Chapter 2

Materials & Methods

2.1 Materials

2.1.1 Reagents (listed in alphabetical order)

Antibodies

5-methylcytidine, purified (Eurogentec)

Chemicals

All chemicals used for this thesis were purchased from Sigma.

Enzymes

AmpliTaq Gold Polymerase (Applied Biosystems)

EcoRI (New England Biolabs)

Proteinase K (Roche)

Ribonuclease A (Sigma)

T4 DNA ligase (New England Biolabs)

T4 DNA polymerase (New England Biolabs)

Fluorophores

Cyanine 3-dCTP (Perkin Elmer)

Cyanine 5-dCTP (Perkin Elmer)

Primer Pairs

All primers were purchased from Sigma. List and sequences of primer pairs are provided in Appendix table 2.1. The table section numbers are referenced in the relevant sections of this thesis.

Other Reagents

Big Dye (Applied Biosystems)

dNTP set - 100mM each (GE Healthcare)

Dynalbeads M-280 Sheep-anti mouse IgG (DynalBiotech)

Human Cot1 DNA (Invitrogen)

Human Genomic DNA (Roche)

SYBR Green MasterMix Plus (Eurogentec)

Yeast tRNA (Invitrogen)

2.1.2 Commercial kits

BD Advantage–GC Genomic PCR BD (Biosciences)

DNeasy Tissue Kit (QIAGEN)

Expand Kit (Roche Diagnostics)

EZ-meth Kit (Genetix)

PCR purification Kit (QIAGEN)

RNeasy Mini Kit (QIAGEN)

TOPO TA cloning kit (Invitrogen)

Transcriptor First Strand cDNA synthesis Kit (Roche)

Zymo DNA clean up concentrator-5 (Genetix)

2.1.3 Solutions & Buffers (listed in alphabetical order)

Note: HPLC water was used to prepare solutions & buffers

10mM dNTP	10mM each dNTP (dCTP, dATP, dGTP, dTTP)
(mix for PCR)	
10 x dNTP	0.5mM dCTP
(mix for DNA labelling)	2mM each of dGTP, dTTP and dATP
2 x IP buffer	20mM sodium phosphate (pH 7.0)
	280mM NaCl
	0.1% Triton X-100
EcoRI buffer	New England Biolabs
GTE buffer	20% Glucose
	1M Tris-HCl, pH 8.0

	0.1M EDTA
Hybridization buffer	2 x SSC
	50% deionised formamide
	10 mM Tris-Cl (pH 7.4)
	5% dextran sulphate
	0.1% Tween-20
Proteinase K buffer	10mM Tris-Cl (pH 7.8)
	5mM EDTA
	0.5% SDS
Wash solution 1	2 x SSC
	0.03% SDS
Wash solution 2	0.2 x SSC
Wash solution 3	1 x PBS
	0.05% Tween 20 (Sigma)
Precipitation Mix	100ml 96% ethanol
	200μl 3M sodium acetate
	400μl 0.1mM EDTA

Sequencing Reaction Buffer (x4)

0.32M Tris Base pH 9.0

0.006M $MgCl_2$

9.9% Tetramethylene Sulfone (Sigma)

0.18% Tween-20 (Sigma)

5.9% glycerol

1.0% formamide

1x Restriction Enzyme Buffer 2 New England Biolabs

50 mM NaCl

10 mM Tris-HCl

10 mM $MgCl_2$

1 mM Dithiothreitol

pH 7.9 at 25°C

25mM $MgCl_2$ Applied Biosystems

10 x PCR Gold buffer Applied Biosystems

100 x BSA New England Biolabs

Phosphate Buffer Saline (PBS) pH 7.4

137mM NaCl

2.7mM KCl

10mM Na_2HPO_4

2mM KH_2PO_4

10 x Tris-borate EDTA electrophoresis buffer (TBE) pH 8.3

0.9M Tris-borate

20mM EDTA

10 x TE (Tris-EDTA) buffer pH: 8.0

100mM Tris-Cl

10mM EDTA

LB medium

10 mg/ml Bacto-tryptone

5 mg/ml yeast extract

10 mg/ml NaCl

pH 7.4

LB plates LB medium

15 g/l agar

75 μg/ml Ampicillin

2.1.4 DNA used for tDMR screen – chapter 4

Human DNA samples from healthy individuals were obtained from AMS Biotechnology (Oxon, UK), Analytical Biological Services (Wilmington DE, USA) and from the MHC Haplotype Project (Turner et al., 2008). Samples included DNA extracted from two tissues (liver and placenta) and two cell types ($CD8^+$ lymphocytes and sperm). Additional information on those samples is summarized in Table 2.1.

