

**Aberrant Hypermethylation of Xq12 in Lymphoma**

by

**DÉSIRÉE QUINT**

**B.Sc., University of British Columbia, 2004**

**A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE  
REQUIREMENTS FOR THE DEGREE OF**

**MASTER OF SCIENCE**

in

**THE FACULTY OF GRADUATE STUDIES**

**(MEDICAL GENETICS)**

**THE UNIVERSITY OF BRITISH COLUMBIA**

**August 2006**

**© Désirée Quint, 2006**

## **Abstract**

It has been well established that abnormal CpG island promoter methylation in cancer is associated with the epigenetic silencing of tumour suppressor genes, genes that when inactivated predispose a cell to malignancy. Such aberrant methylation was seen for the androgen receptor (AR) located on the X chromosome in 84% of follicular lymphoma patient samples. Methylation has been shown to spread several megabases from its origin and this has led us to hypothesize that a candidate tumour suppressor gene is located in the vicinity of the AR. AR and the nearby oligophrenin (OPHN1) gene are both considered to be poor tumour suppressor gene candidates since deletions of these genes exist and they do not result in lymphoma. In addition, OPHN1 is mainly expressed in brain tissues. In order to identify candidate genes, the extent of the abnormal CpG island methylation within the region surrounding the AR was first established in two different lymphoma cell lines (SUDHL3, DoHH2). A region of approximately 8 Mb was examined which led to the focus on a 1.5 Mb highly methylated sub-region located downstream of the AR. The methylation status of CpG islands within this sub- region was examined in three additional lymphoma cell lines (HBL-2, JVM-2 and Z138) as well as in patient samples. STARD8, a GTPase-activating protein, is located within this region and is abnormally silenced in at least two of the lymphoma cell lines making it a possible candidate.

## **Table of Contents**

<b>Abstract</b>	ii
<b>List of Tables</b>	vi
<b>List of Figures</b>	vii
<b>List of Abbreviations</b>	ix
<b>Acknowledgements</b>	x
<b>Chapter 1: Introduction to Methylation and Lymphoma</b>	1
1.1) Thesis Background	1
1.1.1) Androgen Receptor and Abnormal Methylation in Lymphoma	1
1.1.2) Thesis Goal	2
1.2) Epigenetic Modifications	3
1.3) DNA Methylation	4
1.3.1) Cytosine and CpG Islands	4
1.3.2) DNA Methyltransferases	6
1.3.3) DNA Methylation Pattern Establishment in Early Development	7
1.3.4) Link Between Methylation and Chromatin Modifications	9
1.3.4.1) Methylated DNA Binding Protein Complexes	10
1.3.4.2) Inactive Chromatin	11
1.3.4.2.1) Histone Modifications Directing DNA Methylation	11
1.3.4.2.2) DNA Methylation Directing Histone Modifications	11
1.3.5) Methylation and Repetitive Elements	12
1.4) X Chromosome Inactivation	13
1.4.1) Mouse/Human Hybrid Cell Lines	15
1.5) DNA Methylation and Cancer	15
1.5.1) Tumour Suppressor Genes Silenced by Methylation	16
1.5.2) Potential Sources of Abnormal Methylation	17
1.5.2.1) Seeds of Methylation	17
1.5.2.2) Aging	18
1.5.2.3) Chronic Inflammation	19
1.5.2.4) Viral Infection	19
1.5.3) CpG Island Methylator Phenotype (CIMP)	20
1.5.4) Gene Inactivation and Methylation	21
1.5.5) Models of Gene Silencing Associated with Abnormal Methylation	23
1.5.6) Usefulness of Abnormal Methylation as Markers	24
1.5.7) Methylation as a Target of Cancer Therapy	26
1.6) Lymphoma	27
1.6.1) Non-Hodgkin's Lymphoma	28
1.6.2) Lymphoma and the X Chromosome	30

<b>Chapter 2: Materials and Methods</b>	33
2.1) Tissue Culture	33
2.1.1) Thawing Cell Lines	33
2.1.2) Freezing Down Cell Lines	33
2.2) Methylation of CpG Islands	35
2.2.1) DNA Extraction	35
2.2.2) Double Restriction Enzyme Digestion	37
2.2.2.1) Double Restriction Enzyme Digestion of Cell Line DNA	37
2.2.2.2) Double Restriction Enzyme Digestion of Patient Sample DNA	38
2.2.3) Primer Design	39
2.2.3.1) Primer Table	39
2.2.4) Polymerase Chain Reaction (PCR)	40
2.2.5) Gel Electrophoresis	40
2.2.5.1) Agarose Gels	47
2.2.5.2) Polyacrylamide Gels	47
2.3) Gene Expression	47
2.3.1) RNA Extraction	48
2.3.1.1) Acid-Guanidinium-Phenol-Chloroform RNA Extraction	48
2.3.1.2) QIAGEN's RNeasy Mini Kit	48
2.3.2) DNase Treatment of RNA	49
2.3.3) Reverse Transcriptase PCR (RT-PCR)	49
2.3.4) Primer Design	50
2.3.4.1) Expression Primer Table	50
2.4) Quantifying Gene Expression	53
2.4.1) Quantitative PCR (qPCR)	53
2.4.2) Gel Extraction	54
<b>Chapter 3: CpG Island Methylation Pattern and Gene Expression on the X Chromosome in the Vicinity of Xq11.2-q12 in Lymphoma</b>	55
3.1) Introduction	55
3.2) Results – A.) Methylation	55
3.2.1) CpG Islands and Associated Genes	55
3.2.2) Methylation in an 8 Mb Region Around Xq11.2-q12	63
3.2.2.1) Methylation in Two Lymphoma Cell Lines	63
3.2.2.1.1) Gene Promoter Methylation	77
3.2.2.2) Methylation in Additional Control Lymphoblast and Mouse/Human Hybrid Cell Lines	77
3.2.3) Methylation in the 1.4 Mb Sub-Region in Three Additional Lymphoma Cell Lines	89
3.2.4) Methylation in the 4 Mb Sub-Region in Lymphoma Patient Samples	91
3.3) Results – B.) Gene Expression	103
3.3.1) Genes in the Region	103
3.3.2) Expression of Genes in Lymphoblast and Lymphoma Cell Lines	109
3.3.3) Quantitation of STARD8 Gene Expression	117

3.4) Discussion of DNA Methylation and Gene Expression Results	119
3.4.1) Methylation in Lymphoblast and Mouse/Human Hybrid Cell Lines	119
3.4.1.1) Assay Variability	119
3.4.1.2) Interindividual Variability	120
3.4.1.3) Methylation of the Xa	120
3.4.1.4) Hypomethylation of the Xi	121
3.4.2) Methylation in the 1.4 Mb Sub-Region in Three Additional Cell Lines	121
3.4.3) Methylation in the 1.4 Mb Sub-Region in Lymphoma Patient Samples	121
3.4.4) Summary of DNA Methylation and Gene Expression Analysis	122
<b>Chapter 4: Discussion</b>	128
<b>References</b>	134

## List of Tables

<b>Table 2.1</b>	Table of cell lines and their tissue culture conditions	34
<b>Table 2.2</b>	Table of primers used to check methylation status	41
<b>Table 2.3</b>	Table of primers used to check gene expression	51
<b>Table 3.1</b>	CpG islands and their associated genes in an 8 Mb region on the X chromosome located in the vicinity of AR (Xq11.2-12)	56
<b>Table 3.2</b>	Location of CpG islands in an 8Mb region on Xq11.2-12 relative to genes	59
<b>Table 3.3</b>	CpG islands located in the vicinity of AR on the X chromosome and their associated repetitive elements	61
<b>Table 3.4</b>	Methylation status of CpG islands located in an 8 Mb region in the vicinity of AR (Xq11.2-12) in two control and two male lymphoma cell lines	65
<b>Table 3.5</b>	CpG islands with unexpected methylation status in female and male control cell lines	76
<b>Table 3.6</b>	Summary of the methylation results of selected CpG islands located in the 8 Mb region in the vicinity of AR in additional female and male control cell lines and Xa and Xi mouse/human hybrid cell lines	81
<b>Table 3.7</b>	Methylation assay results in the sub-region and for AR in three mantle cell lymphoma cell lines, Z138, HBL-2 and JVM-2	92
<b>Table 3.8</b>	Summary of the CpG islands methylation results in the 1.4 Mb sub-region for all patient samples	98
<b>Table 3.9</b>	Summary of the methylation status of AR and all the CpG islands assessed in the MCL and DBCL patient DNA	102
<b>Table 3.10</b>	Genes located in the 8 Mb region in Xq11.2-q12	105
<b>Table 3.11</b>	Genes and their associated CpG islands located in the 8 Mb region located at Xq11.2-q12	110
<b>Table 3.12</b>	Summary of the gene expression results for 10 selected genes located in the 8 Mb region at Xq11.2-12	113
<b>Table 3.13</b>	Percentage of lymphoma patient samples and lymphoma cell lines showing methylation of the CpG islands found in the 2 sub-regions of interest and the AR	123
<b>Table 3.14</b>	Summary of the gene expression status of 10 genes and the methylation status of their associated CpG islands in lymphoblast and lymphoma cell lines	125

## List of Figures

<b>Figure 2.1</b> Description of the methylation assay	36
<b>Figure 3.1</b> Distribution of CpG island gene association	60
<b>Figure 3.2</b> CpG island methylation results for an 8 Mb region located at Xq11.2-12	67
<b>Figure 3.3</b> Examples of CpG islands that are abnormally methylated in lymphoma cell lines, SUDHL3 and DoHH2	69
<b>Figure 3.4</b> Five examples of CpG islands that are not abnormally methylated in lymphoma cell lines in the 8 Mb region examined on the X chromosome	70
<b>Figure 3.5</b> CpG islands showing variation in methylation status between the two follicular lymphoma cell lines, SUDHL3 and DoHH2	71
<b>Figure 3.6</b> CpG islands showing a lack of methylation in the female control cell lines	73
<b>Figure 3.7</b> Five examples of CpG islands located within the examined 8 Mb region on the X chromosome showing partial methylation	74
<b>Figure 3.8</b> Examples of CpG islands located in the 8 Mb region in the vicinity of AR on the X chromosome showing partial methylation	75
<b>Figure 3.9</b> Methylation of the promoter and first exon region of four genes located in a CpG islands sparse area upstream of AR	78
<b>Figure 3.10</b> Selected CpG islands whose methylation status was assessed in additional female and male control cell lines as well as Xa and Xi mouse/human hybrid cell lines	80
<b>Figure 3.11</b> Methylation assay in female and male control and Xa/Xi retaining mouse/human hybrid cell lines of CpG islands previously showing methylation in a male control cell line	83
<b>Figure 3.12</b> Methylation assay in female and male control and Xa/Xi retaining mouse/human hybrid cell lines of CpG islands that previously showed a lack of methylation in a female control cell line	85
<b>Figure 3.13</b> Methylation assay in female and male control and Xa/Xi retaining mouse/human hybrid cell lines of CpG islands that previously showed variation in methylation in lymphoma cell lines	87
<b>Figure 3.14</b> Methylation assay in female and male control and Xa/Xi retaining mouse/human hybrid cell lines of 5 selected CpG islands	88
<b>Figure 3.15</b> Map of the 1.4 Mb sub-region of interest	90
<b>Figure 3.16</b> Methylation results of AR and the 1.4 Mb sub-region in the three additional lymphoma cell lines, Z138, HBL-2 and JVM-2	93
<b>Figure 3.17</b> Methylation results for AR and CpG islands located in the 1.4 Mb sub-region in the first group of patient samples	95
<b>Figure 3.18</b> Methylation results for AR and CpG islands located in the 1.4 Mb sub-region in the second group of patient samples	97
<b>Figure 3.19</b> Methylation status of CpG islands located upstream of AR in patient samples	101
<b>Figure 3.20</b> Map of second sub-region of interest	104

<b>Figure 3.21</b> Expression status of 10 different gene located in the region around the AR on the X chromosome	111
<b>Figure 3.22</b> Quantitation of STARD8 expression in lymphoblast and lymphoma cell lines and lymphoma patients	118
<b>Figure 3.23</b> CpG island methylation and gene expression results for an 8 Mb region located at Xq11.2-12	124
<b>Figure 4.1</b> Distribution of ‘normally’ methylated CpG islands relative to the location of the two sub-regions of interest	131

## **List of Abbreviations**

- 5-azaCdR: 5-aza-2'-deoxycytidine  
5-aza-CR: 5-azacytidine  
AIS: Androgen Insensitivity Syndrome  
CIMP: CpG Island Methylator Phenotype  
CIN: Chromosomal Instability  
DEPC: Diethyl Pyrocarbonate  
DLBCL: Diffuse Large B-Cell Lymphoma  
DMSO: Dimethyl Sulfoxide  
Dnmt/DNMT: DNA Methyltransferases  
dpc: days post-coitum  
DTT: Dithiothreitol  
E (*i.e.* E11.5): Embryonic Day 11.5  
EBV: Epstein-Barr Virus  
FADD: Fas-Associated Death Domain  
FCS: Fetal Calf Serum  
FL: Follicular Lymphoma  
HCC: Hepatocellular Carcinoma  
HCG: High CpG Content  
HDAC: Histone Deacetylases  
HMT: Histone Methyltransferases  
HNPCC: Hereditary Non-Polyposis Colorectal Cancer  
HP1: Heterochromatin Protein 1  
IAP: Intracisternal A-Particle  
IDT: Integrated DNA Technologies  
ICM: Inner Cell Mass  
K: Lysine  
LCG: Low CpG Content  
LINE: Long Interspersed Nuclear Elements  
LOH: Loss of Heterozygosity  
LRES: Long-Range Epigenetic Silencing  
Mb: Megabase  
MBD: Methyl-CpG-Binding Domain  
MDS: Myelodysplastic Syndrome  
MCL: Mantle Cell Lymphoma  
MIS: Microsatellite Instability  
M-MLV: Moloney Murine Leukemia Virus  
NCBI: National Centre for Biotechnology Information  
NHL: Non-Hodgkin's Lymphoma  
NK: Natural Killer Cell  
OPHN1: Oligophrenin 1  
PCR: Polymerase Chain Reaction  
PGC: Primordial Germ Cells  
PP: Protein Phosphatase  
qPCR: Quantitative Polymerase Chain Reaction

RhoGAP: Rho-like GTPase Activating Protein  
RNAi: RNA Interference  
RNasin: RNAase Inhibitor  
RT: Reverse Transcriptase  
SBMA: Kennedy Spinal and Bulbar Muscular Atrophy  
SDS: Sodium Dodecyl Sulfate  
SINE: Short Interspersed Nuclear Elements  
siRNA: small interfering RNA  
TCC: Transitional Cell Carcinoma  
TE: Trophoectoderm  
TGF- $\beta$ : Transforming Growth Factor-Beta  
TRD: Transcriptional Repression Domain  
TSA: Trichostatin A  
UCSC: University of California Santa Cruz  
Xa: Active X Chromosome  
Xi: Inactive X Chromosome

## Acknowledgements

I would like to thank all members of the Brown Lab for their support, help and a great experience. I would especially like to thank Dr. Carolyn Brown for her guidance and the opportunity to do research in her lab.

I also appreciate the time and feedback my advisory committee, Dr. Dixie Mager and Dr. Angela Brooks-Wilson, put into this thesis.

I would also like to acknowledge Helen McDonald for the initial studies; Catherine Tucker and Richard Klasa for the mantle cell lymphoma cell lines; Joe Connors, Randy Gascoyne, Doug Horsman and Ron deLeeuw for patient samples; and all members of the UBC Centre for Lymphoid Cancer for discussions.

## **Introduction to Methylation and Lymphoma**

### **1.1) Thesis Background**

#### ***1.1.1) Androgen Receptor and Abnormal Methylation in Lymphoma***

Numerical chromosomal abnormalities are common in cancers, reflecting both the consequence of chromosomal instability (CIN) experienced by malignant cells and selective pressure (reviewed in [1, 2]). The frequent occurrence of a specific aneuploidy in a cancer suggests that the cell is under selective pressure and a gene(s) potentially involved in tumorigenesis might be located on the chromosome in question. The acquisition of an extra X chromosome is one such frequently observed aneuploidy in lymphomas, occurring in 7-33% cases [3]. Individuals born with an extra copy of an X are however not known to be predisposed to developing lymphoma. In addition, an increase in X aneuploidy has been observed with aging (reviewed in [4]). This correlation might be caused by an age related increase in non-disjunction and/or a reduction in cell cycle control protecting against aneuploidy due to limited gene expression from the inactive X chromosome. In mammalian cells, all but one X chromosome is transcriptionally inactivated for dosage compensation between males and females (reviewed in [5, 6]) and in both constitutional and acquired X chromosome aneuploidy, the inactivated X is most commonly observed to be involved (reviewed in [4]). Therefore, to determine whether the extra copies of X-linked genes confer a selective growth advantage to the malignant cells, or are a consequence of premature aging, the inactive status of the acquired copy of the X chromosome in lymphomas needs to be determined. In 2000, McDonald *et al.* undertook this task in both female and male lymphoma patient samples and the results are discussed below in section 1.6.2 [4].

Two features of the inactive X chromosome are hypermethylation and transcriptional silencing (described in more detail below and reviewed in [5]). McDonald *et al.* (2000) used multiple assays based on either the expression of polymorphisms or the methylation of CpG dinucleotides located near polymorphisms on the X chromosome to determine the activity state of the acquired X chromosome. The expression and methylation status provide information on the activity of the X chromosome and the polymorphism distinguishes the two X chromosomes in females. One of these assays

looked at the methylation state of the (CAG)<sub>n</sub> repeat polymorphism in the androgen receptor (AR) located at Xq11.2-q12. McDonald *et al.* (2000) observed abnormal methylation in 84% of the lymphoma patients at this locus. Nine out of the ten informative female patients and seven out of the nine male patients showed hypermethylation at AR. Abnormal hypermethylation is associated with the silencing of tumour suppressor genes in malignancies (reviewed in [7-11]) and this led to the hypothesis of a candidate tumour suppressor gene located on the X chromosome in the vicinity of the AR locus.

### **1.1.2) Thesis Goal**

The work in this thesis was done to follow up on the McDonald *et al.* (2000) results and investigate the hypothesis of a potential candidate tumour suppressor gene located on the X chromosome in the vicinity of AR. AR codes for a steroid-hormone activated transcription factor which upon binding androgen hormone ligands, like progesterone, glucocorticoids and mineralocorticoids, stimulates the transcription of androgen responsive genes [12]. AR itself is not a good candidate lymphoma tumour suppressor gene for several reasons. It is normally expressed in fibroblasts and immature hematopoietic lineages (reviewed in [4]), but not in lymphocytes, and patients suffering from androgen insensitivity syndrome (AIS) as a result of AR mutations are not known to be predisposed to lymphoma. Abnormalities in AR have been observed in some cancers. For example, AR germline mutations were identified in male breast cancer [13] and AR hypermethylation [14] and a correlation between the length of the polymorphic repeat and the risk of developing aggressive cancer were seen in prostate cancer [15]. However, such abnormalities have not been observed in lymphomas.

Methylation can spread several Mb from its origin [9], therefore, a candidate tumour suppressor gene could be located near AR. In this situation, the aberrant methylation of AR is hypothesized to be a consequence of the selection for silencing of a nearby tumour suppressor gene. Oligophrenin 1 (OPHN1), located immediately downstream of AR, is also a poor candidate tumour suppressor gene since it is predominantly expressed in brain tissues and mutations, some of which also involve the AR, result in mental retardation, but not in a predisposition to lymphoma [16].

Loss of heterozygosity (LOH) of this region has been observed in 25% of cervical cancers and 30% of ovarian tumours (reviewed in [4]), strengthening the argument for the existence of a candidate tumour suppressor gene in the vicinity of AR. The ideal candidate is one that is abnormally methylated and silenced or expressed at reduced levels in lymphoma compared to lymphoblast cells. In addition, an ideal candidate would likely code for a protein involved in cell cycle control or apoptosis regulation like other classic tumour suppressors (reviewed in [17]).

The extent of abnormal methylation spreading was examined in two male lymphoma cell lines by assessing the methylation status of CpG islands. This led to the identification of two smaller, highly hypermethylated regions of interest. The methylation patterns of these smaller sub-regions were then determined in additional male lymphoma cell lines as well as patient samples. Expression of some genes of interest within the region around the AR was also examined to check for potential changes in gene expression between control lymphoblast cell lines and lymphoma cell lines and patient samples.

The methylation status of almost all CpG islands in an 8 Mb region on the X chromosome was examined in lymphoma cell lines as well as in female and male control lymphoblast cell lines to determine the extent of abnormal methylation. The examination of the female and male control cell lines allowed for inferences to be made on the differences in CpG island methylation patterns between the inactive and active X chromosome with the assumption that any methylation observed in the male comes from the active X chromosome and that any methylation in the female controls is representative of the inactive X chromosome.

## **1.2) Epigenetic Modifications**

Genetic mutations like point mutations, duplications and deletions have been well studied, but recently attention has shifted to alterations that do not involve changes to the genetic code itself but still influence gene expression levels. These heritable modifications of cellular information are termed epigenetic and include DNA methylation and histone modifications like methylation, acetylation and phosphorylation (reviewed in [18]). These epigenetic modifications are involved in genomic imprinting (involving

parent of origin specific allele expression) and X inactivation (reviewed in [18]). These modifications all affect chromatin organization, marking active and inactive chromatin. Histone modifications are very well conserved between all eukaryotic organisms, whereas DNA methylation is more commonly seen in higher eukaryotic organisms with more complex genomes (reviewed in [19]). Epigenetic modifications play an important role in tissue-specific gene expression regulation, global gene silencing, and genome reprogramming during embryogenesis and gametogenesis. Changes in normal epigenetic marks have been associated with pathophysiological conditions.

### **1.3) DNA Methylation**

DNA methylation functions in gene silencing, chromatin remodeling and genomic stability. In mammals, it plays an important role in normal development (see section 1.3.3 DNA Methylation Pattern Establishment in Early Development) and is involved in silencing expression of repetitive elements, imprinted genes and genes on the inactive X chromosome. The evolutionary origin of methylation is unknown, but a popular hypothesis involves the development of methylation as a defense mechanism against the expression of endogenous retroviruses, like transposable elements, that could potentially integrate into and disrupt genes (reviewed in [20]).

#### ***1.3.1) Cytosine and CpG Islands***

DNA methylation principally involves the covalent addition of a methyl group from S-adenosylmethionine to the carbon 5' of cytosines located 5' to guanosines in CpG dinucleotides. The methylated cytosine is often referred to as DNA's fifth base and accounts for 0.75 – 1 % of all the bases, and 70 % of all the CpG dinucleotides in the genome are methylated (reviewed in [7]).

CpG dinucleotides are overall underrepresented in the vertebrate genome relative to the GC content. This is a consequence of deamination of methylated cytosines. When unmethylated cytosines become deaminated, the base becomes uracil. The DNA repair machinery easily recognizes this as an error since uracil is not one of the bases found in DNA. However, when methylated cytosines become deaminated, they become thymines. Thymine is one of the bases found in DNA, and thus the DNA repair machinery has

difficulty recognizing this as an error. The overall result is a depletion of methylated CpGs in the germline [21].

CpG dinucleotides are found in the expected or greater than expected number in regions called CpG islands which span 0.5-4 kb. The majority of these islands are located in the promoter region of almost half of all transcribed genes (reviewed in [22]) and estimations suggest the human genome has approximately 29,000 CpG islands (reviewed in [7]). The association between promoter and CpG islands has led to the use of CpG islands to predict promoters [23]. Promoters with high CpG content are presumed to be expressed in almost all tissues, and in general, these islands are associated with 'housekeeping' genes [24]. The majority of the CpG islands remain unmethylated (this is why they still exist) with the exception of CpG islands associated with the inactive X chromosome in females and imprinted genes (reviewed in [22]).

There is however no definitive definition of CpG islands and most depend on *ad hoc* thresholds of length, CpG fraction and GC content. Recently, Saxonov *et al.* (2006) looked at the distribution of CpG dinucleotides with respect to promoters in order to overcome the problem of studying the relationship between promoters and CpG islands without a standard definition of CpG islands and what constitutes CpG island – promoter association. They identified two classes of promoters based on CpG content. 72% of promoters belonged to the high CpG content (HCG) class and 28% to the low CpG content (LCG) class. The CpG content of the LCG class of promoters was approximately the same as the rest of the human genome. The authors also observed a peak in CpG content approximately 15 bp upstream of the transcription start site in the HCG class promoters and they believe this represents the presence of CpG islands. The genes associated with HCG promoters had gene ontology terms associated with housekeeping functions, whereas the LCG promoter associated genes had terms associated with more specific functions. This agrees with the previous findings that CpG islands are generally associated with housekeeping genes. In addition the genes associated with HCG promoters are expressed in all or almost all tissues whereas the LCG associated promoter genes are only expressed in a few tissues.

Previously it was believed that approximately 60% of the genes in the human genome are associated with CpG islands, based on the definitions of *ad hoc* thresholds.

Saxonov *et al.* (2006) results suggest that this number is much higher; however, the results are potentially biased since the promoters they investigated might have been identified by their association with a CpG island.

### 1.3.2) DNA Methyltransferases

DNA methyltransferases (DNMT) are families of trans-acting enzymes responsible for the transfer of methyl groups from S-adenosylmethionine to the cytosine base of CpG dinucleotides (reviewed in [25]). There are two types of Dnmt activity, *de novo* methylation and maintenance methylation. *De novo* methylation refers to the transfer of methyl groups to cytosines that were previously unmethylated, whereas maintenance methylation refers to the addition of methyl groups to hemimethylated DNA following replication (reviewed in [25]). There are three different families of Dnmts, Dnmt1, Dnmt2 and Dnmt3, which are distantly related, diverging before the plant and animal kingdom separated (reviewed in [26]). However, only the Dnmt1 and Dnmt3 families have been shown to have methyltransferase activity and both have maintenance and *de novo* methylation capabilities (reviewed in [25]). The only known sequence specificity, which is shared between all the families, is CpG dinucleotides (reviewed in [27]).

The most prominent mammalian DNMT is DNMT1, homologs of which can be found in almost all eukaryotes that methylate cytosines [26]. This enzyme has different affinities for maintenance and *de novo* methylation activity, the latter being 1-2 orders of magnitude lower than the former [25]. The disruption of Dnmt1 in mice results in global demethylation, bi-allelic expression of imprinted genes and embryonic lethality [28, 29]. The ability to methylate newly integrated retroviral DNA is however not affected (reviewed in [26]). The Dnmt1 protein localizes to the replication foci, which indicates maintenance methylation and DNA replication are linked [30]. Together, these pieces of evidence support the idea that Dnmt1 is the major Dnmt responsible for maintaining the methylation pattern of proliferating cells in a replication-dependent manner [31].

The Dnmt3 family consists of Dnmt3a and Dnmt3b. These enzymes have approximately equal affinities for maintenance and *de novo* methylation activity [27]. Dnmt3a and Dnmt3b are highly expressed in undifferentiated mouse embryonic stem

(ES) cells and are poorly expressed in differentiated adult somatic cells. When these enzymes are inactivated, *de novo* methylation is blocked in ES cells and early postimplantation embryos but maintenance methylation or methylation of imprinted genes is not altered [31]. The majority of the *de novo* methylation takes place early in embryonic development and the above suggests both Dnmt3a and Dnmt3b are required for *de novo* methylation and mouse development [31]. In addition to its role in development, Dnmt3b also appears to have a function separate from Dnmt3a. Mutations in human DNMT3b result in ICF syndrome, a rare autosomal recessive disorder characterized by immunodeficiency and facial anomalies. Another feature of ICF is hypomethylated pericentromeric satellites of chromosomes 1, 9, and 16, resulting in centromeric instability [32]. Global methylation in ICF patients appears to be normal, but other heterochromatic regions like the inactive X chromosome in females might also be hypomethylated. Dnmt3b is therefore specifically needed for the methylation of the pericentromeric satellite repeats [31].

### **1.3.3) DNA Methylation Pattern Establishment in Early Development**

There are two rapid DNA methylation reprogramming phases during early development, one occurs in primordial germ cells (PGCs) and the other in preimplantation embryos, just after fertilization [33, 34]. During these times, old epigenetic marks are erased and replaced by new marks. The epigenetic reprogramming of PGCs is an opportunity for the developing germ cells to reset imprints. The functional role that erasure of DNA methylation patterns plays in preimplantation embryos is not known, but one model suggests that it erases the gametic methylation marks, allowing for chromatin decondensation and transcription of zygotic genes needed for early development (reviewed in [19]).

Mice models were used to study the methylation pattern establishment during early development. PGCs, derived from epiblast cells, have highly methylated DNA. When these cells reach the developing germinal ridge at E11.5 (embryonic day 11.5), they start to differentiate and expand (reviewed in [33], [34]). At the same time, the cells experience genome-wide demethylation and at E12.5, the majority of the methylation marks have been erased. This rapid demethylation is believed to be active demethylation

since it takes place in the presence of the maintenance DNA methyltransferase, Dnmt1[33]. The mechanism behind active demethylation is not known, but several models exist (see below). The demethylation phase is followed by a *de novo* methylation phase which establishes new imprint marks. The timing of *de novo* methylation is not exactly known in females, but is thought to initiate at ~ 14.5 dpc (days post-coitum) in males [34]. The final result is highly methylated mature gametes.

The next phase of epigenetic reprogramming occurs after fertilization. At this time, the maternal and paternal genomes are in different stages of the cell cycle [33]. The paternal genome is present in single copy and is tightly compacted with protamines in place of histones. The maternal genome, on the other hand, is present in diploid, as the oocyte was arrested at metaphase II. Following fertilization, histones replace the protamines in the paternal genome and the maternal genome finishes meiosis. Both parental genomes then experience a round of demethylation, but not by the same mechanism.

The paternal genome undergoes active demethylation over a short period of time whereas the maternal genome experiences passive replication-dependent demethylation over several cleavage divisions [35]. The active demethylation of the paternal genome occurs between 6-8 hours following fertilization, and is completed before transcription and DNA replication is initiated. Several regions of the genome are protected from demethylation including imprinted genes, centromeric heterochromatin and IAP retrotransposons (reviewed in [33]). As mentioned above, the mechanism for active demethylation is not known, nor is it known whether or not the mechanism is the same between PGCs and the paternal genome in the preimplantation embryo. There is evidence supporting that this is an enzyme driven reaction occurring in the oocyte as oocytes are capable of actively demethylating transferred somatic nuclei (reviewed in [33]). Both direct and indirect demethylation mechanisms have been proposed (reviewed in [34]). In direct demethylation, the C-C bond is broken in order to remove the methyl group. Indirect demethylation, on the other hand, involves the removal of the methylated cytosine base and replacing it with an unmethylated cytosine or the CpG dinucleotide is removed by nucleotide excision. Another indirect mechanism is hydrolytic deamination

of the methylated cytosine, resulting in thymine. In the next replication cycle, the thymine is replaced with cytosine by the DNA repair machinery.

The passive demethylation in the maternal genome results from exclusion of the oocyte form of the maintenance DNA methyltransferase, Dnmt1o, from the nucleus (reviewed in [33, 34]). Methylation is progressively lost from the genome over cell divisions as methylation is unable to be maintained. However, imprinted genes do not lose their methylation marks, similar to what occurs in the paternal genome. By the blastocyst stage, both the maternal and the paternal genome have lost the majority of the gametic methylation marks (reviewed in [19]).

Lineage specific DNA methylation patterns are re-established following implantation [19]. The inner cell mass (ICM), which goes on to form the embryo proper, experiences *de novo* methylation perhaps as early as the late morula stage and the trophoectoderm (TE), which differentiates into the trophoblast and primitive endoderm, is relatively undermethylated (reviewed in [33]). Methylation still occurs in the TE, however it is not maintained in the trophoblast and primitive endoderm lineages [19].

#### **1.3.4) Link Between Methylation and Chromatin Modifications**

As previously mentioned, DNA methylation is an epigenetic mark of inactive, transcriptionally repressed chromatin. There are two popular models explaining how DNA methylation is able to cause silencing. In the first mechanism methylation blocks the direct binding of the basal transcriptional machinery or transcription factors (reviewed in [19, 36]). This is, however, not believed to be the primary mechanism of DNA methylation silencing as it is rare *in vivo* and transcription of heavily methylated genes has been observed in the absence of methyl-CpG binding proteins.

The second mechanism connects DNA methylation with other epigenetic marks of inactive chromatin. DNA methylation is associated with histone modifications, like histone deacetylation and histone H3 lysine 9 (H3K9) methylation, to produce heterochromatic regions [37]. Many of the chromatin modification proteins are found in large complexes supporting the idea that there is communication between the marks to establish repressive chromatin together (reviewed in [19]). For example, Dnmt1 and Dnmt3a have both been shown to interact with histone deacetylases (HDACs). Proteins

that specifically bind to methylated CpG dinucleotides and the complexes they form are discussed below.

Even though DNA methylation does not act on its own to cause epigenetic silencing, there is evidence suggesting that it is the dominant event stabilizing transcriptional repression (reviewed in [9, 37]). For example, treatment of tumour cells with trichostatin (TSA), an inhibitor of HDAC, was not sufficient to reactivate the aberrant silencing of tumour suppressor genes. However, when the cells were first treated with demethylating agents like 5-aza-cytidine, and then treated with TSA, reactivation of the silenced genes was observed.

#### ***1.3.4.1) Methylated DNA Binding Protein Complexes***

Proteins capable of binding methylated CpG dinucleotides were first identified by Adrian Bird's group and were later shown, along with Dnmts, to directly interact with HDACs (reviewed in [18]). The first methyl-CpG-binding protein identified was MeCP2 which is made up of a methyl-CpG-binding domain (MBD) and a transcriptional repression domain (TRD) (reviewed in [36]). The TRD domain has been shown to interact with Sin3A, which in turn binds HDAC. The overall result is a complex that is able to cause transcriptional silencing in a methylation dependent manner [19]. Five other methyl-CpG-binding proteins have been identified, MBD1, MBD2, MBD3, MBD4 and Kaiso. MBD1-4 were found based on their possession of a domain closely related to MeCP2's MBD domain [36]. Kaiso, on the other hand, does not have a MBD domain, but is able to bind methylated CGCG (reviewed in [19]). MBD1, 2 and 4 all preferentially bind methylated CpG dinucleotides. However, MBD4 is not involved in transcriptional silencing but is thought to play a role in DNA mismatch repair. MBD2 and MBD3 are closely related proteins, showing sequence similarity beyond the MBD domain, and both interact with the Mi2/NuRD deacetylase complex (reviewed in [36]). However, since MBD3 does not have significant affinity for methylated CpG dinucleotides, it is believed that MBD3 is recruited to site of methylation by MBD2. The MBD protein and chromatin modifying protein complexes provide a link between DNA methylation, histone deacetylation and transcriptional repression.

#### **1.3.4.2) Inactive Chromatin**

Features of inactive chromatin, specifically DNA methylation associated silencing, have been discussed above. It is well established that there is a mechanistic link between DNA methylation and histone acetylation, but the sequential order of events is not known. Does DNA methylation precede histone deacetylation or vice versa? Both models are discussed below.

##### **1.3.4.2.1) Histone Modifications Directing DNA Methylation**

There are two models of histone modification directing DNA methylation. The first involves H3K9 methylation and the second, histone deacetylation. There is evidence to support both models and perhaps both occur under different situations/conditions. Support for the latter model, histone deacetylation directing DNA methylation comes from the treatment of *Neurospora* cells with TSA which results in selective loss of cytosine methylation (reviewed in[37]).

DNA methyltransferases have been shown to interact with a H3K9 methyltransferase, Suv39h, potentially through the aid of an adaptor molecule like HP1 (heterochromatin protein 1) (reviewed in [37]). Therefore, according to the first model, H3K9 methylation marks inactive chromatin and recruits HP1. HP1 then either directly or indirectly attracts DNA methyltransferases to maintain the DNA methylation and silencing. The observed decrease in Dnmt3b-dependent CpG methylation at centromere satellites in Suv39h-knockout embryonic stem cells supports this model. In addition, mutations in *Neurospora crassa*'s dim-5 ( defective in methylation 5), a H3-K9 methyltransferase, results in loss of all DNA methylation (reviewed in [19].)

##### **1.3.4.2.2) DNA Methylation Directing Histone Modification**

In this model, DNA methylation is the primary event leading to histone deacetylation and H3K9 methylation. The DNA is *de novo* methylated by Dnmt3a and Dnmt3b which then triggers the recruitment of the CpG-binding-protein and HDAC complexes mentioned above. The methylated DNA, deacetylated histones and inactivated chromatin recruit H3-K9 histone methyltransferases (HMT), again potentially through the aid of HP1. The final result is stably inactivated chromatin (reviewed in [19]).

### **1.3.5) Methylation and Repetitive Elements**

As previously mentioned, one theory on the evolution of DNA methylation involves host defense against selfish mobile elements, DNA sequences that are capable of moving and inserting into new locations within their host's genome. There are four different kinds of transposable elements, DNA transposons, and three kinds of retrotransposons, long interspersed nuclear elements (LINEs), short nuclear interspersed nuclear elements (SINEs), and retrovirus-like elements. The classes can be further divided based on whether they can transpose independently (autonomous) or not (nonautonomous). DNA transposons move by a 'cut and paste' mechanism using a transposase enzyme. Retrotransposons move through an RNA intermediate that is reverse transcribed back into DNA and inserted elsewhere in the host genome. It is estimated that approximately half of the mammalian genome is made up of transposable elements (reviewed in [38]).

These elements pose a threat to the host genome as they can integrate into and disrupt genes, or they can cause rearrangements through homologous recombination. Despite this destructive potential of transposable elements, only 1 in 500 new germline mutations are caused by these elements due to the host's mobility control mechanisms (reviewed in [39]). There are two known mechanisms, cosuppression, usually through small interfering RNAs (siRNA), and cytosine methylation (reviewed in [38-40]).

There is considerable evidence supporting a role of methylation in regulating transposable element activity. As mentioned above, there is a well established association between DNA methylation and gene inactivation and inactive condensed chromatin. The majority of the genome's methylated cytosines lie within transposable elements (reviewed in [39]). Correlation between the demethylation of mouse intracisternal A particles (IAP), retroviral-like retrotransposons, and an increase in IAP expression has also been observed (reviewed in [38]). In addition, *Drosophila*, who lack DNA methylation, experience a high frequency of mutations caused by transposons (reviewed in [39]). Methylation can also indirectly regulate mobile elements by the induction of methylated C – T mutations over time.

There is a point in development when transposable elements are particularly active, and this corresponds with the time the host genome experiences a wave of

demethylation followed by remethylation (reviewed in [38-40]). This occurs during gametogenesis and, from the selfish perspective of the mobile element, this is the best time to be active (reviewed in [40]). Transposable element insertions in somatic cells would fail to be passed onto the next generation and the insertion could harm the host, and therefore the fitness of the transposable element. Active mobile elements have also been observed in some malignant cells (reviewed in [39]). Tumours are often globally hypomethylated, resulting in a lack of suppression of transposable element activity, and this might lead to the chromosomal instability observed in cancers.

#### **1.4) X Chromosome Inactivation**

Dosage compensation between mammalian females and males is obtained through the silencing of one of the two X chromosomes in females during early embryogenesis. In other organisms, dosage equivalency is achieved via different mechanisms. For example, in *Drosophila melanogaster*, expression of genes located on the single X chromosomes in the males is upregulated and in *Caenorhabditis elegans*, gene expression on both X chromosomes in females is down regulated (reviewed in [5]). Even though the whole X chromosome is inactivated as high as 25% of the genes can escape silencing in humans [41]. The X chromosome is approximately 155 Mb in size and contains 1098 annotated genes. This makes the X chromosome relatively gene poor, but it does have a higher number of repetitive elements (reviewed in [5]). There is an abundance of LINE 1 (L1) elements, and relatively reduced number of SINEs, except for Alu sequences. The high number of LINE 1 elements on the X compared to autosomes, and the limited spread of inactivation into the autosome that occurs in X/autosome translocations, led to the Lyon “repeat” hypothesis (reviewed in [5]). In this hypothesis, proposed by Mary Lyon in 1998, L1 elements act as way stations to amplify and spread the inactivation signal.

Prior to, or concomitant with X inactivation initiation, counting and choice decisions are made. Only one X chromosome per diploid autosome set remains active in female cells. Both counting and choice of which X chromosome to inactivate are controlled by the X-inactivation centre (XIC) located at Xq13 (reviewed in [5, 6, 42]). The exact elements responsible are not known, but the factor responsible for counting has been localized to a 20 kb bipartite domain located 3' to Xist (reviewed in [42]). Choice is

also controlled by XIC, but it appears that more than one element in this region is responsible. In mice, the antisense of Xist, Tsix, is believed to play a role (reviewed in [42]).

Initiation of X inactivation is caused by the *cis*-expression of and coating of the future inactive X chromosome by the X-inactive-specific transcript (XIST). XIST is a 17-kb untranslated RNA expressed from the inactive X chromosome (reviewed in [5, 42]). Following the coating of the inactive X chromosome by XIST, sequential epigenetic features are recruited resulting in facultative heterochromatin. Histone 3 hypoacetylation occurs after Xist coating, and this is followed by hypomethylation of histone 3 lysine 4 (H3K4), methylation at histone 3 lysine 9, 20 and 27 (H3K9, 20, 27), and ubiquitinylation of histone H2A (reviewed in [5, 42]). This leads to gene inactivation, which is stabilized by further chromatin modifications like the incorporation of the histone variant macro H2A, and lastly by DNA methylation. DNA methylation is key to ‘locking’ the inactive state of the X chromosome, as treatment with agents that inhibit DNA methylation, like 5-aza-2'-deoxycytidine (5-azaCdR) result in the reactivation of previously silenced X-linked genes (reviewed in [18]). Dnmt1 activity is required to maintain the inactive state of the X chromosome, further supporting a stabilizing role of DNA methylation [43]. In addition, the genes known to escape inactivation were found to have relatively fewer associated CpG islands and SINE MIR elements, suggesting that the CpG islands provide targets for DNA methylation in stabilizing X chromosome inactivation [44]. Other features of the inactive X include late replication timing, peripheral nuclear localization and the formation of a condensed Barr body (reviewed in [5]).

