.55	6	4.319381599	20	21.29072398	3	3.389894419	29	0.678072908	
0.45-0.5	15	12.21342383	30	60.20135747	37	9.585218703	82	3.5119E-21	***
0.4-0.45	16	6.702488688	27	33.03733032	2	5.260180995	45	0.000331935	***
0.35-0.4	11	9.830316742	49	48.45475113	6	7.714932127	66	0.76854173	
0.3-0.35	15	12.36236802	55	60.93552036	13	9.702111614	83	0.322722756	
0.25-0.3	23	18.91591252	89	93.23868778	15	14.8453997	127	0.583889081	
0.2-0.25	18	21.59690799	109	106.4536199	18	16.9494721	145	0.695907568	
0.15-0.2	32	32.1719457	165	158.5791855	19	25.24886878	216	0.405061147	
0.1-0.15	41	45.87481146	227	226.1221719	40	36.00301659	308	0.617189357	
0.05-0.1	70	72.23793363	365	356.0690045	50	56.69306184	485	0.581726745	
0-0.05	125	148.0505279	773	729.7579186	96	116.1915535	994	0.00798653	**
Total	395	395	1947	1947	310	310	2652		

EC

Delta beta	Genes >2 fold Observed	Genes >2 fold Expected	No difference Observed	No difference Expected	Genes >2 fold Observed	Genes >2 fold Expected	Total	Pvalue Chi squared(2df)	Significance
0.8-0.85	0	0.135459425	1	0.757183454	0	0.10735712	1	0.851852447	
0.75-0.8	1	0.406378276	2	2.271550363	0	0.322071361	3	0.542896456	
0.7-0.75	4	0.948215977	2	5.300284181	1	0.751499842	7	0.002529912	**
0.65-0.7	5	1.760972529	7	9.843384907	1	1.395642564	13	3.19E-02	*
0.6-0.65	3	1.760972529	8	9.843384907	2	1.395642564	13	4.77E-01	
0.55-0.6	6	3.657404484	18	20.44395327	3	2.898642248	27	0.407352114	
0.5-0.55	7	5.960214714	33	33.31607199	4	4.723713293	44	0.862744788	
0.45-0.5	14	7.179349542	34	40.13072308	5	5.689927376	53	0.02351642	*
0.4-0.45	9	9.482159773	54	53.00284181	7	7.514998421	70	0.961476752	
0.35-0.4	16	14.35869908	82	80.26144616	8	11.37985475	106	0.540884382	
0.3-0.35	21	17.33880644	94	96.91948216	13	13.7417114	128	0.637290309	
0.25-0.3	31	24.24723713	129	135.5358383	19	19.21692453	179	0.333164149	
0.2-0.25	45	34.00031576	182	190.053047	24	26.9466372	251	0.121114841	
0.15-0.2	41	46.32712346	263	258.9567414	38	36.71613514	342	0.697472107	
0.1-0.15	58	58.78939059	331	328.6176192	45	46.59299021	434	0.959668584	
0.05-0.1	69	77.61825071	438	433.8661194	66	61.51562993	573	0.516023276	
0-0.05	99	125.0290496	720	698.8803284	104	99.09062204	923	0.042846011	*
Total	429	429	2398	2398	340	340	3167		

TERT

Delta beta	Genes >2 fold Observed	Genes >2 fold Expected	No difference Observed	No difference Expected	Genes >2 fold Observed	Genes >2 fold Expected	Total	Pvalue Chi squared(2df)	Significance
0.8-0.85	0	0.150537634	1	0.73655914	0	0.112903226	1	0.836246279	
0.75-0.8	1	0.451612903	1	2.209677419	1	0.338709677	3	0.269925116	
0.7-0.75	3	1.204301075	2	5.892473118	3	0.903225806	8	0.006357208	**
0.65-0.7	6	2.709677419	10	13.25806452	2	2.032258065	18	9.09E-02	
0.6-0.65	7	3.462365591	16	16.94086022	0	2.596774194	23	4.36E-02	*
0.55-0.6	10	4.064516129	16	19.88709677	1	3.048387097	27	0.004507816	**
0.5-0.55	5	4.967741935	27	24.30645161	1	3.725806452	33	0.317762357	
0.45-0.5	12	8.88172043	44	43.45698925	3	6.661290323	59	0.210771587	
0.4-0.45	14	14.30107527	73	69.97311828	8	10.72580645	95	0.660340824	
0.35-0.4	12	15.80645161	82	77.33870968	11	11.85483871	105	0.532794679	
0.3-0.35	33	20.47311828	82	100.172043	21	15.35483871	136	0.001476024	*
0.25-0.3	23	25.74193548	126	125.9516129	22	19.30645161	171	0.716099964	
0.2-0.25	38	34.47311828	168	168.672043	23	25.85483871	229	0.71222175	
0.15-0.2	43	48.32258065	238	236.4354839	40	36.24193548	321	0.610697778	
0.1-0.15	79	68.64516129	328	335.8709677	49	51.48387097	456	0.393318198	
0.05-0.1	78	86.55913978	431	423.5215054	66	64.91935484	575	0.607627183	
0-0.05	84	107.7849462	547	527.3763441	85	80.83870968	716	0.045206305	*
Total	448	448	2192	2192	336	336	2976		

North-Shelves

YST

Delta beta	Genes >2 fold Observed	Genes >2 fold Expected	No difference Observed	No difference Expected	Genes >2 fold Observed	Genes >2 fold Expected	Total	Pvalue Chi squared(2df)	Significance
0.75-0.8	0	0.606593407	5	4.689230769	1	0.704175824	6	0.686786125	
0.7-0.75	0	0.606593407	6	4.689230769	0	0.704175824	6	0.43232283	
0.65-0.7	1	0.808791209	7	6.252307692	0	0.938901099	8	0.584639611	
0.6-0.65	2	1.314285714	9	10.16	2	1.525714286	13	7.27E-01	
0.55-0.6	2	1.415384615	11	10.94153846	1	1.643076923	14	0.781345734	
0.5-0.55	6	1.617582418	8	12.50461538	2	1.877802198	16	0.001168719	***
0.45-0.5	5	1.415384615	9	10.94153846	0	1.643076923	14	0.003954061	**
0.4-0.45	8	3.336263736	22	25.79076923	3	3.872967033	33	0.026339914	*
0.35-0.4	4	4.145054945	28	32.04307692	9	4.811868132	41	0.124901245	
0.3-0.35	7	5.863736264	40	45.32923077	11	6.807032967	58	0.180009467	
0.25-0.3	14	7.582417582	49	58.61538462	12	8.802197802	75	0.016816679	
0.2-0.25	8	10.21098901	77	78.93538462	16	11.85362637	101	0.372210412	
0.15-0.2	19	17.79340659	131	137.5507692	26	20.65582418	176	0.411381569	
0.1-0.15	26	28.30769231	228	218.8307692	26	32.86153846	280	0.366956001	
0.05-0.1	51	57.62637363	449	445.4769231	70	66.8967033	570	0.626949177	
0-0.05	77	87.34945055	699	675.2492308	88	101.4013187	864	0.147138273	
Total	230	230	1778	1778	267	267	2275		

