

Patterns of DNA methylation on the human X chromosome and use in analyzing
X-chromosome inactivation

by

Allison Marie Cotton

B.Sc., The University of Guelph, 2005

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate Studies

(Medical Genetics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

January 2012

© Allison Marie Cotton, 2012

Abstract

The process of X-chromosome inactivation achieves dosage compensation between mammalian males and females. In females one X chromosome is transcriptionally silenced through a variety of epigenetic modifications including DNA methylation. Most X-linked genes are subject to X-chromosome inactivation and only expressed from the active X chromosome. On the inactive X chromosome, the CpG island promoters of genes subject to X-chromosome inactivation are methylated in their promoter regions, while genes which escape from X-chromosome inactivation have unmethylated CpG island promoters on both the active and inactive X chromosomes.

The first objective of this thesis was to determine if the DNA methylation of CpG island promoters could be used to accurately predict X chromosome inactivation status. The second objective was to use DNA methylation to predict X-chromosome inactivation status in a variety of tissues. A comparison of blood, muscle, kidney and neural tissues revealed tissue-specific X-chromosome inactivation, in which 12% of genes escaped from X-chromosome inactivation in some, but not all, tissues. X-linked DNA methylation analysis of placental tissues predicted four times higher escape from X-chromosome inactivation than in any other tissue. Despite the hypomethylation of repetitive elements on both the X chromosome and the autosomes, no changes were detected in the frequency or intensity of placental Cot-1 holes.

The third objective of this thesis was to use DNA methylation to investigate X-chromosome inactivation in female samples with chromosomally abnormal karyotypes. The spread of X-chromosome inactivation into the autosomal portion of six unbalanced X;autosome translocations revealed similarities between X;autosome translocations involving the same autosome and therefore suggested a role for DNA sequence in influencing X-chromosome inactivation status of genes. Autosomal genes that escaped from inactivation were found to have significantly lower L1 and LTR but higher Alu content than genes which were subject to inactivation. Lastly, DNA methylation was used to predict the number of inactive X chromosomes in triploid placental samples. Triploid samples provide an excellent system in which to study the counting step of X-chromosome inactivation and DNA methylation analysis provides a means to determine the number of inactive X chromosomes using only a DNA sample.

Preface

Parts of this thesis were performed by collaborators:

- Chapter 2
 - Sample collection was performed by Dr. Deborah McFadden. DNA extraction and sample preparation was performed by the Robinson lab (Ruby Jiang and Dr. Maria Peñaherrera). Illumina Infinium HumanMethylation27 array was run by the Robinson (Ruby Jiang, Dr. Maria Peñaherrera and Dr. Ryan Yuen) and Kobor labs (Sarah Neumann and Lucia Lam). Autosomal DNA methylation analysis of fetal tissues included in this chapter has previously been analyzed [1].
- Chapter 3
 - Sample recruitment and collection as well as DNA extraction and Illumina Infinium GoldenGate array processing was performed by Dr. Deborah McFadden members of the Robinson lab (Luana Avila and Dr. Maria Peñaherrera).
- Chapter 4
 - Sample recruitment and collection as well as DNA extraction and Illumina Infinium HumanMethylation27 array processing was performed by members of the Robinson lab (Luana Avila, Ruby Jiang, Dr. Maria Peñaherrera and Dr. Ryan Yuen). DNA methylation analysis of placental samples included in this chapter has previously been analyzed [2].
- Chapter 5
 - Illumina Infinium HumanMethylatio450 array processing was performed by members of the Kobor lab (Sarah Neumann and Lucia Lam).
- Chapter 6
 - Sample recruitment was performed by Dr. Deborah McFadden. Parent of origin analysis as well as DNA extraction and Illumina Infinium HumanMethylation27 array processing was performed by members of the Robinson lab (Ruby Jiang, Dr. Maria Peñaherrera and Dr. Ryan Yuen). DNA methylation analysis from the Illumina Infinium HumanMethylation27 array of the triploid placental samples included in this chapter has previously been published [3].

Ethics approval was obtained from the University of British Columbia and the Children's and Women's Health Centre of British Columbia. Ethics certificates were as follows:

- Samples used in chapter 2: H08-02773 and H06-70085
- Samples used in chapter 3: H06-70085
- Samples used in chapter 4: H06-70085
- Samples used in chapter 6: H06-70085

Parts of this thesis were previously published in:

1. **Cotton AM**, Lam L, Affleck JG, Wilson IM, Peñaherrera MS, McFadden DE, Kobor MS, Lam WL, Robinson WP and CJ Brown (2011). "Chromosome-wide DNA methylation analysis predicts human tissue-specific X inactivation." Human Genetics. 130(2):187-201
 - The data and text published in this paper are contained in chapter 2. The candidate (A. Cotton) wrote the manuscript and performed all experiments and analysis except for sample collection and processing on the Illumina Infinium HumanMethylation27 array.
2. **Cotton AM**, Avila L, Penaherrera MS, Affleck JG, Robinson WP and CJ Brown (2009). "Inactive X chromosome-specific reduction in placental DNA methylation." Human Molecular Genetics 18(19): 3544-3552
 - All data and part of the text published in this paper are contained in chapter 3. The remainder of the text is contained in the introduction of chapter 4. The candidate (A. Cotton) wrote the manuscript and performed all experiments and analysis except for sample collection and processing on the Illumina Infinium GoldenGate array.

Table of Contents

Abstract	ii
Preface.....	iii
Table of Contents.....	v
List of Tables.....	viii
List of Figures	ix
List of Abbreviations	xi
List of Gene Names.....	xii
List of Histone Modifications and Variants	xiii
Acknowledgements	xiv
Dedication	xv
1 Introduction.....	1
1.1 Thesis overview	2
1.2 Dosage compensation in humans	2
1.3 Features of the mammalian X_i	3
1.3.1 XIST expression and localization	3
1.3.2 Cot-1 holes	4
1.3.3 Histone modifications.....	5
1.3.4 DNA methylation.....	6
1.3.5 Other features of the X_i	9
1.4 Escape from XCI	9
1.4.1 Features associated with escape from XCI.....	10
1.5 XCI in extra-embryonic tissues.....	11
1.5.1 <i>XIST</i> and <i>TSIX</i> expression in extra-embryonic tissues.....	12
1.5.2 DNA methylation in extra-embryonic tissues.....	13
1.6 XCI in human X;autosome translocations.....	13
1.6.1 XCI in balanced and unbalanced X;autosome translocations.....	14
1.6.2 Features of XCI in X;autosome translocations	14
1.7 Spread of XCI.....	15
1.8 Human triploidy	16
1.8.1 Origins and phenotypes of human triploidy	17
1.8.2 XCI ratios and features of XCI in human triploidy.....	18
1.9 Thesis objective	19
2 Chromosome-wide DNA methylation trends on the human X	25
2.1 Introduction	26
2.2 Methods	27
2.2.1 Sample collection and DNA extraction	27
2.2.2 Illumina Infinium HumanMethylation27 array	27
2.2.3 CpG density definitions.....	27
2.2.4 Illumina Infinium HumanMethylation27 array composition and probes removed from analysis	28
2.2.5 Statistical analysis	28
2.2.6 Decision tree to predict XCI status.....	29
2.2.7 RNA extraction and Q-PCR	29
2.2.8 MeDIP and whole genome amplification	29
2.2.9 NimbleGen array processing and analysis.....	30

2.2.10	Expression data.....	30
2.3	Results.....	30
2.3.1	X-linked promoters show differences in DNA methylation dependent on sex and CpG density.....	30
2.3.2	Promoter DNA methylation effectively predicts XCI status.....	31
2.3.3	Tissue-specific XCI is observed at 12% of genes.....	33
2.3.4	Non-island DNA methylation is a poor predictor of XCI status.....	34
2.3.5	X-linked HC and IC promoters show the strongest sex-specific DNA methylation difference.....	35
2.3.6	X-linked genes with higher expression show higher gene-body DNA methylation.....	36
2.4	Discussion.....	37
3	X_i chromosome-specific reduction in placental DNA methylation.....	60
3.1	Introduction.....	61
3.2	Methods.....	62
3.2.1	Sample collection.....	62
3.2.2	DNA extraction and bisulfite conversion.....	62
3.2.3	Illumina GoldenGate panel.....	62
3.2.4	Pyrosequencing.....	62
3.2.5	Statistical analysis.....	63
3.2.6	CpG density definitions used.....	63
3.3	Results.....	64
3.3.1	Placental hypomethylation is specific to X-linked promoters and repetitive elements.....	64
3.3.2	CpG density influences DNA methylation patterns of X-linked promoters.....	64
3.3.3	DNA methylation is female-specific at X-linked island promoters but is consistently high at intragenic and intergenic regions.....	65
3.3.4	Placental hypomethylation is greater on the X _i than the X _a	66
3.4	Discussion.....	67
4	X-linked DNA methylation predicts increased escape from XCI in human placenta...83	83
4.1	Introduction.....	84
4.2	Methods.....	85
4.2.1	Sample collection, DNA extraction and Illumina Infinium HumanMethylation27 array.....	85
4.2.2	CpG density definitions.....	85
4.2.3	Statistical analysis.....	85
4.2.4	RNA FISH.....	86
4.2.5	Post RNA FISH imaging.....	86
4.3	Results.....	87
4.3.1	Placenta gestational age does not significantly affect X-linked promoter DNA methylation.....	87
4.3.2	Placental hypomethylation is found at non-island promoters on both the X chromosome and the autosomes.....	87
4.3.3	X-linked CpG island DNA methylation predicts increased escape from XCI in the placenta.....	88
4.3.4	All repetitive elements examined are hypomethylated in the placenta compared to blood.....	89
4.3.5	Placental cells demonstrate the same frequency of Cot-1 holes as somatic cells.....	90
4.3.6	The X _i of Cot-1 hole positive cells is found most frequently at the nuclear periphery.....	90

4.3.7	The lowest Cot-1 transcription occurs in the Barr body and XIST overlap region of Cot-1 hole positive cells.....	91
4.4	Discussion.....	92
5	DNA methylation demonstrates spread of XCI into X;autosome translocations	106
5.1	Introduction	107
5.2	Methods	108
5.2.1	Sample preparation and bisulfite conversion	108
5.2.2	Illumina Infinium HumanMethylation450 array	108
5.3	Results.....	109
5.3.1	CpG island DNA methylation changes with distance from TSS.....	109
5.3.2	DNA methylation analysis refines the minimal region of translocation breakpoints by 50%	110
5.3.3	DNA methylation analysis predicts varied degrees of spread of inactivation between X;autosome translocations.....	111
5.3.4	Domains which escape from inactivation are depleted for L1s and LTRs but enriched for Alus	112
5.4	Discussion.....	113
6	XCI in human triploid placentas.....	122
6.1	Introduction	123
6.2	Methods	124
6.2.1	Sample collection, processing and Illumina Infinium HumanMethylation27 array.....	124
6.2.2	Predicting %X _i in triploids with DNA methylation from the Illumina Infinium HumanMethylation27 array	124
6.2.3	Statistical analysis	125
6.2.4	Predicting %X _i from pyrosequencing.....	125
6.2.5	XIST RNA FISH.....	126
6.3	Results.....	126
6.3.1	DNA methylation can accurately predict the percent of X chromosomes which are X _i s in control placentas	126
6.3.2	%X _i determined by Illumina agrees with %X _i determined by pyrosequencing	126
6.3.3	Triploid placentas tend to have more than the expected average number of X _i s.....	127
6.3.4	XIST RNA FISH confirms the %X _i predictions from DNA methylation in cultured triploid somatic cell line.....	128
6.4	Discussion.....	128
7	Discussion.....	137
	References	144

List of Tables

Table 1.1: Histone modifications associated with XCI.	20
Table 1.2: Summary of inactivation status as determined by RT-PCR in unbalanced X;autosome translocations.....	21
Table 1.3: Frequency of sex chromosomes in human triploids.	22
Table 1.4: XCI patterns observed in 69, XXX triploids.	23
Table 1.5: XCI patterns observed in 69, XXY triploids.	24
Table 2.1: Q-PCR primer sequences and conditions used to determine the XCI status of <i>TSR2</i> and <i>ZRSR2</i> along with control assay <i>ZFX</i>	42
Table 2.2: Genes in which different probes for the same gene show DNA methylation conflicts in blood.	43
Table 2.3: Novel XCI status as predicted by DNA methylation.	44
Table 2.4: Genes predicted to show tissue-specific XCI.....	45
Table 2.5: Predicted XCI status of X-linked genes in females with multiple fetal tissues.	47
Table 2.6: Tissue combinations of genes predicted to escape from XCI and to be subject to XCI for tissue-specific XCI in females with multiple tissues.	48
Table 2.7: Genes with LC probes that have DNA methylation predictive of an XCI status.	49
Table 3.1: Primer sequences and cycling conditions used in pyrosequencing assays.....	70
Table 3.2: DNA methylation assays for island (HC and IC) promoters failing to show MeXIP or containing features believed to interfere with MeXIP or showing discordant DNA methylation results.	73
Table 3.3: Average percent DNA methylation as determined by pyrosequencing at promoter, intragenic and intergenic regions across the X chromosome in blood and placenta for high (HC), intermediate (IC) and low (LC) CpG density.	75
Table 4.1: Control placental samples used in chapter 4.	96
Table 4.2: Degree of placental hypomethylation calculated as a delta (blood-placenta) at different types of repetitive elements.	97
Table 5.1: Samples used in chapter 5.	116
Table 5.2: Training set of genes which are subject to XCI.....	117
Table 5.3: Training set of genes which escape from XCI.....	117
Table 6.1: Triploid placental samples used in chapter 6.	132
Table 6.2: X-linked probes from the Illumina Infinium HumanMethylation27 array determined to be most accurate at determined %X _i	132

List of Figures

Figure 2.1: Decision tree used to predict the XCI status of X-linked probes.	51
Figure 2.2: Promoter DNA methylation analysis in blood (female: n=59 and male: n=36) reveals X chromosome sex-specific DNA methylation differences as well as differences based on CpG density.	52
Figure 2.3: The XCI status predicted using DNA methylation in blood corresponds with previously determined XCI status.....	53
Figure 2.4: DNA methylation in female somatic tissues is similar to blood.....	54
Figure 2.5: Most genes show the same predicted XCI status in all tissues examined while 12% of genes show tissue-specific XCI.....	55
Figure 2.6: Genes predicted to show tissue-specific XCI based on DNA methylation from the Illumina Infinium HumanMethylation27 array as found across the X chromosome.	56
Figure 2.7: DNA methylation histograms reveal X-linked HC promoters show the largest X-linked sex-specific DNA methylation difference.....	57
Figure 2.8: IC promoters as well as HC non-promoters show a sex-specific DNA methylation difference on the X chromosome but not on the autosomes.....	58
Figure 2.9: X-linked introns but not exons show differences in DNA methylation based on expression level.	59
Figure 3.1: Reduced placental DNA methylation found at L1 repetitive elements and promoters on the X chromosome.....	76
Figure 3.2: Heatmap illustrating DNA methylation levels at 84 sites across the X chromosome from the Illumina GoldenGate promoter DNA methylation array.....	77
Figure 3.3: Effect of distance from transcription start site on DNA methylation.....	78
Figure 3.4: X-linked CpG island promoters show female-specific DNA methylation whereas DNA methylation is high in both X-linked intragenic and intergenic regions in females and males.....	79
Figure 3.5: The X_i shows less placental DNA methylation compared with blood than the X_a at the majority of regions examined across the X chromosome	80
Figure 3.6: The X_i shows less placental DNA methylation compared with blood than the X_a at the majority of promoters examined on the X chromosome.	81
Figure 3.7: Summary of DNA methylation analyses showing placental reduction in DNA methylation predominately on the X_i	82
Figure 4.1: Male and female placentas show minimal X-linked DNA methylation differences across gestational ages.	98
Figure 4.2: X-linked CpG density influences placental hypomethylation differently in males and females.....	99
Figure 4.3: Increased placental escape from XCI predicted using DNA methylation.....	100
Figure 4.4: Non-CpG island repetitive elements are hypomethylated in the placenta on both the X chromosome and the autosomes.....	101
Figure 4.5: The frequency of Cot-1 hole positive and negative cells is not significantly different in placental and somatic cells.	102
Figure 4.6: The X_i of Cot-1 hole positive and negative cells differ in nuclear location.	103

Figure 4.7: Cot-1 intensity is lowest in the Barr body and XIST overlap region of Cot-1 hole negative cells.	104
Figure 4.8: Summary of differences between Cot-1 hole positive and negative cells.	105
Figure 5.1: X-linked CpG island promoter DNA methylation is influenced by distance from the TSS.	118
Figure 5.2: Karyograms and breakpoint location analysis of X;autosome translocations.	119
Figure 5.3: Autosomal CpG island promoter DNA methylation suggests different degrees of spread of inactivation.	120
Figure 5.4: Repetitive element frequency differs based on inactivation status.	121
Figure 6.1: Steps used to calculate %X _i in placenta.	133
Figure 6.2: %X _i in control and triploid placentas as predicted using X-linked DNA methylation.	134
Figure 6.3: Illumina Infinium HumanMethylation27 array and pyrosequencing show correlated levels of predicted %X _i	135
Figure 6.4: XIST RNA FISH confirms %X _i predicted using DNA methylation.	136

List of Abbreviations

$\%X_i$ = percentage of X chromosomes which are inactive X chromosomes
BATMAN = bayesian tool for methylation analysis
bp = base pairs
BLAST = basic local alignment search tool
BrdU = Bromodeoxyuridine
CT = cancer-testis
cDNA = complementary deoxyribonucleic acid
ChIP = chromatin immunoprecipitation
Cot = concentration over time
Cot-1 = concentration over time fraction 1
DAPI = 4'-6-Diamidino-2-phenylindole
DNA = deoxyribonucleic acid
DNMT/Dnmt = DNA methyltransferase
dNTPs = deoxynucleosides
ES = embryonic stem
FISH = fluorescence in situ hybridization
GFP = green fluoresce protein
HC = high CpG density
IC = intermediate CpG density
IN = input
IP = immunoprecipitation
kb = kilo base pairs
L1 = long interspersed elements 1
LC = low CpG density
LINE = long interspersed elements
LTR = long terminal repeats
Mb = mega base pairs
MeDIP = methyl-DNA immunoprecipitation
MeXIP = methylation of X-linked island promoters
MIR = mammalian interspersed repetitive
NCBI = National Center for Biotechnology Information
PBS = phosphate buffered saline
PCR = polymerase chain reaction
PBMCs = peripheral blood mononuclear cell
PRC2 = polycomb repressive complex 2
Q-PCR = quantitative-polymerase chain reaction
RNA = ribonucleic acid
RNA pol II = RNA polymerase II
RT-PCR = reverse transcription polymerase chain reaction
SINE = short interspersed elements
SNP = single-nucleotide polymorphism
TSS = transcription start site
UCSC = University of California – Santa Cruz
UTR = untranslated region
 X_a = active X chromosome
 X_i = inactive X chromosome
XCI = X chromosome inactivation
XIC/Xic = X inactivation centre

List of Gene Names

Note that only genes which are discussed in the body of the thesis are included in this list. Genes in all capital letters refer to human gene names whereas mouse and vole gene names are in lowercase letters.

AR = androgen receptor
ARHGAP6 = Rho GTPase activating protein 6
ATRX/Atrx = alpha thalassemia/mental retardation syndrome X-linked
BHLHB9 = basic helix-loop-helix domain containing, class B, 9
CHM = choroideremia (Rab escort protein 1)
CTCF/Ctcf = CCCTC-binding factor (zinc finger protein)
DDX3X/Ddx3x = DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, X-linked
DKC1 = dyskeratosis congenita 1
DNMT1/Dnmt1 = DNA methyltransferase 1
DNMT3A/Dnmt3a = DNA methyltransferase 3 alpha
DNMT3B/Dnmt3b = DNA (cytosine-5)-methyltransferase 3 beta
EED/Eed = embryonic ectoderm development
EZH2/Ezh2 = enhanced of zests homolog 2
FANCB = Fanconi anemia, complementation group B
FOS = FBJ murine osteosarcoma viral oncogene homolog (previously known as *cFOS*)
G6PD = glucose-6-phosphate dehydrogenase
GPKOW = G patch domain and KOW motifs
HPRT = hypoxanthine guanine phosphoribosyl transferase
Kdm5c = lysine (K)-specific demethylase 5C (previously known as *Jarid1c*)
MeCP2 = methyl CpG-binding protein 2
MID1/Mid1 = midline 1
NDP = Norrie disease (pseudoglioma)
OTC/Otc = ornithine transcarbamylase
PDK3 = pyruvate dehydrogenase kinase, isozyme 3
PNP = purine nucleoside phosphorylase (previously known as *NP*)
Smc1a = structural maintenance of chromosomes 1A (previously known as *Sb1.8*)
SPRY3 = sprout homolog 3 (*Drosophila*) (previously known as *HSPRY3*)
SUZ12/Suz12 = suppressor of zeste 12 homolog
TBL1X = transducin (beta)-like 1X-linked
TIMP1 = tissue inhibitor of metalloproteinase 1
TSIX/Tsix = X (inactive)-specific transcript, antisense
TSR2 = 20S rRNA accumulation
Uba1 = ubiquitin-like modifier activating enzyme 1 (previously known as *Ube1*)
XIST/Xist = X_i specific transcript
VAMP7 = vesicle-associated membrane protein 7 (previously known as *SYBL1*)
ZFX = zinc finger protein, X-linked
ZRSR2 = zinc finger (CCCH type), RNA-binding motif and serine/arginine rich 2

List of Histone Modifications and Variants

Only histone modifications and variants discussed in the thesis are listed here.

acH4 = acetylated histone 4
H2AK9ac = acetylated histone 2 alpha lysine 9
H2A.Z = H2A histone family, member Z
H2K9ac = acetylated histone 2 lysine 9
H2BK5ac = acetylated histone 2 beta lysine 5
H2BK12ac = acetylated histone 2 beta lysine 12
H2BK20ac = acetylated histone 2 beta lysine 20
H2BK120ac = acetylated histone 2 beta lysine 120
H3K4ac = acetylated histone 3 lysine 4
H3K4me2 = dimethylated histone 3 lysine 4
H3K9ac = acetylated histone 3 lysine 9
H3K9me1 = monomethylated histone 3 lysine 9
H3K9me2 = dimethylated histone 3 lysine 9
H3K9me3 = trimethylated histone 3 lysine 9
H3K18ac = acetylated histone 3 lysine 18
H3K4me1 = monomethylated histone 3 lysine 4
H3K4me2 = dimethylated histone 3 lysine 4
H3K4me3 = trimethylated histone 3 lysine 4
H3K20me3 = trimethylated histone 3 lysine 20
H3K27me3 = trimethylated histone 3 lysine 27
H3K36ac = acetylated histone 3 lysine 36
H3R2me = monomethylated histone 3 arginine 2
H3R17me2 = dimethylated histone 3 arginine 17
H3R26me2 = dimethylated histone 3 arginine 26
H3K36me2 = monomethylated histone 3 lysine 36
H4K5ac = acetylated histone 4 lysine 5
H4K8ac = acetylated histone 4 lysine 8
H4K12ac = acetylated histone 4 lysine 12
H4K16ac = acetylated histone 4 lysine 16
H4K20me1 = monomethylated histone 4 lysine 20
HP1 = heterochromatin protein 1
macroH2A = macro histone 2 alpha
meH3 = methylated histone 3
ubH2A = ubiquitinated histone 2 alpha

Acknowledgements

This thesis would not have been possible without many important people. First and foremost, I'd like to thank my supervisor Dr. Carolyn Brown for her guidance and mentorship. I would also like to thank my committee including Dr. Michael Kobor, Dr. Deborah McFadden and Dr. Wendy Robinson for their time and feedback.

The contributions of four laboratories have been indispensable. I am extremely thankful of all the help provided by past and present members of the Brown lab most especially Joslynn Affleck, Sarah Baldry and Nancy Thorogood. Thanks to the Robinson lab for the collection and processing of placental samples, to Dr. Wan Lam and members of his lab for all their insight into MeDIP and to the members of the Kobor lab for their collaboration with the processing of the Illumina arrays.

Lastly I would like to acknowledge my friends and family for all their support. Thanks to my roommate Michelle Decker for making grad school such a wonderful experience. To my sister Cecilia, thank you for all your advice, statistical and otherwise. Lastly and most importantly, I cannot thank my parents enough for all their encouragement and support. Without you this PhD would not have been possible.

Dedication

To my family.

1 Introduction

1.1 Thesis overview

Epigenetic regulation is important in maintaining appropriate gene expression. The process of X-chromosome inactivation (XCI) provides an opportunity to study the role of epigenetic features in preserving gene silencing in a naturally occurring system. Many epigenetic features differ between active and inactive genes however this thesis will focus on DNA methylation. The ability of DNA methylation to predict XCI statuses will be discussed as well as a comparison of X-linked DNA methylation between a number of tissues (blood, muscle, kidney, neural and placenta). DNA sequences are thought to play a role in the spread of XCI but studies of DNA sequences on the X chromosome are confounded by the complex evolutionary history of the sex chromosomes. As an alternative, X;autosome translocations, in which inactivation spreads from the X chromosome onto the autosome, allow for DNA sequences involved in the spread of inactivation to be examined without introducing an evolutionary bias. One of the earliest events in the process of XCI is the counting step in which the number of X chromosomes to be silenced to be established. The study of human triploids allows for the consequences of the counting step to be examined but in order to do so it is necessary to accurately determine the number of inactive X chromosomes (X_i) in a cell. DNA methylation will be used to establish an assay to count the number of X_i s thereby facilitating future investigation of the counting step in XCI. Overall the study of DNA methylation and its role in XCI not only provides insight into a process which occurs in every mammalian female but it also offers a model for better understanding of other systems which rely on DNA methylation to maintain gene silencing.

1.2 Dosage compensation in humans

Human males and females normally have 22 pairs of autosomes but differ in their complement of sex chromosomes. While males have one Y chromosome and one X chromosome, human females have two X chromosomes and no Y chromosome. It is due to this difference in sex chromosomes that chromosomally normal females undergo the process of XCI to achieve dosage compensation with males [4]. In human 46, XX cells, one X chromosome is silenced and becomes the X_i while the other X chromosome remains transcriptionally active and is referred to as the active X chromosome (X_a). In humans, XCI is random and occurs early in development with either X chromosome having the potential to be inactivated and become the X_i [5]. Females are mosaic for two cell populations with different X chromosomes being capable of becoming the X_i because XCI is random. Once XCI occurs, it is stably maintained throughout future divisions. The timing of XCI in human is currently unclear, however in mice, it appears that XCI is reversed in oogenesis to allow for a new round of random XCI to occur in female offspring [6].

1.3 Features of the mammalian X_i

After XCI occurs there are many epigenetic features which differ between the X_a and the X_i . The X_a is composed mostly of loosely packed chromatin known as euchromatin, whereas the X_i is generally composed of more densely packed heterochromatin. The majority of features which differ between the X_a and the X_i are associated with these differences in chromatin structure. The differences between the X_a and the X_i provide a model of how epigenetic modifications work together to create a transcriptionally silent domain. Most features associated with the X_i are also found elsewhere throughout the genome, but the association of the X_i Specific Transcript (*XIST* in humans, *Xist* in mice) with the X_i is unique to the X_i and is the first step in the process of XCI [7].

1.3.1 *XIST* expression and localization

XCI is initiated by the expression of the non-coding RNA (ribonucleic acid) *XIST* which is uniquely expressed from the X_i and in humans is located at Xq13.2 in the X inactivation center (XIC) [7]. Studies in mice have shown that *Xist* is necessary for XCI by demonstrating that deletions of *Xist* result in non-random XCI in which the wild type X chromosome is always inactivated [8, 9]. *XIST* has also been found to be sufficient to cause XCI by the ectopic insertion of *XIST* which leads to inactivation [10-12]. Despite the clear role of *XIST* in the initiation of XCI, once XCI has occurred *XIST* is generally not required to maintain XCI [12-14]. It should, however, be noted that the deletion of *Xist* in mice after the establishment of XCI does result in a higher degree of reactivation of at least two genes (X-linked endogenous *Hprt* and *GFP* (green fluorescent protein) transgene) implying that *Xist* may have role in maintaining the stability of XCI [14].

Once transcribed, the approximately 2000 molecules of *Xist/XIST* RNA [15] are spliced and polyadenylated [7] and coat the X_i in *cis* [16]. In humans, *XIST* RNA does not associate with the X_i during metaphase, however, at interphase *XIST* RNA localizes in a cloud-like structure that covers 80-85% of the X_i domain [16, 17]. The region of the X_i covered by *XIST* RNA is slightly larger than the dense heterochromatic core of the X_i known as the Barr body (discussed further in section 1.3.5) [16, 18]. The association of *XIST* RNA with the X_i domain is of special importance to the study of XCI because all other features associated with the X_i are simply enriched, or depleted, on the X_i but are also present in other locations across the genome [7].

XIST expression can be detected in both male and female human preimplantation embryos [19, 20], however, there are conflicting reports as to the form that *XIST* RNA accumulation takes and even the status of XCI in early human preimplantation embryos. van den Berg [21] found human male preimplantation embryos (8-cell to blastocyst stage) tended to lack *XIST* RNA

signals but did find pinpoint signals in a minority of cells. Human females generally had a single pinpoint XIST RNA signal at the 8-cell stage but a single full XIST RNA cloud by the blastocyst stage [21]. Transcriptional silencing was correlated with the presence of the XIST RNA cloud by both the exclusion of Cot-1 (concentration over time fraction 1) (see section 1.3.2 for more details) and the lack of biallelic X-linked gene expression, suggesting that XCI had indeed begun in human female blastocysts [21]. These results are challenged by a more recent study of human blastocysts which found that both males and females show XIST RNA accumulation without transcriptional silencing, suggesting that although the features of the X_i may be accumulating, XCI has not fully occurred [5]. The presence of XIST RNA without actual transcriptional silencing has also previously been observed in somatic cell hybrids [22]. Therefore, although XIST RNA expression is the most likely first step in XCI, the exact time at which XCI begins in humans is still unknown and further studies are needed.

1.3.2 Cot-1 holes

Genome complexity can be demonstrated through the use of Cot (concentration over time) curves which use the rate at which DNA (deoxyribonucleic acid) sequences reanneal to determine how commonly that repetitive pair of DNA sequences is found throughout the genome [23]. The rate at which DNA anneals is directly related to how frequently the DNA sequence is found throughout the genome with more common DNA sequences annealing more quickly. Eukaryotic genomes typically show complex Cot curves composed of three general levels of genome complexity. These three classes are present due to the differences in the reannealing kinetics demonstrated by each class. The first, highly repetitive class is composed sequences derived from telomeres and centromeres while the slowest annealing class represents unique genome sequences. The middle class contains moderately repetitive DNA and has a Cot value equal to one and is therefore known as the Cot-1 fraction [23]. Included in the Cot-1 fraction of the genome are the LINE (long interspersed elements) and SINE (short interspersed elements) families of repetitive elements which together represent approximately 28.4% (LINE: 16.7%, SINE: 11.7%) of the human genome [24]. Cot-1 DNA can be used as a fluorescence in situ hybridization (FISH) probe to detect transcription of the repetitive elements in the genome. Once the XIST RNA has coated the X_i , one of the first features to differ between the X_a and the X_i is the creation of a nuclear compartment in which transcription is silenced. This can be observed by both the lack of RNA polymerase II (RNA pol II) and by the lack of RNA transcription of the Cot-1 fraction [25-27].

The lack of Cot-1 transcription within this confined nuclear compartment is referred to as a Cot-1 hole. The Cot-1 hole is slightly smaller than the domain covered by XIST RNA and generally overlaps with the Barr body [16]. Cot-1 DNA is composed of LINE and SINE sequences

suggesting that it is these elements which are being transcriptionally silenced to form a Cot-1 hole. The presence of Cot-1 holes may help explain the structure that the X_i takes in the nucleus with the highly repetitive sequences of the X_i , along with the X centromere, coming together to form a dense heterochromatic core while genic sequences remain more peripheral. During early XCI, the presence of a Cot-1 hole does not correspond with gene silencing as transcription of X-linked genes can still be detected. Once XCI occurs, silent and active genes are located peripheral to the Cot-1 hole but silenced genes move more within this inactive domain while genes which escape from XCI remain outside [28, 29]. The relationship between early XCI and transcriptional silencing appears to be complex with different types of elements becoming silent at different times. Whether or not the timing of silencing plays a role in the process of XCI remains to be determined.

1.3.3 Histone modifications

Histones, the proteins around which DNA wraps to form chromatin, can be modified by the processes of ADP ribosylation, acetylation, deamination, methylation, N-acetylglucosamine, phosphorylation, proline isomerization, sumoylation, ubiquitylation and even by the removal of part of the histone tail (histone tail clipping) (reviewed in [30]). The study of histone modifications involved in XCI has generally focused on the acetylation and methylation of histone tails as well as the presence of the histone variant macroH2A and how these marks differ between the X_a and the X_i . In general, the human X_i is characterized by an enrichment of histone marks typically associated with inactive chromatin while histone marks associated with active chromatin tend to be depleted (see Table 1.1). The X_a shows a different pattern of histone modifications, which is to be expected given its active transcriptional state.

In mice, histone modifications can differentiate the X_i and X_a as early as the 8 cell stage with the X_i showing hypoacetylation of H3K9 and hypomethylation of H3K4 [31]. By the mid blastocyst stage, the mouse X_i is enriched for nearly all the marks typically associated with the X_i [31, 32]. However, as with *XIST* expression, there are conflicting results as to when histone marks associated with the X_i appear in human preimplantation embryos. At the blastocyst stage, one study detected an enrichment for the inactive marks H3K27me3 and macroH2A, along with a lack of H3K9ac [21] while another study did not see an enrichment of H3K27me3 [5].

Intriguingly, marsupials, which undergo XCI but lack *XIST*, show an exclusion of active histone marks from the X_i but do not show the degree of enrichment of most inactive histone marks observed on the X_i of placental mammals. Two of the inactive histone marks typically enriched on the human and mouse X_i , H3K9me3 and H4K20me3, are present on the marsupial X_i despite the lack of *XIST*, suggesting that at least some histone modifications are recruited to the X_i independent of *XIST* expression [33].

The role that histone modifications play in maintaining the random XCI in the embryo and the imprinted XCI in the extra-embryonic tissue (discussed further in section 1.5) appears to differ. Female mice with mutations in *Eed* maintain random XCI in the embryo but do not maintain imprinted XCI in extra-embryonic tissues [34]. *Eed*, along with *Ezh2* and *Suz12*, is part of the Polycomb Repressive Complex 2 (PRC2) which is responsible for the trimethylation of H3K27 [35-37]. The PRC2 is recruited to the X_i through the presence of *Xist* RNA and once the PRC2 has laid down H3K27me₃, other proteins including heterochromatin protein 1 (HP1) are recruited and in turn bind to H3K9me₃ [38, 39]. The interaction between H3K9me₃ and HP1 can be observed in the non-overlapping bands of histone modifications found along the human X_i at metaphase [40]. The fact that some histone modifications appear critical to the maintenance of imprinted, but not random, XCI demonstrates that histone modifications likely act along with other epigenetics factors, such as DNA methylation, to maintain transcriptional silencing.

1.3.4 DNA methylation

Of the many different types of epigenetic marks now known, DNA methylation was the first to be identified [41, 42]. In human somatic cells, DNA methylation is found almost exclusively at the cytosine in CpG dinucleotides [43], which are underrepresented across the human genome due to the process of deamination in which the methylated cytosine of a CpG mutates into a thymine [44]. CpGs are enriched at the promoters of 60% of genes, resulting in regions known as CpG islands [44, 45]. Definitions of what constitutes a CpG island differ [46, 47]; however, the use of three levels of CpG density: high CpG density (HC), intermediate CpG density (IC) and low CpG density (LC) allows for the unique properties of CpG islands to be dissected [45]. Of the approximately 1.2 million CpGs on the human X chromosome only 5% are located within CpG islands, and approximately 50% of these islands are associated with known promoters, while the remaining islands are equally distributed within gene bodies and between genes across the X chromosome [48].

DNA methylation is controlled by a group of proteins called DNA methyltransferases (*DNMTs* in human, *Dnmts* in mice). *De novo* DNA methylation is established by *Dnmt3A* and *Dnmt3B* [49]. Mutations of *DNMT3B* in humans give rise to ICF (immunodeficiency, centromere instability and facial abnormalities) [50] and these patients show hypomethylation of satellite DNA [51] as well as hypomethylation of X-linked CpG island promoters [52]. *Dnmt1* is the maintenance methyltransferase and methylates hemi-methylated DNA which exists after DNA replication, thereby ensuring that daughter cells have the same DNA methylation patterns as their parent cell [53, 54]. Mutations in *Dnmt1* can result in ectopic expression of *Xist* in male mouse embryos leading to XCI [55, 56]. The link between DNA methylation and XCI was first proposed

by Riggs in 1975 [42]. Multiple lines of evidence support the importance of DNA methylation in maintaining XCI including the reactivation of human X-linked genes upon treatment with DNA methyltransferase inhibitors [57] and the high reactivation frequency for X-linked genes from the hypomethylated marsupial X_i during cell culture [58-60].

1.3.4.1 DNA methylation at promoters

The presence of DNA methylation at CpG island promoters is generally associated with the transcriptional silencing of the associated gene [61, 62]. Given that the X_i is largely heterochromatic, and therefore not expressed, it is not surprising that the CpG island promoters of genes on the X chromosome that are subject to XCI typically have CpG islands that are unmethylated in males, but show partial DNA methylation in females [63, 64]. This reflects that CpG island promoters on the X_a are unmethylated, similar to autosomal CpG island promoters, while CpG island promoters on the X_i are methylated. We term this pattern of low male DNA methylation in combination with moderate female DNA methylation, MeXIP (DNA methylation of X-linked island promoters). Those genes with promoter CpG islands that escape from XCI (discussed further in section 1.4) tend to be unmethylated on both the X_a and X_i [65] which is indicative of expression from both the X_a and the X_i . The consistent relationship between X-linked CpG island promoter DNA methylation and XCI status has been shown in many individual gene studies [66-68] and in a study of neutrophils to propose novel genes which escape from XCI [69].

Most studies examining the properties of CpG island promoters have examined autosomes and have excluded the X chromosome as to not be confounded by differences in DNA methylation caused by sex. Autosomal promoters associated with CpG islands are typically unmethylated in all tissues, however, a subset of autosomal promoters has been found to show tissue-specific DNA methylation. The regions surrounding CpG islands, named CpG island shores, show the largest DNA methylation differences between tissues [70-73]. Aberrant DNA methylation of CpG island promoters has been associated with cancer [74], and mutations in the proteins associated with DNA methylation can also result in disease (for example MeCP2 mutations cause Rett syndrome, reviewed in [75]).

Methylated cytosines in CpGs undergo deamination to become thymines which results in a mutation in the DNA sequence whereas the deamination of unmethylated cytosines results in uracil and is readily repaired [44]. Cytosines in CpGs which are methylated in the germline are therefore likely to be lost over time, the remaining cytosines in CpGs must be unmethylated in the germline for a biological reason. Although roughly 60% of genes have promoters associated with CpG islands [44, 45], only approximately 50% of CpG islands overlap a known

transcription start site (TSS) [72, 76]. The remaining 50% of CpG islands, called orphan CpG islands, are distributed throughout the intragenic and the intergenic regions of the genome [72, 76]. Despite not being associated with a known gene, these orphan CpG islands are often associated with promoter histone modifications and RNA Pol II, suggesting that at least 40% to 60% of orphan CpG islands are acting as promoters for currently unknown genes [76]. RNA sequence data also suggests that there are transcripts which initiate at many orphan CpG islands [77]. The remaining orphan CpG islands, as well as CpGs not found in islands, provide an opportunity to study the role that DNA methylation plays in non-promoter regions across the genome.

1.3.4.2 Intergenic and intragenic DNA methylation

CpG island promoters represent less than 1% of the DNA of the X chromosome; therefore, the study of DNA methylation across non-promoters is important in gaining a complete picture of X chromosome-wide DNA methylation levels. It appears that X-linked gene bodies are hypermethylated on the X_a compared to the X_i [78, 79] and it has been proposed that this is due to gene transcription on the X_a [80]. The link between transcription and gene-body DNA methylation has not been detected just on the X chromosome but across the genome, with highly expressed genes showing more gene-body DNA methylation than genes with low expression [81]. There is conflicting evidence that gene-body DNA methylation may down-regulate transcription [82, 83]. The gene-body is not a homogenous region rather introns, exons and UTRs (untranslated region) all contributing unique DNA methylation patterns. Within genes, the 5' most exons show DNA methylation which, like CpG island promoters, correlates with gene silencing, while both internal exons and introns show different DNA methylation [84, 85].

The majority of DNA methylation is found at CpGs not located in CpG islands [43, 70] and approximately half of all CpGs are contained within repetitive elements [86]. DNA methylation is an important means of silencing repetitive elements [87] and, as will be discussed later (see section 1.7), the X chromosome is enriched for the L1s (long interspersed elements 1) [88] which are known to be hypermethylated on both the X_a and X_i in normal cells [89]. Efforts to compare overall DNA methylation levels between the X_a and the X_i have yielded conflicting results depending on the technique being used. DNA methylation-sensitive restriction enzyme analysis showed the X_i to be hypomethylated compared with the X_a , whereas, *in situ* nick translation data demonstrated the X_a to be hypomethylated compared with the X_i , and antibody staining demonstrated no difference between the X chromosomes [90-92]. Because different

regions on the X chromosome have different DNA methylation patterns, it is unwise to compare DNA methylation across the entire X_a and X_i .

1.3.5 Other features of the X_i

In 1949, the presence of a dense region of heterochromatin in the nuclei of female, but not male, cells was the first cytogenetic evidence for the process of XCI. This region was termed the Barr body [93] and was later determined to be the condensed heterochromatin of the X_i [94, 95]. More recently, however, it was found that the Barr body is smaller than the X_i domain and represents only the condensed heterochromatic core of the X_i [16]. Despite the heterochromatic nature of the X_i , comparisons of the X_a and X_i domains have shown that while they differ with respect to shape, they occupy roughly the same amount of space in the nucleus [96, 97]. The X_a and X_i differ with respect to their usual nuclear location with the X_i typically found to localize to the nuclear periphery, or more rarely, to the nucleolus [98]. Another feature of heterochromatin is that it replicates later in S phase than euchromatin [99]. It is therefore not surprising that overall the X_i replicates later in S phase than the X_a and that early studies of XCI used the late replicating nature of the X_i to identify and count X_i s [100-102]. Further examination of the X_i also revealed different degrees of late replication detected as bands across the X chromosome [103-105]. The many features which differentiate the X_a and the X_i illustrate the unique nature of the X_i and the importance of epigenetic features in XCI.

1.4 Escape from XCI

The purpose of XCI is to achieve dosage compensation between males and females and to do so, the majority of genes are subject to XCI. However, using somatic cell hybrids it was found that approximately 15% of human X-linked genes escape from XCI and are expressed from both the X_a and the X_i [106], compared to the 3% of genes which escape from XCI in mice [107]. Escape from XCI is usually defined as expression from the X_i at a minimum of 10% of the X_a but can range as high as X_a levels [106, 107]. Given that more genes escape from XCI in humans than in mice, it is not surprising that genes which escape from XCI are found in different locations in the two species. Human genes which escape from XCI tend to be found in clusters whereas in mice, single escape genes tend to be flanked by genes subject to XCI [106, 107]. At each tip of the X chromosome lie pseudoautosomal regions which correspond to regions of homology between the X and Y chromosomes [108]. As would be expected for X-linked genes which have Y homologs, all of the genes examined in the Xp pseudoautosomal region escape from XCI [106]. Additional genes which escape from XCI are found distributed throughout the X chromosome [106]. The presence of a Y homolog does not always equate to escape from XCI. In the Xq pseudoautosomal region, *VAMP7* (previously known as *SYBL1*) and *SPRY3* are inactivated on both the X and Y chromosomes [109, 110]. While *VAMP7* has a

promoter CpG island and is hypermethylated on both the X_i and Y chromosomes [111] and *SPRY3* lacks a promoter CpG island [112], both genes maintain a single gene dose in both males and females without XCI.

While some genes have been found to escape from XCI in all females, other genes escape from XCI in only a subset of females and are subject to XCI in other females. This expression pattern is referred to as variable escape from XCI and is found at approximately 10% of human X-linked genes [106]. Variable escape from XCI may explain why some females who carry certain mutations are affected by a disease while others are not. For example, a study of female twins where only one was affected with primary biliary cirrhosis found lower expression of two X-linked genes in the affected twin suggesting that these genes escaped from XCI in the non-affected twin but were subject to XCI in the affected twin [113]. There is also evidence that escape from XCI can vary within the same individual as a female ages and that as females age there is an increase in the degree of skewing of XCI [114-117]. As telomerase deficient mice age there is a decrease in X-linked H3K27me3 and the reactivation of at least one X-linked gene [118] while *Otc* is known to reactivate with increased age [119]. Variable expression, suggesting variable escape, of X-linked genes has been observed across tissues [120] indicating that escape from XCI in one tissue does not imply escape from XCI in all tissues.

Genes which escape from XCI may help to explain the phenotypes of those individuals with an abnormal number of X chromosomes. The consequences of genes which escape from XCI are obvious through the comparison of 46, XX females to females who lack an X chromosome (45, X). 39, X mice have minimal negative phenotypes and are fertile [121] whereas, 99% of 45, X human conceptuses abort *in utero* and those which survive into adulthood are rarely fertile [122, 123]. Some features, such as the sterility, of 45, X females are likely due to a lack of chromosome pairing during meiosis [124] while other features, such as short stature, can be attributed to the lack of the second copy of genes which escape from XCI (reviewed in [125]). Conversely, when either males or females have an extra X chromosome, the resulting phenotype would likely be due to an overexpression of the genes which escape from XCI. In humans, the phenotype caused by the presence of an extra X chromosome is much less severe than monosomy X, although there is increased mortality associated with the presence of an extra X chromosome in both males and females [126, 127].

1.4.1 Features associated with escape from XCI

Given that many features associated with the X_i are related to its transcriptionally silent nature, it is not surprising that genes which escape from XCI often show a set of features more similar to the active chromatin of the X_a . One might expect genes which escape from XCI to be located outside of the XIST RNA domain while genes which are subject to XCI to be contained within

the transcriptionally silent domain. However, it appears that regardless of XCI status, all genes are found at the edge of the X_i domain [18]. Despite the similar nuclear location of genes which are subject to XCI and escape from XCI, regions of escape from XCI replicate along with the X_a , not later in S phase as the majority of the X_i does [128]. Additionally, regions which escape from XCI tend to lack inactive histone marks [129, 130] but maintain active histone marks [130-132]. The promoters of genes which escape from XCI are also hypomethylated (see section 1.3.4.1) compared to those genes which are subject to XCI [65]. Genes which escape from XCI are generally surrounded by genes which are subject to XCI and they therefore provide the opportunity to study the interaction between active and inactive epigenetic marks and possibly even the elements which prevent the spread of inactivation into active domains.

In addition to differences in the frequency of epigenetic features, genes which escape from XCI appear to have differences in the surrounding sequence compared with genes which are subject to XCI. Specifically, genes which escape from XCI are associated with fewer LINES than genes subject to XCI [88, 133, 134] but are also depleted in AT-rich sequences [135, 136]. These differences in sequence composition are interesting given that the dense heterochromatic core of the X_i is thought to be composed mainly of repetitive elements (see section 1.3.2). It has therefore been proposed that escape elements may be capable of looping out of the X_i domain due to their lack of these repetitive elements [137]. The mouse gene *Kdm5c* escapes from XCI regardless of where it is located on the X chromosome, suggesting that it carries a specific sequence which ensures escape from XCI [138]. Many questions still exist as to the means by which individual genes escape from XCI. It is unclear whether all genes escape from XCI by the same mechanisms or if some escape from XCI due to an inability to spread XCI (discussed in 1.7) while others escape from XCI due to specific elements which induce escape.

1.5 XCI in extra-embryonic tissues

Extra-embryonic tissues include the amnion, chorion and the placenta and are a unique group of organs that are critical for healthy *in utero* development, but are no longer needed after birth. Although the mouse is typically used as a model of human XCI, studies of extra-embryonic tissues have highlighted many differences between XCI in the mouse and the human. In the mouse, extra-embryonic tissues which are derived from the trophoblast layer of the blastocyst show imprinted XCI in which the paternal X chromosome is always selected to become the X_i [31, 139]. The imprinted XCI is erased in the cells of the inner cell mass to allow for random XCI to occur in the embryo and in the extra-embryonic tissues which are derived from the inner cell mass [31, 140]. A variety of findings on XCI in human extra-embryonic tissues have led to different conclusions including the statement that the XCI status of human extra-embryonic

tissues is not known. Skewed XCI where the paternal X chromosome was preferentially inactivated has been reported [141-145], however, preferential inactivation of the paternal X chromosome represented only 49% of these samples and samples with random XCI (39%) and preferential inactivation of the maternal X chromosome (11%) were also found [141-145]. The over-representation of cells with preferential inactivation of the paternal X chromosome may represent a selective advantage of this XCI pattern but not true imprinted XCI as is seen in the mouse extra-embryonic tissue. The viability of 45, X females as well as 47, XXX females regardless of the parent of origin are evidence against imprinted XCI in human extra-embryonic tissue. If human extra-embryonic tissues underwent true imprinted XCI in which only the paternal X chromosome was capable of becoming the X_i then only 45, X females with a maternal X chromosome would be viable, as 45, X females with a paternal X chromosome would inactivate their single X chromosome resulting in death. Moreover, 47, XXX individuals who carried two maternal X chromosomes and one paternal X would inactivate the single paternal X and would therefore have two X_a s. Clearly neither of these situations occur as 45, X and 47, XXX individuals can be viable regardless of parent of origin [146-148]. The cause for the different observations of XCI of human extra-embryonic tissues may be due to large patches of cells with different XCI patterns [149]. Additionally, when multiple X-linked loci are examined, several different XCI patterns can be found within a single extra-embryonic tissue [149]. Taken together, all these facts demonstrate that human extra-embryonic tissues undergo random XCI, not imprinted XCI as is observed in mouse extra-embryonic tissues.

1.5.1 *XIST* and *TSIX* expression in extra-embryonic tissues

Cells from human extra-embryonic tissues have long been known to be more capable of undergoing X reactivation than cells from somatic tissues [150, 151]. The reactivation of single gene loci has also been detected in human extra-embryonic tissues [152]. The fusion of term placental cells with mouse A9 cells to create somatic cell hybrids has demonstrated that placental cells are also capable of undergoing complete X reactivation [150], a process typically only observed during oogenesis [153]. Those placental cells capable of undergoing X reactivation have the same amount of *XIST* RNA as those placental cells incapable of undergoing X reactivation [151]; upon reactivation of the X chromosome, *XIST* expression is repressed indicating that the X chromosome is now active [154].

In both mice and humans there is an extra-embryonic specific and antisense transcript to *Xist/XIST* called *Tsix/TSIX*. In mice *Tsix* is key in controlling *Xist* expression and therefore in regulating XCI [155, 156]. The role of *TSIX* in humans is much less clear. There are many differences between the mouse and human copies of *Tsix/TSIX* including that *TSIX*, unlike *Tsix*, does not cross the TSS or the A repeats of *XIST* and lacks a promoter CpG island [157-159].

However, the most striking difference between *Tsix* and *TSIX* is that while *Tsix* is expressed from the X_a [155], *TSIX* appears to be expressed from the X_i [158]. Although *TSIX* is expressed from the X_i along with *XIST*, the size of the *TSIX* RNA signal is usually smaller than that of the *XIST* RNA [158]. An additional difference between mice and humans is that *Tsix* expression is lost in differentiated mouse cells [155] whereas *TSIX* expression has been detected in children up to age 8 [158]. In humans, *TSIX* is not paternally imprinted and its expression can be detected from the X_i regardless of parent of origin [158]. All the differences between *TSIX* and *Tsix* are generally thought to indicate that *TSIX* in humans does not have the same function as in mice where it protects the X_a from XCI.

1.5.2 DNA methylation in extra-embryonic tissues

The placenta has been shown by high-performance liquid chromatography to have approximately 20% fewer methylated cytosines compared with the vast majority of other tissues, other than sperm [160-162]. Regions previously shown to have reduced placental DNA methylation include repeat elements (Alus, LINEs, and satellites) and several X-linked genes with CpG island associated promoters [161, 163-166]. As discussed previously (see section 1.3.4), DNA methylation is established through the actions of the DNMTs. Given the hypomethylated nature of the placenta it is not surprising that the promoter of *DNMT1* is hypermethylated in human placentas compared to all other tissues examined. Correspondingly, the level of *DNMT1* expression is lower in the placenta than somatic tissues, possibly explaining the hypomethylated nature of the placenta [167]. The hypomethylated nature of the placental has been proposed to create a permissive environment in which X reactivation can occur [151] and at the single gene level, reactivation is associated with hypomethylation [150]. Given that DNA methylation is thought to lock in XCI (see section 1.3.4) is it not surprising that hypomethylation in the placenta may loosen control on XCI and allow for X reactivation to occur.

1.6 XCI in human X;autosome translocations

In humans, X;autosome translocations occur in approximately one to three in 10,000 live births [168] but the phenotypes range from severe (for example [169, 170]) to mild (for example [171, 172]) and XCI patterns associated with X;autosome translocations are diverse. Initial studies of X;autosome translocations in mice demonstrated that inactivation could spread out of the X chromosome portion of the translocated chromosome into the autosomal portion thereby silencing autosomal genes [173-175]. While any translocation carries the risk of disrupting genes [176] or being subject to position effects (reviewed in [177]), X;autosome translocations carry the additional risk of gene imbalances due to autosomal genes being silenced by the spread of inactivation. X;autosome translocations can also involve the X chromosome not

being in disomy and together these gene imbalances are likely the causes of the many, and diverse, phenotypes associated with X;autosome translocations.

1.6.1 XCI in balanced and unbalanced X;autosome translocations

As discussed above (see section 1.3.5), early studies on XCI used the late replication timing characteristic of heterochromatin to identify the X_i . It was therefore a natural extension to use late replication as a means to predict if, and to what degree, the autosomal portion of the X;autosome translocation underwent inactivation. Although the challenges associated with only using replication timing to mark XCI were well known (reviewed in [178]), the assumption that late replication corresponded to XCI seemed appropriate. Through late replicating timing it was determined that in balanced X;autosome translocations, the normal X chromosome is typically inactivated [179] whereas in unbalanced X;autosome translocations, the derivative X chromosome is usually inactivated (see Table 1.2) [180]. This example of non-random inactivation appears to be due to secondary selection in which those cells which maintain the most normal expression survive. Those cells in which inactivation causes major expression imbalances do not survive. In balanced X;autosome translocations, the inactivation of the X;autosome translocation would result in the silencing of autosomal genes and therefore a decrease in their expression and would also result in a portion of the X chromosome not being silenced. Individuals with balanced X;autosome translocations tend to have mild phenotypes [179], although the selection of the normal X chromosome as the X_i could in theory result in the over-representation of X-linked recessive disorders. In non-translocation carrying females, the disruption of an X-linked gene can result in skewed XCI in which the X chromosome carrying the disrupted gene is always inactivated (eg. *DKC1* [181, 182]). When an X;autosome translocation disrupts such a gene it would appear that the inactivation of the derivative chromosome is selected for, causing the normal X to remain active [183].

