Chapter 2

Materials and Methods

Materials

2.1 Chemical reagents

All common chemicals were purchased from Sigma and Invitrogen unless specified below or in the text.

Amersham Biosciences	2'-deoxynucleoside 5' triphosphates
Invitrogen	Agarose
	Knockout^{TM} DMEM
	Foetal Bovine Serum (FBS)
	Trypsin-EDTA
Melford	5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-Gal)
New England Biolabs	Bovine Serum Albumin (BSA)
Sigma	β-mercaptoethanol
	Dimethyl sulfoxide (DMSO)

Sodium acetate (NaOAc)

Sodium dodecyl sulfate (SDS)

Formamide

Phenol

Phenol:chloroform:isoamylalcohol, 25:24:1

Chloroform:isoamylalcohol, 24:1

Phosphate buffered saline, pH7.2 (PBS)

Minimum Essential Medium Eagle (MEME)

Penicillin/streptomycin/L-glutamine

2.2 Enzymes and commercially prepared kits

All restriction endonucleases were purchased from New England Biolabs (NEB).

Applied Biosystems	AmpliTaq® DNA polymerase
	AmpliTaq Gold® DNA polymerase
BD Biosciences	BD Advantage™ 2 PCR kit
Eppendorf	Phase Lock Gel (PLG)
Invitrogen	Proteinase K
Promega	LigaFast™ rapid DNA ligation system
QIAGEN	DNeasy® blood & tissue kit
	Gel extraction kit
	PCR purification kit
Zymo Research	EZ DNA Methylation Kit™

2.3 Solutions and buffers

Solutions used in this thesis are listed in alphabetical order below. Final concentrations of reagents are given for all solutions. Unless otherwise specified, solutions were made in nanopure water.

2.3.1 Buffers

6x Agarose Gel Loading Dye

0.25% Bromophenol blue

0.25% Xylene cyanol

25% Glycerol

10x NEB PCR buffer

670 mM Tris-HCl (pH 8.8)

166 mM $(NH_4)_2SO_4$ (enzyme grade)

67 mM $MgCl_2$

Adjust pH to 8.8 with HCl

10x TBE

890 mM Tris Base

890 mM Borate

20 mM EDTA (pH 8.0)

1x TE

10 mM Tris-HCl (pH 7.4)

1 mM EDTA

1x $T_{0.1}E$

10 mM Tris-HCl (pH 7.5)

0.1 mM EDTA

1x SET

10 mM Tris-HCl (pH 7.5)

100 mM $NaCl_2$

1mM EDTA (pH 8.0)

28% Sucrose/Cresol red solution

1x $T_{0.1}E$

28% (w/v) Sucrose

0.008% (w/v) Cresol red

2.3.2 Media

All media were prepared in nanopure water and either autoclaved or filter-sterilised prior to use. When used for bacterial growth, 15 mg/ml bacto-agar were added to the appropriate media. Where appropriate ampicillin (dissolved in 1 M sodium bicarbonate, stored at -20 °C) was added to media at a final concentration of 75

μg/ml.

LB

10 mg/ml Bacto-tryptone

5 mg/ml yeast extract

10 mg/ml NaCl

2% w/v dextrose (pH 7.4)

LB/amp/X-Gal plates

LB media

15 g/l agar

75 μg/ml ampicillin

100 μg/ml X-Gal

2.4 Size markers

E-Gel® Low Range Quantitative DNA Ladder (Invitrogen)

Consists of five linear double-stranded DNA fragments of the following sizes in bp: 100, 200, 400, 800, 2000. Electrophoresis of 10 μl of the ladder results in bands containing 5, 10, 20, 40, and 100 ng of DNA, respectively.

1 kb ladder (1 mg/ml) (Invitrogen)

Contains 1 to 12 repeats of a 1018 bp fragment and vector fragments from 75 bp to 1636 bp to produce the following sized fragments in bp: 75, 134, 154, 201, 220, 298, 344, 394, 506/517, 1018, 1636, 2036, 3054, 4072, 5090, 6108, 7125, 8144,

9162, 10180, 11198, 12216. The 1636 bp band contains 10% of the mass applied to the gel.