This thesis focuses on abnormal methylation on the X chromosome in Non-Hodgkin’s lymphoma. It is interesting that a lot of the early work that set the grounds for hypermethylation studies in cancer came from X inactivation research (reviewed in [18]). For example the relationship between methylation and gene inactivation was first noted in studies looking at the inactive X. In addition, the first observation of specific methylation of CpG dinucleotide clusters on the inactive X chromosome later led to the classification of CpG islands by Adrian Bird (reviewed in [18]).

#### ***1.4.1) Mouse/Human Hybrid Cell Lines***

When mouse/human hybrid cell lines are created, most of the human chromosomes are lost. The advantage to creating such cell lines is that under tissue culture selection, hybrids retaining either the inactive or active human X chromosome can be selected for. This allows observations to be made on features of either X chromosome. This is particularly of interest in this thesis since the methods used to assess methylation status (see Methods below) can not distinguish between normal methylation seen on the inactive X chromosome in females and abnormal methylation seen on the active X chromosome in malignant cells. The use of two such hybrid cell lines, AHA-11aB1 (retains the human active X) and t86-B1maz1b-3B (retains the human inactive X), were used to assess the methylation status of CpG islands which had unexpected methylation results such as methylation of the single active X in the male control cell line and a lack of methylation in female control cell line.

#### **1.5) DNA Methylation and Cancer**

Today, cancer is recognized as both a genetic as well as an epigenetic disease. DNA methylation and histone modification, as discussed above, are epigenetic modifications; and abnormalities of both have been identified in neoplasias. Abnormalities in the methylation patterns in cancer include both hypomethylation and hypermethylation. Hypomethylation occurs at the global level whereas hypermethylation is a focal event. Since methylation is associated with gene silencing, hypomethylation is usually considered an oncogenic event, leading to the activation of normally silenced genes and repetitive elements and hypermethylation is considered to silence tumour suppressor genes, which are usually active in the normal cells.

In 1983, hypomethylation was the first epigenetic modification to be associated with the development of cancer (reviewed in [17, 18, 21]). Malignant cells can experience a 20-60% reduction in global methylation relative to normal non-malignant counterparts (reviewed in [45]) and it is now known that global hypomethylation occurs in almost every human cancer (reviewed in [46]). Hypomethylation removes normal methylation marks leading to the reactivation of transposable elements and loss of centromeric methylation. Hypomethylation can therefore contribute to the characteristic

chromosomal instability observed in cancers as well as loss of imprinting (reviewed in [45, 47]).

Site-specific hypermethylation in malignancies was observed later, but has received more attention than hypomethylation. This was mostly due to limitations in experimental design as it was easier to look at specific sites than it was at global events, but this is now changing with the advent of techniques that look at the global genomic picture (reviewed in [18]). The focus on hypermethylation has led to the identification of a growing number of tumour suppressor genes silenced by methylation. Epigenetic abnormalities are recognized today as being just as important as genetic mutations in neoplastic development. The remainder of the introduction will focus on focal aberrant methylation.

### ***1.5.1) Tumour Suppressor Genes Silenced by Methylation***

Tumour suppressor genes have roles in cell-cycle regulation, tumour cell invasion, DNA repair, chromatin remodeling, cell signaling, transcription and apoptosis. (reviewed in [17]). When these genes are silenced, the tumour cells have a growth advantage. The association between abnormal hypermethylation and the silencing of tumour suppressor genes was first identified in retinoblastoma in 1989, when a CpG island located at the 5' end of RB, the first known tumour suppressor gene, was found to be abnormally hypermethylated in a sporadic unilateral retinoblastoma tumour [48]. Since methylation is associated with gene silencing, the observation of abnormal methylation in cancer led to the hypothesis that methylation might be able to cause silencing of tumour suppressor genes. The same group who initially saw the abnormal methylation of RB, later confirmed that a reduction in RB gene expression correlated with abnormal allele specific methylation in retinoblastoma tumours [49].

Since the identification of RB methylation, many more known tumour suppressor genes have been found to be silenced and abnormally methylated. For example, a CpG island located in the 5' region of the von Hippel-Lindau (VHL) tumour suppressor was shown to be abnormally methylated in some spontaneous clear-cell renal carcinomas [50]. In some of these cases, both alleles experienced abnormal hypermethylation, and in others, one allele was lost and one was hypermethylated. In all cases, there was a lack of

VHL expression and treatment with 5-aza-2'-deoxycytidine resulted in re-expression. Another example of abnormal methylation and silencing is the DNA mismatch repair gene, hMLH1, in colon cancer [51]. Treatment with 5-azacytidine resulted in the re-expression of hMLH1, similar to what occurred with VHL.

New tumour suppressor genes have also been identified through methylation studies. For example, the suppressor of cytokine signaling (SOCS-1) was shown to be abnormally methylated and transcriptionally silent in both hepatocellular carcinoma cell lines and primary tumours [52]. When SOCS-1 was re-expressed, the growth of the malignant cells was suppressed. These results suggest that SOCS-1 might have a role as a tumour suppressor.

The Knudson two hit hypothesis states that for the complete inactivation of tumour suppressor genes, both alleles of the gene need to be mutated either by intragenic mutations or deletions. With the evidence for methylation associated silencing of tumour suppressor genes, the Knudson hypothesis has been modified to include abnormal methylation as one of the two hits (reviewed in [8]).

### ***1.5.2) Potential Sources of Abnormal Methylation***

The cause of the observed abnormal methylation is still unknown, but is speculated on in several theories. Seeds of methylation, aging, chronic inflammation, and viral infection potentially all play a role and are not necessarily independent from each other. These are discussed in more detail below.

#### ***1.5.2.1) Seeds of Methylation***

This model was suggested by Clark and Melki in 2002 [53] and is based on the theory that gene inactivation precedes maintenance methylation in malignant cells (discussed further below). Low levels of CpG dinucleotide methylation have been observed in CpG islands in normal cells (reviewed in [53]). Clark and Melki believe that in normal cells, there is a balance between low levels of *de novo* methylation and demethylation as a result of active gene transcription or active chromatin. In the abnormal tumour cells, however, there is a lack of active transcription resulting in a shift in the balance between *de novo* methylation and demethylation. Therefore, random methylated

CpG dinucleotides serve as ‘seeds’ of methylation that is able to spread in the abnormal malignant cells. This model stems from the observation that methylated DNA, full or partial, is able to induce *cis* spreading of methylation better than unmethylated DNA (reviewed in [11]). Turker and his colleagues suggested that methylation centers exist and are sequences like repetitive elements that become methylated (reviewed in [20, 54]). It is hypothesized that in order to retain gene expression, yet to be defined barriers exist that prevent *de novo* methylation spreading into unmethylated CpG rich region (reviewed in [9, 55]). In malignant cells, these barriers are believed to break down, allowing methylation to spread.

### 1.5.2.2) Aging

The role of aging and abnormal methylation is related to the above model of methylation seeding and spread of methylation. DNA methylation appears to be a function of time, slowing progressing from regions outside of CpG islands and spreading to transcription start sites (reviewed in [9, 11]). It has been shown that exonic CpG islands are more susceptible to methylation relative to CpG islands located in promoters because they lack the protection against methylation from transcription factor binding (reviewed in [11]). Over time, this methylation can then act as seeds of methylation, spreading into the gene promoters, perhaps because of a reduction in gene expression with age [20].

The first observed gene with promoter hypermethylation associated with aging was the estrogen receptor (ER), a known tumour suppressor gene, in normal human colonic mucosa cells [56]. ER is methylated in most colorectal tumours where its expression is either reduced or completely silenced. Another example of a gene experiencing an increase in methylation with age is the imprinted insulin-like growth factor II (IGF2) promoter [57]. This gene is normally expressed from the paternal allele and the maternal allele is methylated. Issa *et al.* (1996) noted that over time, the paternal allele also became methylated. IGF2 is transcribed from four different promoters and in many cancers, expression from three of these promoters is reduced or absent and is associated with promoter hypermethylation [57]. Normal methylation increase with aging was also seen for the candidate tumour suppressor gene, DBCCR1, in normal urothelial

cells [58]. Methylation of this gene increased in transitional cell carcinoma (TCC) of the bladder and upper urinary tract.

These observations all provide links between methylation and aging, potentially explaining why one of the most common risk factors for cancer development is aging (reviewed in [9]).

#### ***1.5.2.3) Chronic Inflammation***

A link between chronic inflammation and a relative increase in abnormal methylation compared to patients who do not suffer from chronic inflammation has been observed in several cancers. In hepatocellular carcinoma (HCC), patients with cirrhosis and hepatitis C experience a higher degree of methylation relative to HCC patients without these chronic illnesses [59]. There is geographic variation in the development of HCC, and in high risk populations in the Far East and sub-Saharan Africa, hepatitis B and C account for 50-90% of HCC cases (reviewed in [59]). Another example linking chronic inflammation and methylation occurs in colon cancer. Patients suffering from ulcerative colitis (UC), which causes chronic inflammation and high cell turnover, have a high risk for developing colon cancer (reviewed in [60]). UC patients with colon cancer have a higher degree of methylation in both their dysplastic and nondysplastic mucosa relative to UC patients without colon cancer and non-UC controls [60]. The authors propose that high cell turnover caused by chronic inflammation, prematurely ages the cells which leads to accelerated CpG island methylation.

#### ***1.5.2.4) Viral Infection***

Viral infection and chronic inflammation can be related since some viral infections cause chronic inflammations, like the above example of hepatitis C and cirrhosis. It is not known whether the increase in methylation is associated with proliferative changes caused by chronic inflammation or with the viral exposure [59]. An example of a link between viral infection and cancer is the correlation between Epstein-Barr virus (EBV), a herpes virus, and the development of gastric carcinoma (GC). This association predicts a causal role of EBV in developing GC and it was recently shown that GC patients with EBV experience more abnormal methylation relative to EBV-

negative GC patients [61]. In one study, *de novo* methylation in cells infected with EBV was hypothesized to suppress immunodominant viral antigens in attempt by the virus to prevent detection by the host immune system [62]. These results suggest that viral infection can lead to the activation of *de novo* methylation which in turn can also methylate host genes (reviewed in [59]). The evidence supporting a role of both chronic inflammation and viral infection as sources/triggers of abnormal methylation in cancer also suggests that environmental influences can effect epigenetic modifications [20].

### **1.5.3) CpG Island Methylator Phenotype (CIMP)**

Many questions remain to be answered in the study of abnormal methylation in cancer and one of the main topics concerns the cause of this abnormal methylation. There are two potential possibilities, methylation normally occurs at random (due to various causes as described above) and is selected for in malignant cells because of a selective growth advantage, or there is an underlying mutation in the DNA methylation machinery (reviewed in [20]). The observation of increased DNMT1, DNMT3a and DNMT3b expression in human colon cancer supports the second hypothesis [63]. The increase in DNMT1 expression was confirmed by looking at protein expression in colon cancer [64]. However, other studies did not observe a correlation between expression levels of any of the DNMTs and *de novo* methylation of sampled commonly hypermethylated CpG islands in colon cancer [25].

The identification of multiple genes with abnormal methylation in some tumours led to the suggestion of a CpG island methylator phenotype (CIMP). CIMP implies that there is some defect in the DNA methylation machinery resulting in elevated levels of DNA methylation (reviewed in [20]). The majority of the CIMP research was done on colon cancer with microsatellite instability (MIS) which often experience abnormal methylation of several loci within the same tumour, including INK4A coding gene CDKN2A, thrombospondin 1 (THBS1), hyperplastic polyposis gene 1 (HPP1) and the mismatch repair gene MLH1[65-67]. The restoration of DNA mismatch repair function of MLH1 following treatment with 5-aza 2'-deoxycytidine suggests that the abnormal methylation of the MLH1 promoter is responsible for MIS [68]. The abnormal methylation of the promoters of the genes mentioned above was not observed in inherited

cases of MIS-positive colon cancer, where the MLH1 gene was silenced via genetic mutations [69]. Further observations led to the breakdown of sporadic colon cancer into 4 subsets based on methylation patterns and MIS. This classification led to the identification of colorectal cancer CIMP specific features of prognosis and histology (reviewed in [20]).

CIMP is not isolated to colon cancer. It has also been shown, amongst others, in pancreatic adenocarcinoma [70], gastric cancer [71] and in adult acute lymphocytic leukemia [72]. The identification of CIMP is significant in understanding cancer and methylation changes for two main reasons. First, it led to the additional sub classification of types of cancers, which could lead to advancements in the treatment of CIMP patients. The other significant contribution of these findings is that it points to an abnormality in DNA methylation control as a culprit in aberrant methylation patterns in cancer.

There is, however, still great debate over the existence of CIMP. For example, Anacleto *et al.* (2005) found evidence disputing the existence of CIMP [73]. In their study, they did not observe a discontinuous distribution of methylated genes. They did find an association between methylation of three or more loci and location of the tumour proximal to the splenic flexure. However, this association disappeared when tumours with MLH1 methylation and MSI were removed from the analysis. There is a known association between MSI and tumour location in hereditary non-polyposis colorectal cancer (HNPCC) in which methylation is rare, suggesting that CIMP is not real and that the associations are a result of statistical artifacts caused by confounding of associations. On the other hand, arguments for CIMP include sporadic MSI cancer specific characteristics, like BRAF mutations, which are believed to be a consequence of CIMP (reviewed in [74]). In addition, groups of cancers with the CIMP phenotype that lack MSI have also been identified (reviewed in [74]).

#### **1.5.4) Gene Inactivation and Methylation**

There are two lines of thought when it comes to the sequential order of gene inactivation associated with promoter methylation. One theory is that gene inactivation precedes methylation and the other is the reverse, that methylation precedes gene inactivation. Therefore, methylation either stabilizes silencing or triggers silencing. The

sequence of events is difficult to determine since methylation is an early event in tumorigenesis and when malignant samples are analyzed, gene inactivation and methylation have both already been established (reviewed in [53]).

In the model suggesting that gene inactivation comes before methylation, a process much like ‘seeding’ of methylation (described above) occurs, where there is a shift in the dynamic balance between gene transcription and the spreading of *de novo* methylation from *cis* acting elements (reviewed in [54]). Evidence for this model comes from studies looking at the mouse *Aprt* gene and its associated B1 repetitive elements and Sp1 binding sites. The B1 elements were shown to experience *de novo* methylation and could potentially act as a *cis*-acting element [75]. When one or more of the Sp1 binding sites associated with *Aprt* were lost either through site directed mutagenesis of transcription factor binding sites or deletion, spread of methylation into the promoter was observed [76]. The reverse has also been shown, that Sp1 sites can prevent methylation associated silencing of a transgene [77] and a provirus [78].

The example above however does not occur under malignant situations. It is still unknown if the abnormal methylation in cancer is mediated by the same mechanisms that cause methylation in normal development. Tumours with variation in methylation between cells but show gene silencing no matter of the degree of methylation, provide support for the model predicting gene inactivation prior to methylation in cancer cells (reviewed in [53]). Clark and Melki suggest that the genes are inactivated in these cells, rendering them susceptible to *de novo* methylation that randomly accumulates into variable patterns within and between tumours. In a previous study, the authors looked at the methylation status of *GSTP1*, which is frequently hypermethylated at both alleles in prostate cancer (reviewed in [53]). They found that the removal of transcription factor binding sites, like Sp1 sites, did not cause hypermethylation, but neither did the silencing of the *GSTP1* promoter. However, when the promoter was silenced in combination with random *de novo* seeds of methylation, hypermethylation of *GSTP1*’s associated CpG island was triggered [79].

Arguments for methylation causing gene inactivation comes from *Dnmt1* mutant mice studies and from claims disputing the timing of methylation and gene inactivation during the normal process of X inactivation (reviewed in [8]). Mice with targeted

mutations in Dnmt1 experience disrupted genomic imprinting [80] and reduced tumour formation [81]. These observations suggest that in the absence of functional DNA methyltransferase, methylation associated gene inactivation is inhibited. However, both this observation and the ability of DNA methyltransferase inhibiting agents to re-activate gene expression, do not rule out the possibility of gene inactivation occurring first. They only demonstrate the requirement of DNA methylation for stabilizing gene inactivation.

The sequential order of events in the process of X inactivation have been established with gene inactivation preceding maintenance methylation. However, some dispute this order of events and suggest that methylation and gene inactivation occur at the same time. These arguments stem from the fact that the first studies were done on intronic sites rather than promoters (reviewed in [8, 82]).

#### ***1.5.5) Models of Gene Silencing Associated with Abnormal Methylation***

Robertson (2005) reviewed two models that link normal methylation occurring during development and abnormal methylation in cancer [17]. The first model involves an antagonistic relationship between two zinc finger proteins, CTCF and BORIS [17]. CTCF is a CCCTC-binding protein involved in regulating imprinted gene expression by acting as a chromatin insulator. BORIS is a CCCTC-binding factor-like, or CTCFL protein, that shares homology with the CTCF zinc finger region [83]. Expression of these proteins is mutually exclusive and in cancer, BORIS is upregulated. In the model, methylation is the default state of the genome, and CTCF binding creates methylation free regions. In cancers, competition for binding sites between these two proteins leads to hypermethylated regions because BORIS is present at higher levels. When BORIS binds sites that are normally bound by CTCF, the methylation barrier is lost, allowing Dnmts to methylate the region. In this model, global hypomethylation experienced by cancerous cells is caused by the spreading of the repression machinery out and away from the previously methylated DNA sequences following the breakdown of the methylation barriers.

The second model involves the RNAi pathway (reviewed in [17]). DNA repetitive elements can be transcribed from either strand of DNA and therefore their transcription can lead to the creation of dsRNA molecules [84]. In fission yeast, the RNAi machinery

can recruit histone modifiers to the region to be silenced [85]. In mammalian cells, dsRNA has been shown to cause gene silencing through *de novo* methylation, perhaps through the recruitment of histone modifiers as the RNAi machinery is well conserved [86]. Together, along with evidence of low level of transcription from centromeric repeats in mammals (reviewed in [17]), this suggests that histone modification and DNA methylation of future heterochromatic regions is directed by the RNAi pathway. In cancer, abnormal promoter hypermethylation could result from low level expression from nearby repetitive elements, causing targeted chromatin modifications.

Robertson (2005) also hypothesizes that these mechanisms may act together, with the RNAi pathway directing specific loci methylation and CTCF protecting larger regions against methylation [17].

#### ***1.5.6) Usefulness of Abnormal Methylation as Markers***

The aberrant hypermethylation observed at different tumour suppressor genes can have translational applications. The hypermethylated loci can be used as markers for four different clinical applications, detection of malignant cells, prediction of tumour behaviour, prediction of treatment outcomes and therapy development attempting to reverse the epigenetic mark (reviewed in [45]). The first three are discussed in this section and the development of therapies targeting abnormal methylated loci is reviewed below in section 1.5.7.

Methylation profiles of cancers are being determined to identify cancers based on their methylation patterns [87]. Hypermethylation of genes is considered to be an independent event at each locus (reviewed in [45, 87]) and some cancers will share methylation at some loci, but differ at others, producing a profile unique to cancer types. Cancer cells have been identified in a number of different specimens, including plasma and serum, by methylation analysis. Abnormal methylation is an early event in tumorigenesis, therefore, methylation profiles could be used to for early cancer detection in these specimens (reviewed in [45]). However, detection of malignant cell DNA in patients serum and plasma is likely a result of apoptosis and is an indicator of poor prognosis (reviewed in [88]). For early detection, prior to the appearance of malignant DNA in the serum and plasma, promoter DNA methylation screens of tumour

suppressors could be established in other body fluids like urine, or saliva, and in luminal content, like sputum. Methylation was found to precede the onset of lung cancer by three years in a high-risk group (frequent exposure to carcinogen) [89].

Methylation markers can also be used as prognostic markers in cases where there is a correlation between methylation marks and the aggressiveness of a cancer (reviewed in [45, 88]). Several associations have been documented to date. For example, the methylation of INK4A coding for p16 is linked to poor prognosis in colorectal cancer [90] and high promoter methylation of the adenomatous polyposis coli (APC) gene is associated with poor survival in non-small lung cancer [91]. It is important that the association between methylation and prognosis is observed in the malignant tissue biopsies, since detection of tumour DNA in the bloodstream is thought to be correlated with invasiveness (reviewed in [88]). An association between any tumour specific marker and poor prognosis would likely be seen in plasma or serum samples, but this might not reflect an actual prognostic relationship in the malignant tissues themselves.

The use of abnormal methylation as tumour biomarkers to predict treatment outcomes, aiding in decision of the best treatment course for a patient, is an up and coming field of research. There is some debate on the inferences that can be drawn from retrospective multidrug studies, arguing that prospective randomized clinical trials would provide more substantial evidence, separating any prognostic associations of the marker used to predict treatment outcome (reviewed in [81]). Nonetheless, there have been several studies showing the potential of methylation markers in predicting the outcome of chemotherapy drugs. An example is the association between the abnormal methylation of O6-methylguanine-DNA methyltransferase (MGMT) and the treatment response of gliomas to carmustine [92] and diffuse large B-cell lymphoma (DLBCL) to cyclophosphamide in a multi-drug regime [93]. MGMT codes for an enzyme responsible for removing alkyl groups from guanines, and when this gene is inactive, the buildup of alkylated guanines leads to cell death. Carmustine is an alkylating agent, and gliomas with methylated MGMT were found to be sensitive to this agent. The sensitivity of DLBCL with methylated MGMT to a multidrug regime treatment that included cyclophosphamide, is likely due to MGMT's role in the protection against the toxicity of one of cyclophosphamide's metabolites.

Methylation markers can also be used for the reverse, to predict when certain types of treatment would not work on specific subtypes of malignancies. For example, in breast cancer, if the estrogen receptor (ER) is silenced and associated with methylation, ER methylation can serve as a marker for insensitivity to antisteroidal treatment (reviewed in [45]).

#### ***1.5.7) Methylation as a Target of Cancer Therapy***

Methylation, unlike genetic mutations, is a reversible event and therefore holds great promise as new targets for cancer therapy development. The two classic agents with the capacity to inhibit DNA methylation and re-activate methylation silenced genes, 5-azacytidine (5-aza-CR) and 2'deoxy-5-azacytidine (5-aza-CdR), have been used in tissue culture since the 1980s as cytotoxic agents (reviewed in [45, 94, 95]). These agents are nucleoside analogues that are able to incorporate into replicating DNA in place of cytosines (reviewed in [94]). These analogues cause DNA methyltransferases to covalently attach to the DNA, preventing methylation, and resulting in an overall demethylated state. There are several other nucleoside analogs, as well as procainamide, a DNA methyltransferase inhibitor, antisense oligonucleotides and HDAC inhibitors, some of which have shown some promising results in clinical trials.

Azacytidine has been a particular success story. It was approved by the US FDA for the treatment of chronic myelodysplastic syndrome (MDS) in May 2004 (reviewed in [96]). Clinical trials have shown that azacytidine has a significant effect on MDS, with 54% of MDS imatinib resistant patients showing full or partial hematologic response (reviewed in [96, 97]). Other epigenetic therapy agents are also showing promising results. Treatment with an antisense oligonucleotide, MG<sub>98</sub>, that can down regulate DNMT1, was shown to cause the demethylation of two tumour suppressor genes and methylation of an oncogene in head and neck cancer (reviewed in [97]).

Another area of interest in epigenetic therapy is treatment with a combination of agents (reviewed in [45, 94, 95, 97]). Of particular interest is the coupling of DNA methylation inhibitors with HDAC inhibitors. Several different combinations of these classes of agents has shown a synergistic interaction, causing the reactivation of tumour suppressor genes at lower required doses (reviewed in [45]). In addition to combining

epigenetic therapies, another approach is to treat patients with conventional chemo- or immunotherapy following the sensitization of the cells with epigenetic therapy (reviewed in [94, 95, 97]).

The translation of these agents to human clinical trials has not been easy due to several concerns with the compounds. The first concern is the lack of specificity resulting in global hypomethylation, and nonspecific activation of genes and transposable elements in non-malignant cells (reviewed in [45, 94]). The re-methylation of genes following completion of treatment course with the compound has also been observed (reviewed in [98].) Another concern is the high toxicity of the agents at high doses in normal cells (reviewed in [45]). However, there has been much advancement in this field to improve and address some of these problems. For example, less toxic and orally administered DNMT inhibitor compounds have been developed (reviewed in [99]).

### **1.6) Lymphoma**

Lymphoma is a cancer of the lymphatic system. The lymphatic system has three functions, it returns fluids from tissues back to the blood stream, absorbs fats and some vitamins from the small intestines and transports them to the blood, and is part of the immune system. The lymphatic system is made up of lymphatic vessels, found throughout the body, and lymphoid organs like lymph nodes. The vessels carry lymph, a fluid with lymphocytes (the cells of the immune system). The lymph nodes are responsible for producing and storing lymphoblasts and filtering bacteria, viruses and other foreign particles from the lymph. They are concentrated in the neck, under the arms, and in the groin and abdomen.

In lymphoma, the lymphocytes become malignant, losing growth control and forming tumours. The symptoms of lymphoma are rather non-specific, including enlarged lymph nodes, fatigue, shortness of breath, abdominal swelling, unusual abdominal or back pain, fever, night sweats and unexplained weight loss. Lymphoma is the 5<sup>th</sup> most commonly diagnosed cancer and the 7<sup>th</sup> most common cause of cancer death (<http://www.bccancer.bc.ca/PPI/TypesofCancer/Lymphomas.htm>). Lymphomas are broken into two large categories, Hodgkin's (HL) and Non-Hodgkin's (NHL) lymphoma. Non-Hodgkin's lymphoma is discussed in more detail below in section 1.6.1.

The only known way to differentiate between Hodgkin's and Non-Hodgkin's lymphoma is to examine the malignant cells under the microscope (<http://www.cancercenter.com/non-hodgkins-lymphoma.htm>). In HL, the affected cells are called Reed-Sternberg cells. HL differs from NHL in that it is relatively uncommon, affects younger adults and has a 90% cure rate (<http://www.bccancer.bc.ca/PPI/TypesofCancer/HodgkinsDisease/default.htm>).

### **1.6.1) Non-Hodgkin's Lymphoma**

The incidence of Non-Hodgkin's lymphoma has steadily increased over the past 30 years, but has recently stabilized (reviewed in [100, 101]). In general, there is a higher incidence in males than in females and the highest rates of incidence are found in North America and Australia (reviewed in [101]). In 2000, there was an incidence of 21 males per every 100,000 and 14 females per every 100,000 in Canada and an estimated cause of death for 9 males per every 100,000 and 6 females per every 100,000

(<http://www.bccancer.bc.ca/PPI/TypesofCancer/NonHodgkinsLymphoma/default.htm>).

The lowest incidence rates are found in Asia and Africa and intermediate levels are found in Europe (reviewed in [101]). NHL is more common in adults between the ages 40 and 70, and children are rarely affected. Both the incidence of NHL and death increases with age and peaks at the age of 70 (reviewed in [101]).

The most common risk factor for developing NHL is immune deficiency (reviewed in [100, 101]). Immune deficiencies give a 10-100 greater risk of developing NHL compared to normal. This includes post-transplantation immunosuppression, HIV/AIDS infection, and congenital immune deficiencies in children. Some specific subtypes of NHL are caused by viral infections, for example gastric mucosa-associated lymphoid tissue (MALT) is associated with *Helicobacter pylori* infection. Autoimmune conditions are also associated with an increase in NHL risk, but only moderately and allergic and atopic conditions show no increase in NHL risk.

There are more than 20 different subgroups of Non-Hodgkin's lymphoma (reviewed in [101]) and they all differ in pathogenesis and treatment response (reviewed in [102-104]). Over the years the classification system of NHL has changed with the advancement of new techniques in cytogenetics, molecular genetics and immunology

(reviewed in [102]). For example, previously unidentified subgroups were identified based on gene expression profiling of B-cell lymphomas and normal B-cells (reviewed in [103]). These advancements increased our knowledge of the lymphatic system and provided new diagnostic techniques. There is now a common set of guidelines established by the World Health Organization (WHO) to classify the different types of lymphomas. The use of one standard guideline solved many problems that occurred when each region of the world used their own set of classification, inhibiting the ability to compare cases (reviewed in [102, 104]).

WHOs classification breaks down lymphoma based first on lineage of origin, either B-cell or T-cell/Natural Killer (NK) lymphoma, and second on where in the lymphocytes development the transformation occurred (reviewed in [102, 104]). Malignant lymphocytes appear to freeze in their development at the time of transformation (reviewed in [103]). Lymphomas are further broken down into precursor or mature neoplasms (reviewed in [102, 104]). Lymphomas are also assessed for stage, using the Ann Arbor four staging classification system based on the number of lymph nodes affected, if lymph nodes are found on both sides of the diaphragm and if extralymphatic organs, or sites outside of the lymphatic system are involved. In addition to subgroup classification and staging, lymphomas are also graded using the international prognostic index which predicts risk of disease recurrence and survival. This index is based on five factors associated with survival, age ( $\leq 60$  versus  $> 60$ ), tumour stage (stage I and II versus stage III and IV), number of extranodal sites of disease ( $\leq 1$  versus  $> 1$ ), performance status (general health) (0 or 1 versus  $\geq 2$ ), and serum lactate dehydrogenase (LDH) level (liver function test measuring the amount of serum enzyme released into the bloodstream by the liver) ( $\leq 1$  times than normal versus  $> 1$  times than normal). Risk groups definitions are based on the number of risk factors, 0 to 1 (low), 2 (low-intermediate), 3 (high-intermediate), 4 or 5 (high). There are 9 types of T-cell NHL, all of which are relatively uncommon (reviewed in [105]) and 15 distinct types of B-cell lymphoma (reviewed in [103]).

95% of lymphomas are B-cell lymphoma (reviewed in [103]) and the most frequent subtypes are follicular lymphoma (FL) and diffuse large B-cell lymphoma (DLBCL) (reviewed in [100]). It is not surprising that more B-cells are transformed

relative to T-cells, even though they are present at similar frequencies, because B-cells undergo changes in their genome at three different times during the remodeling of the immunoglobulin genes (reviewed in [103]). Both T- and B-cells undergo V(D)J recombination in the light and heavy chains that make up immunoglobins (Igs) in early development. In addition, B-cells also undergo somatic hypermutation, the introduction of point mutations, deletions, or duplications in the V region of the Ig genes, and class switching, the replacement of heavy chains from one class of Ig to another. These are all points of time in the development of B-cells where errors can occur. Such errors can lead to translocations. One of the hallmarks of lymphoma is the presence of a reciprocal translocation involving the placement of a proto-oncogene under the control of an Ig promoter, which results in constitutive oncogene expression. For example, one of the most commonly seen abnormalities in NHL is a translocation event between chromosomes 14 and 18 ( $t(14;18)(q32;q21)$ ), juxtapositioning the BCL-2 oncogene into the heavy chain region of the Ig loci on chromosome 14 (reviewed in [100, 103]). This results in the inhibition of apoptosis and occurs in 85% of FL and 28% of DLBCL cases (reviewed in [100]). Other transformation events include silencing of tumour suppressor genes, genomic amplifications, translocation involving loci other than Ig and some types of viral infections (mentioned above) (reviewed in [103]).

### ***1.6.2) Lymphoma and the X Chromosome***

X chromosome abnormalities, gains, losses and some structural abnormalities are common occurrences in non-Hodgkin's lymphoma. Incidences between 3-14% have been reported for X chromosomal loss and between 7-33% for X chromosomal gains (reviewed in [3]). Structural changes involving the X chromosome are less common relative to numerical anomalies and one study reported an incidence of 6% (reviewed in [106]). However, of this 6%, 35% involved p22 and these cases were also associated with high grade NHL. This indicates a gene involved in the development of lymphoma might be located there. Harigae *et al.* (2002) identified a case of primary marginal zone lymphoma of the thymus with a duplication involving p22 ( $46,X,dup(X)(p11p22)$ ) [106]. The case in question was progressing slowly unlike the previous cases involving p22, but the high grade p22 cases also had a number of other chromosome abnormalities. This led

the authors to hypothesize that chromosomal changes of p22 are involved in the initial transformation and additional changes need to occur for the progression to high grade disease.

Other studies have indicated the involvement of Xq28, and the putative oncogene MTCP1, in the development of T-cell NHL (reviewed in [107-109]). Renedo *et al.* (2001) did CGH analysis on T-cell NHLs and found the most recurrent abnormality involved either the gain of the whole X chromosome, or Xq26-27 [109]. Balanced translocations involving Xq28 have also been identified in B-cell NHL and in one case of nodal marginal zone B-cell lymphoma, a translocation involving q28 (46,XX,t(X,5)(q28;q22)) was the only detected chromosomal abnormality (reviewed in [107]).

A third region of interest is located at Xq13. A number of cases of NHL with translocations involving Xq13 have been reported, including t(X;4)(q13;p13) in a diffuse mixed NHL, t(X;10)(q13;p13) in a Burkitt NHL case, t(X;12)(q13;q11) and t(X;18)(q13;q11) in two cases of immunoblastic NHL and t(X;18)(q13;p11) in a case of aggressive natural killer (NK) NHL (reviewed in [110]). The cases involving Xq13 translocations are usually associated with high-grade disease. The number of cases with Xq13 breakpoints suggest that a potential oncogene for NHL is located there.

Despite the common occurrence of X chromosomal abnormalities in NHL, there is still uncertainty of the role, if any, the X plays in the development of NHL. X chromosomal gains occur in normal lymphoblasts with aging and therefore the gains seen in NHL might reflect what normally occurs in lymphocytes (reviewed in [4]). In addition, one study showed that of the tumours with chromosome X numerical abnormalities, only an average of 5.2% of the cells in a specimen carried the abnormality [3]. This indicates that the gain of an X is potentially a secondary event.

The study that triggered this thesis' investigation of the X chromosome as a potential location of a candidate tumour suppressor gene, investigated the inactivation status of the acquired X chromosome in NHL patients. The gain of an extra X chromosome associated with normal aging preferentially involves the inactive X and therefore McDonald *et al.* (2000) hypothesized that if a similar mechanism for the gain of the X is involved in NHL as in aging, the inactive X should be over-represented in the tumours. They found that in males, the extra X chromosome did not undergo *de novo*

inactivation, and that there was no preference for which X chromosome was gained in the examined females. One interesting result was the gain of the active X in all four of the informative cases with a translocation involving the heavy chain of Ig and BCL-2 ( $t(14;18)$ ), whereas the inactive X chromosome was gained in four out of the five informative non  $t(14;18)$  cases. The authors hypothesized the gain of an active X chromosome might confer a growth advantage in only some types of lymphomas, like patients with the  $t(14;18)$  translocation [4]. However, the authors offer caution when interpreting this potential association between the gain of an active X and  $t(14;18)$  due to the small numbers of samples.

## **Chapter 2 Materials and Methods**

### **2.1) Tissue Culture**

The majority of the DNA methylation and gene expression analysis was done using cell lines. Table 2.1 is a list of all the cell lines used in this thesis and their tissue culture conditions. There were four control female and male lymphoblast cell lines, five male lymphoma (2 follicular and 3 mantle cell) cell lines, one female human pluripotent embryonal carcinoma cell line and two mouse/human hybrid cell lines. One of the hybrid cell lines retained the human active X chromosome (AHA 11aB1) and the other the inactive X chromosome (t86-B1maz1b-3B). DNA and RNA from these cell lines were isolated from cell pellets. The pellets were made by centrifugating cell suspensions at room temperature and aspirating off the media. The cell pellets were stored at -70°C.

#### ***2.1.1) Thawing Cell Lines***

All cell lines were stored in liquid nitrogen and when needed, underwent a quick thaw procedure. Cells in cryogenic vials were thawed in a 37°C water bath. Then, inside the tissue culture hood, cells were transferred to a t25 flask containing 10 ml of media and placed in CO<sub>2</sub> incubator at 37°C for 4 hours. This allows the cells to settle and the media with the dimethyl sulfoxide (DMSO) (used to freeze down cells, see below) to be aspirated and replaced by 10 ml of fresh media without DMSO.

#### ***2.1.2) Freezing Down Cell Lines***

While cells were thawed quickly, they were frozen slowly. The cells to be frozen down were first pelleted by centrifuging cell suspensions at room temperature. The remaining media was aspirated and the cells were then resuspended in 1 ml of freeze down media. The freeze down media consisted of the cell line's media plus 10% DMSO. The resuspended cells were then transferred to a cryogenic vial which was frozen at -70°C in a holder bathed in isopropanol. After slow freezing for four hours in the -70°C, the vial was transferred and stored in liquid nitrogen.

**Table 2.1** Table of cell lines and their tissue culture conditions. (Xa is an active X chromosome and Xi is an inactive X chromosome.)

Cell Line	Cell Type	Sex	Tissue Culture Conditions
GM 11198	Lymphoblast	female	37°C, RPMI media, 15 % FCS, L-Glut, PenStrep
GM 11199	Lymphoblast	female	37°C, RPMI media, 15 % FCS, L-Glut, PenStrep
GM 11201	Lymphoblast	female	37°C, RPMI media, 15 % FCS, L-Glut, PenStrep
GM 7059	Lymphoblast	female	37°C, RPMI media, 15 % FCS, L-Glut, PenStrep
GM 7033	Lymphoblast	male	37°C, RPMI media, 15 % FCS, L-Glut, PenStrep
GM 7009	Lymphoblast	male	37°C, RPMI media, 15 % FCS, L-Glut, PenStrep
GM 11200	Lymphoblast	male	37°C, RPMI media, 15 % FCS, L-Glut, PenStrep
GM 7057	Lymphoblast	male	37°C, RPMI media, 15 % FCS, L-Glut, PenStrep
SUDHL3	Follicular Lymphoma	male	37°C, RPMI media, 15 % FCS, L-Glut, PenStrep
DoHH2	Follicular Lymphoma	male	37°C, RPMI media, 15 % FCS, L-Glut, PenStrep
HBL-2	Mantle Cell Lymphoma	male	37°C, RPMI media, 15 % FCS, L-Glut, PenStrep
JVM-2	Mantle Cell Lymphoma	male	37°C, RPMI media, 15 % FCS, L-Glut, PenStrep
Z138	Mantle Cell Lymphoma	male	37°C, RPMI media, 15 % FCS, L-Glut, PenStrep
Nteral	human pluripotent embryonal carcinoma, ovarian teratoma	female	37°C, DMEM, 10% FCS
AHA 11aB1	Xa human mouse hybrid	NA	37°C, α-MEM, 7.5% FCS, PenStrep
t86-B1maz1b-3B	Xi human mouse hybrid	NA	39°C, α-MEM, 7.5% FCS, PenStrep

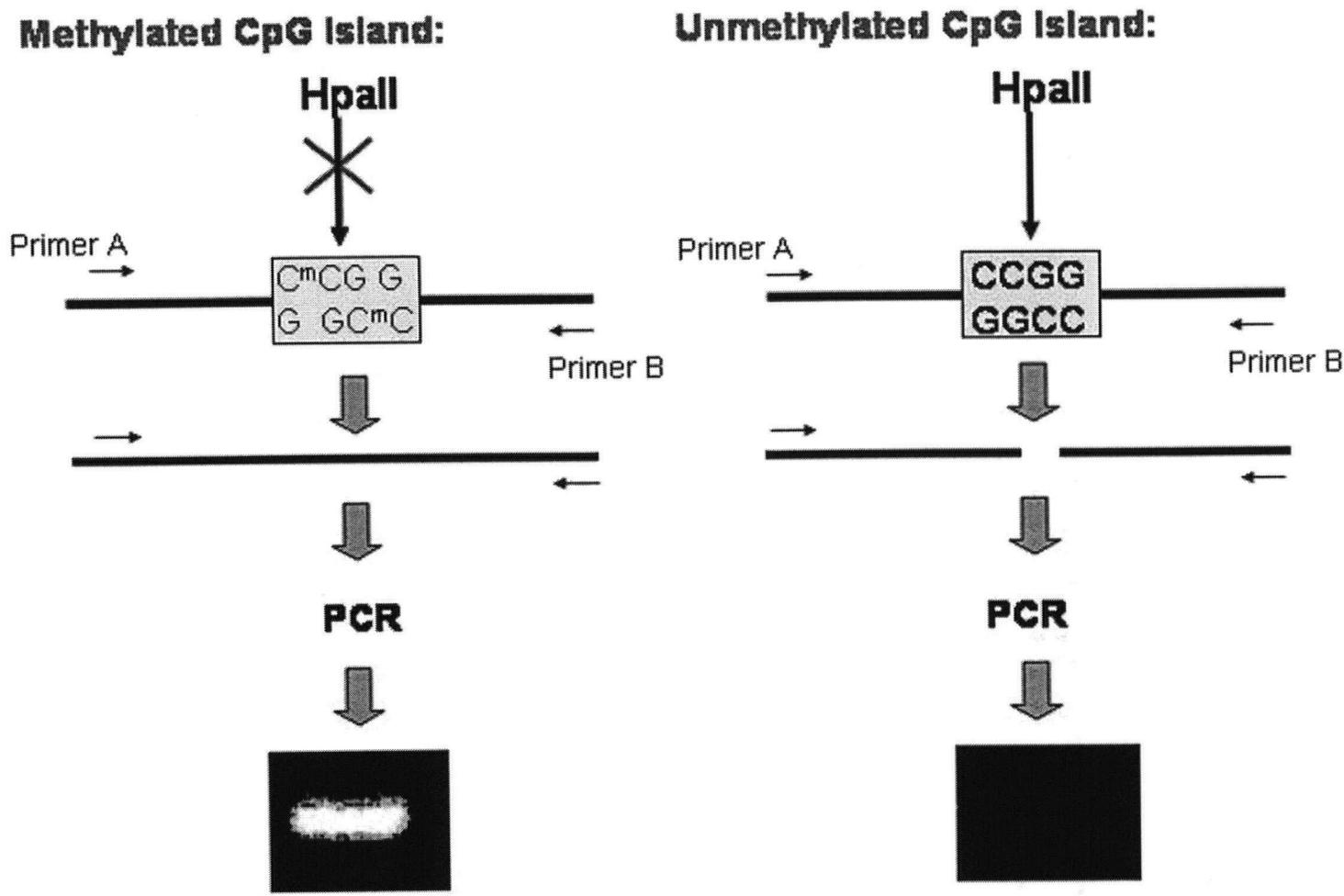
## **2.2) Methylation of CpG Islands**

The methylation status of CpG islands was determined using a restriction enzyme assay. The DNA to be assessed was first digested with a methylation-sensitive restriction enzyme, HpaII. This was followed by polymerase chain reaction (PCR) to amplify a specific region of interest, usually a CpG island, flanking a number of *Hpa*II cut sites. The PCR reactions were then run out on an agarose gel and methylation status was inferred based on either the presence or absence of a band. See figure 2.1 for an illustration of these methods. Each of the steps are described in more detail below.