EC

Delta beta	Genes >2 fold Observed	Genes >2 fold Expected	No difference Observed	No difference Expected	Genes >2 fold Observed	Genes >2 fold Expected	Total	Pvalue Chi squared(2df)	Significance
0.7-0.75	0	0.49517415	3	3.038187159	1	0.466638691	4	0.575429124	
0.65-0.7	0	0.618967688	5	3.797733949	0	0.583298363	5	4.53E-01	
0.6-0.65	2	0.9903483	6	6.076374318	0	0.933277381	8	3.75E-01	
0.55-0.6	3	2.104490138	9	12.91229543	5	1.983214436	17	0.046063733	*
0.5-0.55	4	2.847251364	16	17.46957616	3	2.683172472	23	0.730613786	
0.45-0.5	6	2.723457826	15	16.71002937	1	2.566512799	22	0.079139312	
0.4-0.45	5	4.704154427	28	28.86277801	5	4.433067562	38	0.943225539	
0.35-0.4	11	7.303818716	38	44.8132606	10	6.882920688	59	0.115440656	
0.3-0.35	5	7.551405791	47	46.33235418	9	7.116240034	61	0.50400863	
0.25-0.3	11	11.63659253	67	71.39739824	16	10.96600923	94	0.270287652	
0.2-0.25	25	16.34074696	93	100.2601762	14	15.39907679	132	0.072746869	
0.15-0.2	18	23.14939152	137	142.0352497	32	21.81535879	187	0.04786579	*
0.1-0.15	31	33.3004616	199	204.3180864	39	31.38145195	269	0.341824667	
0.05-0.1	64	64.24884599	391	394.2047839	64	60.54637012	519	0.894036817	
0-0.05	110	116.984893	756	717.7717163	79	110.2433907	945	0.003504226	**
Total	295	295	1810	1810	278	278	2383		

TERT

Delta beta	Genes >2 fold Observed	Genes >2 fold Expected	No difference Observed	No difference Expected	Genes >2 fold Observed	Genes >2 fold Expected	Total	Pvalue Chi squared(2df)	Significance
0.7-0.75	1	0.541598695	2	2.969004894	1	0.489396411	4	0.538759604	
0.65-0.7	2	1.08319739	5	5.938009788	1	0.978792822	8	6.30E-01	
0.6-0.65	3	1.895595432	6	10.39151713	5	1.712887439	14	1.22E-02	**
0.55-0.6	4	2.572593801	12	14.10277325	3	2.324632953	19	0.521591653	
0.5-0.55	6	3.24959217	14	17.81402936	4	2.936378467	24	0.171206796	
0.45-0.5	6	4.33278956	21	23.75203915	5	3.915171289	32	0.532331286	
0.4-0.45	4	6.634584013	33	36.37030995	12	5.995106036	49	0.025057022	*
0.35-0.4	9	6.634584013	35	36.37030995	5	5.995106036	49	0.588563742	
0.3-0.35	16	10.69657423	51	58.63784666	12	9.665579119	79	0.12318465	
0.25-0.3	13	12.59216966	61	69.02936378	19	11.37846656	93	0.048503698	*
0.2-0.25	22	22.74714519	123	124.6982055	23	20.55464927	168	0.844261227	
0.15-0.2	26	27.07993475	152	148.4502447	22	24.46982055	200	0.82810255	
0.1-0.15	33	39.13050571	220	214.5106036	36	35.3588907	289	0.573339358	
0.05-0.1	64	72.84502447	404	399.3311582	70	65.82381729	538	0.498194677	
0-0.05	123	119.9641109	681	657.634584	82	108.4013051	886	0.025513848	*
Total	332	332	1820	1820	300	300	2452		

South-Shores

YST

Delta beta	Genes >2 fold Observed	Genes >2 fold Expected	No difference Observed	No difference Expected	Genes >2 fold Observed	Genes >2 fold Expected	Total	Pvalue Chi squared(2df)	Significance
0.85-0.9	2	0.476163769	0	2.233595064	1	0.290241167	3	0.011998893	**
0.8-0.85	2	0.793606282	3	3.722658441	0	0.483735278	5	0.292598212	
0.75-0.8	5	2.698261357	11	12.6570387	1	1.644699944	17	0.296244953	
0.7-0.75	11	3.650588895	10	17.12422883	2	2.225182277	23	0.000137579	***
0.65-0.7	7	4.28547392	15	20.10235558	5	2.612170499	27	0.07437474	
0.6-0.65	7	5.713965227	22	26.80314077	7	3.482893999	36	9.53E-02	
0.55-0.6	12	6.825014021	24	32.01486259	7	4.160123388	43	0.019555492	*
0.5-0.55	13	7.30117779	28	34.24845766	5	4.450364554	46	0.059130716	
0.45-0.5	8	8.094784072	33	37.9711161	10	4.934099832	51	0.053580468	*
0.4-0.45	16	11.11048794	47	52.11721817	7	6.772293887	70	0.26422887	
0.35-0.4	21	16.18956814	71	75.94223219	10	9.868199663	102	0.416294172	
0.3-0.35	24	18.8878295	84	88.59927089	11	11.51289961	119	0.439274921	
0.25-0.3	39	23.64946719	92	110.9352215	18	14.41531127	149	0.000872967	***
0.2-0.25	48	31.42680875	132	147.4172743	18	19.15591699	198	0.005455832	**
0.15-0.2	61	51.26696579	235	240.4837353	27	31.24929893	323	0.279335385	
0.1-0.15	59	63.17106001	306	296.3236119	33	38.5053281	398	0.501948439	
0.05-0.1	79	107.136848	534	502.5588895	62	65.30426248	675	0.008549959	**
0-0.05	152	203.3219293	1008	953.7450925	121	123.9329781	1281	0.000317463	***
Total	566	566	2655	2655	345	345	3566		

EC

Delta beta	Genes >2 fold Observed	Genes >2 fold Expected	No difference Observed	No difference Expected	Genes >2 fold Observed	Genes >2 fold Expected	Total	Pvalue Chi squared(2df)	Significance
0.85-0.9	1	0.420770128	1	2.27747958	1	0.301750292	3	0.209122004	
0.8-0.85	1	0.701283547	3	3.7957993	1	0.502917153	5	0.675231987	
0.75-0.8	5	1.402567095	5	7.5915986	0	1.005834306	10	0.003853449	**
0.7-0.75	2	1.262310385	5	6.83243874	2	0.905250875	9	0.325225503	
0.65-0.7	5	3.225904317	16	17.46067678	2	2.313418903	23	5.65E-01	
0.6-0.65	7	3.927187865	19	21.25647608	2	2.816336056	28	2.37E-01	
0.55-0.6	8	8.134889148	44	44.03127188	6	5.833838973	58	0.99651034	
0.5-0.55	13	11.08028005	58	59.97362894	8	7.946091015	79	0.819584071	
0.45-0.5	15	11.36079347	60	61.49194866	6	8.147257876	81	0.413154142	
0.4-0.45	20	14.30618436	73	77.43430572	9	10.25950992	102	0.262543559	
0.35-0.4	27	20.757993	104	112.3556593	17	14.88634772	148	0.246782322	
0.3-0.35	26	25.80723454	138	139.6854142	20	18.50735123	184	0.931387233	
0.25-0.3	37	35.6252042	185	192.8266044	32	25.54819137	254	0.367873322	
0.2-0.25	55	42.21726954	220	228.5071179	26	30.2756126	301	0.091129902	
0.15-0.2	66	59.74935823	319	323.4021004	41	42.84854142	426	0.672476518	
0.1-0.15	84	75.45810968	401	408.4280047	53	54.11388565	538	0.56979259	
0.05-0.1	94	109.4002334	608	592.1446908	78	78.45507585	780	0.273204667	
0-0.05	135	176.1624271	994	953.5047841	127	126.3327888	1256	0.00344539	**
Total	601	601	3253	3253	431	431	4285		