100 bp DNA Ladder (Invitrogen)

Consists of 15 blunt ended fragments between 100 and 1500 bp in multiples of 100 with an additional fragment at 2072 bp.

2.5 Tissues and cell lines

Human fibroblasts

Female human fibroblast line GM01122 and male human fibroblast line GM01237 (Coriell Cell Repositories: NIGMS Human Genetic Cell Repository), obtained from a brother-sister pair, were cultured in MEME supplemented with 100 U penicillin, 0.1 mg/ml streptomycin, 2 mM L-glutamine, and 15% FBS.

Mouse embryonic fibroblasts

Female mouse embryonic fibroblast line MEF4 and male mouse embryonic fibroblast line MEF2, prepared by Dr Colette Johnston, were cultured in Knockout™ DMEM supplemented with 100 U penicillin, 0.1 mg/ml streptomycin, 2 mM L-glutamine, and 7% FBS.

Opossum tissue samples

Opossum tissue samples were obtained from Dr Sarah Mackay in Glasgow University. Liver, spleen, kidney and heart samples were snap frozen in liquid nitrogen immediately post-mortem, shipped in dry ice and stored at -70 °C.

2.6 Primer sequences

All primer sequences used in this thesis are listed in Appendices I and II. Primers were synthesised by either Sigma or Invitrogen.

2.7 Key World Wide Web addresses

Table 2.1: Key World Wide Web addresses used in this study

Website	Address
EBI- ClustalW2	http://www.ebi.ac.uk/Tools/clustalw2/index.html
EBI - MUSCLE	http://www.ebi.ac.uk/Tools/muscle/index.html
Ensembl genome browser	http://www.ensembl.org
Gap4	http://staden.sourceforge.net/manual/gap4_unix_toc.html
GeneDoc	http://www.nrbsc.org/
NCBI	http://www.ncbi.nlm.nih.gov/
Primer3	http://frodo.wi.mit.edu/primer3/input.htm
RepeatMasker	http://www.repeatmasker.org/cgi-bin/WEBRepeatMasker
The Human Epigenome Project	http://www.epigenome.org
The Wellcome Trust Sanger Institute	http://www.sanger.ac.uk/
UCSC genome browser	http://genome.ucsc.edu/
Vega	http://vega.sanger.ac.uk/
VISTA	http://genome.lbl.gov/vista/index.shtml

Methods

2.8 Mammalian cell culture

The mammalian cells used in this study and their culture conditions are detailed in section 2.5.

2.8.1 Resuscitating frozen cells

1. A vial containing a frozen aliquot of the cells was taken from liquid nitrogen and rapidly defrosted in a 37 °C water bath.
2. Using a plastic Pasteur pipette, contents of the vial were quickly transferred to a 75 cm^2 flask with warm media.
3. Cells were monitored for growth and split after 2-3 days (see section 2.8.2).

2.8.2 Growing and harvesting cells

All cell lines were adherent and were grown at 37°C, 5% CO_2 in 75 cm^2 flasks (Cornings). Cells were passaged when their density reached greater than 80% confluency. These fibroblasts were typically split at 1:3 and required splitting every 3-4 days.

1. The media from the culture flask was aspirated into 1% Virkon.
2. The cells were washed twice in 10 ml PBS.
3. 2 ml of pre-warmed Trypsin-EDTA were added.

4. The flask was incubated at 37 °C for 2 minutes and the cells were dislodged by sharply tapping the flask 2-3 times.

5. Trypsinisation was stopped by addition of 8 ml of pre-warmed growth media with gentle pipetting to resuspend the cells.

6. 4 ml of cell suspension were transferred to each fresh flask containing 12 ml pre-warmed growth media.

If necessary, cells were counted using a haemocytometer (Sigma) with a 0.1 mm sample depth and light microscope (Olympus).