1.6.2 Features of XCI in X;autosome translocations

The spread on inactivation into the autosomal portion of X;autosome translocations was shown through the functional deletion of Retinoblastoma on X;autosome translocations [184, 185] and later confirmed through the use of RT-PCR (reverse transcription polymerase chain reaction) [186]. This was the first time that autosomal genes on the X;autosome translocations were shown to both escape and be subject to inactivation by direct expression analysis [186] rather than through replication timing. It had previously been observed that when the autosomal portion of X;autosome translocations did show late replication timing it was often detected in bands which did not extend across the entire autosomal portion of the translocated chromosome [187, 188]. The correlation between late replication timing and expression of autosomal genes was examined in 2001 by Sharp *et al.* [189]. Sharp *et al.* found that although

no spread of late replication was found in the autosomal portion of the X;autosome translocation, four out of five genes were found by RT-PCR to be subject to inactivation [189].

Despite the lack of late replication timing, the detection of silenced autosomal genes led to examination other features of XCI, such as XIST localization, histone modifications and DNA methylation. Again, it was found that silencing of autosomal genes occurred despite the lack of cytogenetic features typically associated with XCI [190, 191]. XIST RNA has been found to associate with a higher proportion of the X chromosome portions of X;autosome translocation compared to the autosomal portions, despite the fact that other marks associated with transcriptionally silent chromatin were present across the autosomal portion [191]. The DNA methylation at a limited number of autosomal genes has also been examined and showed good agreement with the inactivation status predicted based on expression [190, 192]. The examined histone modifications (H3K14ac, H4K8ac and H3K4me2) were better associated with inactivation than late replication and genes within late replicating regions were always inactive although inactivation was also possible in non-late replicating regions [190]. The combination of some, but not all, marks of XCI in X;autosome translocations suggests a complicated relationship between the spread of silencing into the autosomal portion of X;autosome translocations as well as the maintenance of XCI.

1.7 Spread of XCI

The different degree of spread of inactivation into the autosomal portions of X;autosome translocation led to the hypothesis that the X chromosome must be more receptive to the spread, and possibly maintenance, of XCI than autosomal DNA. The original proposal was that “way stations” spread XCI along the X chromosome and that these elements, while not unique to the X, were enriched on the X chromosome [193]. In 1998, L1 elements were proposed to be such elements based on their enrichment on the X chromosome [194]. The human X chromosome is almost twice as rich in L1s as the human autosomes but is especially enriched at the XIC [88, 195, 196]. Studies examining the frequency of repetitive elements on the X chromosome have found the examination of larger regions of DNA more accurate at predicting XCI status than smaller regions [133, 135], suggesting that XCI is determined at the level of large domains rather than a small single locus level. L1s have been found to be enriched at genes which are subject to XCI [133, 135, 197]. LINEs are thought to aid in the spread of XCI through the formation of the Cot-1 hole (see section 1.3.2). The early creation of the Cot-1 hole corresponds to the silencing of the highly repetitive fraction of the genome and is thought to pull nearby sequences into this transcriptionally silent domain [27, 197]. A high frequency of LINEs at the breakpoint in an X;autosome translocation patient with minimal phenotype has been

proposed to be the reason for the efficient spread of inactivation into the autosomal portion of the X;autosome translocation [172].

In addition to L1s, MIRs (mammalian interspersed repetitive) and AT-rich sequences are most consistently enriched at genes subject to XCI whereas Alus and GC-rich sequences are enriched at genes which escape from XCI [133, 135, 197]. Dinucleotide repeats, which are capable of forming unique secondary structures, are also enriched across the entire X chromosome compared to the autosomes. It has been proposed that these repeats are key to forming the heterochromatin of the X_i [18]. The largest region of escape is found at the Xp pseudoautosomal region and in this region there is a 10 fold enrichment for (GATA)_n repeats [18]. Hall and Lawrence propose that these repeats may prevent heterochromatinization thus allowing a region to escape from XCI.

A possible confounding factor in any study of the genomic content of the X chromosome is that the X chromosome is composed of two pseudoautosomal regions as well as five blocks of DNA of different evolutionary age known as strata. The youngest stratum (S5) is closest to pseudoautosomal region one and contains the highest number of X-linked genes which still have Y homologs while stratum one is the most divergent from the Y chromosome and is therefore considered the oldest [198]. The different strata show varying degrees of escape from XCI [106], however, these regions also show varying genome compositions and it is unclear if escape from XCI is a product of these differences or if it is related to their evolutionary ages. Another possible confounding factor is that while some genes may be subject to XCI or escape from XCI due to large domain based sequence differences, the XCI status of other genes may be determined on a smaller scale. Given the differences between the size of escape domains in mice and humans, this suggests that the mechanism of escape may not always be the same. Evidence from mice suggests that genes can be in close proximity but show different XCI patterns due to the presence of a CTCF (CCCTC-binding factor (zinc finger protein)) boundary element [199].

1.8 Human triploidy

The study of triploids provides insight into the counting step of XCI which determines the X:autosome ratio and ensures that one X chromosome remains active per diploid autosome set [200-202](and reviewed in [203]). Triploidy occurs when an individual has three complete sets of haploid chromosomes (instead of the normal two haploid sets) resulting in 69 chromosomes. Roughly 1-3% of human conceptions are triploid and are usually spontaneously aborted [204-206]. Triploids account for about 10% of all spontaneous abortions [207] and are the most common chromosomal abnormality found in first trimester spontaneous abortions [208]. Very rarely, one in every 10,000 live births, a triploid can be carried to term [209, 210]. However,

liveborn triploids do not generally show long term survival (a summary of long surviving triploids can be found in [211]) with the longest survival found to be 312 days [212]. Those liveborn triploids which do show long term survival have an increased likelihood of being mosaic for a diploid cell line highlighting that pure triploidy is generally not viable [213].

1.8.1 Origins and phenotypes of human triploidy

Triploids can be divided into two groups based on the parent of origin of the extra haploid set of chromosomes. In diandric triploids, the extra haploid set is from the father and can in theory result from two sperm fertilizing the same egg (dispermy) or from a nondisjunction event during either meiosis I or meiosis II of spermatogenesis. Digynic triploids receive the extra haploid set from the mother and usually result from a nondisjunction event during either meiosis I or meiosis II of oogenesis and very rarely from the fusion of two eggs (dieggy) [204, 206].

Differences in ages and parent of origin of triploid samples have greatly affected conclusions made in triploid research. Early studies indicated that over two thirds of triploids were diandric with the majority resulting from dispermy [204, 210, 214-217], while more recent studies point to the majority of triploids being digynic in origin [218-220]. These conflicts may be explained by examining the different phenotypes associated with the parent of origin of the extra haploid set. When all studies are taken together it would appear that digynic triploids tend to either abort during early development (<8.5 weeks) or develop into a well formed fetus and survive until late gestation; on the other hand, diandric triploids tend to be older but lack an embryo ([221, 222] summarized in [204, 206]). Since early studies tended to examine spontaneous abortion samples (regardless of phenotype) which were between 8 and 20 weeks [204, 210, 214, 215] and later studies tended to limit samples, either based on gestational age greater than 10 weeks or the presence of a well defined fetus [218-220, 223], the parent of origin effect on triploid phenotypes may have misled previous conclusions on triploid frequency.

Triploids can possess three different sex chromosome combinations. If dispermy and dieggy occur at the same frequency and then the expected frequencies of each sex chromosome combination are: 37.5% 69, XXX, 50% 69, XXY and 12.5% 69, XYY. However, as can be seen in Table 1.3, an analysis of liveborn and spontaneously aborted triploids reveals that the observed frequency of sex chromosomes is not as expected. Specifically, there is a lack of 69, XYY triploids which is possibly due to either the uncommonness of process by which 69, XYY triploids occur or by a significant decrease in the viability of 69, XYY triploids. The number of 69, XYY embryos found by preimplantation genetic diagnosis was significantly higher than the number found in the first trimester and none were found in the second trimester suggesting that although 69, XYY triploids do occur, they typically abort before detection [224].

1.8.2 XCI ratios and features of XCI in human triploidy

In diploids, the counting step of XCI results in one X_i and one X_a ; in cells with three X chromosomes, but an otherwise normal complement of chromosomes (47, XXX), it results in two X_i s [158, 225]. Triploids do not have the normal X:autosome ratio observed in either 46, XX or 46, XY individuals and, as such, human triploids have a variety of XCI patterns, not only differing between individuals but even within different cells from the same individual [226](see Table 1.4 and Table 1.5 for a summary of observed XCI patterns in human triploids).

Early studies of human fetal triploid cells showed more 69, X_aX_aY cells than 69, X_aX_iY cells in 69, XXY triploids and more 69, $X_aX_aX_i$ cells than 69, $X_aX_iX_i$ cells in 69, XXX triploids [225-227]. The cells of some liveborn triploids, however, tend to show more 69, X_aX_iY cells in 69, XXY triploids [228] and more 69, $X_aX_iX_i$ cells in 69, XXX triploids [229, 230]. While there may be a survival advantage to having only one X_a , there appears to be a growth advantage (faster or better growth) to cells with more X_a s since the longer both 69, XXX and 69, XXY triploid cells are cultured the more X_a s are present [225, 228]. Studies on triploid human ES (embryonic stem) cells have also found that the number of X_a s increases with passage [231]. The increase in the number of X_a s in older triploids does not appear to be a reactivation of X_i s into X_a s since clonal analyses of triploid cells results in true breeding clones [225]. The lack of reactivation of X_i s in triploids is important as it demonstrates that XCI in triploids is stable despite the diverse number of X_i s observed in different triploids.

The majority of data on XCI in human triploids have been based on the late replicating nature of the X_i (see Table 1.4 and Table 1.5), although there is additional evidence to support that the detected late replicating triploid X chromosomes are the X_i s. First, the level of XIST RNA observed on the triploid X_i is similar to the levels of XIST RNA found on the X_i in 46, XX cell lines. This holds true for triploid cells which contain only one X_i as well as those that have two X_i s and therefore, twice the XIST RNA [225]. Second, DNA methylation at the *G6PD* promoter shows the appropriate levels of DNA methylation based on the number of X_i and X_a found in varying triploid cultures [225]. Third, overall X:autosome expression is higher in triploid cells with two X_a s than in triploid cells with only one X_a [232] suggesting that the previously observed features of the X_i are in fact resulting in XCI.

Only one study has examined the XCI patterns of human extra-embryonic triploid tissues; it showed that triploid extra-embryonic tissues have a significantly higher number of X_i s than triploid embryonic tissue [226]. An age specific difference in the number of X_i was also observed within the triploid extra-embryonic tissue samples; the older the tissue, the more X_i s were found [226]. Tissue-specific differences in the degree of XCI have also been observed between other tissues. In two liveborn triploids, distinct XCI patterns were found in separate

tissues suggesting that tissues may respond differently to the proposed growth advantages conferred by the presence of more X_a s [230, 233]. The varying number of X_i seen in human triploids illustrates that the counting step of XCI truly does rely on the number of autosomes and may provide insight into how this process works.

1.9 Thesis objective

DNA methylation is thought to be an important lock in maintaining XCI. In my thesis I investigated whether the DNA methylation level of X-linked CpG islands could be used to predict XCI status. A variety of tissues were examined including extra-embryonic tissues to determine the nature of placental hypomethylation and how this might have influenced XCI. DNA methylation was also used to predict the degree of spread of inactivation into the autosomal portions of X;autosome translocations. Human triploid samples underwent DNA methylation analysis as a means to determine the variable number of X_i s.

Table 1.1: Histone modifications associated with XCI.

		mouse	human
enriched on X _i	macroH2A	[234]	[40, 235]
	H3K9me1	[236, 237]	-
	H3K9me2	[33]	[129, 235]
	H3K9me3	[33]	[40]
	H3K27me3	[35, 37]	[40]
	ubH2A	[238, 239]	-
	H3K20me3	-	[40]
	HP1	* NOT enriched in mouse [237]	[235]
	H4K20me1	[37]	[240]
depleted on X _i	H3R2me	-	[235]
	H3R17me2	[241]	[235]
	H3R26me2	-	[235]
	H3K4me2	[129, 241]	[40, 235]
	H3K36me2	[241]	[235]
	acH4	[131, 241, 242]	[243]
	H3K9ac	[241]	[21]
	meH3	[241]	[235]

Table 1.2: Summary of inactivation status as determined by RT-PCR in unbalanced X;autosome translocations.

Note that studies of inactivation in unbalanced X;autosome translocations which did not include RT-PCR are not listed in this table.

	Techniques used to detect spread of inactivation	Inactivation status				Distance from TSS of farthest autosomal gene subject to inactivation from:		DNA methylation data	Reference
		genes	subject	escape	unknown	autosomal breakpoint	XIC		
46,X,der(X) t(X;4)(q22;q24)	RT-PCR of somatic cell hybrids	20	70%	30%	0%	~ 81 Mb	~ 113 Mb	-	[186]
46,X,der(X) t(X;10)(q26.3;q23.3)	BrdU labeling and RT-PCR	5	80%	20%	0%	~ 23 Mb	~ 85 Mb	-	[189]
46,X,der(X) t(X;6)(p11.2;p21.1)	RT-PCR of lymphoblasts	9	56%	33%	11%	~ 43 Mb	~ 64 Mb	DNA methylation agrees with XCI status	[190]
46,X,der(X) t(X;6)(q28;p12)	RT-PCR of lymphoblasts	7	43%	57%	0%	~ 35 Mb	~ 113 Mb	DNA methylation agrees with XCI status	[190]
46,X,der(X) t(X;7)(q27.3;q22.3)	RT-PCR of lymphoblasts	3	33%	33%	33%	~ 9 Mb	~ 80 Mb	-	[190]
46,X,der(X) t(X;11)(q26.3;p12)	RT-PCR of lymphoblasts	12	67%	17%	17%	~ 37 Mb	~ 100 Mb	DNA methylation agrees with XCI status	[190]
46,X,der(X) t(X;5)(q22.1;q31.1)	RT-PCR of somatic cell hybrids	20	45%	40%	15%	~ 42 Mb	~ 69 Mb	DNA methylation agrees with XCI status	[192]

Table 1.3: Frequency of sex chromosomes in human triploids.

Study details		69, XXX	69, XXY	69, XYY
Niebuhr 1974 [244]	spontaneous abortions and liveborn	92/255= 36%	153/255= 60%	10/255= 3%
Jacobs <i>et al.</i> 1979 [226]	spontaneous abortions	9/36= 25%	27/36= 75%	0/36= 0%
Jacobs <i>et al.</i> 1982 [210]	spontaneous abortions	31/96= 31%	64/96= 67%	1/96= 1%
Ohno <i>et al.</i> 1991 [245]	spontaneous abortions	3/8= 38%	5/8= 63%	0/8= 0%
Neuber <i>et al.</i> 1993 [246]	spontaneous abortions	26/63= 41%	36/63= 57%	1/63= 2%
Zaragoza <i>et al.</i> 2000 [204]	spontaneous abortions	8/27= 30%	17/27= 63%	2/27= 7%
Baumer <i>et al.</i> 2000 [223]	spontaneous abortions and liveborn	14/25= 56%	11/25= 44%	0/25= 0%
McFadden and Langlois 2000 [222]	spontaneous and elective abortions	20/40= 50%	19/40= 48%	1/40= 3%
Daniel <i>et al.</i> 2001 [247]	spontaneous abortions and liveborn	8/17= 47%	9/17= 53%	0/17= 0%
McFadden and Robinson 2006 [206]	spontaneous abortions	9/23= 39%	14/23= 61%	0/23= 0%
McWeeney <i>et al.</i> 2009 [224]	preimplantation embryos and clinically recognized pregnancies	264/549= 48%	248/549= 45%	37/549= 7%
	Total avg ->	484/1139= 42%	603/1139= 53%	52/11395= 5%

Table 1.4: XCI patterns observed in 69, XXX triploids.

study	sample age (ID)	sample size			cell type examined	technique to determine XCI status	XCI findings
		diandric	digynic	unknown parent of origin			
Fryns <i>et al.</i> 1977 [229]	Liveborn (2 months survival)	0	0	1	lymphocytes	radioautography	69, X _a X _i X _i
Jacobs <i>et al.</i> 1979 [226]	141 and 198 days	2	0	0	Fetus, cord	BrdU labeling	Mainly 69, X _a X _a X _i
Jacobs <i>et al.</i> 1979 [226]	65-133 days	5	2	0	Amnion, chorion, villi	BrdU labeling	Mainly 69, X _a X _a X _i but more X _i s than fetus/cord
Maraschio <i>et al.</i> 1984 [230]	Liveborn (45 days)	1	0	0	Lymphocytes , fibroblasts	BrdU labeling	Mainly 69, X _a X _i X _i but degree differs across tissues
Gartler <i>et al.</i> 2006 [225]	14-25 fetal weeks (GM04376, GM07744, GM10013 and GM10006) Liveborn (1-2 days survival: GM04939 and 75-29)	0	0	6	Embryonic fetal fibroblasts (GM04376, GM07744, GM10013 and GM10006) Fibroblasts (GM04939 and 75-29)	XIST RNA FISH, DNA methylation	Mainly 69, X _a X _a X _i , increase in X _a s with passage

Table 1.5: XCI patterns observed in 69, XXY triploids.

study	sample age (ID)	sample size			cell type examined	technique to determine XCI status	XCI findings
		diandric	digynic	unknown parent of origin			
Migeon <i>et al.</i> 1979 [227]	28 week abortus (GM1322)	0	0	1	Skin fibroblasts	G6PD variants	69, X _a X _a Y
Jacobs <i>et al.</i> 1979 [226]	54-187 days	5	5	0	Fetus, cord	BrdU labeling	Mainly 69, X _a X _a Y
Jacobs <i>et al.</i> 1979 [226]	51-220 days	15	6	0	Amnion, chorion, villi	BrdU labeling	Mainly 69, X _a X _a Y but more X _i s than fetus/cord
Willard and Berg 1980 [248]	Liveborn (1 hour survival: GM1672), 28 week abortus (GM1322), SA (FB530)	0	0	3	Biopsy (GM1672), skin fibroblast (GM1322), fetal membranes (FB530)	Barr bodies	69, X _a X _a Y
Yu <i>et al.</i> 1983 [233]	Liveborn (25 hour survival)	0	1	0	Buccal mucosal Cultured lymphocytes, cartilage, skin and gonadal cells	Barr bodies BrdU labeling	69, X _a X _a Y but degree differs across tissues
Vogel <i>et al.</i> 1983 [228]	Liveborn	0	0	1	fibroblast	BrdU labeling	Mainly 69, X _a X _i Y, increase in X _a s with passage

2 Chromosome-wide DNA methylation trends on the human X

2.1 Introduction

The facultative heterochromatin of the X chromosome provides an excellent system to study epigenetic silencing, and one of the first epigenetic marks proposed to play a role in XCI was DNA methylation [42]. On the autosomes, the DNA methylation of CpG island promoters is known to be associated with silencing [61, 62]. On the X chromosome, the CpG island promoters of genes subject to XCI have previously been determined to be unmethylated in males and partially methylated in females, while genes which escape from XCI are unmethylated in both males and females [63-65]. The gold standard technique to determine XCI status would be one based directly on expression analysis. However, because females are mosaic for XCI, it would be necessary to have a female with totally skewed XCI who is also a heterozygote for a SNP (single-nucleotide polymorphism) at the gene being examined. In addition, a gene must be expressed in the tissue being examined or the XCI cannot be determined. DNA methylation analysis has the advantage that a female need not be skewed for XCI, need not be heterozygous for SNPs in that gene, and XCI can be determined even in tissues where the gene is not expressed.

X-linked DNA methylation has previously been used to establish XCI status [69] in neutrophils, however, a comparison of X-linked DNA methylation patterns across tissues has not been performed. The examination of XCI status in brain tissues is of particular interest given the over-representation of X-linked mental retardation [249-251] in males. While there appears to be a clear relationship between X-linked CpG island promoter DNA methylation and XCI status, the relationship between gene-body DNA methylation and transcription is less clear. Across the genome, highly expressed genes show higher gene-body DNA methylation than genes with low expression, and for the X chromosome, previous reports have suggested that the gene-bodies of the X_a are hypermethylated compared to those of the X_i [78, 79, 81].

In this chapter, the Illumina Infinium HumanMethylation27 array was used to analyze DNA methylation at 777 X-linked promoters in human blood and fetal somatic tissues (muscle, brain, spinal cord and kidney). DNA methylation was used to predict an XCI status for each gene and the predicted XCI status was compared between tissues. MeDIP (Methylated DNA ImmunoPrecipitation) and hybridization to a NimbleGen 2.1M array allowed for chromosome-wide DNA methylation analysis in human blood. Establishing the pattern of normal X-linked DNA methylation was critical to allow for further analysis of X-linked DNA methylation in other tissues (placenta: see chapters 3 and 4) and in chromosomally abnormal situations (X;autosome translocation: see chapter 5, triploids: see chapter 6).

2.2 Methods

2.2.1 Sample collection and DNA extraction

Collection of samples was approved by the ethics committees of the University of British Columbia and the Children's and Women's Health Centre of British Columbia with collection and preparation performed by the Robinson lab. Whole blood samples (female n=59 and male n=36) were collected from anonymous donors (ethics approval number H08-02773) and PBMCs (peripheral blood mononuclear cell) isolated using BD Cell Preparation Tubes as per manufacturer's instructions. DNA was extracted using Qiagen AllPrep DNA/RNA mini kits as per standard conditions. Fetal tissues (muscle: female n=6 and male n=4, spinal cord: female n=2 and male n=1, brain: female n=4 and male n=4, kidney: female n=6 and male n=5) were chromosomally normal and collected from biopsied abortuses from anonymous pregnancies terminated for medical reasons (ethics approval number H06-70085). Genomic DNA was extracted using a standard salting-out method as outline in Papageorgiou *et al.* [252].

2.2.2 Illumina Infinium HumanMethylation27 array

Bisulfite conversions and array processing were performed by the Robinson and Kobor labs. Genomic DNA was bisulfite modified with the EZ DNA methylation Kit (Zymo Research) as per the manufacturer's instructions and 180-200 ng of bisulfite DNA was then amplified, fragmented and hybridized to Illumina Infinium HumanMethylation27 beadarray chips (Illumina, Inc) using Illumina supplied reagents and conditions. The arrays were scanned on the Illumina iScan system and imported into GenomeStudio for further analysis (2010.2). Results were subjected to a background normalization using BeadStudio (versions 3.1.3.0 Illumina, Inc) and probes with p-values greater than 0.05 were removed. Quantile normalization was performed in R 2.11.0 using the limma package [253]. Although beta-values are compressed when less than 0.2 and greater than 0.8 and both of these ranges show high heteroscedasticity, since we were interested in large DNA methylation differences between males and females, for the purposes of this paper, beta-value was considered equivalent to percent DNA methylation [254].

2.2.3 CpG density definitions

We used CpG density classifications based on those used by Weber *et al.* [45] to define three CpG densities; HC, IC and LC. The program CpGIE [255] was used to define and locate HC and IC islands on the X chromosome and chromosomes 20, 21 and 22. HCs had a GC content greater than 55%, an $\text{Observed}_{\text{CpG}}/\text{Expected}_{\text{CpG}}$ greater than 0.75 and were at least 500 bp (base pairs) in length. ICs had a GC content greater than 50%, an $\text{Observed}_{\text{CpG}}/\text{Expected}_{\text{CpG}}$ greater than 0.48 and were at least 200 bp in length. Those ICs which overlapped with an HC were excluded from the IC category but their HC component remained in the HC category. Additionally, all ICs which overlapped with repetitive elements, as defined by RepeatMasker

[256, 257], were not included in the IC category. LCs were all those regions which were not HC or IC.

2.2.4 Illumina Infinium HumanMethylation27 array composition and probes removed from analysis

The Illumina Infinium HumanMethylation27 array is a promoter array with all probes located in close proximity to an annotated TSS. Approximately 45% of X-linked promoters are represented on the array and those promoters which overlap CpG islands (HC and IC) represent nearly three quarters of the probes on the array, despite the fact that only 5% of CpGs on the X chromosome are located in islands. The basic local alignment search tool (BLAST) program [258] from NCBI (National Center for Biotechnology Information) was used to determine if a probe sequence mapped to a single unique location in the genome or to multiple sites. Due to the large number of genes on the X chromosome which have homologs on the Y chromosome, probes which mapped to the Y chromosome as well the X chromosome were not removed from the analysis. 153 X-linked and 134 autosomal (chromosomes 20,21 and 22) probes were removed from analysis due to mapping to more than one location in the genome. 137 X-linked probes located in the promoters of the cancer-testis (CT) family of genes were removed from analysis since they are known to be methylated in all tissues except testis regardless of CpG density [259]. To determine if probes were located in repetitive elements, probe locations were compared against the location of known repetitive elements from RepeatMasker for UCSC (University of California – Santa Cruz) [256, 257] which resulted in the removal of 88 X-linked and 220 autosomal probes.

2.2.5 Statistical analysis

Statistical analysis of the Illumina Infinium HumanMethylation27 array was performed using the Mann-Whitney test as calculated by Graphpad Prism. Statistical analysis of MeDIP data was calculated in R [260] using the Kolmogorov-Smirnov Test. Intrasex variation was calculated for each sex by comparing all combinations of samples using the Kolmogorov-Smirnov test. Due to the large sample size, p-values less than 0.0001 were considered significant and p-values less than 1.0×10^{-10} highly significant. In order for the results to be considered significant and to ensure the difference between the average male and average female DNA methylation was larger than any differences within the sexes, we required that the p-value resulting from the comparison of the average male and average female DNA methylation was smaller than the intrasex p-values.

2.2.6 Decision tree to predict XCI status

Probes were predicted to escape XCI when the male average and female average were unmethylated (<15% methylated) and when males and females showed a similar range of DNA methylation (either the range of male and female DNA methylation overlapped or if the ranges did not overlap the difference between the male and female average was less than 10%).

Probes were predicted to be subject to XCI when males and females showed a different range of DNA methylation where the difference between the male average and female average was greater than 10%. Probes were predicted to variably escape from XCI when the difference between the male and female average was greater than 10% and there was also an overlap in the range of male and female DNA methylation. When the male and female averages were greater than 15% and/or the difference between the male and female average was less than 10% probes were defined as unclassifiable. Table 2.1 outlines the decision tree.

2.2.7 RNA extraction and Q-PCR

RNA from four somatic cell hybrids containing a human X_i (t75-2maz34-4a, t48-1a-1Daz4A, t86-B1maz1b-3a and t11-4Aaz5), two somatic cell hybrids containing a human X_a (AHA-11aB1 and t60-12) and a control female cell line (GM7350) was extracted using Trizol (Invitrogen) as per the manufacturer's protocol and 5 μ g converted to complementary deoxyribonucleic acid (cDNA) via a standard RT-PCR reaction using M-MLV (Invitrogen) at 42°C for two hours followed by a five minute incubation at 95°C. Real-time quantitative PCR (Q-PCR) was performed using the StepOnePlus™ Real-Time PCR System (Applied Biosystems, Darmstadt, Germany) on each sample in triplicate with the following conditions: 95°C (5 mins), [95°C (30s), 60°C (30s), assorted annealing temperature (30s)] for 40 cycles then melting curve analysis [95°C (15s), 60°C (60s) then fluorescence was measured every 0.3°C per until 95°C]. Primer sequences and annealing temperatures are listed in Table 2.1. The average of three triplicate Ct values were corrected based on the average efficiency for each assay, as calculated by LinReg [261, 262] and delta Ct values calculated for *TSR2* and *ZRSR2* compared to *ZFX*. The negative and positive error was calculated based on the sum of the standard deviation for the test (*TSR2* or *ZRSR2*) and the control (*ZFX*) assay for each sample (in triplicate). All assays were found to not amplify mouse gDNA (data not shown).

2.2.8 MeDIP and whole genome amplification

MeDIP of male (n=3) and female (n=3) blood was performed as outlined in Vucic *et al.* (2009). Briefly, 3 reactions of 1 μ g of genomic DNA were sonicated then 200 ng of input removed. The remaining 800 ng of DNA was denatured (95°C, 10 minutes) then 5 μ g of anti-5'-methylcytosine mouse mAb (CalBiochem) added before incubating at 4°C for two hours. 30 μ L of Dynabeads M-280 Sheep anti-Mouse IgG (DynaL Biotech, Invitrogen) were then added followed by a two

hour incubation at 4°C. Two rounds of washing was performed to remove the Dynabeads then 100 µg of proteinase K added and left overnight at 50°C. A phenol:chloroform clean-up was performed the next day and DNA resuspended in 10 µL H₂O. Whole genome amplification was performed using the GenomePlex Complete Whole Genome Amplification Kit (Sigma) according to the manufacturer's instructions.

2.2.9 NimbleGen array processing and analysis

Three reactions using 1 µg of whole genome amplified DNA for each sample were labeled using Cy3-9mer primers for input and Cy5-9mer primers (TriLink Biotechnologies, Inc.) for IP (immunoprecipitation). Labeling was performed as outlined in the NimbleGen Arrays User's Guide: ChIP-chip Analysis v3.0 (Roche NimbleGen, Inc) then samples sent to the Fred Hutchinson Cancer Research Center (Seattle, WA, USA) for hybridization to a Human ChIP-chip 2.1M Whole-Genome Tiling, array number 10 (Roche NimbleGen, Inc). Files of the scanned arrays were processed according to the NimbleGen Arrays User's Guide: DNA methylation Analysis v5.0 and the resulting ratio files subjected to Bayesian Tool for DNA methylation Analysis (BATMAN) [263] to correct the effect of CpG density of MeDIP efficiency. The average standard deviation of the three samples was 0.05 in males and females. To ensure that samples were more similar within a sex than between, the six blood samples were combined in all 18 possible combinations of three and the average standard deviation (0.06) of all 18 combinations was always greater than that observed within each sex. Galaxy [264, 265] was then used to calculate the frequency of probes in the various genomic elements examined.

2.2.10 Expression data

Expression ratios on log base 2 scale from whole blood were downloaded from <http://symatlas.gnf.org> [266]. Genes were divided by chromosome and ranked from lowest to highest expression. The top and bottom 20% of genes from each chromosome were used to represent those genes with the lowest and highest expression levels in blood.

2.3 Results

2.3.1 X-linked promoters show differences in DNA methylation dependent on sex and CpG density

To determine how X-linked promoter DNA methylation differed between the sexes we applied DNA from 36 male and 59 female bloods to the Illumina Infinium HumanMethylation27 array containing 1085 X-linked probes. A total of 308 X-linked probes were removed from analysis as they were located in repetitive elements, mapped to more than one location in the genome and/or were located in the promoters of the CT family of genes (summarized in Figure 2.1). In order to detect the large DNA methylation differences between males and females previously

reported at most X-linked promoters, we created three broad DNA methylation classes: unmethylated (0-0.15 beta value), intermediate (0.15-0.60 beta value) and methylated (0.60-1.00 beta value) and then separated the remaining 777 X-linked probe results into one of the three DNA methylation classes. The majority (67%) of the X-linked probes in male blood were shown to be unmethylated whereas the majority of (66%) of probes in female blood had intermediate methylation. As CpG density is known to influence DNA methylation we subdivided probes based on their location within HC and IC islands or LCs. X-linked promoter probes in HC and IC islands were generally those which were unmethylated in male blood and intermediately methylated in female blood, whereas X-linked promoter probes in LCs were usually methylated in both sexes (Figure 2.2A).

To ensure that the observed X-linked DNA methylation differences between males and females were not simply due to differences in overall DNA methylation between the sexes, autosomal DNA methylation levels were compared between the same male and female blood samples. Male and female probes (1843 probes located on chromosomes 20, 21 and 22) were compared after removal of 307 autosomal probes that were located in repetitive elements and/or mapped to multiple locations in the genome. The majority (98%) of all autosomal probes showed the same DNA methylation level in males and females regardless of CpG density. Furthermore, those probes for which males and females had different DNA methylation classes (unmethylated, intermediate or methylated) showed only a 2% difference in DNA methylation and were not significantly different in their DNA methylation, making it unlikely that this difference was biologically functional. We therefore conclude that the difference between male and female DNA methylation is mostly attributable to the X chromosome and occurred primarily at X-linked promoter probes located in HC and IC islands (Figure 2.2B). Given that 10% of X-linked probes in HC and IC islands were unmethylated in females, and that unmethylated X-linked promoters have previously been found at genes which escape from XCI [67] we wanted to examine whether the detection of unmethylated probes in HC and IC islands in males and females reflected escape from XCI as previously proposed by Yasukochi *et al.* [69][69].

2.3.2 Promoter DNA methylation effectively predicts XCI status

Before we could evaluate how effective X-linked promoter DNA methylation might be at predicting the XCI status of a gene, we had to establish a consistent method of translating male and female DNA methylation results from multiple probes into a single genic XCI status. We developed a decision tree examining the average male and average female DNA methylation levels, the difference between these averages and the range of DNA methylation observed in males and females in order to assign the XCI status of each probe (Figure 2.1). The majority of probes in genes with HC and IC islands (77%) predicted the gene was subject to XCI, 11%

predicted escape from XCI, 3% variable escape and 8% were unclassifiable. The 560 probes in HC and IC islands were found in 343 X-linked genes (145 genes were represented by only one probe), therefore probes from the 198 genes that contained more than one probe in an HC or IC island were combined to create a single predicted XCI status for each gene. For 90% of genes, all probes (if multiple probes were present) predicted the same XCI status. Of the remaining 10% of genes, over half had one probe where an XCI status was predicted with the other probe being unclassifiable. These genes were therefore given the XCI status of the probe which had predicted an XCI status (subject, escape or variable escape). In only five genes out of the 343 X-linked genes was there a conflict in which one probe predicted an XCI status of escape and the other subject to XCI was found (Table 2.2). Interestingly, all five of these were found in genes that had previously been reported to escape, or variably escape from XCI [106]. We predicted that the majority (81%) of genes with probes in HC or IC islands were subject to XCI, 10% escaped XCI, 2% variably escaped XCI and 5% of genes remained unclassifiable (Figure 2.3A). Of the 19 unclassifiable genes, 13 were methylated in both males and females. Overall we were able to use DNA methylation to predict an XCI status for 95% of examined X-linked genes with probes in HC or IC islands and could therefore compare these predictions to the XCI status of the same genes previously determined by expression.

To examine whether X-linked promoter DNA methylation was effective at predicting XCI status, we analyzed genes at which the XCI status had previously been established and determined the degree to which our predicted XCI status agreed. We compared our predicted XCI status in blood with the XCI status derived from reported studies of somatic cell hybrids by Carrel and Willard [106]. Since the bulk of X-linked genes are subject to XCI we first examined the genes with probes in HC or IC islands which we predicted to be subject to XCI. After the removal of genes not examined by Carrel and Willard [106], 83% (n=192) of the genes predicted by X-linked promoter DNA methylation to be subject to XCI were also found by Carrel and Willard [106] to be subject to XCI. Given our interest in using X-linked promoter DNA methylation to predict escape from XCI we also examined those genes with probes in HC or IC islands for which our DNA methylation data had predicted escape from XCI. Here we found that 72% of genes predicted by X-linked promoter DNA methylation to escape from XCI were also shown by Carrel and Willard [106] to escape from XCI (Figure 2.3B). When DNA methylation was used to predict an XCI status using only one probe, the same degree of discordance (18% discordant) with the XCI statuses found by Carrel and Willard [106] was found compared to genes represented by more than more probe (19% discordant). To further address the ability of promoter DNA methylation to predict XCI status, the expression patterns of two genes (*TSR2* and *ZRSR2*) not examined by Carrel and Willard [106] were examined by Q-PCR in somatic cell

hybrids (four hybrids containing a human X_i and two containing a human X_a) as well as a control female cell line. Based on promoter DNA methylation, *TSR2* was predicted to be subject to XCI and this was confirmed, as none of the X_i hybrids showed expression comparable to the X_a hybrids or the female cell line. *ZRSR2* was predicted to escape from XCI and this, too, was confirmed with all X_i hybrids showing expression at least as high as the X_a hybrids and the female cell line (Figure 2.3C and D).

This validation, along with the high degree of agreement between previously determined XCI status and our prediction using X-linked promoter DNA methylation, led us to conclude that the DNA methylation of probes in HC and IC islands X-linked promoters can be used to predict XCI status and therefore we can propose an XCI status for 62 genes (see Table 2.3). A few of these genes have been described in other studies and in these cases our prediction of XCI status is in agreement with previous reports [69, 267, 268]. While our results suggest that DNA methylation is an effective predictor of XCI status, 59 genes were shown by Carrel and Willard [106] to have a different XCI in somatic cell hybrids than was predicted by our analysis of DNA methylation in blood. Tissue-specific escape from XCI has been reported in mouse [269] and therefore we wished to investigate the extent to which tissue-specific differences could be a contributor to the 15% discordance we observed between the XCI status in somatic cell hybrids and blood [106].

2.3.3 Tissue-specific XCI is observed at 12% of genes

We extended our Illumina Infinium HumanMethylation27 array analysis to fetal tissues (muscle, kidney, brain and spinal cord) to determine if all tissues showed the same male and female DNA methylation, and therefore the same predicted XCI status. We first confirmed that, as with blood, there was a sex-specific DNA methylation difference that was limited to the X chromosome and not the autosomes (Figure 2.4). We then examined fetal muscle and fetal kidney and combined fetal brain and fetal spinal cord into one fetal “neural” tissue category. We used the same process of predicting XCI status as in blood and again demonstrated that although the level of DNA methylation was significantly different between tissues (p -value <0.0001), the majority of X-linked CpG island genes showed a pattern of DNA methylation consistent with being subject to XCI (unmethylated males and intermediate females) regardless of the tissue examined (blood=81%, fetal muscle=74%, fetal neural=66%, fetal kidney=73%). Interestingly, a larger proportion of X-linked genes showed a pattern of DNA methylation that we considered predictive of escape from XCI in fetal tissues (muscle=15%, neural=17%, kidney=15%) compared to blood (10%). The fetal tissues had a considerably smaller sample size than blood which led us to compare the predicted XCI in all tissues to determine how often the same XCI status was predicted in all tissues (Figure 2.5A).

We compared the predicted XCI status across all tissues and found that the majority (78%) of X-linked genes showed the same predicted XCI status in all tissues examined. An additional 6% of genes showed the same predicted XCI status in all but one tissue (which was designated unclassifiable). However, at 12% of X-linked genes, promoter DNA methylation resulted in a different predicted XCI status in different tissues (Figure 2.5B) and we designated these genes as showing tissue-specific XCI. Of the genes which showed tissue-specific XCI, nearly half (48%) showed more escape in the fetal tissues compared to blood. Table 2.4 lists genes which displayed tissue-specific XCI; the locations of these genes are shown in Figure 2.6 along with the location of genes that showed consistent XCI patterns. The distribution of these genes is influenced by the choice of probes on the array, and notably no pseudoautosomal probes were included. Our finding that X-linked promoter DNA methylation differs across 12% of genes examined suggests tissue-specific XCI in these genes.

To investigate if tissue-specific XCI was consistent between females, we examined six different females each with at least two different fetal tissues. We compared the predicted XCI status in fetal tissues (muscle, neural tissue and kidney) from four females, fetal muscle and fetal kidney from one female, and fetal neural and fetal kidney from another female. We used the individual female's DNA methylation value along with the average male DNA methylation in the same tissue to predict XCI status. In each female examined, 84% to 86% (see Table 2.5) of the total X-linked genes examined were predicted to escape or be subject to XCI across all tissues, while 8% to 14% of genes were unclassifiable and 1% to 7% of genes showed tissue-specific XCI. In females with multiple tissues, fetal muscle showed the fewest genes with tissue-specific escape while fetal neural tissue showed the most. We found that when escape from XCI was predicted by X-linked promoter DNA methylation in one tissue, it was generally predicted in all tissues (listed in Table 2.6). Overall, DNA methylation based evidence for tissue-specific XCI was found in all females examined, with the highest degree of tissue-specific escape observed in fetal neural tissue but with a great deal of variability between females.

2.3.4 Non-island DNA methylation is a poor predictor of XCI status

Having established that the X-linked promoter DNA methylation of probes in HC and IC islands was highly predictive of XCI status, we were interested if the same methodology could be applied to LC probes (those not located in HC and IC islands). It should be noted that the majority of probes on the Illumina Infinium HumanMethylation27 array are located in CpG islands (65%) associated with promoters. Some X-linked LC promoters have been reported to exhibit DNA methylation that correlates with gene silencing on the X_i (such as *TIMP1* [68], *CHM* [270], and *OTC* [271]). The same decision tree Figure 2.1 used to predict the XCI status of probes in HC and IC islands was applied to LC probes to evaluate what proportion of LC probes

showed a DNA methylation status which could predict an XCI status. Approximately one quarter (27%) of all LC probes examined were located in promoters (+/- 1kb (kilo base pairs) around the TSS) which also included an HC or IC island while the remaining LC probes were located in promoters which lacked an HC or IC island. LC probes were generally unclassifiable (82%) due to high DNA methylation regardless of whether a CpG island was present within the promoter region. The remaining probes showed DNA methylation patterns classifiable as escape (4%), variable escape (4%) or subject to XCI (genes with CpG islands: 21%, genes without CpG islands: 7%). Those LC probes with a DNA methylation status which predicted an XCI status of subject, variable escape or escape, along with any HC or IC probes in the same gene, are listed in Table 2.7. We compared the XCI status predicted using DNA methylation to that determined by Carrel and Willard [106] and found that approximately 40% of LC probes predicted the same XCI status as Carrel and Willard [106] regardless of whether the LC probe was in the promoter of a gene with a CpG island or not. Given the low concordance between the predicted XCI status based on LC probe DNA methylation and that previously determined by Carrel and Willard [106], we conclude that LC probes are not usually reliable as a predictor of XCI status.

2.3.5 X-linked HC and IC promoters show the strongest sex-specific DNA methylation difference

The data we analysed from the Illumina Infinium HumanMethylation27 array examined only approximately 45% of X-linked promoters and did not examine any non-promoter elements such as the intragenic and intergenic regions of the chromosome. To expand the study of X-linked DNA methylation beyond promoters, MeDIP was performed on DNA isolated from male (n=3) and female (n=3) blood followed by hybridization to a NimbleGen 2.1M array to analyse chromosome-wide DNA methylation of chromosome 20, 21 and 22 along with the X chromosome. To correct for the effect of CpG density on MeDIP efficiency, BATMAN [263] was used to convert the ratio of IP:IN into a DNA methylation value from zero to one. BATMAN was performed on all samples and the resulting scores averaged to create one average male score and one average female score, in subsequent analyses only male versus female differences that were greater than intrasex differences were considered for statistical significance. DNA methylation histograms were compiled to assess the distribution of DNA methylation on different DNA elements of interest.

To determine if the X-linked sex-specific DNA methylation difference found using the Illumina Infinium HumanMethylation27 array could also be observed via MeDIP, the first elements we examined were X-linked promoters. Promoters were defined as the probes within 1kb up and downstream of all TSS, therefore the presence of an HC island in a promoter, or the presence

of an IC but not an HC, resulted in the classification of HC or IC promoter respectively. LC promoters were those promoters which had neither an HC nor IC island in the region 1kb upstream and downstream of the TSS. X-linked HC promoters showed a higher frequency of unmethylated probes in the male than the female with a significantly different ($D=0.12$, $p\text{-value}<2.2 \text{ E-}16$) distribution between the sexes (Figure 2.7A). IC promoters on the X chromosome showed a significantly different ($D=0.07$, $p\text{-value}=2.5 \text{ E-}10$) distribution between the sexes and were slightly more unmethylated on the male X chromosome than on the female X chromosomes. Additionally, X-linked IC promoters also had a higher percentage of both male and female probes being intermediate or fully methylated than was observed in HC promoters (Figure 2.8A). On the autosomes, neither HC nor IC promoters were significantly different between males or females, however, HC promoters were mostly unmethylated while IC promoters also showed intermediate methylation. X-linked LC promoters were mostly methylated and were not significantly different between males and females while autosomal LC promoters were slightly less methylated (Figure 2.7B). By examining all known X-linked promoters we were able to show that sex-specific DNA methylation differences were highly significant at X-linked HC promoters, slightly significant at X-linked IC promoters but not significant at X-linked LC promoters or on the autosomes.

Our definition of promoter elements comprised only approximately 2% of base pairs on the X chromosome; therefore, determining the DNA methylation status at non-promoter elements was of critical importance if an overview of chromosome-wide DNA methylation was to be established. Intragenic and intergenic regions showed similar DNA methylation in males and females; however, on the X chromosome these regions were bimodally distributed whereas on the autosomes they were not (Figure 2.7C and D). Across the genome there are CpG islands not currently associated with known genes. Males and females displayed significantly different DNA methylation at HC islands not associated with a known TSS on the X chromosome ($D=0.09$, $p\text{-value}=2.4 \text{ E-}9$) but not on the autosomes nor at X-linked or autosomal IC islands not associated with a known TSS. The IC islands were more methylated in both sexes than IC promoters on either the X chromosome or the autosomes (Figure 2.8B and C). Having compared elements across the entire X chromosome we confirmed that although HC promoters compose a small fraction of the X chromosome they are the element which showed the strongest degree of X-linked sex-specific DNA methylation.

2.3.6 X-linked genes with higher expression show higher gene-body DNA methylation

It has previously been shown that the intragenic regions of highly expressed genes are more methylated than those of genes with low expression [81] and on the X chromosome gene-

bodies of the X_a have been found to be more methylated than on the X_i [79]. We therefore used published expression data [266] to separate genes with high expression levels (top ranking 20%) from those with low expression (bottom 20%) to allow for a male:female comparison of gene-body DNA methylation levels relative to expression levels. No significant differences between the distribution of male and female DNA methylation were found at exons or introns on either the X chromosome or the autosomes. However, X-linked introns of genes with high expression were more methylated than those gene with low in both males and females (Figure 2.9A and B). Interestingly, X-linked exons did not show this difference based on expression. Overall, the division of genes based on expression did not demonstrate a significant difference in the distribution of male:female DNA methylation although X-linked intronic DNA methylation was greater in genes with high expression compared to genes with low expression.

2.4 Discussion

The presence of DNA methylation at X-linked CpG island promoters on the X_i is classically associated with genes subject to XCI [272, 273]. We found that in all tissues examined (blood, fetal muscle, fetal kidney, fetal brain and fetal spinal cord), the majority of X-linked promoter probes in HC and IC islands were unmethylated in males and intermediately methylated in females, which is the pattern of DNA methylation typically associated with genes subject to XCI. In support of this sex difference being reflective of XCI, nearly all autosomal probes (over 95%) showed the same DNA methylation status in males and females, regardless of CpG density. Genes which escape from XCI have previously been found to be unmethylated in both males and females [67] and this unique property has been used to propose novel genes which escape from XCI [69] in neutrophils. We extended the search for genes which escape from XCI to blood, fetal muscle, fetal kidney, and fetal neural tissue using the DNA methylation pattern for genes with probes in HC and IC islands. We found a high degree of concordance (81%) with the XCI status previously determined by Carrel and Willard [106], however, for 19% of genes there was discordance between our DNA methylation-based prediction in blood and results from expression in somatic cell hybrids.

Previous studies which have examined XCI status have typically used either somatic cell hybrids or females with clonal XCI who are heterozygous for known SNPs [106]. Similar to our results, previous comparisons between expression in hybrids and in female tissues have shown discordancies [106, 274]. We propose several different reasons for the differences between the XCI status we predicted using DNA methylation and that of Carrel and Willard [106]. First, DNA methylation may not always be an accurate predictor of XCI status. This might occur in regions where other epigenetic marks, such as histone modifications, are more important to maintaining XCI. A second possible explanation is that due to the proposed decrease in stability of XCI of

somatic cell hybrids [274, 275] genes which are typically subject to XCI in blood now escape from XCI in somatic cell hybrids. If the differences in XCI were caused by a decrease in stability of XCI in somatic cell hybrids then any conflicts in XCI status should involve a higher degree of escape from XCI in the somatic cell hybrids. 15% of genes examined showed more escape from XCI in the somatic cell hybrids than in blood supporting this hypothesis, however, 4% of genes showed more escape in blood than the somatic cell hybrids. These conflicts cannot be explained by a decrease in the stability of XCI in somatic cell hybrids, suggesting that hybrid instability is not the full explanation.

A third possibility is that somatic cells hybrids and blood actually have a different XCI status at a subset of genes. We attempted a direct comparison of DNA methylation in hybrids (data not shown) with expression status for individual genes; however, we observed considerable variability of DNA methylation between hybrids, even in X_a hybrids (data not shown) and were thus not able to compare DNA methylation to expression in the hybrids. We therefore examined male and female DNA methylation levels in different human tissues to determine if tissue-specific DNA methylation changes were frequent. While most genes had the same predicted XCI status in all tissues examined, we detected potential tissue-specific XCI in 12% of genes, the majority of which reflected genes being subject to XCI in blood while at least one other tissue was not subject to XCI. We also found that over 50% of genes showing tissue-specific XCI were found within 1Mb (mega base pairs) of each other suggesting a possible regional effect causing tissue-specific XCI. We caution that when examining X-linked genes, the XCI status should always be confirmed in the tissue of interest. The degree of predicted tissue-specific XCI differed between the six examined females and between tissues, with neural tissue showing the highest degree of predicted tissue-specific escape from XCI. Studies examining expression amongst all X-linked genes have consistently shown brain to have one of the highest X:autosome gene expression ratios, regardless of the technique being used [276]. The X chromosome contains an over-representation of genes expressed in the brain [277, 278] and many X-linked genes are known to play a role in X-linked mental retardation (reviewed in [251]) which is significantly more common in males than in females [249, 250]. Expression of genes from the X_i when the Y homolog is no longer functional could lead to a dosage difference between males and females, and might contribute to sex-specific differences in disease susceptibility.

For autosomes, tissue-specific DNA methylation differences in CpG islands have previously been detected across a number of tissues [70-73, 279] and it has been proposed that the majority of tissue-specific differentially methylated regions are located in the regions surrounding CpG islands known as CpG island shores [73]. We found that the majority of

probes which showed tissue-specific XCI (83%) were located in HC islands rather than shores. The criteria we used to detect sex-differences on the X chromosome were designed to identify large changes in DNA methylation associated with the XCI status of the gene which may explain why the tissue-specific DNA methylation we observed on the X chromosome was mostly located in the CpG islands and not in the shores as was previously reported on the autosomes [73]. Our analysis of X-linked HC non-promoters (HC islands not associated with a known promoter) revealed a similar hypomethylation in male blood compared to female blood. The presence of a sex-specific DNA methylation difference is evidence that these HC islands may be the promoters of unannotated X-linked genes that are subject to XCI. This is in agreement with a previous report in which the majority of genome-wide orphan CpG islands were predicted to be associated with the promoters of unknown genes based on histone modifications and the presence of RNA Pol II [76]. The X-linked IC non-promoter we examined lacked a significant sex-specific DNA methylation difference suggesting that it is less likely that these CpG islands are associated with unknown genes. To confirm that the X-linked non-promoters islands we were predicting to be promoters were in fact not enhancers, we examined the histone modifications typically associated with enhancers [280] and did not find any significant enrichment (data not shown).

Across the genome, the most widely expressed genes tend to have a promoter CpG island as do a smaller subset of tissue-specific genes [46]. On the X chromosome, some genes, notably androgen receptor which has been widely used to examine XCI skewing, also have tissue-specific expression [266] yet show consistent DNA methylation (males: unmethylated, females: ~50% methylated) even in tissues in which they are not expressed [281]. Consistent with this observation, data from the Illumina Infinium HumanMethylation27 array showed that female X-linked promoters had no differences in DNA methylation between genes with high and low expression at any CpG density while males showed a slight significance at X-linked IC promoters. Chromosome-wide DNA methylation analysis showed that the HC promoters of highly expressed X-linked genes maintained a significant difference between males and females (data not shown), where males were more hypomethylated than females. On both the X chromosome and the autosomes, all other promoters showed no significant difference between the distribution of males and female DNA methylation at genes with high and low expression.

The association between promoter DNA methylation and transcriptional silencing is well established [46, 282], [283]; however, the interaction between gene-body DNA methylation and transcription, as well as the DNA methylation status of intergenic regions, is less clear. In general, the distribution of DNA methylation in intragenic and intergenic regions of the X

chromosome is different from the autosomes, likely reflecting the unique sequence composition of the sex chromosomes. This difference is less apparent at exons where the distribution of DNA methylation on the X chromosome is more similar to the autosomes. When X-linked introns are examined, the DNA methylation of the top 20% of expressing genes differs from the bottom 20% of genes, whereas on the autosomes, expression does not greatly affect the distribution of DNA methylation. The shift of DNA methylation of highly expressed X-linked introns yields a distribution of DNA methylation very similar to that found at all X-linked exons. Although we do not observe a significant difference between the distributions of male and female DNA methylation, we do see that the introns of X-linked genes with high expression are more methylated than the introns of genes with low expression. A role for transcription in gene-body DNA methylation is supported by a recent genome-wide study which showed that early replicating genes have more gene-body DNA methylation than late replicating genes [81] while another study showed DNA methylation of the gene-body was more likely to be found in highly expressed genes [85]. Differences in autosomal exon and intron DNA methylation have previously been found, with first exons typically being unmethylated (especially if the gene is expressed) [84, 85] while internal exons and introns tend to show different DNA methylation [85]. While the difference between X-linked male and female gene-body DNA methylation is small, this is consistent with our previous analysis [273] and suggests that X-linked male:female gene-body DNA methylation differences may not be as large as other studies have suggested.

There are several features of exons and introns which may explain the observed differences in DNA methylation. Although exons typically make up a smaller portion of genes than introns, exons have a higher GC content and CpG fraction than any region of the genome other than promoters [282]. The difference in size between exons and introns may have also influenced the observed DNA methylation as the smaller exons will be more affected by the surrounding DNA methylation than the larger introns. CpG density is known to have an effect on the pull down success of techniques such as MeDIP and while BATMAN is designed to correct for the effect of CpG density on pull down efficiency [263] the DNA methylation differences observed between exons and introns may in part be due to differences in CpG density. Exons have been shown to be enriched, compared to introns, for histone modifications associated with the transcription of active genes [284]; our data suggest that X-linked exons maintain their DNA methylation status regardless of expression while introns show increased DNA methylation with higher expression, suggesting that transcription may affect exons and introns differently.

The nature of the X_a and X_i provides a unique system to compare DNA methylation between active and inactive chromatin domains. We conclude that the largest difference in X-linked DNA methylation between males and females is at CpG island promoters. Therefore, we proposed

that DNA methylation differences between the sexes could be used to predict XCI status and overall found good concordance with the XCI status previously determined by expression analysis. Most genes showed similar DNA methylation, and therefore the same predicted XCI status across tissues. Thus, our results support that discrepancies between the XCI status we predicted using DNA methylation and those previously determined may be due to tissue-specific XCI, as 12% of genes showed DNA methylation patterns suggestive of tissue-specific XCI in the four tissues we examined. Using DNA methylation to predict XCI status would allow for examination of genes that are not expressed and would not require extraction of RNA or restrict studies to females with clonal XCI. Outside of CpG islands, chromosome-wide DNA methylation analysis showed differences between exons and introns that suggest that the effects of transcription on gene-body DNA methylation may affect exons and introns differently.

Table 2.1: Q-PCR primer sequences and conditions used to determine the XCI status of *TSR2* and *ZRSR2* along with control assay *ZFX*.

Assay	Sequence (5' to 3')	Annealing Temperature
ZFX	F: TTCAGTGCCCAGATATCATGGA R: GGACTGTGCAATGTGCTAAAGAA	80°C
TSR2	F: TGGCCTCCTGCATCACTCA R: CCACATCATCTTCATCCTCATCAG	79°C
ZRSR2	F: CAACATCCAGTCCTACCCTTCTTATT R: TGGTAGGTTTCTTCCTCGCTGTA	80°C

Table 2.2: Genes in which different probes for the same gene show DNA methylation conflicts in blood.