TERT

Delta beta	Genes >2 fold Observed	Genes >2 fold Expected	No difference Observed	No difference Expected	Genes >2 fold Observed	Genes >2 fold Expected	Total	Pvalue Chi squared(2df)	Significance
0.9-0.95	1	0.154014599	0	0.737469586	0	0.108515815	1	0.064155497	
0.85-0.9	0	0.462043796	1	2.212408759	2	0.325547445	3	0.007677235	**
0.8-0.85	2	0.924087591	3	4.424817518	1	0.651094891	6	0.387040744	
0.75-0.8	7	2.002189781	6	9.587104623	0	1.410705596	13	0.000493433	***
0.7-0.75	6	3.080291971	12	14.74939173	2	2.170316302	20	0.192687841	
0.65-0.7	7	4.158394161	16	19.91167883	4	2.929927007	27	2.12E-01	
0.6-0.65	4	5.698540146	28	27.2863747	5	4.015085158	37	6.82E-01	
0.55-0.6	16	8.778832117	37	42.03576642	4	6.18540146	57	0.025793127	*
0.5-0.55	18	10.4729927	43	50.14793187	7	7.379075426	68	0.039795029	*
0.45-0.5	25	15.24744526	60	73.00948905	14	10.74306569	99	0.008465725	**
0.4-0.45	21	17.86569343	71	85.54647202	24	12.58783455	116	0.001249713	***
0.35-0.4	28	22.33211679	100	106.93309	17	15.73479319	145	0.369769286	
0.3-0.35	26	26.18248175	135	125.3698297	9	18.44768856	170	0.061436964	
0.25-0.3	49	41.27591241	174	197.6418491	45	29.08223844	268	0.00151413	**
0.2-0.25	52	49.74671533	238	238.2026764	33	35.05060827	323	0.894847233	
0.15-0.2	72	67.30437956	331	322.2742092	34	47.42141119	437	0.112903644	
0.1-0.15	79	77.31532847	372	370.2097324	51	54.47493917	502	0.875013891	
0.05-0.1	76	109.350365	550	523.6034063	84	77.04622871	710	0.002323043	**
0-0.05	144	170.6481752	854	817.1163017	110	120.2355231	1108	0.035127037	*
Total	633	633	3031	3031	446	446	4110		

South-Shelves

YST

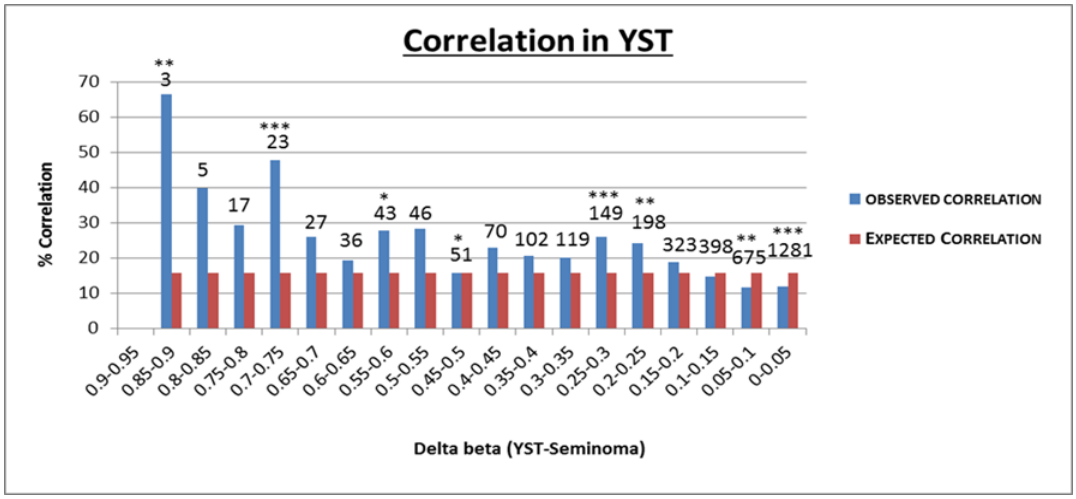
Delta beta	Genes >2 fold Observed	Genes >2 fold Expected	No difference Observed	No difference Expected	Genes >2 fold Observed	Genes >2 fold Expected	Total	Pvalue Chi squared(2df)	Significance
0.8-0.85	0	0.22639821	2	1.540044743	0	0.233557047	2	0.741808936	
0.75-0.8	2	0.339597315	1	2.310067114	0	0.35033557	3	0.009993984	**
0.7-0.75	1	0.565995526	3	3.850111857	1	0.583892617	5	0.664634336	
0.65-0.7	2	1.018791946	6	6.930201342	1	1.051006711	9	0.584983964	
0.6-0.65	3	1.131991051	4	7.700223714	3	1.167785235	10	2.09E-02	*
0.55-0.6	5	2.037583893	11	13.86040268	2	2.102013423	18	0.086195441	
0.5-0.55	3	1.697986577	7	11.55033557	5	1.751677852	15	0.012187268	**
0.45-0.5	1	2.150782998	17	14.63042506	1	2.218791946	19	0.434091684	
0.4-0.45	3	3.395973154	21	23.10067114	6	3.503355705	30	0.364872584	
0.35-0.4	7	5.886353468	40	40.04116331	5	6.072483221	52	0.818669502	
0.3-0.35	3	5.093959732	34	34.65100671	8	5.255033557	45	0.315558561	
0.25-0.3	12	8.829530201	56	60.06174497	10	9.108724832	78	0.472288894	
0.2-0.25	10	10.07472036	68	68.53199105	11	10.39328859	89	0.980149269	
0.15-0.2	24	19.58344519	126	133.2138702	23	20.20268456	173	0.411886265	
0.1-0.15	32	30.56375839	201	207.9060403	37	31.53020134	270	0.536391943	
0.05-0.1	50	58.29753915	400	396.5615213	65	60.1409396	515	0.448566725	
0-0.05	95	102.1055928	724	694.560179	83	105.3342282	902	0.039205069	*
Total	253	253	1721	1721	261	261	2235		