For harvesting, cells were transferred to a 50 ml Falcon tube after step 5 and pelleted by centrifugation at 1200 rpm for 5 minutes. Cell pellets were washed with 5 ml PBS, re-pelleted, and resuspended in appropriate buffers or stored at -70 °C.

2.8.3 Freezing cells for storage

1. Cells from one 75 cm^2 flask were harvested as described in section 2.8.2 and resuspended in 2 ml of freezing medium containing 75% growth medium, 20% FBS, and 5% DMSO.

2. 1 ml of cell suspension was aliquoted into each cryo vial (Nunc).

3. Vials were frozen at 1°C per minute down to -70°C for 24 hours.

4. Once frozen the cells were transferred into liquid nitrogen for long-term storage.

2.9 DNA extraction

2.9.1 Phenol/chloroform extraction of genomic DNA from animal tissues

The DNA solutions were mixed with organic solvents using a blood tube rotator (Stuart) at 20 rpm for 10 minutes. All centrifugation steps were carried out in a Sorvall Legend RT centrifuge at 3000 rpm for 5 minutes. All steps were performed at room temperature unless otherwise stated.

1. A slice of tissue was weighed in a 1.5 ml eppendorf tube. Up to 500 mg of tissue was used for one preparation.

2. The tissue slice was ground to a fine powder in liquid nitrogen using a mortar and pestle and gently resuspended in 4.5 ml 1x SET, 250 μl 10% SDS and 100 μl proteinase K (10 mg/ml) in a 50 ml Falcon tube.

3. The mixture was shaken at 90 rpm at 55 °C overnight in a shaking incubator.

4. 5 ml phenol (TE saturated) were added to the mixture and mixed by gentle inversion. The sample was rolled for 10 minutes to mix, followed by 5 minutes of centrifugation to separate the two phases.

5. Using a Pasteur pipette, the aqueous layer was carefully transferred to a fresh 50 ml Falcon tube containing 10 ml phenol:chloroform:isoamylalcohol (25:24:1).

6. 5 ml 1x SET were added to the organic layer in the original tube, mixed by gentle inversion and submitted to centrifugation.

7. The aqueous layer was added to the phenol:chloroform:isoamylalcohol mix from the previous step, mixed by gentle inversion and submitted to centrifugation.

8. The aqueous layer was combined with 10 ml chloroform:isoamylalcohol (24:1), mixed by gentle inversion and submitted to centrifugation.

9. The extraction with chloroform:isoamylalcohol was repeated (steps 7 and 8), and the final aqueous layer was transferred to a fresh tube containing 25 ml of ice cold absolute ethanol and 1 ml 3M NaOAC (pH 6).

10. The tube was hold horizontally and gently tipped from side to side to mix, and then incubated at -20 °C for one hour or longer.

11. Glass hooks for spooling DNA were made by holding an unplugged glass Pasteur pipette in a gas flame until the glass melts and bends into a hook.

12. DNA was spooled out using a sterile glass hook, dipped into a 1.5 ml eppendorf tube containing 70% ethanol to rinse, then rinsed again in a second tube containing fresh 70% ethanol.

13. The excess liquid was drained off and the DNA was dissolved in 500 μl $T_{0.1}E$.

2.9.2 Phenol/chloroform extraction of DNA from a reaction mix

The phase-lock gel (PLG) heavy was used to facilitate extraction. DNA was extracted from 100-500 μl of reaction mix in one 1.5 ml PLG tube.

1. PLG tubes were spun at 13000 rpm for 1 minute immediately prior to use.
2. The sample and an equal volume of phenol:chloroform:isoamylalcohol (25:24:1) were added to the top of the gel and mixed by vigorously shaking the tube.
3. The tube was spun at 13000 rpm for 5 minutes to separate the phases.
4. The aqueous layer above gel was transferred to a fresh eppendorf tube for ethanol precipitation (section 2.10.1).

2.9.3 Mini-preps of genomic DNA from cultured cells

Genomic DNA was extracted from cultured cells using the DNeasy blood & tissue kit (QIAGEN) in accordance with the manufacturer's protocol. One fresh or frozen pellet containing 1x 10^6 cells (section 2.8.2) was used for each preparation.

