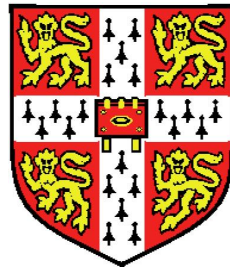


Evolutionary Landscape of CpG Island Methylation in X Chromosome Inactivation



Mengning Liu

Corpus Christi College, University of Cambridge

& The Wellcome Trust Sanger Institute

This dissertation is submitted for the degree of

Doctor of Philosophy

October 2008

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text.

Mengning Liu

To my dearest mum, for giving me unconditional love and the
curiosity genes...

Abstract

Dosage compensation in mammals depends upon X Chromosome Inactivation (XCI), the transcriptional silencing of one X in female cells. In human and mouse, CpG island methylation on the inactive X (X_i) has been implicated in the maintenance of XCI and has been used previously to indicate a gene's XCI status. There is evidence that XCI is more complete in mouse than in human, and so I speculated that CpG island methylation on the X_i might also be more extensive in the mouse. In marsupials, by contrast, the small amount of available evidence points to absence of methylation on the X_i . I have studied the methylation of CpG islands on the X chromosomes of human, mouse and opossum in order to provide further evidence for these suggested species differences.

Ninety-one human genes, including several thought to escape from XCI, were assayed together with the mouse orthologues for 52 of these genes. Female and male genomic DNA was digested with *MspI* or its methylation-sensitive isoschizomer *HpaII* then tested the ability of the DNA to support amplification of PCR products containing multiple *MspI/HpaII* sites. Only the three known mouse escapees tested (*Eif2s3*, *Ddx3x*, *Utx*) were hypomethylated in the female, compared to 13 genes in human. This is consistent with the suggestion that gene

silencing is more effective in the mouse. Furthermore, the assay indicated that partial methylation is common among the human genes but rare in the mouse. Bisulphite sequencing of CpG islands from eight human-mouse orthologous pairs has confirmed this difference.

The restriction-PCR assay was then applied to 37 X-linked genes in opossum and CpG island hypermethylation was found to be rare on the female X chromosomes. However, for at least six genes, there was a greater level of methylation in the female sample, which was subsequently confirmed by bisulphite sequencing.

Results from this study support the view that CpG island methylation as a maintenance mechanism for XCI is common in the eutherian mammals. The lower level of methylation in human than mouse is consistent with the suggestion that escape from XCI is associated with a failure to maintain the inactive state. If this is correct, a high level of escape from XCI in marsupials might be anticipated. This study has also provided the first evidence of a possible role of CGI methylation in marsupial XCI.

Acknowledgements

I would firstly like to thank my supervisor Mark Ross, for his encouragement, support, and guidance throughout this project. It was extremely fortunate for me to have such a supervisor, who always had time for me, and taught me so much in every aspect of science, from experimental design to data analysis, from critical thinking to scientific writing. I can not have asked for a better supervisor.

I would also like to thank my second supervisor Nigel Carter, for all the help and support he has given me, without which I could not have finished this project so smoothly after my team left. And thanks to Nigel as the most active member of my thesis committee, who has injected many helpful comments during the entire project. I am also grateful to have the other two members of my thesis committee, Malcolm Ferguson-Smith and Chris Tyler-Smith, for the many lively discussions we had.

Special thanks also to Colette Johnston, for all the experimental skills, precious experience, and inspiring discussions she shared with me; and Alison Coffey, for taking valuable time to proof-read my thesis and the great advice she has offered me. Thanks to Christina Hedberg-Delouka, without her superb temper and organisation skills my PhD would not have progressed so efficiently.

I am indebted to many people at the Sanger Institute for their expertise and kindness: Val Curwen and Steve Searle for computational help; Matthew Francis, Nancy Holroyd, Ruth Taylor, Zahra Abdellah, Carl Skuce, and other members of the faculty small sequencing projects team, for fulfilling my massive sequencing requests with great patience and top quality.

The X team, regrettably no longer existing, will always be fondly remembered. Thanks to my fantastic team-mates Frances Lovell, Christine Burrows, Helen Albert, Daniel Leongamornlert and Yujun Zhang, for all the help and fun. Also thanks to the wonderful people in Team 70 and office F207, for giving me the joy of belonging to a team again. My fellow Sanger PhD students have been a source of delight at all times. Special thanks to Matthew Davis, for sitting next to me in many stressful evenings and weekends; and to Katherine Fawcett, for keeping my spirits up in the final days of my PhD.

Last but not least, I would like to thank my family. Thanks to my parents, for their ongoing love and support and opening my eyes to science; and thanks to my beloved husband, Li Liu, for, really, everything.