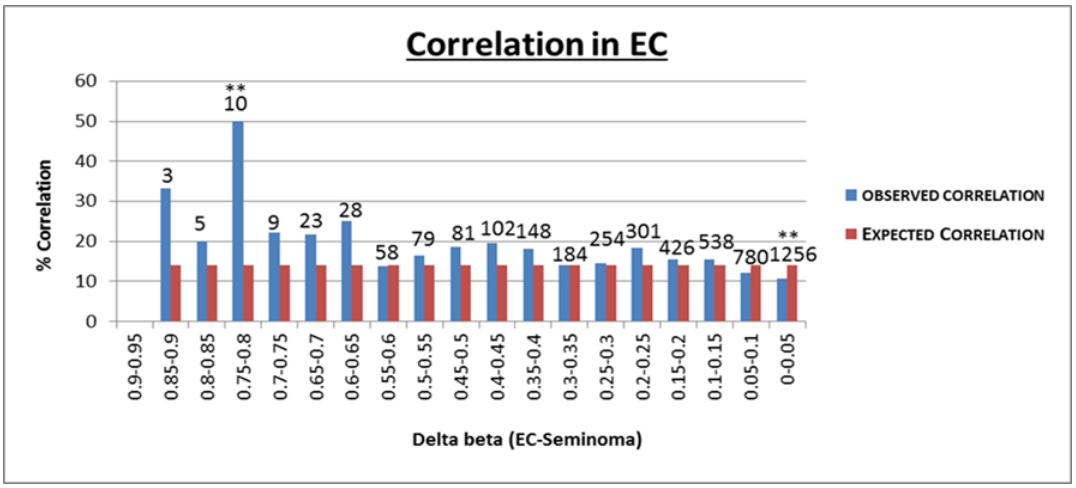
EC

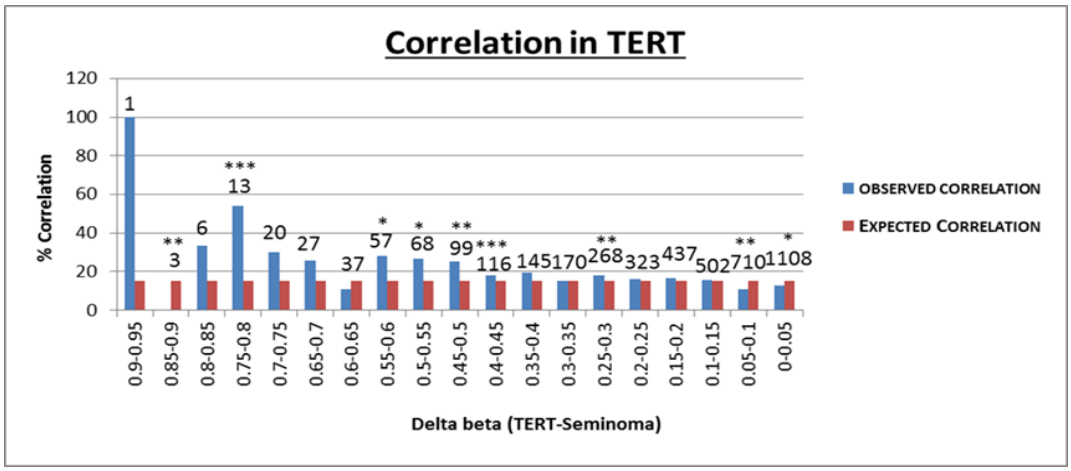
Delta beta	Genes >2 fold Observed	Genes >2 fold Expected	No difference Observed	No difference Expected	Genes >2 fold Observed	Genes >2 fold Expected	Total	Pvalue Chi squared(2df)	Significance
0.85-0.9	0	0.123874839	1	0.763823403	0	0.112301757	1	0.856756466	
0.8-0.85	1	0.123874839	0	0.763823403	0	0.112301757	1	0.029119938	*
0.75-0.8	0	0.247749679	2	1.527646807	0	0.224603515	2	0.734031642	
0.7-0.75	1	0.495499357	2	3.055293613	1	0.44920703	4	0.459894973	
0.65-0.7	2	0.990998714	6	6.110587227	0	0.898414059	8	3.81E-01	
0.6-0.65	2	1.114873553	6	6.87441063	1	1.010715817	9	6.66E-01	
0.55-0.6	0	1.61037291	10	9.929704243	3	1.459922846	13	0.198344201	
0.5-0.55	7	2.477496785	9	15.27646807	4	2.246035148	20	0.002238426	**
0.45-0.5	6	3.963994856	18	24.44234891	8	3.593656237	32	0.017021215	*
0.4-0.45	6	5.574367767	34	34.37205315	5	5.053579083	45	0.98162368	
0.35-0.4	5	7.060865838	40	43.53793399	12	6.401200171	57	0.055410133	
0.3-0.35	6	7.432490356	51	45.8294042	3	6.738105444	60	0.230703733	
0.25-0.3	13	10.77711102	61	66.45263609	13	9.770252893	87	0.372776176	
0.2-0.25	11	15.85597943	99	97.76939563	18	14.37462495	128	0.298644495	
0.15-0.2	19	22.79297042	134	140.5435062	31	20.66352336	184	0.04720667	*
0.1-0.15	30	34.18945564	212	210.8152593	34	30.99528504	276	0.666543636	
0.05-0.1	70	63.30004286	388	390.3137591	53	57.38619803	511	0.589158759	
0-0.05	110	110.8679811	709	683.621946	76	100.5100729	895	0.031337914	*
Total	289	289	1782	1782	262	262	2333		

TERT

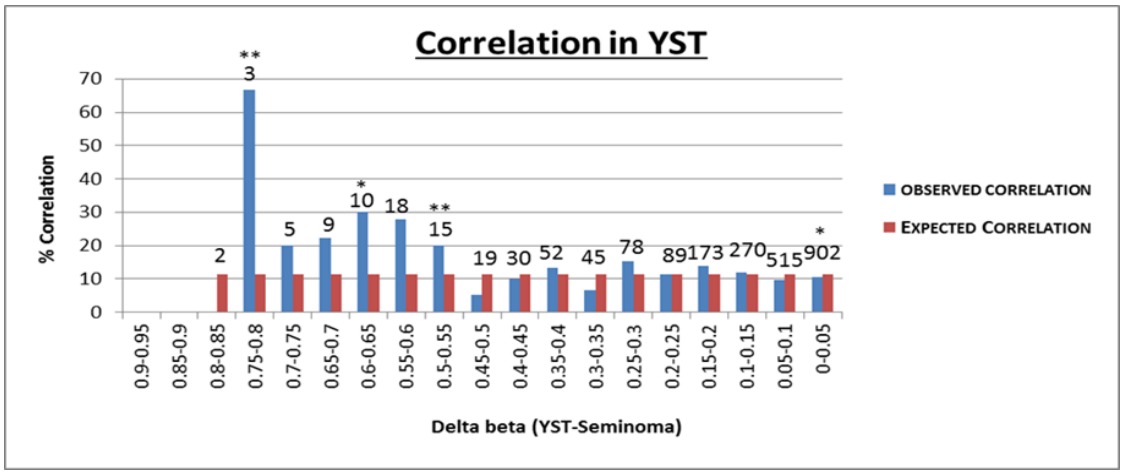
Delta beta	Genes >2 fold Observed	Genes >2 fold Expected	No difference Observed	No difference Expected	Genes >2 fold Observed	Genes >2 fold Expected	Total	Pvalue Chi squared(2df)	Significance
0.8-0.85	1	0.125	0	0.754071661	0	0.120928339	1	0.030197383	*
0.75-0.8	0	0.5	3	3.016286645	1	0.483713355	4	0.591217871	
0.7-0.75	1	0.75	4	4.524429967	1	0.725570033	6	0.883416442	
0.65-0.7	1	0.625	4	3.770358306	0	0.604641694	5	0.65585585	
0.6-0.65	3	1.75	10	10.55700326	1	1.692996743	14	5.47E-01	
0.55-0.6	7	2.75	10	16.58957655	5	2.660423453	22	3.62E-03	**
0.5-0.55	3	2.875	15	17.34364821	5	2.781351792	23	0.351354538	
0.45-0.5	3	4	20	24.13029316	9	3.86970684	32	0.020666401	*
0.4-0.45	1	5.625	36	33.93322476	8	5.441775244	45	0.076869112	
0.35-0.4	5	8.375	50	50.5228013	12	8.102198697	67	0.197835718	
0.3-0.35	17	9.875	55	59.57166124	7	9.553338762	79	0.045636433	*
0.25-0.3	13	13	80	78.42345277	11	12.57654723	104	0.891667837	
0.2-0.25	12	15.625	98	94.25895765	15	15.11604235	125	0.609458451	
0.15-0.2	19	24.125	142	145.5358306	32	23.33916938	193	0.11143964	
0.1-0.15	33	38.125	229	229.9918567	43	36.88314332	305	0.425781639	
0.05-0.1	79	68.375	419	412.4771987	49	66.1478013	547	0.045060353	*
0-0.05	109	110.5	677	666.5993485	98	106.9006515	884	0.630111955	
Total	307	307	1852	1852	297	297	2456		