2.10 DNA manipulation

2.10.1 Purification of DNA

Ethanol precipitation

1. 0.1 volumes of 3M sodium acetate and 2.5 volumes of 100% ethanol were added to the DNA sample.
2. The samples were well mixed by vortexing and incubated at -20 °C for 30 minutes.
3. The DNA was pelleted by centrifugation for 20 minutes at 13000 rpm using a bench top centrifuge.

4. The pellet was washed with 1 volume of 70% ethanol, followed by centrifugation for 5 minutes at 13000 rpm.

5. DNA pellets were air-dried and resuspended in $T_{0.1}E$.

Direct PCR product purification

PCR products (10 to 50 μl) were purified for subsequent analysis using the PCR purification kit (QIAGEN) in accordance with the manufacturer's protocol.

Agarose gel purification

1. The appropriately sized DNA fragment was excised from the gel using a clean scalpel.

2. The gel slice was weighed in a 1.5 ml eppendorf tube.

3. Purification proceeded using the gel extraction kit (QIAGEN) in accordance with the manufacturer's instructions.

2.10.2 Quantification

DNA was quantified by applying the Beer-Lambert equation relating absorbance, and extinction co-efficient to DNA concentration. Absorbance readings were measured at 260 nm and the extinction coefficients used were 50 for dsDNA and 33 for ssDNA. Absorbance readings were taken on either Eppendorf Biophotometer (Eppendorf) or NanoDrop spectrophotometer (NanoDrop Technologies).

2.10.3 Restriction enzyme digests

1. Restriction digests of genomic DNA (up to 20 μg) were completed using the appropriate buffer and 10-50 units of enzyme per μg DNA (10 units per μg DNA for bisulphite treatment, 50 units per μg DNA for the restriction-PCR methylation assay).

2. Samples were incubated at 37°C overnight.

2.10.4 Agarose gel electrophoresis

1. An agarose gel was prepared (2.5% for most PCR amplified products and 1% for fragments over 1 kb) by melting agarose in 1x TBE. Ethidium bromide was added to a concentration of 250 ng/ml.

2. The appropriate amount of loading buffer was added to the DNA depending on volume.

3. The appropriate amount of samples and appropriate size markers (see section 2.4) were loaded into each well.

4. Electrophoresis was performed in 1x TBE with a voltage ranging between 50-200 volts (5 volts per cm from electrode to electrode) for the time required to obtain satisfactory separation, typically 45 minutes.

5. DNA was visualised under UV light on a transilluminator and digitally photographed using LabWorks Image Acquisition and Analysis Software (UVP Bioimaging Systems).

2.10.5 Bisulphite modification

Bisulphite modification of DNA was performed in a 96-well microtitre plate (Costar Thermowell^{TM} M-type) in a PTC-225 (MJ Research) thermal cycler. All reagents (except $T_{0.1}E$) were from the EZ DNA methylation kit (Zymo Research). Samples were processed in batches of 10.

1. 20 μg of genomic DNA was digested with HindIII (section 2.10.3) and purified using phenol/chloroform extraction (section 2.9.2) followed by ethanol precipitation (section 2.10.1). Digested DNA was diluted to 27 ng/μl in $T_{0.1}E$.

2. One tube of CT Conversion Reagent was dissolved in 750 μl water and 210 μl M-Dilution buffer, and mixed thoroughly by vortexing every 1-2 minutes for 10 minutes.

3. In each of wells 1-10 of row A of a 96-well plate, 45 μl of digested DNA were combined with 5 μl M-Dilution buffer.

4. The plate was incubated in a pre-heated thermal cycler at 37 °C for 15 minutes.

5. The 50 μl of sample in row A were split into five aliquots of 10 μl in rows A-E, to which 20 μl of prepared CT Conversion Reagent were quickly added.

6. Bisulphite modification was performed with the cycling profile: 20 cycles at 95 °C for 30 seconds and 50 °C for 15 minutes.