A conflict was defined as one probe being predicted to escape from XCI while another probe in the same gene was predicted to be subject to XCI. DNA methylation classes: unmethylated (0 to 15% methylated), intermediate DNA methylation (15 to 60% methylated) and methylated (60 to 100% methylated).

gene name	Carrel and Willard [106] XCI status	probe location (bp)	probe CpG density	distance from TSS (bp)	female blood average	male blood average	average female DNA methylation class	average male DNA methylation class	predicted XCI status (by probe)
<i>UBA1</i>	escape	46,934,973	IC	-231	8%	1%	unmethylated	unmethylated	escape
		46,935,424	IC	220	13%	1%	unmethylated	unmethylated	subject
<i>DIAPH2</i>	variable escape	95,826,287	HC	-78	22%	4%	intermediate	unmethylated	subject
		95,826,366	HC	1	11%	3%	unmethylated	unmethylated	escape
<i>GLA</i>	variable escape	100,549,431	IC	176	7%	3%	unmethylated	unmethylated	escape
		100,549,707	IC	-100	46%	2%	intermediate	unmethylated	subject
<i>PAK3</i>	variable escape	110,225,969	IC	-287	4%	1%	unmethylated	unmethylated	escape
		110,226,518	IC	262	25%	2%	intermediate	unmethylated	subject
<i>L1CAM</i>	escape	152,794,259	IC	246	6%	3%	unmethylated	unmethylated	escape
		152,794,653	IC	-148	16%	2%	intermediate	unmethylated	subject

Table 2.3: Novel XCI status as predicted by DNA methylation.

Genes with no data on XCI status from Carrel and Willard [106] but where an XCI status was predicted based on DNA methylation (subject: no formatting, variable escape: underlined and escape: bold)

<i>AFF2</i>	<i>CXorf40B</i>	<i>GPC3</i> ^a	<i>MBTPS2</i>	<i>POU3F4</i>	<i>UPRT</i> ^b
<i>APBC2</i>	<i>CXorf41</i>	<i>GPR101</i>	<i>MMGT1</i>	<i>RAB33A</i>	<i>XK</i>
<i>ARHGAP36</i>	<i>CXorf58</i>	<i>GRIA3</i>	<i>NCRNA00086</i>	<i>RPL13</i>	<i>XKRX</i>
<i>ARHGEF9</i>	<i>DACH2</i>	<i>HDAC8</i> ^a	<i>NDUFA1</i>	<i>SLC7A3</i>	<i>ZC4H2</i>
<i>ARX</i> ^a	<i>EGFL6</i> ^b	<i>IGSF1</i>	<i>NDUFB11</i>	<i>SLITRK4</i>	<i>ZCCHC12</i>
<i>BAD</i>	<i>ERAS</i>	<i>IL1RAPL2</i>	<i>NKAPP1</i>	<i>SOX3</i>	<i>ZDHHC15</i>
<i>BCAP31</i>	<i>ESX1</i>	<i>IRS4</i>	<i>NR0B1</i>	<i>TBX22</i>	<i>ZIC3</i>
<i>BRWD3</i>	<i>F8A1</i>	<i>KLHL34</i>	<i>NSBP1</i>	<i>TCEAL3</i>	<i>ZNF711</i>
<u><i>CDX4</i></u>	<i>FAM120C</i>	<i>LOC389852</i>	<i>PCDH11X</i> ^b	<i>TMSB4X</i> ^a	<i>ZNK280C</i>
<i>CNKSR2</i>	<i>FAM123B</i>	<i>MAOB</i> ^b	<i>PNCK</i>	<i>TSR2</i>	<i>ZRSR2</i> ^b
<i>CXorf22</i>	<i>GABRQ</i>				

^a genes with other evidence for being subject to XCI

^b genes with other evidence for escaping to XCI

Table 2.4: Genes predicted to show tissue-specific XCI.

Tissue-specific XCI was defined as a gene in which are least one tissue (blood, fetal muscle, fetal neural or fetal kidney) had a different XCI status from the other tissues. The XCI status in each tissue is given for probes and genes. XCI status defined as escape, variable escape, subject or unclassifiable based on the decision tree in Figure 2.1.

gene name	Carrel and Willard [106] XCI status	probe location (bp)	probe CpG density	distance from TSS (bp)	Blood		Muscle		Neural		Kidney	
					predicted XCI status (by probe)	predicted XCI status (by gene)	predicted XCI status (by probe)	predicted XCI status (by gene)	predicted XCI status (by probe)	predicted XCI status (by gene)	predicted XCI status (by probe)	predicted XCI status (by gene)
<i>ARHGAP6</i>	subject	11,594,040	HC	-298	S	S	E	E	E	E	E	E
<i>FRMPD4</i>	variable escape	12,066,420	HC	-86	VE	VE	E	E	E	E	E	E
<i>TMSB4X</i>	no data	12,902,888	HC	-260	S	S	U	U	VE	VE	VE	VE
<i>PIGA</i>	subject	15,263,175	HC	391	S	S	E	E	E	E	S	S
<i>PHKA2</i>	variable escape	18,913,164	HC	-1067	S	S	E	E	U	U	E	E
<i>PDHA1</i>	subject	19,271,661	HC	-311	S	S	S	S	VE	VE	S	S
<i>SH3KBP1</i>	variable escape	19,817,043	IC	-1403	S	S	S	S	U	U	VE	VE
<i>RP11-450P7.3</i>	no data	21,585,788	HC	581	S	S	U	E	U	E	U	E
		21,586,973	HC	-604	S	S	E	E	E	E	E	E
<i>SAT</i>	subject	23,710,977	HC	-248	S	S	E	E	E	E	E	E
<i>CXorf22</i>	no data	35,847,870	HC	51	S	S	E	E	E	E	E	E
<i>TSPAN7</i>	subject	38,305,270	HC	-413	U	S	U	S	U	VE	U	S
		38,306,022	HC	339	S	S	S	S	VE	VE	S	S
<i>MAOB</i>	no data	43,626,770	HC	-145	S	S	E	E	E	E	E	E
<i>RP2</i>	subject	46,581,295	HC	-25	S	S	S	S	VE	VE	S	S
		46,581,370	HC	50	S	S	S	S	U	VE	S	S
<i>UXT</i>	subject	47,403,713	HC	-209	S	S	U	S	U	S	U	VE
		47,403,793	HC	-289	S	S	S	S	S	S	VE	VE
<i>RRAGB</i>	subject	55,760,794	IC	-110	S	S	S	S	VE	VE	S	S
		55,760,916	IC	12	S	S	S	S	VE	VE	U	S
<i>MTMR8</i>	subject	63,532,108	HC	-72	S	S	S	S	VE	VE	S	S
<i>MSN</i>	subject	64,804,464	HC	228	S	S	S	S	E	E	S	S
<i>CDX4</i>	no data	72,583,795	IC	-20	U	VE	U	U	U	VE	U	U
		72,583,820	IC	5	VE	VE	U	U	VE	VE	U	U
<i>RNF12</i>	subject	73,751,321	HC	-153	S	S	S	S	VE	VE	S	S
<i>ABCB7</i>	variable escape	74,292,691	HC	166	S	S	S	S	E	E	S	S
		74,292,976	HC	-119	S	S	S	S	U	E	S	S
<i>POU3F4</i>	no data	82,649,903	IC	-54	S	S	S	S	U	U	S	conflict
		82,650,174	IC	217	S	S	U	S	U	U	E	

gene name	Carrel and Willard [106] XCI status	probe location (bp)	probe CpG density	distance from TSS (bp)	Blood		Muscle		Neural		Kidney	
					predicted XCI status (by probe)	predicted XCI status (by gene)						
<i>FSHPRH1</i>	subject	100,240,009	HC	-1445	S	S	S	S	VE	VE	S	S
<i>ARMCX2</i>	subject	100,800,794	IC	697	S	S	VE	VE	S	S	S	S
		100,801,504	IC	-13	S	S	U	VE	U	S	S	S
<i>TCEAL2</i>	subject	101,267,674	HC	358	S	S	E	E	E	E	E	E
<i>TCEAL3</i>	no data	102,749,660	HC	170	S	S	E	E	U	U	U	U
<i>MORF4L2</i>	variable escape	102,828,154	IC	98	S	S	E	conflict	E	E	E	conflict
		102,828,325	IC	-73	S	S	S	conflict	E	E	S	conflict
<i>IL1RAPL2</i>	no data	103,697,114	IC	-538	S	S	E	conflict	E	E	E	E
		103,697,681	IC	29	S	S	S	conflict	U	E	U	E
<i>FLJ11016</i>	variable escape	106,248,228	HC	462	S	S	E	E	E	E	E	E
<i>PRPS1</i>	subject	106,758,134	HC	-281	S	S	E	E	E	E	E	E
<i>PSMD10</i>	subject	107,221,590	HC	-86	S	S	E	E	E	E	E	E
<i>TRPC5</i>	variable escape	111,212,151	IC	509	VE	VE	U	U	E	E	S	S
<i>AMOT</i>	subject	111,970,237	IC	462	S	S	S	conflict	VE	E	S	conflict
		111,971,104	HC	-405	S	S	E	conflict	E	E	E	conflict
<i>WDR44</i>	subject	117,363,720	HC	-350	S	S	S	S	VE	VE	VE	VE
		117,363,780	HC	-290	S	S	VE	S	VE	VE	VE	VE
<i>CXorf56</i>	variable escape	118,583,511	IC	-119	S	S	S	S	VE	VE	S	S
<i>UPF3B</i>	variable escape	118,870,975	HC	21	S	S	S	S	VE	VE	S	S
		118,871,191	HC	-195	S	S	U	S	U	VE	U	S
<i>ZBTB33</i>	subject	119,268,412	HC	-223	S	S	S	S	VE	VE	S	S
<i>COVA1</i>	variable escape	129,864,838	HC	51	S	S	E	E	E	E	E	E
<i>BGN</i>	subject	152,414,100	IC	495	VE	VE	S	S	S	S	VE	VE
<i>DUSP9</i>	subject	152,561,503	IC	321	S	S	S	S	U	U	VE	VE
<i>STK23</i>	subject	152,699,674	IC	-30	VE	VE	S	S	S	S	S	S
		152,699,961	IC	257	VE	VE	S	S	U	S	S	S
<i>ARD1A</i>	variable escape	152,854,219	HC	-557	VE	VE	S	S	S	S	S	S
<i>RENBP</i>	variable escape	152,863,416	IC	10	S	S	E	E	E	E	E	E
<i>CXorf12</i>	variable escape	152,889,981	HC	-1454	S	S	E	E	E	E	E	E
		152,891,773	HC	338	S	S	U	E	U	E	U	E

Table 2.5: Predicted XCI status of X-linked genes in females with multiple fetal tissues.

All three fetal tissues (muscle, neural and kidney) were available for four (mT4, FT13, FT4 and FT20) of six females while the remaining two females (FT3 and FT16) only had two fetal tissues available.

tissues examined	Females with multiple fetal tissues					
	FT3	FT16	mT4-5	FT13	FT4	FT20
	muscle - kidney	- neural kidney	muscle neural kidney	muscle neural kidney	muscle neural kidney	muscle neural kidney
subject in all tissues	60%	45%	50%	55%	53%	54%
escape in all tissues	14%	15%	12%	12%	13%	11%
subject (but not escape) in some tissues ^a	8%	15%	17%	13%	13%	15%
escape (but not subject) in some tissues ^b	4%	9%	7%	7%	6%	5%
unclassifiable in all tissues	14%	12%	8%	9%	10%	9%
tissue-specific XCI	1%	4%	7%	4%	5%	6%

^a at least one tissue was deemed unclassifiable based on the average male and female DNA methylation, however, all classifiable genes were predicted to be subject to XCI.

^b at least one tissue was deemed unclassifiable based on the average male and female DNA methylation, however, all classifiable genes were predicted to escape to XCI.

Table 2.6: Tissue combinations of genes predicted to escape from XCI and to be subject to XCI for tissue-specific XCI in females with multiple tissues.

All three fetal tissues (muscle, neural and kidney) were available for four (mT4, FT13, FT4 and FT20) of six females while the remaining two females (FT3 and FT16) only had two fetal tissues available.

Predicted XCI status of escape	Predicted XCI status of subject	Females with multiple fetal tissues					
		FT3 muscle - kidney	FT16 - neural kidney	mT4-5 muscle neural kidney	FT13 muscle neural kidney	FT4 muscle neural kidney	FT20 muscle neural kidney
neural	muscle	-	-	19	8	7	13
neural	kidney	-	12	1	1	5	1
neural	muscle and kidney	-	-	0	0	0	0
kidney	neural	-	3	0	0	1	1
kidney	muscle	2	-	3	3	2	3
kidney	muscle and neural	-	-	1	3	3	1
muscle	kidney	1	-	0	0	0	0
muscle	neural	-	-	0	0	0	1
muscle and kidney	muscle	-	-	0	0	0	1
muscle and neural	kidney	-	-	0	0	0	0

Table 2.7: Genes with LC probes that have DNA methylation predictive of an XCI status.

Any HC or IC probes located in the same gene are also listed.

gene name	Carrel and Willard [106] XCI status	probe location (bp)	probe CpG density	distance from TSS (bp)	female blood average	male blood average	predicted XCI status (by probe)	predicted XCI status (by gene)
<i>TLR7</i>	no data	12,794,895	LC	-228	0.31	0.30	unclassifiable	subject
		12,796,042	LC	919	0.64	0.46	subject	
<i>GRPR</i>	variable escape	16,051,530	LC	185	0.03	0.02	escape	escape
<i>IL1RAPL1</i>	no data	28,514,582	LC	-898	0.12	0.06	escape	conflict
		28,515,663	LC	183	0.44	0.07	subject	
<i>CXorf21</i>	no data	30,505,611	LC	224	0.35	0.02	subject	subject
<i>TIMP1</i>	variable escape	47,326,723	LC	89	0.07	0.01	escape	escape
<i>WAS</i>	variable escape	48,426,376	IC	-777	0.84	0.77	unclassifiable	subject
		48,427,342	LC	189	0.43	0.04	subject	
<i>PCSK1N</i>	subject	48,578,609	HC	295	0.55	0.04	subject	subject
		48,579,311	LC	-407	0.66	0.47	variable escape	
<i>KCND1</i>	subject	48,713,014	LC	181	0.14	0.04	escape	escape
<i>GPR173</i>	subject	53,095,229	LC	-2	0.59	0.41	variable escape	variable escape
		53,095,356	LC	125	0.36	0.31	unclassifiable	
<i>FLJ31204</i>	variable escape	57,329,759	LC	-105	0.29	0.02	subject	subject
		57,329,800	LC	-64	0.28	0.09	variable escape	
<i>EDA</i>	subject	68,752,214	LC	-422	0.20	0.04	subject	subject
		68,752,833	IC	197	0.55	0.02	subject	
<i>IL2RG</i>	no data	70,248,040	LC	88	0.22	0.03	subject	subject
<i>ZMYM3</i>	subject	70,390,897	HC	251	0.49	0.01	subject	subject
		70,391,508	LC	-360	0.40	0.02	subject	
<i>CXCR3</i>	subject	70,754,715	LC	377	0.74	0.59	variable escape	variable escape
<i>HDAC8</i>	no data	71,709,270	HC	108	0.53	0.09	subject	conflict
		71,709,900	LC	-522	0.07	0.03	escape	
<i>COX7B</i>	variable escape	77,041,800	LC	183	0.39	0.03	subject	subject
<i>P2RY10</i>	no data	78,086,697	LC	-788	0.69	0.79	unclassifiable	escape
		78,087,577	LC	92	0.06	0.03	escape	
<i>NSBP1</i>	no data	80,263,404	LC	368	0.30	0.02	subject	subject
		80,264,064	IC	-292	0.28	0.01	subject	
<i>SH3BGRL</i>	escapes	80,344,705	LC	427	0.43	0.01	subject	subject
<i>CHM</i>	variable escape	85,189,965	LC	-743	0.79	0.67	variable escape	variable escape
		85,190,083	LC	-861	0.64	0.70	unclassifiable	
<i>TMEM35</i>	subject	100,220,434	LC	-85	0.24	0.08	subject	subject
		100,220,606	LC	87	0.41	0.23	variable escape	
<i>WBP5</i>	variable escape	102,497,909	LC	-127	0.17	0.04	subject	subject

<i>gene name</i>	Carrel and Willard [106] XCI status	probe location (bp)	probe CpG density	distance from TSS (bp)	female blood average	male blood average	predicted XCI status (by probe)	predicted XCI status (by gene)
<i>FLJ11016</i>	variable escape	106,248,228 106,249,029	HC LC	462 -339	0.15 0.44	0.01 0.18	subject subject	subject
<i>COL4A6</i>	variable escape	107,569,195 107,569,422	LC LC	172 -55	0.79 0.12	0.84 0.06	unclassifiable escape	escape
<i>DCX</i>	no data	110,542,022	LC	40	0.13	0.16	escape	escape
<i>LHFPL1</i>	variable escape	111,809,867 111,811,026	LC LC	71 -1088	0.21 0.62	0.06 0.86	variable escape unclassifiable	variable escape
<i>AGTR2</i>	no data	115,215,546	LC	-485	0.72	0.60	variable escape	variable escape
<i>SH2D1A</i>	no data	123,307,985	LC	110	0.18	0.05	variable escape	variable escape
<i>CXorf9</i>	subject	128,741,352 128,741,661	LC LC	-289 20	0.53 0.57	0.07 0.14	subject subject	subject
<i>ZNF75</i>	subject	134,304,692	LC	931	0.43	0.03	subject	subject
<i>CD40LG</i>	variable escape	135,558,079	LC	77	0.68	0.49	variable escape	variable escape
<i>ARHGEF6</i>	subject	135,690,989 135,691,446	LC LC	180 -277	0.13 0.13	0.02 0.06	variable escape escape	escape
<i>GABRA3</i>	no data	151,370,046 151,370,895	LC LC	440 -409	0.39 0.05	0.13 0.02	subject escape	conflict
<i>AVPR2</i>	escapes	152,823,017 152,823,584	LC LC	-547 20	0.75 0.48	0.57 0.43	subject unclassifiable	subject
<i>F8</i>	subject	153,904,199	LC	-7	0.16	0.14	escape	escape
<i>CXorf53</i>	no data	153,951,795	LC	-1109	0.46	0.02	subject	subject
<i>CLIC2</i>	variable escape	154,217,162	LC	18	0.38	0.21	variable escape	variable escape

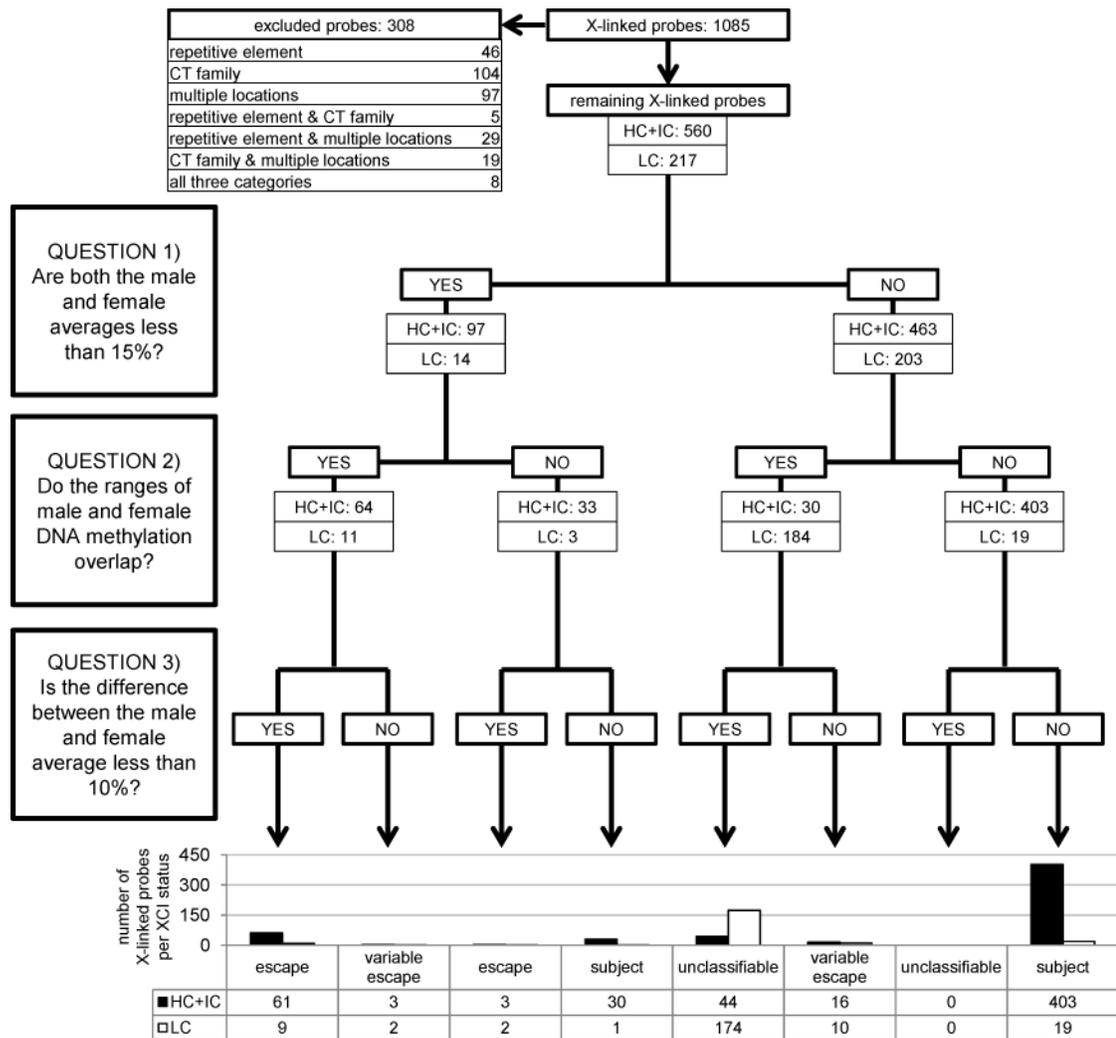


Figure 2.1: Decision tree used to predict the XCI status of X-linked probes.

308 X-linked probes were removed from analysis for three possible reasons: being located in a repetitive element, being located in the promoter of a CT family gene, or mapping to one than one location in the genome. Several probes were removed for multiple reasons and the number of probes removed in each category is given. To predict XCI status, the remaining probes were first divided based on whether both the male average and female average DNA methylation levels were less than 15% (unmethylated). Next, the range of male and female DNA methylation were compared to see if they overlapped. Lastly, the difference between the average male and female DNA methylation were compared to see if they were less than 10%. Based on these decisions probes were classified as escaping from XCI, variably escaping from XCI, being subject to XCI or being unclassifiable. The number of probes in HC or IC islands or LC probes which satisfy each criteria are given at all steps in the tree as well as at the final predicted XCI status. Probes located in the same gene were combined to give a predicted XCI for each gene as shown in Figure 2.3.

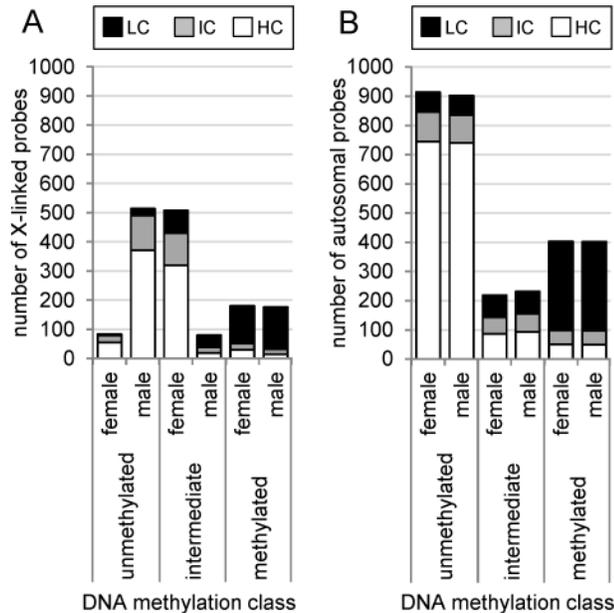


Figure 2.2: Promoter DNA methylation analysis in blood (female: n=59 and male: n=36) reveals X chromosome sex-specific DNA methylation differences as well as differences based on CpG density.

Probes were divided by CpG density (LC: black, IC: grey, HC: white) and classified as unmethylated (0% to 15% methylated), intermediate (15% to 60% methylated) or methylated (60% to 100% methylated) in males and females. **(A)** DNA methylation levels in males and females were significantly different (p -value<0.0001, Mann-Whitney test) across all X-linked probes. The majority of HC and IC promoter probes ($n=560$) on the X chromosome were unmethylated in males and intermediate in females. X-linked LCs probes ($n=217$) were mostly methylated regardless of sex. **(B)** DNA methylation levels in males and females were not significantly different (p -value=0.2779, Mann-Whitney test) across autosomal probes. The majority of promoter probes on chromosomes 20, 21 and 22 were unmethylated. Probes in HC and ICs ($n=1088$) were mostly unmethylated whereas LC ($n=448$) probes were mostly methylated. Males and females showed no differences in their DNA methylation classes.

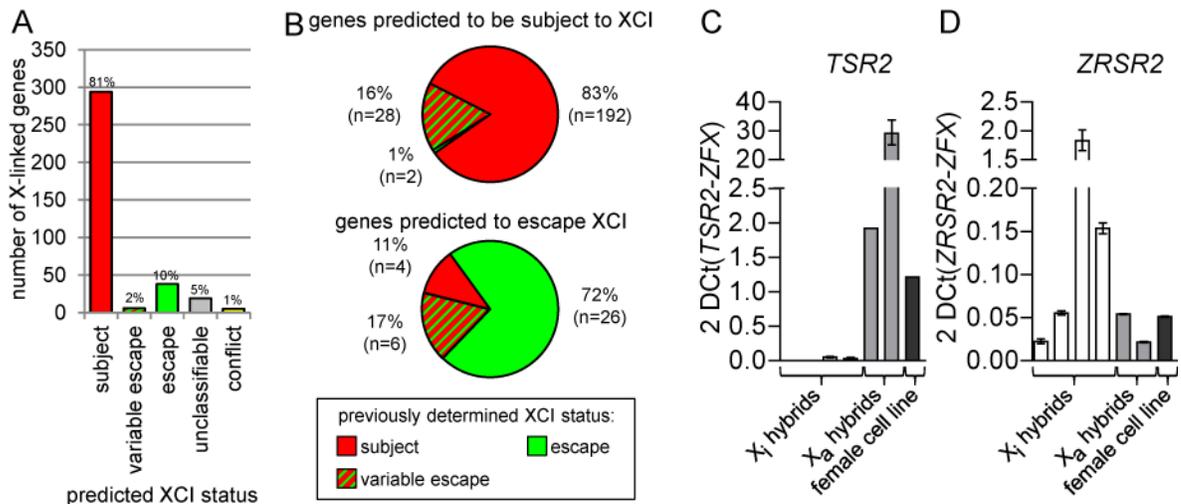


Figure 2.3: The XCI status predicted using DNA methylation in blood corresponds with previously determined XCI status.

(A) The XCI status (subject: red, variable escape: red and green diagonal stripes, escape: green, unclassifiable: grey, conflicts: yellow) of 372 X-linked genes with probes in HC and IC islands was predicted using DNA methylation. The percentage of the total X-linked genes with probes in HC and IC islands is given for each predicted XCI status. In blood, the majority (81%) of genes are predicted to be subject to XCI. (B) The XCI status previously determined by Carrel and Willard [106] in somatic cell hybrids (subject: red, variable escape: red and green diagonal stripes, escape: green) for those genes predicted by DNA methylation to be subject to XCI (red bar in (A) and top pie chart) and those genes predicted to escape from XCI (green bar in (A) and bottom pie chart). (C-D) Q-RT-PCR confirmation in somatic cell hybrids of predicted XCI status based on DNA methylation. The expression level of two genes (*TSR2* and *ZRSR2*) in four somatic cell hybrids containing a human *X_i* (white: t75-2maz34-4a, t48-1a-1DAZ4A, t86-B1maz1b-3a and t11-4Aaz5), two somatic cell hybrids containing a human *X_a* (light grey: AHA-11aB1 and t60-12) and a control female cell line (dark grey: GM7350) were compared to confirm that DNA methylation could predict XCI status. Test genes (*TSR2* and *ZRSR2*) were normalized against a gene known to escape from XCI (*ZFX*). Error bars represent the positive and negative error between three replicate PCRs. (C) *TSR2* was unmethylated in male blood and intermediate in female blood and was predicted to be subject to XCI. (D) *ZRSR2* was unmethylated in male and female blood and was predicted to escape from XCI.

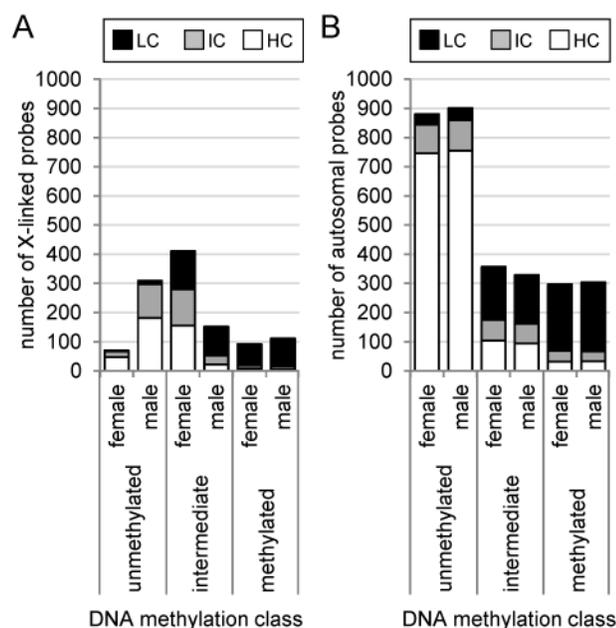


Figure 2.4: DNA methylation in female somatic tissues is similar to blood.

Promoter DNA methylation analysed in fetal somatic tissues (muscle: female n=6 and male n=4, kidney: female n=6 and male n=5, brain: female n=4 and male n=4 and spinal cord: female n=2 and male n=1) revealed X chromosome sex-specific DNA methylation differences as well as differences based on CpG density similar to those observed in blood. Probes were divided by CpG density (LC: black, IC: grey, HC: white) and classified as unmethylated (0% to 15% methylated), intermediate (15% to 60% methylated) or methylated (60% to 100% methylated) in males and females. **(A)** DNA methylation levels in males and females were significantly different (p -value<0.0001, Mann-Whitney test) across all X-linked probes. The majority of HC and IC promoter probes (n=560) on the X chromosome were unmethylated in males and intermediate in females. X-linked LCs probes (n=217) were mostly methylated regardless of sex. **(B)** DNA methylation levels in males and females were not significantly different (p -value=0.9821, Mann-Whitney test) across all autosomal probes. The majority of promoter probes on chromosomes 20, 21 and 22 were unmethylated. Probes in HC and ICs (n=1088) are mostly unmethylated whereas LC (n=448) probes were mostly methylated. Males and females showed no difference in their DNA methylation classes.

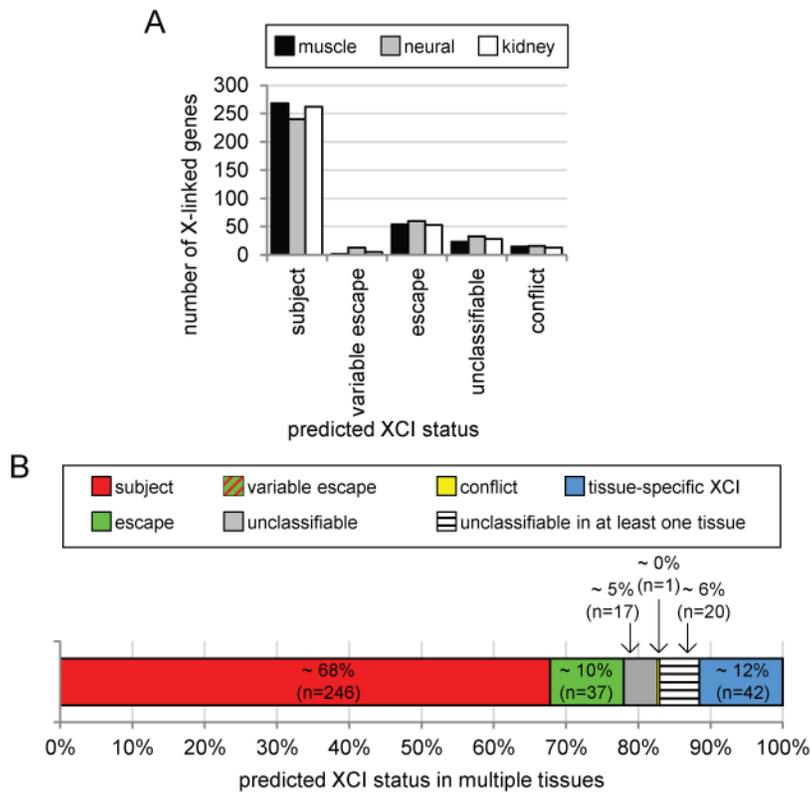


Figure 2.5: Most genes show the same predicted XCI status in all tissues examined while 12% of genes show tissue-specific XCI.

(A) Male and female DNA methylation was used to predict XCI status (as outlined in Table 2.1) of genes with probes in HC and IC islands in fetal muscle (black: female n=6, male n=4), fetal neural tissue (grey: female n=6, male n=5) and fetal kidney (white: female n=6, male n=5). (B) The combined predicted XCI status in all four tissues examined (blood, fetal muscle, fetal neural and fetal kidney). The majority of genes showed the same XCI status (subject: black, variable escape: diagonal stripes, escape: white, unclassifiable: grey, conflicts: dotted) in all tissues. No genes were found to variably escape from XCI in all tissues. ~6% of genes were unable to predict an XCI status in at least one tissue but predicted same XCI status in all other tissues (horizontal stripes). ~12% of genes showed tissue-specific DNA methylation differences which resulted in at least one tissue having a different predicted XCI from the other tissues (blue).

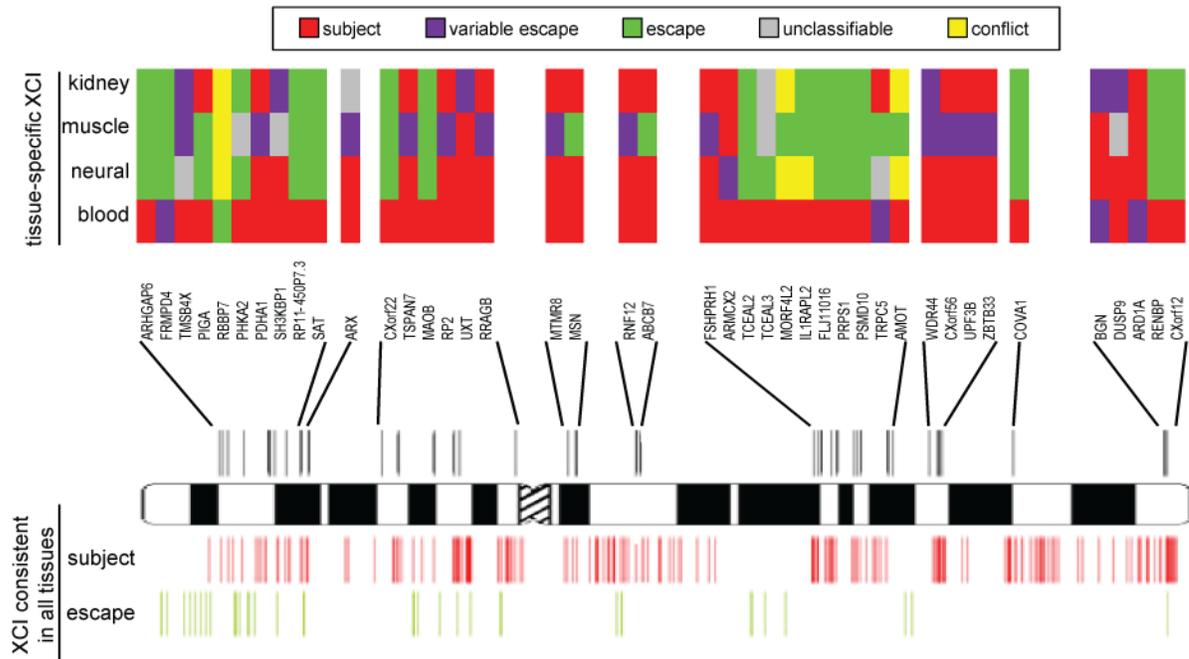


Figure 2.6: Genes predicted to show tissue-specific XCI based on DNA methylation from the Illumina Infinium HumanMethylation27 array as found across the X chromosome.

The genomic locations of genes which showed the same predicted XCI status in all tissues examined (escape: green, subject: red) are shown below the X chromosome ideogram. On the top are the genomic locations of genes in which at least one tissue had a predicted XCI status different from the other tissues. The predicted XCI status (subject: red, variable escape: purple, escape: green, unclassifiable: grey, conflict: yellow) in each tissue examined (blood, fetal neural, fetal muscle and fetal kidney) is shown along with the names of all genes which show tissue-specific XCI.

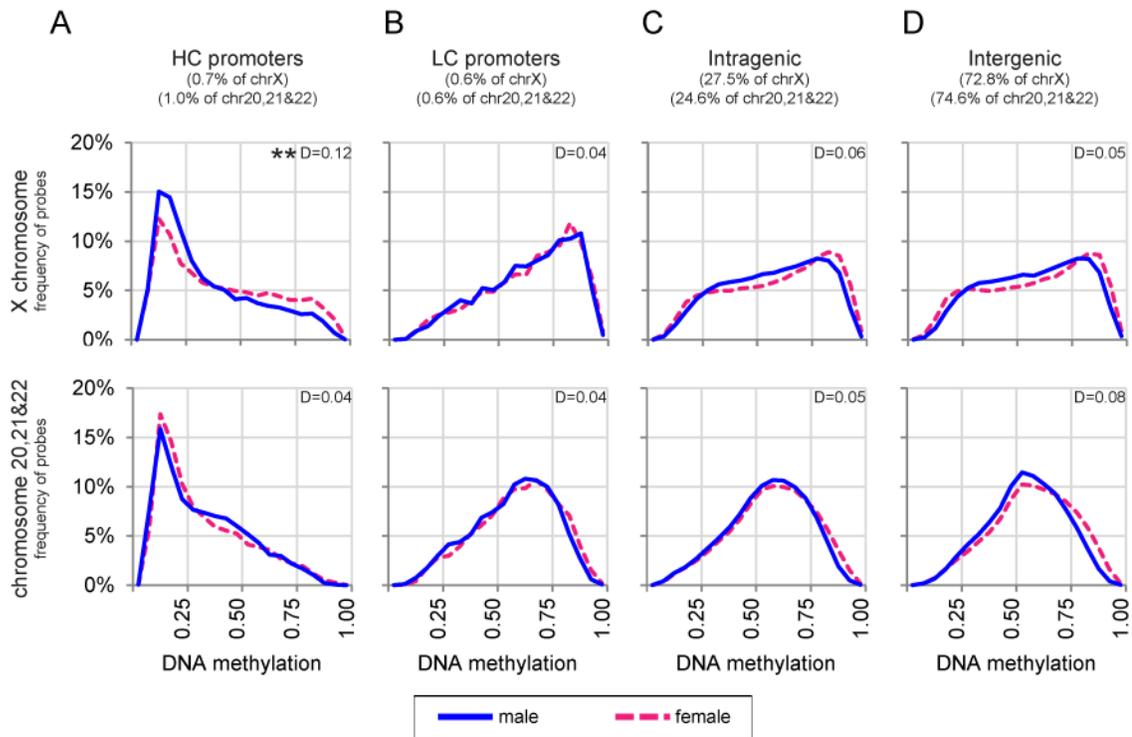


Figure 2.7: DNA methylation histograms reveal X-linked HC promoters show the largest X-linked sex-specific DNA methylation difference.

The average male and average female DNA methylation from probes representing four different genomic elements was used to create DNA methylation histograms by determining the frequency at which probes were at a specific level of DNA methylation (20 bins from 0 to 1.00 methylated). Female DNA methylation frequencies are shown as pink dotted lines and males as solid blue lines with DNA methylation frequencies from the X chromosome displayed on the upper row and the autosomal average from chromosomes 20, 21 and 22 on the bottom. The percentage of the total chromosomal DNA each element represents is given for the X chromosome and the autosomes (chromosomes 20, 21 and 22). Significance was calculated comparing the distribution of average male and average female DNA methylation using the Kolmogorov-Smirnov Test. When p-values were greater than 0.0001 they were not significant, however, p-values between 0.0001 and 1.0 E-10 (*) and p-values <1.0 E-10 (**) were considered significantly different. **(A-B)** Promoter elements (the one kb up and downstream of all TSS) showed differences in DNA methylation frequencies based on CpG density. HC promoters **(A)** showed males were hypomethylated compared to females on the X chromosome but not the autosomes. LC promoters **(B)** (contained neither an HC nor IC island) showed no sex-specific DNA methylation difference on either the X chromosome or the autosomes. **(C-D)** Non-promoter elements tended to be methylated on the male and female X chromosome and intermediately methylated on the autosomes in both intragenic **(C)** and intergenic **(D)** regions.

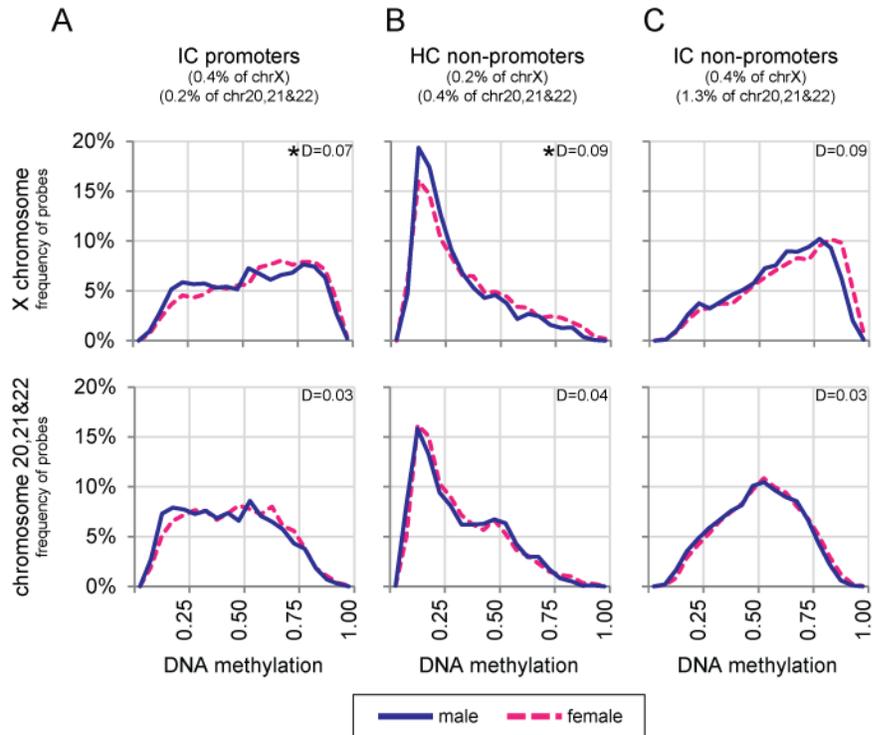


Figure 2.8: IC promoters as well as HC non-promoters show a sex-specific DNA methylation difference on the X chromosome but not on the autosomes.

The average male and average female DNA methylation from probes representing different genomic elements was used to create DNA methylation histograms by determining the frequency at which probes were at a specific level of DNA methylation (20 bins from 0 to 1.0 methylated). Female DNA methylation frequencies are shown as pink dotted lines and males as solid blue lines with DNA methylation frequencies from the X chromosome displayed on the upper row and the autosomal average from chromosomes 20, 21 and 22 on the bottom. The percentage of the total chromosomal DNA each element represents is given for the X-chromosome and the autosomes (chromosomes 20, 21 and 22). Significance was calculated comparing the distribution of male and female DNA methylation using the Kolmogorov-Smirnov Test. When p -values > 0.0001 were not significant (ns), p -values between 0.0001 and 1.0×10^{-10} (*) and p -values $< 1.0 \times 10^{-10}$ (**) were considered significantly different. **(A)** IC promoters (the one kb up and downstream of all TSS which contained an IC but not an HC) showed a slight X-linked DNA methylation differences between the sexes. **(B-C)** HC and IC islands not associated with known promoters revealed HC non-promoter islands **(B)** were similarly methylated to HC promoters while IC non-promoter islands **(C)** were more methylated.

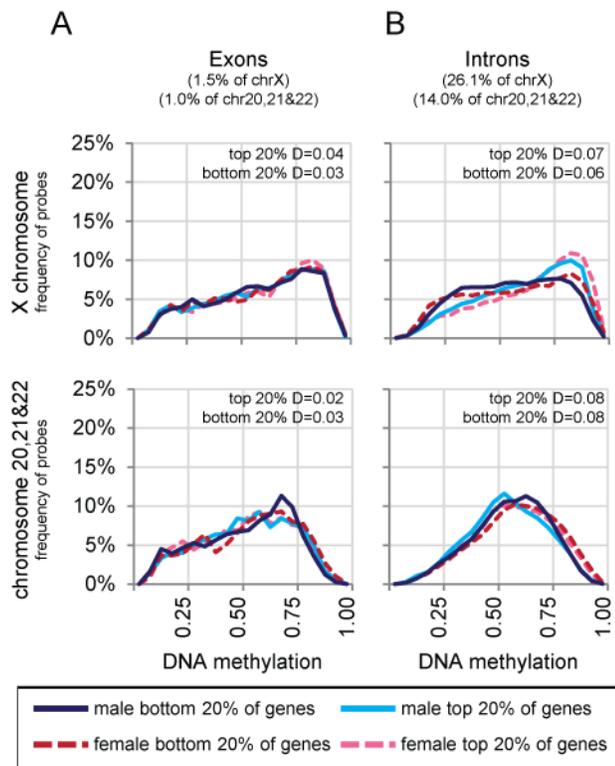


Figure 2.9: X-linked introns but not exons show differences in DNA methylation based on expression level.

The average male and average female DNA methylation from probes representing four different genomic elements was used to create DNA methylation histograms by determining the frequency of probes at a specific level of DNA methylation (20 bins from 0 to 1.0 methylated). Female DNA methylation frequencies are shown as dotted lines and males as solid lines with DNA methylation frequencies from the X chromosome displayed on the upper row and the autosomal average from chromosome 20, 21 and 22 on the bottom. X-linked and autosomal (chromosomes 20, 21 and 22) genes were separated based on expression (determined in [266]) and the top (male: light blue, female: light pink) and bottom (male: dark blue, female: dark pink) 20% divided into those which correspond to either the exons (**A**) or introns (**B**). Significance was calculated comparing the distribution of male and female DNA methylation using the Kolmogorov-Smirnov Test. When p-values were greater than 0.0001 they were not significant (ns), however, p-values between 0.0001 and 1.0 E-10 (*) and p-values <1.0 E-10 (**) were considered significantly different. While exons (**A**) were similarly methylated regardless of sex or expression level on both the X chromosome and the autosomes, introns (**B**) were more methylated in X-linked genes with high expression than X-linked genes with low expression in both sexes. Autosomal introns showed no DNA methylation difference in either sex.

3 X_i chromosome-specific reduction in placental DNA methylation

3.1 Introduction

The tissue-specific differences in XCI detected in chapter 2 were based on comparisons of DNA methylation in only somatic tissues. The placenta is composed of numerous villi, the outer most layer of these villi is derived from the trophoblast of the blastocyst whereas the inner layers are derived from the inner cell mass. While many studies use the mouse as a model of human XCI, XCI in the extra-embryonic tissues of the mouse does not accurately represent XCI in human extra-embryonic tissues. An important difference between mouse and human XCI in extra-embryonic tissues is that mice show imprinted XCI while humans have random XCI [31, 139, 141-145]. These species-specific differences in XCI highlight the importance of studying XCI in the human placenta. Specifically, the placenta provides a unique opportunity to study the role that DNA methylation may play in maintaining XCI.

Previous studies have shown a global reduction in DNA methylation of placenta compared with the vast majority of somatic tissues and a specific placental hypomethylation of repetitive elements and at a limited number of X-linked promoters [160-162, 285]. Once XCI is established *XIST* is not necessary to maintain the silent nature of the X_i [12-14], however, the removal of *Xist* in mice results in a higher frequency of genic reactivation, suggesting that other factors, including DNA methylation, may act with the *XIST/Xist* RNA to maintain silencing [14]. The best example of XCI reactivation in the placenta comes from studies on the *G6PD* gene. When an individual has one *G6PD* variant on the X_a and the other on the X_i (*G6PD* A and *G6PD* B) a heterodimer will be formed if *G6PD* escapes from XCI [152]. *G6PD* heterodimers are more common in placental cells and later it was determined that the CpG island promoter of *G6PD* was hypomethylated in the placenta leading to the hypothesis that placental hypomethylation created a permissive environment in which X reactivation of single genes could occur [150, 152].

In this chapter DNA methylation levels were investigated to determine how X-linked placental DNA methylation differed from that of somatic tissues. The comparison of DNA methylation between placenta and blood revealed global L1 placental hypomethylation, as well as placental hypomethylation of X-linked, but not autosomal, promoters. Using a combination of Illumina GoldenGate DNA methylation analysis and pyrosequencing, three different types of X-linked regions (promoter, intragenic and intergenic) were shown to have less DNA methylation in placenta compared with blood. This hypomethylation was greater in female than male placenta implying the difference is predominantly due to the X_i . These findings suggest a different system for the establishment or maintenance of DNA methylation on the X_a and X_i .

3.2 Methods

3.2.1 Sample collection

Placentas were collected by the Robinson lab, with consent, from pregnancies delivered at British Columbia's Women's Hospital. Placentas from females (n=11) were an average of 37.8 months gestational age and had an average maternal age of 34.6 years. Placentas from males (n=15) were an average of 38.9 months gestational age and had an average maternal age of 35.2 years. Blood samples were from anonymous males (n=6) and females (n=6) ranging in age from 2 to 49 years for males and 22 to 44 years for females. Ethics approval was obtained from the University of British Columbia clinical research ethics board (H06-70085).

3.2.2 DNA extraction and bisulfite conversion

DNA was extracted from fresh whole blood samples following a standard salting out DNA extraction [286]. After removal of amniotic and chorionic membranes, DNA was extracted from two separate samples of chorionic villi from the fetal side of the placenta as outlined in Penaherrera *et al.* [287]. 500 ng of DNA was then used for a bisulfite conversion following the instructions in the EZ DNA methylation Gold Kit (Zymo Research Corporation). Complete conversion was confirmed using the internal bisulfite conversion controls in each pyrosequencing assay and the bisulfite conversion control summary graph for the Illumina GoldenGate panel.

3.2.3 Illumina GoldenGate panel

Samples were applied to an Illumina GoldenGate bead array that assays 1505 CpG sites located in promoters across the genome. 84 of these sites were located on the X chromosome and associated with 39 X-linked genes. The location of autosomal CpGs assayed on this panel ranged from 1499 bp upstream of the TSS to 497 bp downstream. The X chromosome assays were located up to 1681 bp upstream of the TSS and 472 bp downstream. Data underwent average normalization using the DNA methylation Module (version 3.2.0) in BeadStudio (version 3.1.3.0 Illumina, Inc.) to ensure that the background intensities of each array were comparable. The heatmap and dendrogram in Figure 3.2 were generated using the Manhattan Hierarchical Cluster Metric in BeadStudio.

3.2.4 Pyrosequencing

Pyrosequencing was performed using a Pyromark ID machine and the PyroGoldSQA reagent kit (Biotage). Each 25 μ L pyrosequencing PCR contained the following final concentration of reagents: 1X PCR Buffer (Qiagen), 0.2 mM dNTPs (deoxynucleosides), 0.025U HotStart Taq DNA Polymerase (Qiagen), 0.25 mM forward primer, 0.25 mM reverse primer and ~25 ng bisulfite converted DNA. Cycling conditions for each assay were the same for all primers

except for the annealing temperature which is listed for each assay individually in Table 3.1. Cycling conditions were: 95°C for 15 minutes, 50 cycles of 94°C for 30 seconds, annealing temperature (listed in Table 3.1) for 30 seconds, 72°C for 60 seconds, followed by a final step of 72°C for 10 minutes. Template preparation and pyrosequencing was then done according to Tost and Gut with sequencing primers listed in Table 3.1 [288]. Global DNA methylation of L1 elements was performed using the L1 assay from Biotage with PCR and cycling conditions as specified by the supplier. The UCSC Genome Browser was used to compare non-promoter regions against 4 histone modifications (H2AK9ac, H2BK5ac, H3K18ac, H3K36ac) that are concentrated around the TSS and eight histone modifications (H2BK12ac, H3K4ac, H4K5ac, H4K8ac, H4K12ac, H2BK20ac, H2BK120ac, H4K16ac) that are elevated in promoters and the transcribed regions of active genes [48, 289]. Four histone methylation modifications (H3K4me1, H3K4me2, H3K4me3, H2A.Z) previously found to be in promoter regions were also examined [290].

3.2.5 Statistical analysis

Statistical analysis of the Illumina GoldenGate panel was performed using the Mann-Whitney test as calculated by GraphPad prism. For the comparison of Illumina GoldenGate results the data were analyzed separately for CpG islands and non-islands based on information from the manufacturer. Male DNA methylation levels were used as the DNA methylation level of the X_a . Since the DNA methylation level obtained from females is the average of both X chromosomes this was multiplied by two and then the X_a DNA methylation was subtracted resulting in the calculated amount of DNA methylation on the X_i .

3.2.6 CpG density definitions used

The definition of a CpG island used by Illumina and UCSC to define a CpG island (GC content greater than 50%, a ratio of greater than 0.6 for $\text{Observed}_{\text{CpG}}/\text{Expected}_{\text{CpG}}$ and a length greater than 200 bp) is the same definition first proposed by Gardiner-Garden and Frommer in 1987 [46]. In a recent genome-wide analysis of promoter DNA methylation, Weber *et al.* introduced three CpG density classes (HC, IC and LC) [45]. To ensure that all potential regions of CpG islands were recognized in this study we have also classified CpG islands based on these three density classes. Each class was defined as follows, HC, greater than 55% GC content, greater than 0.75 $\text{Observed}_{\text{CpG}}/\text{Expected}_{\text{CpG}}$ and a length greater than 500 bp; LC had a $\text{Observed}_{\text{CpG}}/\text{Expected}_{\text{CpG}}$ ratio less than 0.48 and were shorter than 500 bp. Regions which were neither HC or LC were classified as IC. Using this system most regions classified as CpG islands by UCSC are defined as HCs whereas only approximately one quarter of ICs would be classified by UCSC as islands [45].

3.3 Results

3.3.1 Placental hypomethylation is specific to X-linked promoters and repetitive elements

The average DNA methylation at the subset of L1s examined by pyrosequencing, displayed significant (p -value=0.0009) placental hypomethylation compared with blood in both females (32% lower DNA methylation) and males (36% lower DNA methylation) (Figure 3.1A). To determine if the placental hypomethylation previously reported at X-linked promoters also extended to autosomal promoters, the Illumina GoldenGate panel was used to assess the level of promoter DNA methylation at 1421 promoter sites on the autosomes and 84 sites on the X chromosome in both blood and placenta (Figure 3.1B). Average autosomal DNA methylation levels showed no significant difference (p -value=0.05) between male and female blood and placenta, whereas dramatic sex and tissues differences were observed for the X chromosome (p -value<0.001). The X chromosomes were examined in more detail to determine the extent of the X-linked placental hypomethylation.

