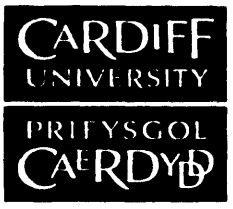


BINDING SERVICES
Tel +44 (0)29 2087 4949
Fax +44 (0)29 20371921
e-mail bindery@cardiff.ac.uk

"*In vivo* Consequences of Removal of Methyl CpG Binding Domain Proteins."

Kathryn Maddison

Supervisor : Prof Alan Clarke

PhD

Cardiff University

2005

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ACKNOWLEDGEMENTS.

Thanks to Nathan Hill for maintaining animal stocks, to Derek Scarborough for histology and to Yvonne Hay for array processing.

Special thanks must go to my supervisor Prof. Alan Clarke and to Dr. Owen Sansom, whose advice, support and optimism has been invaluable.

Many thanks also to my long-suffering family, who have been wonderful throughout, and also to my friends and Andy, who always managed to cheer me up when things went wrong.

TABLE OF CONTENTS.

DECLARATION. i
ACKNOWLEDGEMENTS. ii
TABLE OF CONTENTS. iii
ABBREVIATIONS. viii
LIST OF FIGURES. xiv
LIST OF TABLES. xvii

Chapter 1: Introduction. 1
1.1. DNA methylation and epigenetics. 1
1.1.1 Genetic and Epigenetic information. 1
1.1.2. Methylation of DNA. 2
1.1.3. Patterns of methylation: CpGs and methylated CpGs are non-randomly distributed in the mammalian genome. 3
1.1.4. Methylation of DNA in Other Species. 5
1.1.5. Functions of DNA methylation 6
1.2. Genomic Imprinting. 6
1.2.1. Evolution of Imprinting in Mammals. 9
1.2.2. X-Chromosome inactivation. 13
1.3. The Chromatin Environment and the Histone Code. 15
1.3.1. The histone code. 16
1.4. DNA methylation and transcriptional silencing. 18
1.4.1. Reading and writing the methylation marks: methyl CpG binding domain proteins and DNA methyltransferases. 19
1.4.2. DNA methyltransferases. 19
1.4.3. Dnmt1. 20
1.4.4. Dnmt2 21
1.4.5. Dnmt3a, Dnmt3b and Dnmt3l 21
1.4.6. Establishment and Maintenance of Methylation Marks. 23
1.4.7. How are the sex-specific methylation marks reset correctly? 25
1.5. Reading the methylation marks: methyl binding domain proteins. 26
1.5.1. Mbd1. 28
1.5.2. Mbd2. 29
1.5.3. Is Mbd2 a demethylase? 31
1.5.4. Mbd3. 31
1.5.5. Mbd4. 33
1.5.6. Kaiso. 36
1.6. MECP2 and Rett Syndrome. 37
1.6.1. Mutations in Methyl CpG Binding Protein 2 *(MECP2)* cause Rett syndrome. 39
1.6.2. Rett syndrome is phenotypically diverse. 40
1.6.3. Genotype/Phenotype correlation in Rett syndrome. 42
1.6.4. Several neurodevelopmental syndromes have considerable phenotypic and genetic overlap. 45
1.6.5. Rett Syndrome in Males. 47

1.6.6. Mouse models of Rett syndrome. 49
1.6.7. Conditional alleles of *Mecp2*. 51
1.6.8. Elevated levels of MeCP2 also cause a progressive neurological disorder. 55
1.6.9. Neuropathological difference in the Rett syndrome brain. 56
1.6.10. Rett Syndrome Causes Defects in Synaptic Elaboration and Pruning. 58
1.6.11. MeCP2 and neurogenesis – narrowing down the phenotype. 59
1.7. MECP2 – A member of the Methyl Binding Domain Family. 60
1.7.1. MECP2A and MECP2B: alternative splicing of MECP2. 61
1.7.2. MECP2 – Function *in vivo*. 64
1.7.3. MECP2 links DNA methylation to histone methylation. 66
1.7.4. Is MeCP2 Really a Global Transcriptional Repressor? 67
1.7.5. MeCP2 Directly regulates *Bdnf* and *xHairy2a*. 69
1.7.6. MeCP2 and *Bdnf*. 69
1.7.7. Is Rett syndrome a disorder of altered imprinting? 73
1.8. Examining the role of MBD proteins in the murine intestine. 74

Chapter 2. Materials and Methods. 76

2.1. Immunohistochemistry/stains. 76
2.1.1 H&E (Haematoxylin and Eosin) staining 76
2.1.2. GIP (Gastrointestinal polypeptide) staining. 77
2.1.3.Grimelius (Enteroendocrine cell stain.) 79
2.1.4. Alcian Blue. 80
2.1.5.Alkaline phosphatase. 81
2.1.6. BrdU (Bromodeoxyuridine staining) 81
2.2. Genotyping. 82
2.2.1. DNA Purification: Puregene method 82
2.2.2. *Mecp2* PCR. 83
2.2.3. *Blg-Cre* PCR. 84
2.2.4. *Ah-Cre* PCR 85
2.2.5. *Mbd4* PCR 86
2.2.6. *Mlh1* PCR 86
2.2.7. Apc^{flox} PCR 87
2.2.8. *LacZ (Rosa26R)* PCR. 88
2.3. LacZ Staining of small intestinal wholemounts. 89
2.3.1. DTT Demucifying Solution. 90
2.3.2. Wax Plates. 90
2.3.3. Mammary Wholemounts. 90
2.3.4. 2% Paraformaldehyde. 91
2.4. Scoring Apoptosis. 91
2.5. Preparation of Murine Colonic / Small Intestinal Crypts. 91
2.5.1. Short-Term Culture of Colonic Crypts. 92
2.6. Cell Culture. 92
2.6.1. Preparation of murine embryonic kidney cells. 92
2.6.2. To trypsinise the cells (for passaging) 93
2.7. RNA extraction and preparation for Affymetrix Microarray. 94
2.7.1. RNA Extraction From Intestinal Tissue. 94

2.7.2. Fragmentation buffer 95
2.7.3. Obtaining RNA from Tissue Samples. 95
2.7.4. First Strand cDNA Synthesis. 96
2.7.5. Second Strand Synthesis. 97
2.7.5. Clean-Up of Double-Stranded cDNA. 98
2.7.6. *In Vitro* Transcription. 99
2.7.7. Clean-Up of cRNA 100
2.7.8. Ethanol Precipitation. 101
2.7.9. cRNA Fragmentation. 101
2.8. RNA Gel. 102
2.8.2. 10x MOPS. 102
2.8.3. Checking the RNA Quality. 103
2.9. RNA extraction and preparation for RT-PCR 103
2.9.1. Obtaining RNA from Tissue Samples. 103
2.10. Making cDNA from total cellular RNA. 105
2.11. Checking for genomic DNA contamination. 105
2.12. RT-PCR. 105
2.12.1. RT-PCR primers 106

Chapter 3: Investigating the Role of Mbd4 in the Intestine. 108

3.1. Apoptosis and the response to DNA damage. 108
3.1.1. Mismatch Repair 109
3.1.2. Defective MMR and human disease. 112
3.1.3. Mbd4 and Apoptosis. 114
3.1.4. Apoptosis and anoikis. 114
3.1.5. Fas-Mediated Apoptosis. 115
3.1.6. Anoikis. 116
3.2. Results 119
3.2.1. Mbd4 influences anoikis in isolated murine small intestinal crypts. 119
3.2.2. The role of Mbd4 in Fas-mediated apoptosis. 123
3.2.3. Investigating the interaction between Mlh1 and Mbd4. 127
3.2.4. Characterising the apoptotic response in *Mbd4/Mlh1* double null mice. 133
3.3. Discussion. 135
3.3.1. Interaction between Mlh1 and Mbd4. 137

Chapter 4: Investigating the Role of MeCP2 in the development and involution of the murine mammary gland. 141

4.1. Mammary growth and development in the mouse. 141
4.2. Methyl Binding Proteins in the Developing Mammary Gland. 144
4.3. Investigating the role of methyl CpG binding domain proteins in the mammary gland. 145
4.3.1. Investigating the role of MeCP2 in the developing murine mammary gland. 146
4.4. Investigating the Role of MeCP2 in the Involution of the Mammary Gland. 149
4.5. Conclusions. 153

Chapter 5: Characterising the Phenotype of Loss of MeCP2 in the Murine Small Intestine. 154

5.1. *MECP2* and Rett syndrome. 154
5.1.1 Methyl CpG binding proteins and the intestine. 155
5.1.2. The small intestine as a model system; development of the gut. 155
5.1.3. Normal structure of the small intestinal epithelium. 156
5.2. Examining the consequences of loss of MeCP2 in the murine intestine. 160
5.2.1. Removing MeCP2 in the murine intestine; quantitative real-time RT-PCR. 162
5.3. Characterisation of the intestinal phenotype of MeCP2-deficient mice. 163
5.4. Further characterisation of the MeCP2 deficient phenotype. 165
5.4.1. Crypt size remains constant in the absence of MeCP2. 165
5.4.2. Loss of MeCP2 does not alter apoptosis in the crypt. 169
5.4.3. Loss of MeCP2 in the murine intestine results in faster migration of epithelial cells along the villus. 170
5.5. Immunohistochemistry of MeCP2-deficient and wild-type intestinal tissue. 173
5.5.1. The number of enteroendocrine cells is not altered in the absence of MeCP2. 175
5.5.2. Loss of MeCP2 does not alter the number of K-cells in the small intestine. 178
5.6. Loss of MeCP2: Summary of the intestinal phenotype. 180
5.7. Discussion. 181
5.7.1. Which factors can alter proliferation in the intestine? 181
5.7.2. The Proglucagon-Derived Peptides. 183
5.7.3. Glp-2 Synthesis in the Intestine. 185
5.7.4. Glp-2 and MeCP2. 186
5.7.5. Gip. 187
5.7.6. Igf-I. 188

Chapter 6: Molecular characterisation of the intestinal phenotype induced by loss of MeCP2. 190

6.1. Loss of MeCP2 in the murine intestine results in a stable phenotype. 190
6.2. Microarrays: vital tools in examining the transcriptome. 192
6.2.1. Microarray Analysis of *Mecp2*$^{flox.y}$ *Cre*$^{+}$ and *Mecp2*$^{+/y}$ *Cre*$^{+}$ intestinal tissue. 194
6.3. Analysis of Microarray Data. 196
6.3.1. Method 1: Analysis using SAM. 197
6.4. Method 2. Analysis using a T-test. 202
6.5. Identification of targets for follow-up using quantitative RT-PCR. 205
6.6. Confirmation of Array Targets: Quantitative Real-time RT-PCR. 206
6.7. Using QRT-PCR to examine array targets. 208
6.8. Quantitative real-time RT-PCR analysis of targets suggested by microarray analysis of 211
Mecp2$^{+/+}$ and *Mecp2*$^{-/-}$ deficient murine small intestine. 211
6.8.1. Example of QRT-PCR calculation: Gip. 212
6.9. Discussion. 216
6.9.1. How does loss of MeCP2 cause the villus lengthening phenotype? 216
6.9.2. Bdnf is involved in energy balance and metabolism. 218

6.10. What can the data reveal about the mechanisms underlying Rett syndrome? 222
6.10.1. Rett syndrome, synaptic plasticity and neuronal maturation; Bdnf and Liprin-α 222
6.10.2. Rett syndrome and bone abnormalities. 225
6.10.3. Does Rett syndrome cause metabolic stress via defects in fatty acid metabolism? 225

Chapter 7: Summary. 229

7.1. The role of MBD proteins is diverse. 229

Published Work. 233

REFERENCES. 234

ABBREVIATIONS.

- 5-AzaC – 5- Aza-cytidine
- 5-FU – 5-fluorouracil
- 5mC – 5- methyl cytosine
- Acetyl CoA – Acetyl coenzyme A.
- AdoMet – Adenosyl – L – methionine
- AhCre – Aryl hydrolase linked Cre recombinase

Amino acids.

G = Glycine (Gly)
P = Proline (Pro)
A = Alanine (Ala)
V = Valine (Val)
L = Leucine (Leu)
I = Isoleucine (Ile)
M = Methionine (Met)
C= Cysteine (Cys)
F = Phenylalanine (Phe)
Y = Tyrosine (Ty)
W = Tryptophan (Trp)
H = Histidine (His)
K = Lysine (Lys)
R = Arginine (Arg)
Q = Glutamine (Gln)
N = Asparagine (Asn)
E = Glutamic acid (Glu)
D = Aspartic acid (Asp)
S = Serine (Ser)
T = Threonine (Thr)

- AMPA – α-amino-3-hydroxyl-S-methyl-4-isoxazolepriopionate
- A-myb – Myeloblastosis viral oncogene
- Apaf1 – Apoptotic protease activating factor
- APC – Adenomatous polyposis coli
- ARF-GEP – ADP ribosylation factor guanine nucleotide exchange factor.
- ARP1 – Actin related protein 1.
- ARX – Aristaless-related homeobox gene
- BDNF – Brain derived neurotrophic factor.
- BER – Base excision repair
- BID – BH3 interacting death domain agonist
- BLGCre - β-lactoglobulin linked Cre recombinase.
- BMP2 – Bone morphogenetic protein 2.

- BrdU - Bromodeoxyuridine.
- BTB/POZ – Broad complex, tramtrack, bric-a-brac/poxvirus and zinc finger
- CAD – Carbamoyl-phosphatase synthetase 2, aspartate transcarbmylase, and dioorotase.
- CCFi – Complement component factor i.
- CCKA – Cholecystokinin receptor A.
- CD44 – CD44 antigen (Homing function and Indian blood group system.)
- CDKL5/STK9 – Cyclin dependent kinase like 5.
- Cdx1 – Caudal type homeobox transcription factor 1
- ChIP – Chromatin immunoprecipitation.
- cII – Reporter locus in the BigBlue mouse.
- CMT3 – Chromodomain containing methyltransferase 3
- CNS – Central nervous system.
- Col2a – Colipase 2a.
- Co-REST – Co-repressor element silencing transcription factor
- CRE – cAMP response element.
- Cre – Causes Recombination (a bacterial recombinase.)
- CREB – CRE binding protein.
- CSF1 – Colony stimulating factor 1.
- c-SKI – c-ski sarcoma viral oncogene homologue
- Ct – Threshold cycle.
- CTCF – CCCTC-binding factor (Zinc finger-like)
- CTD – C-terminal domain
- CytC – Cytochrome C
- CytP450 – Cytochrome P450.
- DeLGEF – Deafness locus associated putative guanine nucleotide exchange factor.
- DEPC – Diethyl pyrocarbonate
- DISC – Death-inducing signalling complex.
- Dlb1 – Dolicos bifluoros 1.
- Dlx5 – Distaless homeobox factor 5
- Dlx6 – Distaless homeobox factor 6
- DMD – Differentially methylated domain
- DMEM – Dulbecco's Modified Eagle Medium
- DNA – Deoxyriobnucleic acid
- DNMT – DNA methyltransferase
- Dpgat1 – GlcNAc-1- P transferase
- DPX – Distrene plasticizer and xylene
- DRE1 – Damage response element binding protein 1
- DRM – Dnmt – related methylase
- E11 – Embryonic day 11.
- EBI-1 – (CCR7) Chemokine (C-C motif) receptor 7
- ECM – Extracellular matrix.
- Eed – Embryonic ectoderm development
- EGF – Epidermal growth factor.

- ENS – Enteric nervous sytem.
- Enx (Hox11L1) – Murine homologue of *Drosophila* Enhancer of Zeste
- ER – Oestrogen recptor.
- ES – Embryonic stem (cell)
- Evl – Ena-vasodilator stimulated phosphoprotein.
- EXO1 – Exonuclease 1
- FADD – Fas associated death domain protein.
- FAP – Familial adenomatous polyposis
- Fas (CD95) – Fas (TNF receptor superfamily, member 6)
- FDR – False discovery rate.
- Fkh6 – Forkhead6.
- FLICE – Caspase 8
- FMR2 – FRAXE mental retardation 2 gene.
- FRAXE – Fragile X.
- GABA – γ-aminobutyric acid
- GABRB3 - γ-aminobutyric acid receptor B3
- Galgt2 – UDP-GalNac:Neu5Ac α 2-3 Gal β 1,4 acetylgalactosaminyltransferase
- GAPDH –Glyceraldehyde-3-phosphate dehydrogenase (reduced)
- GIP – Gastrointestinal polypeptide / glucose dependent insulinotrophic polypeptide.
- GIPR – GIP receptor.
- GLP-1 – Glucagon like peptide 1.
- GLP-2 – Glucagon like peptide 2.
- GLP2R – Glucagon like peptide 2 receptor.
- GRPP – Glicentin related polypeptide.
- GTP – Guanidine tri phosphate
- H&E – Haematoxylin and eosin.
- H19 – Non translated RNA
- H2A – Histone H2A
- H2B – Histone H2B
- H3K27 – Histone 3 lysine 27
- H3K9 – Histone 3 lysine 9
- HBSS – Hanks Balanced Salt Solution
- HDAC – Histone deacetylase
- HNFα/β - Hepatocyte nuclear factor α/β
- HNPCC – Hereditary non-polyposis colorectal cancer
- HPRT – Hypoxanthine phosphoribosyltransferase 1
- IAP – Intracisternal A particle
- ICAD – Form of carbamoyl-phosphatase synthetase 2, aspartate transcarbmylase, and dioorotase.
- ICF – Immunodeficiency, centromeric instability, facial abnormalities syndrome
- ICM –Inner cell mass
- IDL – Insertion/deletion loops
- Igf2 – Insulin like growth factor 2
- IGFBP10 – Insulin-like growth factor binding protein 10.

- IGFBP3 – Insulin-like growth factor binding protein 3.
- IGF-I – Insulin-like growth factor I.
- IGF-II – Insulin like growth factor II.
- IP1/2 – Intervening peptide 1/ 2.
- IPW – Imprinted in Prader-Willi syndrome
- ISSX – Infantile spasm syndrome X-linked
- kb – Kilobase.
- kDa – kilodalton
- KF-1 – Zinc finger protein expressed in hippocampus
- LAR – Leukocyte common antigen related receptor
- LCFA – Long chain fatty acid.
- LDL – Low density lipoprotein.
- LINES - Long interspersed nuclear elements
- Lit1 – (KVQT1-AS/KCMQ1OT1) paternally expressed antisense RNA
- LoxP – Locus x of crossover P1 site.
- M46 - Ghrelin.
- MAGE – Melanoma associated antigen
- MAPK – Mitogen associated protein kinase
- MBD – Methyl CpG binding domain
- MBD1 – Methyl CpG binding domain protein 1
- MBD2 – Methyl CpG binding domain protein 2
- MBD3 – Methyl CpG binding domain protein 3
- MBD4 – Methyl CpG binding domain protein 4
- MBT – Mid blastula transition
- MeCP1 – Methyl CpG binding protein 1
- MeCP2 – Methyl CpG binding protein 2
- MEF – Murine embryonic fibroblasts.
- Mf3 – Forkhead 5
- Mi2/NURD – Nucleosome remodelling and histone deacetylation
- Min – Multiple intestinal neoplasia
- Mlh1 – MutLhomologue 1
- MM – Mismatch.
- MMR – Mismatch repair
- MPGF – Major proglucagon fragment.
- MRI – Magnetic resonance imaging.
- mRNA - Messenger RNA.
- Msh2 – MutS homologue 2
- Msh6 – MutS homologue 6.
- MSI – Microsatellite instability
- NADH – Nicotinamide adenine dehydrogenase
- NDN – Necdin homolgue
- NER – Nucleotide excision repair
- Nkx2-3 – NK2 class homeobox gene 3
- NLS – Nuclear localisation signal.
- NPY – Neuropeptide Y.

- NTD – N-terminal domain
- O/N - overnight
- ORF – Open reading frame.
- ORN – Olfactory neuron.
- OTB – Ovarian time bomb (theory)
- P14/p16 – Promoter 14/16
- P19 – Mouse embryonal carcinoma cell line
- p53 – Protein of molecular weight 53kDA
- PA1 – Human ovarian teratoma cell line
- PARP – Poly (ADP-ribose) polymerase
- PBS – Phosphate buffered saline
- PCNA – Proliferating cell nuclear antigen
- PCR – Polymerase chain reaction.
- PE – Primitive ectoderm
- PEG3 – Paternally expressed 3
- PGC – Primordial germ cell
- PI3K – phosphatidyinositol 3 kinase
- PM – Perfect match
- PMS2 – Postmeiotic segregation increased 2
- PR – Progesterone receptor.
- PRL – Prolactin.
- PRLR – Prolactin receptor.
- PYY – Peptide YY
- QRT-PCR – Quantitative (real-time) reverse transcription polymerase chain reaction.
- QT interval – (The time for electrical activation and inactivation of the ventricles)
- Rasgrf1 – Ras protein specific guanine nucleotide releasing factor 1
- Rb – Retinoblastoma
- RbAp46 – Retinoblastoma binding protein 46
- RbAp48 – Retinoblastoma binding protein 48
- Rgr – Retinal G-protein coupled receptor
- RNA – Ribonucleic acid.
- RNA Pol II – RNA polymerase II
- RT-PCR – Reverse transcription polymerase chain reaction
- RTT – Rett syndrome
- SAM – Statistical analysis of microarrays.
- Sap 18 – Sin3 associated protein 18
- Sap30 – Sin3 associated protein 30
- Sin3A – Sin3 homologue A, transcription regulator (yeast)
- Snrpn – Small nuclear ribonucleoparticle associated protein
- Sp1 – Specificity protein 1
- SRY – Sex determining region Y
- Ssc2/Elovl2 – Elongation of very long chain fatty acids like 2 (also called FEN1/Elo2, SUR4/Elo3)
- Stat5a – Signal transducer and activator of transcription 5a

- Suv39H1 – Suppressor of variegation 39 homologue 1.
- *Syd2* – synapse deficient 2
- TA – Transit amplifying.
- TAM3 – Methyl binding domain-R2
- TBS – Tris buffered saline
- TCA cycle – Tricarboxylic acid cycle
- TCF/LEF – T Cell factor/ lymphocyte enhancer factor
- TCF4 – Transcription factor 4
- TCR – T-cell receptor.
- TDG – Thymine glycosylase.
- TE – Trophoectoderm
- TEB – Terminal end bud.
- TNFα - Tumour necrosis factor α.
- TRD – Transcriptional repression domain.
- TrkB – Tyrosine kinase receptor B.
- TTX – Tetrodotoxin.
- UBE3A – Ubiquitin protein ligase 3A
- UBE3A/E6-AP – Ubiquitin protein ligase/ human papilloma virus E6 associated protein.
- UPD – Uniparental disomy
- UTR – Untranslated region.
- VIP – Vasoactive intestinal polypeptide.
- VLCFA – Very long chain fatty acid
- WDNM1 – Whey associated protein four disulphide core domain protein 1.
- Wnt – Wingless type MMTV integration site family member
- XCI – X-chromosome inactivation
- Xi – The inactive X
- Xic – X-inactivation centre
- Xist – X inactivation specific transcript
- Xm – Maternal X-chromosome
- Xp – Parental X chromosome
- Znf127 – Zinc finger protein 127.

LIST OF FIGURES.

Fig 1.1. The formation of 5-methylcytosine.....p.2 .

Fig 1.2. Imprinting, Insulators and enhancer blocking at the H19/Igf2 locus.....p.9.

Fig 1.3. X-chromosome inactivation..... p.14.

Fig 1.4. The structure of chromatin.....p.15.

Fig 1.5. Some histone modifications.....p.17.

Fig 1.6. Levels of methylation in the germ line and developing embryo.....p.24.

Fig 1.7. Structure of the methyl CpG binding domain.....p.27.

Fig 1.8a. Conditional allele of MeCP2.....p.52.

Fig 1.8b. Using Cre-Lox technology to conditionally delete MeCP2.....p.53.

Fig 1.9. Schematic representation of the process of synaptic elaboration and pruning in Rett syndrome and normal brain.....p.57.

Fig 1.10.Alternative splicing generates two isoforms of MECP2.....p.61.

Fig 1.11. MECP2 and transcriptional repression.....p.64.

Fig 1.12. Possible mechanism of action of MeCP2 at the *Bdnf* promoter IV in the mouse.....p.70.

Fig 3.1. Prokaryotic and eukaryotic repair.....p.109.

Fig 3.2. Apoptosis triggered through the fas receptor pathway.....p.114.

Fig 3.3. Anoikis in isolated murine small intestinal crypts.....p.118

Fig 3.4. Loss of MBD4 reduces anoikis in isolated small intestinal crypts.....p.120.

Fig 3.5. Apoptosis in the small intestine of MBD4-deficient mice in response to Fas Ligand..... p.122

Fig 3.6. Apoptosis in the large intestine of MBD4-deficient mice in response to Fas ligand.....p.123.

Fig 3.7. Apoptosis in the liver of Mbd4-deficient mice in response to Fas ligand.....p.124

Fig 3.8. Dose curve for intestinal apoptosis for 5μg, 8μg and 10μg Fas ligand.....p.127.

Fig 3.9. Loss of Mlh1 reduces the apoptotic response to Fas ligand.....p.128

Fig 3.10. Loss of Mlh1 increases the mitotic response to Fas ligand.....p.129

Fig 3.11. 8μg Fas ligand induces extensive liver apoptosis at 6 hours.....p.130

Fig 3.12. *Mlh1*$^{-/-}$ and *Mlh1*$^{+/+}$ mice show high variability in rates of apoptosis; small intestine.....p.132

Fig 3.13. *Mlh1*$^{-/-}$ *Mbd4*$^{-/-}$ mice show high variability in rates of apoptosis: large intestine. p.134

Fig 3.14. Putative mechanism of Mbd4/Fadd interaction. .....p.138.

Fig 3.15. Putative model for the interaction between Fadd, Mbd4 and Mlh1.....p.138.

Fig 4.1. Growth and development of the mouse mammary gland..... .p.140

Fig 4.2. Levels of *MECP2* and *MBD2* mRNA in the developing human mammary gland. p.143

Fig 4.3. Loss of MeCP2 does not alter the development of the mammary gland at 4 weeks.....p.145.

Fig 4.4. Loss of MeCP2 does not alter the mature adult appearance of the virgin mammary gland. p.146

Fig 4.5. Preliminary investigation of the involution of the *Mecp2*$^{flox/flox}$ *Blg-Cre* $^{+}$ mammary gland. p.147

Fig 4.6. 21-day involution; loss of MeCP2 does not grossly affect the involution process.....p.150.

Fig 4.7. Some abnormal areas are seen in *Mecp2*$^{flox/flox}$ *BLG-Cre* $^{+}$ mammary glands after 21 days involution. p.151

Fig 5.1. The structure of the small intestine; the crypt-villus unit. p.157

Fig 5.2. Removing MeCP2 in the murine intestine; LacZ reporter staining.....p.160

Fig 5.3. Quantitative real time RT-PCR analysis of *Mecp2* mRNA levels in *Mecp2*$^{flox/flox}$ *Blg-Cre* $^{+}$ mice after 4 injections of 80mg/kg β-napthoflavone.....p.161

Fig 5.4. Villi become 12.5% longer when MeCP2 is removed from the murine small intestine.....p.163

Fig 5.5. Distribution of villus sizes as quantified by cell number..... p.163

Fig 5.6. Crypt size remains constant in the absence of MeCP2.....p.165

Fig 5.7a. Proliferation in the crypt is increased in the absence of MeCP2 .....p.167

Fig 5.7b. The proliferative zone expands upwards in the absence of MeCP2.....p.167

Fig 5.8. Apoptosis is not altered in the absence of MeCP2.....p.168

Fig 5.9. Migration along the villus is faster in the absence of MeCP2.....p.170

Fig 5.10. Loss of MeCP2 does not alter the crypt-villus boundary.....p.172

Fig 5.11. Loss of MeCP2 does not alter the number or distribution of goblet cells... p.173

Fig 5.12. Loss of MeCP2 does not alter the number of enteroendocrine cells.....p.175

Fig 5.13. Enteroendocrine cells in the small intestine of *Mecp2*$^{flox/flox}$*Blg-Cre* $^{+}$ and *Mecp2*$^{flox/flox}$*Blg-Cre* $^{+}$ and *Mbd2*$^{-/-}$ mice.....p.176

Fig 5.14. No change in the number of K-cells in the absence of MeCP2.....p.178

Fig 5.15. Phenotype summary. p.179

Fig 5.16. Proglucagon-like hormones and Glp-2.....p.183

Fig 6.1. Using the Affymetrix GeneChip expression array to analyse the transcriptome of MeCP2 deficient mice.....p.192

Fig 6.2. Sample Affymetrix data from the MeCP array.....p.195.

Fig 6.3. SAM plot; normalised data analysed using SAM.....p.197

Fig 6.4. *Hprt* confirmation.....p.198

Fig 6.5. Real-time Quantitative RT-PCR.....p.205

Fig 6.6a. Sample data for quantitative Realtime RT-PCR: amplification curve for *Igfbp3*. p.208

Fig 6.6b. Melting curve for *Igfbp3*.....p.208

Fig 6.7. Example of QRT-PCR calculation for Gip.....p.210

Fg 6.8. QRT-PCR results.....p.211

Fig 6.9a. Mice carrying a polymorphism at the *Galgt2* locus have long villi.....p.213

Fig 6.9b. The *Galgt2* locus. p.213

Fig 6.10. Model for villus lengthening in the absence of MeCP2.....p.219

Fig 6.11. Model for action of MeCP2 deficiency in Rett syndrome..... p.222.

LIST OF TABLES.

Table 1. Phenotypic variation in Rett syndrome.....p.41

Table 2. Significantly upregulated genes (SAM.).....p.199

Table 3. Significantly downregulated genes (SAM.).....p.201

Table 4. Significantly downregulated genes (T-test.).....p.202

Table 5. Significiantly upregulated genes (T-test.).....p.203

Table 6. Targets suggested by analysis of microarray and phenotype.....p.204

Table 7. Significant changes in gene expression.....p.212

Chapter 1: Introduction.

1.1. DNA methylation and epigenetics.

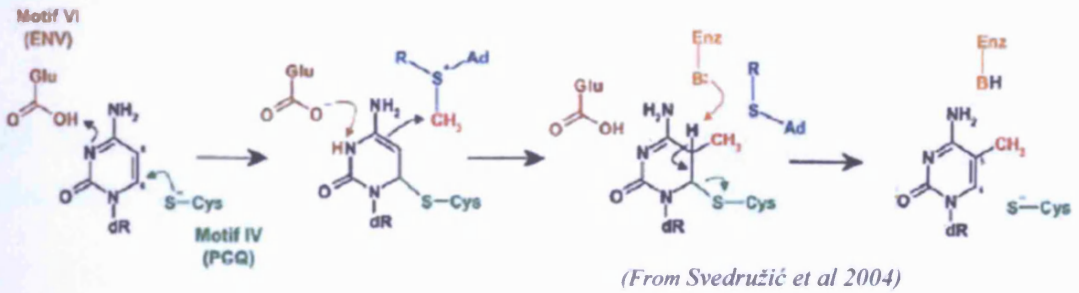
1.1.1 Genetic and Epigenetic information.

The function of DNA is to encode the information needed to enable an organism to grow, live and reproduce itself. Our knowledge of how this information is used in the cell, to create the complex regulatory networks which govern processes such as day-to-day functioning and the reproduction of the cell has undergone an extraordinary expansion in the half century since the structure of DNA was elucidated by Watson and Crick (Watson and Crick 1953.) However, the more we discover about the extraordinary complexity of the cell, the more we realise how much is still unknown. In particular, the understanding of the process of gene regulation has undergone a major paradigm shift over the last few decades. It was initially thought that all the information encoded by DNA was genetic, i.e. contained within the primary sequence of the DNA molecule. As our knowledge has expanded, however, it has become clear that there is much more information encoded within the structure of DNA, at levels which vary from alterations of the bases themselves, such as 5-methylcytosine (5mC) through alterations in the proteins which package the DNA, to the higher-order topology of the DNA molecule as it resides in the nucleus. These extra levels of information are described as **epigenetic**, in that they lie 'above and outside of' the primary structure. Epigenetic alterations allow stable, heritable changes in gene expression to persist through the life of the organism and across generations, with no detectable change in the primary structure of the DNA. Epigenetic effects are involved in processes as diverse as X-chromosome silencing, DNA repair and genomic imprinting and can be thought of as creating a number of extra, regulatory 'layers' within the genome (see fig 1.1.)

1.1.2. Methylation of DNA.

DNA can be methylated by the covalent addition of a methyl (CH_3) group at the 5 position of cytosines to form 5 methyl cytosine (5mC); in mammals the bulk of this methylation occurs at cytosines 5' to guanines, to form a symmetrical CpG dinucleotide (fig. 1.1.) and these methylated CpGs are the most common epigenetic modification of the vertebrate genome (Ballestar and Wolffe 2001) and are formed by the actions of a family of DNA methyltransferase (DNMT) enzymes, using S-adenosyl-L-methionine (AdoMet) as the methyl donor (Svedružic *et al* 2004.)

Fig 1.1. The formation of 5-methycytosine.



(From Svedružić et al 2004)

Fig 1.1. Formation of 5-mC

The positive charge on the sulphonium atom of S-adenosyl-L-methionine (AdoMet) makes it a powerful alkylating agent. The methyl group of AdoMet is bound to a sulphonium atom, which thermodynamically destabilises the molecule and makes the otherwise rather inert methylthiol group of the methionine moiety susceptible to nucleophilic attack by nitrogen, sulphur and oxygen atoms and carbanions. Nucleophilic attack by the Dnmt enzyme on the C6 of the target cytosine is carried out by the thiol group of the cysteine residue in a conserved PCQ motif (motif IV) within the active site. The formation of the covalent bond thus activates the C5 atom and leaves it susceptible to electrophilic attack. The DMNT enzyme is known to 'flip' the base out of its regular configuration in order to proceed with the reaction.The target base is rotated on its sugar phosphate bonds so that it projects into the concave catalytic site. This process requires no breaking of covalent bonds, although the H-bonds between bases and the stacking π interactions between adjacent bases are lost (Klimavsaukas et al 1994, Svedružić et al 2004.)

Although methylation can occur at the N6 position of adenosines and the N4 and C5 position of cytosines, only the latter modification is seen in mammals (Hendrich and Tweedie 2003.) These methylation marks are key mediators of many epigenetic processes; they can recruit proteins to the DNA to remodel chromatin and silence genes, they are involved in genomic imprinting, X-chromosome inactivation and genomic stability, and are implicated as having roles in DNA repair and apoptosis (Holiday and Ho 2002.) We are beginning to understand how methylation relates to normal and abnormal ageing (Liu *et al* 2003), and how errors and defects in the methylation patterns or the proteins which form or read them are involved in diseases such as cancer (Feinberg and Tycko 2004), as well as many congenital disorders such as Rett syndrome, ICF syndrome and perhaps even autism (Levin *et al* 2003.)

1.1.3. Patterns of methylation: CpGs and methylated CpGs are non-randomly distributed in the mammalian genome.

Methylated CpG dinucleotides are non-randomly distributed in the mammalian genome due to the hypermutability of the methylated CpG, which is prone to undergo deamination to form a T•A pair, leading to a depletion of CpGs from the genome (Jones and Takai 2001.) This leads to an under-representation of CpG in the genome (Jones and Takai 2001.) Methylated CpG dinucleotides are also non-randomly distributed: the vertebrate genome is globally methylated, i.e. methylated cytosine residues are found over the entire genome apart from short (around 1 kb) unmethylated regions where CpG is found at its expected ratio, called CpG islands (Bird 1987.) CpG islands are thought to account for around 1% of the genome (Hendrich and Tweedie 2003) and are frequently found in association with genes, most often in the promoters, first exons and regions towards the 3' end of a gene (Jones and Takai 2001.) CpG islands coincide with the promoters of an estimated 60% of RNA Pol II transcribed genes (Svedružić *et al* 2004,

Zuckercandl 2002.) The original definition of a CpG island, proposed by Gardinier-Garden and Frommer was of a region greater than 200bp region with a high GC content and an observed/expected CpG ratio of >0.6 (Gardinier-Garden and Frommer 1987) but this has now been made more stringent in terms of length and CpG content to exclude a number of small exonic regions and various parasitic DNAs (Jones and Takai 2001.) CpG islands are unmethylated in the germline and thus are not subjected to the mutagenic pressure put upon methylated cytosines (Jones and Takai 2001.) However, the reason why these particular sequences are not methylated, and the mechanisms which keep them unmethylated have not been fully elucidated. A physical barrier to methylation of CpG islands may be formed by the protein Sp1 (specificity protein 1) – when Sp1 binding sites are removed from DNA flanking a CpG island, the sequence becomes methylated (Macleod *et al* 1994, Brandeis *et al* 1994.) It appears, therefore, that CpG islands act to keep frequently transcribed promoters in a state conducive to transcription. Since so much of the genome is non-coding sequence, and due to the heavy burden of repetitive elements such as Alu, LINES (Long interspersed nuclear elements) and SINES (short interspersed nuclear elements) it is thought that the genome is kept in a default state of repression (Bird 1995.) This would prevent transcriptional 'noise' caused by random transcripts as well as chromosomal re-arrangements between repeated sequences (Carter and Segal 2001.)

Although repression of potentially deleterious rogue sequences is a function of DNA methylation, it really tells us little of how, when and why methylation evolved (Hendrich and Tweedie 2003.) Also, although CpG islands are generally unmethylated, there do exist exceptions; a number of CpG islands have been found to be methylated in a tissue-specific manner, namely the testis-specific H2b gene in the rat and the germline-specific *MAGE* (melanoma associated antigen) gene family in the human germline (Newell-Price *et al* 2000.)

1.1.4. Methylation of DNA in Other Species.

DNA methylation is found in a wide range of eukaryotes, including insects, mammals, birds, fish and amphibians; it is also found in some plants, such as Arabidopsis (Adams 1996.) Plant genomes contain relatively high levels of 5-methylcytosine (5mC) around 5-25% depending on the species examined (Hendrich and Tweedie 2003, Adams 1996.) Unlike mammals, where 5mC is mainly found in the context of CpG dinucleotides, the major 5mC content of the plant genome can be found in three different contexts; CpG, symmetrical CpNpG sites and asymmetric CpHpH sites (where N is any nucleotide and H is A,C or T) (Rangwala and Richards 2004.)

In plants, 5mC is concentrated in heterochromatic regions of DNA such as long repeated arrays (e.g. the rRNA genes) (Rangwala and Richards 2004.) Plant methylation patterns appear to be highly compartmentalised with expressed genes being generally unmethylated and embedded within a heavily methylated region; studies in maize have shown that exons are generally unmethylated whereas the majority of transposons are methylated (Rabinowicz *et al* 2003, Palmer *et al* 2003.) Methylation patterns in plants are laid down by Dnmt1 class methyltransferases (such as Met1 in Arabidopsis) at CpG sites, and plants also contain distinct enzyme activities that can methylate cytosine at other, non CpG nucleotide sequences. The chromodomain containing methyltransferase Cmt3 and the Dnmt3-related methyltransferases Drm1 and Drm2 are responsible for these activities and are directed by DNA methylation independent cues (Rangwala and Richards 2004, Cao *et al* 2002.)

Although it was known that *Drosophila melanogaster* carried genes with homology to the vertebrate methylation machinery, for a long time it was believed that insect species such as *Drosophila* had no methylated bases in their genomes (Lyko 2001.) The overall methylation levels in *D. melanogaster* are extremely low, and are thought to be less than 1% of cytosines (Lyko *et al*

2000, Gowher *et al* 2000.) Most 5-Me-C modifications are not based within CpG pairs and were thus not spotted by traditional assays (Field *et al* 2004.) Methylation in *D. melanogaster* is formed by a *DNMT2*-like enzyme, mainly at CpT and CpA dinucleotides (Kunert *et al* 2003.) It is not clear how this methylation is maintained in the absence of the symmetrical CpG dinucleotides which direct the mammalian machinery (Field *et al* 2004.)

1.1.5. Functions of DNA methylation

DNA methylation has been found to be connected to many diverse functions in the normal cell, including transcriptional repression, genomic imprinting, X-chromosome inactivation, and roles in the control of gene expression in fundamental embryonic development, normal tissue function and DNA repair (Urnov and Wolffe 2001, Zuckercandl 2002.) Disregulation of the chromatin environment and defects in the methylation machinery are now known to be the cause of a number of human pathologies such as ICF syndrome (Immunodeficiency, centromere instability and facial anomlies syndrome) (Xu *et al* 1999) and disregulation of epigenetic mechanisms have been found to play an important role in the formation of cancer (Verma *et al* 2004.) With so much still unknown about the origins and functions of DNA methylation, this is a rich and diverse field that undoubtedly still holds many surprises.

1.2. Genomic Imprinting.

Imprinted genes are those which are expressed differentially depending on their parent of origin, not due to any significant difference (beyond normal allelic variation) in primary sequence, but due to different sex-specific imprint 'marks' which generally lead to silencing of one allele and expression of the other (Verona *et al* 2003.) Genomic imprinting is therefore an epigenetic

mechanism of transcriptional regulation in which the expression of a small subset of genes is restricted to or predominantly expressed from one parental allele (Verona *et al* 2003.) In order to be differentially expressed, the alleles must be different in some way; these differences are not in the primary sequence of the DNA (beyond normal allelic variation) but are due to epigenetic differences. The two parental alleles of an imprinted gene often show different chromatin structure, DNA methylation, histone modifications and asynchronous replication (Bartolomei and Tighman 1997.) The exact ways in which the epigenetic differences result in differential expression are well-characterised for some genes (e.g. the *Igf2/H19* system) (Arney 2003.)

Imprinting was first identified in mammals in the early 1980s, when mouse embryos were created carrying chromosomes derived solely from the male or female parent; although they were diploid, these embryos failed to develop (McGrath and Solter 1984.) This ran counter to the then - accepted dogma that the function of genetic material was independent of the parent it was derived from, and implied that alleles could be differentially expressed depending on their parent of origin (Verona *et al* 2003.) Further experiments in the early 1990s identified the first paternally-expressed gene (*Igf2*) (DeChiara *et al* 1991) and the first maternally expressed genes (*Igf2r*) (Wilson 1991.)

Since then, many more imprinted genes have been discovered (www.mgu.har.mrc.ac.uk/research/imprinting) and the often complex mechanisms which underlie epigenetic control mechanism have been partially unravelled. For example; imprinted genes are unusually enriched in CpG islands, with 88% of mouse imprinted genes having a CpG island compared with the average figure of 47% (Reik and Walter 2001.) This implies that methylation may play a role in the regulation of imprinted regions (fig 1.2.) That methylation is important for imprinting has been proven (Li *et al* 1993.)

Imprinted genes also often cluster on the chromosome, to form large, imprinted domains (Verona *et al* 2003, www.mgu.har.mrc.ac.uk/research/imprinting.) This implies that a common mechanism or group of mechanisms may control whole imprinting domains and that there may be *cis*-acting imprinting control regions that function across the domain to regulated imprinting for all the genes in that cluster (Verona *et al* 2003.)

However, many intriguing questions remain: how and in response to what pressures did imprinting evolve? What advantages does it confer? Why are only a small fraction of our genes imprinted? Perhaps the most currently socially relevant effect of imprinting relates to the emergent technologies of reproductive and therapeutic cloning. Many mammalian embryos are extremely difficult to clone, possibly in part due to the complexity of the epigenetic 'reprogramming' events which occur around fertilisation, development and germ cell formation (Verona *et al* 2003.) An increased understanding of the role of epigenetics in normal development would have massive implications for the efficiency of reproductive and therapeutic technologies, and therefore human health.

The study of imprinted genes such as *Igf2* has also allowed us to understand epigenetic mechanisms such as chromatin boundary elements, enhancer blockers and insulators (Engel *et al* 2004, Thorvaldson and Bartolomei 2000.) Imprinted expression at the *H19/Igf2* locus, for example, depends on a 2 kb, differentially-methylated domain (DMD) that acts as a maternal-specific insulator which is maternally-expressed and is paternally hypermethylated (Engel *et al* 2004, Thorvaldson and Bartolomei 2000.) (Fig 1.2.)

Fig 1.2. Imprinting, insulators and enhancer blocking at the H19/Igf2 locus.

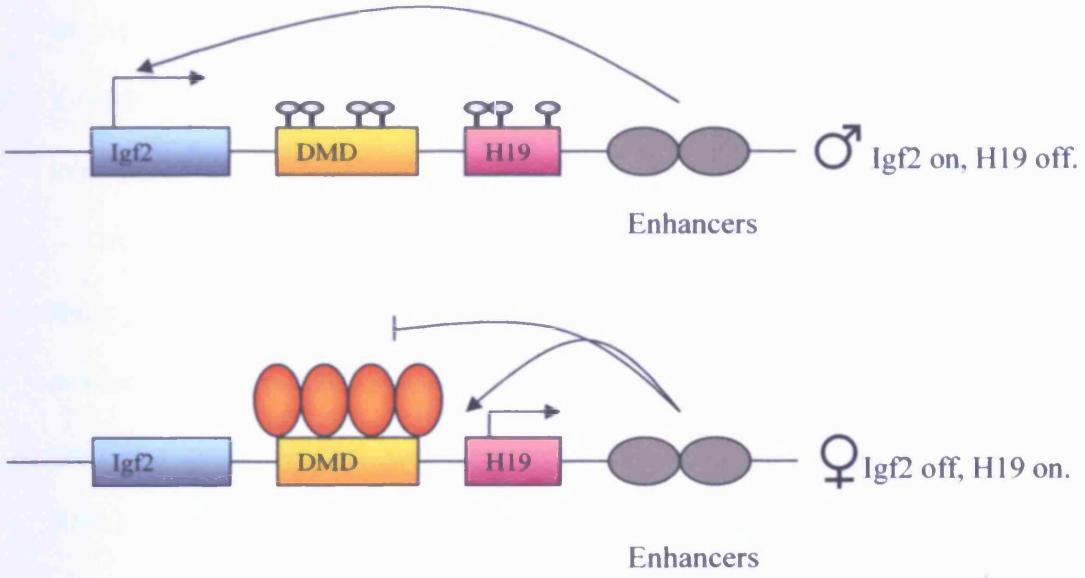


Fig 1.2. Model for Igf2/H19 regulation. (from Bell et al 2001)

Igf2 and H19 are located on mouse chromosome 7. Both genes require a set of enhancers located downstream of H19 (grey ovals) for high-level expression. The maternal allele of H19 is unmethylated and recruits CTCF, functioning as an insulator and preventing Igf2 accessing the downstream enhancer. H19 gains enhancer access and is expressed. On the paternal chromosome, the H19 DMD is methylated, and the CTCF cannot bind, so there is no insulator activity. The downstream enhancers are now free to activate expression of paternal Igf2. Inactivation of the paternal H19 is accomplished by a separate mechanism (Bell et al 2000, Drewell et al 2000.)

1.2.1. Evolution of Imprinting in Mammals.

An imprinted gene is expressed only from one parental allele, which leads to it being functionally haploid (Murphy and Jirtle 2003.) How such a mechanism could have evolved is a

puzzle, as diploidy prevents the deleterious effects of recessive alleles in the heterozygous state and must therefore be a selectable advantage (Murphy and Jirtle 2003.) Functional haploidy 'unmasks' these effects, which would be expected to lead to a drop in fitness for the individual organism (Murphy and Jirtle 2003.) Clearly then, the benefits gained by having a certain locus imprinted must be greater than the risks of haploidy at that locus. This may also explain why such a small proportion of our genes are imprinted - the number of loci at which haploidy can be tolerated in a diploid organism is probably very low (Murphy and Jirtle 2003.)

A number of theories have been advanced to explain how imprinting may have arisen; the three leading theories are the conflict (also called the kinship) theory (Wilkins and Haig 2003), the ovarian time bomb (OTB) theory (Varmuza and Mann 1994) and the evolvability theory (Beaudet and Jiang 2002.) Of the three, the most accepted is the conflict theory (Wrzeska and Rejduch 2004.)

The evolvability theory, proposed by McGaudet and Martin, and Beaudet and Jiang, (Beaudet and Jiang 2002, Wilkins and Haig 2003) states that imprinting has evolved because functional haploidy at imprinted loci increases 'evolvability' i.e. the ability of a genetic system to generate novel adaptations to selection pressures (Wilkins and Haig 2003.) In other words, an imprinted allele is 'hidden' from the pressure of natural selection pressure and can generate new mutations which may be beneficial if that allele is then not in a silenced state (Wilkins and Haig 2003.) The theory does have shortcomings; the time an allele spends silenced may not be enough to accumulate new mutations, and the theory does not explain why specific, small numbers of loci should be imprinted or in which sex, or in which phyla, since in the evolvability model, the benefits are identical for all diploid organisms, whereas imprinting is only seen in a subset of mammals (Wilkins and Haig 2003.)

The ovarian time bomb (OTB) theory put forward by Varmuza and Mann proposes that imprinting has evolved to prevent female mammals from ovarian trophoblastic disease (Varmuza and Mann 1994.) If an unfertilised oocyte spontaneously develops, it may form a teratoma, (a tumour consisting of multiple tissue types) but because the paternal genes are needed for full development of an invasive trophoblast, the ovarian teratoma is relatively benign (Varmuza and Mann 1994.) However, the theory does not explain how genes which are not involved in trophoblast development come to be imprinted, nor does it explain why imprinting (of *Igf2*) is present in marsupials, which lack a fully invasive placenta, or in eutherians with non-invasive placentas such as sheep (Wilkins and Haig 2003.)

The leading theory for the evolution of genomic imprinting is the 'conflict theory', sometimes also called the 'kinship theory' (Moore and Haig 1991.) This states that differential imprinting of genes depending on their parent of origin depends on the differing strategies each gene needs to maximise its likelihood of being carried to the next generation (Moore and Haig 1991.) These strategies are carried out within the context of maternal provisioning of resources for the foetus. It is generally more advantageous for the foetus to take maximal nourishment from the mother, but this may weaken the mother and reduce her own chances of survival and future reproductive rate.

Maternally and paternally derived genes in an offspring may be in conflict thus; a mother is equally related to all her offspring, with each maternal gene having a 50% probability on average of being in any of her other offspring (Moore and Haig 1991.) In a species with any degree of multiple paternity, the father cannot guarantee that the mother is not unfaithful and thus the chances of a paternally derived gene being represented in any other offspring from the same mother are less than 50%. Hence it is advantageous, the conflict theory states, for paternally-derived genes to drive foetal growth in order to obtain maximal resource provisioning from a (potentially unfaithful) mother (Moore and Haig 1991.) A large foetus however, is not optimal for

the mother, as it is both a drain on her resources and a potential danger to give birth to. Therefore, the conflict theory states that maternally-derived genes should serve to check foetal growth (Moore and Haig 1991.) One imprinted system which supports this theory is the paternally-expressed *Igf2*, which encourages foetal growth and the maternally expressed *H19* transcript, which is reciprocally expressed (Wrzeska and Rejduch 2004, Murphy and Jirtle 2003.) Imprinting defects in chromosome 11 around the *IGF2* locus in humans are associated with Beckwith-Wiedemann syndrome, a human foetal overgrowth syndrome (Reik and Mayer 1997, Nicholls *et al* 1998.) There is also some data which does not support the conflict theory of imprinting – one study of UPDs (uniparental disomies) in humans found that some paternal UPDs were in fact growth retarding (Miozzo and Simoni 2002.) However, the duplication of an entire chromosome is an extremely blunt way in which to extrapolate data concerning individual genes or gene clusters, and so is by no means a strong refutation of the conflict theory. Other, gene specific studies have produced results which support the conflict theory; Peg3 (paternally expressed 3) is a paternally-expressed imprinted gene and so would be expected to be a growth promoter and indeed the Peg3 knockout mouse is smaller than normal wild type mice (Li *et al* 1999.)

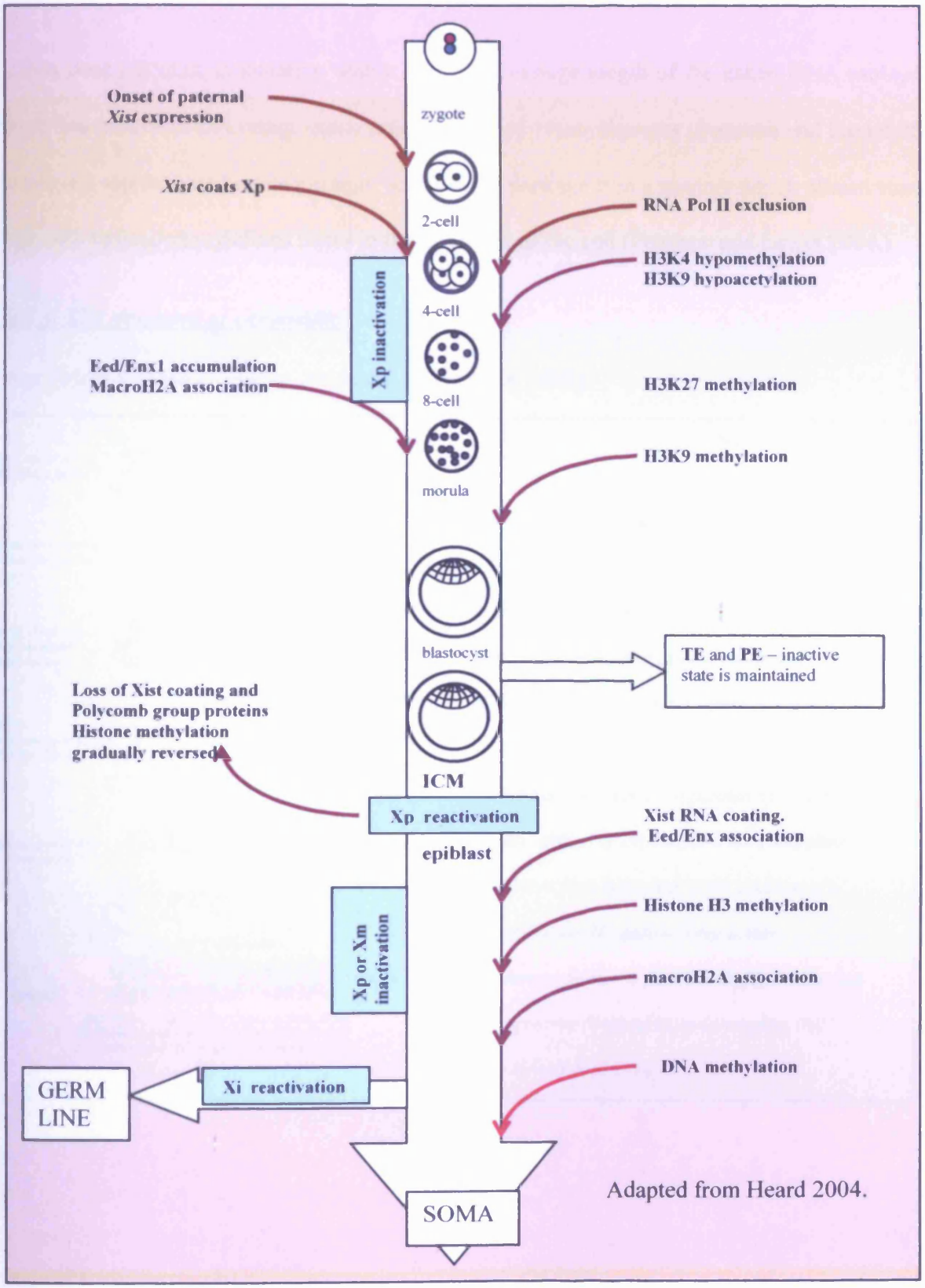
Imprinted genes have been found in marsupial (metatherian) and placental (eutherian) mammals, but as yet, no imprinted genes have been found in monotremes (prototherians) implying that imprinting arose after the divergence of the common ancestor of the marsupials and placental mammals from the ancestor of the monotremes (Murphy and Jirtle 2003, John and Surani 2000.) Since many imprinted genes are involved in foetal growth and development, this does back the conflict theory as it would allow the mother to exercise control over foetal resource provisioning.

1.2.2. X-Chromosome Inactivation.

In mammals, males have a single X chromosome and a single Y chromosome and females have two X chromosomes (Heard 2004.) The X-chromosome carries around 1500 genes and if both Xs were expressed in the female, there would be significant problems with gene dosage effects – female mammals have resolved this problem by inactivating one X chromosome (Lyon 1961.) The silencing occurs early in development and thus female mammals are mosaic for their X-linked gene expression (Heard 2004.) This has obvious advantages in that random inactivation leading to random mosaicism will to a certain extent mask the effects of being haploid for either X-chromosome (Heard 2004.) Males have no such advantage and so suffer more frequent and severe X-linked defects, such as colour blindness (Deeb 2005.) Not all X-linked genes are subject to inactivation, with an estimated 15% escaping X-inactivation (Carrel and Willard 2005.)

The initiation of X silencing is dependent on a master control locus, the X-inactivation centre (Xic) (Heard 2004.) At the heart of the Xic is the Xist (X-inactivation specific transcript) gene, which does not produce protein, but instead makes a 17kb long untranslated RNA that coats the entire X in *cis* and triggers its inactivation (see fig 1.3) (Heard 2004.) Initially this phase is reversible, but it becomes 'locked in' through a cascade of protein interactions and DNA methylation (Heard 2004.) The Polycomb group proteins Eed (embryonic ectoderm expression) and Enx1 (The mouse homologue of enhancer of zeste) seem to be implicated in the process and their association is accompanied by H3K27 methylation (Plath *et al* 2003.) H3K9 di-methylation is also observed (Heard *et al* 2001.) These histone modifications are followed by DNA methylation, which may keep the X-chromosome in its final, inactive state (see fig 1.3) (Silva *et al* 2003, Plath *et al* 2003.)

Fig 1.3. X-Chromosome Inactivation.



1.3. The Chromatin Environment and the Histone Code.

DNA does not exist in isolation within the cell; the huge length of the naked DNA molecule - almost two metres in an average eukaryotic nucleus of 10μm diameter (Peterson and Laniel 2004) - requires a sophisticated nucleoprotein 'scaffold' to package it in a manner which allows storage, usage and replication at defined times in the life cycle of the cell (Peterson and Laniel 2004.)

Fig 1.4. The structure of chromatin

(from Felsenfeld and Groudine 2003 and Alberts et al 1998.)

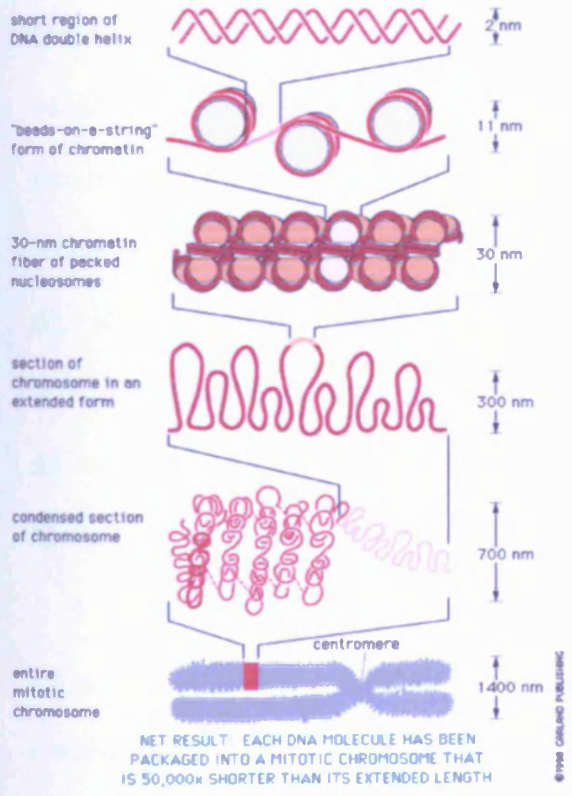


Fig 1. 4. The structure of chromatin.

The basic structural unit of chromatin is the nucleosome, which consists of 147bp of DNA wrapped almost twice around an octomer of histone proteins (A central heterotetramer of H3 and H4, and two heterodimers of H2A and H2B.) Nucleosomes are separated by a 10-60bp 'linker' region of DNA to form the 10nm fibre. The 30nm fibre is formed by the addition of a linker histone H1 and the fibre is then condensed further to form the highly compacted chromosome (Felsenfeld and Groudine 2003 and Alberts et al 1998.)

This scaffold is far from being an inert support for DNA; as our knowledge of the structure of chromatin expands, it is clear that the components of the chromatin environment play a key part in the modulation of gene expression (Peterson and Laniel 2004.)

1.3.1. The histone code.

Each histone has an amino terminal 'tail' consisting of 20-35 residues and histone H2A also has a ~37 residue carboxy terminal tail (Peterson and Laniel 2004.) The histones can undergo a range of post translational modifications, mainly on the tails, although new modifications in the central histone domains are being discovered; given the massive number of potential modifications, some believe that virtually all accessible histone residues may be modified (Peterson and Laniel 2004.) Modification of histone tails changes the way they interact with each other to form higher-order nucleosomal arrays, and so modification of histones is a key part of the chromatin remodelling process (Peterson and Laniel 2004.) Modifications include acetylation of lysines (K) and arginines (R,) phosphorylation of serines (S) and threonines (T,) the ubiquitination and sumoylation of lysines and the ribosylation of lysines (Peterson and Laniel 2004.) Each lysine residue can be mono-, di- or tri-methylated and arginine residues can be mono- or di-methylated (Peterson and Laniel 2004.)

In recent years, the patterns of histone modifications have begun to be correlated with specific chromatin and transcriptional states, which has lead to the concept of the so-called 'histone code' (Jenuwein and Allis 2001.) For example, transcriptionally active chromatin is often associated with H4 K8 acetylation, H3 S10 phosphorylation and H3 K14 acetylation (Peterson and Laniel 2004.) Transcriptional repression in higher eukaryotes, on the other hand, is frequently associated with H3 K9 tri-methylation and the loss of H3 and H4 acetylation (Peterson and Laniel 2004.) The

histone modification patterns also seem to change with global changes in chromatin structure; for example, H2A S1 and H2A T119 phosphorylation and H3 T3, H3 S10 and H3 S28 phospohorylation are associated with highly condensed mitotic chromatin (Peterson and Laniel 2004.) Some histone modifications are shown in fig 1.5.

Fig 1.5. Some histone modifications

(Adapted from Peterson and Laniel 2004.)

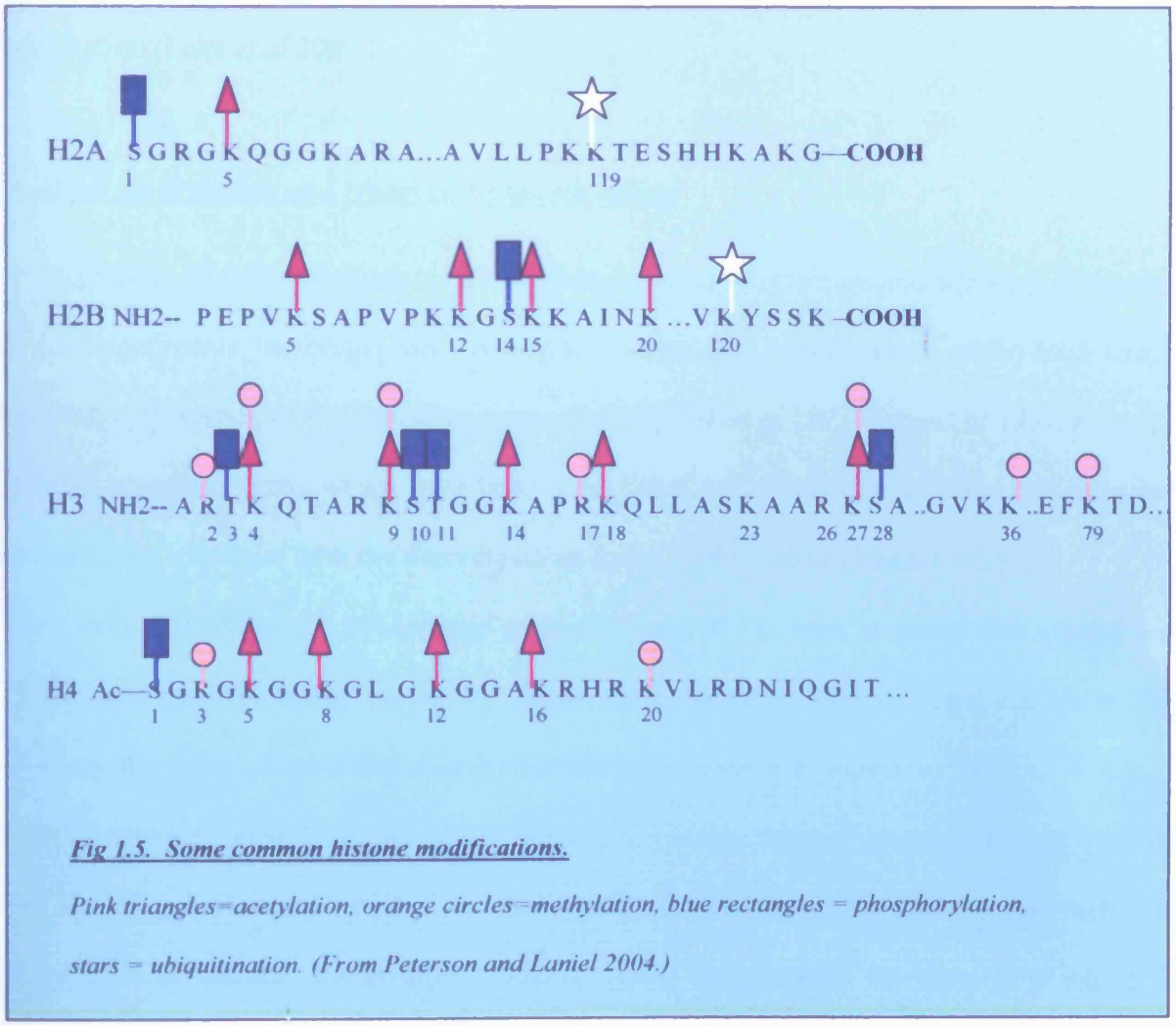


Fig 1.5. Some common histone modifications.

Pink triangles=acetylation, orange circles=methylation, blue rectangles = phosphorylation, stars = ubiquitination. (From Peterson and Laniel 2004.)

The code itself appears to be highly flexible rather than an absolute correlation of state-to-modification (Peterson and Laniel 2004.) Since MeCP2, Mbd2 and some Dnmts recruit chromatin remodelling activities such as Hdac1 and Hdac2, there are clear mechanistic links between histone modifications and the methylation machinery of the cell. For example, MeCP2 has been shown to be linked to histone methylation (Fuks *et al* 2002); MeCP2 associates with histone methyltransferase activity *in vivo* to methylate H3 K9 (Fuks *et al* 2002.) This clearly shows that DNA methylation can recruit factors which remodel chromatin into a form refractory to transcription (Fuks *et al* 2002.)

1.4. DNA methylation and transcriptional silencing.

A large body of evidence connects DNA methylation and transcriptional silencing (Curradi *et al* 2002.) Experiments performed with Xenopus oocytes and microinjected methylated templates show that methylation of DNA inhibits transcription (Kass *et al* 1993, Kass *et al* 1997, Keshet *et al* 1985.) Conversely, genes which have been methylated and silenced in cultured cell lines can be reactivated by treatment with the demethylating agent 5-azacytidine (Jones 1985.)

How does methylation of DNA cause gene silencing? It has been proposed that methylation of DNA causes steric blocking, preventing access of transcription factors (Eden and Cedar 1994.) However, the bulk of methylation-induced silencing does not appear to be this simple. The currently accepted model for the majority of methylation-induced silencing is that methylated DNA recruits proteins (the methyl CpG binding domain proteins); these proteins then recruit other factors such as histone deacetylases (Hdacs) which can remodel the chromatin into a form refractory to transcription (Jones *et al* 1998, Nan *et al* 1998) Transcriptionally active euchromatin is less condensed than the densely-packed and transcriptionally-inactive heterochromatin

(Jenuwein and Allis 2001.) In accordance with this model, the histones assembled on methylated DNA are less acetylated than those on an unmethylated template and methylated transfected genes can be reactivated by treatment with Trichostatin A (Curradi *et al* 2002, Hseih *et al* 1994.) A distinctive chromatin structure, which is insensitive to nuclease digestion (a mark of repressive chromatin) is formed on methylated DNA (Keshet *et al* 1996, Curradi *et al* 2002) Methylation and histone remodelling are thought to act together: naturally hypermethylated genes can only be reactivated by Trichostatin A after they have been demethylated with 5-azacytidine (Cameron *et al* 1999.)

1.4.1. Reading and writing the methylation marks: methyl CpG binding domain proteins and DNA methyltransferases.

The pattern of methylation within the genome carries an extra 'layer' of information over and above the primary nucleotide sequence. The marks themselves are laid down by a family of DNA methyltransferase proteins (Dnmts) and interpreted by a family of proteins capable of recognising methylated DNA; the Methyl-CpG-Binding Domain (Mbd) proteins.

1.4.2. DNA methyltransferases.

The DNA methyltransferases (Dnmts) are responsible for introducing methylation into the DNA. There are several members of the Dnmt family in mammals, each with its own specificity and function. All the Dnmts contain ten highly conserved motifs which make up a 500 amino acid catalytic domain catalytic domain (Cheng *et al* 1995.)

1.4.3. Dnmt1.

Dnmt1 was the first mammalian Dnmt to be identified, and is responsible for so-called maintenance methylation, i.e replenishing the symmetrical methylation marks on DNA which has been replicated and is thus hemi-methylated (Bestor 1992.) Dnmt1 has a higher affinity for hemimethylated DNA *in vivo*, up to 40-fold higher than for unmethylated DNA (Fatemi *et al* 2001) and is thought to be targeted to replication foci by Pcna (Chuang *et al* 1997) where it can methylate DNA immediately after replication. Dnmt1 is a large protein of 184kDa, which has a large multi-functional N-terminal domain which contains a nuclear localisation sequence, a replication foci targeting domain and a cysteine-rich Zn^{2+} binding domain (Hermann *et al* 2004.) Dnmt1 appears to be essential for development; the *Dnmt1*$^{-/-}$ mouse exhibits embryonic lethality; the embryos are stunted and have delayed development and reduced methylation levels (Li *et al* 1992.) Interestingly, *Dnmt1* null ES cells are viable, and contain low levels of methylation and a methyltransferase activity, which led to speculation that there may be another Dnmt activity in ES cells (Lei *et al* 1996.)

Dnmt1 does not function in isolation and many proteins have been reported to bind Dnmt1 through its N-terminal domain (Hermann *et al* 2004.) Dnmt1 is implicated in cell cycle regulation through interactions with Pcna, and Rb, as well as several other proteins (Hermann *et al* 2004.) Dnmt1 also interacts directly with histone modifying activities such as the histone methyltransferase Suv39H1 (Fuks *et al* 2003) the histone deacetylases Hdac1 and Hdac2 (Rountree *et al* 2000, Robertson *et al* 2000) and also with methyl CpG binding proteins Mbd2, Mbd3 and MeCP2 (Tatematsu *et al* 2000, Kimura *et al* 2003) and the *de novo* methyltransferases Dnmt3a and Dnmt3b (Kim *et al* 2002.) These diverse associations show that Dnmt1 is not only

maintaining the methylation marks, but is actively involved in methylation-dependent silencing at different levels, and provides a possible feedback link between the proteins which create and maintain the methylation and those which read and act upon it.

1.4.4. Dnmt2

The *Dnmt2* gene is conserved among eukaryotes, even in those which show no detectable methylation (such as *S. pombe*) (Hermann *et al* 2004.) The enzyme does not have the large N-terminal regulatory domain and is therefore more like bacterial methyltranseferases (Hermann *et al* 2004.) Dnmt2 was thought for a long time to have no catalytic activity, but has recently been shown to have a weak methyltransferase activity *in vivo* and *in vitro* (Hermann *et al* 2003.) Dnmt2 is ubiquitously expressed at low levels in most human and mouse tissue and in mouse ES cells (Okano *et al* 1998) and ES cells null for *Dnmt2* are viable and show normal levels of methylation at endogenous sequences (Okano *et al* 1998.) The biological function of Dnmt2 in mammals is not yet known, although a Dnmt2-like protein has been shown to mediate DNA methylation in *Drosophila* (Kunert *et al* 2003.)

1.4.5. Dnmt3a, Dnmt3b and Dnmt3l

Dnmt3a and Dnmt3b are *de novo* methyltransferases, i.e. they methylate previously unmethylated templates, and are therefore involved in laying down the methylation patterns in the developing embryo (Kaneda *et al* 2004.) *Dnmt3a* knockout mice were born appeared to be normal at birth. However, most homozygous mutant mice became runted and died at about 4 weeks of age. In contrast, no viable *Dnmt3b*$^{-/-}$ mice were recovered at birth (Okano *et al* 1999.) A recent conditional deletion of *Dnmt3a* has shown that it may have roles in genomic imprinting, as males carrying the

conditional allele show defective spermatogenesis and lack of methylation at some paternally imprinted loci (Kaneda *et al* 2004.) The two genes are highly similar (Hermann *et al* 2004) and each can partially compensate for the loss the other, as shown by a double mouse knockout (*Dnmt3a*$^{-/-}$ *Dnmt3b*$^{-/-}$) which has a more severe phenotype than the individual deletions (Okano *et al* 1999.) The fact that patients with the disorder ICF (immunodeficiency, centromeric instability, facial abnormalities) (who have mutations in *DNMT3b*) are affected despite having normal *DNMT3a* levels, however, shows that the two proteins cannot fully compensate for the loss of the other, as do the individual mouse knockouts (Hermann *et al* 2004.) Dnmt3a and Dnmt3b also exist as multiple, differently spliced isoforms, allowing a greater diversity of function (Hermann *et al* 2004.)

A third member of the Dnmt3 family, Dnmt3L, has been shown to interact with Dnmt3a and Dnmt3b and although it has no catalytic activity itself, it has been shown to act as a stimulatory factor in methylation (Chedin *et al* 2002) and also as a transcription repressor that recruits Hdac activity (Deplus *et al* 2002.) Dnmt3a and Dnmt3L have also been shown to be essential for the establishment of imprinting in the oocyte (Hata *et al* 2002.)

Despite the recent advances in this field, it is still not fully understood how the complex patterns of methylation are established during development (Herman *et al* 2004.) It is not known how some regions escape the global methylation which occurs in embryogenesis or how chromosomes are marked to 'remember' which regions to methylate or leave unmethylated. The answers to these questions are important, not just for basic research, but for their implications in human disease, development, and in processes such as cloning, in which defective imprinting is a major hurdle.

1.4.6. Establishment and Maintenance of Methylation Marks.

The patterns and overall level of methylation within the genome varies over the life cycle of an organism, with two 'waves' of demethylation and epigenetic reprogramming occurring, one as the germ cells develop, and one when the egg is fertilised and begins to develop into an embryo (Reik *et al* 2001.) In the developing mouse embryo, between 10.5 and 12.5 days post coitum, the primordial germ cells (PGCs) which will become the progenitors of future gametes, migrate, and enter the germinal ridge, which will form gonadal tissue (Reik *et al* 2001.) During this time their sex-specific methylation marks are erased (Reik *et al* 2001.) During subsequent gametogenesis, methylation levels rise as sex-specific marks are re-established, until methylation levels are high in fully mature gametes (Reik *et al* 2001.) Initiation of *de novo* methylation in sperm occurs around 14.5 days post coitum and continues postnatally in diploid gametocytes before they undergo meiosis (Reik *et al* 2001.) In females, *de novo* methylation occurs in oocytes which are arrested in the diplotene stage of meiosis (Reik *et al* 2001.) (See fig 1.5.)

The second wave of reprogramming is initiated at fertilisation; both egg and sperm carry sex-specific methylation marks and these must be 'reset' to allow for correct development of the embryo (Mayer *et al* 2000.) There is a marked asymmetry between the way maternal and paternal pronuclei undergo demethylation; the paternal genome undergoes a rapid 'active' demethylation within 4 hours of fertilisation, in which methylation is stripped off in an active process by an as yet unidentified demethylase, in the absence of transcription or translation (Reik *et al* 2001.)

By contrast, the maternal genome appears to undergo a slower, 'passive' demethylation. The methylation marks are effectively diluted out as there is no primary *de novo* methylase (Dnmt1) at this stage, so when strands are replicated, no *de novo* methylation of the hemimethylated daughter product occurs (Reik *et al* 2001.)

Fig1.6. Levels of Methylation in the Germ line and Developing Embryo

(Adpated from Reik et al 2001.)

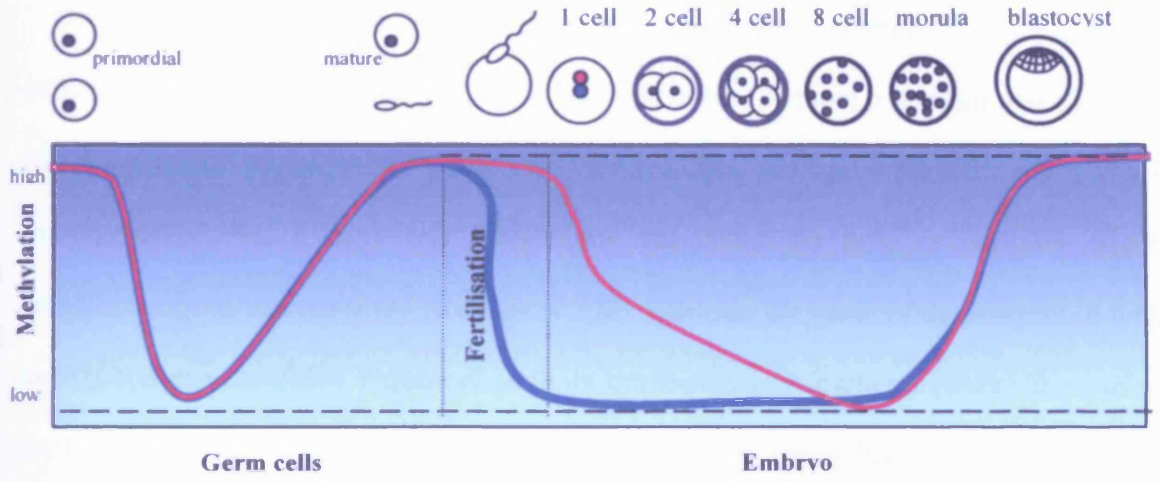


Fig 1.6. Levels of methylation in the maternal and paternally-derived genomes through the life cycle of the mouse.

Maternal levels are represented by a pink line, paternal levels by a blue line. The dashed lines indicate imprinted genes (top) which are not demethylated at fertilisation, and unmethylated sequences, which are not methylated (bottom.) Note the two waves of demethylation and re-methylation, effectively 'resetting' the imprint marks as they pass through the germ line, and the asymmetry between the rates of demethylation for the maternal and paternally-derived genomes. (Reik et al 2001.)

Around the fifth cell division, *de novo* methylation is initiated (Reik *et al* 2001.) At this point, the first two cell lineages emerge; the inner cell mass (ICM) and the trophoectoderm (TE); the ICM becomes highly methylated and gives rise to adult tissues, whilst the TE becomes

hypomethylated and gives rise to the extraembryonic tissues (Reik *et al* 2001.) The PGCs arise from the ICM and migrate to the germinal ridge and the cycle begins anew (Reik *et al* 2001.)

1.4.7. How are the sex-specific methylation marks reset correctly?

Despite the almost global demethylation, each cycle sees sex-specific methylation marks reset on the correct parts of the genome; how does this occur? Is there a residual mark left on the DNA which directs the methylation machinery? This question, and a possible mechanism of action, is far from resolved, but it does seem that some sequences are fully or partially protected from demethylation at various times (Reik *et al* 2001.) During the wave of demethylation which occurs in PGCs, some repetitive sequences such as intracisternal-A particles (IAPs) seem to retain their methylation (Reik *et al* 2001.) However, at this stage, using bisulfite genomic sequencing to examine methylation patterns, Hajkova *et al* have found that some imprinted genes, such as Peg3, Lit1, Snrpn and H19 are demethylated between 10.5 and 13.5 days gestation in the mouse (Hajokova *et al* 2002.) How imprinted genes are identified and 're-imprinted' remains unknown.

During the pre-implantation wave of demethylation, imprinted genes appear to retain their gamete-derived methylation patterns, presumably to allow the asymmetric monoallelic expression patters necessary for embryonic development (Reik *et al* 2001.)

Recent work by Bourc'his and Bestor has shown that Dnmt3L (DNA methyl transferase 3-like protein) may aid in the *de novo* methylation of dispersed repeated sequences in the male germ cells (Bourc'his and Bestor 2004.) Dnmt3L does not have the catalytic motifs of the other Dnmt family members, but it is related structurally to the other Dnmts (Hata *et al* 2002.) Male mice which lack Dnmt3L are viable but produce no sperm and are thus sterile (Hata *et al* 2002.) Female mice are fertile, but their heterozygous offspring die before mid-gestation due to the biallelic expression of genes which are normally silenced on the maternal allele (Hata *et al* 2002, Bourc'his and Bestor

2004.) In early germ cells, the loss of Dnmt3L caused meiotic failure, due to non-homologous synapsis (Bourc'his and Bestor 2004, Hata *et al* 2002.)

1.5. Reading the methylation marks: methyl binding domain proteins.

The methyl CpG binding domain (MBD) proteins 'read' and interpret the methylation marks on DNA, and thus are critical mediators of many epigenetic processes. At the present time, the MBD family consists of Mbd1, Mbd2, Mbd3, Mbd4 and MeCP2. Another methyl binding protein, Kaiso, has also been characterised (Daniel and Reynolds 1999, Prokhortchouk *et al* 2001.) Although it binds CpGs in a different manner, it acts as a transcriptional repressor and so can be considered a member of the MBD family (Prokhortchouk *et al* 2001.)