7. Following bisulphite modification the plate was immediately incubated on ice for 10 minutes.

8. Samples from each column (*e.g.* well 1 from rows A-E) were pooled into a 1.5 ml eppendorf tube, to which 400 μl of M-Binding buffer were added.

9. Desulphonation was carried out in accordance with the manufacturer's instructions.

10. 50 μl $T_{0.1}E$ were added to the 10 μl eluted DNA sample.

11. Concentrations of the samples were measured using a NanoDrop spectrophotometer set at ssDNA (section 2.10.2). Samples were diluted to 10 ng/μl with $T_{0.1}E$ prior to PCR (section 2.11.3). If the concentration was lower than 10 ng/μl, the sample was used directly in PCR.

2.11 Polymerase Chain Reaction (PCR)

PCR was performed in a 96-well microtitre plate or in 0.2 ml tubes (ABgene) in a PTC-225 thermal cycler (MJ Research).

1. A premix sufficient for the number of planned reactions was prepared, allowing for a 1x reaction mix once the DNA template was added. The volume and contents of the final reaction are described below.

2. PCR amplications were performed with the specific cycling profile as described below.

3. Reaction products were visualised by agarose gel electrophoresis and stained with ethidium bromide (section 2.10.4).

2.11.1 Primer design

Primers were designed using Primer3. Where possible primers were selected using the following parameters:

- Melting temperature between 57°C and 63°C
- (G+C) content between 30-80%
- Length between 18 - 22 bp
- Less than 2°C difference in melting temperature between the two primers

Wherever necessary, the parameters were relaxed to the following:

- Melting temperature between 55°C and 65°C
- (G+C) content between 30-80%
- Length between 18 - 27 bp
- Less than 2°C difference in melting temperature between the two primers

Primer pairs that would amplify fragments containing HindIII cleavage sites (AAGCTT) were excluded. BLAST searches (section 2.14.2) were performed to test the specificity of the primer pair.

2.11.2 Restriction-PCR Methylation Assay (RPMA)

RPMA protocol a)

1. PCR mix was prepared containing the following: 1x Advantage 2 PCR buffer, 5% DMSO, 0.4 mM of each deoxyribonucleoside triphosphate (dATP,

dCTP, dGTP, dTTP), 0.67 μM of each primer, 0.5x Advantage 2 polymerase mix and 0.75 U of AmpliTaq Gold DNA polymerase, in a final reaction volume of 10 μl and aliquoted into wells of a 96-well microtitre plate.

The primers used were designed to amplify 150-250 bp fragments containing (ideally 2-3) HpaII/MspI cleavage sites (CCGG) within 5' CpG islands. The expected amplified sequences were checked for number of CCGG sites and absence of HindIII cleavage (AAGCTT) sites using a Perl script (see Appendix I for sequences).

2. Initially the PCR mix was made without the deoxyribonucleoside triphosphate and the polymerases. 5 μl of restriction-enzyme-digested genomic DNA (10 ng/μl) (section 2.10.3) was added to the previously prepared PCR mix and incubated at 100 °C for 5 minutes.

3. Then the rest of ingredients were added and PCR was performed using the cycling profile: 94 °C for 2 minutes, followed by 35 cycles at 94 °C for 30 seconds, 65 °C for 30 seconds and 72 °C for 1 minute.

4. Success of amplification was assessed using gel electrophoresis and ethidium bromide staining (section 2.10.4).

RPMA protocol b)

1. PCR mix was prepared containing the following: 1x GeneAmp PCR Gold buffer, 1.5 mM $MgCl_2$, 0.2 mM of each deoxyribonucleoside triphosphate, 0.4 μM of each primer, and 0.75 U of AmpliTaq Gold DNA polymerase,

in a final reaction volume of 20 μl and aliquoted into wells of a 96-well microtitre plate.

The primers used were as described in protocol a).

2. 5 μl of restriction-enzyme-digested genomic DNA (10 ng/μl) (section 2.10.3) was added to the previously prepared PCR mix.