Contents

Abstract	iii
Acknowledgements	v
1 Introduction	1
1.1 Dosage compensation and X chromosome inactivation (XCI)	1
1.2 The mechanism of XCI	5
1.3 Escape from XCI	12
1.4 XCI in other mammals	19
1.5 Aims of the thesis	24
2 Materials and Methods	26
2.1 Chemical reagents	26
2.2 Enzymes and commercially prepared kits	27
2.3 Solutions and buffers	28
2.3.1 Buffers	28
2.3.2 Media	29
2.4 Size markers	30
2.5 Tissues and cell lines	31
2.6 Primer sequences	32

CONTENTS

2.7	Key World Wide Web addresses	32
2.8	Mammalian cell culture	33
2.8.1	Resuscitating frozen cells	33
2.8.2	Growing and harvesting cells	33
2.8.3	Freezing cells for storage	34
2.9	DNA extraction	35
2.9.1	Phenol/chloroform extraction of genomic DNA from animal tissues	35
2.9.2	Phenol/chloroform extraction of DNA from a reaction mix	36
2.9.3	Mini-preps of genomic DNA from cultured cells	37
2.10	DNA manipulation	37
2.10.1	Purification of DNA	37
2.10.2	Quantification	38
2.10.3	Restriction enzyme digests	39
2.10.4	Agarose gel electrophoresis	39
2.10.5	Bisulphite modification	40
2.11	Polymerase Chain Reaction (PCR)	41
2.11.1	Primer design	42
2.11.2	Restriction-PCR Methylation Assay (RPMA)	42
2.11.3	Bisulphite PCR	44
2.11.4	Colony PCR	45
2.12	Bacterial cloning	46
2.12.1	Ligation of vector and insert	46
2.12.2	Transformation	46
2.13	Plasmid sequencing	47

2.14 Computational analysis	47
2.14.1 cpg	47
2.14.2 BLAST	47
2.14.3 RepeatMasker	48
2.14.4 GAP4	48
2.14.5 Alignment of DNA sequences	48
2.14.6 Multi-species comparative sequence analysis	49
2.14.7 MethTools	49
2.14.8 Perl scripts	49
3 Comparison of CpG island methylation on the mouse and human	
X chromosomes	51
3.1 Introduction	51
3.1.1 Identification of CGIs	54
3.1.2 Approaches to studying CGI methylation	56
3.1.3 Aims	61
3.2 Identification of target CpG islands	61
3.2.1 Region of interest	61
3.2.2 Identification of target human and mouse CGIs	62
3.3 Assessing CpG island methylation on mouse and human X chromosomes by Restriction-PCR Methylation Assay (RPMA)	63
3.3.1 Optimisation of experimental conditions	65
3.3.2 Primer design	69
3.3.3 Analysis of CGI methylation on the human X chromosome	69
3.3.4 Analysis of CGI methylation on the mouse X chromosome	76

3.4 Discussion	79
4 Comparison of detailed CGI methylation patterns on the human and mouse X chromosomes	84
4.1 Introduction	84
4.1.1 Aims of this chapter	89
4.2 High resolution analysis of methylation levels on the mouse and human X chromosomes by bisulphite sequencing	91
4.2.1 Identification of target CGIs	91
4.2.2 Primer design and PCR	93
4.2.3 Cloning and sequencing	94
4.2.4 Methylation analysis	95
4.2.5 Comparison between human and mouse methylation profiles	97
4.2.6 Comparison between bisulphite sequencing and RPMA results	110
4.3 Discussion	113
5 Investigation of CpG island methylation on the opossum X chromosome	118
5.1 Introduction	118
5.1.1 The grey short-tailed opossum (<i>M. domestica</i>)	121
5.1.2 The opossum X chromosome	122
5.1.3 Aims of this chapter	124
5.2 Identification of CGIs on the opossum X chromosome	124
5.3 Assessing CGI methylation on the opossum X chromosomes by RPMA	125

5.3.1	Analysis of CGI methylation on the opossum X chromosome	126
5.4	Assessing CGI methylation on the opossum X chromosomes by bisulphite sequencing	132
5.5	Comparison of methylation status across multiple tissues	143
5.6	Discussion	153
6	Comparative analysis of features of 5' CGIs in human, mouse and opossum	158
6.1	Introduction	158
6.1.1	Aims of this chapter	160
6.2	CGIs in the conserved region of human, mouse and opossum X chromosome	160
6.3	CGIs associated with orthologous genes on human, mouse and opossum XCR	163
6.4	Conservation of 5' CGIs associated with orthologous genes	165
6.5	Discussion	172
7	Discussion	175
7.1	Future directions	179
7.2	Conclusions	183
References		184