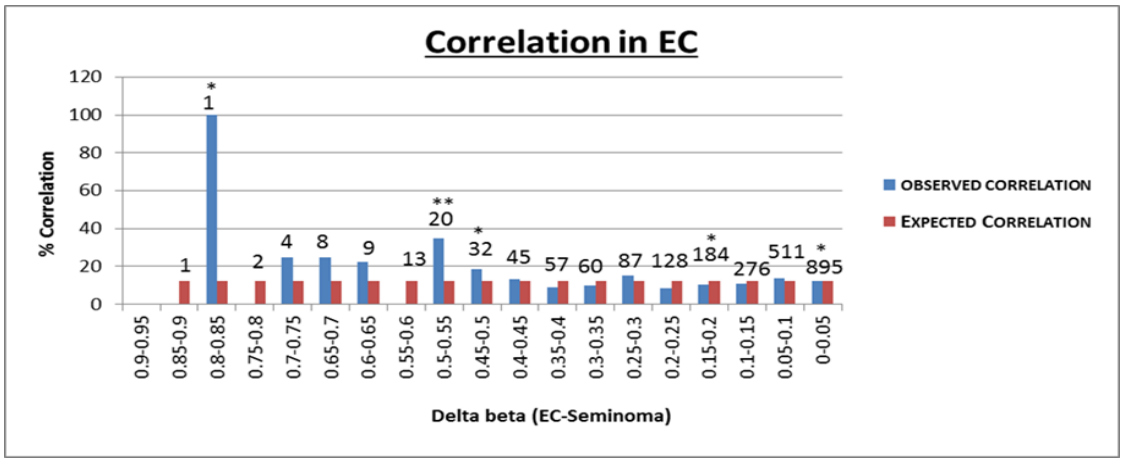

Correlation in YST
% Correlation
Delta beta (YST-Seminoma)
OBSERVED CORRELATION
EXPECTED CORRELATION
0.9-0.95
0.85-0.9
0.8-0.85
0.75-0.8
0.7-0.75
0.65-0.7
0.6-0.65
0.55-0.6
0.5-0.55
0.45-0.5
0.4-0.45
0.35-0.4
0.3-0.35
0.25-0.3
0.2-0.25
0.15-0.2
0.1-0.15
0.05-0.1
0-0.05

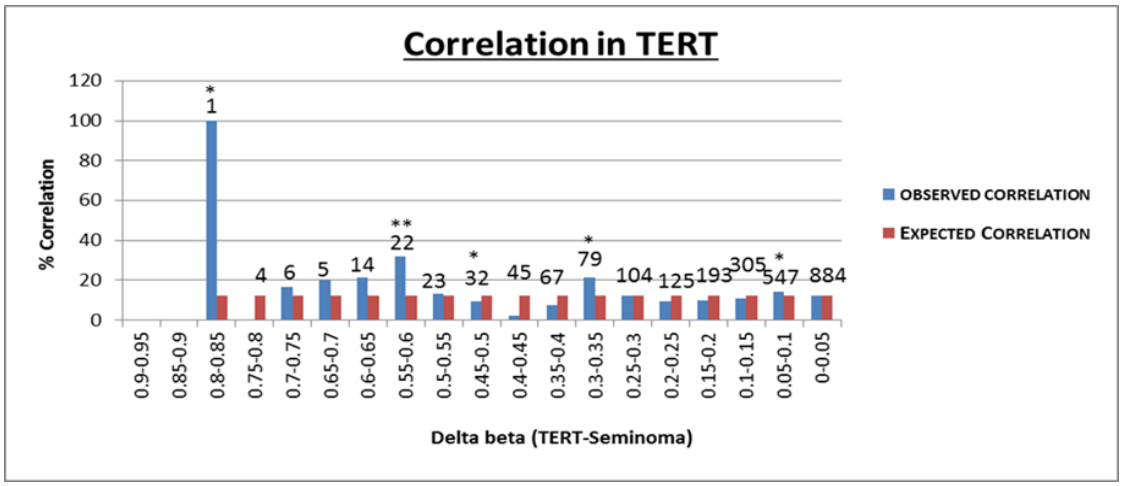

Correlation in EC
% Correlation
Delta beta (EC-Seminoma)
OBSERVED CORRELATION
EXPECTED CORRELATION
0.9-0.95
0.85-0.9
0.8-0.85
0.75-0.8
0.7-0.75
0.65-0.7
0.6-0.65
0.55-0.6
0.5-0.55
0.45-0.5
0.4-0.45
0.35-0.4
0.3-0.35
0.25-0.3
0.2-0.25
0.15-0.2
0.1-0.15
0.05-0.1
0-0.05


Correlation in TERT
% Correlation
Delta beta (TERT-Seminoma)
OBSERVED CORRELATION
EXPECTED CORRELATION
0.9-0.95
0.85-0.9
0.8-0.85
0.75-0.8
0.7-0.75
0.65-0.7
0.6-0.65
0.55-0.6
0.5-0.55
0.45-0.5
0.4-0.45
0.35-0.4
0.3-0.35
0.25-0.3
0.2-0.25
0.15-0.2
0.1-0.15
0.05-0.1
0-0.05

South-Shelves







Appendix II. R commands (performed by Claire Wallace)

Input commands in R are shown in blue writing, preceded by >. R outputs are shown in black writing, and commentary explanations of the process are shown in green writing, preceded by #

> source("http://bioconductor.org/biocLite.R") .

Bioconductor version 2.11 (BiocInstaller 1.8.3), ?biocLite for help

#This instructs R to use the Bioconductor website to install packages to use with Affymetrix .CEL files

> biocLite()

BioC_mirror: http://bioconductor.org

Using Bioconductor version 2.11 (BiocInstaller 1.8.3), R version 2.15.

Installing package(s) 'Biobase' 'IRanges' 'AnnotationDbi'

The downloaded binary packages are in

/var/folders/jy/jy3gx2IrGiCkg48a5WZgH++++TI/-Tmp-//Rtmp2DsOlq/downloaded_packages

Old packages: 'cluster', 'foreign', 'KernSmooth', 'lattice', 'Matrix', 'nnet', 'R.oo', 'rpart', 'survival', 'XML', 'xtable'

Update all/some/none? [a/s/n]:

> a

The downloaded binary packages are in

/var/folders/jy/jy3gx2IrGiCkg48a5WZgH++++TI/-Tmp-//Rtmp2DsOlq/downloaded_packages

#This downloads and updates all relevant packages

> biocLite("affy")

BioC_mirror: http://bioconductor.org

Using Bioconductor version 2.11 (BiocInstaller 1.8.3), R version 2.15.