Sequence alignments of MBD family members reveal an intron in a conserved position in all members of the family (with the exception of the recently discovered Kaiso), indicating that they are evolutionarily related, and defined two subgroups; the MBDs of MeCP2 and Mbd4 being similar to each other, while the MBDs of Mbd1, 2 and 3 are more similar to each other than those of MeCP2 and Mbd4 (Ballestar and Wolffe 2001, Hendrich and Tweedie 2003.) The structure of the methyl binding domain itself has been solved by NMR (Okhi *et al* 1999, Wakefield *et al* 1999.) The domain contains around 70 residues and forms a wedge shaped α/β structure, in which four antiparallel β strands form one face of the wedge (see fig 1.7) (Ohki *et al* 1999.) The longest two strands are thought to interact with the major groove, where the methyl group would be located (Wakefield *et al* 1999, Ohki *et al* 1999.)

Fig 1.7. Structure of the methyl CpG binding domain as it binds to methylated DNA.

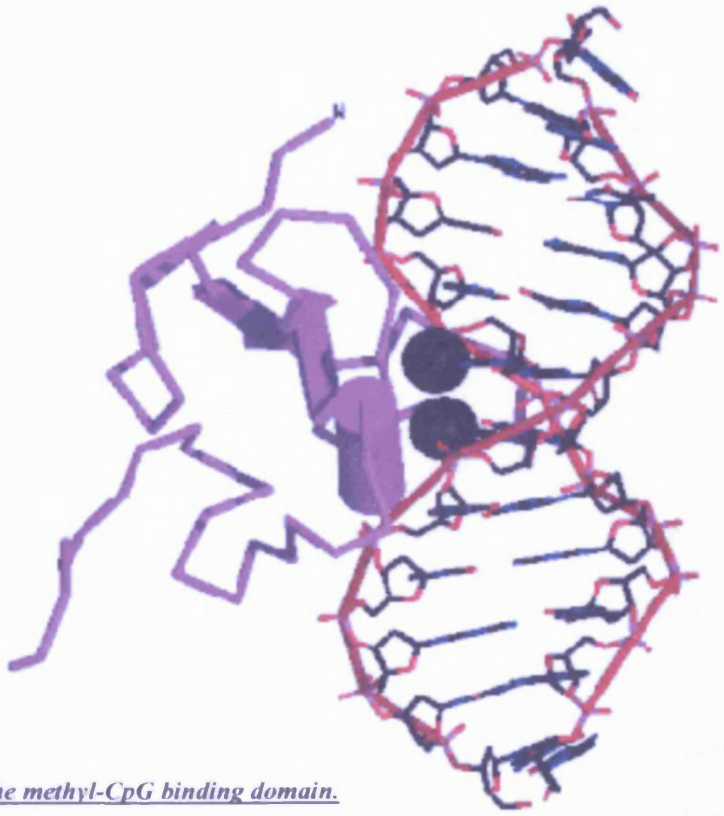


Fig 1.7. The methyl-CpG binding domain.

The domain contains around 70 residues and forms a wedge-shaped α/β structure. In which four antiparallel β strands form one side of the wedge. The longest two strands are thought to interact with the major groove, where the methyl group would be located. (Wakefield et al 1999.)

Mbd1, Mbd2 and MeCP2 have been demonstrated to repress transcription from methylated promoters *in vitro* and *in vivo*, via the association of their transcription repression domain (TRD) with a co-repressor complex (Nan and Bird 1999.) Mbd4, on the other hand, is a thymine

glycosylase and is not associated with transcriptional repression (Hendrich *et al* 1999.) It is thought that Mbd4 may play a part in limiting the mutagenicity of methylcytosine (which readily undergoes deamination) (Hendrich *et al* 1999, Bellacosa 2001, Petronzelli 2000.) MeCP2 is a transcriptional repressor; mutations in human MECP2 cause Rett syndrome, a neurodevelopmental syndrome which affects mainly females and is a major cause of mental retardation (Hagberg *et al* 1985, Amir *et al* 1999.)

1.5.1. Mbd1.

Mbd1 is the largest member of the MBD family (Wade *et al* 2001.) It is not essential for development (i.e. $Mbd1^{-/-}$ null mice survive) but $Mbd1^{-/-}$ adults have defects in adult neurogenesis and hippocampal function and have an elevated rate of aneuploidy in neurons (Zhao *et al* 2003.) Mbd1 contains a TRD at the C-terminal end (Wade 2001.) Yeast 2-hybrid screening has revealed that Mbd1 interacts with the Suv39H1 histone methyltransferase (Fujita *et al* 2003) and p150, a subunit of chromatin assembly factor 1 (Reese *et al* 2003.) Mbd1 has also been shown to interact with the methylpurine-DNA glycosylase (Mpg) which removes 3mA and 7mG from DNA by base excision repair (BER) (Watanabe *et al* 2003) providing an interesting link between transcriptional repression and DNA repair functions via an MBD protein. Watanabe *et al* proposed that an Mbd1/Mpg complex sits on the promoter; when challenged with the alkylating agent MMS, Mbd1 dissociates, leaving Mpg to repair the damage. After repair, the complex reforms (Watanabe *et al* 2003.)

Uniquely among the MBD family, Mbd1 also contains three zinc-coordinating CXXC motifs (Reese *et al* 2003.) The third of these domains, CXXC-3, undergoes alternative splicing in mouse and human and splice variants containing CXXC-3 are capable of DNA methylation independent

transcriptional repression (Stancheva *et al* 2004, Reese *et al* 2003.) *Mbd1* localises to CpG sites *in vivo* via the CXXC-3 domain (Stancheva *et al* 2004.) Only the MBD is required for binding to methylated DNA and the CXXC-3 motif can bind to unmethylated DNA (Jørgenson *et al* 2004.) Several mechanisms may account for this; firstly that the dual binding capacity of Mbd1 may allow it to target CpGs regardless of their methylation state; secondly, the different isoforms may confer a degree of flexibility, allowing tissue specific gene regulation, and thirdly that Mbd1 may use both domains at once to allow binding to sites which have both methylated and unmethylated CpGs (Jørgenson *et al* 2004.) Clearly, the role of Mbd1 in gene expression is more complex than originally thought.

1.5.2. Mbd2.

Mbd2 is the methyl-binding component of the Mecp1 complex, which represses transcription from densely methylated genes (Ng *et al* 1999, Feng and Zhang 2001.) Mecp1 also contains Hdac1, Hdac2 and RbAp46/48 proteins which allow Mbd2 to target histone deactylase/chromatin remodelling activity to methylated templates (Feng and Zhang 2001.) Mbd2 can also associate with Mbd3, which is part of the Mi2/NuRD co-repressor complex. Mbd2 and Mbd3 can form a complex with Dnmt1 on hemimethylated DNA at replication foci which may help to retain the repressive state after replication (Feng *et al* 2002.) Mbd2 can also act synergistically with other proteins; for example, the recently identified Mizf protein interacts with Mbd2 and acts as a transcriptional repressor in an HDAC-dependent manner (Sekimata *et al* 2001.)

Mbd2 null (*Mbd2*$^{-/-}$) mice have been created and are found to be viable and fertile, with normal imprinting patterns and methylation levels (Hendrich *et al* 2001.) The knockout is not entirely benign, however, as these mice have a number of problems in carrying and delivering offspring,

and show impaired nurturing behaviour (Hendrich *et al* 2001.) Significantly reduced repression of exogenous methylated promoters was also seen in cells cultured from *Mbd2*$^{-/-}$ mice, (which cannot form a normal MeCP1 complex) which were capable of being rescued by re-introduction of Mbd2. It is not known what the effects of this reduced silencing would be *in vivo*, although it is interesting to speculate that the consequences of 'leaky' repression could include a predisposition to neuronal defects (Hendrich *et al* 2001, Carter and Segal 2001.)

It appears that the activity and function of Mbd2 is highly dependent on its various binding partners. For example, one binding partner, the GTP-ase Mbd2in is capable of binding to Mbd2 and reactivating Mbd2-repressed promoters without altering their methylation status (Lembo *et al* 2003.)

Perhaps the most exciting recent developments surrounding Mbd2, however, have been clear evidence that the gene plays a role in tumourigenesis. A paper by Sansom *et al* used the *Apc*^{Min} mouse crossed onto *Mbd2* heterozygous or null backgrounds to show that Mbd2 deficiency strongly suppresses intestinal tumourigenesis in the mouse (Sansom *et al* 2003c.) The *Apc*^{Min} mouse is strongly predisposed to intestinal tumourigenesis, and is used as a mouse model of the human intestinal cancer syndrome FAP (Familial Adenomatous Polyposis.) When the *Mbd2* gene is deleted on an *Apc*^{Min} background, mice live significantly longer and have a greatly reduced tumour burden (Sansom *et al* 2003c.) Mbd2 has also been shown to have a protective role in bladder carcinoma (Zhu *et al* 2004.) Mbd2 may also be involved in other tumourigenic phenotypes; the human p14 promoter lies within a CpG island which shows abberant methylation in colon carcinomas and Mbd2 also binds to the 5' regulatory ends of the p16 and p14 tumour suppressor genes in human cancer cell lines, contributing to their transcriptional repression. (Magdinier *et al* 2001, Zhu *et al* 2004.) Since *Mbd2* null mice appear both viable and fertile,

MBD2 inhibitors could potentially be a productive avenue to explore for therapeutic purposes. Although work is at an early stage, *MBD2* antisense inhibitors have been used to suppress tumourigenesis of human cancer cell lines *in vitro* and *in vivo*, providing valuable proof-of-principle (Campbell *et al* 2004, Zhu *et al* 2004.)

1.5.3. Is Mbd2 a demethylase?

Mbd2 has been described as being a demethylase (Battacharya 1999) although other groups have disputed this (Ng *et al* 1999, Hendrich *et al* 2001, Santos *et al* 2002.) If reduced methylation inhibits intestinal tumourigenesis in the Apc^{Min} mouse, then the loss of Mbd2 would be expected to increase the amount of methylation, thus increasing tumour formation, whereas it in fact reduces it (Sansom *et al* 2003c.) This argues that Mbd2 is not a demethylase in the mouse.

1.5.4. Mbd3.

Mbd3 is highly similar to Mbd2 all along its length, (Hendrich and Bird 1998) and is a component of the Mi2/Nurd chromatin remodelling complex (Zhang *et al* 1999.) The binding properties of Mbd3 appear to vary between species; murine Mbd3 does not bind methylated DNA (Hendrich and Bird 1998, Saito and Ishikawa 2002) whereas *Xenopus* Mbd3 binds strongly to methylated DNA (Wade *et al* 1999) with a similar affinity to the isolated MBD domain from MeCP2 (Wade *et al* 1999.) Intriguingly, *Xenopus* also produces a long form of Mbd3 via alternative splicing, xMbd3lf, which has an extra 20 aa inserted into the MBD and is thus incapable of binding methylated DNA (Wade *et al* 1999.)

Because *Mbd2* and *Mbd3* are similar in sequence, it had been suggested that they may have a number of redundant functions *in vivo*. Studies by Hendrich *et al* on $Mbd3^{-/-}$ mice show that Mbd3

is absolutely required for embryonic development (Hendrich *et al* 2001.) *Mbd3*$^{-/-}$ blastocysts were detected, but no normal embryos were present past the implantation stage, with embryos being severely abnormal and in the process of being reabsorbed by 8.5 dpc (Hendrich *et al* 2001.) These studies show conclusively that Mbd2 and Mbd3 are not functionally redundant *in vivo* (Hendrich *et al* 2001.) The vital nature of Mbd3 to the developmental process has also been reiterated in the *Xenopus* system. *xMbd3* was found to be highly expressed in the nascent eye regions, in the brain and the branchial arches of *Xenopus* embryos (Iwano *et al* 2004.) Injection of *xMbd3* antisense morpholino oligonucleotides caused severe defects in eye and brain formation, and over--expression of *xMbd3* and *xMbd3lf* (which cannot bind methylated DNA) also caused a similar phenotype in a dominant negative manner (Iwano *et al* 2004.)

In 2002, two novel genes were identified on human chromosome 19 that are 30-42% identical to human *MBD3* and *MBD2* but which lack the MBD itself (Jiang *et al* 2002.) These genes were named *MBD3L1* and *MBD3L2* (Jiang *et al* 2002.) Neither protein was found to be capable of binding methylated DNA (Jiang *et al* 2002.) Mouse orthologues of both proteins were identified (although at only 31% and 36% identical, one may be somewhat sceptical that they are true orthologues) and characterised and *MBBD3L* expression was found only weakly in normal human tissue and stronger expression was found only in abnormal tissues such as germ cell tumours, the human ovarian teratocarcinoma cell line PA1 and the mouse embryonal carcinoma cell line P19 (Jiang *et al* 2002.)

Expression patterns of *MBD3L1* were examined, and MBD3L1 was found to be localised in the nucleus, in large foci, and was capable of a degree of transcriptional repression which was independent of HDAC activity (Jiang *et al* 2002.) Interestingly, the expression of *Mbd3l1* was extremely tissue-specific; occurring at significant quantities only in mouse testis at stages VII-XI of mouse spermatogenesis (Jiang *et al* 2002.) Surprisingly, Jiang *et al* did not make a connection

between the expression of *Mbd3l1* (which has similarity to a methyl CpG binding protein but does not bind methylated DNA) in the process of spermatogenesis, when methyl marks are being altered in the male gamete

1.5.5. Mbd4.

Mbd4 contains a methyl CpG binding domain (Hendrich and Bird 1998) but unlike many other members of the MBD family, it does not appear to be involved in transcriptional repression (Bellacosa 2001.) Instead, Mbd4 functions to maintain the integrity of the genome, and has roles in base excision repair (BER) mismatch repair (MMR) and the cell cycle response to DNA damage (Parsons 2003.) *Mbd4*$^{-/-}$ mice have been created and are viable and fertile with no apparent increase in tumour susceptibility (Wong *et al* 2002, Millar *et al* 2002.)

Mbd4 has two functional domains; the MBD, which directs binding to hemi- or fully-methylated DNA, and a C-terminal domain which acts as a thymine DNA glysosylase to remove T or U from mismatched CpG sites *in vitro* (Hendrich *et al* 1999, Petronzelli *et al* 2000) The CpG dinucleotide is hypermutable, with G•T and G•U mismatches being formed by the hydrolytic deamination of 5 methyl-cytosine and cytosine to thymine and uracil. These deamination events are frequent, occurring at a rate of 2-300 per cell per day and if left unrepaired would form G•C and A•T transition mutations at the next round of DNA replication (Bellacosa 2001, Hendrich *et al* 1999.) By preventing these mutations, Mbd4 may act as a caretaker of genomic fidelity at hypermutable CpG sites. Such mutations are known to contribute to tumourigenesis; nearly 50% of somatic P53 mutations in colorectal cancers arise at 'hotspots' where cytosines in CpGs are deaminated to form transition mutations (Schmutte *et al* 1995, Petronzelli *et al* 2000.) Loss of functional Mbd4 may accelerate the formation of

transition mutations and therefore Mbd4 can be considered to have tumour suppressor activity. The thymine glycosylase activity of Mbd4 has been shown to reduce the mutability of 5mC *in vivo*, as the frequency of mutation of 5mC in a mouse transgene is significantly increased in *Mbd4*$^{-/-}$ mice (Millar *et al* 2002, Wong *et al* 2002.) Mbd4 is also capable of repairing other lesions, albeit with a lower specificity than for T or U mismatches. Mbd4 can remove 5-flurouracil (5-FU) from 5-FU base pairs (Petronzelli *et al* 2000.) Mbd4 can also remove thymidine from O6 methylguanine•T base pairs (Cortellino *et al* 2003.) This indicates that Mbd4 may play an important role *in vivo* in reducing mutation damage and thus might be expected to play a role in cancer prevention.

The *MBD4* gene has been shown to be mutated in a number of mismatch-repair dependent tumours (Riccio *et al* 1999, Bader *et al* 1999, Bader *et al* 2000.) Moreover, deficiency of Mbd4 has been shown to accelerate tumourigenesis on an Apc^{Min} background in transgenic mice (Millar *et al* 2002.) The tumours formed in these mice showed an increase in the number of mutations at CpG sites in the *Apc* allele. This effect was not fully penetrant, which implied that Mbd4 may be doing more *in vivo* than just initiating repair of T•G mismatches (Sansom *et al* 2003a, Millar *et al* 2002.)

In addition to a role in base excision repair, Mbd4 has been shown to interact with Mlh1, a key component of the mismatch repair (MMR) pathway (Bellacosa *et al* 1999.) During the major pathway of post-replicative MMR, Mlh1 and Pms2 form heterodimers which interact with Msh2/Msh6 heterodimers bound to mismatched bases (Jiricny 1998.) Mbd4 was found to interact with Mlh1 and also with Msh2 – components of the mammalian mismatch repair (MMR) system, in yeast 2-hybrid screens, prompting speculation that it may itself be a part of the MMR machinery (Bellacosa 2001.) However, other co-immuoprecipitation-based studies

have shown that Mbd4 does not interact with Msh6 or Pms2, which argues that Mbd4 is not part of the MMR machinery as it is understood (Sansom *et al* 2004b.)

Many sporadic colorectal tumours and those arising in HNPCC (Hereditary non-polyposis colorectal cancer) patients are characterised by MSI (microsatellite instability) and are unable to repair slippage-induced insertion/deletion loops (IDLs) at simple repetitive sequences (Riccio *et al* 1999.) Bellacosa *et al* showed that the Mbd4 gene contains 4 such sequences (Bellacosa 2001.) Bellacosa and Bader detected Mbd4 mutations in approximately 20-25% of human colorectal and gastric carcinomas (Bellacosa 2001, Bader *et al* 1999.) Virtually all of these mutations were in a poly$(A)_{(10)}$ tract in the central portion of the gene, resulting in frameshifts which produce a truncated protein with no C-terminal domain and no catalytic ability (Bader *et al* 1999.)

Sansom *et al* have crossed Mbd4 deficient ($Mbd4^{-/-}$) mice to mice lacking MMR and have shown that in the context of MMR deficiency, the additional loss of Mbd4 does not accelerate spontaneous mutation frequency (as measured at the endogenous Dlb-1b locus) or alter tumour onset, tumour spectrum or levels of MSI (Sansom *et al* 2004.) This suggests that Mbd4 does not affect MMR-dependent tumourigenesis (Sansom *et al* 2004.) Mice doubly null for Mbd4 and Mlh1 show no increase in apoptosis when exposed to 5FU or temozolamide, implying that Mbd4 and Mlh1 lie on the same pathway and that MMR-dependent apoptosis may be mediated through Mbd4.

Mbd4 also has a key role in mediating the apoptotic response to DNA damage (Sansom *et al* 2003a.) Sansom *et al* have characterised this response, and showed that mice null for Mbd4 show a significantly reduced apoptotic response to a range of DNA-damaging agents such as 5-FU, temozolamide, γ-irradiation and cisplatin (Sansom *et al* 2003a.) Mice deficient in MMR have

previously been shown to have a similarly reduced response to DNA damaging agents (Toft *et al* 1999, Sansom *et al* 2001, Sansom and Clarke 2002) and so Sansom *et al* postulate that the interaction between Mbd4 and MMR is associated with the normal damage response, and that the response to DNA damage is somehow mediated through Mbd4 (Sansom *et al* 2003a.)

1.5.6. Kaiso.

Kaiso is a newly discovered member of the Methyl CpG binding domain protein family (Prokhortchouk *et al* 2001.) Kaiso belongs to the BTB/POZ (broad complex, tramtrack, bric a brac/ pox virus zinc finger) family of proteins and unlike other methyl CpG binding domain proteins, it does not contain a classical MBD, but instead binds CpG dinucleotides through a zinc finger motif (Prokhortchouk *et al* 2001.) Kaiso recognises and binds to sequences which contain at least two methyl CpG dinucleotides and is capable of repressing transcription from methylated templates in a methylation-dependent manner (Prokhortchouk *et al* 2001.) Kaiso was first identified as a binding partner of the p120 catenin, $p120^{ctn}$ (Daniel and Reynolds 1999.) $p120^{ctn}$ is an armadillo repeat protein and a component of the E-cadherin-catenin cell adhesion complex (Daniel and Reynolds 1999.) It binds $p120^{ctn}$ at a distinct juxtamembrane site that has been implicated in the regulation of cell adhesion and motility (Daniel and Reynolds 1999.) This raises the possibility that Kaiso provides a link between cell-surface signalling and methylation-dependent gene regulation. $p120^{ctn}$ and Kaiso are also associated in the nucleus, where it is thought that $p120^{ctn}$ may negatively regulate Kaiso-mediated transcriptional repression (Ruzov *et al* 2004, Kim *et al* 2004.)

Kaiso has been shown to be a methyl CpG-specific repressor; Prokhortchouk *et al* showed that ectopic expression of Kaiso in *Mbd2*$^{-/-}$ fibroblasts was capable of restoring repression of a methylated construct (Prokhortchouk *et al* 2001, Kim *et al* 2004.)

A *Xenopus* homologue of Kaiso, xKaiso, has been studied *in vivo* proving that in *Xenopus*, xKaiso is essential for development (Ruzov *et al* 2004.) *Xenopus* embryos are an excellent model in which to study the mid-blastula-transition (MBT) the point where the zygote's own genome, previously repressed, begins to be transcribed (Ruzov *et al* 2004.) The mechanisms whereby zygotic genes are repressed pre-MBT are still unknown, but a recent paper by Ruzov *et al* suggests that xKaiso may be responsible (Ruzov *et al* 2004.) An *xKaiso* knockdown recapitulates the phenotype of an *xDnmt1* knockdown: the embryos exhibit premature activation of gene expression, apoptosis and ultimately, developmental arrest (Ruzov *et al* 2004.) Excessive expression of *xKaiso* did not appear to have any negative effects on development (Ruzov *et al* 2004.) As well as the initiation of zygotic transcription, the MBT is associated with the slowing of cell division, and the beginning of cell migration. Ruzov *et al* note that loss of cell adhesion is seen with both *xDnmt1* and *xKaiso* knockdowns (Ruzov *et al* 2004.)

Although Kaiso is clearly vital for development in *Xenopus*, the same is not true for mice; *Kasio*$^{-/-}$ mice have been generated and viable, fertile and appear healthy (Clarke AR, personal communication 2004.) This indicates that either the primary roles of the protein are different in the two species, or that mammals have another activity which can compensate for the loss of Kaiso.

1.6. MECP2 and Rett Syndrome.

The importance of the MBD family of proteins is underlined by the effects of mutations in one member of the family; *MECP2*, which cause the human disorder, Rett syndrome (RTT OMIM#

312750, Neul and Zogbhi 2004.) This thesis will examine the consequences of removal of MeCP2 in the intestine and mammary gland to attempt to shed light on the complex etiology of this syndrome.

Rett syndrome is a complex neurodevelopmental syndrome which affects an estimated 1 in 10,000 to 22,000 live births (Hagberg 1985, Miyamoto *et al* 1997.) Most affected patients are female, although there are rare male cases (Coleman 1990, Clayton-Smith *et al* 2000, Topcu *et al* 2002.)

Rett syndrome was first described in 1966 by Andreas Rett, an Austrian physician who noticed two girls with a strikingly similar phenotype in his waiting room (Rett 1966.) The condition was not widely recognised, however, until 1983, when Hagberg *et al* described 35 girls with similar symptoms (Hagberg *et al* 1983.) A classic Rett patient will develop apparently normally for the first 6 to 18 months of life, and then undergo a rapid period of regression, losing acquired speech and motor skills and acquiring a range of symptoms which define the syndrome (Hagberg *et al* 1983.) This regression has been characterised as comprising four stages (Hagberg and Witt-Engerstrom 1986.) In the first stage of classical Rett syndrome (stage I) the normal acquisition of new learned skills stops. Head growth decelerates, leading to microcephaly. Indeed, one of the earliest clinical signs of Rett syndrome is a marked deceleration of growth, with 85 to 94% of girls showing growth deceleration across all measurements (Motil *et al* 1994.) Affected girls become emotionally withdrawn and avoid eye contact, classic signs of autism (Neul and Zoghbi 2004.)

In the second stage, which generally lasts from around 1 to 4 years of age) affected girls lose skills they have learned, such as any speech they may have developed, and the ability to use the hands purposefully (Neul and Zoghbi 2004.) In stage II, apraxia and breathing irregularities develop, along with the stereotypical 'hand-washing' movements which are so characteristic of Rett syndrome (Neul and Zoghbi 2004.) During stage II, around half of girls also develop seizures

(Neul and Zoghbi 2004.) During stage III (generally 4-7 years) most acquired skills are lost and there is a period of stabilisation (Neul and Zoghbi 2004.) Epileptic seizures become more common (Neul and Zoghbi 2004.)

Stage IV is considered to be from ages 5-15 and onwards; during this time, seizure frequency commonly decreases, but motor impairment and deterioration continues (Zoghbi *et al* 2002.) Many girls develop scoliosis which can leave them wheelchair-bound (Hageber and Witt-Engerstrom 1996.)

The life expectancy of Rett patients is decreased, and although many do live well into adulthood (Zoghbi *et al* 2002) there is an elevated incidence of sudden death of 1.2% per year (Kerr *et al* 1997) which is thought to be attributable to cardiac and breathing abnormalities, such as the long, uncorrected QT intervals suffered by patients (Sekul *et al* 1994, Guideri *et al* 1999.)

Because the mechanism of action of MECP2 is still unknown, treatment for Rett patients is based around the alleviation of symptoms to improve quality of life, rather than an attempt to reverse the primary (unknown) causes; physical, occupational and music therapies are used to improve general well-being (Christodolou *et al* 2003.) L-carnitine is also used, as Rett patients have lowered carnitine and very long chain fatty acid levels (Christodolou *et al* 2003.) A ketogenic diet can be used (to suppress seizures) along with opiate agonists to help with breathing disturbances, seizures and stereotyped involuntary movements (Christodolou *et al* 2003.)

1.6.1. Mutations in Methyl CpG Binding Protein 2 *(MECP2)* cause Rett syndrome.

The actual cause of Rett syndrome was unknown until 1999, when Amir *et al* identified mutations in *MECP2* as being responsible (Amir *et al* 1999.) Although the almost total lack of affected males suggested an X-linked disorder which was lethal to the male foetus, conventional

linkage analysis was hampered by the paucity of familial cases of Rett. A small number of familial cases were found, however, which allowed the target region to be narrowed down. Transcripts in this region were studied, and eventually, in 1999, Amir *et al* found mutations in human *MECP2* to be responsible for Rett syndrome (Amir *et al* 1999.) The case for *MECP2* as the causative mutation was further strengthened by the discovery of *MECP2* mutations in 80% of sporadic Rett cases (Neul and Zoghbi 2004.) The other 20% may well be mutations which lie outside the sequenced portions, in as-yet unknown control elements (Neul and Zoghbi 2004.)

1.6.2. Rett syndrome is phenotypically diverse.

Rett syndrome is not tightly defined by a set of clinical criteria, nor by a single type of mutation. Instead, a range of *MECP2* mutations have been found in Rett patients which lead to a wide spectrum of symptoms (Hagberg and Skjeldal 1994.) In the most severe female forms of Rett, the so called 'congenital forms' the normal period of development is absent (Hagberg and Skjeldal 1994) and girls often have infantile spasms and congenital hypotonia (Neul and Zoghbi 2004, Hagberg 1995) whereas in the milder 'form fruste' ('worn down') form, motor functions and speech may be preserved and patients do not have seizures (Zapella *et al* 1998, Hagberg *et al* 1995.) Diagnosis is reached clinically by using a set of criteria similar to that shown in table 1 below. As is clear from table 1, the phenotypic range of Rett syndrome is extremely broad. (Zoghbi *et al* 2002.)

Table 1. Phenotypic Variation in Rett syndrome. *(Females)*

(From Young and Zoghbi 2002.)

	Phenotype		
	Classic RTT	**Mild RTT variants**	**Severe RTT variants**
Genotypes	Null/severe inactivating alleles. Balanced XCI	Hypomorphic alleles (late truncations) with balanced XCI or null alleles with favourable XCI.	Null alleles or severe inactivation, possibly with skewed XCI.
Age at onset	6-18 months	Later onset	Congenital
Head/body size	Small head and body	May have small head/ body	Small head/body
Seizures	Yes	No	Yes. Early onset
Speech	Loss of speech	Speech preserved	No speech
Motor function	Motor deficiencies	Usually ambulatory	Hypotonia, motor deficiencies
Hand use	Stereotypical hand use	Retain hand use	---
Social interactions	Autistic features	variable	---
Intelligence	Mental retardation	Mild/no retardation	Severe mental retardation
Spinal curvature	Scoliosis and/or kyphosis	Rarely seen	Scoliosis and/kyphosis
Respiration	Breathing dysfunction	Rarely seen	Breathing dysfunction

1.6.3. Genotype/Phenotype correlation in Rett syndrome.

The correlation of genotype to phenotype is complex and still at an early stage, and is further complicated by the fact that *MECP2* is carried on the X chromosome, so that skewed X-inactivation may contribute to severity (Neul and Zoghbi 2004.) However, some genotype/phenotype correlation has been done, albeit with somewhat contradictory results. Two studies have found that the phenotypes associated with truncating mutations are more severe than those associated with missense mutations (Cheadle *et al* 2000 and Monros *et al* 2001.) However,

other studies have failed to find such a link (Young and Zoghbi 2004, Amir *et al* 2000, Bienenue *et al* 2000, Huppke *et al* 2000, Guinti *et al* 2001, Yamada *et al* 2001 Chae *et al* 2002.) Other work appears to show a correlation between mutations affecting the NLS and phenotype, (Huppke *et al* 2002.) Huppke *et al* found four relationships between severity and mutation; mutations affecting the NLS were more severe than those not affecting it, mutations in the CTD were less severe than other types of mutation, truncations were found to be more severe than missense mutations (unless they occurred in the CTD) and missense mutations in the TRD and MBD were equivalent (Zoghbi *et al* 2004, Huppke *et al* 2002.)

The relationship between individual mutations and phenotype has been the subject of a study by Leonard *et al*, who studied the R133C mutation. *In vitro* models suggested that the R133C variant has a relatively normal functioning, and indeed, patients with R133C were less likely to have severe symptoms relating to speech, hand movements and ability to walk (Leonard *et al* 1999.) Clearly, a great deal more research is needed in this area to correlate individual mutations with specific phenotypes. The Edinburgh-based database www.MeCP2.org.uk is collecting detailed information provided by parents and carers on *MECP2* mutations and their associated symptoms. The RetBASE database http://MeCP2.chw.edu.au also collects information from clinical sources such as paediatricians and laboratories that test for mutations in *MECP2.*

Correlation of genotype to phenotype is complicated by the fact that *MECP2* lies on the X-chromosome. Because *MECP2* is X-linked, and *MECP2* is subject to X chromosome inactivation (XCI) affected females are a mosaic of wild-type and mutant cells (Neul and Zoghbi 2004.) If the X-inactivation patterns are skewed, the severity of the disease can be modulated (Neul and Zoghbi 2004.) Whilst studying a mouse model of Rett syndrome in which the MECP2 protein is truncated at amino acid 308 ($Mecp2^{308/X}$) which results in a hypomorphic allele similar to a mutation found in Rett patients, Young and Zoghbi noted that there was much more phenotypic variation in the

symptoms of the female mice than would have been expected (Young and Zogbhi 2004.) They used this mouse model to study XCI patterns in the brain and found that XCI patterns were in fact non-random, with the wild type allele being favoured (Young and Zogbhi 2004.) This suggests that there is a survival advantage for the wild-type neurons, or a differentiation defect in mutant neuronal precursors (Young and Zogbhi 2004.) Interestingly, no females were found in which the mutant allele was favoured, which favours the non-random survival theory over stochastic XCI skewing (Young and Zogbhi 2004.) How relevant is this to the human form of the disease? Mice with the *Mecp2*$^{308/Y}$ mutation show a much greater phenotypic variability than humans with a comparable mutation, which implies that directed skewed XCI may be less important in humans (Neul and Zoghbi 2004.) The majority of human Rett patients show balanced XCI in brain tissue (Shabazian 2002) but examples of non-random XCI have been found (Ranieri *et al* 2003) Most interestingly, skewed XCI which favours the expression of the wild-type allele has been found in an asymptomatic carrier of a classic Rett mutation (Siranni *et al* 1998) and in patients who have much milder symptoms than would be expected due to favourable XCI (Wan *et al* 1999, Bienvenue *et al* 2000, Ishii *et al* 2001, Zapella *et al* 2001, Huppke *et al* 2002) This raises the intriguing possibility that females with mild mental retardation of unknown cause could in fact be carrying *MECP2* mutations and that the spectrum of the disease may be much wider than was previously thought.

1.6.4. Several neurodevelopmental syndromes have considerable phenotypic and genetic overlap.

There are a number of syndromes with considerable clinical overlap in which diagnosis may well be down to the (subjective) diagnosis of a clinician rather than the genotype of the individual. Several syndromes, such as RTT, Angelman and other autism-spectrum disorders may share common mechanistic pathways (Zoghbi *et al* 2003.) The autistic spectrum of disorders shows a complex etiology (Samaco *et al* 2005.) A genetic component is highly likely, since autism exhibits a high degree of heritablility in families (Volkmar *et al* 2003) but there are probably multiple gene loci involved, as well as a considerable degree of environmental influence (Samaco *et al* 2005, Volkmar *et al* 2003.)

Angelman syndrome (AS, OMIM #105830) is an imprinted disorder which can be caused by deficiency of maternal chromosome 15q11-q13, methylation defects or mutations in the maternal copy of *UBE3A,* which encodes the ubiquitin ligase UBE3A/E6-AP (Zoghbi *et al* 2003.) Angelman syndrome and RTT have a degree of phenotypic overlap (Zoghbi *et al* 2003) as both present with a high frequency of autistic traits, developmental delay, seizures, stereotypic behaviours and language impairment (Zoghbi *et al* 2003.) One study has suggested that up to 8% of patients who show an Angelmann-like phenotype but have no mutations or abnormalities in chromosome 15, may have mutations in *MECP2* (Watson *et al* 2001) and *MECP2* mutations have also been reported in patients diagnosed primarily with autism (Lam *et al* 2000, Beyer *et al* 2002.) The 15q11-q13 duplications have also been found in ~1% of autism case (Schroer *et al* 1998.) This suggests that the boundaries that clinicians draw between these syndromes are not necessarily set in stone, and hint at a common pathology for several disorders.

One study has shown that epigenetic overlap between RTT and Angelman syndrome can occur because MECP2 deficiency actually causes reduced expression of *UBE3A*, the gene which is often mutated in Angelman syndrome, and *GABARB3*, which encodes the β-subunit of the $GABA_A$ receptor and is another gene associated with autism (Samaco *et al* 2005, Cook *et al* 1998, Buxbaum *et al* 2002.) Samaco *et al* also found reduced *UBE3A* and *GABARB3* expression in samples of autistic brain, consistent with another report (Veenstra-VanderWeele *et al* 1999.) Indeed they go as far as to posit the *$GABA_A$* receptor genes as potential candidate genes for autism, either as direct mutations, or as indirectly affected genes. There is enough evidence to make this an attractive theory: several studies have demonstrated linkage to the $GABA_A$ receptor family (Samaco *et al* 2005, Cook *et al* 1998, Buxbaum *et al* 2002, Nurmi *et al* 2003.) Autoradiography has shown reduced $GABA_A$ binding in autistic brain (Blatt *et al* 2001) and elevated circulating $GABA_A$ levels have been found in autistic children (Moreno-Fuenmayor *et al* 1996, Dhossche *et al* 2002, Aldred *et al* 2003.) Perhaps other primary lesions act upstream of an eventual $GABA_A$ effect?

One recent paper has opened up debate on the subject of phenotypic overlap by finding that one severe variant of RTT, the Hanefield variant, which presents with early onset of seizures, is associated with mutations, not in *MECP2* but in the cyclin dependent kinase-like 5, *CDKL5/STK9* (Scala *et al* 2005.) The Hanefield variant of RTT has considerable phenotypic overlap with another syndrome, West syndrome, also called Infantile Spasm Syndrome X-linked (ISSX OMIM #308350.) ISSX causes infantile spasms, hypsarrhythmia and severe to profound mental retardation (Cook *et al* 1998.) Some families with ISSX have been found to have mutations in the aristaless related homeobox gene *ARX* (Cook *et al* 1998, Buxbaum *et al* 2002.)

Mutations in *ARX* have been documented to cause several forms of epilepsy, inclding infantile seizures, myoclonic seizures and peripheral dystonia, as well as syndromic and nonsyndromic X-lined mental retardation (Cook *et al* 1998, Buxbaum *et al* 2002.) The *CDKL5/STK9* gene has also been found to be involved in ISSX (Scala *et al* 2005.) Scala *et al* reasoned that the phenotypic overlap seen between the early-onset of seizures from of RTT and ISSX warranted screening two early-onset RTT cases for ISSX-causing *CDKL5/STK9* mutations (Scala *et al* 2005.) These patients were found to have no point mutations or re-arrangements in the *MECP2* gene (including the exon 1 region) but both had mutations in various parts of the *CDKL5/STK9* gene, leading to non-functional protein being expressed (Scala *et al* 2005.) Scala *et al* then extended their analysis to 19 classic RTT and 15 PSV cases, and found no mutations in *CDKL5/STK9*, implying that these mutations are responsible for a rare subset of severe, early-onset RTT of the type described by Hanefield (Scala *et al 2005.*) This version of RTT is extremely rare, and Scala *et al* acknowledge that only a few cases have ever been described (Zoghbi *et al* 2003, Mount *et al* 2003) and so it may well be a distinct syndrome, or a variant of ISSX.

1.6.5. Rett Syndrome in Males.

Because Rett syndrome is an X-linked disorder, it was initially assumed that it was an exclusively female disease and that Rett syndrome would be universally lethal to the male foetus. However, rare male cases are found (Wan *et al* 1999, Villard *et al* 2000, Hoffbuhr *et al* 2001, Geerdink *et al* 2002, Zeev *et al* 2002) which tend to fall into 3 categories; boys with the classical Rett syndrome, boys with severe encephalopathy and boys with les severe neurological phenotypes (Neul and Zogbhi 2004.)

Boys who have classical Rett syndrome mutations tend to be born with severe encephalopathy and die within one or two years of age (Wan *et al* 1999, Villard *et al* 2000, Hoffbuhr *et al* 2001, Geerdink *et al* 2002 Zeev *et al* 2002.) The fact that these mutations are also found in viable female children (Zeev *et al* 2002) shows both the effects of mosaicism in the body and the hazard of unmasking alleles on the X-chromosome. In a male with an XY karyotype, all cells carry the mutant allele, which accounts for the massively increased severity of the phenotype in males; unlike females they have no mosaicism to ameliorate the effects of the mutant allele and so the male MeCP2 allele is functionally haploid. However, in males with Klinefelter's syndrome, who are genetically XXY but phenotypically male, the extra X can be inactivated, leading to classical Rett mosaicism and viability (Zoghbi *et al* 2003.)

Some males are also found who have acquired an MeCP2 mutation early in development, leading to somatic mosaicism (Clayton-smith *et al* 2000, Armstrong *et al* 2001, Topcu *et al* 2002.) Other rare events can also cause Rett syndrome, such as a case of a male with a translocation of the SRY (sex-determining region Y) carrying part of the Y chromosome onto a paternal X chromosome with an MeCP2 mutation, leading to a phenotypically male but karyotypically female (XX) subject (Zoghbi *et al* 2003, Neul and Zogbhi 2004.) There are even cases of MeCP2 mutations with none of these mitigating factors; Ravn *et al* reported a boy who has no apparent mosaicism or karyotypic abnormalities, but who has classic Rett syndrome caused by a mutation within the TRD (Ravn *et al* 2003.)

One exciting development in the field is the discovery of boys who carry *MECP2* mutations but do not appear to be severely disabled or have classic Rett syndrome. These boys were found by examining families with X-linked mental retardation and boys who show features of Angelmann syndrome, mental retardation or other neurodevelopmental abnormalities (Meloni *et al*,2000, Couvert *et al* 2001, Dotti *et al* 2002, Klauck *et al* 2002, Kleefstra *et al* 2002, Winnepenninckx *et al*

2002, Ynemtra *et al* 2002, Imessaoudee *et al* 2001, Cohen *et al* 2002.) The mutations found in these boys are never picked up in girls, possibly because in the heterozygous form their phenotypic effects are so mild that they may well escape detection, even though they may well exist (Neul and Zogbhi 2004.)

The spectrum of symptoms caused by *MECP2* mutations is therefore much wider than was previously thought, although almost all include mental retardation of some kind which reinforces the idea that MeCP2 has some critical role to play in the developing brain (Neul and Zogbhi 2004.) The actual mechanisms of that role, however, are far from clear.

1.6.6. Mouse models of Rett syndrome.

Mouse models of Rett syndrome have been generated by targeting the *Mecp2* allele. Guy *et al* created an *Mecp2* null mouse by creating a conditional 'floxed' allele of *Mecp2* and using a ubiquitous Cre recombinase (the so-called 'deleter' Cre) to create mice in which *Mecp2* is floxed out in the early embryo (Guy *et al* 2001.) Mice develop to birth in a normal fashion, indicating that MeCP2 is not absolutely required for viability (Guy *et al* 2001.) After birth, these mice develop normally until around 3-8 weeks of age, when they develop tremors and breathing irregularities, as well as other abnormalities such as misalignment of the jaw (Guy *et al* 2001.) The mice tend to die at around 54 days.

Mecp2 null mice were also generated by other groups, who also observed a short period of normal development followed by the onset of tremors at around 5 weeks of age and death at around 10 weeks (Chen *et al* 2001.) The reason for the disparities in lifespan may be due to the exact targeting of the allele or background effects. Chen *et al* found that the brains of their null

mice were grossly normal, although they did report smaller, and more closely packed neurons, closely mirroring findings in human Rett patients (Chen *et al* 2001.)

Both Guy *et al* and Chen *et al* also used a neural and glial cell precursor specific Cre (nestin fused) to remove *Mecp2* from the neuronal and glial populations. Mice produced in this way are reported to recapitulate the phenotype of whole-body null mice, and from this, both groups concluded that Rett syndrome was a purely nervous-system-confined disorder (Guy *et al* 2001, Chen *et al* 2001.)

MeCP2 is expressed in most bodily tissues, however (D'Esposito *et al* 1996), and so it may be expected that it plays a role, albeit perhaps a subtle one, in other tissues of the body. If *Mecp2* is deleted in neural progenitor cells, mice develop a Rett-like phenotype by 3-8 weeks of age and die at 6-10 weeks of age (Chen *et al* 2001, Guy *et al* 2001) however, when *Mecp2* is deleted in post-mitotic neurons, using a Cam-kinase linked Cre, a similar phenotype occurs, but much delayed (Chen *et al* 2001.) This indicates that *Mecp2* plays a role in maintenance and maturation of neurons, rather than their initial formation (Neul and Zogbhi 2004.)

The constitutive (whole-body) null mice are analogous to male cases of Rett syndrome in that they have no functioning MeCP2, unlike the mosaicism seen in most human female cases. In order to more closely recreate a human mutation, Shabazian *et al* created a mouse with a stop codon after codon 308 (Shabazian *et al* 2002.) The resulting MeCP2 protein contains the MBD, the TRD and the NLS but removes the C terminal domain (Shabazian *et al* 2002.) The male mice carrying this mutation appear normal until around 6 weeks of age, when they develop a mild tremor (Shabazian *et al* 2002.) The tremor becomes worse as the mice age, and at around 5 months of age, mice develop kyphosis, mimicking the scoliosis seen in human patients (Shabazian *et al* 2002.)

Some but not all mice develop seizures, and all exhibit anxiety, abnormal social behaviour and decreased grooming (Shabazian *et al* 2002.) Again, these mice develop forepaw stereotypies; many mice also lived to a year or more (Shabazian *et al* 2002.)

What have these mouse models shown us about the molecular complexities of Rett syndrome? They show that Rett syndrome can be modelled in mice, which opens up a range of promising avenues of research. Transcriptome analysis of the truncation mutant provided no significant gene changes, although histone H3 was found to be hyperacetylated, indicating abnormal chromatin structure (Shabazian *et al* 2002.) This implies that the normal processes linking methylation and chromatin remodelling may be disrupted.

1.6.7. Conditional alleles of *Mecp2.*

When modelling a human disease in the mouse, the advantages of using conditional alleles are clear; many genes, if knocked out constitutively, cause either embryonic lethality, or morbidity (such as the constitutive model of RTT.) Using a conditional allele, however, the problem of embryonic lethality can be circumvented, as mice can be born and develop with functioning copies of the gene. The gene may be switched on or removed in a specific tissue at a defined point in the life cycle, allowing the effect of removal in a specific tissue to be distinguished from the effects of whole-body loss (Baranda and Dymecki 2004.)

Cre (Causes recombination) recombinase is a site specific recombinase of the integrase family, isolated from bacteriophage P1 (Sauer and Henderson 1988.) Cre catalyses site-specific recombination between defined 34bp '*Lox P*' (locus χ of crossover P1) sites (Baranda and Dymecki 2004.) If a gene is placed between two Lox P sites and exposed to Cre, then the gene will

be excised or 'floxed out' (Baranda and Dymecki 2004.) The Cre-Lox system can be used to create tissue-specific conditional deletions of an allele, overcoming the problem of embryonic lethality or developmental defects (Baranda and Dymecki 2004.) Mice carrying the Cre recombinase under the control of an inducible, tissue-specific promoter are crossed with mice carrying the gene of interest that has been flanked by Lox P sites (Baranda and Dymecki 2004.) When the mice are given the appropriate inducer, the Cre is expressed in a spatially defined manner and the gene of interest is floxed out only in a specific tissue (fig 1.8) (Baranda and Dymecki 2004.)

Fig 1.8a. The conditional allele of Mecp2.

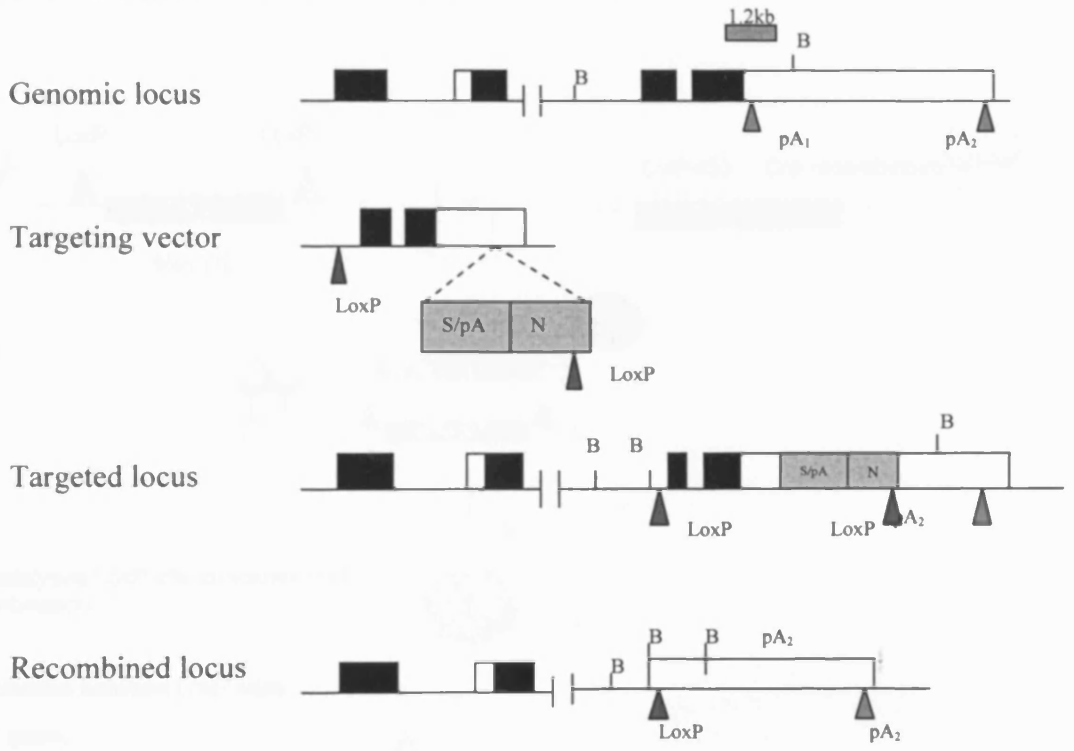


Fig 1.8a. Conditional allele of Mecp2.

The conditional allele of Mecp2 was created by Guy et al using Cre-Lox technology. Exons 3 and 4 were replaced with the same exons flanked by LoxP sites. When crossed to a line carrying an inducible Cre recombinase, the resulting progeny can be used to conditionally delete the gene in a specific tissue.

Fig 1.8b. Using Cre Lox technology to conditionally delete Mecp2.

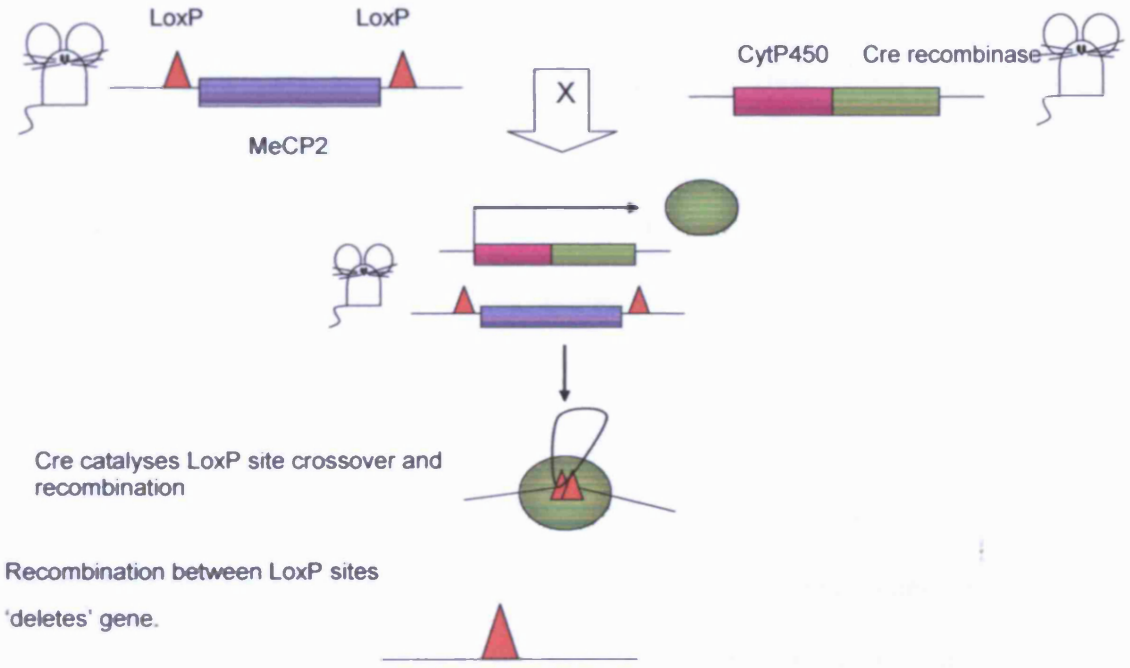


Fig 1.8b. Cre-Lox Technology.

Two mouse lines, one carrying the tissue-specific inducible Cre recombinase and one carrying the floxed allele, are bred to create a line carying both transgenes. When this mouse is exposed to the inducer (in this case, β-napthoflavone) the Cre is induced and recombination occurs between the LoxP sites, removing the gene.

1.6.8. Elevated levels of MeCP2 also cause a progressive neurological disorder.

A therapeutic approach to Rett syndrome might take the form of introducing the protein during an appropriate developmental window, in order to try to alleviate the effects of endogenous MECP2 loss or malfunction. Dose levels may well be absolutely crucial; recent work by Collins *et al* and Luikenhuis *et al* shows that slight over-expression of *Mecp2* in the mouse can have damaging effects (Collins *et al* 2004, Luikenhius *et al* 2004.) Collins *et al* over-expressed wild-type human MECP2 in transgenic mice to create levels around twice that of wild-type. At around 10 weeks of age, mice actually showed enhanced synaptic plasticity in the hippocampus, accompanied by enhanced motor and contextual learning. At 20 weeks of age, however, mice developed seizures, aggressiveness (shown by increased propensity to bite) kyphosis and hypoactivity (freezing.) The severity and onset of this phenotype depended largely on the levels of expression of the *MECP2* construct (Collins *et al* 2004.) Mice which had the highest expression levels developed more severe symptoms sooner than lower-expressing lines (Collins *et al* 2004.) Interestingly, mice which survived for a year appeared to live normal lifespans (Collins *et al* 2004.)

Collins *et al* also crossed their transgenic line onto the *Mecp2* null mouse generated by Guy *et al* and observed that human MECP2 can completely rescue the Rett-like phenotype of these mice (Collins *et al* 2004.) This argues that the effects seen in their mildly overexpressing mice are indeed due solely to the fact that MECP2 is being overexpressed, not the presence of human MECP2 in a murine environment (Collins *et al* 2004.)

Luikenhuis *et al* have also created mice in which *Mecp2* is under the control of a neuron-specific promoter (Tau) and found that 2 to 4-fold over-expression of *Mecp2* in postmitotic neurons lead to severe motor dysfunction, including side-to-side swaying, tremors and ataxia

(Luikenhius *et al* 2004.) They also found that the Rett phenotype was rescued when a transgenic construct was used to express MeCP2 in *Mecp2* mutant mice, with rescued mice showing no obvious signs of any Rett symptoms (Luikenhius *et al* 2004.) Both studies show that *MECP2* levels must be tightly regulated *in vivo*, and raise the interesting possibility that some human cases of Rett syndrome may in fact be due to overexpression of MECP2. Increased MECP2 levels have in fact been recorded in a single case of autism and a case of pervasive developmental disorder as well as a case of the preserved speech variant of Rett (Arani *et al* 2004, Samaco *et al* 2004.)

1.6.9. Neuropathological difference in the Rett syndrome brain.

Rett syndrome is considered to be a neurodevelopmental disorder, rather than a neurodegenerative disorder (Shabazian and Zoghbi 2002.) Postmortem analysis of Rett syndrome brains reveals that they weigh 14%-34% less than normal (Jellinger and Seitelberger 1986) a difference which persists even when the smaller size of Rett patients is accounted for (Armstrong *et al* 1999.)

Studies of cerebral blood flow show that Rett patients have an infantile blood flow pattern (Nielsen *et al* 1990) and *in vivo* MRI scans reveal a reduction in the volume of both grey and white matter in the Rett syndrome brain (Reiss *et al* 1993), with the average brain weight being around 950g, around that of a one-year-old child. However, no atrophy or degeneration is visible (Armstrong 1992, Armstrong *et al* 2003), indicating that the brain has failed to develop properly, rather than having developed and degenerated. The effects of Rett syndrome on the brain appear to vary with region; reduced neuron size has been reported in the hippocampus-entorrhinal cortex and in the speech-related cortex (areas 40, 45 and 22) (Baumann *et al* 1995, Belichenko 2001.) Post-mortem studies of RTT brains reveal smaller cortexes and cerebellums, accompanied by

decreased cell size, packing density and synapse formation (Armstrong *et al* 1999.) Reduced levels of nerve growth factors, along with neurotransmitters and their receptors have been reported (Armstrong *et al* 1999.)

If lack of functional MECP2 does result in incorrect execution of the brain's developmental plan, then the effects are by no means seen in all brain regions. The dendrites of the superior temporal and occipital neurons do not show the reduction in dendritic branching seen in the limbic, frontal and motor neurons, for example (Armstrong *et al* 1995.) Also, the neuroendocrine system appears to mature as normal, with no difference in the age of menarche between Rett patients and normal controls (Engerstrom *et al* 1992.)

One significant difference between Rett and normal brains occurs in the process of elaboration and pruning of neuronal synapses which occurs during the early postnatal period (see fig 1.9) (Johnston *et al* 2003) with the density of synapses being decreased in the Rett syndrome brain (Belichenko *et al* 1994.) These data suggest that Rett syndrome may be caused by alterations in dendritic organisation and problems at the synapse. This possibility will be discussed in greater detail later in the light of various experimental data.

1.6.10. Rett Syndrome Causes Defects in Synaptic Elaboration and Pruning.

Fig 1.9. Synaptic pruning in Rett and normal brain.

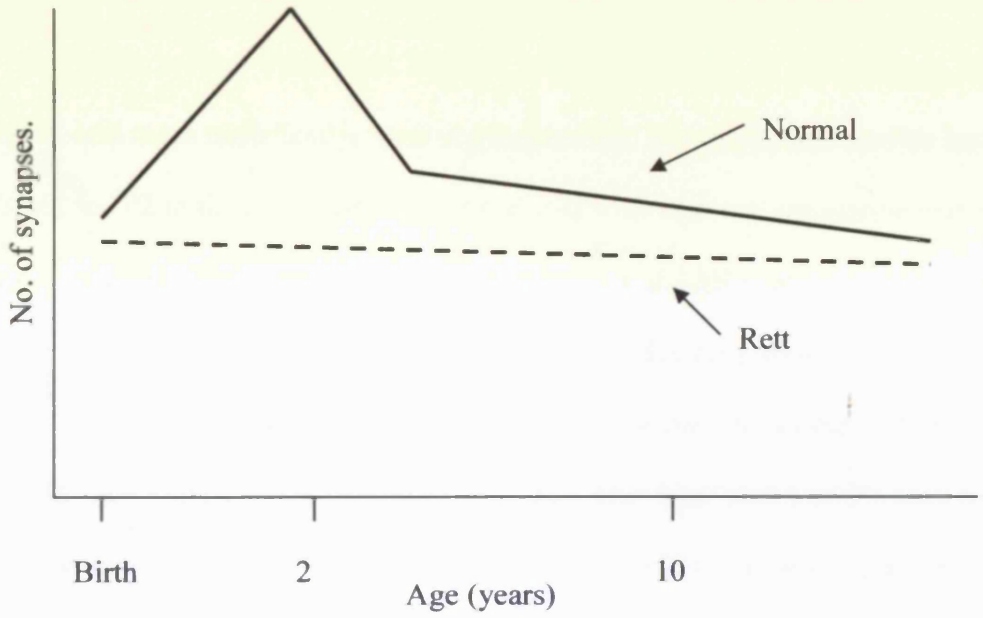


Fig 1.9. Synaptic elaboration and pruning in the RTT brain.

In the normal brain, the number of synapses increases from birth until around 2 years of age; after this point, synaptic pruning takes place to ensure the correct connections between neurons, in effect 'fine tuning' the connections within the brain. In the Rett brain, this process appears to be radically altered, with no elaboration or pruning occurring. The consequences of this are unknown, but it is logical to assume that incorrect or inefficient neuronal connectivity would result; could this alteration in brain architecture be partly responsible for the symptoms of RTT?

1.6.11. MeCP2 and neurogenesis – narrowing down the phenotype.

Could Rett syndrome be a disease of abnormal synapse function or development? Patients with Rett syndrome have a high incidence of seizures, hyperkinetic movements and abnormal breathing patterns which would suggest a defect in excitatory synapse function (Jarrar *et al* 2003.) A role for Rett syndrome in neuronal maturation and synaptic development has also been reported in other studies and MECP2 has been located to the postsynaptic compartment (Aber *et al* 2003)

There have been several clues that MECP2 might somehow be involved in the process of neurogenesis, and more specifically, neuronal maturation. Developmental studies have shown that the levels of MeCP2 in the rodent brain are correlated with neuronal maturation and synaptognesis (Mullaney *et al* 2004, Shabazian *et al* 2002, Akbarian *et al* 2001 Coy *et al* 1999) with structures that develop sooner, such as the brain stem, becoming MeCP2 positive before later developing structures such as the hippocampus and cerebral cortex. In cell cultures, MeCP2 levels also correlate with increasing maturity (Jung *et al* 2003.) This hints that MeCP2 may be involved in neuronal maturation. However, microarray studies show little or no changes in the mouse CNS in response to the loss of *Mecp2* (Tudor *et al* 2002, Matarazzo *et al* 2004)

Others have used the olfactory system as a model for the developing nervous system; as olfactory neurons (ORNs) mature, they travel up through the epithelium which allows the ageing of neurons based on position and the expression of stage-specific markers (Matazzaro *et al* 2004) The olfactory epithelium can also be accessed via biopsy in humans, and so provides a valuable opportunity to examine human tissue. The olfactory epithelia of Rett patients had fewer terminally-differentiated ORNs and far more immature neurons, supporting the theory that MECP2 is needed for neuronal maturation. Studies on the mouse olfactory system have already shown changes in the olfactory proteome (Matarazzo and Ronnet 2004) and so this system was examined

to determine whether loss of MeCP2 in the mouse lead to changes in the neurodevelopmental program of olfactory neurons (Matarazzo and Ronnet 2004.) Lack of MeCP2 resulted in delays in the terminal differentiation of developing neurons and also their glomerular organisation (Matarazzo and Ronnet 2004.) This correlates with post-mortem examinations of Rett brains, which have smaller, more densely-packed neurons with reduced dendritic arborisation (Zoghbi *et al* 2003, Armstrong *et al* 2003.) If *Mecp2* is deleted in neural progenitor cells, mice develop a Rett-like phenotype by 3-8 weeks of age and die at 6-10 weeks of age (Chen *et al* 2001, Guy *et al* 2001.) However, when *Mecp2* is deleted in post-mitotic neurons, using a Cam-kinase linked Cre, a similar phenotype occurs, but much delayed (Chen *et al* 2001.)

1.7. MECP2 – A member of the Methyl Binding Domain Family.

MECP2 is an abundant chromosomal protein (Nan *et al* 1999) located at Xq28 in humans and is a member of the MBD family (Ballestar and Wolffe 2001.) The protein is rich in basic amino acids and has several potential phosphorylation sites (Lewis *et al* 1992.) MeCP2 contains four functional domains; an 85 amino acid methyl CpG binding domain (MBD) similar in structure to those found in other MBD proteins, a 104 amino acid transcriptional repressor domain (TRD), a nuclear localisation signal (NLS) and a C-terminal portion which mediates binding to the nucleosome core (Nan *et al* 1999.) The methyl binding domain (MBD) (Nan *et al* 1996) contains around 70 residues and forms a wedge shaped α/β structure, in which four antiparallel β strands form one face of the wedge (Okhi *et al* 1999.) The longest two strands are thought to interact with the major groove, where the methyl group would be located (Okhi *et al* 1999, Wakefield *et al* 1999) (See fig 1.7.)

MeCP2 is localised to the CpG-rich heterochromatic pericentromic regions of mouse chromosomes (Nan *et al* 1999.) This localisation can be disrupted, becoming diffuse in mouse embryonic stem (ES) cells lacking Dnmt1 (Nan *et al* 1996) which have lower levels of

methylation; this implies that binding of MeCP2 to chromatin appears to be dependent on the presence of methylated DNA (Nan and Bird 2001.)

1.7.1. MECP2A and MECP2B: alternative splicing of MECP2.

MECP2 was originally thought to be a 3-exon gene, and was later found to have an extra, upstream non-coding exon (exon 1) (Reichwald *et al* 2000.) It was also known that the MECP2 transcript had two different polyadenylation sites which gave rise to two mRNA variants with different 3'UTRs of 2kb and 10 kb (Coy *et al* 1999.) However, two recent simultaneous reports have been published, by Kriaucionis *et al* and Mnatzakanian *et al*, which show that *Mecp2* undergoes alternative splicing to form two distinct isoforms with different N-termini (fig 1.10.) (Kriaucionis *et al* 2004, Mnatzakanian *et al* 2004.)

Fig 1.10. Alternative splicing generates two isoforms of MECP2.

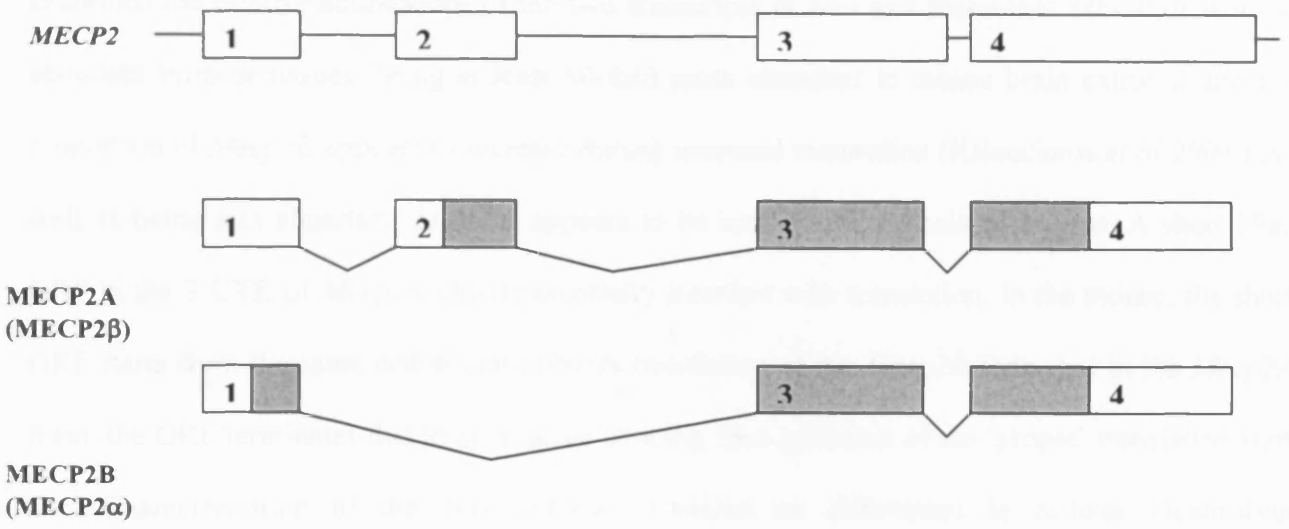


Fig 1.10. The exon/intron structure of MECP2.

Mecp2 undergoes alternative splicing to form 2 distinct isoforms; MECP2A *and* MECP2B.

Shaded regions indicate protein coding.

(from Kriaucionis et al 2004 and Mnatzakanian et al 2004.)

A search of EST databases showed that *MECP2* sequences grouped into those containing, and those without, the second exon (Kriaucionis *et al* 2004.) (Perhaps rather confusingly, Kriaucionis *et al* called their new transcript *MECP2α,* whereas Mnatzakanian *et al* called theirs *MECP2B* (Kriaucionis *et al* 2004, Mnatzakanian *et al* 2004.) The A/B notation will be used herein. Exon 2

normally contains the ATG START site, but if exon 1 is included, it contains an ATG which can act as the START site for an ORF of 501aa in mice and 498aa in humans (Kriaucionis *et al 2004.*)

Alignment of the two isoforms shows that *Mecp2B* is more closely related to ancestral forms of *Mecp2* such as those found in *X.leavis* and *D.rario* (Kriaucionis *et a 2004l*) Kriaucionis *et al* also examined the relative abundances of the two transcripts *in vivo* and found that *MECP2B* is more abundant in most tissues, being at least 10-fold more abundant in mouse brain extracts, and the proportion of *Mecp2B* appears to increase during neuronal maturation (Kriaucionis *et al 2004.*) As well as being less abundant, *Mecp2A* appears to be inefficiently translated *in vivo*. A short 39aa ORF in the 3'UTR of *Mecp2A* could potentially interfere with translation. In the mouse, the short ORF starts from the same codon that initiates translation of the *Mecp2B* form, but in the *Mecp2A* form, the ORF terminates due to alternative splicing 55nt upstream of the 'proper' translation start site. Characterisation of the two isoforms revealed no differences in cellular localisation (Kriaucionis *et al 2004.*)

The realisation that exon 1 does contain coding sequences could have major implications for Rett patient screening. Exon 1 has previously been excluded from many screening programs (which could explain why not all clinically diagnosed Rett patients have defined *MECP2* mutations.) (Kriaucionis *et al 2004.*) However, because *MECP2B* is the more abundant isoform, mutations in exon 1 could potentially remove over 90% of total MECP2 from the cell (Kriaucionis *et al 2004.*) Exon 1 contains repeated GCC and GCA codons which may be susceptible to expansion such as that seen in the FMR2 gene, causing FRAXE mental retardation (OMIM #309548) (Kriaucionis *et al* 2004.) The finding of the MECP2A isoforms may also explain why no exon 2 mutations are found in Rett patients; they may have such a mild effect that the MECP2B isoform can compensate for them (Stancheva *et al* 2003.)

A key question which remains unanswered is whether the two isoforms differ in function; for example, the MECP2B N-terminus contains a potential phosphorylation site which is absent from the MECP2A isoform, which could lead to functional differences (Kriaucionis *et al* 2004.) On the other hand, human MECP2A was able to rescue MeCP2 deficiency in *Xenopus* embryos, which argues against large functional differences, although one also could argue that it may well not rescue human cells (Kriaucionis *et al* 2004.)

1.7.2. MECP2 – Function *in vivo.*

Despite intensive effort by many groups around the world, the mechanism of action of MECP2 is still unknown. Very few target genes for MECP2 have been identified and the theory that MECP2 is somehow responsible for neuronal maturation has been strengthened, but in many ways, we are still far from identifying the precise molecular pathways which lead to such a devastating neurological phenotype, much less being at a stage of being able to offer therapy for Rett patients.

It was widely presumed that loss of MECP2 would lead to significant changes in a small number of genes that could be linked to the phenotype. Studies by Nan *et al* have showed that MeCP2 can repress transcription *in vitro* from methylated promoters, but not non-methylated promoters (Nan *et al* 1997.) Nan *et al* also found that MeCP2 appears to bind the Sin3A repressor complex, which acts to remodel chromatin into a state refractory to transcription via the recruitment of histone deacetylases (Hdacs) (fig 1.11) which deacteylate histone tails (Nan *et al* 1998.)

Fig 1.11.* MeCP2 *and Transcriptional Repression.

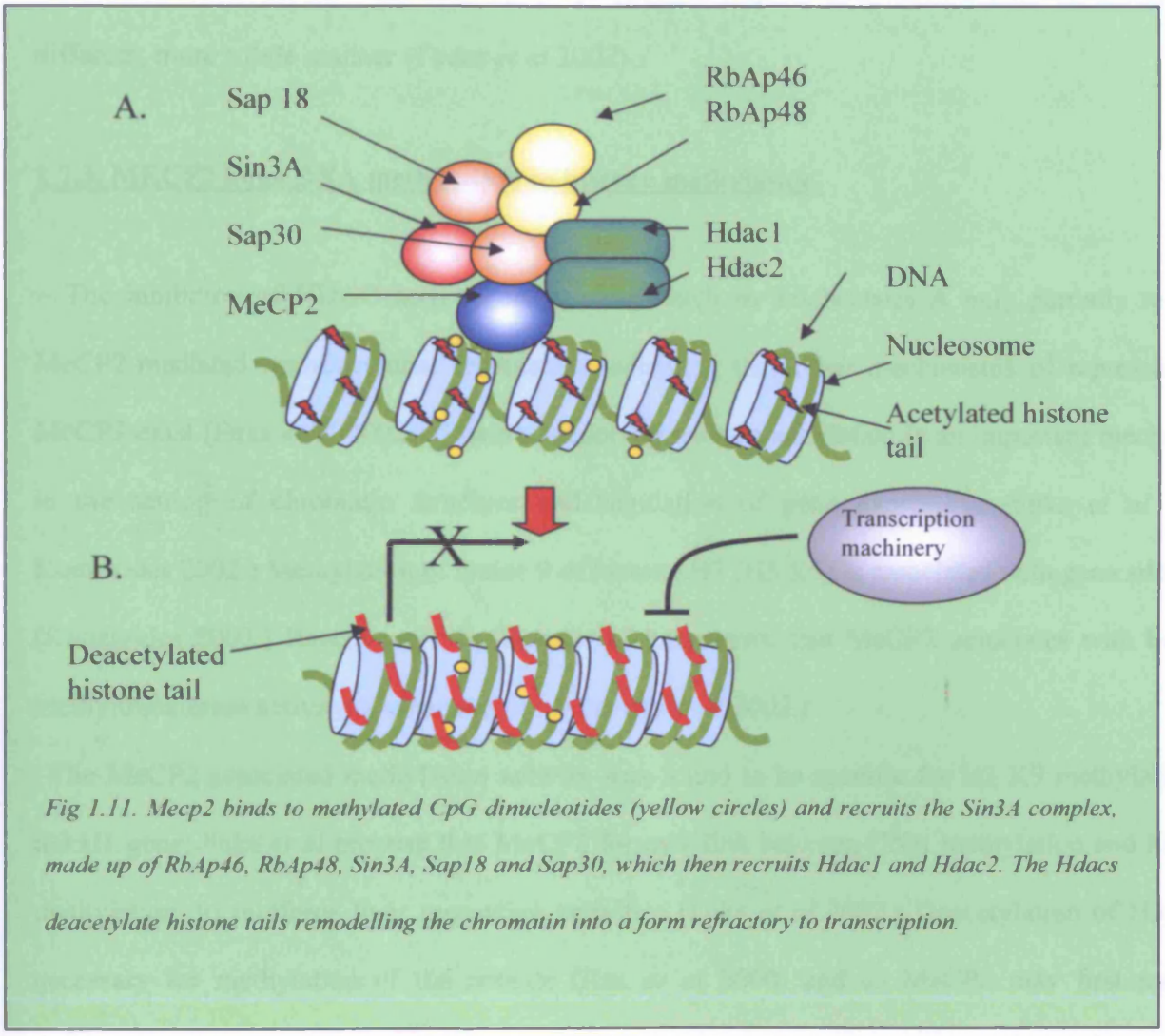


Fig 1.11. Mecp2 binds to methylated CpG dinucleotides (yellow circles) and recruits the Sin3A complex, made up of RbAp46, RbAp48, Sin3A, Sap18 and Sap30, which then recruits Hdac1 and Hdac2. The Hdacs deacetylate histone tails remodelling the chromatin into a form refractory to transcription.

Indeed, the transcriptional repression induced by MeCP2 *in vitro* can be relieved by Trichostatin A, a HDAC inhibitor (Nan *et al* 1999.) This evidence, and the fact that MeCP2 binds to a single symmetrically methylated CpG pair, which are widespread in the genome, lead Nan *et al* to hypothesise that MeCP2 is a global transcriptional repressor *in vivo* (Nan *et al* 1997.) As such, mutations in MeCP2 were expected to cause significant changes in the transcriptional profile

of the cell, which would lead to the symptoms of Rett syndrome. However, the lack of large transcriptional differences in the absence of MeCP2 implies that MeCP2 may be acting in a different, more subtle manner (Tudor *et al* 2002)

1.7.3. MECP2 links DNA methylation to histone methylation.

The inhibition of HDAC activity, using drugs such as Trichostatin A only partially relieves MeCP2-mediated transcriptional repression, indicating that other mechanisms of repression by MeCP2 exist (Fuks *et al* 2002.) Post-translational histone methylation is an important mechanism in the setting of chromatin structure and regulation of gene expression (Fuks *et al* 2002, Kouzarides 2002.) Methylation of lysine 9 of histone H3 (H3 K9) is associated with gene silencing (Kouzarides 2002.) Recent work by Fuks *et al* has shown that MeCP2 associates with histone methyltransferase activity *in vitro* and *in vivo (Fuks et al 2002.)*

The MeCP2 associated methylation activity was found to be specific for H3 K9 methylation of the H1 gene. Fuks *et al* propose that MeCP2 forms a link between DNA methylation and histone methylation, to reinforce their repressive activities (Fuks *et al* 2002.) Deacetylation of H3K9 is necessary for methylation of the residue (Rea *et al* 2000) and so MeCP2 may first mediate deacetylation, then methylation of H3 K9 which may in turn attract proteins such as HP1 (Fuks *et al* 2002) setting up a self-reinforcing cycle of repression.

For such a vital process, the lack of significant transcriptional differences seen in the absence of MeCP2 suggests that other factors may be able to partially compensate for the loss of MeCP2; perhaps the activity is mediated by a large protein complex which can use overlapping protein activities to make up for the loss of MeCP2.

1.7.4. Is MeCP2 Really a Global Transcriptional Repressor?

In vivo, it is thought that MeCP2 may exist as part of a larger chromatin-remodelling complex; proteins which have been reported to bind mammalian MeCP2 include Dnmt1, Co-Rest, Suv39H1 and c-Ski (Kimura and Shiota 2003, Kokura *et al* 2001, Lunyak *et al* 2002.) MeCP2 has been shown to interact with the Sin3A/Hdac (Nan *et al* 1998) chromatin remodelling complex, which remodels chromatin into a form refractory to transcription. In *X.leavis*, xMeCP2 was found to partially co-fractionate with Sin3A and transcriptional repression by MeCP2 can be partially alleviated by the Hdac inhibitor Trichostatin A (TSA) (Jones 1998.) However, native MeCP2 has recently been purified from rat brain and found to not bind stably to the Sin3A complex, or indeed any other proteins in nuclear extracts (Klose and Bird 2004.) Size exclusion chromatography of MeCP2 revealed that its apparent molecular weight was 400 to 500kDa, much higher than the predicted weight based on the amino acid sequence (52.4 to 53 kDa, depending on the species) (Klose and Bird 2004.) This finding could result from MeCP2 existing *in vivo* as a larger complex, as a protein with an unusual shape, or as a homo-multimer (Klose and Bird 2004.)

Klose and Bird used a combined size exclusion chromatography and sucrose gradient sedimentation to determine that MeCP2 appears to behave like an elongated monomer of approximately 6.15nm (Klose and Bird 2004.) Of course, these studies are carried out *in vitro*, and the properties of purified MeCP2 may well be different from MeCP2 which is interacting with DNA in a nucleoprotein context. Further work is needed to ascertain binding partners *in vivo* and to find out exactly what MeCP2 is interacting with in the cell.

The survival of *Mecp2*$^{-/y}$ and *Mecp2*$^{-/-}$ mice (Guy *et al* 2001, Chen *et al* 2001) and the lack of major developmental abnormalities in RTT patients demonstrates that MeCP2 is not absolutely

required for execution of the developmental process. However, the creation of mice null for MeCP2 confirmed that loss of the protein is necessary and sufficient for the development of Rett-like symptoms. Clearly, whilst not essential for viability, MeCP2 is an important part of at least some stage of the brain's developmental program.

With the advent of microarray technology, the transcriptome of a cell can be easily examined, and huge amounts of data generated in a way that was unthinkable only a few years ago. Since MeCP2 is thought to be a global transcriptional repressor, microarray analysis of cells with and lacking MeCP2 would be expected to reveal widespread and significant changes in the transcriptome. However, transcriptome analysis of *Mecp2*$^{-/-}$ brains showed very few differences, although computer programs were able to pick out wild type and mutant samples from their expression profiles (Tudor *et al* 2002.) Whatever the effects of removing MeCP2 in the brain, they clearly do not cause a lasting large alteration in the transcriptional activity of one or two specific genes, (which would allow the function of those genes, and thus the mechanisms behind Rett pathology, to be elucidated.) This may be due to the window of MeCP2 action being limited; the degenerative period in Rett is relatively short compared to lifespan and it may well be that analysis of brain tissue at one particular time point in the disease progression will yield larger transcriptional differences. It could also be that removal of MeCP2 results in a subtle transcriptional alteration which neurons may be very sensitive to; it has been suggested that MeCP2 damps down the inherent 'noise' in the system, of 'leaky' transcripts which neurons might be very sensitive to; this observation lead to the proposal that neurons may be more sensitive to transcriptional 'noise' (Carter and Segal 2001.) However, if neurons *per se* are more sensitive to spurious transcripts, this does not explain the differences seen in various brain regions; why would some brain regions, containing similar neuron types, react so differently? This is an intriguing idea; spurious transcription may be toxic to the cell at a threshold level determined by cell type, or

perhaps cell position. Experiments which have examined the transcriptome of the MeCP2 null brain (Tudor *et al* 2002) and found no significant changes warrant an examination of the role of MeCP2 as a global transcriptional repressor.

1.7.5. MeCP2 Directly regulates *Bdnf* and *xHairy2a.*

Despite initial microarray studies finding no significant changes in the MeCP2 null transcriptome, more recently, groups using different methodologies have found two genes to be controlled directly by MeCP2; *Bdnf* in the mouse, and *xHairy2a* in the frog (Chen *et al* 2003, Vetter 2003, Stancheva *et al* 2004, Martinowich *et al* 2003.)

1.7.6. MeCP2 and *Bdnf.*

Bdnf (brain derived neurotrophic factor) was the first gene found in the mouse to be regulated by MeCP2. Bdnf is synthesised in response to neuronal activity, and is thought to be involved in changing short-lived neuronal stimuli into more lasting alterations in the brain – a process thought to be at the root of memory and learning (West *et al* 2000.) Bdnf therefore plays a vital role in neuronal plasticity, learning and memory (West *et al* 2000.)

The *Bdnf* locus is controlled in a complex manner, with at least four transcriptional start sites that produce a range of mRNA products (Timmusk *et al* 1994.) In the rat the *Bdnf* exon III promoter (the equivalent promoter in the mouse is the *Bdnf* exon IV promoter) is activated in response to the wave of calcium that floods into the cell on depolarisation

When calcium is absent, the promoter is inactive, but when depolarisation leads to calcium uptake, the promoter is activated around 100-fold (Chen *et al* 2003, Tao *et al* 2002.)