### ***2.2.1) DNA Extraction***

The DNA extraction protocol used was a salting out technique. This method requires the addition of large amounts of salt in order to precipitate out, or salt out, proteins. This is believed to work by dehydrating the environment surrounding the proteins. This protocol is for cell pellets containing approximately  $5 \times 10^7$  cells from a t75 flask. Cells, from frozen cell pellets, were first resuspended in 2.5 ml of Tris-EDTA (TE) buffer (10 mM Tris pH 7.5 – 8.0 and 1 mM EDTA). Next, 1/20<sup>th</sup> of the volume 20% sodium dodecyl sulfate (SDS) and approximately 1  $\mu$ l of Proteinase K were added. The SDS lyses the cell membranes to release the nucleic acids and the Proteinase K digests any proteins found in the sample. The cells were then incubated at room temperature over night.

The next day, 200  $\mu$ l of 5 M NaCl was added to the lysed cells which were then further incubated at 37°C for another two hours or until in solution. Once in solution, an additional 825  $\mu$ l 5M NaCl was added and followed by vigorous shaking. The solution was then centrifuged at room temperature for 15 minutes at 2,500 rpm. The supernatant was transferred to a new 15 ml falcon tube and an additional 75  $\mu$ l of 20% SDS detergent and 825  $\mu$ l 5M NaCl were added. Once again, the sample was vigorously shaken, centrifuged at room temperature for 15 min at 2,500 rpm and the supernatant was transferred to a new tube. To precipitate the DNA, 2 volumes of 100% ethanol were added followed by gentle rocking. The DNA was then removed from the tube using a glass pipette, and resuspended in 1 ml of water (or TE). The sample was left overnight to slowly dissolve or incubated at 55°C for 2 hours.



**Figure 2.1** Description of the methylation assay

This methylation assay is based on the presence or absence of a PCR product following genomic DNA digestion with the methylation sensitive restriction enzyme, HpaII, and PCR amplification of a region within a CpG island containing HpaII recognition sites.

### **2.2.2) Double Restriction Enzyme Digestion**

The double restriction enzyme digestion protocols used for DNA isolated from cell lines and patient samples differed. The procedure for the cell line DNA spanned four days and involved a DNA phenol: chloroform clean up step to remove the primary cutter's buffer. The starting amount of DNA from patient samples was very low, therefore to reduce the chance of losing DNA during the phenol: chloroform clean up, a protocol lacking this step was used. *EcoRI* was used as the primary cutter and *HpaII*, a methylation sensitive restriction enzyme, was used as the secondary cutter for both protocols. The advantage of using two enzymes is that the *EcoRI* digests the DNA into smaller fragments, increasing the efficiency of *HpaII* in finding it's recognition sequence.

#### **2.2.2.1) Double Restriction Enzyme Digestion of Cell Line DNA**

On the first day of the four-day procedure, the DNA from the cell lines was digested into smaller fragments by the common cutter *EcoRI*. 40 µl of genomic DNA was incubated overnight at 37°C with 5 µl of React 3 buffer (50 mM Tris-HCl (pH8.0), 10 mM MgCl<sub>2</sub>, 100 mM NaCl) and 50 units (U) of Invitrogen EcoRI enzyme.

The following day, 2 µl of 1 mg/ml RNase was added to the sample to remove any RNA which may interfere with the spectrophotometer reading done in a later step to determine the concentration of DNA. After the sample has been incubated at 37°C for 15 minutes, a phenol: chloroform clean up was performed to remove the React 3 buffer and RNase enzyme. A fresh solution of 1:1 Tris-buffered phenol: chloroform was made before use. The reaction volume was brought to 200 µl with water for easy handling. An equal volume of phenol: chloroform (200 µl) was added to the sample which was then vortexed and centrifuged for 5-7 minutes at 13,000 rpm. The top layer was removed to a fresh eppendorf tube and an equal volume of chloroform was added. Again, the sample was vortexed and centrifuged for 5-7 minutes at 13,000 rpm. Finally, the top layer was removed to a new eppendorf tube and 1/10 volume of 3M KOAc salt and twice the volume of 100% EtOH was added. The sample was gently rocked and stored overnight at -20°C.

On the third day of the protocol, the sample was first centrifuged for 10-15 minutes at 13,000 rpm. The supernatant was removed and the pellet allowed to air dry to

eliminate any remaining traces of EtOH which may interfere with the next enzyme digestions. 35 µl of ddH<sub>2</sub>O were added to the dried cell pellet which was then left to sit for 4-6 hours to allow the DNA to resuspend. The DNA concentration of the sample was then determined using the Ultrospec 2000 UV/Visible Spectrophotometer from Pharmacia Biotech. For a 20 µl digestion reaction, 2 µg of *EcoRI* digested DNA was added to 20 U of NEB *HpaII* enzyme and 2 µl of 1X NEB buffer1 (10 mM Bis Tris Propane-HCl, 10mM MgCl<sub>2</sub>, 1mM DTT, pH 7.0). The reaction volume was brought up to 20 µl with ddH<sub>2</sub>O for a final DNA concentration of 100 ng/µl. A mock digest, containing all the reagents of the digested sample except for *HpaII*, was made for every digested sample. Both the mock and digested samples were then incubated overnight at 37°C.

On the last day, the enzyme was heat killed by incubating the digestions for 15 minutes at 65°C. Lastly, PCR (see below for protocol) using two different primer pairs was done to check for complete cutting and the presence of DNA in the digested samples. A primer pair amplifying a region of the pseudoautosomal region on the X chromosome (MIC2 A:B), which is known to be unmethylated, was used to check for cutting. A primer pair amplifying XIST DNA (XIST 3':5') was used to check for the presence of DNA in the digests. For primer conditions, see Table 2.2 below.

#### **2.2.2.2) Double Restriction Enzyme Digestion for Patient Sample DNA**

The double restriction enzyme digestion procedure used on the patient DNA only took two days and did not involve a phenol: chloroform clean up step. Both enzymes were added at the same time using the secondary cutter's, *HpaII*, buffer. For a 20 µl digestion reaction, 2 µl NEB1 buffer, 2 µl *HpaII*, 1 µl *EcoRI* and a total of 15 µl of DNA and ddH<sub>2</sub>O were incubated at 37°C overnight. A mock digest was also made for each sample and was treated the same as the digested sample, except no enzyme was added. The next day the reactions were incubated at 65°C for 15 minutes to heat kill the enzymes. Like the protocol for digesting cell line DNA, PCRs with MIC2 A:B and XIST 3':5' primer pairs were performed to check for complete cutting and the presence of DNA.

A limited amount of patient DNA was available to us courtesy of Joe Connors, Randy Gascoyne, Doug Horsman, and Ron deLeeuw from the BC Cancer Research

Centre, therefore, the smallest quantity of DNA required for the assay needed to be determined. To do this, a gradient of different amounts of cell line DNA was digested using the above protocol. These digests were then amplified with ‘finicky’ primers to see what the lowest quantity of DNA was that could still produce a band on a 2% agarose gel. This was found to be as little as 200ng of DNA in a 20 µl digest reaction (10 ng/µl).

### **2.2.3) Primer Design**

All primers for the methylation analysis were designed using sequence and CpG island information from the University of California Santa Cruz (UCSC) genome browser (<http://genome.ucsc.edu/>). The sequence information came from the May 2004 freeze of the human genome assembly. The algorithm the database used to search for CpG islands had the following criteria: a GC content greater than or equal to 50%, longer than 200 bp and a ratio of observed to expected CpG dinucleotides (dependent on the number of guanosines and cytosines in the segment) greater than 0.6.

The primers were designed with the aid of bioinformatics tools like Webcutter 2.0 (<http://rna.lundberg.gu.se/cutter2/>) and Applied Biosystems Primer Express Version 1.5 program. Webcutter was used to find where within the CpG island of interest HpaII and EcoRI cut. The primers were designed to flank as many *Hpa*II cut sites as possible while avoiding EcoRI cut sites. Primer Express was used to check for primer dimers, hairpins, GC content and melting temperature. The ideal primers were 20 bp long, amplified a 200-500 bp segment and had a theoretical melting temperature between 58°C and 60°C. The National Center for Biotechnology Information (NCBI) BLAST (<http://www.ncbi.nlm.nih.gov/BLAST/>) program was used to check the specificity of the primers. The primers were then ordered from either the Nucleic Acid Protein Service (NAPS) unit or Integrated DNA Technologies (IDT), Inc..

#### **2.2.3.1) Primer Table**

Table 2.2 is a list of all the primers and their PCR conditions used to analyze CpG island methylation status in the 8 Mb region of interest on the X. The name of the primers refer to the UCSC name of the CpG islands. This nomenclature system is based on the number of CpG dinucleotides in the island and this has resulted in some ambiguity.

Whenever there were two or more islands with the same number of CpG dinucleotides, the name of the primer was followed by a lower case letter dependent on the order the CpG islands were examined (*i.e.* looked at the methylation status of CG27b A:B before CG27c A:B)

#### **2.2.4) Polymerase Chain Reaction (PCR)**

The PCR conditions varied for each primer pair and are listed in table 2.2. Optimization of new primer pairs was done by modifying the following PCR conditions, 30 cycles of a 1 minute denaturation step at 95°C (for methylation analysis) or 94°C (for gene expression analysis), 1 minute annealing step at 54°C, and a 2 minute elongation step at 72°C. All PCR reactions consisted of 1 µl 10X PCR buffer (200mM Tris HCl (pH 8.4), 500mM KCl), 1.5mM MgCl<sub>2</sub>, 0.2mM dNTPs, 1µM primer pair, 100ng template DNA, and 0.625 U Invitrogen Taq DNA Polymerase in a total volume of 25 µl. The reactions were run in either Biometra or Techne Genius Thermal Cylcers.

Primer optimization was done on genomic DNA (gDNA) diluted in NEB1 buffer, the buffer used in the HpaII digestions. Several different modifications of the above PCR conditions were tried to find the optimal PCR conditions for each primer pair. These modifications included: changing the annealing temperature (increasing to remove unspecific bands and decreasing to visualize the expected band), changing the Mg<sup>+2</sup> concentration, adding 4 or 8% DMSO, adding 1 or 2 M Betaine, and when all else has failed, using a different Taq DNA polymerase called platinum Taq Polymerase from Invitrogen. Betaine and DMSO were often added to PCR reactions because the regions of interest are GC rich. These reagents destabilize secondary structures that could otherwise cause the DNA polymerase to fall off the DNA strand.

#### **2.2.5) Gel Electrophoresis**

The PCR products for both DNA methylation and gene expression assays were visualized using gel electrophoresis. Both agarose and polyacrylamide gels were utilized depending on the expected size of the PCR product. Products equal to or greater than 200 bp were run out on agarose gels, whereas smaller expected products were run on

**Table 2.2** Table of primers used to check methylation status

Locus	Primer Pair	Product Size (bp)	Primer Pair Sequence	PCR Conditions	# of flanked HpaII cut sites
AR	AR C:D	287	<b>C:</b> TGCATTTGCTCTCCACCT <b>D:</b> CCGAGTCTTAGCAGCTT	(95°1',56°1',72°2')X40 1.5mM MgCl <sub>2</sub>	NA
HEPH	cv f:r	512	<b>f:</b> GCTATTTGGCGAGAGGAAC <b>r:</b> CCTGAGATCCTAACCAAGGAT	(95°1',60°1',72°1')X40 1.5mM MgCl <sub>2</sub> Platinum Taq Pol'ase	NA
EDA2R	EDA2R A:B	305	<b>A:</b> AGATGTGTGCTCTGCGCTGA <b>B:</b> AATGGCTGCCAAGATCCTAG	(95°1',56°1',72°2')X35 1.5mM MgCl <sub>2</sub>	3
MSN	MSN f:r	343	<b>f:</b> GGCAAGGCCAGCGGTCGG <b>r:</b> GAGGCTCAGCCAGAGCCA	(95°1',60°1',72°1')X40 1.5mM MgCl <sub>2</sub> 4% DMSO	NA
STARD8	STAR A:B	328	<b>A:</b> AGCTTTCCCCTCCGTGG <b>B:</b> CACTGATTCAAACCTCGGCACC	(95°1',54°1',72°2')X35 1.5mM MgCl <sub>2</sub>	4
CpG 18	CG18 A:B	436	<b>A:</b> TGGACCTATGTCATCAGCGTGG <b>B:</b> TCGCTTTTCAGCCAGGCC	(95°1',60°1',72°2')X35 1.5mM MgCl <sub>2</sub>	3
CpG 19 (1)	CG19 A:B	387	<b>A:</b> ATTGTCCCATTCCCTGGTCCC <b>B:</b> GGGTGGATCACAGGGAAACC	(95°1',58°1',72°2')X35 1.5mM MgCl <sub>2</sub>	2
CpG 19 (2)	CG19b A:B	405	<b>A:</b> TAACGTTCTGGCCTTCGCTG <b>B:</b> TGGCTGTGCTCTGAGTTGGAG	(95°1',58°1',72°2')X35 1.5mM MgCl <sub>2</sub>	4
CpG 20	CG20 A:B	210	<b>A:</b> GAAGGAAGCCTCAGGGTTCTAA <b>B:</b> GAAACGCGCAGGCATACAC	(95°1',60°1',72°2')X40 1.0mM MgCl <sub>2</sub> 8% DMSO	5
CpG 22	CG22 A:B	240	<b>A:</b> TAGGGACTCCGAGCATGGG <b>B:</b> GCTCAGTTTCACCCCCAGCA	(95°1',56°1',72°2')X40 1.0mM MgCl <sub>2</sub> 2M Betaine	5

Locus	Primer Pair	Product Size (bp)	Primer Pair Sequence	PCR Conditions	# of flanked HpaII cut sites
CpG 23	CG23 A:B	339	<b>A:</b> GCTGTCGACCCCAATGGCT <b>B:</b> GGAAAGTGTGTTCTGGACCAAG	(95°1',54°1',72°2')X40 1.0mM MgCl <sub>2</sub>	2
CpG 24 (2)	CG24b A:B	347	<b>A:</b> GGCAAGTCCCCGACCCCTA <b>B:</b> CGACCCCTACTGCATCCA	(95°1',54°1',72°2')X40 1.0mM MgCl <sub>2</sub> 1M Betaine	5
CpG 24 (1)	CG24c A:B	382	<b>A:</b> CCACACTATTCCAAGTCACCTGC <b>B:</b> CGTGAGTGACAGTGCAGGGA	(95°1',54°1',72°2')X35 1.0mM MgCl <sub>2</sub>	5
CpG 26	CG26 A:B	262	<b>A:</b> GCCGATCCTGTTGCCCTCT <b>B:</b> TTACCTTCCATCCTGCTGAGCT	(95°1',56°1',72°2')X40 1.5mM MgCl <sub>2</sub>	5
CpG 27 (2)	CG27 A:B	151	<b>A:</b> ACAATAGGGCTTGGCTGGC <b>B:</b> GGGACTCAGCAGCTTGTACC	(95°1',54°1',72°2')X40 1.5mM MgCl <sub>2</sub>	2
CpG 27 (5)	CG27b A:B	388	<b>A:</b> GAAATGGTTGGAGCTGGGG <b>B:</b> TGTGTAGCGAGGAGGATTGGG	(95°1',54°1',72°2')X40 1.5mM MgCl <sub>2</sub>	3
CpG 27 (6)	CG27c A:B	211	<b>A:</b> TCTGCCTTTCCGATTCAA <b>B:</b> GTGTGAAGAATTGAGGGATGAG	(95°1',50°1',72°2')X40 2.5mM MgCl <sub>2</sub>	2
CpG 27 (3)	CG27d A:B	324	<b>A:</b> CAGAGAACAACTACTGCCCCCAC <b>B:</b> GGGAAGGCACAAGAGCATGG	(95°1',60°1',72°2')X35 1.5mM MgCl <sub>2</sub>	3
CpG 27 (4)	CG27e A:B	192	<b>A:</b> CCTTCCTTCCCAGAGCTCCT <b>B:</b> GGACCAACTGCATCCTAGGC	(95°1',58°1',72°2')X35 1.5mM MgCl <sub>2</sub>	5
CpG 27 (1)	CG27f A:B	232	<b>A:</b> AAGGGAGGTTACACCAAAGGGC <b>B:</b> TCCAGCTTGATGCGAGCGT	(95°1',52°1',72°2')X40 0.5mM MgCl <sub>2</sub>	3
CpG 31	CG31 A:B	493	<b>A:</b> TTGAAGCTGGTGGAGGTATGC <b>B:</b> TAGGGTATTGGGGTCCAAGC	(95°1',54°1',72°2')X40 1.0mM MgCl <sub>2</sub> 1M Betaine Platinum Taq Pol'ase	5

Locus	Primer Pair	Product Size (bp)	Primer Pair Sequence	PCR Conditions	# of flanked HpaII cut sites
CpG 32	CG32 A:B	438	<b>A:</b> GCTCTCCTGCCGCCGTATC <b>B:</b> GCAGTTCCCTCCCCCTTTAA	(95°1',54°1',72°2')X40 0.5mM MgCl <sub>2</sub> 1M Betaine	5
CpG 34 (1)	CG34 A:B	247	<b>A:</b> GAAGGCTGAGGAGGTGTGTCT <b>B:</b> CGAGGATCAGGTGACGAGAG	(95°1',58°1',72°2')X40 1.5mM MgCl <sub>2</sub>	4
CpG 34 (3)	CG34b A:B	326	<b>A:</b> ATAAACGCCGCACCACCTC <b>B:</b> CCCCCGTAACTCACCAA <del>C</del> TAA	(95°1',54°1',72°2')X40 1.5mM MgCl <sub>2</sub>	5
CpG 34 (2)	CG34c A:B	413	<b>A:</b> TATTCAGGAAGCAGGGTCC <b>B:</b> AAGAGGC GTTGGCTGTCGC	(95°1',56°1',72°2')X40 0.5mM MgCl <sub>2</sub>	5
CpG 41	CG41 A:B	278	<b>A:</b> AAACACAGCCATCATCCGCT <b>B:</b> TGGCTCTCAGTTCACCTGGAT	(95°1',54°1',72°2')X40 1.5mM MgCl <sub>2</sub> Platinum Taq Pol'ase	5
CpG 43 (2)	CG43 A:B	347	<b>A:</b> TTTCCGTACCACGC ACTCC <b>B:</b> GGCGGAACCTAGAGGCTGA	(95°1',58°1',72°2')X40 1.5mM MgCl <sub>2</sub>	4
CpG 43 (1)	CG43b C:D	237	<b>A:</b> GTCCAGCATTCTCTCTCGGG <b>B:</b> GGATCGAACGAGTGTCTGGC	(95°1',60°1',72°2')X40 1.5mM MgCl <sub>2</sub>	4
CpG 43 (3)	CG43c A:B	274	<b>A:</b> GC GG CAGGGTTAGTGCAAA <b>B:</b> CATTCCACTCCGCAATTCC	(95°1',54°1',72°2')X35 1.5mM MgCl <sub>2</sub>	3
CpG 44 (1)	CG44 A:B	493	<b>A:</b> GGAAGAGCCCTGACTGCTGT <b>B:</b> TTAATGGCTCTGGTCCCTCC	(95°1',56°1',72°2')X40 1.0mM MgCl <sub>2</sub>	6
CpG 44 (1)	CG44 C:D	312	<b>C:</b> GTTCTCCTCGTCGGTTCCCTTC <b>D:</b> TTTCAGCCCTGAGCTCCCTTAG	(95°1',54°1',72°2')X40 1.0mM MgCl <sub>2</sub> Platinum Taq Pol'ase	5
CpG 44 (2)	CG44b A:B	416	<b>A:</b> GCCCACCGTCTGAGGATTAAA <b>B:</b> AGGGTATTGGGTACTGCGTATTG	(95°1',54°1',72°2')X35 1.5mM MgCl <sub>2</sub> 1M Betaine	5

Locus	Primer Pair	Product Size (bp)	Primer Pair Sequence	PCR Conditions	# of flanked HpaII cut sites
CpG 45 (2)	CG45b A:B	411	<b>A:</b> TGGTCATGCAAATAAAGGCG <b>B:</b> CGTAGAGAGGTAGGGGGACT	(95°1',52°1',72°2')X40 1.5mM MgCl <sub>2</sub>	3
CpG 45 (1)	CG45c A:B	233	<b>A:</b> CGCCTTCTCCGAAATCAAACTC <b>B:</b> CTGCTGCTGAGCGAGGCA	(95°1',54°1',72°2')X40 0.5mM MgCl <sub>2</sub> 1M Betaine	3
CpG 47	CG47 A:B	239	<b>A:</b> AGAACAGGCCGATGGAGGC <b>B:</b> CTTCTCTGAAAAGCGTGTGGC	(95°1',54°1',72°2')X40 1.0mM MgCl <sub>2</sub>	3
CpG 50	CG50 A:B	275	<b>A:</b> CCTCCTTCTTCCCTGGTTGC <b>B:</b> GTGGTTGAGAGGACGGCGG	(95°1',56°1',72°2')X40 1.5mM MgCl <sub>2</sub>	5
CpG 51	CG51 A:B	288	<b>A:</b> CAGAGGAGACGACGGGAC <b>B:</b> AGGGCGCGTGTCAATAC	(95°1',56°1',72°2')X40 1.5mM MgCl <sub>2</sub> 1M Betaine	2
CpG 60	CG60 A:B	389	<b>A:</b> CATTAGGTGACGCCGGCGT <b>B:</b> TCCCTTTGTCGCAGATCCCA	(95°1',60°1',72°2')X40 1.5mM MgCl <sub>2</sub> 1M Betaine	3
CpG 62 (2)	CG62 A:B	376	<b>A:</b> GAGACAAGGAAGGTTGCACAGA <b>B:</b> AGCGAGTGCCTTGTAGGAGC	(95°1',54°1',72°2')X40 1.0mM MgCl <sub>2</sub>	6
CpG 62 (1)	CG62b A:B	449	<b>A:</b> TGACCTCCTGCACCTGCG <b>B:</b> AGAACCAAGCCCTACAAGCTGGA	(95°1',54°1',72°2')X35 1.5mM MgCl <sub>2</sub>	6
CpG 71 (2)	CG71 A:B	310	<b>A:</b> GGAAGCTCATAGCCTCCGTC <b>B:</b> AGGGACATGCTGCTTGCTAG	(95°1',56°1',72°2')X40 2.0mM MgCl <sub>2</sub> 1M Betaine	3
CpG 71 (1)	CG71b C:D	303	<b>C:</b> GGGGTTTGAGCACTTCTAGG <b>D:</b> GTTACCAGTCCCCTCAGCGA	(95°1',54°1',72°2')X40 1.5mM MgCl <sub>2</sub>	2
CpG 76	CG76 A:B	252	<b>A:</b> ACGTCAAGAAGCCGATCATCC <b>B:</b> GGGGCGGGTAGTTGATT	(95°1',54°1',72°2')X40 1.5mM MgCl <sub>2</sub>	3

Locus	Primer Pair	Product Size (bp)	Primer Pair Sequence	PCR Conditions	# of flanked HpaII cut sites
CpG 79 (1)	CG79 A:B	395	<b>A:</b> AGATCCAGACAAACGGGG <b>B:</b> CCTCGAAAGCAGAGAAAGGAGA	(95°1',52°1',72°2')X40 1.0mM MgCl <sub>2</sub> 4% DMSO Platinum Taq Pol'ase	5
CpG 79 (1)	CG79 C:D	305	<b>C:</b> CTGCTCTTCTGGCTGGCAC <b>D:</b> CCCAATTCTGGCGATTCTGA	(95°1',54°1',72°2')X35 1.0mM MgCl <sub>2</sub>	3
CpG 81	CG81 A:B	458	<b>A:</b> CAATGAAGGACTCGCGGA <b>B:</b> AAGACTGTGGGGTCACCACAGA	(95°1',60°1',72°2')X35 1.5mM MgCl <sub>2</sub>	3
CpG 104 and CpG 79 (2)	CG104 A:B	306	<b>A:</b> AGCAGCACCAAGGAAGAACGTC <b>B:</b> GGTTCCAGCCGAGCCTCTC	(95°1',54°1',72°2')X40 1.5mM MgCl <sub>2</sub> 1M Betaine	4
CpG 109	CG109 A:B	236	<b>A:</b> AGGCAGAGGCGAACCTCA <b>B:</b> ATGTGCAGCTCTACTCCGAGGG	(95°1',56°1',72°2')X40 1.5mM MgCl <sub>2</sub> 2M Betaine	3
CpG 109	CG109 C:D	413	<b>C:</b> GGATGAACGACAGCGCCAGT <b>D:</b> TCAGCCGTTACGACCTCTGA	(95°1',54°1',72°2')X40 1.5mM MgCl <sub>2</sub> 1M Betaine	3
CpG 128	CG128 A:B	385	<b>A:</b> AAAGCGGGTTCCAAGGAGA <b>B:</b> ATATGCCTGCCAATCAGGACG	(95°1',54°1',72°2')X40 1.5mM MgCl <sub>2</sub> 4% DMSO	5
CpG 145	CG145 A:B	452	<b>A:</b> AGGAAGAAGGGGCTCGGAA <b>B:</b> TCCTCTCGGGTTTCCCC	(95°1',50°1',72°2')X40 1.5mM MgCl <sub>2</sub> 2M Betaine	8
CpG 154	CG154 A:B	218	<b>A:</b> CTAAGAGAGCAGGGTTGGCCT <b>B:</b> CCAGTCCCATTAGTTGCGC	(95°1',56°1',72°2')X40 1.0mM MgCl <sub>2</sub>	3

<b>Locus</b>	<b>Primer Pair</b>	<b>Product Size (bp)</b>	<b>Primer Pair Sequence</b>	<b>PCR Conditions</b>	<b># of flanked HpaII cut sites</b>
CpG 186	CG186 A:B	356	<b>A:</b> GACAAACCTTGC GCC GACTC <b>B:</b> TTGGTGCAAGCTGCGTCC	(95°1',56°1',72°2')X40 1.5mM MgCl <sub>2</sub> 2M Betaine	8

polyacrylamide gels. 10 µl of the PCR product combined with 6x loading dye were run per lane. The loading dye consisted of bromophenol blue, xylene cyanol and 40% sucrose. All gels were stained using ethidium bromide, an intercalating agent, and viewed using a transilluminator emitting UV light at 302nm.

#### ***2.2.5.1) Agarose Gels***

All primers, except for one, amplified sequences that were within the length range required for visualization on 2% gels. All gels were made with 1X TAE (0.4 M Tris/HCl, 0.013 M NaOAc and 0.002 M EDTA, pH 8.0) buffer and standard high-melting-temperature agarose. The gels were run in 1X TAE buffer.

#### ***2.2.5.2) Polyacrylamide Gels***

The primers used to look at gene expression of TMEM28 amplified a gDNA product of 522 bp and a cDNA product of 116 bp. To visualize the cDNA product, the PCR reactions needed to be run out on a 12% polyacrylamide gel. The gel consisted of 12% acrylamide, 1X TBE (0.045 M Tris, 0.175 M Boric Acid, and 0.002 M EDTA), 0.07% ammonium persulfate and TEMED (a polymerization catalyst added just prior to pouring the gel). The gels were run in 1X TBE buffer.

### **2.3) Gene Expression**

The gene expression status of ten different genes was tested in two male lymphoma cell lines, SUDHL3 and DoHH2, two control lymphoblast cell lines, one female and one male, and a human pluripotent embryonal carcinoma cell line (control for the primer PCR conditions in the cases where the gene is not expressed in any of the lymphoblast cell lines). The expression status was determined using reverse transcriptase (RT) PCR. The techniques used to isolate RNA, remove DNA contamination from the isolated RNA, and make cDNA from the RNA templates are described in more detail below.

### **2.3.1) RNA Extraction**

Two different protocols were used to extract RNA. The majority of the RNA was extracted using the acid-guanidinium-phenol-chloroform method and a few extractions were done using QIAGEN's RNeasy Mini Kit. Both techniques are described below.

#### ***2.3.1.1) Acid-Guanidinium-Phenol-Chloroform RNA Extraction***

RNA was extracted from frozen cell pellets using an acid-guanidinium-phenol-chloroform RNA extraction protocol. The cells were first dissolved in 5 ml of solution D (4M guanidinium thiocyanate, 25mM sodium citrate (pH 7), 0.5% sarcosyl, and 0.1M 2-mercaptoethanol) and vortexed to disrupt the cell membranes. Guanidinium thiocyanate is a strong denaturant that is able to disrupt the cells, solubilize the components and denature RNases. The guanidinium, along with water, forms a complex with the RNA, preventing any hydrophilic interactions with DNA and protein. After the addition of chloroform, RNA remains in the aqueous layer in the phenol: chloroform extraction, whereas the DNA and protein are found in the interphase. An equal volume of diethyl pyrocarbonate (DEPC)-treated water saturated phenol and 1/10<sup>th</sup> volume of 2M sodium acetate were added to the cells. The cells were vortexed, and 2 ml of chloroform was added. The cells were vortexed again and placed on ice for 5-15 minutes. They were then centrifuged for 10 minutes at 12500 rpm, after which the upper aqueous layer containing the RNA was removed to a new tube. An equal volume of isopropanol was then added and the extraction stored overnight at -20°C . The following day, the extraction was centrifuged for 10 minutes at 12500 rpm. The precipitate was rinsed with 70% ethanol centrifuged for 10 minutes at 12500 rpm. All traces of ethanol were removed and the pellet was re-dissolved in 10-100 µl TE.

#### ***2.3.1.2) QIAGEN's RNeasy Mini Kit***

The lymphoma patient RNA we received from our colleagues, Joe Connors, Randy Gascoyne, Doug Horsman, and Ron deLeeuw, from the BC Cancer Research Centre was extracted using the RNeasy Mini Kit from QIAGEN. To standardize the methods, this kit was also used to extract RNA from four cell lines, two control lymphoblasts (GM7059 and GM7009) and two follicular lymphoma (SUDHL3 and

DoHH2). The kit uses a silica-gel based membrane and centrifugation. This RNA was used to quantitate the expression of STARD8, a candidate tumour suppressor gene.

### ***2.3.2) DNase Treatment of RNA***

To remove all traces of DNA, the RNA was first treated with RNase-free DNase, followed by a phenol: chloroform clean up to remove proteins and buffer. 1/20<sup>th</sup> (of the total volume of RNA) porcine RNase inhibitor (RNasin) and 1/10<sup>th</sup> volume RNase-free DNase were added to the extracted RNA and incubated for an hour at 37°C. After the incubation, the volume was brought up to 200 µl with DEPC-treated water for ease of handling. An equal volume of 1:1 phenol: chloroform was then added, and the RNA extraction was vortexed for one minute and placed on ice for another minute. Next, the RNA was centrifuged for 5-10 minutes at 4°C at 12500 rpm. Afterwards, the aqueous phase was transferred to a new eppendorf tube and an equal volume of chloroform was added. The RNA was once again vortexed, and centrifuged at 12500 rpm for 5-10 minutes at 4°C. The aqueous layer was removed to a new eppendorf tube, but this time 0.15 volume 2M NaOAc and an equal volume of isopropanol were added. The RNA was gently mixed and incubated overnight at -20°C. The following day, the RNA extraction was centrifuged for 10-15 minutes at 12500 rpm at 4°C. The supernatant was removed and the pellet allowed to air dry to remove traces of the isopropanol. Once dry, the pellet was resuspended in 10-100 µl DEPC-treated water.

### ***2.3.3) Reverse Transcriptase PCR (RT-PCR)***

cDNA was made from RNA templates to assess the gene expression in the cell lines and patient samples. Each reverse transcriptase reaction consisted of 5µg RNA, 1X first-strand buffer, 0.01M Dithiothreitol (DTT), 0.0625 mM dNTPs, 1 µl random hexamers, 1 U RNasin, and 1 U M-MLV reverse transcriptase. The volume of the reaction was brought up to 20 µl with DEPC-treated water. For each sample, both a positive RT reaction and a negative RT (consisting of all the same reagents as the RT, except the reverse transcriptase) reaction was made. All the reactions were left to sit at room temperature for five minutes, after which they were incubated for two hours at 42°C and at 95°C for another five minutes. The cDNA was stored at -20°C.

#### **2.3.4) Primer Design**

The features of the methylation and expression primers were very similar with the optimal length and melting temperature of both being 20 bp and 58-60°C respectively. The sequence of the genes were obtained from the May 2004 freeze of the human genome assembly found on the University of California Santa Cruz (UCSC) genome browser (<http://genome.ucsc.edu/>). Where possible, primers were designed to amplify PCR products that differed in length depending on the template, either DNA or cDNA. This was done by designing primers that flank introns. In this way, it is possible to differentiate between the presence of a band as a result of positive gene expression and DNA contamination. The primers, with the aid of the Applied Biosystems Primer Express Version 1.5 program, were designed to have the fewest number of potential hairpin structures and primer dimers. The optimal length of the cDNA product was between 200-400 bp. The specificity of the primer sequences were checked using the National Center for Biotechnology Information (NCBI) BLAST (<http://www.ncbi.nlm.nih.gov/BLAST/>) program. All primers were ordered from either the Nucleic Acid Protein Service (NAPS) unit or Integrated DNA Technologies (IDT), Inc..

##### **2.3.4.1) Expression Primer Table**

Table 2.3 is a list of all the primers and their PCR conditions used to test the gene expression status of ten different genes located in the 8 Mb region. Five out of the ten genes are located in the 1.4 Mb sub-region of interest and three are found in the second sub-region.

**Table 2.3** Table of primers used to check gene expression

Gene	Primer Pair	Primer Pair Sequence	gDNA Product Size (bp)	cDNA Product Size (bp)	PCR Conditions
HEPH	cv 3:4	<b>3:</b> GCTCCTGGGTTCCAGATAC <b>4:</b> CCAGTGGCCAGACAGTAGT	~200	1050	(94°1', 60°1', 72°1') x 40 1.5mM MgCl <sub>2</sub>
EDA2R	XDR f:r	<b>f:</b> GATTGTGGTTATGGAGAGGGTGG <b>r:</b> TGCACGGGATGCACTCTTG	235	776	(94°1', 54°1', 72°1') x 35 1.5mM MgCl <sub>2</sub>
AR	AR f:r	<b>f:</b> CTCCTTGAGCCTGCTCT <b>r:</b> CAGATCAGGGCGAAGTAGAG	220		(94°1', 54°1', 72°1') x 40 1.5mM MgCl <sub>2</sub>
MGC21416	MGC f:r	<b>f:</b> TTGTCTGGTTGGTGCAGTTAC <b>r:</b> AGCCGAACCATGAAGTTACA	188	1440	(94°1', 56°1', 72°1') x 40 2.5mM MgCl <sub>2</sub>
STARD8	STAR f:r	<b>f:</b> GATCAAGAGCAAACGCAGCCT <b>r:</b> TTGAAGCGCTCAGCAGCATC	291	548	(94°1', 56°1', 72°1') x 40 1.5mM MgCl <sub>2</sub>
STARD8	STARb f:r	<b>f:</b> GCTGGACTCCAGGCATCAAT <b>r:</b> ATGAACCTGGGCATTGACCAG	266		(94°1', 58°1', 72°1') x 35 1.5mM MgCl <sub>2</sub>
PJA1	PJA1 f:r	<b>f:</b> CCGAGCCAAAGTACCCCTGAAG <b>r:</b> GGTTTGCTCTCGGCTTCGA	206	206	(94°1', 54°1', 72°1') x 35 1.5mM MgCl <sub>2</sub>
TMEM28	TED f:r	<b>f:</b> CCCCGACAATGAGGAAATGG <b>r:</b> AGGAGACCCACTGCACGTCA	116	522	(94°1', 54°1', 72°1') x 35 1.5mM MgCl <sub>2</sub>
DGAT2L4	DG2L4 A:B	<b>A:</b> CCTTACATACTCACACTGGGAGCC <b>B:</b> AAGAACCTGGCAGGCTGTATCT	187	421	(94°1', 54°1', 72°1') x 35 1.5mM MgCl <sub>2</sub>
OTUD6A	HIN6 f:r	<b>f:</b> ATCTTCCAGGCTGAGATGTCGG <b>r:</b> AGTCGCTGGTCTCGGGGTT	262	262	(94°1', 54°1', 72°1') x 30 1.5mM MgCl <sub>2</sub>
IGBP1	IGBP1 f:r	<b>f:</b> AAACCGTGGGAGTGGTGC <b>r:</b> CTTGGGGAGAGAGGAAACCCG	290	472	(94°1', 54°1', 72°1') x 40 1.5mM MgCl <sub>2</sub>
β-ACTIN	ACTIN 1:2	<b>1:</b> ATGATATGCCCGCGCTCG <b>2:</b> CGCTCGGTGAGGATCTTCA	580	580	(94°1', 54°1', 72°1') x 30 1.5mM MgCl <sub>2</sub>

<b>Gene</b>	<b>Primer Pair</b>	<b>Primer Pair Sequence</b>	<b>gDNA Product Size (bp)</b>	<b>cDNA Product Size (bp)</b>	<b>PCR Conditions</b>
G6PDH*	G6PDH f:r	<b>f:</b> CCATGACCACTTCTCAGCCC <b>r:</b> CTTCAGCATCCACGGTCTTT	215		(94°1', 54°1', 72°1') x 35 1.5mM MgCl <sub>2</sub>

\* used for qPCR to normalize for amount of RNA present in each sample

## **2.4) Quantifying Gene Expression**

The expression of a candidate tumour suppressor gene, STARD8, was quantitated. Initial RT-PCR results showed that this gene was expressed in lymphoblast control cell lines, but not in the two male follicular lymphoma cell lines. The level of STARD8 RNA was quantitated in these cell lines and normalized to G6PDH expression to see if the change in expression levels was significant.

### ***2.4.1) Quantitative PCR (qPCR)***

The dye SYBR Green 1 from Sigma was used to quantitate levels of cDNA in each sample. SYBR Green 1 binds and brightly fluoresces when bound to double stranded DNA and it is unable to bind to single stranded DNA. The qPCR reactions were prepared in the same way as regular PCR reactions (described above) with the addition of 1X SYBR Green 1. All samples were run in triplicate. The optimal primers for qPCR amplify a product between 200-400bp (the smaller the better), produce a single band and have minimal primer dimers since the quantitation method can not discriminate between the different bands. The DNA Engine Opticon Monitor qPCR machine from MJ Research and Opticon Monitor Software 2.02 were used to quantitate STARD8 expression.

The standard curve method was used to quantitate the levels of mRNA. In this method, a genomic DNA dilution series of known amounts was run at the same time as the samples. The software created a standard curve of the DNA dilution series, plotting C(T) cycle against log quantities, which was then used to calculate the quantities of cDNA and therefore mRNA in each of the samples. The quantitation of STARD8 expression was corrected for the amount of total RNA in each sample by normalizing to the housekeeping gene G6PDH expression levels.

The primers used to quantitate STARD8 gene expression produced primer dimers in samples lacking an amplified product (blanks, -RTs and RTs lacking gene expression). To account for this background expression, the difference between the average cDNA quantity in all the samples and the average quantity of STARD8 cDNA in the blanks (which in reality was zero, but the software calculated a value that actually reflected the amount of primer dimer in the sample instead of the quantity of STARD8 cDNA) was taken. This results in slightly lower expression levels for samples expressing the gene

since the amount of primer dimer is lower in these samples. The new standard of deviation of the corrected STARD8 quantitation was calculated by taking the difference between the top and bottom range of ng STARD8 cDNA and the calculated ng cDNA found in the blank (*i.e.* (average ng STARD8 + standard of deviation) – (average ng blank + standard deviation) and (average ng STARD8 – standard of deviation) – (blank – standard deviation)). A similar calculation was done to determine the standard deviation of normalized STARD8 expression in each sample, but instead of taking the difference, the quotient was taken.

The VassarStats: website for statistical computation (<http://faculty.vassar.edu/lowry/VassarStats.html>) was used to determine whether the change in gene expression levels between the patient samples and the male control lymphoblast cell line was significant. Using this website, I performed a one-tailed t-test on the normalized quantities of STARD8 gene expression in the lymphoma patients and the lymphoblast control cell line.

#### **2.4.2) Gel Extraction**

The primers used to determine the gene expression status and mRNA quantity of STARD8 and G6PDH were cDNA specific. This meant that genomic DNA could not be used to create a standard curve since no genomic product could be amplified with these primers. Instead, cDNA was PCR amplified with the primer pairs, run on a gel, extracted from the gel and then used to make a 10X dilution series. Qiagen QIAquick Gel Extraction Kit was used to extract the PCR product from the gel.