Donor Information						
Index	**Tissue**	**Replicate**	**Age (yrs)**	**Sex**	**Ethnicity**	**Supplier**
1	Liver	1	37	M	Caucasian	ABS, Wilmington, DE, USA
2	Liver	2	29	M	Caucasian	BCI, Haywatd, CA, USA
3	Placenta	1	29 (mother)	F	Caucasian	ABS, Wilmington, DE, USA
4	Placenta	2	31 (mother)	F	Caucasian	ABS, Wilmington, DE, USA
5	Sperm	1	20-49	M	Caucasian	MHC Haplotype Project (Turner et al., 2008)
6	Sperm	2	20-49	M	Caucasian	MHC Haplotype Project (Turner et al., 2008)
7	T-cells $CD8^+$	1	41	M	Caucasian	ABS, Wilmington, DE, USA
8	T-cells $CD8^+$	2	27	F	African American	ABS, Wilmington, DE, USA

Table 2.1. **Tissues and cell types used in this study.**
ABS: Analytical Biological Services
BCI: BioChain Institute

2.1.5 Cell Lines used for pDMR screen – chapter 5

Cancer cell lines K562, MCF7, 578T, H69, CCRF-CEM, Colo-205, MDA-MB-231, MDA-MB-361 and T47D were provided by the Cancer Genome Project (Wellcome Trust Sanger Institute). The EBV-transformed B-lymphoblastoid cell lines GM10851 and GM15510 were provided by Nigel Carter (The Wellcome Trust Sanger Institute).

2.1.6 Cell Culture Media and Reagents

Cell Culture Media

All media listed in table 2.2 apart from Iscoves Modified DM were purchased from Invitrogen. Iscoves Modified DM was purchased from LGC Promochem.

Cell Culture Reagents

Insulin solution from bovine pancreas (Sigma-Aldrich)

5-aza-2'-deoxycytidine (Sigma-Aldrich)

Foetal Bovine Serum - FBS (Invitrogen)

Penicillin/Streptomycin (Invitrogen)

Dimethyl Sulphoxide - DMSO (Sigma-Aldrich)

Non-essential amino acids (Invitrogen)

D-(+)-Glucose Solution (Sigma-Aldrich)

Sodium biocarbonate (Sigma-Aldrich)

Trypsin/EDTA solution (Invitrogen)

2.1.7 Bacterial Clones

Recombinant pUC plasmid clones were used for the construction of the MHC array (section 2.2.7). These clones were generated at the Wellcome Trust Sanger Institute as part of the HapMap project (The International HapMap Project, 2003). In total 1662 clones were selected. These clones cover the entire MHC (approximately 4Mb). Clones corresponding to gaps and controls were generated as described in section 2.2.7.2. Clone names and genome coordinates of their respective inserts can be found in appendix table 2.2.

2.1.8 Key World Wide Web addresses

Website	Address
Ensembl	http://www.ensembl.org/index.html
GNF - Atlas of Gene Expression	http://expression.gnf.org/cgi-bin/index.cgi
HUGO gene nomencalture	http://www.genenames.org/index.html
Primer3	http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi
Reverse Complement	http://www.bioinformatics.org/sms/rev_comp.html
The Wellcome Trust Sanger Institute	http://www.sanger.ac.uk
UCSC genome browser	http://genome.ucsc.edu/
Vega	http://vega.sanger.ac.uk/index.html
zPicture	http://zpicture.dcode.org/

2.2 Methods

2.2.1 Tissue Culture

2.2.1.1 Culturing of Cell Lines

All cell lines were cultured in media with 10% foetal bovine serum and 1% penicillin-streptomycin solution. Table 2.2 provides the information about the media and supplements used for each cell line. All cell lines had the media changed every two days.

Cell Lines	Media	Supplements	Growth Properties
K562	Iscove's Modified DM		suspension
MCF7	Eagle's MEM	0.01mg/ml bovine insulin 0.1 x non essential amino acids	adherent
T47D	RPMI-1640	0.01mg/ml bovine insulin 2mM L-glutamine 1.5 g/L sodium bicarbonate 4.5 g/L glucose 10 mM HEPES 1.0 mM sodium puruvate	adherent
578T	DMEM	0.01mg/ml bovine insulin	adherent
Colo 205	RPMI-1640		mixed
CCRF-CEM	RPMI-1640	2mM L-glutamine 1.5 g/L sodium bicarbonate 4.5 g/L glucose 10 mM HEPES 1.0 mM sodium puruvate	suspension
MDA-MB-231	Leibovitz's L-15		adherent
MDA-MB-361	Leibovitz's L-15		loosely adherent
H69	RPMI-1640	2mM L-glutamine	suspension
GM10851	RPMI-1640	10 mM HEPES	suspension
GM15510	RPMI-1640	10 mM HEPES	suspension

Table 2.2 **List of all cell lines used.** Note: All cell lines were grown under 5% CO_2 at 37°C in flasks with vented caps (Corning). Only exceptions were MDA-MB-231 and MDA-MB-361 which were cultured in 100% air and in flasks with plug seal caps (Corning).