In 2003, two groups independently discovered the link between *Bdnf* and MeCP2; Martinowich *et al* created *Bdnf* promoters with reduced methylation levels which they then transfected into cultured neurons (Martinowich *et al* 2003.) Upon depolarisation, reduced levels of transcription were seen in templates with lower methylation levels (Chen *et al* 2003, Martinowich *et al* 2003.) They then directly analysed the levels of methylation in cultured neuronal cells and found that there was a difference in methylation density at the exon IV promoter before and after depolarisation (Chen *et al* 2003, Martinowich *et al* 2003.) This implies that altered methylation states may be at least in part responsible for controlling *Bdnf* promoter IV. Martinowich *et al* went on to examine the proteins which associated with the *Bdnf* locus using chromatin immuoprecipitation (Martinowich *et al* 2003.)

These experiments showed that MeCP2 was associated with the locus in the absence of depolarisation; when depolarisation occurred, MeCP2 was largely lost from the locus (Chen *et al* 2003, Martinowich *et al* 2003.) Martinowich *et al* also reported a major redistribution of MeCP2 within the cell in response to depolarisation, becoming more punctate (Martinowich *et al* 2003.) MeCP2 was only lost from specific loci, as an imprinted gene, Rasgrf1, did not lose associated MeCP2 (Martinowich *et al* 2003.) This implies that depolarisation incurs both global and specific changes in MeCP2 distribution. Martinowich *et al* found that the CRE binding protein Creb was also associated with a CRE element in the *Bdnf* promoter and that on depolarisation, Creb had a higher affinity for the promoter (Martinowich *et al* 2003.) They predict that this creates a transcriptional activation complex (see fig 1.12) and that this acts in concert with the loss of the Hdacs to remodel chromatin into a more permissive form (Martinowich *et al* 2003.)

Fig 1.12. Possible mechanism of action of* MeCP2 *at Bdnf promoter IV in the mouse.

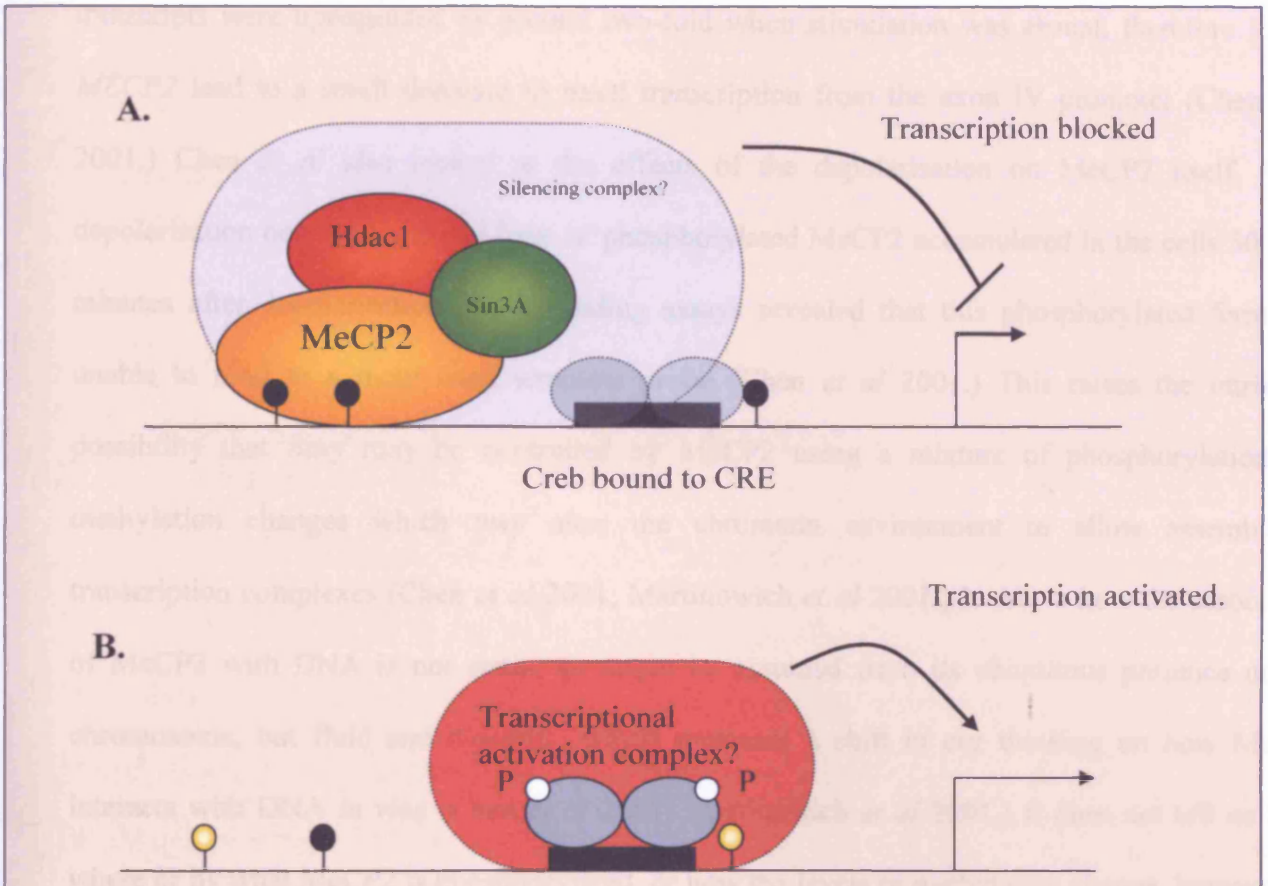


Fig.1.12. Possible mechanism of action of* MeCP2 *at Bdnf promoter IV in the mouse.

In the resting cell (A) the promoter is highly methylated (black circles) and MeCP2 *is bound to the promoter. CREB binds with lower affinity for the CRE element.* MeCP2 *recruits* Sin3A *and Hdacs and remodels the chromatin into a repressive state. Upon depolarisation (B) the promoter loses some methylation (yellow circles.)* MeCP2 *is phosphorylated and has lower affinity for the promoter. Creb becomes phosphorylated and becomes tightly bound to the CRE, possibly assembling a transcriptional activation complex. In the presence of this, and in the absence of the Mecp2-recruited Hdacs, the chromatin is remodelled into a permissive state. (Adapted from Matinowich et al 2001.)*

Chen *et al* also found that depolarisation of neurons leads to the dissociation of MeCP2 from the *Bdnf* exon III promoter (Chen *et al* 2001.) In neuronal cultures from *Mecp2* mutant mice, *Bdnf* transcripts were upregulated by around two-fold when stimulation was absent, therefore loss of *MECP2* lead to a small decrease in basal transcription from the exon IV promoter (Chen *et al* 2001.) Chen *et al* also looked at the effects of the depolarisation on MeCP2 itself. When depolarisation occurred, a novel form of phosphorylated MeCP2 accumulated in the cells 30 to 60 minutes after depolarisation. DNA binding assays revealed that this phosphorylated form was unable to bind to a methylated template probe (Chen *et al* 2001.) This raises the intriguing possibility that *Bdnf* may be controlled by MeCP2 using a mixture of phosphorylation and methylation changes which may alter the chromatin environment to allow assembly of transcription complexes (Chen *et al* 2001, Martinowich *et al* 2001.) It shows how the association of MeCP2 with DNA is not static, as might be assumed from its ubiquitous presence on the chromosome, but fluid and dynamic, which represent a shift in our thinking on how MeCP2 interacts with DNA *in vivo* (Chen *et al* 2001, Martinowich *et al* 2001.) It does not tell us how, where or by what MeCP2 is phosphorylated, or how the levels of methylation change, however, so there must be more protein players in the MeCP2/*Bdnf* story.

It is also noteworthy that the loss of MeCP2 does not lead to the same 100-fold increase in *Bdnf* transcription that occurs on depolarisation, as only a 2 fold-change is seen (Chen *et al* 2001.) This also indicates that there may well be more proteins involved in the MeCP2-mediated repression mechanism.

The *Xenopus* transcriptional repressor, *xHairy2a* has also been found to be a target gene for MeCP2 (Stancheva *et al* 2003) and MeCP2 appears to have a similar mode of action at the xHairy2a promoter; MeCP2 appears to be bound to the methylated upstream flank of the promoter

and leaves the promoter upon activation of the gene (Stancheva *et al* 2003, Klose and Bird . Hairy2a is a transcriptional repressor which represses the expression of pro-neuronal genes in the non-neuronal cells which surround developing neurons (Klose and Bird 2003, Stancheva *et al* 2004.) A deficiency of *Xenopus* MeCP2 leads to inappropriate activation of the x*Hairy2a* gene, with severe consequences for the embryonic nervous system i.e abnormal patterning of primary neurons during neuronal differentiation (Stancheva *et al* 2003.)

1.7.7. Is Rett syndrome a disorder of altered imprinting?

Although there is *in vivo* evidence that MeCP2 binds upstream of the differentially methylated region upstream of *H19* (Drewell *et al* 2002) most researchers have rejected the idea that Rett syndrome is a disorder of altered imprinting. A number of imprinted genes such as *H19*, Igf2, Snrpn, Ipw and Ndn have been examined (Balmer *et al* 2002) and normal imprinting patterns found in Rett patients. Recent work however, claims that Rett syndrome could be due to alterations in imprinting patterns (Horike *et al* 2005.) Horike *et al* examined whether MeCP2 was involved in repressing only one allele of an imprinted pair. They sequenced 100 MeCP2 binding sites identified by ChIP and found that several were contained in the imprinted cluster on mouse chromosome 6 that contains the *Dlx5* and *Dlx6* genes (Horike *et al* 2005.) *Dlx5* is a member of the Distal-less homeobox gene family and play roles in several developmental processes, including limb formation and neurogenesis (Panganiban *et al* 2002.)

In humans, DLX5 is maternally expressed; in mice however, this imprinting is 'relaxed' and there is biallelic expression, although expression from the maternal allele is higher (Horike *et al* 2005.) (This is an important caveat of this work – it is not clear how a fully imprinted and a

relaxed imprinting state compare *in vivo*.) Horike *et al* found that the CpG island of Dlx5 is, surprisingly, normally unmethylated and the MeCP2 binding sequence 52kb from the 3' end is not differentially methylated (Horike *et al* 2005.) In wild type mice, this region is enriched in histone H3 methylated at lysine 9 and is associated with a higher order loop characteristic of silenced chromatin (Horike *et al* 2005.) In the absence of MeCP2, this region still somehow loses its ability to bind *Hdac1* and becomes enriched in acetylated histone H3, losing the loop (Horike *et al* 2005.) This implies that these distinct loop structures may be induced by the presence of MeCP2 and may control longer-term transcriptional activity via a mechanism which is dependent on histone methylation rather than DNA methylation (Pescucci *et al* 2005) (Horike *et al* 2005.)

1.8. Aims and Objectives

1.8.1. Examining the role of MBD proteins in the murine intestine.

The methyl binding domain proteins 'read' and act upon the methylation marks created on DNA (Hendrich and Tweedie 2003.) The importance of these proteins is demonstrated by the effects of their removal on various systems; removal of MECP2 in the human causes Rett syndrome, which can be modelled in the mouse (Guy *et al* 2001, Chen *et al* 2001.) Removal of Mbd2 reduces intestinal tumourigenesis whilst Mbd4 appears to regulate the apoptotic response to a variety of DNA damaging agents (Sansom *et al* 2003a, Sansom *et al* 2004.) This demonstrates that MBD proteins are involved in a wide range of cellular processes.

This project will use a combined transgenic and biochemical approach to further characterise the roles of MeCP2 and Mbd4 in the murine intestine and the removal of MeCP2 in the mammary gland. A conditional allele of *Mecp2*, combined with an intestinal-specific or mammary-specific

Cre-recombinase will allow removal of MeCP2 only in the intestine or mammary gland, in order to study the effects of removal of the gene in a non-neuronal setting.

Studying the consequences of MeCP2 removal in a system other than the nervous system may be informative; changes in the proliferation, migration or differentiation in the intestine may be more easily scored than behavioural changes which occur as a result of MeCP2 removal in the brain or nervous system. In such a system as the intestine, it may be simpler to dissect out specific molecular pathways for the action of MeCP2 which may in turn shed light on the the mechanisms underlying Rett syndrome pathology.

Constitutive null alleles of *Mbd4* and *Mlh1* will be used to investigate the role of Mbd4 and Mlh1 in apoptotic signalling in response to Fas ligand and anoikis, and the interaction of Mbd4 with the MMR system. These experiments will be designed to shed light on the roles of MBD proteins in the intestine and mammary gland, and provide insights into the elusive mechanism of action of MECP2 in Rett syndrome.

Chapter 2. Materials and Methods.

2.1. Immunohistochemistry/stains.

- All procedures use sections obtained from formalin-fixed, paraffin-embedded samples, at a thickness of 5-10μm
- Sections must not be allowed to dry out at any point in the procedure.
- Xylene and ethanol clearing/ dehydration steps must be carried out in a fume hood.
- Incubations are carried out in a humidified slide box, at room temperature unless otherwise stated.

2.1.1 H&E (Haematoxylin and Eosin) staining

Clearing.

- Wash slides 2 x 10 mins in xylene
- Wash 2 x 5 mins in 100% ethanol
- Wash 1 x 5 mins in 95% ethanol
- Wash 1 x 5 mins in 75% ethanol
- Wash in tap water 1x 2 mins.

- Immerse slides in haematoxylin for 1 min
- Rinse in tap water (with a continuous flow-through) for 5 mins
- Immerse in eosin for 30 seconds
- Rinse in tap water for 5 mins

Dehydration

- Wash slides 1 x 5 mins in 75% ethanol
- Wash 1 x 5 mins in 95% ethanol
- Wash 2 x 5 mins in 100% ethanol
- Wash 2 x 10 mins in xylene
- Mount in DPX (Distrene plasticiser and xylene) (Sigma) Place the coverslip on a flat surface and apply a small amount of DPX, making sure there are no bubbles. Pick up the slide and gently place it, section side down, onto the coated coverslip. Turn the assembly over and gently press down on the coverslip with a blunt pencil until the DPX spreads out to cover the whole section and the whole of the underneath of the coverslip.
- Leave on a level surface to dry for 48 hours.

2.1.2. GIP (Gastrointestinal polypeptide) staining.

Use Dako kit # K4006 or K4007 and GIP antibody (Monosan)

Clearing.

- Wash slides in 2 x 10 mins in xylene
- Wash 2 x 5 mins in 100% ethanol
- Wash 1 x 5 mins in 95% ethanol
- Wash 1 x 5 mins in 75% ethanol

Antigen retrieval

- Immerse the slides in preheated 1x citrate buffer (Sigma) at 99°C for 20 mins – preheat the buffer by placing the slide container in a cool water bath and bringing the container and buffer up to 99°C. Do not place a cold slide container/glass slides into hot buffer/water bath as they are likely to shatter.
- Remove the slide container from the water bath using gloves, and allow to cool to room temperature for ½ hour.

Peroxidase block

- Draw around the section with a DAKO immunohistochemistry pen, to leave a hydrophobic line around the section – this keeps the liquid on the section and prevents it from drying out.
- Apply enough peroxidase block from bottle 1(supplied with kit) to cover the section.
- Incubate in a humidified chamber for 4-6 mins.
- Gently rinse the section with buffer (1xPBS – phosphate buffered saline (Sigma))
- Immerse the sections in buffer

Primary antibody

- Tap off the excess buffer
- Apply the primary antibody (diluted 1:400 in buffer.) Use just enough to cover each section.
- Incubate for 50 mins at room temp.
- Rinse with buffer.
- Tap off excess buffer - do not allow the sections to dry out.
- Apply enough labelled polymer from bottle 2 (supplied with kit) to cover the section.
- Incubate for 30 mins
- Rinse with buffer.

Substrate chromogen.

- Add one drop of bottle 3b to one ml of bottle 3a, (supplied with kit) mix and use immediately.
- Apply enough of the mixture to just cover each section.
- Incubate for 5-10 mins
- Rinse gently with buffer.

Counterstain.

- Immerse the slides in Haematoxylin for 30 sec to 1 min.
- Rinse the slides in distilled water
- Dip each slide 10 times into a bath of 37mM ammonia
- Rinse the slides in deionised water for 10 mins
- Dehydrate and clear (by immersion in increasing concentrations of ethanol followed by xylene) and mount in DPX (as in protocol 2.1.1.)

GIP positive cells appear dark brown, background appears blue/brown.

2.1.3.Grimelius (Enteroendocrine cell stain.)

Solutions needed:

1. **Silver staining solution** (100 ml aceteate buffer pH 5.6, 87ml distilled water, 3ml 1% silver nitrate solution.)
2. **Reducing solution** (2.5g hydrated sodium sulphite, 1g hydroquinone, 50ml distilled water.)
3. **Acetate buffer** (4.8ml 0.2M acetic acid, 45.2 ml 0.2Msodium acetate, 50ml distilled water.)

All solutions must be made up in glass containers only, which must be thoroughly rinsed with distilled water. Contact with plastic will result in the stain failing.

De-waxing and Rehydration.

- Wash the slides for 2 x 10 mins in xylene
- Wash 2 x 5 mins in 100% ethanol
- Wash 1 x 5 mins in 95% ethanol
- Wash 1 x 5 mins in 75% ethanol

- Wash the slides for 5 mins in tap water.
- Incubate for 3 hours at 65°C in silver nitrate solution – bring the container with the slides and silver nitrate solution up to 65°C in a water bath, do not place a cold container into a hot water bath or the container and slides may shatter.
- Remove the container carefully using heatproof gloves.
- Drain the silver solution from the slides
- Treat with freshly prepared reducing solution for 1 min at 45°C.
- Repeat steps 4 and 5 until the desired effect is achieved (typically 3-5 repeats.)
- Dehydrate, clear and mount in DPX, as in protocol 2.1.1.
- Enteoendocrine cells will appear black with black granules, the background appears yellow/brown.

2.1.4. Alcian Blue.

Solutions needed;

1. **Alcian blue pH 2.5** (2g alcian blue (or as needed) 100ml 0.5% acetic acid.)
2. **0.1% nuclear fast red** (0.1g nuclear fast red, 2.5g ammonium sulphate, 100ml distilled water.)

Method.

- Wash 2 x 10 mins in xylene
- Wash 2 x 5 mins in 100% ethanol
- Wash 1 x 5 mins in 95% ethanol
- Wash 1 x 5 mins in 75% ethanol
- Wash (and transport if necessary) in tap water.
- Stain in alcian blue solution for 5 mins
- Wash well in tap water for 5 mins
- Counterstain in 0.1% nuclear fast red for 5 min
- Wash well in tap water

- Dehyrate, clear and mount in DPX.
- Mucins appear bright blue, nuclei appear red. Background appears red/pink.

2.1.5.Alkaline phosphatase.

- Alklaline phosphatase staining for the villus was carried out on paraffin sections of intestine, using Dako kit #KO699, according to the supplied instructions. (http://www.dakocytomation.co.uk/prod_downloadpackageinsert.pdf?objectid=100608001)

2.1.6. BrdU (Bromodeoxyuridine staining)

- Wash the slides for 2 x 10 mins in xylene
- Wash 2 x 5 mins in 100% ethanol
- Wash 1 x 5 mins in 95% ethanol
- Wash 1 x 5 mins in 75% ethanol
- Wash for 5 mins in tap water.
- Add 5ml of antigen unmasking solution (DAKO to 500mls of DDW
- Place in the microwave and microwave on full power for 3 x 5 mins for antigen unmasking. Shake the container carefully between 5 minute incubations to remove bubbles.
- Allow the slides and unmasking solution to cool for 30 mins.
- Place the slides in 5m HCl for 10 mins.
- Remove the slides and wash for 2 x 5 mins in PBS.
- Draw around the section with a DAKO pen.

- Block non-specific staining by incubating in PBS/BSA (Bovine serum album 1%) for 20 mins.
- Apply a 1:50 dilution of BrdU antibody (MCA2060 Serotec) in 1% BSA/PBS.
- Incubate for 1 hour.
- Wash the slides 3 x 5 mins in PBS.
- Develop in Dab chromogen (Sigma) (2 drops of bottle A from the Dab kit to 1 ml of bottle B from the kit) for 5-8 mins until a dark brown colour develops
- Wash the slides for 5 mins in PBS
- Counterstain with a 30 second immersion in haematoxylin.
- Wash for 5 mins in tap water
- Dehydrate and mount in DPX as in protocol 2.1.1.
- BrdU positive cells appear dark brown/black, background appears blue.

2.2. Genotyping.

2.2.1. DNA Purification: Puregene method

- Take a small (3-5mm) section of mouse tail tip from the dead animal (or use appropriate anaesthetic to take a sample from a live animal.)
- Add 500µl of cell lysis solution (Gentra) and 10µl of proteinase K 20mg/ml to each tube containing mouse tail and shake overnight at 37oC
- Allow the tubes to cool to room temp
- Add 200µl of protein precipitation solution (Gentra) to each tube. Vortex and then spin in microfuge at 13,000rpm for 5 mins
- Remove the supernatant into clean tube containing 500µl of isopropanol.
- Vortex and spin at 13,000 rpm for 5 mins
- Pour off the supernatant and leave the tubes upside down to dry overnight.

- Add 500µl of nuclease-free water to each sample.
- Vortex gently for 2 seconds and leave at 37°C shaking overnight to resuspend the DNA.
- Use 2µl of the resulting DNA suspension in the following PCR protocols.

2.2.2. *Mecp2* PCR.

Primers p5 (forward) 5' **GGT AAA GAC CCA TGT GAC CC** 3'

p7 (reverse) 5' **GGC TTG CCA CAT GAC AA** 3'

For one reaction....

- 5µl 10 x Buffer (Sigma)
- 3µl 25mM $MgCl_2$ (Sigma)
- 1µl 10mM DNTP mix
- 1µl of each primer
- 2µl RedTaq (Sigma)
- 35µl distilled H2O
- 2µl DNA.

PCR Cycle.

- 95°C for 3 mins
- 95°C for 30 sec
- 56°C for 1 min
- 72°C for 1 min

x 30

- 72°C for 5 mins
- END

The PCR products are then run on 2% agarose gel containing ethidium bromide (4µl per 100ml gel) and bands visualised under UV light. Bands appear around 416 bp (wild type) and 470 bp (floxed allele)

2.2.3. *Blg-Cre* PCR.

Primers. *Blg* forward; CTTCTGGGGTCTACCAGGA

Blg reverse: TCGTGCTTCTGAGCTCTGCA

For one reaction...

- 5µl 10 x Buffer (Sigma)
- 3µl 25mM $MgCl_2$ (Sigma)
- 1µl 10 mM DNTP mix (Sigma)
- 1µl of each primer
- 2µl RedTaq (Sigma)
- 35µl distilled H_2O
- 2µl DNA.

PCR Cycle.

- 95°C for 3 mins
- 95°C for 30 sec ←
- 55°C for 1 min | x 30
- 72°C for 1 min —
- 72°C for 5 mins
- END

The PCR products are then run on 2% agarose gel containing ethidium bromide (4µl per 100ml gel) and bands visualised under UV light.

2.2.4. *Ah-Cre* PCR

Primers: P1= TGACCGTACACCAAAATTTG
P2= ATTGCCCCTGTTTCACTATG

For one reaction..

- 5µl 10 x Buffer (Sigma)
- 3µl 25mM $MgCl_2$ (Sigma)
- 1µl 10 mM DNTP mix (Sigma)
- 1µl of each primer
- 2µl RedTaq (Sigma)
- 35µl distilled H_2O
- 2µl DNA.

PCR program.

- 94°C for 2 mins
- 94°C for 1 min ←┐
- 55°C for 1 min ├ x 35
- 72°C for 1 min ←┘
- 72°C for 5 mins
- END

Run the PCR products on a 2% agarose gel containing ethidium bromide. The AH-cre appears as a band just below 1kb.

2.2.5. *Mbd4* PCR

Primers: P1: CGT GTG GAT GGG AAA GAG TT

P2: GGA AGT CAG AGC TGC AAA C

For one reaction...

- 35.25 µl distilled H_2O
- 5 µl Buffer
- 1 µl dNTPs
- 1 µl Primers (x4)
- 2 .5 µl Mg
- 0.25µl Taq
- 2 µl DNA

- PCR Program:

- 94°C for 2 mins
- 45 sec 94°C
- 45 sec 65°C
- 1 min 72°C
- 5 min 72°C

- The PCR products are the run on a 4% agarose gel. Product: wt = 322 bp. KO (knockout) = 469 bp

2.2.6. *Mlh1* PCR

Primers P1: GAC AGC ACC AGA CCA AGC TA

P2: AGG ACT GTC TAA GGC AGC A

For 1 reaction....

- 5µl 10x PCR buffer
- 2µl of MgCl2
- 1µl P1 (40 pmols per reaction)
- 1µl P2 (40pmol per reactiom)
- 1µl P3 (40pmol per reaction)
- 4 µl dNTPS (undiluted)
- 0.25 µl of TaQ polymerase
- 393µl of distilled H_2O

PCR protocol.

- 94°C 1 min
- 94°c 1 min
- 60°C 1 min
- 72 °C for 1 min
- 72°C for 5 mins

- Run the PCR products on a 2% agarose gel, the wild type band = 258bp, mutant band = 198bp

2.2.7. *Apc*^{flox} PCR

Primers;

- *Apc-P3* GTTCTGTATCATGGAAAGATAGGTGGTC

- *Apc-P4* CACTCAAAACGCTTTTGAGGGTTGATTC

- Reaction x 1 5µl 10x Buffer
    - 2.5µl 50mM $MgCl_2$
    - 1µl 10mM dNTP MIX
    - 1µl both primers (10pmol/µl; 10µM)
    - 0.25µl Taq
    - 37.25µl distilled H_2O
    - 2µl DNA

- PCR Cycle 95°C 3 minutes
    - 95°C 30 secs
    - 60°C 30 secs
    - 72°C 1 min
    - 72°C 5 mins

- The products are run on a 2% agarose gel. The wild type band appears at 296bp and the floxed allele at 314 bp.

2.2.8. *LacZ (Rosa26R)* PCR.

Primers;

- LacZ1 TACCACAGCGGATGGTTCGG
- LacZ2 GTGGTGGTTATGCCGATCGC

- Reaction x 1 5µl 10x Buffer
    - 2.5µl 50mM $MgCl_2$
    - 1µl 10mM dNTP MIX
    - 1µl both primers (40pmol/µl; 40µM)
    - 0.25µl Taq

- 37.25µl dH_2O
- 2µl DNA

- PCR Cycle 95°C 3 minutes
  - 95°C 30 secs
  - 55°C 30 secs
  - 72°C 40 secs

  x 35.
  - 72°C 5 mins

- Run the PCR product on a 2% agarose gel, a band will appear at 351bp if the allele is present.

2.3. LacZ Staining of small intestinal wholemounts.

- Immediately after killing the mouse, remove the intestinal tract and flush with ice-cold PBS.
- Flush the gut with X-gal fix (2% formaldehyde, 0.1%glutaraldehyde in PBS kept on ice)
- Cut the gut into 7cm sections
- Pin down on wax plate with the mesenteric line uppermost, trim off any excess mesentery.
- Pin down the other end of the gut so that it is stretched slightly. Begin opening the gut and pin out.
- Fix in 10% formalin for 1 hr.
- Wash with PBS
- Demucify by incubation in DTT demucifying solution for 30 mins.
- Flood the plate and pipette off the excess mucus.
- Add approx. 40 mls X-gal stain (1ml of solution A to 100ml solution B.)
- Solution A = 2% X-gal in DMF (Promega)
- Solution B= 0.1g MgCl, 0.48g K ferricyanide, 0.64g K ferrocyanide in 500mls

PBS.

- Incubate overnight in the dark at room temperature with gentle agitation.
- Fix with formalin or X-gal fix.

2.3.1. DTT Demucifying Solution.

- 1 vol glycerol (5mls)
- 1 vol 0.1M Tris pH 8.2 (5 mls)
- 2 vols 100% ethanol (10mls)
- 6 vols saline (30ml)
- 170mgs DTT per 50 mls solution.

2.3.2. Wax Plates.

- Melt Ralwax (BDH# 36154) in a pyrex beaker over a bunsen flame. Once melted add 0.1vols of paraffin oil. Pour into plates and allow to cool.

2.3.3. Mammary Wholemounts.

- Remove the entire mammary gland from the mouse
- Place on a glass slide and gently flatten the gland as much as possible.
- Place the slide into freshly made up 2% paraformaldehyde and leave for at least 2 hours, and preferably overnight.
- Remove the gland from the slide and rinse in PBS and then distilled water (at least 5 mins)
- Place the gland in carmine red overnight. (1g carmine red dye (Sigma) +2.5g potassium alum in 500ml distilled water.)
- Dehydrate the gland as follows;

- 15 mins in 75% ethanol
- 15 mins in 95% ethanol
- 2x15 mins in 100% ethanol
- Clear the gland in benzyl benzoate overnight
- Mount the gland on a glass slide by placing it on the slide, coveing it in a minimal amount of DPX and placing a coverslip over it.
- Photograph immediately.

2.3.4. 2% Paraformaldehyde.

- Makes 500ml.

- Dissolve 1g of paraformaldehyde (Sigma) in 500ml distilled water.
- Gently heat (on a hot plate set to around 50°C) and stir and add NaOH dropwise until the solution is clear.
- Use immediately or freeze aliquots and store at -20°C)

2.4. Scoring Apoptosis.

Apoptosis was scored in H&E stained paraffin sections. Apoptotic cells can be identified by their distinctive morphology. Apoptotic cells are rounded and stain redder than surrounding cells, and are often surrounded by a lighter 'halo'. Chromatin can be seen to be condensed within the cell, and there may be membrane blebbing.

2.5. Preparation of Murine Colonic / Small Intestinal Crypts.

- Immediately after culling the mouse, remove whole colon/small intestine.
- Flush through with water and open up the intestine along the long axis.

- Rinse in HANKS balanced salt solution (HBSS, Sigma) at room temp.
- Shake gently in 10ml of 2mM EDTA (pH8) in HBSS for 20 mins at 37°C.
- Discard the supernatant.
- Vortex the remaining material for around 20 seconds in 10ml HBSS and transfer the supernatant containing complete crypts (and some single cells) into a 15ml tube
- Repeat the vortexing until the supernatant is almost clear.
- Allow the crypts to settle out for 5 mins at room temperature
- Pipette (using a glass pipette) some of the crypts into 5 small microfuge tubes.
- Label the tubes with the appropriate timepoints.
- Spin the tubes at 500rpm for two minutes.
- Discard the supernatant and wash the pellet twice with Dulbecco's-MEM (DMEM, Sigma)
- Add 1ml DMEM and resuspend the by gently inverting the tubes a few times and incubate at 37°C in 5%CO2.

2.5.1. Short-Term Culture of Colonic Crypts.

- Transfer the isolated crypts to culture dishes (inserts coated with 1% heat-denatured BSA overnight at 4°C or collagen)
- Incubate at 37°C, 5% CO_2 for the time needed.
- Take the tubes at the relevant timepoints, spin down gently and replace the supernatant with 10% formalin.
- Resuspend the crypts by gently inverting the tube.
- Spin down (3 mins at 500rpm) and replace the supernatant with fresh formalin.
- Embed in paraffin, section and stain with H&E.

2.6. Cell Culture.

2.6.1. Preparation of murine embryonic kidney cells.

Solutions needed: 15ml PBS + 37.5mg trypsin +37.5μg collagenase A, (Sigma) filter sterilised.
PBS, filter sterilised.
DMEM +5% FBS (Foetal bovine serum) (Sigma)
DMEM +5% FBS +0.1g gentamicin. (Sigma)

All solutions are 0.2μm filter sterilised.

- Cull pregnant female at day E19.
- Remove embryos and kill by decapitation after immersion in ice-cold PBS.
- Remove kidneys and process each one separately.
- Dip each kidney briefly in 70% ethanol and then wash/transport in PBS.
- Place into 5ml of PBS +trypsin/collagen solution.
- Mechanically disrupt the kidney, mince finely using crossed sterile scalpel blades.
- Shake or stir gently for 30 mins at 37°C.
- Add 5 ml of the DMEM+5%FBS solution and gently mix by pipetting.
- Allow the clumps to settle.
- Take off the supernatant and spin for 5 mins at 500rpm.
- Resuspend the pellet in DMEM/FBS/gentamicin
- Incubate at 37°C in 5%CO_2.
- Check, feed and wash as necessary.

2.6.2. To trypsinise the cells (for passaging)

Solutions needed.

- Sterile PBS
- PBS+0.05% trypsin+0.053mM EDTA
- 47.5 ml DMEM + 2.5mls FBS + 2.5mg gentamicin

All solutions are warmed to 37°C

- Aspirate the medium and discard it.

- Wash cells gently with PBS, aspirate and discard.
- Repeat.
- Add enough 1xtrypsin/EDTA solution to cover the cells, then gently rock the culture vessel 4-5 times to coat the monolayer
- Watch the cells until around 75% have rounded up and detached.
- Resuspend the cells in culture medium
- Centrifuge at 500rpm for 4 mins.
- Resuspend in medium and transfer to a new culture dish.

2.7. RNA extraction and preparation for Affymetrix Microarray.

2.7.1. RNA Extraction From Intestinal Tissue.

Reagents.

- RNeasy mini-kit Qiagen cat no. 74104
- TRIzol reagent Invitrogen cat no. 15596-018
- Superscript Double-stranded cDNA synthesis kit (custom) (10 reactions) Invitrogen cat no. 11917-010
- T7(dT_{24}) primer . GENSET Custom Primer. HPLC-purified supplied by Helena Biosciences.
- Sequence: 5'-GGCCAGTGAATTGTAATACGACTCACTATAGGGAGGCGG-(dT_{24})-3'
- Phase Lock Gel (light) tubes. Helena Biosciences Cat. No. 0032-007961
- Enzo BioArray High Yield Transcript Labelling kit supplied by Affymetrix (10 reactions) cat no. 900182
- Trizma base. Sigma cat no. T1503
- Magnesium Acetate . Sigma cat no. M2545
- Potassium acetate. Sigma cat no. P5708

- 7.5M Ammonium acetate. Sigma cat no. A2708
- β-mercaptoethanol . Sigma cat no. M3148
- 0.5M EDTA Sigma cat. No. E7889
- Phenol/chloroform/IAA. Ambion cat no. 9730
- Glycogen (5mg/ml) Ambion cat no. 9510
- 1Kb plus DNA ladder Invitrogen cat no. 10787-018
- 0.24-9.5Kb RNA ladder Invitrogen cat no. 15620-010
- RNA century size markers Ambion cat no. 7140
- RNAlater Sigma cat no. R0901

2.7.2. Fragmentation buffer

(5X) fragmentation buffer: (200mM Tris-acetate, pH 8.1, 500mM KOAc, 150mM MgOAc)

The fragmentation buffer is made using RNase-free reagents. Tris-containing solutions should NOT be treated with DEPC. However, once dH_20 has been DEPC treated and autoclaved, it can be used for making Tris solutions.

Buffer: 4ml 1M Tris acetate pH.1 (Trizma base, pH adjusted with glacial acetic acid.)
0.64g magnesium acetate
0.98g potassium acetate
DEPC treated dH_20 to 20 ml.

2.7.3. Obtaining RNA from Tissue Samples.

- Dissect out the tissue of interest to a maximum thickness in one direction of 5mm. (For the intestinal MeCP2 array, the very top 3cm of the proximal small intestine was used.)
- Transfer immediately to cold RNA later™ (Sigma cat no. R0901) to inactivate endogenous RNase activity.
- Tissue stored in RNA later™ can keep for 1 day at 37°C, 1 week at 25°C, 1 month at 4°C and can be archived indefinitely at -20°C.

To process the tissue;

- Pour off the RNA later™ and note the wet weight of the tissue.
- Add 1ml cold Trizol reagent (Gibco cat no. 15596-018) per 100mg tissue.
- Homogenise the tissue.
- Clean the homogeniser before first use, between each use and after final use with a 15 second pulse in 5ml 4M NaOH followed by a 15 second pulse in 5ml H20 for 15 sec, then a 15 second pulse in 5 ml 75% ethanol (all solutions made with nuclease-free water)
- Aliquot 1 ml samples into clean, labelled RNase-free centrifuge tubes. Keep on ice.
- Centrifuge for 10 mins at max speed in a bench-top centrifuge at 4°C.
- Take off the supernatant (avoid the surface layer of fat) and transfer to new, labelled, sterile RNase-free tubes.
- Add 200µl chloroform, shake for 5-10 mins, then chill for 5-10 mins.
- Centrifuge samples for 15mins at max. speed (bench-top centrifuge at 4°C)
- Carefully remove the top aqueous layer which will contain the RNA and transfer to new tubes. Avoid taking material from the interface, which will contain DNA. Dispose of the lower layer in phenol waste.
- To precipitate the RNA, add an equal volume of isopropanol, mix by inversion and leave for 10 mins at 4°C. Leaving the tubes for longer may maximise RNA yields.
- Centrifuge samples for 15 mins at max. speed (bench-top centrifuge at 4°C)
- Discard the supernatant and air-dry the pellet for 5-10 mins.
- Resuspend the RNA in 100µl (or more if needed) nuclease-free water.
- Determine the total yield of RNA spectrophotometrically. 1 A260 unit = 40µg/ml of single-stranded RNA.

2.7.4. First Strand cDNA Synthesis.

- Place the following in a microfuge tube;

10μg (Xμl) total RNA in DEPC (Diethyl pyrocarbonate, Sigma)-treated water.

1μl T_7 (dt_{24}) primer (100pmol/μl)

Yμl DEPC-treated water

= 11μl total.

- Incubate at 65-70°C in a heated block for 10 mins.
- Place tubes on ice
- Prepare a master mix on ice. Quantities per reaction are;

4μl (5x) first strand buffer (thaw at 37°C then keep on ice)

2μl (0.1M) DTT

1μl dNTPs (10mM)

=7μl total per reaction.

- Add 7μl mix to each tube.
- Incubate at 42°C for 2 mins
- Add 2μl superscript II reverse transcriptase
- Incubate for 1 hr on a heated block at 42°C.

- Place the tubes on ice and proceed to the second strand synthesis or freeze on dry ice and store at -80°C.

2.7.5. Second Strand Synthesis.

NB.Use EXACTLY 2U RNase H per reaction

- Place all the reaction tubes on ice. For each reaction use:

91μl DEPC treated water

30μl (5x) second strand buffer

3μl (10mM) dNTPs

1μl *E.coli* DNA ligase (10U/μl)

4µl *E.coli* DNA Pol I (10U/µl)

1µl *E.coli* RNase H (2 U/µl)

=130µl per reaction total.

- Mix by vortexing briefly, centrifuge down briefly if necessary.
- Add 130µl of master mix to each of the first strand reaction tubes.
- Mix by pipetting, centrifuge down briefly if necessary.
- Incubate at 16°C for 2 hrs.
- Add 2µl T4 DNA polymerase.
- Incubate at 16°C for 5 mins.
- Place tubes on ice and proceed to 'clean-up of double-stranded cDNA or freeze on dry ice and store at -80°C.

2.7.5. Clean-Up of Double-Stranded cDNA.

- Centrifuge phase-lock tubes for 30 sec at maximum speed.
- Add an equal volume of RT buffer saturated Phenol/chloroform/IAA to the cDNA reaction (≈160µl)
- Vortex briefly to mix
- Place the mixture in a Phase-lock tube
- Centrifuge at maximum speed for 2 mins
- Transfer the upper (aqueous) phase to a new eppendorf tube.

- Add: 0.5 vols 7.5M Ammonium acetate (≈80µl)

  4µl glycogen

  2.5 vols 100% ethanol (room temperature) (≈600µl)

- Mix by tapping the tube
- Centrifuge at maximum speed for 20 mins – NO LONGER.
- Remove the supernatant, being careful not to disturb or discard the pellet.

- Add 160µl cold 80% ethanol.
- Centrifuge at room temperature for 5 mins at maximum speed.
- Discard the supernatant.
- Repeat the cold 80% ethanol wash, removing all the ethanol carefully with a pipette.
- Allow the pellet to air-dry for 5-10 mins.
- Resuspend in 12µl of DEPC-treated water.
- Proceed to 'in-vitro transcription' or freeze on dry ice and store at -80°C.

2.7.6. *In Vitro* Transcription.

- Use reagents from the Enzo Bio-array high yield transcript labelling kit.
- Once thawed, reagents should be kept at room temperature until the reaction is ready to be incubated to reduce the precipitation of DTT.
- The reaction should be kept in an incubator/warm room to reduce condensation on the inside of the tube cap.
- Thaw all reagents and the dsDNA at room temperature.
- Prepare a master mix;

10µl DEPC-treated water
4µl (10x) HY reaction buffer
4µl biotin-labelled ribonucleotides.
4µl DTT
4µl RNase inhibitor mix
2µl T7 RNA polymerase.

= 28 µl total per reaction

- Add 28 µl of master mix to each of the 12µl samples of cDNA.
- Mix by pipetting, spin briefly if necessary
- Incubate at 37°C for 5 hrs. Gently mix the reaction every hour and spin briefly if necessary to remove condensation.
- Proceed to 'clean-up of cRNA' or freeze on dry ice and store at -80°C.

2.7.7. Clean-Up of cRNA

- Do not touch the membrane of the spin columns
- Centrifuge at 10,000 rpm.
- The capacity of the RNeasy™ mini-column is 100μg. If the expected yield is more than 100μg, split the reaction and use 2 columns
- To the *in vitro* transcription reactions, add 60μl DEPC-treated water and 350 μl RLT buffer.
- Mix by pipetting
- Add 250μl (100%) ethanol.
- Apply to RNeasy™ mini-column, centrifuge for 15 sec.
- Empty collection tube.
- Add 500μl RPE buffer
- Spin at for 15 sec.
- Empty collection tube
- Add 500μl RPE buffer
- Spin for 2 mins
- Transfer spin column to a fresh collection tube
- Spin for 1 min
- Transfer spin column to a new collection tube.
- Elute cRNA by placing 50μl DEPC-treated water on the middle of the membrane and leave at room temperature for four minutes.
- Spin for 1 min
- Take the eluate and reapply to the same spin column
- Leave at room temperature for 4 mins
- Spin for 1 min.
- Determine the concentration of the cRNA spectrophotometrically. Measure the absorbance of an aliquot of the eluate at 260nm and 280nm. The aliquot must be diluted in dH_20 (not DEPC treated)
- Ensure the sample is diluted sufficiently to give a reading in the linear range for the spectrometer (usually in the range of 0.1-0.4 absorbance units.) Calculate the

concentration of RNA given that an A260 reading of 1 is equivalent to 40μg/ml of single-stranded RNA.

- The expected yield of cRNA starting from 10μg total RNA is 40μg to 110μg.
- If you have less than 40μg of cRNA per reation then do not use the target as there is likely to be a problem with the cRNA preparation.
- The cRNA fragmentation procedure requires the RNA to be at a minimum concentration of 0.6μg/μl. if necessary the RNA can be concentrated by ethanol precipitation.

2.7.8. Ethanol Precipitation.

- To the reactions add: 0.5 vols 7.5M ammonium acetate.
  2.5 vols **cold** 100% ethanol
- Vortex to mix thoroughly.
- Precipitate at -20°C for 1 hour or overnight.
- Centrifuge at full speed for 30 mins at 4°C.
- Wash the pellet twice with 0.5ml **cold** 80% ethanol
- Air dry the pellet and resuspend in an appropriate volume of DEPC-treated water.

2.7.9. cRNA Fragmentation.

- Fragment 25μg of cRNA (20μg are needed for hybridisation and 1-2μg for visualisation on a gel.)
- Do not change the ratio of cRNA: fragmentation buffer.
- Prepare a mastermix; Xμl cRNA (25μg)
  Yμl DEPC-treated water
  10μl (5x) fragmentation buffer
  = 50μl total.
- Incubate at 94°C in a heating block for EXACTLY 35 minutes
- Put tubes on ice.
- Run samples on a denaturing formaldehyde gel to check quality

- Store at -80°C until hybridisation.

2.8. RNA Gel.

- Add 2.5g agarose to 150ml DEPC-treated water. Boil and cool.
- Work in a fume cupboard.
- Add; 6µl ethidium bromide
  17.5mls 10x MOPS
  5mls (37%) formaldehyde
- mix well and allow to cool and set.
- Age the gel in 1x MOPS overnight, or for at least ½ hour.

2.8.1. Loading the Gel.

- Place 10µg RNA in a clean RNase-free tube and make up to 20µl with 1x RNA loading buffer.
- Denature at 65°C for 15 mins.
- Cool on ice.
- Load from ice, run alongside RNA markers, (Sigma) at 100v for 1-2 hours. Photograph frequently as the ethidium bromide does not intercalate into fragmented cRNA with high efficiency.

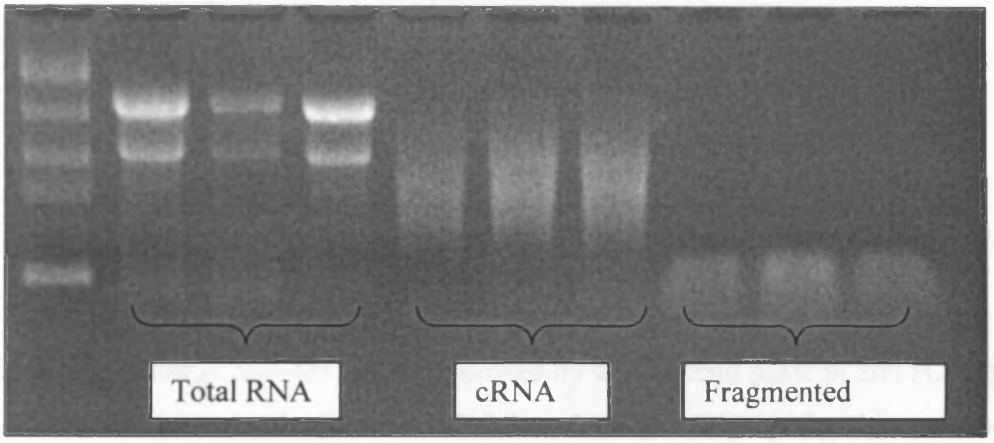
2.8.2. 10x MOPS.

- 10x MOPS (makes 1l)
  41.8g MOPS (free acid, Sigma)
  pH to 7 with NaOH
  Add 16.6ml 3M NaOAc
  Add 20ml 0.5ml EDTA (pH8)
  Make up to 1l with distilled water

Autoclave.

2.8.3. Checking the RNA Quality.

Ladder.



- 0.5-1µg total RNA should show 2 bands with the upper band at roughly twice the intensity of the lower band.
- 500ng-1µg cRNA should show a smear from 100bp-2kb with a brighter region from 500bp-1kb.
- 1-2µg cRNA should show a smear from 35bp-200bp.
- It is important to photograph the gel several times whilst it is running, as the fragmented cRNA does not intercalate ethidium bromide at the same efficiency as total RNA.

2.9. RNA extraction and preparation for RT-PCR

2.9.1. Obtaining RNA from Tissue Samples.

- Dissect out the tissue of interest to a maximum thickness in one direction of 5mm. (For the intestinal MeCP2 array, the very top 3cm of the proximal small intestine was used.)

- Transfer immediately to cold RNA later™ (Sigma cat no. R0901) to inactivate endogenous RNase activity.

- Tissue stored in RNA later™ can keep for 1 day at 37°C, 1 week at 25°C, 1 month at 4°C and can be archived indefinitely at -20°C.

To process the tissue;

- Pour off the RNA later™ and note the wet weight of the tissue.
- Add 1ml cold Trizol reagent (Gibco cat no. 15596-018) per 100mg tissue.
- Homogenise the tissue.
- Clean the homogeniser before first use, between each use and after final use with a 15 second pulse in 5ml 4M NaOH followed by a 15 second pulse in 5ml H20 for 15 sec, then a 15 second pulse in 5 ml 75% ethanol (all solutions made with nuclease-free water)
- Aliquot 1 ml samples into clean, labelled RNase-free centrifuge tubes. Keep on ice.
- Centrifuge for 10 mins at max speed in a bench-top centrifuge at 4°C.
- Take off the supernatant (avoid the surface layer of fat) and transfer to new, labelled, sterile RNase-free tubes.
- Add 200µl chloroform, shake for 5-10 mins, then chill for 5-10 mins.
- Centrifuge samples for 15mins at max. speed (bench-top centrifuge at 4°C)
- Carefully remove the top aqueous layer which will contain the RNA and transfer to new tubes. Avoid taking material from the interface, which will contain DNA. Dispose of the lower layer in phenol waste.
- To precipitate the RNA, add an equal volume of isopropanol, mix by inversion and leave for 10 mins at 4°C. Leaving the tubes for longer may maximise RNA yields.
- Centrifuge samples for 15 mins at max. speed (bench-top centrifuge at 4°C)
- Discard the supernatant and air-dry the pellet for 5-10 mins.
- Resuspend the RNA in 100µl (or more if needed) nuclease-free water.
- Determine the total yield of RNA spectrophotometrically. 1 A260 unit = 40µg/ml of single-stranded RNA.

2.10. Making cDNA from total cellular RNA.

- Make up 1μg RNA in 9μl DEPC-treated water
- Heat to 70°C for 10 mins
- Heat to 42°C for 2 mins
- Add: 2μl random hexamer primers
  2μl DTT
  4μl buffer
  1μl 10mM dNTPs
  1μl DEPC-treated H_20.
- Add 1μl Superscript reverse transcriptase to half the samples, and 1μl DEPC-treated water to the other half.
- Heat to 42°C for 50mins-1 hr.
- Heat to 70°C for 15 mins
- Store at -20°C until needed.
- Use 2μl of this cDNA in each PCR reaction.

2.11. Checking for genomic DNA contamination.

Run an Hprt PCR on each sample (+Reverse transcriptase and – reverse transcriptase.) No bands should be seen in the absence of reverse transcriptase. The presence of a band indicates genomic DNA contamination and the sample should be re-prepared.

2.12. RT-PCR.

- For quantitative realtime RT-PC, the BioRad SYBR-green system was used.
- Each sample was run in duplicate.
- For each sample, use: 12.5μl SYBRgreen
  8.5μl distilled water
  1μl each primer (1μM)

2μl cDNA

- Place the samples in duplicate into a 96 well plate (or tube strips) and seal with clear optical quality caps or sealing film. With each run, run 4 'blank' samples, in which the cDNA is replaced with water, to control for contamination.
- Genomic DNA contamination can be controlled for with a single run (to save sample.) A single run using β-actin primers on sample with/without reverse transcriptase will control for genomic DNA contamination.

2.12.1. RT-PCR primers

Hprt: P1: TGTTGTTGGATATGCCCTTG
P2: TTGCGCTCATCTTAGGCTT

Glucagon: P1: GATTGCTTATTAATGCTGGTGT
P2: TCTTCATTCATCTCATCAGGG

Cd44 P1: GTGGCAGAAGAAAAAGCTG
P2: TTGTTCACCAAATGCACCAT

Mecp2 P1: ACCTCTAACCCTGCCTGGAT
P2: AGGCCGTGCTAGCAAAGTAA

m46: P1: AGAAGCCACCAGCTAAACTG
P2: TCCTGACAGCTTGATGCCAA

Gip P1: GCAAGATCCTGAGAGCCAAC
P2: TTAGCATGGGATCGGAACTC

Pedgf: P1:ACCGTGACCCAGAACTTGAC
P2:CACGGGTTTGCCAGTAATCT

Wdnm1 P1: CGCAGTTTTGGAACTTGTGA
P2: CAAGCATGGGGTCTGTAGGT

Igfbp3 P1: CAAAGCACAGACACCCAGAA
P2: CTGCTTTCTGCCTTTGGAAG

Igfbp10 P1: AGACCCGGATCTGTGAAGTG
P2: TTCTGGTCTGCAGAGGTGT

Neuromedin P1: CCGAGGGACCAGAGACTACA
P2: CATTGTCAGATTCCCTGGAT

β-Actin P1: TACAGCTTCACCACCACAGC
P2:AAGGAAGGCTGGAAAAGAGC

Ccf-i P1: TATGTTCCAACCGAATGACA
P2: GAGTCTCCTTTGCAGGCATC

Liprinα P1:TGATGTGGATGAGGATGAGC
P2: GGTCACACGGGTCTCAATCT

Dlx5 P1: CAGAAGAGTCCCAAGCATCC
P2: GAGCGCTTTGCCATAAGAAG

Dlx6 P1:CACAGCCCTTACCTCCAGTC
P2: AGTCTGCTGAAAGCGATGGT

SSc2 P1: GCTCTCGATATGGCTGGGTA
P2: TACCCGGACATCACCTTCTC

Chapter 3: Investigating the Role of Mbd4 in the Intestine.

Mbd4 (also known as Med1) is unique in the MBD family in that it appears to play no role in transcriptional repression (Bellacosa 2001.) However, it is becoming clear that Mbd4 plays a vital role in maintaining genomic fidelity, suppressing tumourigenesis and mediating the apoptotic response to DNA damage (Sansom *et al* 2003a. Millar *et al*)

3.1. Apoptosis and the response to DNA damage.

DNA damage is caused by a wide variety of agents, from environmental carcinogens such as PAHs in tobacco smoke (De Marini *et al* 2004) to endogenous damage caused by by-products of metabolism (Evans *et al* 2004.) Since damage to DNA would otherwise lead to damaging mutations, a number of mechanisms operate to detect and repair damage, such as base excision repair (BER), nucleotide excision repair (NER) and mismatch repair (MMR) (Christmann *et al* 2003.)

Mbd4, a member of the methyl CpG binding domain family, is involved in protecting DNA by acting as a G•T and G•U mismatch specific thymine or uracil glycolsylase (Bellacosa 2001, Petronzelli 2000.) Furthermore, Mbd4 acts preferentially on these mismatches when they are present in the context of methylated or unmethylated CpG sites (Bellacosa 2001, Petronzelli 2000.) The CpG dinucleotide is hypermutable, with G•T and G•U mismatches being formed by the hydrolytic deamination of 5 methyl-cytosine and cytosine to thymine and uracil. These deamination events are frequent, occurring at a rate of 2-300 per cell per day and if left unrepaired would form G•C and A•T transition mutations at the next round of DNA replication (Bellacosa 2001, Petronzelli 2000.) By preventing these mutations, Mbd4 may act as a caretaker

of genomic fidelity at hypermutable CpG sites (Bellacosa *et al* 2001.) Such mutations are known to contribute to tumourigenesis; nearly 50% of somatic P53 mutations in colorectal cancers arise at 'hotspots' where cytosines in CpGs are deaminated to form transition mutations (Bellacosa 2001, Petronzelli 2000.) Loss of functional Mbd4 may accelerate the formation of transition mutations and therefore Mbd4 can be considered to have tumour suppressor activity.

Mbd4 may also have a role in the MMR pathway of DNA repair. Having found that Mbd4 interacted with Mlh1 (a MutL homologue) in yeast 2-hybrid screens, Bellacosa proposed that Mbd4 is part of a multimeric protein complex involved in mismatch repair (Bellacosa 2001.) Initial characterisation of Mbd4 supported this hypothesis. The protein is composed of three domains; an N-terminal methyl binding domain (MBD) a central region and a C-terminal catalytic domain with homology to bacterial DNA damage specific base excision repair glycoslyases / lyases (Bellacosa 2001, Petronzelli 2000.) During the major pathway of post-replicative MMR, Mlh1 and Pms2 form heterodimers which interact with Msh2/Msh6 heterodimers bound to mismatched bases (Jiricny *et al* 1998, Buermeyer *et al* 1999.) Mbd4 was found to interact with Mlh1 (Bellacosa 2001) and also with Msh2 – components of the mammalian mismatch repair (MMR) system, in yeast 2-hybrid screens, prompting speculation that it may itself be a part of the MMR machinery. However, other co-immuoprecipitation-based studies have shown that Mbd4 does not interact with Msh6 or Pms2, which argues that Mbd4 is not part of the MMR machinery as it is currently understood (Sansom *et al* 2003b.)

3.1.1. Mismatch Repair

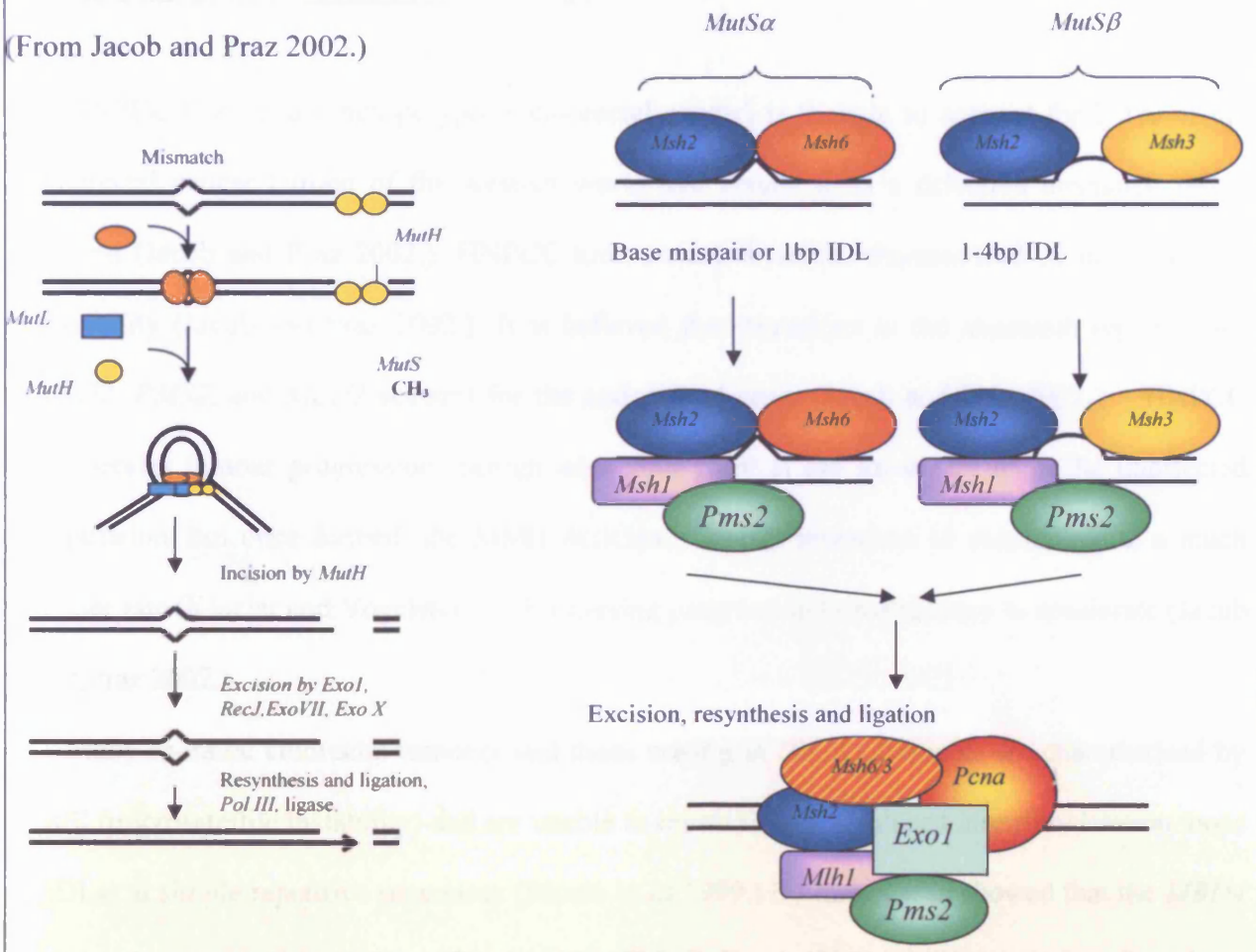
The mismatch repair system helps to maintain genomic stability by repairing insertion and deletion loops (IDLs), mismatched base pairs and heteroduplexes that can form during DNA

replication and recombination (Jacob and Praz 2002.) Although the bacterial MMR system is well characterised, less is known about how MMR works in mammalian systems (Jacob and Praz 2002, Hart *et al* 2002.) In bacterial MMR (see fig 3.1) the MutS protein detects the mismatch and interacts with MutL, activating MutH, a single-strand endonuclease. MutH can recognise the newly synthesised strand that carries the mismatched base, due to the transient lack of methylation on the adenines of the palindromic sequence GATC (Jacob and Praz 2002.) Since both bases in the mismatch appear 'wrong' the lack of methylation marks the strand on which the endonuclease should act (Jacob and Praz 2002.) The endonuclease then excises a stretch of DNA up to 1-2kb in length from between the hemi-methylated GATC sites. The tract is then resynthesised correctly by DNA Pol III and ligated by DNA ligase (Jacob and Praz 2002.)

Most of the key MMR components are highly conserved from prokaryotes to mammals and homologues have been found for most of the bacterial genes involved (Jacob and Praz 2002.) However, to date, no mammalian MutH homologues have been found, and the methylation recognition system must be subtly different since mammalian DNA is methylated on cytosines in CpG dinucleotides, not adenines in GATC sequences (Jacob and Praz 2002.) Four homologues of MutL (Mlh1-Mlh3 and Pms1) have been identified in yeast and mammals along with six homologues of MutS (Msh1-Msh6) in yeast and five in mammals, which have no Msh1 (Jacob and Praz 2002.) Msh2 can form heterodimers with Msh6 (to form a MutSα complex) or with MSH3 to form a MutSβ complex. The MutSα and MutSβ complexes recognise the mismatch (Jacob and Praz 2002.)

Fig 3.1. Mismatch repair.

(From Jacob and Praz 2002.)



A. In prokaryotic MMR, MutS detects the mismatch and interacts with MutL, activating MutH, a single-strand endonuclease. MutH can recognise the newly synthesised strand that carries the mismatched base, due to the transient lack of methylation on the adenines of the palindromic sequence GATC. The unmethylated strand is acted upon. The endonuclease then excises a stretch of DNA up to 1-2kb in length from between the hemi-methylated GATC sites and the tract is then resynthesised correctly by DNA Pol III and ligated by DNA ligase (Jacob and Praz 2002.)

B. Several MutL homologue activities are present in eukaryotes, but the major mismatch repair activity is a heterodimer of Mlh1 and Pms1 (MutLα) which interacts with MutSα and MutSβ. Once bound, the MutLα complex can interact with other factors to form a complex involved in the processing of the mismatch. In eukaryotes, MutSα (Msh2/Msh6) recognises base-base mismatches and 1bp IDLs and MutSβ (Msh2/Msh3) recognises larger IDLs. Both then recruit Pms2/Mlh1. Later steps are less well characterised but may involve recruitment of Exo1 (a 3'-5' exonuclease) and Pcna to excise, resynthesised and ligate the repair (Jacob and Praz 2002.)

3.1.2. Defective MMR and human disease.

HNPCC (Hereditary non-polyposis colorectal cancer) is thought to account for 2-4% of the colorectal cancer burden of the western world and results from a defective mismatch repair system (Jacob and Praz 2002.) HNPCC tumours are therefore characterised by microsatellite instability (Jacob and Praz 2002.) It is believed that mutations in the mismatch repair genes *MSH2, PMS2*, and *MLH1* account for the majority of cases (Jacob and Praz 2002.) HNPCC accelerates tumour progression; benign adenomas form at the same rate as in the unaffected population, but once formed, the MMR deficiency causes mutations to accumulate at a much higher rate (Kinzler and Vogelstein 1996) causing progression to malignancy to accelerate (Jacob and Praz 2002.)

Many sporadic colorectal tumours and those arising in HNPCC patients are characterised by MSI (microsatellite instability) and are unable to repair slippage-induced insertion/deletion loops (IDLs) at simple repetitive sequences (Riccio *et al* 1999.) Bellacosa *et al* showed that the *MBD4* gene contains 4 such sequences (Riccio *et al* 1999, Bellacosa 2001.) *MBD4* mutations have been detected in approximately 20-25% of human colorectal and gastric carcinomas (Bader *et al* 1999.) Virtually all of these mutations were in a poly$(A)_{(10)}$ tract in the central portion of the gene, resulting in frameshifts which produce a truncated protein with no C-terminal domain and no catalytic ability (Bader *et al* 1999.) This suggests that MBD4 is a tumour suppressor, albeit perhaps not a major suppressor, in the human intestine. Interestingly, no studies have yet looked at possible epigenetic inactivations of *MBD4* in human colorectal cancers.

If Mbd4 is indeed such an important part of the DNA repair machinery, it would be expected that loss of Mbd4 would accelerate tumourigenesis. *Mbd4* null mice have been generated and crossed to mice carrying the Apc^{Min} mutation, which predisposes to multiple intestinal neoplasia (Sansom *et al* 2003a.) $Mbd4^{-/-}Apc^{Min/+}$ mice showed a significant increase in intestinal tumour burden when compared with their $Mbd4^{+/-}/Apc^{Min/+}$ littermates, further reinforcing the status of Mbd4 as a tumour suppressor gene (Sansom *et al* 2003a.)

Because Mbd4 plays a role in DNA repair, it might be possible that Mbd4 deficiency would lead to an increase in mutation rates. Sansom *et al* found no increase in mutation rates in $Mbd4^{-/-}$ mice using the Dlb-1 assay (Monroe *et al* 1995); however, using the BigBlue® assay (Millar *et al* 2002) a small but significant increase in mutation was noted (The Dlb-1 assay can be used to score mutation rates in the intestinal crypts (Millar *et al* 2002) while the BigBlue® reporter locus, *cII*, is highly methylated and would respond to an increase in C→T transitions.)

Although the negative Dlb1 result might at first argue that loss of Mbd4 does not increase the mutation rate, it is not a highly sensitive assay, and the increased rate of C→T transitions along with the significantly increased tumour burden of the $Apc^{Min}/Mbd4^{-/-}$ mouse equates to an increase in mutation as it represents a shift in the mutation type to a deamination-like event. Loss of Mbd4 activity may not greatly compromise cell viability, as mice possess another thymine glycosylase activity, Tdg which may partially compensate for its loss (Nedderman *et al* 1996.)

Mbd4 can also act to remove 5-Fluorouracil (5-FU) mismatches in the context of G•5-FU mismatches (Sansom *et al* 2003a.) 5-FU is a uracil analogue used as a chemotheraputic agent. This raises the possibility that Mbd4 may be involved in tumour resistance to 5-FU therapy (Sansom *et al* 2003a.)

3.1.3. Mbd4 and Apoptosis.

The possible involvement of Mbd4 in the MMR system of DNA repair suggests a role in the signalling of apoptosis. Data presented by Sansom *et al* points to a role for Mbd4 in the mediation of the apoptotic response (Sansom *et al* 2003a.) *Mbd4*$^{-/-}$ mice showed a significantly reduced apoptotic response after treatment with cisplatin, ionising radiation or the alkylating agent temozolamide, all of which induce apoptosis in the intestinal epithelium via DNA damage (Sansom *et al* 2003a.) RNA damaging agents such as 5-FU (5-flurouracil), did not cause a differential response, implying that Mbd4 is needed to signal or mediate the apoptotic response following DNA damage but not the response to RNA damage (Sansom *et al* 2003a.)

Following on from this observation, this project will assess the response of Mbd4 deficient (*Mbd4*$^{-/-}$) mice to Fas ligand, which can trigger an apoptotic response via a cell-surface receptor, and the role of Mbd4 in anoikis (apoptosis driven by loss of cellular attachment.)

3.1.4. Apoptosis and anoikis.

Apoptosis is a regulated, programmed process that results in cellular 'suicide' (Brown and Attardi 2005.) Unlike necrosis, where the cell essentially 'bursts' and releases its contents into the surrounding tissue, an apoptotic cell is broken down in a highly regulated series of stages and is finally engulfed by other cells types (Brown and Attardi 2005.) Apoptosis plays a central role in growth and development; for example, in the developing human embryo, cells between the nascent fingers undergo apoptotic death to shape the embryonic hand (Hardy 1999.) Apoptotic death also serves to remove damaged cells which may otherwise carry dangerous genetic lesions (Brown and Attardi 2005.) Apoptosis is triggered by a variety of cues, such as

DNA damage, growth factor withdrawal or detachment from the matrix (Brown and Attardi 2005.) The process of apoptosis is carried out via two central pathways; one which involves the activation of caspases and a second, mitochondrial pathway (Brown and Attardi 2005.) One such pathway, mediated by the death receptor, Fas, is shown in fig 3.2.)

One hallmark of the cancer cell is reduced sensitivity to apoptotic signals, which increases their ability to proliferate uncontrollably. Loss of sensitivity can be caused by mutations in genes encoding members of the apoptotic machinery, or by loss or mutations of proteins such as the cellular sentinel P53, which normally transduces cellular stress signals into growth arrest or apoptosis. Since the conditions encountered in a typical tumour (hypoxia etc) are precisely those which would normally trigger apoptosis, a tumour may have higher rates of apoptosis than surrounding normal tissue, but it may also select for cells which can survive under these conditions (Hanahan and Weinberg 2000.)

3.1.5. Fas-Mediated Apoptosis.

Fas (also known as CD95 or APO-1) is one of specialised subset of TNF receptors known as 'death receptors' (Nagata 1999.) Interaction of Fas ligand (FasL) with its receptor induces apoptosis in Fas-bearing cells (Fig 3.2.) Upon ligand binding, the receptor trimerises (Nagata 1999.) The cytoplasmic regions of the receptors contain 'death domains'; trimerisation of the death domains recruits FADD, and adaptor molecule, to form the death-inducing signalling complex (DISC) (Nagata 1999.) FADD then recruits caspase 8, which auto activates and triggers a caspase cascade (Nagata 1999.) Caspases are the effectors of the apoptotic pathway and cleave a number of cellular substrates leading to the functional and morphological consequences of apoptosis (Nagata 1999.)

Fig 3.2. Apoptosis Triggered through the Fas Receptor Pathway.

Adapted from Algeciras-Schimich et al 2002 and Joza et al 2002.)

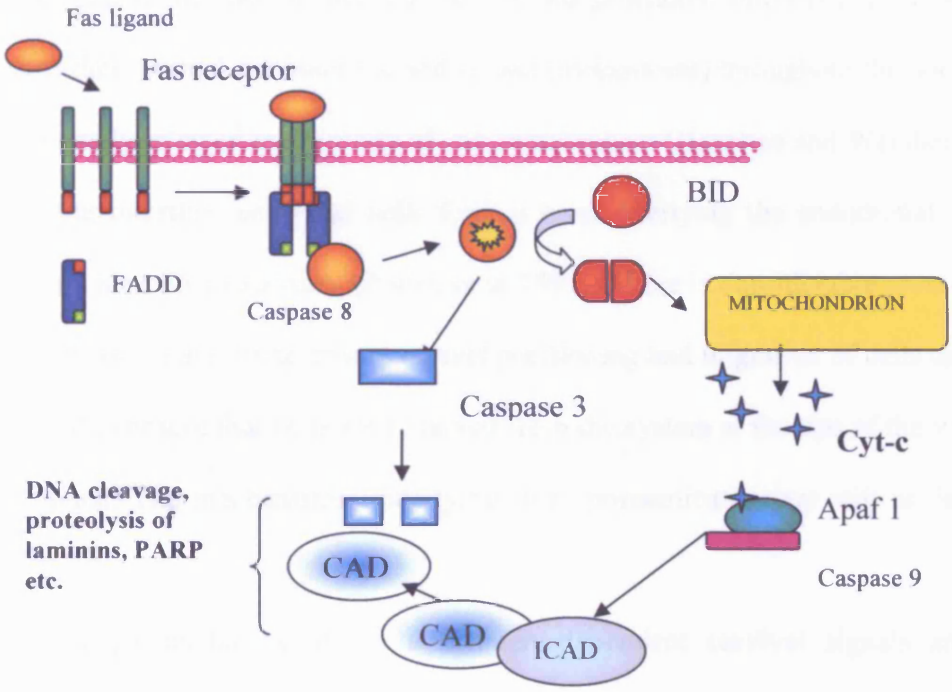


Fig 3.2. Fas-mediated apoptosis.

Interaction of Fas ligand (FasL) with its receptor induces apoptosis in FasL-bearing cells. Upon ligand binding, the receptor trimerises. The cytoplasmic 'death' domains trimerise and recruit FADD, anadaptor molecule, to form the death-inducing signalling complex (DISC.) FADD then recruits caspase 8, which auto-activates to trigger a caspase cascade that leads to apoptosis

3.1.6. Anoikis.

Since the default state of the cell is, effectively, apoptotic death, in a normal tissue, cells require a constant input of survival signals to remain alive. These signals can be both chemical (e.g. growth factors) and physical (attachment signals) (Hanahan and Weinberg 2000.) Cell death caused by loss of attachment is called 'anoikis', from the Greek for 'falling

off' (Valentijn *et al* 2004.) Correct attachment signals ensure that the cell is in its proper place within the tissue microenvironment (Valentijn *et al* 2004.) One hallmark of a typical cancer cell is the loss of this requirement for positional information - this allows cells to escape their normal environment and spread (metastasise) throughout the body, hence anoikis is extremely relevant to the study of colorectal cancer (Hanahan and Weinberg 2000.)

In the intestine, epithelial cells form a layer overlying the endothelial 'scaffold' which supports the villus structures (Potten *et al* 1997.) There is considerable cross-talk between all elements of the system to ensure correct positioning and migration of cells up the crypt-villus axis and to ensure that cells are removed from the system at the tips of the villi and shed into the lumen. The mechanisms underlying this communication are still unclear (Potten *et al* 1997.)

The major mediators of the attachment-dependent survival signals are the integrins. Integrins are transmembrane cell surface receptors which are composed of heterodimers of various combinations of α and β chains (Valentijn *et al* 2004.) At least 22 α/β combinations are known, giving a large and diverse family of molecules which are capable of binding various different components of the extracellular matrix (ECM.) Since each component of the ECM is also capable of binding different integrins, the mechanisms of attachment and signalling are complex and not fully understood (Valentijn *et al* 2004.)

It has recently been shown that the Fas-mediated apoptotic pathway plays a role in triggering anoikis (Valentijn *et al* 2004, Grossman 2002.) It is thought that changes in cell shape caused by detachment may bring the Fas receptors into close enough proximity to trimerise and trigger the pathway (Valentijn *et al* 2004, Grossman 2002.) However, since cells can be rescued by addition of soluble matrix components, this may not be the major route of Fas-mediated anoikis (Valentijn *et al* 2004.) Another theory is that detachment

triggers upregulation of components of the Fas pathway; detatched endothelial cells have been shown to produce a 3-fold upregulation of Fas receptor, a 1.5 fold upregulation of Fas ligand and a massive downregulation of FLICE, an inhibitor of the pathway (Valentijn *et al* 2004.)

This project assesses the role of Mbd4 in anoikis by examining how quickly wildtype *Mbd4*$^{+/+}$ and *Mbd4*$^{-/-}$ null isolated intestinal epithelial crypts undergo detachment-mediated death. This project also assess the requirement for Mbd4 in the apoptotic response of the small and large intestine and the liver, by challenging mice with wild type or a null allele of *Mbd4* with the pro-apoptotic Jo-2 Anti-Fas antibody, which acts in a similar manner to Fas ligand (FasL) and triggers the Fas-mediated apoptotic death pathway.

3.2. Results

3.2.1. Mbd4 influences anoikis in isolated murine small intestinal crypts.

Anoikis is the term given to programmed cell death induced by a loss of detachment signals (Valentijn 2004.) Mbd4 has been shown to modulate the anoikis response to cell detachment (Screaton *et al* 2003.) Fadd is also reported to be localised to the nucleus in adherent cells lines (Screaton *et al* 2003) and adherence is seen to protect to some degree against Fas-ligand induced cell death. Taken together, these results imply a role for Mbd4 and Fadd in the modulation of the anoikis response to cell detachment.

I first wanted to determine whether it was possible to isolate small intestinal crypts from the murine small intestine, to define the process of death occurring in wild-type crypt populations. Isolated small intestinal crypts were derived from wild-type mice and then observed at 0,1,2 and 3 hours from the time of isolation. Once taken out of the *in vivo* environment, crypts are deprived of the normal attachment and survival signals they receive and, as would be expected, gradually undergo a mixture of apoptotic (stress-induced) and apoptotic (anoikis-induced) death.

Fig 3.3. Anoikis in Isolated Murine Small Intestinal Crypts.

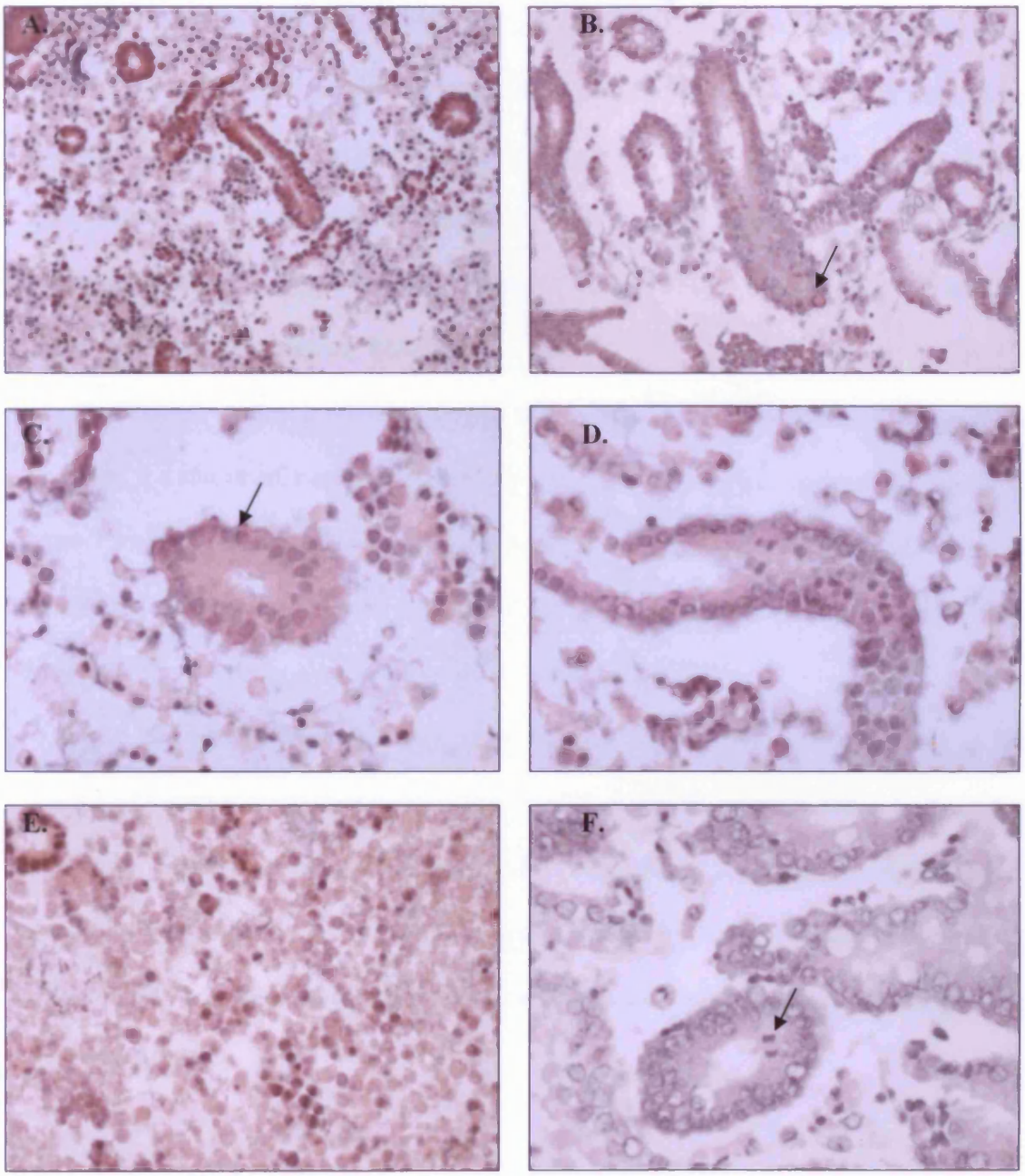


Fig 3.3. Anoikis in isolated murine small intestinal crypts.

A.) Isolated crypts on a background of single-cell debris. B) Arrow = Apoptosis/anoikis occurring in a crypt cell. C.) Cells are extruded from intact crypts, then undergo apoptosis/anoikis. D.) At 2 hours, many crypts are still intact. E.) By 3hours, most crypt structures have gone, leaving only single-cell debris. F.) Even at 3 hours some cells remain viable and cell division can still be seen occurring in the intact crypt.

H&E stained section visually; an apoptotic cell is rounded, and tends to stain more pin its surrounding cells (fig 3.3b.) The chromatin in the nucleus is condensed and darkly while the membrane may be 'blebbing.'

In contrast to the short-lived single-cell debris, those cells enclosed in crypt st survived a number of hours. Even at 3 hours, cells were seen undergoing mitosis (arrow) implying that they were still viable. Cell which did undergo apoptosis within structure often appeared to be 'squeezed out' of the crypt (fig 3.3c arrow), althou possible that this could also be a sign of membrane blebbing.

Given the ability to isolate murine small intestinal crypts, I next wanted to de whether the presence or absence of Mbd4 would affect the rate of anoikis in this assa matched 3 month old male and female *Mbd4*$^{+/+}$ and *Mbd4*$^{-/-}$ outbred mice segregating f and C57/Bl6 genomes were culled and crypt preparations made from the small intestin crypt populations were examined at 0, 1, 2 and 3 hours post-isolation for apoptosis. Re shown in fig 3.4.