## **Chapter 3: CpG Island Methylation Pattern and Gene Expression on the X Chromosome in the Vicinity of Xq11.2-q12 in Lymphoma**

### **3.1) Introduction**

The acquisition of an extra X chromosome is a common occurrence in lymphoma and in 2000, McDonald *et al.* investigated the role the X chromosome plays in the neoplastic development of lymphoma [4]. To determine the activity state of the acquired X chromosome, McDonald *et al.* (2000) used several assays including examining the methylation status of the (CAG)<sub>n</sub> repeat polymorphism located in the androgen receptor (AR) at Xq11.2-q12. They observed aberrant hypermethylation of AR in 84% of the examined patient samples. Abnormal hypermethylation has been associated with the silencing of tumour suppressor genes in malignancies (reviewed in [17]). In addition, methylation is able to spread several Mb from its origin, perhaps in cancer due to the breakdown of boundaries separating hetero- and euchromatin (reviewed in [9, 111]). This led to the hypothesis of the existence of a candidate tumour suppressor gene in the vicinity of AR. Mutations involving AR result in androgen insensitivity syndrome and mental retardation if the nearby oligophrenin 1 (OPHN1) gene is also involved, but not in a known predisposition to lymphoma (reviewed in [4]). In addition, both AR and OPHN1 are not normally expressed in lymphoblast cells and therefore, both genes make poor candidate tumour suppressor genes. In order to identify possible candidate tumour suppressor gene(s) and a region of focus, the extent of the abnormal methylation spread first needed to be determined, as described in part A. To follow in part B, the expression status of some of the genes located in the abnormally hypermethylated region was examined.

### **3.2) Results – A.) Methylation**

#### **3.2.1) CpG Islands and Associated Genes**

The status of almost all of the CpG islands located in an 8 Mb interval spanning 4.2 Mb upstream and 3.7 Mb downstream of AR was examined using a methylation sensitive restriction enzyme digestion assay. Table 3.1 lists all the CpG islands located in the examined region as well as their associated genes. There are 46 CpG islands in this

**Table 3.1** CpG islands and their associated genes in an 8Mb region on the X chromosome located in the vicinity of AR (Xq11.2-12). CpG islands with no listed associated gene indicates that no known associated gene has been identified.

CpG Island	Location on the X Chromosome	Associated Gene
CpG 79 (1)	57,501,424-57,502,349	within 5' ZXDB
CpG 104	57,818,700-57,819,954	within 5' ZXDA
CpG 51	62,353,974-62,354,447	within 5' LOC139886 <sup>1</sup>
CpG 43 (1)	62,757,735-62,758,239	within 5' ARHGEF9
CpG 16	63,046,926-63,047,150	
CpG 128	63,207,220-63,209,211	within 5' RP11-403E24.2 <sup>1</sup>
CpG 34 (1)	63,397,979-63,398,444	within 5' MTMR8
CpG 71 (1)	64,037,430-64,038,312	
CpG 145	64,409,765-64,411,066	
CpG 43 (2)	64,537,338-64,537,826	within 5' FLJ12525/LASIL
CpG 62 (1)	64,554,428-64,555,214	within 5' end of FKSG43 <sup>2</sup>
CpG 23	64,670,571-64,670,858	within 5' to MSN
CpG 45 (1)	64,824,918-64,825,325	
CpG 34 (2)	66,546,706-66,547,098	5' to AR
CpG 27 (1)	66,549,059-66,549,300	within 5' AR
CpG 27 (2)	67,135,672-67,135,944	within 3' OPHN1
CpG 71 (2)	67,435,968-67,436,928	within 5' OPHN1
CpG 60	67,501,449-67,502,092	within 5' Y1PF6
CpG 41	67,689,133-67,689,496	middle of STARD8
CpG 79 (2)	67,696,321-67,697,155	mid to 3' end of STARD8
CpG 30	67,786,275-67,786,634	5' to SERBP1P <sup>2</sup>
CpG 154	67,831,681-67,833,273	within 5' EFNB1
CpG 27 (3)	67,843,200-67,843,579	within 3' EFNB1
CpG 20	67,897,462-67,897,731	
CpG 31	68,131,421-68,131,796	
CpG 22	68,152,840-68,153,106	
CpG 47	68,167,935-68,168,425	5' to PJA1
CpG 24 (1)	68,224,407-68,224,715	
CpG 27 (4)	68,308,572-68,308,912	
CpG 186	68,506,753-68,509,072	within 5' TMEM28
CpG 44 (1)	68,541,213-68,541,668	
CpG 109	68,618,670-68,620,084	within 5' EDA
CpG 76	69,065,406-69,066,306	within 5' OTUD6A
CpG 62 (2)	69,136,126-69,136,925	within 5' IGBP1
CpG 34 (3)	69,292,666-69,292,983	within 5' of both KIF4A and PDZDII
CpG 81	69,436,729-69,437,682	3' to GPD2
CpG 24 (2)	69,448,124-69,448,376	within 5' DLG3
CpG 45 (2)	69,457,821-69,458,326	within 5' DLG3

<b>CpG Island</b>	<b>Location on the X Chromosome</b>	<b>Associated Gene</b>
CpG 27 (5)	69,933,697-69,934,023	within 5' SLC7A3
CpG 19 (1)	70,056,031-70,056,276	within SNX12
CpG 44 (2)	70,070,912-70,071,380	within 5' SNX12
CpG 32	70,098,518-70,098,874	5' to MLLT7
CpG 26	70,099,371-70,099,692	within 5' MLLT7
CpG 50	70,121,302-70,121,887	within 5' MED12
CpG 19 (2)	70,172,557-70,172,821	within 3' NLGN3
CpG 27 (6)	70,184,884-70,185,234	
CpG 18	70,227,068-70,227,291	middle of GJB1
CpG 43 (3)	70,256,834-70,257,308	within 5' ZMYM3

<sup>1</sup>predicted gene

<sup>2</sup>pseudogene

region according to the University of California, Santa Cruz (UCSC) Genome Browser. The names of the islands are based on the number of CpG dinucleotides located within the CpG island. Unfortunately there is much ambiguity in using such a nomenclature system and to lessen the confusion about which CpG island is being discussed, I have numbered the island names that appear more than once based on their sequential order on the chromosome.

The majority of the literature states that CpG islands are located in the promoter and first exon region of genes [8]. However, looking at these 46 CpG islands, only 27 (59%) are associated with the 5' region of known or predicted genes. Table 3.2 breaks down the CpG islands according to their location with respect to the genes in the region of interest. Even though only 59% of the CpG islands are associated with the promoter and first exon region, 17.4% are associated with other regions of genes, and 13% are associated with ESTs. These numbers might be biased since CpG islands are used to identify predicted genes [23]. Overall, the majority of the CpG islands appear to somehow be associated with a gene, and could potentially play a role in their regulation. Figure 3.1 shows the distribution of the CpG islands with respect to genes along the X chromosome. The figure is not drawn to scale, but shows the relative location of the CpG islands in the 8 Mb region.

One hypothesis states that methylation evolved as a defense mechanism against repetitive elements. These repetitive elements pose a threat since they could potentially interfere with normal gene function by integrating into the host DNA and directly disrupting the coding region, introns and splice sites or regulatory sequences [39]. Methylation of these elements is believed to be capable of spreading to nearby CpG dinucleotides (reviewed in [112]). Table 3.3 lists the location of repetitive elements, like LINEs, SINEs and LTRs, relative to the CpG islands found in the vicinity of AR. 54% of the CpG islands are associated with at least some kind of repetitive element within a 500 bp window, and this number increases to 78% if the distance is extended to 1000 bp. The majority of the repetitive elements are SINEs, accounting for 64% of the elements. After SINEs, the next most common elements are LINEs at 30%, followed by LTRs at a mere 6%. The most frequent type of SINE associated with the CpG islands in this region are

**Table 3.2** Location of CpG islands in an 8Mb region on Xq11.2-12 relative to genes.

CpG islands associated with genes < 500 bp away	CpG islands associated with genes btwn 1000-500 bp away	CpG islands not associated with genes	CpG islands associated with gene but not 5'
CpG 79 (1)	CpG 34 (2)	CpG16 <sup>1</sup>	CpG 27 (2) <sup>2</sup>
CpG 104	CpG30 <sup>6</sup>	CpG 71 (1) <sup>1</sup>	CpG 41 <sup>8</sup>
CpG 51 <sup>3</sup>		CpG 45 (1) <sup>4</sup>	CpG 79 (2) <sup>5</sup>
CpG 43 (1)		CpG 20 <sup>1</sup>	CpG 27 (3) <sup>2</sup>
CpG 128 <sup>3</sup>		CpG 31	CpG 81 <sup>7</sup>
CpG 34 (1)		CpG 22	CpG 19 (1) <sup>11</sup>
CpG 43 (2)		CpG 24 (1) <sup>9</sup>	CpG 19 (2) <sup>2</sup>
CpG 62 (1) <sup>12</sup>		CpG 27 (4)	CpG 18 <sup>8</sup>
CpG 23		CpG 44 (1)	
CpG 27 (1)		CpG 27 (6) <sup>10</sup>	
CpG 71 (2)		CpG 145	
CpG 60 <sup>3</sup>			
CpG 154			
CpG 47			
CpG 186			
CpG 109			
CpG 76			
CpG 62 (2)			
CpG 34 (3)			
CpG 24 (2)			
CpG 45 (2)			
CpG 27 (5)			
CpG 44 (2)			
CpG 32			
CpG 26			
CpG 50			
CpG 43 (3)			

<sup>1</sup>nearby EST

<sup>2</sup>within 3' end

<sup>3</sup>associated with predicted gene

<sup>4</sup>nearby EST and segmental duplication

<sup>5</sup>middle to 3' end of gene

<sup>6</sup>5' to predicted gene and segmental duplication

<sup>7</sup>downstream of gene but closer to an EST

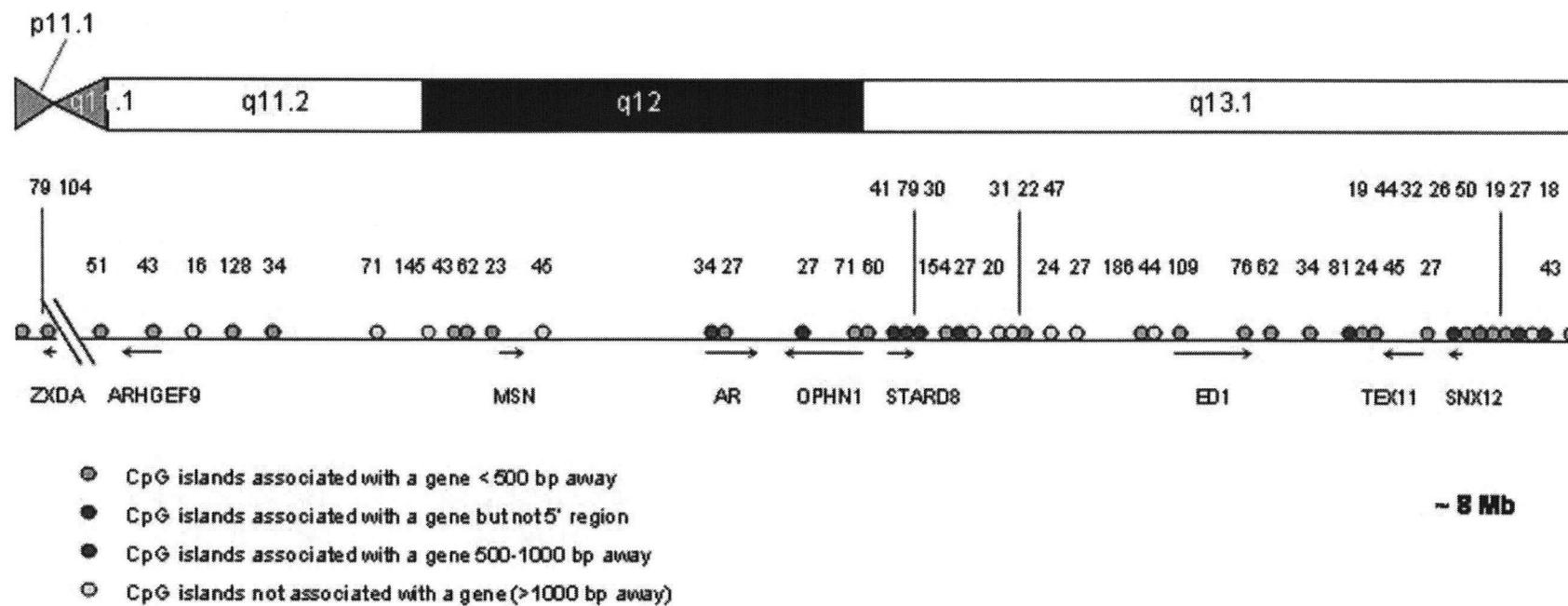
<sup>8</sup>middle of gene

<sup>9</sup>nearby Ensemble predicted gene

<sup>10</sup>within EST

<sup>11</sup>middle to 5' end of gene

<sup>12</sup>within pseudogene



**Figure 3.1** Distribution of CpG island gene association.

The map shows the relative location of CpG islands and their gene association in an 8 Mb region on the X chromosome in the vicinity of the AR. CpG islands 79 and 104 are located on Xp and are not included in the 8 Mb region of interest, even though their methylation status was examined. Below the CpG islands, the location of a few genes are noted. These are not all the genes in the region, and are only meant to provide reference of the approximate location on the chromosome. The map is only relatively drawn to scale due to the large region it covers and the small size of each CpG island.

**Table 3.3** CpG islands located in the vicinity of AR on the X chromosome and their associated repetitive elements.

CpG Island	Repetitive element < 500 bp away	Repetitive element btwn 1000-500 bp away
CpG 79 (1)	LINE (L1)	
CpG 104	LINE (L1)	
CpG 51	LINE (L1)	LINE (L1) SINE (MIR)
CpG 43 (1)		SINE (MIR) LINE (L1)
CpG 16	2 x SINE (Alu)	2 x LINE (L1)
CpG 128	SINE (MIR)	SINE (MIR)
CpG 34 (1)	2 x SINE (MIR)	SINE (MIR) LTR (MaLR) LINE (L2)
CpG 71 (1)	LINE (L2) SINE (MIR) SINE (Alu)	SINE (MIR)
CpG 145	LTR (ERV1) LINE (L1)	
CpG 43 (2)	SINE (Alu)	SINE (MIR)
CpG 62 (1)	LTR (MaLR)	SINE (Alu)
CpG 23	LINE (CR1)	
CpG 45 (1)	LINE (L1)	LINE (CR1) LTR (MaLR)
CpG 34 (2)		
CpG 27 (1)		
CpG 27 (2)	LINE (L1)	2 x LINE (L1) SINE (Alu)
CpG 71		LTR (MaLR)
CpG 60	SINE (Alu)	SINE (MIR) SINE (Alu)
CpG 41		3 x SINE (MIR) SINE (Alu)
CpG 79 (2)	SINE (Alu)	SINE (MIR) SINE (Alu)
CpG 30	SINE (Alu) 3 x LINE (L1)	SINE (Alu) SINE (MIR)
CpG 154		
CpG 27 (3)		SINE (MIR)
CpG 20		SINE (MIR)
CpG 31	SINE (MIR)	2x SINE (MIR)
CpG 22	SINE (MIR) SINE (Alu)	SINE (MIR) LINE (L2)

CpG Island	Repetitive element < 500 bp away	Repetitive element btwn 1000-500 bp away
CpG 47		
CpG 24 (1)	SINE (MIR) SINE (Alu)	SINE (Alu) SINE (MIR)
CpG 27 (4)	3 x SINE (MIR)	LINE (L1) 2 x SINE (MIR)
CpG 186	SINE (MIR)	SINE (MIR) LINE (L2)
CpG 44 (1)	2 x LINE (L2)	
CpG 109		
CpG 76	SINE (Alu)	2 x SINE (Alu) LINE (L1)
CpG 62 (2)		SINE (MIR)
CpG 34 (3)	SINE (MIR)	LINE (L2) SINE (Alu)
CpG 81		
CpG 24 (2)		
CpG 45 (2)		
CpG 27 (5)		2 x SINE (MIR)
CpG 19 (1)		2 x LINE (L1) 2 x SINE (Alu)
CpG 44 (2)		LINE (L2)
CpG 32		SINE (Alu) LINE (L2)
CpG 26	LINE (L2)	
CpG 50	SINE (MIR)	SINE (MIR) SINE (Alu)
CpG 19 (2)	LTR (ERV1)	SINE (Alu)
CpG 27 (6)		2 x SINE (Alu)
CpG 18		
CpG 43 (3)		

MIRs, comprising 37% of the total number of associated repetitive elements and 59% of the SINEs. In general the X chromosome is relatively rich in repetitive sequences, accounting for 56% of the sequence compared to the genome average of 45% [113]. Of these repetitive elements, Alu repeats are found at levels below average, LTRs at approximately average levels and LINE L1 elements at above average levels. LINE L1 elements account for 29% of the X chromosome sequence while the genome average is only 17%. MIR elements are found at approximately average levels with a coverage of 1.8% of the X chromosome compared to an average of 2.2% of the genome.

A recent study looked at the distribution of repetitive elements with respect to genes with or without CpG islands, and the length of the CpG islands with respect to the distribution of the repetitive elements [112]. Genes associated with CpG islands have a greater number of Alu elements in a 5' flanking 10 kb region relative to genes lacking CpG islands. L1 elements, however, showed the opposite trend, being more common in the 10 kb upstream region of genes lacking CpG islands. The LTR distribution was similar between the two groups of genes. On the other hand, the density of all three of these retroelement types was found to be the lowest within a 1 kb region upstream of the transcription start site of CpG island associated genes. The levels of all three elements rose steadily and in the case of Alus and LTRs, plateaued in a region 2-3 kb upstream of the transcription start, while numbers of L1 elements continued to increase. The plateau levels of L1 elements were higher for the genes lacking CpG islands while the Alu plateau levels were higher for the genes with CpG islands.

### ***3.2.2) Methylation in an 8 Mb Region Around Xq11.2-q12***

#### ***3.2.2.1) Methylation in Two Lymphoma Cell Lines***

Two follicular lymphoma cell lines, SUDHL3 and DoHH2, were used to establish the extent of the aberrant hypermethylation. It is important to note that both of these cell lines are male, since normal and abnormal methylation can not be distinguished in females when using the methylation restriction enzyme methylation assay. In females, it is expected to always observe methylation since the region of interest is located on the X chromosome and methylation is involved in stabilizing the inactive state of one of the X

chromosomes. In addition to the two lymphoma cell lines, the methylation status of the CpG islands was also examined in a female and male lymphoblast cell line as controls for normal methylation.

The results of the methylation analysis for almost all the CpG islands in the 8 Mb region located at Xq11.2-12 are summarized in table 3.4 and in figure 3.2. The methylation status of two CpG islands located on Xp, CpG 79 and 104, was also tested to investigate whether the abnormal methylation had spread across the centromere. These islands are located in a large tandem segmental duplication and one primer pair designed from the sequence of CpG 104 was used to infer the methylation status of both CpG islands. Amplification of the digested lymphoma cell line DNA with this primer pair showed a lack of methylation, indicating methylation had not spread across the centromere into the p arm of the X chromosome. This also meant that the primer pair could be used to look at the methylation status of both islands. If the assay indicated that the region is methylated, then the assumption that both islands are methylated is not valid since the methylation of only one of the islands is sufficient to obtain a positive result.

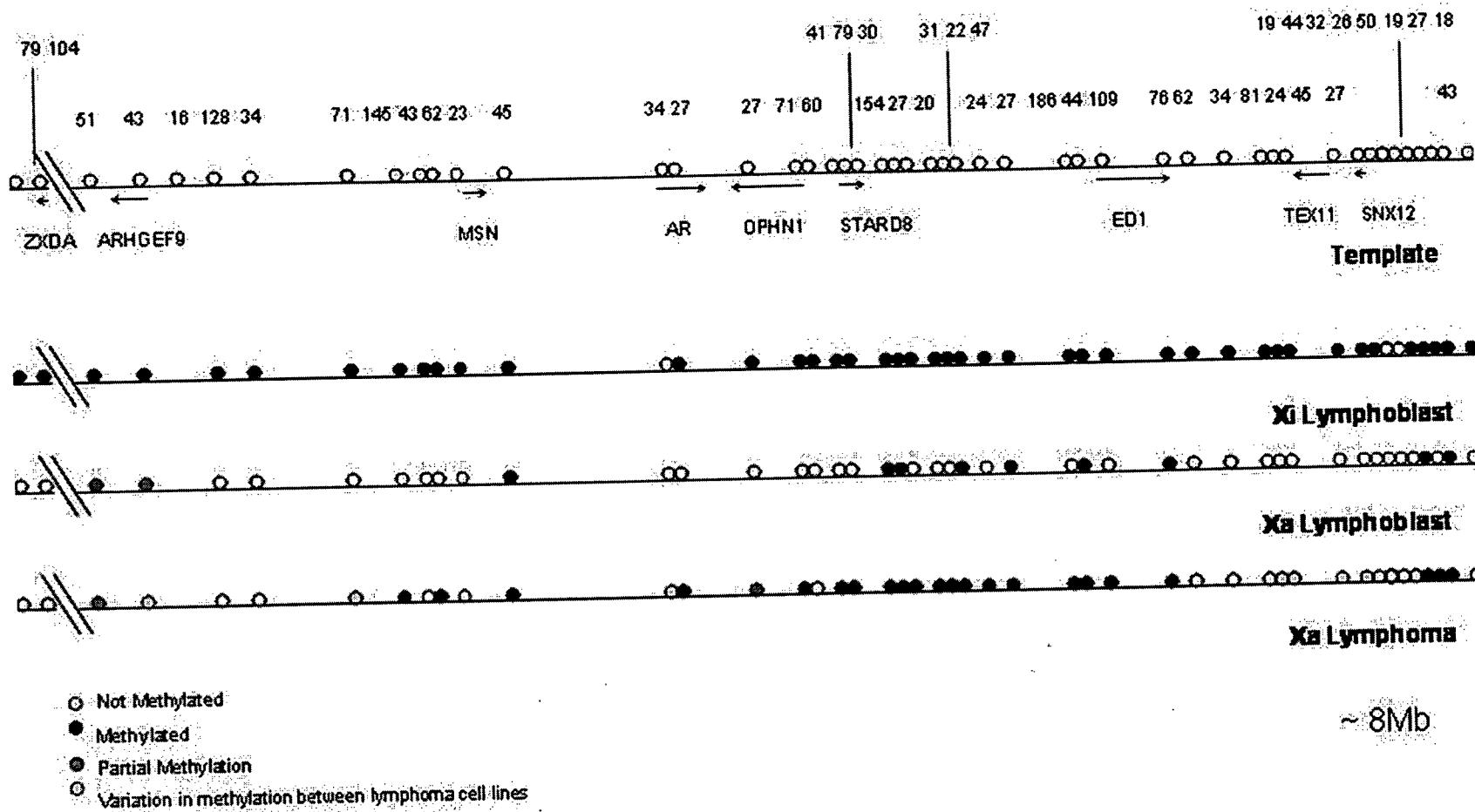
The methylation status of two CpG islands, CpG 16 and 30, located within the 8 Mb region was not examined. CpG 16 only contains one HpaII recognition sequence (CCGG) and analysis performed with primers flanking this site, might not reflect the true methylation status of the CpG island. The majority of the primers were designed to flank at least three HpaII cut sites due to the assay's reliance on complete cutting. There is a higher chance that at least one of the three HpaII recognition sites are cut if they are not methylated relative to a single site. Therefore, there is a greater chance that primers flanking the single HpaII cut site in CpG 16 might indicate that the island is methylated when in fact, there was incomplete cutting. CpG 16 is not located in a region with a high degree of aberrant methylation and therefore was not further examined. In the future, for a more complete picture of the methylation status of all the CpG islands located in this region, DNA can be digested with a different methylation sensitive restriction enzyme, for example, with AciI which has 17 recognition sites within CpG 16.

Difficulty with primer design is also the reason why the methylation status of CpG 30 has not yet been determined in the lymphoma cell lines. In this case, like with

**Table 3.4** Methylation status of CpG islands located in an 8 Mb region in the vicinity of AR (Xq11.2-12) in two control and two male lymphoma cell lines. The colours represent different CpG island methylation status. Black represents methylation, white represents no methylation and red means partial methylation (presence of a band in the digested sample that is weaker than the mock digested sample of the same cell line). Boxes filled in with diagonal bars are CpG islands whose methylation status was not investigated. CpG islands are listed in linear order.

CpG Island	Methylation Status			
	Control Female	Control Male	Lymphoma SUDHL2	Lymphoma DoHH2
79 (1)				
104				
51				
43 (1)				
16				
128				
34 (1)				
71 (1)				
145				
43 (2)				
62 (1)				
23				
45 (1)				
34 (2)				
27 (1)				
27 (2)				
71 (2)				
60				
41				
79 (2)				
30				
154				
27 (3)				
20				
31				
22				
47				
24 (1)				
27 (4)				
186				
44 (1)				
109				

CpG Island	Methylation Status			
	Female Control	Male Control	Lymphoma SUDHL3	Lymphoma DoHH2
76				
62 (2)				
34 (3)				
81				
24 (2)				
45 (2)				
27 (5)				
19 (1)				
44 (2)				
32				
26				
50				
19 (2)				
27 (6)				
18				
43 (3)				



**Figure 3.2** CpG island methylation results for an 8 Mb region located at Xq11.2-12.

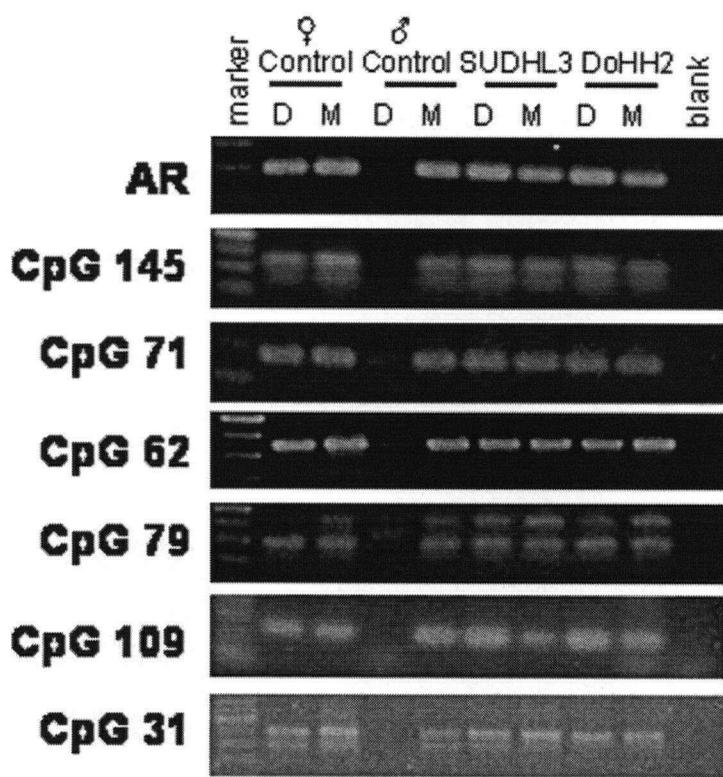
The top map represents a template of the region of interest and shows the approximate location of all the CpG islands examined as well as CpG 16 and 30. The location of a few, but not all of the genes in this region, are also shown to serve as reference points. Due to the size of the region of interest, the map is only drawn to relative scale. The two hatch marks represent the crossing of the centromere and CpG 79 and 104 are located on Xp. The second map represents the methylation assay results for the inactive X chromosome as inferred from the female control cell line and the third map represents the active X methylation results taken from the male control cell line. The fourth and last map is a summary of the methylation results of the CpG islands in the two lymphoma cell lines, SUDHL3 and DoHH2.

CpG 79 and 104, the CpG island is located within a segmental duplication. Unlike the situation above, the duplication is found on chromosome 4 and specific primers could not be designed. In the future it might be useful to examine the methylation status of CpG 30 as it is located in a region showing both normal and aberrant methylation.

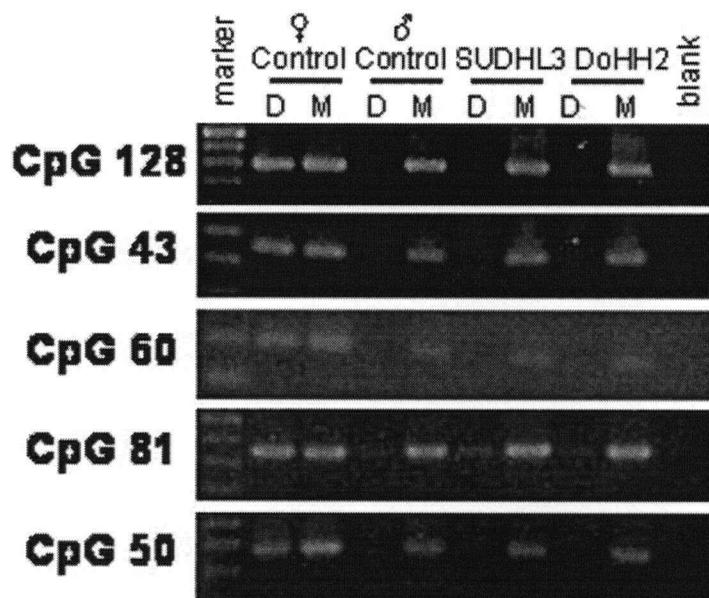
The methylation status of AR was first determined in the two lymphoma cell lines since the abnormal methylation of this locus in lymphoma patients initiated the study. Abnormal methylation is any methylation seen in the lymphoma cell lines or patient samples that is not observed in the male lymphoblast control. Both SUDHL3 and DoHH2 showed aberrant methylation of AR (as seen in figure 3.3). The methylation results of five other CpG islands that experienced abnormal hypermethylation in SUDHL3 and DoHH2 are also shown in figure 3.3. These are not the only abnormally methylated CpG islands, but serve as examples. In comparison, figure 3.4 shows examples of CpG islands that were not abnormally methylated in the lymphoma cell lines. Table 3.4 is a complete summary of the methylation results. 16 out of the 46 CpG islands in the 8 Mb region are fully or partially abnormally hypermethylated in both follicular lymphoma cell lines, and four more CpG islands are aberrantly methylated in at least one of the lymphoma cell lines. The data for the CpG islands that showed variation in methylation between the lymphoma cell lines are shown in figure 3.5. SUDHL3 and DoHH2 both appear to experience CpG island methylation skipping resulting in the lack of a definite aberrant methylation boundary. The size of the region to be assayed was determined by finding at least 3-4 CpG islands at both ends of the region that were not abnormally methylated.

There is a cluster of hypermethylated CpG islands in a smaller sub-region located within the 8 Mb region. This 1.4 Mb sub-region is located at Xq13.1, spans from CpG 41 to CpG 76 and is flanked by CpG islands with variation in methylation between SUDHL3 and DoHH2. This sub-region is discussed in more detail below.

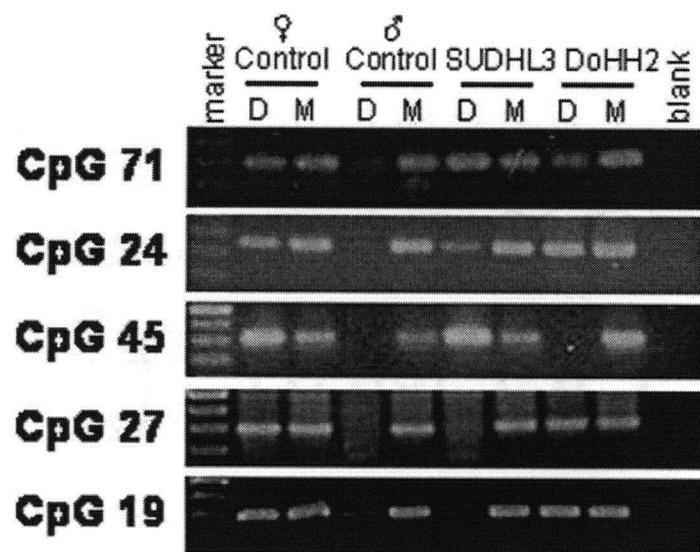
Overall the female lymphoblast control cell line showed methylation at the majority of the investigated CpG islands, which is presumably due to the inactive X chromosome. On the other hand, the male lymphoblast control cell line lacked methylation at most of the CpG islands. There were, however, some exceptions to this observation. Three CpG islands were not methylated in the female control as shown in



**Figure 3.3** Examples of CpG islands that are abnormally methylated in lymphoma cell lines, SUDHL3 and DoHH2. (D = HpaII digested sample, M = HpaII mock digested sample). The presence of a band in the digested samples of SUDHL3 and DoHH2 and the lack of a band in the digested male control cell line indicates that these islands are abnormally methylated in the lymphomas. The weak bands in the male control digested samples for CpG 62, 71 and 79 are most likely due to incomplete cutting.



**Figure 3.4** Five examples of CpG islands that are not abnormally methylated in lymphoma cell lines in the 8 Mb region examined on the X chromosome. (D = HpaII digested sample, M = HpaII mock digested sample.) SUDHL3 and DoHH2 are two follicular lymphoma cell lines. The lack of bands in the digested samples of the lymphoma cell lines indicate that these islands are unmethylated. The weak band in the digested samples for CpG 81 is most likely due to incomplete cutting.



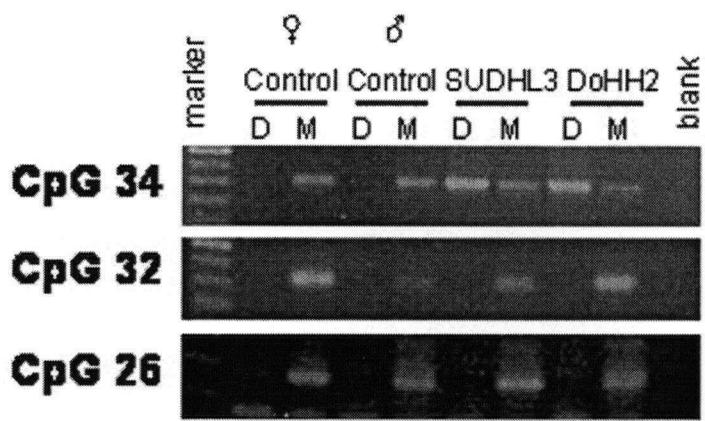
**Figure 3.5** CpG islands showing variation in methylation status between the two follicular lymphoma cell lines, SUDHL3 and DoHH2. (D = HpaII digested sample, M = HpaII mock digested sample.) All these islands are normally not methylated, therefore any methylation seen is abnormal (band in digested male control for CpG 19 is likely a result of incomplete cutting by HpaII).

figure 3.6. CpG 32 and 26 are located upstream and within the 5' region of MLLT7, a gene that was shown to variably escape X chromosome inactivation [41]. The remaining CpG island, CpG 34, is associated with AR, but is located over 500bp upstream of the gene.

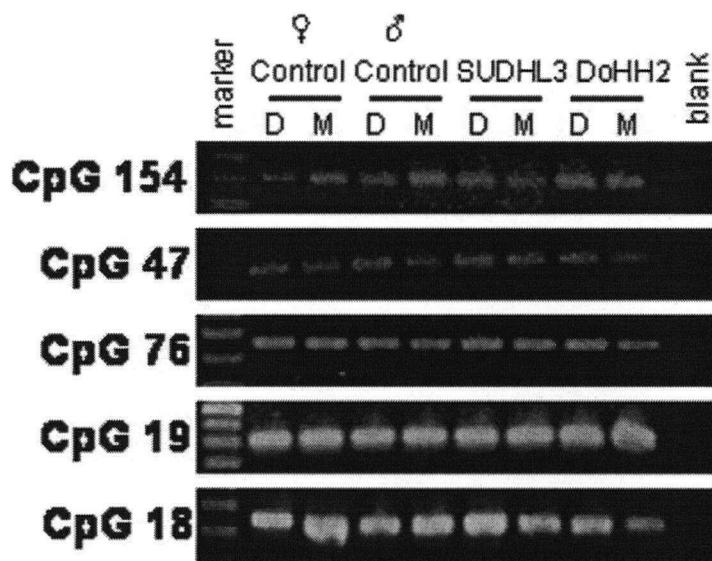
Nine CpG islands were ‘normally’ methylated and two more showed ‘partial’ methylation in the male control cell line. Five examples of CpG islands with normal methylation as defined by the presence of a PCR product for the HpaII digested male control cell line DNA are shown in figure 3.7. ‘Partial’ methylation is defined by the presence of a PCR product for the digested sample whose band is weaker in strength than the mock digested sample band. Examples of CpG islands showing partial methylation in any of the examined cell lines are shown in figure 3.8.

The CpG islands with unexpected methylation results, either methylation in the male control or a lack of methylation in the female control, were examined more closely with respect to their location relative to genes and repetitive elements (table 3.5). There appears to be no obvious relationship to explain the methylation observed in the male controls. Some CpG islands were not associated with any known genes and others were found in the 5' region or somewhere within a gene. The same observations were made for association with repeats, some CpG islands were associated with repeats while others were not. The abnormal methylation of CpG 45 (1), 27 (4) and 44 (1) could potentially be a result of their close proximity to multiple repetitive elements and their lack of association with a gene.

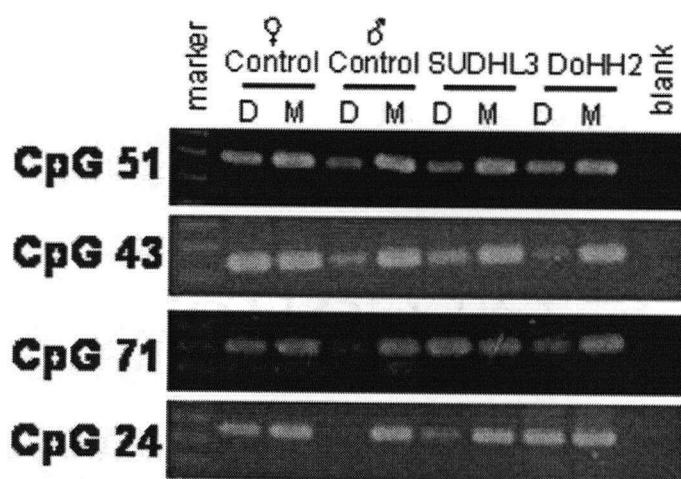
A recent paper looked at the relationship between some DNA-related features and CpG islands that are more prone to methylation in normal tissues, specifically in lymphocytes [114]. The authors found three DNA attributes of CpG islands, sequence, repeats, and structure, that led to a predisposition to methylation. CpG islands rich in CpGs and CpG islands associated with promoters appeared to be correlated with a lack of methylation. CpG islands that overlapped tandem repeats or segmental duplication and CpG islands with predicted DNA structures like high DNA rise and low DNA twist were associated with methylation. However, in this study, not all of the CpG islands that showed normal methylation followed these predictors of normal methylation.



**Figure 3.6** CpG islands showing a lack of methylation in the female control cell lines. (D = Hpall digested sample, M = Hpall mock digested sample.) SUDHL3 and DoHH2 are two follicular lymphoma cell lines. The presence of a band represents methylation and all three of these islands appear to be unmethylated in the female control cell lines.



**Figure 3.7** Five examples of CpG islands located within the examined 8 Mb region on the X chromosome that show methylation in the male control cell line. (D = HpaII digested sample and M = HpaII mock digested sample.) The presence of a band in the digested male control cell line samples indicate that these islands are methylated and indicate that the methylation seen in the lymphoma cell lines, SUDHL3 and DoHH2, is not abnormal.



**Figure 3.8** Examples of CpG islands located in the 8 Mb region in the vicinity of AR on the X chromosome showing partial methylation. (D = HpaII digested sample and M = HpaII mock digested sample.) SUDHL3 and DoHH2 are two follicular lymphoma cell lines. Partial methylation is classified in this thesis, as any band in the digested sample that is noticeably weaker than the band in the mock digested sample of the same cell line. The male control shows partially methylation of CpG 51 and 43. SUDHL3 has partial methylation of CpG 51 and 24, while DoHH2 has partial methylation of CpG 43 and 71.

**Table 3.5** CpG islands with unexpected methylation status in female and male control cell lines. The table looks at the association of these CpG islands with genes and repetitive elements in an attempt to explain the results. CpG islands associated with the 5' region of genes are believed to remain unmethylated whereas islands located near repetitive elements and satellite DNA are methylated.

CpG Island	Type of unexpected methylation result	Location of CpG island relative to nearest gene	Location of CpG islands relative to nearest repeat <sup>1</sup>
CpG 51	Partial methylation in ♂ control	5' region of gene	2 x LINE (L1) SINE (MIR)
CpG 43 (1)	Partial methylation in ♂ control	5' region of gene	SINE (MIR) LINE (L1)
CpG 45 (1)	Methylation in ♂ control	NOT associated with a gene	LINE (L1 and CR1) LTR (MaLR)
CpG 154	Methylation in ♂ control	5' region of a gene	
CpG 34 (2)	No methylation in ♀ control	749 bp upstream of gene	
CpG 27 (3)	Methylation in ♂ control	Located in 3' end of gene	SINE (MIR)
CpG 47	Methylation in ♂ control	5' region of a gene	
CpG 27 (4)	Methylation in ♂ control	NOT associated with a gene	5 x SINE (MIR) LINE (L1)
CpG 44 (1)	Methylation in ♂ control	NOT associated with a gene	2 x LINE (L2)
CpG 76	Methylation in ♂ control	5' region of a gene	3 x SINE (Alu) LINE (L1)
CpG 32 <sup>2</sup>	No methylation in ♀ control	357 bp upstream of gene	SINE (Alu) LINE (L2)
CpG 26 <sup>2</sup>	No methylation in ♀ control	5' region of a gene	LINE (L2)
CpG 19 (2)	Methylation in ♂ control	Located within the 3' end of a gene	LTR (ERV1) SINE (Alu)
CpG 18	Methylation in ♂ control	Located in the middle of a gene	

<sup>1</sup>repeats located within a 1 kb window of the CpG island

<sup>2</sup>associated with MLLT7, which, according to Carrel and Willard (2005), shows variation in inactivation.

The absence of methylation of CpG 32 and 26 in females can be explained by the variable inactivation of MLLT7 and it would be interesting to see if MLLT7 escapes methylation in the female control cell lines used here. The methylation status of the CpG islands with unexpected results, as well as CpG islands with variable methylation in the lymphoma cell lines, were assessed in four additional female and male control cell lines. The results from these assays are discussed below in section 3.2.2.2.