3.3.2 CpG density influences DNA methylation patterns of X-linked promoters

The Manhattan Hierarchical Cluster Metric (Illumina BeadStudio) separated the 29 samples analyzed into four clusters that corresponded to male blood, female blood, male placenta and female placenta (Figure 3.2). To assess the impact of CpG density we utilized the expanded nomenclature of Weber *et al.* (HC, IC, LC) [45]. In blood, 51 of 84 X-linked promoter assays examined demonstrated moderate DNA methylation in females (average beta-value of 0.67) and negligible DNA methylation in males (average beta-value of 0.08), the anticipated pattern for X-linked CpG island promoters of genes subject to inactivation, MeXIP (labeled group 1 on Figure 3.2). MeXIP assays tended to have less DNA methylation in the placenta. One third of all assays demonstrated high DNA methylation in both male and female blood (group 3 on Figure 3.2) and could be subdivided into those with generally less placental DNA methylation (group 3a on Figure 3.2) and those that were also highly methylated in placenta (group 3b on Figure 3.2). A small number of assays demonstrated extremely low DNA methylation in both males and females (group 2 on Figure 3.2). As shown in the pie charts on Figure 3.2, LC assays were found in each group. However, the majority of assays in LC regions were located within group 3 while the promoters displaying MeXIP were generally in HC and IC regions. A complete list of the HC and IC genes that do not display MeXIP is in Table 3.2. The majority of these exceptions can be explained by expression pattern or proximity to repetitive elements; however, there are some CpG island promoter assays with no obvious reason for deviation from the MeXIP pattern.

For regions that display MeXIP, the DNA methylation provides a means to examine the distance over which promoter DNA methylation is correlated with inactivation status. Plotting DNA methylation levels against the distance from the TSS for each promoter CpG density class, sex and tissue (Figure 3.3) demonstrated that a majority of HC assays outside the -700 to +200 bp promoter window had nearly complete DNA methylation in both tissues [45]. The HC and IC regions examined were not significantly different from each other in either sex or tissue and thus, both will subsequently be referred to as CpG islands. Non-island (LC) assays showed no relationship between DNA methylation and the distance from the TSS.

3.3.3 DNA methylation is female-specific at X-linked island promoters but is consistently high at intragenic and intergenic regions

The Illumina GoldenGate panel only provided data on CpGs in promoters, so pyrosequencing was used to confirm the level of DNA methylation at promoters of all three CpG densities, as well as to determine DNA methylation at intragenic and intergenic regions (10 assays each). Another advantage to pyrosequencing is that a larger number of CpGs within a small region can be examined. The set of 30 assays analyzed here examined 5076 CpG sites for 6 blood samples and 6 placenta samples (two sites sampled per placenta). There was more variation in DNA methylation levels between placental samples and between sites of the same placenta than for blood samples across all males and females. An individual CpG in a region was, on average, only 7% different from the average of all CpGs assayed from that region, with only one assay (rs1212068) showing an average difference of more than 12% due to a single outlier CpG. Therefore, the average percent DNA methylation for each assay was compared (see Figure 3.4) for the six male and six female blood as well as all placenta samples. Assays were subdivided into panels for location in promoters, intragenic, or intergenic regions and ordered according to the CpG density (HC/IC/LC) of the region. Consistent with the Illumina GoldenGate promoter DNA methylation data, female placenta showed an average of 16% less DNA methylation compared with female blood, whereas male placenta showed an average of only 4% less DNA methylation compared with male blood.

Ten promoter regions of varying CpG density were examined by pyrosequencing to determine if they followed MeXIP (Figure 3.4A). One assay, NDP, stood out as an outlier with low DNA methylation (below 20%) in all samples despite carriers manifesting Norrie's disease with X chromosome rearrangements suggesting that NDP undergoes X inactivation [291]. For other assays the DNA methylation patterns of the island promoter assays were very similar, with the highest DNA methylation being detected in female blood which averaged 38%, followed by female placenta which averaged 21%. Male blood and placenta generally showed low DNA methylation averaging 10%. One of the three non-island promoters examined (ILRAPL1)

showed DNA methylation levels comparable to island promoters in blood, while the others showed higher DNA methylation levels than the island promoter assays in all samples. Therefore, in both the Illumina GoldenGate panel and the pyrosequencing data, the HC and IC promoter assays examined showed MeXIP as did a subset of LC promoters.

Intragenic regions analyzed included both exons and introns, which showed similar patterns of DNA methylation. For LC intragenic and intergenic regions there was an average of 80% and 70% DNA methylation respectively, across all assays with less difference observed between sexes for blood than for placenta. While five of the intragenic and intergenic island assays showed a similar high DNA methylation, others (ARHGAP6, BHLHB9, CpG145, CpG36 and CpG70) were much more reminiscent of the MeXIP pattern of DNA methylation of the promoter assays. Examination of the histone modifications in the regions of CpG36 and BHLHB9, using the histone modification tracks of UCSC Genome Browser, showed that they possess histone modification pattern reminiscent of a promoter [48, 135, 290]. Thus, the MeXIP pattern may provide an additional approach to determine the location of unannotated promoters on the X chromosome.

Two assays, ARHGAP6 and BHLHB9, showed hypermethylation of the placenta compared with blood in males and females and are located within two kb of an alternative promoter for an isoform of their respective genes and thus may reflect tissue-specific DNA methylation. One male blood sample showed a DNA methylation level very different from the others for ARHGAP6 at all five CpG sites in the assay, despite sex-normal DNA methylation at all other loci examined. This could potentially reflect allele-specific DNA methylation in this individual [292]. Overall, with the exception of likely unannotated promoters, intragenic and intergenic assays were heavily methylated independent of CpG density. Less DNA methylation was still observed in the placenta, particularly in female samples, suggesting that the X_i shows more placental hypomethylation than the X_a .

3.3.4 Placental hypomethylation is greater on the X_i than the X_a

To distinguish how the X_a and X_i differed in DNA methylation the assumption was made that males and females would have equivalent DNA methylation on their respective X_a s and, therefore, the DNA methylation of the male X was used as the value for the X_a in both males and females. The X_i DNA methylation level was then calculated as described in section 3.2.5. The difference in DNA methylation levels between blood and placenta for each pyrosequencing assay is shown for the X_a and the X_i in Figure 3.5 and summarized according to location and CpG density in Table 3.3. The assays located in promoters had, on average, approximately 3% less DNA methylation on the X_a in placenta compared with blood and a significant (p -value<0.05) decrease in DNA methylation on the X_i of 35%. The island-containing promoters

had limited X_a DNA methylation and thus would not be anticipated to differ in placenta. Therefore, we can assume the difference in female placenta is due to the X_i . The non-island promoters, as well as the intragenic and intergenic regions, however, showed equivalent X_a and X_i DNA methylation in blood. In contrast, the X_i showed two-fold less DNA methylation in placenta when compared with blood, relative to the X_a . In placenta, the X_a did appear to show a slight decrease in DNA methylation compared to blood, however, this was not significant whereas the decrease in DNA methylation from the X_i was significant in all regions (p -value <0.01). The DNA methylation difference was also greatest from females overall for X-linked assays on the Illumina GoldenGate panel (Figure 3.6:), further supporting that the majority of the DNA methylation decrease observed in placenta is from differences in DNA methylation of the X_i .

3.4 Discussion

In this study, a comparison of autosomal and X-linked promoters showed that DNA methylation of promoters in placenta was strikingly reduced only on the X chromosome, particularly in females, with the placenta showing an average of 27% less promoter DNA methylation than blood in females and 8% less in males. The greater placental hypomethylation observed in females implicates an X_i -specific DNA methylation decrease. To calculate X_i DNA methylation, we make the assumption that the male and female X_a are equivalently methylated, a common assumption in the study of X inactivation. For DNA methylation, this assumption has been supported by studies of the X_i isolated in mouse/human somatic cell hybrids, or distinguished by SNPs in clonal female population of cells; however, differences at individual loci could arise due to hormonal differences or the different gene content of the sex chromosomes [79, 293].

In general, IC regions are not classified as CpG islands, however, in this study HC and IC regions were not significantly different in their DNA methylation levels suggesting that, functionally, IC regions on the X chromosome behave as CpG islands. For promoters demonstrating MeXIP, we propose that the low level of DNA methylation seen in males (averaging 8%) is not biologically relevant in preventing expression from the X_a , whereas the higher level observed in females reflects the additional DNA methylation on the X_i which is associated with gene silencing. DNA methylation levels varied between promoters and were also dependent on the assay technology. Nonetheless, consistent with a recent study, it is clear that for individual gene promoters, silencing on the X_i can be maintained with substantially less than 100% DNA methylation [294].

Figure 3.7 summarizes the changes in placental X-linked DNA methylation and illustrates that not only is placenta less methylated than blood but the majority of this difference is due to the X_i . The X_i appeared to show the greatest decrease in DNA methylation at promoters, however,

the limited number of assays and the combination of two DNA methylation detection techniques precludes a definitive conclusion as to the degree of X_i placental hypomethylation between regions. MeXIP is the clear pattern for X-linked CpG island promoters, as well as a subset of non-island promoters. This trend is maintained for the intragenic and intergenic regions for CpG islands, likely due to the presence of unannotated gene promoters. The exclusion of possible unannotated promoters (regions demonstrating MeXIP) resulted in higher intragenic and intergenic DNA methylation on the X_a than the X_i for both blood and placenta regardless of CpG density.

The observation of MeXIP at CpG island promoters is consistent with previous array-based studies that have shown an inverse correlation between DNA methylation and expression at island promoters [45]. The regulatory nature of the promoter DNA methylation has been demonstrated by the removal of DNA methylation through 5-azacytidine treatment resulting in the reactivation of genes on the X chromosome [295]. In contrast, it has been suggested that a consequence of transcription may be subsequent gene-body DNA methylation, a finding supported by higher intragenic X_a DNA methylation detected by array based technologies [79, 80, 296]. If gene-body DNA methylation were reflecting transcription then the DNA methylation patterns of intragenic and intergenic regions should be very different. We observe, however, a similar reduction in DNA methylation in intergenic regions on the X_i compared with the X_a , consistent with the observation by Hellman and Chess that 5 of the 17 most consistently X_a methylated SNPs examined were located outside of the gene bodies [79]. Therefore, the relative $X_a:X_i$ hypermethylation cannot be attributed solely to the transcription of currently annotated genes. Hansen has previously proposed that the X_a and X_i are methylated by different *de novo* methyltransferases based on the hypomethylation of L1 elements on the X_i but not the X_a or autosomes in ICF syndrome cells [89, 194]. While ICF syndrome cells, which have a mutation in *DMNT3B*, also show hypomethylation of X-linked promoters and several classes of satellite elements, no reduction in DNA methyltransferases has been observed in the placenta [51, 52, 297].

CpG islands are generally unmethylated, but genome-wide studies have demonstrated tissue-specific hypermethylation of some island promoter regions as well as hypermethylation in cancer [279, 298, 299]. While we did not observe a significant DNA methylation difference in autosomal promoter DNA methylation between blood and placenta, a recent study comparing blood and placenta indicates that many regions show tissue-specific differences on chromosomes 13, 18 and 21, with hypomethylation being more common than hypermethylation [252]. In a minority of our X-linked assays (4 pyrosequencing assays and 14 Illumina GoldenGate assays out of a total of 114) the level of DNA methylation was significantly (p -

value<0.05) higher in placenta than in blood for at least one sex, perhaps reflecting sex and tissue-specific gene silencing on the X. Little is known about the DNA methylation at tissue-specific island promoters on the X except for *AR* which has been shown to maintain MeXIP in a variety of tissues, forming the basis of a commonly used DNA methylation based X inactivation skewing assay [300].

Here we have shown that reduced DNA methylation in the human placenta is not consistent across the genome; rather, it occurs in repetitive elements and across the X chromosome. The X_a consistently showed less decrease in placental DNA methylation than the X_i , even when the X_a was equivalently or more methylated than the X_i . This suggests that the facultative heterochromatin of the X_i behaves similarly to repetitive elements in the placenta. As previously reported, we observe that X-linked island promoters are methylated in females, not males, a pattern which we have termed MeXIP. Intriguingly, MeXIP is also seen for 20% of non-island promoters as well as 50% of non-promoter islands, which we attribute to unannotated promoters. Outside of promoters, the X_a is slightly more methylated than the X_i in both intragenic and intergenic regions, with intragenic regions tending to be more methylated than intergenic regions. Further study is required to determine the contribution of transcription or other processes to the establishment and/or maintenance of such DNA methylation patterns.

Table 3.1: Primer sequences and cycling conditions used in pyrosequencing assays.

PCR product size, assay class and CpG class of each assay also listed.

Name ^a	Sequence (5' to 3')	Annealing Temperature	PCR product size (bp)	Assay Class	CpG class
PDK3_78_F1 PDK3_78_R1_B PDK3_78_S1	GGTTGTAAAATTTAAGTGTTAGGA /Biotin/AACCCAACCCAACAAATACAA AAAATTTAAGTGTTAGGATG	57°C	211	Promoter	HC
UXT_89_F1 UXT_89_R1B UXT_89_S1	GTTAATGGGGGATTGTAAAAAG /Biotin/TCACTTCCTCTACCTCCACCTAT ATGGGGGATTGTAAAA	57°C	130	Promoter	HC
FANCB_93_F1B FANCB_93_R1 FANCB_93_S1	/Biotin/TTTGGGGAGTGTTGTGAAAGTA AACCAAACCTCAACCTAAATC CCTCAACCTAAATCCCAT	57°C	167	Promoter	HC
PGRMC1_95_F1 PGRMC1_95_R1_B PGRMC1_95_S1	GGGGAAGGGTTATTAAGGAGAG /Biotin/CCCATTCTAAAACCCCTCATCT GGGAAGGGTTATTAAGGA	57°C	164	Promoter	HC
CpG186_89_F1_B CpG186_89_R1 CpG186_89_S1	/Biotin/TGTAGTTTGGATATTTTGTATGGG AACCAATCCTTACCTTACAACCT TCCTTACCTTACAACCTTT	57°C	220	Promoter	HC
NDP_F1 NDP_R1 NDP_S1	AGAGAGAGAATGTTAAATGGAAAAGTGTTA /Biotin/ATTTAACCTCTTATTAATTCCATAATACCA AGAGAATGTTAAATGGAAAA	57°C	255	Promoter	IC
CHM_F1 CHM_R1 CHM_S1	GTGGGAGATTTGGATATTTTTTGTAT /Biotin/AAATAAAAATCTCCTTTATTCACAAAAC GATAATATTGAAGTAAAATTGTTAG	57°C	111	Promoter	IC
PHEX_F1 PHEX_R1 PHEX_S1	AGTTTTTTAAAGTGTTGGGATTATAGG /Biotin/ACTTCAACAAATTCCCCAAAATAAA AAAGTGTTGGGATTATAGG	57°C	93	Promoter	LC
II1RAPL1_F1 II1RAPL1_R1 II1RAPL1_S1	/Biotin/TTGGGGAGATAGTGATGGG CACACTCTTAATAACCTCCTTTTCATC ATCTCTTCTCTTTAAAACAAAT	55°C	91	Promoter	LC
ODZ1_F1 ODZ1_R1 ODZ1_S1	GTATTAAGGATTAAGTTGGAGGTTGTAGT /Biotin/TTATACTCCTCACCCTTTCAAATCTAAT ATAGTTTTTAAAAATATTTGTATTG	57°C	193	Promoter	LC
ARHGAP6_F1 ARHGAP6_R41 ARHGAP6_S1	/Biotin/ATTTGATTGAAGGTTGAATGAG CCAACCCTAAATTC AATATTTCTT CAATATTTCTTTACCCCA	64.5°C	149	Within Gene	HC
MXA5_E_F1 MXA5_E_R1 MXA5_E_S1	TTTTTTTGATGGAAAGGGTT /Biotin/TCTTCCCTAACAAAAAATAACAAACT TTTTTTTGATGGAAAGG	57°C	90	Within Gene	HC

Name ^a	Sequence (5' to 3')	Annealing Temperature	PCR product size (bp)	Assay Class	CpG class
SHROOM2_F1 SHROOM2_R1 SHROOM2_S1	GGTGGAGAATGTTTTTAATAATTTG /Biotin/CCCCCATTTCCAAATCAA GGTGGAGAATGTTTTTA	53°C	86	Within Gene	IC
BHLHB9_F2 BHLHB9_R2 BHLHB9_S2	/Biotin/GGGGTTTTTTGAGGTAGTTTGGTGT CCCCTCTCAAACCCACCTTAATT TCTCAAACCCACCTTAAT	57°C	102	Within Gene	IC
IRCH2_int_F1 IRCH2_int_R1 IRCH2_int_S1	GAGTAGGAGGTTATTATGAGGAGAA /Biotin/ACTAAACTACTATAACCCCACTATAAAT GGAGGTATTATGAGGAGA	57°C	101	Within Gene	IC
POLA_F3 POLA_R2 POLA_S5	/Biotin/GGGGGGTAGTGTTTTATGTATATTTAAAT ACCACATAAAACCCACACATATAAT ATAAACTAACTTTTCCTATC	57°C	115	Within Gene	LC
GLRA2_Int_F3 GLRA2_Int_R4 GLRA2_Int_S2	/Biotin/GAATTTTTTGATGGATTGGATATGG CCTTCTATTAACCTCCACACTCCTATATCA ATCTCATAACTATCTACATTAACC	57°C	183	Within Gene	LC
GLRA2_E4_F3 GLRA2_E4_R1 GLRA2_E4_S3	TGTAAATAGAATTTTTGTGTTAGGGTAAT /Biotin/ATAGAATTTTTGTGTTAGGG ATAGAATTTTTGTGTTAGGG	57°C	135	Within Gene	LC
WNK5_Int_F1 WNK5_Int_R1 WNK5_Int_S1	/Biotin/TAAAAATTAGTTGGGAGTGGTGGTAGG CTCATTTACATTTTCTCCCTCATCA CCAAATTAATAACAATAACACA	57°C	217	Within Gene	LC
TBL1X_int_F1 TBL1X_int_R1 TBL1X_int_S1	TGTGTTAAGTTTGGATTGTAGAAATGAAT /Biotin/CCCTAAATAATAATCTCAATTTTCCCTATA GTAGAAATGAATTTGAAGAAG	55°C	147	Within Gene	LC
CpG145_89_F1 CpG145_89_R1_B CpG145_89_S1	TTGGATTTGTTTGGTTAGGATTG /Biotin/CAAACCCAACCTACTTCAATAACCT GGATTTGTTTGGTTAGGAT	57°C	182	Between Genes	HC
CpG46_F1 CpG46_R1 CpG46_S1	/Biotin/GGTTTTAGTGGTTTTGATTTTATAGAGT CTCCTCTTACTAAAAACAACCTACC TCCTCTTACTAAAAACAACCT	57°C	108	Between Genes	HC
CpG36_F1 CpG36_R20 CpG36_S18	GGAAAGGAAAAGGGAGAATT /Biotin/CCCTCACCCTAAACAATTA GGAAAGGAAAAGGGAGAATT	57°C	80	Between Genes	HC
rs5960421_1_F1 rs5960421_1_R1 rs5960421_1_S1	GGTTTGTAGAGTGTGGTAGAGG /Biotin/CCCTCCCACCAAAATCAAAT AGTGTGGTAGAGGTGTT	57°C	143	Between Genes	IC
CpG70_F1 CpG70_R1 CpG70_S1	GTTTGAAGTAGGAGGTTTGGATGTA /Biotin/CTAAACTCCTATTTCTCCAATTTATACAAC GAAGTAGGAGGTTTGGAT	55°C	169	Between Genes	IC

Name ^a	Sequence (5' to 3')	Annealing Temperature	PCR product size (bp)	Assay Class	CpG class
rs36021843_F1 rs36021843_R1 rs36021843_S1	/Biotin/ATGGTTGGTTTATATGGTTATTTAGAGTT CCCTAAAAAATAACCTCCTACTTAACTAT AATAATATTCCACCTCCC	57°C	185	Between Genes	LC
rs1212068_1_F1 rs1212068_1_R1 rs1212068_1_S1	/Biotin/TGAGAGATGAGTGTTATGGAGAAA CAAAAAACAAACTCTCCAAATTCA TCTCCAAATTCAAATCAAT	64.5°C	183	Between Genes	LC
rs34350719_F1 rs34350719_R1 rs34350719_S1	GTTTTGGGTTTGGAAAAATTAGAGT /Biotin/CCCATAAAATTCAAAAACTTCTTACCT TGGGTTTGGAAAAATTAG	57°C	79	Between Genes	LC
rs17308229_F1 rs17308229_R1 rs17308229_S1	GGGTTTTTTATTTTTTTGAGATTTGTTAG /Biotin/AACCACTCAAACCTATATCTACAAACAATA TATTTATAAGTTATTGTATTTAGGG	64°C	214	Between Genes	LC
rs4825396_F1 rs4825396_R1 rs4825396_S1	/Biotin/TTTTTTGATGGGGGAGAAGGGT CCCATCCTAATCTTCCTATTTTCTTATCC TTCCTATTTTCTTATCCACA	57°C	113	Between Genes	LC

a) F: Forward primers, R: Reverse primer, S: Sequencing primers

Table 3.2: DNA methylation assays for island (HC and IC) promoters failing to show MeXIP or containing features believed to interfere with MeXIP or showing discordant DNA methylation results.

Assay name	Average DNA methylation				Features interfering with MeXIP ^a				Discordant Assay DNA methylation Patterns ^d	Previous DNA methylation data ^e	Evaluation
	Female Blood	Female Placenta	Male Blood	Male Placenta	CpG Density	Repetitive Element ^b	Escapes X inactivation	Distance to TSS (bp) ^c			
AR_P54_R	0.09	0.04	0.00	0.01	HC	-	0/6	-54	Yes	[160]	The 2 Illumina assays for AR are discordant and P189 does not match previous DNA methylation results.
AR_P189_R	0.45	0.33	0.07	0.10	HC	-	0/6	-189			
BCAP31_P1072_F	0.99	0.99	0.99	0.99	HC	-	3/9	-1074	-	[161]	Also near a duplication on chr16 ^f .
BCAP31_P1131_F	0.99	0.99	0.99	0.99	HC	-	3/9	-1133			
BGN_E282_R	0.72	0.59	0.20	0.13	IC	-	0/5	282	Yes	-	The 2 Illumina assays for BGN are discordant, however, the other assay is an LC.
CTAG1B_P4_R	1.00	0.90	0.99	0.92	IC	-	-	-4	-	[162, 165]	CT gene family highly methylated, additionally many other CTs on the array were LCs ^g .
CTAG1B_P77_F	1.00	0.99	1.00	0.99	IC	-	-	-77			
DLG3_E340_F	0.84	0.67	0.02	0.27	IC	-	-	340	-	-	Shows MeXIP despite distance to TSS.
FHL1_E229_R	0.46	0.34	0.00	0.05	HC	-	1/9	229	-	-	Shows MeXIP despite distance to TSS and/or presence of a LINE.
FHL1_P768_F	0.50	0.39	0.04	0.26	IC	LINE	1/9	-768			
FMR1_P484_R	0.97	0.60	0.68	0.57	IC	-	1/9	-484	Yes	[163, 164, 166]	The 2 Illumina assays for FMR1 are discordant and P484 does not match previous DNA methylation results.
FMR1_P62_R	0.69	0.08	0.01	0.01	HC	-	1/9	-62			
G6PD_E190_F	0.74	0.15	0.01	0.04	HC	-	0/5	-783	-	[282, 283]	Shows MeXIP despite distance to TSS and/or presence of a LINE.
G6PD_P1065_R	0.87	0.37	0.13	0.07	IC	LINE	0/5	472			
L1CAM_P19_F	0.62	0.56	0.30	0.18	IC	-	-	-19	Yes	-	The 2 Illumina assays for L1CAM are discordant, however, the other assay is an LC.
MKRN4_E249_R	0.99	0.31	0.99	0.15	HC	-	-	249	Yes	-	Pseudogene
MKRN4_P1320_R	0.99	0.90	0.99	0.98	HC	-	-	-1320	Yes		
PCTK1_E77_R	0.72	0.54	0.81	0.66	HC	-	6/6	77	-	[45]	Assay does not match previous DNA methylation results.
PLS3_E70_F	0.29	0.61	0.11	0.11	IC	-	5/9	70	-	[46]	Previous DNA methylation results only examined DNA methylation in males.
PLS3_P94_R	0.30	0.49	0.26	0.20	IC	-	5/9	-94			
SLC6A8_P193_R	0.86	0.17	0.84	0.10	IC	-	-	-193	Yes	[161]	The 3 Illumina assays for SLC6A8 are discordant, however, the other assay is an LC. Previous DNA methylation results only examined DNA methylation in males and there is also has a pseudogene on chr 16 ^f .
SLC6A8_seq_28_S227_F	0.73	0.63	0.00	0.09	HC	-	-	-1681			

Assay name	Average DNA methylation				Features interfering with MeXIP ^a				Discordant Assay DNA methylation Patterns ^d	Previous DNA methylation data ^e	Evaluation
	Female Blood	Female Placenta	Male Blood	Male Placenta	CpG Density	Repetitive Element ^b	Escapes X inactivation	Distance to TSS (bp) ^c			
STK23_E182_R	0.99	0.72	0.87	0.26	HC	-	-	182	Yes	[161]	The 2 Illumina assays for STK are discordant and Previous DNA methylation results only examined DNA methylation in males.
STK23_P24_F	0.87	0.63	0.45	0.21	HC	-	-	-24			
SYBL1_E23_R	0.75	0.21	0.72	0.07	HC	-	0/5	23	-	[48, 282]	Silent on both X and Y chromosomes. P349 shows MeXIP despite presence of a LINE.
SYBL1_P349_F	0.50	0.11	0.57	0.01	IC	LINE	0/5	-349	-		
Xist_seq_80_S47_R	0.85	0.83	0.97	0.97	IC	-	-	-31	Yes	[90, 91, 272]	The 2 Illumina assays for XIST are discordant, however, the other assay is an LC. Is also expressed only from the X _i ^h .

a) Grey shading represents possible features which may interfere with MeXIP.

b) Three LC assays (CTAG2_P1426_F, MAGEC3_P903_F, TIMP1_P615_R) were also located within repetitive elements

c) Seven LC assays (CDM_seq_21_S260_R, CTAG2_P1426_F, MAGEA1_P926_F, MAGEC3_E307_F, MAGEC3_P903_F, MCF2_P1024_R, TIMP1_E254_R) were also beyond 700 bp upstream or 200 bp downstream

d) Grey shading indicates a gene with multiple Illumina assays which show different DNA methylation patterns. Three genes (BTK, MCF2, TIMP1) had only LC assays and also were discordant between assays within the same gene.

e) Grey shading indicates that Illumina DNA methylation results conflict with previous DNA methylation results.

f) Recent genome-wide studies suggest a hypermethylation of pseudogenes and duplicated regions thus it is possible that the presence of a tandem duplication or pseudogene may predispose genes to hypermethylation which may explain the high DNA methylation seen for BCAP and SLC6A8 [92, 161].

g) Members of CT antigen family of genes are often found in palindromic repeats as multicopy genes and pseudogenes and have typically been shown to be highly methylated in all tissues except the germline – a pattern generally found for genes with germline-specific expression [161, 162, 301]. Consistent with high levels of DNA methylation in all tissues other than testis, all MAGEs and CTAGs showed hypermethylation in blood and placenta regardless of CpG density emphasizing that gene function as well as CpG density is important in determining DNA methylation status [161, 162].

h) Both XIST assays on the Illumina GoldenGate panel showed nearly 100% methylation in males, however, females showed DNA methylation levels up to 95%. While the trend of these DNA methylation levels was as expected the level of DNA methylation in females appears to have been substantially overestimated by the Illumina assay.

Table 3.3: Average percent DNA methylation as determined by pyrosequencing at promoter, intragenic and intergenic regions across the X chromosome in blood and placenta for high (HC), intermediate (IC) and low (LC) CpG density.

Location and CpG density	Number of assays ^a	Blood				Placenta				Blood-Placenta	
		Male (X _a) ^b	Female (46, X _a X _i)	X _i ^c	X _a :X _i ratio ^d	Male (X _a) ^b	Female (46, X _a X _i)	X _i ^c	X _a :X _i ratio ^d	Δ X _a ^e	Δ X _i ^e
Promoter	9	23%	44%	64%	0.35*	20%	24%	29%	0.68	-3%	-35%**
HC	5	8%	38%	68%	0.12	12%	24%	35%	0.35	4%	-33%
IC ^f	1	12%	37%	62%	0.20	3%	8%	13%	0.19	-10%	-49%
LC	3	50%	55%	59%	0.85	38%	31%	25%	1.52	-13%	-34%
Intragenic	8	89%	86%	83%	1.08*	82%	73%	63%	1.30**	-7%	-19%**
HC ^g	1	91%	90%	90%	1.01	65%	62%	60%	1.09	-26%	-30%
IC ^g	2	92%	91%	90%	1.03	88%	81%	73%	1.21	-4%	-16%
LC	5	88%	83%	78%	1.12	83%	71%	60%	1.39	-5%	-19%
Intergenic	10	58%	62%	65%	0.88	43%	38%	33%	1.29*	-15%	-32%***
HC	3	29%	42%	56%	0.52	11%	18%	25%	0.43	-18%	-31%
IC	2	46%	59%	71%	0.65	32%	36%	41%	0.77	-15%	-30%
LC	5	79%	74%	69%	1.15	66%	50%	34%	1.92	-13%	-34%

a) Number of independent regions assessed as shown on Figure 3.4, excluding the genes noted below (NDP, ARHGAP6 and BHLHB9) and discussed in the text.

b) DNA methylation in males was used as X_a DNA methylation level.

c) X_i DNA methylation calculated assuming that X_a in female is equivalent to X_a in males.

d) X_a and X_i were compared to determine if they differed in blood and placenta. Significance calculated using Mann-Whitney test with significance shown as p-value=0.01 to 0.05 (*) and p-value=0.001 to 0.01 (**).

e) Blood and placenta were compared to determine if the X_a and X_i differed in their tissue specific DNA methylation levels. Significance calculated using Mann-Whitney test with significance shown as p-value=0.001 to 0.01 (**) and p-value<0.001 (***).

f) NDP removed from the average as it showed low DNA methylation in all samples.

g) ARHGAP6 and BHLHB9 removed from the average as both appeared to be an alternative promoter for an isoform of their respective genes.

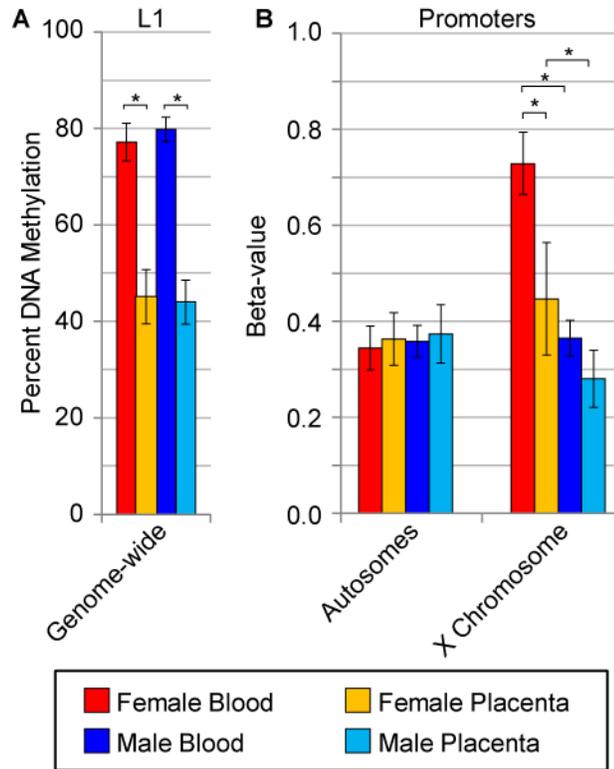


Figure 3.1: Reduced placental DNA methylation found at L1 repetitive elements and promoters on the X chromosome.

Average level of DNA methylation for female blood (red), female placenta (orange), male blood (dark blue) and male placenta (light blue) are shown with error bars (one standard deviation) based on the average sample deviation at a single site. Significance calculated using Mann-Whitney test with p -value < 0.001 (*). **(A)** L1 percent DNA methylation as determined by pyrosequencing at L1 repetitive elements across the genome. **(B)** Illumina GoldenGate Promoter DNA methylation array data averaged separately for 1421 sites on the autosomes and 84 X-linked sites. Beta-values represent average percent DNA methylation.

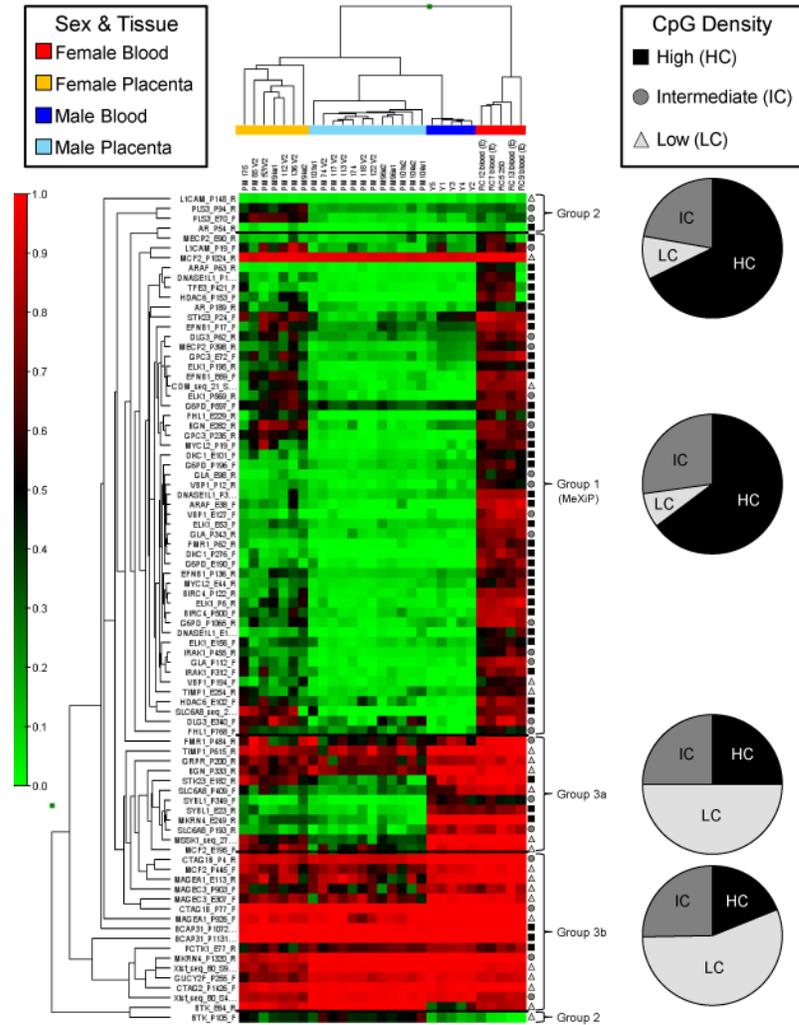


Figure 3.2: Heatmap illustrating DNA methylation levels at 84 sites across the X chromosome from the Illumina GoldenGate promoter DNA methylation array.

The majority of X-linked promoter assays demonstrate MeXIP and are of high and intermediate CpG density while low CpG density assays tend to be highly methylated. DNA methylation levels are represented as a gradient from red (high DNA methylation) to green (low DNA methylation). BeadStudio software used the Manhattan Hierarchical Cluster Metric to group samples which were separated by tissue and sex (coding of samples as follows: yellow = female placenta, blue = male placenta, dark blue = male blood, orange = female blood). Assays were visually divided into 4 groups based on DNA methylation trends. Group 1 shows high female DNA methylation and low male in blood (MeXIP), group 2 shows low DNA methylation in all samples and group 3 shows high DNA methylation in both male and female blood. Group 3a had high and low placenta DNA methylation while group 3b had high DNA methylation in the placenta. The CpG density of each assay, high (HC) (black square), intermediate (IC) (dark grey circle) or low (LC) (light grey triangle), is shown to the right and the assay names to the left of the heatmap. The percent of assays within each group based on CpG density is shown as a pie chart to the far right of the heatmap.

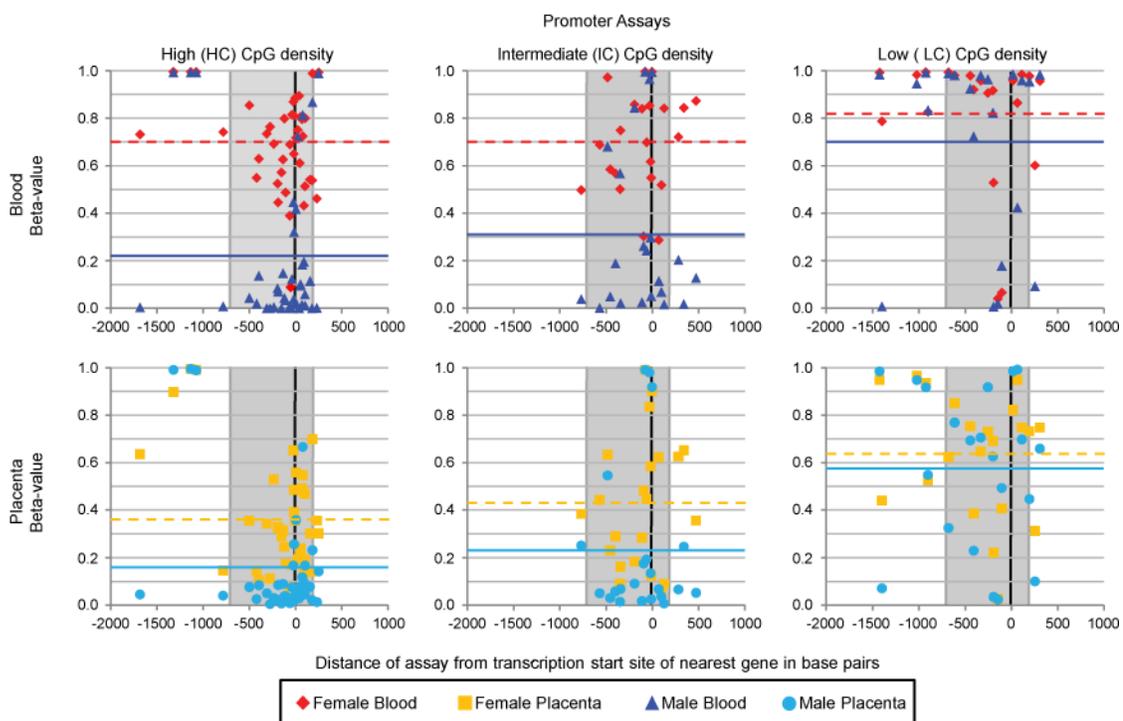


Figure 3.3: Effect of distance from transcription start site on DNA methylation

Beta-values for female blood (red diamond), female placenta (orange square), male blood (dark blue triangle) and male placenta (light blue circle) versus distance from the TSS for each X-linked assay present on Illumina GoldenGate panel. The average beta-value for each sex and tissue is shown as a dashed horizontal line for females and a solid horizontal line for males. Black vertical line marks the TSS (0 bp) and the grey area contains the promoter region as defined by Weber *et al.* (700 bp upstream to 200 bp downstream of the TSS) [45]. Assays were separated based on the CpG density (HC, IC and LC) of the 500 bp around each assay with MeXIP being observed in HC and IC in both blood (upper panels) and placenta (lower panels).

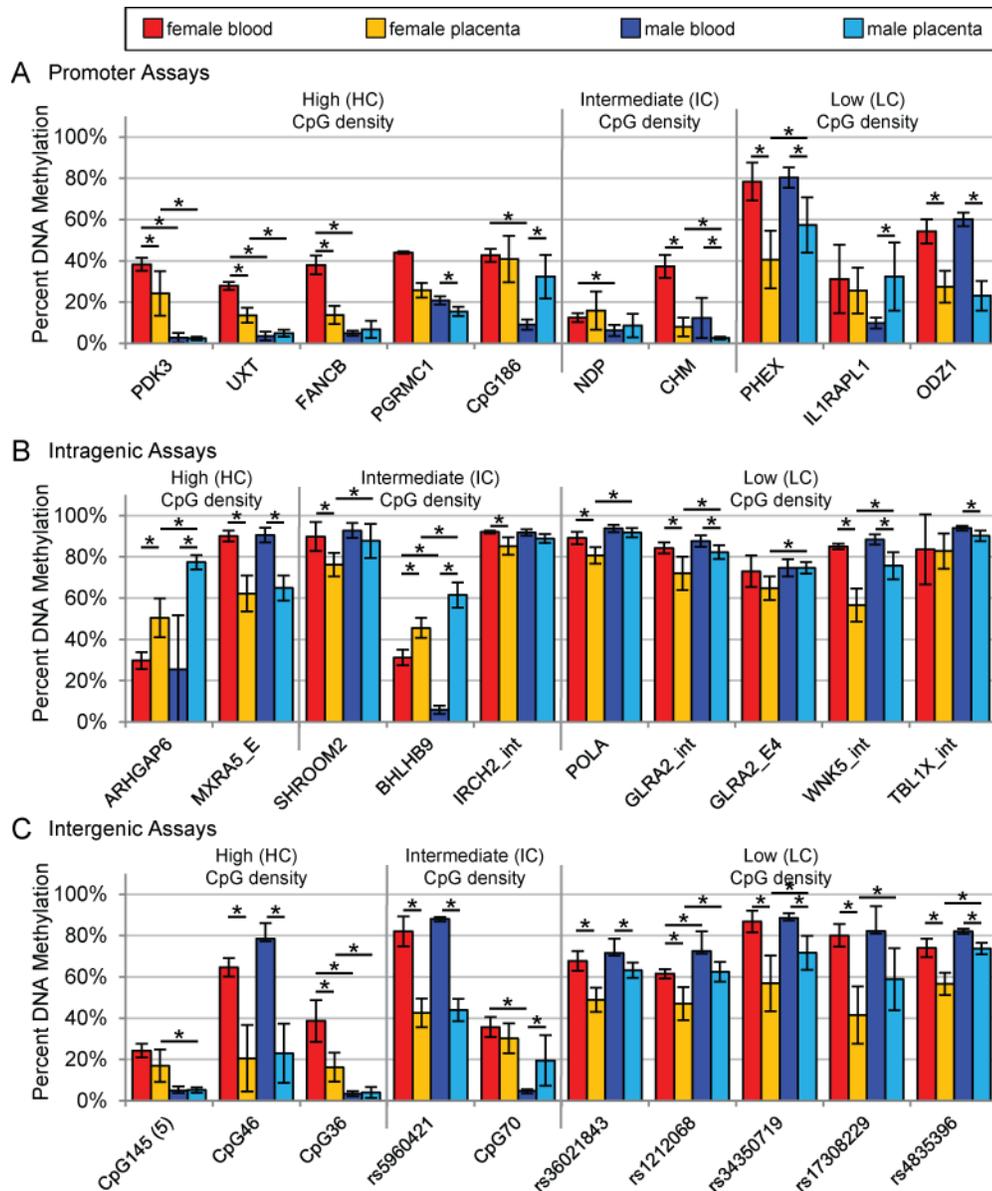


Figure 3.4: X-linked CpG island promoters show female-specific DNA methylation whereas DNA methylation is high in both X-linked intragenic and intergenic regions in females and males.

Average percent DNA methylation from 30 pyrosequencing assays for 6 female blood (red), 6 female placenta (orange), 6 male blood (dark blue) and 6 male placenta (light blue). Each placenta was sampled from two sites within a single placenta for a total of 12 placental samples. Assays are separated into CpG density, high (HC), intermediate (IC) and low (LC), from the left to the right, by vertical lines. (A) Promoter assays (B) intragenic assays (C) intergenic assays. The region assayed is listed below each set of averages. Significance calculated using Mann-Whitney test with p -value <0.01 (*). Error bars are one standard deviation.

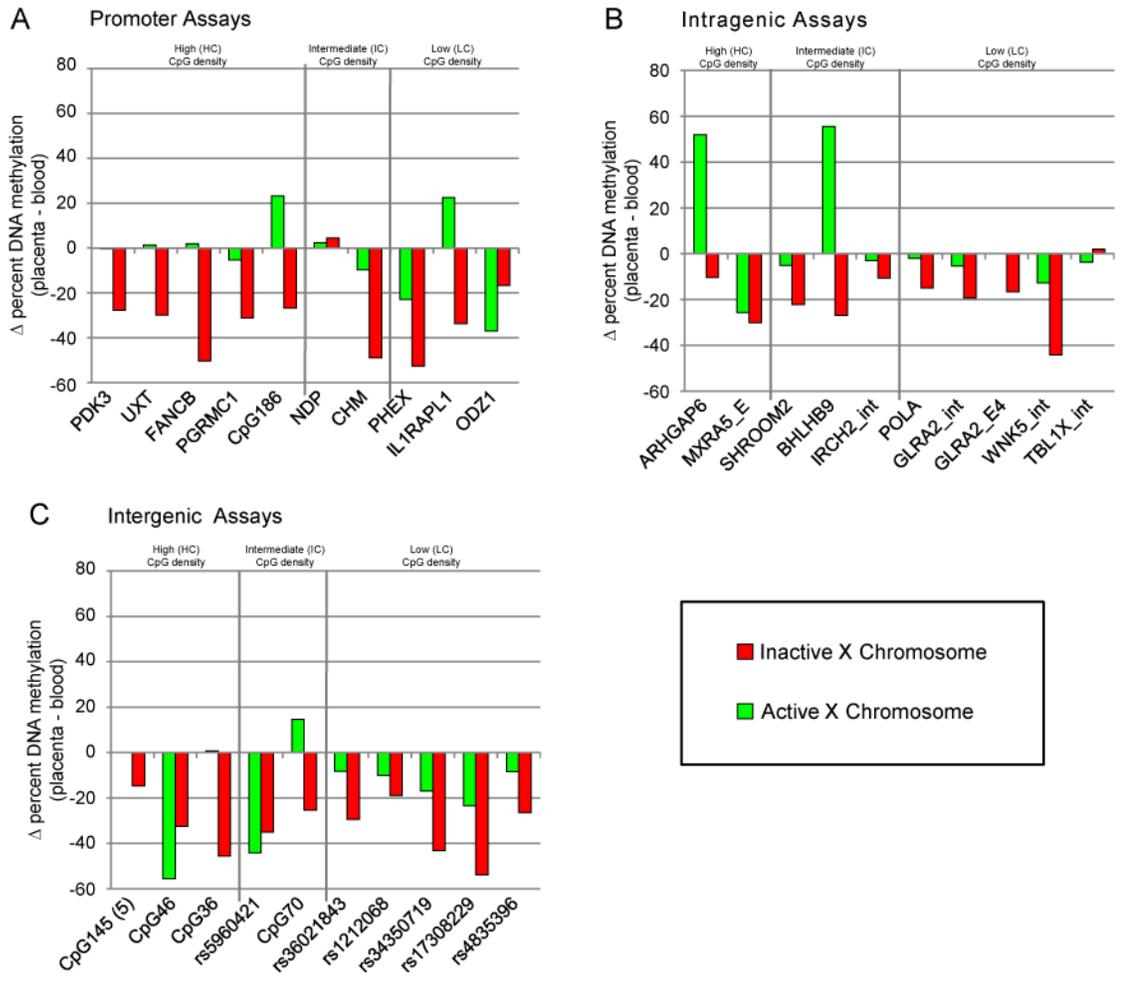


Figure 3.5: The X_i shows less placental DNA methylation compared with blood than the X_a at the majority of regions examined across the X chromosome

Percent DNA methylation change from blood to placenta for X_a (green) and X_i (red) at 30 pyrosequencing assays. Negative percent change DNA methylation indicates that blood is more methylated than placenta while positive shows that placenta is more methylated than blood. Assays are separated into CpG density classes, HC, IC and LC, by vertical lines. (A) promoter assays (B) intragenic assays (C) intergenic assays. X_a value is the level of DNA methylation observed in male, X_i value is calculated by subtracting the X_a from the female DNA methylation level multiplied by two.

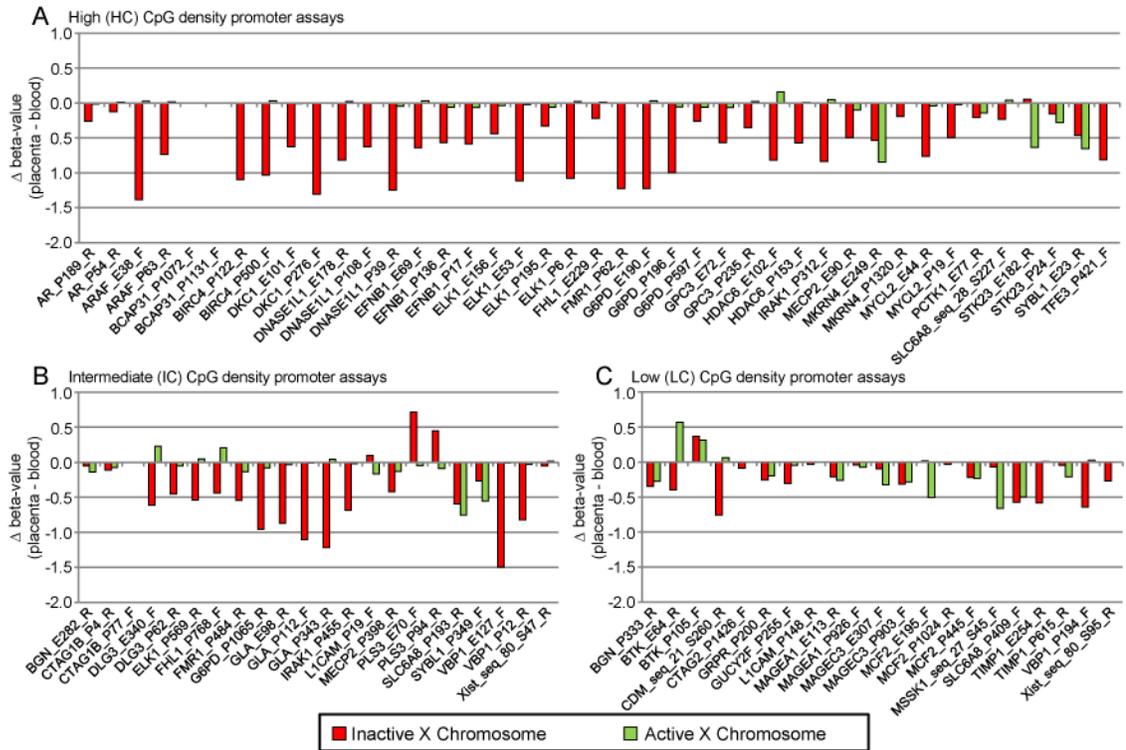


Figure 3.6: The X_i shows less placental DNA methylation compared with blood than the X_a at the majority of promoters examined on the X chromosome.

Beta-value DNA methylation change from blood to placenta for X_a (red) and X_i (green) at Illumina GoldenGate panel assays. Negative percent change DNA methylation indicates that blood is more methylated than placenta while positive shows that placenta is more methylated than blood. Assays are separated into CpG density classes, HC, IC and LC, by vertical lines. (A) promoter assays (B) intragenic assays (C) intergenic assays. X_a value is the level of DNA methylation observed in male, X_i value is calculated by subtracting the X_a from the female DNA methylation level multiplied by two.

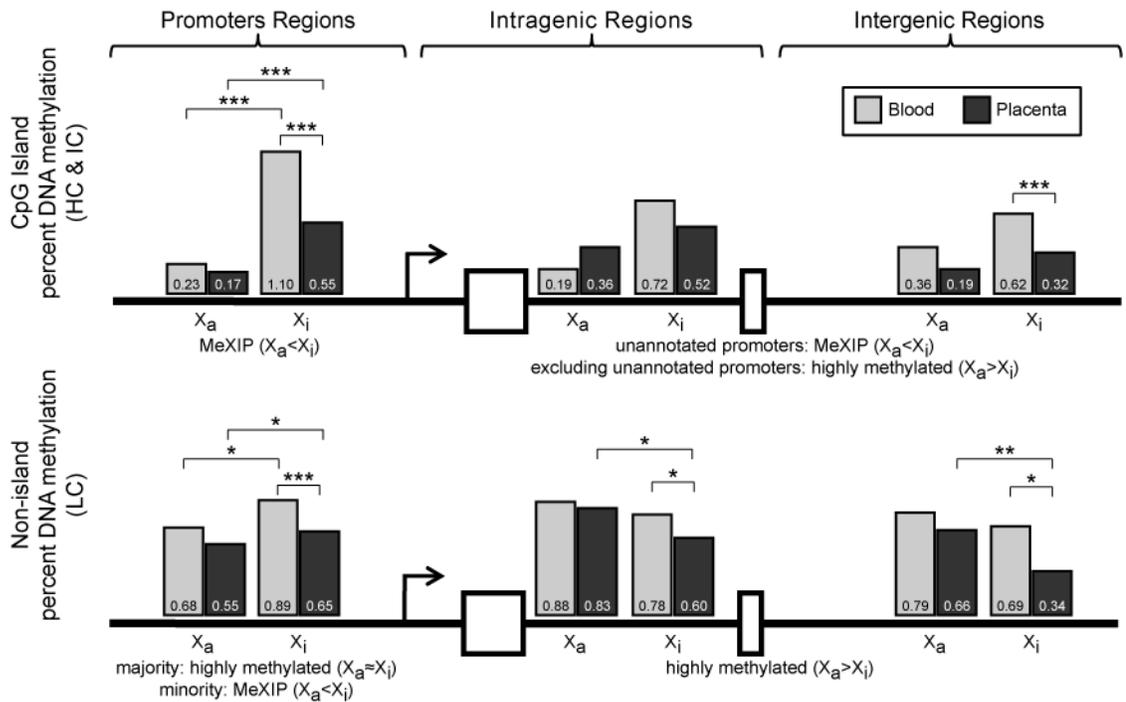


Figure 3.7: Summary of DNA methylation analyses showing placental reduction in DNA methylation predominately on the X_i.

Data from both Illumina and pyrosequencing is combined and shown separately for CpG island (HC and IC) assays and non-island (LC) assays in promoter regions, intragenic regions (includes both introns and exons) and intergenic regions. In order to combine Illumina GoldenGate data (which is only for promoter regions) we converted beta-values to percent DNA methylation. These values were consistent with pyrosequencing data at the low range, but generally higher than pyrosequencing in the midrange, accounting for the X_i value over one for promoters. Percent DNA methylation is the average of all CpGs in the indicated region. Percent DNA methylation is divided into X_a and X_i with grey bars representing DNA methylation in blood and black bars for placenta with the average percent DNA methylation value written in each bar. The summary of the DNA methylation trends for the different regions is described below each bar graph. Significance calculated using Mann-Whitney test with significance shown as p-value=0.01 to 0.05 (*), p-value=0.01 to 0.001 (**), and p-value<0.001 (***).

4 X-linked DNA methylation predicts increased escape from XCI in human placenta

4.1 Introduction

In chapter 3 it was established that placental hypomethylation occurs predominately at L1s and CpG island promoters on the X chromosome. This chapter aims to further investigate placental DNA methylation by looking at nine times the number of X-linked probes. The role that DNA methylation plays in maintaining the epigenetic stability of the placenta is unclear, however, the hypomethylation of different placental regions may influence the ability to form the transcriptionally silenced domain of the X_i . Previous work demonstrated that it is not necessary that a CpG island promoter be 100% methylated in order to silence a gene [294], which leads to the question of how hypomethylated a gene needs to be to become and or to stay transcriptionally active. An overall trend in which higher DNA methylation corresponds to lower expression has been observed across placentas of varying gestational age [2] and in somatic tissues the DNA methylation of CpG island promoters corresponds to transcriptional silencing [61, 62]. Indeed, some normally silent repetitive elements, imprinted genes and individual X-linked genes are expressed in placenta [151, 277, 302, 303]. Expression of both alleles of the X-linked gene *G6PD* was observed in a subset of placental cells and when chorionic villi cells were used to create somatic cell hybrids. This is an example of global X chromosome reactivation, a process which normally occurs only during oogenesis [151, 152, 304]. As there is not a general over-expression of placental X-linked genes compared with other somatic tissues, the synergistic silencing action of chromatin changes and the non-coding XIST RNA are apparently sufficient to maintain some placental X inactivation [151, 277, 303].