Fig 3.4. Loss of* Mbd4 *reduces anoikis in isolated small intestinal crypts

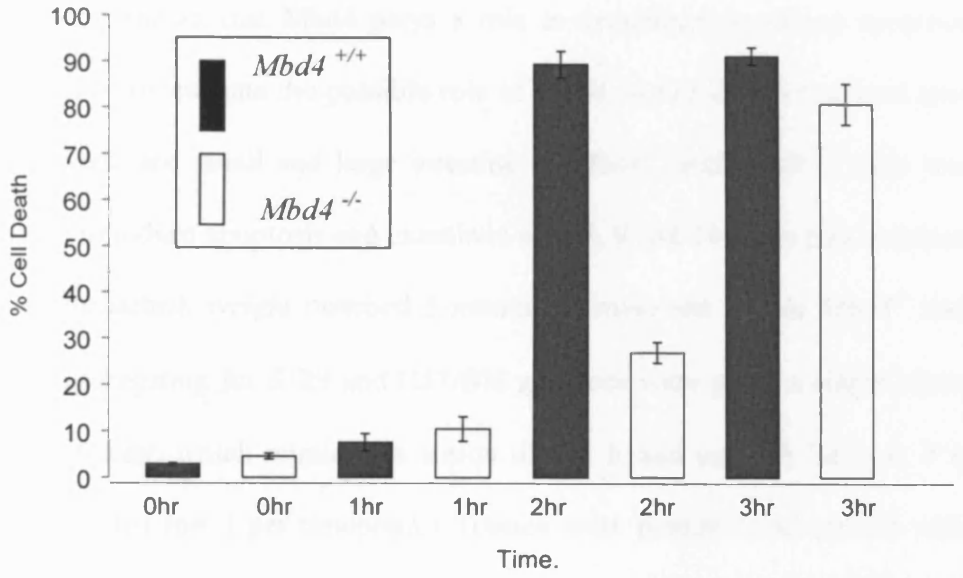


Fig 3.4. Loss of* Mbd4 *results in reduced anoikis in isolated murine small intestinal crypts.

At 2 and 3 hours, $Mbd4^{-/-}$ crypts undergo significantly less anoikis than $Mbd4^{+/+}$ crypts (p= <0.005.) The difference is largest at 2 hours, when around 25% of $Mbd4^{-/-}$ crypts have died compared to over 90% of $Mbd4^{+/+}$ crypts (p=0.0404.)

Fig 3.4 shows the differences in anoikis between the $Mbd4^{+/+}$ and $Mbd4^{-/-}$ crypts. At 0 and 1 hours, no significant difference in the number of apoptotic cells is seen. At 2 hours, however, 25% of cells in $Mbd4^{-/-}$ crypts have died compared to over 90% of cells in $Mbd4^{+/+}$ crypts, a significant difference (p=<0.005.) By 3 hours, the difference is still significant (p=0.0404) , although the difference is much smaller, with around 82% of cells in $Mbd4^{-/-}$ crypts having died compared to over 95% of cells in $Mbd4^{+/+}$ crypts. This shows that there is a significant difference between the rate at which $Mbd4^{-/-}$ and $Mbd4^{+/+}$ crypts undergo detachment-mediated death, implying a role for Mbd4 in anoikis.

3.2.2. The role of Mbd4 in Fas-mediated apoptosis.

Having shown that Mbd4 plays a role in detachment-mediated apoptosis (anoikis) I then wanted to investigate the possible role of Mbd4 in Fas ligand-mediated apoptosis. To do this, the livers, and small and large intestine of *Mbd4*$^{-/-}$ and *Mbd4*$^{+/+}$ mice were exposed to Fas ligand to induce apoptosis and examined at 3, 6, 9 and 24 hours post exposure.

Age-matched, weight matched 3 month old male and female *Mbd4*$^{-/-}$ and *Mbd4*$^{+/+}$ outbred mice segregating for S129 and C57/Bl6 genomes were given a single injection of 10μg Anti-jo2 antibody, which mimics the action of Fas ligand and left for 3, 6, 9 or 24 hours before being culled (n= 3 per timepoint.) Tissues were prepared and stained with H&E. Apoptosis was scored as a percentage of cells in the liver and as the number of apoptotic cells per 50 half crypts. Results are shown in figs 3.5, 3.6 and 3.7.

Fig 3.5. Apoptosis in the Small Intestine of* Mbd4*-deficient Mice - 10 μg Fas ligand.

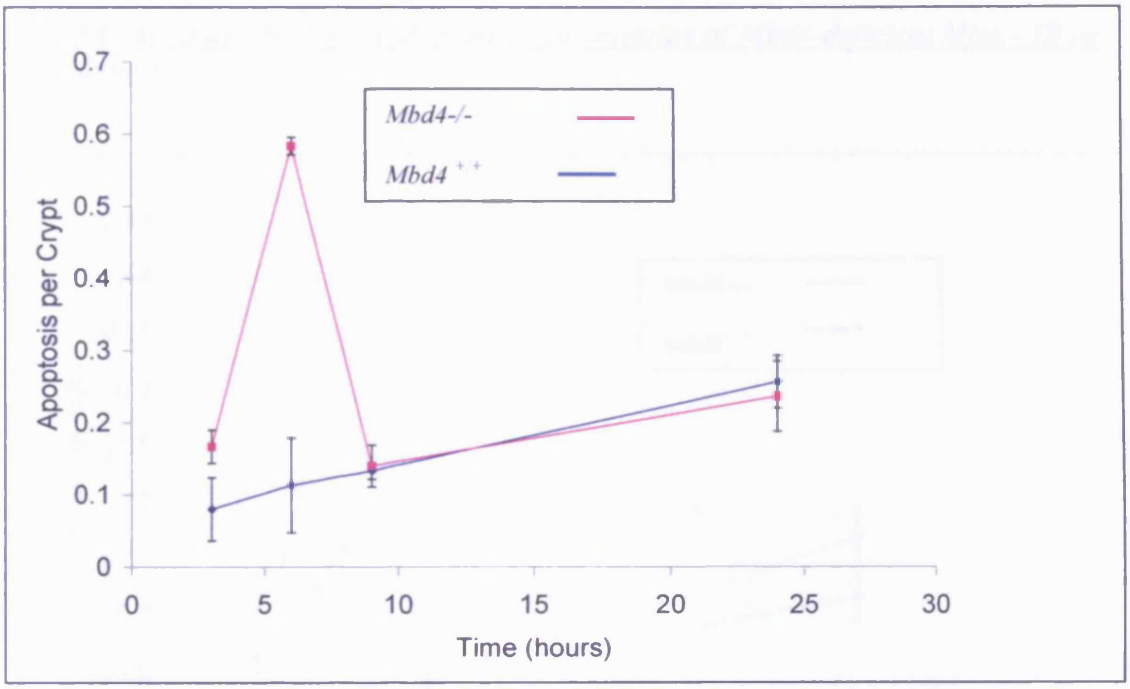


Fig 3.5. Apoptosis in the small intestine at 3, 6, 9 and 24 hours after administration of Jo-2 anti-Fas.

Apoptosis 'peaks' at around 6 hours and there is a significant difference in levels of apoptosis between wild-type and Mbd4 *deficient mice at the 6 hour time point (Mann Whitney 2 tailed test p=0.0048.) No other time points show a significant difference in levels of apoptosis between the two genotypes.*

Fig 3.6. Apoptosis is increased in the large intestine of Mbd4-deficient Mice - 10 μg Fas ligand.

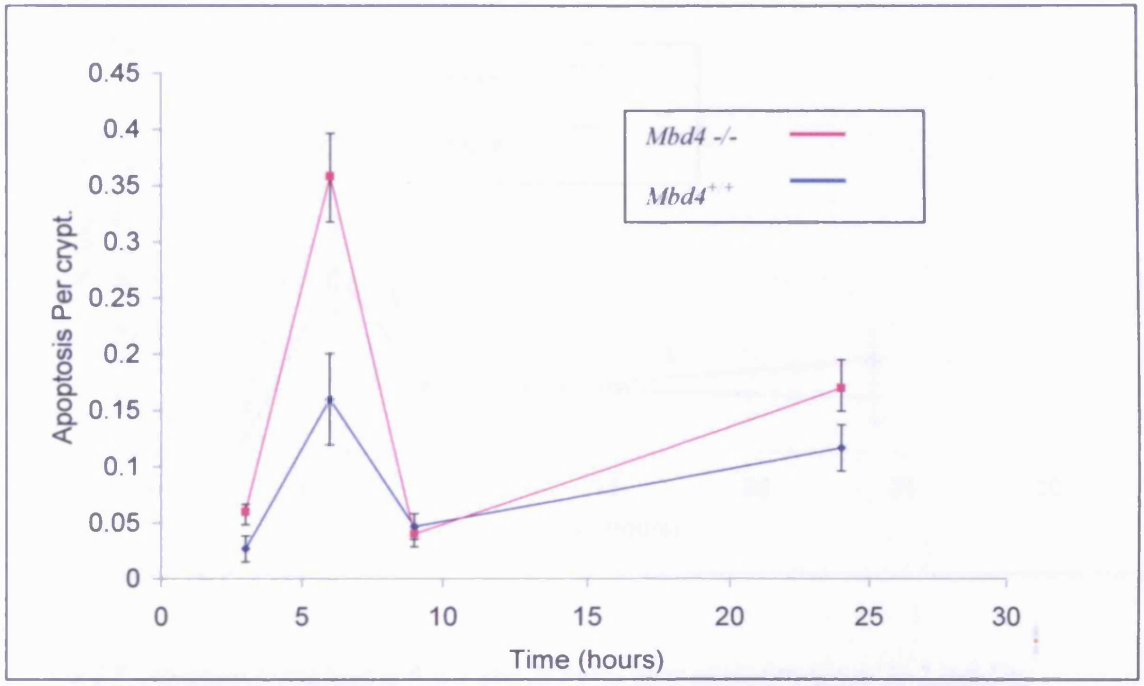


Fig 3.6. Apoptosis in the large intestine at 3, 6, 9 and 24 hours after administration of Jo-2 anti-Fas.

Apoptosis 'peaks' at around 6 hours and there is a significant difference in levels of apoptosis between wild-type and Mbd4 *deficient mice at the 6 hour time point (Mann Whitney 2 tailed test p=0.0127.) Unlike the small intestine, the large intestine does show a small but significant difference at 24 hours (p=0.05)*

Fig 3.7. Apoptosis in the Liver of* Mbd4-*deficient Mice -10 μg Fas ligand.

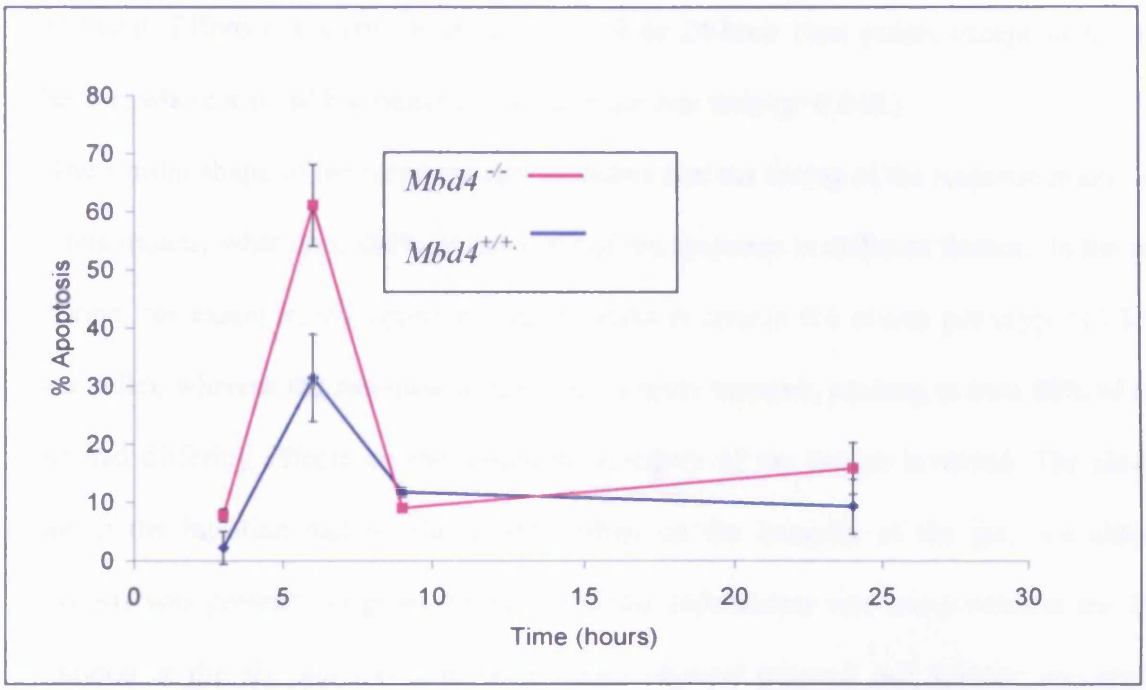


Fig 3.7. Apoptosis in the liver at 3, 6, 9 and 24 hours after administration of Jo-2 anti-Fas.

Note how levels of apoptosis are much higher than in the intestine. Apoptosis 'peaks' at around 6 hours and there is a significant difference in levels of apoptosis between wild-type and Mbd4 *deficient mice at the 6 hour time point (Mann Whitney 2 tailed test p=0.0543.) No difference is observed at any other time point.*

Figs 3.5, 3.6 and 3.7 show an increased apoptotic response in *Mbd4*$^{-/-}$ mice compared to wild-type (*Mbd4*$^{+/+}$) mice following Fas ligand administration. Note how the shape of the response curve is identical in all three tissues tested (small and large intestine and liver) with a peak in apoptosis occurring at the six hour time point in all tissues. The difference in

apoptosis at the six-hour time point is significant in all tissues (Mann-Whitney 2-tailed test, n=6, p=0.0048 (small intestine), p=0.0217 (large intestine) p=0.0543 (liver)) No other significant differences were found at the 3, 9 or 24-hour time points except in the large intestine, where a small but significant difference was seen (p=0.048.)

The similar shape of the response curves shows that the timing of the response is similar in various tissues; what does differ is the scale of the response in different tissues. In the small intestine, for example, the apoptotic index peaks at around 0.6 events per crypt (13.2% of crypt cells), whereas the response in the liver is more extreme, peaking at over 60% of cells. This had differing effects on the structural integrity of the tissues involved. The six-hour peak in the intestine had no discernible effect on the integrity of the gut, (i.e although apoptosis was present, no gross change in tissue architecture was seen) whereas the livers examined at the six and 24- hour time points showed reduced cell number and obvious widespread tissue damage.

The Fas ligand may be absorbed and reach the liver more efficiently than the gut, (so the liver is more affected by the fas-mediated apoptosis, and may also metabolise much of the Fas ligand, meaning that less is available to reach the intestine), or hepatocytes may be more susceptible to fas ligand-induced apoptosis than enterocytes.

From the above data, it can therefore be concluded that **Mbd4 deficient mice have an increased apoptotic response to Fas when compared with wild-type mice across a range of different tissue types.**

3.2.3. Investigating the interaction between Mlh1 and Mbd4.

Since Mbd4 and Mlh1 are known to interact, the effect of Mlh1 on apoptosis and mitosis in the small intestinal crypt and the liver was examined. Fas ligand induces extensive apoptosis

in both intestine and liver and so a range of doses were used in an intital attempt to determine a dose which would allow accurate measurement of apoptosis and mitosis with minimal tissue damage.

Age-matched, weight matched 3 month-old *Mlh1*$^{-/-}$ and *Mlh1*$^{+/+}$ mice were given 5μg, 8μg or 10μg of anti-jo2 Fas ligand as an intraperitoneal injection and left for 6 hours (as the previous study showed that the maximal apoptotic response occurs at 6 hours) At 6 hours, animals were sacrificed, and their small intestines removed and prepared for H&E sections as previously described. Apoptosis and mitosis were scored visually. Results are shown in fig 3.8.

Fig 3.8. Dose curve; small intestinal apoptosis for 5μg, 8μg and 10μg Fas ligand at 6 hours.

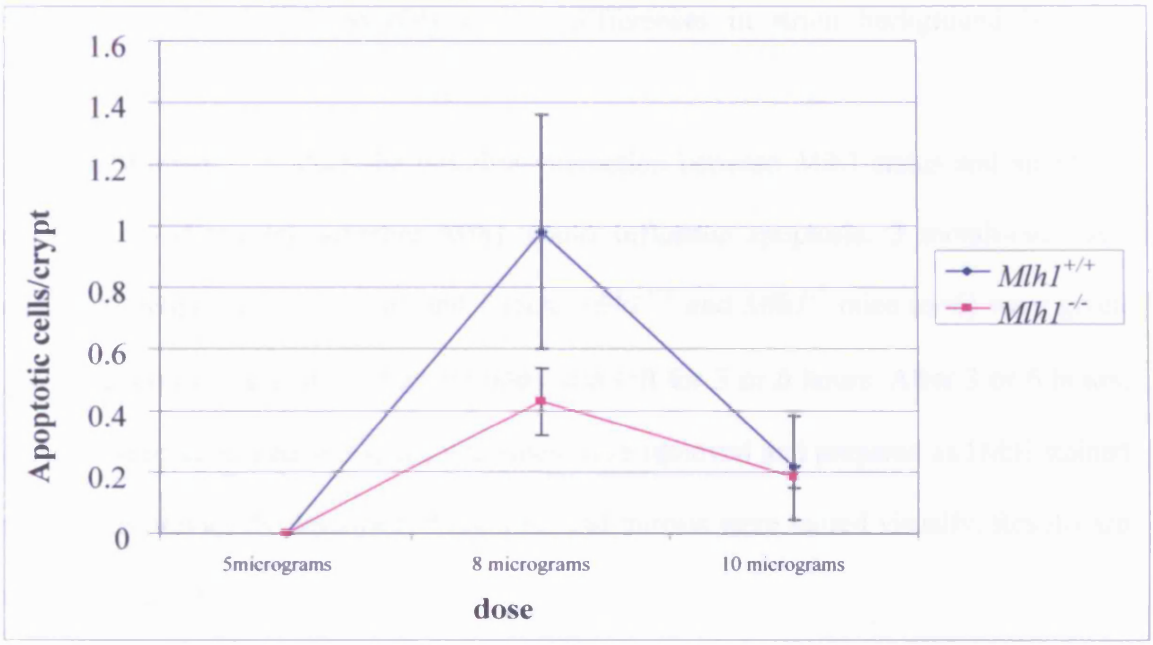


Fig 3.8. Dose curve of apoptosis at different levels of Fas administration in Mlh1 $^{+/+}$ and Mlh1$^{-/-}$ mice.

AT 5 μg Fas ligand, almost no apoptosis is seen, whereas at 10μg, tissue damage swaps the apoptotic signal. A significant difference in apoptosis between Mlh1$^{+/+}$ and Mlh1$^{-/-}$ mice is seen at 8μg for 6 hours, (p=0.038 Mann-Whitney) with loss of Mlh1 *causing a lowered apoptotic response to Fas ligand.*

Fig 3.8 shows the apoptotic response at varying doses of Fas ligand. At the 5μg dose, almost no apoptotic response was observed, making this dose too low for useful apoptotic index studies. At 10μg, tissue damage swamped the apoptotic response (the apparent low numbers of apoptoses observed in fig 3.8 at 10μg is due to the large amount of tissue damage reducing cell number to an extent where apoptosis is not seen.) Since the 8μg dose produced a significant difference in apoptosis between the genotypes (p= 0.038 Mann-Whitney) without causing tissue damage, the dose of 8μg Fas ligand was used in

the subsequent experiment. A dose of 10μg Fas ligand had been used in the previous experiment (on *Mbd4*$^{-/-}$ mice) but the *Mlh1*$^{+/+}$ and *Mlh1*$^{-/-}$ mice appear to be unusually sensitive to Fas ligand, possibly due to differences in strain background between colonies.

I next wanted to explore the possible interaction between Mlh1 status and apoptotic response, to determine whether Mlh1 could influence apoptosis. 3 month-old age matched, weight matched male and female *Mlh1*$^{+/+}$ and *Mlh1*$^{-/-}$ mice (n=6) were given one injection of 8μg anti-jo2 Fas antibody and left for 3 or 6 hours. After 3 or 6 hours, animals were sacrificed and small intestines were removed and prepared as H&E stained sections as previously described. Apoptosis and mitosis were scored visually. Results are shown in fig 3.9.

Fig 3.9. Loss of* Mlh1 *reduces the apoptotic response to Fas.

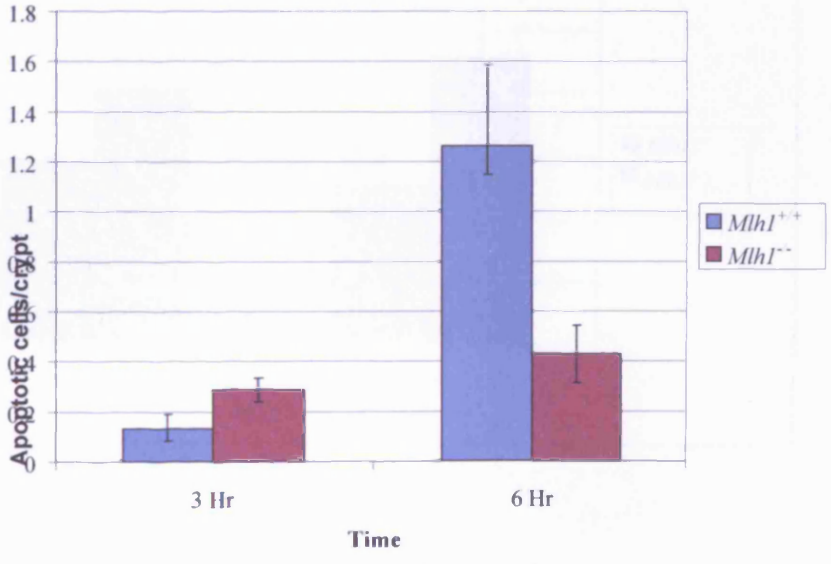


Fig 3.9. Loss of* Mlh1 *reduces the apoptotic response to Fas at 6 hours.

At 6 hours, a significantly lower apoptotic response to Fas is seen in Mlh1$^{-/-}$ *cells (p=0.0187 Mann-Whitney.) At 3 hours, a small but significant difference (p=0.019 Mann-Whitney) is seen.*

Fig 3.9 shows the lowered apoptotic response of *Mlh1*$^{-/-}$ cells to Fas ligand, scored as the number of apoptotic cells per crypt. A significant difference in apoptosis was seen at 6 hours. This difference is significant (p=0.0187 Mann-Whitney) and large (approximately 3-fold.) A smaller but still significant (p=0.0190) difference is seen at 3 hours.

Since *Mlh1* had affected the apoptotic response to Fas ligand, the mitotic response to Fas ligand was investigated using the same samples as the apoptosis experiment. Mitosis was scored visually as the number of mitotic figures per crypt. Results are shown in fig 3.10 below.

Fig 3.10. Loss of Mlh1 increases the mitotic response to Fas ligand.

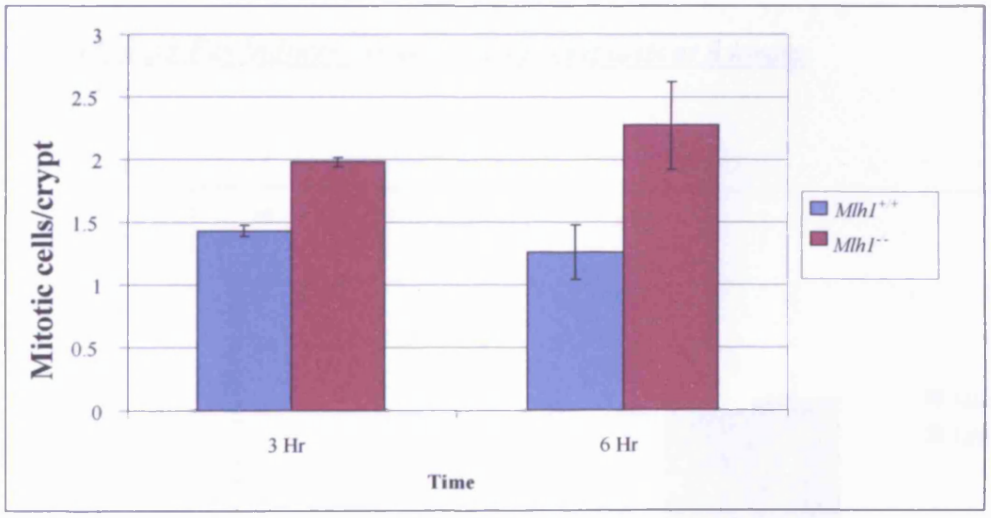


Fig 3.10. Loss of* Mlh1 *increases the mitotic response to Fas ligand.

Loss of Mlh1 *leads to a significant difference in mitosis in response to Fas at both 3 hours (p= 0.0259) and 6 hours (p= 0.0304 Mann-Whitney.)*

Mlh1$^{-/-}$ mice were found to have an increased mitotic response to Fas at both 3 and 6 hours. Both the increases were significant (3 hours, p=0.05, 6 hours p=0.018.) The loss of Mlh1 in the murine small intestine therefore, leads to a greater mitotic response to Fas, but a smaller apoptotic response. This implies that *MLH1* is involved in both cell proliferation and apoptosis signalling.

Since the loss of Mbd4 increased the amount of apoptosis seen in both the intestine and the liver at 6 hours after Fas ligand administration, I next wanted to determine whether the different (lower) apoptotic response to Fas ligand in the absence of Mlh2 was also seen in the liver. Liver samples were taken from the same mice as the Mlh1/Fas ligand intestinal experiment and stained with H&E. Apoptosis was scored visually as a percentage of cells in the liver undergoing apoptosis. Results are shown in fig 3.11.

Fig 3.11. 8 µg Fas induces extensive liver apoptosis at 6 hours.

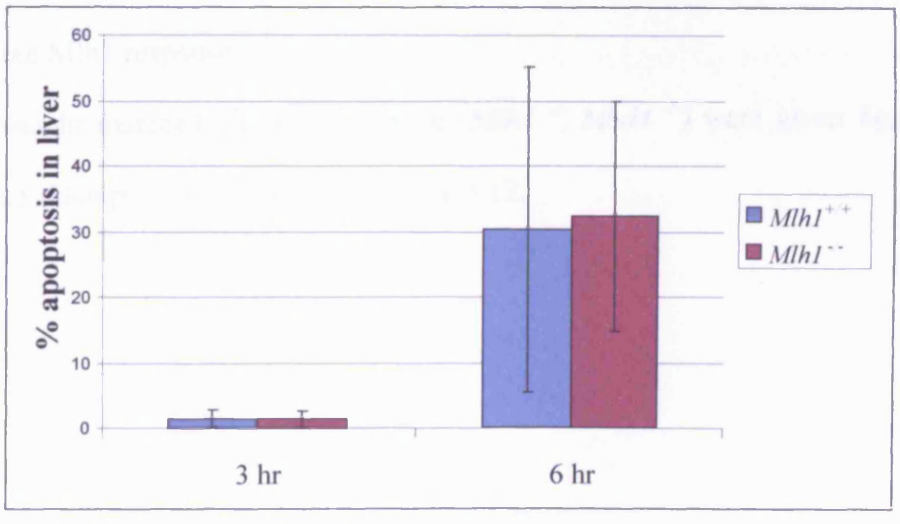


Fig 3.11. 8µg Fas induces extensive apoptosis in the liver at 6 hours, but there is no difference between Mlh1$^{-/-}$ *and Mlh1* $^{+/+}$ *mice (p=0.885 Mann-Whitney.)*

Fas ligand did induce apoptosis in *Mlh1*$^{-/-}$ and *Mlh1*$^{+/+}$ mice, with lower levels being seen at 3 hours and higher levels at 6 hours. However, no difference was seen between genotypes (p=0.885 Mann-Whitney) and there was a high degree of variability within the sample. This could be due to background effects of the outbred colony.

3.2.4. Characterising the apoptotic response in *Mbd4/Mlh1* double null mice.

The above results indicate that both Mbd4 and Mlh1 are mediating the apoptotic response to Fas ligand in the intestine. Therefore, in order to test if these genetic elements exist in the same pathway, the phenotype of mice deficient for both Mlh1 and Mbd4 was investigated. Examining the apoptotic response to Fas in mice doubly null for Mbd4 and Mlh1 would be highly informative. It would be expected that the double null mouse would have an increased apoptotic response to Fas ligand, due to loss of Mlh1 and Mbd4 which would leave large amounts of free Fadd in the cell. It might also be expected that the lack of Mbd4 would be dominant to the Mlh1 response.

Age and weight matched doubly null mice (*Mlh1* $^{-/-}$ *Mbd4* $^{-/-}$) were given 8μg Fas and sacrificed after 6 hours. Results are shown in fig 3.12.

Fig 3.12. Mlh1$^{-/-}$ Mbd4$^{-/-}$ mice show high variability in apoptosis rates; small intestine

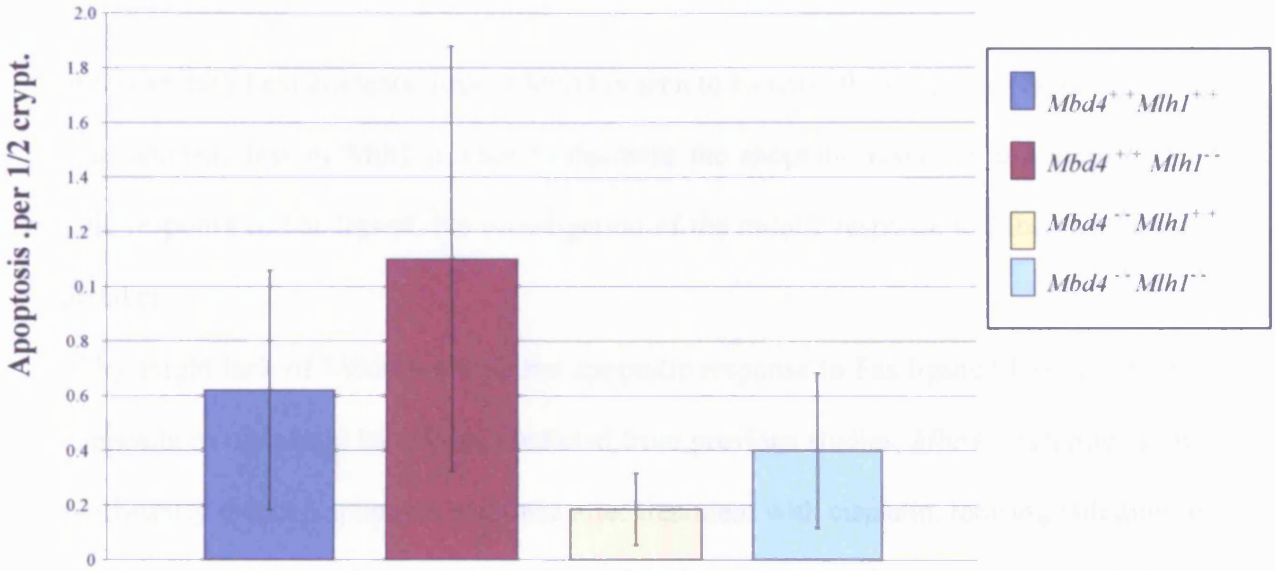


Fig 3.12. Administration of Fas induces apoptosis in the small intestine. However, no significant differences are seen between any of the genotypes.

Fig 3.12 shows the highly variable response to Fas in *Mbd4/Mlh1* double null mice. No significant differences were seen between any of the genotypes. This may be due to strain background effects of the double cross – each strain was outbred segregating C57/Bl6 and S129 genomes. Repeating and refining this experiment with a more inbred strain might eliminate this variation and could lead to greater insight into the interactions between Mlh1, Mbd4 and apoptotic signalling.

3.3. Discussion.

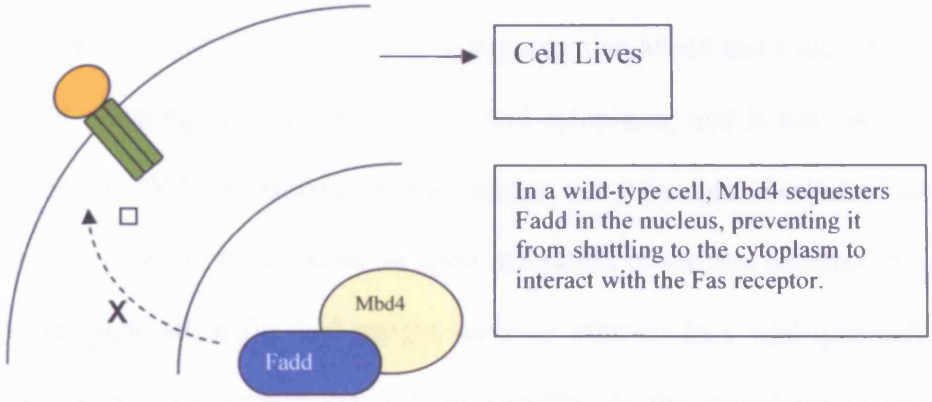
In this series of experiments, loss of Mbd4 is seen to increase the apoptotic response to Fas ligand, whereas loss of Mlh1 is seen to decrease the apoptotic response and increase the mitotic response to Fas ligand. No investigation of the mitotic response to Mbd4 status was undertaken.

Why might lack of Mbd4 increase the apoptotic response to Fas ligand? Indeed, exactly the opposite result would have been predicted from previous studies; *Mbd4*$^{-/-}$ null mice show a significantly reduced apoptotic response after treatment with cisplatin, ionising radiation or the alkylating agent temozolamide, all of which induce apoptosis in the intestinal epithelium via DNA damage (Sansom *et al* 2003a.) RNA damaging agents such as 5-FU, did not cause a differential response (Sansom *et al* 2003a) which implies that the reduced apoptotic response depends on the mode of damage, i.e. DNA-damaging agents induce an MBD4-dependent apoptotic response whereas RNA damaging agents do not (Sansom *et al* 2003a.) The nature of the signal is therefore likely to be an important determinant of the response; DNA damage can be viewed as an 'internal' signal, coming from within the cell. Fas ligand, on the other hand, is an 'external' signal, originating outside the cell and being transduced in a different manner (via the caspase cascade pictured in fig 3.2.)

New evidence from Screaton *et al* has shown that Fadd is primarily a nuclear protein, and is not based in the cytoplasm as previously thought (Screaton *et al* 2003.) They also show that Fadd interacts with Mbd4, suggesting a link between genome surveillance (the primary role of Mbd4) and apoptosis (Screaton *et al* 2003.) This suggests that Mbd4 and Fadd may act together to signal apoptosis, perhaps as pictured in fig 3.13.

Fig 3.13. Putative Mechanism of* Fadd/Mbd4 *Interaction.

a. Wild Type Cell.



b. MBD4-Deficient Cell.

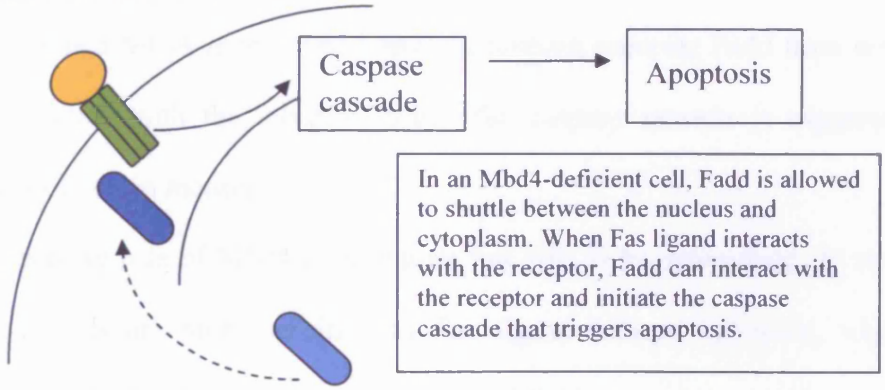


Fig 3.13. A putative mechanism to explain the higher apoptotic response in* Mbd4 *deficient cells.

Fadd is capable of shuttling between the nucleus and cytoplasm, and is held in the nucleus by Mbd4. *Evidence suggests that* Mbd4 *and* Fadd *interact (Screaton et al 2003.) When Fas ligand interacts with the receptor, a wild type cell is less likely to undergo apoptosis;* Fadd *is prevented from leaving the nucleus and so cannot initiate the apoptotic process. When* Mbd4 *is removed however, nothing prevents* Fadd *from entering the cytoplasm and interacting with the receptor, hence the caspase cascade is triggered and the apoptotic machinery is set in motion.*

This implies that Mbd4 may sequester Fadd in the nucleus until an appropriate signal induces it to release Fadd providing further evidence that Md4 plays a role in modulating the apoptotic response (Screaton *et al* 2003.) Evidence suggests that Mbd4 and Fadd interact and that Fadd is capable of shuttling between the nucleus and cytoplasm, and is held in the nucleus by Mbd4 (Screaton *et al* 2003.) Screaton *et al* also suggest that a fraction of cellular Fadd is exported from the nucleus (possibly by exportin 5) upon activation of the Fas receptor (Screaton *et al* 2003.) The model pictured in fig 3.13. might work as follows: In a wild-type cell, Mbd4 sequesters Fadd in the nucleus, preventing it from shuttling to the cytoplasm to interact with the Fas receptor. When Fas ligand interacts with the receptor, a wild type cell is less likely to undergo apoptosis; Fadd is prevented from leaving the nucleus and so cannot initiate the apoptotic process. When Mbd4 is removed however, nothing prevents Fadd from entering the cytoplasm and interacting with the receptor, hence the caspase cascade is triggered and the apoptotic machinery is set in motion.

The precise role of Mbd4 in apoptosis has still to be determined. *In vivo*, *Mbd4*$^{-/-}$ liver and intestinal cells are more sensitive to Fas ligand-induced apoptosis, whereas MEF (Murine embryonic fibroblast) cells over expressing Mbd4 were less sensitive to Fas ligand-induced apoptosis (Screaton *et al* 2003.) Isolated small intestinal crypts are also less sensitive to anoikis. Perhaps the levels of free Mbd4 are critical to the nature of the response, but the data do show unequivocally that Mbd4 influences the apoptotic response.

3.3.1. Interaction between Mlh1 and Mbd4.

As well as playing a major role in removing common DNA lesions such as O6-MeG, MMR also modulates the response to a range of other DNA damaging agents such as 5-FU,

topoisomerase inhibitors and cisplatin (Sansom *et al* 2003a, Cortellino *et al* 2003.) It is thought that repeated attempts by Mbd4 to repair a lesion may in some cases trigger a long patch repair requiring the MMR machinery, and that this may be recruited via the association of Mlh1 with Mbd4 (Cortellino *et al* 2003.)

The interaction between MMR and Mbd4 is still largely unclear; loss of Mbd4 does not increase tumour onset, spectrum or MSI in MMR-deficient animals, showing that Mbd4 does not affect MMR-dependent tumorigenesis (Sansom *et al* 2004.) Intruigingly, however, MMR proteins have been found to be reduced in Mbd4-deficient cells, and this appears to be a post-translational effect; could they be unstable in the absence of Mbd4?

Mlh1 has been found to be associated with Fadd (Screaton *et al* 2003.) Could Mbd4 be associated with Fadd and Mlh1 in a large multi-protein complex that signals to the apoptotic pathway in response to a certain level of DNA damage, or Fas-mediated death? One putative model for such an interaction is shown in figure 3.14. If Mbd4, Fadd and Mlh1 are complexed together, via death domain interaction, there could be competition between the three proteins such that they are held together in an unstable triumvirate, with the stability and strength of the interactions being determined by Mbd4. In the presence of Mlh1, the Mbd4/Mlh1/Fadd complex is less stable, and there is more free Fadd in the cell. When the Fas receptor is activated, this Fadd can bind the receptor and activate the caspase cascade. This in turn would create double strand breaks in DNA which would be sensed further down the pathway by Mlh1 and the cell would undergo apoptosis. In the absence of Mlh1, the Mbd4/Fadd interaction would be strengthened, and less free Fadd would be available, reducing the chances of an activated Fas receptor being able to initiate an apoptotic cascade. The cell would be more likely to live.

The increased apoptosis seen in *Mbd4* $^{-/-}$ mice implies that the Mlh1/Fadd combination is weaker than the whole complex (freeing more Fadd.) Conversely, the decreased apoptosis seen in *Mlh1* $^{-/-}$ mice implies that the Mbd4/Fadd complex is stronger than the whole complex, freeing less Fadd and resulting in lowered levels of apoptosis.

Investigation of the binding strengths of these proteins *in vitro* and *in vivo* could provide evidence for or against this model. It would also be informative to use immunohistochemistry to attempt to localise Mbd4, Mlh1 and Fadd in the cell as it is exposed to the Fas ligand signal; it might be expected that the proteins would shuttle in and out of the nucleus in response to the signal.

Fig 3.14. A putative model for the interaction between Fadd, Mbd4 and Mlh1 in response to Fas ligand.

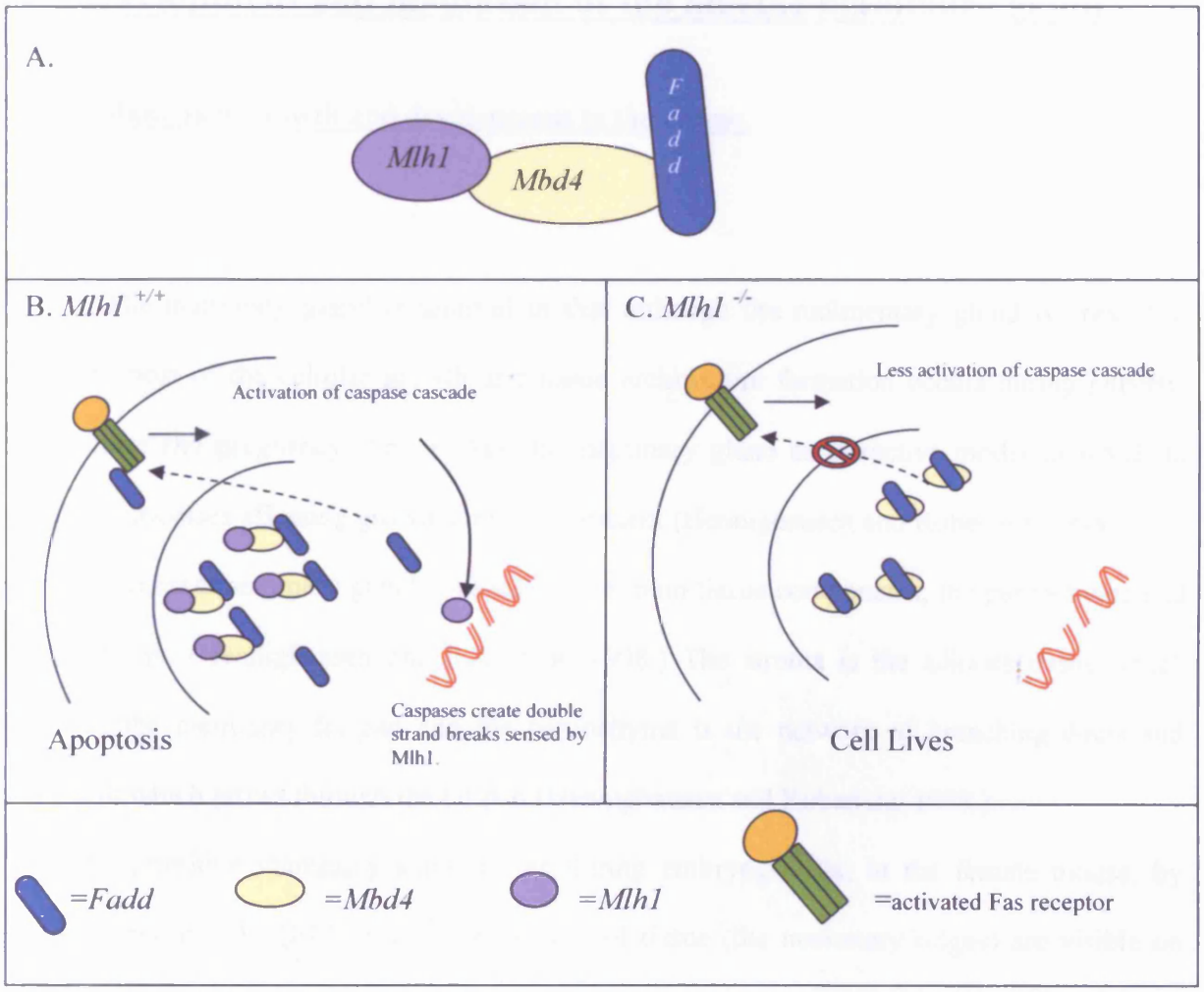


Fig 3.14. A putative model for the interaction between Mbd4, Mlh1 and Fadd in death receptor signalling.

A. Fadd, Mbd4 and Mlh1 are known to interact via death domains.

B. In the presence of Mlh1, the Mbd4/Mlh1/Fadd complex is less stable, and there is more free Fadd in the cell. When the Fas receptor is activated, this Fadd can bind the receptor and activate the caspase cascade. This in turn creates double strand breaks in DNA which are sensed further down the pathway by Mlh1.

C. In the absence of Mlh1, the Fadd/Mbd4 interaction is more stable. Less free Fadd is present, and so less Fadd is available to activate the caspase cascade which leads to apoptosis if the Fas receptor is stimulated and so cell is more likely to survive.

Chapter 4: Investigating the Role of MeCP2 in the development and involution of the murine mammary gland.

4.1. Mammary growth and development in the mouse.

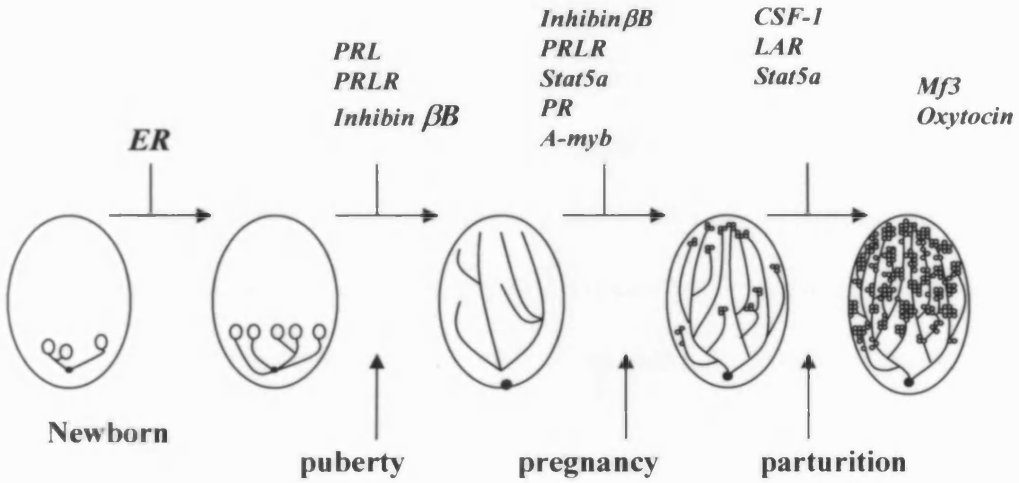
The mammary gland is unusual in that although the rudimentary gland is present at birth, most of the cellular growth and tissue architecture formation occurs during puberty, lactation and pregnancy. This makes the mammary gland an attractive model in which to study processes affecting growth and development (Hennighausen and Robertson 1998.)

The murine mammary gland consists of two main tissue components; the parenchyma and the stroma (Hennighausen and Robertson 1998.) The stroma is the adipose tissue which forms the mammary fat pad and the parenchyma is the network of branching ducts and alveoli which grows through the fat pad (Hennighausen and Robertson 1998.)

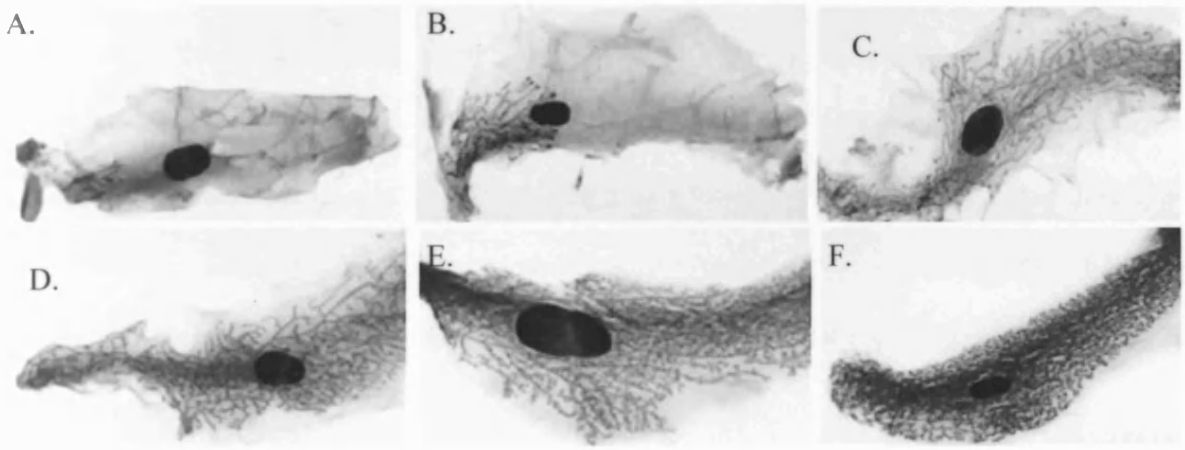
The primitive mammary gland forms during embryogenesis; in the female mouse, by embryonic day 11 (E11) Two raised ridges of tissue (the mammary ridges) are visible on either side of the ventral midline (Hennighausen and Robertson 1998.) By day 12, cells within the ridges have concentrated to form two rows of five mammary buds (Hennighausen and Robertson 1998.) These buds proliferate and grow rapidly just before birth, to form the mammary cord, which opens into the nipple externally, and branches into the rudimentary gland internally (Hennighausen and Robertson 1998.) so that by birth, several ducts are present in the gland (Hennighausen and Robertson 1998) (Fig 4.1.)

Fig 4.1. Growth and Development of the Mouse Mammary Gland.

i. *(Adapted from Henninghausen and Robinson 1998)*



ii.



(Taken from http://mammary.nih.gov/atlas/wholemounts/normal/index.html)

Fig 4.1.i Stages in the development and growth of the mouse mammary gland, showing the gene products required for each stage of development. ER=oestrogen receptor, PRL=prolactin, PRLR= prolactin receptor, PR= progesterone receptor, CSF-1= colony stimulating factor 1.

Fig 4.1. ii. Wholemount analysis of mammary gland growth in the mouse. A = newborn, B= 4 week virgin, C= 6 week virgin, D= 10 week virgin, E = pregnancy day 9, F = pregnancy day 16.

During the first few weeks of life, the gland grows isometrically (i.e. it keeps pace with and does not exceed the growth of the animal and the mammary ducts grow and branch slowly (Hennighausen and Robinson 1998.) At the leading end of each duct is a structure called the terminal end bud (TEB.) TEBs are highly dynamic and proliferative structures and are made of up of two cell types; cap cells and body cells (Humphreys *et al* 1987, Daniel and Silberstein 1987.) Cap cells form a thin outer layer and interact with the surrounding stroma through a thin basal lamina, and around 6-10 layers of body cells form the interior of the TEB (Humphreys *et al* 1987, Daniel and Silberstein 1987.) Cap cells are the precursors of myoepithelial cells in the mature gland, while body cells give rise to epithelial cell lineages (Humphreys *et al* 1987.)

At around four weeks of age, the mouse enters puberty, and the rate of growth increases markedly, exceeding the rate of growth of the animal. The TEBs drive through the stroma, influenced by various growth factors and steroid hormones, principally oestrogen and EGF (epidermal growth factor), to create the branching pattern of ducts seen at this stage of development (Humphreys *et al* 1997, Cunha *et al* 1997.) During puberty, the ductal tree continues to grow until it has reached the periphery of the gland, whereupon the TEBs form terminal ductal structures with low mitotic activity (Humphreys *et al* 1987, Daniel and Silberstein 1987, Hennighausen and Robinson 1998.)

Between adolescence and pregnancy, the cells of the mammary gland are mitotically dormant and undifferentiated (Hennighausen and Robinson 1998.) However, during pregnancy, the gland must be remodelled to produce and secrete the milk that nourishes the young. To accomplish this, a huge amount of cellular growth is needed. Oestrogen, progesterone and other, placental hormones stimulate extensive branching of the ductal tree – the mammary epithelium expands to fill the stroma between the ducts and extensive lobulo-

alveolar proliferation and differentiation occurs (Humphreys *et al* 1987, Daniel and Silberstein 1987, Hennighausen and Robinson 1998.) This proliferation extends into the early period of lactation, with about 20% of total mammary growth occurring during this time (Hennighausen and Robertson 1998.) When the young mice are weaned, the gland is remodelled once again to resemble the virgin-like state, in the process of involution (Hennighausen and Robertson 1998.) Various signals, including stasis of the milk in the ducts, trigger involution which occurs in two distinct phases; firstly, apoptosis of the secretory epithelial cells, then apoptotic degradation of the lobular alveolar structure and the mammary basement membrane, accompanied by degradation of the extracellular matrix. The epithelium is replaced by adipose tissue and the gland resembles that of a mature virgin female. The process of growth, lactation and involution repeats itself for each pregnancy (Hennighausen and Robertson 1998.)

The dynamic changes which occur in the mammary gland during development, pregnancy and involution are under tight genetic and hormonal control. Disruption of the fine control of mammary morphogenesis can manifest itself as alterations in the growth and structure of the gland, as well as altering the timing and manner of involution. This project will assess the role of MeCP2 in the developing mammary gland and the role of MeCP2 in remodelling the gland during involution.

4.2. Methyl Binding Proteins in the Developing Mammary Gland.

Levels of two MBD proteins; MECP2 and MBD2, have been measured in the human foetal mammary gland (Billard *et al* 2002) (see fig 4.2.) Levels of *MECP2* RNA were found to increase gradually through gestation and reach adult levels at 31 weeks - around the same time as mammary morphogenesis is complete (33 weeks.) This suggests that MECP2 may

play a role in the developing mammary gland (Billard *et al* 2002.) Levels of MBD2 mRNA rise gradually through foetal mammary development, but remain lower at birth than in the adult tissue, which suggests that MBD2 does not play a role in mammary development (Billard *et al* 2002.)

Fig 4.2. Levels of MECP2 and MBD2 RNA in the Developing Human Mammary Gland.

From (Billard *et al* 2002.)

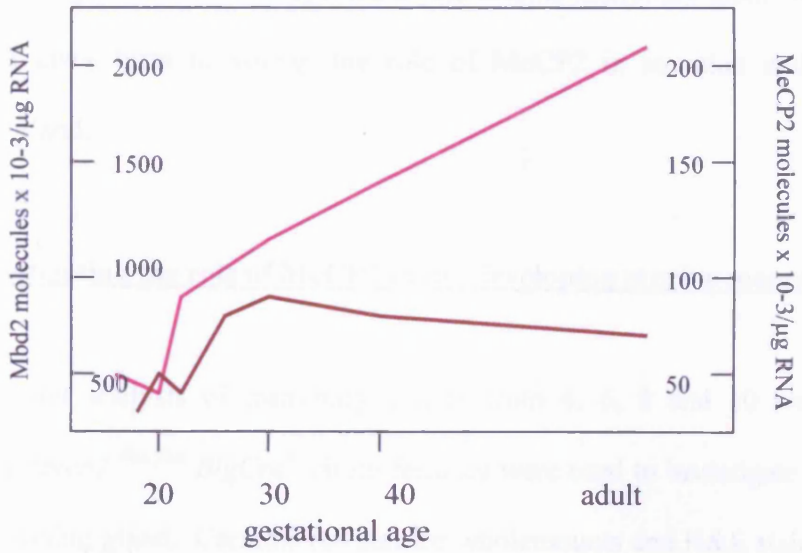


Fig 4.2. Levels of MBD2 and MECP2 mRNA during mammary development in the mouse.

Levels of MECP2 mRNA (brown line) in the mammary gland peak during mammary development, whereas levels of Mbd2 mRNA (pink line) continue to rise after birth into adulthood. This suggests that MECP2, but not MBD2, might play a role in mammary growth and development.

4.3. Investigating the role of methyl CpG binding domain proteins in the mammary gland.

In order to investigate the role of MeCP2 in the developing gland, a line of mice carrying both a conditional, LoxP-flanked allele of *Mecp2* and a mammary-specific Cre recombinase were generated by crossing *Mecp2*$^{flox/flox}$ and *Mecp2* $^{wt/wt}$ females onto a mouse line carrying

the mammary-specific β-lactoglobulin Cre (Blg-Cre (Selbert *et al* 1998)) to form a line carrying various alleles of the two transgenes. Blg is expressed during development at a lower level (7% Cre mediated deletion) and at a higher level during lactation (70-80% Cre-mediated deletion.) (Selbert *et al* 1998.) This means that the Cre-driven recombination and removal of MeCP2 will occur only in the mammary gland at a slightly lower level during development and at a high level during lactation. This will allow the analysis of the role of MeCP2 in the growth and development of the murine mammary gland, and, if the mouse is mated and gives birth to young, the role of MeCP2 in lactation and involution of the mammary gland.

4.3.1. Investigating the role of MeCP2 in the developing murine mammary gland.

Wholemount analysis of mammary glands from 4, 6, 8 and 10 week old *Mecp2* $^{+/+}$ *BlgCre*$^{+}$ or *Mecp2* $^{flox/flox}$ *BlgCre*$^{+}$ virgin females were used to investigate the role of MeCP2 in the developing gland. Carmine red stained wholemounts and H&E stained sections of 6, 8 and 10 week old *Mecp2* $^{flox/flox}$ *BlgCre*$^{+}$ and *Mecp2* $^{+/+}$ *BlgCre*$^{+}$ glands were examined for developmental criteria such as the extent of penetration of the ductal system through the fat pad and the extent of branching of the ductal system. Results are shown in fig 4.3.

Fig 4.3. Loss of Mecp2 does not alter the development of the mammary gland at 4 weeks.

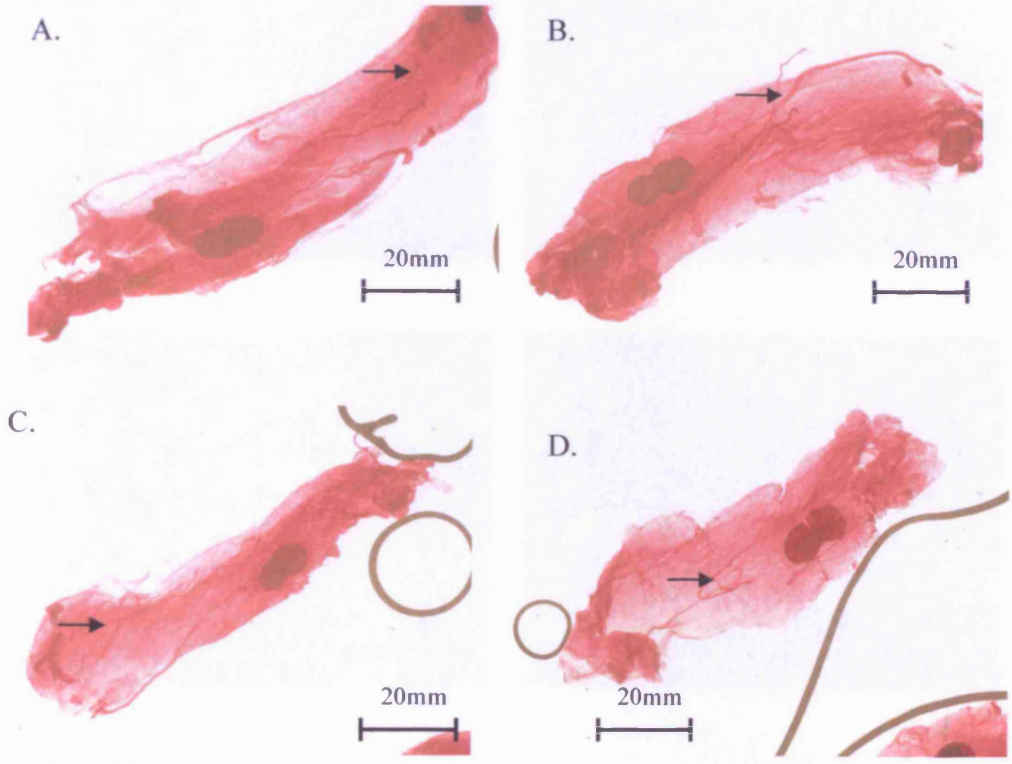


Fig 4.3. Loss of Mecp2 does not affect the growth of the mammary gland.

A and B = $Mecp2^{wt/wt}Blg\text{-}Cre^{+}$. C and D = $Mecp2^{flox/flox}Blg\text{-}Cre^{+}$. At 4 weeks of age there is no visible difference between $Mecp2^{flox/flox}Blg\text{-}Cre^{+}$ and $MecpP2^{wt/wt}Blg\text{-}Cre^{+}$ female mammary glands. Both wild type and Mecp2 null glands contain similar epithelial trees with a similar degree of branching (arrows) which have extended into the fat pad.

At 4 weeks of age, no difference was seen between the mammary development of $Mecp2^{flox/flox}Blg\text{-}Cre$ and $Mecp2^{wt/wt}Blg\text{-}Cre^{+}$ mice. The extent of penetration of the epithelial tree into the fat pad was similar, and no differences were seen in gross morphology (e.g the degree of branching of the tree.)

Fig 4.4. Loss of* MeCP2 *does not alter the mature adult appearance of the virgin mammary gland.

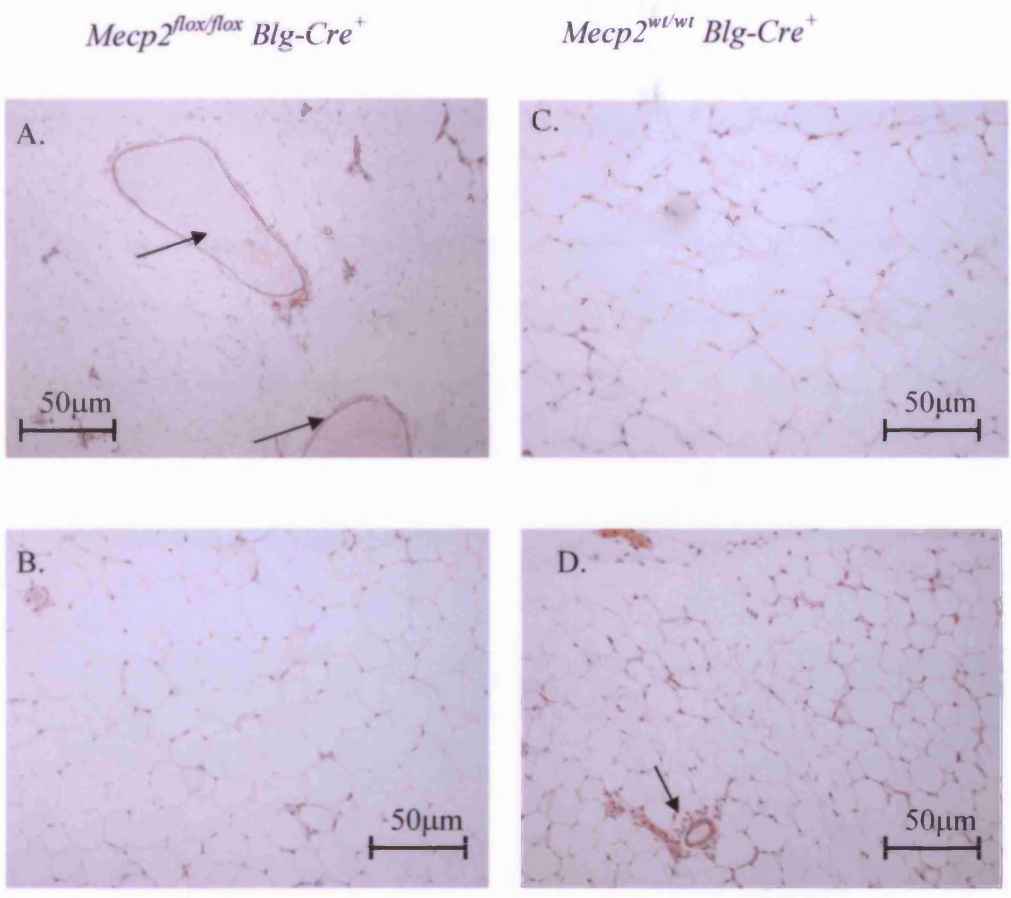


Fig 4.4. Loss of Mecp2 does not appear to affect the final appearance of the murine mammary gland.

A,C= Mecp2$^{flox/flox}$ Blg-Cre^{+} , B,D = Mecp2$^{wt/wt}$ Blg-Cre^{+} , both virgin glands are composed mainly of adipose tissue, with the epithelial ductal tree running through it. A shows that very rare anomalous structures were seen (in two out of three Mecp2$^{flox/flox}$ Blg-Cre^{+} mice.)

The mature adult appearance of both *Mecp2$^{flox/flox}$ Blg-Cre^{+}* and *Mecp2$^{wt/wt}$ Blg-Cre^{+}* mammary glands was largely similar, being composed mainly of adipose tissue containing a small amount of epithelial tissue (the epithelial ducts) running through the fat pad (fig 4.4d, arrow.) The loss of MeCP2 does not appear to grossly affect the development of the mammary gland; very occasional areas of slight abnormality were seen but these were not present in all mice examined. In the light of this result, no further developmental studies were carried out.

4.4. Investigating the Role of MeCP2 in the Involution of the Mammary Gland.

Preliminary investigation of a multiparous *Mecp2*$^{flox/flox}$ *Blg-Cre*$^{+}$ female mouse that had undergone several rounds of pregnancy, lactation and natural involution, appeared to imply that involution may be delayed in *Mecp2* $^{flox/flox}$ *BLG-Cre*$^{+}$ females (Fig 4.5.)

Fig 4.5. Preliminary investigation of the involution of the Mecp2-null mammary gland.

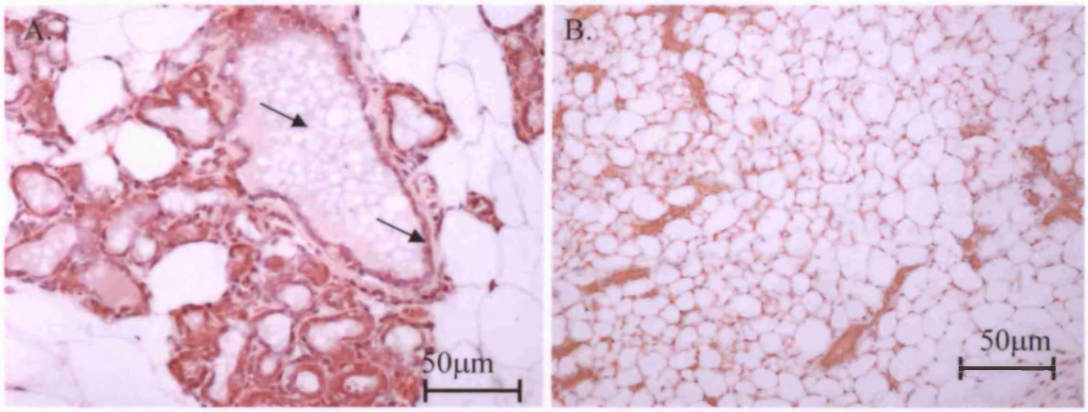


Fig 4.5. *A= multiparous Mecp2* $^{flox\ flox}$ *Blg-Cre*$^{+}$ *female mammary gland, 19 days after weaning of last litter.*

B = Mecp2$^{wt\ wt}$ *Blg-Cre*$^{+}$ *fully-involuted female mammary gland (21 days.) The wild-type gland is fully involuted and mainly consists of adipose tissue with only small amounts of residual epithelium. By contrast, the Mecp2-null gland contains thickened epithelium and milk remaining in the ductal structures (arrows.)*

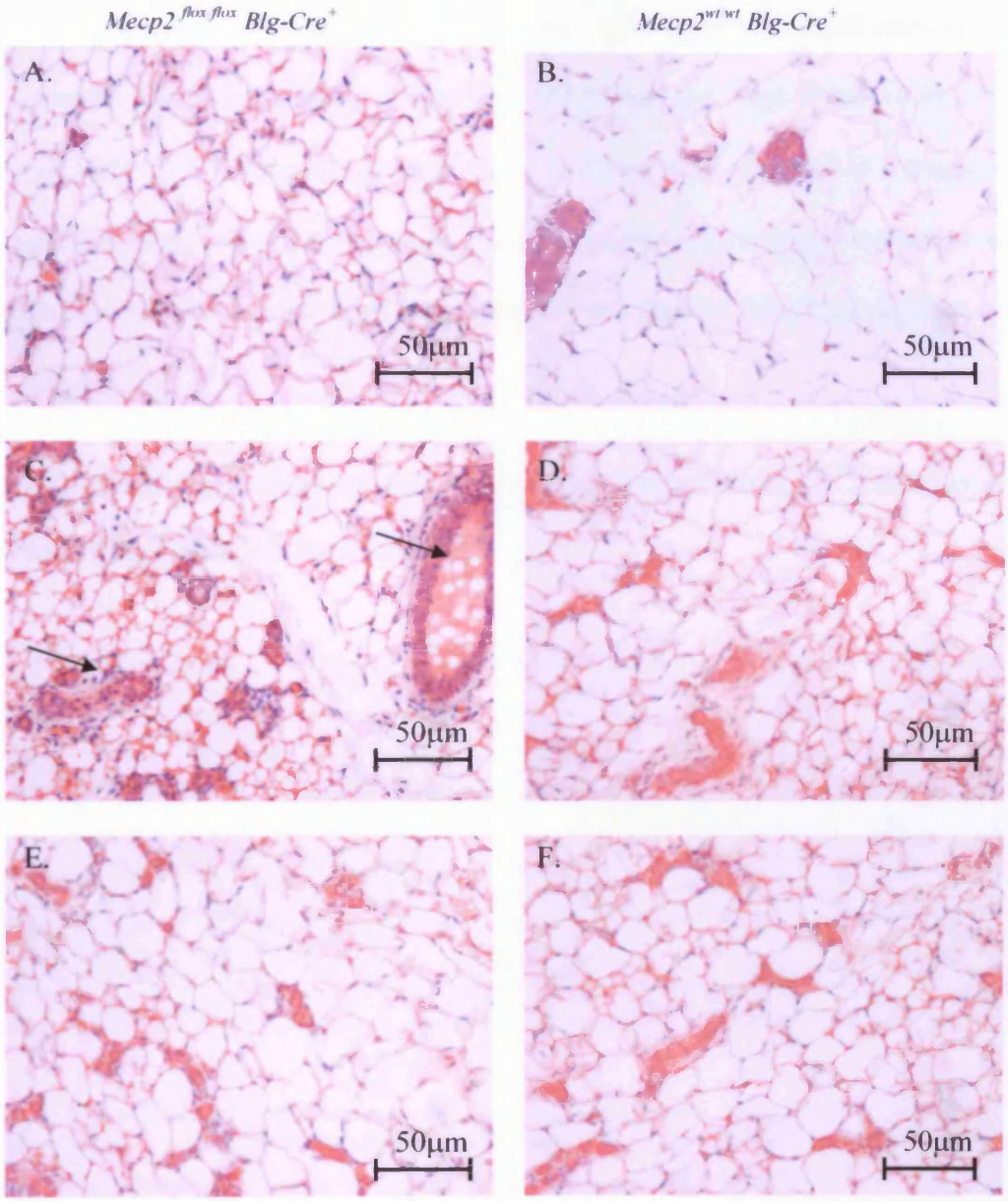
The *Mecp2* $^{flox/flox}$ *Blg-Cre*$^{+}$ gland does not appear to have fully involuted, and still clearly showed thickened epithelial ductal structures which contained static milk remaining in the duct. The *Mecp2*$^{wt/wt}$ *Blg-Cre*$^{+}$ gland, by contrast, is fully involuted and has remodelled to a

The *Mecp2* $^{flox/flox}$ *Blg-Cre*$^{+}$ gland does not appear to have fully involuted, and still clearly showed thickened epithelial ductal structures which contained static milk remaining in the duct. The *Mecp2*$^{wt/wt}$ *Blg-Cre*$^{+}$ gland, by contrast, is fully involuted and has remodelled to a virgin-like state; the majority of the epithelial tissue has been removed, with only small areas remaining and with adipose tissue dominating the gland. The apparent perturbation in involution implied that MeCP2 could play a role in the involution procees and that the apparent mammary phenotype of the *Mecp2* $^{flox/flox}$ *BLG-Cre*$^{+}$ mouse warranted further investigation.

In order to investigate the possible role of MeCP2 in the involution of the mammary gland, the glands of *Mecp2* $^{flox/flox}$ *Blg-Cre*$^{+}$ *and Mecp2*$^{wt/wt}$ *Blg-Cre*$^{+}$ mice were observed at 21 days post removal of pups. Involution is triggered by the removal of the litter which has been suckling normally for 10 days. 21 days is considered to be 'full involution.' This allows the final outcome of involution to be examined. If any changes were observed, further timepoints at 3, 6 and 10 days involution would be added.

3 *Mecp2 f*$^{lox/flox}$ *Blg-Cre*$^{+}$ and 3 *Mecp2*$^{wt/wt}$ *Blg-Cre*$^{+}$ 3 month old age matched female mice, segregating for C57/Bl6 and S129 genomes were mated to *Mecp2*$^{flox/y}$ *BLG-Cre*$^{+}$ male mice. Male mice were removed from the mating cage when the first signs of pregnancy were visible and female mice were allowed to give birth and suckle young naturally for 10 days. Ideally, each female should suckle the same number of pups (6-8) and so a number of 'foster mothers' were used to foster excess pups and to provide pups to mothers whose litters were less than 6. After 10 days, the pups were removed from the experimental female to trigger the process of involution (pups were transferred to foster mothers for weaning.) Mice were then left for 21 days, culled and the mammary glands removed and prepared as H&E stained sections for analysis. Results are shown in Fig 4.6.

Fig 4.6. 21 Day involution; loss of Mecp2 does not grossly affect the involution process.



A,B,C = Mecp2 $^{flox\,flox}$ Blg-Cre^{+}. D,E,F = Mecp2$^{wt\,wt}$ Blg-Cre^{+} mice. H&E stained slides show that most of the mammary gland has been remodelled in both Mecp2 $^{flox\,flox}$ Blg-Cre^{+}. D,E,F = Mecp2$^{wt\,wt}$ Blg-Cre^{+} mice by 21 days after removal of pups. D,E and F (wild-type) show that most of the epithelium has been replaced by adipose tissue. Most Mecp2 deficient glands had the appearance of A, some showed small areas of remaining epithelium and milk remaining in the ducts (arrows,C.)

In both genotypes, the gland had reverted to a virgin-like state, dominated by adipose tissue and with only small areas of epithelium remaining. Very occasional small areas of epithelium were observed, but these were at a very low frequency and were observed in 2 of 3 mice examined. Only 2 or 3 areas of approximately 30-50μm were observed at a frequency of 2-3 per gland. Although this could represent a mild phenotype occurring a low penetrance in the absence of Mecp2, the rare nature of these areas did not warrant frther investigation.

Fig 4.7. Some abnormal areas are seen in* Mecp2-*deficient (Mecp2$^{flox/flox}$ Blg-Cre^{+}) mammary glands after 21-days involution.

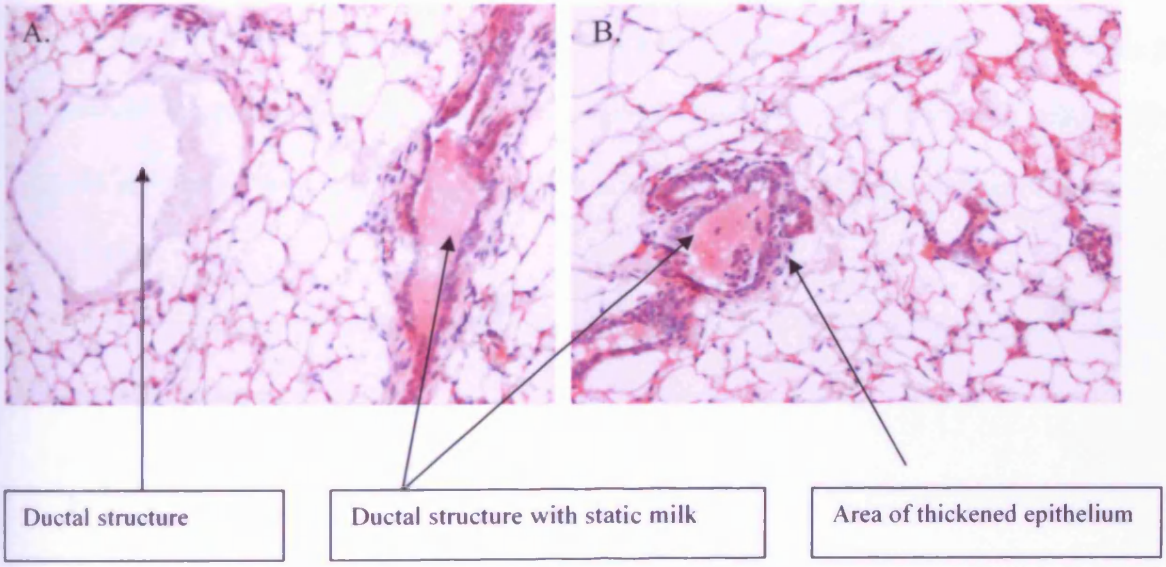


Fig 4.7. Abnormalities seen in* Mecp2-*deficient (Mecp2$^{flox/flox}$ Blg-Cre^{+}) mammary glands.

Very occasional areas of not fully involuted gland were observed in 2 Mecp2 $^{flox/flox}$ Blg-Cre animals. However, these areas were not present in all animals and occurred at very low frequency.

Chapter 5: Characterising the Phenotype of Loss of MeCP2 in the Murine Small Intestine.

5.1. *MECP2* and Rett syndrome.

Rett syndrome (OMIM #312750) is a neurodevelopmental disorder which mainly affects girls, caused by mutations in the methyl CpG binding protein 2 (*MECP2)* gene (Amir *et al* 1999.) After a short (6-18 months) period of normal development, patients enter a regressive phase, losing acquired speech and motor skills and developing a range of symptoms including autistic traits, scoliosis, and characteristic 'hand wringing' movements (Neul and Zoghbi 2004.)

The condition is considered to be purely confined to the nervous system and a mouse model of Rett (RTT) in which *Mecp2* is conditionally deleted in the nervous system produces a virtually identical phenotype to a model with a whole body *Mecp2* deletion (Chen *et al* 2001.)

The exact mechanisms whereby loss of functional MECP2 causes the symptoms of RTT are not known, although the bulk of current research suggests a fault in neuronal maturation, patterning and survival (Neul and Zoghbi 2004.) This places the primary mechanism of RTT pathology firmly within the CNS; however, patients with RTT can have a wide spectrum of symptoms, and frequently present with symptoms which are difficult to explain if RTT is exclusively CNS-based, such as bone abnormalities (Leonard *et al* 1995) and gastrointestinal problems (Motil *et al* 1999.) Despite these studies providing a sound rationale for examining the role of MeCP2 in non-nervous tissue, little research on the extra-CNS roles of MeCP2 has been carried out.

5.1.1 Methyl CpG binding proteins and the intestine.

Other methyl CpG binding proteins have been found to have roles in the intestine, in modulating the tumourigenic response (Sansom *et al* 2003a,c, Millar *et al* 2002.) Loss of the methyl CpG binding domain protein, Mbd2 in an intestinal-tumour prone $Apc^{Min/+}$ background results in a significant reduction in tumour burden and mortality, even though the *Mbd2* null ($Mbd2^{-/-}$) allele alone does not produce a discernible phenotype in the intestine (Sansom *et al* 2003c.) Mbd4 deficiency has been shown to accelerate tumourigenesis on an Apc^{Min} background (Millar *et al* 2002.) This shows that the MBD proteins have a potential role in the normal intestine; mouse models of RTT are available (Guy *et al* 2001) and this, along with the observed gastrointestinal problems seen in RTT patients, provides a further rationale for examining the role of MeCP2 in the murine intestine.

5.1.2. The small intestine as a model system; development of the gut.

The small intestine is an ideal model in which to study the growth, development and normal turnover of a system as it consists of a number of repeating basic units, the crypt-villus unit, which are well characterised and easy to study (Potten *et al* 1997.) Patterns of tissue organisation, cell proliferation, migration and apoptosis are all well-studied, and so if the loss of MeCP2 in the gut results in changes to any of these factors, they can be observed and quantified.

The development of the small intestine is also well characterised; the embryonic ectoderm forms the lining of the digestive tract and the liver, gall bladder and pancreas (Clatworthy *et al* 2001.) These endodermal cells form only the lining of the digestive tube – the musculature and connective tissue that surround it are derived from the lateral plate mesoderm (Clatworthy *et al*

2001.) From embryonic day E14.5 to E18.5 the murine small intestine develops from a stratified epithelium into a simple columnar epithelium, accompanied by the formation of nascent villi in a proximal to distal progression (Clatworthy *et al* 2001.) At this stage of development there are no crypts, only villi separated by an intervillus epithelium that gives rise to the crypt structures after birth (Kaestner *et al* 1997.)

Crypts develop after birth (days 1-5) from the flat intervillus epithelium with mitotic cells segregating in these structures (Clatworthy *et al* 2001.) As the crypts elongate, they also undergo replication by bifurcating from the base (Clatworthy *et al* 2001.) The epithelium differentiates along the cephalo-caudal axis resulting in morphological differences along the length of the duodenum, ileum and colon (Clatworthy *et al* 2001.)

5.1.3. Normal structure of the small intestinal epithelium.

The small intestine needs to maximise its internal surface area for optimal absorption of nutrients and water. This is achieved by the inner epithelium being folded and lined with millions of tiny projections (villi) and pits (crypts of leiberkuhn) (fig 5.1.) The surface area is expanded even further by the presence of microvilli on the lumenal surface of the cells which form the villi (Clatworthy *et al* 2001.)