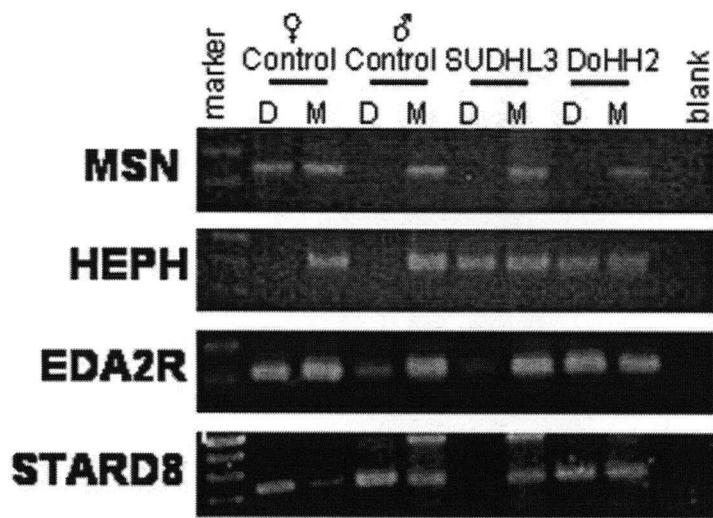
### ***3.2.2.1.1) Gene Promoter Methylation.***

In addition to determining the methylation status of the CpG islands in the 8 Mb region, the methylation status of the promoter and first exon region of 5 genes, including AR, were also examined (data is shown in figure 3.9). Immediately upstream of AR, there are relatively few CpG islands. The methylation status of the promoter region of three genes, moesin (MSN), hephaestin (HEPH) and ectodysplasin A2 receptor (EDA2R), located within this CpG island sparse area were determined to see if their methylation status is altered in the lymphoma cell lines. MSN appeared to remain unmethylated in the lymphoma cell lines while HEPH became methylated. EDA2R's methylation results were more complicated. The male control showed partial methylation and there was variation in methylation between the lymphoma cell lines.

The methylation status of the promoter region of STARD8 was also determined. This gene is located downstream of AR, at the 5' end of the 1.4 Mb highly methylated sub-region. STARD8 is associated with two CpG islands, but both are found within the gene. The promoter of STARD8, similar to the EDA2R promoter, showed variation in its methylation pattern. The male control was fully methylated and only one of the two lymphoma cell lines appeared to be methylated.

### ***3.2.2.2) Methylation in Additional Control Lymphoblast and Mouse/Human Hybrid Cell Lines***

The initial survey of the methylation status of CpG islands located in the vicinity of the AR produced some interesting and unexpected results, including methylation skipping, methylation in the male control, lack of methylation in the female control and abnormal methylation of a greater than expected region in the follicular lymphoma cell lines. The unexpected methylation patterns in the female and male controls were further



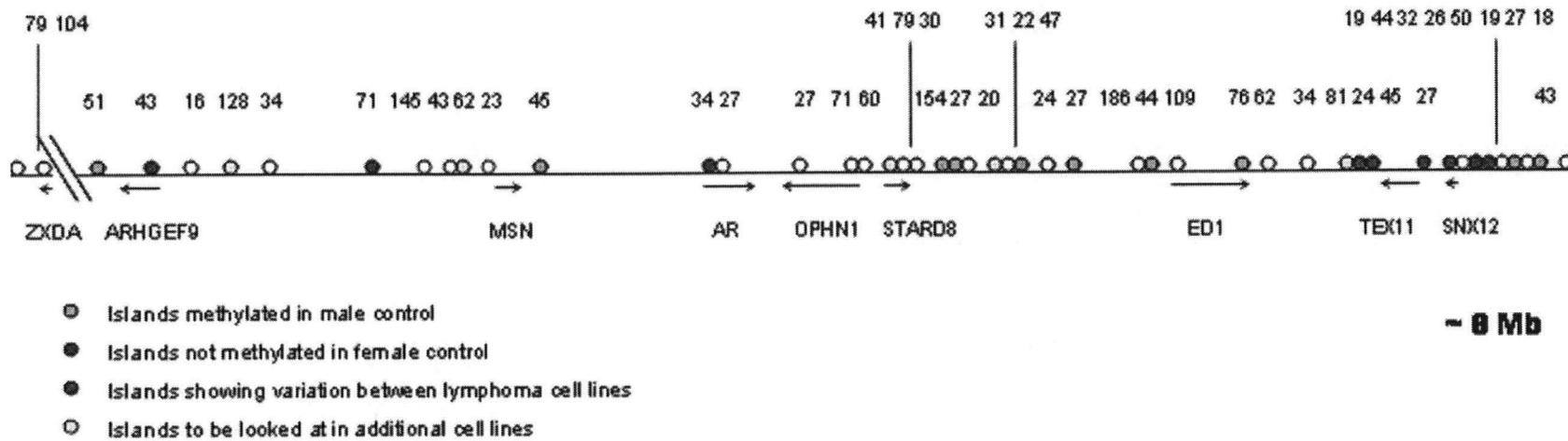
**Figure 3.9** Methylation of the promoter and first exon region of four genes located in a CpG island sparse area upstream of AR. (D = HpaII digested sample and M = HpaII mock digested sample.) SUDHL3 and DoHH2 are follicular lymphoma cell lines. MSN is the only gene which appears to unmethylated in the lymphoma cell lines as inferred from the lack of bands in the digested lymphoma samples. HEPH is abnormally methylated in both the lymphoma cell lines, and the female control is unexpectedly found to be unmethylated. EDA2R (also known as XEDAR) shows partial methylation of the male control and methylation in one of the lymphoma cell lines, DoHH2. STARD8 is also methylated in DoHH2 and the male control cell line looks to be fully methylated.

investigated by looking at the methylation status of selected CpG islands in additional female (GM11198, GM1119, GM11201, and GM7059) and male (GM7033, GM7009, GM11200, and GM7057) control lymphoblast cell lines. The selected CpG islands are shown in figure 3.10. If similar methylation patterns are observed for the selected CpG islands in the additional control cell lines, then it is unlikely that the results are due to an artifact of the cell line.

The assumption that any methylation seen in the male control represents normal methylation on the active X (X<sub>a</sub>) chromosome was made since males only possess a single copy of the X chromosome. As follows, it was assumed lack of methylation seen in the female control represents CpG islands that remain unmethylated on the inactive X (X<sub>i</sub>) chromosome. The methylation status of the selected CpG islands examined in the additional lymphoblast controls (figure 3.10) were also determined in the mouse/human hybrid cell lines retaining either the human X<sub>a</sub> (AHA 11aB1) or X<sub>i</sub> (t86-B1maz1b-3B). The hybrid cell lines provided information regarding which X chromosome experienced methylation at a specific CpG island as long as the primers were human specific. Table 3.6 is a summary of the methylation results for the selected CpG islands in all the additional cell lines.

The female control cell lines and the X<sub>i</sub> mouse/human hybrid cell line behaved as expected (methylated) for the majority of the ‘normally’ methylated CpG islands with a few exceptions. GM11198 and GM11199 were only partially methylated at CpG 44(1) and t86 was partially methylated at both CpG44 (1) and CpG 27 (4) (data is shown in figure 3.11). The male control cell lines, however, showed a great deal of variation in methylation. The normal methylation originally seen in one of the male control cell lines was not observed again in several cases. For example, none of the four male control cell lines showed methylation of CpG 45, whereas GM11200 initially showed methylation of CpG45 and the X<sub>a</sub> mouse/human hybrid cell line showed partial methylation. The CpG 47 methylation pattern was also unusual. The CpG island appeared to be methylated in all the male control cell lines, but unmethylated in the X<sub>a</sub> mouse/human hybrid cell line.

Three CpG islands, CpG 76, 19 (2) and 18, showed methylation in all of the examined cell lines, strongly suggesting that these CpG islands, in at least cell lines, are ‘normally’ methylated. The remaining five CpG islands which initially appeared to be



**Figure 3.10** Selected CpG islands whose methylation status was assessed in additional female and male control cell lines as well as Xa and Xi mouse/human hybrid cell lines. The selected CpG islands were chosen on the basis of their unexpected methylation status result in the initial methylation survey of the region to find the extent of the abnormal methylation in lymphoma. The CpG islands coloured in green are those who showed methylation in the male control and the CpG islands in red are those who lacked methylation in the female control cell line. The few CpG islands that varied in their methylation status between the two follicular lymphoma cell lines were also examined in the additional cell lines and are represented by the blue circle in the above map. As a control, as well as to get an even distribution of CpG islands along the chromosome segment, a few CpG islands whose methylation pattern was the expected methylation in the female control and unmethylated in the male control, were also looked at (CpG islands in yellow). The map is drawn relatively to scale and also shows all the remaining CpG islands (unfilled circles) and a few of the genes located in the 8 Mb region. The two hatch marks at the 5' end represent the crossing of the centromere and is not included in the 8 Mb region.

**Table 3.6** Summary of the methylation results of selected CpG islands located in the 8 Mb region in the vicinity of AR in additional female and male control cell lines and Xa and Xi mouse/human hybrid cell lines. The blocks filled in black represent methylated islands, in white, unmethylated islands, in red, partially methylated islands and with diagonal lines, CpG islands not examined due to primer optimization difficulties.

	♀ Control Cell Lines				♂ Control Cell Lines				Xa Hybrid	Xi Hybrid
CpG Island <sup>1</sup>	GM11198	GM11199	GM11201	GM7059	GM7033	GM7009	GM11200	GM7057	AHA	t86
<b>Islands methylated in male control:</b>										
51										
45(1)										
154										
27(3) <sup>2</sup>										
47										
27(4) <sup>3</sup>										
44(1)										
76										
19(2)										
18										
<b>Islands not methylated in female control:</b>										
34(2)										
32 <sup>4</sup>									NA	NA
26										
<b>Islands showing variation in methylation in lymphoma:</b>										
43(1)										
71(1)										
24(2)										
45(2)										
27(5) <sup>2</sup>										
19(1)										

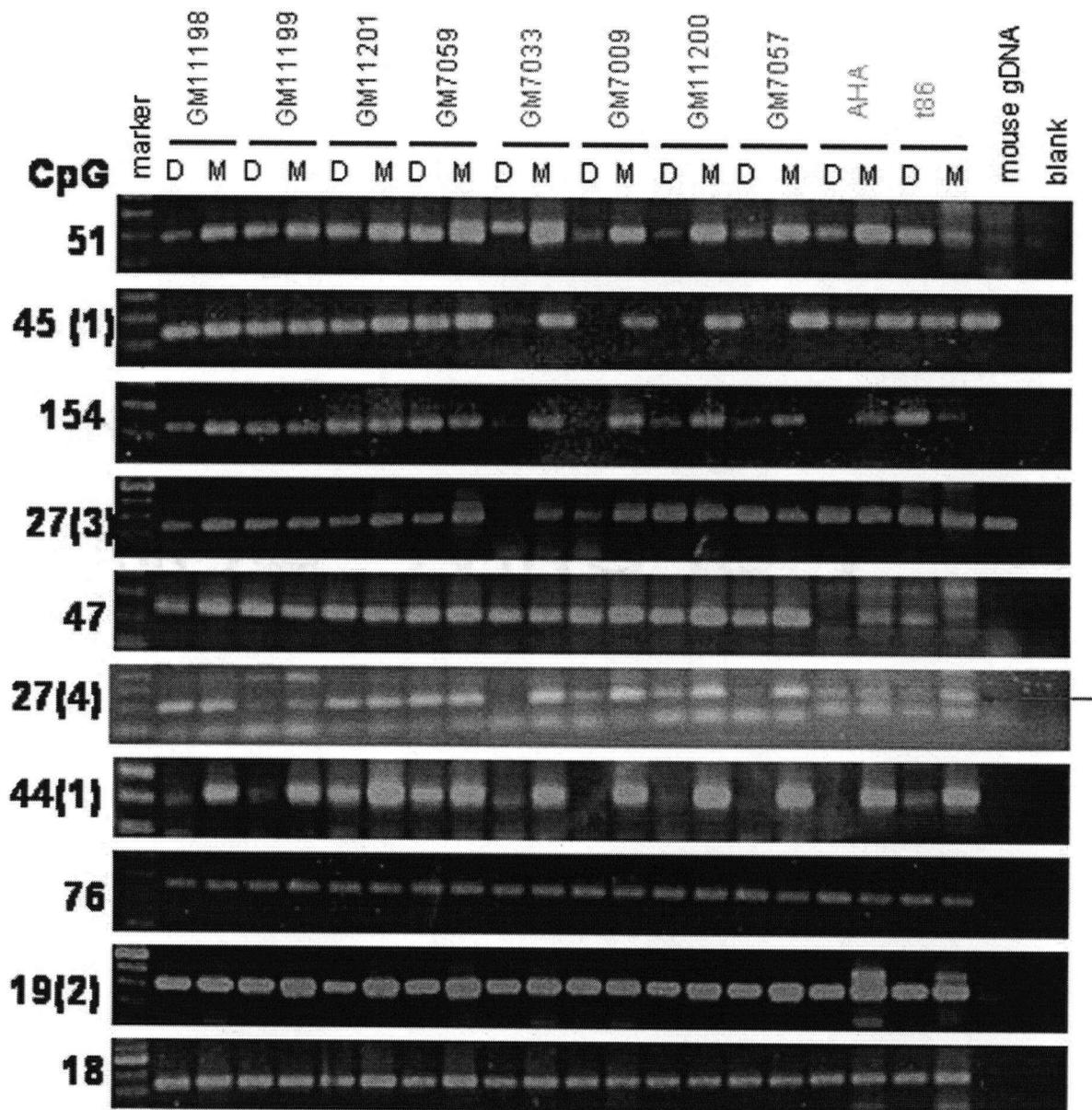
	♀ Control Cell Lines				♂ Control Cell Lines				Xa Hybrid	Xi Hybrid
CpG Island <sup>1</sup>	GM11198	GM11199	GM11201	GM7059	GM7033	GM7009	GM11200	GM7057	AHA	t86
<b>Islands looked at in additional cell lines:</b>										
128										
62(1)										
41										
31										
81										

numbers in () used to distinguish CpG islands since nomenclature is based on the number of CpG dinucleotides as per the UCSC Genome Browser

<sup>2</sup> Primers amplify same product size in mouse gDNA

<sup>3</sup> Primers amplify a larger product in mouse genome and bands for expected human products are very weak

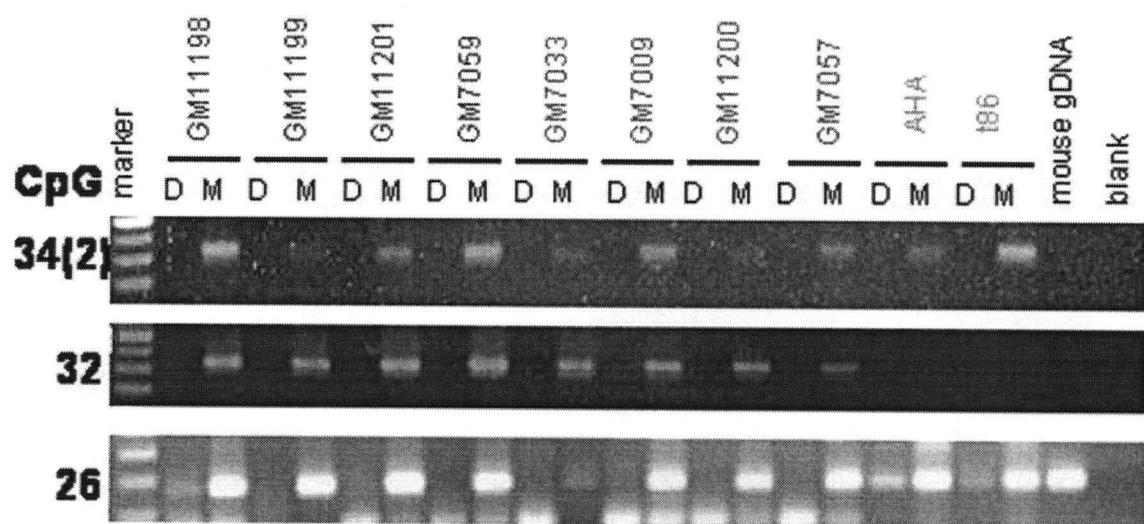
<sup>4</sup> no methylation results for hybrid cell lines since DNA did not amplify



**Figure 3.11** Methylation assay in female and male control and Xa/Xi retaining mouse/human hybrid cell lines of CpG islands previously showing methylation in a male control cell line. (D = HpaII digested sample and M = HpaII mock digested sample.) Cell line names in pink are female controls, in blue, male controls and in green mouse/human hybrids. AHA is a Xa mouse/human cell line and t86 is an Xi cell line. The larger band for CpG 27 (4) is the expected band. The methylation seen in the male control in the original methylation assays is not always seen in the additional control cell lines (CpG islands 51, 45, 154, and 44). The primers used to check the methylation status of CpG 27(3) are not human specific (PCR product amplified in mouse gDNA) and therefore the hybrid cell lines are uninformative.

methylated in the male control have at least one male control that is partially methylated in the follow up study. CpG 51 only lacked methylation in GM7033, and the other three male cell lines and the Xa mouse/human hybrid all showed partial methylation. CpG 154 was partially methylated in three of the male cell lines, and unmethylated in one male and AHA. CpG 27 (3) experienced a lot of methylation variation between the male control cell lines, ranging from unmethylated in GM7033, partially methylated in GM7009, and methylated in GM11200, and GM7057. AHA also appeared to be methylated, based on the presence of a PCR product, but the primers also amplified mouse genomic DNA and therefore no conclusions could be drawn about the methylation status of the mouse/human hybrid cell lines. Methylation of CpG 27 (4) also varied between cell lines. It was unmethylated in GM7033 and GM7057, partially methylated in GM7009 and GM11200 and methylated in AHA. CpG 44 (1) was only partially methylated in GM7033, and unmethylated in all the remaining male control and Xa mouse/human hybrid cell lines. The observation of at least some methylation in a few of the male cell lines suggests that methylation of CpG islands 51, 154, 27(3), 47, 27(4), 76, 19(2) and 18 is potentially normal but also variable. The ‘normal’ methylation initially seen in CpG islands 45 and 44(1) was likely a consequence of incomplete cutting by HpaII.

The three islands that were unmethylated in the original survey of the methylation status of CpG islands in the vicinity of AR, also showed a lack of methylation in the additional investigated female control cell lines (data shown in figure 3.12). CpG 34 (2) was unmethylated in all ten of the examined cell lines. CpG 32 was also unmethylated in all the female and male control cell lines, but the methylation status of this island could not be determined in the hybrid cell lines as the DNA failed to amplify. CpG 26 showed partial methylation in GM11198 and in both of the hybrid cell lines, but was unmethylated in the remaining cell lines. As was previously mentioned, CpG 32 and 26 are associated with a gene that is known to escape X inactivation and perhaps the CpG islands are therefore not required to be methylated. CpG 34 (2) is associated with AR, which is normally susceptible to silencing, but the CpG island is located more than 500 bp upstream of the gene and CpG 27 (1) is more closely associated with AR. This suggests that methylation of CpG 34 (2) is perhaps not necessary to maintain stable silencing of AR on the inactive X chromosome.

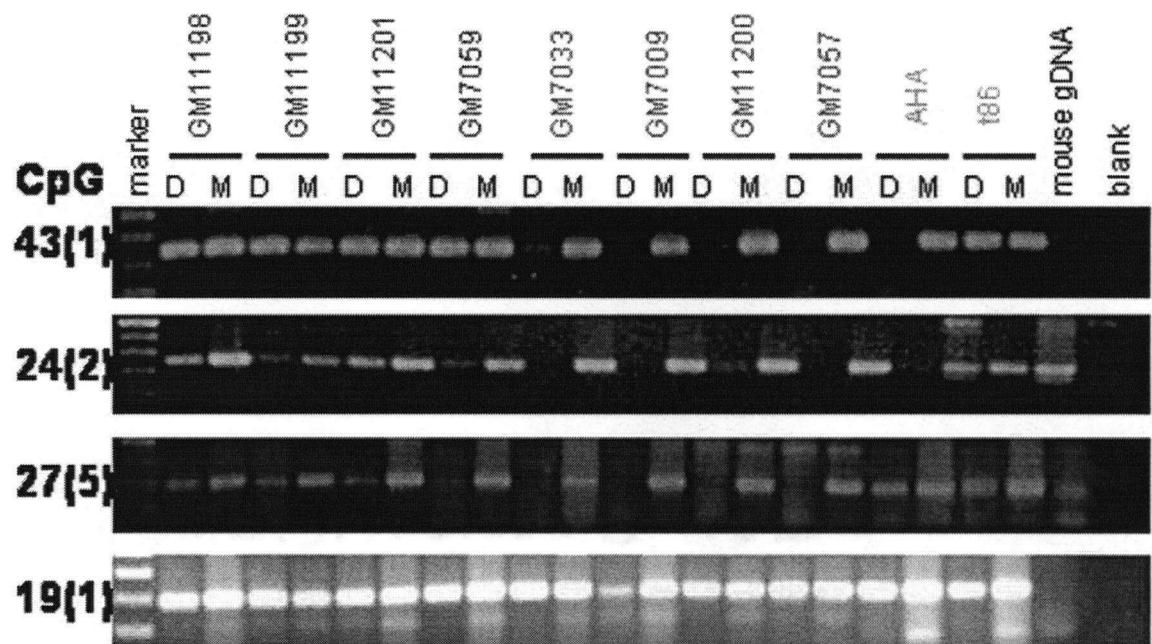


**Figure 3.12** Methylation assay in female and male control and Xa/Xi retaining mouse/human hybrid cell lines of CpG islands that previously showed a lack of methylation in a female control cell line. (D = HpaII digested DNA and M = HpaII mock digested sample.) Cell line names in pink are female controls, in blue, male controls and in green mouse/human hybrids. AHA is a Xa mouse/human cell line and t86 is an Xi cell line. Mouse/human hybrid cell line DNA did not amplify with CG32 primer pair. The lack of methylation at these loci is very consistent across all examined female control cell lines with the exception of the partial methylation seen in GM11198 at CpG island 26. The primers used to look at the methylation status of CpG 26 also amplify mouse genomic DNA therefore, bands seen in digested hybrid cell lines are not informative.

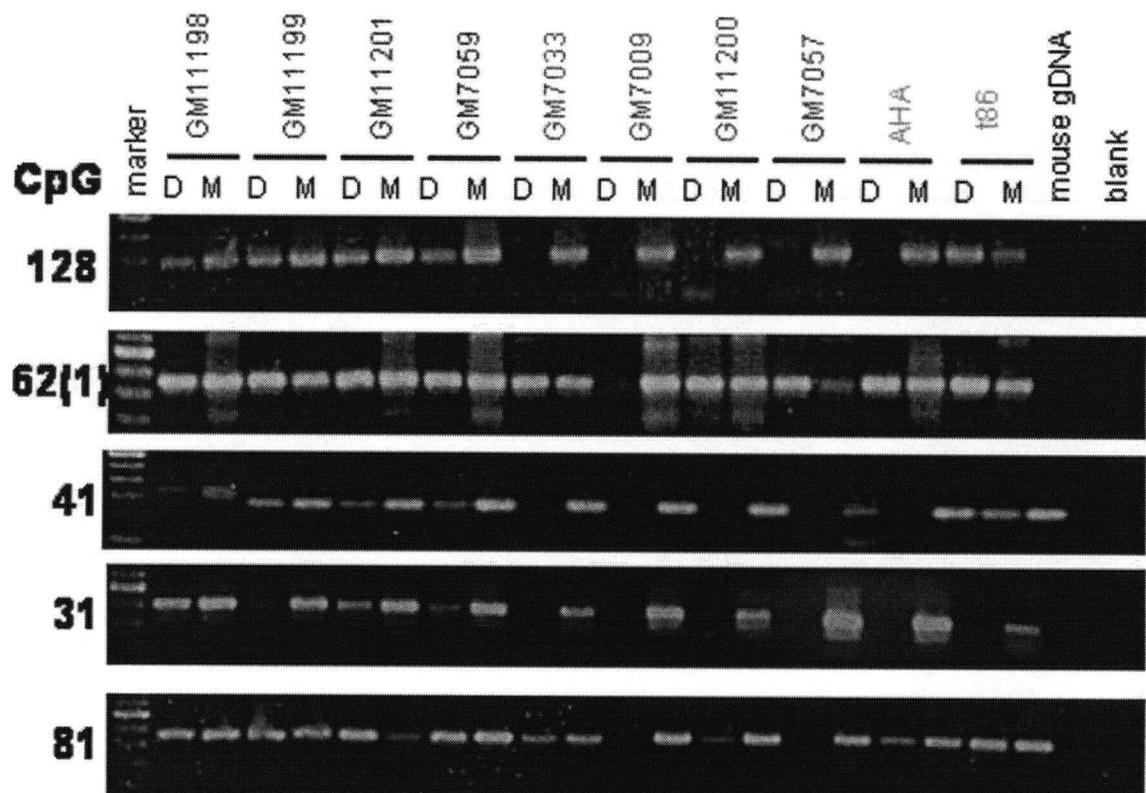
Six CpG islands showed variation in methylation between the two follicular lymphoma cell lines, SUDHL3 and DoHH2 (table 3.6). The methylation status of these islands were also determined in the additional control cell lines, but results were only obtained for four of the CpG islands. There were difficulties in obtaining consistent PCR products for all the cell lines with the primers used to assess the methylation status of CpG 71 (1) and CpG 45 (2). The primers were optimized on genomic DNA diluted in the digestion buffer. However, when the optimized PCR conditions were used to amplify the digested DNA, there was a lot of inconsistency within and between cell lines and a clear picture of the methylation status of these CpG islands could not be determined.

The results of the methylation analysis of the four CpG islands which showed methylation variation between the lymphoma cell lines are shown in figure 3.13. The methylation patterns for CpG 43 (1), 24 (2) and 27 (5) followed, for the most part, the expected pattern of methylation in the female and no methylation in the male control cell lines. Some of the female control cell lines only showed partial methylation of CpG 24 (2) and CpG 27 (5). One of the female control cell lines, GM7059, appeared to be unmethylated at the CpG 27 (5). Unfortunately, the methylation status of the active X on its own in the hybrid cell line could not be established for this CpG island since the primers also amplified mouse genomic DNA. The lack of methylation in GM7059 could be the result of methylation variation between cell lines (either natural variation of methylation of this CpG island, or variation induced by tissue culture conditions). Quite unexpected, CpG 19 (1) was methylated in all the control cell lines, suggesting that it is normally methylated. In the original methylation assay of the region, the male control was unmethylated. It is possible that this CpG island has become methylated in tissue culture over time or the result is a false positive caused by the limitations of the assay.

The methylation status of an additional five CpG islands was also determined in the additional control cell lines in order to obtain an even distribution of re-examined CpG islands along the chromosome. The data is shown in figure 3.14 and summarized in table 3.6. All the female control cell lines showed some degree of methylation, but it was surprising to observe methylation in some of the male control cell lines. CpG 62 (1) was methylated in all the cell lines, including the Xa mouse/human hybrid (AHA), with the exception of a lone male control cell line, GM7009. CpG 81 was also methylated in some



**Figure 3.13** Methylation assay in female and male control and Xa/Xi retaining mouse/human hybrid cell lines of CpG islands that previously showed variation in methylation in lymphoma cell lines. (D = HpaII digested DNA and M = HpaII mock digested sample.) Cell line names in pink are female controls, in blue, male controls and in green mouse/human hybrids. AHA is a Xa mouse/human cell line and t86 is an Xi cell line. The methylation results indicate that the majority of these cell lines have the expected methylation pattern, methylation in female control and no methylation in the male control. CpG 19 is unusual since it appears to be methylated in all the examined male control cell lines, when in the original methylation assay done at this locus, the male was not methylated.



**Figure 3.14** Methylation assay in female and male control and Xa/Xi retaining mouse/human hybrid cell lines of 5 selected CpG islands. (D = HpaII digested DNA and M = HpaII mock digested sample.) Cell line names in pink are female controls, in blue, male controls and in green mouse/human hybrids. AHA is a Xa mouse/human cell line and t86 is an Xi cell line. There is some variation in methylation status between the cell lines, but the majority of the cell lines follow the expected pattern of methylation in the female control and a lack of methylation in the male control. The exceptions include, methylation in the male GM7033, GM11200 and GM7057 cell lines at CpG 62, no methylation in the female GM1199 and t86 Xi mouse/human hybrid cell line at CpG 31 and methylation in male GM7033 and partial methylation in GM11200 and AHA at CpG 81.

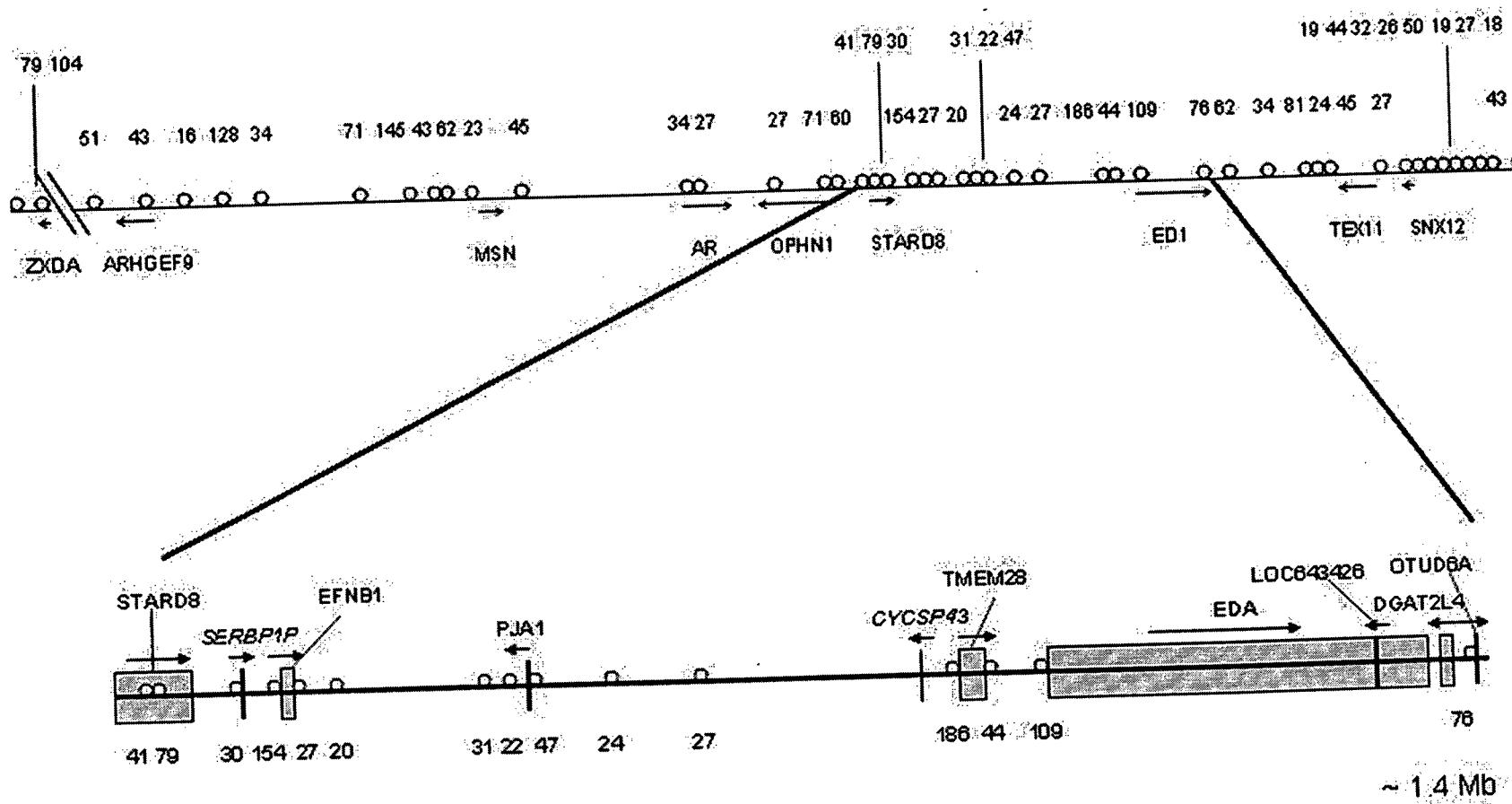
of the male control cell lines and AHA was partially methylated. The other unexpected result was the lack of methylation of the Xi mouse/human hybrid at CpG 31. Two of the female control cell lines were partially methylated and the other two were fully methylated at this CpG island, but none showed a lack of methylation.

### ***3.2.3) Methylation in the 1.4 Mb Sub-Region in Three Additional Lymphoma Cell Lines***

The first goal of the thesis was to determine the extent of the abnormal hypermethylation spread along the X chromosome in lymphoma. This was done using two follicular lymphoma cell lines, SUDHL3 and DoHH2, as described in section 3.2.2.1. In order to meet the ultimate goal of the study, to identify candidate tumour suppressor gene(s), it was necessary to reduce the region of interest from the large 8 Mb region to a more focused sub-region. Figure 3.15 depicts a 1.4 Mb sub-region which showed a high degree of hypermethylation in the two follicular lymphoma cell lines. In fact, all of the 14 examined CpG islands (CpG 41 to 76) in this region located at Xq13.1, were hypermethylated. The methylation status of only one CpG island, CpG 30, in this region was not assessed for methylation due to problems designing specific primers (previously mentioned in section 3.2.2.1).

In order to further narrow down on a smaller region of interest, the methylation status of the 14 CpG islands were examined in additional lymphoma cell lines. The optimal outcome of this analysis would be that the additional lymphoma cell lines also show methylation in this region, but perhaps smaller segments which can be overlapped to further reduce the region of interest. We received three mantle cell lymphoma (MCL) cell lines, Z138, HBL-2 and JVM-2, from Catherine Tucker and Richard Klasa from the BC Cancer Research Centre and proceeded to look at the methylation status of the CpG islands in these cell lines. Z138 and HBL-2 are male cell lines and JVM-2 is a female cell line. Methylation positive results for JVM-2 can not be called abnormal due to the involvement of DNA methylation in X inactivation.

The methylation status of AR was tested first since the abnormal methylation of AR in lymphoma patient samples motivated this study. The three MCL cell lines' methylation results of AR and the CpG islands in the 1.4 Mb region are summarized in



**Figure 3.15:** Map of the 1.4 Mb sub-region of interest. This region located downstream of AR at Xq13.1 showed abnormal hypermethylation of all 14 examined CpG islands in the follicular lymphoma cell lines, SUDHL3 and DoHH2. The methylation status of these CpG islands has also been determined in three mantle cell lymphoma cell lines, which also showed a high degree of hypermethylation. In addition to showing the location of all the CpG islands, the map also gives the location of all the known genes according to the UCSC May 2004 hg17 genome assembly freeze (pseudogenes are italicized). The genes located in this region differ slightly when comparing information taken from the UCSC genome browser and NCBI Entrez Gene (see Section 3.2.1).

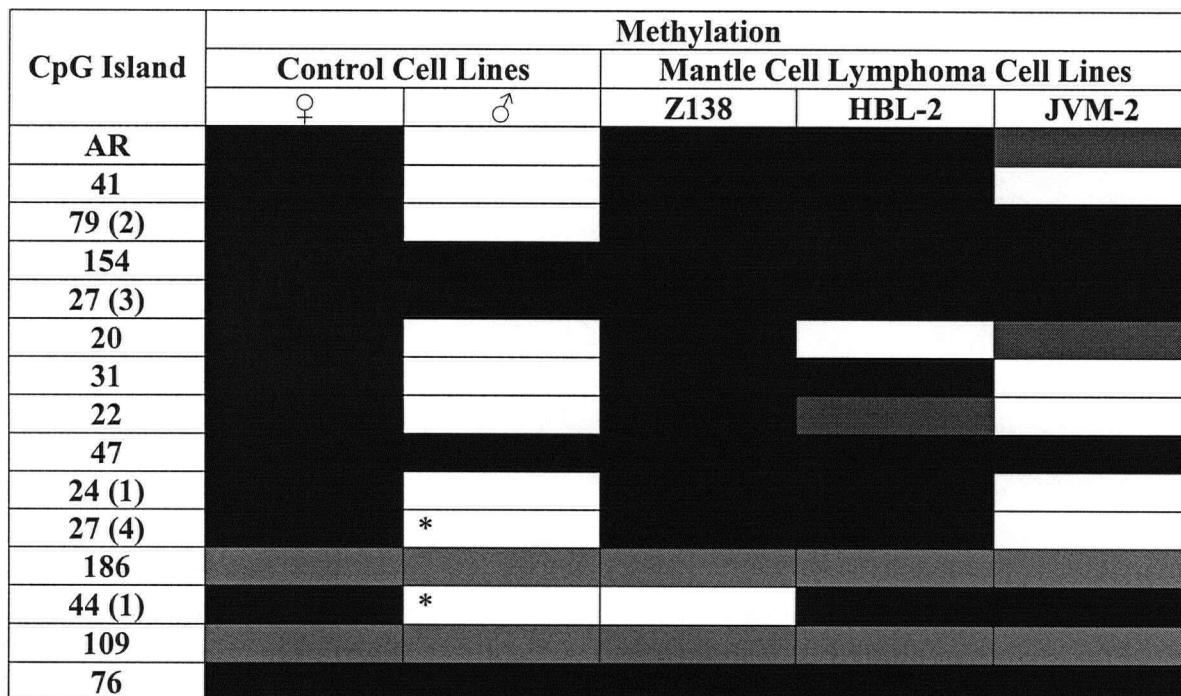
table 3.7 and the data is shown in figure 3.16. Two of the cell lines, Z138 and HBL-2, showed abnormal methylation of AR, whereas the remaining cell line, JVM-2, appeared to be partially methylated. The JVM-2 cell line also differed from the other two in cell growth and overall methylation pattern of the CpG islands found in the 1.4 Mb sub-region. The JVM-2 cells were relatively slow growing and formed clumps of cells on the bottom of the t75 flask whereas HBL-2 and Z138 were relatively fast growing and formed an even layer of cells.

JVM-2 was only methylated in 8 out of the 12 examined CpG islands (the methylation status of CpG 186 and 109 were not determined due to problems with primer PCR conditions). Of these eight islands, two showed partial methylation, and four were CpG islands that were ‘normally’ methylated in at least one male control cell line. These results suggests JVM-2 is relatively hypomethylated, which might result from the loss of the inactive X chromosome. However, genome array CGH profiles of this cell line suggests otherwise [115]. Both Z138 and HBL-2 showed a high degree of methylation in the 1.4 Mb sub-region. Z138 was methylated in all but one CpG island. HBL-2 lacked methylation of one CpG island, was partially methylated in another CpG island, but showed abnormal methylation for the remaining CpG islands.

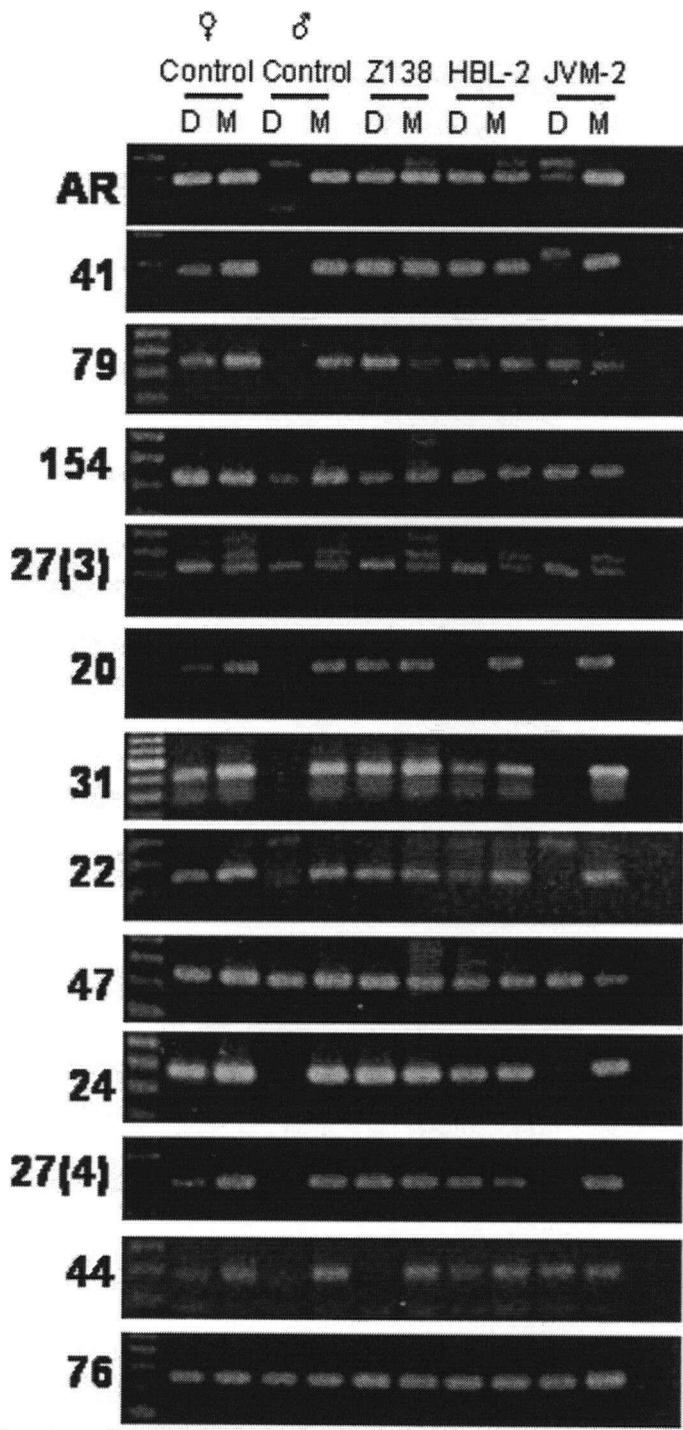
### ***3.2.4) Methylation in the 1.4 Mb Sub-Region in Lymphoma Patient Samples***

The focus so far has been to establish the extent of the abnormal methylation in lymphoma cell lines rather than in lymphoma patients where the aberrant methylation was first observed [4]. There are a couple of reasons for this. The lymphoma patient samples are often contaminated with normal non-malignant cells which could interfere with the methylation assessment. The other reason concerns the rarity of the samples and the amount of DNA required to complete the overview of the methylation pattern in the 8 Mb region. It was unfortunate that assessing the methylation status of the CpG islands located in the 1.4 Mb sub-region in the MCL cell lines did not provide information to further reduce the size of the region of interest, however it did confirm that abnormal methylation is present in at least 4 lymphoma cell lines. The next step was to determine if the hypermethylation in the 1.4 Mb sub-region is also present in lymphoma patients.