3. PCR was performed using the cycling profile: 95 °C for 10 minutes, followed by 35 cycles at 95 °C for 30 seconds, 60 °C for 30 seconds and 72 °C for 1 minute, and finally followed by 1 extension cycle at 72 °C for 2 minutes.

4. Success of amplification was assessed using gel electrophoresis and ethidium bromide staining (section 2.10.4).

2.11.3 Bisulphite PCR

1. PCR mix was prepared containing the following: 1x GeneAmp PCR Gold buffer, 1.5 mM $MgCl_2$, 0.2 mM of each deoxyribonucleoside triphosphate, 0.4 μM of each primer, and 0.75 U of AmpliTaq Gold DNA polymerase, in a final reaction volume of 10 μl and aliquoted into wells of a 96-well microtitre plate.

   The primers used were designed to amplify 400-800 bp fragments within 5' CpG islands. Primers were designed to *in silico* bisulphite-converted sequences and contained no CpGs (see Appendix II for sequences).

2. 5 μl of bisulphite modified genomic DNA (section 2.10.5) was used as template and added to the previously prepared PCR mix.

3. PCR was performed using the cycling profile: 95 °C for 10 minutes, followed by 3 cycles at 94 °C for 45 seconds, 60 °C for 1 minute and 72 °C for 1 minute 30 seconds, 3 cycles at 94 °C for 45 seconds, 59 °C for 1 minute and 72 °C for 1 minute 30 seconds, 3 cycles at 94 °C for 45 seconds, 58 °C for 1 minute and 72 °C for 1 minute 30 seconds, and 35 cycles at 94 °C for 45 seconds, 57 °C for 1 minute and 72 °C for 1 minute, where 5 seconds was added to the 72 °C extension step each cycle, and finally followed by 1 extension cycle at 72 °C for 15 minutes.

4. Success of amplification was assessed using gel electrophoresis and ethidium bromide staining (section 2.10.4).

2.11.4 Colony PCR

1. PCR mix was prepared containing the following: 1x NEB PCR buffer, 7 ng/μl of β-mercaptoethanol, 330 μg/ml of BSA, 0.5 mM of each deoxyribonucleoside triphosphate, 0.75 μM of each primer, and 1 U of AmpliTaq DNA polymerase, in a final reaction volume of 15 μl and aliquoted into wells of a 96-well microtitre plate.

   The primers used were M13 forward and reverse (see Appendix II for sequences).

2. Following bacterial transformation (section 2.12.2) individual white colonies were picked using a sterile inoculation loop (VWR) and dipped into 15 μl of the previously prepared PCR mix.

3. PCR was performed using the cycling profile: 94 °C for 5 minutes, followed

by 30 cycles at 94 °C for 30 seconds, 60 °C for 30 seconds and 72 °C for 1 minute 30 seconds, and finally followed by 1 extension cycle at 72 °C for 5 minutes.

4. Success of amplification was assessed using gel electrophoresis and ethidium bromide staining (section 2.10.4).

2.12 Bacterial cloning

2.12.1 Ligation of vector and insert

The pGEM®-T Easy vector was used in all cloning in this thesis. The inserts were all PCR products amplified using Taq polymerase so no A-tailing was necessary.

1. A 10 μl ligation reaction was performed containing 3 units of T4 DNA ligase, 1x ligation buffer, 20 ng vector, and 2 μl insert DNA.

2. The reaction was mixed and incubated for 20 minutes at room temperature or overnight at 4 °C.

2.12.2 Transformation

1. 50 μl of competent Mach1^{TM} cells (Invitrogen) were thawed on ice.

2. 2 μl of ligation mix (section 2.12.1) were added to the cells and mixed by gently flicking the tube.

3. The mixture was incubated on ice for 30 minutes.

4. The cells were heat-shocked at 42 °C in a water bath for 45 seconds, followed by a 2 minute incubation on ice.

5. 500 μl of LB broth were added to each transformation reaction, which was then shaken at 225 rpm at 37 °C for an hour.

6. The transformation mix was then plated out in varying quantities (typically 50 μl) onto agar plates containing ampicillin and X-Gal.