The intestinal epithelium is a dynamic tissue with a large proportion of cycling cells. In the mouse, each one of the approximately 1.1 million crypts contains around 250 epithelial cells (Potten *et al* 1997.) A small number of stem cells are situated around position four in the crypt (i.e. the 4^{th} cells from the base of the crypt.) Asymmetric division of these cells forms two sub-populations; cells which replenish the stem cell compartment and transit amplifying (TA) cells (Hermiston *et al* 1996.) There are about 150 TA cells per crypt, which cycle roughly once every 24 hours (Hermiston *et al* 1996) forming the so-called proliferative or transit amplifying (TA)

zone in the middle of the crypt (Potten *et al* 1997.) TA cells differentiate into one of four cell lineages (absorptive enterocytes, goblet cells, enteroendocrine cells and paneth cells) and migrate up the villi, with the exception of Paneth cells, which migrate down to the base of the crypts to complete their differentiation (Mills and Gordon 2001.) Recently, it has been proposed that the active multipotent stem cell divides to two different types of cell in the proliferative zone; a C_o lineage which forms the absorptive enterocytes, and an M_o lineage from which arises the secretory cells (paneth, goblet and enteroendocrine) (Bjerknes and Cheng 2001.)

Fig 5.1. The structure of the small intestine: the crypt-villus unit.

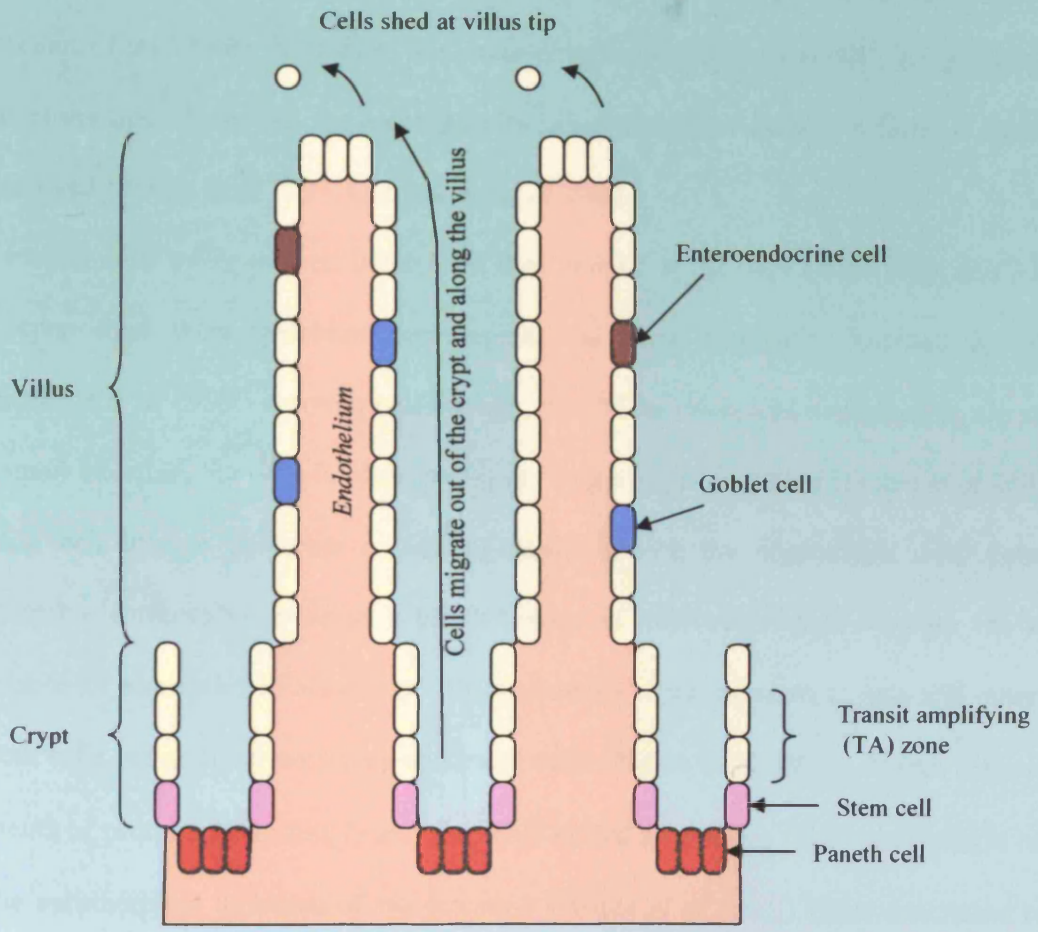


Fig 5.1. The crypt-villus unit of the small intestine.

The crypt contains 22 cells along each side from base to villus junction, and the villus contains around 80 cells from crypt junction to tip (Potten et al 1997.) Cells are 'born' in the crypt, originating from division of stem cells at position 4, proliferate in the transit amplifying (TA) zone, and then differentiate as they move up out of the crypt and along the villus. Paneth cells move down into the bases of the crypts as they differentiate. Cells move up the villus and are lost by shedding or apoptosis at the tip (Potten et al 1997.)

Cells migrate up the villus and are eventually sloughed off into the lumen of the gut (Potten *et al* 1997, Clatworthy *et al* 2001.) The effect of the rapidly cycling epithelium is to create a continuous flow of new cells from the lower crypts to the tips of the villi; the question of whether cells at the tips of the villi are mechanically sloughed off or undergo a form of apoptosis is still unresolved (Potten *et al* 1997, Clatworthy *et al* 2001.)

Paneth cells differ in their behaviour; they remain in the base of the crypt and have a longer life span then their epithelial counterparts, and are eventually removed by phagocytosis (Hermiston *et al* 1996.) Another cell lineage, the 'M' or microfold lineage cells, are also found in the small intestine, but only in close proximity to the Peyer's patches (Potten *et al* 2001.)

Each cell lineage performs a distinct function with the crypt-villus unit. Epithelial cells (absorptive enterocytes) possess a brush border of microvilli which increase the surface area available for absorption (Potten *et al* 2001) allowing rapid transport of ions and other molecules. Paneth cells are granular secretory epithelial cells (Potten *et al* 2001.) Paneth cells secrete large amounts of proteins including lysozyme, TNF-α and cryptidins, which are thought to contribute to the antimicrobial defences of the intestine (Potten *et al* 2001.) Enteroendocrine cells secrete neuropeptides, whereas goblet cells are flask-shaped and have no apical microvilli but instead produce and secrete the mucins that lubricate and protect the gut (Potten *et al* 2001.)

Rates of proliferation and loss must be tightly controlled; too little proliferation would compromise the absorptive ability of the gut and decrease the integrity of the epithelial barrier, whereas too much proliferation would result in neoplasia. Any genetic, epigenetic or environmental disturbance to the fine balance of proliferation, migration and shedding has the potential to modify tumourigenesis (Potten *et al* 1997.)

5.2. Examining the consequences of loss of MeCP2 in the murine intestine.

In order to characterise the effects of loss of functional MeCP2 on the murine small intestine, a conditional allele of MeCP2 and an intestinal-specific Cre recombinase were used to remove MeCP2 only in the small intestinal epithelium. The conditional *Mecp2* allele used was that created by Guy *et al*, and replaces exons 3 and 4 of *Mecp2* with the same exons flanked by LoxP sites (Guy *et al* 2001, see fig 1.8a.) To delete the conditional *Mecp2* allele only in the intestinal epithelium, an intestinal specific Cre recombinase driven by the *Cyp1A* promoter (Ah-Cre) was used (Ireland *et al* 2004, Sansom *et al* 2004.) This Ah-Cre recombinase allows virtually 100% recombination in the small intestine using a regime of four daily 80mg/kg injections of the β-napthoflavone inducer (Sansom *et al* 2004.)

Age matched male *Mecp2*$^{wt/y}$*Cre*$^{+}$ (wildtype, n=6) *Mecp2*$^{flox/y}$*Cre*$^{+}$ (n=6) and female *Mecp2*$^{wt/wt}$*Cre*$^{+}$ (n=6) and *Mecp2*$^{flox/flox}$*Cre*$^{+}$ (n=6) mice were used on an outbred background segregating for 129 and C57/Bl6 genomes. Mice were given one daily injection of 80mg/kg β-napthoflavone each day at the same time (+/- 1 hour) for 4 consecutive days to induce the Cre recombinase and recombine out the conditional *Mecp2* allele. At the appropriate time point, mice were killed by cervical dislocation and the small intestine was removed, prepared, sectioned and stained with haematoxylin and eosin and examined.

5.2.1. Mecp2 is removed from intestinal tissue after induction of the Cre recombinase.

In order to determine whether MeCP2 had been deleted in response to the β-napthoflavone inducer, mice carrying the Ah-Cre allele, the floxed *Mecp2* allele and the *Rosa26R* reporter locus

were subjected to the same injection regime as the experimental mice (the colony used for the main experiments does not carry the *Rosa26R* allele.) Results are shown in fig 5.2.

Fig 5.2. Removing Mecp2 in the murine intestine: LacZ reporter staining

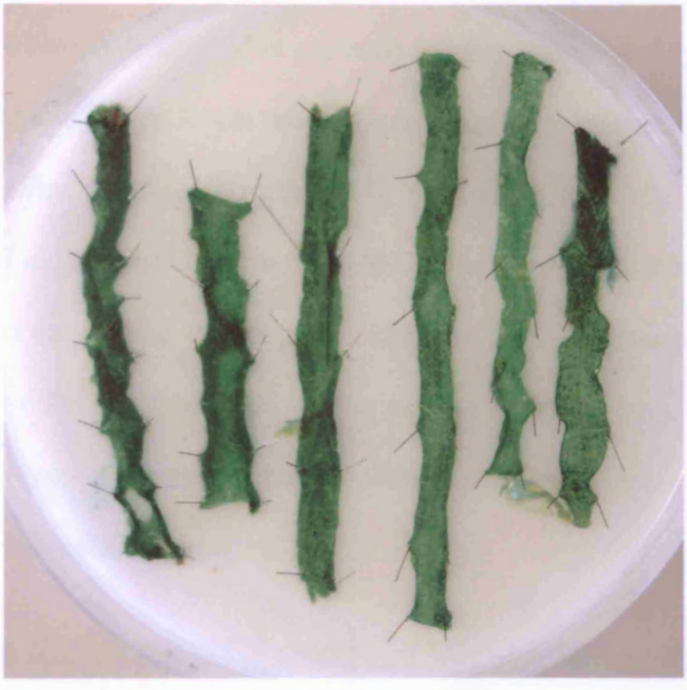


Fig 5.2. LacZ reporter staining shows that recombination has taken place at high frequency in the small intestine.

The intestine has been opened longitudinally and sectioned, with the proximal end at the top right. Recombination causes a floxed STOP cassette in a LacZ reporter gene to be removed. The gene is then transcribed and can be assayed by treatment with X-gal. The blue colour corresponds with areas in which Mecp2 has been removed.

Induction of the Ah-Cre recombinase results in excision of a stop codon in the *Rosa26R* allele and the subsequent expression of a *LacZ* reporter gene. This allows staining with X-gal to reveal areas of recombination as blue (Soriano 1999.)

5.2.1. Removing MeCP2 in the murine intestine; quantitative real-time RT-PCR.

The *LacZ* reporter allele shows that the proximal intestine undergoes almost complete recombination after 4 injections of β-napthoflavone. The half-life of the MeCP2 protein is between 3 and 4 hours (Reichwald *et al* 2000) so after 7 days all the MeCP2 should be removed. Quantitative realtime RT-PCR was then used to examine the amount of *Mecp2* RNA in intestinal tissues 7 days after the final injection. Fig 5.4 shows that *Mecp2* mRNAlevels are greatly reduced (2.13-fold, p=<0.005 Mann Whitney.) Note that some *Mecp2* activity remains; this is primarily due to the diverse cell population of the small intestine. Whole intestine is used to prepare RNA, but the recombination occurs only in the epithelium, which makes up a minority of the total cell types of the small intestine. Residual *Mecp2* is most likely due to that present in endothelial and muscular lineages etc.

Fig 5.4. Quantitative real time PCR analysis of Mecp2 mRNA levels in Mecp2$^{flox/y}$Cre^{+} mice after 4 injections of 80mg/kg β- napthoflavone.

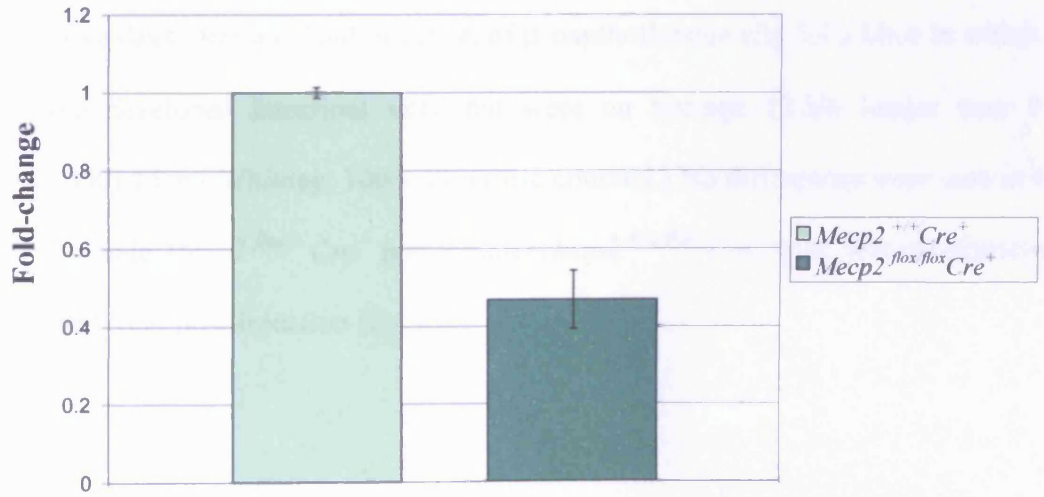


Fig 5.4. Mecp2 is removed in the small intestine.

Mecp2 is found to be significantly downregulated (2.13-fold, p=<0.005 Mann-Whitney) residual Mecp2 can be attributed to that remaining in non-epithelial tissues.

5.3. Characterisation of the intestinal phenotype of MeCP2-deficient mice.

Having shown that the recombination procedure removes MeCP2 from the small intestinal epithelium, I then proceeded to characterise the phenotype induced by loss of MeCP2. 3 month old age-matched male *Mecp2*$^{flox/y}$*Cre*$^{+}$ and *Mecp2*$^{+/y}$*Cre*$^{+}$ mice were induced and killed and their small intestines histologically examined 5, 7 and 9 days after the final injection of the β-napthoflavone inducer. The length of villi and the size of crypts were quantified in order to detect and changes in intestinal morphometry. Villus length was counted in the proximal 5cm of the small intestine, as the number of cells from the tip of the villus down one side to the crypt/villus boundary. Crypt size was counted as the number of cells from the base of the crypt up one side to the crypt villus boundary. The sizes of the crypt and the villus have been well-defined in the mouse, with a normal crypt being 22 cells long and a normal villus being ≈80 cells long (Mahmood *et al* 1998.)

Upon removal of MeCP2 from the intestine, a phenotype appeared that was visible and stable five days after the final injection of β-napthoflavone (fig 5.4.) Mice in which MeCP2 was removed developed intestinal villi that were on average 12.5% longer than the wild-type ($p=<0.0001$ Mann-Whitney, 100 viili/mouse counted.) No differences were seen in villus lengths between male *Mecp2* $^{flox/y}$ *Cre*$^{+}$ and female *Mecp2* $^{flox/flox}$ *Cre*$^{+}$ mice, and the phenotype remained stable at 1 year post-injection (fig 5.4.)

Fig 5.4. Villi become 12.5% longer when Mecp2 is removed from the murine small intestine.

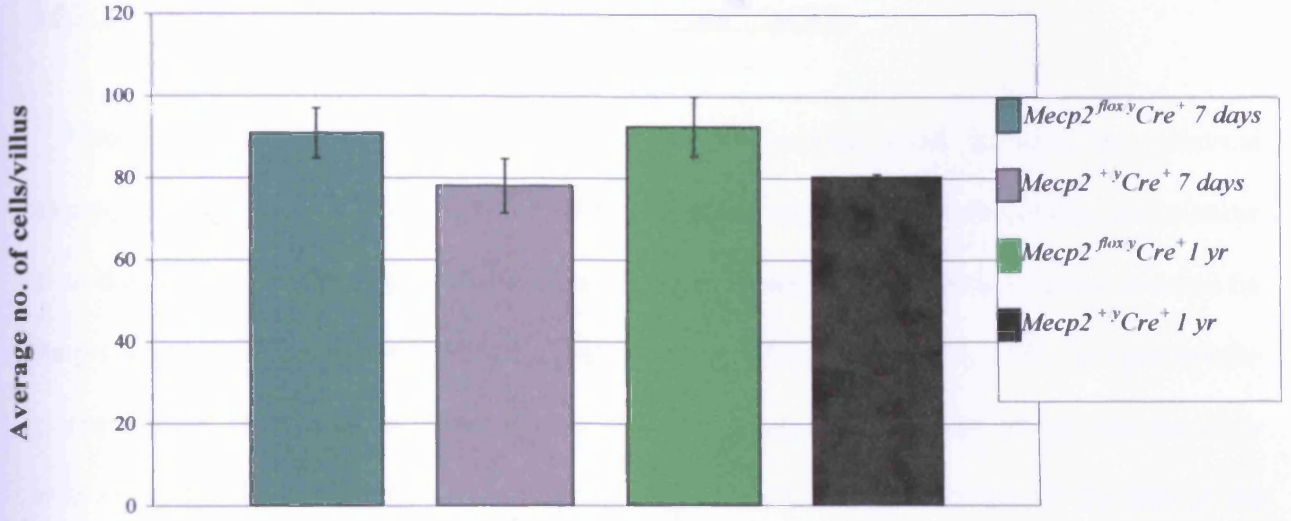


Fig 5.4. Villi become longer in the absence of Mecp2.

Villi become 12.5% longer and the effect is stable, lasting at least 1 year.

Fig 5.5. Distribution of villus sizes as quantified by cell number.

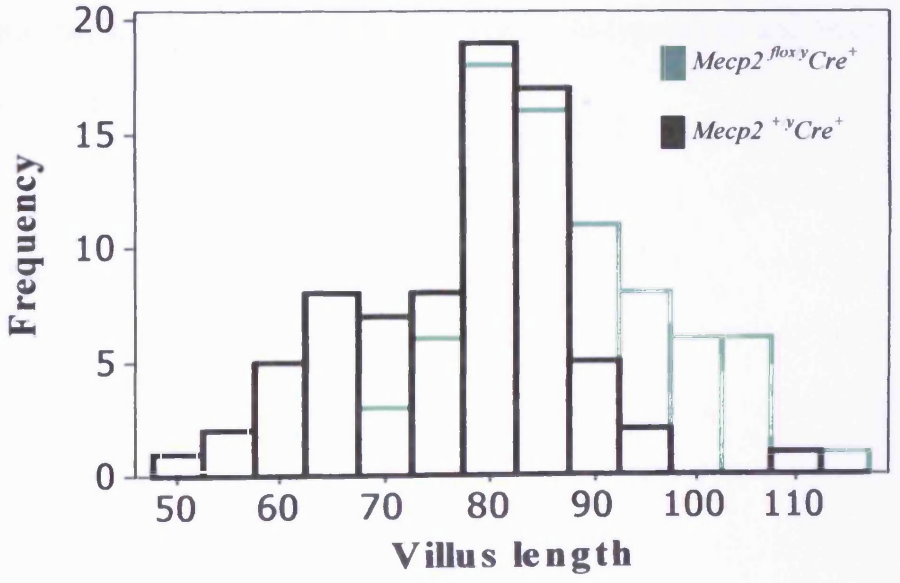


Fig 5.5. Altered distribution of villus sizes in the absence of MECP2.

The histogram shows how the distribution of variously-sized villi differs in MeCP2$^{flox/y}$ Cre^{+} and MeCP2 $^{+/y}$Cre^{+} mice. MeCP2$^{flox/y}$ show an altered distribution of villus sizes, with more larger villi being found.

5.4. Further characterisation of the MeCP2 deficient phenotype.

5.4.1. Crypt size remains constant in the absence of MeCP2.

When MeCP2 is conditionally removed from the murine small intestine, an intestinal phenotype develops with villi becoming 12.5% longer as quantified by cell number (cell number is used as measuring the length of the villus down the microscope is prone to errors induced by the preparation of the section, which can distort tissues.) This phenotype is stable and persists for at least 1 year. This villus lengthening may be due to a number of factors, which may act alone or in concert; cellular proliferation and proliferation rates may have increased, the size of the crypt compartment may have increased, or cellular apoptosis rates may have decreased. In order to determine which factors are important in the villus-lengthening phenotype, a range of experiments examining proliferation, migration and apoptosis were carried out.

In order to determine if both the crypt and the villus compartments had expanded, the number of cells in the crypts of $Mecp2^{flox/y}Cre^{+}$ and $Mecp2^{+/y}Cre^{+}$ mice were counted (50 crypts/mouse). No significant differences were found between wild-type crypts and *Mecp2* null crypts ($p=0.387$ Mann-Whitney.) Results are shown in fig 5.6.

Fig 5.6. Crypt size remains constant in the absence of Mecp2.

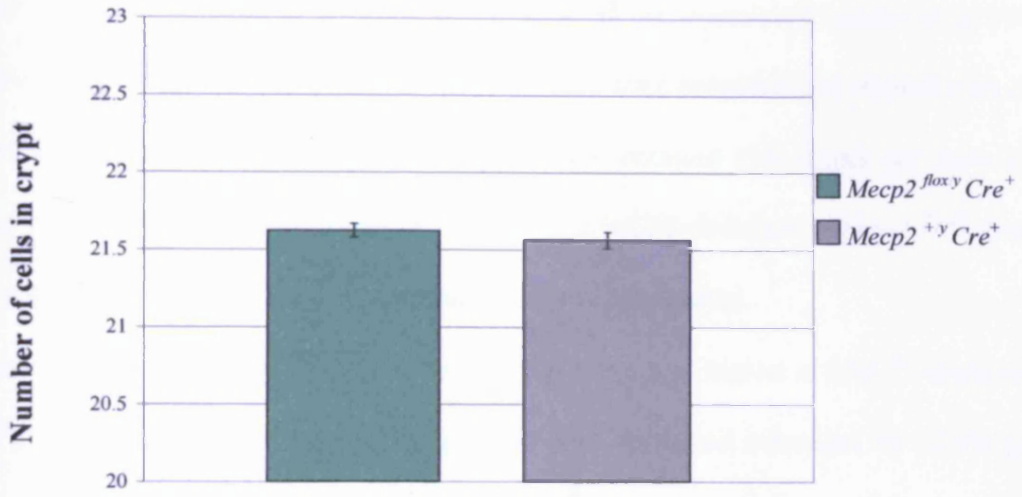


Fig 5.6. Crypt size remains constant in the absence of Mecp2.

The number of cells in each crypt remains constant at around 22 cells. No differences are seen between $Mecp2^{flox/y}$ Cre^{+} *and* $Mecp2^{+/y}$ Cre^{+} *mice.*

Both the Mecp2 deficient crypts, and those wild-type for Mecp2 contain around 22 cells (counted from the base of the crypt up one side of the crypt.) If the crypt was found to be larger, than it would be expected to produce more dividing cells to populate a longer villus without a change in proliferation rates. However, no significant differences were found in crypt size, in the absence or presence of Mecp2. Since no change was found in the size of the crypt, this implies that there are either more cycling cells in the crypt, or that the existing dividing population is dividing faster.

To examine proliferation rates, 3 month old age-matched male $Mecp2^{flox/y}$ Cre^{+} and 3 $Mecp2^{+/y}$ Cre^{+} mice were exposed to the β-napthoflavone inducer to induce the Cre recombinase

remove *Mecp2* from the small intestine and left for 7 days. Mice were then given a single injection of BrdU to label proliferating cells. BrdU is bioavailable for less than 2 hours, and effectively labels those cells in S-phase at the time of the injection (Sansom *et al* 2004.) Mice were culled 3 hours after the BrdU injection, and guts were prepared and stained with anti-BrdU. The number of proliferating cells in the crypt was counted (50 crypts per mouse) and the positions of the BrdU labelled cells were noted. MeCP2-deficient (*Mecp2* $^{flox/y}$ *Cre*$^{+}$) crypts showed a number of alterations in proliferation rates and patterns.

The absolute number of proliferating cells in the crypt was higher in MeCP2-deficient (*Mecp2* $^{flox/y}$ *Cre*$^{+}$) crypts (fig 5.7a) showing that proliferation rates had increased by 12.9% ($p= 0.0404$ Mann-Whitney.) The figure of a 12.9% increase in dividing cells correlates with the 12.5% increase in villus length. Hence the number of dividing cells in the crypt has increased with no concurrent increase in crypt size. This implies that the TA (transit amplifying) zone may have expanded. To determine if the TA zone had expanded, the position of BrdU-labelled cells in the crypt was quantified. Fig 5.7b shows that the transit amplifying (TA) zone was altered in the absence of MeCP2. The lower boundary of the TA zone remains stable, but the upper edge of the zone is expanded upwards.

Fig 5.7a. Proliferation in the crypt is increased in the absence of Mecp2

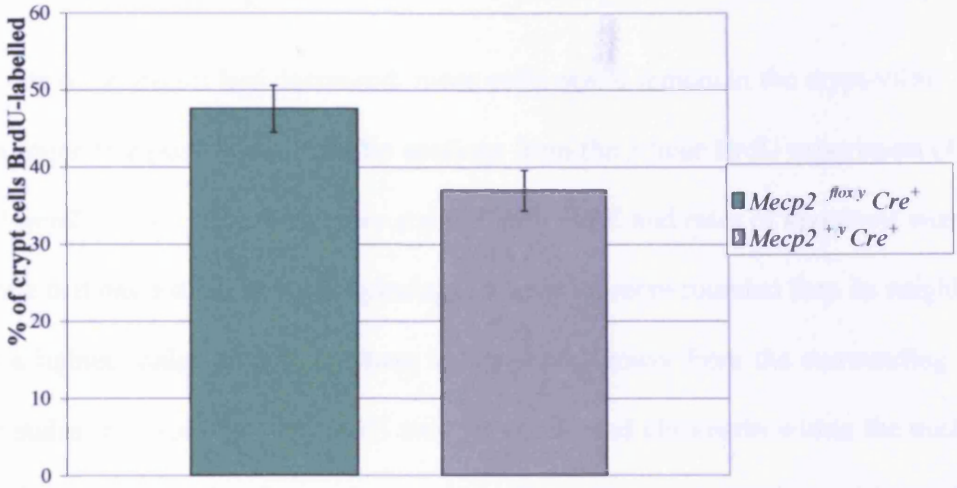


Fig 5.7b. The proliferative zone (TA zone) expands upwards in the absence of Mecp2.

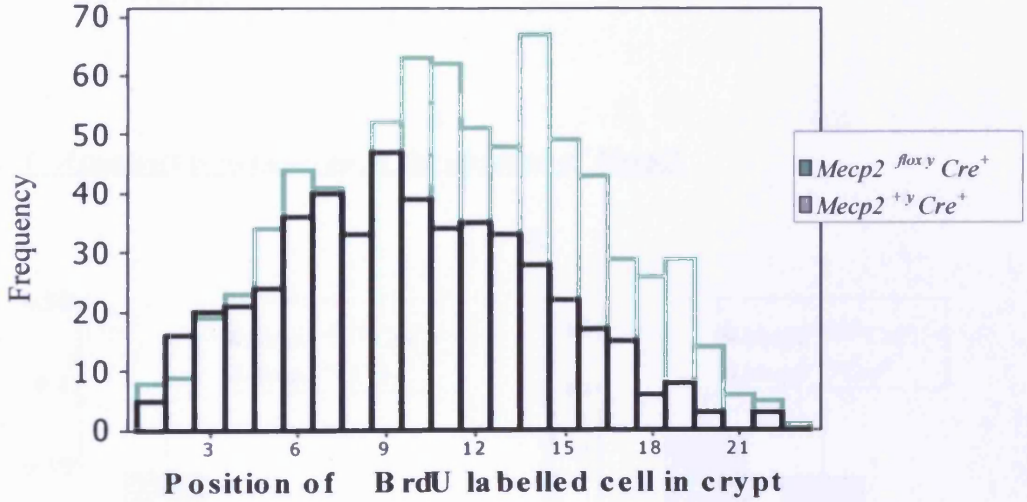


Fig 5.7. *7 days after induction of AH-Cre, mice were given BrdU to label cells in S-phase and then killed 3 hours later. A = Levels of proliferation increase in the Mecp2$^{flox/y}$Cre^{+} intestine. B = Histogram showing the expansion of the proliferative zone. The lower boundary remains constant but the upper boundary is expanded upwards.*

5.4.2. Loss of Mecp2 does not alter apoptosis in the crypt.

If the rate of apoptosis had decreased, more cells would remain in the crypt-villus system. In order to examine this possibility, paraffin sections from the 3 hour BrdU experiment (*Mecp2* $^{flox/y}$ *Cre*$^{+}$ and *Mecp2*$^{+/y}$ *Cre*$^{+}$ mice n=6) were stained with H&E and rates of apoptosis were counted. An apoptotic cell has a distinctive morphology; it appears more rounded than its neighbours, and may have a lighter 'halo' around it where it has shrunk away from the surrounding cells. The cytoplasm stains more pinkly with H&E and the condensed chromatin within the nucleus stains darkly. Both the absolute numbers of apoptotic cells in the crypt and the positions of apoptotic cells remained constant (fig 5.8) therefore altered apoptosis does not contribute to the villus-lengthening phenotype.

Fig 5.8. Apoptosis is not altered in the absence of Mecp2.

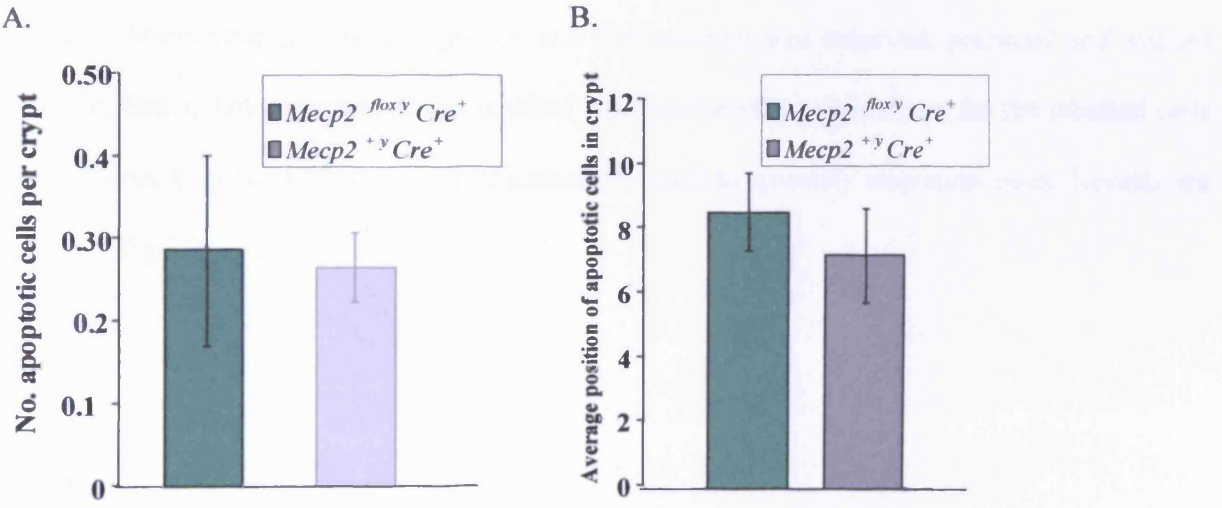


***Fig 5.8. Apoptosis is not altered in the absence of* Mecp2.**

A. Absolute numbers of apoptotic cells are not altered (p=0.381 Mann-Whitney.)

B. Positions of apoptotic cells within the crypt are not altered (p=0.396 Mann-Whitney.)

No alteration was seen in the rates of apoptosis in the absence of MeCP2. This shows that the increase in villus length is not due to decreased cell death.

5.4.3. Loss of MeCP2 in the murine intestine results in faster migration of epithelial cells along the villus.

In the absence of MeCP2, villi become 12.5% longer than those found in wild-type mice. Characterisation of the crypt showed that proliferation was increased, due to the TA zone being expanded up the crypt. If more cells are feeding onto the villus, the rate of migration of cells up the villus may be altered. In order to characterise this, proliferating cells were labelled with BrdU. Age matched 3 month old male *Mecp2*$^{flox/y}$*Cre*$^{+}$ and three *Mecp2*$^{+/y}$*Cre*$^{+}$ mice (n=6) were given the β-napthoflavone inducer and left for 7 days. They were then given a single injection of BrdU to label cells which were proliferating in the crypt at the time of the injection. Mice were killed 65 hours after the BrdU injection and the intestine was removed, prepared and stained with anti-BrdU. The positions of the labelled cells on the villus shows how far the labelled cells have migrated up the villus and can therefore be used to quantify migration rates. Results are shown in fig 5.9.

Fig 5.9. Migration along the villus is faster in the absence of Mecp2.

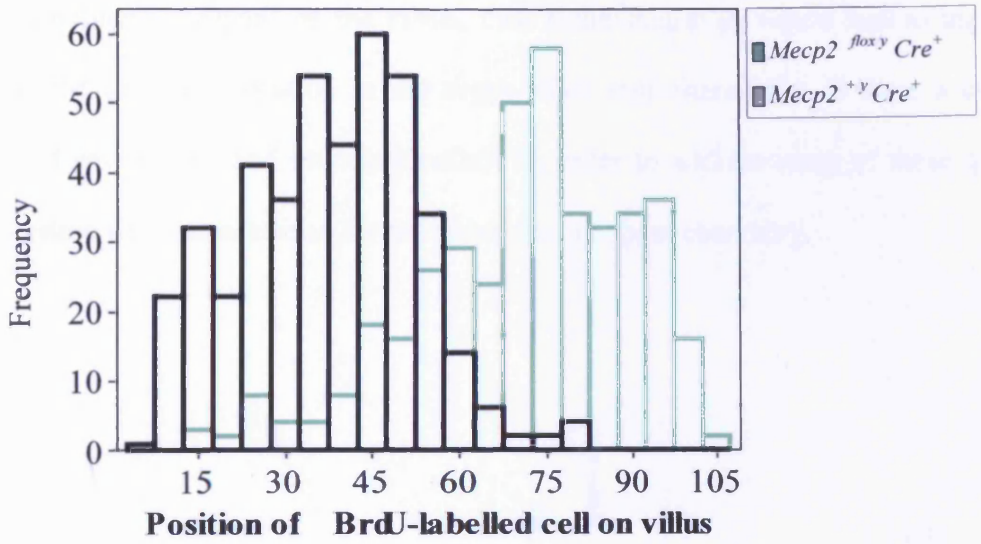


Figure 5.9. The distribution of labelled cells on the villus differs between Mecp2 $^{flox/y}$ Cre^{+} and Mecp2$^{+/y}$ Cre^{+} mice.

The peak of the Mecp2$^{+/y}$ Cre^{+} distribution lies at around position 45, whereas the peak of the Mecp2$^{flox/y}$ Cre^{+} distribution lies at around position 75, 60% further. The majority of BrdU-labelled cells have travelled further along the villus in the absence of Mecp2, *proving that migration rates are increased (p=<0.0001 Mann-Whitney.)*

The distribution of labelled cells differs between the two genotypes; in *Mecp2 $^{flox/y}$ Cre^{+}* mice the peak of the *Mecp2$^{+/y}$ Cre^{+}* distribution lies at around position 45, whereas the peak of the *Mecp2$^{flox/y}$ Cre^{+}* distribution lies at around position 75, 60% further. The majority of BrdU-labelled cells have travelled further along the villus in the absence of Mecp2, proving that migration rates are increased (p=<0.0001 Mann-Whitney.)

Is the increase in migration rate caused by the greater numbers of cells feeding onto the villus or is the increased proliferation in the crypt a response to the lengthening of the villus? If each cell has a defined 'lifespan' on the villus, then faster migration would lead to increased villus length. Is the lineage allocation in the crypt-villus unit altered (i.e. is there a change in the numbers of enterocytes and secretory cells?) In order to address some of these questions, the small intestine was characterised further using immunohistochemistry.

5.5. Immunohistochemistry of MeCP2-deficient and wild-type intestinal tissue.

Fig 5.10. Loss of* MeCP2 *does not alter the crypt-villus boundary.

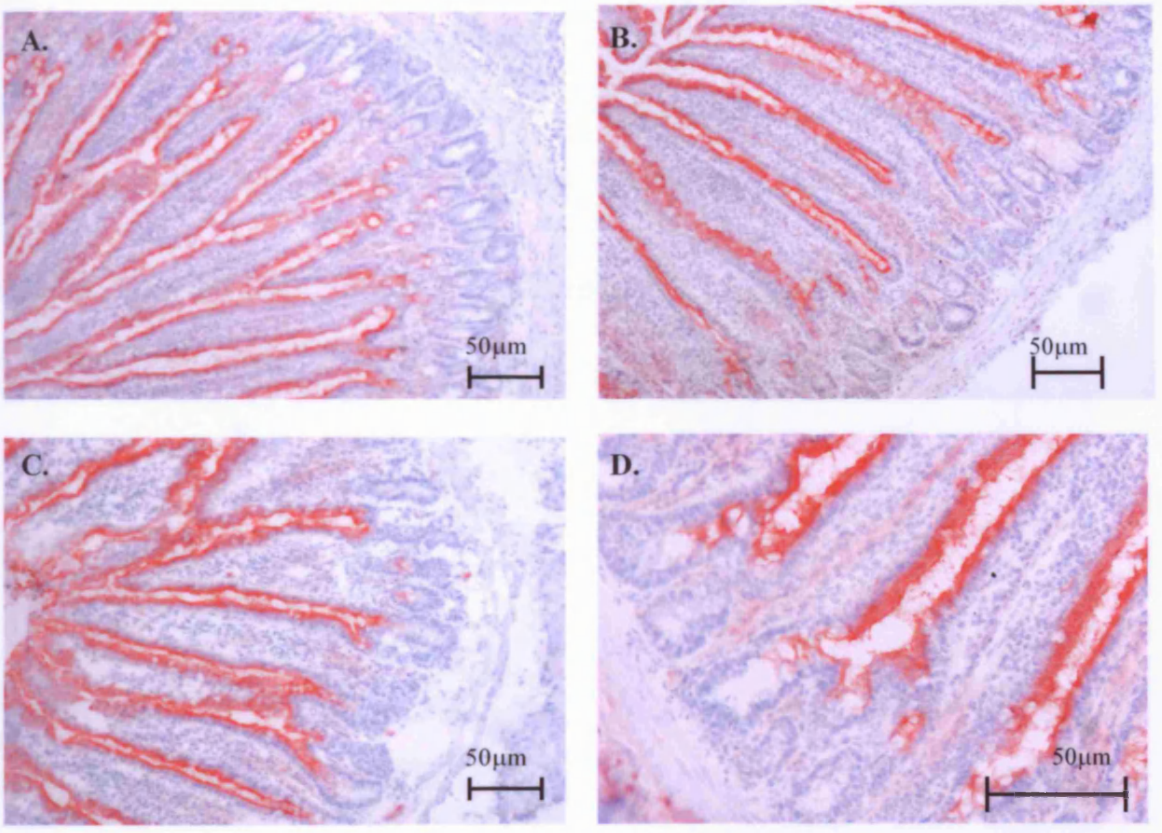


Fig 5.10. Loss of* MeCP2 *does not alter the crypt-villus boundary.

An alkaline phosphatase stain was used on cryofrozen sections; alkaline phosphatase stains the only the villus and not the crypt compartment and thus can show if the crypt-villus boundary has moved. Red stain shows alkaline phosphatase. A and C = $Mecp2^{+/y}Cre^{+}$, B and D = $Mecp2^{flox/y}Cre^{+}$. Loss of MeCP2 *does not alter this boundary and thus the identities of crypt and villus are not changed.*

Since alkaline phosphatase stains only the villus, staining can reveal the boundary between crypt and villus to determine if the boundaries of either compartment have moved in their relative positions. As fig 4.11 shows, no change in the boundaries of crypt and villus was seen using the alkaline phosphatase stain (red.)

Next, a selection of the various cell populations within the small intestine were stained, to determine whether their numbers or distribution were grossly altered.

Fig 5.11. Loss of* Mecp2 *does not alter the number of goblet cells.

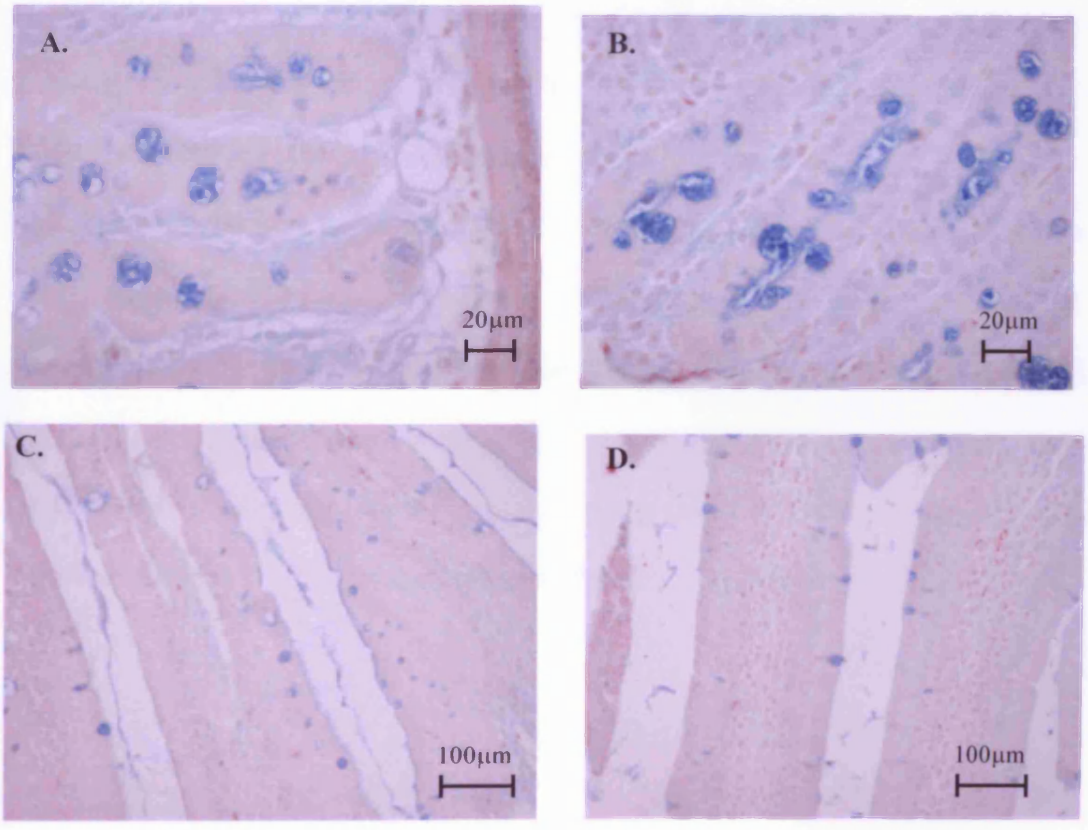


Fig 5.11. No alteration in the number or distribution of goblet cells.

A and C = Mecp2$^{+/y}$Cre^{+}, B and D = Mecp2$^{flox/y}$Cre^{+}. No differences were seen in the number or distribution of goblet cells in the two genotypes. Goblet cells and the mucins they secrete are stained blue with alcian blue.

Goblet cells produce mucins which lubricate the intestine. A change in goblet cell number or distribution would indicate that the process of cell lineage allocation had been altered (i.e. more/fewer cells being allocated to secretory lineages.) However, no gross changes in goblet cell number or distribution were seen (fig 5.11), implying cell lineage allocation has not been altered.

5.5.1. The number of enteroendocrine cells is not altered in the absence of MeCP2.

Next, the total enteroendocrine cell population was examined. Enteroendocrine cells secrete a variety of endocrine peptides which have affect the proliferation and growth of the gut (Potten *et al* 1997) and thus an alteration in the number of enteroendocrine cells might be expected to lead to changes in gut morphology. An alteration in number or distribution of enteroendocrine cells might also indicate a change in cell lineage allocation in the gut.

The total enteroendocrine population of the small intestine was stained with a Grimelius stain (Grimelius 2004) and the number of enteroendocrine cells in the crypt and villus was counted. No significant differences were found in the numbers of enteroendocrine cells in *Mecp2* $^{flox/y}$ Cre^{+} and *Mecp2* $^{+/y}$ Cre^{+} mice (fig 5.12 and 5.13.) As a comparison, the number of enteroendocrine cells in an *Mbd2* $^{-/-}$ mouse is also shown (fig 5.12b) the *Mbd2* $^{-/-}$ mouse shows a significant reduction in the total enteroendocrine cell population.

Fig 5.12. Loss of* Mecp2 *does not alter the number of Enteroendocrine cells in the small intestine.

Fig 5.12a. Enteroendocrine cell number in the crypt.

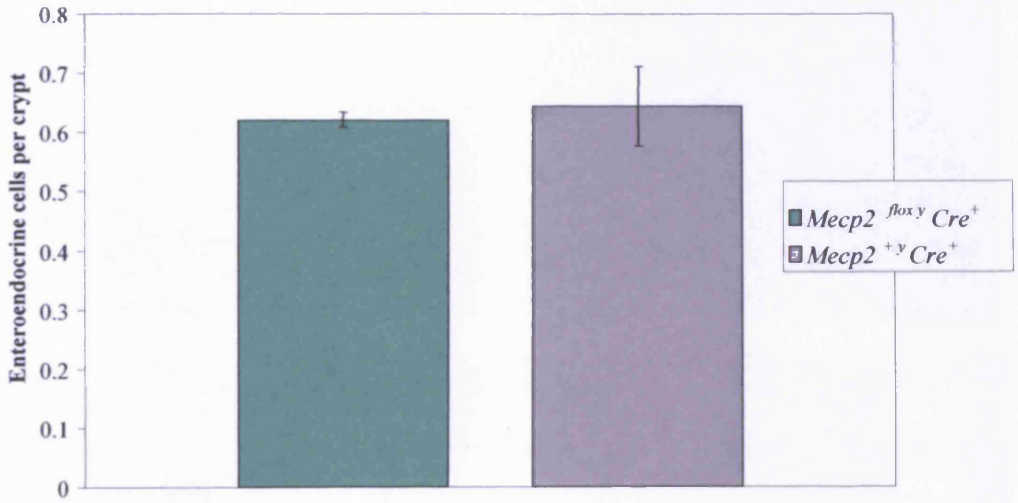


Fig 5.12b. Enteroendocrine cell number in the villus.

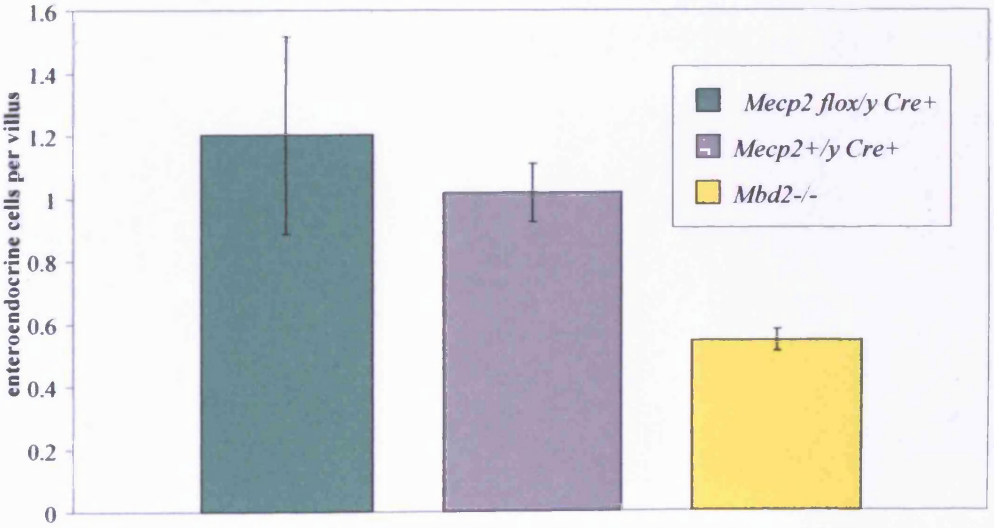


Fig 5.12. Enteroendocrine cell number is not altered in either the crypt (a) or the villus (b) in the absence of Mecp2. *By contrast, the Mbd2*$^{-/-}$ *mouse shows a significant reduction in the number of enteroendocrine cells in the small intestinal villi.*

Fig 5.13. Enteroendocrine cells in the small intestine of Mecp2$^{flox/y}$Cre^{+}, Mecp2$^{+/y}$Cre^{+} and Mbd2$^{-/-}$ mice.

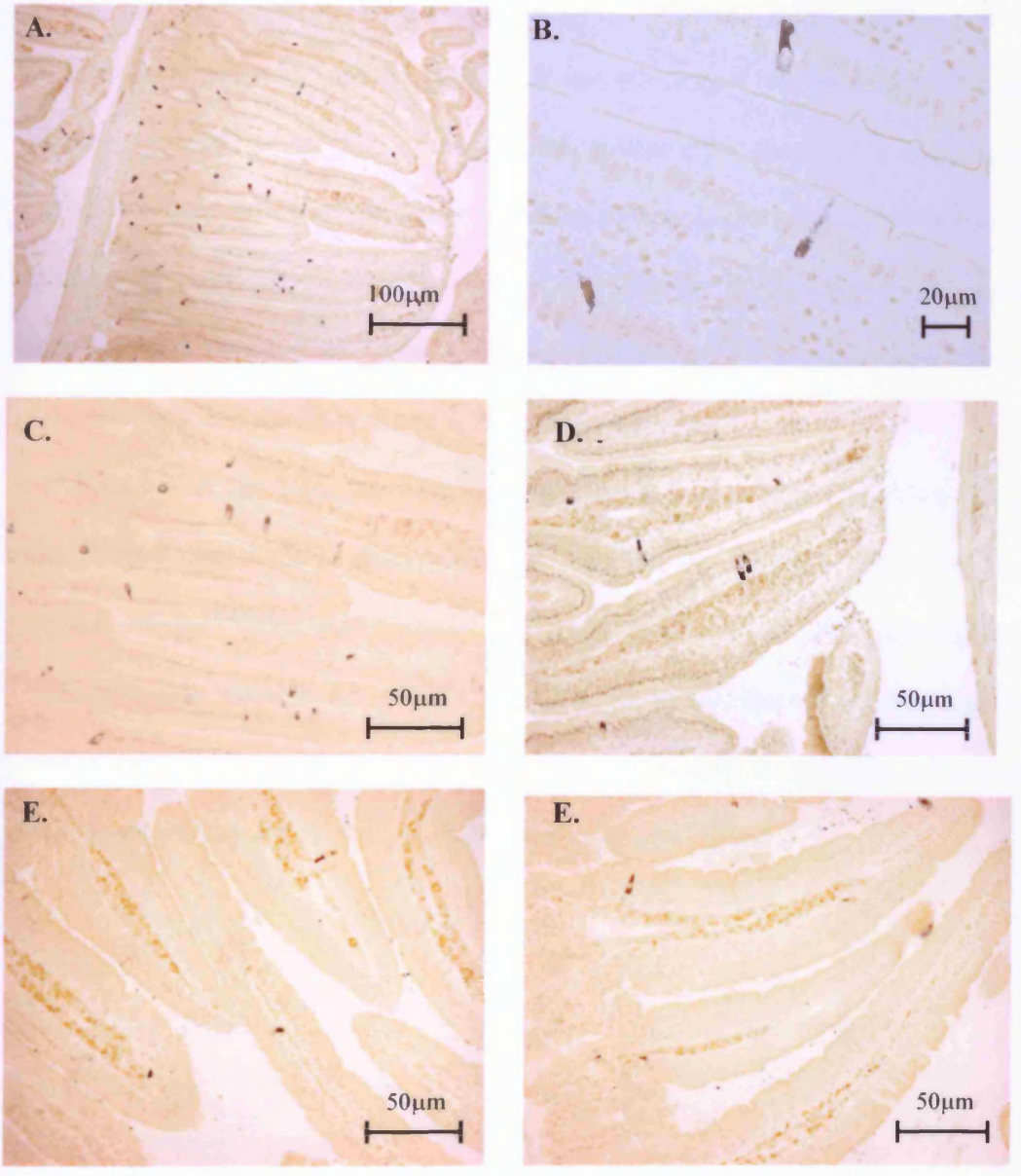


Fig 5.13. Grimelius stain showing total enteroendocrine cell population.

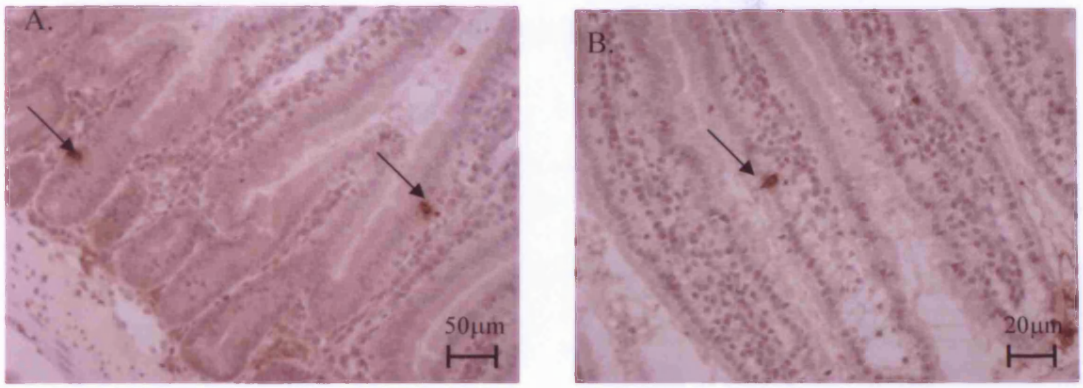
A and C = Mecp2 $^{flox\,y}$Cre^{+}, B and D = Mecp2$^{+\,y}$Cre^{+}, E and F = Mbd2$^{-/-}$. No difference is seen in the numbers or distribution of enteroendocrine cells in the presence or absence of MeCP2. *In contrast, loss of* Mbd2 *leads to significantly fewer enteroendocrine cells.*

5.5.2. Loss of MeCP2 does not alter the number of K-cells in the small intestine.

If the gross number of enteroendocrine cells is constant in the intestine, the proportions of the various subtypes may still be altered. A specific subset of enteroendocrine cells, the K-cells, secrete a peptide called gastrointestinal polypeptide (Gip) involved in nutrient sensing, a process which has been linked to alterations in villus morphology, and so it is possible that the observed phenotype derives from an expansion of the K-cells. To investigate this possibility, intestinal samples from the induced *Mecp2* $^{flox/y}$ *Cre*$^{+}$ and *Mecp2*$^{+/y}$ *Cre*$^{+}$ mice were stained with and anti-GIP antibody and the numbers of K-cells were counted.

As shown in fig 5.14, no significant differences were found in either the gross number or distribution of K-cells as determined by immunohistochemistry. When the numbers of K-cell cells in *Mecp2*$^{flox/y}$*Cre*$^{+}$ and *Mecp2*$^{+/y}$*Cre*$^{+}$ mice were counted, a slightly higher average number of K-cells was found in *Mecp2*$^{flox/y}$*Cre*$^{+}$ mice, but this difference was not statistically significant and is thus probably due to the slight increase in K-cells on a longer villus (which will not be significant as they are at such a low frequency on the villus.)

***Fig 5.14. No change in the number of K-cells in the absence of* Mecp2.**



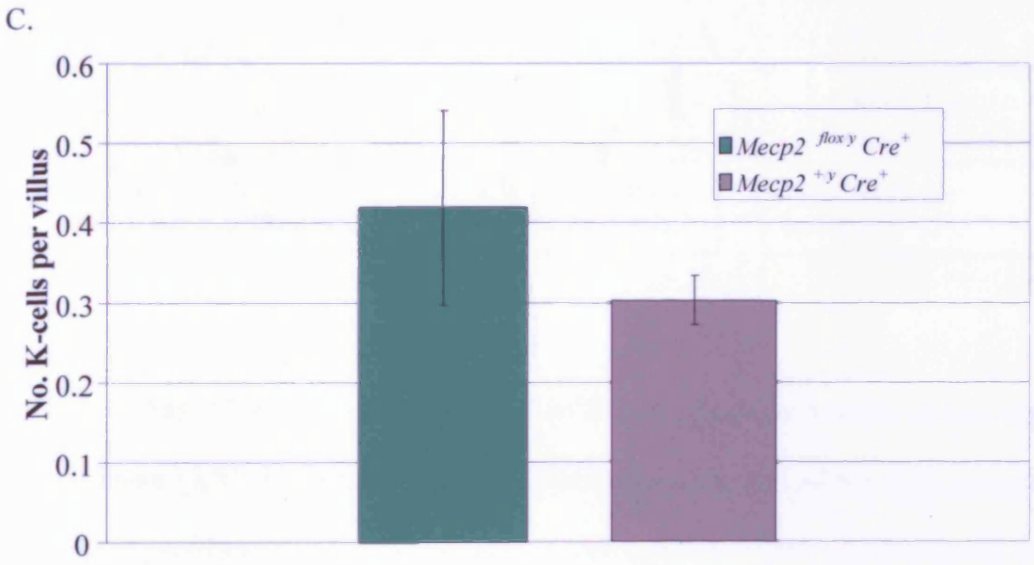


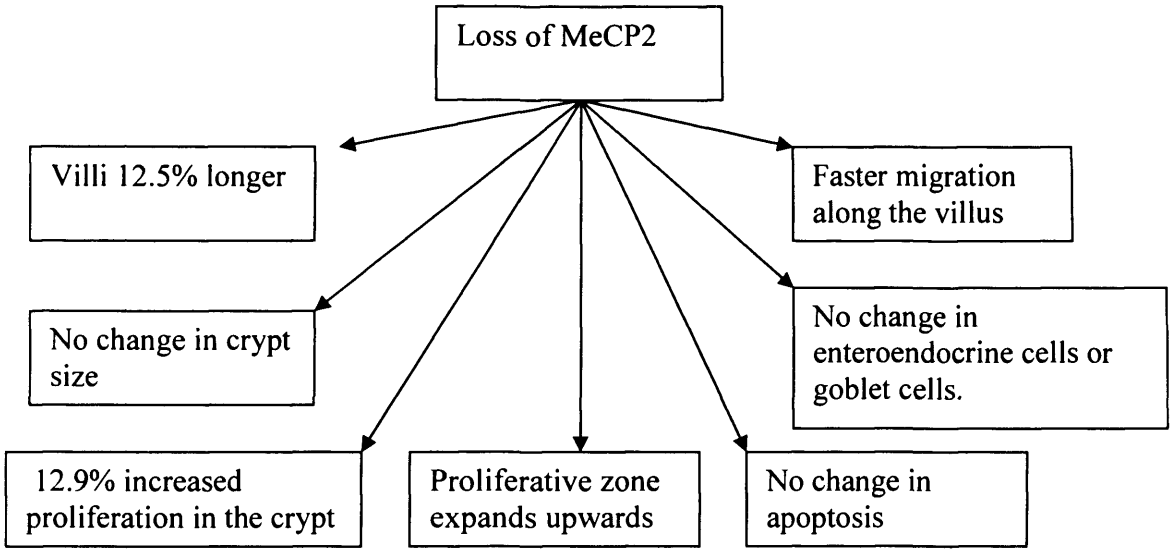
Fig 5.14. Loss of Mecp2 does not cause a significant change in the number of K-cells.

A = Mecp2$^{flox/y}$ Cre^{+} B=Mecp2$^{+/y}$Cre^{+}; no gross differences in K-cell number or distribution are visible (dark brown stain = K-cell, background = light brown)

C. Graph showing no significant difference in K-cell number in the presence or absence of Mecp2 *(p=0.56 Mann-Whitney.) Mecp2 $^{flox/y}$Cre^{+} induced mice do have a slightly higher average number of K-cells, but the difference is not statistically significant.*

5.6. Loss of MeCP2: Summary of the intestinal phenotype.

Fig 5.15. Phenotpye summary.



Loss of MeCP2 in the murine intestinal epithelium results in a stable intestinal phenotype. The villi become 12.5% longer. No change is observed in the size of the crypt compartment, but the number of proliferating cells in the crypt is increased by 12.9%, and the transit amplifying (TA) or proliferative zone is expanded upwards, leading to a larger population of cycling cells in the crypt. No changes in apoptosis are seen in the crypt, either in the number of apoptotic event, or in the position of apoptotic cells within the crypt. Examination of the villus itself reveals faster rates of migration of enterocytes along the length of the villus. No apparent changes in cell lineage allocation are seen, with goblet and enteroendocrine cell numbers showing no significant differences in $Mecp2^{flox/y}\ Cre^{+}$ and $Mecp2^{+/y}Cre^{+}$ intestine.

5.7. Discussion.

5.7.1. Which factors can alter proliferation in the intestine?

A wide variety of factors are known to have effects on the growth of the intestine including hormones, nutrient intake, small peptides and growth factors (Walters 2004.)

Kaestner *et al* described how knocking out the forkhead gene, *Fkh6* in the mouse alters intestinal epithelial proliferation (Kaestner *et al* 1997.) *Fkh6*, also known as *Foxl1*, expresses a winged helix transcription factor in the mesoderm of the intestine directly adjacent to the epithelium (Kaestner *et al* 1997.) Forkhead factors appear to mediate cross-talk between the epithelium and the endothelium (Kaestner *et al* 1997.) Both the embryonic development of the gut and its morphology in the adult are affected by loss of Fkh6. *Fkh6*$^{-/-}$ mice have fewer, poorly formed villi at embryonic day E16.5-E18.5 when compared to wild-type littermates (Kaestner *et al* 1997.) The spatial control of early proliferation is also altered; in normal mice the proliferating cells are limited to the intervillus region, whereas in null mice they are scattered throughout the villi (Kaestner *et al* 1997.) At birth, null mice have fewer, shorter and wider villi and many nulls rapidly lost weight and died before weaning, hinting at severe intestinal dysfunction (Kaestner *et al* 1997.) Mice that survived the weaning period however, rapidly gained weight, and were equal in weight to their siblings by 5 months of age (Kaestner *et al* 1997.)

Abnormalities were also found in adult animals; 50-day-old *Fkh6*$^{-/-}$ mice had expanded, branched crypts and cell-lined, mucus-filled cystic structures (Kaestner *et al* 1997.) The authors also noted that these abnormalities were more pronounced in the proximal intestine, which correlates with our findings that the increased villus length of MeCP2 deficient mice is more pronounced in the proximal intestine. In contrast to the fewer, shorter and poorly formed villi

seen during development, 50-day-old null mice had longer villi than their normal siblings and the effect was stable with age (Kaestner *et al* 1997.) Adult null mice showed a four-fold increase in the number of proliferating cells in the crypt, an increase in the size of the crypts and an increase in the amount of apoptosis within the mucosa (Kaestner *et al* 1997.) The number of goblet cells was slightly increased, indicating an alteration in cell-lineage specification (Kaestner *et al* 1997.) No data was available on migration rates (Kaestner *et al* 1997.)

A range of genes are implicated in the control of epithelial growth and regulation (Drucker *et al* 1996), although there is obviously an extremely complex web of control involving many genes (which in turn may interact with nutritional status, increasing the complexity of the system.) *Tcf-4*, for example, is a member of the *Tcf-Lef* family and expressed in developing and adult epithelium (Clatworthy *et al* 2001.) Knockout mice die soon after birth, have fewer villi in the small intestine and decreased numbers of epithelial cells in the intervillus region accompanied by a depleted stem cell compartment (Clatworthy *et al* 2001.) Tcf is a downstream effector of Wnt signalling (Clatworthy *et al* 2001.) Another gene known to be involved is *Cdx-1*, a *Drosophila Caudal* homologue; major expression is limited to the proliferating zone and is lost as cells differentiate and move out of the zone (Clatworthy *et al* 2001.) Over-expression of *Cdx-1* in an intestinal IEC-6 cell line has been shown to increase proliferation (Clatworthy *et al* 2001.) Wnt/β-catenin is also thought to regulate *Cdx-1* expression; *Tcf-1* knockout mice have lowered *Cdx-1* expression in the small intestine (Clatworthy *et al* 2001.) The *Cdx-1* promoter contains Tcf binding motifs and is regulated by Tcf-1/β-catenin complexes. Stem cell function and proliferation are therefore linked to Wnt/β-catenin and Tcf-4 (Clatworthy *et al* 2001.)

Another gene affecting crypt proliferation is *Nkx2-3,* deletion of which leads to increased crypt proliferation (Clatworthy *et al* 2001.) The phenotype resembles that of the *Fkh-6* knockout but has no effect on lineage allocation with normal numbers of goblet cells present (Clatworthy *et al* 2001.) Increased proliferation in these mice is associated with decreased levels of the bone morphogenetic proteins Bmp2 and Bmp4 (Kaestner *et al* 1997.) RNA protection analysis of *Fkh6*$^{-/-}$ mouse small intestine by Kaestner *et al* showed no change in the expression levels of *Cdx-1*, *Hnf3α*, *Hnf3β* (two other forkhead factors implicated in proliferation) *Tnf-α* or *Tgf-β* (Kaestner *et al* 1997.) Fkh6 null mice did show a 30-40% reduction of expression in Bmp2 in stomach, duodenum and jejunum, indicating that Bmp2 may be a direct or indirect downstream target of Fkh6 (Kaestner *et al* 1997.)

Clearly then, forkhead factors are only part of the story. A more direct cause of villus lengthening, however, can be ascribed to the actions of various other intestinally-active peptides.

5.7.2. The Proglucagon-Derived Peptides.

Several peptides, including bombesin, Gpi, insulin-like growth factors I and II (Igf-I and Igf II) peptide YY (PYY) and transforming growth factor α/β (Tgf-α/β) have been shown to affect intestinal proliferation (Drucker *et al* 1996.) Glp-2 is of particular interest, as it has been shown to have a specific and potent stimulatory effect on the intestinal epithelium (Drucker *et al* 1996, Jaleen *et al* 2000, Drucker 2001.)

Glp-2 (glucagon-like peptide 2) is a member of the proglucagon-derived group of hormones which includes Glp-1 and Glp-2, glicentin and oxyntomodulin (Drucker *et al* 1996, Jaleen *et al*

2000, Drucker 2001.) All these peptides are derived from a single proglucagon transcript (Fig 5.15.) which is post-translationally modified to give a range of peptides (Drucker 2001.)

Fig 5.15. Proglucagon-Like Hormones and GLP-2.

Adapted from Drucker et al 2001.

Fig 5.15a.) Proglucagon.

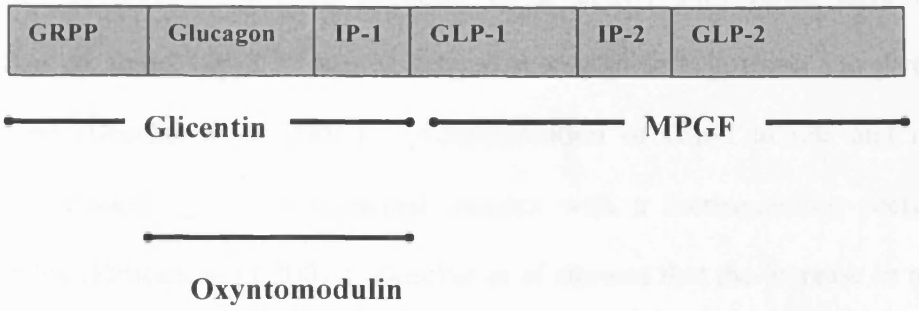


Fig 5.15b.) Alternative Processing.

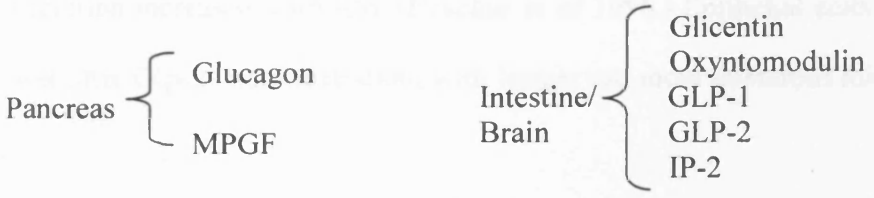


Fig 5.15. Structure of mammalian proglucagon.

IP =intervening peptide. MPGF = major proglucagon fragment. GRPP = glincentin related polypeptide.

In mammals, Glp-2 is found only in certain tissues (e.g. in the brain and intestine but not the pancreas) due to the cell-type specific expression of prohormone convertases (Drucker *et al* 1996, Jaleen *et al* 2000.)

5.7.3. Glp-2 Synthesis in the Intestine.

The type and amount of nutrients present in the gut appear to be a major determinant of both intestinal proglucagon gene expression and Glp-2 secretion. For example, fibre-rich diets stimulate Glp-2 secretion in human and rodent studies (Drucker *et al* 2001.) Intestinal injury also stimulates Glp-2 secretion (Drucker *et al* 2001.) Glp-2 is estimated to have a half-life in circulation of around 7.2 mins (Drucker *et al* 2001.) Two forms have been detected in circulation, an active Glp-2 $^{1-33}$ that is cleaved at an alanine at position 2 to give an inactive Glp-2 $^{3-33}$ form (Drucker *et al* 2001.) Administration of Glp-2 to rats and mice results in a significant thickening of the intestinal mucosa with a corresponding decrease in intestinal permeability (Drucker *et al* 2001.) Drucker *et al* showed that the increase in mucosal thickness was due to an increase in villus height (Drucker *et al* 1996.) Crypt size remained constant, but crypt proliferation increased markedly (Drucker *et al* 1996.) Epithelial cells also appear longer and narrower after Glp-2 administration, with longer and more numerous microvilli (Drucker *et al* 1996.)

Studies using the human intestinal epithelial cell line Caco-2 have shown that GLP-2 stimulates *in vitro* proliferation and have suggested that the effects of GLP-2 may be mediated by PI 3-kinase and the MAPK pathway (Jaleen *et al* 2000.) However, results obtained using immortalised and isolated cells lines must be treated with some caution as they cannot replicate a structurally and biochemically intact *in vivo* system.

Bjerknes and Cheng have demonstrated that administration of Glp-2 specifically increases proliferation of the C_o progenitor (Mills and Gordon 2001, Bjerknes and Cheng 1996.) However, they also demonstrate that the epithelium is devoid of Glp-2 receptors (Glp-2r) (Mills and Gordon 2001, Bjerknes and Cheng 1996.) Instead of eliciting a direst response, the action of

Glp-2 is thought to be mediated via the enteric nervous system (ENS), a complex network that regulates blood flow, secretion and gut motility (Mills and Gordon 2001, Bjerknes and Cheng 1996.) The crypt is richly innervated with enteric neurons lying close to the crypt base. These neurons express Glp-2r and respond to the Glp-2 signal by transmitting a signal to the epithelium (Bjerknes and Cheng 1996.) This signal can be blocked by tetrodotoxin (TTX) which blocks voltage gated sodium channels. The signal then somehow stimulates proliferation in the C_o progenitor (Bjerknes and Cheng 1996.) Glp-2 itself is produced by a subset of enteroendocrine cells, which may act as nutrient sensors (Bjerknes and Cheng 1996.) Hence a feedback loop between the epithelium, nutrient status and the ENS may be formed, which can regulate intestinal proliferation in response to nutrient conditions (Mills and Gordon 2001, Bjerknes and Cheng 1996.) This is a particularly interesting result in relation to Rett syndrome, as the enteric neurons are part of the autonomic system, which is known to be affected by RTT (Motil *et al* 1999) and could provide a mechanism for the gastrointestinal problems seen in RTT patients.

5.7.4. Glp-2 and MeCP2.

The similarity of the MeCP2-null gut phenotype to that produced by an excess of Glp-2 suggests that removal of MeCP2 may somehow increase transcription of the proglucagon transcript or Glp-2 secretion. Several other peptides which promote intestinal growth, such as bombesin, neurotensin and keratinocyte growth factor increase Glp-2 synthesis and/or secretion, suggesting that their effects may be in part mediated via Glp-2 (Drucker 2001.) RNA samples from *Mecp2*$^{flox/y}$*Cre*$^{+}$ and *Mecp2* $^{+/y}$*Cre*$^{+}$ mouse intestine will be used in array studies and

immunohistochemistry of frozen sections from *Mecp2*$^{flox/y}$ *Cre*$^{+}$ and *Mecp2*$^{+/y}$ *Cre*$^{+}$ mice will be used to ascertain if removal of MeCP2 increases Glp-2 levels.

5.7.5. Gip.

Gip, also known as glucose dependent insulinotropic polypeptide or gastric inhibitory polypeptide is a member of the vasoactive intestinal polypeptide (Vip)/secretin/glucagon family of gastrointestinal polypeptides (Meier and Nauk 2004.) Gip is a 42 amino acid peptide hormone that is synthesised and secreted from K cells, a subset of intestinal epithelial cells that are most abundant in the proximal duodenum (Cheung *et al* 2000, www.glucagon.com/gip.html.) K-cells are thought to be the 'nutrient sensors' of the intestine, as they respond to glucose levels in the intestine (Cheung *et al* 2000.) K-cells secrete Gip in response to glucose in a manner that closely parallels the way in which insulin is released from pancreatic β-cells (Cheung *et al* 2000.) Like Glp-1, Gip stimulates glucose-dependent insulin secretion and these two peptides are thought be responsible for up to 60% of the total insulin secretion which occurs after a meal (Cheung *et al* 2000.) Gip receptor negative (*Gipr*$^{-/-}$) mice also show high glucose intolerance and an impaired insulin secretion response to glucose (Miyawaki *et al* 1999.)

Recently, Gip has been linked to obesity, with the observation that Gip receptor negative (*Gipr*$^{-/-}$) mice are resistant to weight gain on a high fat diet (Nyberg *et al* 2005.) This implies that Gip may be involved in regulating food intake and may be part of the gut brain axis. A recent paper has found Gip and Gip receptor expression in the adult rat brain in the hippocampus (Nyberg *et al* 2005.) All the other members of the Vip/secretin/glucagon family have also been found in the brain, where they are classified as neurotransmitters and are involved in a range of brain functions including brain development, cell cycle regulation, differentiation and regulation

of food intake (Nyberg *et al* 2005.) Interestingly, Nyberg *et al* describe a role for Gip in stimulating proliferation of the progenitor cells of the rat hippocampus (Nyberg *et al* 2005.) Gip has also been found to be mitogenic in a range of other cells types (Nyberg *et al* 2005), such as the vascular endothelium (Ding *et al* 2003) quiescent adrenal tumour cells (Chabre *et al* 1998) and the osteoblast-like cell line MG-63 (Zhong *et al* 2003.) Gip is also known to be a growth factor for pancreatic β-cells (Trumper *et al* 2001.) Other members of the Vip/secretin/glucagon family are also known to have growth-stimulatory properties (Nyberg *et al* 2005) which again provides a rationale for examining the *Mecp2* null intestine as it provides a link between the gut and the brain.

5.7.6. Igf-I.

Igf-I is synthesised in the liver and intestine and has a profound effect on small intestinal growth (Ziegler *et al* 1996.) Administration of exogenous Igf-I increases the regeneration of the gut in models of intestinal mucosal atrophy (Zeigler *et al* 1996.) Igfs are extremely potent mitogens, and can alter epithelial cell kinetics by stimulating proliferation and inhibiting mitosis (Howarth 2003.) Igf-I is thought to predominantly target endothelial and epithelial cells and fibroblasts in the intestine (Howarth 2003.) Igf-I administered to rats continuously for 14 days resulted in marked and preferential growth of the gastrointestinal tract, increasing the weight of the gut as a fraction of total body weight by up to 22% (Howarth 2003.) In the same study, crypt size increased by 33% and villus cell number by 20% (Howarth 2003.) Igf-I receptors are found mainly in the basolateral region of the crypt, and Igf-I is thought to exert its effect by accelerating progression of the proliferating crypt cells through G1 to S-phase (Howarth 2003.) The phenotype of Igf-I – mediated intestinal growth has partial similarity to that seen in the *Mecp2* null intestine; unlike the MeCP2 deficient mouse, Igf-I over-expression results in an

increase in crypt size, and the size of both the mucosal and muscular layers is increased. In the MeCP2 deficient mouse, only the villus is seen to increase in size.

None of the factors which are known to increase intestinal proliferation exactly match the phenotype seen in the MeCP2-deficient mouse, and in many cases (such as Gip), the data on proliferation and migration within the crypt-villus unit is incomplete or absent. It may be that another, as yet unknown factor is responsible for the phenotype, or it may be that two or more of these factors are altered in the absence of MeCP2.

Since MeCP2 is considered to be a transcriptional repressor (Nan *et al* 1999), it would be expected that loss of MeCP2 would lead to deregulated gene expression. Microarray analysis will be used to examine the transcriptome of the MeCP2 deficient small intestine. No significant changes in gene expression have been found to date in mouse brain (Tudor *et al* 2002), so any significant changes would be important in our understanding of the *in vivo* role of MeCP2. It is hoped that any changes found will not only be able to shed light on the intestinal phenotype of the MeCP2 conditional knockout, but also on the fundamental problem of how loss of MeCP2 causes the pathology of Rett syndrome.

Chapter 6: Molecular characterisation of the intestinal phenotype induced by loss of MeCP2.

6.1. Loss of MeCP2 in the murine intestine results in a stable phenotype.

Using a conditional allele of *Mecp2* and an intestinal-specific Cre-recombinase (Ah-Cre) an intestinal phenotype caused by the loss of MeCP2 has been characterised in chapter 5.

Removal of MeCP2 in the intestine results in the stable lengthening of intestinal villi by a 12.5% increase in cell number. The crypt compartment is not expanded, but the levels of proliferation within the crypt are increased by 12.9%, with the upward expansion of the transit amplifying zone. Migration along the villus is significantly faster in *Mecp2* null intestinal tissue, although apoptosis does not appear to be affected. The numbers and distribution of cell types in the villus do not appear to be significantly altered. This is the first time a role for MeCP2 outside the central nervous system has been described.

In order to further define the phenotype and to try to elucidate a mechanism of action for MeCP2 in the gut, a microarray-based approach was used to examine the transcriptome of $Mecp2^{flox/y}Cre^{+}$ and $Mecp2^{+/y}Cre^{+}$ null intestinal tissue. Since MeCP2 represses transcription from methylated transcripts, and MeCP2 is widespread in chromatin, it might be assumed that the loss of MeCP2 would lead to a significant and genome-wide alteration in transcription levels (Nan *et al* 1999.) However, transcriptome analysis of *Mecp2* null brain tissue has revealed no significant changes in gene expression (Tudor *et al* 2002.) This implies that MeCP2 is either not a global transcriptional repressor as was originally though (Nan *et al 1999*) or perhaps that MeCP2 acts in concert with other factors to repress transcription, and that in its absence, some degree of compensation takes place, reducing the impact of

MeCP2 loss. There must, however, be effects on the transcriptome, as loss of MECP2 in humans causes RTT, and loss of MeCP2 in the murine intestine causes alterations in proliferation and migration of cells. Microarray analysis of the *Mecp2* null intestine has a number of advantages over previous work examining the *Mecp2* null brain; firstly, the intestinal phenotype is clearly defined in terms of proliferation and migration, which will make identification of targets suggested by the array easier. Secondly, if the loss of MeCP2 causes an effect at a specific point in the development or maturation of the brain, it may have a long-lasting effect on the architecture or function of the brain without the original transcriptional event being still detectable. In the intestine, however, cells are constantly cycling and so the events which lead to villus lengthening through loss of MeCP2 may be more easily defined.

Finally, in the murine intestine we have an extremely well characterised and easily accessible model. Physical changes in the intestine, such as a quantifiable change in migration rates, are easier to compare than behavioural traits such as hypoactivity, which may depend on a number of difficult to quantify variables. The use of microarrays combined with conditional mouse models is an extremely powerful method. The conditional allele allows the gene to be ablated in any tissue at any point in the life cycle of the model organism, while retaining the allele in other tissues. This allows any particular system to be looked at in isolation. RNA extracted from such tissues can be analysed with microarray technology, allowing a vast amount of data to be generated. This can be statistically analysed to reveal specific alterations in the levels of transcription of some genes. These targets can

then be confirmed or rejected using techniques such as immunohistochemistry, or quantitative real-time PCR, which allows highly accurate quantification of changes in expression.