The Cot-1 DNA fraction of the genome represents highly repetitive sequences [23, 305] and traditionally has been used to block non-specific binding and cross-hybridization due to repetitive elements [306]. The lack of Cot-1 RNA transcription under the X_i was first observed by Hall *et al.* and was initially examined as a means to evaluate the overall transcriptional status of the X_i rather than the transcriptional status of individual genes [25]. Mice were used to examine Cot-1 RNA transcription early in development. A Cot-1 hole could be detected under the Xist RNA signal by the two cell stage [307]. However, if the placenta shows hypomethylation of repetitive elements and DNA methylation is important in maintaining the transcriptional silencing of repetitive elements, then one might expect to see a lack of a Cot-1 hole in placental cells resulting in possible changes in the structure of the X_i . The structure that the X_i takes in the nucleus has long been a topic of investigation. In 1996, it was found that the Barr body is approximately 20% smaller than both the XIST RNA domain and the area covered by X chromosome paint [16]. X-linked genes, regardless of XCI status, have been found to associate with the boundary of the Barr body whereas Cot-1 DNA is strongest within the Barr

body. The strongest Cot-1 holes (the weakest Cot-1 RNA hybridization signal) have been found to colocalize with the Barr body as well as with the area with most Cot-1 DNA [29].

To further investigate the DNA methylation status of X-linked loci in the human placenta, 777 X-linked sites were examined compared to the 84 sites in chapter 3. Having previously established (chapter 2, [273, 308]) the ability of DNA methylation at X-linked CpG island promoters to predict the XCI status of genes in somatic tissues, the same process was applied to the placenta; four times as many genes were predicted to escape from XCI in placenta (41%) compared to blood (10%). Included in this chapter are 88 X-linked probes previously excluded from chapter 2 as they were located in repetitive elements. Through the analysis of these probes, X-linked and autosomal repetitive elements were found to be hypomethylated in the placenta compared to blood. Placental repetitive element hypomethylation did not result in any detectable changes in the presence of Cot-1 holes, XIST RNA localization or the presence of the Barr body.

4.2 Methods

4.2.1 Sample collection, DNA extraction and Illumina Infinium

HumanMethylation27 array

Collection of samples was approved by the ethics committees of the University of British Columbia and the Children's and Women's Health Centre of British Columbia with recruitment, collection and preparation performed by the Robinson lab (H06-70085). DNA extraction was performed by the Robinson lab as outlined in chapter 2 with details on placental samples given in Table 4.1. Bisulfite conversions and array processing was performed by the Robinson lab as outlined in chapter 2. The decision tree used to predict XCI in placenta was the same as previously outlined in section 2.2.6.

4.2.2 CpG density definitions

CpG density classifications were based on those used by Weber *et al.* [45] and are outlined in section 2.2.3.

4.2.3 Statistical analysis

Statistical analysis of the Illumina Infinium HumanMethylation27 array was performed using the Mann-Whitney or the Kruskal-Wallis test as calculated by Graphpad Prism. When comparisons resulted in p-values greater than 0.05 these differences were not significant; when p-values were between 0.01 and 0.05 (*), between 0.01 and 0.001 (**), and <0.001 (***) the comparisons were considered significantly different.

4.2.4 RNA FISH

RNA FISH was performed as previously outlined [25]. Briefly, cells (somatic n=300 and placental n=817) were grown as a monolayer on glass coverslips then permeabilized with 0.5% Triton X-100 and fixed with 4% paraformaldehyde. Two different probes were combined to allow detection of XIST RNA and Cot-1 RNA. The probe XIST G1A is approximately 10kb long and represents the genomic DNA from the fourth intron to 3' end of the *XIST* gene. Human Cot-1 DNA (Invitrogen, #15279-011) was used as the probe to detect Cot-1 RNA. Both probes were separately nick translated using the Nick Translation Reagent Kit (Abbott Molecular Inc) with Spectrum red-UTP (Vysis) for Cot-1 DNA probes and Spectrum green-UTP (Vysis) for XIST G1A probes. To perform the RNA FISH 5 µg of each probe were combined along with salmon sperm and tRNA. These reagents were dried then resuspended in formamide (Sigma) and denatured at 80°C for 10 minutes. The denatured probe mixture was then combined 1:1 with a RNA hybridization buffer, applied to dried coverslips and hybridized overnight at 37°C. The next day cells underwent two washes: 50% formamide and 50% 4X SSC at 37°C for 30 minutes, and 2X SSC at 37°C for 30 minutes. Coverslips were then briefly incubated in 4'-6-Diamidino-2-phenylindole (DAPI) and washed in 1X PBS (phosphate buffered saline) before being mounted onto slides with Vectashield (Vector labs).

4.2.5 Post RNA FISH imaging

A Retiga 4000R (Q-Imaging) camera on a Leica inverted microscope (DMI 6000B) was used with OpenLab software (PerkinElmer) to take images of all cells. All colour channels on all images were then independently processed using Adobe Illustrator (Adobe) in order to correct for inter-cellular variations in intensity caused by variations in FISH efficiency. A comparison of the same images adjusted by two independent users demonstrated no user adjustment bias (data not shown). Cells were defined as Cot-1 hole positive or negative based on the Colocalization Highlighter tool (Cot-1 RNA as red channel with threshold=50, XIST RNA as green channel with threshold=0 and a ratio=50%) in the ImageJ software [309]. The nuclear location was determined by eye based on the XIST RNA signal and was done at a different time from the Cot-1 hole identification as to not bias the data. ImageJ [309] was used to define the nuclear and the XIST areas based on thresholds of 1,255 and 20,255 respectively. Adjusting all cells images in Adobe Illustrator (Adobe) allowed for the same nuclear area and XIST area thresholds to be used in all cells. However, the area occupied by the Barr body was set using a different threshold in each cell. This threshold was selected separately for each cell due to the high levels of DAPI variation observed across the nucleus. In all cases the Versatile Wand tool (not including holes and with 8-connected connectedness) was used to select areas to be measured.

4.3 Results

4.3.1 Placenta gestational age does not significantly affect X-linked promoter DNA methylation

The Illumina Infinium HumanMethylation27 array was used to investigate placental hypomethylation at a higher resolution than was performed in chapter 3. Probes located in repetitive elements, those which mapped to the autosomes as well as those in CT genes, were removed from analysis as outlined in section 2.2.4. Placental samples in this section represent whole villi samples which are composed of variety of cell types including trophoblast cells arising from the trophectoderm [310]. Male and female placentas were grouped based on gestational age (first, second or third trimester) and then compared to determine if X-linked promoter DNA methylation differed across gestational ages. As expected, at each gestational age, the average female placental DNA methylation was significantly ($p\text{-value}<0.001$) higher than the average male placental DNA methylation (Figure 4.1A). Within each sex, the only significant difference in DNA methylation across gestational age was found between second and third trimester male placentas ($p\text{-value}= 0.05$). As this was only a 1% difference in average DNA methylation, we concluded that gestational age did not have a substantial effect on average X-linked promoter DNA methylation and placentas of all gestational ages were combined into one category for further analysis. Separation of X-linked promoters based on CpG density confirmed that DNA methylation at CpG island promoters (HC and ICs) was significantly different ($p\text{-value}<0.001$) between male and female placentas but non-island promoter DNA methylation was not (Figure 4.1B).

4.3.2 Placental hypomethylation is found at non-island promoters on both the X chromosome and the autosomes

A comparison of promoter DNA methylation in blood and placenta using the Illumina GoldenGate panel had previously shown that placental hypomethylation was greatest at L1s and female X-linked promoters (Figure 3.1). Because of the larger number of probes on the Illumina Infinium HumanMethylation27 array, the effect of CpG density on promoter DNA methylation could now be examined on both the X chromosome and the autosomes. Only autosomal DNA methylation from chromosomes 20, 21 and 22 was examined to allow comparisons of autosomal DNA methylation between this study and that in chapter 2. Using the Illumina Infinium HumanMethylation27 array, DNA methylation at X-linked probes of all CpG densities were analyzed together. X-linked female probes showed the largest decrease in placental DNA methylation compared to blood (placenta 17% less methylated) and male X-linked and male and female autosomal probes showed between a 0% and 4% decrease in placental DNA methylation. However, when probes were separated based on CpG density,

substantial differences in placental hypomethylation were observed. At X-linked CpG islands, the female placenta was significantly hypomethylated compared to blood (blood-placenta=16%, p-value<0.001). Male X-linked CpG islands promoters were significantly hypermethylated in placenta compared to blood (blood-placenta=-6%, p-value<0.001). Both males and females showed similar significant (p-value<0.001) levels of placental hypomethylation at X-linked non-island promoters (Figure 4.2A).

Unexpectedly, placental DNA methylation at autosomal probes was also affected by CpG density. In both males and females, CpG island promoters on the autosomes were significantly (p-value<0.001) hypermethylated in the placenta compared to blood. This difference in DNA methylation ranged between 6% and 7%, similar to the hypermethylation observed at male X-linked CpG island promoters. Autosomal non-island promoters were significantly (p-value<0.001) hypomethylated in the placenta compared to blood (blood-placenta female: 27% and male: 30%), however, the degree of hypomethylation was between 1.4 and 2.0 times greater than that observed on X-linked non-island promoters (blood-placenta female: 19% and male:15%) (Figure 4.2B). The only significant difference (p-value<0.001) in DNA methylation between males and females in blood or placenta was found at X-linked CpG island promoters (Figure 4.2A). In agreement with previous analysis (chapter 3, [273]), CpG island promoters on the female X chromosome were more hypomethylated in placenta compared with blood than on the autosomes. In addition, non-island promoters on both the X chromosome and the autosomes had placental hypomethylation compared with blood in both males and females.

4.3.3 X-linked CpG island DNA methylation predicts increased escape from XCI in the placenta

Placental DNA methylation was next used to predict XCI status, with the expectation that the hypomethylation of X-linked CpG island promoters in females would correspond with the frequency of genes subject XCI and escaping from XCI. The same decision tree (Figure 2.1) as was used in chapter 2 to predict the XCI status across various somatic tissues was applied to 25 placental samples (male: n=15, female: n=10). In blood, 81% of genes were predicted to be subject to XCI while in placenta only 8% (n=28) of genes showed that same predicted XCI status. 24% (n=85) of genes in placenta showed variable escape from XCI, which is substantially greater than the 2% of genes in blood. More than four times as many genes (41%, n=147) were predicted to escape from XCI in placenta than the 10% that were predicted in blood. Due to the combination of decreased female DNA methylation and increased male DNA methylation, there were more genes that could not be predicted in placenta (27%, n=96) compared to blood (5%). Genes in which one probe predicted the gene was subject to XCI while another predicted escape from XCI occurred at the same frequency in placenta and blood

(1% n=5). Figure 4.3 compares the predicted XCI status in placenta to that previously determined in blood and somatic tissues. All genes which were predicted to be subject to XCI in placenta were also predicted to be subject to XCI in all other tissues examined. Placenta hypomethylation therefore predicts that this organ will display the highest degree of escape from XCI of any tissue examined.

4.3.4 All repetitive elements examined are hypomethylated in the placenta compared to blood

Previously, only the global DNA methylation status of LINEs was examined. Therefore the DNA methylation at other repetitive elements was now evaluated. Repetitive element DNA methylation on the X chromosome and the autosomes was examined using DNA methylation at those probes which had been removed from analysis of Illumina Infinium HumanMethylation27 array due to being located in repetitive elements. One-third of probes located in repetitive elements were also located in CpG islands (HC or IC) and demonstrated the typical X-linked CpG island DNA methylation pattern of MeXIP (data not shown). These were therefore excluded from the analysis to avoid a CpG density bias. All repetitive elements, regardless of type, showed a significantly (p -value <0.001) lower average level of DNA methylation in placenta compared to blood. At X-linked repetitive elements the placenta was 20% less methylated in males and 23% less methylated in females compared with blood (Figure 4.4A). X-linked hypomethylation was greater than that for the autosomes, with a 17% decrease in placental DNA methylation compared to blood (Figure 4.4B).

To determine if different repetitive elements showed different degrees of placental hypomethylation, the three most common types of repetitive elements on the Illumina Infinium HumanMethylation27 array were examined. These repetitive elements were SINEs (47% of repetitive elements probes), LINEs (26% of repetitive elements probes) and LTRs (long terminal repeats) (13% of repetitive elements probes). Male and females showed similar patterns of placental hypomethylation compared to blood, however, on the X chromosome LINEs showed the largest decrease in DNA methylation (blood-placenta=26%, male and female p -value <0.001) whereas on the autosomes, LTRs showed the largest decrease in DNA methylation (blood-placenta=19%, male and female p -value <0.001) (Table 4.2). Repetitive elements on the X chromosome showed a greater decrease in DNA methylation from blood to placenta in both males and females compared to DNA methylation at LC probes not in repetitive elements. On the autosomes, the opposite trend was observed: DNA methylation at LC probes not in repetitive elements showed a greater decrease in DNA methylation from blood to placenta in both males and females. The analysis of the Illumina Infinium HumanMethylation27 array data demonstrated that repetitive elements on the X chromosome are more

hypomethylated (blood-placenta) than those on the autosomes and that this hypomethylation occurs across all examined types of repetitive elements.

4.3.5 Placental cells demonstrate the same frequency of Cot-1 holes as somatic cells

Cot-1 holes are caused by a decrease in Cot-1 RNA transcription in a small area compared with the level of Cot-1 transcription observed across the nucleus. Given that repetitive elements in the placenta are hypomethylated compared to blood and somatic tissues (see sections 3.3.1 and 4.3.4) there may be a higher degree of transcription from repetitive elements which would, in turn, result in a decrease in the frequency or intensity of Cot-1 holes. In order to determine the frequency at which Cot-1 holes occur, RNA FISH was performed, combining probes to detect XIST RNA as well as Cot-1 RNA. Analysis using RNA FISH requires cells be cultured, therefore, while derived from whole villi, only cells from the chorionic mesoderm are expected to grow in culture. These cells are derived from the inner cell mass rather than the whole villi samples used in previous sections [310]. Cells were individually scored as being Cot-1 hole positive or Cot-1 hole negative (an example of each is shown in Figure 4.5A). While the placental cells examined had a lower frequency of Cot-1 hole positive cells (69%) compared to somatic cells (80%) the difference was not statistically significant (Figure 4.5B). The intensity of Cot-1 hybridization was compared between placental and somatic cells to investigate the possibility that placental Cot-1 hole positive cells had less Cot-1 transcription than somatic Cot-1 hole positive cells. No statistical difference was detected in the intensity of Cot-1 hybridization or size of the Cot-1 hole between placental and somatic cells, therefore other differences between Cot-1 hole positive and Cot-1 negative cells were investigated.

4.3.6 The X_i of Cot-1 hole positive cells is found most frequently at the nuclear periphery

Consistent with being transcriptionally silent, the X_i is typically located at the nuclear periphery. Therefore, the location of the X_i , as marked by presence of XIST RNA, was compared in Cot-1 hole positive and negative cells to determine if the nuclear location of the X_i was influenced by the presence of a Cot-1 hole. One of three possible nuclear locations was assigned to each X_i : peripheral (touching the nuclear periphery), semi-peripheral (close to, but not touching the nuclear periphery) or central (far from the nuclear periphery); an example of each location is shown in Figure 4.6A. No significant differences in location were found between placental and somatic cells, so all cells were combined to allow for an overall comparison between Cot-1 hole positive and negative cells. Cot-1 hole positive cells were most often (72%) found at the nuclear periphery whereas Cot-1 hole negative cells were most often (47%) semi-peripheral (Figure 4.6B). The Cot-1 holes in Cot-1 positive cells were as strong when found at the nuclear

periphery as when found in the center of the nucleus precluding imaging issues as the cause of these differences. The proportion of cells in which the X_i was centrally located was nearly three times higher (29% vs 10%) in Cot-1 hole negative cells compared to positive cells. The area of the X_i in Cot-1 hole positive and negative cells was next compared to determine if the loss of a peripheral nuclear location coincided with an increase in area.

The dense DAPI stained Barr body region and the XIST RNA area were used to measure the area occupied by the X_i . Three different regions were compared based on the presence of these two marks: the Barr body only region, the XIST only region and the Barr body and XIST overlap region. A comparison between placental and somatic cells did not yield any significant differences between the areas occupied by any of the three regions, so all data were combined to examine whether Cot-1 hole positive and negative cells differ in areas across these three regions. There were no significant differences between Cot-1 hole positive and negative cells with respect to any specific region (Figure 4.7A). However, in both Cot-1 hole positive and negative cells, the area covered by the XIST only region was significantly (p -value<0.001) larger than the area covered by the Barr body only region. In Cot-1 hole positive cells the Barr body and XIST overlap region was also significantly (p -value<0.05) larger than the Barr body only region. Classically, either the Barr body or the presence of XIST RNA is used to identify the X_i . Approximately 25% (Cot-1 hole positive: 29%, Cot-1 hole negative: 25%) of the area of the Barr body does not overlap with XIST RNA. Conversely, just over 50% (Cot-1 hole positive: 58%, Cot-1 hole negative: 52%) of the area occupied by XIST RNA does not overlap the Barr body. The differences in the areas occupied by the Barr body and XIST led to the assessment of other features of XCI between these regions.

4.3.7 The lowest Cot-1 transcription occurs in the Barr body and XIST overlap region of Cot-1 hole positive cells

The hypomethylation of repetitive elements in the placenta was hypothesized to result in an increase in repetitive element transcription which would in turn result in a decrease in the frequency of Cot-1 hole positive cells. Although the frequency of Cot-1 hole positive cells did not decrease in the placenta (see section 4.3.5), the strength of Cot-1 holes, represented by the intensity of Cot-1 hybridization, was compared between somatic and placental cells to determine if the hypomethylation of repetitive elements in the placenta instead resulted in “weaker” Cot-1 holes. As expected, Cot-1 hole positive cells had a significantly (p -value<0.001) lower Cot-1 intensity than Cot-1 hole negative cells in all regions examined (Figure 4.7B). In no region was the intensity of Cot-1 hybridization significantly different between placental and somatic cells. While Cot-1 hole negative cells showed no significant differences in Cot-1 intensity across the examined regions, the Barr body and XIST overlap region of Cot-1 hole

positive cells had significantly (p -value <0.001) lower Cot-1 intensity than the Barr body only region.

The presence of a Barr body or XIST RNA has traditionally been used to mark the X_i . However as discussed in section 4.3.6, these two regions do not fully overlap. The level of XIST intensity was not significantly different between Cot-1 hole positive and negative cells, suggesting that it is not the level of XIST RNA which determines the presence of Cot-1 holes. In both Cot-1 hole positive and negative cells, XIST and DAPI intensity was (p -value <0.001) strongest in the Barr body and XIST overlap region (Figure 4.7C and D). The DAPI intensity of the Barr body only region of Cot-1 hole positive cells was significantly (p -value <0.001) weaker than in Cot-1 hole negative cells (Figure 4.7D). Therefore, as summarized in Figure 4.8, the region in which XIST RNA and DAPI intensity are the strongest corresponds to where the Cot-1 hybridization intensity is weakest suggesting that the features of XCI are highly inter-related.

4.4 Discussion

The X_i is characterized by more than differential DNA methylation, it is also associated with a peripheral nuclear location [98], the presence of XIST RNA [16, 17] and the presence of a Barr body [16, 18]. The X_i in Cot-1 hole negative cells was on the nuclear periphery less often than in Cot-1 hole positive cells (Figure 4.6B) but no significant differences in the areas occupied by XIST RNA and/or the Barr body were detected. The difference in nuclear location alone does not account for the lack of Cot-1 holes in Cot-1 hole negative cells. In Cot-1 hole positive cells the Cot-1 hole was the same intensity and size regardless of location (data not shown). The X_i , as marked by XIST RNA, in Cot-1 hole positive cells can therefore be located off the nuclear periphery and still result in a repression of Cot-1 transcription. Although both XIST RNA and the Barr body are often used to identify the X_i , the overlap in these two marks was only one third of the total X_i area (as identified by either XIST RNA or the Barr body). Different histone modifications associate with different portions of the X chromosome (see Table 1.1). XIST RNA co-localizes with macroH2A and H3K27me3 while H3K9me3 and H3K20me3 co-localize with HP1 but not XIST RNA [40]. Therefore the Barr body only region in this study most likely represents the HP1/H3K9me3/H3K20me3 positive regions of the X_i and not the XIST RNA/macroH2A/H3K27me3 positive regions. In Cot-1 hole positive cells the Cot-1 hybridization was significantly stronger in the Barr body only region suggesting that these regions of the X_i have more Cot-1 transcription than the XIST only and XIST and Barr body overlap regions. The distinct combinations of histone marks across the X_i may create domains in which silencing is maintained in different manners, which may in turn affect factors such as the degree of repetitive element silencing.

Although recent comparisons of autosomal promoter DNA methylation in placentas of varying gestational ages have shown an overall increase in DNA methylation with increased gestational age [2, 311], no biologically relevant DNA methylation changes on the X chromosome with gestational age were detected here (Figure 4.1). By combining placentas of varying gestational age, it was demonstrated that placental hypomethylation is not consistent across the genome. Data from chapters 3 and 4 agree that the largest decrease in placental DNA methylation (blood-placenta) occurs at X-linked CpG island promoters in females. A novel finding in this study is that three regions (X-linked male CpG islands promoters, autosomal male and female CpG islands promoters) were found to be significantly hypermethylated in placenta compared to blood. Studies on autosomal t-DMR have previously shown that placental t-DMRs have the third highest proportion (62%) of hypermethylated t-DMRs of the examined tissues (tissue n=16) [279]. Although the statement that the placenta is hypomethylated is commonly used, it is misleading and should be amended to state that in both sexes some, but not all, non-island promoters on the X chromosome and autosomes as well as a subset of X-linked island promoters in females are hypomethylated but that autosomal non-island promoters in males and females as well as X-linked non-island promoters in males are generally hypermethylated. Given that female X-linked CpG island promoters were found to have the largest decrease in placental DNA methylation (blood-placenta) it is not surprising that this would translate into differences in the predicted XCI status of these genes. Previous studies have identified single genes that escape from XCI in placenta and not in other tissues but nowhere near the 65% of genes that are predicted to escape from XCI by DNA methylation (Figure 4.3) [150, 152]. One gene *G6PD*, is known to be subject to XCI in somatic tissues but to escape from XCI in the placenta [152], and in agreement with previous reports [150] the *G6PD* promoter CpG island demonstrated MeXIP in blood (male=3%, female=38%) but was hypomethylated in placenta (male=8%, female=13%). Although more genes were predicted to escape than variably escape from XCI in the placenta, the variable escape category showed the largest increase from the XCI status predicted in blood with 12 times more genes predicted to variably escape from XCI in the placenta compared to blood. The variable escape category of genes is of particular interest in the placenta given the role that environmental factors are thought to play in influencing placental DNA methylation levels [312]. Because DNA methylation predicts that a gene escapes from XCI in the placenta, does not mean that it is expressed in the placenta. Genes which are not expressed in the placenta may be more likely to be unmethylated specifically because it does not matter if the stability of XCI is lost in the placenta.

DNA methylation analysis has shown L1 elements to be hypomethylated in the placenta while Alus were not hypomethylated compared to blood [313]. In this study, all examined classes of

repetitive elements (LINEs, SINEs and LTRs) showed placental hypomethylation compared to blood in both males and females. This was true on the X chromosome as well as the autosomes. LINEs were more hypomethylated than SINEs and for both types of elements the X chromosome was more hypomethylated than the autosomes (Table 4.2). When examining repetitive element DNA methylation by any technique it is important to consider exactly which classes of elements are being examined. The differences in placental hypomethylation between this and previous reports are likely due to differences in the groups of repetitive elements examined. L1 elements have long been proposed to play a role in XCI [194] and, more recently, silent LINEs were shown to play a role in the creation of the X_i domain in mouse [197]. If certain types of repetitive elements are hypomethylated in the placenta there may be consequences for XCI such as the loss of repetitive element silencing causing the loss of the Cot-1 hole.

It is important to note that the placental cells used in the DNA methylation analysis in this chapter were a mixture of different cell types than the placental cells used in the Cot-1 hole analysis. Because these cell types have different origins, future DNA methylation analysis of cultured placental cells will provide a better system to evaluate the effects of DNA methylation on Cot-1 holes. However, while the comparison of DNA methylation from whole villi (used in sections 4.3.1 to 4.3.4) may not be ideal to represent DNA methylation from chorionic mesoderm (used in sections 4.3.5 to 4.3.7), it does provide a preliminary comparison of the effects of placental hypomethylation of repetitive elements on Cot-1 hole frequency. Despite the observed hypomethylation of repetitive elements in the placenta, there was no significant difference in the frequency of cells which were Cot-1 hole positive (Figure 4.5B and Figure 4.8). Beside the differences in cell composition, one possible explanation is that although the repetitive elements tested here show placental hypomethylation, these are not the repetitive elements which make up the Cot-1 fraction. A more likely possibility is that while the same repetitive elements are represented in both the DNA methylation and Cot-1 hole experiments, the degree of hypomethylation is not sufficient to result in an increase in transcription and hence a loss of the Cot-1 hole. Examples of repetitive element hypomethylation which drive the expression of specific genes are associated with much lower placental DNA methylation than was, on average, observed here. For example, the average hypomethylation from blood-placenta found by Macaulay *et al.* in the repetitive element driven genes *KCNH5*, *INSL4*, *ERVWE1*, *EDNRB*, *PTN* and *MID1* was approximately 60% [314]. The largest average decrease in placental DNA methylation (blood-placenta) we observed in repetitive elements was less than half of this (23%) in female X-linked repetitive elements. An examination of all X-linked repetitive elements found only four in which both the male and female placenta showed a

decrease in DNA methylation compared to blood of greater than 45% (the smallest decrease from *KCNH5*, *INSL4*, *ERVWE1*, *EDNRB*, *PTN* and *MID1*). It is also possible that not all placental cells show the same level of repetitive element hypomethylation. In this situation, those placenta cells in which there was minimal hypomethylation would not demonstrate a change in Cot-1 hole frequency whereas cells where hypomethylation was more extensive would show a decrease in Cot-1 hole frequency. Overall while placentas do show an overall hypomethylation at repetitive elements, we postulate that the degree of hypomethylation is not sufficient to result in an increased repetitive element transcription.

These studies of placental DNA methylation have revealed a complex pattern in which certain regions, including female X-linked CpG island promoters, are hypomethylated compared to blood, while others, such as male X-linked CpG island promoters, are hypermethylated. The hypomethylation of female X-linked CpG island promoters translates into a predicted degree of escape from XCI not previously observed in somatic tissues. The autosomes also demonstrate unique patterns of placental DNA methylation with CpG island promoters being surprisingly hypermethylated compared to blood, whereas non-island promoters are hypomethylated. Despite the hypomethylation of all types of repetitive elements in the placenta compared to blood, there was no detected effect on the presence or intensity of Cot-1 holes. We conclude therefore that the decrease in DNA methylation observed at repetitive elements in the placenta is not sufficient to result in an overall increase in repetitive element transcription.

Table 4.1: Control placental samples used in chapter 4.

First and second trimester placentas were from chromosomally normal, aborted pregnancies. Third trimester samples were from successfully term pregnancies.

Sample name	Sex	Type	Gestational age (weeks)	Trimester	Tissue type
PZET1*	Female	control	8-14	first trimester	villi
PZET2*	Male	control	8-14	first trimester	villi
PZET5*	Male	control	8-14	first trimester	villi
PZET7*	Male	control	8-14	first trimester	villi
PZET10*	Male	control	8-14	first trimester	villi
PZET12*	Male	control	8-14	first trimester	villi
PZET13*	Female	control	8-14	first trimester	villi
PZET15*	Male	control	8-14	first trimester	villi
PZET16*	Male	control	8-14	first trimester	villi
PZET24*	Male	control	8-14	first trimester	villi
FT3 vil*	Female	control	19.7	second trimester	villi
FT13 vil*	Female	control	17	second trimester	villi
FT5 vil*	Male	control	23.7	second trimester	villi
FT18 vil*	Male	control	20.4	second trimester	villi
mT4-5 vil*	Female	control	20.3	second trimester	villi
PM144V2*	Female	control	41	third trimester	villi
PM155V2*	Female	control	41.6	third trimester	villi
PM172V2*	Female	control	40.7	third trimester	villi
PM181V2*	Male	control	39	third trimester	villi
PM135V2*	Male	control	39	third trimester	villi
PM143V2*	Male	control	39.3	third trimester	villi
PM182V2*	Female	control	39.8	third trimester	villi
PM190V2*	Male	control	39.1	third trimester	villi
PM201V2*	Male	control	38.6	third trimester	villi
PM202V2*	Female	control	39.6	third trimester	villi

* analyzed by Illumina Infinium HumanMethylation27 array

Table 4.2: Degree of placental hypomethylation calculated as a delta (blood-placenta) at different types of repetitive elements.

Blood and placental DNA methylation were compared using the Kruskal-Wallis test. When p-values were greater than 0.05 they were not significant (ns), however, p-values between 0.01 and 0.05 (*), between 0.01 and 0.001 (**), and p-values <0.001 (***) were considered significantly different.

			X chromosome	autosomes
Female	average blood – average placenta	LINEs	26% ***	17% **
		SINEs	20% **	15% ***
		LTRs	23% ^{ns}	19% ***
Male	average blood – average placenta	LINEs	26% ***	17% ***
		SINEs	17% **	14% ***
		LTRs	21% ^{ns}	19% ***

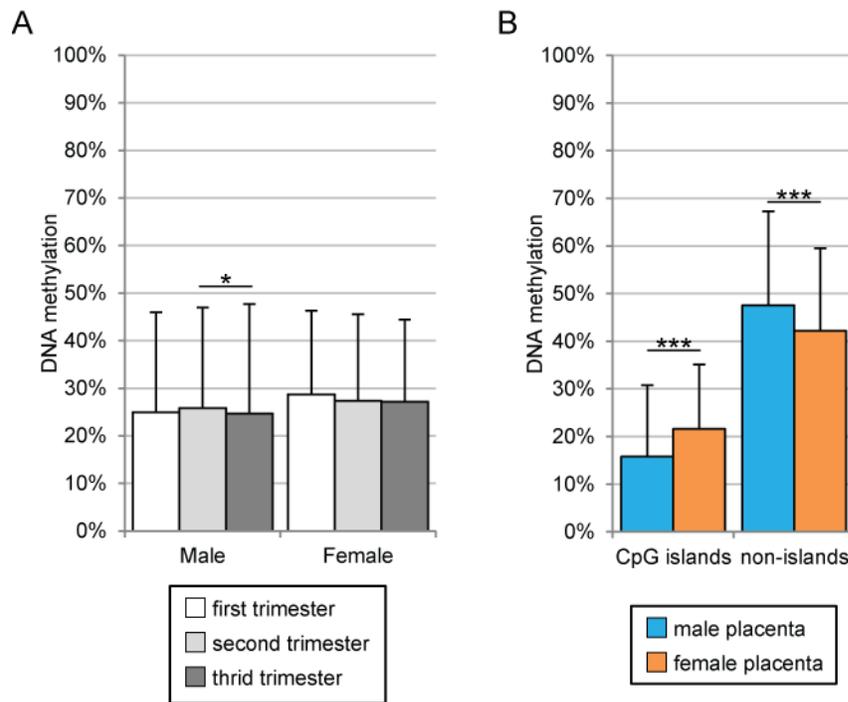


Figure 4.1: Male and female placentas show minimal X-linked DNA methylation differences across gestational ages.

(A) Male (n=15) placentas and female (n=10) placentas were separated by gestational age (first trimester male n=8, female n=2: white, second trimester male n=2, female n=3: light grey, third trimester male n=5, female n=5: dark grey) and the only significant difference in X-linked DNA methylation found within sex but between gestational ages was in males between the second and third trimester (p-value<0.05). (B) X-linked probes were grouped by CpG density (CpG islands: HCs and ICs, non-islands: LCs) with the average male placental DNA methylation shown in light blue and the average female placental DNA methylation in orange. As in somatic tissues, X-linked CpG islands were significantly (p-values <0.001) more in female placentas compared to male placentas. Significances were calculated using the Kruskal-Wallis test. When p-values were greater than 0.05 they were not significant, however, p-values between 0.01 and 0.05 (*), between 0.01 and 0.001 (**), and p-values <0.001 (***) were considered significantly different. Error bars are one standard deviation.

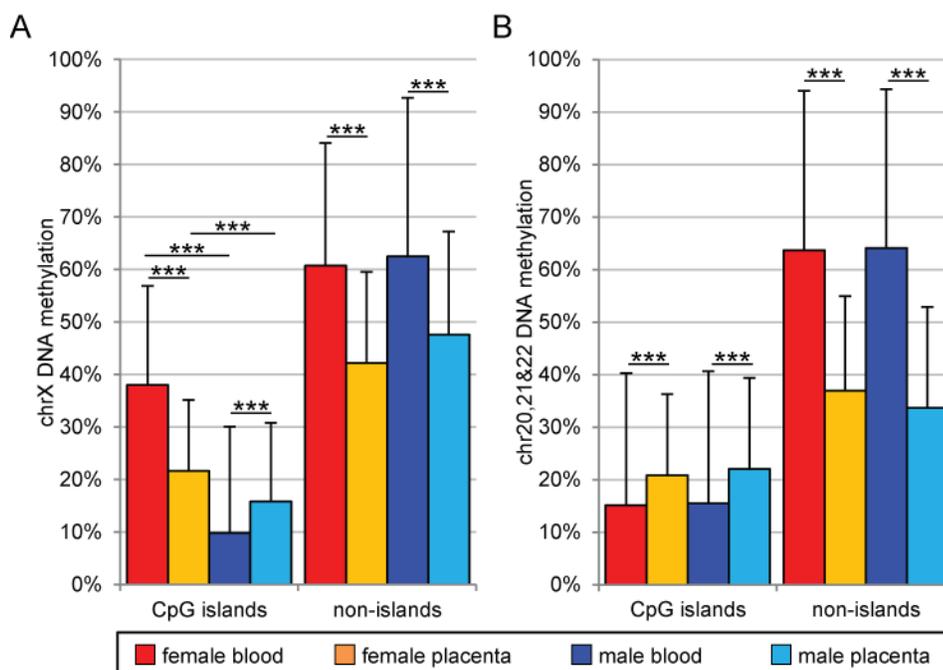


Figure 4.2: X-linked CpG density influences placental hypomethylation differently in males and females.

Average DNA methylation levels between tissues (female blood: red, female placenta: orange, male blood: dark blue, male placenta: light blue) with probes grouped by CpG density (CpG islands: HCs and ICs, non-islands: LCs). **(A)** X-linked promoter probes reveal similar significant (p -value <0.001) patterns of placental hypomethylation at CpG island promoters in females and non-island promoters in males and females while male CpG island promoters show significant (p -value <0.001) placental hypermethylation. **(B)** Autosomal (chromosomes 20, 21 and 22 only) promoters demonstrated CpG density dependent DNA methylation with CpG island promoters showing significant (p -value <0.001) placental hypermethylation compared to blood and somatic tissues while non-island promoters were significantly (p -value <0.001) hypomethylated in placenta compared to blood and somatic tissues. Significance calculated using Kruskal-Wallis test. When p -values were greater than 0.05 they were not significant, however, p -values between 0.01 and 0.05 (*), between 0.01 and 0.001 (**), and p -values <0.001 (***) were considered significantly different. Error bars are one standard deviation.

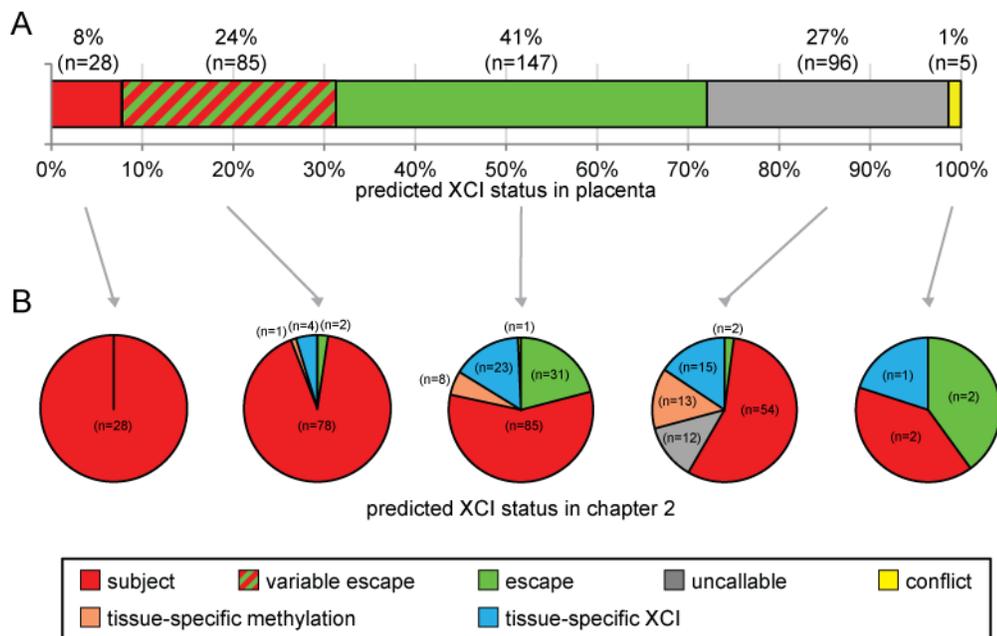


Figure 4.3: Increased placental escape from XCI predicted using DNA methylation.

(A) Male and female DNA methylation levels were used to predict XCI status (as outlined in Figure 2.1) of genes with probes in CpG islands (HC and IC) in placenta. The percentage of genes in each XCI status are listed above the bar graph. (B) The XCI statuses (subject: red, variable escape: diagonal red and green stripes, escape: green, unclassifiable: grey, conflicts: yellow, tissue-specific DNA methylation: peach, tissue-specific XCI: blue) found in chapter 2 for each of the five XCI statuses found in placenta are shown as pie charts below each XCI status. Although no genes were found to variably escape from XCI in all somatic tissues in chapter 2, the XCI status is included to allow comparison with other thesis figures.

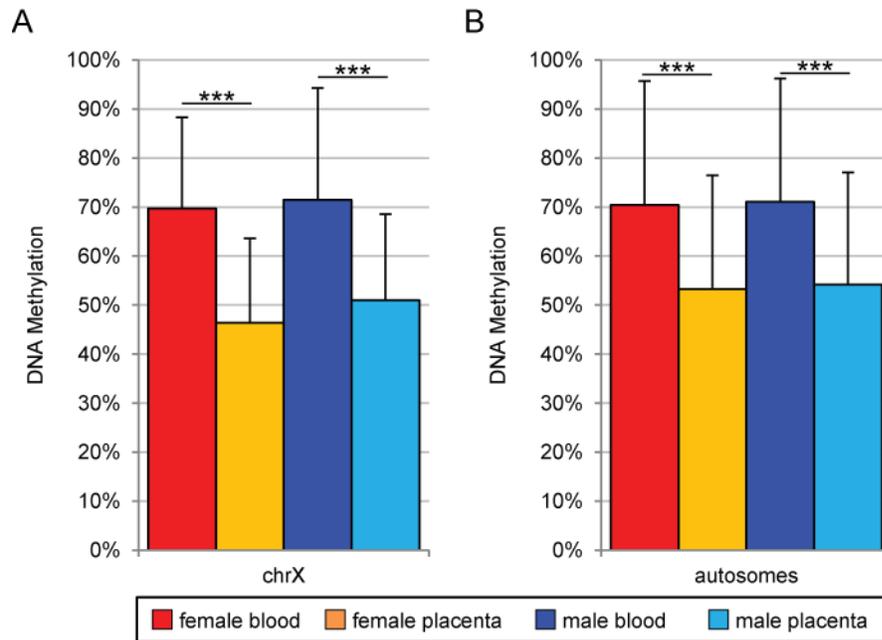


Figure 4.4: Non-CpG island repetitive elements are hypomethylated in the placenta on both the X chromosome and the autosomes.

Average DNA methylation levels at repetitive elements in non-CpG islands across tissues (female blood: red, female placenta: orange, male blood: dark blue, male placenta: light blue) **(A)** X-linked repetitive element probes showed similar patterns of significant (p -value <0.001) placental hypomethylation in males and females. **(B)** Autosomal (chromosomes 20, 21 and 22 only) repetitive element probes showed similar patterns of significant (p -value <0.001) placental hypomethylation in males and females. Significance calculated using Kruskal-Wallis test. When p -values were greater than 0.05 they were not significant, however, p -values between 0.01 and 0.05 (*), between 0.01 and 0.001 (**), and p -values <0.001 (***) were considered significantly different. Error bars are one standard deviation.

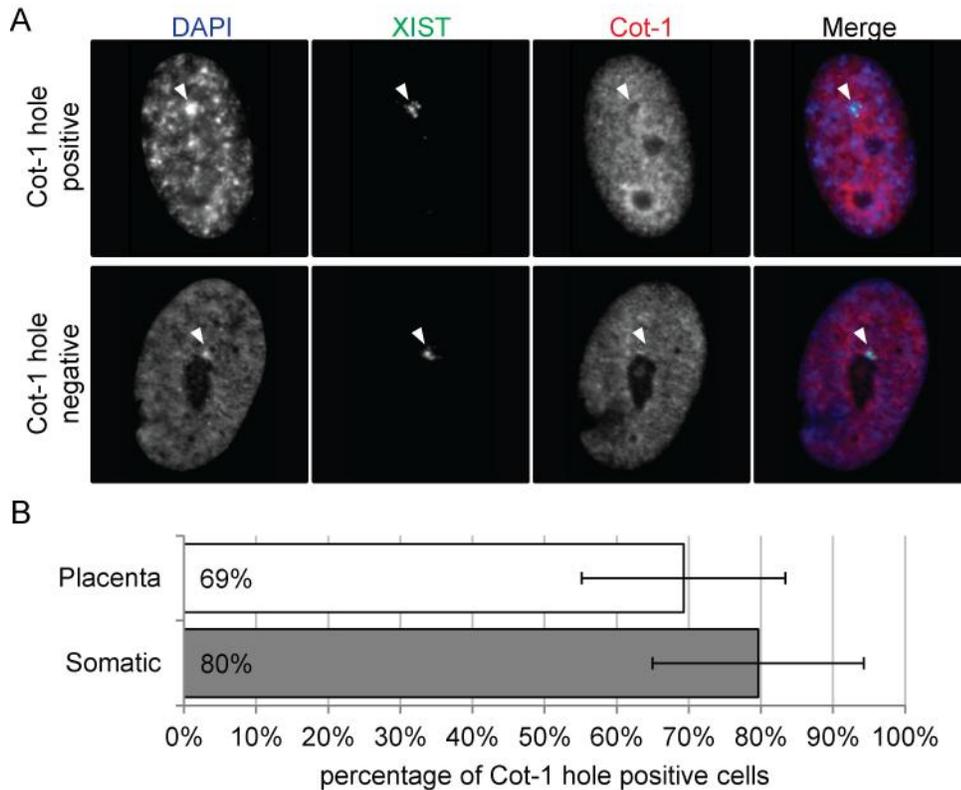


Figure 4.5: The frequency of Cot-1 hole positive and negative cells is not significantly different in placental and somatic cells.

(A) RNA FISH combining XIST RNA (green in merged image) and Cot-1 RNA (red in merged image) along with DAPI (blue in merged image) for nuclear staining on somatic and placental cells. An example of a Cot-1 hole positive cell is shown above an example of a Cot-1 hole negative cell. White arrow head marks the location of the X_i (as determined by the presence of XIST RNA). (B) Somatic (n=282) and placenta (n=741) cells were individually scored as Cot-1 hole positive or Cot-1 hole negative through the use of the co-localization highlighter tool in Image J [309]. No statistically significance difference between the frequency of Cot-1 holes in placental and somatic cells was detected using Mann-Whitney test. Error bars are one standard deviation.

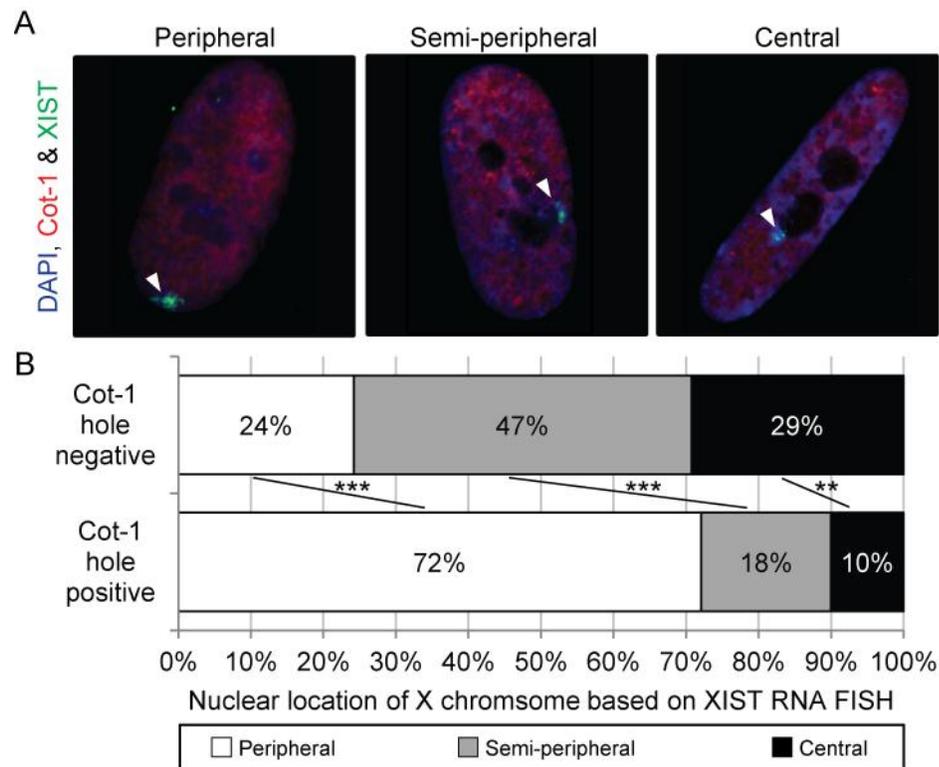


Figure 4.6: The X_i of Cot-1 hole positive and negative cells differ in nuclear location.

(A) Examples of the three possible nuclear locations (peripheral, semi-peripheral or central) assigned to all examined X_i s. A merged image of an RNA FISH with XIST RNA (green), Cot-1 RNA (red) and DAPI (blue) is shown for each location with a white arrowhead to mark the X_i (as determined by the presence of XIST RNA). (B) No significant difference was detected between placental and somatic cells therefore both cell types were combined and then divided into Cot-1 hole positive and negative cells. The percentage of cells in each location (peripheral (white), semi-peripheral (grey) or central (black)) is given for Cot-1 hole positive and negative cells with significant differences between location frequencies in Cot-1 hole positive and negative cells shown between. Significance calculated using Kruskal-Wallis test. When p-values were greater than 0.05 they were not significant, however, p-values between 0.01 and 0.05 (*), between 0.01 and 0.001 (**), and p-values <0.001 (***) were considered significantly different.

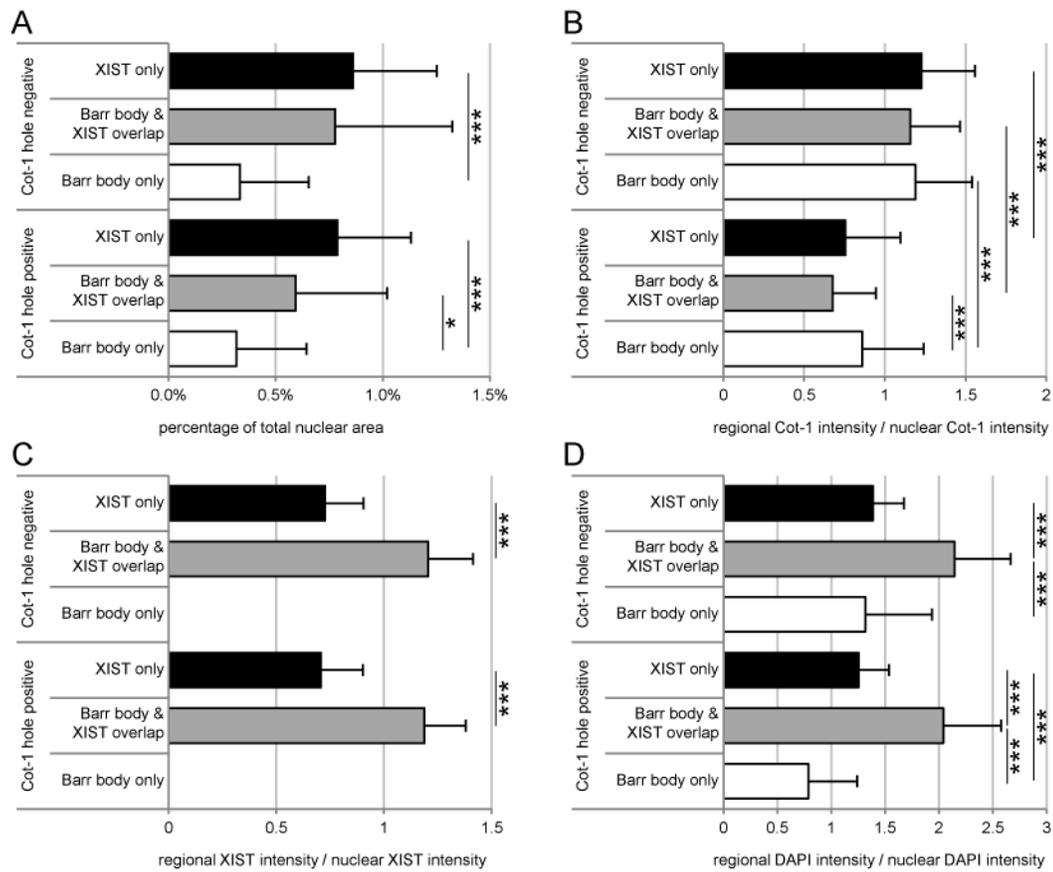


Figure 4.7: Cot-1 intensity is lowest in the Barr body and XIST overlap region of Cot-1 hole negative cells.

No significant difference was detected between placental and somatic cells therefore both cell types were combined and then divided into Cot-1 hole positive and negative cells. **(A)** The area occupied by each region is calculated as a percentage of the overall nuclear area (as determined by DAPI staining). **(B)** The Cot-1 intensity in each region is calculated as a percentage of the average nuclear Cot-1 intensity. **(C)** The XIST intensity in the Barr body and XIST overlap region and the XIST only region was calculated as a percentage of the average nuclear XIST intensity. The Barr body only region is not shown as by definition there was no XIST RNA detected in this region. **(D)** The DAPI intensity in each region is calculated as a percentage of the average nuclear DAPI intensity. Although the XIST only region was not found to overlap the Barr body, the DAPI intensity at this region could still be analyzed. **(A-D)** Comparisons of the area occupied by each region were calculated using the Kruskal-Wallis test. When p-values were greater than 0.05 they were not significant, however, p-values between 0.01 and 0.05 (*), between 0.01 and 0.001 (**), and p-values <0.001 (***) were considered significantly different. Error bars are one standard deviation.

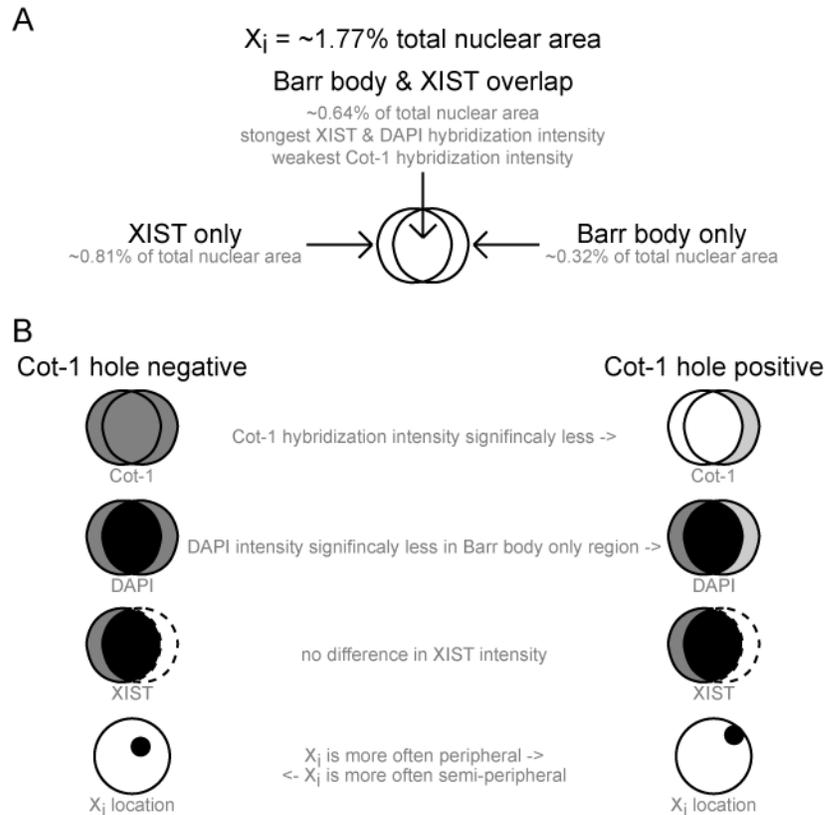


Figure 4.8: Summary of differences between Cot-1 hole positive and negative cells.

(A) Differences in area occupied by the XIST only region (far left), Barr body and XIST overlap region (center) and the Barr body only regions (far right). The average percent of the total nuclear area (defined as the maximum DAPI stain) is given for each region. (B) Graphically summary of the difference between Cot-1 hole positive and negative cells. For Cot-1, DAPI and XIST, white shading represents the weakest hybridization signal and black the strongest hybridization signal. For X_i location the X_i is shown as a black circle inside a white circle representing the nucleus.

5 DNA methylation demonstrates spread of XCI into X;autosome translocations

5.1 Introduction

The different levels of spread of inactivation into the autosomal portion of X;autosome translocations provides a system to identify DNA elements involved in XCI. These include: regions which are protected from inactivation through the presence of escape elements, regions which are subject to inactivation due to the presence of way stations and regions in which boundary elements act between active and inactivated genes (see section 1.6 for further details). It is not clear how the expression of XIST RNA from a single X-linked locus can lead to the silencing of an entire chromosome. Through the examination of how inactivation spreads into the autosomal portions of X;autosome translocations it was proposed that way station elements, found along the chromosome, help propagate inactivation [193]. Since inactivation was known to spread more effectively on the X chromosome portion of X;autosome translocations, it was proposed that way stations would be more common on the X chromosome than the autosomes, and L1 elements were put forward as the best way station candidate [194]. In mice, evidence for the existence of escape elements has been shown through the integration of the *Kdm5c* (previously known as *Jarid1c*) into four different locations on the X chromosome. *Kdm5c* escapes from XCI in its normal location as well as at each of the four other X-linked integration sites [138]. The consistent ability of *Kdm5c* to escape from XCI suggests that within the immediate *Kdm5c* region there must be an element which causes escape from XCI rather than escape from XCI being the result of a larger, euchromatic domain.

Unbalanced X;autosome translocations typically show non-random XCI in which the translocated chromosome is always inactivated, whereas balanced X;autosome translocations typically show non-random XCI in which the normal X chromosome is inactivated [179, 180]. Although XCI in unbalanced X;autosome translocations is non-random, it is not always clonal. Different cells from a single individual with an X;autosome translocation can show varying degrees of the spread of late replication into the autosomal portion [170], suggesting that the spread of inactivation can vary between cells. Unbalanced X;autosome translocations show non-random XCI in which the translocated chromosome is always the X_i , providing the opportunity to study the ability of inactivation to spread out of the X chromosome and into autosomal DNA.