**Table 3.7** Methylation assay results in the sub-region and for AR in three mantle cell lymphoma cell lines, Z138, HBL-2 and JVM-2. The rectangles filled in black represent methylated CpG islands, in white, unmethylated CpG islands, in red, partially methylated CpG islands and in diagonal lines, unexamined CpG islands. Z138 and HBL-2 mantle cell lymphoma cell lines behave most similarly to the two examined follicular lymphoma cell lines, SUDHL3 and DoHH2. Of note, CpG 27 (4) and CpG 44 (1) both appeared to be unmethylated in the male control whereas they previously showed methylation in a different male lymphoblast control cell line.



\*CpG islands that previously appeared to be methylated in a different male control cell line



**Figure 3.16** Methylation results of AR and the 1.4 Mb sub-region in the three additional lymphoma cell lines, Z138, HBL-2 and JVM-2. (D = HpaII digested DNA and M = mock digested DNA). All three of the cell lines are mantle cell lymphomas. The methylation status of three CpG islands could not be determined within the sub-region. CpG 30 is located within a segmental duplication and specific primers could not be designed. The methylation status of CpG 186 and 109 could not be determined due to primer problems. Overall the methylation patterns of Z138 and HBL-2 most closely resemble that of the two follicular lymphoma cell lines, SUDHL3 and DoHH2, showing a high degree of aberrant methylation.

We obtained DNA from a total of 13 male lymphoma patients from our collaborators at the BC Cancer Research Centre. Due to limitations in the availability of tumour DNA, the smallest quantity of DNA required to assess the methylation status of the 14 CpG islands was determined. The lowest DNA quantity that still managed to produce a visible PCR product on an agarose gel was the smallest quantity of DNA required from the patient samples to carry out the study. 200 ng of genomic DNA was found to meet these requirements and a total of 400 ng DNA was needed from each patient sample, 200 ng for each a digested and mock digested 20 µl reaction. This gave a final concentration of 10 ng/µl of digested DNA, from which 1 µl was used per PCR reaction. This is a 10 fold reduction in the amount of DNA usually used in the PCRs described above, and this caused some problems as described below.

The patient samples came from two different kinds of lymphoma, mantle cell (MCL) and diffuse large B cell lymphoma (DLBC). Samples were received and analyzed on two separate occasions and the methylation results are shown in two figures, 3.17 and 3.18 and summarized in table 3.8. In these figures, the patient samples were not separated based on type of lymphoma, but on when the sample was received. However, the first group of patient samples were all MCL cases and the second were all DLBC with the exception of a lone MCL case, 6965. The separation of the samples based on when they were received was done because the same CpG islands were not examined in all the patient samples due to problems with the primers and/or limitations in the quantity of DNA.

The methylation status of the AR was first assessed in all the patients, just as was done with the additional MCL cell lines. Twelve out of the thirteen (92%) patients showed aberrant methylation of AR. This frequency is higher than what was first observed by McDonald *et al.* (2000) who saw abnormal methylation in 84% of lymphoma patients.

The patients' CpG island methylation pattern was not as dense as the methylation seen in the two FL cell lines. Of the 10 CpG islands whose methylation status was successfully determined in the first group of patients, the majority of the CpG islands that showed methylation were ones 'normally' methylated in the male lymphoblast control cell line. CpG islands 27(3), 47, 27(4) and 76 all showed either full or partial methylation

**A)**

marker	9988	9215	8033	8487	8435	8610	blank
D	M	D	M	D	M	D	M

**AR**

**CpG 41**



**CpG 154**



**CpG 27(3)**

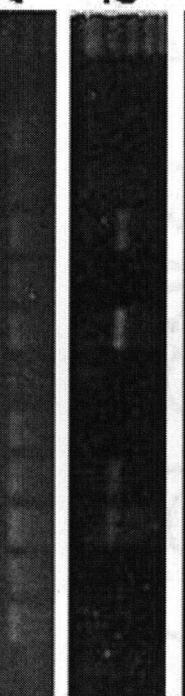
**CpG 20**



**CpG 31**



**CpG 22**



**CpG 47**



**CpG 24(1)**



**CpG 76**



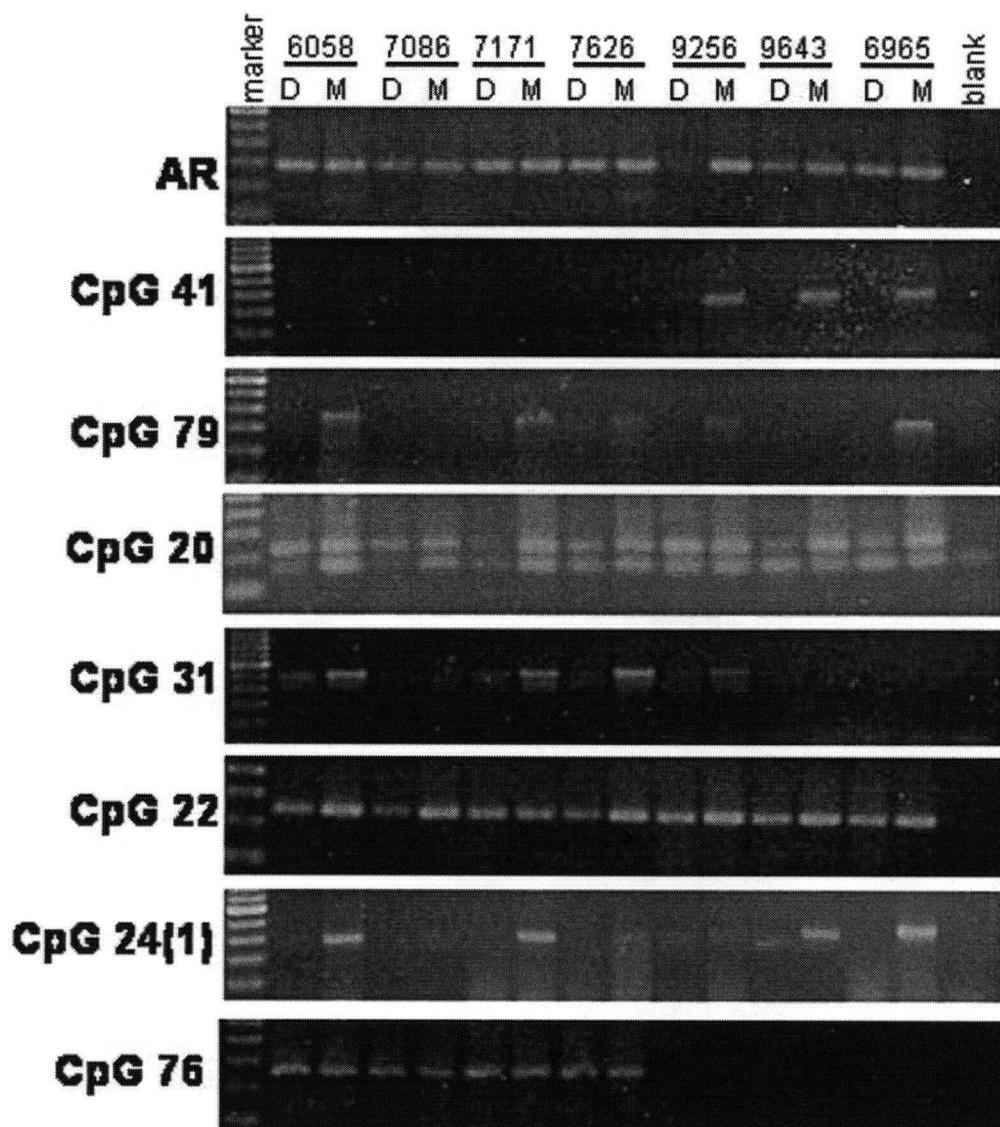
**B)**

	CpG 31		CpG 22		CpG 47		
marker	9988	8435	8610	blank	9988	8487	blank
D	M	D	M	D	M	D	M

**C)**

	CpG 24(1)	
marker	9988	8435
D	M	D

**Figure 3.17** Methylation results for AR and CpG islands located in the 1.4 Mb sub-region in the first group of patient samples. (D = HpaII digested sample and M = mock digested sample). All six of the primary tumour patient samples have mantle cell lymphoma. The methylation status of five CpG islands located within this region have not been assessed. CpG 30 is located within a segmental duplication and specific primers could not be designed to investigate its methylation status. The methylation status of CpG islands 79, 186, 44 (1) and 109 could not be determined for the patient samples due to primer problems. Part A) shows the methylation status for all examined CpG islands as well as AR in all patients. There were some dropouts, and the methylation status of four CpG islands (B) CpG 31, 22, and 47 and C) 24(1)) were re-examined for some of the patient samples. The bands are very weak for many of the results, for example CpG 27(3) and 20, and this is likely due to the low quantities of DNA in the digested and mock digested samples.



**Figure 3.18** Methylation results for AR and CpG islands located in the 1.4 Mb sub-region in the second group of patient samples. (D = HpaII digested sample and M = mock digested sample) Six out of the seven primary tumour samples (6058, 7086, 7171, 7626, 9256 and 9643) are cases of DLBC and the seventh primary tumour sample (6965) is a MCL case. The methylation status of AR and only seven of the CpG islands in the 1.4 Mb sub-region were analyzed in the second group due to primers difficulties and limited amounts of digested patient sample DNA. Four of the CpG islands not examined were ones that showed 'normal' methylation (CpG 154, 27(3), 47 and 27(4)). The methylation status of CpG 186, 44(1) and 109 could not be determined due to primer problems. The upper band in the results for CpG 20 is the expected band. The band in the digested sample of 9643 for CpG 24(1) is lower than the expected band and therefore the island is not methylated. In addition, there are very weak bands present for both the digested and mock digested 9256 sample amplified with the CpG 24(1) primers, suggesting this island is methylated.

**Table 3.8** Summary of the CpG islands methylation results in the 1.4 Mb sub-region for all patient samples. Boxes coloured in black represent methylation, in red, partial methylation, in white, no methylation, in diagonal lines, unexamined CpG islands and boxes with NA represent patients whose DNA could not be amplified. Two types of lymphoma are represented by the patient samples below, mantle cell lymphoma (9988, 9215, 8033, 8610, 8487, 8435 and 6965) and diffuse large B-cell lymphoma (6058, 7086, 7171, 7626, 9256, 9643). In addition to the CpG islands located in the 1.4 Mb sub-region, the methylation status of AR in the patient samples is also shown.

	♀ Control	♂ Control	Patient Samples												
CpG Island	GM11198	GM7009	9988	9215	8033	8610	8487	8435	6058	7086	7171	7626	9256	9643	6965
AR															
41									NA	NA	NA	NA			
79(2)										NA				NA	
154															
27(3)															
20															
31			NA			NA		NA		NA				NA	NA
22							NA								
47															
24(1)										NA		NA			
27(4)															
186															
44(1)															
109															
76													NA	NA	NA

in the patients and full methylation in the male control cell line. One CpG island, CpG 154, was not methylated in the patients even though it was methylated in one of the male control cell lines. This is not surprising because when the methylation status of this CpG island was re-assessed in the additional control cell lines, none of the males showed full methylation (see table 3.6). Two CpG islands in the first group of patients showed some abnormal methylation. CpG 20 appeared to be partially methylated in all six of the patient samples and CpG 41 was fully methylated in four patients and partially methylated in the remaining two. It is interesting that CpG 41 shows abnormal methylation in the patients since it is associated with a potential candidate tumour suppressor gene, STARD8 (further discussed in section 3.2).

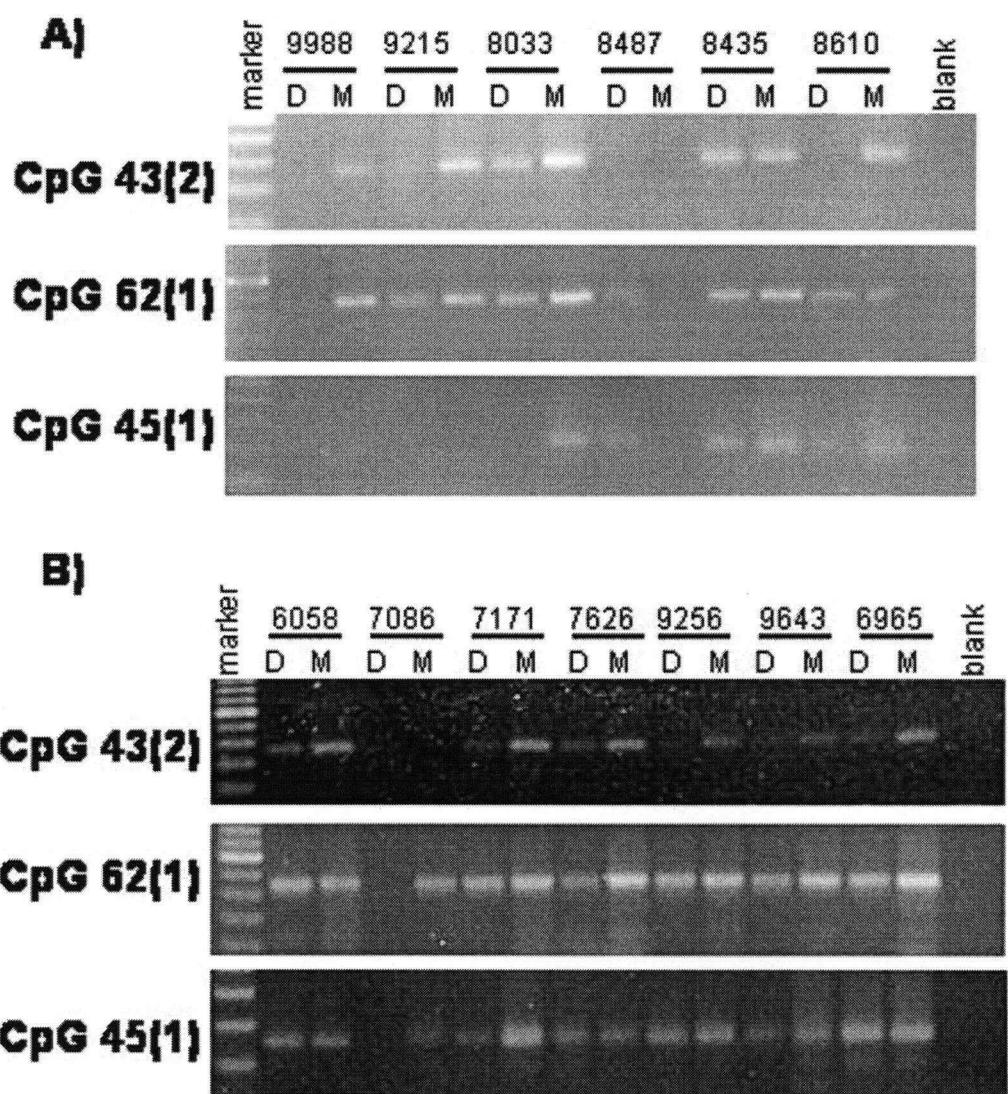
The methylation status of four CpG islands was not determined in the first group of patients due to PCR amplification difficulties. The region of interest is CG rich and therefore designing and optimizing the PCR conditions for primers was difficult. The primers originally used to assess the methylation status of CpG islands 79(2), 186, 44(1) and 109 in the FL cell lines produced either inconsistent results or failed to amplify DNA when used to determine the methylation status in patient samples. Re-optimizing the PCR conditions failed to solve the problem, and in the future, to get a complete picture of the methylation pattern, it might be beneficial to design new primers for these CpG islands. The other problem experienced while looking at the methylation patterns in the patients was frequent failure of PCR amplification. This is likely due to the low DNA quantities used per PCR reaction.

The methylation status of only six CpG islands was assessed in the second group of patients with a disappointing number of dropouts. There was not enough DNA to investigate the methylation status of the remaining CpG islands even though the same quantity of DNA was received for both sets of patients. In the second group, however, difficulties were encountered in digesting the DNA, and more DNA was needed to check for complete cutting after several rounds of HpaII re-digestions. The CpG islands whose methylation status was not determined were ones who showed ‘normal’ methylation and therefore, any methylation seen in the patients would not be as informative in identifying a candidate tumour suppressor gene.

Only two CpG island showed abnormal methylation in the second group of patients. CpG 20, the same CpG island that was partially methylated in all of the first group of MCL patients, was fully methylated in four of the second group of patients, partially methylated in an additional two patients and not methylated in one patient. The other CpG island that showed abnormal methylation in the patients, CpG 22, was only methylated in one of the five informative patients in the first group. In the second group of patients, this CpG island was fully methylated in five patients and partially methylated in the other two. Two other CpG islands, CpG 31 and 24(1), showed some degree of methylation in one patient (but not the same patient). CpG 41, which showed such a high degree of methylation in the first group of patients, was not methylated in any of the three informative patients in the second group.

At first glance, it is easy to conclude that the differences in methylation observed between the two groups of patients is due to the type of lymphoma as the first group is made up of all MCL cases and the second, mainly of DLBC cases. However, both MCL and DLBC patients had abnormally methylated AR. This observation, together with the high number of islands experiencing normal methylation (50%) suggests that the sub-region is a potentially normally hypermethylated region that does not house a candidate tumour suppressor gene. CpG islands located upstream of AR also showed abnormal methylation in the two FL cell lines and thus could not be ruled out as a region of interest.

The methylation status of three CpG islands located upstream of AR was assessed in both patient groups to see if this region showed any more abnormal hypermethylation than what was observed for the sub-region. The data is shown in figure 3.19 and the methylation results for all the CpG islands assessed in the patient samples are summarized in table 3.9. All three of the CpG islands showed abnormal methylation in many of the patients. CpG 45(1) was methylated to some degree in 82% of the patients. This CpG island originally showed ‘normal’ methylation in the male control cell line, but when the methylation status was re-assessed in additional control cell lines, the methylation was not observed again, even in the original male control cell line (see section 3.1.1.3). The next CpG island, CpG 62(1), was fully methylated in 85% of the informative patients and CpG 43(2) was methylated in 73% of the informative patients.



**Figure 3.19** Methylation status of CpG islands located upstream of AR in patient samples. (D = HpaII digested samples and M = mock digested samples.) **A)** shows the methylation results for the first group of patient samples, all of which are MCL cases. **B)** shows the methylation results for the second group of patient samples, of which six (6058, 7086, 7171, 7626, 9256 and 9643) are DBCL cases and one (6965) is a MCL case.

**Table 3.9** Summary of the methylation status of AR and all the CpG islands assessed in the MCL and DBCL patient DNA. Boxes filled in black represent methylation, in red, partial methylation, in white, no methylation, in diagonal lines, CpG islands not examined and boxes with NA represent cases where patient DNA did not amplify. There are seven cases of MCL (9988, 9215, 8033, 8610, 8467, 8435, and 6965) and six cases of DBCL (6058, 7086, 7171, 7626, 9256 and 9643).

	♀ Control	♂ Control	♂ Patient Samples												
CpG Island	GM11198	GM7009	9988	9215	8033	8610	8487	8435	6058	7086	7171	7626	9256	9643	6965
43(2)							NA			NA					
62(1)								**							
45(1)*			NA	NA			**								
AR															
41									NA	NA	NA	NA			
79(2)										NA				NA	
154															
27(3)															
20															
31			NA			NA		NA		NA				NA	NA
22							NA								
47															
24(1)										NA		NA			
27(4)															
186															
44(1)															
109															
76													NA	NA	NA

\* primers used to assess methylation status only flank a single HpaII restriction cut site

\*\* mock digest did not amplify

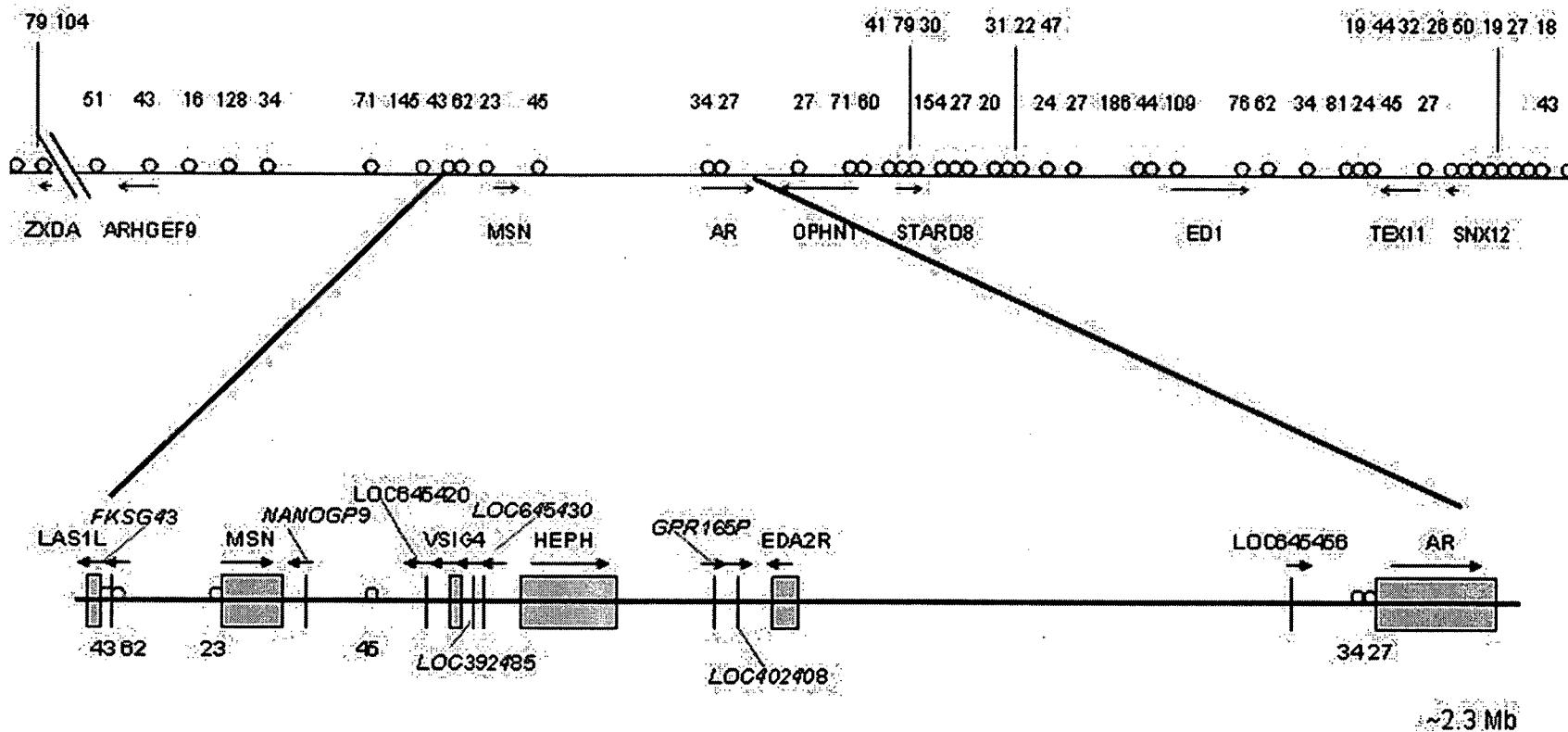
The methylation seen for CpG 43(2) was mainly partial methylation and this CpG island was not found to be methylated in the two FL cell lines. CpG 62(1) is associated with a pseudogene, therefore, the methylation is unlikely to be the result of targeted methylation. There are other genes located near CpG 62(1) and the abnormal methylation of this CpG island is likely the consequence of methylation spreading. Figure 3.20 shows a more detailed map of this second 2.3 Mb sub-region of interest.

### **3.3) Results – B.) Gene Expression**

The goal of the thesis was to identify a lymphoma candidate tumour suppressor gene(s). The first part of the study determined the extent of the abnormal hypermethylation on the X chromosome in the vicinity of the AR and identified two sub regions of interest. However, the regions were relatively large, at 2.3 and 1.4 Mb, and examination of the methylation status in additional lymphoma cell lines and lymphoma primary tumour samples failed to decrease the size of the region of interest. The next step in accomplishing the thesis goal was to examine the expression status of genes located in the sub region. An ideal candidate tumour suppressor gene is a gene that is normally expressed in lymphoblasts but is silenced or experienced reduced levels in lymphomas. Expression and methylation analysis were for the most part done simultaneously and therefore not all the genes whose expression status was determined, were located in the two identified sub-regions of interest. In total, the expression status of ten genes was investigated and one potential candidate tumour suppressor gene was identified.

#### ***3.3.1) Genes in the Region***

Table 3.10 is a complete list of all the genes found in the 8Mb region surrounding the AR. The base pair locations of the genes do not correspond with the base pair location of the associated CpG islands shown in table 3.1. This is because the information was taken from two different databases, the gene information came from NCBI's Entrez Gene (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene>) and the CpG island information from the UCSC Genome Browser. The sequence information used by these two databases was derived from different genome assemblies, the UCSC used the May 2004 (hg17) freeze and the NCBI used build 36 of the human genome assembly.



**Figure 3.20** Map of second sub-region of interest. This region spans ~2.3 Mb upstream of AR at Xq11.2-q12. This region was not as densely abnormally methylated as the first sub-region in the lymphoma cell lines (CpG 43 and 23 were not methylated, CpG 62 and 27 were abnormally methylated, CpG 34 showed variation in methylation and was not methylated in female control, and CpG 45 was normally methylated in SUDHL3 and DoHH2 lymphoma cell lines). However, the few CpG islands in this region whose methylation status was determined in lymphoma patients, showed a high degree of abnormal methylation, with CpG 43 showing abnormal methylation in 73% of patients, CpG 62 with abnormal methylation in 85% and CpG 45 was abnormally methylated in 82% of patients. The map is not drawn to scale but shows the approximate location of all the CpG islands (as defined by UCSC Genome Browser) and known genes (taken from Entrez Gene) (pseudogenes are italicized) relative to each other.

**Table 3.10** Genes located in the 8Mb region in Xq11.2-q12. The RefSeq status definitions are as follows: model RefSeq records are predicted by genome sequence analysis; predicted RefSeq records have evidence for a protein based on existence of cDNA clones, ESTs or homology, but the record has not been individually reviewed; provisional RefSeq records are one that have not yet been individually reviewed, but the existence of a protein and transcript are well supported; reviewed RefSeq records are ones whose sequence data and the literature have been reviewed by NCBI staff or collaborators; and validated RefSeq records have been reviewed but have not undergone a final review (<http://www.ncbi.nlm.nih.gov/RefSeq/key.html>).

Gene ID	Gene	Gene Description	Orientation	Location	RefSeq Status
653588	LOC653588	similar to auto-antigen La	+	61,915,503–61,916,526	model, pseudogene
645251	LOC645251	similar to chromobox protein homolog 1	-	62,449,075–62,435,721	model, pseudogene
139886	LOC139886	hypothetical protein	-	62,487,936–62,483,832	predicted, protein coding
23229	ARHGEF9	Cdc42 guanine nucleotide exchange factor (GEF) 9	-	62,891,718–62,771,573	provisional, protein coding
8251	HNRPDP	heterogeneous nuclear ribonucleoprotein D	+	63,180,849–63,181,884	provisional pseudogene
139285	RP11-403E24.2	hypothetical protein FLJ39827	-	63,342,349–63,321,723	predicted protein coding
142689	ASB12	ankyrin repeat and SOCS box containing 12	-	63,362,228–63,360,801	reviewed protein coding
55613	MTMR8	myotubularin related protein 8	-	63,532,036–63,404,686	provisional protein coding
645338	LOC645338	similar to SHC transforming protein 1	+	63,570,542–63,570,754	model protein coding
392481	LOC392481	similar to src homology 2 domain – containing transforming protein C	+	63,569,503–63,570,102	model protein coding
442455	LOC442455	similar to keratin, type II cytoskeletal 8	+	63,759,574–63,761,180	model pseudogene
55906	KIAA1166		-	64,113,061–64,052,986	provisional protein coding
645374	LOC645374	similar to UPF0308 protein C9orf21	-	64,494,149–64,493,491	model pseudogene
645381	LOC645381	similar to transducin – like enhancer protein 1	+	64,543,777–64,545,904	model pseudogene
645388	LOC645388	similar to Adaptor – related protein complex 1, mu 2 subunit	+	64,605,667–64,607,052	model pseudogene
340554	ZC3H12B	zinc finger CCCH-type containing 12B	+	64,625,431–64,644,492	predicted protein coding

<b>Gene ID</b>	<b>Gene</b>	<b>Gene Description</b>	<b>Orientation</b>	<b>Location</b>	<b>RefSeq Status</b>
81887	LAS1L	LAS1-like	-	64,671,392-64,649,191	provisional protein coding
83957	FKSG43		-	64,689,026-64,607,227	provisional pseudogene
4478	MSN	Moesin, membrane-organizing extension spike protein 1	+	64,804,236-64,878,518	reviewed protein coding
349386	NANOGP9	Nanog homeobox pseudogene 9	-	64,910,395-64,908,883	provisional pseudogene
645420	LOC645420	similar to RNA binding motif protein, X linked	-	65,093,763-65,093,347	model protein coding
11326	VSIG4	V-set and immunoglobulin domain containing 4	-	65,176,610-65,158,305	provisional protein coding
392485	LOC392485	similar to ataxin 7-like 3	-	65,193,149-65,192,154	model pseudogene
645430	LOC645430	hypothetical protein LOC645430	-	65,212,912-65,211,001	model pseudogene
9843	HEPH	hephaestin	+	65,299,388-65,403,956	reviewed protein coding
392486	GPR165P	G protein-coupled receptor orphanA7	+	65,577,304-65,578,015	provisional pseudogene
402408	LOC402408	similar to Pyruvate kinase, isozyme M1/M2	+	65,634,366-65,635,991	model pseudogene
60401	EDA2R	ectodysplasin A2 receptor	-	65,752,598-65,732,204	reviewed protein coding
645456	LOC645456	similar to zinc finger protein 681	+	66,510,894-66,511,828	model, protein coding
367	AR	androgen receptor	+	66,680,599-66,860,844	reviewed protein coding
4983	OPHN1	oligophrenin 1	-	67,570,372-67,179,440	reviewed protein coding
5231	PGK1P1	phosphoglycerate kinase 1 pseudogene 1	-	67,208,413-67,206,644	provisional pseudogene
643374	LOC643374	hypothetical protein LOC643374	-	67,464,166-67,463,792	model protein coding
286451	YIPF6	Yip 1 domain family member 6	+	67,635,611-67,669,026	provisional protein coding
9754	STARD8	START domain containing 8	+	67,784,236-67,862,403	provisional protein coding
389866	SERBP1P	SERPINE 1 mRNA binding protein 1 pseudogene	+	67,919,941-67,922,999	provisional pseudogene
1947	EFNB1	ephrin- B1	+	67,965,556-67,978,726	reviewed protein coding

Gene ID	Gene	Gene Description	Orientation	Location	RefSeq Status
64219	PJA1	praja 1	-	68,301,997-68,297,429	validated protein coding
360187	CYCSP43	cytochrome c, somatic pseudogene	-	68,622,220-68,621,909	provisional pseudogene
27112	TMEM28	transmembrane protein 28	+	68,641,803-68,669,076	validated protein coding
1896	EDA	ectodysplasin A	+	68,752,636-69,176,047	reviewed protein coding
643426	LOC643426	similar to CCR4-NOT transcription complex subunit 7	-	69,074,756-69,073,743	model protein coding
158835	DGAT2L4	diacylglycerol O-acyltransferase 2 like 4	-	69,186,513-69,177,117	validated protein coding
139562	OTUD6A	OTU domain containing 6A	+	69,199,066-69,200,754	provisional protein coding
3476	IGBP1	immunoglobulin (CD79A) binding protein 1	+	69,270,043-69,302,899	reviewed protein coding
347516	DGAT2L6	diacylglycerol O-acyltransferase 2-like 6	+	69,314,061-69,342,276	provisional protein coding
158833	DGAT2L3	diacylglycerol O-acyltransferase 2-like 3	+	69,371,271-69,376,865	provisional protein coding
5030	P2RY4	pyrimidinergic receptor P2Y, G-protein coupled, 4	-	69,396,379-69,394,741	reviewed protein coding
407	ARR3	arrestin 3, retinal (X-arrestin)	+	69,404,927-69,418,415	provisional protein coding
347517	RAB41	member RAS oncogene family	-	69,418,793-69,421,577	provisional protein coding
51248	PDZD11	PDZ domain containing 11	-	69,426,597-69,422,686	provisional protein coding
24137	KIF4A	kinesin family member 4A	+	69,426,687-69,567,546	provisional protein coding
54857	GDPD2	glycerolphosphodiesterter phosphodiesterase domain containing 2	+	69,559,716-69,569,956	provisional protein coding
174	DLG3	discs, large homolog 3	+	69,581,544-69,639,258	provisional protein coding
56159	TEX11	testis expressed sequence 11	-	70,045,292-69,665,515	reviewed protein coding
84889	SLC7A3	solute carrier family 7	-	70,067,655-70,062,154	provisional protein coding
653658	LOC653658	similar to ribosomal protein S23	+	70,092,066-70,099,924	model protein coding

<b>Gene ID</b>	<b>Gene</b>	<b>Gene Description</b>	<b>Orientation</b>	<b>Location</b>	<b>RefSeq Status</b>
29934	SNX12	sorting nexin 12	-	70,204,956-70,197,513	reviewed protein coding
4303	MLLT7	myeloid/lymphoid or mixed lineage leukemia	+	70,232,935-70,240,109	provisional protein coding
158830	LOC158830	similar to Ab2-183	-	70,243,145-70,240,562	predicted protein coding
3561	IL2RG	interleukin 2 receptor, gamma	-	70,248,128-70,243,984	reviewed protein coding
9968	MED12	mediator of RNA polymerase II transcription, subunit 12 homolog	+	70,255,298-70,278,872	provisional protein coding
54413	NLGN3	neuroligin 3	+	70,281,436-70,307,776	validated protein coding
2705	GJB1	gap junction protein, beta 1, 32kDa	+	70,359,801-70,361,769	provisional protein coding
9203	ZMYM3	zinc finger, MYM-type 3	-	70,391,148-70,376,199	validated protein coding

There are a total of 65 genes in the 8 Mb region, 16 (25%) of which are pseudogenes. Twenty-nine (45%) of the genes have at least one associated CpG island and seven (11%) of these genes have two associated CpG islands (shown in table 3.11). The majority (89%) of the genes with associated CpG islands are protein coding. The first identified sub-region is 1.4 Mb, spans from STARD8 to OTUD6A and encompasses 10 genes (8 genes and 2 pseudogenes) as shown in figure 3.15. Seven of the ten genes are associated with a CpG island. The larger 2.3 Mb sub-region is very CpG island poor and spans from LAS1L to AR, flanking a total of 14 genes (8 genes and 6 pseudogenes) as shown in figure 3.20. Only 4 out of the 14 genes have an associated CpG island.

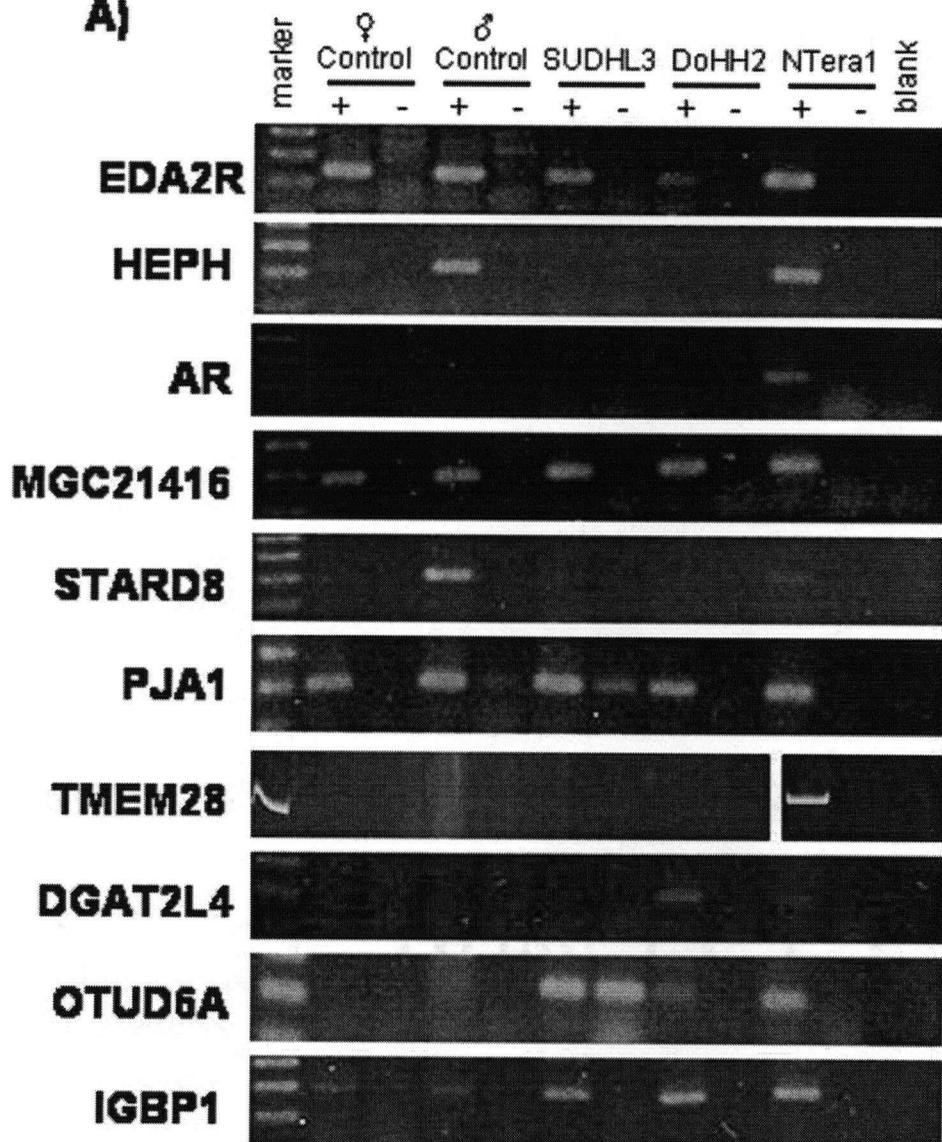
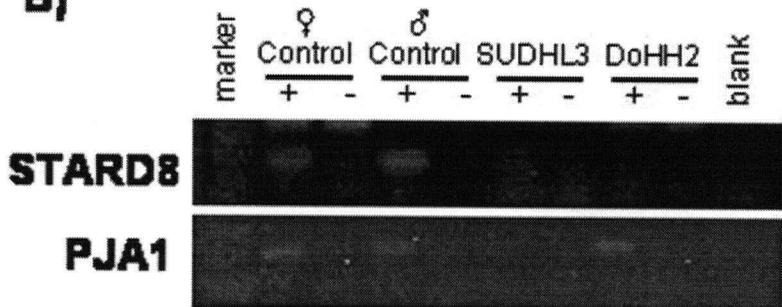
### ***3.3.2) Expression of Genes in Lymphoblast and Lymphoma Cell Lines***

The expression status of 10 genes was determined in two lymphoblast control cell lines, one male and one female, and in two male follicular lymphoma cell lines, SUDHL3 and DoHH2, using reverse-transcriptase PCR (RT-PCR). The data is shown in figure 3.21 and summarized in table 3.12. Three of the genes (HEPH, EDA2R and AR) are located in the second sub-region of interest, and five genes (STARD8, PJA1, TMEM28, DGAT2L4 and OTUD6A) are found in the first sub-region of interest located downstream of AR. The remaining two genes flank either side of the first 1.4 Mb sub-region.

One of the 10 genes was the AR itself. The AR was one of the few loci that consistently showed abnormal hypermethylation in both the lymphoma cell lines and in the patient samples (92%). AR, as mentioned above in the introduction, is a steroid activated transcription factor [12]. Mutations of AR have been documented and result in androgen insensitivity syndrome and in Kennedy spinal and bulbar muscular atrophy (SBMA). AR mutations have also been shown to have a role in prostate [14, 15] and male breast cancers [13]. Despite the AR involvement in some cancers, it is believed to be a poor lymphoma candidate tumour suppressor gene because it is not normally expressed in lymphoblasts. This is consistent with what was found in the lymphoblast cell lines. AR appeared to be expressed in the pluripotent embryonal carcinoma Nteral1 cell line, but not in any of the lymphoblast or lymphoma cell lines.