6.2. Microarrays: vital tools in examining the transcriptome.

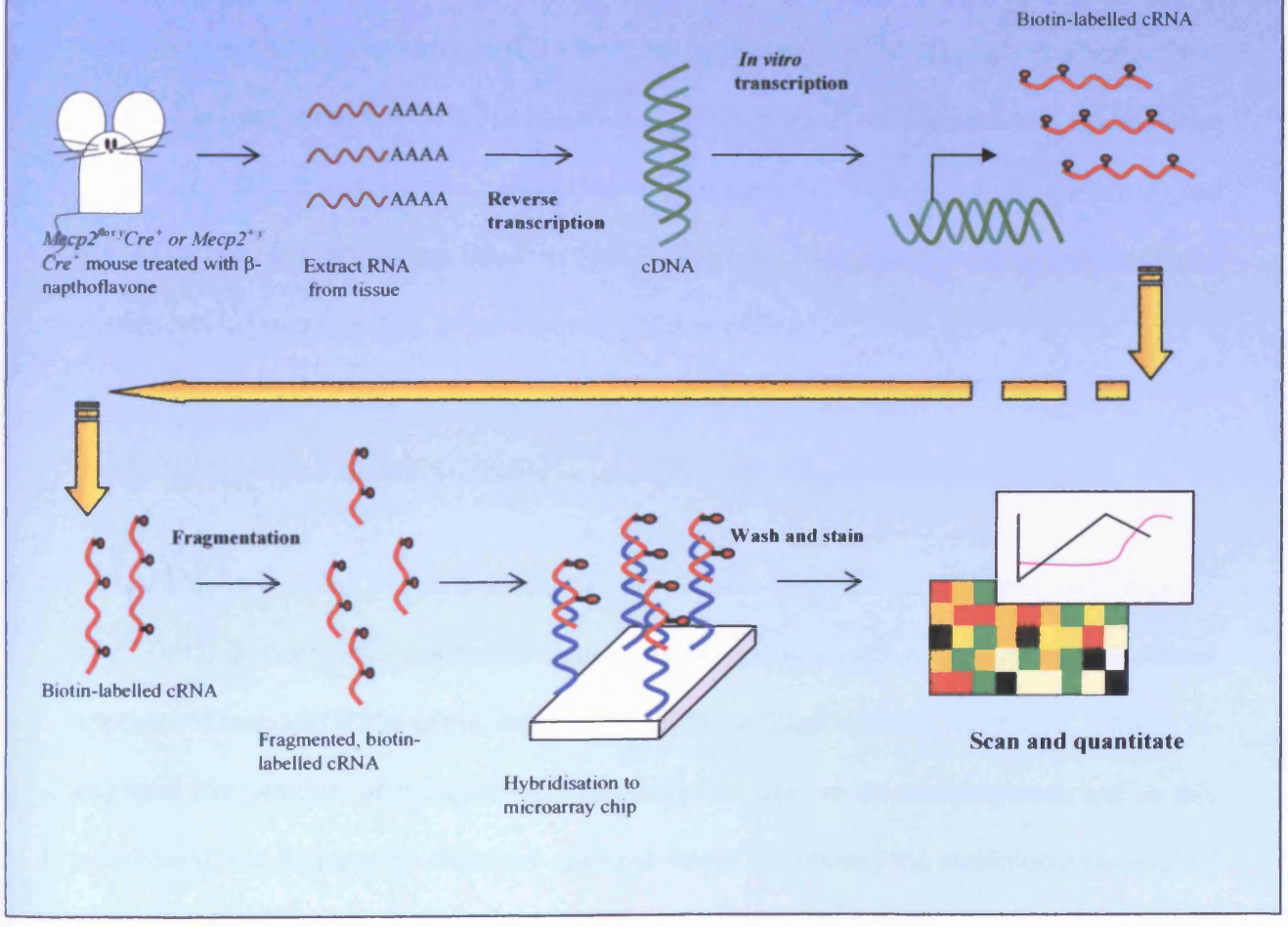
Microarray technology allows the entire transcriptome of an organism to be examined in a single experiment, and is capable of revealing not just expression changes in single genes but also patterns of change across an entire genome (Clarke *et al* 2004.) Rapid advances in genome sequencing programs have gone hand in hand with microarray research, allowing gene expression array chips to be developed for a wide range of model organisms (Clarke *et al* 2004.)

Microarrays are an especially useful tool for an experiment in a mouse model, in which the consequences of removal of a gene are unknown. A simple analysis of a microarray set will provide a vast amount of data which will allow targets to be followed up. It is worth stating, however, that microarrays are just one of a number of tools with which to investigate a phenotype. They do not in themselves provide absolute proof of gene expression changes as the data can be interpreted in a number of ways, and must be followed up using quantitative RT-PCRs, immunohistochemistry, western blots and other, more conventional techniques, in order to give definite answers to a particular question. Indeed, the huge amount of data generated by a typical microarray can be overwhelming if it is not carefully managed and analysed, and may well lead to a number of false avenues of exploration.

The Affymetrix GeneChip array (mouse U74) used in this experiment are manufactured using a combination of photolithography and combinatorial chemistry (www.affymetrix.com.)

The designs for each probe are derived from a unique sequence at the 3' end of each transcript. Two probes are prepared for each sequence; a perfect match (PM) and a mismatch (MM) which has a single incorrect nucleotide in the middle of the sequence. These paired probes act as a quality control mechanism for non-specific hybridisation.

Fig 6.1. Using the Affymetrix © GeneChip expresison array to analyse the transcriptome of the Mecp2 deficient murine intestine.



Biotin labelled, fragmented cRNA samples derived from the RNA from a tissue can be hybridised to the chip, and the chip scanned to quantify the amount of signal at each probe pair. Internal controls on the chip are formed by comparison with house-keeping gene signals, and 'spiked' controls (www.affymetrix.com/technology/index.affx.) The housekeeping genes, *β-actin* and *Gapdh* are used to measure the quality of the prepared sample RNA; probes are designed to the 3', middle and 5' portions of the transcript (www.affymetrix.com/technology/index.affx.) The ratio of the 3':5' signals should be 1. Deviation from a value of 1 indicates poor quality starting material and the sample should be re-derived. Other exogenous controls are also 'spiked' into the array to monitor washing, hybridisation and staining (www.affymetrix.com/technology/index.affx.) The transcript BioA is added at a concentration of 1.5pM; as this is at the limits of the sensitivity of the detection procedure, this transcript will be detected 70% of the time. BioB (5pM) BioC (25pM) and CreX (100pM) are also added. BioB, at 5pM, is designed to represent a transcript present at 3 copies/cell (www.affymetrix.com/technology/index.affx.)

6.2.1. Microarray Analysis of *Mecp2*$^{flox/y}$ *Cre*$^{+}$ and *Mecp2*$^{+/y}$ *Cre*$^{+}$ intestinal tissue.

Intestinal tissue from 3 month old age-matched male *Mecp2*$^{flox/y}$ *Cre*$^{+}$ (n=6) and *Mecp2*$^{+/y}$ *Cre*$^{+}$ (n=5) 3 month old age-matched male mice was analysed using a microarray-based approach. Since MeCP2 is considered to play a role in transcriptional repression, it might be expected that removal of the gene would result in changes to the transcriptome, and so this experiment was designed to show any changes which accompany the conditional removal of MeCP2 in the intestine.

Previous experiments such as those carried out by Tudor *et al* have failed to find any significant differences in expression in brain tissue in the absence of *Mecp2* (Tudor *et al* 2002.) However, given the strength of the phenotype in the intestine, it is reasonable to assume that transcriptional changes may be observed.

Mecp2$^{flox/y}$ *Cre*$^{+}$ (n=6) and *Mecp2*$^{+/y}$ *Cre*$^{+}$ (n=6) 3 month old age-matched male mice were given 4 intraperitoneal injections of β-napthoflavone (80mg/kg) and left for 7 days to allow the Cre to be induced and MeCP2 to be removed. Mice were then killed and the small intestine removed. Intestinal tissue was placed into RNAlater™ within 1 minute of the animal being killed. Peyer's patches were removed because of the high levels of immune tissue (which would have a different transcriptional profile) within them. Total cellular RNA was extracted with Trizol reagent, cleaned, quantified and reverse transcribed to make cDNA. *In vitro* transcription was used to create biotin-labelled cRNA which was then fragmented. The fragmented cRNA was sent to the Patterson Institute for hybridisation to the Affymetrix ® U74 mouse microarray chip containing ≈12500 genes. Sample data is shown in fig 6.2.

Fig 6.2. Sample Affymetrix data from the Mecp2 array.

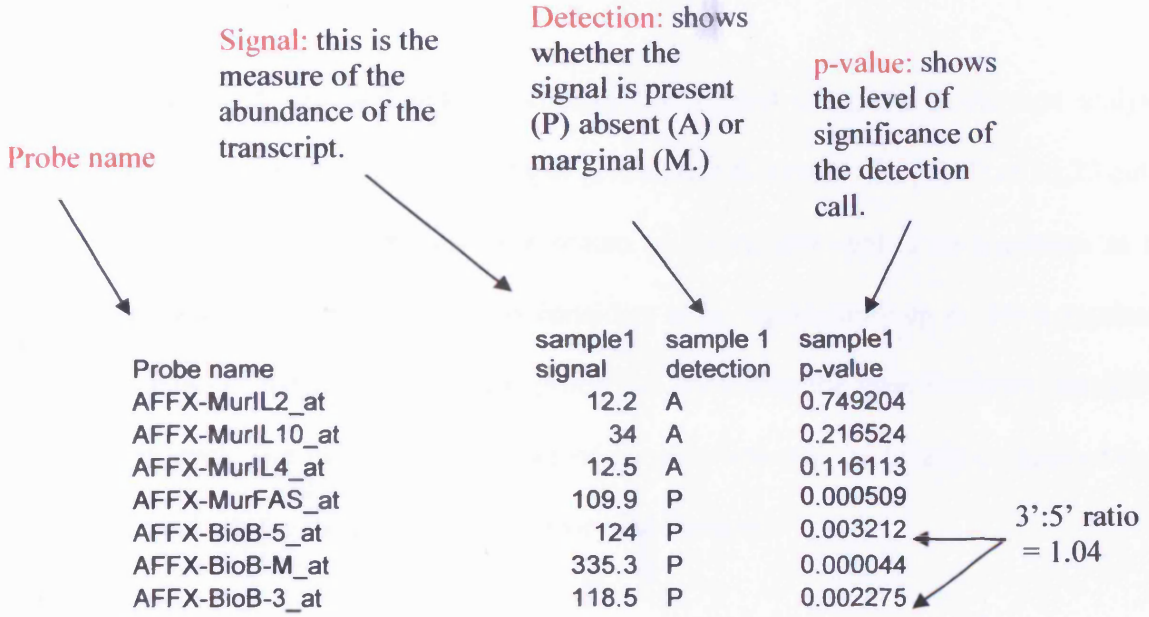


Fig 6.2. A set of sample data from the U74 gene chip array.

Data is presented in a format compatible with Word and Excel, to allow manipulation.

6.3. Analysis of Microarray Data.

Two methods of analysis were used; as microarray data can be analysed in a number of different way, it was decided to use two different analysis methods and compare the results from each. The first method involved normalising the data, then running it through the SAM (statistical analysis of microarrays) program (http://www-stat.stanford.edu/~tibs/SAM/) The second method was to take the array data, normalise it using MaxD and then perform a T-test on each set of data points. Targets provided by these methods were then followed up using quantitative real-time PCR.

6.3.1. Method 1: Analysis using SAM.

Array data was normalised against the laboratory dataset using MaxD and then analysed using SAM, using a delta value of 0.5144 to give a false discovery rate (FDR) of 33.23 out of 79 genes. The SAM software runs as a macro in Excel, and applies an algorithm to the normalised data to show which genes it considers to be significantly up or down-regulated. The user can define the stringency of the procedure, by setting the false discovery rate (FDR) at a desired level. The FDR is the number of genes which may be falsely considered to be significant, and represents the amount of error in the process.

Fig 6.3. SAM plot; normalised data analysed using SAM.

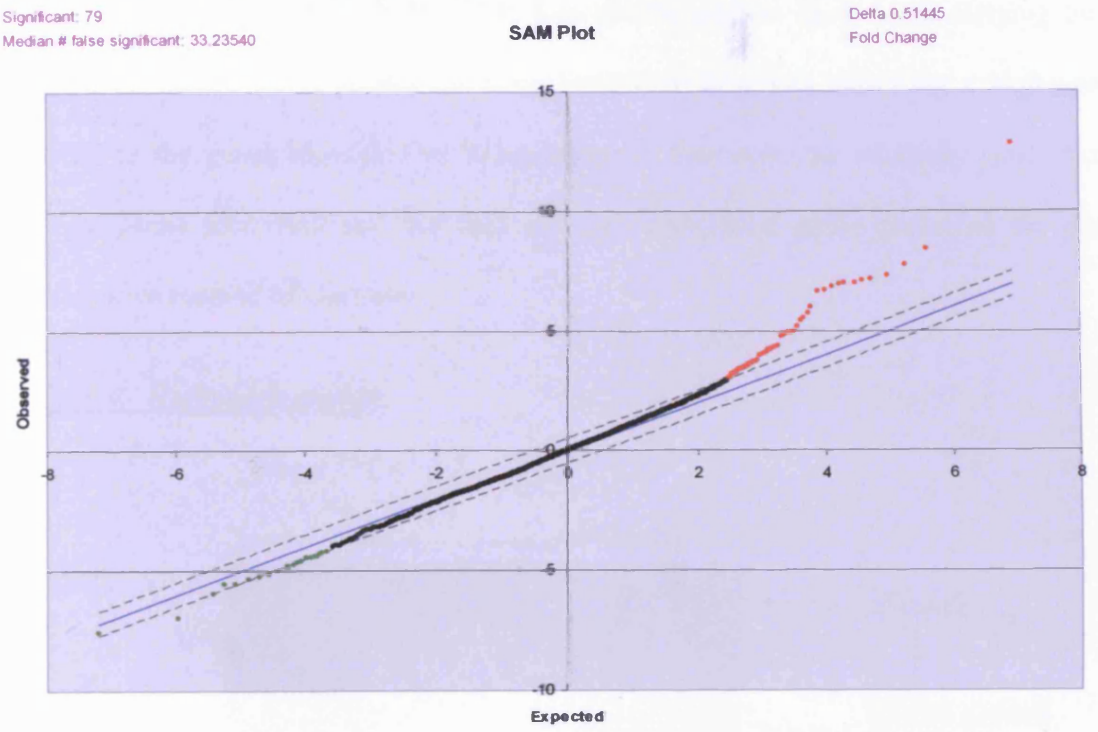


Fig 6.3. SAM plot from normalised data analysed with SAM.

This plot shows both up (red) and down (green) regulated genes, using a value of delta of 0.51445 (which sets the stringency of the algorithm and determines the FDR of 33.24.)
SAM assigns each data point a d value, analogous to a p value in a T-test, it then plots the observed set of d values against the expected set of d values for the array data. Values which deviate from the diagonal line are significantly different; downregulated genes (green, bottom left quadrant) have an observed d value lower than expected and upregulated calls (red, upper right quadrant) have observed d values which are higher than expected.

Among the targets identified as being downregulated are *Mecp2*, which is being removed from the intestine, and *Hprt*. The presence of *Hprt* is significant as the *Mecp2* conditional allele was created using *Hprt*-deficient ES cells. And thus a mouse carrying the floxed allele

would be expected to be Hprt-deficient. To further confirm the accuracy of this method of analysis, a semi-quantitative RT-PCR was carried out for *Hprt*. Mice carrying the floxed allele do indeed show greatly downregulated *Hprt* (fig 6.4) indicating a high confidence levels in the genes identified as being changed. However, the relatively small number of target genes identified and the lack of any upregulated genes prompted the use of an alternative method of analysis.

Fig 6.4. Hprt confirmation.

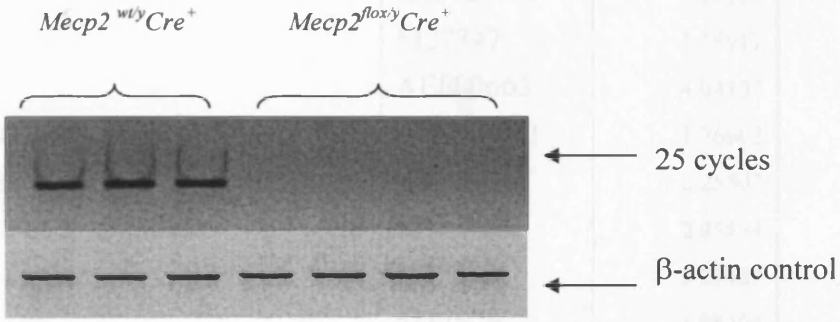


Fig 6.4. Hprt RNA levels are greatly reduced in mice carrying the conditional allele.

Since the Mecp2 conditional allele was created using Hprt-deficient ES cells, both an array and semi-quantitative RT-PCR should be able to detect a significant decrease in the expression of Hprt in mice carrying the floxed allele. Hprt RNA is undetectable at 25 cycles in Mecp2 flox/y Cre+ mice and only becomes detectable when the reaction is nearing saturation, at around 33-35 cycles. This validates the array, as Hprt is among the downregulated targets.

The SAM plot in fig 6.3 shows both upregulated (red) and downregulated (green) genes. The FDR is 33.24 of 79 genes, which means that just under half of the genes SAM considers to be significant are in fact false calls. Reducing delta by even a small amount dropped the number of significant genes to around 4, and so a higher FDR was used to give a larger

number of targets, which will be validated using quantitative real-time PCR. The lists of up and down regulated genes are shown below in tables 2 and 3.

Table2. List of upregulated significant genes.

Gene	Accession	Fold-change
EGFbp type C	M17962	2.77658
Elastase 2	X04573	7.72598
kallikrein	V00829	1.78947
Pancreatic ribonuclease I	X60103	4.49655
Uknown	AW120579	1.87238
P6-5 gene	M27347	5.56917
TCR-β	AE000663	4.04397
Unknown clone	AW209561	1.76602
Similar to a splicing factor	AV360487	2.25502
Partial sequence of a kallikrein	M133500	2.05834
Carboxyl ester lipase	U37586	5.53439
EGFbp type A	M17979	1.96394
Chymotrypsin like protease precursor1	AA244542	4.44722
Glutamine:fructose-6-phosphate	CC77609	2.25440
Kallikrein 5	Y00500	1.95515
Colipase presursor	AA710635	2.91832
TCR beta	AE000663	4.11208
Insulin II	X04724	4.14591
Chymotrypsinogen B1	AA590358	5.32154
Kallikrein 5	AV053185	1.79643
syncollin	AA607809	6.97226
ARF-GEP1	AV378443	2.32600
unknown	AV059956	3.68188
Cholecystokinin receptor A CCKA	D85605	3.50179
Triacyglycerol lipase precursor	AA674409	8.81336
Lipoxygenase 3	Y14695	1.48666
Possible sodium channel	AV336781	1.84022
glycine N-methyltransferase	D89664	2.63638
makorin	AW124942	1.33024
TCR beta locus	AE000664	5.66377

kallikrein	V00829	1.99685
IgE heavy chain constant.	X0857	1.71576
α-amylase I	J00356	3.93422
Myocyte enhancer factor 2C	L13171	1.90137
Glandular kallikrein	J00389	2.01177
Col2a-1	M63709	1.40886
unknown	C78889	1.42932
Possible tyrosine phosphatase	AW046257	1.81661
Reg1	D14011	4.88281
unknown	AV344835	1.32866
Mrp/plf3	X16009	2.03988
NGF-γ	X00472	2.14776
γ -renin	J03877	1.62172
Androgen binding protein-α	AF039064	1.74335
Nitotinic acetylcholine receptor δ-subunit	L10076	1.76499
Glucose-regulate protein 78	AV351546	2.08820
Phospholipase A2 presursor	AI327450	2.03943
TCR	AE000664	7.71579
T-complex protein 1	M3579	1.36447
Ena-vasodilator stimulated phosphoprotein (Evl)	AV371846	1.68127
Cytotoxic T lymphocyte lipase	M30687	1.93626
preprotrypsin	X04574	7.22785
unknown	C80068	1.36033
propionyl Coenzyme A carboxylase, beta polypeptide	AA882332	1.51082
EBI-1	L31580	1.45606
Interleukin-6	X54542	1.68617
AMPA1 glutamate receptor	AW048549	1.52968
unknown	AV096879	1.46605
Testis-specific protein (Mc2)	AF092208	2.24796

Table 3. Significant downregulated genes.

Gene	Accession	Fold change
unknown	AI850090	0.55229
unknown	AA544871	0.59134
Keratin 16	AF053235	0.50347
Similar to DRE1	AA691628	0.47028
Ssc2	AI317360	0.51715
Cis-retinol androgen dehydrogenase 1	AF030513	0.55407
	AI586160	0.58678
Paired IG-like receptor	U96682	0.48791
Similar to NADH dehydrogenase	AV089517	0.47444
MeCP2	AJ13922	0.36807
Neurotrophin 3	X53257	0.52872
unknown	AI850270	0.50437
Znf127	U19106	0.53017
Down syndrome critical region 3	AI848178	0.73263
unknown	AV294852	0.47640
SRY-box containing gene 17	D49473	0.55726
LDL-related protein 6	AF074265	0.55307
Aristaless 3	U96109	0.53903

6.4. Method 2. Analysis using a T-test.

The second method of analysis used involved normalising the data using MaxD, then using an Excel spreadsheet to calculate the fold-changes and then run a T-test on them. Data points were considered potentially significant if they showed a fold-change of greater than 1.5 fold up or down regulation and a p-value of ≤ 0.05. The SAM procedure is loosely based on a T-

test, but it also applies its own algorithm to 'judge' which genes are significantly altered in expression. The T-test alone would therefore be expected to give a much larger number of significant genes, and possibly a larger FDR. As can be seen from table 4,even genes which had large fold changes were often not significant when analysed using the T-test (e.g. *Dpgat1*, which has a fold change of 0.09, but a p-value of 0.33, which is not significant.) While the statistical procedures are objective, a degree of human judgement is required to decide if a gene with a marginal p-value will be followed up with QRT-PCR.

Table 4. List of significant downregulated genes found using a T-test

Gene	Accession	Fold-change	p-value
Hypoxanthine guanine phosphoribosyl transferase	K01515:	0.023543	0.014522
Complement component factor i	U47810:	0.093198	0.15889
Dpagt1	AV324170:	0.09816	0.337963
m46 protein	AJ243503	0.099342	0.196816
unknown	AV067753	0.117647	0.4621
WDNM1	X93037	0.118427	0.241908
Small inducible cytokine A2	M19681:	0.120076	0.4449
KF-1	AV336118	0.127451	0.016171
methyl-CpG-binding protein 2	AJ132922	0.132566	0.039909
pigment epithelium-derived factor (PEDF)	AF036164	0.16934	0.174152
unknown	AI837838	0.174888	0.181826
(clone- pMAT1)	L31958	0.17757	0.234558
unknown	AV294852	0.179012	0.030881
SH3 binding glutamic acid-rich (SH3BGR) gene	AJ239082	0.184416	0.007537
RGR opsin (Rgr)	AF076930	0.184676	0.302487
ARP-1	AV364118	0.186813	0.190737
unknown	AV023068	0.189236	0.409202

ADP-ribosaltransferase	X87612:M	0.194836	0.285373
alpha-albumin protein	AJ011080	0.195777	0.289661
deafness locus associated putative guanine nucleotide factor (DeLGEF gene)	AJ243952	0.196035	0.358125
unknown	AV254726	0.196891	0.100122
Serum amyloid A 3	X03505:	0.199657	0.880511
unknown	AI504260	0.20283	0.269253
unknown	AV333272	0.208145	0.339565

How accurate is the T-test? Whereas method 1 (SAM analysis of normalised data) picked out many significant upregulated genes, only one of the top 25 upregulated genes is significant according to its p-value from the T-test. Of the 322 genes which were 1.5 fold upregulated or higher, only 10 were found to be significant. They are shown in table 5.

Table 5. Significant upregulated genes (T-test method.)

Gene	Accession	Fold-change	p-value
Plasmacytoma variant translocation 1	Z11981	2.571429	0.024318
Unknown	C79525	2.148649	0.018962
Unknown	AV357533	2.103604	0.03235
Unknown	AV244683	2.085271	0.028013
Synaptonemal complex protein 3	Y08486	2.077922	0.031393
Unknown	AI851685	1.989071	0.022
Unknown	C78889	1.831395	0.025714
heat shock transcripton factor 2	X61754:	1.8223	0.017426
Unknown	AJ133428	1.700637	0.016744
Unknown	C76472	1.585091	0.046822

Similarly, only 5 of the top 25 downregulated genes were significant according to their p-values. Of the 5885 genes found to be downregulated by 1.5 fold or greater, 838 (14.2%) are significant according to their T-test p-value.

6.5. Identification of targets for follow-up using quantitative RT-PCR.

The following genes were identified as targets for further investigation using quantitative RT-PCR. These targets are mainly those suggested by the array, with the addition of other targets suggested by the literature (such as *Bdnf*, *Dlx5* and *Dlx6*) to be altered in the absence of MeCP2. Other targets, such as glucagon, were chosen as they may be altered in a villus lengthening phenotype. *Gip* was chosen as other members of the Vip/secretin/glucagon family are upregulated by *Bdnf*, and play a significant role in the interaction of brain and gut to modulate food intake. Dlx5 and Dlx6 are included in both columns as they are described in the literature as being downregulated (Horike *et al* 2005) but appear to be downregulated on the array.

Table 6. Targets suggested by analysis of microarray data and literature/villus lengthening models.

Upregulated	Downregulated
Galgt2	*Mecp2*
Liprin α-4	*Ssc2*
CCF-i	*Neuromedin*
M46 (ghrelin)	*Igfbp3*
CCK-A	*Igfbp10*
Collagen	*PEDF*
GIP	*Dlx5*
WDNM1	*Dlx6*
Dlx5	
Dlx6	

6.6. Confirmation of Array Targets: Quantitative Real-time RT-PCR.

Quantitative real time RT-PCR is a highly sensitive tool which can reliable detect extremely small changes in gene expression. The method relies on a system in which the amount of a fluorescent signal is proportional to the amount of DNA product in the reaction and the fluorescence levels in the reaction are measured after every cycle, allowing a quantifiable read-out of the amount of product present in realtime.

Fig 6.5. Quantitative realtime PCR.

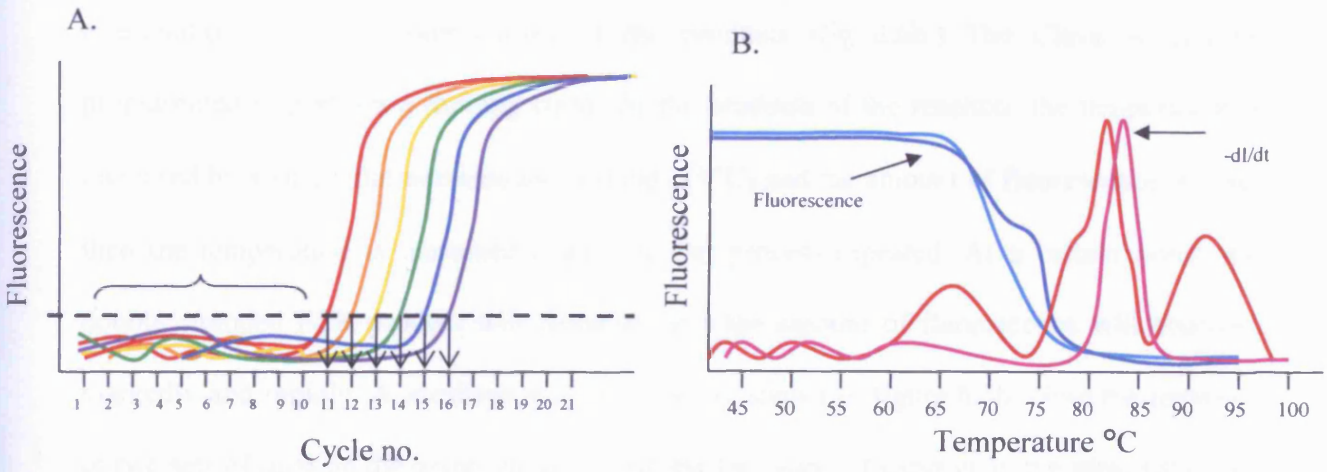


Fig 6.5. Quantitative Real time PCR

Fluorescent traces for a series of serially-diluted RNA samples. The amount of fluorescence increases as the reaction proceeds according to normal PCR kinetics (i.e a typical sigmoid curve.) The threshold cycle (Ct) is measured when the reaction is in the log phase; the user sets a threshold (dotted line) above which the measurement is made, avoiding the noise at the lag phase of the reaction. Each cycle represents a two-fold increase in the starting number of copies of RNA (cDNA) in the sample, i.e. the red sample, which becomes visible at a Ct of 11, has twice as many copies of the RNA (cDNA) as the orange trace, which in turn has twice as many as the yellow trace and so on. B= melting curves (explanation in text.)

The fluorescent molecule used in this experiment, SYBRgreen, emits fluorescence when it is complexed in double stranded DNA, hence as the reaction proceeds, more double stranded DNA is produced and the fluorescence signal is greater. Fluorescence is measured by a laser after every cycle, allowing a quantitative real-time image of the reaction to be built up (fig 6.5.)

Advantages of SYBR green are its low cost, ease of use and the ability to use non-modified oligonucleotides as primers. The major disadvantage of SYBR green is that it is non-discriminatory; double stranded DNA could potentially be formed from primer dimers or

non-specific amplification, and the signal from such DNA would not be distinguished from that emitted by legitimate product. This problem can largely be circumvented by careful examination of the melting curves of the products (fig 6.6b.) The Chromo4 can be programmed to perform a melting curve on the products of the reaction; the temperature is increased by a small, defined amount (around 0.5°C) and the amount of fluorescence is read, then the temperature is increased again and the process repeated. At a certain point, the double stranded PCR product will denature, and the amount of fluorescence will decrease markedly and rapidly. A idealised melting curve is shown in figure 6.5b. Note the presence of two sets of lines on the graph; these are merely two ways of expressing the same data. The blue lines show the total amount of fluorescence, and the red/pink lines show –dl/dt where l=luminescence and t is temperature. The red/dark blue lines on fig 6.5b represent an ideal melting curve; a single peak or drop representing a single PCR product. The pink/light blue lines represents a non-ideal melting curve, characterised by the presence of a stepped drop in fluorescence, or 'shoulders' on the peak of the melting curve this can imply the presence of multiple products, and means the reaction conditions must be changed. The peak of the melting curve should lie between 82°C and 90°C. If the peak lies at a lower temperature, it may indicate that the fluorescence of primer dimers is being examined.

6.7. Using QRT-PCR to examine array targets.

Quantitative real-time RT-PCR was carried out on cDNA samples derived from RNA samples from 3 month old age-matched male *Mecp2*$^{flox/y}$*Cre*$^{+}$ and *Mecp2*$^{+/y}$*Cre*$^{+}$ mice which had been given 4 daily injections of 80mg/kg β-napthoflavone and left for 7 days. Primers to

targets suggested by the array data were designed using Primer3© software (http://frodo.wi.mit.edu/cgi-bin/primer3_www.cgi) and the mRNA sequence data supplied through a search of Pubmed; (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=Search&DB=Pubmed.) A full list of primers can be found in section? *β-actin* primers were obtained (J. Zabkiewicz, personal communication) and used as a control. Reactions were carried out on an MJ Research Chromo4 thermal cycler and data was analysed using the Opticon Monitor 2 © software. Sample data for a run is shown in figures 6.6a and 6.6b. All reactions were carried out at least in duplicate with no-template controls. Controls for the absence of genomic DNA were carried out by running samples in which no reverse transcriptase was present in the cellular RNA→cDNA reaction. Only samples in which both controls were negative were used.

The reaction is measured by selecting a threshold value (Ct) at which the increase in fluorescence is still in the log phase (dotted line of fig 6.6a.) This gives a Ct value for each sample. Sample data for one of the genes examined, *Igfbp3,* are shown in fig 6.6a and 6.6b.

Fig 6.6a. Sample data for quantitative realtime RT-PCR; amplification curve for Igfbp3.

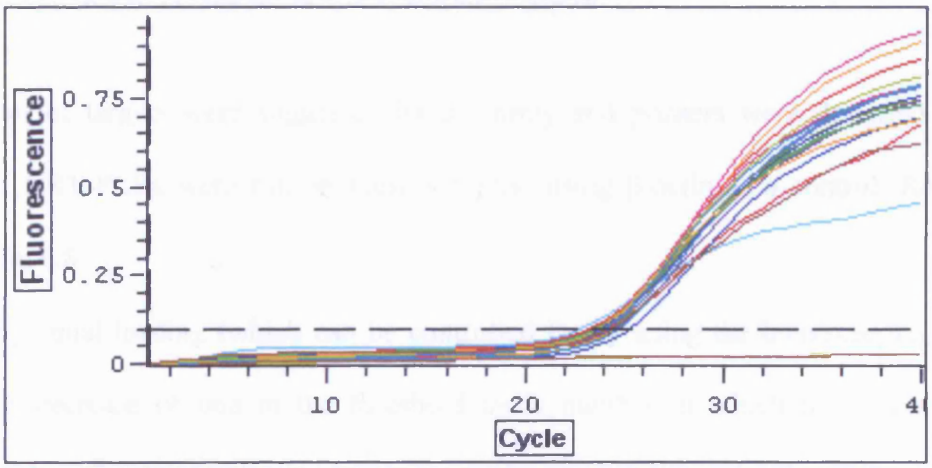


Fig 6.6b. Melting curve for Igfbp3.

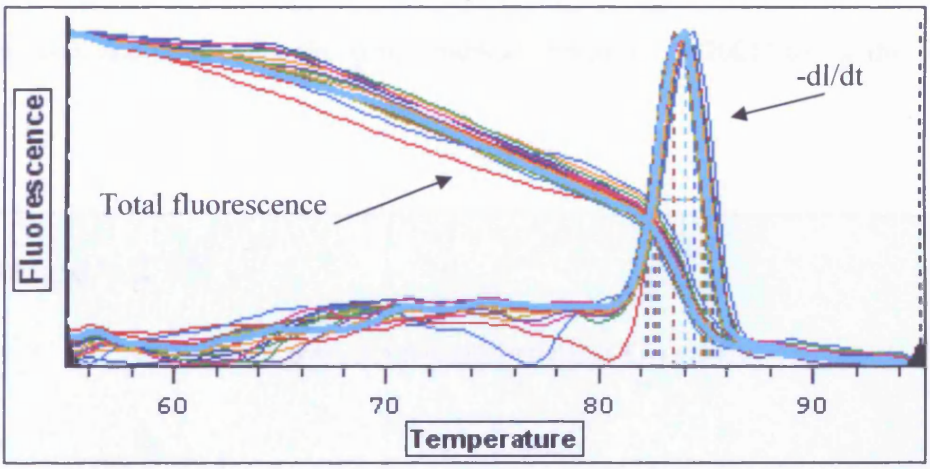


Fig 6.6b. Melting curve for Igfbp3.

The high quality of the reaction is shown by the shape of the melting curve of the products; a single sharp peak with no 'shoulders' that would indicate other products. The peak is centred at 84℃, too high a temperature to be formed by spurious product such as primer dimers.

6.8. Quantitative real-time RT-PCR analysis of targets suggested by microarray analysis of *Mecp2*$^{+/+}$ and *Mecp2*$^{-/-}$ deficient murine small intestine.

A number of targets were suggested by the array and primers were designed to these sequences. QRT-PCRs were run on these samples, using β-actin as a control. Results are shown in fig 6.8.

Assuming equal loading (which can be controlled for by using the housekeeping gene β-actin) each decrease of one in the threshold cycle number at which the signal appears corresponds to a 2-fold increase in abundance of the original transcript. i.e. if gene A has an average Ct of 14 and gene B has an average Ct of 15, then transcripts of gene A are twice as abundant as those of gene B. Because of this relationship, data from the QRT-PCR experiments was analysed using the ΔΔCt method (Livak *et al* 2001) using the following equation:

$$\textbf{Fold change} = 2^{-\Delta\Delta Ct}$$

$$= 2^{-((\text{Ct floxed gene})-(\text{Ct floxed } \beta\text{-actin}))-((\text{Ct wt gene})-(\text{Ct wt } \beta\text{-actin}))}$$

Where; Ct = threshold cycle

Ct floxed gene = the Ct of the gene being examined in the *Mecp2*$^{flox/y}$*Cre*$^{+}$ mouse

Ct floxed β-actin = the Ct of the β-actin control in the *Mecp2*$^{flox/y}$ *Cre*$^{+}$ mouse

Ct wt gene = the Ct of the gene being examined in the *Mecp2*$^{+/y}$ *Cre*$^{+}$ mouse

Ct wt β-actin = the Ct of the β-actin control in the *Mecp2*$^{+/y}$ *Cre*$^{+}$ mouse.

6.8.1. Example of QRT-PCR calculation: Gip.

The expression of Gip (gastrointestinal polypeptide) gene was examined in *Mecp2*$^{flox/y}$ *Cre*$^{+}$*mice and Mecp2* $^{+/y}$*Cre*$^{+}$ mice (n=6), using *β-actin* as a control. Results were as follows;

Fig 6.7. Sample data for Gip.

Ct gene (floxed)	Ct β-actin (floxed)	Ct gene (WT)	Ct β-actin (WT)	Δ Ct floxed (Ct gene floxed)- (Ct β-actin floxed)	ΔCt wt (Ct gene wt)- (Ct β-actin wt)	Δ ΔCT (Δ Ct floxed) -(ΔCt wt)	$2^{-\Delta\Delta CT}$
21.44	15.43	22.28	15.17	6.01	7.107	-1.093	2.133

Gip, therefore, has been upregulated 2.134-fold by the loss of MeCP2.

Fig 6.8. QRT-PCR Results.

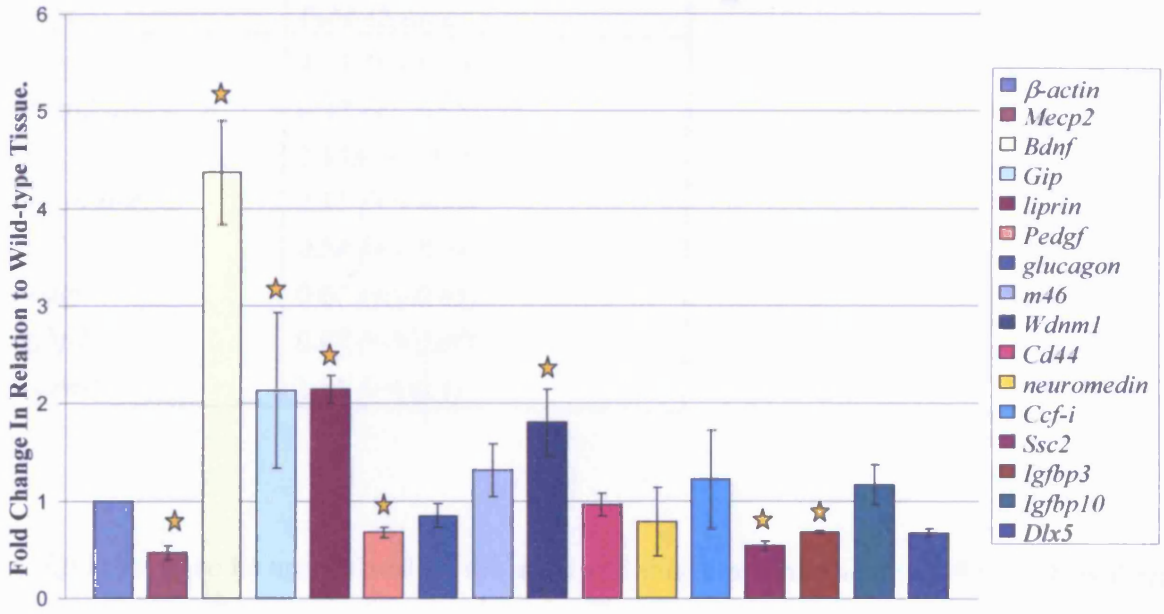


Fig 6.8. Quantitative realtime PCR analysis of selected array targets.

Fold-changes observed using QRT-PCR follow-up of array targets. Significant changes are marked with stars. Of the genes examined, Bdnf, Gip, Liprinα-4 and Wdnm1 were found to be significantly upregulated (stars.) and Mecp2, Pedgf, Ssc2 and Igfbp3 were found to be downregulated. Errors are calculated as the standard deviation of the means of expression levels in $Mecp2^{flox/y}\ Cre^{+}$ *versus* $Mecp2^{+/y}\ Cre^{+}$ *mice.*

Of the genes examined, *Bdnf, Gip, Liprinα-4* and *Wdnm1* were found to be significantly upregulated, and *Mecp2, Pedgf, Ssc2* and *Igfbp3* were found to be downregulated. A list of significantly altered genes is shown in table 7.

Table 7. Significant changes in gene expression.

Gene	Fold-change
Bdnf	4.37 *(+/- 0.53)*
Mecp2	0.47 *(+/- 0.08)*
Gip	2.133 *(+/- 0.8)*
Liprin α-4	2.15 *(+/- 0.14)*
Ssc2	0.54 *(+/- 0.044)*
Pedgf	0.67 *(+/- 0.05)*
Igfbp3	0.67 *(+/- 0.006)*
Wdnm1	1.81 *(+/- 0.3)*

Gip is found to be upregulated on the array and this was confirmed by QRT-PCR; is there any evidence that upregulation of Gip alone can cause the lengthening of villi? If a direct cause-and-effect is to be established, then Gip would need to be shown to be both necessary and sufficient to cause the villus lengthening phenotype. Since Gip has such a short half-life *in vivo*, this would be difficult to accomplish using exogenously administered Gip, but it could be accomplished by over-expressing Gip *in vivo*. One interesting observation which provides further evidence that Gip is involved in villus overgrowth was provided by a strain of mice which were found to have even longer villi than the MeCP2 deficient mice. The Dlb *(Dolicos biflorus)* mouse has a polymorphism at the Galgt2 locus, which lies within 0.2CM of the GIP locus. The Dlb1 mouse was found to have villi which averaged ≈95 cells, significantly different from the wild-type mice (fig 6.9.)

Fig 6.9a. Mice carrying a polymorphism at the Galgt2 locus have long villi.

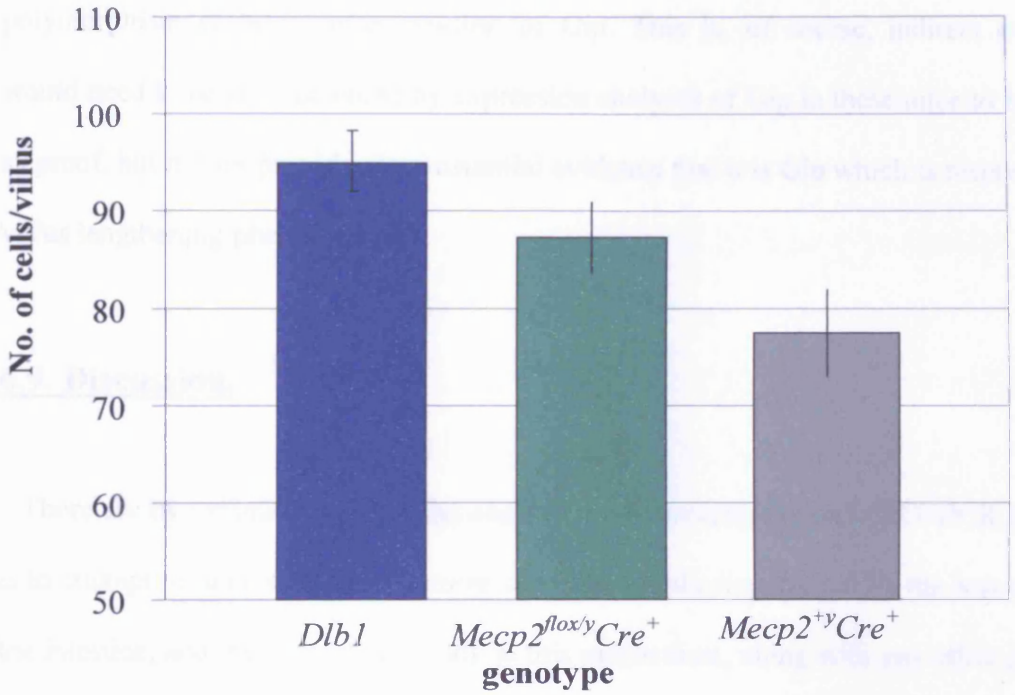
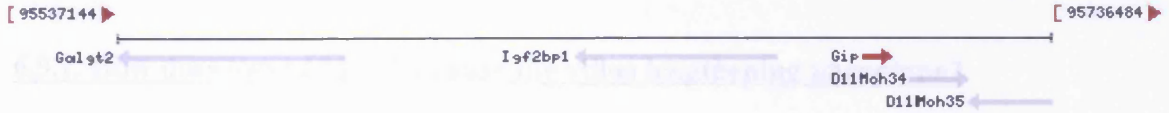


Fig 6.9b. The Galgt2 locus (mouse.)



Taken from :

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene&cmd=retrieve&dopt=default&list_uids=14607

Fig 6.9. Mice carrying a polymorphism at the Galgt2 locus, which lies extremely close to the Gip locus, have longer villi than wild-type mice. All other aspects of the phenotype were identical to Mecp2$^{flox/y}$Cre^{+} mice

The longer villi seen in the *Dlb1* mice provide some circumstantial evidence that Gip is involved in the lengthening phenotype. It is not unreasonable to suppose that the polymorphism causes over-expression of Gip. This is, of course, indirect evidence, and would need to be supplemented by expression analyses of Gip in these mice to be considered as proof, but it does provide circumstantial evidence that it is Gip which is responsible for the villus lengthening phenotype.

6.9. Discussion.

There are two primary aims in the analysis of the microarray and QRT-PCR data; the first is to attempt to unravel the mechanism of villus lengthening caused by the loss of MeCP2 in the intestine, and the second is to look at this mechanism, along with any other gene changes seen in the array/QRT-PCR data to reflect on the mechanisms which underlie the pathology of Rett syndrome.

6.9.1. How does loss of MeCP2 cause the villus lengthening phenotype?

A variety of peptides have been previously documented to cause alterations in the morphology of intestinal villi, for example, Igf-I, Glp-2 and bombesin. Glp-2 is derived from glucagon; no change in the expression of glucagon was observed. However, since Glp-2 is a post-translationally-derived product, this does not exclude the involvement of Glp-2 in the phenotype, as alterations in the processing of glucagon or breakdown of Glp-2 could lead to altered Glp-2 levels.

Igf-I is also not itself transcriptionally altered. Igf-I is known to be a potent intestinotrophic factor (Zeigler *et al* 1996.) Although the expression levels of *Igf-I* did not appear to be significantly different on the array (the fold-change is 1.44, but the high p-value of 0.56 rules this out as being significant) the expression of *Igfbp3* is changed on both the array (fold-change = 0.55, p = 0.029 T-test) and when examined using QRT-PCR (fold-change = 0.67+/-0.006.) *Igfbp10* is also downregulated both on the array (fold-change = 0.33, p=0.013 T-test.) Igfbp3 is the chief binding partner for Igf-I and is antagonistic to its action. Decreased Igfbp3 levels could therefore cause an increase in circulating Igf-I levels with no change in expression levels, and the raised levels of Igf-I could cause a degree of villus lengthening. The increase in Igf-I is unlikely to be the main cause of the phenotype, however; when rats are given exogenous Igf-I, both the crypt and the villus compartment expand (Zeigler *et al* 1996), and the MeCP2 deficient phenotype does not show any increase in the size of the crypt compartment. Whilst increased circulating Igf-I could certainly be a contributor to this phenotype, therefore, it seems unlikely that it is the major protagonist. It would be interesting to measure whether there is an increase in circulating Igf-I and if so, to exogenously administer this amount to wild-type mice; in this way the contribution of Igf-I to the phenotype could be quantified.

Surprisingly, *Bdnf* showed no n*et al*teration on the array (fold change = 1, p=0.89 T-test) possibly due to the fact that the levels of *Bdnf* appeared to be highly variable within classes and the SAM algorithm seems to be unable to deal effectively with highly variable data. However, investigation using quantitative RT-PCR showed that *Bdnf* was upregulated 4.36-fold+/-0.53 in $Mecp2^{flox/y}Cre^{+}$ mice compared with $Mecp2^{+/y}Cre^{+}$ mice. The magnitude of this change is surprisingly large; in *Mecp2*-null brain, basal transcription levels of *Bdnf* in non-activated neurons raise by approximately 2-fold, from 1% to 2% of the levels found in

activated neurons. The fold-change seen in the gut is considerably higher – could this indicate a role for *Bdnf in* the gut?

Bdnf is already known to be part of the brain-gut axis through its role in controlling food intake (Mattson *et al* 2004.) Many gastrointestinal hormones are also found in the brain and many of these hormones serve to mediate interactions between the brain and the gut (Strader and Woods 2005.) The possibility that the Ah-Cre has been induced in enteric neurons causing *Bdnf* upregulation to affect the neuronal plasticity of autonomic neurons in the gut cannot be entirely ruled out; it is known that enteric neurons can affect intestinal growth and development through other gut hormones such as Glp-2 (Bjerknes and Cheng 2001.) The possibility of extra-intestinal effects via non-specific recombination (i.e. a 'leaky' Cre) cannot be entirely ruled out. However, the transcriptional changes seen would remain valid even if there had been whole-body recombination.

6.9.2. Bdnf is involved in energy balance and metabolism.

Bdnf is already known to be involved in energy metabolism (Mattson *et al* 2004.) The level of expression of *Bdnf* is sensitive to alterations in energy balance; Bdnf levels rise in response to hypoglycaemia and intermittent fasting, and are lowered by hyperphagia (overeating) (Lindvall *et al* 1992, Lee *et al* 2002.) The *Bdnf* heterozygous knockout mouse is obese and has hyperglycaemia (Lyons *et al* 1999.) If these mice are subjected to dietary restriction, their Bdnf levels, body weights and blood glucose levels return to normal (Duan *et al* 2003.) The receptor for Bdnf is TrkB (tyrosine kinase receptor B;) *TrkB* knockout mice are also hyperphagic and become obese when fed a high fat diet (Xu *et al* 2003.) If *Bdnf* is knocked out conditionally in the brain after birth, mice become obese, develop elevated

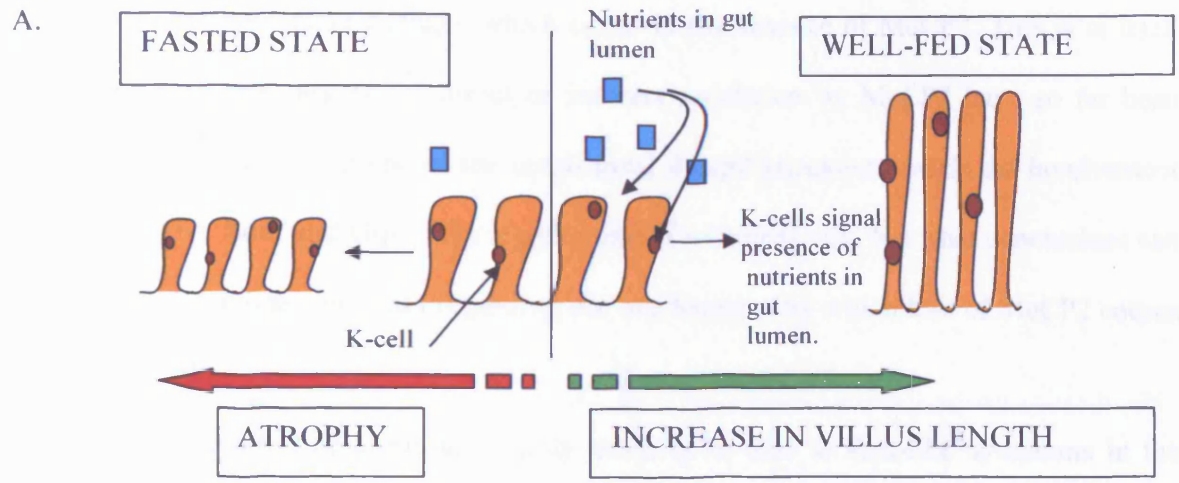
insulin and blood glucose levels and are, interestingly, hypersensitive to stress (Rios *et al* 2001.) Since Bdnf is so intimately involved in the cellular regulation of energy balance, knocking out *Mecp2* and thus deregulating *Bdnf* in the gut may interfere with the mechanisms which recognise nutrient status in the intestine.

One set of nutrient sensors in the gut are thought to be the K-cells, a specialised subset of enteroendocrine cells which release Gip (gastrointestinal polypeptide or glucose-dependent insulinotrophic polypeptide) in response to fat and carbohydrate in the gut (Cheung *et al* 2001.) The release of Gip stimulates the release of insulin, and so Gip is also involved in the regulation of energy balance, specifically blood glucose levels and adiposity. Gip was found to be upregulated by 2.13 +/- 0.6 fold in $Mecp2^{flox/y}Cre^{+}$ mice. Is this upregulation due to direct repression by MeCP2 or by an indirect mechanism? To definitively answer this question would require studies such as those carried out by Chen *et al* to examine the effects of MeCP2 mediated repression on Gip expression. Some evidence, however, does suggest that the rise in Gip levels may be a secondary effect, mediated through elevated Bdnf. Gip is a member of the Vip/secretin/glucagon family, and Bdnf has been shown to regulate the expression of Vip in retinal amarcrine cells (Cellerino *et al* 2003.) Bdnf also increases the level of somatostatin (another gastrointestinal hormone) in the brain (Villuendas *et al* 2001.) Bdnf can also induce expression of neuropeptide Y (NPY), which is a potent stimulator of feeding (Leibowitz *et al* 2005.) This shows that Bdnf is capable of directly regulating the expression of genes which have highly similar functions to Gip and are involved in the regulation of food intake controlled by the brain-gut axis. The upregulation of Bdnf may in turn cause an upregulation of Gip, disrupting the normal feedback mechanisms which inform the brain of the nutrient status of the gut. It is known that nutrient status and villus length are interrelated, and so this may lead to a lengthening of the villi in the small intestine. It is

interesting that the region of the K-cells and the region of the highest floxing density (the proximal small intestine) overlap. The number of enteroendocrine cells in the MeCP2 deficient mouse does not increase; this implies that any effects on the enteroendocrine cells are not due to altered lineage allocation but instead may be due altered transcription patterns within the enteroendocrine cells. A number of enteroendocrine markers are altered in the MeCP2-deficient phenotype; *somatostatin* is downregulated (fold-change = 0.51, p=0.028 T-test) as is *secretin* (fold-change = 0.51, p=0.019.) Neither *Vip* (fold-change = 0.93, p= 0.184 T-test) or *Cck-A* (fold-change = 0.43, p=0.128) were found to be altered, (*Npy* and *Pyy* were not found on the array.) This implies that a generalised deregulation of enteroendocrine cells is not occurring and hints that the effects are quite specific.

Bdnf is known to directly regulate proliferation in at least one cell type, in the rat hippocampal progenitors, and so may also regulate cell proliferation in the gut directly, and through effects on Gip. If *Gip*, like its family member *Vip* and closely functionally related genes *somatostatin* and *Npy*, is also regulated by Bdnf, then a mechanism for villus lengthening mediated via MeCP2, Bdnf and Gip, and augmented by Igf-I, could be proposed (fig 6.10.)

Fig 6.10. Model for villus lengthening in the absence of MeCP2



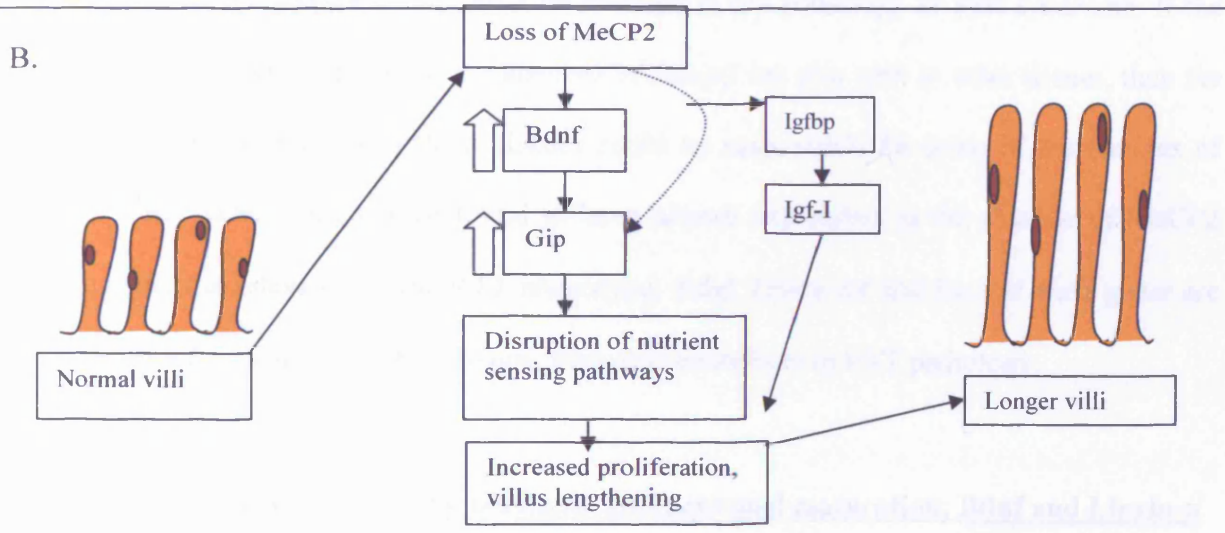


Fig 6.10. A model for villus lengthening in the absence of Mecp2.

A = K-cells act as nutrient sensors in the gut by releasing Gip in response to carbohydrate and fat in the gut lumen. Fasting leads to villus atrophy and re-feeding to villus growth.

B = A model for the villus lengthening seen in Mecp2-deficient mice; Reduced Mecp2 leads to increased expression of Gip, either directly or via upregulation of Bdnf. Igf-I levels are increased by the decreased expression of its chief binding partner, Igfbp3. Igf-I causes some villus lengthening directly, and the raised Gip levels mimic the 'well fed' state, leading to greater proliferation in the crypt compartment and villus lengthening.

6.10. What can the data reveal about the mechanisms underlying Rett syndrome?

The combination of microarray and quantitative real time RT- PCR analysis has revealed a number of significant gene changes which occur in the absence of MeCP2. This is in itself important, as so few targets for direct or indirect regulation by MeCP2 have so far been found. The intestinal phenotype of the conditional *Mecp2* knockout reveals the involvement of MeCP2, Igf-I, Bdnf and Gip in the lengthening of intestinal villi, but what conclusions can be drawn from the intestinal data regarding the mechanisms by which loss of MeCP2 causes Rett syndrome?

Changes in the intestine alone are highly unlikely to lead to Rett-like symptoms in the mouse. However, intestinal data may be relevant to the pathology of Rett syndrome. If the changes seen in the intestine in the absence of *Mecp2* are also seen in other tissues, then the transcriptional alterations in those tissues could be responsible for some of the features of RTT. Three genes which were found to have altered expression in the absence of MeCP2 have particular relevance to the RTT phenotype; *Bdnf*, *Liprin-α4* and *Ssc2*. If these genes are altered in RTT patients in other tissues, they may contribute to RTT pathology.

6.10.1. Rett syndrome, synaptic plasticity and neuronal maturation; Bdnf and Liprin-α

As has been discussed previously, many of the symptoms of RTT are thought to be due to defects in neuronal maturation and/or survival (Neul and Zogbhi 2004.) *Bdnf* has previously been shown to be directly controlled by MeCP2-mediated transcriptional repression, with a higher basal level of transcription found in *Mecp2* null neurons (Chen *et al* 2003.) MeCP2 binds selectively to the promoter and represses expression of *Bdnf*; on depolarisation,

calcium influx triggers calcium-dependent phosphorylation and release of *Mecp2* and the gene is transcribed. *Bdnf* was first described as a factor which promoted survival and differentiation of selected neuronal populations (Phillips *et al* 1990, Branchi *et al* 2004.) Bdnf is now considered to be a synaptic morphogen (Poo 2001, Chao 2003) whose expression and release is controlled by neuronal stimulation (Branchi *et al* 2004, Zara *et al* 1990, Ghosh *et al* 1994 Tao *et al* 2002) Bdnf levels have been found to be high in the hippocampus and cerebral cortex, areas which show reduced dendritic complexity in Rett syndrome brain. One recent study has shown that *Bdnf* over-expression profoundly alters the form and stability of basal dendritic branches, which correlates with the disordered arborisation seen in Rett patients; clearly, it is not unreasonable to suppose that disregulated Bdnf levels are the cause of at least some of the symptoms of Rett pathology.

Another gene whose product is known to be involved in synaptic plasticity is *Liprin-α4*, which is upregulated 2.15 fold +/- 0.141 according to the QRT-PCR data. *Liprin-α4,* also known as Ppfia4, is a multidomain protein which interacts with the Lar family of receptor protein tyrosine kinases (Serra-pages *et al* 1995, Ko *et al* 2003) and the Grip/Abp family of Ampa receptor interacting proteins (Wyszynski *et al* 2002) Mutations in the *C.elegans* homologue *syd-2* (synapse defective 2) and the *D.melanogaster* homologue *Dliprin-α* lead to impaired synaptic transmission (Zhen and Jin 1999) and defective axon terminal branching (Kaufmann *et al* 2002) respectively (Ko *et al 2003*,) Liprin-α4 is localised to the synapse and reduced levels result in reduced neuromuscular bouton number. MeCP2 is also localised to the synapse, in the post synaptic compartment (Jarrar *et al* 2003.) Perhaps many of the symptoms of RTT are caused by defects in brain architecture at the level of the connectivity of individual neurons or by defects in synaptic transmission?

Fig 6.11. Model for action of MeCP2 deficiency in Rett syndrome

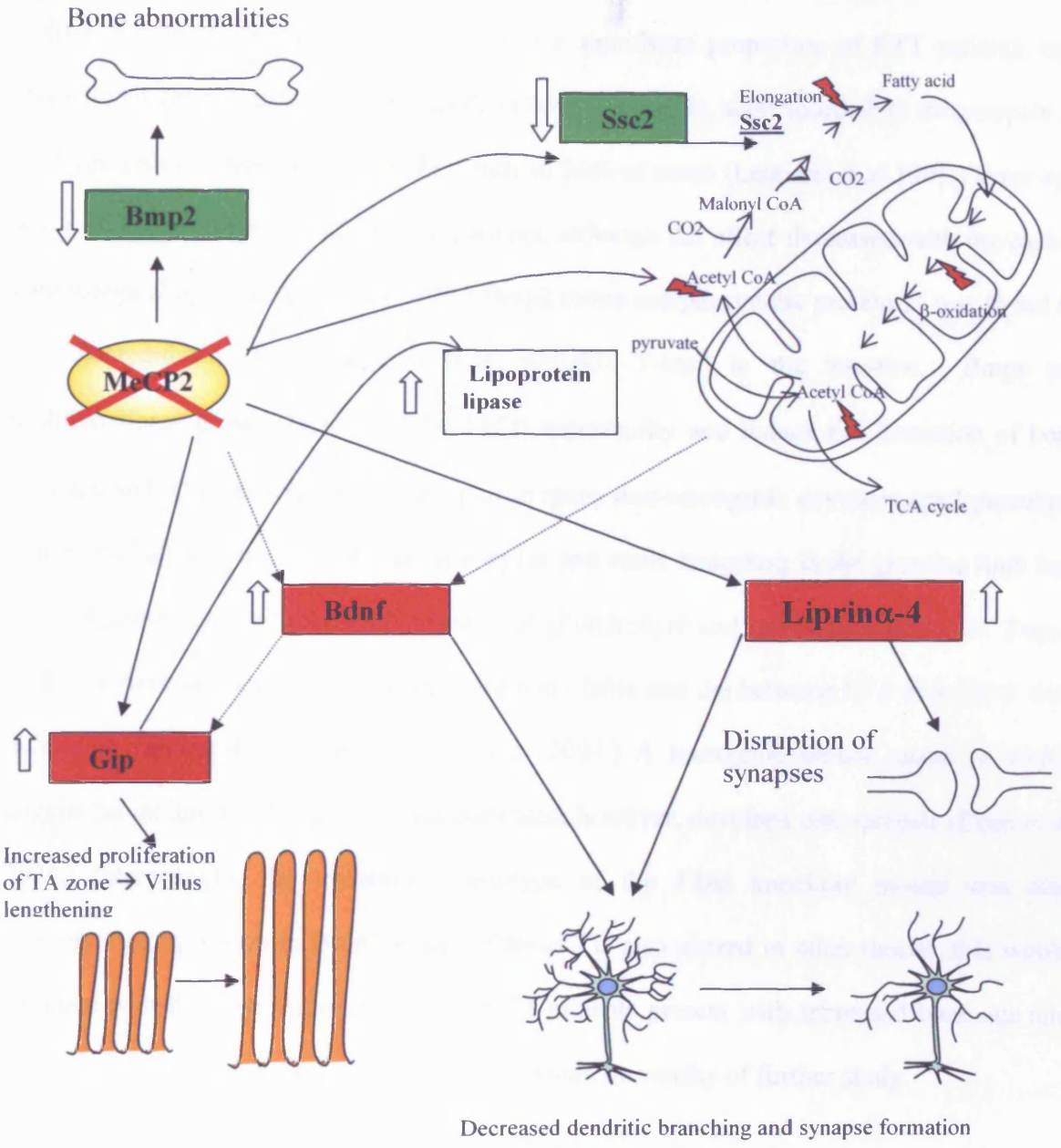


Fig 6.11. A putative model for how some of the changes observed in the intestine might, if seen in other tissues, contribute to the pathology of RTT.

Loss of MeCP2 *may deregulate Bdnf, which may have direct effects on neuronal plasticity and dendritic branching, altering brain architecture, and secondary effects on neuroendocrine peptides such as* Gip. *Increased* Gip *may cause increased proliferation of the crypt cells and therefore increased villus length. Altered* Bmp2 *levels may lead to bone abnormalities. Alterations in the metabolism of fatty acids may lead to metabolic stress, causing a further rise in* Bdnf. Liprin-α *levels are also disrupted, leading to changes in synaptic plasticity.*

6.10.2. Rett syndrome and bone abnormalities.

Bone abnormalities have been recorded in a significant proportion of RTT patients; one study found short fourth/fifth metatarsals in 65% of patients, short fourth/fifth metacarpals in 57% and reduced bone density in the hands in 86% of cases (Leonard *et al* 1995.) Bone age was also found to be altered in RTT patients, although the effect decreased with increasing chronological age (Leonard *et al* 1995.) Bmp2 (bone morphogenetic protein 2) was found to be downregulated (fold-change = 0.34, p=0.002 T-test) in the intestine. Bmps are multifunctional growth factors of the Tgf-β superfamily and induce the formation of both cartilage and bone, as well as plying a part in many non-osteogenic developmental processes (Chen *et al* 2004.) Bmp-2 interacts with Fgf-4 and sonic hedgehog in the growing limb bud to limit growth and induce the expression of chondrocyte and osteocyte precursors. *Bmp-2* null mice have not been studied as they are non-viable and die between E7.5 and E10.5 with defects in cardiac development (Chen *et al* 2004.) A transgenic mouse model in which noggin (an inhibitor of Bmp-2) is overexpressed, however, develops osteoporosis (Chen *et al* 2004.) Interestingly, the intestinal phenotype of the *Fkh6* knockout mouse was also accompanied by a drop in Bmp2 levels. If Bmp-2 is also altered in other tissues, this would correlate with the observation that some RTT patients present with increased bone age and bone abnormalities (Leonard *et al* 1995), and would be worthy of further study.

6.10.3. Does Rett syndrome cause metabolic stress via defects in fatty acid metabolism?

If *Bdnf* knockout mice are obese, then elevated Bdnf levels might be expected to have an opposite effect on body mass. RTT patients are known to have lower energy balance per kg/body weight (Motil *et al* 1994.) Could Bdnf be responsible for the lower energy useage of

RTT patients? The hypoactivity observed in these patients could then be seen as an attempt to conserve energy. A number of studies in Rett patients have indicated defects in basic metabolism (Motil *et al* 1994) and other studies have hinted that mitochondria may somehow be involved (Lappalainen and Riikonen 1994.) BDNF levels in the brain are known to increase in response to metabolic stress (Mattson *et al* 2002.) The fact that *Ssc2* is downregulated may also be significant; Ssc2, also known as Elovl2 in the mouse, is a component of the long chain fatty acid elongation pathway (Tvrdik *et al* 2000.) A 1999 study showed that Rett patients have reduced levels of long chain fatty acids (VLCFAs) and carnitine, which is slightly ameliorated by administration of exogenous L-carnitine, a common theraputant in Rett syndrome (Stradomska *et al* 1999.) Lowered VLCFAs have been considered a secondary effect of Rett syndrome (Stradomska *et al* 1999), and the reasons for the low levels were not known; if lowered Ssc2 levels are found in other tissues apart from the intestine, this may well provide part of the answer.

The lowered VLCFA levels are certainly an investigative avenue worth pursuing as it may lead to more effective methods of raising VLCFA levels in Rett patients. Proof of principle of the contribution of VLCFA dysfunction to brain disorders can be found in the mouse 'quaking' and 'jimpy' mutants, which both have mutations in other components of the LCFA elongation machinery. Recent research has explored the links between VLCFA depletion and brain function and has concluded that VLCFAs are essential for normal brain growth and function (Ruxton *et al* 2004.)

A generalised defect in fatty acid metabolism is hinted at by other gene changes seen on the array. Acetyl Coenzyme A dehydrogenase is downregulated (fold change = 0.425, p=0.01 T-test) on the array. Acetyl CoA is used as a substrate for fatty acid synthases in the cytoplasm. GIP is also known to increase the activity of lipoprotein lipase (Eckel *et al* 1979),

a gene seen to be upregulated on the array; lipoprotein lipase liberates fatty acids from triglycerides received from the gut (Keiffer *et al* 2003.) These fatty acids can then be converted to triglycerides for storage in adipose tissue.

If there are indeed defects in Rett mitochondria, carnitine levels, fatty acid metabolism and Ssc2 levels, then there may be fewer VLCFAs synthesised, the delivery of LCFAs to the mitochondria for oxidation may be impeded, and VLCFAs may be used for β-oxidation instead, further depleting VLCFA levels. Both the production of fatty acids and their breakdown may be affected. This may cause a degree of metabolic stress; and since Bdnf is known to rise in response to such a stress signal, the altered Bdnf levels may be both a direct consequence of the removal of the MeCP2-mediated repression and also a marker of this metabolic stress response, and may even reinforce each other in a positive feedback loop.

Could Rett syndrome be a disease characterised by metabolic stress caused by defects in fatty acid metabolism? Much further work would need to be done, including a thorough examination of mitochondrial function. The availability of both constitutive and conditional models of RTT should greatly aid this process.

Further avenues of study could involve examining other tissues in the constitutive knockout mouse to determine whether the transcriptional changes seen in the intestine are also seen in other tissues. Human samples could also be investigated, although this is a difficult area as it involves acquisition of human tissue for which ethical consent must be obtained. Other avenues could involve challenging constitutive *Mecp2* null mice with a high-glucose or high-fat diet to examine the effects of raised Gip levels, characterising the link between Bdnf and Gip (i.e. is Gip directly upregulated by MeCP2 loss, or is it an indirect effect caused by Bdnf?) Examining body weights, food intake and eating behaviour of MeCP2 null mice throughout life would also help to elucidate the effects of Gip and Bdnf on

the gut brain axis. Finally, a thorough characterisation of fatty acid metabolism in MeCP2 null mice would be valuable to examine the theory that Rett syndrome may be predominantly caused by alterations in neuronal growth and connectivity, perhaps caused by neuronal metabolic stress caused by defects in fatty acid metabolism and mitochondrial defects.

Chapter 7: Summary.

7.1. The role of MBD proteins is diverse.

The methyl CpG binding domain (MBD) proteins are a family of proteins which 'read' the methylation marks on DNA and act upon them. As our knowledge of the MBD proteins increases, it is becoming apparent that they are linked to a wide variety of cellular processes, from tumour suppression to neuronal development. The breadth of function of these proteins underlines how epigenetic processes are fundamental to the cell.

The MBD family consists of 6 known members; Mbd1, Mbd2, Mbd3, Mbd4, MeCP2 and Kaiso. Mbd1 is involved in transcriptional repression (Wade 2001) and also links DNA methylation to DNA repair. Mbd2 has a variety of binding partners which appear to modulate its activity (Lembo *et al* 2003.) Mbd3 does not appear to bind methylated DNA (Hendrich and Bird 1998) but is thought to be crucial to the developmental process, as reduced levels of Mbd3 cause developmental abnormalities in a model system (*Xenopus*) (Iwano *et al* 2004.)

Mbd4 does not play a role in transcriptional repression, but instead appears to maintain the integrity of the genome (Sansom *et al* 2003a) through its involvement with base excision repair (BER) mismatch repair (MMR) and the cell cycle response to DNA damage (Parsons 2003, Sansom *et al* 2003a) MeCP2 is a transcriptional repressor which can recruit chromatin remodelling factors to methylated DNA, remodelling chromatin into a form refractory to transcription and thereby silencing the gene (Nan *et al* 1998.) Kaiso binds p120 catenin and represses genes in a methylation-specific manner (Kim *et al* 2004.)

The correct interpretation of the methylation signal is vital to the normal growth and development of the organism, and this is illustrated by various knockout and conditional

mouse models which have been created. *Mbd1*$^{-/-}$ mice, although viable and fertile, have defects in hippocampal function and neurogenesis and have an elevated rate of aneuploidy in neurons (Zhao *et al* 2003.) *Mbd2*$^{-/-}$ mice are again viable and fertile, but cannot form a normal MeCP1 complex, leading to reduced repression from methylated promoters (Hendrich *et al* 2001.) The *Mbd2*$^{-/-}$ mouse also has problems in carrying, delivering and caring for offspring (Hendrich *et al* 2001.) The *Mbd3*$^{-/-}$ knockout is far less benign; the *Mbd3*$^{-/-}$ allele is embryonic lethal, and embryos are severely abnormal and in the process of being reabsorbed by day E8.5 (Hendrich *et al* 2001.) Reducing Kaiso levels in Xenopus leads to premature activation of genes before the mid-blastula transition and developmental arrest (Ruzov *et al* 2004.)

The MBD proteins are also involved in tumourigenesis; *Mbd2*$^{-/-}$ mice crossed onto an intestinal tumour prone *Apc*^{Min} background (a model of the human cancer syndrome FAP) develop fewer, smaller tumours (Sansom *et al* 2003c.) The role of Mbd4 in tumourigenesis is less clear; Mbd4 repairs C→T transitions, and thus might be expected to be protective against tumourigenesis (Bellacosa 2001.) Deficiency of Mbd4 has been shown to accelerate tumourigenesis on an *Apc*^{Min} background (Millar *et al* 2002.) MBD4 mutations are mainly found in a subset of (MMR-deficient) tumours (Riccio *et al* 1999) and so may be that MBD4 is not a major tumour suppressor in the human. No study has yet looked at possible epigenetic inactivation of MBD4 in human gastrointestinal tumours.

This project has shown that Mbd4 mediates the apoptotic response to two signals; Fas ligand and anoikis. Anoikis is apoptotic death triggered through the loss of normal attachment signals (Nagata *et al* 1999.) One property of a cancer cell is a loss of sensitivity to such signals, making a cancer cell more resistant to apoptosis (Hanahan and Weinberg 2000.) I have demonstrated that loss of Mbd4 results in significantly less apoptotic death in response

to loss of attachment signals in a model using isolated murine intestinal crypts. This also implies that Mbd4 may have additional tumour suppressor functions through mediation of anoikis. In contrast, I have shown that *Mbd4*$^{-/-}$ mice have an increased apoptotic response to Fas ligand. Since Mbd4 is known to interact with Mlh1, the apoptotic response to Fas ligand was also examined in *Mlh1*$^{-/-}$ mice, which showed a lowered apoptotic response to Fas ligand. These results lead to a putative model of the interaction between Mlh1, Mbd4 and Fas signalling, in which both proteins are part of a larger signalling complex (see model in fig 3.14.)

The role of MeCP2 in the murine mammary gland and intestine were also examined using conditional knockouts. Although RTT is considered to primarily affect the CNS, the presence of gastrointestinal problems (Motil *et al* 1999) and bone abnormalities (Leonard *et al* 1995) in RTT patients provided a rationale for examining extra-CNS tissues. Loss of MeCP2 in the murine mammary gland appeared to cause no changes in either development or involution. However, loss of MeCP2 in the intestine resulted in a marked phenotype; intestinal villi become 12.5% longer, and the proliferation and migration dynamics of the crypt-villus unit were perturbed. Increased levels of proliferation in the crypt were observed, characterised by an expansion of the transit amplifying (TA) zone with no increase in the size of the crypt. Migration rates were 12.9% increased and no change in apoptosis was observed.

No gross changes in cell lineage allocation were seen, with normal numbers and distribution of major cells types being observed. Microarray and QRT-PCR analysis revealed a number of transcriptional differences in the *Mecp2* null transcriptome; *Bdnf*, a known target of MeCP2 was over 4-fold upregulated, and smaller but significant increases were also seen in *Gip*, *Liprinα4*, and *Wdnm1*. *Igfbp3*, *Ssc2* and *Dlx5* were significantly downregulated. Increased Gip levels are thought to be responsible for the villus lengthening phenotype. If the

analysis can be extended to other tissues, the results also offer possible insights into the mechanisms underlying RTT pathology. Bdnf may play a role in creating the neuronal abnormalities seen in RTT patients. Lowered Ssc2, which is involved in very long chain fatty acid (VLCFA) metabolism, may explain the lowered VLCFA levels seen in RTT patients. The bone abnormalities may be explained by the upregulation of *Bmp2* seen on the array, although this was not followed up with QRT-PCR.

This work provides further evidence that epigenetic processes mediated by methyl CpG binding domain proteins are important in a number of developmental processes and in the development of cancer.

APPENDIX 1.

Published Work.

1. Maddison K, Clarke A R. "New Approaches for modelling cancer in the mouse." J.Pathol. 2005 205: 181-193.
2. Screaton R A, Keissling S, Sansom OJ, Millar C B, Maddison K, Bird A P. "Fas-associated death domain protein interacts with methyl CpG binding domain protein 4: A potential link between genome surveillance and apoptosis." PNAS 2003 100 (9) 5211-5216.

Journal of Pathology
J Pathol 2005; **205**: 181–193
Published online in Wiley InterScience (www.interscience.wiley.com). **DOI:** 10.1002/path.1698

Review Article

New approaches for modelling cancer mechanisms in the mouse

Kathryn Maddison and Alan R Clarke*
School of Biosciences, Cardiff University, Cardiff, CF10 3US, UK

*Correspondence to:
Alan R Clarke, School of Biosciences, Cardiff University, Cardiff, CF10 3US, UK.
E-mail: clarkear@cf.ac.uk*

Abstract

Mouse models of human cancer are vital to our understanding of the neoplastic process, and to advances in both basic and clinical research. Indeed, models of many of the major human tumours are now available and are subject to constant revision to more faithfully recapitulate human disease. Despite these advances, it is important to recognize that limitations do exist to the current range of models. The principal approach to modelling has relied upon the use of constitutive gene knockouts, which can often result in embryonic lethality, can potentially be affected by developmental compensation, and which do not mimic the sporadic development of a tumour expanding from a single cell in an otherwise normal environment. Furthermore, simple knockouts are usually designed to lead to loss of protein function, whereas a subset of cancer-causing mutations clearly results in gain of function. These drawbacks are well recognized and this review describes some of the approaches used to address these issues. Key amongst these is the development of conditional alleles that precisely mimic the mutations found *in vivo*, and which can be spatially and tissue-specifically controlled using 'smart' systems such as the tetracycline system and Cre-Lox technology. Examples of genes being manipulated in this way include Ki-Ras, Myc, and p53. These new developments in modelling mean that any mutant allele can potentially be turned on or off, or over- or under-expressed, in any tissue at any stage of the life-cycle of the mouse. This will no doubt lead to ever more accurate and powerful mouse models to dissect the genetic pathways that lead to cancer.

Received: 8 October 2004
Accepted: 17 October 2004

Keywords: **mouse models of cancer; p53; Rb; Ki-Ras; Myc; conditional alleles**

Introduction

Mouse models of tumourigenesis have been vital in our attempts to unravel the complex, multistage processes which confer a neoplastic phenotype. Mouse models of many of the major human cancers are available and are widely used in both basic research and clinical and therapeutic trials. Mice have many advantages as a model; they are small, easy to house, and have a short gestation time. Our increased understanding of the mouse genome over the last few years has also been a powerful force, allowing precise manipulation of the mouse genome to produce ever more sophisticated models.

It is also important to recognize the limitations of the mouse in modelling human pathologies. Mice have shorter lifespans and many differences in basic cellular processes (see Rangarajan and Weinberg [1].) The spectrum of common sporadic tumours in mouse and man is also different. Mice tend to develop sarcomas and lymphomas, derived from mesenchymal tissues, whereas humans are more likely to develop carcinomas derived from epithelial tissues, such as carcinomas of the colon, breast, lung, skin, and pancreas [1,2]. Immortalization kinetics in mice and humans are also different, due to the differences in telomeres and telomerase expression [1,3].

Certain mouse models recapitulate human disease extremely well. For example, overexpression of *c-Myc* in the mouse leads to similar pathologies (B-cell lymphomas), as it does in man [4,5]. However, identical genetic lesions do sometimes produce very different pathologies in the two species. A good example of this is the retinoblastoma gene product Rb. Rb transduces anti-proliferative signals and is an important tumour suppressor [6]. In humans, loss of the retinoblastoma tumour suppressor gene *RB* leads to the development of retinoblastoma at an early age, followed by osteosarcomas and small cell lung cancer [7]. In mice, however, loss of *Rb* very rarely causes retinoblastoma [5,7]; *Rb* null mice exhibit embryonic lethality, and heterozygotes develop pituitary carcinomas and thyroid tumours at high frequency [5,7,8]. *Prima facie*, then, the mouse model of retinoblastoma may appear to be of little use. However, even if the tumour spectra are different, the underlying cellular processes that lead to the human disease may well be similar. For example, it was hypothesized that there may be redundancy or compensation between *Rb* and family members such as *p107* and *p130* [9– 11]. Studies by Dyer *et al* in chimeric models [10,11] found that

p107 was up-regulated in the developing mouse retina and that $Rb^{+/}$ $p107^{-/-}$ null mice do develop bilateral retinoblastomas at high frequency [11]. This suggests a compensatory mechanism in the mouse which does not exist in the human retina. Such experiments show the complexity of signalling pathways and provide important insights into their mechanisms of action in both the mouse and the human.