Installing package(s) 'affy'

trying URL 'http://bioconductor.org/packages/2.11/bioc/bin/macosx/leopard/contrib/2.15/affy_1.36.1.tgz'

Content type 'application/x-gzip' length 1482408 bytes (1.4 Mb)

opened URL

```
==========================================
downloaded 1.4 Mb
The downloaded binary packages are in
        /var/folders/jy/jy3gx2IrGiCkg48a5WZgH++++TI/-Tmp-
//Rtmp2DsOlq/downloaded_packages
#This downloads the required package “affy”
> library(affy)
Loading required package: BiocGenerics
Loading required package: Biobase
Welcome to Bioconductor
#This loads the specified package “affy”
> norm <- exprs(justRMA())
#This reads the .CEL files present in the working directory and performs RMA
background correction of the expression intensity levels.
> dim(norm)
[1] 54675     4
#This shows that there are 54 675 probes with expression intensity values for 4 cell
lines
> head(norm)
           EC.CEL Seminoma.CEL Teratoma.CEL  YST.CEL
1007_s_at 10.516547    10.449916     9.873835 9.447998
1053_at    9.763713     8.540985     9.519655 9.788232
117_at     5.420071     5.416785     5.376868 5.389875
121_at     7.699719     7.634446     7.724840 7.731468
1255_g_at  4.900915     7.890857     6.144012 4.219233
1294_at    5.226387     5.435574     5.087242 5.353925
#The command ‘head()’ shows the first 6 lines of the table of data
This table shows the probe label and the expression values of each probe in each
cell line after RMA background correction. Expression intensities are given as log
(base 2) values.
> require(simpleaffy)
Loading required package: simpleaffy
Loading required package: genefilter
```

```
Loading required package: gcrma
No methods found in "Biobase" for requests: geneNames
#This downloads the relevant package “simpleaffy”
> raw <- ReadAffy()
#This reads the .CEL files again, without first background correcting them. We performed this step in order to determine which probes give values below the background level.
> call <- detection.p.val(raw)$call
#This determines whether each probe gives a value above or below a baseline intensity value, determined by control probes within the Affymetrix gene chip. A value above baseline is labelled “P” (present), and a value below baseline is labelled “A” (absent) - see table below.
> head(call)
        EC.CEL.present Seminoma.CEL.present Teratoma.CEL.present
YST.CEL.present
1007_s_at "P"           "P"                 "P"                 "P"
1053_at   "P"           "P"                 "P"                 "P"
117_at    "A"           "A"                 "A"                 "A"
121_at    "P"           "P"                 "P"                 "P"
1255_g_at "P"           "P"                 "P"                 "A"
1294_at   "A"           "A"                 "A"                 "A"
> calldet <- rowSums(call=="A")
#This counts, for each probe, the number of cell lines which show an intensity level below baseline (see table below)
> head(calldet)
1007_s_at   1053_at    117_at    121_at 1255_g_at   1294_at
 0                0                4         0        1                4
> sum(calldet == 4)
[1] 24733
#24 733 probes (out of 54 675) were labelled as “absent” in all four cell lines
> filt.probe <- norm[calldet < 4, ]
#This creates a new table of genes which were “present” in at least one cell line
```

```
> dim(filt.probe)
[1] 29942     4
#29 942 probes remained in our dataset after this filtering process
> head(filt.probe)
           EC.CEL Seminoma.CEL Teratoma.CEL  YST.CEL
1007_s_at 10.516547    10.449916     9.873835 9.447998
1053_at    9.763713     8.540985     9.519655 9.788232
121_at     7.699719     7.634446     7.724840 7.731468
1255_g_at  4.900915     7.890857     6.144012 4.219233
1316_at    5.336027     5.095768     5.298631 5.617900
1431_at    4.092673     4.500429     4.860048 4.206779
>  sEC <- filt.probe[, 2] - filt.probe[, 1]
> sTER <- filt.probe[, 2] - filt.probe[, 3]
> sYST <- filt.probe[, 2] - filt.probe[, 4]
> FC <- data.frame(sEC=sEC, sTER=sTER, sYST=sYST)
#The above four commands create a new table showing the fold change
differences in expression intensity values between the EC, Teratoma and YST cell
lines relative to the Seminoma cell line (see table below)
> head(FC)
                sEC       sTER       sYST
1007_s_at -0.06663077  0.57608158  1.00191860
1053_at   -1.22272832 -0.97867038 -1.24724738
121_at    -0.06527352 -0.09039394 -0.09702187
1255_g_at  2.98994200  1.74684446  3.67162365
1316_at   -0.24025831 -0.20286311 -0.52213145
1431_at    0.40775594 -0.35961902  0.29365011
> require("hgu133plus2.db")
Loading required package: hgu133plus2.db
Loading required package: org.Hs.eg.db
Loading required package: DBI
#This downloads the package required for probe annotation with gene information.
```

```
> ann <- merge(toTable(hgu133plus2SYMBOL),
merge(toTable(hgu133plus2ENTREZID), toTable(hgu133plus2GENENAME),
by=1), by=1)
#This creates a table of gene information for each probe, detailing the gene
symbol, Entrez ID and full gene name (see table below)
> head(ann)
  probe_id symbol gene_id                              gene_name
1 1007_s_at   DDR1     780 discoidin domain receptor tyrosine kinase 1
2   1053_at   RFC2    5982 replication factor C (activator 1) 2, 40kDa
3    117_at  HSPA6    3310        heat shock 70kDa protein 6 (HSP70B')
4    121_at   PAX8    7849                                paired box 8
5 1255_g_at GUCA1A    2978     guanylate cyclase activator 1A (retina)
6   1294_at   UBA7    7318 ubiquitin-like modifier activating enzyme 7
> all <- cbind(FC, filt.probe)
#This merges the tables detailing the intensity level expression value and the fold
changes for each probe (all given as log (base 2)) - see table below:
> head(all)
                sEC       sTER       sYST    EC.CEL Seminoma.CEL Teratoma.CEL
YST.CEL
1007_s_at -0.06663077  0.57608158  1.00191860 10.516547    10.449916
9.873835 9.447998
1053_at   -1.22272832 -0.97867038 -1.24724738  9.763713     8.540985
9.519655 9.788232
121_at    -0.06527352 -0.09039394 -0.09702187  7.699719     7.634446
7.724840 7.731468
1255_g_at  2.98994200  1.74684446  3.67162365  4.900915     7.890857
6.144012 4.219233
1316_at   -0.24025831 -0.20286311 -0.52213145  5.336027     5.095768
5.298631 5.617900
1431_at    0.40775594 -0.35961902  0.29365011  4.092673     4.500429
4.860048 4.206779
> all.ann <- merge(ann, all, by.x=1, by.y=0, sort=FALSE)
#This merges the gene annotations to the above table (see table below)
```

```
> head(all.ann)
  probe_id symbol gene_id                                    gene_name        sEC
sTER       sYST    EC.CEL Seminoma.CEL Teratoma.CEL  YST.CEL
1 1007_s_at   DDR1     780          discoidin domain receptor tyrosine kinase 1 -
0.06663077  0.57608158  1.00191860 10.516547    10.449916     9.873835
9.447998
2   1053_at   RFC2    5982          replication factor C (activator 1) 2, 40kDa -
1.22272832 -0.97867038 -1.24724738  9.763713     8.540985     9.519655
9.788232
3    121_at   PAX8    7849                                paired box 8 -0.06527352 -
0.09039394 -0.09702187  7.699719     7.634446     7.724840 7.731468
4 1255_g_at GUCA1A    2978             guanylate cyclase activator 1A (retina)
2.98994200  1.74684446  3.67162365  4.900915     7.890857     6.144012
4.219233
5   1316_at   THRA    7067                  thyroid hormone receptor, alpha -
0.24025831 -0.20286311 -0.52213145  5.336027     5.095768     5.298631
5.617900
6   1431_at CYP2E1    1571 cytochrome P450, family 2, subfamily E, polypeptide 1
0.40775594 -0.35961902  0.29365011  4.092673     4.500429     4.860048
4.206779
> write.table(all.ann, file="results.txt", sep="\t", quote=FALSE)
```