**Table 3.11** Genes and their associated CpG islands located in the 8Mb region located at Xq11.2-12.

Gene	Associated CpG Island
LOC139886	CpG 51
ARHGEF9	CpG 43 (1)
RP11-403E24.2	CpG 16
MTMR8	CpG 128
LAS1L	CpG 43 (2)
FKSG43	CpG62 (1)
MSN	CpG 23
AR	CpG 34 (2), CpG 27 (1)
OPHN1	CpG 27 (2), CpG 71 (2)
YIPF6	CpG 60
STARD8	CpG 41, CpG 79 (2)
SERBP1P	CpG 30
EFNB1	CpG 154, CpG 27 (3)
PJA1	CpG 47
TMEM28	CpG 186
EDA	CpG 109
OTUD6A	CpG 76
IGBP1	CpG 62 (2)
PDZD11	CpG 34 (3)
KIF4A	CpG 34 (3)
GDPD2	CpG 81
DLG3	CpG 24 (2), CpG 45 (2)
SLC7A3	CpG 27 (5)
SNX12	CpG 19 (1), CpG 44 (2)
MLLT7	CpG 32, CpG 26
MED12	CpG 50
NLGN3	CpG 19 (2)
GJB1	CpG 18
ZMYM3	CpG 43 (3)

**A)****B)**

**Figure 3.21** Expression status of 10 different genes located in the region around the AR on the X chromosome. (+ represents +RT reactions and - represents -RT reactions.) Expression was looked at in a female and male lymphoblast control, positive control pluripotent embryonal carcinoma (NTera1) and two lymphoma (SUDHL3 and DoHH2) cell lines. **A)** Expression results for 10 genes are shown. There is a weak band for the female control cell line sample with the HEPH primers suggesting that *HEPH* is expressed in all control cell lines. *DGAT2L4* and *OTUD6A* appear to only be expressed in one of the lymphoma cell lines (DoHH2) and in the positive control pluripotent embryonal carcinoma (NTera1) (weak band). The band in both the RT and -RT of SUDHL3 for *OTUD6A* suggests that there is potential DNA contamination and expression status can not be determined. Weak bands in the -RT for *PJA1* in the male control and SUDHL3 also indicate genomic DNA contamination; however an earlier RT-PCR done with this primer (shown in part B) indicate that this gene is expressed in both of these cell lines. *JGBP1* is expressed in all examined cell lines. Bands representing the lymphoblast controls appear to be relatively weak. There was some gDNA contamination in these samples and a larger gDNA band was also amplified, and may explain the weaker bands. **B)** Expression results for *STARD8* and *PJA1* gene expression assay. In these cases, expression in a positive control was not looked at. There is a very faint band for SUDHL3 for *PJA1* suggesting that this gene is expressed in all of the examined cell lines. The expression results of *STARD8* represent the initial results that suggested the gene was expressed in control lymphoblast cell lines and silenced in lymphoma cell lines. However, in part A) of this figure, later results showed that this gene is only expressed in the male lymphoblast control.

**Table 3.12** Summary of the gene expression results 10 selected genes located in the 8Mb region at Xq11.2-12.

Gene	Expression			
	Lymphoblast Control Cell Lines		Follicular Lymphoma Control Cell Lines	
	♀	♂	SUDHL3	DoHH2
HEPH	yes	yes	no	no
EDA2R	yes	yes	yes	yes
AR	no	no	no	no
YIPF6	yes	yes	yes	yes
STARD8	yes	yes	no	no
PJA1	yes	yes	yes	yes
TMEM28	no	no	no	no
DGAT2L4	no	no	no	yes
OTUD6A	no	no	NA	yes
IGBP1	yes	yes	yes	yes

Yip1 domain family, member 6 (YIPF6), also known as MGC21416, is a provisional protein coding gene. The expression status of this gene was not examined for its potential role as a candidate tumour suppressor gene, instead, it was tested because of the methylation results of its associated CpG island, CpG 60. In this study, CpG 60 was the first island that showed methylation skipping as CpG islands located both up and downstream of CpG 60 appeared to be abnormally methylated in the lymphoma cell lines. The expression status of YIPF6 in both the control lymphoblast and the lymphoma cell lines was determined to see if there was an association between the methylation status of the associated CpG island and gene expression. YIPF6 appeared to be expressed in the lymphoblast cell lines and remained expressed in the lymphoma cell lines and therefore gene expression correlated with the lack of methylation.

Of the remaining eight genes, two have an expression pattern of what is expected of a tumour suppressor gene. The remaining six genes are either not expressed in the control lymphoblasts (TMEM28, DGAT2L4 and OTUD6A) or are expressed and stay expressed in the lymphoma cell lines (EDA2R, PJA1, and IGBP1).

EDA2R, also known as XEDAR, codes for an ectodysplasin A2 receptor. This protein is a type III transmembrane receptor and is part of the tumour necrosis factor receptor (TNFR) family. The encoded protein binds to the EDA-A2 isoform of ectodysplasin, a gene which when mutated results in a loss of hair, sweat glands and teeth (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene>). EDA2R has not been implicated in any cancers, however it is a death receptor that is able to induce apoptosis through a caspase 8- and Fas-associated death domain (FADD)- dependent manner [116]. EDA2R is highly expressed during embryonic development, where it is believed to have a role in apoptosis. This gene appeared to be expressed in the lymphoblast cell lines as well as in the two lymphoma cell lines.

Praja 1 (PJA1) encodes a RING-H2 protein with ubiquitin E3 ligase activity [117]. PJA1 plays a role in cell proliferation and apoptosis (reviewed in [118]) and has been shown to interact with the Smad4 adaptor protein, ELF (embryonic liver fordin) in a TGF- $\beta$  (transforming growth factor-beta) dependent manner (reviewed in [118]). Smads are mediators of the TGF- $\beta$  signaling pathway and both Smad4 and ELF are tumour suppressor genes. Ubiquitination of ELF by PJA1 leads to ELF degradation and PJA1

overexpression in gastrointestinal cancers results in reduced ELF expression [118]. In this study, PJA1 appeared to be expressed in all the examined cell lines. This expression profile as well as its potential oncogenic-like function makes PJA1 a poor candidate tumour suppressor gene.

Similar gene expression results were found for Immunogloblin (CD79A) binding protein 1 (IGBP1). The encoded phosphoprotein is part of Ig receptor-mediated signal transduction in B-cells [119] and is a noncatalytic subunit of protein phosphatases 2A (PP2A), PP4, and PP6 complexes (reviewed in [120]). Studies in mice suggest that IGBP1 has a role in repressing p53 mediated apoptosis [120] and is potentially involved in the transformation of growth factor dependent to growth factor independent lymphoid tumours [119]. IGBP1 is not a good candidate tumour suppressor gene for the same reasons as PJA1.

TMEM28, DGAT2L4 and OTUD6A all lacked gene expression in the control lymphoblast cell lines. TMEM28 is a validated transmembrane protein coding gene whose function is unknown. DGAT2L4 codes for a protein that is a member of the diacylglycerol O-acyltransferase 2 family and mediates the synthesis of triacylglycerol and long chain esters [121]. As mentioned above, this gene does not appear to be expressed in the two lymphoblast cell lines, but surprisingly, it does show gene expression in one of the lymphoma cell lines, DoHH2. OTUD6A, OTU domain containing 6A, is a provisional protein coding gene that is also known as HIN-6 protease (HSHIN6). Similar to the expression results for DGAT2L4, OTUD6A appeared to be expressed in the lymphoma cell line DoHH2, but was silent in the remaining cell lines. The RT for SUDHL3 also amplified with the primers used to check for OTUD6A expression, but the -RT also amplified. This suggests there is genomic DNA contamination rather than a cDNA product. There is no evidence to date showing any involvement of these genes in cancer.

The two genes that were observed to have gene expression patterns expected of a tumour suppressor gene are hephaestin (HEPH) and START domain containing 8 (STARD8). Both of these genes appeared to be normally expressed in the lymphoblast cell lines but silenced in the lymphoma cell lines. STARD8 is a good potential candidate tumour suppressor gene and is discussed in more detail below. HEPH on the other hand is

a rather poor candidate due to its function. HEPH codes for a protein involved in iron transport from the epithelial cells of the intestinal lumen into the circulatory system. Reduction of HEPH expression has been shown to be associated with colon cancer where risk of development is associated with iron levels [122]. Loss or reduction of HEPH expression was associated with more advanced colon cancer. One hypothesis is that an increase in the importation of iron and a reduction in the exportation results in high intracellular iron concentrations that could induce cell proliferation and prevent cell adhesion [122]. However, HEPH remains a poor lymphoma tumour suppressor gene candidate because of its role in dietary iron transport is unlikely to be important in the development of lymphoma. The abnormal silencing of this gene is most likely a consequence of the abnormal methylation of the region.

STARD8 codes for a RhoGAP protein (Rho-like GTPase activating protein). The initial gene expression analysis of this gene showed expression in both the male and female lymphoblast cell lines (data shown in part B of figure 3.21), however more recent results have only shown expression in the male lymphoblast cell line. STARD8 expression was also examined at in 2 additional male and 3 additional female cell lines and expression was observed in 2 out of the 3 males and in only one of the female cell lines. This suggests that STARD8 expression is variable and further investigation is needed to determine STARD8's gene expression profile.

STARD8 has not been associated with any cancers, however other RhoGAP protein family members have. RhoGAP proteins targets include Rho/Rac/Cdc42-like small GTPases. Rho family G proteins, which are active when bound by GTP and inactive when bound by GDP, are involved in cytoskeleton formation, cell proliferation and the JNK signaling pathway (<http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>). RhoGAP proteins negatively regulate G proteins by catalyzing the reaction converting G proteins from their active to inactive state (reviewed in [123]). Therefore, abnormal silencing of RhoGAP proteins in cancer can lead to increased activity of the G proteins due to limitations in GTP hydrolysis. An example of a RhoGAP protein acting as a tumour suppressor gene is the DLC-1 gene in multiple myeloma, hepatocellular carcinoma (HCC) and breast cancer ([123-125]). DLC-1 is located at 8p21.3-22, a region that is frequently deleted in HCC, and is involved in inhibiting cell proliferation in HCC

[125] and suppressing metastasis in breast cancer [123]. Both genetic and epigenetic alterations of DLC-1 have been identified [124, 125].

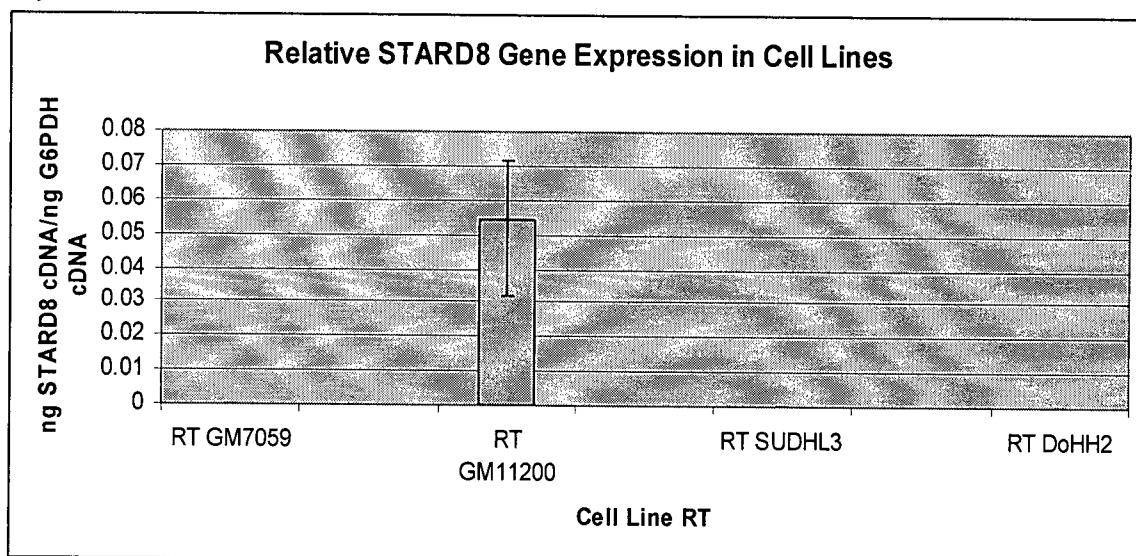
### ***3.3.3) Quantitation of STARD8 Gene Expression***

STARD8 gene expression was quantified using the DNA Engine Opticon 2 qPCR machine from MJ Research and the dye SYBR Green 1, in both lymphoblast and lymphoma cell lines and primary tumour samples. Quantitation of gene expression was done using the standard curve method, where known amounts of DNA are amplified along with the unknown samples. The Opticon Monitor 2 Software formulated a standard curve plotting C(T) cycle against log quantities from the samples with known quantities of DNA and then used this curve to calculate the amount of cDNA present in the unknown samples. STARD8 gene expression was then normalized to G6PDH (hexose-6-phosphate dehydrogenase (glucose 1-dehydrogenase)) expression, a widely expressed housekeeping gene located at 1p36, to account for the difference in total RNA quantities between samples. The quantitation results are shown in figure 3.22. This figure shows STARD8 gene expression above background level normalized to G6PDH.

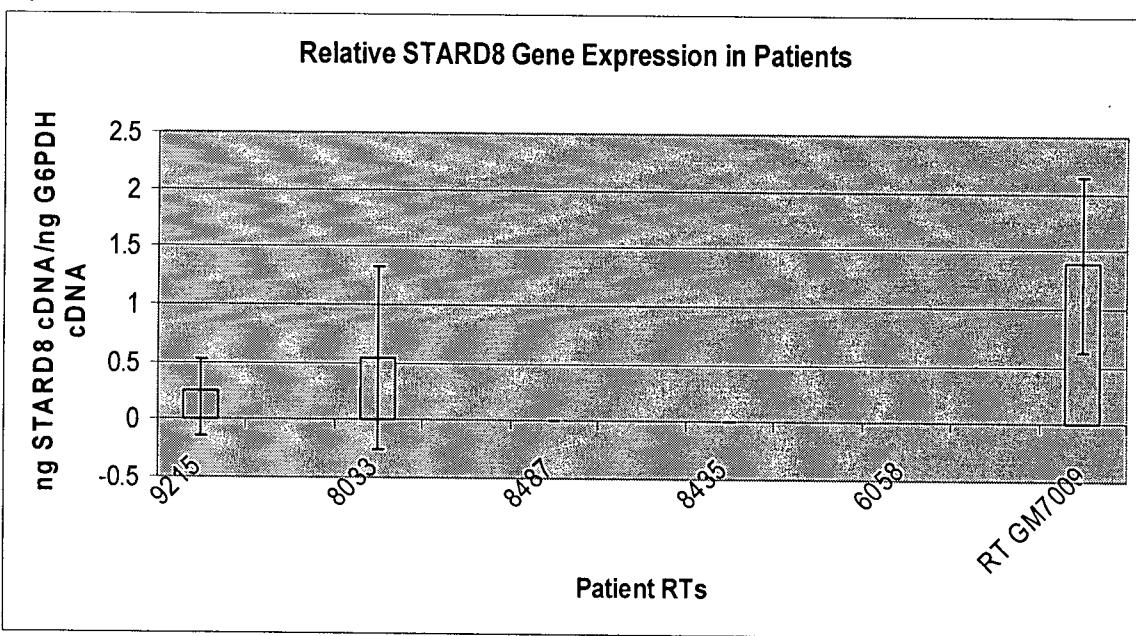
Primer dimers proved to be very problematic in quantifying STARD8 gene expression since the qPCR software is unable to differentiate between primer dimers and amplified product. Primer dimers were more often seen in cases where there was no amplified PCR product. In these situations, the software calculated the levels of primer dimer as gene expression and therefore these readings do not accurately reflect expression levels in the cell line/patient samples in question. For example, according to the calculations made by the software, low quantities of STARD8 cDNA were detected in the negative blank controls, but when the qPCR products were run out on a gel, no product was seen. Two different primer pairs were used to quantitate STARD8 expression, however, both produced a large amount of primer dimer. To analyze the qPCR data, the amount of 'cDNA' located in the blank and negative RT samples were called background expression and the difference between the quantity of cDNA detected in the blank negative control and each of the RT samples was taken.

STARD8 appeared to be expressed in the male control cell lines GM11200 and GM7009, and lacked expression in the female control cell line and the lymphoma cell

A)



B)



**Figure 3.22** Quantitation of STARD8 expression in lymphoblast and lymphoma cell lines and in lymphoma patients.

**A)** Quantitation of STARD8 in a female (GM7059) and male (GM11200) lymphoblast cell line and two lymphoma cell lines (SUDHL3 and DoHH2). The expression shown is above the background and normalized to G6PDH expression. **B)** Quantitation of STARD8 in five lymphoma patient samples and a control male lymphoblast cell line (GM7009). The expression shown is above background levels and has been normalized to G6PDH expression.

lines, SUDHL3 and DoHH2. cDNA was successfully obtained from five lymphoma patient samples (9215, 8033, 8487, 8435 and 6058) and STARD8 expression was only detected in two of these primary tumour tissues. STARD8 expression, however, seems to be reduced relative to the male control cell line in 9215. The other three patients do not express STARD8. Overall, the difference in expression between the patients and the lymphoblast cell line was significant ( $t=4.65$ ,  $p=0.0048305$ ).

The expression observed in the male control cell line might not reflect the expression levels found in normal lymphoblast cells. The optimal situation, in order to determine whether there is a change in STRAD8 expression in lymphomas, would be to compare gene expression in primary tumour tissue with patient matched normal lymphoblast or germinal center B-cells. Another concern is the large error bars seen in figure 3.22 for patient samples 9215 and 8033. These are likely a result of pipetting error caused by adding the RT of each sample individually to each qPCR reaction. More replicates are needed to see if these results are reproducible and reflect what is occurring in lymphoma.

### **3.4.) Discussion of DNA Methylation and Gene Expression Results**

The methylation and gene expression analysis of the 8 Mb region on the X chromosome has shown both promising and unexpected results and led to the identification of one potential candidate tumour suppressor gene. Discussed below is a summary of all the results.

#### ***3.4.1) Methylation in Lymphoblast and Mouse/Human Hybrid Cell Lines***

As previously mentioned, the CpG islands with unexpected methylation results were re-examined in additional control cell lines. The methylation status of the CpG islands was not always reproduced within a cell line and/or between control cell lines. Two possible reasons for this, assay and interindividual variability, are discussed below.

##### ***3.4.1.1) Assay Variability***

The methods used to assess methylation status rely on complete DNA cutting with the methylation sensitive restriction enzyme HpaII. This was checked by amplifying

the digested DNA with a pair of primers known to flank an unmethylated region (located in the pseudoautosomal region of the X chromosome) and several HpaII restriction cut sites. It was assumed that there was complete cutting when the digested DNA failed to amplify with the cutting control primers; however, this could just indicate that the specific region was completely digested but not the entire genome. This might result in false positives when assessing the methylation status of CpG islands located elsewhere.

The differences in methylation results for a single locus in a cell line might also be caused by the sensitivity of the assay. If 98-99% of the genome was completely cut with the methylation sensitive restriction enzyme, PCR might still be able to amplify the remaining 1-2% that failed to cut.

#### ***3.4.1.2) Interindividual Variability***

Another possible explanation for the variation seen in the methylation status of some of the CpG islands between the lymphoblast control and the hybrid cell lines is tissue specific methylation patterns. The hybrid cell lines were derived from fibroblast cells and the methylation pattern in lymphoblast and fibroblast cells might differ. Another possibility is that the hybrids failed to retain methylation.

In some cases, the methylation seems to be cell line specific, even between the same cell types. This suggest that CpG island methylation patterns within a cell type are heterogeneous or that they differ between individuals since the cell lines were derived from different people. These differences potentially reflect stochastic changes in the cells (*i.e.* gain of methylation by *de novo* methylation or loss of methylation by failure of methylation maintenance) or differential sequence activity.

#### ***3.4.1.3) Methylation of the Xa***

The normal methylation seen in the male control cell line for 10 CpG islands in the first methylation survey does not appear to be the result of cell line artifacts. Full or partial methylation was seen in at least one additional male control cell line for all but one CpG island. The lack of methylation in all male controls is likely the result of interindividual variability as described above. CpG 45, the single CpG island that showed a lack of methylation in all the male control cell lines in the second study, including

GM11200 where the ‘normal’ methylation was originally observed, is likely unmethylated in most cell lines. The methylation first observed is probably a false positive caused by the limitations of the assay as described above.

#### ***3.4.1.4) Hypomethylation of the Xi***

The CpG islands that were found to be unmethylated in the female control cell lines in the original methylation survey are the only ones where the initial results were confirmed in the additional control cell lines. The consistent observation of an unmethylated state of these islands suggests that this is not an artifact of the cell line induced by the tissue culture environment or transformation.

#### ***3.4.2) Methylation in the 1.4 Mb Sub-Region in Three Additional Cell Lines***

Both follicular and mantle cell lymphoma cell lines showed abnormal methylation in the 1.4 Mb sub-region. The FL cell lines showed methylation of all the assessed CpG islands, while there was variation between the less densely methylated MCL cell lines. A possible explanation for the differences observed between JVM-2 and Z138 and HBL-2 is that even though all three are classified as mantle cell lymphoma, they may belong to different sub-classes. Some cancer subtypes have recently been identified by looking at the genetic level of the tumour instead of relying on histology [126-128]. In addition, a recent paper has shown that JVM-2 differs from the other two MCL in other characteristics as well, including growth rate, survival time, and cyclin D1, D2 and p16 expression levels [129]. Overall, four out of the five assessed lymphoma cell lines showed abnormal hypermethylation in the 1.4 Mb sub-region, supporting the hypothesis that this is a region of interest in lymphomas. However, the determination of the methylation status in these cell lines did not complete the outset goal, to reduce the region of interest, since the majority of the CpG islands in the sub-region were methylated.

#### ***3.4.3) Methylation in the 1.4 Mb Sub-Region in Lymphoma Patient Samples***

Methylation skipping was a common occurrence in the 1.4 Mb region in both the patient samples and the MCL cell lines. The assessment of the methylation status of the CpG islands in the patient samples, like in the case of the MCL cell lines, was not able to

reduce the size of the region of interest. Instead, the results indicate that there are two potential regions of interest, separated by 0.6 Mb. Table 3.13 is a summary of the methylation results for both the patient samples and the lymphoma cell lines. Many MCL and DLBC lymphoma patients experienced full abnormal methylation of CpG62(1) and AR, and it is possible that a candidate tumour suppressor gene lies somewhere in the vicinity of these islands rather than in the 1.4 Mb region located downstream of AR.

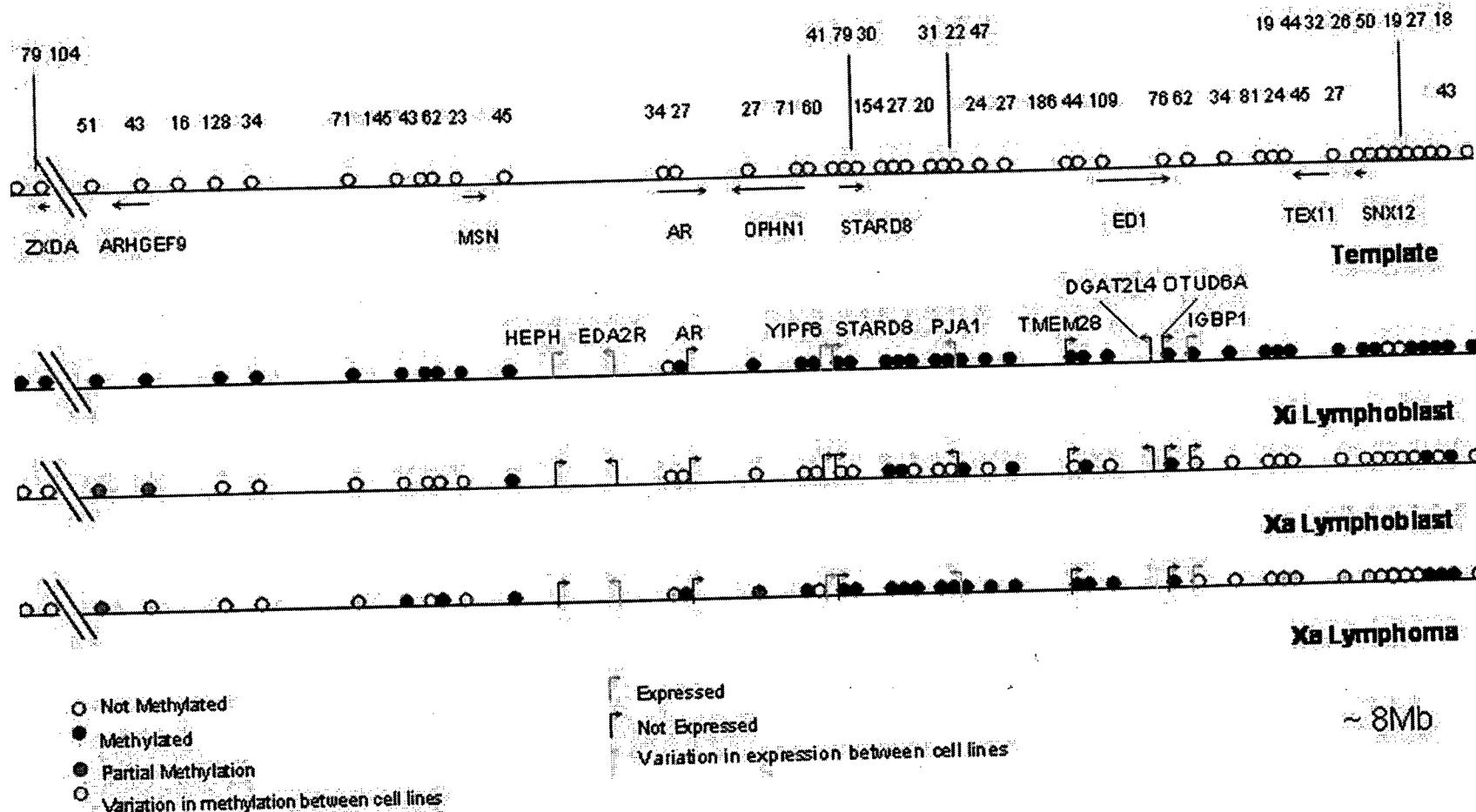
#### ***3.4.4) Summary of DNA Methylation and Gene Expression Analysis***

The methylation and gene expression results are shown together in figure 3.23 and summarized in table 3.14. Of the ten genes whose expression status was determined, seven had associated CpG islands. Two of these CpG islands (CpG 60 and 62 (2)) were found to be unmethylated in both the control lymphoblast as well as the lymphoma cell lines and the associated genes were expressed in all of the examined cell lines. Both AR and TMEM28 are associated with CpG islands that were abnormally methylated, however the genes were not expressed in the control lymphoblast cell lines. Two genes, PJA1 and OTUD6A, are associated with CpG islands that showed ‘normal’ methylation. PJA1 appeared to be expressed in the lymphoblast cell lines as well as in the lymphoma cell lines despite the methylation of its associated CpG 47 island. CpG 47 was ‘normally’ methylated in all four male lymphoblast control cell lines but it lacked methylation in the Xa mouse/human hybrid cell line. CpG 47 spans the promoter region of PJA1, and together, this suggests that gene inactivation might not always be necessary prior to methylation and that methylation alone is not sufficient to cause silencing. It would therefore be interesting to see if PJA1 lacks other marks of inactive chromatin. OTUD6A is associated with CpG 76, which consistently showed ‘normal’ methylation in all examined cell lines. OTUD6A was not expressed in either the lymphoblast cell lines and showed variable expression in the lymphoma cell lines.

The last gene whose expression status was determined is STARD8. STARD8 is a good candidate tumour suppressor gene for three reasons. The CpG island methylation and gene expression patterns as well as its role as negative regulator of G proteins all meet the criteria of a tumour suppressor gene. STARD8 is associated with two CpG islands, CpG 41 and 79, which are located in an intron. Both CpG islands experienced

**Table 3.13** Percentage of lymphoma patient samples and lymphoma cell lines showing methylation of the CpG islands found in the 2 sub-regions of interest and the AR. CpG islands with partial methylation were considered to be methylated in calculating the percent of patient samples and cell lines showing methylation.

CpG Island	% Methylated Patient Samples	n	% Methylated Lymphoma Cell Lines	n
CpG 43 (2)	73	11	0	2
CpG 62 (1)	85	13	100	2
CpG 23	NA	NA	0	2
CpG 45 (1)	82	11	100	2
CpG 34 (2)	NA	NA	50	2
CpG 27 (1)	NA	NA	100	2
CpG AR	92	13	100	5
CpG 41	67	9	80	5
CpG 79 (2)	0	5	100	5
CpG 154	0	6	100	5
CpG 27 (3)	100	6	100	5
CpG 20	92	13	80	5
CpG 31	14	7	80	5
CpG 22	67	12	80	5
CpG 47	100	6	100	5
CpG 24 (1)	18	11	80	5
CpG 27 (4)	100	6	80	5
CpG 186	NA	NA	100	2
CpG 44 (1)	NA	NA	80	5
CpG 109	NA	NA	100	2
CpG 76	100	10	100	5



**Figure 3.23** CpG island methylation and gene expression results for an 8 Mb region located at Xq11.2-12.

The top map represents a template of the region of interest and shows the approximate location of all the CpG islands examined as well as CpG 16 and 30. The location of a few, but not all of the genes in this region, are also shown to serve as reference points. Due to the size of the region of interest, the map is only drawn approximately to scale. The two hatch marks represent the crossing of the centromere and CpG 79 and 104 are located on Xp. The second map represents the methylation assay and gene expression results for the inactive X chromosome as inferred from the female control cell line and the third map represents the active X methylation results taken from the male control cell line. The fourth and last map is a summary of the methylation and gene expression results of the CpG islands in the two lymphoma cell lines, SUDHL3 and DoHH2.

**Table 3.14** Summary of the gene expression status of 10 genes and the methylation status of their associated CpG islands in lymphoblast and lymphoma cell lines.

Gene	Expressed in Lymphoblast	Expressed in Lymphoma	CpG Island	Lymphoma Methylation Status
HEPH	yes	no	NA	NA
EDA2R	yes	yes	NA	NA
AR	no	no	CpG 34 (2), CpG 27 (1)	variable methylation and methylated
YIPF6	yes	yes	CpG 60	not methylated
STARD8	yes	no	CpG 41, CpG 79 (2)	methylated
PJA1	yes	yes	CpG 47	'normal' methylation
TMEM28	no	no	CpG 186	methylated
DGAT2L4	no	variable	NA	NA
OTUD6A	no	variable	CpG 76	'normal' methylation
IGBP1	yes	yes	CpG 62 (2)	not methylated

abnormal methylation in the lymphoma cell lines and CpG 41 was abnormally methylated in 67 % of patient samples. STARD8 appears to be silenced in the lymphoma cell lines, however the expression status needs to be determined in additional control cell lines to see if STARD8 is normally expressed.

HEPH and EDA2R are both not associated with CpG islands, however the methylation status of the promoter region of these two genes was investigated (see section 3.2.2.1.2 Gene Promoter Methylation). The promoter of EDA2R appeared to be partially methylated in the male control and variable methylated between the two lymphoma cell lines. Despite this methylation, the gene was expressed in both the control and lymphoma cell lines. This might indicate that promoter methylation is not enough to silence a gene, potentially because the region is not CpG rich and therefore methylation might be below a certain threshold level needed to induce transcriptional silencing. Perhaps the continuous transcription of the gene prevents the promoter from becoming densely methylated and recruiting other features of inactive chromatin. HEPH's promoter, on the other hand, did experience abnormal methylation and transcriptional silencing in the lymphoma cell lines.

The methylation status of STARD8's promoter region was also determined since STARD8's associated CpG islands are located within an intron of the gene instead of the promoter. CpG islands located in the promoter and first exon region of genes have been shown to have a role in gene regulation, however the ability of intronic CpG islands to regulate gene expression has not been greatly investigated. However, some evidence exists that suggests these CpG islands are also able to induce the formation of inactive chromatin structure and reduce gene expression [130]. The methylation status of STARD8's promoter was found to be variable in the lymphoma cell lines, while both the female and male control lymphoblast cell line showed methylation. The male control cell line (GM7009) used to determine the methylation status of the promoter was not the same male control that lacked STARD8 expression (GM7033). It is possible, just like with EDA2R, that methylation of just the promoter rather than a CpG island within the promoter region, is not sufficient to recruit features of inactive chromatin. The normal methylation in the male control could potentially act as seeds of methylation in malignant cells resulting in methylation spread to the downstream CpG islands that might have a

role in regulating STARD8 expression. In addition, it is also possible that the methylation seen in the male control is an artifact of the cell line as methylation specific to cultured cell lines is known to occur. The increase in cell line methylation is potentially a consequence of selection for rapid cell division in the tissue culture environment by silencing genes inhibiting proliferation [131].

## **Chapter 4: Discussion**

The determination of the methylation status of CpG islands and the expression status of ten genes within an 8 Mb region in the vicinity of AR in lymphoma has led to the identification of one potential candidate tumour suppressor gene, STARD8. Even though STARD8 looks like a good candidate, further investigation is required to provide additional supporting evidence. This includes examining the expression status of STARD8 in additional control cell lines, to determine how variable the gene expression really is, as well as in additional lymphoma patients. As mentioned previously, the ideal situation would be to look at STARD8 expression in paired normal and primary tumour tissues from patients to determine whether there is a change in expression. To investigate if methylation is responsible for the silencing of STARD8 gene expression, lymphoma cells could be treated with agents preventing methylation, like 5-aza-CR, or histone deacetylation inhibitors like trichostatin A (TSA). The re-expression of STARD8 in these cells would provide further evidence for the role of STARD8 as a tumour suppressor. Preliminary results suggests that treatment of lymphoma cell lines with Zebularine is able to reactivate STARD8 expression. Zebularine is a DNA methylation inhibitor that is preferentially taken up by and is better able to inhibit growth and promote gene expression in malignant cells compared to normal healthy cells [98]. It would also be interesting to see if any genetic mutations exist in STARD8 in lymphoma patients. Other follow up experiments include knocking out STARD8 expression in mice or using RNAi to silence STARD8 expression in mouse lymphoid cells, and then monitoring for cancer related characteristics. Lastly, an expression vector containing STARD8 could be cloned into lymphoma cell lines to see what effect the induction of STARD8 expression has on the malignant cells.

Other potential candidate tumour suppressor genes might exist in the region surrounding the AR and additional investigation of genes and ESTs is required. For example moesin's (MSN) expression pattern within a cell, either membranous, mixed or cytoplasmic, has been shown to be a prognostic factor in oral squamous cell carcinoma [132]. MSN, a member of the ERM (ezrin/radixin/moesin) family, codes for cytoskeleton linker proteins. These proteins are involved in cell morphology, adhesion and motility. MSN is a suspected tumour suppressor gene, despite observations of ERM protein

upregulation in some cancers, due to its high homology to and colocalization with a known tumour suppressor gene, Merlin (reviewed in [132]). However, MSN does not appear to be expressed either lymphoblast or lymphoma cells (unpublished data from Dr. Carolyn Brown's lab).

Ephrin-B1 (EFNB1) is located in the first identified sub-region of interest downstream of STARD8. The encoded protein is a membrane-anchored ligand of the EphB receptor protein kinase. EFNB1 has a well established role in the development of the nervous system, but recent reports suggest that members of the ephrin ligand family are also involved in angiogenesis and oncogenesis where they have roles in cell adhesion, morphogenesis, capillary sprouting and chemoattraction (reviewed in [133]). EFNB1 was shown to be overexpressed in gastric cancer [133], and hepatocellular carcinoma [134] and EFNB1 expression in osteosarcoma was associated with poor prognosis [135]. In the case of EFNB1 overexpression in hepatocellular carcinoma, evidence implicates EFNB1 in neovascularization. The function of ephrin-B1 indicates an oncogenic function rather than a tumour suppressor function, however, this gene indicates that this region on the X chromosome is involved in tumorigenesis.

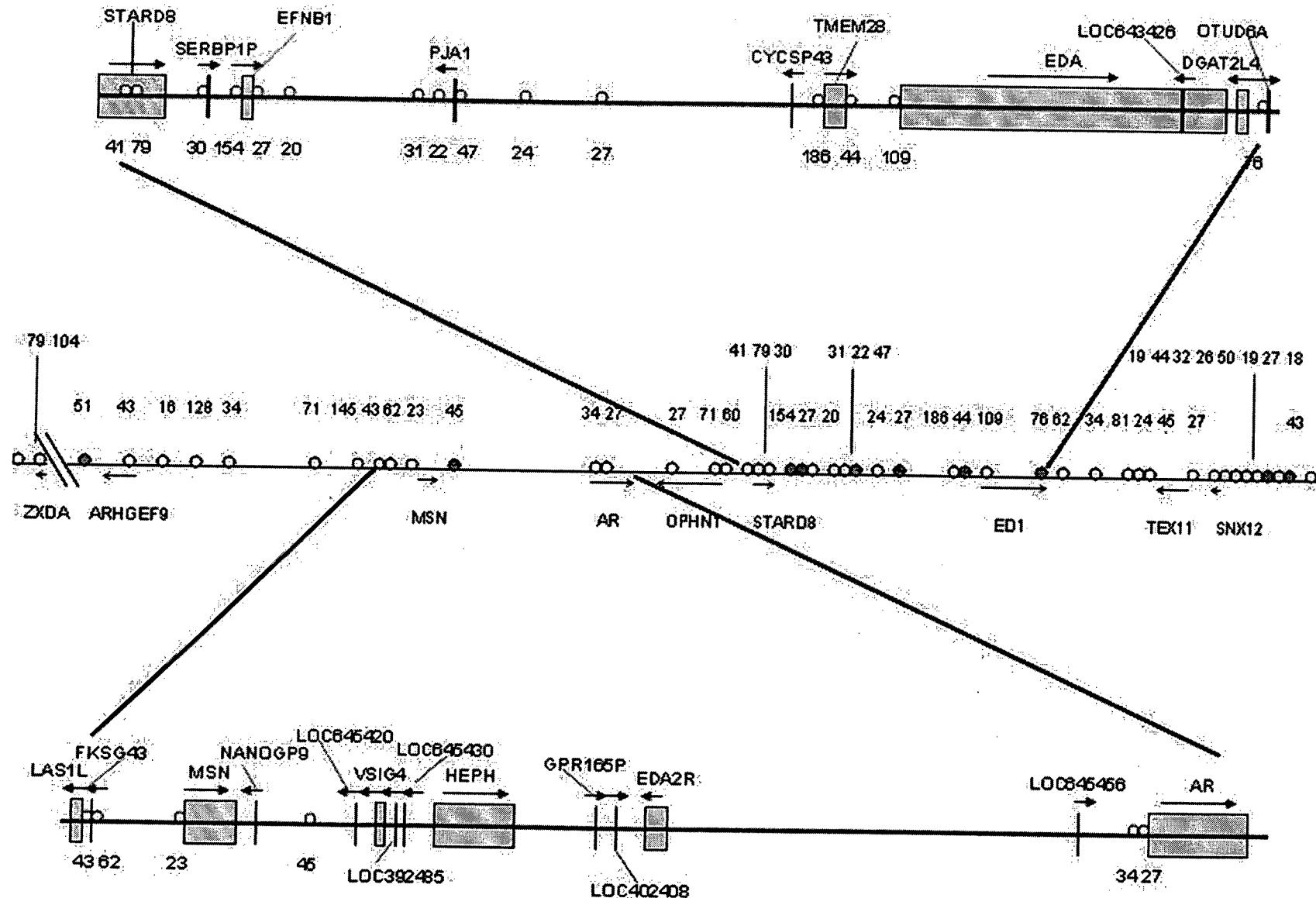
The focus of this thesis was primarily on abnormal methylation and finding the extent of the methylation spread. However, as previously mentioned, methylation analysis alone was not enough to narrow down on a region of interest. Methylation appeared to have spread over several Mb and to complicate the investigation further, methylation skipping, 'normal' methylation, a lack of methylation in the female control (and presumably on the inactive X chromosome) and 'partial' methylation were all observed.

Two possible explanations exist for the presence of partial methylation. The first has to do with limitations of the assay and the second with the composition of cells making up a sample. With respect to the assay, partial methylation might be caused by incomplete cutting of the HpaII recognition sites of some of the DNA strands in the sample, producing bands weaker in intensity relative to the uncut control. Another potential cause of partial methylation is that the digested DNA is obtained from a heterogeneous cell population with respect to methylation status. The initial

determination of methylation spread was done on cell line DNA and therefore the samples are expected to come from a homogeneous population of cells. It is however possible that cells within a cell line have obtained some differential methylation marks. In patient samples, partial methylation might be caused by contamination of the sample with normal cells or it may reflect a DNA methylation heterogeneous cell population in the tumour.

One theory explaining the ability of methylation to spread in malignant cells involves the breakdown of boundaries separating euchromatin from heterochromatin, allowing heterochromatin to spread (reviewed in [9]). As mentioned above in the introduction, seeds of methylation are hypothesized to act as origins of abnormal methylation in cancer. Taken together, these two theories can be applied to the data presented in this thesis to hypothesize the role of observed ‘normal’ CpG island methylation in the spread of abnormal methylation. Figure 4.1 shows the distribution of ‘normally’ methylated CpG islands relative to the location of the two abnormally hypermethylated sub-regions. Six out of the fourteen CpG islands examined in the first sub-region of interest and one of the six CpG islands examined in the second sub-region showed ‘normal’ methylation. It is possible that in lymphoma, the boundary elements compartmentalizing active and inactive domains are lost and methylation from these ‘normally’ methylated islands is able to spread to the surrounding regions. Maintained boundary elements in the lymphoma cells might explain abnormal methylation CpG island skipping.

A recent study looked at global methylation changes in colorectal cancer and identified a 4 Mb region on chromosome 2q14.2 that experienced abnormal hypermethylation [136]. Similar to what is reported in this thesis, within this large region of abnormal methylation, smaller sub-regions with continuous methylation were identified. In the case of colon cancer, three suburbs, as the authors called them, were found with the largest one spanning 1 Mb and flanking 12 CpG islands. These results argue against localized methylation of a target gene and the authors postulate a new mechanism in tumorigenesis involving whole chromosomal region (or ‘neighborhood’) chromatin changes called long-range epigenetic silencing (LRES). The authors also



**Figure 4.1** Distribution of 'normally' methylated CpG islands relative to the location of the two sub-regions of interest. The map in the middle of the figure shows the relative location of all the CpG islands in the 8Mb examined region examined. The top map shows the genes and CpG islands located in the first identified 1.4Mb sub-region and the bottom map shows the genes and CpG islands found in the second 2.3Mb sub-region. CpG islands filled in red represent islands that showed abnormal methylation in at least one cell line. The methylation seen in the male control for CpG 45 located within the the sub-region upstream of AR might be a false negative.

found evidence for the presence of abnormal H3K9 methylation which correlated with the repressed chromosomal regions. Some unmethylated genes that flanked hypermethylated regions were also silenced and Frigola *et al.* (2006) [136] proposed that these genes are silenced by chromatin remodeling including H3K9 methylation. LRES, both DNA methylation and histone modifications, can therefore silence neighborhood gene expression similar to genetic deletions. Interestingly, the results presented in this thesis show that there was a correlation between methylation status and expression, with genes associated with unmethylated CpG islands remaining active (*i.e.* CpG 60 and YIPF6) despite being located in the vicinity of a densely methylated region.