Even models which do not appear to recapitulate the human disease exactly can be valuable tools for understanding the mechanisms of tumourigenesis. Recent advances in gene manipulation such as conditional gene targeting, high throughput screening strategies, and informatics are all facilitating a new generation of mouse models which promise to recapitulate human disease more faithfully and so bring us closer to the goal of curing human neoplasias.

The first generation of mouse models of cancer

The first generation of models was created by constitutively expressing cellular and viral oncogenes such as *c-Myc* [4], or by 'knocking out' tumour suppressor genes such as *Rb* [5,7,12,13]. These models continue to contribute greatly to our understanding of the cellular process underlying cancer development. However, they have a number of drawbacks; first, expression of an exogenous gene, or ablation of a vital tumour suppressor gene, is often incompatible with normal development, leading to embryonic lethality or severe developmental disruption or sterility in the adult [14– 16]. Second, whole body expression or ablation of a gene does not mimic sporadic tumour development *in vivo*, where abnormal cells carrying the genetic lesion are surrounded by normal tissue. Third, with addition transgenesis there is usually little control over site of integration and copy number. In these circumstances, the exogenous gene may affect genes near the insertion site, or be affected by endogenous control elements [14– 16]. Fourth, spatial and temporal control of the transgene is limited, although some models do achieve good tissue specificity by the use of a tissue-specific promoter. One such successful example is a model of mouse pancreatic cancer that uses the SV40 T-antigen under control of the insulin promoter to drive pancreatic transformation [17].

Recent advances in gene targeting technology have led to models in which the expression of single or multiple genes can be tightly controlled, both spatially and temporally. Using this approach, the problem of embryonic lethality can be circumvented, as has been shown for many genes, including the adenomatosis polyposis coli gene (*Apc*) [18]. The gene of interest can then be mutated in a specific tissue at a defined point in the life cycle, allowing the effect of removal to be precisely defined. Cell-targeted alteration of gene expression also recapitulates the *in vivo* development of sporadic cancers [19] and allows the influence of normal surrounding tissue to be examined.

A range of control systems have now been developed which allow the precise spatial and temporal control of gene expression, and are predominantly based on a bitransgenic approach, as described in detail below. Mice carrying a tissue-specific inducible transactivator gene are crossed to mice carrying the allele of interest which has been engineered so as to be controlled by the transactivator. Offspring that carry both transgenic elements can then be treated with an appropriate inducer to express the transactivator gene in a specific tissue, which then acts on the desired allele.

Transgenic tools. Cre-Lox and FLP: site-specific recombinases

Cre (Causes recombination) recombinase is a site-specific recombinase of the integrase family, isolated from bacteriophage P1 [20– 22]. Cre catalyses site-specific recombination between defined 34 bp '*Lox P*' (locus χ of crossover P1) sites. If a gene is placed between two Lox P sites and exposed to Cre, then the gene will be excised or 'floxed' out. A similar recombinase, FLP, isolated from *S cerevisiae*, also catalyses recombination from similar 34 bp FRT (FLP recombination target) sites [20,23,24]. The Cre and FLP systems can be used to create tissue-specific conditional deletions of an allele, overcoming the problem of embryonic lethality or developmental defects [20]. Mice carrying the Cre recombinase under the control of an inducible or tissue-specific promoter are crossed with mice carrying the gene of interest that has been flanked by Lox P sites [20]. When the mice are given the appropriate inducer, Cre is expressed in a spatially defined manner and the gene of interest is 'floxed out' in a specific tissue (Figure 1A). The Cre-Lox and FLP systems can also be used to activate a gene. In this instance, a STOP cassette flanked by Lox P or Frt sites is placed before the gene of interest, and expression of the Cre results in the cassette being excised, allowing expression of the gene [20] (Figure 1B).

Because Cre recombinase is exogenous to the mouse [25,26], it was thought that expression would have no effects other than to target the specific Lox P sites. However, recent reports show that Cre appears to act, albeit with low affinity, with 'pseudo-Lox P sites' in the mouse genome [26] and its expression in mammalian cells can have deleterious effects on the stability of the mouse genome, including chromosome rearrangements in mouse spermatids [25]. Although the implications of this are as yet unknown, this has in part stimulated the development of self-excising Cre vectors [27].

Several forms of Cre delivery have been developed; thus, Cre may be ligand-induced, such as through creation of the fusion protein CreER™ transgene, which is tamoxifen-dependent [20,28]. Cre may also be delivered packaged in adenovirus [20,29], or may

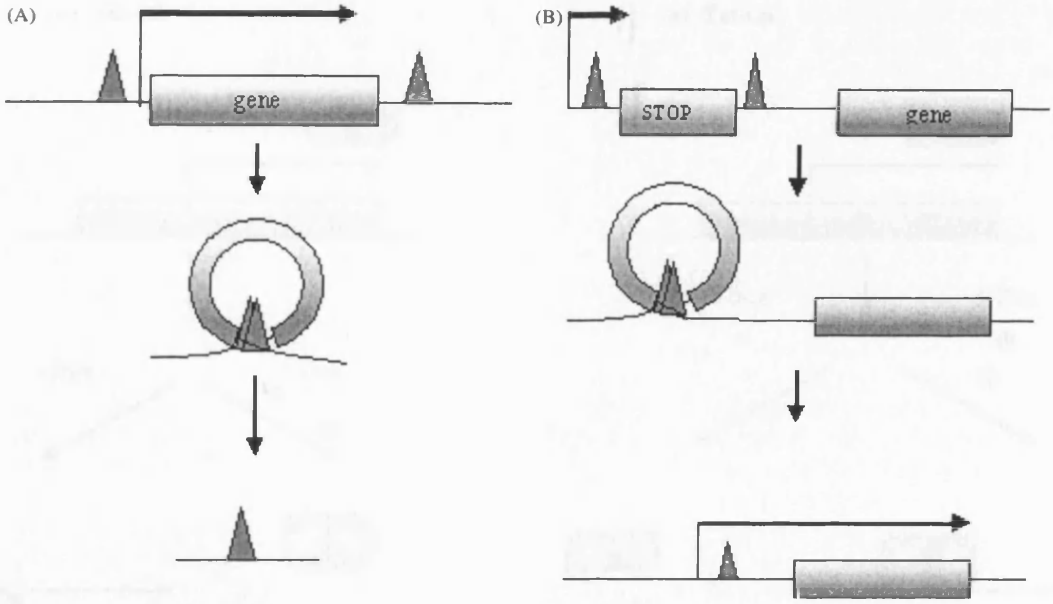


Figure 1. The Cre-Lox system. (A) Gene ablation: mice containing a transgenic construct in which the gene of interest is flanked by two Lox P sites are crossed with mice carrying a Cre recombinase gene under the control of a tissue-specific or inducible promoter. On activation of the Cre recombinase, the Lox P sites recombine, excising the gene of interest. (B) Gene activation: mice carrying a tissue-specific Cre recombinase are crossed with mice carrying a construct where the gene of interest is preceded by a STOP cassette, preventing transcription. Upon activation of the Cre recombinase, the STOP cassette is recombined out ('floxed out') and the gene may be transcribed. The Cre recombinase can be activated in several ways (detailed in the main text), allowing tight spatial and temporal control of gene expression

be placed under the control of a promoter which has a well-defined spatio-temporal expression pattern, such as the use of the *WAP* [30] and *BLG* promotors to deliver Cre expression to the mouse mammary gland [31]. Finally, Cre activity may be controlled through the use of an inducible promoter; for example, through use of the *Cyp1A* promoter [32]. Despite the apparent versatility of these systems, the one caveat that remains is their irreversibility. Thus, if a model requires a gene to be switched on and off, then a more suitable alternative may be constructed with, for example, a tetracycline-responsive system.

Tetracycline-inducible systems

The tetracycline (tet)-dependent system developed by Gossen and Bujard [33,34] allows tight spatial and temporal control through the use of a tissue-specific transactivator and an effector gene (Figure 2). Unlike the Cre and FLP systems, the tet system allows genes to be turned on and off at will, via the administration of the inducer (tetracycline, or more usually doxycycline) The system can also be tuned to either switch on gene expression (the 'tet-on' or tTA system) or switch it off (the tet-off or rtTA system [35]).

The tet-on (tTA) system uses the Tn10-specific tetracycline resistance operon of *E coli* (Figure 2) and is composed of two parts; a transactivator and an effector [35]. The transactivator is composed of the DNA binding domain of the *E coli* tetR gene fused to the transactivation domain of the VP16 herpes simplex virion protein 16 gene, under the control of a tissue-specific promoter which drives expression of the tTA (tetR/VP16) fusion protein in the desired tissue. The second component is a construct containing the gene of interest driven by the minimal promoter of human cytomegalovirus (Pcmv), under the control of the tet operator, tetO. In the absence of doxycycline, the tTA protein binds tetO and activates the minimal promoter, driving expression of the gene. However, when doxycycline is added, it binds to the tTA protein and causes a conformational change which prevents it from binding tetO. The promoter is not activated and the gene is not expressed [35] (Figure 2A).

In the tet-on system (rtTa), the second component of the binary system is identical to the tet-off system, but the transactivator component is made up of a tissue-specific promoter and a mutant version of the tetR DNA binding domain fused to the VP16 transactivator [35]. This produces a mutant rtTA protein: in the absence of doxycycline, rtTA does not bind to the tetO operator, and there is no gene expression. When doxycycline is added, the protein undergoes a conformational change and can bind the operator, leading to expression of the gene of interest. If absolute repression of a gene is required, the transactivator can be constructed with a silencer with the KRAB-AB silencing domain of the *Kid1* gene inserted instead of the VP16 domain. This prevents leaky transcription caused by the basal level of affinity of tTA for tetO in the absence of doxycycline. The tet-on, tet-off, and silencer systems can be used together to control more than one gene at a time.

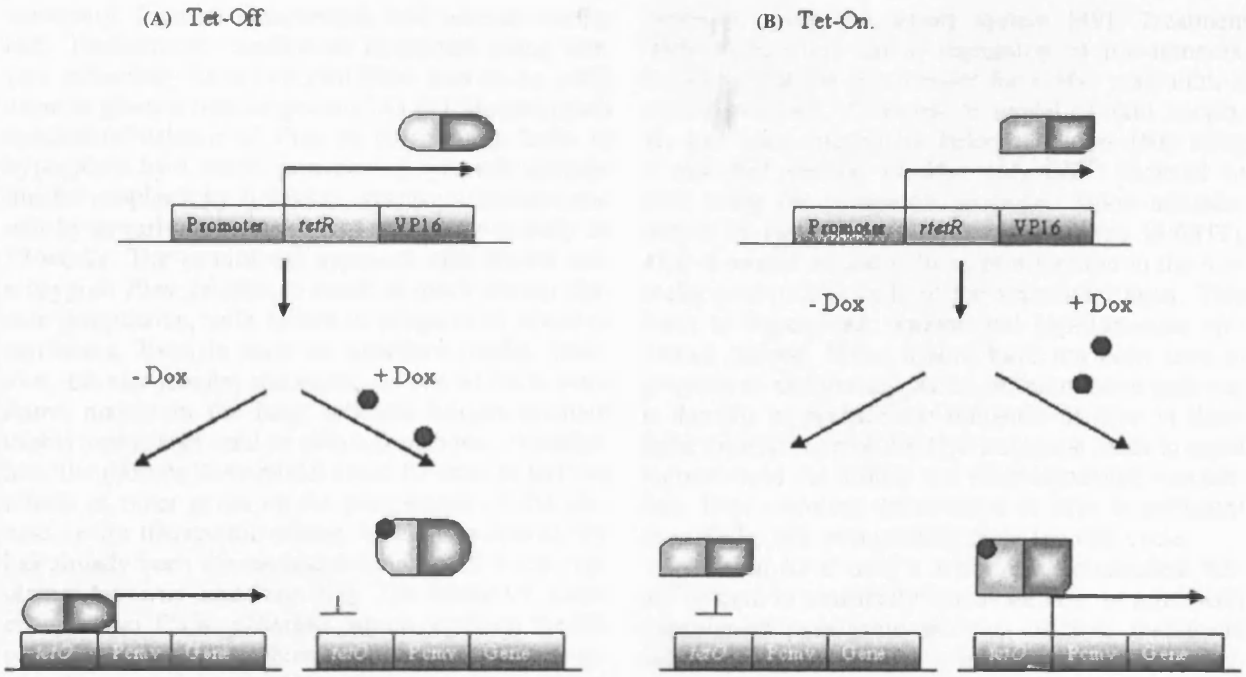


Figure 2. The tetracycline-responsive system. (A) In the tet-off system, the DNA binding domain of the *E coli* tetR gene is fused to the transactivation domain of the herpes simplex virion protein 16 (VP16) gene, under the control of a tissue-specific promoter which drives spatially controlled expression of the tTA (tetR/VP16) protein in the desired tissue. The effector construct contains the transgene of interest, driven by the minimal promoter of human cytomegalovirus (P_{cmv}) under the control of the tet operator tetO. In the absence of doxycycline, tTA binds tetO and drives expression of the transgene. When doxycline is added, it binds tTA and prevents it from binding tetO, stopping expression of the transgene. (B) In the tet-on system (rtTa), the second component of the binary system is identical to the tet-off system, but the transactivator component is made up of a tissue-specific promoter and a mutant version of the tetR DNA binding domain fused to the VP16 transactivator. In the absence of doxycycline, rtTA does not bind to the tetO operator and there is no gene expression. When doxycycline is added, the protein undergoes conformational change and can bind the operator, leading to expression of the gene of interest

Other conditionally inducible systems are under development, although they are not as progressed as the tet and Cre-Lox systems. These include systems based on the insect steroid moulting hormone ecdysone [36], the progesterone analogue mifepristone [37], the Lac operator-repressor system [38], and the GAL4/UAS system [39].

Modelling six of the hallmarks of cancer

In their 2000 paper [40], Hanahan and Weinberg proposed that cancer cells can be seen as having six characteristics, or 'hallmarks:' self-sufficiency in growth signals; non-responsiveness to anti-proliferative signals; resistance to apoptosis; unlimited replicative capacity; angiogenesis; and the ability to metastasize and invade. Aspects of each of these different facets have been modelled in the mouse, providing valuable insights into the relevance of these mechanisms. The remainder of this review will therefore focus on a number of examples relating to each of these 'hallmarks'. We will also briefly discuss the modelling of other factors which assist or trigger carcinogenesis, such as genomic instability or alterations of the epigenetic imprint of the cell.

Self-sufficiency in growth signals

Oncogenes and tumour suppressor genes (TSGs) are involved in regulation of the cell cycle. The core of the cell cycle machinery consists of a family of cyclin-dependent kinases (CDKs), which drive the cell through the cell cycle by phosphorylating key effector substrates [41]. The CDKs are primarily positively regulated by the cyclins and negatively regulated by the members of the Ink4 family and by p21kip and p19Arf. In turn, these primary regulators are themselves regulated by many other molecules in response to signals from inside (eg DNA damage surveillance) and outside (eg growth hormones) the cell. The tight control of cell growth and division is vital to a multicellular organism and mutations in these key cycle regulators can lead to uncontrolled division — one of the hallmarks of the cancer cell.

Mouse models of cell cycle regulators are allowing us to unravel the complex web of control that dictates when and how fast cells divide. For example, *PTEN* (phosphotin and tensin homologue deleted on chromosome 10) is a phosphatase which antagonizes the PI3K pathway [42]. *PTEN* is a tumour suppressor gene and is mutated in a variety of human sporadic tumours. Constitutive *Pten* homozygous knockouts show early embryonic lethality; *Pten* heterozygotes develop a range of tumours of the intestine, prostate,

mammary, thyroid, endometrial, and adrenal tissues [42]. Furthermore, conditional knockouts using Cre-Lox technology have revealed *Pten* loss as an early event in prostate tumourigenesis [43,44]. Homozygous conditional deletion of *Pten* in the prostate leads to hyperplasia by 4 weeks post-floxing, prostate intraepithelial neoplasia by 6 weeks, invasive adenocarcinomas by as early as 9 weeks, and metastasis as early as 12 weeks. The conditional approach also shows heterozygous *Pten* deletion to result in much slower disease progression, with failure to progress to invasive carcinoma. Even in such an excellent model, however, caveats remain; metastatic lesions in mice were found mainly in the lung, whereas human prostate cancer metastases tend to develop in bone. Nonetheless, the prostate *Pten* model could be used to test the effects of other genes on the progression of the disease, or for therapeutic testing. Indeed, co-operativity has already been demonstrated for the cell cycle regulators *Ink4aArf* and *Pten* [45]. The *Ink4a/Arf* locus encodes two TSGs: *p16Ink4a*, which regulates the Rb pathway, and *p19Arf*, which regulates *p53*. The compound mutant *Ink4a/Arf*$^{+/-}$ *Pten*$^{+/-}$ mouse shows shortened tumour latency and *Ink4a/Arf*$^{-/-}$ *Pten*$^{+/-}$ mice shorter latency still (although, intriguingly, progression to invasive carcinoma is not seen in this model.)

Murine models are also illustrating how incompletely we understand cell cycle control. Thus, although the cell cycle regulators Cdk4 and Cdk6 have been considered critical initiators of the cell cycle, Cdk4 has been shown to be dispensable for proliferation in many cell types [46,47]. Furthermore, *Cdk6* knockout mice are viable with only minor defects [48], and no synergism has been observed in mice doubly null for *Cdk2* and *Cdk6*. By contrast, deficiency of both *Cdk4* and *Cdk6* does lead to embryonic lethality, although this is due to severe anaemia, rather than a general cell division defect [46]. MEFs (murine embryonic fibroblasts) from *Cdk4*$^{-/-}$ *Cdk6*$^{-/-}$ mice were capable of division in culture (albeit more slowly than wild-type cells) and underwent immortalization after repeated passaging. Mutant cells also became quiescent in response to serum withdrawal but re-entered the cycle normally with the appropriate stimuli. Cells also had lowered Rb phosphorylation — a serious challenge to the accepted theory that full Rb phosphorylation by both Cdk2 and Cdk4/6 is required for G1/S transition [48]. Taken together, the data from these mice therefore challenge our understanding of the role of the D-type cyclins and suggest that alternative mechanisms are capable of initiating proliferation.

c-Myc

If a cell has activated a gene allowing self-sufficiency in growth factor signals, deactivating that gene may lead to tumour regression. Marinkovic *et al* have used a mouse model in which *c-Myc* is conditionally expressed under the control of a lymphoid-specific promoter using the tet-off system [49]. Treatment with doxycycline led to regression of the tumours, implying that the requirement for *c-Myc* activation is continuous [46]. A reversible model of skin neoplasia has been created by Pelengaris *et al* [50] using a modified version of *Myc* (*MycER*^{TM}) targeted to skin using the involucrin promoter. Upon administration of the inducer 4-hydroxytamoxifen (4-OHT), *Myc* is turned on and induces proliferation in the normally post-mitotic cells of the suprabasal layer. This leads to hyperplastic lesions and papillomatous epidermal lesions. These lesions have not been seen to progress to malignancy, as the differentiation pathway is thought to override the influence of *Myc* in these cells. Deactivation of the *Myc* transgene leads to rapid regression of the lesions and their associated vasculature. Even transient deactivation of *Myc* is sufficient to exclude cells permanently from the cell cycle.

Jain *et al* have used a tetracycline-controlled 'tet-off' system to transiently inactivate *Myc* in mice with transplanted osteogenic sarcoma cells or transgenic tumours [51]. When *Myc* was inactivated, the sarcomas and osteogenic tumours underwent differentiation to mature bone and showed significant regression. Cultured tumour cells in which *Myc* was subsequently reactivated did not resume proliferation but instead underwent apoptosis. In all of these tumour models, it will be interesting to see if the requirement for *c-Myc* continues even after other enabling mutations have occurred in the cell. If so, targeting *c-Myc* could be an extremely productive therapeutic strategy for some tumour types.

Ras

The product of the *Ras* oncogene is involved in a wide range of cellular processes, including progression through the cell cycle, transcription and translation, cell survival, and apoptosis, via interactions with a wide range of effectors [52]. Several mouse models with defects in *Ras* have been developed, both conditional and conventional. *K-Ras*$^{-/-}$ mice die at E12–E14 from liver defects and anaemia [53,54], while *N-Ras*$^{-/-}$ and *H-Ras*$^{-/-}$ mice are viable, with the former having defects in immune and T-cell function and the latter having no obvious phenotype except for decreased carcinogen-induced tumourigenesis. Conditional *Ras* alleles are a relatively recent development. Johnson *et al* have created a mouse model which carries a latent oncogenic allele of *K-Ras*, which can undergo spontaneous activation *in vivo* [55,56,57]. One hundred per cent of animals carrying this allele developed multiple lung tumours, with a high proportion also developing thymic lymphomas and skin papillomas. The lung tumours appear to be similar to human non small-cell lung cancer (NSCLC) progressing through hyperplastic and dysplastic stages before progressing to carcinoma. The spontaneous and stochastic nature of the activation events makes

this mouse an excellent model of spontaneous *K-Ras*-induced lung cancer in humans, and also may mimic the interaction of adjacent mutant and normal tissues seen in human cancer *in vivo*.

K-Ras is widely mutated in human tumours, although this varies tremendously with tumour type, so cellular context may be highly important [52]. A recent paper demonstrates this, using a *K-Ras* allele activated by the Cre-Lox-mediated removal of a STOP cassette to drive *K-ras* expression [58]. Guerra *et al* also engineered a bicistronic colour marker into their system, and thus show that expression of the endogenous *K-Ras* allele appears to have no discernible consequences in most cell and tissue types [58]. A low incidence of sarcomas and anal papillomas was seen in these mice; however, 100% developed multiple lung lesions, which derived from type I pneumocytes, although a small proportion were derived from Clara cells. This replicates the *K-ras*-induced lung tumours found by Johnson *et al* [57], Jackson *et al* [59], and Meuwissen *et al* [60].

If *Ras* oncogenes are expressed under the control of a highly active promoter, then cells enter G1arrest and senescence due to oncogenic stress [61]. However, the models of Johnson *et al* [57] and Guerra *et al* [58] show that under endogenous control, *K-RasV12* does not induce senescence, but instead causes cells to acquire hyperproliferative capacity akin to that of immortalized cells. This shows how important levels of expression are in determining the exact reaction of a system.

Oncogenic *Ras* has also been shown to be important for tumour maintenance. Wong and Chin have created a mouse melanoma model using doxycycline-induced *H-RasV12G* on an *Ink4a* null background [62]. Induction and maintenance of *H-RasV12G* were shown to be absolutely required for induction and maintenance of melanomas. When doxycycline was removed and *H-Ras V12G* was down-regulated, extensive apoptosis and regression of tumours were observed.

Lack of response to anti-proliferative signals

Rb is a major transducer of anti-proliferative signals [63], and mutations in *RB1* in humans produce the rare childhood malignancy retinoblastoma [63]. Efforts to model this disease in mice were hampered by the fact that *Rb* knockout mice do not develop retinoblastoma at high frequencies [7]. The first true phenotypic mouse model of retinoblastoma was created in 2004 by Zhang, Schweers, and Dyer [63], and contains six engineered alleles: *Rb1*, *p53* and *p107* null alleles; floxed *p53* and *Rb* alleles; and a *Chx 10-Cre* which targets retinal progenitors. Retinal progenitor cells lacking Rb and p107 keep proliferating past the time when they should undergo terminal cell-cycle exit, showing that p107 compensates for lack of Rb by preventing deregulated proliferation. Tumours which lacked Rb, p107, and p53 were highly aggressive and metastatic; hence p53-mediated apoptosis may also be a barrier to aggressive retinoblastoma formation in the mouse. This model will no doubt be extremely use for preclinical research and is being refined still further — a key difference between the murine and human diseases is that the number of Rb inactivation events in the human retina is much lower, leading to a small number of focal, clonal tumours. In an attempt to mimic this in the mouse, a retrovirally delivered Cre is being used to target small numbers of cells and increase the accuracy of the model [63].

Evading apoptosis

Endogenous and exogenous DNA damage are considered key initiators of the apoptotic response [40], and many TSGs are involved in either the repair or sensing of DNA damage. It is now widely assumed that cells with damaged DNA will attempt repair, but if this is not possible, they may default to an apoptotic pathway. Avoidance of this pathway is a frequent feature of neoplastic cells, which may progress through the cell cycle regardless of damage, resulting in mutation or karyotypic alterations [40]. One key mediator of the apoptotic response is p53, which integrates cellular stress signals and triggers apoptosis, repair or senescence.

p53

Constitutive defects in human *p53* are associated with the Li–Fraumeni syndrome, which predisposes to a wide range of tumours [64] and the first mouse models of p53 defects were null alleles which developed a variety of neoplasias [65] 69]. Although Li–Fraumeni patients mainly develop epithelial tumours [70], *p53* null mice predominantly develop lymphomas [71]. The short lifespan of *p53* null mice due to early-onset lymphoma has largely precluded the study of tumours with longer latencies. It was also soon realized that most spontaneous and familial human mutations of p53 are missense mutations in the DNA binding domain (DBD) [72] and so the conventional null alleles were not a faithful model of *p53* mutation *in vivo*. Missense *p53* mutations associated with cancer are generally gain-of-function mutations, which have been ascribed dominant negative effects [73].

Several new mouse models of *p53* function are therefore being developed which better mimic the real (and diverse) range of human mutations. For example, Liu *et al* have produced a *p53* DBD missense mutation that carries a common mutation (*R172H*) and also a splice variant that reduces expression levels to almost wild type [74]. Clear differences were found between the $p53^{+/-}$ and $p53^{R172H}$ gain of function mouse, with the latter developing fewer lymphomas and more carcinomas, with a higher degree of metastasis in the tumours observed.

Another engineered version of the p53 protein shows the sheer complexity of the effects each of the mutations in *p53* may have *in vivo*. Lozano's group

have recently created a mouse carrying a version of *p53* which cannot induce apoptosis but can still induce cell cycle arrest [75]. Mice homozygous for this mutation ($p53^{515c/515c}$) have significantly delayed tumour onset when compared with *p53* null mice. Tumours which arose in the $p53^{515c/515c}$ mice showed a longer latency and had more stable genomes than those arising in *p53* null mice. Cells from *p53* null mice had extensive aneuploidy and chromosome breaks and abnormalities, whereas cells from $p53^{515c/515c}$ mice were generally 2N, 4N or 8N with many fewer abnormalities. This suggests that the $p53^{515c}$ mutation is capable of suppressing genomic instability, and also argues strongly that abrogation of the apoptotic programme does predispose to malignancy.

It is not just specific mutations in *p53* that need to be explored, but also its basic functions. It had been questioned whether loss of heterozygosity (LOH) of the remaining wild-type *p53* allele was an important factor in tumourigenesis in Li–Fraumeni patients [70]. In $p53^{+/-}$ models [76], around 50% of tumours showed LOH, but in the $p53^{R172H}$ model, only 9% of tumours exhibited LOH, implying that LOH is not absolutely necessary for tumour formation (although it may still represent a significant route to tumourigenesis in a subset of tumours). Another important mechanistic question answered by new p53 models is whether the transcriptional transactivation activity of p53 is essential for p53-mediated apoptosis and tumour suppression [69]. A mouse model has been developed with missense mutations in the transactivation domain by Jimenez *et al* ($p53^{L25Q/W26S}$) [77]. The mutant p53 protein produced still has a functional DBD. Mice carrying this mutation have lowered p53-mediated apoptosis in thymocytes *in vivo* and their cultured embryonic fibroblasts become neoplastic. This strongly argues that transcriptional transactivation is indeed essential for p53's function as a tumour suppressor gene.

Although these studies therefore support differing roles for p53 in tumour suppression (which includes a role in initiating apoptosis), the precise relevance of p53-mediated apoptosis to tumour suppression remains somewhat unclear. The original observation that p53 mediates the apoptotic response to DNA damage gave rise to the simple hypothesis that p53 mediates its primary tumour suppressive activities through initiation of the apoptotic response, and that failure to delete cells harbouring DNA damage would result in increased mutation and thereby to accelerated tumourigenesis. This hypothesis has, however, been somewhat difficult to prove, with clear tissue-specific differences in the role of p53-mediated apoptosis. For example, it appears that the loss of p53-dependent apoptosis is a good predictor of mutation burden and tumour predisposition within the haematological system but not the intestine [78].

Despite these complexities, conventional models and null alleles still have a central role in p53 research. For example, combinatorial use of alleles null for *p53* and *p19Arf* and ectopically expressed *Mdm2* have been used to explore the *in vivo* interactions between these genes [79]. Increased Arf levels are known to inhibit Mdm2-mediated proteolytic degradation of p53 [80,81], and Arf expression can lead to increased oncogenic signalling, such as overexpression of c-Myc or activation of Ras [82,83]. Various combinations of the three alleles have been used to investigate cooperativity between the three gene products [78]. Thus, *p19Arf* deficiency accelerated tumour development in the *Mdm2*-overexpressing mouse, whilst the principal TSG activity of *Arf* appears to be p53-dependent.

Angiogenesis

As a solid tumour grows, its size exceeds its blood supply and the inner part of the tumour may become hypoxic [84,85]. To maintain its rate of growth, the tumour must establish its own blood supply to ensure a constant supply of oxygen and to remove catabolites. Angiogenesis is difficult to model *in vivo*; many models require tissue to be transplanted into the mouse, which causes injury, one of the normal triggers of apoptosis [84].

Although Ras and Myc are implicated in angiogenesis, the process is more complex than may have first been appreciated. Both oncogenic K-Ras and H-Ras can stimulate VEGF expression [62]. In the mouse model of melanoma developed by Wong and Chin and described earlier [62], decreased expression of H-Ras resulted in vascular regression before the main tumour regression. High rates of apoptosis were seen in cells lining tumour-associated vessels, indicating that continued activation of Ras may be needed for a stable tumour vasculature. Wong and Chin also showed that loss of Ras-stimulated VEGF expression is not responsible for tumour regression and that high VEGF levels are not sufficient to maintain the tumour vasculature if H-Ras is absent. There must therefore be another mechanism by which Ras promotes angiogenesis.

Modulation of angiogenesis can also occur through deregulation of thrombospondin 1 (TSP1), an inhibitor of angiogenesis. Thus, expression of TSP1 can be inhibited by many of the oncogenes associated with malignant progression, namely oncogenic *Ras, c-Myc, v-Src, c-Jun* and *Id1* [85]. Furthermore, expression of TSP1 has recently been shown to be under the control of TSGs such as *p53* and *Pten* [86].

Although the mechanisms which control angiogenesis are still incompletely understood, evidence is growing for the existence of an 'angiogenic switch', whereby the normal balance of inhibitors and promoters of angiogenesis is disrupted, allowing the tumour to form a new vasculature [84,85].

Unlimited replicative potential; avoiding senescence

If a cell acquires mutations that allow it to supply its own growth signals and ignore anti-proliferative

and pro-apoptotic signals, then in theory, that cell should be capable of unlimited growth and division [40]. In reality, somatic cells have a limited number of divisions 'programmed' into them, with the absolute number of divisions varying with tissue and cell type [87]. Cultured cells will proliferate for a number of generations and then enter senescence. This state can be overcome by various mutations, including those in *p53* and *Rb*, leading to further growth and division until cells enter a 'crisis' phase. This is characterized by karyotypic abnormalities such as chromosome fusions and a high proportion of cell death [88]. In a population of such cells, clones will emerge which are immortalized and capable of replicating indefinitely. The mechanisms surrounding immortalization appear to be largely dependent on maintaining telomeres [89]. Telomeres stabilize chromosomes by preventing recombination and fusion events, and also by preventing the cell from recognizing the end of the chromosome as a double-strand break [90]. With every cell division, the ends of chromosomes progressively 'fray off', counting off the cell generations until they reach a certain non-permissive limiting length which triggers senescence or apoptosis [91–93]. In order to overcome this, a cell may either acquire activating mutations in the telomerase gene, which acts to maintain telomeres and is strongly repressed in normal cells, or activate a mechanism that maintains telomeres through recombination [89].

Mice have very long telomeres (40–60 kb as opposed to 10 kb in human) and also have a wider pattern of expression of telomerase than in the human [94]. For a human cell to escape replicative senescence, it must acquire mutations in the telomerase maintenance system, but this restraint apparently does not apply to mouse cells [94]. Mice engineered to have shorter telomeres have a higher rate of spontaneous tumourigenesis [94]. One possible reason why mice and humans have different tumour spectra was hinted at when these short-telomered mice were crossed onto a late-generational *p53* null allele. The resultant mice developed a high rate of epithelial-derived carcinomas, similar to the tumour spectrum in humans [95].

Mice have also been generated which provide insights into the relationship between telomeres, premature ageing, and cancer [96]. Chang *et al* generated mice null for both the telomerase RNA component *Terc*, and *Wrn*. Inactivation of WRN (a *RecQ* helicase family member) causes Werner syndrome, which presents with premature ageing, genomic instability, and an increased incidence of cancer [97,98]. Cells with *WRN* deficiency have an increased rate of telomere loss and undergo premature senescence, which can be rescued by forced expression of Tert [99,100]. Doubly null $Wrn^{-/-}$ $Terc^{-/-}$ mice experience premature ageing, cataracts, and hair loss, as well as increased chromosomal instability and incidence of non-epithelial cancers such as osteosarcomas and soft tissue sarcomas [96]. Interestingly, the effects of the *Wrn* deficiency on *Terc* null mice increased through subsequent generations, with the first and second generations of $Terc^{-/-}$ mice being unaffected by *Wrn* status. In contrast, by the fourth to sixth generations, $Terc^{-/-}$ $Wrn^{-/-}$ mice had lower body weights and a shorter life expectancy than $Terc^{-/-}$ $Wrn^{+/-}$ mice [96]. Later generations of $Terc^{-/-}$ $Wrn^{-/-}$ mice also showed age-related degeneration in mesenchymal tissues, unlike $Terc^{-/-}$ mice, which show phenotypes in the highly proliferative compartments of skin, blood, and intestine. The lower impact of *Wrn* deficiency in epithelial tissues could be due to higher telomerase activity or functional redundancy in the *RecQ* helicase family such as the Bloom helicase.

The nature of the checkpoints which monitor telomere length are still not understood in detail, and no doubt new mouse models such as the *Wrn/Terc* null mouse will be vital in unravelling the mechanisms by which cells acquire unlimited replicative capacity. Furthermore, despite the clear importance of telomeres, they are not solely responsible for inducing senescence, as evidenced by the fact that some mouse cell cultures can undergo senescence with telomeres of 50 kb or more remaining [101].

One gene recognized to induce cellular senescence is p53 [64]. In this light, it is intriguing that mice with a deletion of the first six exons of *p53*, which encodes a truncated RNA forming a carboxy-terminal p53 fragment, have been shown to have an altered lifespan [102]. This mutation confers some properties of activated p53, and the mice display an early ageing phenotype including reduced lifespan, osteoporosis, weight loss, lordokyphosis (hunched spine), and muscle loss. Notably, none of the mice developed tumours, compared with around 45% of wild types. These data argue for a role for p53 in regulating senescence *in vivo* and also suggest that there may be a 'pay-off' between increased lifespan and tumour susceptibility.

Invasion and metastasis

In order to metastasize, a tumour must break down surrounding tissue such as a basement membrane to escape its tissue of origin; it must then travel within the body and re-establish itself at a new site, supplying its own growth factors [40]. It must be resistant to anoikis (apoptosis induced by lack of correct positional information) and establish a new vasculature at the distant site. Metastases are the cause of an estimated 90% of cancer deaths [40] and thus the study of metastatic progression is of vital importance.

Several types of adhesion molecule are altered in many cancer cells, such as the calcium-dependent cadherins, which mediate cell–cell interaction, and the integrins, which mediate cell–extracellular matrix (ECM) interactions [103].E-cadherin mediates cell–cell adhesions in epithelial tissues; the intracellular domain interacts with β-catenin to transduce anti-proliferative signals such as those relayed via the

Lef/Tcf system. Loss of E-cadherin expression is associated with de-differentiation and invasion in a variety of human cancers [104,105] — the forced expression of E-cadherin in a transgenic model of pancreatic β-cell carcinogenesis has already been shown to prevent the transition from adenoma to carcinoma, underlining the importance of correct adhesion and positional signals in invasion and metastasis [106].

Integrins consist of a non-covalently linked α and β subunit (there are 18 known α and eight known β units) which spans the membrane and transduces information about the cell's environment into the cell [107]. Different combinations of α and β subunits have different affinities for ECM components. The expression profiles of some integrins are known to change during tumourigenesis, allowing the cell to alter its migration and adhesion; for example, $\alpha 6\beta 4$ and $\alpha 3\beta 5$ are known to be up-regulated during tumourigenesis, whereas reduced expression of $\alpha 1$, $\alpha 6$, $\beta 1$ or $\beta 4$ is associated with breast epithelial neoplasms [107]. The $\alpha 3\beta 5$ integrin binds a wide range of ECM components and is considered an attractive target for therapeutic intervention, as it is expressed at low levels in resting endothelial cells and at higher levels on vascular cells in tumours [107,108].

Integrins are also thought to play a key role in directing tissue remodelling via their ability to activate MMP precursors. $\alpha 3\beta 5$ interacts with MMP2, a key player in tissue remodelling. MMP2 knockout mice show reduced angiogenesis and tumourigenesis [109] and inhibitors of the MMP2α–$\alpha 3\beta 5$ interaction are potent suppressors of angiogenesis, gliomas, and melanomas [110,111]. However, knockout mice lacking $\beta 3$ or $\beta 3$ and $\beta 5$ subunits, generated by Reynolds *et al*, show enhanced tumour growth and enhanced angiogenesis within the tumours that form [108]. This shows that the $\alpha 3\beta 5$ and $\alpha v\beta 5$ integrins, which were thought to be vital for angiogenesis, are in fact not essential. The apparent discrepancy between the effects of antagonists and knockouts may be due to compensation, or the antagonists being non-specific [108]. However, it seems more likely that we simply do not yet have a complete enough understanding of the complex interactions between cell and stroma that are vital in driving invasion and metastasis.

Genomic instability, a helping hand on the road to cancer

Although not one of the six hallmarks identified by Hanahan and Weinberg, genomic instability is increasingly being identified as a major contributor to many cancer types. Genomic instability is seen in many cancer cells, either as an initiating event or as a later stage. It is estimated that up to 30% of genes in the genome code for proteins that regulate genomic fidelity [112] and many inherited cancer syndromes are due to mutations in these genes [eg Li–Fraumeni (*p53*, *Chk2* [113]; OMIM #151 623), Bloom syndrome (*BLM*; OMIM #210 900), Nijmegen breakage syndrome (*NBS1*; OMIM #251 260), ataxia telangiectasia (*ATR/ATM*; OMIM #208 900), and HNPCC (OMIM #114 500; the MMR system)]. Genomic instability can have many causes: hypomethylation, inefficient MMR (mismatch repair), increased mitotic recombination, chromosomal translocations or defects in genes which code for proteins that monitor and repair lesions, or monitor genomic fidelity through the cell cycle [40].

Chromosomal translocations frequently give rise to fusion proteins which may participate in the cell cycle and drive clonal expansion of the tumour. For example, translocations at human 11q23 can fuse the 5′ end of the MLL (myeloid lymphoid leukaemia) gene to a range of potential genes, giving rise to a diverse number of leukaemia aetiologies [114]. Forster *et al* have developed a highly elegant model which accurately mimics the translocation that fuses the MLL and ENL genes to give rise to myeloid leukaemia and mixed myeloid/lymphoid leukaemia [114]. Forster *et al* placed Lox P sites at defined breakpoints on separate chromosomes. Cre recombinase was then expressed under the control of the haematopoietic promoter *Lmo2*, leading to inter-chromosomal reciprocal translocations and the rapid onset of myeloid tumours with high penetrance. This accurately recapitulates the randomly occurring spontaneous translocations and clonal expansion of the human leukaemia and is the first model to generate both translocation products.

Genomic instability can also be caused by hypomethylation, as recently shown by Gaudet *et al* [115]. The Dnmt1 protein maintains methylation in somatic cells and is required for embryonic development — Dnmt1 null mice die during gestation [116,117], so Gaudet *et al* created a heterozygous Dnmt1 model with one Dnmt1 null allele and one Dnmt1 hypomorphic allele ($Dnmt1^{chip/-}$) These mice are estimated to have only 10% Dnmt1 activity and have substantially reduced levels of genomic methylation. The mice developed aggressive T-cell lymphomas at around 4[] 8 months of age. Array CGH (comparative genome hybridization) showed an increased rate of chromosome gains, most notably gain of chromosome 15 and the *c-Myc* gene, which it carries. This shows that methylation levels can directly contribute to tumourigenesis via genomic instability.

The role of methylation in tumourigenesis is, however, complex. Thus, loss of DNMT1 can also be shown to have a tumour protective effect within the intestine, presumably as a consequence of blocking transcriptional repression of TSGs [118]. Similarly, Sansom *et al* [119] used a mouse model (Apc^{Min}) of the human intestinal cancer syndrome FAP (familial adenomatous polyposis coli) to show that constitutive loss of the *Mbd2* gene (a member of the methyl binding domain family) in the intestine results in greatly reduced tumour burden. Clearly, the role of methylation patterns in cancer development is complex, and

mouse models of altered epigenetic states will no doubt provide insights into these divergent mechanisms.

Large-scale mutagenesis screens: identifying novel mutations

The few thousand mutant mouse strains in existence represent only about 10% of the total genes of the mouse genome [120]. Although we have almost complete sequence information available for the mouse, we still do not know the biological functions of most of these genes [120]. Functional annotation of the mouse genome will necessarily be slower than sequencing, given the volume of work needed to characterize mutant phenotypes. However, several concerted programmes are now underway [121□123] to systematically mutate each gene in the mouse genome using high throughput approaches such as ENU (ethylnitrosourea) mutation. γENU is a potent alkylating agent which causes point mutations at high frequency in many tissues [120]. When male mice are injected with ENU, point mutations occur in the pre-meiotic spermatogonial stem cells [120]. The F1 progeny can then be screened for dominant phenotypic abnormalities (behavioural or biochemical) or bred for a further two generations to expose recessive phenotypes. The ENU approach has a number of advantages: point mutations can often produce both null alleles and hypomorphic alleles, which are more informative than null alleles alone; novel functions in both known and unknown genes can be exposed; and the whole approach is 'phenotype-driven', removing any bias towards known or expected gene functions. A large-scale ENU mutagenesis programme (especially one screening for dominant and recessive mutations) is, however, a massive logistical undertaking, requiring co-ordination of animal resources, personnel, and informatics. Several such projects are now underway and will no doubt provide a vast number of novel mutant lines, many of which will aid in the study of tumourigenesis.

Conclusions and future directions

The current generation of mouse models is extremely sophisticated, yielding insights into the fundamental processes underlying normal cell physiology and cancer. Future mouse models will no doubt provide more precise control over gene expression at various stages of development. As we understand more about the expression profiles of genes through embryogenesis, it is also becoming possible to activate mutant alleles at precise points in the developmental programme, for example through the use of inducible Cre-Lox technology [124,125]. Inevitably, some problems remain with these technologies. For example, certain Cre-expressing lines are 'leaky' and drive recombination in tissues other than the target tissue. This issue will need to be carefully addressed, although it could itself be used as a useful tool. Ultimately, it is possible to envisage the development of a modular 'toolbox' containing a range of defined transgenes and systems which will permit controlled gene expression of many different alleles of a given gene throughout all stages of the life cycle of an organism. With novel reporter systems and imaging technologies [126], it should also be possible to follow the initiation and spread of tumours *in vivo* and to integrate all these data with a bio-informatic database. Indeed, the integration of such informatics will be key in allowing the full potential of the mouse (and human) genomes to be exploited. Finally, it should be stressed that the primary purpose of these systems is the modelling of human disease, and by implication the resolution of human disease. The need to develop cohesive, productive links between the basic murine studies described here and translational and clinical research cannot be overstated.

References

1. Rangarajan A, Weinberg RA. Comparative biology of mouse versus human cells: modelling human cancer in mice. *Nature Rev Cancer* 2003; **3**: 952□959.
2. DePinho RE. The age of cancer. *Nature* 2000; **408**: 248□254.
3. Blasco MA, Lee HW, Hande MP, *et al.* Telomere shortening and tumour formation by mouse cells lacking telomerase RNA. *Cell* 1997; **91**: 25□34.
4. Adams JM, Harris AW, Pinkert CA, *et al.* The c-Myc oncogene driven by immunoglobulin enhancers induces lymphoid malignancy in transgenic mice. *Nature* 1985; **318**: 533□538.
5. McCleod KF, Jacks T. Insights into cancer from transgenic mouse models. *J Pathol* 1999; **187**: 43□60.
6. Dyson N. The regulation of E2F by pRB-family proteins. *Genes Dev* 1998; **12**: 2245□2262.
7. Jacks T, Fazeli A, Schmitt EM, Bronson RT, Goodell MA, Weinberg RA. Effects of an Rb mutation on the mouse. *Nature* 1992; **359**: 259□300.
8. Harrison DJ, Hooper ML, Armstrong JF, Clarke AR. Effects of heterozygosity for the *Rb-1t19 neo* allele in the mouse. *Oncogene* 1995; **10**: 1615□1620.
9. Dyer MA, Cepko CL. Regulating proliferation during retinal development. *Nature Rev Neurosci* 2001; **2**: 333□342.
10. Dyer MA, Cepko CL. $p27^{Kip1}$ and $p57^{Kip2}$ regulate proliferation in distinct retinal progenitor cell populations. *J Neurosci* 2001; **21**: 4259□4271.
11. Robanus-Maandag E, Dekker M, Van Der Walk M, *et al.* p107 is a suppressor of retinoblastoma development in pRb-deficient mice. *Genes Dev* 1998; **12**: 1599□1609.
12. Clarke AR, Maandberg ER, Von Roon M, *et al.* Requirement for a functional Rb1 gene in murine development. *Nature* 1992; **359**: 328□330.
13. Lee EY, Chang CY, Hu N, *et al.* Mice deficient for Rb are nonviable and show defects in neurogenesis and haematopoiesis. *Nature* 1992; **359**: 228□294.
14. Palmiter RD, Brinster RL. Germ-line transformation of mice. *Annu Rev Genet* 1986; **20**: 465□499.
15. Dorer DR. Do transgene arrays form heterochromatin in vertebrates? *Transgenic Res* 1997; **6**: 3□10.
16. Politi K, Kljuic A, Szabolcs M, Fisher P, Ludwig T, Efstratiadis A. Designer tumours in mice. *Oncogene* 2004; **23**: 1558□1565.
17. Hanhan D. Heritable formation of pancreatic B-cell tumours in transgenic mice expressing recombinant insulin/simian virus 40. *Nature* 1985; **315**: 115□122.

18. Shibata H, Toyama K, Shioya H, *et al.* Rapid colorectal adenoma formation initiated by conditional targeting of the Apc gene. *Science* 1997; **278**: 120–123.
19. Jonkers J, Berns A. Conditional mouse models of cancer. *Nature Rev Cancer* 2002; **2**: 251–265.
20. Baranda CS, Dymecki SM. Talking about a revolution : the impact of site-specific recombinases on genetic analyses in mice. *Dev Cell* 2004; **6**: 7–28.
21. Stenberg N, Hamilton D. Bacteriophage P1 site-specific recombination. I. Recombination between LoxP sites. *J Mol Biol* 1981; **150**: 467–486.
22. Kuhn R, Torres RM. Cre/loxP recombination system and gene targeting. *Methods Mol Biol* 2002; **180**: 175–204.
23. Hoess R, Ziese M, Stenberg N. P1 site-specific recombination : nucleotide sequence of the recombining sites. *PNAS* 1982; **79**: 3398–3402.
24. O'Gorman S, Fox DT, Wahl GM. Recombinase mediated gene activation and site-specific integration in mammalian cells. *Science* 1991; **251**: 1351–1355.
25. Schmidt EE, Taylor DS, Prigge JR, Barnett S, Capecchi M. Illegitimate Cre-dependent chromosome re-arrangements in transgenic mouse spermatids. *PNAS* 2000; **97**: 13 702–13 707.
26. Tyagarajan B, Guimaraes MJ, Calos MP. Mammalian genomes contain active recombinase recognition sites. *Gene* 2000; **244**: 47–54.
27. Silver DP, Livingston DM. Self-excising retroviral vectors encoding the Cre recombinase overcome Cre-mediated cellular toxicity. *Mol Cell* 2001; **8**: 233–243.
28. Metzger D, Chambon P. Site- and time-specific gene targeting in the mouse. *Mol Methods* 2001; **24**: 71–80.
29. Prost S, Sheahan S, Rannie D, Harrison DJ. Adenovirus-mediated Cre deletion of floxed sequences in primary mouse cells is an efficient alternative for studies of gene deletion. *Nucleic Acids Res* 2001; **29**: E80.
30. Wagner KU, Wall RJ, St-Onge L, *et al.* Cre-mediated gene deletion in the mammary gland. *Nucleic Acids Res* 1997; **25**: 4323–4330.
31. Selbert S, Bentley DJ, Melton DW, *et al.* Efficient BLG-Cre mediated gene deletion in the mammary gland. *Transgenic Res* 1998; **5**: 387–396.
32. Ireland H, Kemp R, Houghton C, *et al.* Inducible Cre-mediated control of gene expression in the murine gastrointestinal tract: effect of loss of β-catenin. *Gastroenterology* 2004; **126**: 1236–1246.
33. Gossen M, Bujard H. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *PNAS* 1992; **89**: 5547–5551.
34. Baron U, Bujard H. Tet-repressor-based system for regulated gene expression in mammalian cells: principles and advances. *Methods Enzymol* 2000; **327**: 401–421.
35. Zhou Z, Zheng T, Lee CG, Homer RJ, Elias J. Tetracycline-controlled transcriptional regulation systems: advances and applications in transgenic animal modelling. *Cell Dev Biol* 2002; **13**: 121–128.
36. No D, Yao TP, Evans RM. Ecdysone-inducible gene expression in mammalian cells and transgenic mice. *PNAS* 1996; **93**: 3346–3351.
37. Ngan ES, Schillinger K, DeMayo F, Tsai SY. The mifepristone-inducible gene regulatory system in mouse models of disease and therapy. *Semin Cell Dev Biol* 2002; **13**: 143–149.
38. Cronin CC, Gluba W, Scrable H. The Lac operator-repressor system is functional in the mouse. *Genes Dev* 2001; **15**: 1506–1517.
39. Wang XJ, Liefer KM, Tsai S, O'Malley BW, Roop DR. Development of gene-switch transgenic mice that inducibly express transforming growth factor β in the epidermis. *PNAS* 1999; **96**: 8483–8488.
40. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 2000; **100**: 57–70.
41. Park MT, Lee SJJ. Cell cycle and cancer. *Biochem Mol Biol* 2003; **36**: 60–65.
42. Lesli NR, Downes P. PTEN function: how normal cells control it and tumour cells lose it. *Biochem J* 2004; **382**: 1–11.
43. Wu X, Huang J, Powell WC, *et al.* Generation of a prostate epithelial cell-specific Cre transgenic mouse model for tissue-specific gene ablation. *Mech Dev* 2001; **101**: 61–69.
44. Lesche R, Groszner M, Gao J, *et al.* Cre/LoxP-mediated inactivation of the murine Pten tumour suppressor gene. *Genesis* 2002; **32**: 148–149.
45. You MJ, Fukabori Y, McBride G, *et al.* Genetic analysis of Pten and INK4a/Arf in the suppression of tumourigenesis in mice. *PNAS* 2002; **99**: 1455–1460.
46. Rane SG, Dubus P, Mettus RV, *et al.* Loss of Cdk4 expression causes insulin-deficient diabetes and Cdk4 activation results in B-cell hyperplasia. *Nature Genet* 1999; **22**: 44–52.
47. Tsutsui T, Hesabi B, Moons DS, Pandolfi PP, Hansel KS, Koff A, *et al.* Targeted disruption of Cdk4 delays cell cycle entry with enhanced P27Kip1 activity. *Mol Cell Biol* 1999; **19**: 7011–7019.
48. Malumbres M, Sotillo R, Santamaria D, *et al.* Mammalian cells cycle without the D-type cyclin-dependent kinases Cdk4 and Cdk6. *Cell* 2004; **118**: 493–504.
49. Marinkovic D, Marinkovic T, Mahr B, *et al.* Reversible lymphomagenesis in conditionally c-MYC expressing mice. *Int J Cancer* 2004; **110**: 336–342.
50. Pelengaris S, Littlewood T, Khan M, Elia G, Evan G. Reversible activation of c-Myc in skin: induction of a complex neoplastic phenotype by a single oncogenic lesion. *Mol Cell* 1999; **3**: 565–577.
51. Jain M, Arvanitis C, Chu K, *et al.* Sustained loss of a neoplastic phenotype by brief inactivation of MYC. *Science* 2002; **297**: 102–104.
52. Malumbres M, Barbacid M. RAS oncogenes : the first 30 years. *Nature Rev Cancer* 2002; **3**: 7–13.
53. Koera K, Nakamura K, Nakao K, *et al.* K-ras is essential for the development of the mouse embryo. *Oncogene* 1997; **15**: 1151–1159.
54. Johnson L, Greenbaum D, Cichowski K, *et al.* K-ras is an essential gene in the mouse with partial functional overlap with N-ras. *Genes Dev* 1997; **11**: 2468–2481.
55. Umanoff H, Edelmann W, Pellicer A, *et al.* The murine N-Ras gene is not essential for growth and development. *PNAS* 1995; **92**: 1709–1713.
56. Esteban LM, Vicario-Abejon C, Fernandez-Salguero P, *et al.* Targeted genomic disruption of H-ras and N-ras, individually or in combination, reveals the dispensability of both loci for mouse growth and development. *Mol Cell Biol* 2001; **21**: 1444–1452.
57. Johnson L, Mercer K, Greenbaum D, *et al.* Somatic activation of the K-ras oncogene causes early-onset lung cancer in mice. *Nature* 2001; **410**: 1111–1116.
58. Guerra C, Mijimolle N, Dhawahir A, Dubus P, Barrada M, Serrano M, *et al.* Tumour induction by an endogenous K-ras oncogene is highly dependent on cellular context. *Cancer Cell* 2003; **4**: 111–120.
59. Jackson EL, Willis N, Mercer K, *et al.* Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. *Genes Dev* 2001; **15**: 3243–3248.
60. Meuwissen R, Linn SC, van der Valk M, Mooi WJ, Berns A. A mouse model for lung tumorigenesis through Cre/lox controlled sporadic activation of the K-Ras oncogene. *Oncogene* 2001; **20**: 6551–6558.
61. Serrano M, Blasco MA. Putting the stress on senescence. *Curr Opin Cell Biol* 2001; **13**(6): 748–753.
62. Wong AK, Chin L. An inducible melanoma model implicates a role for RAS in tumor maintenance and angiogenesis. *Cancer Metastasis Rev* 2000; **19**: 121–129.
63. Zhang J, Schweers B, Dyer MA. The first knockout mouse model of retinoblastoma. *Cell Cycle* 2004; **3**: 952–959.
64. Hofseth LJ, Hussain SP, Harris CC. p53: 25 years after its discovery. *Trends Pharmacol Sci* 2004; **25**: 177–181.
65. Sansom OJ, Clarke AR. P53 null mice: damaging the hypothesis? *Mutation Res* 2000; **452**(2): 149–162.

66. Lowe SW, Schmitt EM, Smith SW, *et al.* p53 is required for radiation-induced apoptosis in mouse thymocytes. *Nature* 1993; **362**: 847–849.
67. Donehower LA, Harvey M, Slagle BL, *et al.* Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. *Nature* 1992; **356**: 215–221.
68. Purdie CA, Harrison DJ, Peter A, *et al.* Tumour incidence, spectrum and ploidy in mice with a large deletion in the *p53* gene. *Oncogene* 1994; **9**: 611–619.
69. Clarke AR, Hollstein M. Mouse models with modified p53 sequences to study cancer and ageing. *Cell Death Differ* 2003; **10**: 443–450.
70. Evans SC, Lozano G. The Li–Fraumeni syndrome: an inherited susceptibility to cancer. *Mol Med Today* 1997; **3**: 390–395.
71. Donehower L, Harvey M, Slagle BL, *et al.* Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. *Nature* 1992; **356**: 215–221.
72. Lozano G, Zambetti GP. What have animal models taught us about the p53 pathway? *J Pathol* 2005; **205**: 206–220.
73. Hergenhahn M, Luo JL, Hollstein M. p53 designer genes for the modern mouse. *Cell Cycle* 2004; **3**: 738–741.
74. Liu Z, Hergenhahn M, Schmeiser HH, Wogan GN, Hong A, Hollstein M. Human tumour p53 mutations are selected for in mouse embryonic fibroblasts harbouring a humanized p53 gene. *PNAS* 2004; **101**: 2963–2968.
75. Liu G, Parant JM, Lang G, *et al.* Chromosome stability, in the absence of apoptosis, is critical for suppression of tumourigenesis in Trp53 mutant mice. *Nature Genet* 2004; **36**: 63–68.
76. Venkatachalam S, Shi YP, Jones SN, *et al.* Retention of wild-type p53 in tumors from p53 heterozygous mice: reduction of p53 dosage can promote cancer formation. *EMBO J* 1998; **17**: 4657–4667.
77. Jiminez GS, Mister M, Stommel JM, *et al.* A Transactivation-deficient Trp53 mouse model provides insights into Trp53 regulation and function. *Nature Genet* 2000; **26**: 37–47.
78. Zabkiewicz J, Clarke AR. DNA damage induced apoptosis: insights from the mouse. *BBA Rev Cancer* (in press).
79. Moore L, Venkatchalam S, Vogel H, *et al.* Cooperativity of p19ARF, Mdm2 and p53 in murine tumourigenesis. *Oncogene* 2003; **23**: 7831–7837.
80. Kamijo T, Weber JD, Zambetti G, Zindy F, Roussel MF, Sherr CJ. Functional and physical interactions of the ARF tumor suppressor with p53 and Mdm2. *Proc Natl Acad Sci U S A* 1998; **95**: 8292–8297.
81. Zhang Y, Xiong Y, Yarbrough WG. ARF promotes MDM2 degradation and stabilizes p53: ARF–INK4a locus deletion impairs both the Rb and p53 tumor suppression pathways. *Cell* 1998; **92**: 725–734.
82. Palmero I, Pantoja C, Serrano M. p19ARF links the tumour suppressor p53 to Ras. *Nature* 1998; **395**: 125–126.
83. Zindy F, Eischen CM, Randle DH, *et al.* E1A signaling to p53 involves the p19(ARF) tumor suppressor. *Genes Dev* 1998; **12**: 2434–2442.
84. Bergers G, Benjamin LE. Tumourigenesis and the angiogenic switch. *Nature Rev Cancer* 2003; **3**: 401–410.
85. Volpert OV, Alani RM. Wiring the angiogenic switch: Ras, Myc and thrombospondin-1. *Cancer Cell* 2003; **3**: 199–200.
86. Lawler J. Thrombospondin-1 as an endogenous inhibitor of angiogenesis and tumor growth. *Cell Mol Med* 2002; **6**: 1–12.
87. Hayflick L. The limited *in vitro* lifetime of human diploid cell strains. *Exp Cell Res* 1965; **37**: 614–636.
88. Wright WE, Pereira-Smith OM, Shay JW. Reversible cellular senescence: implications for immortalisation of normal human diploid fibroblasts. *Mol Cell Biol* 1989; **9**: 3088–3092.
89. Meeker AK, Hicks JL, Iacobuzio-Donahue CA, *et al.* Telomere length abnormalities occur early in the initiation of epithelial carcinogenesis. *Cancer Res* 2004; **10**: 3317–3326.
90. Blackburn EH. Structure and function of telomeres. *Nature* 1991; **350**: 569–572.
91. Shay JW, Wright WE, Werbin H. Defining the molecular mechanisms of human cell immortalisation. *Biochim Biophys Acta* 1991; **1072**: 1–7.
92. Harley CB, Kim NW, Prowse KR, Weinrich SL, Hirsch KS, West MD, *et al.* Telomerase, cell immortality and cancer. *Cold Spring Harbor Symp Quant Biol* 1994; **59**: 307–315.
93. Wright WE, Shay JW. Telomere dynamics in cancer progression and prevention: fundamental differences in human and mouse telomere biology. *Nature Med* 2000; **6**: 849–851.
94. Rudolph KL, Chang S, Lee H-W, Blasco M, Gottlieb GJ, Greider C, *et al.* Longevity, stress-response and cancer in aging telomerase-deficient mice. *Cell* 1999; **96**: 701–712.
95. Artandi SE, Chang S, Lee S-L, *et al.* Telomere dysfunction promotes non-reciprocal translocations and epithelial cancers in mice. *Nature* 2000; **406**: 641–645.
96. Chang S, Multani AS, Cabrera NG, *et al.* Essential role of limiting telomeres in the pathogenesis of Werner syndrome. *Nature Genet* 2004; **36**: 877–882.
97. Martin GM, Oshima J. Lessons from human progeroid syndromes. *Nature* 2000; **408**: 263–266.
98. Hickson ID. RecQ helicases: caretakers of the genome. *Nature Rev Cancer* 2003; **3**: 169–178.
99. Wyllie FS, Jones CJ, Skinner JW, *et al.* Telomerase prevents the accelerated cell ageing of Werner syndrome fibroblasts. *Nature Genet* 2000; **24**: 16–17.
100. Schulz VP, Zakian VA, Ogburn CE, *et al.* Accelerated loss of telomeric repeats may not explain accelerated replicative decline of Werner syndrome cells. *Hum Genet* 1996; **97**(6): 750–754.
101. Kipling D. Telomere structure and telomerase expression during mouse development and tumorigenesis. *Eur J Cancer* 1997; **33**: 792–800.
102. Tyner SD, Venkatchalam A, Choi J, *et al.* p53 mutant mice that display early ageing associated phenotypes. *Nature* 2002; **415**: 45–53.
103. Aplin AE, Howe A, Alahari SK, Juliano RL. Signal transduction and signal modulation by cell adhesion receptors: the role of integrins, cadherins, immunoglobulin-cell adhesion molecules and selectins. *Pharmacol Rev* 1998; **50**: 197–263.
104. Behrens J, Frixen U, Schipper J, Weidner M, Birchmeiser W, *et al.* Cell adhesion in invasion and metastasis. *Semin Cell Biol* 1992; **3**: 169–178.
105. Takeichi M. Cadherins in cancer: implications for invasion and metastasis. *Curr Opin Cell Biol* 1993; **5**: 806–811.
106. Perl AK, Wilgenbus P, Dahl U. A causal role for E-cadherin in the transition from adenoma to carcinoma. *Nature* 1998; **392**: 190–193.
107. Hood JD, Cheresh D. Role of integrins in cell invasion and migration. *Nature Rev Cancer* 2002; **2**(2): 91–100.
108. Reynolds LE, Wyder L, Lively JC, *et al.* Enhanced pathological angiogenesis in mice lacking β_3 integrin or β_3 and β_5 integrins. *Nature Med* 2002; **8**(1): 27–34.
109. Itoh T, Tanioka M, Yoshida T, Nishimoto H, Itohara S. Reduced angiogenesis and tumor progression in gelatinase A-deficient mice. *Cancer Res* 1998; **58**: 1048–1051.
110. Siletti S, Kessler T, Goldberg J, Boger DL, Cheresh DA. Disruption of matrix metalloproteinase 2 binding to integrin $\alpha v\beta 3$ by an organic molecule inhibits angiogenesis and tumour growth *in vivo*. *PNAS* 2001; **98**: 119–124.
111. Bello L, Lucini V, Carrabba G, *et al.* Simultaneous inhibition of glioma angiogenesis, cell proliferation and invasion by a naturally-occurring fragment of human metalloproteinase-2. *Cancer Res* 2001; **61**: 8730–8736.
112. Grady WM. Genomic instability and colon cancer. *Cancer Metastasis Rev* 2004; **23**: 11–27.
113. Varley J. TP53, hChk2, and the Li–Fraumeni syndrome. *Methods Mol Biol* 2003; **222**: 117–129.
114. Forster A, Pannell R, Drynan LF, *et al.* Engineering *de novo* reciprocal chromosomal translocations associated with Mll to replicate primary events of human cancer. *Cancer Cell* 2003; **3**: 449–458.
115. Gaudet F, Hodgson JG, Eden A, *et al.* Induction of tumors in mice by genomic hypomethylation. *Science* 2003; **300**: 489–492.
116. Lei H, Oh SP, Okano M, *et al.* *De novo* DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development* 1996; **122**: 3195–3205.

117. Li E, Bestor TH, Jeansich R. DNA methylation, genomic imprinting, and mammalian development. *Cold Spring Harbor Symp Quant Biol* 1993; **58**: 297□ 305.
118. Eads CA, Nickel AE, Laird PW. Complete genetic suppression of polyp formation and reduction of CpG-island hypermethylation in $Apc^{(Min/+)}$ Dnmt1-hypomorphic mice. *Cancer Res* 2002; **62**: 1296□ 1299.
119. Sansom OJ, Berger J, Bishop SM, Hendrich B, Bird A, Clarke AR. Deficiency of Mbd2 suppresses intestinal tumorigenesis. *Nature Genet* 2003; **34**: 145□ 147.
120. Brown SD, Hardisty RE. Mutagenesis strategies for identifying novel loci associated with disease phenotypes. *Semin Cell Dev Biol* 2003; **14**: 19□ 24.
121. Beier DR. ENU mutagenesis : a work in progress. *Physiol Genom* 2002; **11**: 111□ 113.
122. Hrabe de Angelis MH, Flaswinkel H, Fuchs H, *et al.* Genome-wide, large-scale production of mutant mice by ENU mutagenesis. *Nature Genet* 2000; **25**: 444□ 447.
123. www.gsf.de/ieg/groups/enu-mouse.html 'The ENLI-mouse mutagenesis screen project', 25 September 2004.
124. Katoh M, Horiya N, Valdivia RP. Mutations induced in male germ cells after treatment of transgenic mice with ethylnitrosourea. *Mutat Res* 1997; **388**: 229□ 237.
125. Sansom OJ, Reed K, Hayes AJ, *et al.* Loss of Apc immediately perturbs differentiation and migration in the small intestine. *Genes Dev* 2004; **18**: 1385□ 1390.
126. Lyons SK. Advances in imaging mouse tumour models *in vivo*. *J Pathol* 2005; **205**: 194□ 205.

Fas-associated death domain protein interacts with methyl-CpG binding domain protein 4: A potential link between genome surveillance and apoptosis

Robert A. Screaton*, Stephan Kiessling*, Owen J. Sansom†, Catherine B. Millar‡, Kathryn Maddison†, Adrian Bird‡, Alan R. Clarke†, and Steven M. Frisch*§

*The Burnham Institute, 10901 North Torrey Pines Road, La Jolla, CA 92037; †Cardiff School of Biosciences, Cardiff University, P.O. Box 911, Cardiff CF10 3US, United Kingdom; and ‡Wellcome Trust Centre for Cell Biology, The King's Buildings, Edinburgh University, Edinburgh EH9 3JR, United Kingdom

Communicated by Erkki Ruoslahti, The Burnham Institute, La Jolla, CA, February 28, 2003 (received for review January 22, 2003)

Fas-associated death domain protein (FADD) is an adaptor protein bridging death receptors with initiator caspases. Thus, its function and localization are assumed to be cytoplasmic, although the localization of endogenous FADD has not been reported. Surprisingly, the data presented here demonstrate that FADD is mainly nuclear in several adherent cell lines. Its accumulation in the nucleus and export to the cytoplasm required the phosphorylation site Ser-194, which was also required for its interaction with the nucleocytoplasmic shuttling protein exportin-5. Within the nucleus, FADD interacted with the methyl-CpG binding domain protein 4 (MBD4), which excises thymine from GT mismatches in methylated regions of chromatin. The MBD4-interacting mismatch repair factor MLH1 was also found in a complex with FADD. The FADD–MBD4 interaction involved the death effector domain of FADD and a region of MBD4 adjacent to the glycosylase domain. The FADD-binding region of MBD4 was downstream of a frameshift mutation that occurs in a significant fraction of human colorectal carcinomas. Consistent with the idea that MBD4 can signal to an apoptotic effector, MBD4 regulated DNA damage-, Fas ligand-, and cell detachment-induced apoptosis. The nuclear localization of FADD and its interaction with a genome surveillance/DNA repair protein that can regulate apoptosis suggests a novel function of FADD distinct from direct participation in death receptor signaling complexes.

DNA damage-induced apoptosis signals originate in the nucleus. For example, double-strand DNA breaks activate the ATM/ATR (1) and c-abl-related kinases (2), leading to activation of the p53-family proteins (3, 4) and subsequent apoptosis. Other forms of DNA damage that are repaired by the mismatch repair (MMR) complex (reviewed in refs. 5 and 6) can promote apoptosis through the MMR component MLH1. Thus, MLH1 defects in tumor cells cause apoptosis resistance with regard to several DNA-damaging drugs (7–11). In fact, all five major types of DNA repair complexes are thought to signal to the apoptotic machinery (reviewed in ref. 12), although the mechanism of signaling is largely unknown.

Interestingly, several apoptosis-regulatory proteins are in the nucleus constitutively or conditionally. These include PEA-15 (13), DEDD 1-3 (14–16), DEDAF (17), p84N5 (18), caspases-2 (19) and -6 (20), Daxx (21), and (under conditions where nuclear export is inhibited) TRADD (22). Possible roles for these proteins in nuclear apoptotic signaling remain to be determined.

Recently, we have identified an additional link between a DNA repair protein and apoptosis (O.J.S. and A.R.C., unpublished data), namely the protein methyl-CpG binding domain protein 4 [MBD4 (23), also known as MED1 (24)]: MBD4 promoted the apoptotic response to DNA-damaging agents. The N-terminal conserved MBD of this protein targets it to bind methylated regions of DNA, whereas the C terminus is a uracil/thymine-*N*-glycosylase. The latter excises spontaneously deaminated cytosine (i.e., uracil) or methylcytosine (i.e., thymine) from G-T/U mismatches (25–27), the major individual source of point mutations in the human genome (28). MBD4 interacts with the mismatch repair/tumor suppressor protein MLH1 (24); both are mutator genes (29, 30). The inactivation of either gene causes increased apoptosis resistance with regard to DNA-damaging agents (refs. 7–11; O.J.S. and A.R.C., unpublished data), implying possible apoptotic signaling functions for both proteins. As with MLH1 (reviewed in ref. 30), MBD4 and also frequently mutated in certain human tumors (6, 31–34). The mechanistic link between MBD4, MLH1, and apoptosis will be important to elucidate.

Fas-associated death domain protein (FADD) is known mainly for its death receptor adaptor function at the cell surface (35–37). Thus, it is widely assumed that FADD is primarily or solely a cytoplasmic protein. However, there are no published data clearly supporting this assumption (see *Discussion*). Indeed, FADD has been implicated in potentially death receptor-independent apoptotic responses such as DNA damage and anoikis (38–42) that would not necessarily require FADD to be cytoplasmic. Moreover, cell-matrix adhesion can protect against Fas ligand (FasL)-induced apoptosis (R.A.S. and S.M.F., unpublished observations; refs. 43 and 44), suggesting that certain Fas-interacting proteins might be subject to relocalization by extracellular signals. We therefore examined the subcellular localization of FADD in the hope of revealing additional functions of the protein. Here, we report that FADD primarily localized to the nucleus, implying that the nuclear-cytoplasmic transport of FADD must be actively regulated and that FADD may have a novel nuclear function. We report that exportin-5 is a candidate nucleocytoplasmic transport protein for FADD and that a genome surveillance protein, MBD4, interacts directly with FADD and modulates apoptosis, suggesting a novel link between genome surveillance and apoptosis.

Materials and Methods

Cell Lines, Cell Culture, and Transfections. The normal human mammary epithelial cell line MCF10a was used for most experiments in this article. *MBD4−/−* mouse embryo fibroblasts (MEFs) have been described (29). Cells were transfected by using lipofection or infected with retroviruses based on MSCV-ires-zeo (45), and expression of transgenes was checked by Western blotting.

Apoptosis Assays (Cell Culture Experiments). Cells were maintained in low-attachment wells for the indicated times (for anoikis); some cell samples were exposed to recombinant FasL in the presence of enhancer antibody (for FasL). Cell lysates were

Abbreviations: MBD4, methyl-CpG binding domain protein 4; FADD, Fas-associated death domain protein; FasL, Fas ligand; YFP, yellow fluorescent protein; HA, hemagglutinin; NLS, nuclear localization signal; MEF, mouse embryo fibroblast; DED, death effector domain.

§To whom correspondence should be addressed. E-mail: sfrisch@burnham.org.

prepared and assayed fluorimetrically for Ac-DEVD-AFC cleavage activity.

Apoptosis Assays (*in Vivo* Experiments). *MBD4+/+* or *MBD4−/−* mice were injected with the Fas-agonistic antibody Jo-2. Histological sections of the indicated tissues were stained with hematoxylin/eosin and analyzed for apoptosis as described (46).

Antibodies. Most antibodies were obtained from commercial sources; additional FADD polyclonal antibodies were prepared against bacterially produced GST-FADD protein, and a sheep anti-MBD4 was prepared against the C-terminal 180 aa of recombinant MBD4 protein. A detailed characterization of antibodies is presented in Fig. 5, which is published as supporting information on the PNAS web site, www.pnas.org.

Nuclear-Cytoplasmic Fractionation. Nuclear-cytoplasmic fractionation was accomplished by homogenization of cells in hypotonic buffer followed by low-speed (800 × *g*) centrifugation.

Indirect Immunofluorescence. Indirect immunofluorescence was performed on formaldehyde-fixed, Triton X-100-permeabilized cells by using fluorescently tagged secondary antibodies.

Short-Interfering RNA. Short-interfering RNA was introduced by using lipofection of commercially synthesized RNA duplexes.

FADD Export Assay. WT or mutant forms of FADD were expressed with N-terminal yellow fluorescent protein (YFP) tags and C-terminal nuclear localization signals (NLSs) derived from simian virus 40 virus. After transfection, cells were detached for 15 min before formaldehyde fixation and cells were scored for nuclear vs. cytoplasmic localization on a fluorescence microscope.

Cell Surface Fas Coimmunoprecipitation/Western Blotting. Cell surface Fas coimmunoprecipitation/Western blotting was performed by incubating cell cultures with the Fas-agonistic antibody 2R2 (Roche Molecular Biochemicals) followed by cell lysis in Triton X-100-containing buffer, precipitation with protein A-Sepharose, and Western blot analysis.

Yeast Two-Hybrid Screening, Protein Interactions *in Vitro*, Western Blotting, and Immunoprecipitation. Yeast two-hybrid screening, protein interactions *in vitro*, Western blotting, and immunoprecipitation were performed by using standard methodology as described in *Supporting Methods*, which is published as supporting information on the PNAS web site.

See *Supporting Methods* for detailed materials and methods.

Results

FADD Localizes to the Nucleus in Several Cell Lines. In the course of examining the basis for protection against FasL-induced apoptosis by cell adhesion (R.A.S. and S.M.F., unpublished work), we observed that FADD was predominantly nuclear by using a mAb (Fig. 1). Nuclear FADD staining was blocked by preincubation of the antibody with recombinant antigen. We then generated two monospecific anti-human FADD polyclonal antibodies, which confirmed that FADD was primarily nuclear (Fig. 1). FADD was also primarily nuclear in several cell lines (HT1080 fibrosarcoma, Caco-2 colorectal carcinoma, HaCat keratinocytes, human umbilical vein endothelial cells) representative of diverse attached cell types (data not shown); lymphocytes were not examined in this study. This staining pattern was not seen with antibodies against another death receptor adaptor protein TRADD (data not shown) in the absence of leptomycin, in agreement with a previous report (22). It was unlikely that fixation and permeabilization selectively leached a non-nuclear

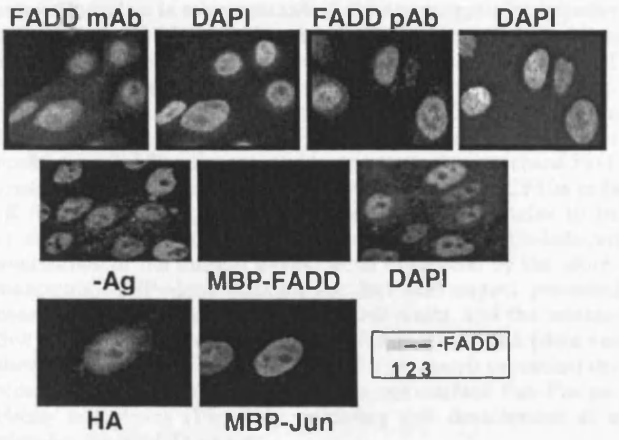


Fig. 1. FADD localizes primarily to the nucleus. (*Top*) MCF10a cells were analyzed for FADD localization by immunofluorescence using a mAb and a polyclonal antibody (pAb); a second pAb produced similar results (data not shown). Nuclear location is indicated by 4′,6-diamidino-2-phenylindole (DAPI) staining. (*Middle*) Immunofluorescent staining was carried out after incubation of the primary antibody (mAb) with recombinant maltose-binding protein fusion proteins of FADD or c-jun. (*Bottom*) MCF10a cells were infected with a retrovirus containing HA-tagged FADD and stained with anti-HA antibody. (*Inset*) Fixation/permeabilization did not extract a detectable pool of FADD. After fixation (lane 2) or fixation plus permeabilization (lane 3) the insoluble fraction was dissolved in SDS sample buffer and analyzed for FADD by Western blotting. Lane 1 contains the soluble fraction from fixation plus permeabilization.

FADD population, because no FADD was found by immunoblotting of the extracted material obtained during these procedures (Fig. 1). Treatment of MCF10a cells with FADD short-interfering RNA reduced both the FADD signal on a Western blot and the average signal intensity of nuclear immunofluorescence (Fig. 6*a*, which is published as supporting information on the PNAS web site). To confirm the specificity of the nuclear signal, we also generated a monospecific anti-mouse FADD antibody and compared the immunofluorescence micrographs by using FADD-knockout MEFs versus FADD-expressing cells (47). The nuclear FADD signal obtained with an anti-mouse FADD polyclonal antibody was reduced to background levels in FADD-knockout MEFs (Fig. 6*a*). Also, certain mutants of FADD (described below) were found in the cytoplasm under the same fixation conditions, further excluding the possibility of a "fixation artifact."

We then tested whether FADD was nuclear by a cell fractionation approach. Hypotonic lysates were subjected to low-speed (800 *g*) centrifugation, yielding a pellet that was highly enriched for nuclei. Equal cell equivalents of nuclear and cytoplasmic fractions were assayed for FADD content by Western blotting (Fig. 6*b*), indicating that the majority of FADD protein was in the nuclear fraction, as was the majority of c-jun protein. By contrast, the cytoplasmic protein procaspase-3 was mainly in the cytoplasmic fraction. It is not yet clear whether the small percentage of FADD found in the cytoplasmic fraction in some experiments was caused by leakage from the nucleus during the extraction procedure, the extent of which varies widely among nuclear proteins (48, 49), or represented a minor cytoplasmic pool that escaped immunofluorescent detection.

Taken together, the data strongly indicated that the primary localization of FADD is in the cell nucleus of adherent cell lines.

Exportin-5 Is a Potential FADD-Shuttling Protein. FADD functions at the membrane as a death receptor adaptor protein, implying that either a minor cytoplasmic pool is maintained constitutively or

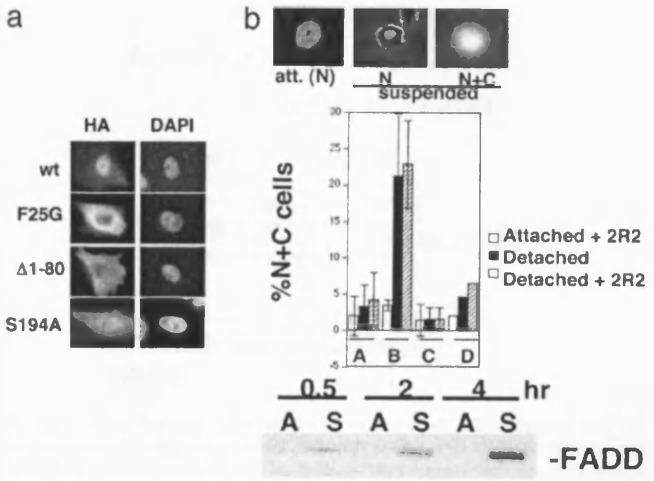


Fig. 2. Characterization of FADD nuclear import/export. (*a*) Mutation of the Ser-194 phosphorylation site or of the DED (amino acids 1–80) of FADD increases its cytoplasmic accumulation. HA–FADD expression constructs containing the indicated FADD mutants were transfected into MCF10a cells, which were then immunostained with anti-HA antibody after being maintained in attached conditions. Representative HA-stained cells and the positions of their nuclei (4′,6-diamidino-2-phenylindole, DAPI) are shown. (*b*) The same mutations of FADD that cause cytoplasmic accumulation also prevent induction of export after nuclear localization has been experimentally enforced. YFP–FADD–NLS expression constructs containing WT FADD or mutant FADD sequences linked to three simian virus 40 NLSs were transfected into MCF10a cells. Attached or (15 min) detached cells treated in the presence or absence of a Fas-agonistic antibody (2R2) were then scored for nuclear (N) or nuclear plus cytoplasmic (N + C) localization of the fluorescent tag; representative images used to score the cells are shown (*Top*), and quantitation of the results is shown in the histogram (*Middle*). The percentages of NC cells under attached conditions (without 2R2) were on average 2% greater for YFP–FADD–NLS than for YFP–NLS. [A, YFP–NLS; B, YFP–FADD (WT)–NLS; C, YFP–FADD (S194A)–NLS; D, YFP–FADD (Δ1–80)–NLS.] (*Bottom*) Detachment of cells from matrix promotes the recruitment of endogenous FADD protein to cell surface Fas–FasL complexes. Cells were incubated with a Fas agonistic mAb (2R2) under attached (A) or suspended (S) conditions for the indicated times; complexes were immunoprecipitated and analyzed for FADD by Western blotting.

that conditions that sensitize the cell to FasL trigger a rapid export of some fraction of FADD from the nucleus. In either event, we suspected that FADD's nucleocytoplasmic transport was likely to be performed actively by a transport protein rather than occurring by diffusion. To investigate the transport mechanism, we noted that FADD is phosphorylated at Ser-194 (50, 51). The function of this phosphorylation is unknown; it is not, however, required for the interaction of FADD with caspase-8 or Fas (50).