#This writes the above table as a .txt file, and saves it in the working directory. The file can then be opened in Excel to view the table and further analyse the data.

Appendix III. Sequencing for Myc-pcDNA3.1-PRDM14

BamH1
BstX 1
EcoR 1
START CODON

NNNNNNNNANNNNNNNNNTNNNCTTACCNTGGGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATG
ACTGGTGGACAGCAAATGGGTCGGGATCTGTACGACGATGACGATAAGGTACCAGGATCTCGACGGTATCGAT
TTAAAGCTATGGAGCAAAAGCTCATTTCTGAAGAGGACTTGAATGAAATGGAGCAAAAGCTCATTTCTGAAGA
GGACTTGAATGAAATGGAGCAAAAGCTCATTTCTGAAGAGGACTTGAATGAAATGGAGCAAAAGCTCATTTCT
GAAGAGGACTTGAATGAAATGGAGCAAAAGCTCATTTCTGAAGAGGACTTGAATGAAATGGAGAGCTTGGGCG
ACCTCACCATGGAGCAAAAGCTCATTTCTGAAGAGGACTTGGGATCCAGTGTGGTGGAATTCATGGCTCTACC
CCGGCCAAGTGAGGCCGTGCCTCAGGACAAGGTGTGCTACCCGCCGGAGAGCAGCCCGCAGAACCTGGCCGCG
TACTACACGCCTTTCCCGTCCTATGGACACTACAGAAACAGCCTGGCCACCGTGGAGGAAGACTTCCAACCTT
TCCGGCAGCTGGAGGCCGCAGCGTCTGCTGCCCCCGCCATGCCCCCCTTCCCCTTCCGGATGGCGCCTCCCTT
GCTGAGCCCGGGTCTGGGCCTACAGAGGGAGCCTCTCTACGATCTGCCCTGGTACAGCAAGCTGCCACCGTGG
TACCCAATTCCCCACGTCCCCAGGGAAGTGCCGCCCTTCCTGAGCAGCAGCCACGAGTACGCGGGTGCCAGCA
GTGAAGATCTGGGCCACCAAATCATTGGTGGCGACAACGAGAGTGGCCCGTGTTGTGGACCTGACACTTTAAT
TCCACCGCCCCCTGCGGATGCTTCTCTGTTACCTGAGGGGCTGAGGACCTCCCAGTTANTACCTTGCTCACCC
AGCAAGCAGTCAGAGGATGGTCCCAAACCCTCCAACCAAGAAGGGAAGTCCCCTGCTCGGTTCCAGTTCACGG
AGGAGGACCTGCACTTCGTTCNGTANGGGGTCACTCCNGCCTNNNCACCCAGCCAGCCTGCACCATGCGATTT
CNNGNTCNGGTNCCCCCAGACAGCTCTGGANNTGATNNTCTTCNCAAACTCTGGANAAAGACTCCCNTNANTT
CNNNNNCTNNNCCTCATGCAGANNNNNNNNNNNANTCCCNNNTTTTNGNNGNGTTCTGCAGNANTTTNNCNNN
NNNNNNNNNNNNNNNNNNTCNNNNAANNGNCATGNCAGNNANNNNACNNCGNANNNNATNNNNNNNANGNGGN
NATCTTNNAANANNGNNCNTTNNNNNNTTNNNNNNNNNNNGNGNTANGGAANNGANNNNNNNTNNNNANNNNN
NNNNNCNGNNCNANNNNNNNNNNNNNNGNNNNNNNNN

Blast

```
Query  1     ATGGCTCTACCCCGGCCAAGTGAGGCCGTGCCTCAGGACAAGGTGTGCTACCCGCCGGAG  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  428   ATGGCTCTACCCCGGCCAAGTGAGGCCGTGCCTCAGGACAAGGTGTGCTACCCGCCGGAG  487

Query  61    AGCAGCCCGCAGAACCTGGCCGCGTACTACACGCCTTTCCCGTCCTATGGACACTACAGA  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  488   AGCAGCCCGCAGAACCTGGCCGCGTACTACACGCCTTTCCCGTCCTATGGACACTACAGA  547

Query  121   AACAGCCTGGCCACCGTGGAGGAAGACTTCCAACCTTTCCGGCAGCTGGAGGCCGCAGCG  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  548   AACAGCCTGGCCACCGTGGAGGAAGACTTCCAACCTTTCCGGCAGCTGGAGGCCGCAGCG  607

Query  181   TCTGCTGCCCCCGCCATGCCCCCCTTCCCCTTCCGGATGGCGCCTCCCTTGCTGAGCCCG  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  608   TCTGCTGCCCCCGCCATGCCCCCCTTCCCCTTCCGGATGGCGCCTCCCTTGCTGAGCCCG  667

Query  241   GGTCTGGGCCTACAGAGGGAGCCTCTCTACGATCTGCCCTGGTACAGCAAGCTGCCACCG  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  668   GGTCTGGGCCTACAGAGGGAGCCTCTCTACGATCTGCCCTGGTACAGCAAGCTGCCACCG  727

Query  301   TGGTACCCAATTCCCCACGTCCCCAGGGAAGTGCCGCCCTTCCTGAGCAGCAGCCACGAG  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  728   TGGTACCCAATTCCCCACGTCCCCAGGGAAGTGCCGCCCTTCCTGAGCAGCAGCCACGAG  787

Query  361   TACGCGGGTGCCAGCAGTGAAGATCTGGGCCACCAAATCATTGGTGGCGACAACGAGAGT  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  788   TACGCGGGTGCCAGCAGTGAAGATCTGGGCCACCAAATCATTGGTGGCGACAACGAGAGT  847
```

```
Query  421   GGCCCGTGTTGTGGACCTGACACTTTAATTCCACCGCCCCCTGCGGATGCTTCTCTGTTA  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  848   GGCCCGTGTTGTGGACCTGACACTTTAATTCCACCGCCCCCTGCGGATGCTTCTCTGTTA  907

Query  481   CCTGAGGGGCTGAGGACCTCCCAGTTATTACCTTGCTCACCCAGCAAGCAGTCAGAGGAT  540
             |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Sbjct  908   CCTGAGGGGCTGAGGACCTCCCAGTTANTACCTTGCTCACCCAGCAAGCAGTCAGAGGAT  967

Query  541   GGTCCCAAACCCTCCAACCAAGAAGGGAAGTCCCCTGCTCGGTTCCAGTTCACGGAGGAG  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  968   GGTCCCAAACCCTCCAACCAAGAAGGGAAGTCCCCTGCTCGGTTCCAGTTCACGGAGGAG  1027

Query  601   GACCTGCACTTCGTTCTGTACGGGGTCACTCCCAGCCTGGAGCACCCAGCCAGCCTGCAC  660
             |||||||||||||||| ||| ||||||||||  ||||    ||||||||||||||||||
Sbjct  1028  GACCTGCACTTCGTTCNGTANGGGGTCACTCCN-GCCTNNN-CACCCAGCCAGCCTGCAC  1085

Query  661   CATGCGATTTCAGGCCTCCTGGTCCCCCCAGACAGCTCTGGATCTGATTCTCTTCCTCAA  720
             ||||||||||  |  ||  ||| ||||||||||||||||  ||||  |||||  |||
Sbjct  1086  CATGCGATTTCNNGN-TCN-GGTNCCCCCAGACAGCTCTGGANNTGATNNTCTTCN-CAA  1142

Query  721   ACTCTGGATAAAGACTCCC  739
             |||||||| ||||||||||
Sbjct  1143  ACTCTGGANAAAGACTCCC  1161
```