Previous DNA methylation analysis of autosomal genes in X;autosome translocations has demonstrated good correlation between transcriptional silencing and the presence of DNA methylation at autosomal CpG island promoters [190, 192]. Therefore, just as inactivated genes on the X chromosome demonstrate MeXIP, autosomal genes which are subject to inactivation are methylated at CpG island promoters on the translocated chromosome, while autosomal genes which escape from inactivation remain unmethylated. Further evidence for

the role of DNA methylation in the inactivation of autosomal genes has been demonstrated through the use of the demethylating agent 5-azacytidine. When cells containing X;autosome translocations were treated with 5-azacytidine there was a decrease in the degree to which the autosomal portion of X;autosome translocation was late replicating [170]. The switch to early replication timing as a result of the loss of DNA methylation supports the hypothesis that the presence of DNA methylation on the autosomal portions of X;autosome translocations is a sign of inactivation and that DNA methylation can be used as a proxy to detect autosomal genes subject to inactivation.

Having established that DNA methylation can accurately predict the XCI status of X-linked genes in chapter 2 [308], the study in this chapter employed DNA methylation analysis of the autosomal portion of six X;autosome translocations to examine the spread of inactivation. A comparison of the frequency of repetitive elements in regions that are subject to inactivation with their frequency in regions that escape from inactivation provided insight into the role repetitive elements play in the spread of inactivation.

5.2 Methods

5.2.1 Sample preparation and bisulfite conversion

All samples, except HFF, were commercially available fibroblast cell lines (see Table 5.1 for sample information). Primary fibroblasts cultured from human fetal foreskin were a gift from Dr. Matthew Lorincz. DNA was extracted using the standard extraction protocol with a Qiagen RNA/DNA Allprep kit. 750 ng of DNA was bisulfite converted using the EZ DNA methylation kit (Zymo Research) with the alternative incubation conditions outlined for use with the Illumina Infinium HumanMethylation450 array.

5.2.2 Illumina Infinium HumanMethylation450 array

Samples were run on the Illumina Infinium HumanMethylation450 array by the Kobor lab. Briefly, 160 ng of bisulfite converted DNA was whole genome amplified, fragmented and hybridized to the Illumina Infinium HumanMethylation450 array following standard protocol as outlined in the user guide. CpG islands were defined as previously outlined (see chapter 2). Probes within HCs and ICs were considered CpG island probes whereas probes which were located in LCs were considered non-island probes. Probes were removed from analysis due to being located either in repetitive elements (chr2: 5324, chr9: 1455, chr14: 2128, chr21: 496, chr22: 1278, chrX: 1390) and/or cancer/testis genes (chr2: 75, chr9: 23, chr14: 16, chr21: 44, chr22: 0, chrX: 585).

5.3 Results

5.3.1 CpG island DNA methylation changes with distance from TSS

Given that the CpG island promoters of X-linked genes subject to XCI demonstrate MeXIP, whereas the CpG island promoters of genes which escape from XCI are unmethylated, it was hypothesized that the CpG island promoters of autosomal genes on X;autosome translocations which were subject to inactivation would also become methylated. From previous X-linked DNA methylation analysis (chapter 2, [308]) CpG island promoters, but not non-island promoters, were known to accurately predict XCI status. However, X-linked genes associated with CpG island promoters on the Illumina Infinium HumanMethylation450 array contain an average of eight probes per gene, compared to the two probes per gene on the Illumina Infinium HumanMethylation27 array.

The first step was to establish at which probes DNA methylation was predictive of XCI status. Two data sets with known XCI status were designed to function as training sets in order to compare how distance from the TSS affected X-linked CpG island promoter DNA methylation. The first training set was composed of 173 genes which were previously found to be subject to XCI in all examined tissues [308] and escaped from XCI in less than or equal to 22% of X_i hybrids [106] (Table 5.2). The second training set was composed of 32 genes which were previously found to escape from XCI in all examined tissues [308] and escaped from XCI in greater than or equal to 78% of X_i hybrids [106] (Table 5.3). When using DNA methylation to predict XCI status it is critical to examine the difference in DNA methylation between the X_a and the X_i . A DNA methylation delta (female – male) will be close to 0% if a gene is unmethylated in both males and females and therefore escapes from XCI. When a gene is subject to XCI, MeXIP translates into a DNA methylation delta (female – male) significantly higher than 0%. Therefore the best location to predict XCI status would be where the DNA methylation delta (female – male) was smallest for genes which escape from XCI and largest for genes subject to XCI.

The training set composed of genes subject to XCI had a consistent DNA methylation delta of approximately 30% whereas the training set composed of genes which escape from XCI had a DNA methylation delta of approximately 10% at the TSS and 30% at those probes located more than 701 bp upstream or 1301 bp downstream of the TSS (Figure 5.1). Probes located 400 bp upstream to 1300 bp downstream of the TSS (shaded grey box on Figure 5.1) were averaged together to create a single DNA methylation value from each gene. The average DNA methylation from only those probes located in CpG island promoters between 400 bp upstream and 1300 bp downstream of the TSS were used to establish rules to predict XCI status because probes in these regions showed the largest difference in methylation between training sets.

Combining the DNA methylation of multiple probes into a single average DNA methylation value was beneficial since it is the overall DNA methylation of a CpG island which should be used to predict the XCI status of a gene rather than individual CpGs. Genes predicted to be subject to XCI had male DNA methylation less than 20%, female DNA methylation greater than 25% and a DNA methylation delta (female – male) between 22% and 60%. Genes which escaped from XCI had male DNA methylation less than 20%, female DNA methylation less than 25% and a DNA methylation delta (female – male) between -10% and 20%. Additionally, an XCI status was only predicted when at least two CpG island probes were between 400 bp upstream and 1300 bp downstream of the TSS. Having established the DNA methylation trends typically associated with XCI status, the DNA methylation status of autosomal CpG island promoters was utilized to investigate the spread of inactivation into the autosomal portions of X;autosome translocations.

5.3.2 DNA methylation analysis refines the minimal region of translocation breakpoints by 50%

When using DNA methylation to predict the XCI status of an X-linked gene, male and female DNA methylation levels are compared. To predict if inactivation had spread into the autosomal portion of X;autosome translocations, the DNA methylation at normal (non-translocated) chromosomes was compared against the DNA methylation of an X;autosome translocation. Therefore in addition to the normal 46, XX female sample, each X;autosome translocation also served as a normal control for the autosomes not involved in the translocation. The criteria used to predict the inactivation status of autosomal genes were based on the levels of DNA methylation observed at X-linked genes known to be subject to XCI or to escape from XCI. An autosomal gene was predicted to be inactivated when the normal DNA methylation was below 20%, the X;autosome translocation DNA methylation was greater than 25% and the DNA methylation delta (X;autosome translocation – normal) was between 22% and 60%. An autosomal gene was predicted to escape from inactivation when the normal DNA methylation was below 20%, the X;autosome translocation DNA methylation was less than 25% and the DNA methylation delta (X;autosome translocation – normal) was between -10% and 20%. When a gene was represented by only a single probe the previously established XCI status and the XCI status predicted by DNA methylation were different at 19% of genes. However, when multiple probes were present in a gene, the previously established XCI status and the XCI status predicted by DNA methylation were different at only 6% of genes. Therefore genes with only one probe were excluded from further analysis. Once DNA methylation had been used to predict an inactivation status for autosomal genes, we examined how inactivation might be used to refine the regions in which the breakpoints of X;autosome translocation are found.

For the six X;autosome translocations studied, the current cytologically defined breakpoint locations range in size from 8.4 Mb to 25.6 Mb. In order to restrict the size of the regions in which the breakpoints were found, DNA methylation was used to predict the inactivation status of genes surrounding and within each breakpoint (Figure 5.2). In five out of six (GM07501, GM01414, GM00074, GM08134 and GM05396) samples the X chromosome portion of the X;autosome translocation was disomic while the remainder of the X chromosome was monosomic. When monosomy occurs on the X chromosome, genes do not undergo XCI and are unmethylated at CpG island promoters. When genes on the X chromosome portion of an X;autosome translocation were unmethylated but were found to be subject to XCI on normal X chromosomes, that region was defined as monosomic and therefore not on the X;autosome translocation. The X chromosome breakpoint portion of GM01730 could not be refined as DNA methylation differences are minimal between disomic and trisomic X chromosomes (Figure 5.2E). The locations of the X chromosome breakpoints in three samples (GM07503, GM01414 and GM08134) were located outside of the currently defined breakpoint regions. Refining the autosomal portion of the X;autosome translocations was more difficult since the absence of DNA methylation, and therefore a prediction of escape from inactivation, is not informative as to whether or not that gene is on the autosomal portion of the X;autosome translocation. However, genes which are predicted to be subject to inactivation (hypermethylated) must be on the portion of the autosome on the X;autosome translocation. Using DNA methylation to refine breakpoint locations reduced the size in which the breakpoint could be found by an average of 50% and in one case (Figure 5.2A) to less than 20%. Having refined the breakpoint locations, the number of genes which DNA methylation predicted were subject to inactivation could be compared across X;autosome translocations.

5.3.3 DNA methylation analysis predicts varied degrees of spread of inactivation between X;autosome translocations

Previous analysis of X;autosome translocations has shown varied degrees of inactivation not only between individuals with X;autosome translocations but also within a single X;autosome translocation [186, 190]. For each gene within the autosomal portion (excluding those within the refined breakpoint) of the X;autosome translocation, an average DNA methylation was calculated for all samples with a normal autosome and for the sample carrying the X;autosome translocation. Using the criteria outlined in the previous section to predict inactivation status, the normal and X;autosome translocation DNA methylation levels were compared and an inactivation status predicted for each gene (Figure 5.3). The false discovery rate for each autosome was calculated by dividing the false positives by total number of genes (false positives plus true positives). False positives were autosomal genes which were predicted to be

subject to inactivation despite not being on the autosome involved in the X;autosome translocation. True positives were autosomal genes which were predicted to escape from inactivation and were not on the autosome involved in the X;autosome translocation. False positives were only present on chromosome 14 (n=3) and resulted in a false discovery rate of 0.0017.

GM01414 (Figure 5.3B) and GM08134 (Figure 5.3D) had the highest percentage of genes subject to inactivation (29%) and GM05396 (Figure 5.3F) the lowest (2%). While the spread of inactivation varied between X;autosome translocations, an average of 38% of autosomal genes were subject to inactivation in the first 10 Mb after the breakpoint compared to only 4% in the 10 Mb farthest from the breakpoint. The four samples (GM07503, GM01414, GM00074 and GM08134) in which the autosomal portion of the X;autosome translocation was trisomic showed a higher percentage of genes subject to inactivation than the samples in which the autosomal portion of the X;autosome translocation was disomic (GM01730 and GM05396). Only 14% of genes predicted to be subject to inactivation in one X;21 translocation were predicted to escape from inactivation in the other, suggesting that DNA sequence plays a substantial role in determining inactivation status. In order to determine the influence of sequence composition on inactivation status, the frequency of repetitive elements in autosomal regions which escaped from inactivation were compared against regions which were subject to inactivation.

5.3.4 Domains which escape from inactivation are depleted for L1s and LTRs but enriched for Alus

Domains in which all the examined genes had the same predicted inactivation status were defined and the repetitive element composition in each was compared against the genome average (Figure 5.4). Additional domains were created around singleton genes which were surrounded by genes showing the opposite inactivation status. Five types of repetitive elements were examined: L1, Alu, LTR, low complexity and simple repeats. The average Alu frequency (17.4%) in domains which contained multiple genes which escaped from inactivation was significantly ($p\text{-value}<0.0001$) higher than the average Alu frequency (11.2%) in domains which were subject to inactivation. The genome average for Alu frequency is 10.0%, therefore domains which contained multiple genes which escaped from inactivation demonstrated an increase in the Alu sequence frequency (Figure 5.4A). Domains which contained multiple genes which escaped from inactivation had a significantly lower L1 ($p\text{-value}=0.0210$) and LTR ($p\text{-value}=0.0109$) frequency (average L1 frequency = 14.1%, average LTR frequency = 7.6%) than domains which were subject to inactivation (average L1 frequency = 17.0%, average LTR frequency = 9.1%). The genome average of 16.3% suggests that L1 elements are depleted in domains which escape from inactivation. However, the genome average of LTR frequency was

8.1%, making it less clear which type of domain is enriched or lacking LTR elements (Figure 5.4A).

Sequence composition analysis of simple repeats and low complexity repeats did not demonstrate any significant differences between domains which contained genes which escape from inactivation compared to domains which contained genes subject to inactivation (Figure 5.4B). Sequence composition analysis of singleton domains revealed no differences between genes subject to inactivation and those escaping from inactivation. Additionally, for L1, Alu and LTR elements, the domains which contained single genes which escaped from inactivation were significantly different (L1 p-value=0.0325, Alu p-value<0.0001, LTR p-value=0.0387) in sequence composition from the domains which contained multiple genes which escaped from inactivation. This suggests that while sequence composition may play a role in the creation of large domains which escape from inactivation, single genes which escape from inactivation may be controlled through another mechanism.

5.4 Discussion

The CpG island promoters of autosomal genes are typically unmethylated [315], therefore, an increase in DNA methylation at the CpG island promoter of an autosomal gene in an X;autosome translocation suggests that the gene has become silenced due to the spread of inactivation. Previous analysis of X;autosome translocations has demonstrated that the DNA methylation status of autosomal genes shows good agreement with inactivation status [190, 192], however DNA methylation has not been used to predict the spread of inactivation into the autosomal portion of X;autosome translocations. Two autosomal genes, *FOS* and *PNP*, were previously examined by expression analysis in GM00074 [191]. *FOS* was found to be subject to inactivation whereas *PNP* was found to escape from inactivation. Using DNA methylation to predict inactivation yielded that same inactivation status for both genes in this study as was previously found using expression.

Because X-linked genes are not subject to XCI when the X chromosome is monosomic, DNA methylation was able to refine X-linked breakpoints. On the autosomal portion of the X;autosome translocations, the presence of DNA methylation also refined breakpoints but not to the same degree as on the X chromosome portion. Although no genes within the current breakpoints of GM01730 and GM05696 could be examined using the Illumina Infinium HumanMethylation450 array, there are genes with CpG island promoters within these regions at which DNA methylation could be examined by other techniques thereby allowing for the possibility of further breakpoint refinement. The inactivation pattern observed in GM01414 is different from the other X;autosome translocations in that the first autosomal gene predicted to be subject to inactivation is found over 14 Mb away from the start of the current breakpoint.

Either this X;autosome translocation is unique or the current cytogenetically defined breakpoint is incorrect. The simplest way to resolve this issue is to determine if the region between the current breakpoint and the first gene subject to inactivation is present in one or two copies. This could be accomplished through either DNA FISH or SNP analysis. If the current breakpoint location is correct, the region between the current breakpoint and the first gene subject to inactivation would be an excellent candidate in which to search for a lack of way stations. It is also possible that the breakpoint location of GM01414 is correct and that the spread of inactivation is simply more distinct in this sample.

A confounding factor to using DNA methylation to predict inactivation status in X;autosome translocations is that not all cells which carry the same X;autosome translocation always show the same pattern of inactivation [170]. If some autosomal genes are subject to inactivation in only a subset of cells then the average DNA methylation might not be high enough to be predicted as subject to inactivation. Therefore, genes with average DNA methylation in the uncallable range could be variably inactivated genes which are only inactivated in a subset of cells. The different degrees of spread of inactivation between the six X;autosome translocations studied here could be the result of four different factors. The first is the ability of XCI to spread across a centromere. In the mouse, Xist RNA does not bind to the centromere [316]; the majority of human genes which escape from XCI are located on the other side of the centromere [317] leading to the suggestion that the centromere may act as barrier to the spread of inactivation on the X chromosome. Therefore, it is interesting to note that GM05396, which showed the lowest degree of inactivation, is a dicentric chromosome in which inactivation would be required to cross not one, but two centromeres to spread inactivation. The second factor which may influence the degree to which inactivation spreads is the distance from *XIST*. There may be a maximum distance that XIST RNA is able to spread. If this were the case then X;autosome translocations with large autosomal portions might not show as strong a spread of inactivation as those involving small autosomes. This, however, seems unlikely since the inactivation in GM01414 spreads over 140 Mb from XIST to the farthest autosomal gene predicted to be subject to inactivation.

The third factor which may be influencing the degree to which inactivation spreads on the autosomal portion of the X;autosome translocation is secondary selection. In order to maintain the most normal expression pattern, when the autosomal portion of the X;autosome translocation is disomic there may be selection against cells in which extensive silencing occurs. Conversely, when the autosomal portion of the X;autosome translocation is trisomic, selection may work against cells in which minimal silencing occurs. When silencing is able to spread extensively into the trisomic autosomal portion of an X;autosome translocation, the

negative phenotype associated with that trisomy can be minimized [172]. Indeed, since most autosomal trisomies are not viable, the ability of X;autosome translocations to exist with trisomic autosomal portions speaks to the ability of inactivation to achieve a more normal expression pattern.

The fourth factor which may determine the degree to which inactivation spreads along the chromosome is the repetitive element composition at the breakpoint and at the region being silenced. Previously, a high level of L1 elements on both the autosomal and X-linked portions of the breakpoint have been identified in an X;autosome translocation with extensive silencing [172]. It is therefore conceivable that repetitive element composition of the breakpoint may prevent the spread of silencing. The 500 kb region known to be on the autosomal portion of the X;autosome translocation closest to the refined autosomal breakpoint of GM07503 has 22.2% Alu and 6.2% L1 elements. This lower than average L1 frequency in combination with the higher than average Alu frequency may explain the minimal spread of silencing in this X;autosome translocation. The relationship between repetitive element composition and inactivation status is complicated. A low frequency of LINEs has previously been associated with escape from XCI [88, 133, 134]. In theory, regions with high LINE frequency could come together and form the X_i domain while the regions with low LINE frequency would then be capable of looping outside of the silent X_i domain [137]. Therefore, while the large domains of escape from XCI could be due to a lack of way stations, the domains which contain only single genes which escape from XCI are more likely to contain escape elements similar to *Kdm5c* [138].

Using DNA methylation to predict the inactivation status of autosomal genes in X;autosome translocations is a natural extension of using DNA methylation to predict the XCI status of X-linked genes. The comparison of six X;autosome translocations has revealed overall variability in the degree of inactivation with the highest degree of inactivation close to the breakpoint. A comparison of the frequency of repetitive elements showed that the domains which contain multiple genes which escape from inactivation are depleted below the genome average for L1 and LTR elements, whereas domains of only single escape genes are not depleted for these elements. The difference in sequence composition between large domains which escape from inactivation and smaller domains which escape suggest differences in the means by which escape from inactivation is accomplished and will require further analysis to determine how escape from inactivation is achieved in each domain.

Table 5.1: Samples used in chapter 5.

ID	Karyotype	Cell type
GM08399	46, XX	fibroblast
HFF	46, XY	fibroblast
GM07503	46,X,der(X)(Xpter>Xq28::2p21>2pter)mat	fibroblast
GM01414	46,X,+der(9)(9pter>9q34::Xq13> Xqter)mat	fibroblast
GM00074	47,Y,t(X;14) (Xpter>Xq13::14q32>14qter;14pter>14q32::Xq13>Xqter), +der(14)(14pter>14q32::Xq13>Xqter)mat	fibroblast
GM01730	46,XX,der(21)(21qter>21p11::Xq11>Xqter)mat	fibroblast
GM08134	46,X,der(X)(Xpter>Xq22.3:: 21q11>21qter)mat	fibroblast
GM05396	45,X,der(22)t(X;22)(Xqter> Xp11::22p12>22qter)	fibroblast

Table 5.2: Training set of genes which are subject to XCI.

All genes were previously found to be subject to XCI by DNA methylation analysis [308] and somatic cell hybrid analysis [106].

<i>ACSL4</i>	<i>DKC1</i>	<i>HDAC6</i>	<i>MOSPD2</i>	<i>PLP2</i>	<i>SLC10A3</i>	<i>TMEM185A</i>
<i>AIFM1</i>	<i>DLG3</i>	<i>HMGB3</i>	<i>MPP1</i>	<i>PLXNA3</i>	<i>SLC16A2</i>	<i>TMEM47</i>
<i>AMMECR1</i>	<i>DNASE1L1</i>	<i>HPRT1</i>	<i>MST4</i>	<i>POLA1</i>	<i>SLC25A14</i>	<i>TMLHE</i>
<i>AR</i>	<i>DYNLT3</i>	<i>HSD17B10</i>	<i>MTMR1</i>	<i>PORCN</i>	<i>SLC35A2</i>	<i>TRMT2B</i>
<i>ARAF</i>	<i>EBP</i>	<i>HTATSF1</i>	<i>NAP1L2</i>	<i>PQBP1</i>	<i>SLC9A6</i>	<i>TRO</i>
<i>ARMCX1</i>	<i>EDA</i>	<i>IDH3G</i>	<i>NDUFA1</i>	<i>PRAF2</i>	<i>SLC9A7</i>	<i>TSC22D3</i>
<i>ARMCX3</i>	<i>EFNB1</i>	<i>IDS</i>	<i>NDUFB11</i>	<i>PRDX4</i>	<i>SLITRK2</i>	<i>TSPAN6</i>
<i>ARMCX5</i>	<i>ELF4</i>	<i>IGBP1</i>	<i>NGFRAP1</i>	<i>PRICKLE3</i>	<i>SMARCA1</i>	<i>TSPYL2</i>
<i>ARMCX6</i>	<i>EMD</i>	<i>IRAK1</i>	<i>NKAP</i>	<i>PRPS2</i>	<i>SMS</i>	<i>TSR2</i>
<i>ATP6AP1</i>	<i>FAM122C</i>	<i>KLF8</i>	<i>NKRF</i>	<i>RAB39B</i>	<i>SNX12</i>	<i>UBE2A</i>
<i>ATP6AP2</i>	<i>FAM127A</i>	<i>LAGE3</i>	<i>NLGN3</i>	<i>RAB9B</i>	<i>SRPK3</i>	<i>UBL4A</i>
<i>ATRX</i>	<i>FAM3A</i>	<i>LAMP2</i>	<i>OCRL</i>	<i>RAP2C</i>	<i>SRPX</i>	<i>UBQLN2</i>
<i>BCORL1</i>	<i>FAM50A</i>	<i>LANCL3</i>	<i>OPHN1</i>	<i>RBM3</i>	<i>STAG2</i>	<i>USP51</i>
<i>BHLHB9</i>	<i>FAM70A</i>	<i>LAS1L</i>	<i>OTUD5</i>	<i>RBMX</i>	<i>STARD8</i>	<i>VBP1</i>
<i>C1GALT1C1</i>	<i>FGD1</i>	<i>LDOC1</i>	<i>PCSK1N</i>	<i>RBMX2</i>	<i>SUV39H1</i>	<i>WDR13</i>
<i>CASK</i>	<i>FLNA</i>	<i>LONRF3</i>	<i>PCYT1B</i>	<i>REPS2</i>	<i>SYN1</i>	<i>WDR45</i>
<i>CCDC22</i>	<i>FTSJ1</i>	<i>MAGT1</i>	<i>PDK3</i>	<i>RGN</i>	<i>SYP</i>	<i>WNK3</i>
<i>CD99L2</i>	<i>FUNDC2</i>	<i>MAP7D2</i>	<i>PDZD4</i>	<i>RPGR</i>	<i>TBC1D8B</i>	<i>WWC3</i>
<i>CDKL5</i>	<i>G6PD</i>	<i>MBTPS2</i>	<i>PGK1</i>	<i>RPL10</i>	<i>TCEAL1</i>	<i>YIPF6</i>
<i>GETN2</i>	<i>GABRE</i>	<i>MCTS1</i>	<i>PGRMC1</i>	<i>RPL36A</i>	<i>TCEAL4</i>	<i>ZC4H2</i>
<i>CHST7</i>	<i>GK</i>	<i>MECP2</i>	<i>PHF16</i>	<i>RPL39</i>	<i>TCEAL8</i>	<i>ZNF41</i>
<i>CSTF2</i>	<i>GLA</i>	<i>MED12</i>	<i>PHF6</i>	<i>RPS6KA3</i>	<i>TFE3</i>	<i>ZNF449</i>
<i>CXorf26</i>	<i>GNL3L</i>	<i>MID1IP1</i>	<i>PHF8</i>	<i>SCML1</i>	<i>TIMM17B</i>	<i>ZNF673</i>
<i>CXorf40A</i>	<i>GPC4</i>	<i>MID2</i>	<i>PHKA1</i>	<i>SCML2</i>	<i>TIMM8A</i>	
<i>CXorf57</i>	<i>GPRASP2</i>	<i>MOSPD1</i>	<i>PIM2</i>	<i>SEPT6</i>	<i>TMEM164</i>	

Table 5.3: Training set of genes which escape from XCI.

All genes were previously found to escape from XCI by DNA methylation analysis [308] and somatic cell hybrid analysis [106].

<i>AP1S2</i>	<i>CXorf15</i>	<i>FUNDC1</i>	<i>KAL1</i>	<i>PNPLA4</i>	<i>SYAP1</i>
<i>ARSD</i>	<i>CXorf38</i>	<i>GEMIN8</i>	<i>KDM5C</i>	<i>PRKX</i>	<i>TBL1X</i>
<i>CA5B</i>	<i>DDX3X</i>	<i>GPM6B</i>	<i>L1CAM</i>	<i>RAB9A</i>	<i>TRAPPC2</i>
<i>CDK16</i>	<i>EIF1AX</i>	<i>GYG2</i>	<i>MED14</i>	<i>RIBC1</i>	<i>UBA1</i>
<i>CTPS2</i>	<i>EIF2S3</i>	<i>HDHD1A</i>	<i>NLGN4X</i>	<i>RPS4X</i>	<i>ZFX</i>

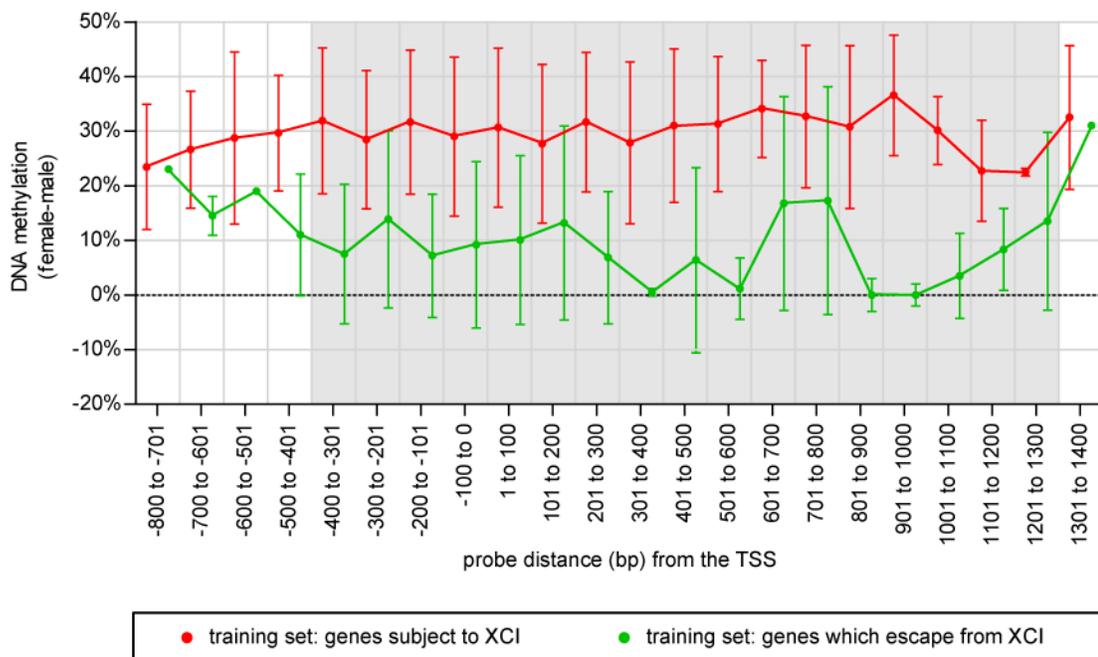


Figure 5.1: X-linked CpG island promoter DNA methylation is influenced by distance from the TSS.

The DNA methylation of normal male and female fibroblasts were subtracted to create a DNA methylation delta for two training sets of genes; those X-linked genes known to be subject to XCI (red) and those X-linked genes which escape from XCI (green). Bins of 100 bp were created surrounding the TSS for all X-linked genes in the training sets and the average DNA methylation delta of all probes located within that bin averaged for each training set. Grey shading highlights probes which most accurately predict XCI. Error bars represent on standard deviation.

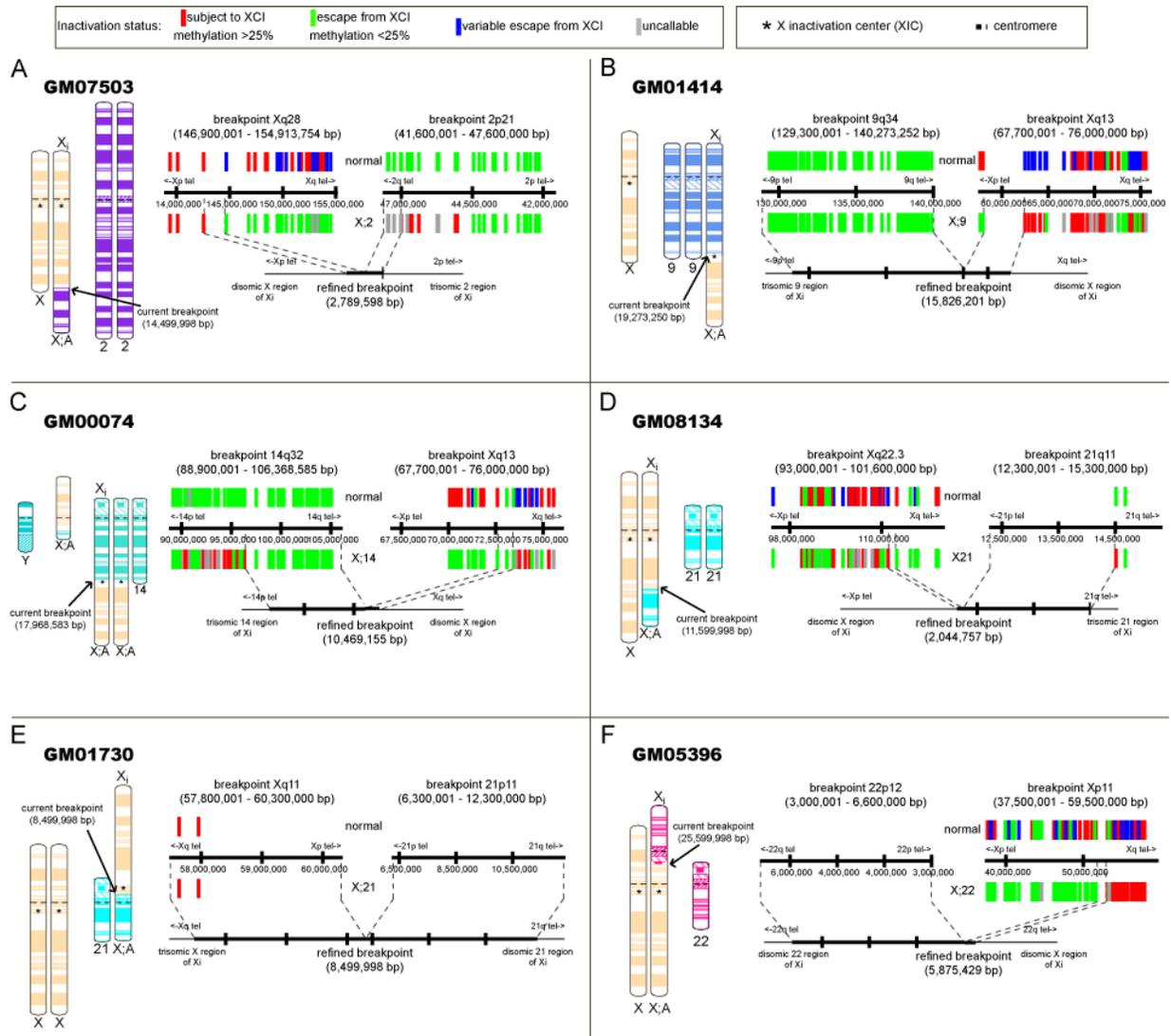


Figure 5.2: Karyograms and breakpoint location analysis of X;autosome translocations.

For each X;autosome translocation (A-F) the karyogram of the of the X chromosome, Y chromosome (if present) and the X;autosome translocation is shown. The chromosome in each sample which is inactivated is labeled X_i and the location of the XIC shown as an asterisk. Current breakpoints were taken from the Coriell website whereas refined breakpoints were determined using DNA methylation of CpG island promoters. The inactivation status of the normal samples is based on the analysis of all samples which did not involve that autosome in the X;autosome translocation. On the X chromosome only samples which were not monosomic were included in the normal XCI status analysis.

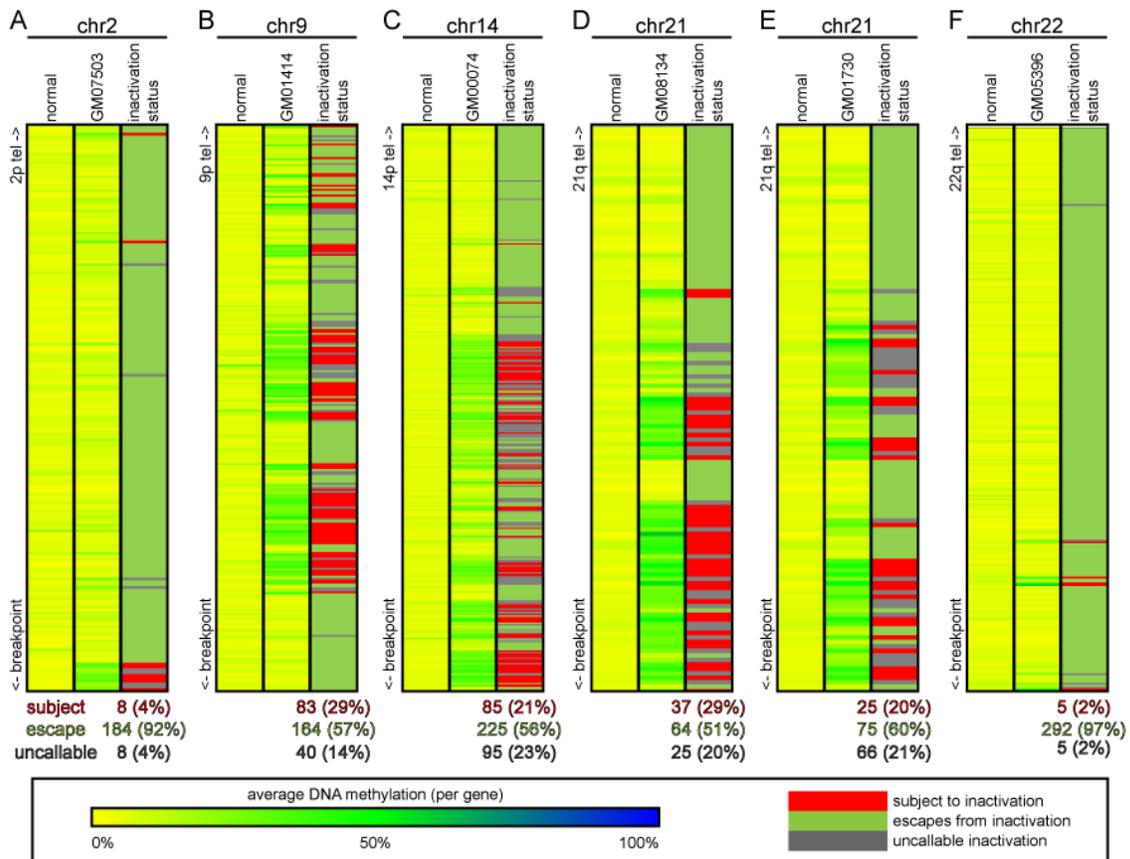


Figure 5.3: Autosomal CpG island promoter DNA methylation suggests different degrees of spread of inactivation.

For each autosome involved in an X;autosome translocation the normal genic average DNA methylation is shown to the left and the average genic DNA methylation of the X;autosome translocation in the middle. The inactivation status for each gene is shown to the right with the number of genes subject to inactivation (red), escaping from inactivation (green) and uncallable (grey) given below each sample. Average genic DNA methylation levels are shown in the order found on the chromosome but the distance between genes is not to scale. DNA methylation is shown on a colour scale from 0% (yellow) to 50% (green) to 100% (blue).

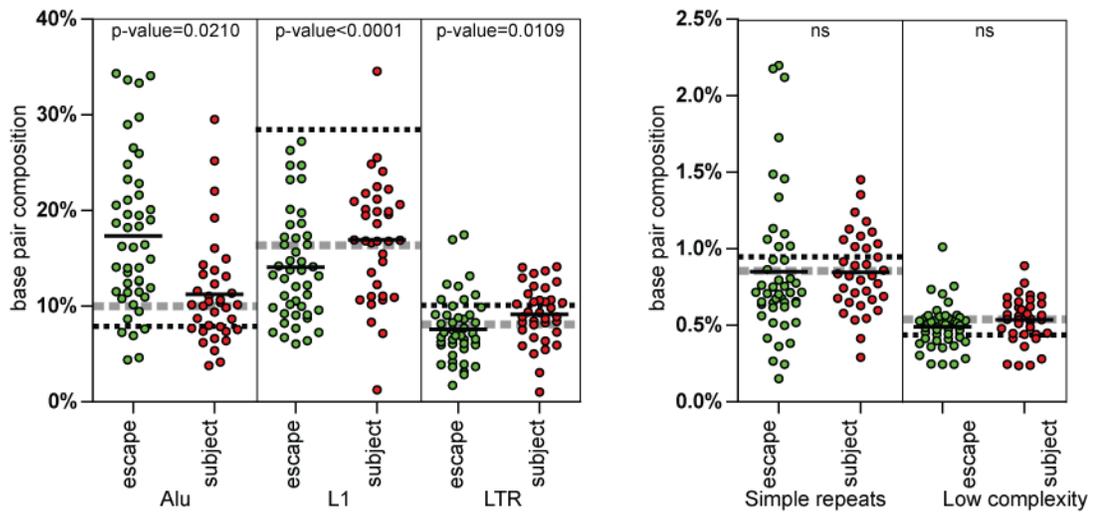


Figure 5.4: Repetitive element frequency differs based on inactivation status.

The frequencies of repetitive elements in domains of genes which escape from inactivation (green, $n=47$) were compared against domains of genes subject to inactivation (red, $n=38$). The genome average for each repetitive element is shown as a dotted grey line and the X chromosome average as a black dotted line. Statistical comparisons were performed using a Mann-Whitney test and p-values of significant differences given above each element. Domains of genes which escape from inactivation were on average 3.2 Mb (standard deviation = 5.1 Mb), domains of genes subject from inactivation were on average 1.8 Mb (standard deviation = 2.2 Mb).

6 XCI in human triploid placentas

6.1 Introduction

Triploidy is the presence of three haploid chromosome sets instead of the two sets present in normal diploid cells. Triploidy is rarely observed in liveborns; however, it accounts for 10% of spontaneous abortions [207]. In order to achieve correct dosage compensation, diploid cells appear to follow the 'n-1' rule in which one X chromosome per haploid autosome set remains active [318]. For example, when one extra X chromosome is present in a cell, as is the case in 47, XXX females and 47, XXY males, one X chromosome remains the X_a and all other X chromosomes become X_i s (for example, 47, $X_aX_iX_i$ and 47, X_aX_iY) [158, 225]. Two sets of autosomes result in one X_a and artificially generated tetraploid mice have two X_a s [319, 320]. Unlike diploids and tetraploids, triploids do not appear to have a "correct" number of X chromosomes to keep active. Therefore triploids offer a unique opportunity to study the means by which a cell counts and inactivates X chromosomes. To study the process of count in triploids it is necessary to determine how many X chromosomes have been inactivated in a triploid sample.

Human triploids demonstrate a variety of XCI patterns (Table 1.4 and Table 1.5), with the majority of somatic cells having two X_a s regardless of sex. The pattern of XCI in triploids appears to be associated with several factors including length of survival, time in culture and tissue examined. While triploids which survive to term tend to have more X_i s than triploids which do not [228-230], the longer triploid cells are in culture, the more X_a s are present [225, 228, 231]. Different patterns of XCI have also been detected across different tissues from the same individual triploid [230, 233]. Only one study has examined XCI patterns in triploid extra-embryonic tissues and this report demonstrated a higher number of X_i s in extra-embryonic tissues compared to somatic triploid tissues, with older extra-embryonic tissues having more X_i s compared to younger extra-embryonic tissues [226]. Historically the number of X_i in a triploid sample has been determined through Bromodeoxyuridine (BrdU) labeling [226, 228, 230, 233] or rarely through XIST RNA FISH [225]. Both of these techniques require actively growing triploid cultures which may not always be easy to obtain and/or maintain. A technique which could determine the number of X_i s in a triploid sample solely based on DNA analysis would therefore make triploids a more desirable system in which to study the process of count and its role in XCI.

The DNA methylation status of only one X-linked CpG island promoter (*G6PD*) has previously been reported in human triploids. The same trend of DNA methylation, that is, unmethylated X_a s and methylated X_i s (MeXIP), was observed as in normal females [225]. In chapter 2, the DNA methylation of X-linked promoters with CpG islands was established to accurately predict the XCI status of X-linked genes. Due to the variety of XCI patterns observed across different

triploid cells it was not possible to determine the XCI status of individual X-linked genes in triploid samples. Instead, we examined whether MeXIP, which is found at genes subject to XCI, could be used to determine the overall number of X_i s present in a triploid sample. As demonstrated in chapters 3 and 4, the placenta is hypomethylated compared to somatic tissues at a variety of elements including repetitive elements and X-linked CpG island promoters in females. Therefore, in order to examine triploid placental samples it was necessary to compare the DNA methylation of triploid placentas with diploid placentas, not diploid somatic tissues. Using X-linked DNA methylation to determine the number of X_i s in a triploid samples would be of benefit in cases where only a DNA sample is available and would not require the laborious work of culturing cells, which also induces culture artifacts, and requires counting the number of X_i s by XIST RNA FISH or replication timing. Knowing the number of X_i s in a triploid cell would allow for comparisons between male and female as well as between diandric and digynic triploids and may provide insight into how the number of X chromosomes to be inactivated is determined.

6.2 Methods

6.2.1 Sample collection, processing and Illumina Infinium HumanMethylation27 array

Collection of samples was approved by the ethics committees (H06-70085) of the University of British Columbia and the Children's and Women's Health Centre of British Columbia. Triploid sample recruitment and collection was performed by Dr. Deborah McFadden. Parent of origin and gestational age were previously published [206, 321]. DNA extraction on whole villi was performed by the Robinson lab as previously described in section 2.2.1. Table 4.1 lists control (average gestational age = 25 weeks) placental samples and Table 6.1 the triploid placentas (average gestational age = 11 weeks) used. Processing of the Illumina Infinium HumanMethylation27 array was performed by the Robinson lab as previously described in section 2.2.2. DNA methylation analysis of these triploids samples also has previously been published [3].

6.2.2 Predicting % X_i in triploids with DNA methylation from the Illumina Infinium HumanMethylation27 array

Since the placenta shows considerable variability in the level of DNA methylation at X-linked island promoters, we took steps to identify maximally informative CpG sites rather than examine all X-linked probes on the Illumina Infinium HumanMethylation27 array. Four steps were taken in order to determine which X-linked probes from the Illumina Infinium HumanMethylation27 array were best at predicting the % X_i (percentage of X chromosomes which are X_i s) in control

placentas. Figure 6.1 outlines these steps as well as the remaining number of probes after each step. First, all probes from the Illumina Infinium HumanMethylation27 array on the X chromosome which were located in repetitive elements, mapped to the autosomes as well as the X chromosome and/or were located in the promoters of CT genes were removed from further analysis. Second, the decision tree (Figure 2.1) established in chapter 2 was used to predict the XCI status of the remaining probes in control placenta. Only probes which predicted that a gene was subject to XCI were kept because these probes demonstrated no overlap between the male and female DNA methylation levels as well as an average difference of 22% between male and female DNA methylation. These probes were the most likely to accurately predict differences in the number of X_i s. The average male and female control placental DNA methylation was then used to calculate the average DNA methylation on the X_a and X_i for these probes. Using the calculated X_a and X_i DNA methylation, $\%X_i$ was calculated for each individual female control placenta. Given the range of DNA methylation levels observed across female controls, the acceptable range of $\%X_i$ s was defined to be between 40% X_i and 60% X_i with the expected range between 45% X_i and 55% X_i . A $\%X_i$ greater than 100% X_i or less than 0% X_i was theoretically impossible and therefore deemed a failure. The third step was to remove any probes in which one or more control female placenta had a $\%X_i$ s in the failure range. After completion of steps one to three, there were 29 probes which remained as candidates for determining the $\%X_i$ in placenta. This third step was to determine which subset of these 29 probes was most accurate at determining that control female placentas had 50% X_i (all cells 46, X_aX_i) and control male placentas had 0% X_i s (all cells 47, X_aY). All possible combinations of probes were tested and it was determined that those seven probes (listed in Table 6.2) in which at least four female control placentas fell within the expected $\%X_i$ range (45% X_i to 55% X_i) were the most accurate at predicting the expected $\%X_i$ in male and female placentas.

6.2.3 Statistical analysis

Statistical comparisons of the predicted $\%X_i$ s were done using the Mann-Whitney test as calculated by GraphPad Prism. The correlations between $\%X_i$ s determined using the Illumina Infinium HumanMethylation27 array versus pyrosequencing were done using the Pearson test as calculated by GraphPad Prism. For both tests, a p-value greater than 0.05 was not significant, however, p-values between 0.01 and 0.05 (*), between 0.01 and 0.001 (**), and p-values <0.001 (***) were all considered significantly different.

6.2.4 Predicting $\%X_i$ from pyrosequencing

Pyrosequencing was performed as previously outlined in section 3.2.4. Primer sequences and conditions for the assays, PDK3, FANCB and TBL1X are given in Table 3.1. The $\%X_i$ for

samples which were analyzed by both pyrosequencing and the Illumina Infinium HumanMethylation27 array were averaged to create a single $\%X_i$ which is shown on Figure 6.2.

6.2.5 XIST RNA FISH

RNA FISH of XIST was performed on GM04939, a female triploid cell line, as outlined in section 4.2.4. The number of XIST signals per cell was counted for 100 cells (300 X chromosomes) by eye.

6.3 Results

6.3.1 DNA methylation can accurately predict the percent of X chromosomes which are X_i s in control placentas

Determining the XCI status of triploids is challenging because, unlike 46, XX females, in which there is an expectation that there will be one X_a and one X_i per cell, triploids can show a variety of XCI patterns within different cells of the sample individual. Therefore, rather than predict XCI patterns we investigated what percentage of triploid X chromosomes were X_i s. To predict the $\%X_i$ we used the knowledge that X-linked promoters with CpG islands subject to XCI are methylated on the X_i but unmethylated on the X_a (MeXIP). Although MeXIP is the general trend at X-linked island promoters subject to XCI, not all individual CpGs are as consistent in the degree to which MeXIP is shown. Therefore, it was important to select those CpGs which were most accurate at predicting the expected percentage (male: 0%, female: 50%) of X_i s in control placenta samples (list in Table 4.1) before the number of X_i s was predicted in triploid placentas. (Figure 6.1 outlines the steps taken to determine the $\%X_i$ as well as the available number of candidate probes after each step). Seven X-linked probes from the Illumina Infinium HumanMethylation27 array were identified as most accurate at predicting 50% X_i in female control placentas and 0% X_i in male placentas (orange and blue triangles in Figure 6.2). As expected, the $\%X_i$ was significantly different between male and female control placentas (male average: 0% X_i & female average: 50% X_i , p-value<0.001). Having determined which X-linked probes were most accurate at determining the $\%X_i$ in control placentas, we went on to predict the percentage of X chromosomes which were X_i s in triploid placentas.

6.3.2 $\%X_i$ determined by Illumina agrees with $\%X_i$ determined by pyrosequencing

The $\%X_i$ in 20 triploid placental samples (listed in Table 6.1) was determined using the seven assays listed in Table 6.2 (leftmost symbols in each triploid category on Figure 6.2). Not enough samples were available for a robust comparison across gestational age, however, samples are marked with different symbols to denote gestational age in Figure 6.2 (first trimester triploids: diamonds, second trimester triploids: circles). Although the Illumina Infinium HumanMethylation27 array examined over 1000 X-linked probes and provided a means to

examine multiple X-linked and autosomal promoters, we were interested if single pyrosequencing assays might also be able to determine the %X_is in triploid samples. To examine this possibility we performed pyrosequencing using three X-linked assays (PDK3, FANCB and TBL1X) previously used in chapter 3. The DNA methylation profiles of six triploid samples also analysed on the Illumina Infinium HumanMethylation27 array were determined and the %X_i calculated as outlined in steps 2A and 2B on Figure 6.1. We next compared the %X_i determined by each pyrosequencing assay against the %X_i determined by the seven X-linked assays from the Illumina Infinium HumanMethylation27 array in section 6.3.1. There was a significant correlation between the %X_i predicted by the seven Illumina assays and the pyrosequencing results for PDK3 ($r^2 = 0.7666$, $p\text{-value}=0.0223$) and FANCB ($r^2=0.7693$, $p\text{-value}=0.0217$) whereas TBL1X was not significantly correlated (Figure 6.3A). Used alone, the PDK3 pyrosequencing assay tended to overestimate the %X_i while the FANCB pyrosequencing assay underestimated the %X_i. The average %X_i predicted using the average of PDK3 and FANCB correlated with the %X_i predicted based on the seven assays from the the Illumina Infinium HumanMethylation27 better and to a higher degree of significance than either the PDK3 or FANCB assay alone ($r^2=0.8950$, $p\text{-value}=0.0043$) (Figure 6.3B). Therefore, we concluded that it was preferable to use the average from two pyrosequencing assays (PDK3 and FANCB) to predict the %X_i in triploid samples rather than use an entire Illumina Infinium HumanMethylation27 array. We were thus able to examine and determine the %X_i in an additional seven triploid placental samples (rightmost symbols in each triploid category on Figure 6.2).

6.3.3 Triploid placentas tend to have more than the expected average number of X_is

Through a combination of the Illumina Infinium HumanMethylation27 array and pyrosequencing, the %X_i was predicted for 27 triploid placentas. All triploid placentas, as well as control female placentas, had a significantly different %X_i from control male placentas ($p\text{-value}<0.001$) meaning that as a whole, each triploid placenta contained at least one X_i. Control female placentas (46, XX) and male triploid placentas (69, XXY) had two X chromosomes, however, while control females are always 46, X_aX_i, male triploids could in theory be 69, X_aX_aY, 69, X_aX_iY or a mosaic of both 69, X_aX_aY and 69, X_aX_iY cells. One male triploid, TP-2, was within the range of %X_i observed in control male placentas suggesting that this triploid has few, if any, X_is. Only one male triploid, TP-23, was above the theoretical maximum of 50% X_i in male triploids. No significant difference in the %X_is between male diandric and digynic triploids was observed. Together these data suggest that the majority of male triploid placentas show varying degrees of mosaicism for 69, X_aX_aY and 69, X_aX_iY cell lines but with more 69, X_aX_iY cells than 69, X_aX_aY

cells, regardless of parent of origin of the extra haploid chromosome set. Unlike control female placentas (46, XX), female triploid placentas (69, XXX) have the potential to have two X_i s and could therefore be pure 69, $X_aX_aX_i$, pure 69, $X_aX_iX_i$ or a mosaic of these two XCI patterns. Diandric female triploids (62% X_i) had a significantly (p -value=0.0238) higher % X_i s than control female placenta (50% X_i). Digynic female triploids were not significantly different in their % X_i (54% X_i) than control female placentas (50% X_i), however, digynic female triploids showed a range of % X_i twice as large as control female placentas. This suggests that while digynic female triploids generally have approximately 50% X_i s, i.e. mosaicism for 69, $X_aX_aX_i$ and 69, $X_aX_iX_i$ cell lines, there is considerable variation between different triploids as to the proportion of 69, $X_aX_aX_i$ and 69, $X_aX_iX_i$ cells.

6.3.4 XIST RNA FISH confirms the % X_i predictions from DNA methylation in cultured triploid somatic cell line

To determine whether the % X_i predicted using X-linked DNA methylation was accurate, we analyzed an additional triploid sample using both the pyrosequencing assays PDK3 and FANCB as well XIST RNA FISH to count the number of X_i s directly. It was not possible to obtain a triploid placental sample for culture so a female triploid somatic sample was used. Therefore, control male and female blood, not placentas, were used to establish the average X_a and X_i DNA methylation and then to calculate the % X_i . The pyrosequencing assays PDK3 and FANCB were used to predict a % X_i of 54% in the triploid somatic samples. XIST RNA FISH was then performed and 100 cells (300 X chromosomes) counted. 48% of cells were 69, $X_aX_aX_i$ and 52% were 69, $X_aX_iX_i$, examples of 69, $X_aX_aX_i$ and 69, $X_aX_iX_i$ cells are shown in Figure 6.4. The % X_i determined by XIST RNA FISH was 51% X_i which is very similar to the 54% X_i found by converting the average DNA methylation found at the PDK3 and FANCB pyrosequencing assays are an accurate means of determining the % X_i in a triploid sample.

6.4 Discussion

The link between the level of DNA methylation and X-linked CpG islands in triploids has previously only been examined at the *G6PD* promoter and only in somatic cells [225]. It was concluded that the DNA methylation at this X-linked CpG island promoter in triploids showed similar X_a (unmethylated) and X_i (fully methylated) DNA methylation patterns to those observed in normal 46, XX females (MeXIP). We exploited the differences in the DNA methylation levels of the X_a and X_i as a means to predict the percentage of X chromosomes which were X_i s (% X_i s) in human triploid placentas. Our analysis predicted that human triploid placentas generally have more X_i s than expected by chance alone. This is in contrast to previous studies of somatic triploid tissues which found mostly 69, $X_aX_aX_i$ cells in females and 69, X_aX_aY cells in males (see Table 1.4 and Table 1.5 for details). Only one other study has examined the XCI

status of extra-embryonic tissues (a combination of amnion, chorion and villi) in human triploids and a translation of those results from into %X_i allows for a more direct comparison with our current study [226]. Jacobs *et al.* [226] determined that triploid fetuses (male: 6% X_i, female: 31% X_i) had a lower %X_i than triploid extra-embryonic tissues (male: 16% X_i, female: 38% X_i). The extra-embryonic tissue in this chapter represent placental villi with a higher %X_i (male: 35% X_i, female: 55% X_i) than Jacobs *et al.* observed in either triploid somatic or extra-embryonic tissues.

Several factors could contribute to the observed differences in the %X_i between studies. The first is the difference in the techniques used to determine the number of X_is. Jacobs *et al.* [226] used the late replicating nature of the X_i to count the number of X_is whereas we used X-linked CpG island promoter DNA methylation. The discordance between studies could be due to an underestimation of %X_i due to poor BrdU staining and/or due to an overestimate of the %X_i by the selection of X-linked promoters which undergo hypermethylation in triploid placentas. A second possible issue is the use of different sample types. We used uncultured whole villi samples, the outer layer of which is derived from the trophoblasts while the inner layer is derived from the inner cell mass. Jacobs *et al.* [226] used a combination of cultured villi, amnion and chorion to represent extra-embryonic tissues. Amnion and chorion are derived from the inner cell mass and would therefore be expected to have DNA methylation patterns similar to somatic tissues [322]. Different tissues are known to have different levels of DNA methylation and, as discussed in chapter 2, DNA methylation on the X chromosome in placenta shows a unique DNA methylation pattern. Even if the same tissue types were initially collected, not all cells grow in culture and the composition of cells, and therefore DNA methylation, would differ before and after culture. One would therefore expect the amnion and chorion samples to have a %X_i more similar to fetal triploid samples, which are also derived from the inner cell mass, which would decrease the average %X_i observed in the Jacobs *et al.* [226] extra-embryonic tissue category, possibly accounting for the differences in %X_i determined in this study.