When hemagglutinin (HA) epitope-tagged FADD was transfected transiently into MCF10a cells, most cells showed mainly nuclear staining. However, the mutation of Ser-194 to alanine caused the FADD to be almost uniformly distributed throughout the cell (Fig. 2*a*). Among other mutants tested, the only other alterations similarly affecting FADD localization were either the deletion of the death effector domain (DED) (amino acids 1–80) or point mutation of Phe-25 to a glycine, which disrupts DED homotypic interactions (52). Despite this finding, there was no obvious basic amino acid-containing NLS present in the region of Ser-194 or in the DED, suggesting an unusual import mechanism (discussed below).

We then established an assay for the nuclear export of FADD in which YFP-labeled FADD was forced to assume a nuclear localization at time 0, regardless of FADD mutations, caused by three simian virus 40 NLSs appended to the C terminus. This assay allowed us to assay mutants that were potentially defective in both import and export for export specifically. We found that a brief period of cell detachment from matrix caused the WT YFP–FADD–NLS to partially export from the nucleus, producing a nuclear-plus-cytoplasmic signal in ≈20–25% of the cells in 15 min, whereas the control YFP–NLS remained in 95% of the nuclei (Fig. 2*b*) This export may contribute to the increased FasL sensitivity of suspended compared with attached MCF10a cells (R.A.S. and S.M.F., unpublished work), which remains to be established. The export was not caused by anoikis-induced breakdown of the nuclear envelope, as evidenced by the aforementioned YFP-alone control, the fact that export preceded detectable caspase activation by several hours, and the insensitivity of export to the pan-caspase inhibitor zVAD-fmk (data not shown.) Moreover, the detachment of from matrix promoted the recruitment of endogenous FADD to cell-surface Fas–Fas antibody complexes (Fig. 2*b*), validating cell detachment as a stimulus for FADD export.

We used this assay to compare the export capabilities of various mutants of FADD. Interestingly, the S194A and DED mutations (that inhibited import) also inhibited export of FADD. Although the DED contained a leucine-rich sequence that was a potential signal for crm1-mediated export LTELKFLCL, mutations of the leucines in this sequence did not affect export; also, the crm1 inhibitor leptomycin B did not affect export of YFP–FADD–NLS at doses that completely prevented export of the crm1-cargo, protein kinase A-inhibitor (data not shown), suggesting that crm1 did not export FADD.

The observation that the same mutations prevented both FADD import and export and that a phosphorylation site was required for both suggested the involvement of a nucleocytoplasmic shuttle protein that interacted preferentially with phosphorylated cargoes. Such a protein has been identified recently in human cells: hmsn5/exportin-5 (53). Although this protein was originally dubbed an exportin, it was subsequently shown to act as an importin as well (54). FADD interacted efficiently with exportin-5/hmsn5 but not with crm1 in cotransfection experiments in 293T cells (Fig. 7, which is published as supporting information on the PNAS web site). Interestingly, the mutants of FADD that did not transport to or from the nucleus efficiently, S194A and Δ1–80, also did not interact efficiently with exportin-5/hmsn5. Furthermore, GST–FADD protein recovered from transfected 293T cells interacted *in vitro* with purified exportin-5/hmsn5 protein in the presence of activated Ran protein [in accordance with the role of activated Ran in nuclear transport by exportin-5 (53)]; however, the same FADD protein first subjected to dephosphorylation did not (Fig. 7). Taken together, exportin-5/hmsn5 is a candidate FADD-shuttling protein, which remains to be confirmed by protein ablation experiments (complicated by the fact that exportin-5 is required for viability). Nevertheless, the data clearly define a role for Ser-194 phosphorylation in determining the localization of FADD protein.

FADD Interacts with MBD4. We hypothesized that FADD may carry out a novel function in the nucleus. Thus, a cDNA library derived from MCF-7 mammary adenocarcinoma cells was screened for FADD-interacting proteins in a yeast two-hybrid system.

Surprisingly, about one-third of the positive clones were fragments of MBD4. MBD4's interaction with FADD was highly specific in yeast, as several unrelated bait proteins failed to interact with it (Fig. 8*a*, which is published as supporting information on the PNAS web site). Using MBD4 deletion mutants expressed in the yeast two-hybrid system, we mapped the interacting region of MBD4 to amino acids 400–455, which is immediately upstream of the glycosylase domain (Fig. 8*b*). The regions of FADD required for MBD4 interaction were mapped similarly, indicating that the DED (amino acids 1–80) was required for the interaction but the carboxyl-terminal tail region

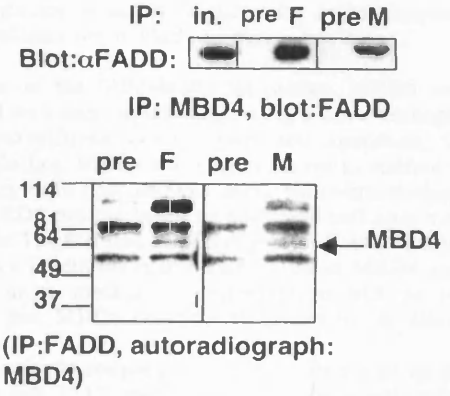


Fig. 3. FADD interacts with MBD4 in mammalian cells. (*Upper*) MCF10a lysates were immunoprecipitated (IP) with sheep anti-MBD4 antibody (M), anti-FADD polyclonal Ab (F), or preimmune serum (pre). Immunoprecipitates were analyzed by Western blotting using FADD mAb and a heavy v chain-specific anti-mouse secondary antibody. (*Lower*) Lysates from ^{32}P-labeled MCF10a cells were immunoprecipitated as above and the products were detected by autoradiography. Note that one of the two major antibody-specific bands precipitated with anti-FADD antibody comigrated with MBD4 (compare lanes M and F). The ≈90-kDa band precipitated with FADD antibody was suspected to be MLH1, which was confirmed below. All of the lanes shown were derived from the same gel; intervening lanes were deleted for clarity.

(needed for exportin-5 interaction as shown above) was not required (Fig. 8*c*). To test the sufficiency of the DED for MBD4 interaction, noting that the DED alone was not expressed well in yeast or 293T cells, the interaction was assayed *in vitro* by using recombinant proteins. Interestingly, the DED of FADD was reproducibly found to interact more efficiently with MBD4 than did full-length FADD (Fig. 8*c*). The involvement of the DED of FADD suggested that MBD4 and caspase-8 might compete against each other for binding to FADD, which was confirmed *in vitro* (data not shown).

To test the FADD–MBD4 interaction in mammalian cells, we first cotransfected epitope-tagged forms of the two proteins into 293T cells, pulling down one protein on glutathione beads and probing a Western blot for the other in both combinations (Fig. 9*a*, which is published as supporting information on the PNAS web site). FADD and MBD4 interacted efficiently and specifically by this criterion. We then assayed for the interaction of the endogenous proteins by coimmunoprecipitation (Fig. 3). A sheep polyclonal antibody was prepared against the C-terminal domain of bacterially expressed human MBD4 and its specificity was confirmed (Fig. 5*d*). This antibody was used to immunoprecipitate MBD4 from MCF10a lysates. These immunoprecipitates (in contrast to those with preimmune sheep IgG) contained a FADD signal that was readily detectable on Western blots with exposure times of <10 s. To perform the converse experiment (noting that the sheep anti-MBD4 did not work well for Western blot detection and that both FADD and MBD4 are phosphoproteins), cellular proteins were metabolically labeled with ^{32}P, and lysates were immunoprecipitated with our polyclonal FADD antiserum. The immunoprecipitates were analyzed by autoradiography of an SDS gel, revealing a specific band that precisely comigrated with MBD4 immunoprecipitated from parallel lysates, as well as a band of the molecular mass of MLH1 (≈87 kDa). Indeed, endogenous FADD immunoprecipitated from MCF10a cells revealed the presence of MLH1 protein by Western blot analysis (Fig. 9*b*), and 293T cotransfection confirmed that the DED was required for this association as well. These data indicate that endogenous FADD and MBD4 proteins interact in MCF10a cells.

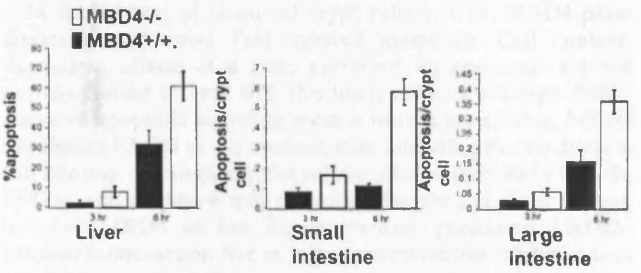


Fig. 4. MBD4 regulates FasL-induced apoptosis. MBD4 knockout or WT mice were injected with the agonistic Fas antibody Jo-2, and at the times indicated, intestinal crypt cells or hepatocytes were analyzed for apoptosis as described (47).

MBD4 Affects Apoptosis. The interaction of MBD4 with the caspase-8-activating protein FADD suggested that MBD4 might be capable of regulating apoptosis. To test this, we used our previously described MBD4 knockout mouse model (29). We recently observed evidence for a proapoptotic role of MBD4 with respect to DNA-damaging agents (O.J.S. and A.R.C., unpublished data). In light of the primary role of FADD in death receptor-induced apoptosis, WT control or MBD4 knockout mice were injected with the Fas-agonistic antibody Jo-2, which induces apoptosis rapidly in intestinal and hepatic epithelial cells that can be scored readily by histological analysis. Interestingly, the MBD4 knockout mice were reproducibly more sensitive (≈5.5-fold in the small intestine) to apoptosis in this context (Fig. 4). To examine the effect of MBD4 on FasL responses in MEFs, we rescued *MBD4*−/− MEFs with an HA–mouse–MBD4 retrovirus, generating mixed populations that expressed MBD4 at 3-fold above the WT MEF level (but substantially lower than the levels in established cell lines such as MCF10a; data not shown). This approach circumvented the comparison of cells from different embryos, which was confounded by variability. Surprisingly, MBD4 sensitized these cells, in contrast to the intestinal cells *in vivo*, to FasL-induced apoptosis (Fig. 10*a*, which is published as supporting information on the PNAS web site). This finding suggested that the effect of MBD4 on FasL-induced apoptosis depended on the cell type or the precise MBD4 expression level (which may in turn affect protein complex composition or MBD4 posttranslational modifications). Interestingly, MBD4 reexpression sensitized these MBD4-knockout cells to anoikis-related caspase activation as well (Fig. 10*b*).

Discussion

Nuclear Localization of FADD. FADD couples death receptors with initiator caspases (35–37). Thus, FADD is assumed to be primarily cytoplasmic. However, the evidence to support this assumption is limited to overexpressed FADD, which aggregates into "death effector filaments" (55, 56). These structures have not been observed for endogenous FADD, and, unlike endogenous FADD (R.A.S. and S.M.F., unpublished data) are Triton-insoluble (56). Thus, the localization of endogenous FADD protein was unexplored to our knowledge until the present report. The localization of FADD reported here does not imply that 100% of FADD is nuclear: clearly, the well-documented death receptor adaptor function of FADD logically implies that this is not the case. Thus, either there is a constitutive but minor cytoplasmic pool of FADD for this purpose or costimulatory signals for Fas signaling such as lack of cell-matrix adhesion (R.A.S. and S.M.F., unpublished data; refs. 43 and 44) stimulate the nuclear export of FADD. The latter is consistent with our results (Fig. 2*b*); thus, it will be interesting to determine how extracellular signals regulate FADD localization, as this could potentially modulate death ligand sensitivity. Control of Ser-194

phosphorylation, which is required for import/export and exportin-5 interaction, is likely to play a key role.

Significance of the FADD–MBD4 Interaction. MBD4 and FADD interacted with each other, suggesting a novel linkage between genome surveillance/DNA repair and apoptosis. Consistent with this finding, MBD4 was shown herein to control apoptosis. Preliminary data (not shown) using short-interfering RNA to reduce MBD4 protein levels in epithelial cell lines support this conclusion. The fact that MBD4 regulates FasL responses clearly implicates a functional connection between MBD4 and FADD, although more mechanistic information will be needed to conclude that MBD4 controls apoptosis by its direct FADD interaction.

Metazoan cells couple genome surveillance to apoptosis (reviewed in refs. 1, 12, and 57) to eliminate genetically aberrant and potentially neoplastic cells. We propose that MBD4-mediated genome surveillance is coupled directly to nuclear FADD, an adaptor protein implicated in multiple apoptotic responses (36–42). MBD4 is thought to specifically repair GT mismatches resulting from the spontaneous deamination of methylcytosine (25), the most frequent cause of point mutations in the human genome (28). In fact, MBD4 knockout mice show a mutator phenotype that leads to enhanced spontaneous carcinoma frequency in the *min* background (29). Thus, one compelling function for the MBD4–FADD complex would be to couple extensive "GT mismatch signals" to apoptosis.

The ability of MBD4 to regulate apoptotic responses to diverse DNA lesions (O.J.S. and A.R.C., unpublished data) or stimuli not apparently causing DNA damage is more challenging to explain. First, MBD4 interacts with MLH1 protein, and the latter promotes apoptosis triggered by a wide variety of DNA damage types (7–11), possibly because of its involvement in a large multifunctional DNA repair complex (58). Second, although FasL and anoikis are not considered DNA-damaging, they rapidly activate caspases, in turn activating nucleases that cause double-strand DNA breaks (59, 60). The latter would elicit DNA damage signals, thus activating caspases further in a feed-forward loop that finalizes the cell's commitment to apoptosis. The possible involvement of FADD and MBD4 in this amplification process, which would explain the proapoptotic effects of MBD4, remains to be determined.

In the context of intestinal crypt cells *in vivo*, MBD4 paradoxically suppressed FasL-induced apoptosis. Cell context-dependent effects of a gene knockout on apoptosis are not unprecedented (cf. ref. 61); this likely reflects cell-type differences in apoptotic signaling mechanisms. Conceivably, MBD4 sequesters FADD in the nucleus, thus inhibiting Fas function in this context. Alternatively, the relative concentrations of FADD, MBD4, and caspase-8 may be critical for the apoptotic output, because MBD4 at low concentrations promoted FADD–caspase-8 interaction but at high concentrations inhibited it *in vitro* (data not shown).

MBD4 is frequently mutated in human colorectal carcinoma cells, most commonly by frameshifts in the A10 repeat at position 301–310 (31–33), which would frameshift the downstream sequences including the glycosylase domain, MLH1 interaction domain, and FADD interaction domain. It will be of interest to mutationally inactivate each of these three functions independently to determine their contributions to apoptosis and understand whether this frameshift is potentially advantageous for tumor growth.

Finally, it has been reported that FADD is required for optimal cell cycle progression of activated T lymphocytes (62) and epithelial cells (R.A.S. and S.M.F., unpublished observations). When combined with the current observation that FADD is a component of a genome surveillance complex, one implication is that this complex may function in a cell cycle checkpoint with MLH1 or MBD4 to recognize the damage and FADD to arrest or promote progression through S phase.

We thank Ed Monosov and Souad Rahmouni for help with microscopy, Jean-Francois Cote and Miya Fujimoto for help with production of baculovirus proteins, Erica Golemis for advice and reagents for the yeast two-hybrid screens, Astar Winoto for the FADD-knockout MEFs, Ian Macara for a sample of exportin-5 protein, Carolyn Ho and Danielle Loughran for technical assistance, and Steve Linke and Robert Abraham for useful critique of the manuscript. The project was supported by National Institutes of Health Grant PO1 CA69381 (to S.M.F.), a Department of Defense Breast Cancer award (DAMD 17-00-0170) (to S.M.F.), a Canadian Institute of Health Research–Institut de Recherche Scientifique sur le Cancer Fellowship (to R.A.S.), and grants from Cancer Research U.K. (to A.R.C. and A.B.).

1. Abraham, R. T. (2001) *Genes Dev.* **15,** 2177–2196.
2. Wang, J. Y. (2000) *Oncogene* **19,** 5643–5650.
3. Khanna, K. K., Keating, K. E., Kozlov, S., Scott, S., Gatei, M., Hobson, K., Taya, Y., Gabrielli, B., Chan, D., Lees-Miller, S. P. & Lavin, M. F. (1998) *Nat. Genet.* **20,** 398–400.
4. Gong, J. G., Costanzo, A., Yang, H. Q., Melino, G., Kaelin, W. G., Jr., Levrero, M. & Wang, J. Y. (1999) *Nature* **399,** 806–809.
5. Kolodner, R. D. & Marsischky, G. T. (1999) *Curr. Opin. Genet. Dev.* **9,** 89–96.
6. Bellacosa, A. (2001) *J. Cell Physiol.* **187,** 137–144.
7. Hardman, R. A., Afshari, C. A. & Barrett, J. C. (2001) *Cancer Res.* **61,** 1392–1397.
8. Meyers, M., Wagner, M. W., Hwang, H. S., Kinsella, T. J. & Boothman, D. A. (2001) *Cancer Res.* **61,** 5193–5201.
9. Zhang, H., Richards, B., Wilson, T., Lloyd, M., Cranston, A., Thorburn, A., Fishel, R. & Meuth, M. (1999) *Cancer Res.* **59,** 3021–3027.
10. Humbert, O., Fiumicino, S., Aquilina, G., Branch, P., Oda, S., Zijno, A., Karran, P. & Bignami, M. (1999) *Carcinogenesis* **20,** 205–214.
11. Aebi, S., Fink, D., Gordon, R., Kim, H. K., Zheng, H., Fink, J. L. & Howell, S. B. (1997) *Clin. Cancer Res.* **3,** 1763–1767.
12. Bernstein, C., Bernstein, H., Payne, C. M. & Garewal, H. (2002) *Mutat. Res.* **511,** 145–178.
13. Formstecher, E., Ramos, J. W., Fauquet, M., Calderwood, D. A., Hsieh, J. C., Canton, B., Nguyen, X. T., Barnier, J. V., Camonis, J., Ginsberg, M. H. & Chneiweiss, H. (2001) *Dev. Cell* **1,** 239–250.
14. Roth, W., Stenner-Liewen, F., Pawlowski, K., Godzik, A. & Reed, J. C. (2002) *J. Biol. Chem.* **277,** 7501–7508.
15. Zhan, Y., Hegde, R., Srinivasula, S. M., Fernandes-Alnemri, T. & Alnemri, E. S. (2002) *Cell Death Differ.* **9,** 439–447.
16. Stegh, A. H., Schickling, O., Ehret, A., Scaffidi, C., Peterhansel, C., Hofmann, T. G., Grummt, I., Krammer, P. H. & Peter, M. E. (1998) *EMBO J.* **17,** 5974–5986.
17. Zheng, L., Schickling, O., Peter, M. E. & Lenardo, M. J. (2001) *J. Biol. Chem.* **276,** 31945–31952.
18. Doostzadeh-Cizeron, J., Yin, S. & Goodrich, D. W. (2000) *J. Biol. Chem.* **275,** 25336–25341.
19. Paroni, G., Henderson, C., Schneider, C. & Brancolini, C. (2002) *J. Biol. Chem.* **277,** 15147–15161.
20. Schickling, O., Stegh, A. H., Byrd, J. & Peter, M. E. (2001) *Cell Death Differ.* **8,** 1157–1168.
21. Torii, S., Egan, D. A., Evans, R. A. & Reed, J. C. (1999) *EMBO J.* **18,** 6037–6049.
22. Morgan, M., Thorburn, J., Pandolfi, P. P. & Thorburn, A. (2002) *J. Cell Biol.* **157,** 975–984.
23. Hendrich, B. & Bird, A. (1998) *Mol. Cell. Biol.* **18,** 6538–6547.
24. Bellacosa, A., Cicchillitti, L., Schepis, F., Riccio, A., Yeung, A. T., Matsumoto, Y., Golemis, E. A., Genuardi, M. & Neri, G. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 3969–3974.
25. Hendrich, B., Hardeland, U., Ng, H. H., Jiricny, J. & Bird, A. (1999) *Nature* **401,** 301–304.
26. Petronzelli, F., Riccio, A., Markham, G. D., Seeholzer, S. H., Genuardi, M., Karbowski, M., Yeung, A. T., Matsumoto, Y. & Bellacosa, A. (2000) *J. Cell Physiol.* **185,** 473–480.
27. Petronzelli, F., Riccio, A., Markham, G. D., Seeholzer, S. H., Stoerker, J., Genuardi, M., Yeung, A. T., Matsumoto, Y. & Bellacosa, A. (2000) *J. Biol. Chem.* **275,** 32422–32429.
28. Krawczak, M., Ball, E. V. & Cooper, D. N. (1998) *Am. J. Hum. Genet.* **63,** 474–488.

29. Millar, C. B., Guy, J., Sansom, O. J., Selfridge, J., MacDougall, E., Hendrich, B., Keightley, P. D., Bishop, S. M., Clarke, A. R. & Bird, A. (2002) *Science* **297,** 403–405.
30. Boland, C. R. (2000) *Ann. N.Y. Acad. Sci.* **910,** 50–59; discussion 59–61.
31. Bader, S., Walker, M., Hendrich, B., Bird, A., Bird, C., Hooper, M. & Wyllie, A. (1999) *Oncogene* **18,** 8044–8047.
32. Bader, S., Walker, M. & Harrison, D. (2000) *Br. J. Cancer* **83,** 1646–1649.
33. Riccio, A., Aaltonen, L. A., Godwin, A. K., Loukola, A., Percesepe, A., Salovaara, R., Masciullo, V., Genuardi, M., Paravatou-Petsotas, M., Bassi, D. E., *et al.* (1999) *Nat. Genet.* **23,** 266–268.
34. Bellacosa, A. (2001) *Cell Death Differ.* **8,** 1076–1092.
35. Kischkel, F. C., Hellbardt, S., Behrmann, I., Germer, M., Pawlita, M., Krammer, P. H. & Peter, M. E. (1995) *EMBO J.* **14,** 5579–5588.
36. Chinnaiyan, A. M., O'Rourke, K., Tewari, M. & Dixit, V. M. (1995) *Cell* **81,** 505–512.
37. Boldin, M. P., Varfolomeev, E. E., Pancer, Z., Mett, I. L., Camonis, J. H. & Wallach, D. (1995) *J. Biol. Chem.* **270,** 7795–7798.
38. Micheau, O., Solary, E., Hammann, A. & Dimanche-Boitrel, M. T. (1999) *J. Biol. Chem.* **274,** 7987–7992.
39. Frisch, S. M. (1999) *Curr. Biol.* **9,** 1047–1049.
40. Rytomaa, M., Martins, L. M. & Downward, J. (1999) *Curr. Biol.* **9,** 1043–1046.
41. Shao, R. G., Cao, C. X., Nieves-Neira, W., Dimanche-Boitrel, M. T., Solary, E. & Pommier, Y. (2001) *Oncogene* **20,** 1852–1859.
42. Yuan, X. J. & Whang, Y. E. (2002) *Oncogene* **21,** 319–327.
43. Shain, K. H., Landowski, T. H. & Dalton, W. S. (2002) *J. Immunol.* **168,** 2544–2553.
44. Aoudjit, F. & Vuori, K. (2001) *J. Cell Biol.* **152,** 633–643.
45. Bhattacharya, D., Logue, E., Bakkour, S., DeGregori, J. & Sha, W. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 8838–8843.
46. Toft, N. J., Winton, D. J., Kelly, J., Howard, L. A., Dekker, M., te Riele, H., Arends, M. J., Wyllie, A. H., Margison, G. P. & Clarke, A. R. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 3911–3915.
47. Kuang, A. A., Diehl, G. E., Zhang, J. & Winoto, A. (2000) *J. Biol. Chem.* **275,** 25065–25068.
48. Gurney, T., Jr., & Foster, D. N. (1977) *Methods Cell Biol.* **16,** 45–68.
49. Gurney, T., Jr., & Collard, M. W. (1984) *Anal. Biochem.* **139,** 25–34.
50. Scaffidi, C., Volkland, J., Blomberg, I., Hoffmann, I., Krammer, P. H. & Peter, M. E. (2000) *J. Immunol.* **164,** 1236–1242.
51. Rochat-Steiner, V., Becker, K., Micheau, O., Schneider, P., Burns, K. & Tschopp, J. (2000) *J. Exp. Med.* **192,** 1165–1174.
52. Eberstadt, M., Huang, B., Chen, Z., Meadows, R. P., Ng, S. C., Zheng, L., Lenardo, M. J. & Fesik, S. W. (1998) *Nature* **392,** 941–945.
53. Brownawell, A. M. & Macara, I. G. (2002) *J. Cell Biol.* **156,** 53–64.
54. Yoshida, K. & Blobel, G. (2001) *J. Cell Biol.* **152,** 729–740.
55. Perez, D. & White, E. (1998) *J. Cell Biol.* **141,** 1255–1266.
56. Siegel, R. M., Martin, D. A., Zheng, L., Ng, S. Y., Bertin, J., Cohen, J. & Lenardo, M. J. (1998) *J. Cell Biol.* **141,** 1243–1253.
57. Wahl, G. M. & Carr, A. M. (2001) *Nat. Cell Biol.* **3,** E277–E286.
58. Wang, Y., Cortez, D., Yazdi, P., Neff, N., Elledge, S. J. & Qin, J. (2000) *Genes Dev.* **14,** 927–939.
59. Nagata, S. (2000) *Exp. Cell Res.* **256,** 12–18.
60. Kovacsovics, M., Martinon, F., Micheau, O., Bodmer, J. L., Hofmann, K. & Tschopp, J. (2002) *Curr. Biol.* **12,** 838–843.
61. Xu, Y. & Baltimore, D. (1996) *Genes Dev.* **10,** 2401–2410.
62. Zhang, J., Kabra, N. H., Cado, D., Kang, C. & Winoto, A. (2001) *J. Biol. Chem.* **276,** 29815–29818.

REFERENCES.

Aber KM, Nori P, MacDonald SM, Bibat G, Jarrar MH, Kaufmann WE. "Methyl-CpG-binding protein 2 is localized in the postsynaptic compartment: an immunochemical study of subcellular fractions." Neuroscience. (2003) 116(1):77-80.

Adams RL. "Eukaryotic DNA methyltransferases--structure and function." Bioessays. (1995) 17(2):139-45.

Akbarian S. "The neurobiology of Rett syndrome." Neuroscientist. (2003) 9(1):57-63

Alberts, B. *et al.* Essential Cell Biology: An Introduction to the Molecular Biology of the Cell (Garland, New York, 1998).

Aldred S, Moore KM, Fitzgerald M, Waring RH. "Plasma amino acid levels in children with autism and their families." J Autism Dev Disord. (2003) 33(1):93-7.

Algeciras-Schimnich A, Shen L, Barnhart BC, Murmann AE, Burkhardt JK, Peter ME. "Molecular ordering of the initial signaling events of CD95." Mol Cell Biol. (2002) 22(1):207-20.

Amir RE, Zoghbi HY. "Rett syndrome: methyl-CpG-binding protein 2 mutations and phenotype-genotype correlations." Am J Med Genet. (2000) 97(2):147-52.

Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. "Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2." Nat Genet. (1999) 23(2):185-8.

Ariani F, Mari F, Pescucci C, Longo I, Bruttini M, Meloni I, Hayek G, Rocchi R, Zappella M, Renieri A. "Real-time quantitative PCR as a routine method for screening large rearrangements in Rett syndrome: Report of one case of MECP2 deletion and one case of MECP2 duplication."Hum Mutat. (2004) 24(2):172-7.

Armstrong DD, Deguchi K, Antallfy B. "Survey of MECP2 in the Rett syndrome and the non-Rett syndrome brain." J Child Neurol. (2003) 18(10):683-7.

Armstrong DD. "Rett syndrome neuropathology 2000." Brain Dev. (2001) 23 Suppl 1:S72-6.

Armstrong DD, Dunn JK, Schultz RJ, Herbert DA, Glaze DG, Motil KJ. "Organ growth in Rett syndrome: a postmortem examination analysis." Pediatr Neurol. (1999) 20(2):125-9.

Armstrong DD. "The neuropathology of the Rett syndrome." Brain Dev. (1992) 14 Suppl:S89-98.

Arney KL. "H19 and Igf2--enhancing the confusion?" Trends Genet. (2003) 19(1):17-23.

Ausio, J., D. B. Levin, *et al.* "Syndromes of disordered chromatin remodelling." Clin Genet (2003) 64(2): 83-95.

Bader S, Walker M, Harrison D. "Most microsatellite unstable sporadic colorectal carcinomas carry MBD4 mutations." Br J Cancer. (2000) 83(12):1646-9.

Bader S, Walker M, Hendrich B, Bird A, Bird C, Hooper M, Wyllie A. "Somatic frameshift mutations in the MBD4 gene of sporadic colon cancers with mismatch repair deficiency." Oncogene. (1999) 23;18(56):8044-7.

Ballestar, E. and A. P. Wolffe "Methyl-CpG-binding proteins. Targeting specific gene repression." Eur J Biochem (2001) 268(1): 1-6.

Balmer D, Arredondo J, Samaco RC, LaSalle JM. "MECP2 mutations in Rett syndrome adversely affect lymphocyte growth, but do not affect imprinted gene expression in blood or brain." Hum Genet. (2002) 110(6):545-52.

Baranda CS, Dymecki SM. "Talking about a revolution: the impact of site specific recombinases in mice." Dev Cell. (2004) 6(1) 7-28.

Bartolomei, M. S. and S. M. Tilghman "Genomic imprinting in mammals." Annu Rev Genet (1997) 31: 493-525.

Bauman ML, Kemper TL, Arin DM. "Microscopic observations of the brain in Rett syndrome." Neuropediatrics. (1995) 26(2):105-8.

Beaud*et al* and Jiang YH. "A rheostat model for a rapid and reversible form of imprinting-dependent evolution." Am J Hum Genet (2002) 70 (6) 1389-1397.

Belichenko PV, Oldfors A, Hagberg B, Dahlstrom A. "Rett syndrome: 3-D confocal microscopy of cortical pyramidal dendrites and afferents." Neuroreport. (1994) 5(12):1509-13.

Bell AC, Felsenfeld G. "Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene." Nature. (2000) 405(6785):482-5.

Bellacosa A. "Role of MED1 (MBD4) Gene in DNA repair and human cancer." J Cell Physiol. (2001) 187(2):137-44. .

Bellacosa A, Cicchillitti L, Schepis F, Riccio A, Yeung AT, Matsumoto Y,
Golemis EA, Genuardi M, Neri G. "MED1, a novel human methyl-CpG-binding endonuclease, interacts with DNA mismatch repair protein MLH1." Proc Natl Acad Sci U S A. (1999) 96(7):3969-74.

Bestor TH. "Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain."
EMBO J. (1992) 11(7):2611-7.

Beyer KS, Blasi F, Bacchelli E, Klauck SM, Maestrini E, Poustka A; "International Molecular Genetic Study of Autism Consortium (IMGSAC.) Mutation analysis of the coding sequence of the MECP2 gene in infantile autism." Hum Genet. (2002) 111(4-5):305-9.

Bienvenu T, Carrie A, de Roux N, Vinet MC, Jonveaux P, Couvert P, Villard L, Arzimanoglou A, Beldjord C, Fontes M, Tardieu M, Chelly J. "MECP2 mutations account for most cases of typical forms of Rett syndrome." Hum Mol Genet. (2000) 9(9):1377-84.

Bird A P. "CpG islands as gene markers in the vertegrate nucleus." TIG (1987) 3 342-347.

Bird A P. " Gene number, noise reduction and biological complexity." TIG (1995) 11(3) 94-100.

Brandeis M, Frank D, Keshet I, Siegfried Z, Mendelsohn M, Nemes A, Temper V, Razin A, Cedar H. "Sp1 elements protect a CpG island from de novo methylation." Nature. (1994) 371 435-8.

Blatt GJ, Fitzgerald CM, Guptill JT, Booker AB, Kemper TL, Bauman ML. "Density and distribution of hippocampal neurotransmitter receptors in autism: an autoradiographic study." J Autism Dev Disord. (2001) 31(6):537-43.

Bourc'his D, Bestor TH. "Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L." Nature. (2004) 431(7004):96-9.

Baranda CS, Dymecki SM. "Talking about a revolution: The impact of site-specific recombinases on genetic analyses in mice."
Dev Cell. (2004) 6(1):7-28.

Brown CJ, Greally JM. "A stain upon the silence: genes escaping X inactivation." Trends Genet. (2003) 19(8):432-8.

Brown JM, Attardi LD. "The role of apoptosis in cancer development and treatment response." Nat Rev Cancer. (2005) 5(3):231-7.

Buermeyer AB, Deschenes SM, Baker SM, Liskay RM. "Mammalian DNA mismatch repair." Annu Rev Genet. (1999) 33:533-64.

Buxbaum JD, Silverman JM, Smith CJ, Greenberg DA, Kilifarski M, Reichert J,
Cook EH Jr, Fang Y, Song CY, Vitale R. "Association between a GABRB3 polymorphism and autism." Mol Psychiatry. (2002) 7(3):311-6.

Cameron EE, Bachman KE, Myohanen S, Herman JG, Baylin SB. "Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer."
Nat Genet. (1999) 21(1):103-7.

Campbell PM, Bovenzi V, Szyf M. "Methylated DNA-binding protein 2 antisense inhibitors suppress tumourigenesis of human cancer cell lines *in vitro* and *in vivo*."
Carcinogenesis. (2004) 25(4):499-507.

Cao, X., W. Aufsatz, *et al.* "Role of the DRM and CMT3 methyltransferases in RNA-directed DNA methylation." (2003) Curr Biol 13(24): 2212-7.

Carter, A. R. and R. A. Segal "Rett syndrome model suggests MECP2 gives neurons the quiet they need to think. " Nat Neurosci (2001) 4(4): 342-3.

Chae JH, Hwang YS, Kim KJ. "Mutation analysis of MECP2 and clinical characterization in Korean patients with Rett syndrome." J Child Neurol. (2002) 17(1):33-6.

Cheadle JP, Gill H, Fleming N, Maynard J, Kerr A, Leonard H, Krawczak M, Cooper DN, Lynch S, Thomas N, Hughes H, Hulten M, Ravine D, Sampson JR, Clarke A. "Long-read sequence analysis of the MECP2 gene in Rett syndrome patients: correlation of disease severity with mutation type and location." Hum Mol Genet. (2000) 9(7):1119-29.

Chedin F, Lieber MR, Hsieh CL. "The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a." Proc Natl Acad Sci U S A. (2002) 99(26):16916-21.

Chen RZ, Akbarian S, Tudor M, Jaenisch R. "Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice." Nat Genet. (2001) 27(3):327-31.

Chen D, Zhao M, Mundy GR. "Bone morphogenetic proteins." Growth Factors. (2004) 22(4):233-41.

Chen WG, Chang Q, Lin Y, Meissner A, West AE, Griffith EC, Jaenisch R, Greenberg ME. "Derepression of BDNF transcription involves calcium-dependent phosphorylation of MECP2." Science. (2003) 302(5646):885-9.

Cheng X. " Structure and function of DNA methyltransferases." Annu Rev Biophys Biomol Struct. (1995) 24:293-318.

Christmann M, Tomicic MT, Roos WP, Kaina B. "Mechanisms of human DNA repair: an update." Toxicology. (2003) 193(1-2):3-34.

Chuang LS, Ian HI, Koh TW, Ng HH, Xu G, Li BF. "Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1." Science. (1997) 277(5334):1996-2000.

Clarke AR, Sansom OJ. "Analyzing tumor suppressor activities in the murine small intestine." Oncol Res. (2003) 13(6-10):333-7.

Clarke PA, te Poele R, Workman P. "Gene expression microarray technologies in the development of new therapeutic agents." Eur J Cancer (2004) 40(17) 2560-2591.

Clayton-Smith J, Watson P, Ramsden S, Black GC. "Somatic mutation in MECP2 as a non-fatal neurodevelopmental disorder in males." Lancet. (2000) 356(9232):830-2.

Cohen D, Lazar G, Couvert P, Desportes V, Lippe D, Mazet P, Heron D. "MECP2 mutation in a boy with language disorder and schizophrenia." Am J Psychiatry. (2002) 159(1):148-9.

Coleman, M. "Is classical Rett syndrome ever present in males?" Brain Dev. (1990.) 12: 31-32,

Collins AL, Levenson JM, Vilaythong AP, Richman R, Armstrong DL, Noebels JL, David Sweatt J, Zoghbi HY. "Mild overexpression of MeCP2 causes a progressive neurological disorder in mice." Hum Mol Genet. (2004) 13(21):2679-89.

Cook EH Jr, Courchesne RY, Cox NJ, Lord C, Gonen D, Guter SJ, Lincoln A, Nix K, Haas R, Leventhal BL, Courchesne E. "Linkage-disequilibrium mapping of autistic disorder, with 15q11-13 markers." Am J Hum Genet. (1998) 62(5):1077-83.

Cortellino S, Turner D, Masciullo V, Schepis F, Albino D, Daniel R, Skalka AM, Meropol NJ, Alberti C, Larue L, Bellacosa A. "The base excision repair enzyme MED1 mediates DNA damage response to antitumor drugs and is associated with mismatch repair system integrity." Proc Natl Acad Sci U S A. (2003) 100(25):15071-6.

Couvert P, Bienvenu T, Aquaviva C, Poirier K, Moraine C, Gendrot C, Verloes A, Andres C, Le Fevre AC, Souville I, Steffann J, des Portes V, Ropers HH, Yntema HG, Fryns JP, Briault S, Chelly J, Cherif B. "MECP2 is highly mutated in X-linked mental retardation." Hum Mol Genet. (2001) 10(9):941-6.

Cunha, G.R., Young, P., Christov, K., Guzman, R., Nandi, S., Talamantes, F. and Thordarson,

G, (1995.) Mammary phenotypic expression induced in epidermal cells by embryonic mammary mesenchyme. Acta Anat. 152, 195–204.

Curradi M, Izzo A, Badaracco G, Landsberger N. "Molecular mechanisms of gene silencing mediated by DNA methylation." Mol Cell Biol. (2002) 22(9):3157-73.

Daniel JM, Reynolds AB. "The catenin p120(ctn) interacts with Kaiso, a novel BTB/POZ domain zinc finger transcription factor." Mol Cell Biol. (1999) 19(5):3614-23.

Deeb S. "The molecular basis of variation in human color vision." Clin Genet. (2005) 67(5):369-77.

DeMarini DM. "Genotoxicity of tobacco smoke and tobacco smoke condensate: a review." Mutat Res. (2004) 567(2-3):447-74.

Deplus R, Brenner C, Burgers WA, Putmans P, Kouzarides T, de Launoit Y,
Fuks F. "Dnmt3L is a transcriptional repressor that recruits histone deacetylase." Nucleic Acids Res. (2002) 30(17):3831-8.

D'Esposito M, Quaderi NA, Ciccodicola A, Bruni P, Esposito T, D'Urso M,
Brown SD. "Isolation, physical mapping, and northern analysis of the X-linked human gene encoding methyl CpG-binding protein, MECP2." Mamm Genome. (1996) 7(7):533-5.

Detich N, Theberge J, Szyf M. "Promoter-specific activation and demethylation by

MBD2/demethylase." J Biol Chem. (2002) 277(39):35791-4.

Dhossche D, Applegate H, Abraham A, Maertens P, Bland L, Bencsath A,
Martinez J. "Elevated plasma gamma-aminobutyric acid (GABA) levels in autistic youngsters: stimulus for a GABA hypothesis of autism." Med Sci Monit. (2002) 8(8):PR1-6.

Dotti MT, Orrico A, De Stefano N, Battisti C, Sicurelli F, Severi S, Lam
CW, Galli L, Sorrentino V, Federico A. "A Rett syndrome MECP2 mutation that causes mental retardation in men." Neurology. (2002) 58(2):226-30.

Drewell RA, Goddard CJ, Thomas JO, Surani MA. "Methylation-dependent silencing at the H19 imprinting control region by MeCP2." Nucleic Acids Res. (2002) 30(5):1139-44.

Drewell RA, Brenton JD, Ainscough JF, Barton SC, Hilton KJ, Arney KL,
Dandolo L, Surani MA. "Deletion of a silencer element disrupts H19 imprinting independently of a DNA methylation epigenetic switch."
Development. (2000) 127(16):3419-28.

Eden S, Cedar H. "Role of DNA methylation in the regulation of transcription." Curr Opin Genet Dev. (1994) 4(2):255-9.

Engel N, West AG, Felsenfeld G, Bartolomei MS. "Antagonism between DNA hypermethylation and enhancer-blocking activity at the H19 DMD is uncovered by CpG

mutations." Nat Genet. (2004) 36(8):883-8.

Engerstrom IW. "Rett syndrome: the late infantile regression period--a retrospective analysis of 91 cases." Acta Paediatr. (1992) 81(2):167-72.

Evans MD, Dizdaroglu M, Cooke MS. "Oxidative DNA damage and disease: induction, repair and significance." Mutat Res. (2004) 567(1):1-61.

Fatemi M, Hermann A, Pradhan S, Jeltsch A. "The activity of the murine DNA methyltransferase Dnmt1 is controlled by interaction of the catalytic domain with the N-terminal part of the enzyme leading to an allosteric activation of the enzyme after binding to methylated DNA." J Mol Biol. (2001) 309(5):1189-99.

Feinberg, A. P. and B. Tycko "The history of cancer epigenetics." Nat Rev Cancer (2004) 4(2): 143-53.

Feinberg, A. P. "The epigenetics of cancer etiology." Semin Cancer Biol (2004) 14(6): 427-32.

Felsenfeld G, Groudine M. "Controlling the double helix." Nature. (2003) 421(6921):448-53.

Feng Q, Zhang Y. "The MeCP1 complex represses transcription through preferential binding, remodeling, and deacetylating methylated nucleosomes." Genes Dev. (2001) 15(7):827-32.

Field, L. M., F. Lyko, *et al.* "DNA methylation in insects." Insect Mol Biol (2004) 13(2): 109-15.

Fujita N, Watanabe S, Ichimura T, Tsuruzoe S, Shinkai Y, Tachibana M, Chiba T, Nakao M. "Methyl-CpG binding domain 1 (MBD1) interacts with the Suv39h1-HP1 heterochromatic complex for DNA methylation-based transcriptional repression." J Biol Chem. (2003) 278(26):24132-8.

Fuks F, Hurd PJ, Wolf D, Nan X, Bird AP, Kouzarides T. "The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation." J Biol Chem. (2003) 278(6):4035-40.

.

Gardiner-Garden, M. and M. Frommer "CpG islands in vertebrate genomes." J Mol Biol (1987) 196(2): 261-82.

Geerdink N, Rotteveel JJ, Lammens M, Sistermans EA, Heikens GT, Gabreels FJ, Mullaart RA, Hamel BC. "MECP2 mutation in a boy with severe neonatal encephalopathy: clinical, neuropathological and molecular findings." Neuropediatrics. (2002) 33(1):33-6.

Giunti L, Pelagatti S, Lazzerini V, Guarducci S, Lapi E, Coviello S, Cecconi A, Ombroni L, Andreucci E, Sani I, Brusaferri A, Lasagni A, Ricotti G, Giometto B, Nicolao P, Gasparini P, Granatiero M, Uzielli ML. "Spectrum and distribution of MECP2 mutations in 64 Italian Rett syndrome girls: tentative genotype/phenotype correlation." Brain Dev. (2001) 23 Suppl 1:S242-

Gowher H, Leismann O, Jeltsch A. "DNA of Drosophila melanogaster contains 5-methylcytosine." EMBO J. (2000) 19(24):6918-23.

Grossman D, Kim PJ, Blanc-Brude OP, Brash DE, Tognin S, Marchisio PC, Altieri DC. "Transgenic expression of survivin in keratinocytes counteracts UVB-inducedapoptosis and cooperates with loss of p53." J Clin Invest. (2001) 108(7):991-9

Guideri F, Acampa M. "Sudden death and cardiac arrhythmias in Rett syndrome." Pediatr Cardiol. (2005) 26(1):111.

Guy J, Hendrich B, Holmes M, Martin JE, Bird A. "A mouse MeCP2-null mutation causes neurological symptoms that mimic Rett syndrome." Nat Genet. (2001) 27(3):322-6.

Hagberg B. "Clinical delineation of Rett syndrome variants." Neuropediatrics. (1995) 26(2):62.

Hagberg B, Witt-Engerstrom I. "Rett syndrome: a suggested staging system for describing impairment profile with increasing age towards adolescence." Am J Med Genet Suppl. (1986) 1:47-59.

Hagberg B, Aicardi J, Dias K, Ramos O. "A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases." Ann Neurol. (1983) 14(4):471-9.

Hagberg BA, Skjeldal OH. "Rett variants: a suggested model for inclusion criteria." Pediatr Neurol. (1994) 11(1):5-11.

Hagberg, B.; Aicardi, J.; Dias, K.; Ramos, O. "A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases." Ann. Neurol. (1983.) 14: 471-479,

Haig D. "Genomic imprinting and kinship: how good is the evidence?" Annu Rev Genet. (2004) 38:553-85.

Hajkova P, Erhardt S, Lane N, Haaf T, El-Maarri O, Reik W, Walter J, Surani MA. "Epigenetic reprogramming in mouse primordial germ cells." Mech Dev. (2002) 117(1-2):15-23.

Hanahan D, Weinberg RA. "The hallmarks of cancer." Cell. (2000) 100(1):57-70.

Hardy K. "Apoptosis in the human embryo." Rev Reprod. (1999) 4(3):125-34.

Hata K, Okano M, Lei H, Li E. "Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice." Development. (2002) 129(8):1983-93.

Heard E. "Recent advances in X-chromosome inactivation." Curr Opin Cell Biol. (2004) 16(3):247-55.

Heard, C. Rougeulle, 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), 107 727–738

Hendrich B, Guy J, Ramsahoye B, Wilson VA, Bird A. "Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development." Genes Dev. (2001) 15(6):710-23.

Hendrich B, Hardeland U, Ng HH, Jiricny J, Bird A. "The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites." Nature. (1999) 401(6750):301-4.

Hendrich B, Bird A. "Identification and characterization of a family of mammalian methyl-CpG binding proteins." Mol Cell Biol. (1998) 18(11):6538-47.

Hendrich, B. and S. Tweedie "The methyl-CpG binding domain and the evolving role of DNA methylation in animals." Trends Genet (2003) 19(5): 269-77.

Hendrich B, Abbott C, McQueen H, Chambers D, Cross S, Bird A. "Genomic structure and chromosomal mapping of the murine and human Mbd1, Mbd2,Mbd3, and Mbd4 genes." Mamm Genome. (1999) 10(9):906-12.

Hennighausen L, Robinson GW. "Think globally, act locally: the making of a mouse mammary gland." Genes Dev. (1998) 12(4):449-55.

Hermann A, Gowher H, Jeltsch A. "Biochemistry and biology of mammalian DNA methyltransferases." Cell Mol Life Sci. (2004) 61(19-20):2571-87.

Hoffbuhr K, Devaney JM, LaFleur B, Sirianni N, Scacheri C, Giron J,Schuette J, Innis J, Marino M, Philippart M, Narayanan V, Umansky R, Kronn D, Hoffman EP, Naidu S. "MeCP2 mutations in children with and without the phenotype of Rett syndrome." Neurology. (2001) 56(11):1486-95.

Holliday, R. and T. Ho "DNA methylation and epigenetic inheritance." Methods (2002) 27(2): 179-83.

Horike S, Cai S, Miyano M, Cheng JF, Kohwi-Shigematsu T. "Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome."
Nat Genet. (2005) 37(1):31-40.

Hsieh CL. "Dependence of transcriptional repression on CpG methylation density." Mol Cell Biol. (1994) 14(8):5487-94.

Humphreys RC. "Programmed cell death in the terminal endbud." J Mammary Gland Biol Neoplasia. (1999) 4(2):213-20.

Humphreys RC, Lydon J, O'Malley BW, Rosen JM. "Mammary gland development is mediated by both stromal and epithelial progesterone receptors." Mol Endocrinol. (1997) 11(6):801-11.

Huppke P, Laccone F, Kramer N, Engel W, Hanefeld F. "Rett syndrome: analysis of MECP2 and clinical characterization of 31 patients." Hum Mol Genet. (2000) 9(9):1369-75.

Imessaoudene B, Bonnefont JP, Royer G, Cormier-Daire V, Lyonnet S, Lyon G, Munnich A, Amiel J. "MECP2 mutation in non-fatal, non-progressive encephalopathy in a male." J Med Genet. (2001) 38(3):171-4.

Ishii T, Makita Y, Ogawa A, Amamiya S, Yamamoto M, Miyamoto A, Oki J.
"The role of different X-inactivation pattern on the variable clinical phenotype with Rett syndrome." Brain Dev. (2001) 23 Suppl 1:S161-4.

Iwano H, Nakamura M, Tajima S. "Xenopus MBD3 plays a crucial role in an early stage of development." Dev Biol. (2004) 268(2):416-28.

Jacob S, Praz F. "DNA mismatch repair defects: role in colorectal carcinogenesis." Biochimie. (2002) 84(1):27-47.

Jarrar MH, Danko CG, Reddy S, Lee YJ, Bibat G, Kaufmann WE. "MECP2 expression in human cerebral cortex and lymphoid cells: immunochemical characterization of a novel higher-molecular-weight form." J Child Neurol. (2003) 18(10):675-82.

Jellinger K, Seitelberger F. "Neuropathology of Rett syndrome." Am J Med Genet Suppl. (1986) 1:259-88.

Jiang CL, Jin SG, Lee DH, Lan ZJ, Xu X, O'Connor TR, Szabo PE, Mann JR, Cooney AJ, Pfeifer GP. "MBD3L1 and MBD3L2, two new proteins homologous to the methyl-CpG-binding proteins MBD2 and MBD3: characterization of MBD3L1 as a testis-specific transcriptional repressor." Genomics. (2002) 80(6):621-9.

Jin SG, Jiang CL, Rauch T, Li H, Pfeifer GP. "MBD3L2 Interacts with MBD3 and Components of the NuRD Complex and Can Oppose MBD2-MeCP1-mediated Methylation Silencing." J Biol Chem. (2005) 280(13):12700-9.

Jiricny J. "Eukaryotic mismatch repair: an update." Mutat Res. (1998) 409(3):107-21.

John RM, Surani MA. "Genomic imprinting, mammalian evolution, and the mystery of egg-laying mammals." Cell. (2000) 101(6):585-8.

Johnston MV, Mullaney B, Blue ME. "Neurobiology of Rett syndrome." J Child Neurol. (2003) 18(10):688-92.

Jones PA. "Altering gene expression with 5-azacytidine." Cell. (1985) 40(3):485-6.

Jones, P. A. and D. Takai (2001.) "The role of DNA methylation in mammalian epigenetics." Science 293(5532): 1068-70.

Jones PL, Veenstra GJ, Wade PA, Vermaak D, Kass SU, Landsberger N, Strouboulis J, Wolffe

AP. "Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription." Nat Genet. (1998) 19(2):187-91.

Jorgensen HF, Ben-Porath I, Bird AP. "Mbd1 is recruited to both methylated and nonmethylated CpGs via distinct DNA binding domains."
Mol Cell Biol. (2004) 24(8):3387-95.

Joza N, Kroemer G, Penninger JM. "Genetic analysis of the mammalian cell death machinery." Trends Genet. (2002) 18(3):142-9.

Jung BP, Jugloff DG, Zhang G, Logan R, Brown S, Eubanks JH. "The expression of methyl CpG binding factor MeCP2 correlates with cellular differentiation in the developing rat brain and in cultured cells." J Neurobiol. (2003) 55(1):86-96.

Kaneda M, Okano M, Hata K, Sado T, Tsujimoto N, Li E, Sasaki H. "Essential role for *de novo* DNA methyltransferase Dnmt3a in paternal and maternal imprinting." Nature. (2004) 429(6994):900-3.

Kass SU, Pruss D, Wolffe AP. "How does DNA methylation repress transcription?" Trends Genet. (1997) 13(11):444-9.

Kass SU, Goddard JP, Adams RL. "Inactive chromatin spreads from a focus of methylation."
Mol Cell Biol. (1993)13(12):7372-9.

Kerr AM, Belichenko P, Woodcock T, Woodcock M. "Mind and brain in Rett disorder." Brain Dev. (2001) 23 Suppl 1:S44-9.

Keshet I, Yisraeli J, Cedar H. "Effect of regional DNA methylation on gene expression." Proc Natl Acad Sci U S A. (1985) 82(9):2560-4.

Kim GD, Ni J, Kelesoglu N, Roberts RJ, Pradhan S. "Co-operation and communication between the human maintenance and *de novo* DNA (cytosine-5) methyltransferases." EMBO J. (2002) 21(15):4183-95.

Kim SW, Park JI, Spring CM, Sater AK, Ji H, Otchere AA, Daniel JM, McCrea PD. "Non-canonical Wnt signals are modulated by the Kaiso transcriptional repressor and p120-catenin." Nat Cell Biol. (2004) 6(12):1212-20.

Kimura H, Shiota K. "Methyl-CpG-binding protein, MeCP2, is a target molecule for maintenance DNA methyltransferase, Dnmt1." J Biol Chem. (2003) 278(7):4806-12.

Kinzler KW, Vogelstein B. "Lessons from hereditary colorectal cancer." Cell. (1996) 87(2):159-70.

Klauck SM, Lindsay S, Beyer KS, Splitt M, Burn J, Poustka A. "A mutation hot spot for nonspecific X-linked mental retardation in the MECP2 gene causes the PPM-X syndrome." Am J Hum Genet. (2002) 70(4):1034-7.

Kleefstra T, Yntema HG, Oudakker AR, Romein T, Sistermans E, Nillessen W, van Bokhoven H, de Vries BB, Hamel BC. "*De novo* MECP2 frameshift mutation in a boy with moderate mental retardation, obesity and gynaecomastia." Clin Genet. (2002) 61(5):359-62.

Klose R, Bird A. "Molecular biology. MeCP2 repression goes nonglobal." Science. (2003) 302(5646):793-5.

Klose RJ, Bird AP. "MeCP2 behaves as an elongated monomer that does not stably associate with the Sin3a chromatin remodelling complex." J Biol Chem. (2004) 279(45):46490-6.

Kouzarides T. "Histone methylation in transcriptional control."
Curr Opin Genet Dev. (2002) 12(2):198-209.

Kriaucionis S, Bird A. "The major form of MeCP2 has a novel N-terminus generated by alternative splicing." Nucleic Acids Res. (2004) 32(5):1818-23.

Kriaucionis S, Bird A. "DNA methylation and Rett syndrome."
Hum Mol Genet. (2003) 12 Spec No 2:R221-7.

Kunert N, Marhold J, Stanke J, Stach D, Lyko F. "A Dnmt2-like protein mediates DNA methylation in Drosophila." Development. (2003) 130(21):5083-90.

Lam CW, Yeung WL, Ko CH, Poon PM, Tong SF, Chan KY, Lo IF, Chan LY, Hui J, Wong V, Pang CP, Lo YM, Fok TF. "Spectrum of mutations in the MECP2 gene in patients with

infantile autism and Rett syndrome." J Med Genet. (2000) 37(12):E41.

Lappalainen R and Reiknonen RS. "Elevated CSF lactate in the Rett syndrome: Cause or consequence?" Brain Dev (1994) 16(5) 399-401.

Lei H, Oh SP, Okano M, Juttermann R, Goss KA, Jaenisch R, Li E. "*De novo* DNA cytosine methyltransferase activities in mouse embryonic stem cells." Development. (1996) 122(10):3195-205.

Lembo F, Pero R, Angrisano T, Vitiello C, Iuliano R, Bruni CB, Chiariotti L. "*MBDin*, a novel MBD2-interacting protein, relieves MBD2 repression potential and reactivates transcription from methylated promoters." Mol Cell Biol. (2003) 23(5):1656-65.

Leonard H, Fyfe S, Dye D, Leonard S. "Familial aggregation in Rett syndrome: what is the evidence for clustering of other disorders in families of affected girls?" Am J Med Genet. (1999) 82(3):228-34.

Li E, Bestor TH, Jaenisch R. "Targeted mutation of the DNA methyltransferase gene results in embryonic lethality." Cell. (1992) 69(6):915-26.

Li E, Beard C, Jaenisch R. "Role for DNA methylation in genomic imprinting." Nature. (1993) 366(6453):362-5.

Li L, Keverne EB, Aparicio SA, Ishino F, Barton SC, Surani MA. "Regulation of maternal

behavior and offspring growth by paternally expressed Peg3." Science. (1999) 284(5412):330-3.

Livak KJ, Schmittgen TD. "Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-(\Delta\Delta Ct)}$ method." Methods (2001) 25(4) 402-408.

Luikenhuis S, Giacometti E, Beard CF, Jaenisch R. "Expression of MeCP2 in postmitotic neurons rescues Rett syndrome in mice." Proc Natl Acad Sci U S A. (2004) 101(16):6033-8.

Lunyak VV, Burgess R, Prefontaine GG, Nelson C, Sze SH, Chenoweth J, Schwartz P, Pevzner PA, Glass C, Mandel G, Rosenfeld MG. "Corepressor-dependent silencing of chromosomal regions encoding neuronal genes." Science. (2002) 298(5599):1747-52.

Lyko F. " DNA methylation learns to fly." Trends Genet. (2001) Apr;17(4):169-72.

Lyko F, Ramsahoye BH, Jaenisch R. "DNA methylation in Drosophila melanogaster." Nature. (2000) 408(6812):538-40.

Macleod D, Charlton J, Mullins J, and Bird AP. "Sp1 sites in the mouse aprt gene promoter are required to prevent methylation of the CpG island" Genes & Dev. (1994) 8: 2282 - 2292

Martinowich K, Hattori D, Wu H, Fouse S, He F, Hu Y, Fan G, Sun YE. "DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation." Science. (2003) 302(5646):890-3.

Matarazzo V, Ronnett GV. "Temporal and regional differences in the olfactory proteome as a consequence of MeCP2 deficiency." Proc Natl Acad Sci U S A. (2004) 101(20):7763-8.

McGrath J and Solter D, "Development of mouse embryogenesis requires both the maternal and paternal genomes." Cell (1984), 37 179–183

Meloni I, Bruttini M, Longo I, Mari F, Rizzolio F, D'Adamo P, Denvriendt K, Fryns JP, Toniolo D, Renieri A. "A mutation in the Rett syndrome gene, MECP2, causes X-linked mental retardation and progressive spasticity in males." Am J Hum Genet. (2000) 67(4):982-5.

Millar CB, Guy J, Sansom OJ, Selfridge J, MacDougall E, Hendrich B, Keightley PD, Bishop SM, Clarke AR, Bird A. "Enhanced CpG mutability and tumorigenesis in MBD4-deficient mice." Science. (2002) 297(5580):403-5.

Miyamoto, A.; Yamamoto, M.; Takahashi, S.; Oki, J. "Classical Rett syndrome in sisters: variability of clinical expression." (1997.) Brain Dev. 19: 492-494.

Monroe JJ, Kort KL, Miller JE, Marino DR, Skopek TR. "A comparative study of *in vivo* mutation assays: analysis of hprt, lacI and cII/cI as mutational targets for N nitroso-N-methylurea and benzo[a]pyrene in Big Blue mice." Mutat Res. (1998) 421(1):121-36.

Monros E, Armstrong J, Aibar E, Poo P, Canos I, Pineda M. "Rett syndrome in Spain: mutation analysis and clinical correlations." Brain Dev. (2001) 23 Suppl 1:S251-3.

Moreno-Fuenmayor H, Borjas L, Arrieta A, Valera V, Socorro-Candanoza L. "Plasma excitatory

amino acids in autism." Invest Clin. (1996) 37(2):113-28.

Motil KJ, Schultz R, Brown B, Glaze DG, Percy AK. "Altered energy balance may account for growth failure in Rett syndrome." J Child Neurol. (1994) 9(3):315-9.

Mullaney BC, Johnston MV, Blue ME. "Developmental expression of methyl-CpG binding protein 2 is dynamically regulated in the rodent brain." Neuroscience. (2004) 123(4):939-49.

Murphy SK, Jirtle RL. "Imprinting evolution and the price of silence." Bioessays. (2003) 25(6):577-88.

Nagata S. "Fas ligand-induced apoptosis." Annu Rev Genet. (1999) 33:29-55.

Nan X, Bird A. "The biological functions of the methyl-CpG-binding protein MeCP2 and its implication in Rett syndrome." Brain Dev. (2001) 23 Suppl 1:S32-7. .

Nan X, Campoy FJ, Bird A. "MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin." Cell. (1997) 88(4):471-81.

Nan X, Campoy FJ, Bird A.

Nan X, Tate P, Li E, Bird A. "DNA methylation specifies chromosomal localization of MeCP2." Mol Cell Biol. (1996) 16(1):414-21.

Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, Bird A. "Transcriptional

repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex." Nature. (1998) 393(6683):386-9.

Nan X, Tate P, Li E, and Bird A "DNA methylation specifies chromosomal localization of MeCP2." Mol. Cell. Biol. 1996 16: 414-421.

Neddermann P, Gallinari P, Lettieri T, Schmid D, Truong O, Hsuan JJ, Wiebauer K, Jiricny J. "Cloning and expression of human G/T mismatch-specific thymine-DNA glycosylase." J Biol Chem. (1996) 271(22):12767-74.

Neul JL, Zoghbi HY. "Rett syndrome: a prototypical neurodevelopmental disorder." Neuroscientist. (2004) 10(2):118-28.

Newell-Price, J., A. J. Clark, *et al.* "DNA methylation and silencing of gene expression." Trends Endocrinol Metab (2000) 11(4): 142-8.

Ng HH, Zhang Y, Hendrich B, Johnson CA, Turner BM, Erdjument-Bromage H, Tempst P, Reinberg D, Bird A. "MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex." Nat Genet. (1999) 23(1):58-61.

Nielsen JB, Friberg L, Lou H, Lassen NA, Sam IL. "Immature pattern of brain activity in Rett syndrome." Arch Neurol. (1990) 47(9):982-6.

Nurmi EL, Dowd M, Tadevosyan-Leyfer O, Haines JL, Folstein SE, Sutcliffe JS. "Exploratory

subsetting of autism families based on savant skills improves evidence of genetic linkage to 15q11-q13." J Am Acad Child Adolesc Psychiatry. (2003) 42(7):856-63.

Ohki I, Shimotake N, Fujita N, Nakao M, Shirakawa M. "Solution structure of the methyl-CpG-binding domain of the methylation-dependent transcriptional repressor MBD1." EMBO J. (1999) 18(23):6653-61.

Okano M, Xie S, Li E. "Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases." Nat Genet. (1998) 19(3):219-20.

Palmer, L. E., P. D. Rabinowicz, *et al.* (2003.) "Maize genome sequencing by methylation filtration." Science 302(5653): 2115-7.

Panganiban G, Rubenstein JL. "Developmental functions of the Distal-less/Dlx homeobox genes." Development. (2002) 129(19):4371-86.

Panganiban G. "Distal-less function during Drosophila appendage and sense organ development." Dev Dyn. (2000) 218(4):554-62.

Panganiban G, Irvine SM, Lowe C, Roehl H, Corley LS, Sherbon B, Grenier JK, Fallon JF, Kimble J, Walker M, Wray GA, Swalla BJ, Martindale MQ, Carroll SB. "The origin and evolution of animal appendages." Proc Natl Acad Sci U S A. (1997) 94(10):5162-6.

Pescucci C, Meloni I, Renieri A. "Is Rett syndrome a loss-of-imprinting disorder?"

Nat Genet. (2005) 37(1):10-1.

Peterson CL, Laniel MA. "Histones and histone modifications." Curr Biol. (2004) 14(14):R546-51.

Petronzelli F, Riccio A, Markham GD, Seeholzer SH, Stoerker J, Genuardi M, Yeung AT, Matsumoto Y, Bellacosa A. "Biphasic kinetics of the human DNA repair protein MED1 (MBD4), a mismatch-specific DNA N-glycosylase." J Biol Chem. (2000) 275(42):32422-9.

Potten CS, Booth C, Pritchard DM. "The intestinal epithelial stem cell: the mucosal governor." Int J Exp Pathol. (1997) 78(4):219-43.

Prokhortchouk A, Hendrich B, Jorgensen H, Ruzov A, Wilm M, Georgiev G, Bird A, Prokhortchouk E. "The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor." Genes Dev. (2001) 15(13):1613-8.

Rabinowicz, P. D., L. E. Palmer, *et al.* "Genes and transposons are differentially methylated in plants, but not in mammals." Genome Res (2003) 13(12): 2658-64.

Rangwala, S. H. and E. J. Richards "The value-added genome: building and maintaining genomic cytosine methylation landscapes." Curr Opin Genet Dev (2004) 14(6): 686-91.

Ravn K, Nielsen JB, Uldall P, Hansen FJ, Schwartz M. "No correlation between phenotype and genotype in boys with a truncating MECP2 mutation."

J Med Genet. (2003) 40(1):e5.

Reese BE, Bachman KE, Baylin SB, Rountree MR. "The methyl-CpG binding protein MBD1 interacts with the p150 subunit of chromatin assembly factor 1." Mol Cell Biol. (2003) 23(9):3226-36.

Reichwald K, Thiesen J, Wiehe T, Weitzel J, Poustka WA, Rosenthal A, Platzer M, Stratling WH, Kioschis P. "Comparative sequence analysis of the MECP2-locus in human and mouse reveals new transcribed regions." Mamm Genome. (2000) 11(3):182-90.

Reik W, Dean W, Walter J. "Epigenetic reprogramming in mammalian development." Science. (2001) 293(5532):1089-93.

Reiss AL, Faruque F, Naidu S, Abrams M, Beaty T, Bryan RN, Moser H. "Neuroanatomy of Rett syndrome: a volumetric imaging study." Ann Neurol. (1993) 34(2):227-34.

Renieri A, Meloni I, Longo I, Ariani F, Mari F, Pescucci C, Cambi F. "Rett syndrome: the complex nature of a monogenic disease." J Mol Med. (2003) 81(6):346-54.

Rett, A. "Ueber ein eigenartiges hirnatrophisches Syndrom bei Hyperammoniamie in Kindesalter." (1966.) Wien. Med. Wschr. 116: 723-738,

Riccio A, Aaltonen LA, Godwin AK, Loukola A, Percesepe A, Salovaara R, Masciullo V, Genuardi M, Paravatou-Petsotas M, Bassi DE, Ruggeri BA, Klein-Szanto AJ, Testa JR, Neri G,

Bellacosa A. "The DNA repair gene MBD4 (MED1) is mutated in human carcinomas with microsatellite instability." Nat Genet. (1999) 23(3):266-8.

Robertson KD, Ait-Si-Ali S, Yokochi T, Wade PA, Jones PL, Wolffe AP. "DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters." Nat Genet. (2000) 25(3):338-42.

Rosen JM, Humphreys R, Krnacik S, Juo P, Raught B. "The regulation of mammary gland development by hormones, growth factors, and oncogenes." Prog Clin Biol Res. (1994) 387:95-111.

Rountree MR, Bachman KE, Baylin SB. "DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci." Nat Genet. (2000) 25(3):269-77.

Ruxton CH, Reed SC, Simpson MJ, Millington KJ. "The health benefits of omega-3 polyunsaturated fatty acids: a review of the evidence." J Hum Nutr Diet 17(5) 449-459.

Ruzov A, Dunican DS, Prokhortchouk A, Pennings S, Stancheva I, Prokhortchouk E, Meehan RR. "Kaiso is a genome-wide repressor of transcription that is essential for amphibian development." Development. (2004) 131(24):6185-94.

Saito M, Ishikawa F. "The mCpG-binding domain of human MBD3 does not bind to mCpG but interacts with NuRD/Mi2 components HDAC1 and MTA2." J Biol Chem. (2002) 277(38):35434-9.

Samaco RC, Hogart A, LaSalle JM. "Epigenetic overlap in autism-spectrum neurodevelopmental disorders: MECP2 deficiency causes reduced expression of UBE3A and GABRB3." Hum Mol Genet. (2005) 14(4):483-92.

Sansom OJ, Bishop SM, Bird A, Clarke AR. "MBD4 deficiency does not increase mutation or accelerate tumorigenesis in mice lacking MMR."
Oncogene. (2004) 23(33):5693-6.

a. Sansom OJ, Berger J, Bishop SM, Hendrich B, Bird A, Clarke AR. "Deficiency of Mbd2 suppresses intestinal tumorigenesis." Nat Genet. (2003) 34(2):145-7.

b. Sansom OJ, Bishop SM,Court H, Dudley S, Liskay RM, Clarke AR. "Apoptosis and mutation in the murine small intestine: loss of Mlh1 and Pms2-dependent apoptosis leads to increased mutation in vivo." DNA Repair (2003) 2(9): 1029-1039.

c. Sansom OJ, Berger J, Bishop SM, Hendrich B, Bird A, Clarke AR. "Deficiency of Mbd2 supresses tumourigenesis." Nat Genet. (2003.) 34(2) 145-147.

Sansom OJ, Toft NJ, Winton DJ, Clarke AR. "Msh-2 suppresses *in vivo* mutation in a gene dose and lesion dependent manner." Oncogene. (2001) 20(27):3580-4.

Sansom OJ, Zabkiewicz J, Bishop SM, Guy J, Bird A, Clarke AR. "MBD4 deficiency reduces the apoptotic response to DNA-damaging agents in the murine small intestine."

Oncogene. (2003) 22(46):7130-6.

Santos F, Hendrich B, Reik W, Dean W. "Dynamic reprogramming of DNA methylation in the early mouse embryo." Dev Biol. (2002) 241(1):172-82.

Scala E, Ariani F, Mari F, Caselli R, Pescucci C, Longo I, Meloni I, Giachino D, Bruttini M, Hayek G, Zappella M, Renieri A. "CDKL5/STK9 is mutated in Rett syndrome variant with infantile spasms." J Med Genet. (2005) 42(2):103-7.

Sekimata M, Takahashi A, Murakami-Sekimata A, Homma Y. "Involvement of a novel zinc finger protein, MIZF, in transcriptional repression by interacting with a methyl-CpG-binding protein, MBD2." J Biol Chem. (2001) 276(46):42632-8.

Sekul EA, Moak JP, Schultz RJ, Glaze DG, Dunn JK, Percy AK. "Electrocardiographic findings in Rett syndrome: an explanation for sudden death?" J Pediatr. (1994) 125(1):80-2.

Shahbazian M, Young J, Yuva-Paylor L, Spencer C, Antalffy B, Noebels J, Armstrong D, Paylor R, Zoghbi H. "Mice with truncated MeCP2 recapitulate many Rett syndrome features and display hyperacetylation of histone H3." Neuron. (2002) 35(2):243-54.

Shahbazian MD, Sun Y, Zoghbi HY. "Balanced X chromosome inactivation patterns in the Rett syndrome brain." Am J Med Genet. (2002) 111(2):164-8.

Silberstein GB, Daniel CW. "Investigation of mouse mammary ductal growth regulation using

slow-release plastic implants." J Dairy Sci. (1987) 70(9):1981-90.

Silva J, Mak W, Zvetkova I, Appanah R, Nesterova TB, Webster Z, Peters AH, Jenuwein T, Otte AP, Brockdorff N. "Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes." Dev Cell. (2003) 4(4):481-95.

Sirianni N, Naidu S, Pereira J, Pillotto RF, Hoffman EP. "Rett syndrome: confirmation of X-linked dominant inheritance, and localization of the gene to Xq28." Am J Hum Genet. (1998) 63(5):1552-8.

Slack A, Bovenzi V, Bigey P, Ivanov MA, Ramchandani S, Bhattacharya S, tenOever B, Lamrihi B, Scherman D, Szyf M. "Antisense MBD2 gene therapy inhibits tumorigenesis." J Gene Med. (2002) 4(4):381-9.

Soriano P. "Generalized lacZ expression with the ROSA26 Cre reporter strain." Nat Genet (1999) 21:70–1.

Stradomska TJ, Tylki-Syzmanska A, Bentkowski Z. " Very long chain fatty acids in Rettt syndrome." Eur J Paediatr. (1999) 158(3) 226-229.

Stancheva I, Collins AL, Van den Veyver IB, Zoghbi H, Meehan RR. "A mutant form of MeCP2 protein associated with human Rett syndrome cannot be displaced from methylated DNA by notch in Xenopus embryos." Mol Cell. (2003) 12(2):425-35.

Svedružic, Z. M. and N. O. Reich "The mechanism of target base attack in DNA cytosine carbon 5 methylation." Biochemistry (2004) 43(36): 11460-73.

Tao J, Van Esch H, Hagedorn-Greiwe M, Hoffmann K, Moser B, Raynaud M, Sperner J, Fryns JP, Schwinger E, Gecz J, Ropers HH, Kalscheuer VM. "Mutations in the X-linked cyclin-dependent kinase-like 5 (CDKL5/STK9) gene are associated with severe neurodevelopmental retardation." Am J Hum Genet. (2004) 75(6):1149-54.

Tariq, M. and J. Paszkowski "DNA and histone methylation in plants." Trends Genet (2004) 20(6): 244-51.

Tatematsu KI, Yamazaki T, Ishikawa F. "MBD2-MBD3 complex binds to hemi-methylated DNA and forms a complex containing DNMT1 at the replication foci in late S phase." Genes Cells. (2000) 5(8):677-88.

Thorvaldsen JL, Bartolomei MS. "Molecular biology. Mothers setting boundaries." Science. (2000) 288(5474):2145-6.

Timmusk T, Persson H, Metsis M. "Analysis of transcriptional initiation and translatability of brain-derived neurotrophic factor mRNAs in the rat brain." Neurosci Lett. (1994) 177(1-2):27-31.

Toft NJ, Arends MJ. "DNA mismatch repair and colorectal cancer." J Pathol. (1998)

185(2):123-9.

Topcu M, Akyerli C, Sayi A, Toruner GA, Kocoglu SR, Cimbis M, Ozcelik T. "Somatic mosaicism for a MECP2 mutation associated with classic Rett syndrome in a boy." Eur J Hum Genet. (2002) 10(1):77-81.

Tudor M, Akbarian S, Chen RZ, Jaenisch R. "Transcriptional profiling of a mouse model for Rett syndrome reveals subtle transcriptional changes in the brain." Proc Natl Acad Sci U S A. (2002) 99(24):15536-41.

Tvrdik P, Westerberg R, Sile S, Asadi A, Jakobsson A, Cannon B, Loison G, Jacobsson A. " Role of a new mammalian gene family in the biosynthesis of very long chain fatty acids and sphingolipids." J Cell Biol (2000) 149 (3) 707-718.

Urnov, F. D. and A. P. Wolffe " Above and within the genome: epigenetics past and present." J Mammary Gland Biol Neoplasia (2001) 6(2): 153-67.

Valentijn AJ, Zouq N, Gilmore AP. "Anoikis." Biochem Soc Trans. (2004) 32(Pt3):421-5.

Varmuza S, Mann M. "Genomic imprinting--defusing the ovarian time bomb." Trends Genet. (1994) 10(4):118-23.

Veenstra-VanderWeele J, Gonen D, Leventhal BL, Cook EH Jr. "Mutation screening of the

UBE3A/E6-AP gene in autistic disorder." Mol Psychiatry. (1999) 4(1):64-7.

Verma, M., P. Maruvada, *et al.* "Epigenetics and cancer." Crit Rev Clin Lab Sci (2004) 41(5-6): 585-607.

Verona RI, Mann MR, Bartolomei MS. "Genomic imprinting: intricacies of epigenetic regulation in clusters." Annu Rev Cell Dev Biol. (2003) 19:237-59.

Vetter ML. "Methylation gets SMRT. Functional insights into Rett syndrome." Dev Cell. (2003) 5(3):359-60.

Villard L, Kpebe A, Cardoso C, Chelly PJ, Tardieu PM, Fontes M. "Two affected boys in a Rett syndrome family: clinical and molecular findings." Neurology. (2000) 55(8):1188-93.

Volkmar F, Chawarska K, Klin A. "Autism in infancy and early childhood." Annu Rev Psychol. (2005) 56:315-36.

Volkmar FR, Pauls D. "Autism." Lancet. (2003) 362(9390):1133-41

Wade PA. "Methyl CpG-binding proteins and transcriptional repression." Bioessays. (2001) 23(12):1131-7

Wakefield RI, Smith BO, Nan X, Free A, Soteriou A, Uhrin D, Bird AP, Barlow PN. "The solution structure of the domain from MeCP2 that binds to methylated DNA."

J Mol Biol. (1999) 291(5):1055-65.

Watanabe S, Ichimura T, Fujita N, Tsuruzoe S, Ohki I, Shirakawa M, Kawasuji M, Nakao M. "Methylated DNA-binding domain 1 and methylpurine-DNA glycosylase link transcriptional repression and DNA repair in chromatin." Proc Natl Acad Sci U S A. (2003) 100(22):12859-64.

Watson P, Black G, Ramsden S, Barrow M, Super M, Kerr B, Clayton-Smith J. "Angelman syndrome phenotype associated with mutations in MECP2, a gene encoding a methyl CpG binding protein." J Med Genet. (2001) 38(4):224-8.

Wilkins JF, Haig D. "What good is genomic imprinting: the function of parent-specific gene expression." Nat Rev Genet. (2003) 4(5):359-68.

Winnepenninckx B, Errijgers V, Hayez-Delatte F, Reyniers E, Frank Kooy R. "Identification of a family with nonspecific mental retardation (MRX79) with the A140V mutation in the MECP2 gene: is there a need for routine screening?" Hum Mutat. (2002) 20(4):249-52.

Wong E, Yang K, Kuraguchi M, Werling U, Avdievich E, Fan K, Fazzari M, Jin B, Brown AM, Lipkin M, Edelmann W. "Mbd4 inactivation increases C→T transition mutations and promotes gastrointestinal tumor formation." Proc Natl Acad Sci U S A. (2002) 99(23):14937-42.

Wrzeska M, Rejduch B. "Genomic imprinting in mammals." J Appl Genet. (2004) 45(4):427-33.

Xu GL, Bestor TH, Bourc'his D, Hsieh CL, Tommerup N, Bugge M, Hulten M, Qu X, Russo JJ, Viegas-Pequignot E. "Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene." Nature. (1999) 402(6758):187-91.

Yamada Y, Miura K, Kumagai T, Hayakawa C, Miyazaki S, Matsumoto A, Kurosawa K, Nomura N, Taniguchi H, Sonta SI, Yamanaka T, Wakamatsu N. "Molecular analysis of Japanese patients with Rett syndrome: Identification of five novel mutations and genotype-phenotype correlation." Hum Mutat. (2001) 18(3):253.

Yntema HG, Kleefstra T, Oudakker AR, Romein T, de Vries BB, Nillesen W, Sistermans EA, Brunner HG, Hamel BC, van Bokhoven H. "Low frequency of MECP2 mutations in mentally retarded males." Eur J Hum Genet. (2002) 10(8):487-90.

Young JI, Zoghbi HY. "X-chromosome inactivation patterns are unbalanced and affect the phenotypic outcome in a mouse model of Rett syndrome." Am J Hum Genet. (2004) 74(3):511-20.

Yuasa, Y. "DNA methylation in cancer and ageing." Mech Ageing Dev (2002) 123(12): 1649-54.

Zappella M, Gillberg C, Ehlers S. "The preserved speech variant: a subgroup of the Rett complex: a clinical report of 30 cases." J Autism Dev Disord. (1998) 28(6):519-26.

Zeev BB, Yaron Y, Schanen NC, Wolf H, Brandt N, Ginot N, Shomrat R, Orr-Urtreger A. "Rett syndrome: clinical manifestations in males with MECP2 mutations." J Child Neurol. (2002) 17(1):20-4.

Zhang Y, Ng HH, Erdjument-Bromage H, Tempst P, Bird A, Reinberg D. "Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation." Genes Dev. (1999) 13(15):1924-35.

Zhao X, Ueba T, Christie BR, Barkho B, McConnell MJ, Nakashima K, Lein ES, Eadie BD, Willhoite AR, Muotri AR, Summers RG, Chun J, Lee KF, Gage FH. "Mice lacking methyl-CpG binding protein 1 have deficits in adult neurogenesis and hippocampal function." Proc Natl Acad Sci U S A. (2003) 100(11):6777-82.

Zhu Y, Spitz MR, Zhang H, Grossman HB, Frazier ML, Wu X. "Methyl-CpG-binding domain 2: a protective role in bladder carcinoma." Cancer. (2004) 100(9):1853-8.

Zoghbi HY. "Introduction: Rett syndrome." Ment Retard Dev Disabil Res Rev. (2002) 8(2):59-60.

Zuckerkandl, E. (2002.) "Why so many noncoding nucleotides? The eukaryote genome as an epigenetic machine." Genetica 115(1): 105-29.