7. The plates were incubated overnight at 37 °C.

8. Success of transformation was determined by colony PCR (section 2.11.4).

2.13 Plasmid sequencing

All plasmid end-sequencing was performed by the Sanger Institute Faculty Small Sequencing Projects (FSSP) team.

2.14 Computational analysis

2.14.1 cpg

The presence of CpG islands were predicted by Val Curwen using the program cpg written by Gos Micklem (unpublished, www.ccbi.cam.ac.uk/software/cpg).

2.14.2 BLAST

Web-based BLAST analysis was performed at Ensembl (`http://www.ensembl.org/`) or Vega (`http://vega.sanger.ac.uk/`).

2.14.3 RepeatMasker

In the bisulphite sequencing analysis, prior to *in silico* bisulphite conversion and primer design, human and mouse repeat sequences were masked using RepeatMasker (Smit and Green, unpublished) (`http://www.repeatmasker.org`).

The repeats were soft-masked with lower case letters in the output file, and marked with underline for subsequent *in silico* bisulphite conversion and primer design.

2.14.4 GAP4

In the bisulphite sequencing analysis, the quality of DNA sequences was assessed using Gap4, a program for sequence assembly and editing. The contig editor used phred confidence values to calculate the confidence of the consensus sequence and identify places requiring visual trace inspection. Vector sequences were automatically located and tagged in the process of assembly. Additional information about Gap4 can be found at: `http://staden.sourceforge.net/overview.html`.

2.14.5 Alignment of DNA sequences

DNA sequences were aligned using the program ClustalW2 (Larkin *et al.*, 2007) or MUSCLE (Labarga *et al.*, 2007) via a web-based server at the EBI (`http://www.ebi.ac.uk/`). User-defined parameters were left at their default settings. For bisulphite sequencing results the alignments were then manually edited and presented using the program GeneDoc (Nicholas 1997, unpublished) (`http://www.nrbsc.org/`).

2.14.6 Multi-species comparative sequence analysis

Web-based multi-species comparative sequence analysis was performed using the VISTA server (Frazer *et al.*, 2004) (`http://genome.lbl.gov/vista/`).

2.14.7 MethTools

Methylation analysis was carried out using MethTools, a collection of open source Perl scripts that make graphical representation of bisulphite sequencing results and calculate methylation densities (Grunau *et al.*, 2000) (`http://genome.imb-jena.de/methtools/`). Customised modifications were made to the original scripts as appropriate and noted in the text. The modified scripts are included in Appendix IV.

2.14.8 Perl scripts

The following Perl scripts were generated to aid analysis of large datasets. These scripts were included in Appendix V.

extract_cpg_seq Extracts sequence of CpG island from the Ensembl or Vega database

parse_primers Checks RPMA primers for absence of HindIII cleavage sites and counts number of HpaII/MspI cleavage sites

primer_invitrogen Converts FASTA format primer file to a primer list compatible with Invitrogen primer plate order form

CGI_CGcount Calculates GC content of the XCR of human, mouse and opossum

CGI_compare_1 Searches for 5' CpG island and orthologous genes for

human, mouse, and opossum XCR genes

CGI_compare_2 Extracts sequence of desired CpG islands from a list

Chapter 3

Comparison of CpG island methylation on the mouse and human X chromosomes

3.1 Introduction

The vertebrate genomes are depleted of the dinucleotide CpG, which is only observed at 0.2-0.25 of the frequency expected from base composition (Russell *et al.*, 1976; Swartz *et al.*, 1962). The majority of the CpGs are methylated at the carbon-5-position of the cytosine (Bird and Taggart, 1980; Ehrlich *et al.*, 1982) and their rarity has been attributed to the tendency for 5mCpG to mutate by deamination to TpG (and CpA in the complementary strand) (Coulondre *et al.*, 1978). However, when genomic DNA from a number of vertebrates was digested with 5mCpG-sensitive restriction enzymes, it was found that about 1% of the genome, the HTF (*Hpa*II tiny fragments) fraction, is highly rich in unmethylated