The third possible confounding factor in comparing this and other studies involves the age of samples used. Roughly two thirds of the placentas in this study were first trimester and the remaining third were second trimester (see Table 6.1). This is opposite to the Jacobs *et al.* [226] study where 29% were first trimester and the majority (71%) were second trimester. Jacobs *et al.* [226] found that older extra-embryonic tissues had more X_is than younger tissues. Although neither this study nor Jacobs *et al.* [226] observed a significant difference between the first and second trimester %X_is, in this study, second trimester male triploid placentas had, on average, over 1.6 times higher %X_i than first trimester placentas but this difference was not

statistically significant (p -value > 0.01). Interestingly, neither Jacobs *et al.* [226] nor this study observed a difference between first and second trimester in the average $\%X_i$ in female triploid placentas. Although similar trends were observed between this and previous studies, taken together these three possible confounding factors help explain the observed differences in the predicted $\%X_i$.

In somatic triploid cells, Gartler *et al.* found that increased time in culture resulted in an increase in the number of X_a s. Jacobs *et al.* previously found that two X_a s conferred a growth advantage to the triploid fetus [225, 226]. To investigate whether the presence of more X_a s conferred a growth advantage to triploid placentas, we examined the placentas of diandric triploids, which tend to be large and cystic, might have more X_a s (a lower $\%X_i$) than the placentas of digynic triploids. In both male (69, XXY) and female (69, XXX) placentas no significant difference in the $\%X_i$ between observed based on parent of origin suggesting that if the presence of X_a s confers a growth advantage in the placenta it is not large. If triploid cells have different growth potentials as a result of their XCI patterns, then there might be a selective advantage for cells in which an X_i is reactivated and becomes an X_a . Somatic triploid cells maintain a stable XCI pattern [225]; however, given the decreased stability of XCI in placenta (discussed in chapter 4), the placentas of triploids may also show a higher degree of escape from XCI than somatic triploid tissues.

Since triploid cells do not always inactivate the same number of X chromosomes, male and female triploids can possess a number of different XCI patterns. For example, although female triploids are capable of inactivating two X chromosomes (69, $X_aX_iX_i$), male triploids never appear to be 69, X_iX_iY . Previous studies have examined over 1500 69, XXY cells and an 69, X_iX_iY cell has never been observed [226, 248]. If an autosomal factor is involved in the counting step of XCI and helps to determine how many X chromosomes should be inactivated, then given that male and female triploids have the same autosomal complement, the lack of 69, X_iX_iY cells must be due to selection against cells which contain no X_a s. Similarly, no female triploid cells with three X_i s have been observed. Interestingly, female triploids cells which contain only X_a s do appear to be viable as Jacobs *et al.* found 4% of female triploid cells contained only X_a s [226]. If male triploids never have 69, X_iX_iY cells and 69, X_aX_aY and 69, X_aX_iY cells occur at the same frequency that would give a theoretical average of 25% X_i . If female triploids never have 69, $X_iX_iX_i$ cells, but do have 69, $X_aX_aX_a$ cells, then if all combinations of X_a and X_i were present in equal amounts then theoretical average $\%X_i$ would be 33%. In this study, the predicted $\%X_i$ was greater than the theoretical average $\%X_i$ in both male and female triploid placentas, highlighting a selective advantage associated with the presence of more X_i in the triploid placenta.

Of the 27 triploid placentas examined only one male triploid, and two female triploids were above the theoretical maximum $\%X_i$. The male triploid (TP23) had a high $\%X_i$ due to unusually high DNA methylation at FANCB. Excluding FANCB, TP23 is 43% X_i which is within the range observed for the other male triploid placentas. The two females (TP74 and TP9) which were above the theoretical maximum $\%X_i$ were above by 3% and 8%. Given that some control female placentas showed predicted $\%X_i$ s more than 10% above 50% we conclude that TP74 and TP9 are likely composed of mostly 69, $X_aX_iX_i$ cells and the assessment above 66% is due to the sensitivity of the system. The range of $\%X_i$ in control placentas (male: 21% X_i , female: 22% X_i) shows that even when all samples are known to have the same actual $\%X_i$ (male: 0% X_i , female: 50% X_i) a wide range of predicted $\%X_i$ are found. Therefore, we caution that while the general trends identified here are useful, the exact $\%X_i$ may vary +/- 10% for any individual sample.

Predicting $\%X_i$ based on the DNA methylation at only a subset of CpGs may seem less accurate than using the average DNA methylation level across a larger number of CpGs. Further analysis of the bisulfite sequencing of the *G6PD* promoter examined by Gartler *et al.* demonstrated that seven of the 51s CpG were unmethylated on all X_a s and methylated on all X_i s [225]. It is these CpG that would, therefore, be most accurate at predicting the $\%X_i$ at the *G6PD* promoter. A comparison of $\%X_i$ predicted by these seven individual CpGs and the $\%X_i$ predicted using the average across all CpGs, reveals that the $\%X_i$ from the seven individual CpGs was, on average, only 5% off the $\%X_i$ determined by counting XIST signals whereas $\%X_i$ from all CpGs was, on average, 13% off. Clearly there is an advantage to finding those CpGs which show the strongest MeXIP, as was done here (Figure 6.1), and using only these CpGs to predict the $\%X_i$.

In this chapter, the study of X-linked DNA methylation was extended to triploid placentas to illustrate that X-linked DNA methylation can accurately predict the number of X_i s present in triploid samples. DNA methylation analysis at two X-linked pyrosequencing assays was demonstrated to be as accurate at predicting the number of X_i s in triploid samples as analysis on the Illumina Infinium HumanMethylation27 array. It also established that triploid placental samples have more X_i s than would be expected by chance alone despite the high degree of mosaicism for XCI patterns.

Table 6.1: Triploid placental samples used in chapter 6.

The type (based on parent or origin analysis) and gestation age of triploids have been previously described [206, 321].

sample name	sex	type	gestational age (weeks)	trimester	tissue type
TP1*#	Female	Digynic triploidy	<10	first trimester	villi
TP2#	Male	Diandric triploidy	<12	first trimester	villi
TP3*#	Female	Digynic triploidy	<10	first trimester	villi
TP4#	Female	Digynic triploidy	<12	first trimester	villi
TP5#	Female	Diandric triploidy	<12	first trimester	villi
TP6*#	Female	Diandric triploidy	8	first trimester	villi
TP7*#	Male	Diandric triploidy	<10	first trimester	villi
TP9*#	Female	Diandric triploidy	13	second trimester	villi
TP11#	Male	Diandric triploidy	19	second trimester	villi
TP20*	Female	Digynic triploidy	<10	first trimester	villi
TP23#	Male	Digynic triploidy	>12	second trimester	villi
TP24*#	Female	Diandric triploidy	13	second trimester	villi
TP25#	Female	Digynic triploidy	>12	second trimester	villi
TP26#	Female	Digynic triploidy	>12	second trimester	villi
TP49*	Male	Diandric triploidy	9	first trimester	villi
TP54*	Male	Diandric triploidy	14.3	first trimester	villi
TP56*	Male	Digynic triploidy	8	first trimester	villi
TP57*	Female	Diandric triploidy	8.4	first trimester	villi
TP58*	Female	Digynic triploidy	8	first trimester	villi
TP60*	Male	Digynic triploidy	9	first trimester	villi
TP61*	Female	Digynic triploidy	12	first trimester	villi
TP69*	Female	Digynic triploidy	9.8	first trimester	villi
TP74*	Female	Diandric triploidy	17	second trimester	villi
TP76*	Female	Diandric triploidy	15	second trimester	villi
TP84*	Female	Digynic triploidy	6-8	first trimester	villi
TP85*	Female	Digynic triploidy	7.4	first trimester	villi
TP86*	Male	Diandric triploidy	15	second trimester	villi

* analyzed by Illumina Infinium HumanMethylation27 array

analyzed by pyrosequencing

Table 6.2: X-linked probes from the Illumina Infinium HumanMethylation27 array determined to be most accurate at determined %X_i.

XCI status in chapter 2 as defined based on the decision tree in Figure 2.1.

gene name	probe location (bp)	probe CpG density	distance from TSS (bp)	Carrel and Willard [106] XCI status	chapter 2 XCI status
<i>SCML1</i>	17,665,226	HC	-284	subject	subject
<i>OTUD5</i>	48,700,069	HC	-232	subject	subject
<i>EFNB1</i>	67,965,634	HC	78	subject	subject
<i>RAB9B</i>	102,973,809	HC	5	subject	subject
<i>ELF4</i>	129,073,190	HC	-1037	subject	subject
<i>CXorf40A</i>	148,430,930	HC	467	subject	subject
<i>PDZD4</i>	152,748,915	HC	282	subject	subject

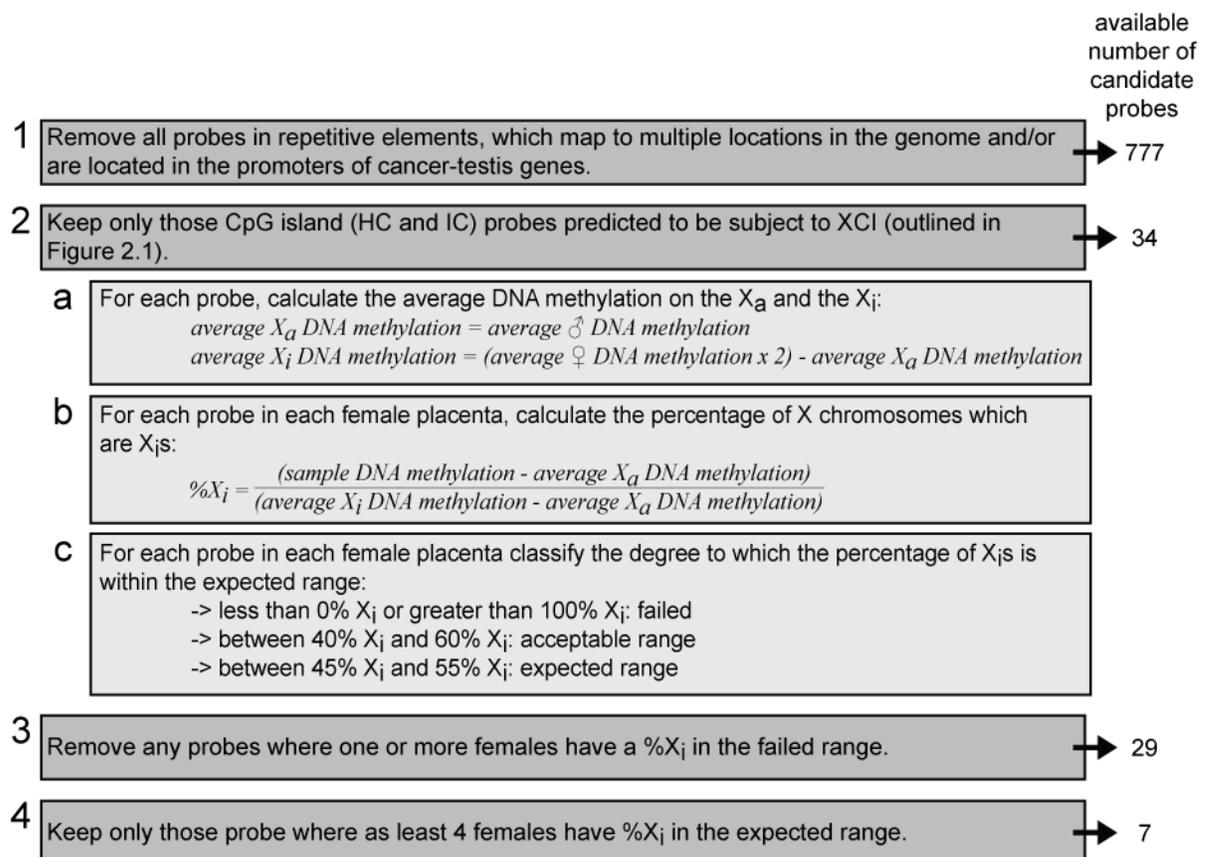


Figure 6.1: Steps used to calculate $\%X_i$ in placenta.

Descriptions of the four steps taken to determine which X-linked probes from the Illumina Infinium HumanMethylation27 array were most accurate at determining the $\%X_i$ in control placentas. To the right of each step is the number of candidate X-linked probes after the step taken to the right.

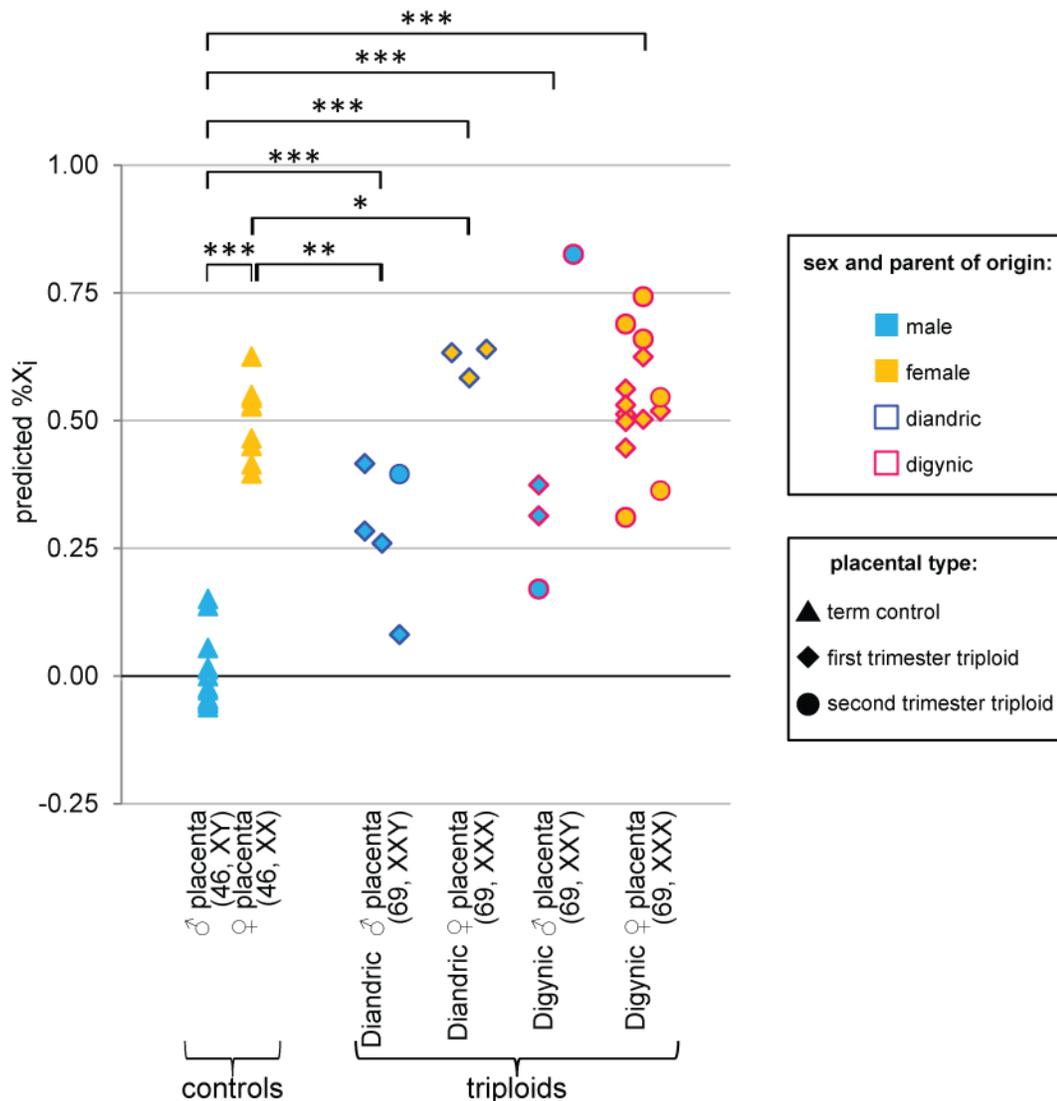


Figure 6.2: %X_i in control and triploid placentas as predicted using X-linked DNA methylation.

The parent of origin of the extra haploid set is shown in the outline of each symbol with blue outlines denoting diandric triploids and pink outlines digynic triploids. The sex of each sample is shown as the internal colour of the symbol (male: blue, female: orange). The age of triploid samples is shown as the shape of the symbol (first trimester: diamond, second trimester: circle). The triploid %X_is predicted using only the Illumina Infinium HumanMethylation27 array (chosen as outline in section 6.3.1) are shown to the left in each category whereas the %X_i predicted using the pyrosequencing average of PDK3 and FANCB are shown to the right of each category. Those triploids which were examined by both techniques are shown as an average %X_i in the middle of p-values between 0.01 and 0.05 (*), between 0.01 and 0.001 (**), and p-values <0.001 (***).

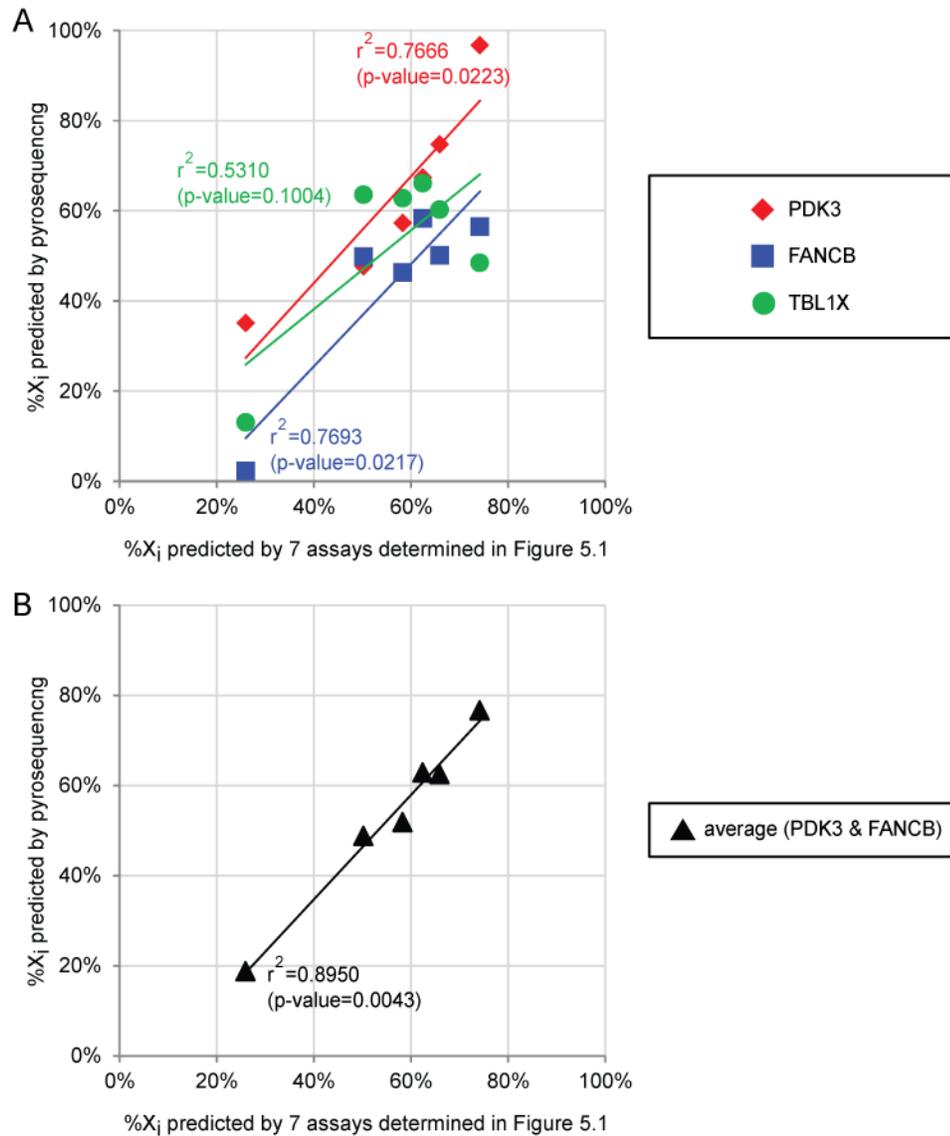


Figure 6.3: Illumina Infinium HumanMethylation27 array and pyrosequencing show correlated levels of predicted %X_i.

Correlation was calculated using the Pearson test with r^2 and p-values given for each pyrosequencing assay. **(A)** The %X_i predicted using the seven probes from the Illumina Infinium HumanMethylation27 array (chosen as outlined in section 6.3.1) was compared against the %X_i predicted using individual pyrosequencing assays. The six triploid samples examined are shown as different symbols for each assay (Red diamonds: PDK3, green circles: TBL1X and blue squares: FANCB). **(B)** The %X_i predicted using the seven probes from the Illumina Infinium HumanMethylation27 array (chosen as outlined in section 6.3.1) was compared against the %X_i predicted using the average of the pyrosequencing assays PDK3 and FANCB. The six triploid samples are shown as black triangles.

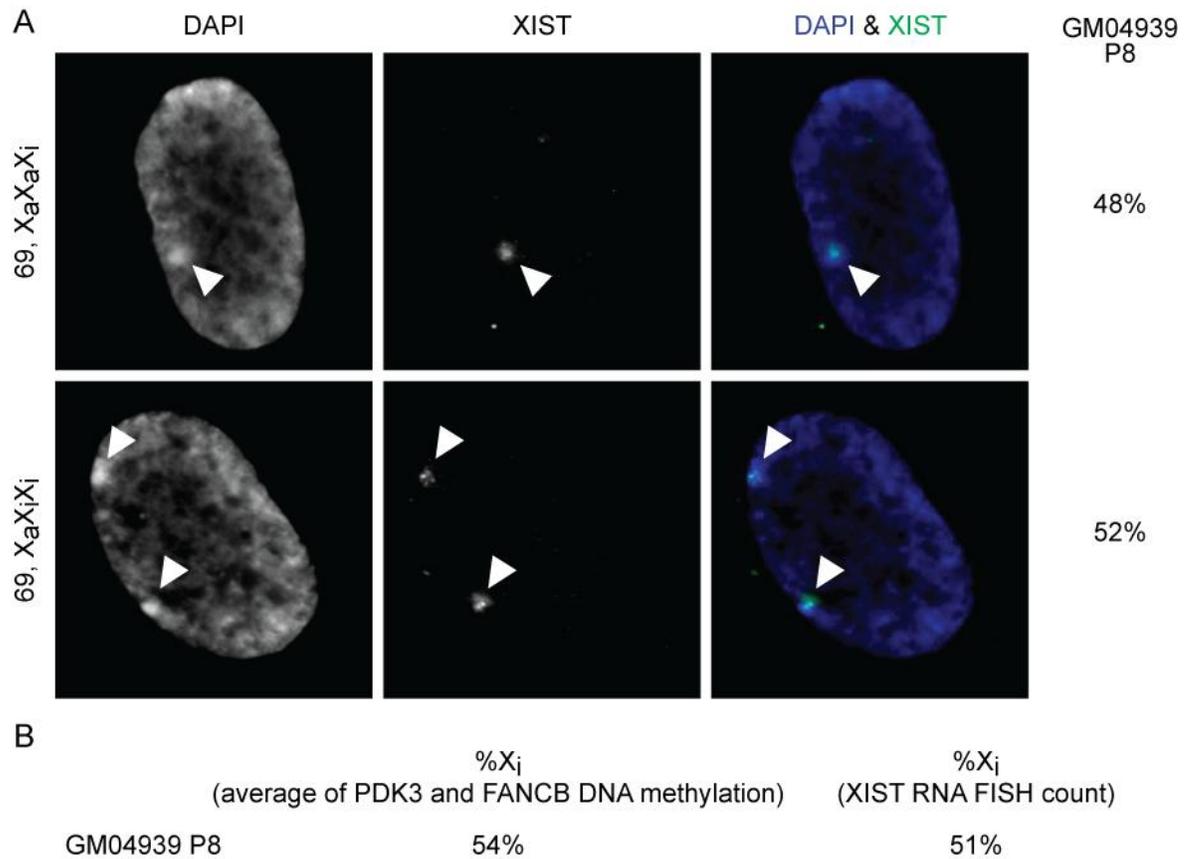


Figure 6.4: XIST RNA FISH confirms %X_i predicted using DNA methylation.

(A) The presence of XIST RNA (green) was used to count the number of X_is present in female triploid cells (69, XXX). An example of a 69, X_aX_aX_i and 69, X_aX_iX_i cell are shown on the top and bottom row of images respectively. DAPI is shown in blue and the number of cells which showed each XCI pattern given below the FISH image. The final %X_i is shown to the far right. (B) The %X_i as calculated by the average of the pyrosequencing assays PDK3 and FANCB is given along with the %X_i calculated by a direct count of X_is using XIST RNA FISH.

7 Discussion

In this thesis, the assumption was made that the DNA methylation of an X-linked CpG island promoter is representative of the XCI status of that gene. In chapter 2 the comparison between the XCI status predicted in somatic cell hybrids [106] and the XCI status predicted by X-linked CpG island promoter DNA methylation (Figure 2.3) provided good evidence that CpG island promoter DNA methylation could accurately predict XCI status. Despite this, a number of genes showed a different XCI status in the somatic cell hybrids compared to blood. These differences were attributed to differences in tissue-specific XCI rather than the inability of DNA methylation to predict XCI status. Cell type-specific differences in XCI have also been observed using expression analysis of XCI status. Of the 92 genes studied in somatic cell hybrids and in fibroblasts [106], 70% demonstrated the same XCI status between the two cell types. Combined with the DNA methylation data in chapter 2, in which tissue-specific DNA methylation differences were identified, this is strong evidence for differences in XCI status occurring between different cell types and hence different tissues. The best way to confirm the accuracy of DNA methylation as a predictor of XCI status would be large scale DNA methylation analysis combined with expression analysis. Future expression analysis should also investigate allele-specific expression, in which one allele is preferentially expressed over the other [323] since X-linked allele-specific expression which may contribute to genes which show different XCI statuses between females. Any future expression analysis should use multiple samples with skewed XCI to increase the likelihood of having heterozygous samples thereby maximizing the number of X-linked genes for which an XCI status can be determined. To further explore the extent to which tissue-specific XCI occurs, it would be ideal to collect multiple tissues from the same female to compare both expression and DNA methylation data. The use of direct tissue samples rather than cultured cells would also minimize the possibility of aberrant DNA methylation caused by culture conditions rather than tissue-specific differences.

Although X-linked CpG island promoter DNA methylation was established in chapter 2 as an accurate means of predicting XCI status, chapter 5 demonstrated that not all CpGs in a CpG island are equal with respect to DNA methylation. Figure 5.1 emphasized that the distance from the TSS greatly affects the level of DNA methylation of X-linked CpG island promoters. Therefore, the more CpGs that can be examined from within the ideal 400 bp upstream and 1300 bp downstream window, the less likely a single CpG with aberrant DNA methylation would affect the prediction of XCI status. For example, in chapter 2 the gene *GPKOW* was predicted to escape from XCI based on DNA methylation at a single CpG, despite the fact that *GPKOW* had previously been found to be subject to XCI in somatic cell hybrids [106]. In chapter 5, *GPKOW* was examined at 6 CpGs and was found to be subject to XCI. In the future, the more CpGs at which DNA methylation can be examined, the more accurate the predictions of XCI

status will be. Further analysis of X-linked CpG island promoter DNA methylation may refine the ideal region in which DNA methylation should be examined. It may also be possible to use gene-body DNA methylation to predict XCI status. Analysis of gene-body DNA methylation, which has also been reported to differ between the X_a and the X_i [79], could allow the XCI status to be predicted for genes which lack promoter CpG islands. Many epigenetic marks other than DNA methylation also differ between active and inactive genes. In the future, the XCI status of genes which lack CpG island promoters may be analyzed through the histone modifications. The combination of multiple techniques will maximize the number of genes for which an XCI status can be determined.

In chapter 5, CpG island DNA methylation was used to predict the inactivation status of genes on the autosomal portion of six X;autosome translocations. Other marks such as late replication timing and histone modifications have shown inconsistent association with the inactivation status of genes on the autosomal portion in X;autosome translocations [189, 191], whereas current DNA methylation analysis in X;autosome translocations has correctly reflected the inactivation status [190, 192]. An important advantage of unbalanced X;autosome translocations is that there is skewed XCI allowing for ready expression analysis at informative markers. In the future, the utility of DNA methylation to predict inactivation could be confirmed either through a large scale analysis of expression of X;autosome translocations or through single gene expression analysis at genes of interest.

DNA methylation and expression analysis are excellent partners and maximize the number of genes which can be examined. DNA methylation analysis allows for genes not expressed in a specific tissue to be examined but currently cannot predict an inactivation status for genes which lack a promoter CpG island. Expression analysis can only determine an inactivation status when a gene is informative as well as expressed, but does not depend on the presence of a promoter CpG island. Through a combination of DNA methylation and expression analysis it will be possible to construct a more complete map of XCI on the X chromosome. Due to the presence of tissue-specific XCI (chapter 2), the tissue in which XCI should be examined is unclear. Blood, in particular, is an excellent candidate tissue due to its easy accessibility, however, DNA methylation analysis in chapter 2 found that blood had the highest degree of genes subject to XCI. The presence of tissue-specific XCI should be confirmed through expression analysis, but if it is as common as outlined in chapter 2 (12%) it will be critical to always examine genes of interest in the tissue of interest to confirm the XCI status. Tissue-specific XCI is of particular importance with respect to the placenta.

DNA methylation data were analyzed in chapters 3 and 4 as a means of predicting XCI status in the placenta. These DNA methylation data predicted that 41% of genes would escape from XCI

- much higher than the 10% of genes predicted to escape from XCI in blood (chapter 2). In voles, there is evidence for increased escape from XCI in the placenta, wherein several genes (*Atrx*, *Chm*, *Ddx3x*, *Mid1*, *Smc1a* and *Ube1x*) which are subject to XCI in somatic tissues are found to escape from XCI in extraembryonic tissues [324]. In humans, only two (*ATRX* and *MID1*) of the six genes are subject to XCI (confirmed by somatic cell hybrid expression [106] and DNA methylation analysis in chapter 2) in all examined somatic tissues, but were predicted by DNA methylation to escape from XCI in the human placenta (chapter 4). Tissue-specific XCI, in which escape from XCI is specific to the placenta, has also been observed in mice. When EGFP was inserted onto the X chromosome, escape from XCI in the placenta was detected despite the fact that the transgene was found to be subject to XCI in somatic tissues [325].

In order to determine if 41% of genes do escape from XCI in placenta (compared to the 18% in somatic tissues), it will be necessary to again directly compare expression and DNA methylation analysis. The difficulties associated with collecting placental RNA [326] may mean that the best way to confirm escape from XCI is through the use of cultured cells on which RNA FISH expression analysis can be combined with DNA methylation analysis. If the hypomethylation of many X-linked CpG island promoters detected in the placenta does not translate into increased escape from XCI, then DNA methylation may not be as critical for maintenance of XCI as previously thought [57-60]. If this is the case, then an examination of other marks associated with XCI might reveal which marks are not depleted in the placenta and are therefore candidates for a role in maintenance of XCI. If high escape from XCI is detected in the placenta, does this represent increased escape from XCI or does it represent reactivation? This distinction is important as it reflects the difference between the establishment and maintenance of XCI. If the increased escape from XCI in placenta was due to reactivation caused by a lack of maintenance of XCI, then older placentas would be expected to show a higher degree of escape from XCI. The first trimester placentas used in chapter 4 predicted 34% of genes escaped from XCI whereas the third trimester placentas predicted 38%. It seems unlikely that if genes were escaping from XCI due to a lack of maintenance that there would be such a small difference between first and third trimester placentas. Rather, it is more likely that this is simply further evidence of the highly heterogeneous nature of the XCI in the placenta.

XCI is thought to spread along the X chromosome through the presence of way stations [193, 194]. This implies that the majority of the X chromosome is way station rich and subject to XCI, while regions lacking way stations would escape from XCI. In addition to a lack of way stations, XCI might also occur due to the presence of yet to be defined escape elements. The similarities in the patterns of inactivation between the two X;autosome translocations involving

chromosome 21 (GM01730 and GM08134) strongly support the role of DNA sequence in determining inactivation status. Through the examination of X;autosome translocations in chapter 5, domains which contained multiple genes which escaped from XCI were confirmed to be depleted for L1s and LTRs but enriched for Alus. While previous data have suggested that LINEs are enriched at domains subject to XCI [88, 133, 134], the analysis of X;autosome translocations suggests that there is a depletion of LINEs in domains that escape from inactivation. The oldest stratum of the X chromosome has both the highest proportion of genes subject to XCI [106, 198] and the highest L1 content [88, 135, 327]. Studies which have examined the sequence composition of regions on the X chromosome which are subject to XCI compared to regions which escape from XCI are confounded by the evolutionary pressures which the X chromosome has undergone. Examination of domains subject to inactivation compared to domains which escape from inactivation on the autosomal portion of X;autosome translocations avoids the issue of evolutionary pressure and thus provides a better system in which to study the role that sequence composition plays in determining XCI status.

The differences in sequence composition are thought to allow domains which escape from XCI to loop out of the X_i domain [137] while domains rich in repetitive elements come together to form the dense heterochromatic core of the X_i [28]. Hypomethylation of repetitive elements in the placenta may, therefore, have consequences in the establishment of the silent core of repetitive elements. For example, LINEs are critical to the formation of the silent X_i domain [197] and L1s were hypomethylated in the placenta (chapter 3 and 4), possibly affecting the formation of the silent X_i domain. Despite the hypomethylated nature of placental repetitive elements, no change in the frequency or intensity of Cot-1 holes was detected (chapter 4), likely due to differences in cell type and because repetitive element DNA methylation did not decrease to the level necessary to result in an increase in repetitive element transcription.

Early during the process of XCI, a choice is made in each cell as to which X chromosome will be inactivated and become the X_i . During random XCI both X chromosomes have the same potential to become the X_i , however, mutations in the X_{ic} can result in one X chromosome having a higher likelihood of becoming the X_i [328]. This process is known as primary selection and differs from secondary selection in which both X chromosomes have equal chances of becoming the X_i , but after XCI occurs cells with one of X chromosomes are selected for. The studies involving the X;autosome translocations in chapter 5 and triploids in chapter 6 are both likely affected by the effects of secondary selection. The X;autosome translocations in chapter 5 are expected to demonstrate selection of the translocated X chromosome as the X_i which would result in skewed XCI (GM01414 has previously been confirmed to have skewed XCI [187]). While most females show random XCI, selection in X;autosome translocations makes

them ideal to use in studies which require skewed XCI. The issue of selection is especially confounding to the study of human triploids, as outlined in chapter 6. Triploid cells in which there are three complete sets of autosomes as well as three sex chromosomes provide an opportunity to study the interaction between autosome number and the number of X_i s. However, potential differences caused by selection confound conclusions regarding XCI in triploids. For example, liveborn triploids generally have more X_i s than triploids which abort before birth [228-230]. There are two possible explanations for the difference in the number of X_i s. First, triploids in which more X chromosomes are inactivated during the process of XCI are more likely to survive to birth and there is little change in the number of X_i s throughout gestation. The second possibility is that after the process of XCI, some triploids undergo selection for cells in which XCI has resulted in more X_i s. These triploids are able to survive to term while triploids in which the selection process is not as strong are not carried to term. It is difficult to differentiate between these two possibilities and the best system to resolve this question in humans will likely come from analysis of human triploid cells in culture or the use of a model organism. The mouse offers excellent potential as a model system to study XCI in triploids however differences in XCI such as the imprinted XCI of the mouse placenta mean that studies on human triploids will need to continue to investigate whether all conclusions are applicable in humans.

DNA methylation is an important epigenetic mark critical for the maintenance of XCI. In this thesis, DNA methylation was initially studied as a means to predict XCI status. Once it was established that DNA methylation usually predicted the same XCI status for genes as was determined through expression analysis, the DNA methylation in a variety of sample types were examined. The comparison of XCI status across tissues using DNA methylation demonstrated the presence of tissue-specific XCI which may have important implications on the expression status of X-linked genes in diseases which affect specific tissues. DNA methylation analysis of the placenta revealed not only hypomethylation suggestive of increased escape from XCI but also hypermethylation. Repetitive element hypomethylation in the placenta did not change the structure of the X_i but did emphasize the hypomethylated nature of the placenta and therefore its usefulness in the study of the consequences of hypomethylation on the maintenance of XCI. The spread of XCI into the autosomal portion of X;autosome translocations was detected through DNA methylation analysis. The examination of X;autosome translocations provided the opportunity to examine the role DNA sequence plays in not only the spread of XCI but also in the ability of certain genes to escape from XCI. Lastly, DNA methylation was used to examine triploids and to predict the number of X_i s in each triploid sample. Triploids provide a unique opportunity to study the consequences of the X:autosome ratio on the counting step of XCI. In

diploid females, the correct number of X chromosomes must be inactivated to achieved dosage compensation therefore the study of triploids is valuable to understanding how the counting step occurs. Although this thesis demonstrated the usefulness of DNA methylation as a tool in the study of XCI, there still remains much to be studied to fully understand the role of DNA methylation in XCI.

References

1. Yuen, R.K., S.M. Neumann, A.K. Fok, M.S. Penaherrera, D.E. McFadden, W.P. Robinson, and M.S. Kobor, *Extensive epigenetic reprogramming in human somatic tissues between fetus and adult*. *Epigenetics & chromatin*, 2011. **4**: p. 7.
2. Novakovic, B., R.K. Yuen, L. Gordon, M.S. Penaherrera, A. Sharkey, A. Moffett, J.M. Craig, W.P. Robinson, and R. Saffery, *Evidence for widespread changes in promoter methylation profile in human placenta in response to increasing gestational age and environmental/stochastic factors*. *BMC genomics*, 2011. **12**(1): p. 529.
3. Yuen, R.K., R. Jiang, M.S. Penaherrera, D.E. McFadden, and W.P. Robinson, *Genome-wide mapping of imprinted differentially methylated regions by DNA methylation profiling of human placentas from triploidies*. *Epigenetics & chromatin*, 2011. **4**(1): p. 10.
4. Lyon, M.F., *Gene action in the X-chromosome of the mouse (*Mus musculus* L.)*. *Nature*, 1961. **190**: p. 372-373.
5. Okamoto, I., et al., *Eutherian mammals use diverse strategies to initiate X-chromosome inactivation during development*. *Nature*, 2011. **472**(7343): p. 370-4.
6. Kratzer, P. and V. Chapman, *X chromosome reactivation in oocytes of *Mus caroli**. *Proc. Natl. Acad. Sci. U.S.A.*, 1981. **78**: p. 3093-3097.
7. Brown, C.J., A. Ballabio, J.L. Rupert, R.G. Lafreniere, M. Grompe, R. Tonlorenzi, and H.F. Willard, *A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome*. *Nature*, 1991. **349**: p. 38-44.
8. Penny, G.D., G.F. Kay, S.A. Sheardown, S. Rastan, and N. Brockdorff, *Requirement for *Xist* in X chromosome inactivation*. *Nature*, 1996. **379**: p. 131-137.
9. Marahrens, Y., B. Panning, J. Dausman, W. Strauss, and R. Jaenisch, **Xist*-deficient mice are defective in dosage compensation but not spermatogenesis*. *Genes & Dev.*, 1997. **11**: p. 156-166.
10. Lee, J.T. and R. Jaenisch, *Long-range cis effects of ectopic X-inactivation centres on a mouse autosome*. *Nature*, 1997. **386**: p. 275-279.
11. Herzing, L.B.K., J.T. Romer, J.M. Horn, and A. Ashworth, **Xist* has properties of the X-chromosome inactivation centre*. *Nature*, 1997. **386**: p. 272-275.
12. Wutz, A. and R. Jaenisch, *A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation*. *Mol. Cell*, 2000. **5**: p. 695-705.
13. Brown, C.J. and H.F. Willard, *The human X inactivation center is not required for maintenance of X inactivation*. *Nature*, 1994. **368**: p. 154-156.
14. Csankovszki, G., A. Nagy, and R. Jaenisch, *Synergism of *Xist* RNA, DNA methylation, and histone hypoacetylation in maintaining X chromosome inactivation*. *The Journal of cell biology*, 2001. **153**(4): p. 773-84.
15. Buzin, C.H., J.R. Mann, and J. Singer-Sam, *Quantitative RT-PCR assays show *Xist* RNA levels are low in mouse female adult tissue, embryos and embryoid bodies*. *Development*, 1994. **120**: p. 3529-3536.
16. Clemson, C.M., J.A. McNeil, H.F. Willard, and J.B. Lawrence, **XIST* RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure*. *J. Cell Biol.*, 1996. **132**(3): p. 259-275.
17. Brown, C.J., B.D. Hendrich, J.L. Rupert, R.G. Lafreniere, Y. Xing, J. Lawrence, and H.F. Willard, *The human *XIST* gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus*. *Cell*, 1992. **71**: p. 527-542.
18. Hall, L.L. and J.B. Lawrence, **XIST* RNA and Architecture of the Inactive X Chromosome: Implications for the Repeat Genome*. *Cold Spring Harbor symposia on quantitative biology*, 2010. **75**: p. 345-56.
19. Daniels, R., M. Zuccotti, T. Kinis, P. Serhal, and M. Monk, **XIST* expression in human oocytes and preimplantation embryos*. *Am. J. Hum. Genet.*, 1997. **61**: p. 33-39.

20. Ray, P.F., R.M. Winston, and A.H. Handyside, *XIST expression from the maternal X chromosome in human male preimplantation embryos at the blastocyst stage*. Hum Mol Genet, 1997. **6**(8): p. 1323-7.
21. van den Berg, I.M., J.S. Laven, M. Stevens, I. Jonkers, R.J. Galjaard, J. Gribnau, and J.H. van Doorninck, *X chromosome inactivation is initiated in human preimplantation embryos*. American journal of human genetics, 2009. **84**(6): p. 771-9.
22. Clemson, C.M., J.C. Chow, C.J. Brown, and J.B. Lawrence, *Stabilization and localization of Xist RNA are controlled by separate mechanisms and are not sufficient for X inactivation*. J. Cell Biol., 1998. **142**: p. 13-23.
23. Britten, R.J. and D.E. Kohne, *Repeated sequences in DNA. Hundreds of thousands of copies of DNA sequences have been incorporated into the genomes of higher organisms*. Science, 1968. **161**(841): p. 529-40.
24. Smit, A.F., *The origin of interspersed repeats in the human genome*. Current opinion in genetics & development, 1996. **6**(6): p. 743-8.
25. Hall, L.L., M. Byron, K. Sakai, L. Carrel, H.F. Willard, and J.B. Lawrence, *An ectopic human XIST gene can induce chromosome inactivation in postdifferentiation human HT-1080 cells*. Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**(13): p. 8677-82.
26. Chaumeil, J., S. Augui, J.C. Chow, and E. Heard, *Combined immunofluorescence, RNA fluorescent in situ hybridization, and DNA fluorescent in situ hybridization to study chromatin changes, transcriptional activity, nuclear organization, and X-chromosome inactivation*. Methods in molecular biology, 2008. **463**: p. 297-308.
27. Namekawa, S.H., B. Payer, K.D. Huynh, R. Jaenisch, and J.T. Lee, *Two-step imprinted X inactivation: repeat versus genic silencing in the mouse*. Molecular and Cellular Biology, 2010. **30**(13): p. 3187-205.
28. Chaumeil, J., P. Le Baccon, A. Wutz, and E. Heard, *A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced*. Genes & development, 2006. **20**(16): p. 2223-37.
29. Clemson, C.M., L.L. Hall, M. Byron, J. McNeil, and J.B. Lawrence, *The X chromosome is organized into a gene-rich outer rim and an internal core containing silenced nongenic sequences*. Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(20): p. 7688-93.
30. Bannister, A.J. and T. Kouzarides, *Regulation of chromatin by histone modifications*. Cell research, 2011. **21**(3): p. 381-95.
31. Okamoto, I., A.P. Otte, C.D. Allis, D. Reinberg, and E. Heard, *Epigenetic dynamics of imprinted X inactivation during early mouse development*. Science, 2004. **303**(5658): p. 644-9.
32. Heard, E., *Delving into the diversity of facultative heterochromatin: the epigenetics of the inactive X chromosome*. Current opinion in genetics & development, 2005. **15**(5): p. 482-9.
33. Chaumeil, J., P.D. Waters, E. Koina, C. Gilbert, T.J. Robinson, and J.A. Graves, *Evolution from XIST-independent to XIST-controlled X-chromosome inactivation: epigenetic modifications in distantly related mammals*. PLoS ONE, 2011. **6**(4): p. e19040.
34. Wang, J., J. Mager, Y. Chen, E. Schneider, J.C. Cross, A. Nagy, and T. Magnuson, *Imprinted X inactivation maintained by a mouse Polycomb group gene*. Nat Genet, 2001. **28**(4): p. 371-5.
35. Plath, K., et al., *Role of histone H3 lysine 27 methylation in X inactivation*. Science, 2003. **300**(5616): p. 131-5.
36. Silva, J., et al., *Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes*. Developmental cell, 2003. **4**(4): p. 481-95.

37. Kohlmaier, A., F. Savarese, M. Lachner, J. Martens, T. Jenuwein, and A. Wutz, *A chromosomal memory triggered by Xist regulates histone methylation in X inactivation*. PLoS Biol, 2004. **2**(7): p. E171.
38. Bannister, A.J., P. Zegerman, J.F. Partridge, E.A. Miska, J.O. Thomas, R.C. Allshire, and T. Kouzarides, *Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain*. Nature, 2001. **410**(6824): p. 120-4.
39. Lachner, M., D. O'Carroll, S. Rea, K. Mechtler, and T. Jenuwein, *Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins*. Nature, 2001. **410**(6824): p. 116-20.
40. Chadwick, B.P. and H.F. Willard, *Multiple spatially distinct types of facultative heterochromatin on the human inactive X chromosome*. Proc Natl Acad Sci U S A, 2004. **101**: p. 17450-17455.
41. Holliday, R. and J.E. Pugh, *DNA modification mechanisms and gene activity during development*. Science, 1975. **187**: p. 226-232.
42. Riggs, A.D., *X inactivation, differentiation, and DNA methylation*. Cytogenet. Cell Genet., 1975. **14**: p. 9-25.
43. Lister, R., et al., *Human DNA methylomes at base resolution show widespread epigenomic differences*. Nature, 2009. **462**(7271): p. 315-22.
44. Bird, A., *DNA methylation and the frequency of CpG in animal DNA*. Nuc. Acids Res., 1980. **8**: p. 1499-1504.
45. Weber, M., I. Hellmann, M.B. Stadler, L. Ramos, S. Paabo, M. Rebhan, and D. Schubeler, *Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome*. Nat. Genet., 2007. **39**(4): p. 457-466.
46. Gardiner-Garden, M. and M. Frommer, *CpG islands in vertebrate genomes*. J. Mol. Biol., 1987. **196**: p. 261-282.
47. Takai, D. and P.A. Jones, *Comprehensive analysis of CpG islands in human chromosomes 21 and 22*. Proc Natl Acad Sci U S A. , 2002. **99**(6): p. 3740-5.
48. Karolchik, D., et al., *The UCSC Genome Browser Database: 2008 update*. Nucleic Acids Res., 2008. **36**: p. D773-D779.
49. 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. 247-57.
50. Hansen, R.S., C. Wijmenga, P. Luo, A.M. Stanek, T.K. Canfield, C.M. Weemaes, and S.M. Gartler, *The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome*. Proc. Natl. Acad. Sci. USA, 1999. **96**: p. 14412-14417.
51. Jeanpierre, M., C. Turleau, A. Aurias, M. Prieur, F. Ledest, A. Fischer, and E. Viegas-Pequignot, *An embryonic-like methylation pattern of classical satellite DNA is observed in ICF syndrome*. Hum. Mol. Genet., 1993. **2**: p. 731-735.
52. Hansen, R.S., et al., *Escape from gene silencing in ICF syndrome: evidence for advanced replication time as a major determinant*. Hum. Mol. Genet., 2000. **9**(18): p. 2575-2587.
53. Bestor, T., A. Laudano, R. Mattaliano, and V. Ingram, *Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases*. Journal of molecular biology, 1988. **203**(4): p. 971-83.
54. Yoder, J.A., N.S. Soman, G.L. Verdine, and T.H. Bestor, *DNA (cytosine-5)-methyltransferases in mouse cells and tissues. Studies with a mechanism-based probe*. Journal of molecular biology, 1997. **270**(3): p. 385-95.
55. Beard, C., E. Li, and R. Jaenisch, *Loss of methylation activates Xist in somatic but not in embryonic cells*. Genes & Dev., 1995. **9**: p. 2325-2334.
56. Panning, B. and R. Jaenisch, *DNA hypomethylation can activate Xist expression and silence X-linked genes*. Genes & Dev., 1996. **10**: p. 1991-2002.

57. Venolia, L., S.M. Gartler, E.R. Wassman, P. Yen, T. Mohandas, and L.J. Shapiro, *Transformation with DNA from 5-azacytidine-reactivated X chromosomes*. Proc. Natl. Acad. Sci. U.S.A., 1982. **79**: p. 2352-2354.
58. Loebel, D.A. and P.G. Johnston, *Methylation analysis of a marsupial X-linked CpG island by bisulfite genomic sequencing*. Genome research, 1996. **6**(2): p. 114-23.
59. Kaslow, D.C. and B.R. Migeon, *DNA methylation stabilizes X chromosome inactivation in eutherians but not in marsupials: evidence for multistep maintenance of mammalian X dosage compensation*. Proc. Natl. Acad. Sci. USA, 1987. **84**: p. 6210-6214.
60. Rens, W., M.S. Wallduck, F.L. Lovell, M.A. Ferguson-Smith, and A.C. Ferguson-Smith, *Epigenetic modifications on X chromosomes in marsupial and monotreme mammals and implications for evolution of dosage compensation*. Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(41): p. 17657-62.
61. Stein, R., A. Razin, and H. Cedar, *In vitro methylation of the hamster adenine phosphoribosyltransferase gene inhibits its expression in mouse L cells*. Proceedings of the National Academy of Sciences of the United States of America, 1982. **79**(11): p. 3418-22.
62. Lorincz, M.C., D. Schubeler, S.R. Hutchinson, D.R. Dickerson, and M. Groudine, *DNA methylation density influences the stability of an epigenetic imprint and Dnmt3a/b-independent de novo methylation*. Molecular and Cellular Biology, 2002. **22**(21): p. 7572-80.
63. Wolf, S.F., S. Dintzis, D. Toniolo, G. Persico, K.D. Lunnen, J. Axelmann, and B.R. Migeon, *Complete concordance between glucose -6- phosphate dehydrogenase activity and hypomethylation of 3' CpG clusters: implications for X chromosome dosage compensation*. Nucl. Acids Res., 1984. **12**(24): p. 9333-9348.
64. Wolf, S.F., D.J. Jolly, K.D. Lunnen, T. Friedmann, and B.R. Migeon, *Methylation of the hypoxanthine phosphoribosyltransferase locus on the human X chromosome: implications for X-chromosome inactivation*. Proc. Natl. Acad. Sci. USA, 1984. **81**: p. 2806-2810.
65. Goodfellow, P.J., C. Mondello, S.M. Darling, B. Pym, P. Little, and P.N. Goodfellow, *Absence of methylation of a CpG-rich region at the 5' end of the MIC2 gene on the active X, and inactive X, and the Y chromosome*. Proc. Natl. Acad. Sci. USA, 1988. **85**: p. 5605-5609.
66. Hansen, R.S. and S.M. Gartler, *5-Azacytidine-induced reactivation of the human X chromosome-linked PGK1 gene is associated with a large region of cytosine demethylation in the 5' CpG island*. Proc Natl Acad Sci U S A, 1990. **87**(11): p. 4174-8.
67. Carrel, L., C.M. Clemson, J.M. Dunn, A.P. Miller, P.A. Hunt, J.B. Lawrence, and H.F. Willard, *X inactivation analysis and DNA methylation studies of the ubiquitin activating enzyme E1 and PCTAIRE-1 genes in human and mouse*. Hum. Mol. Genet., 1996. **5**(3): p. 391-402.
68. Anderson, C.L. and C.J. Brown, *Variability of X chromosome inactivation: effect on levels of TIMP1 RNA and role of DNA methylation*. Hum. Genet., 2002. **110**: p. 271-278.
69. Yasukochi, Y., et al., *X chromosome-wide analyses of genomic DNA methylation states and gene expression in male and female neutrophils*. Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(8): p. 3704-9.
70. Eckhardt, F., et al., *DNA methylation profiling of human chromosomes 6, 20 and 22*. Nat Genet, 2006. **38**(12): p. 1378-85.
71. Schilling, E. and M. Rehli, *Global, comparative analysis of tissue-specific promoter CpG methylation*. Genomics, 2007. **90**(3): p. 314-23.
72. Illingworth, R., et al., *A novel CpG island set identifies tissue-specific methylation at developmental gene loci*. PLoS biology, 2008. **6**(1): p. e22.
73. Irizarry, R.A., et al., *The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores*. Nature Genetics, 2009. **41**(2): p. 178-86.

74. Jones, P.A. and S.B. Baylin, *The epigenomics of cancer*. Cell, 2007. **128**(4): p. 683-92.
75. Gadalla, K.K., M.E. Bailey, and S.R. Cobb, *MeCP2 and Rett syndrome: reversibility and potential avenues for therapy*. The Biochemical journal, 2011. **439**(1): p. 1-14.
76. Illingworth, R.S., et al., *Orphan CpG islands identify numerous conserved promoters in the mammalian genome*. PLoS genetics, 2010. **6**(9).
77. Maunakea, A.K., et al., *Conserved role of intragenic DNA methylation in regulating alternative promoters*. Nature, 2010. **466**(7303): p. 253-7.
78. Weber, M., J.J. Davies, D. Wittig, E.J. Oakeley, M. Haase, W.L. Lam, and D. Schubeler, *Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells*. Nat. Genet., 2005. **37**(8): p. 853-862.
79. Hellman, A. and A. Chess, *Gene body-specific methylation on the active X chromosome*. Science, 2007. **315**(5815): p. 1141-1143.
80. Jones, P.A., *The DNA methylation paradox*. Trends Genet., 1999. **15**(1): p. 34-37.
81. Aran, D., G. Toperoff, M. Rosenberg, and A. Hellman, *Replication timing-related and gene body-specific methylation of active human genes*. Human molecular genetics, 2011. **20**(4): p. 670-80.
82. Lorincz, M.C., D.R. Dickerson, M. Schmitt, and M. Groudine, *Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells*. Nature structural & molecular biology, 2004. **11**(11): p. 1068-75.
83. Klose, R.J. and A.P. Bird, *Genomic DNA methylation: the mark and its mediators*. Trends in biochemical sciences, 2006. **31**(2): p. 89-97.
84. Edwards, J.R., et al., *Chromatin and sequence features that define the fine and gross structure of genomic methylation patterns*. Genome research, 2010. **20**(7): p. 972-80.
85. Brenet, F., M. Moh, P. Funk, E. Feierstein, A.J. Viale, N.D. Socci, and J.M. Scandura, *DNA methylation of the first exon is tightly linked to transcriptional silencing*. PLoS ONE, 2011. **6**(1): p. e14524.
86. Rollins, R.A., F. Haghghi, J.R. Edwards, R. Das, M.Q. Zhang, J. Ju, and T.H. Bestor, *Large-scale structure of genomic methylation patterns*. Genome research, 2006. **16**(2): p. 157-63.
87. Hata, K. and Y. Sakaki, *Identification of critical CpG sites for repression of L1 transcription by DNA methylation*. Gene, 1997. **189**(2): p. 227-34.
88. Bailey, J.A., L. Carrel, A. Chakravarti, and E. Eichler, *Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: the Lyon repeat hypothesis*. Proc. Natl. Acad. Sci., USA, 2000. **97**: p. 6634-6639.
89. Hansen, R.S., *X inactivation-specific methylation of LINE-1 elements by DNMT3B: implications for the Lyon repeat hypothesis*. Hum. Mol. Genet., 2003. **12**(19): p. 2559-2567.
90. Viegas-Pequignot, E., B. Dutrillaux, and G. Thomas, *Inactive X chromosome has the highest concentration of unmethylated Hha I sites*. Proc. Natl. Acad. Sci. USA, 1988. **85**: p. 7657-7660.
91. Miller, D.A., E. Okamoto, B.F. Erlanger, and O.J. Miller, *Is DNA methylation responsible for mammalian X chromosome inactivation?* Cytogenet. Cell Genet., 1982. **33**: p. 345-349.
92. Prantera, G. and M. Ferraro, *Analysis of methylation and distribution of CpG sequences on human active and inactive X chromosomes by in situ nick translation*. Chromosoma, 1990. **99**(1): p. 18-23.
93. Barr, M.L. and E.G. Bertram, *A morphological distinction between neurones of the male and female, and the behaviour of the nucleolar satellite during accelerated nucleoprotein synthesis*. Nature, 1949. **163**: p. 676-677.
94. Ohno, S., W.D. Kaplan, and R. Kinosita, *Formation of the sex chromatin by a single X-chromosome in liver cells of Rattus norvegicus*. Exp. Cell Res., 1959. **18**: p. 415-418.