5'3' Frame 1

```
atggctctaccccggccaagtgaggccgtgcctcaggacaaggtgtgctacccgccggag
 M  A  L  P  R  P  S  E  A  V  P  Q  D  K  V  C  Y  P  P  E
agcagcccgcagaacctggccgcgtactacacgcctttcccgtcctatggacactacaga
 S  S  P  Q  N  L  A  A  Y  Y  T  P  F  P  S  Y  G  H  Y  R
aacagcctggccaccgtggaggaagacttccaacctttccggcagctggaggccgcagcg
 N  S  L  A  T  V  E  E  D  F  Q  P  F  R  Q  L  E  A  A  A
tctgctgcccccgccatgccccccttccccttccggatggcgcctcccttgctgagcccg
 S  A  A  P  A  M  P  P  F  P  F  R  M  A  P  P  L  L  S  P
ggtctgggcctacagagggagcctctctacgatctgccctggtacagcaagctgccaccg
 G  L  G  L  Q  R  E  P  L  Y  D  L  P  W  Y  S  K  L  P  P
tggtacccaattccccacgtccccagggaagtgccgcccttcctgagcagcagccacgag
 W  Y  P  I  P  H  V  P  R  E  V  P  P  F  L  S  S  S  H  E
tacgcgggtgccagcagtgaagatctgggccaccaaatcattggtggcgacaacgagagt
 Y  A  G  A  S  S  E  D  L  G  H  Q  I  I  G  G  D  N  E  S
ggcccgtgttgtggacctgacactttaattccaccgcccccctgcggatgcttctctgtta
 G  P  C  C  G  P  D  T  L  I  P  P  P  P  A  D  A  S  L  L
cctgaggggctgaggacctcccagttattaccttgctcacccagcaagcagtcagaggat
 P  E  G  L  R  T  S  Q  L  L  P  C  S  P  S  K  Q  S  E  D
ggtcccaaaccctccaaccaagaagggaagtcccctgctcggttccagttcacggaggag
 G  P  K  P  S  N  Q  E  G  K  S  P  A  R  F  Q  F  T  E  E
gacctgcacttcgttctgtacggggtcactcccagcctggagcacccagccagcctgcac
 D  L  H  F  V  L  Y  G  V  T  P  S  L  E  H  P  A  S  L  H
catgcgatttcaggcctcctggtccccccagacagctctggatctgattctcttcctcaa
 H  A  I  S  G  L  L  V  P  P  D  S  S  G  S  D  S  L  P  Q
actctggataaagactcccttcaacttccagaaggtctatgcctcatgcagacggtgttt
 T  L  D  K  D  S  L  Q  L  P  E  G  L  C  L  M  Q  T  V  F
ggtgaagtcccacattttggtgtgttctgcagtagttttatcgccaaaggagtcaggttt
 G  E  V  P  H  F  G  V  F  C  S  S  F  I  A  K  G  V  R  F
gggccctttcaaggtaaagtggtcaatgccagtgaagtgaagacctacggagacaattct
 G  P  F  Q  G  K  V  V  N  A  S  E  V  K  T  Y  G  D  N  S
gtgatgtgggagatctttgaagatggtcatttgagccactttatagatggaaaaggaggt
```

```
 V  M  W  E  I  F  E  D  G  H  L  S  H  F  I  D  G  K  G  G
acggggaactggatgtcctatgtcaactgtgcccgcttccccaaggagcagaacctagtt
 T  G  N  W  M  S  Y  V  N  C  A  R  F  P  K  E  Q  N  L  V
gctgtgcagtgtcaagggcatatattttatgagagctgcaaagagatccatcagaaccaa
 A  V  Q  C  Q  G  H  I  F  Y  E  S  C  K  E  I  H  Q  N  Q
gagctccttgtgtggtatggagactgctatgagaaatttctggatattcctgtgagcctt
 E  L  L  V  W  Y  G  D  C  Y  E  K  F  L  D  I  P  V  S  L
caggtcacagagccggggaagcagccatctgggccctctgaagagtctgcagaaggctac
 Q  V  T  E  P  G  K  Q  P  S  G  P  S  E  E  S  A  E  G  Y
agatgtgaaagatgtgggaaggtatttacctacaaatattacagagataagcacctcaag
 R  C  E  R  C  G  K  V  F  T  Y  K  Y  Y  R  D  K  H  L  K
tacacccccctgtgtggacaagggcgataggaaatttccctgttctctctgcaaacgatcc
 Y  T  P  C  V  D  K  G  D  R  K  F  P  C  S  L  C  K  R  S
tttgagaagcgggaccggcttcggatccacattcttcatgttcatgagaagcaccggcct
 F  E  K  R  D  R  L  R  I  H  I  L  H  V  H  E  K  H  R  P
cacaagtgttctacatgtgggaaatgtttctctcaatcttccagcctaaacaaacacatg
 H  K  C  S  T  C  G  K  C  F  S  Q  S  S  S  L  N  K  H  M
cgagtccactctggagacagaccataccagtgtgtgtattgtactaagaggttcacagcc
 R  V  H  S  G  D  R  P  Y  Q  C  V  Y  C  T  K  R  F  T  A
tccagcatactccgcacacacatcaggcagcactccggggagaagcccttcaaatgcaag
 S  S  I  L  R  T  H  I  R  Q  H  S  G  E  K  P  F  K  C  K
tactgtggtaaatcttttgcatcccatgctgcccatgacagccatgtccggcgttcacac
 Y  C  G  K  S  F  A  S  H  A  A  H  D  S  H  V  R  R  S  H
aaggaggatgatggctgctcatgcagcatctgtgggaaaatcttctcagatcaagaaaca
 K  E  D  D  G  C  S  C  S  I  C  G  K  I  F  S  D  Q  E  T
ttctactcccacatgaagtttcatgaagactactagcc
 F  Y  S  H  M  K  F  H  E  D  Y  -
```

A-G Q-Q

T-C D-D

origin	mutation	Type of mutation
Caa- Gln	cag - Gln	silent
Gat- Asp	gac - Asp	silent