List of Tables

2.1	Key World Wide Web addresses used in this study	32
3.1	RPMA results of human CGIs	74
3.2	RPMA results of mouse CGIs	78
4.1	Human and mouse genes chosen for bisulphite sequencing analysis of CGI methylation	92
4.2	Comparison between bisulphite sequencing and RPMA results . .	112
5.1	RPMA results of opossum CGIs	131
6.1	Comparison of XCR gene and CGI features in opossum, human, and mouse.	162
6.2	Comparison of 5' CGIs associated with the orthologous gene set. .	165

List of Figures

1.1	Tortoiseshell cat	5
1.2	Evolutionary strata of human X chromosome	15
3.1	Approaches to studying DNA methylation.	60
3.2	Restriction-PCR Methylation Assay (RPMA)	65
3.3	Optimisation of restriction enzyme digestion conditions	68
3.4	RPMA results of human CGIs with hypomethylation or female-specific methylation	72
3.5	RPMA results of human CGIs with other methylation patterns	73
3.6	RPMA results of mouse CGIs with hypomethylation or female-specific methylation of CGI	77
4.1	Bisulphite sequencing	86
4.2	Hypothesised model of CGI methylation patterns on the human and mouse X chromosomes	90
4.3	Graphical representation of CGI methylation.	96
4.4	CGI methylation profiles of human <i>DDX3X</i> and mouse <i>Ddx3x</i>	98
4.5	CGI methylation profiles of human <i>UTX</i> and mouse <i>Utx</i>	99
4.6	CGI methylation profiles of human <i>EIF2S3</i> and mouse <i>Eif2s3x</i>	100

LIST OF FIGURES

4.7	CGI methylation profiles of human <i>OFD1</i> and mouse <i>Ofd1</i>	101
4.8	CGI methylation profiles of human <i>POLA</i> and mouse <i>Pola1</i>	103
4.9	CGI methylation profiles of human <i>MSL3L1</i> and mouse <i>Msl3l1</i>	104
4.10	CGI methylation profiles of human <i>ATP6AP2</i> and mouse <i>Atp6ap2</i>	105
4.11	CGI methylation profiles of human <i>PRPS2</i> and mouse <i>Prps2</i>	106
4.12	CGI methylation profiles of human <i>RAB9A</i>	107
4.13	CGI methylation profiles of human <i>CDKL5</i>	108
4.14	CGI methylation profiles of mouse <i>Syap1</i>	109
4.15	CGI methylation profiles of mouse <i>Hccs</i>	110
4.16	Variation of CGI methylation on the human and mouse X chromosomes - revised model	117
5.1	South American, grey short-tailed opossum (<i>Monodelphis domestica</i>)	122
5.2	RPMA results of three opossum genes with female-specific methylation of CGI	130
5.3	Distribution of assayed CGIs on the opossum X chromosome	132
5.4	CGI methylation profiles of opossum <i>G6PD</i>	134
5.5	CGI methylation profiles of opossum <i>HPRT1</i>	135
5.6	CGI methylation profiles of opossum <i>FAAH2</i>	138
5.7	CGI methylation profiles of opossum <i>FGF16</i>	139
5.8	CGI methylation profiles of opossum <i>IAG2</i>	140
5.9	CGI methylation profiles of opossum <i>PGK1</i>	141
5.10	CGI methylation profiles of opossum <i>ITM2A</i>	142
5.11	CGI methylation profiles of opossum <i>F8A~</i>	143

LIST OF FIGURES

5.12	Graphical representation of average methylation levels in individual CGI	145
5.13	CGI methylation profiles of opossum <i>G6PD</i> in four tissues	146
5.14	CGI methylation profiles of opossum <i>HPRT1</i> in three tissues	147
5.15	CGI methylation profiles of opossum <i>FGF16</i> in four tissues	148
5.16	CGI methylation profiles of opossum <i>PGK1</i> in four tissues	149
5.17	CGI methylation profiles of opossum <i>ITM2A</i> in four tissues	150
5.18	CGI methylation profiles of opossum <i>IAG2</i> in four tissues	151
5.19	CGI methylation profiles of opossum <i>F8A~</i> in four tissues	152
6.1	Comparison of sizes of 5' CGIs in opossum, human, and mouse XCR.	163
6.2	Orthologous genes on the opossum, human, and mouse X chromosomes	164
6.3	CGIs associated with the orthologous gene set	165
6.4	Homology between 5' CGIs of the orthologous genes.	166
6.5	VISTA plot of human, mouse and opossum <i>PGK1</i> and its CGI	168
6.6	VISTA plot of human, mouse and opossum <i>SOX3</i> and its CGI	169
6.7	VISTA plot of human, mouse and opossum <i>PGRMC1</i> and its CGI	170
6.8	VISTA plot of human, mouse and opossum <i>PIM2</i> and its CGI	171
7.1	Phylogeny and time scale of mammalian species with available sequence information	181