In the same study, methylation skipping within the large hypermethylated 4 Mb region was also observed [136]. These islands were shown to be less susceptible to *de novo* methylation in normal cells. Some CpG islands associated with genes susceptible to abnormal silencing experienced low levels of methylation in normal samples [136]. These results support the theory of the involvement of seeds of methylation in abnormal hypermethylation. The authors also observed an inverse correlation between the susceptibility of DNA methylation and normal gene expression levels. This provides strength for the argument of gene inactivation preceding DNA methylation and that active transcription is able to protect against methylation and promote demethylation (reviewed in [136]).

The similarities between the results shown in this thesis and the Frigola *et al.* (2006) [136] study prompts some future directions in the investigation of the abnormal hypermethylation seen on the X chromosome in lymphoma. It would be interesting to see if this region also experiences abnormal histone methylation and lacks histone acetylation in lymphoma and if this aids in the identification of candidate tumour suppressor gene(s). Another question to be addressed is whether or not the X chromosome is more prone to such LRES in cancer, because a similar process is involved in X inactivation. For example, according the Lyon repeat hypothesis, the enrichment of L1 elements on the X chromosome may have a role in the spread and amplification of the inactive signal by acting as booster stations.

## References:

1. Kops, G.J.P.L., B.A.A. Weaver, and D.W. Cleveland, *On the road to cancer: aneuploidy and the mitotic checkpoint*. Nature Reviews Cancer, 2005. **5**(10): p. 773.
2. Rajagopalan, H. and C. Lengauer, *Aneuploidy and cancer*. Nature, 2004. **432**(7015): p. 338.
3. Younes, A., D. Jendiroba, R. Katz, D. Hill, F. Cabanillas, and M. Andreeff, *Chromosome X numerical abnormalities in patients with Non-Hodgkin's Lymphoma: A study of 59 patients using fluorescence in situ hybridization*. Cancer Genetics and Cytogenetics, 1995. **82**(1): p. 23.
4. McDonald, H.L., R. D. Gascoyne, D. Horsman, and C. J. Brown, *Involvement of the X chromosome in non-Hodgkin lymphoma*. Genes, Chromosomes and Cancer, 2000. **28**(3): p. 246.
5. Chow, J.C., Z. Yen, S. Ziesche, and C. J. Brown, *Silencing of the mammalian X chromosome*. Annual Review Genomics and Human Genetics, 2005. **6**: p. 69.
6. Avner, P. and E. Heard, *X-chromosome inactivation: counting, choice and initiation*. Nature Reviews Genetics, 2001. **2**(1): p. 59.
7. Worm, J. and P. Guldberg, *DNA methylation: an epigenetic pathway to cancer and a promising target for anticancer therapy*. Journal of Oral Pathology and Medicine, 2002. **31**(8): p. 443.
8. Jones, P.A. and P.W. Laird, *Cancer-epigenetics comes of age*. Nat Genet, 1999. **21**(2): p. 163.
9. Jones, P.A. and S.B. Baylin, *The fundamental role of epigenetic events in cancer*. Nature Reviews Genetics, 2002. **3**(6): p. 415.
10. Baylin, S.B., *DNA methylation and gene silencing in cancer*. Nature Clinical Practice Oncology, 2005. **2**: p. S4.
11. Garinis, G., G. Patrinos, N. Spanakis, and P. Menounos, *DNA hypermethylation: when tumour suppressor genes go silent*. Human Genetics, 2002. **111**(2): p. 115.
12. Mhatre, A.N., M.A. Trifiro, M. Kaufman, P. Kazemi-Esfarjani, D. Figlewicz, G. Rouleau, and L. Pinsky, *Reduced transcriptional regulatory competence of the androgen receptor in X-linked spinal and bulbar muscular atrophy*. Nat Genet, 1993. **5**(2): p. 184.
13. Wooster, R., J. Mangion, R. Eeles, S. Smith, M. Dowsett, D. Averill, P. Barrett-Lee, D.F. Easton, B.A.J. Ponder, and M.R. Stratton, *A germline mutation in the androgen receptor gene in two brothers with breast cancer and Reifenstein syndrome*. Nat Genet, 1992. **2**(2): p. 132.
14. Jarrard, D.F., H. Kinoshita, Y. Shi, C. Sandefur, D. Hoff, L.F. Meisner, C. Chang, J.G. Herman, W.B. Isaacs, and N. Nassif, *Methylation of the androgen receptor promoter CpG island is associated with loss of androgen receptor expression in prostate cancer cells*. Cancer Res, 1998. **58**(23): p. 5310.
15. Giovannucci, E., M.J. Stampfer, K. Krishivas, M. Brown, A. Brufsky, J. Talcott, C.H. Hennekens, and P.W. Kantoff, *The CAG repeat within the androgen receptor gene and its relationship to prostate cancer*. PNAS, 1997. **94**(7): p. 3320.
16. Tentler, D., P. Gustavsson, J. Leisti, M. Schueler, J. Chelly, E. Timonen, G. Anneren, H.F. Willard and N. Dahl, *Deletion including the oligophrenin-1 gene*

- associated with enlarged cerebral ventricles, cerebellar hypoplasia, seizures and ataxia.* European Journal of Human Genetics, 1999. **7**: p. 541.
17. Robertson, K.D., *DNA methylation and human disease*. Nature Reviews Genetics, 2005. **6**(8): p. 597.
  18. Feinberg, A.P. and B. Tycko, *The history of cancer epigenetics*. Nature Reviews Cancer, 2004. **4**(2): p. 143.
  19. Li, E., *Chromatin modification and epigenetic reprogramming in mammalian development*. Nature Reviews Genetics, 2002. **3**(9): p. 662.
  20. Issa, J.-P., *CpG island methylator phenotype in cancer*. Nature Reviews Cancer, 2004. **4**(12): p. 988.
  21. Lund, A.H. and M. van Lohuizen, *Epigenetics and cancer*. Genes Dev., 2004. **18**(19): p. 2315.
  22. Farrell, W.E., *Epigenetic Mechanisms of Tumorigenesis*. Hormone and Metabolic Research, 2005(6): p. 361.
  23. Bajic, V.B., S.L. Tan, Y. Suzuki, and S. Sugano, *Promoter prediction analysis on the whole human genome*. Nat Biotech, 2004. **22**(11): p. 1467.
  24. Saxonov, S., P. Berg, and D.L. Brutlag, *A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters*. PNAS, 2006. **103**(5): p. 1412.
  25. Eads, C.A., K.D. Danenberg, K. Kawakami, L.B. Saltz, P.V. Danenberg, and P.W. Laird, *CpG Island Hypermethylation in Human Colorectal Tumors Is Not Associated with DNA Methyltransferase Overexpression*. Cancer Res, 1999. **59**(10): p. 2302.
  26. Bestor, T.H., *The DNA methyltransferases of mammals*. Hum. Mol. Genet., 2000. **9**(16): p. 2395.
  27. Okano, M., S. Xie, and E. Li, *Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases*. Nat Genet, 1998. **19**(3): p. 219.
  28. Lei, H., S.P. Oh, M. Okano, R. Juttermann, K.A. Goss, R. Jaenisch, and E. Li, *De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells*. Development, 1996. **122**(10): p. 3195.
  29. Li, E., T.H. Bestor, and R. Jaenisch, *Targeted mutation of the DNA methyltransferase gene results in embryonic lethality*. Cell, 1992. **69**(6): p. 915.
  30. Leonhardt, H., A.W. Page, H.-U. Weier, and T.H. Bestor, *A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei*. Cell, 1992. **71**(5): p. 865.
  31. Okano, M., D.W. Bell, D.A. Haber, and E. Li, *DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development*. Cell, 1999. **99**(3): p. 247.
  32. Hansen, R.S., C. Wijmenga, P. Luo, A.M. Stanek, T.K. Canfield, C.M.R. Weemaes, and S.M. Gartler, *The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome*. PNAS, 1999. **96**(25): p. 14412.
  33. Morgan, H.D., F. Santos, K. Green, W. Dean, and W. Reik, *Epigenetic reprogramming in mammals*. Hum. Mol. Genet., 2005. **14**(suppl\_1): p. R47.
  34. Santos, F. and W. Dean, *Epigenetic reprogramming during early development in mammals*. Reproduction, 2004. **127**(6): p. 643.

35. Mayer, W., A. Niveleau, J. Walter, R. Fundele, and T. Haaf, *Embryogenesis: Demethylation of the zygotic paternal genome*. Nature, 2000. **403**(6769): p. 501.
36. Bird, A.P. and A.P. Wolffe, *Methylation-Induced Repression-- Belts, Braces, and Chromatin*. Cell, 1999. **99**(5): p. 451.
37. Fuks, F., *DNA methylation and histone modifications: teaming up to silence genes*. Current Opinion in Genetics & Development, 2005. **15**(5): p. 490.
38. Kazazian, H.H., Jr., *Mobile Elements: Drivers of Genome Evolution*. Science, 2004. **303**(5664): p. 1626.
39. Yoder, J.A., C.P. Walsh, and T.H. Bestor, *Cytosine methylation and the ecology of intragenomic parasites*. Trends in Genetics, 1997. **13**(8): p. 335.
40. Dale J. Hedges, M.A.B., *From the margins of the genome: mobile elements shape primate evolution*. BioEssays, 2005. **27**(8): p. 785.
41. Carrel, L. and H.F. Willard, *X-inactivation profile reveals extensive variability in X-linked gene expression in females*. Nature, 2005. **434**(7031): p. 400.
42. Heard, E., *Recent advances in X-chromosome inactivation*. Current Opinion in Cell Biology, 2004. **16**(3): p. 247.
43. Jaenisch, R. and A. Bird, *Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals*. Nat Genet., 2003. **33**(3): p.245
44. Ke, X. and A. Collins, *CpG Islands in Human X-Inactivation*. Annals of Human Genetics, 2003. **67**(3): p. 242.
45. Esteller, M., *Relevance of DNA methylation in the management of cancer*. The Lancet Oncology, 2003. **4**(6): p. 351.
46. Plass, C., *Cancer epigenomics*. Hum. Mol. Genet., 2002. **11**(20): p. 2479.
47. Kisseljova, N.P. and F.L. Kisseljov, *DNA Demethylation and Carcinogenesis*. Biochemistry (Moscow), 2005. **70**(7): p. 743.
48. Greger, V., E. Passarge, W. Höpping, E. Messmer, and B. Horsthemke, *Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma*. Human Genetics, 1989. **83**(2): p. 155.
49. Greger, V., N. Debus, D. Lohmann, W. Höpping, E. Passarge, and B. Horsthemke, *Frequency and parental origin of hypermethylated RB1 alleles in retinoblastoma*. Human Genetics, 1994. **94**(5): p. 491.
50. Herman, J.G., F. Latif, Y. Weng, M.I. Lerman, B. Zbar, S. Liu, D. Samid, D.R. Duan, J.R. Gnarra, W.M. Linehan, and S.B. Baylin, *Silencing of the VHL Tumor-Suppressor Gene by DNA Methylation in Renal Carcinoma*. PNAS, 1994. **91**(21): p. 9700.
51. Veigl, M.L., L. Kasturi, J. Olechnowicz, A. Ma, J.D. Lutterbaugh, S. Periyasamy, G.-M. Li, J. Drummond, P.L. Modrich, W.D. Sedwick, and S.D. Markowitz, *Biallelic inactivation of hMLH1 by epigenetic gene silencing, a novel mechanism causing human MSI cancers*. PNAS, 1998. **95**(15): p. 8698.
52. Yoshikawa, H., K. Matsubara, G.-S. Qian, P. Jackson, J.D. Groopman, J.E. Manning, C.C. Harris, and J.G. Herman, *SOCS-1, a negative regulator of the JAK/STAT pathway, is silenced by methylation in human hepatocellular carcinoma and shows growth-suppression activity*. Nat Genet, 2001. **28**(1): p. 29.
53. Clark, S.J.a.M., J., *DNA methylation and gene silencing in cancer: which is the guilty party?* Oncogene, 2002. **21**(35): p. 5380.

54. Turker, M.S., *Gene silencing in mammalian cells and the spread of DNA methylation*. Oncogene, 2002. **21**: p. 5388.
55. Turker, M.S. and T.H. Bestor, *Formation of methylation patterns in the mammalian genome*. Mutation Research/Reviews in Mutation Research, 1997. **386**(2): p. 119.
56. Issa, J.P.J., Y.L. Ottaviano, P. Celano, S.R. Hamilton, N.E. Davidson, and S.B. Baylin, *Methylation of the oestrogen receptor CpG island links ageing and neoplasia in human colon*. Nat Genet, 1994. **7**(4): p. 536.
57. Issa, J.P.J., P.M. Vertino, C.D. Boehm, I.F. Newsham, and S.B. Baylin, *Switch from monoallelic to biallelic human IGF2 promoter methylation during aging and carcinogenesis*. PNAS, 1996. **93**(21): p. 11757.
58. Habuchi, T., T. Takahashi, H. Kakinuma, L. Wang, N. Tsuchiya, S. Satoh, T. Akao, K. Sato, O. Ogawa, M.A. Knowles and T. Kato, *Hypermethylation at 9q32-33 tumour suppressor region is age-related in normal urothelium and an early and frequent alteration in bladder cancer*. Oncogene, 2001. **20**(4): p. 531.
59. Shen, L., N. Ahuja, Y. Shen, N.A. Habib, M. Toyota, A. Rashid, and J.P.J. Issa, *DNA Methylation and Environmental Exposures in Human Hepatocellular Carcinoma*. J Natl Cancer Inst, 2002. **94**(10): p. 755.
60. Issa, J.P.J., N. Ahuja, M. Toyota, M.P. Bronner, and T.A. Brentnall, *Accelerated Age-related CpG Island Methylation in Ulcerative Colitis*. Cancer Res, 2001. **61**(9): p. 3573.
61. Kang, G.H., S. Lee, W.H. Kim, H.W. Lee, J.C. Kim, M.G. Rhyu, and J.Y. Ro, *Epstein-Barr Virus-Positive Gastric Carcinoma Demonstrates Frequent Aberrant Methylation of Multiple Genes and Constitutes CpG Island Methylator Phenotype-Positive Gastric Carcinoma*. Am J Pathol, 2002. **160**(3): p. 787.
62. Ambinder, R.F., K.D. Robertson, and Q. Tao, *DNA methylation and the Epstein-Barr virus*. Seminars in Cancer Biology, 1999. **9**(5): p. 369.
63. Issa, J.P., P.M. Vertino, J. Wu, S. Sazawal, P. Celano, B.D. Nelkin, S.R. Hamilton and S.B. Baylin, *Increased Cytosine DNA-Methyltransferase Activity During Colon Cancer Progression*. Journal of the National Cancer Institute, 1993. **85**: p. 1235.
64. De Marzo, A.M., V.L. Marchi, E.S. Yang, R. Veeraswamy, X. Lin, and W.G. Nelson, *Abnormal Regulation of DNA Methyltransferase Expression during Colorectal Carcinogenesis*. Cancer Res, 1999. **59**(16): p. 3855.
65. Shibata, D.M., F. Sato, Y. Mori, K. Perry, J. Yin, S. Wang, Y. Xu, A. Olaru, F. Selaru, K. Spring, J. Young, J.M. Abraham, and S.J. Meltzer, *Hypermethylation of HPP1 Is Associated with hMLH1 Hypermethylation in Gastric Adenocarcinomas*. Cancer Res, 2002. **62**(20): p. 5637.
66. Shen, L., Y. Kondo, S.R. Hamilton, A. Rashid, and J.-P.J. Issa, *p14 methylation in human colon cancer is associated with microsatellite instability and wild-type p53*. Gastroenterology, 2003. **124**(3): p. 626.
67. Ahuja, N., A.L. Mohan, Q. Li, J.M. Stolker, J.G. Herman, S.R. Hamilton, S.B. Baylin, and J.P. Issa, *Association between CpG island methylation and microsatellite instability in colorectal cancer*. Cancer Res, 1997. **57**(16): p. 3370.
68. Herman, J.G., A. Umar, K. Polyak, J.R. Graff, N. Ahuja, J.-P.J. Issa, S. Markowitz, J.K.V. Willson, S.R. Hamilton, K.W. Kinzler, M.F. Kane, R.D.

- Kolodner, B. Vogelstein, T.A. Kunkel, and S.B. Baylin, *Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma*. PNAS, 1998. **95**(12): p. 6870.
69. Yamamoto, H., Y. Min, F. Hoh, A. Imsumran, S. Horiuchi, M. Yoshida, S. Iku, H. Fukushima, K. Imai., *Differential involvement of the hypermethylator phenotype in hereditary and sporadic colorectal cancers with high-frequency microsatellite instability*. Genes, Chromosomes and Cancer, 2002. **33**(3): p. 322.
70. Ueki, T., M. Toyota, T. Sohn, C.J. Yeo, J.-P.J. Issa, R.H. Hruban, and M. Goggins, *Hypermethylation of Multiple Genes in Pancreatic Adenocarcinoma*. Cancer Res, 2000. **60**(7): p. 1835.
71. Toyota, M., N. Ahuja, H. Suzuki, F. Itoh, M. Ohe-Toyota, K. Imai, S.B. Baylin, and J.-P.J. Issa, *Aberrant Methylation in Gastric Cancer Associated with the CpG Island Methylator Phenotype*. Cancer Res, 1999. **59**(21): p. 5438.
72. Garcia-Manero, G., J. Daniel, T.L. Smith, S.M. Kornblau, M.-S. Lee, H.M. Kantarjian, and J.-P.J. Issa, *DNA Methylation of Multiple Promoter-associated CpG Islands in Adult Acute Lymphocytic Leukemia*. Clin Cancer Res, 2002. **8**(7): p. 2217.
73. Anacleto, C., Andreia M. Leopoldino, Benedito Rossi, Fernando A. Soares, Ademar Lopes, Jose Claudio C. Rocha, Otavia Caballero, Anamaria A. Camargo, Andrew J. G. Simpson, and Sergio D. J. Pena, *Colorectal Cancer "Methylator Phenotype": Fact or Artifact?* Neoplasia, 2005. **7**(4): p. 331.
74. Issa, J.-P.J., L. Shen, and M. Toyota, *CIMP, at Last*. Gastroenterology, 2005. **129**(3): p. 1121.
75. Yates, P.A., R.W. Burman, P. Mummaneni, S. Krussel, and M.S. Turker, *Tandem B1 Elements Located in a Mouse Methylation Center Provide a Target for de Novo DNA Methylation*. J. Biol. Chem., 1999. **274**(51): p. 36357.
76. Macleod, D., J. Charlton, J. Mullins, and A.P. Bird, *Sp1 sites in the mouse aprt gene promoter are required to prevent methylation of the CpG island*. Genes Dev., 1994. **8**(19): p. 2282.
77. Siegfried, Z., S. Eden, M. Mendelsohn, X. Feng, B.-Z. Tsuberi, and H. Cedar, *DNA methylation represses transcription in vivo*. Nat Genet, 1999. **22**(2): p. 203.
78. Hejnar, J., P. Hajkova, J. Plachy, D. Elleder, V. Stepanets, and J. Svoboda, *CpG island protects Rous sarcoma virus-derived vectors integrated into nonpermissive cells from DNA methylation and transcriptional suppression*. PNAS, 2001. **98**(2): p. 565.
79. Song, J.Z., C. Stirzaker, J. Harrison, J.R. Melki, and S.J. Clark, *Hypermethylation trigger of the glutathione-S-transferase gene (GSTP1) in prostate cancer cells*. Oncogene, 2002. **21**: p. 1048.
80. Li, E., C. Beard, and R. Jaenisch, *Role for DNA methylation in genomic imprinting*. Nature, 1993. **366**(6453): p. 362.
81. Laird, P.W., L. Jackson-Grusby, A. Fazeli, S.L. Dickinson, W. Edward Jung, E. Li, R.A. Weinberg, and R. Jaenisch, *Suppression of intestinal neoplasia by DNA hypomethylation*. Cell, 1995. **81**(2): p. 197.
82. Grant, M., M. Zuccotti, and M. Monk, *Methylation of CpG sites of two X-linked genes coincides with X-inactivation in the female mouse embryo but not in the germ line*. Nat Genet, 1992. **2**(2): p. 161.

83. Loukinov, D.I., E. Pugacheva, S. Vatolin, S.D. Pack, H. Moon, I. Chernukhin, P. Mannan, E. Larsson, C. Kanduri, A.A. Vostrov, H. Cui, E.L. Niemitz, J.E.J. Rasko, F.M. Docquier, M. Kistler, J.J. Breen, Z. Zhuang, W.W. Quitschke, R. Renkawitz, E.M. Klenova, A.P. Feinberg, R. Ohlsson, H.C. Morse, III, and V.V. Lobanenkov, *BORIS, a novel male germ-line-specific protein associated with epigenetic reprogramming events, shares the same 11-zinc-finger domain with CTCF, the insulator protein involved in reading imprinting marks in the soma.* PNAS, 2002. **99**(10): p. 6806.
84. Motamedi, M.R., A. Verdel, S.U. Colmenares, S.A. Gerber, S.P. Gygi, and D. Moazed, *Two RNAi Complexes, RITS and RDRC, Physically Interact and Localize to Noncoding Centromeric RNAs.* Cell, 2004. **119**(6): p. 789.
85. Verdel, A., S. Jia, S. Gerber, T. Sugiyama, S. Gygi, S.I.S. Grewal, and D. Moazed, *RNAi-Mediated Targeting of Heterochromatin by the RITS Complex.* Science, 2004. **303**(5658): p. 672.
86. Castanotto, D., S. Tommasi, M. Li, S. Yanow, G. P. Pfeifer and J. J. Rossi., *Sjort hairpin RNA-directed cytosine (CpG) methylation of the RASSF1A gene promoter in HeLa cells.* Molecular Therapy, 2005. **12**(1): p. 179.
87. Esteller, M., P.G. Corn, S.B. Baylin, and J.G. Herman, *A Gene Hypermethylation Profile of Human Cancer.* Cancer Res, 2001. **61**(8): p. 3225.
88. Laird, P.W., *The power and the promise of DNA methylation markers.* Nature Reviews Cancer, 2003. **3**(4): p. 253.
89. Palmisano, W.A., K.K. Divine, G. Saccomanno, F.D. Gilliland, S.B. Baylin, J.G. Herman, and S.A. Belinsky, *Predicting Lung Cancer by Detecting Aberrant Promoter Methylation in Sputum.* Cancer Res, 2000. **60**(21): p. 5954.
90. Esteller, M., S. Gonzalez, R.A. Risques, E. Marcuello, R. Mangues, J.R. Germa, J.G. Herman, G. Capella, and M.A. Peinado, *K-ras and p16 Aberrations Confer Poor Prognosis in Human Colorectal Cancer.* Journal of Clinical Oncology, 2001. **19**(2): p. 299.
91. Brabender, J., H. Usadel, K.D. Danenberg, R. Metzger, P.M. Schneider, R.V. Lord, K. Wickramasinghe, C.E. Lum, J.M. Park, D. Salonga, J. Singer, D. Sidransky, A.H. Hölscher, S.J. Meltzer and P.V. Danenberg, *Adenomatous polyposis coli gene promoter hypermethylation in non-small cell lung cancer is associated with survival.* Oncogene, 2001. **20**(27): p. 3528.
92. Esteller, M., J. Garcia-Foncillas, E. Andion, S.N. Goodman, O.F. Hidalgo, V. Vanaclocha, S.B. Baylin, and J.G. Herman, *Inactivation of the DNA-Repair Gene MGMT and the Clinical Response of Gliomas to Alkylating Agents.* N Engl J Med, 2000. **343**(19): p. 1350.
93. Esteller, M., G. Gaidano, S.N. Goodman, V. Zagonel, D. Capello, B. Botto, D. Rossi, A. Gloghini, U. Vitolo, A. Carbone, S.B. Baylin, and J.G. Herman, *Hypermethylation of the DNA Repair Gene O6-Methylguanine DNA Methyltransferase and Survival of Patients With Diffuse Large B-Cell Lymphoma.* J Natl Cancer Inst, 2002. **94**(1): p. 26.
94. Egger, G., G. Liang, A. Aparicio, and P.A. Jones, *Epigenetics in human disease and prospects for epigenetic therapy.* Nature, 2004. **429**(6990): p. 457.
95. Brown, R. and G. Strathdee, *Epigenomics and epigenetic therapy of cancer.* Trends in Molecular Medicine, 2002. **8**(4): p. S43.

96. Issa, J.P., H.M. Kantarjian, and P. Kirkpatrick, *Azacitidine*. Nature Reviews Drug Discovery, 2005. **4**(4): p. 275.
97. Rodenhiser, D. and M. Mann, *Epigenetics and human disease: translating basic biology into clinical applications*. Canadian Medical Association Journal, 2006. **174**(3): p. 341.
98. Yoo, C.B., J.C. Cheng and P.A. Jones, *Zebularine: a new drug for epigenetic therapy*. Biochemical Society Transactions, 2004. **32**(6): p. 910.
99. Esteller, M., *DNA methylation and cancer therapy: new development and expectations*. Current Opinion in Oncology, 2004. **17**: p. 55.
100. Ansell, S.M., and Armitage, J., *Non-Hodgkin Lymphoma: Diagnosis and Treatment*. Mayo Clinic Proceedings, 2005. **80**(8): p. 1087.
101. Grulich, A. and C. Vajdic, *The epidemiology of non-Hodgkin lymphoma*. Pathology, 2005. **37**(6): p. 409.
102. Lu, P., *Staging and Classification of Lymphoma*. Seminars in Nuclear Medicine, 2005. **35**(3): p. 160.
103. Kuppers, R., *Mechanisms of B-cell lymphoma pathogenesis*. Nature Reviews Cancer, 2005. **5**(4): p. 251.
104. Armitage, J.O., *Staging Non-Hodgkin Lymphoma*. CA Cancer J Clin, 2005. **55**(6): p. 368.
105. Rizvi, M.A., A.M. Evens, M.S. Tallman, B.P. Nelson, and S.T. Rosen, *T-cell non-Hodgkin lymphoma*. Blood, 2006. **107**(4): p. 1255.
106. Harigae, H., R. Ichinohasama, I. Miura, J. Kameoka, K. Meguro, K. Miyamura, O. Sasaki, I. Ishikawa, S. Takahashi, M. Kaku, and T. Sasaki, *Primary marginal zone lymphoma of the thymus accompanied by chromosomal anomaly 46,X,dup(X)(p11p22)*. Cancer Genetics and Cytogenetics, 2002. **133**(2): p. 142.
107. Cook, J.R., M.E. Sherer, S. Shekhter-Levin, and S.H. Swerdlow, *Nodal marginal zone B-cell lymphoma with a novel t(X;5)(q28;q22): conventional and molecular cytogenetic analysis*. Cancer Genetics and Cytogenetics, 2003. **143**(2): p. 154.
108. Vineis, P., G. Masala and A. Seniori Costantini, *Does a gene in the Xq28 region increase the risk of non-Hodgkin's lymphomas?* Annals of Oncology, 1999. **10**(4): p. 471.
109. Renedo, M., B. Martinez-Delgado, E. Arranz, M.J. Garcia, M. Urioste, A. Martinez-Ramirez, C. Rivas, J.C. Cigudosa and J. Benitez, *Chromosomal changes pattern and gene amplification in T cell non-Hodgkin's lymphomas*. Leukemia, 2001. **15**: p. 1627.
110. Callet-Bauchu, E., R. Rimokh, S. Gazzo, J. Pages, Y. Bastion, F. Berger, P. Coeur, and P. Felman, *Unbalanced X;autosome translocation (X;18)(q13;p11) in a case of aggressive natural killer non-Hodgkin lymphoma*. Cancer Genetics and Cytogenetics, 1997. **98**(1): p. 16.
111. Jones, P.A. and D. Takai, *The Role of DNA Methylation in Mammalian Epigenetics*. Science, 2001. **293**(5532): p. 1068.
112. Kang, M.I., M.G. Rhyu, Y.H. Kim, Y.C. Jung, S.-J. Hong, C.S. Cho, and H.S. Kim, *The length of CpG islands is associated with the distribution of Alu and L1 retroelements*. Genomics, 2006. **87**(5): p. 580.
113. Ross, M.T., D.V. Grafham, A.J. Coffey, S. Scherer, K. McLay, D. Muzny, M. Platzer, G.R. Howell, C. Burrows, C.P. Bird, A. Frankish, F.L. Lovell, K.L.

- Howe, J.L. Ashurst, R.S. Fulton, R. Sudbrak, G. Wen, M.C. Jones, M.E. Hurles, T.D. Andrews, C.E. Scott, S. Searle, J. Ramser, A. Whittaker, R. Deadman, N.P. Carter, S.E. Hunt, R. Chen, A. Cree, P. Gunaratne, P. Havlak, A. Hodgson, M.L. Metzker, S. Richards, G. Scott, D. Steffen, E. Sodergren, D.A. Wheeler, K.C. Worley, R. Ainscough, K.D. Ambrose, M.A. Ansari-Lari, S. Aradhya, R.I.S. Ashwell, A.K. Babbage, C.L. Bagguley, A. Ballabio, R. Banerjee, G.E. Barker, K.F. Barlow, I.P. Barrett, K.N. Bates, D.M. Beare, H. Beasley, O. Beasley, A. Beck, G. Bethel, K. Blechschmidt, N. Brady, S. Bray-Allen, A.M. Bridgeman, A.J. Brown, M.J. Brown, D. Bonnin, E.A. Bruford, C. Buhay, P. Burch, D. Burford, J. Burgess, W. Burrill, J. Burton, J.M. Bye, C. Carder, L. Carrel, J. Chako, J.C. Chapman, D. Chavez, E. Chen, G. Chen, Y. Chen, Z. Chen, C. Chinault, A. Ciccodicola, S.Y. Clark, G. Clarke, C.M. Clee, S. Clegg, K. Clerc-Blankenburg, K. Clifford, V. Cobley, C.G. Cole, J.S. Conquer, N. Corby, R.E. Connor, R. David, J. Davies, C. Davis, J. Davis, O. Delgado, D. DeShazo, P. Dhami, Y. Ding, H. Dinh, S. Dodsworth, H. Draper, S. Dugan-Rocha, A. Dunham, M. Dunn, K.J. Durbin, I. Dutta, T. Eades, M. Ellwood, A. Emery-Cohen, H. Errington, K.L. Evans, L. Faulkner, F. Francis, J. Frankland, A.E. Fraser, P. Galgoczy, J. Gilbert, R. Gill, G. Glockner, S.G. Gregory, S. Gribble, C. Griffiths, R. Grocock, Y. Gu, R. Gwilliam, C. Hamilton, E.A. Hart, A. Hawes, P.D. Heath, K. Heitmann, S. Hennig, J. Hernandez, B. Hinzmann, S. Ho, M. Hoffs, P.J. Howden, E.J. Huckle, J. Hume, P.J. Hunt, A.R. Hunt, J. Isherwood, L. Jacob, D. Johnson, S. Jones, P.J. de Jong, S.S. Joseph, S. Keenan, S. Kelly, J.K. Kershaw, Z. Khan, P. Kioschis, S. Klages, A.J. Knights, A. Kosiura, C. Kovar-Smith, G.K. Laird, C. Langford, S. Lawlor, M. Leversha, L. Lewis, W. Liu, C. Lloyd, D.M. Lloyd, H. Loulseged, J.E. Loveland, J.D. Lovell, R. Lozano, J. Lu, R. Lyne, J. Ma, M. Maheshwari, L.H. Matthews, J. McDowall, S. McLaren, A. McMurray, P. Meidl, T. Meitinger, S. Milne, G. Miner, S.L. Mistry, M. Morgan, S. Morris, I. Muller, J.C. Mullikin, N. Nguyen, G. Nordsiek, G. Nyakatura, C.N. O'Dell, G. Okwuonu, S. Palmer, R. Pandian, D. Parker, J. Parrish, S. Pasternak, D. Patel, A.V. Pearce, D.M. Pearson, S.E. Pelan, L. Perez, K.M. Porter, Y. Ramsey, K. Reichwald, S. Rhodes, K.A. Ridler, D. Schlessinger, M.G. Schueler, H.K. Sehra, C. Shaw-Smith, H. Shen, E.M. Sheridan, R. Shownkeen, C.D. Skuce, M.L. Smith, E.C. Sotheran, H.E. Steingruber, C.A. Steward, R. Storey, R.M. Swann, D. Swarbreck, P.E. Tabor, S. Taudien, T. Taylor, B. Teague, K. Thomas, A. Thorpe, K. Timms, A. Tracey, S. Trevanion, A.C. Tromans, M. d'Urso, D. Verduzco, D. Villasana, L. Waldron, M. Wall, Q. Wang, J. Warren, G.L. Warry, X. Wei, A. West, S.L. Whitehead, M.N. Whiteley, J.E. Wilkinson, D.L. Willey, G. Williams, L. Williams, A. Williamson, H. Williamson, L. Wilming, R.L. Woodmansey, P.W. Wray, J. Yen, J. Zhang, J. Zhou, H. Zoghbi, S. Zorilla, D. Buck, R. Reinhardt, A. Poustka, A. Rosenthal, H. Lehrach, A. Meindl, P.J. Minx, L.W. Hillier, H.F. Willard, R.K. Wilson, R.H. Waterston, C.M. Rice, M. Vaudin, A. Coulson, D.L. Nelson, G. Weinstock, J.E. Sulston, R. Durbin, T. Hubbard, R.A. Gibbs, S. Beck, J. Rogers and D.R. Bentley, *The DNA sequence of the human X chromosome*. Nature, 2005. 434(7031): p. 325.
114. Bock, C., M. Paulsen, S. Tierling, T. Mikeska, T. Lengauer, J. Walter, ouml, and rn, *CpG Island Methylation in Human Lymphocytes Is Highly Correlated with*

- DNA Sequence, Repeats, and Predicted DNA Structure.* PLoS Genetics, 2006. 2(3): p. e26.
115. de Leeuw, R.J., J.J. Davies, A. Rosenwald, G. Bebb, R.D. Gascoyne, M.J.S. Dyer, L.M. Staudt, J.A. Martinez-Climent, and W.L. Lam, *Comprehensive whole genome array CGH profiling of mantle cell lymphoma model genomes.* Hum. Mol. Genet., 2004. 13(17): p. 1827.
116. Sinha, S.K. and P.M. Chaudhary, *Induction of Apoptosis by X-linked Ectodermal Dysplasia Receptor via a Caspase 8-dependent Mechanism.* J. Biol. Chem., 2004. 279(40): p. 41873.
117. Yu, P., Y. Chen, D.A. Tagle, and T. Cai, *PJA1, Encoding a RING-H2 Finger Ubiquitin Ligase, Is a Novel Human X Chromosome Gene Abundantly Expressed in Brain.* Genomics, 2002. 79(6): p. 869.
118. Saha, T., D. Vardhini, Y. Tang, V. Katuri, W. Jogunoori, E.A. Volpe, D. Haines, A. Sidawy, X. Zhou, I. Gallicano, R. Schlegel, B. Mishra, and L. Mishra, *RING finger-dependent ubiquitination by PRAJA is dependent on TGF-[beta] and potentially defines the functional status of the tumor suppressor ELF.* Oncogene, 2005. 25(5): p. 693.
119. Too, C.K.L., *Differential expression of elongation factor-2, [alpha]4 phosphoprotein and Cdc5-like protein in prolactin-dependent/independent rat lymphoid cells.* Molecular and Cellular Endocrinology, 1997. 131(2): p. 221.
120. Kong, M., C.J. Fox, J. Mu, L. Solt, A. Xu, R.M. Cinalli, M.J. Birnbaum, T. Lindsten, and C.B. Thompson, *The PP2A-Associated Protein {alpha}4 Is an Essential Inhibitor of Apoptosis.* Science, 2004. 306(5696): p. 695.
121. Turkish, A.R., A.L. Henneberry, D. Cromley, M. Padamsee, P. Oelkers, H. Bazzi, A.M. Christiano, J.T. Billheimer, and S.L. Sturley, *Identification of Two Novel Human Acyl-CoA Wax Alcohol Acyltransferases: members of the diacylglycerol acyl transferase 2 (DGAT2) gene superfamily.* J. Biol. Chem., 2005. 280(15): p. 14755.
122. Brookes, M.J., S. Hughes, F.E. Turner, G. Reynolds, N. Sharma, T. Ismail, G. Berx, A.T. McKie, N. Hotchin, G.J. Anderson, T. Iqbal, and C. Tselepis, *A modulation of iron transport proteins in human colorectal carcinogenesis.* Gut, 2006: p. gut.2006.094060.
123. Goodison, S., J. Yuan, D. Sloan, R. Kim, C. Li, N.C. Popescu, and V. Urquidi, *The RhoGAP Protein DLC-1 Functions as a Metastasis Suppressor in Breast Cancer Cells.* Cancer Res, 2005. 65(14): p. 6042.
124. Song, Y.F., R. Xu, X.H. Zhang, B.B. Chen, Q. Chen, Y.M. Chen, and Y. Xie, *High frequent promoter hypermethylation of DLC-1 gene in multiple myeloma.* J Clin Pathol, 2006: p. jcp.2005.031377.
125. Wong, C.M., J.M.F. Lee, Y.P. Ching, D.Y. Jin, and I.O.. Ng, *Genetic and Epigenetic Alterations of DLC-1 Gene in Hepatocellular Carcinoma.* Cancer Res, 2003. 63(22): p. 7646.
126. Alizadeh, A.A., M.B. Eisen, R.E. Davis, C. Ma, I.S. Lossos, A. Rosenwald, J.C. Boldrick, H. Sabet, T. Tran, X. Yu, J.I. Powell, L. Yang, G.E. Marti, T. Moore, J. Hudson, L. Lu, D.B. Lewis, R. Tibshirani, G. Sherlock, W.C. Chan, T.C. Greiner, D.D. Weisenburger, J.O. Armitage, R. Warnke, R. Levy, W. Wilson, M.R. Grever, J.C. Byrd, D. Botstein, P.O. Brown, and L.M. Staudt, *Distinct types of*

- diffuse large B-cell lymphoma identified by gene expression profiling.* Nature, 2000. **403**(6769): p. 503.
127. Thieblemont, C., V. Nasser, P. Felman, K. Leroy, S. Gazzo, E. Callet-Bauchu, B. Loriod, S. Granjeaud, P. Gaulard, C. Haioun, A. Traverse-Glehen, L. Baseggio, F. Bertucci, D. Birnbaum, F. Magrangeas, S. Minvielle, H. Avet-Loiseau, G. Salles, B. Coiffier, F. Berger, and R. Houlgate, *Small lymphocytic lymphoma, marginal zone B-cell lymphoma, and mantle cell lymphoma exhibit distinct gene-expression profiles allowing molecular diagnosis.* Blood, 2004. **103**(7): p. 2727.
128. Poulsen, C.B., R. Borup, F.C. Nielsen, N. Borregaard, M. Hansen, K. Gronbaek, M.B. Moller, and E. Ralfkiaer, *Microarray-based classification of diffuse large B-cell lymphoma.* European Journal of Haematology, 2005. **74**(6): p. 453.
129. Tucker, C.A., G. Bebb, R.J. Klasa, M. Chhanabhai, V. Lestou, D.E. Horsman, R.D. Gascoyne, A. Wiestner, D. Masin, M. Bally, and M.E. Williams, *Four human t(11;14)(q13;q32)-containing cell lines having classic and variant features of Mantle Cell Lymphoma.* Leukemia Research, 2006. **30**(4): p. 449.
130. Lorincz, M.C., D.R. Dickerson, M. Schmitt, and M. Groudine, *Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells.* Nat Struct Mol Biol, 2004. **11**(11): p. 1068.
131. Bestor, T.H., *Unanswered Questions about the Role of Promoter Methylation in Carcinogenesis.* Ann NY Acad Sci, 2003. **983**(1): p. 22.
132. Kobayashi, H., J. Sagara, H. Kurita, M. Morifushi, M. Ohishi, K. Kurashina, and S.i. Taniguchi, *Clinical Significance of Cellular Distribution of Moesin in Patients with Oral Squamous Cell Carcinoma.* Clin Cancer Res, 2004. **10**(2): p. 572.
133. Kataoka, H., M. Tanaka, M. Kanamori, S. Yoshii, M. Ihara, Y.J. Wang, J.P. Song, Z.Y. Li, H. Arai, Y. Otsuki, T. Kobayashi, H. Konno, H. Hanai, and H. Sugimura, *Expression profile of EFNB1, EFNB2, two ligands of EPHB2 in human gastric cancer.* Journal of Cancer Research and Clinical Oncology, 2002. **128**(7): p. 343.
134. Sawai, Y., S. Tamura, K. Fukui, N. Ito, K. Imanaka, A. Saeki, S. Sakuda, S. Kiso, and Y. Matsuzawa, *Expression of ephrin-B1 in hepatocellular carcinoma: possible involvement in neovascularization.* Journal of Hepatology, 2003. **39**(6): p. 991.
135. Varelias, A., S.A. Koblar, P.A. Cowled, C.D. Carter and M. Clayer, *Human osteosarcoma expresses specific ephrin profiles.* Cancer, 2002. **95**: p. 862.
136. Frigola, J., J. Song, C. Stirzaker, R.A. Hinshelwood, M.A. Peinado, and S.J. Clark, *Epigenetic remodeling in colorectal cancer results in coordinate gene suppression across an entire chromosome band.* Nat Genet, 2006. **38**(5): p. 540.