95. Lyon, M.F., *Sex chromatin and gene action in the mammalian X-chromosome*. Am. J. Hum. Genet., 1962. **14**: p. 135-145.
96. Eils, R., S. Dietzel, E. Bertin, E. Schrock, M.R. Speicher, T. Ried, M. Robert-Nicoud, C. Cremer, and T. Cremer, *Three-dimensional reconstruction of painted human interphase chromosomes: active and inactive X chromosome territories have similar volumes but differ in shape and surface structure*. J. Cell Biol., 1996. **135**: p. 1427-1440.
97. Visser, A.E., R. Eils, A. Jauch, G. Little, P.J.M. Bakker, T. Cremer, and J.A. Aten, *Spatial distributions of early and late replicating chromatin in interphase chromosome territories*. Exp. Cell Res., 1998. **243**: p. 398-407.
98. Rappold, G.A., T. Cremer, H.D. Hager, K.E. Davies, C.R. Müller, and T. Yang, *Sex chromosome positions in human interphase nuclei as studied by in situ hybridization with chromosome specific DNA probes*. Hum Genet., 1984. **67**: p. 317-325.
99. Lima-de-Faria and H. Jaworksa, *Late DNA synthesis in heterochromatin*. Nature, 1968. **217**: p. 138-142.
100. Taylor, J.H., *Asynchronous duplication of chromosomes in cultured cells of chinese hamster*. Biophysic. Biochem. Cytol., 1960. **7**(3): p. 455-463.
101. Morishima, A., M.M. Grumbach, and J.H. Taylor, *Asynchronous duplication of human chromosomes and the origin of sex chromatin*. Proc. Natl. Acad. Sci. USA, 1962. **48**: p. 756-763.
102. Takagi, N. and M. Oshimura, *Fluorescence and giemsa banding studies of the allocyclic X chromosome in embryonic and adult mouse cells*. Exptl. Cell Res., 1973. **78**: p. 127-135.
103. Willard, H.F. and S.A. Latt, *Analysis of deoxyribonucleic acid replication in human X chromosomes by fluorescence microscopy*. Am. J. Hum. Genet., 1976. **28**: p. 213-227.
104. Willard, H.F., *Tissue-specific heterogeneity in DNA replication patterns of human X chromosomes*. Chromosoma, 1977. **61**: p. 61-73.
105. Babu, A., S. Chemitiganti, and R.S. Verma, *Heterochromatinization of human X-chromosomes: classification of replication profile*. Clin. Genet., 1986. **30**: p. 108-111.
106. Carrel, L. and H.F. Willard, *X-inactivation profile reveals extensive variability in X-linked gene expression in females*. Nature, 2005. **434**(7031): p. 400-4.
107. Yang, F., T. Babak, J. Shendure, and C.M. Disteché, *Global survey of escape from X inactivation by RNA-sequencing in mouse*. Genome research, 2010. **20**(5): p. 614-22.
108. Rouyer, F., M.C. Simmler, C. Johnsson, G. Vergnaud, H.J. Cooke, and J. Weissenbach, *A gradient of sex linkage in the pseudoautosomal region of the human sex chromosomes*. Nature, 1986. **319**: p. 291-295.
109. D'Esposito, M., A. Ciccodicola, F. Gianfrancesco, T. Esposito, L. Flagiello, R. Mazzarella, D. Schlessinger, and M. D'Urso, *A synaptobrevin-like gene in the Xq28 pseudoautosomal region undergoes X inactivation*. Nature Genetics, 1996. **13**(2): p. 227-9.
110. Ciccodicola, A., et al., *Differentially regulated and evolved genes in the fully sequenced Xq/Yq pseudoautosomal region*. Hum. Mol. Genet., 2000. **9**: p. 395-401.
111. Huber, R., et al., *DNA methylation in transcriptional repression of two differentially expressed X-linked genes, GPC3 and SYBL1*. Proc Natl Acad Sci U S A, 1999. **96**(2): p. 616-21.
112. De Bonis, M.L., A. Cerase, M.R. Matarazzo, M. Ferraro, M. Strazzullo, R.S. Hansen, P. Chiurazzi, G. Neri, and M. D'Esposito, *Maintenance of X- and Y-inactivation of the pseudoautosomal (PAR2) gene SPRY3 is independent from DNA methylation and associated to multiple layers of epigenetic modifications*. Hum Mol Genet, 2006. **15**(7): p. 1123-32.
113. Mitchell, M.M., et al., *Epigenetic investigation of variably X chromosome inactivated genes in monozygotic female twins discordant for primary biliary cirrhosis*. Epigenetics : official journal of the DNA Methylation Society, 2011. **6**(1): p. 95-102.

114. Champion, K.M., J.G. Gilbert, F.A. Asimakopoulos, S. Hinshelwood, and A.R. Green, *Clonal haemopoiesis in normal elderly women: implications for the myeloproliferative disorders and myelodysplastic syndromes*. British journal of haematology, 1997. **97**(4): p. 920-6.
115. Gale, R.E., A.K. Fielding, C.N. Harrison, and D.C. Linch, *Acquired skewing of X-chromosome inactivation patterns in myeloid cells of the elderly suggests stochastic clonal loss with age*. British journal of haematology, 1997. **98**(3): p. 512-9.
116. Tonon, L., et al., *Unbalanced X-chromosome inactivation in haemopoietic cells from normal women*. British journal of haematology, 1998. **102**(4): p. 996-1003.
117. Hatakeyama, C., C.L. Anderson, C.L. Beever, M.S. Penaherrera, C.J. Brown, and W.P. Robinson, *The dynamics of X-inactivation skewing as women age*. Clin Genet, 2004. **66**(4): p. 327-32.
118. Schoeftner, S., R. Blanco, I. Lopez de Silanes, P. Munoz, G. Gomez-Lopez, J.M. Flores, and M.A. Blasco, *Telomere shortening relaxes X chromosome inactivation and forces global transcriptome alterations*. Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(46): p. 19393-8.
119. Wareham, K.A., M.F. Lyon, P.H. Glenister, and E.D. Williams, *Age related reactivation of an X-linked gene*. Nature, 1987. **327**: p. 725-727.
120. Talebizadeh, Z., S.D. Simon, and M.G. Butler, *X chromosome gene expression in human tissues: male and female comparisons*. Genomics, 2006. **88**(6): p. 675-81.
121. Ashworth, A., S. Rastan, R. Lovell-Badge, and G. Kay, *X-chromosome inactivation may explain the difference in viability of XO humans and mice*. Nature, 1991. **351**: p. 406-408.
122. Hook, E.B. and D. Warburton, *The distribution of chromosomal genotypes associated with Turner's syndrome: livebirth prevalence rates and evidence for diminished fetal mortality and severity in genotypes associated with structural X abnormalities or mosaicism*. Human Genetics, 1983. **64**(1): p. 24-7.
123. Abir, R., B. Fisch, R. Nahum, R. Orvieto, S. Nitke, and Z. Ben Rafael, *Turner's syndrome and fertility: current status and possible putative prospects*. Human reproduction update, 2001. **7**(6): p. 603-10.
124. Ogata, T. and N. Matsuo, *Turner syndrome and female sex chromosome aberrations: deduction of the principal factors involved in the development of clinical features*. Human Genetics, 1995. **95**(6): p. 607-29.
125. Binder, G., *Short stature due to SHOX deficiency: genotype, phenotype, and therapy*. Hormone research in paediatrics, 2011. **75**(2): p. 81-9.
126. Swerdlow, A.J., M.J. Schoemaker, C.D. Higgins, A.F. Wright, and P.A. Jacobs, *Mortality and cancer incidence in women with extra X chromosomes: a cohort study in Britain*. Human Genetics, 2005. **118**(2): p. 255-60.
127. Swerdlow, A.J., C.D. Higgins, M.J. Schoemaker, A.F. Wright, and P.A. Jacobs, *Mortality in patients with Klinefelter syndrome in Britain: a cohort study*. The Journal of clinical endocrinology and metabolism, 2005. **90**(12): p. 6516-22.
128. Schempp, W. and B. Meer, *Cytologic evidence for three human X-chromosomal segments escaping inactivation*. Hum. Genet., 1983. **63**: p. 171-174.
129. Boggs, B.A., P. Cheung, E. Heard, D.L. Spector, A.C. Chinault, and C.D. Allis, *Differentially methylated forms of histone H3 show unique association patterns with inactive human X chromosomes*. Nature Genetics, 2002. **30**(1): p. 73-6.
130. Goto, Y. and H. Kimura, *Inactive X chromosome-specific histone H3 modifications and CpG hypomethylation flank a chromatin boundary between an X-inactivated and an escape gene*. Nucleic Acids Research, 2009. **37**(22): p. 7416-28.
131. Jeppesen, P. and B.M. Turner, *The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression*. Cell, 1993. **74**(2): p. 281-9.

132. Khalil, A.M. and D.J. Driscoll, *Trimethylation of histone H3 lysine 4 is an epigenetic mark at regions escaping mammalian X inactivation*. Epigenetics : official journal of the DNA Methylation Society, 2007. **2**(2): p. 114-8.
133. Carrel, L., C. Park, S. Tyekucheveva, J. Dunn, F. Chiaromonte, and K.D. Makova, *Genomic environment predicts expression patterns on the human inactive X chromosome*. PLoS Genet, 2006. **2**(9): p. e151.
134. Ross, M.T., et al., *The DNA sequence of the human X chromosome*. Nature, 2005. **434**(7031): p. 325-37.
135. Wang, Z., H.F. Willard, S. Mukherjee, and T.S. Furey, *Evidence of Influence of Genomic DNA Sequence on Human X Chromosome Inactivation*. PLoS Comput Biol, 2006. **2**(9).
136. Nguyen, D.K., F. Yang, R. Kaul, C. Alkan, A. Antonellis, K.F. Friery, B. Zhu, P.J. de Jong, and C.M. Disteche, *Cln4-2 genomic structure differs between the X locus in Mus spretus and the autosomal locus in Mus musculus: AT motif enrichment on the X*. Genome research, 2011. **21**(3): p. 402-9.
137. Heard, E. and W. Bickmore, *The ins and outs of gene regulation and chromosome territory organisation*. Current opinion in cell biology, 2007. **19**(3): p. 311-6.
138. Li, N. and L. Carrel, *Escape from X chromosome inactivation is an intrinsic property of the Jarid1c locus*. Proceedings of the National Academy of Sciences of the United States of America, 2008. **105**(44): p. 17055-60.
139. Takagi, N. and M. Sasaki, *Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse*. Nature, 1975. **256**: p. 640-642.
140. Mak, W., T.B. Nesterova, M. de Napoles, R. Appanah, S. Yamanaka, A.P. Otte, and N. Brockdorff, *Reactivation of the paternal X chromosome in early mouse embryos*. Science, 2004. **303**(5658): p. 666-9.
141. Ropers, H.-H., G. Wolff, and H.W. Hitzeroth, *Preferential X inactivation in human placenta membranes: is the paternal X active in early embryonic development of female mammals*. Hum. Genet., 1978. **43**: p. 265-273.
142. Harrison, K.B. and D. Warburton, *Preferential X-chromosome activity in human female placental tissues*. Cytogenet. Cell Genet., 1986. **41**: p. 163-168.
143. Harrison, K.B., *X-chromosome inactivation in the human cytotrophoblast*. Cytogenet. Cell Genet., 1989. **52**: p. 37-41.
144. Goto, T., E. Wright, and M. Monk, *Paternal X-chromosome inactivation in human trophoblastic cells*. Mol. Hum. Reprod., 1997. **3**(1): p. 77-80.
145. Uehara, S., M. Tamura, M. Nata, J. G, Y. N, K. Okamura, and A. Yajima, *X-chromosome inactivation in the human trophoblast of early pregnancy*. J. Hum. Genet., 2000. **45**: p. 119-126.
146. MacDonald, M., T. Hassold, J. Harvey, L.H. Wang, N.E. Morton, and P. Jacobs, *The origin of 47,XXY and 47,XXX aneuploidy: heterogeneous mechanisms and role of aberrant recombination*. Human molecular genetics, 1994. **3**(8): p. 1365-71.
147. Uematsu, A., T. Yorifuji, J. Muroi, M. Kawai, M. Mamada, M. Kaji, C. Yamanaka, T. Momoi, and T. Nakahata, *Parental origin of normal X chromosomes in Turner syndrome patients with various karyotypes: implications for the mechanism leading to generation of a 45,X karyotype*. American journal of medical genetics, 2002. **111**(2): p. 134-9.
148. Monroy, N., M. Lopez, A. Cervantes, D. Garcia-Cruz, G. Zafra, S. Canun, J.C. Zenteno, and S. Kofman-Alfaro, *Microsatellite analysis in Turner syndrome: parental origin of X chromosomes and possible mechanism of formation of abnormal chromosomes*. American journal of medical genetics, 2002. **107**(3): p. 181-9.
149. Moreira de Mello, J.C., E.S. de Araujo, R. Stabellini, A.M. Fraga, J.E. de Souza, D.R. Sumita, A.A. Camargo, and L.V. Pereira, *Random X inactivation and extensive mosaicism in human placenta revealed by analysis of allele-specific gene expression along the X chromosome*. PLoS ONE, 2010. **5**(6): p. e10947.

150. Migeon, B.R., M. Schmidt, J. Axelman, and C.R. Cullen, *Complete reactivation of X chromosomes from human chorionic villi with a switch to early DNA replication*. Proc. Natl. Acad. Sci. USA, 1986. **83**: p. 2182-2186.
151. Migeon, B.R., J. Axelman, and P. Jeppesen, *Differential X reactivation in human placental cells: implications for reversal of X inactivation*. Am. J. Hum. Genet., 2005. **77**(3): p. 355-364.
152. Migeon, B.R., S.F. Wolf, J. Axelman, D.C. Kaslow, and M. Schmidt, *Incomplete X chromosome dosage compensation in chorionic villi of human placenta*. Proc. Natl. Acad. Sci. USA, 1985. **82**: p. 3390-3394.
153. Andina, R., *A study of X chromosome regulation during oogenesis in the mouse*. Exp. Cell Res. III, 1978: p. 211-218.
154. Luo, S., B.S. Torchia, and B.R. Migeon, *XIST expression is repressed when X inactivation is reversed in human placental cells: a model for study of XIST regulation*. Somat. Cell Mol. Genet., 1995. **21**(1): p. 51-60.
155. Lee, J.T., L.S. Davidow, and D. Warshawsky, *Tsix, a gene antisense to Xist at the X-inactivation center*. Nat. Genet., 1999. **21**: p. 400-404.
156. Sado, T., Z. Wang, H. Sasaki, and E. Li, *Regulation of imprinted X-chromosome inactivation in mice by Tsix*. Development, 2001. **128**: p. 1275-1286.
157. Debrand, E., C. Chureau, D. Arnaud, P. Avner, and E. Heard, *Functional analysis of the DXPas34 locus, a 3' regulator of Xist expression*. Mol. Cell. Biol., 1999. **19**: p. 8513-8525.
158. Migeon, B.R., C.H. Lee, A.K. Chowdury, and H. Carpenter, *Species differences in TSIX/Tsix reveal the roles of these genes in X-chromosome inactivation*. Am J Hum Genet, 2002. **71**: p. 286-293.
159. Migeon, B.R., A.K. Chowdury, J.A. Dunston, and I. McIntosh, *Identification of TSIX, encoding an RNA antisense to human XIST, reveals differences from its murine counterpart: implications for X inactivation*. Am. J. Hum. Genet., 2001. **69**: p. 951-960.
160. Ehrlich, M., M.A. Gama-Sosa, L.-H. Huang, R.M. Midgett, K.C. Kuo, R.A. McCune, and C. Gehrke, *Amount and distribution of 5-methylcytosine in human DNA from different types of tissues or cells*. Nucleic Acids Res., 1982. **10**(8): p. 2709-2721.
161. Gama-Sosa, M.A., R.M. Midgett, V.A. Slagel, S. Githens, K.C. Kuo, C.W. Gehrke, and M. Ehrlich, *Tissue-specific differences in DNA methylation in various mammals*. Biochim. Biophys. Acta., 1983. **740**(2): p. 212-219.
162. Fuke, C., et al., *Age related changes in 5-methylcytosine content in human peripheral leukocytes and placentas: an HPLC-based study*. Ann. Hum. Genet., 2004. **68**(Pt 3): p. 196-204.
163. Driscoll, D. and B. Migeon, *Sex differences in methylation of single-copy genes in human meiotic germ cells: implications for X chromosome inactivation, parental imprinting, and origin of CpG mutations*. Somat. Cell Mol. Genet., 1990. **16**: p. 267-282.
164. Luo, S., J.C. Robinson, A.L. Reiss, and B.R. Migeon, *DNA methylation of the fragile X locus in somatic and germ cells during fetal development: relevance to the fragile X syndrome and X inactivation*. Somat. Cell Mol. Genet., 1993. **19**: p. 393-404.
165. Hellmann-Blumberg, U., M.F. Hintz, J.M. Gatewood, and C.W. Schmid, *Developmental differences in methylation of human Alu repeats*. Mol. Cell Biol., 1993. **13**(8): p. 4523-4530.
166. Shen, H.M., A. Nakamura, J. Sugimoto, N. Sakumoto, T. Oda, Y. Jinno, and Y. Okazaki, *Tissue specificity of methylation and expression of human genes coding for neuropeptides and their receptors, and of a human endogenous retrovirus K family*. J. Hum. Genet., 2006. **51**(5): p. 440-450.
167. Novakovic, B., et al., *DNA methylation-mediated down-regulation of DNA methyltransferase-1 (DNMT1) is coincident with, but not essential for, global hypomethylation in human placenta*. The Journal of biological chemistry, 2010. **285**(13): p. 9583-93.

168. Sanlaville, D., C. Schluth-Bolard, and C. Turleau, *Distal Xq duplication and functional Xq disomy*. Orphanet journal of rare diseases, 2009. **4**: p. 4.
169. Crandall, B.F., R.E. Carrel, J. Howard, W.A. Schroeder, and H. Muller, *Trisomy 13 with a 13-X translocation*. Am. J. Hum. Genet., 1974. **26**: p. 385-392.
170. Schanz, S. and P. Steinback, *Investigation of the "variable spreading" of X inactivation into a translocated autosome*. Hum. Genet., 1989. **82**: p. 244-248.
171. Gustashaw, K.M., V. Zurcher, L.H. Dickerman, R. Stallard, and H.F. Willard, *Partial X chromosome trisomy with functional disomy of Xp due to failure of X inactivation*. American journal of medical genetics, 1994. **53**(1): p. 39-45.
172. Stankiewicz, P., et al., *Minimal phenotype in a girl with trisomy 15q due to t(X;15)(q22.3;q11.2) translocation*. American journal of medical genetics. Part A, 2006. **140**(5): p. 442-52.
173. Lee, J.T., W.M. Strauss, J.A. Dausman, and R. Jaenisch, *A 450 kb transgene displays properties of the mammalian X-inactivation center*. Cell, 1996. **86**: p. 83-94.
174. Russell, L.B. and C.S. Montgomery, *Comparative studies on X-autosome translocations in the mouse. II. Inactivation of autosomal loci, segregation, and mapping of autosomal breakpoints in five T (X;1) S*. Genetics, 1970. **64**(2): p. 281-312.
175. Rastan, S., *Non-random X-chromosome inactivation in mouse X-autosome translocation embryos-location of the inactivation centre*. J. Embryol. Exp. Morph., 1983. **78**: p. 1-22.
176. Tommerup, N., *Mendelian cytogenetics. Chromosome rearrangements associated with mendelian disorders*. Journal of medical genetics, 1993. **30**(9): p. 713-27.
177. Kleinjan, D.J. and V. van Heyningen, *Position effect in human genetic disease*. Human molecular genetics, 1998. **7**(10): p. 1611-8.
178. Therman, E. and K. Patau, *Abnormal X chromosomes in man: origin, behavior and effects*. Humangenetik, 1974. **25**: p. 1-16.
179. Schmidt, M. and D. Du Sart, *Functional disomies of the X chromosome influence the cell selection and hence the inactivation pattern in females with balanced X;autosome translocations: a review of 122 cases*. Am. J. Med. Genet., 1992. **42**: p. 161-169.
180. Disteche, C.M., K. Swisshelm, S. Forbes, and R.A. Pagon, *X-inactivation patterns in lymphocytes and skin fibroblasts of three cases of X-autosome translocations with abnormal phenotypes*. Hum. Genet., 1984. **66**: p. 71-76.
181. Ferraris, A.M., G.L. Forni, R. Mangerini, and G.F. Gaetani, *Nonrandom X-chromosome inactivation in hemopoietic cells from carriers of dyskeratosis congenita*. American journal of human genetics, 1997. **61**(2): p. 458-61.
182. Vulliamy, T.J., S.W. Knight, I. Dokal, and P.J. Mason, *Skewed X-inactivation in carriers of X-linked dyskeratosis congenita*. Blood, 1997. **90**(6): p. 2213-6.
183. Cottrell, C.E., A. Sommer, G.D. Wenger, S. Bullard, T. Busch, K.N. Krahn, A.C. Lidral, and J.M. Gastier-Foster, *Atypical X-chromosome inactivation in an X;1 translocation patient demonstrating Xq28 functional disomy*. American journal of medical genetics. Part A, 2009. **149A**(3): p. 408-14.
184. Ejima, Y., M.S. Sasaki, A. Kaneko, H. Tanooka, Y. Hara, T. Hida, and Y. Kinoshita, *Possible inactivation of part of chromosome 13 due to 13qXp translocation associated with retinoblastoma*. Clinical genetics, 1982. **21**(6): p. 357-61.
185. Mohandas, T., R.S. Sparkes, and L.J. Shapiro, *Genetic evidence for the inactivation of a human autosomal locus attached to an inactive X chromosome*. American journal of human genetics, 1982. **34**(5): p. 811-7.
186. White, W.M., H.F. Willard, D.L. Van Dyke, and D.J. Wolff, *The spreading of X inactivation into autosomal material of an X;autosome translocation: evidence for a difference between autosomal and X-chromosomal DNA*. Am. J. Hum. Genet., 1998. **63**: p. 20-28.
187. Keitges, E.A. and C.G. Palmer, *Analysis of spreading of inactivation in eight X autosome translocations utilizing the high resolution RBG technique*. Hum. Genet., 1986. **72**: p. 231-236.

188. Mohandas, T., B.F. Crandall, R.S. Sparkes, M.B. Passage, and M.C. Sparkes, *Late replication studies in a human X/13 translocation: correlation with autosomal gene expression*. Cytogenet. Cell Genet., 1981. **29**: p. 215-220.
189. Sharp, A., D.O. Robinson, and P.A. Jacobs, *Absence of correlation between late-replication and spreading of X inactivation in an X;autosome translocation*. Hum. Genet., 2001. **109**: p. 295-302.
190. Sharp, A.J., H.T. Spotswood, D.O. Robinson, B.M. Turner, and P.A. Jacobs, *Molecular and cytogenetic analysis of the spreading of X inactivation in X;autosome translocations*. Human molecular genetics, 2002. **11**(25): p. 3145-56.
191. Hall, L.L., C.M. Clemson, M. Byron, K. Wydner, and J.B. Lawrence, *Unbalanced X;autosome translocations provide evidence for sequence specificity in the association of XIST RNA with chromatin*. Human molecular genetics, 2002. **11**(25): p. 3157-65.
192. Giorda, R., M.C. Bonaglia, G. Milani, A. Baroncini, F. Spada, S. Beri, G. Menozzi, M. Rusconi, and O. Zuffardi, *Molecular and cytogenetic analysis of the spreading of X inactivation in a girl with microcephaly, mild dysmorphic features and t(X;5)(q22.1;q31.1)*. European journal of human genetics : EJHG, 2008. **16**(8): p. 897-905.
193. Gartler, S.M. and A.D. Riggs, *Mammalian X-chromosome inactivation*. Ann. Rev. Genet., 1983. **17**: p. 155-190.
194. Lyon, M.F., *X-chromosome inactivation: a repeat hypothesis*. Cytogenet. Cell Genet., 1998. **80**: p. 133-137.
195. Boyle, A.L., S.G. Ballard, and D.C. Ward, *Differential distribution of long and short interspersed element sequences in the mouse genome: chromosome karyotyping by fluorescence in situ hybridization*. Proceedings of the National Academy of Sciences of the United States of America, 1990. **87**(19): p. 7757-61.
196. Abrusan, G., J. Giordano, and P.E. Warburton, *Analysis of transposon interruptions suggests selection for L1 elements on the X chromosome*. PLoS genetics, 2008. **4**(8): p. e1000172.
197. Chow, J.C., et al., *LINE-1 activity in facultative heterochromatin formation during X chromosome inactivation*. Cell, 2010. **141**(6): p. 956-69.
198. Lahn, B.T. and D.C. Page, *Four evolutionary strata on the human X chromosome*. Science, 1999. **286**: p. 964-967.
199. Filippova, G.N., M.K. Cheng, J.M. Moore, J.P. Truong, Y.J. Hu, D.K. Nguyen, K.D. Tsuchiya, and C.M. Distèche, *Boundaries between chromosomal domains of X inactivation and escape bind CTCF and lack CpG methylation during early development*. Dev. Cell, 2005. **8**: p. 31-42.
200. Boumil, R.M. and J.T. Lee, *Forty years of decoding the silence in X-chromosome inactivation*. Hum. Mol. Genet., 2001. **10**: p. 2225-2232.
201. Rougeulle, C. and P. Avner, *Controlling X-inactivation in mammals: what does the centre hold?* Sem. Cell & Dev. Biol., 2003. **14**: p. 331-340.
202. Heard, E., *Recent advances in X-chromosome inactivation*. Curr Opin Cell Biol, 2004. **16**(3): p. 247-55.
203. Barakat, T.S., I. Jonkers, K. Monkhorst, and J. Gribnau, *X-changing information on X inactivation*. Experimental cell research, 2010. **316**(5): p. 679-87.
204. Zaragoza, M.V., U. Surti, R.W. Redline, E. Millie, A. Chakravarti, and T.J. Hassold, *Parental origin and phenotype of triploidy in spontaneous abortions: predominance of diandry and association with the partial hydatidiform mole*. American journal of human genetics, 2000. **66**(6): p. 1807-20.
205. Brancati, F., R. Mingarelli, and B. Dallapiccola, *Recurrent triploidy of maternal origin*. European journal of human genetics : EJHG, 2003. **11**(12): p. 972-4.
206. McFadden, D.E. and W.P. Robinson, *Phenotype of triploid embryos*. Journal of medical genetics, 2006. **43**(7): p. 609-12.

207. Hassold, T., et al., *A cytogenetic study of 1000 spontaneous abortions*. Annals of human genetics, 1980. **44**(Pt 2): p. 151-78.
208. Schinzel, A., *Catalogue of unbalanced chromosome aberrations in man*. 2nd rev. and expanded ed2001, Berlin ; New York: Walter de Gruyter. xvii, 966 p.
209. De Vigan, C., N. Baena, E. Cariati, M. Clementi, and C. Stoll, *Contribution of ultrasonographic examination to the prenatal detection of chromosomal abnormalities in 19 centres across Europe*. Annales de genetique, 2001. **44**(4): p. 209-17.
210. Jacobs, P.A., A.E. Szulman, J. Funkhouser, J.S. Matsuura, and C.C. Wilson, *Human triploidy: relationship between parental origin of the additional haploid complement and development of partial hydatidiform mole*. Annals of human genetics, 1982. **46**(Pt 3): p. 223-31.
211. Takabachi, N., et al., *Long-term survival in a 69,XXX triploid premature infant*. American journal of medical genetics. Part A, 2008. **146A**(12): p. 1618-21.
212. Sherard, J., et al., *Long survival in a 69,XXY triploid male*. American journal of medical genetics, 1986. **25**(2): p. 307-12.
213. Tantravahi, U., D.W. Bianchi, C. Haley, M.M. Destrempes, A.T. Ricker, B.R. Korf, and S.A. Latt, *Use of Y chromosome specific probes to detect low level sex chromosome mosaicism*. Clinical genetics, 1986. **29**(5): p. 445-8.
214. Jacobs, P.A., R.R. Angell, I.M. Buchanan, T.J. Hassold, A.M. Matsuyama, and B. Manuel, *The origin of human triploids*. Annals of human genetics, 1978. **42**(1): p. 49-57.
215. Kajii, T. and K. Ohama, *Androgenetic origin of hydatidiform mole*. Nature, 1977. **268**(5621): p. 633-4.
216. Procter, S.E., E.S. Gray, and J.L. Watt, *Triploidy, partial mole and dispermy. An investigation of 12 cases*. Clinical genetics, 1984. **26**(1): p. 46-51.
217. Uchida, I.A. and V.C. Freeman, *Triploidy and chromosomes*. American journal of obstetrics and gynecology, 1985. **151**(1): p. 65-9.
218. McFadden, D.E., L.C. Kwong, I.Y. Yam, and S. Langlois, *Parental origin of triploidy in human fetuses: evidence for genomic imprinting*. Human Genetics, 1993. **92**(5): p. 465-9.
219. Dietzsch, E., M. Ramsay, A.L. Christianson, B.D. Henderson, and T.J. de Ravel, *Maternal origin of extra haploid set of chromosomes in third trimester triploid fetuses*. American journal of medical genetics, 1995. **58**(4): p. 360-4.
220. Miny, P., et al., *Parental origin of the extra haploid chromosome set in triploidies diagnosed prenatally*. American journal of medical genetics, 1995. **57**(1): p. 102-6.
221. Redline, R.W., T. Hassold, and M.V. Zaragoza, *Prevalence of the partial molar phenotype in triploidy of maternal and paternal origin*. Human pathology, 1998. **29**(5): p. 505-11.
222. McFadden, D.E. and S. Langlois, *Parental and meiotic origin of triploidy in the embryonic and fetal periods*. Clinical genetics, 2000. **58**(3): p. 192-200.
223. Baumer, A., D. Balmer, F. Binkert, and A. Schinzel, *Parental origin and mechanisms of formation of triploidy: a study of 25 cases*. European journal of human genetics : EJHG, 2000. **8**(12): p. 911-7.
224. McWeeney, D.T., S. Munne, R.C. Miller, N.A. Cekleniak, S.A. Contag, J.R. Wax, W.J. Polzin, and W.J. Watson, *Pregnancy complicated by triploidy: a comparison of the three karyotypes*. American journal of perinatology, 2009. **26**(9): p. 641-5.
225. Gartler, S.M., K.R. Varadarajan, P. Luo, T.H. Norwood, T.K. Canfield, and R.S. Hansen, *Abnormal X: autosome ratio, but normal X chromosome inactivation in human triploid cultures*. BMC genetics, 2006. **7**: p. 41.
226. Jacobs, P.A., A.M. Matsuyama, I.M. Buchanan, and C. Wilson, *Late replicating X chromosomes in human triploidy*. American journal of human genetics, 1979. **31**(4): p. 446-57.
227. Migeon, B.R., J.A. Sprenkle, and T.T. Do, *Stability of the "two active X" phenotype in triploid somatic cells*. Cell, 1979. **18**: p. 637-641.

228. Vogel, W., T. Trautmann, H. Horler, and S. Pentz, *Cytogenetic and biochemical investigations on fibroblast cultures and clones with one and two active X chromosomes of a 69,XXY triploidy*. Human Genetics, 1983. **64**(3): p. 246-8.
229. Fryns, J.P., A. van de Kerckhove, P. Goddeeris, and H. van den Berghe, *Unusually long survival in a case of full triploidy of maternal origin*. Human Genetics, 1977. **38**(2): p. 147-55.
230. Maraschio, P., C. Danesino, F. Lo Curto, O. Zuffardi, T. Dalla Fior, and D. Pedrotti, *A liveborn 69,XXX triploid. Origin, X chromosome activity and gene dosage*. Ann. Génét., 1984. **27**: p. 96-101.
231. Liu, W. and X. Sun, *Skewed X chromosome inactivation in diploid and triploid female human embryonic stem cells*. Human reproduction, 2009. **24**(8): p. 1834-43.
232. Deng, X., D.K. Nguyen, R.S. Hansen, D.L. Van Dyke, S.M. Gartler, and C.M. Disteche, *Dosage regulation of the active X chromosome in human triploid cells*. PLoS genetics, 2009. **5**(12): p. e1000751.
233. Yu, C.W., H. Chen, and M. Fowler, *Specific terminal DNA replication sequence of X chromosomes in different tissues of a live-born triploid infant*. American journal of medical genetics, 1983. **14**(3): p. 501-11.
234. Mermoud, J.E., C. Costanzi, J.R. Pehrson, and N. Brockdorff, *Histone macroH2A1.2 relocates to the inactive X chromosome after initiation and propagation of X-inactivation*. J. Cell Biol., 1999. **147**: p. 1399-1408.
235. Chadwick, B.P. and H.F. Willard, *Chromatin of the Barr body: histone and non-histone proteins associated with or excluded from the inactive X chromosome*. Hum. Mol. Genet., 2003. **12**: p. 2167-2178.
236. Heard, E., C. Rougelle, D. Arnaud, P. Avner, C.D. Allis, and D.L. Spector, *Methylation of histone H3 at Lys-9 is an early mark on the X chromosome during X inactivation*. Cell, 2001. **107**: p. 727-738.
237. Peters, A.H., J.E. Mermoud, D. O'Carroll, M. Pagani, D. Schweizer, N. Brockdorff, and T. Jenuwein, *Histone H3 lysine 9 methylation is an epigenetic imprint of facultative heterochromatin*. Nat Genet, 2002. **30**(1): p. 77-80.
238. de Napoles, M., et al., *Polycomb Group Proteins Ring1A/B Link Ubiquitylation of Histone H2A to Heritable Gene Silencing and X Inactivation*. Dev Cell, 2004. **7**(5): p. 663-76.
239. Fang, J., T. Chen, B. Chadwick, E. Li, and Y. Zhang, *Ring1b-mediated H2A ubiquitination associates with inactive X chromosomes and is involved in Initiation of X-inactivation*. J Biol Chem, 2004.
240. Shen, Y., Y. Matsuno, S.D. Fouse, N. Rao, S. Root, R. Xu, M. Pellegrini, A.D. Riggs, and G. Fan, *X-inactivation in female human embryonic stem cells is in a nonrandom pattern and prone to epigenetic alterations*. Proceedings of the National Academy of Sciences of the United States of America, 2008. **105**(12): p. 4709-14.
241. Chaumeil, J., I. Okamoto, M. Guggiari, and E. Heard, *Integrated kinetics of X chromosome inactivation in differentiating embryonic stem cells*. Cytogenet Genome Res, 2002. **99**(1-4): p. 75-84.
242. Keohane, A.M., L.P. O'Neill, N.D. Belyaev, J.S. Lavender, and B.M. Turner, *X-inactivation and histone H4 acetylation in embryonic stem cells*. Dev. Biol., 1996. **180**: p. 618-630.
243. Casas-Delucchi, C.S., A. Brero, H.P. Rahn, I. Solovei, A. Wutz, T. Cremer, H. Leonhardt, and M.C. Cardoso, *Histone acetylation controls the inactive X chromosome replication dynamics*. Nature communications, 2011. **2**: p. 222.
244. Niebuhr, E., *Triploidy in man. Cytogenetical and clinical aspects*. Humangenetik, 1974. **21**(2): p. 103-25.
245. Ohno, M., T. Maeda, and A. Matsunobu, *A cytogenetic study of spontaneous abortions with direct analysis of chorionic villi*. Obstetrics and gynecology, 1991. **77**(3): p. 394-8.
246. Neuber, M., H. Rehder, C. Zuther, R. Lettau, and E. Schwinger, *Polyploidies in abortion material decrease with maternal age*. Human Genetics, 1993. **91**(6): p. 563-6.

247. Daniel, A., et al., *Karyotype, phenotype and parental origin in 19 cases of triploidy*. Prenatal diagnosis, 2001. **21**(12): p. 1034-48.
248. Willard, H.F. and W.R. Breg, *Human X chromosomes: synchrony of DNA replication in diploid and triploid fibroblasts with multiple active or inactive X chromosomes*. Somat. Cell Mol. Genet., 1980. **6**(2): p. 187-198.
249. Turner, G. and B. Turner, *X-linked mental retardation*. Journal of medical genetics, 1974. **11**(2): p. 109-13.
250. Croen, L.A., J.K. Grether, and S. Selvin, *The epidemiology of mental retardation of unknown cause*. Pediatrics, 2001. **107**(6): p. E86.
251. Ropers, H.H., *X-linked mental retardation: many genes for a complex disorder*. Current opinion in genetics & development, 2006. **16**(3): p. 260-9.
252. Papageorgiou, E.A., H. Fiegler, V. Rakyan, S. Beck, M. Hulten, K. Lamnissou, N.P. Carter, and P.C. Patsalis, *Sites of Differential DNA Methylation between Placenta and Peripheral Blood: Molecular Markers for Noninvasive Prenatal Diagnosis of Aneuploidies*. Am. J. Pathol., 2009. **174**(5): p. 1609-1618.
253. Bolstad, B.M., R.A. Irizarry, M. Astrand, and T.P. Speed, *A comparison of normalization methods for high density oligonucleotide array data based on variance and bias*. Bioinformatics, 2003. **19**(2): p. 185-93.
254. Du, P., X. Zhang, C.C. Huang, N. Jafari, W.A. Kibbe, L. Hou, and S.M. Lin, *Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis*. BMC bioinformatics, 2010. **11**: p. 587.
255. Wang, Y. and F.C. Leung, *An evaluation of new criteria for CpG islands in the human genome as gene markers*. Bioinformatics, 2004. **20**(7): p. 1170-7.
256. Smit, A.F.A., R. Hubley, and P. Green. *RepeatMasker Open-3.0*. 1996-2010; Available from: <http://www.repeatmasker.org>.
257. Fujita, P.A., et al., *The UCSC Genome Browser database: update 2011*. Nucleic Acids Research, 2011. **39**(Database issue): p. D876-82.
258. Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman, *Basic local alignment search tool*. Journal of Molecular Biology, 1990. **215**: p. 403-410.
259. De Smet, C., C. Lurquin, B. Lethe, V. Martelange, and T. Boon, *DNA methylation is the primary silencing mechanism for a set of germ line- and tumor-specific genes with a CpG-rich promoter*. Molecular and Cellular Biology, 1999. **19**(11): p. 7327-35.
260. Team, R.D.C., *R: A Language and Environment for Statistical Computing*, R.F.f.S. Computing, Editor 2010: Vienna, Austria.
261. Ramakers, C., J.M. Ruijter, R.H. Deprez, and A.F. Moorman, *Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data*. Neuroscience letters, 2003. **339**(1): p. 62-6.
262. Ruijter, J.M., C. Ramakers, W.M. Hoogaars, Y. Karlen, O. Bakker, M.J. van den Hoff, and A.F. Moorman, *Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data*. Nucleic Acids Research, 2009. **37**(6): p. e45.
263. Down, T.A., et al., *A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis*. Nature biotechnology, 2008. **26**(7): p. 779-85.
264. Goecks, J., A. Nekrutenko, and J. Taylor, *Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences*. Genome Biology, 2010. **11**(8): p. R86.
265. Blankenberg, D., G. Von Kuster, N. Coraor, G. Ananda, R. Lazarus, M. Mangan, A. Nekrutenko, and J. Taylor, *Galaxy: a web-based genome analysis tool for experimentalists*. Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al.], 2010. **Chapter 19**: p. Unit 19 10 1-21.
266. Su, A.I., et al., *Large-scale analysis of the human and mouse transcriptomes*. Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**(7): p. 4465-70.

267. Brinkman, A.B., T. Roelofsen, S.W. Pennings, J.H. Martens, T. Jenuwein, and H.G. Stunnenberg, *Histone modification patterns associated with the human X chromosome*. EMBO reports, 2006. **7**(6): p. 628-34.
268. Lopes, A.M., N. Ross, J. Close, A. Dagnall, A. Amorim, and T.J. Crow, *Inactivation status of PCDH11X: sexual dimorphisms in gene expression levels in brain*. Human Genetics, 2006. **119**(3): p. 267-75.
269. Yang, X., E.E. Schadt, S. Wang, H. Wang, A.P. Arnold, L. Ingram-Drake, T.A. Drake, and A.J. Lusis, *Tissue-specific expression and regulation of sexually dimorphic genes in mice*. Genome research, 2006. **16**(8): p. 995-1004.
270. Carrel, L. and H.F. Willard, *Heterogeneous gene expression from the inactive X chromosome: an X-linked gene that escapes X inactivation in some human cell lines but is inactivated in others*. Proc. Natl. Acad. Sci. USA, 1999. **96**: p. 7364-7369.
271. Yorifuji, T., J. Muroi, A. Uematsu, K. Tanaka, K. Kiwaki, F. Endo, I. Matsuda, H. Nagasaka, and K. Furusho, *X-inactivation pattern in the liver of a manifesting female with ornithine transcarbamylase (OTC) deficiency*. Clinical genetics, 1998. **54**(4): p. 349-53.
272. Jamieson, R.V., P.P. Tam, and M. Gardiner-Garden, *X-chromosome activity: impact of imprinting and chromatin structure*. Int. J. Dev. Biol. , 1996. **40**(6): p. 1065-1080.
273. Cotton, A.M., L. Avila, M.S. Penaherrera, J.G. Affleck, W.P. Robinson, and C.J. Brown, *Inactive X chromosome-specific reduction in placental DNA methylation*. Human molecular genetics, 2009. **18**(19): p. 3544-52.
274. Stabellini, R., J.C. de Mello, L.M. Hernandez, and L.V. Pereira, *MAOA and GYG2 are submitted to X chromosome inactivation in human fibroblasts*. Epigenetics : official journal of the DNA Methylation Society, 2009. **4**(6): p. 388-93.
275. Gartler, S.M. and M.A. Goldman, *Reactivation of inactive X-linked genes*. Dev. Genet., 1994. **15**: p. 504-514.
276. Xiong, Y.Y., X.S. Chen, Z.D. Chen, X.Z. Wang, S.H. Shi, X.Q. Wang, J.Z. Zhang, and X.L. He, *RNA sequencing shows no dosage compensation of the active X-chromosome*. Nature Genetics, 2010. **42**(12): p. 1043-U29.
277. Nguyen, D.K. and C.M. Disteche, *Dosage compensation of the active X chromosome in mammals*. Nat. Genet., 2006. **38**(1): p. 47-53.
278. Vicoso, B. and B. Charlesworth, *Evolution on the X chromosome: unusual patterns and processes*. Nature reviews. Genetics, 2006. **7**(8): p. 645-53.
279. Rakyan, V.K., et al., *An integrated resource for genome-wide identification and analysis of human tissue-specific differentially methylated regions (tDMRs)*. Genome Res., 2008. **18**(9): p. 1518-1529.
280. Heintzman, N.D., et al., *Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome*. Nat Genet, 2007. **39**(3): p. 311-8.
281. Bittel, D.C., M.F. Theodoro, N. Kibiryeve, W. Fischer, Z. Talebizadeh, and M.G. Butler, *Comparison of X-chromosome inactivation patterns in multiple tissues from human females*. Journal of medical genetics, 2008. **45**(5): p. 309-13.
282. Saxonov, S., P. Berg, and D. Brutlag, *A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters*. Proc. Natl. Acad. Sci. USA, 2006. **103**(5): p. 1412-1417.
283. Bird, A.P., *CpG-rich islands and the function of DNA methylation*. Nature, 1986. **321**: p. 209-213.
284. Hodges, E., et al., *High definition profiling of mammalian DNA methylation by array capture and single molecule bisulfite sequencing*. Genome research, 2009. **19**(9): p. 1593-605.
285. Estecio, M.R., et al., *LINE-1 hypomethylation in cancer is highly variable and inversely correlated with microsatellite instability*. PLoS ONE 2007. **2**(5): p. e399.
286. Miller, S.A., D.D. Dykes, and H.F. Polesky, *A simple salting out procedure for extracting DNA from human nucleated cells*. Nucleic Acids Res., 1988. **16**: p. 1215.

287. Penaherrera, M.S., S. Ma, B. Ho Yuen, C.J. Brown, and W.P. Robinson, *X-chromosome inactivation (XCI) patterns in placental tissues of a paternally derived bal t(X;20) case*. Am. J. Med. Genet., 2003. **118A**(1): p. 29-34.
288. Tost, J. and I.G. Gut, *DNA methylation analysis by pyrosequencing*. Nat. Protoc., 2007. **2**(9): p. 2265-2275.
289. Wang, Z., et al., *Combinatorial patterns of histone acetylations and methylations in the human genome*. Nat. Genet., 2008. **40**(7): p. 897-903.
290. Barski, A., S. Cuddapah, K. Cui, T.Y. Roh, D.E. Schones, Z. Wang, G. Wei, I. Chepelev, and K. Zhao, *High-resolution profiling of histone methylations in the human genome*. Cell, 2007. **129**(4): p. 823-837.
291. Ohba, N. and T. Yamashita, *Primary vitreoretinal dysplasia resembling Norrie's disease in a female: association with X autosome chromosomal translocation*. Brit. J. Ophthal., 1986. **70**: p. 64-71.
292. Kerkel, K., et al., *Genomic surveys by methylation-sensitive SNP analysis identify sequence-dependent allele-specific DNA methylation*. Nat. Genet., 2008. **40**(7): p. 904-908.
293. Hansen, R.S., T.K. Canfield, A.M. Stanek, E.A. Keitges, and S.M. Gartler, *Reactivation of XIST in normal fibroblasts and a somatic cell hybrid: abnormal localization of XIST RNA in hybrid cells*. Proc. Natl. Acad. Sci. USA, 1998. **95**: p. 5133-5138.
294. Zeschnigk, M., et al., *Massive parallel bisulfite sequencing of CG-rich DNA fragments reveals that methylation of many X-chromosomal CpG islands in female blood DNA is incomplete*. Hum. Mol. Genet., 2009. **18**(8): p. 1439-1448.
295. Mohandas, T., R.S. Sparkes, and L.J. Shapiro, *Reactivation of an inactive human X chromosome: evidence for X inactivation by DNA methylation*. Science, 1981. **211**(23 Jan): p. 393-396.
296. Ball, M.P., J.B. Li, Y. Gao, J.H. Lee, E.M. LeProust, I.H. Park, B. Xie, G.Q. Daley, and G.M. Church, *Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells*. Nat. Biotechnol., 2009. **27**(4): p. 361-368.
297. Tsien, F., E.S. Fiala, B. Youn, T.I. Long, P.W. Laird, K. Weissbecker, and M. Ehrlich, *Prolonged culture of normal chorionic villus cells yields ICF syndrome-like chromatin decondensation and rearrangements*. Cytogenet. Genome Res., 2002. **98**(1): p. 13-21.
298. Suzuki, M.M. and A. Bird, *DNA methylation landscapes: provocative insights from epigenomics*. Nat. Rev. Genet., 2008. **9**(6): p. 465-476.
299. Straussman, R., D. Nejman, D. Roberts, I. Steinfeld, B. Blum, N. Benvenisty, I. Simon, Z. Yakhini, and H. Cedar, *Developmental programming of CpG island methylation profiles in the human genome*. Nat. Genet., 2009. **41**(2): p. 178-186.
300. Kubota, T., S. Nonoyama, H. Tonoki, M. Masuno, K. Imaizumi, M. Kojima, K. Wakui, M. Shimadzu, and Y. Fukushima, *A new assay for the analysis of X-chromosome inactivation based on methylation-specific PCR*. Hum. Genet., 1999. **104**: p. 49-55.
301. Yen, P.H., P. Patel, A.C. Chinault, T. Mohandas, and L.J. Shapiro, *Differential methylation of hypoxanthine phosphoribosyltransferase genes on active and inactive human X chromosomes*. Proc. Natl. Acad. Sci. USA, 1984. **81**: p. 1759-1763.
302. Reiss, D., Y. Zhang, and D.L. Mager, *Widely variable endogenous retroviral methylation levels in human placenta*. Nucleic Acids Res., 2007. **35**(14): p. 4743-4754.
303. Lambertini, L., A.I. Diplas, M.J. Lee, R. Sperling, J. Chen, and J. Wetmur, *A sensitive functional assay reveals frequent loss of genomic imprinting in human placenta*. Epigenetics, 2008. **3**(5): p. 251-269.
304. Singer-Sam, J., L. Goldstein, A. Dai, S.M. Gartler, and A.D. Riggs, *A potentially critical Hpa II site of the X chromosome-linked PGK1 gene is unmethylated prior to the onset of meiosis of human oogenic cells*. Proc. Natl. Acad. Sci. USA, 1992. **89**(4): p. 1413-1417.
305. Britten, R.J., D.E. Graham, and B.R. Neufeld, *Analysis of repeating DNA sequences by reassociation*. Methods in enzymology, 1974. **29**(0): p. 363-418.

306. Kang, S.H., C.H. Park, H.C. Jeung, K.Y. Kim, S.Y. Rha, and H.C. Chung, *Integrated in silico and biological validation of the blocking effect of Cot-1 DNA on Microarray-CGH*. International journal of molecular medicine, 2007. **19**(6): p. 901-8.
307. Huynh, K.D. and J.T. Lee, *Inheritance of a pre-inactivated paternal X chromosome in early mouse embryos*. Nature, 2003. **426**(6968): p. 857-62.
308. Cotton, A.M., et al., *Chromosome-wide DNA methylation analysis predicts human tissue-specific X inactivation*. Human Genetics, 2011.
309. Abramoff, M.D., P.J. Magalhaes, and S.J. Ram, *Image Processing with ImageJ*. Biophotonics International, 2004. **11**(7): p. 36-42.
310. Baergen, R.N. and SpringerLink (Online service), *Manual of Pathology of the Human Placenta Second Edition*, 2011, Springer Science+Business Media, LLC: Boston, MA.
311. Chavan-Gautam, P., D. Sundrani, H. Pisal, V. Nimbargi, S. Mehendale, and S. Joshi, *Gestation-dependent changes in human placental global DNA methylation levels*. Molecular reproduction and development, 2011. **78**(3): p. 150.
312. Roberts, C.T., *IFPA Award in Placentology Lecture: Complicated interactions between genes and the environment in placentation, pregnancy outcome and long term health*. Placenta, 2010. **31 Suppl**: p. S47-53.
313. Gama-Sosa, M.A., R.Y. Wang, K.C. Kuo, C.W. Gehrke, and M. Ehrlich, *The 5-methylcytosine content of highly repeated sequences in human DNA*. Nucleic Acids Research, 1983. **11**(10): p. 3087-95.
314. Macaulay, E.C., R.J. Weeks, S. Andrews, and I.M. Morison, *Hypomethylation of functional retrotransposon-derived genes in the human placenta*. Mammalian genome : official journal of the International Mammalian Genome Society, 2011.
315. Antequera, F. and A. Bird, *Number of CpG islands and genes in human and mouse*. Proc. Natl. Acad. Sci., 1993. **90**: p. 11995-11999.
316. Duthie, S.M., T.B. Nesterova, E.J. Formstone, A.M. Keohane, B.M. Turner, S.M. Zakian, and N. Brockdorff, *Xist RNA exhibits a banded localization on the inactive X chromosome and is excluded from autosomal material in cis*. Hum. Mol. Genet., 1999. **8**: p. 195-204.
317. Disteché, C.M., *Escapees on the X chromosome*. Proc. Natl. Acad. Sci., USA, 1999. **96**: p. 14180-14182.
318. Grumbach, M.M., A. Morishima, and J.H. Taylor, *Human sex chromosome abnormalities in relation to DNA replication and heterochromatinization*. Proc. Natl. Acad. Sci. USA, 1963. **49**(5): p. 581-588.
319. Webb, S., T.J. de Vries, and M.H. Kaufman, *The differential staining pattern of the X chromosome in the embryonic and extraembryonic tissues of postimplantation homozygous tetraploid mouse embryos*. Genetical research, 1992. **59**(3): p. 205-14.
320. Monkhorst, K., I. Jonkers, E. Rentmeester, F. Grosveld, and J. Gribnau, *X inactivation counting and choice is a stochastic process: evidence for involvement of an X-linked activator*. Cell, 2008. **132**(3): p. 410-21.
321. McFadden, D.E. and D.K. Kalousek, *Two different phenotypes of fetuses with chromosomal triploidy: correlation with parental origin of the extra haploid set*. American journal of medical genetics, 1991. **38**(4): p. 535-8.
322. Robinson, W.P., et al., *Origin of amnion and implications for evaluation of the fetal genotype in cases of mosaicism*. Prenatal diagnosis, 2002. **22**(12): p. 1076-85.
323. Palacios, R., E. Gazave, J. Goni, G. Piedrafita, O. Fernando, A. Navarro, and P. Villoslada, *Allele-specific gene expression is widespread across the genome and biological processes*. PLoS ONE, 2009. **4**(1): p. e4150.
324. Dementyeva, E.V., A.I. Shevchenko, O.V. Anopriyenko, N.A. Mazurok, E.A. Elisaphenko, T.B. Nesterova, N. Brockdorff, and S.M. Zakian, *Difference between random and imprinted X inactivation in common voles*. Chromosoma, 2010. **119**(5): p. 541-52.

325. Hadjantonakis, A.K., L.L. Cox, P.P. Tam, and A. Nagy, *An X-linked GFP transgene reveals unexpected paternal X-chromosome activity in trophoblastic giant cells of the mouse placenta*. *Genesis*, 2001. **29**(3): p. 133-40.
326. Avila, L., R.K. Yuen, D. Diego-Alvarez, M.S. Penaherrera, R. Jiang, and W.P. Robinson, *Evaluating DNA methylation and gene expression variability in the human term placenta*. *Placenta*, 2010. **31**(12): p. 1070-7.
327. McNeil, J.A., K.P. Smith, L.L. Hall, and J.B. Lawrence, *Word frequency analysis reveals enrichment of dinucleotide repeats on the human X chromosome and [GATA]_n in the X escape region*. *Genome Res*, 2006. **16**(4): p. 477-84.
328. Cattanaach, B.M. and C.E. Williams, *Evidence of non-random X chromosome activity in the mouse*. *Genet. Res.*, 1972. **19**: p. 229-240.