Once cell growth was confluent, the following steps were taken:

Adherent Cell Lines

1. Culture medium was removed using sterile 2ml aspirating pipette attached to vacuum trap.

2. Monolayer of cells was washed with PBS

3. Cells were trypsinised using trypsin/EDTA solution. Flask was placed in 37°C incubator for 5 min. Equal volume of 10% FBS medium was added to inactivate trypsin.

4. Cells were transferred into 50 ml Falcon tubes and harvested at 1700 rpm for 5 min.

5. Cell pellet was washed once with PBS.

6. Cells were counted with a haemocytometer with a 0.1 mm sample depth under a light microscope (Olympus).

7. Cell pellet was suspended in medium and moved to a bigger flask, split into more flasks, cryo-preserved or used for DNA/RNA extractions as described later in this chapter.

Suspension Cell Lines

A similar procedure, as for adherent cell lines, was followed but excluding steps 2 and 3.

2.2.1.2 Cell cryo-preservation

Cell pellet was resuspended to 10^6 – 10^7 cells/ml in 10% (v/v) DMSO in 10% FCS culture medium, and transferred into polypropylene cryo-tubes. Cryo-tubes were placed in a freezing vessel primed with 250 ml Isopropanol and stored at -70°C overnight. Finally cryo-tubes were transferred to the gas phase of a liquid nitrogen vessel (-180°C) for permanent storage.

2.2.1.3 5-Aza-2'-Deoxycytidine Treatment

Cell lines MCF7 and 578T were treated with 5-aza-2'-deoxycytidine as described below. 5-aza-2'-deoxycytidine was always added fresh to the media.

MCF7 cells

1 x 10^5 MCF7 cells were plated into a 100mm dish (Corning) and, 24h later (day 1), they were treated with 0, 0.8, 2.4, 4.8 and 9.6 μM 5-aza-2'-deoxycytidine (Sigma-Aldrich). The culture was then replenished with fresh drug-containing medium every 48h. DNA and RNA were isolated from the drug treated culture on day 6 as described in sections 2.2.1.4 and 2.2.1.5

578T cells

4 x 10^5 MCF7 cells were plated into a 100mm dish (Corning) and, 24h later (day 1), they were treated with 0, 4 and 8 μM 5-aza-2'-deoxycytidine (Sigma-Aldrich). The culture was then replenished with fresh drug-containing medium every 48h. DNA and RNA were isolated from the drug treated culture on day 6 as described in sections 2.2.1.4 and 2.2.1.5

2.2.1.4 DNA extraction and manipulation

Total genomic DNA was extracted from all cell lines listed in table 2.2. DNA extraction was performed using the DNeasy Tissue Kit in accordance with the manufacturer's protocol. Approximately 5x10^6 cells were used for each DNA extraction. The concentration of the DNA was determined using a Nanodrop (using 1 OD_{260}=50μg ds DNA).

The integrity of DNA was confirmed by visualization on 1.5% agarose gels using ethidium bromide staining.

2.2.1.5 RNA manipulation

All reagents for RNA work were prepared with Diethylene Pyrocarbonate (DEPC) treated water. Bench surfaces and lab ware were cleaned before use with RNAseZap (Ambion).

2.2.1.5.1 RNA extraction

Total RNA was prepared from all cell lines listed in table 2.2 using the RNeasy Mini kit in accordance with the manufacturer's protocol. RNA was eluted with 35 μl of DECP-treated water.

The integrity of the RNA was confirmed by visualization on 1.5% agarose gels using ethidium bromide staining. The concentration of RNA was determined by using a Nanodrop (using 1 OD_{260}= 40μg RNA). A_{260}/A_{280} ratios were also calculated for each sample. Samples with ratios smaller than 1.7 or greater than 2.1 were discarded.

2.2.1.5.2 cDNA synthesis

cDNA was synthesised from total RNA (1μg) using Transcriptor First Strand cDNA synthesis Kit. Anchored-oligo$(dT)_{18}$ primers were used to prime the cDNA synthesis. The synthesis was completed in accordance with the manufacturer's instructions. The resulting cDNA was diluted to 10ng/μl and was stored at -20°C.

2.2.2 Methylated DNA Immunoprecipitation (MeDIP)

MeDIP was done essentially as described before (Keshet et al., 2006; Weber et al., 2005) with some modifications as described below.

2.2.2.1 Sonication of genomic DNA

10μg of genomic DNA resuspended in 100μl of water, was randomly sheared to fragments of 300 to 1000 bp using a Virtis sonicator on full power. DNA was sonicated twice for 75 sec with 1 min incubation on ice in between.

The size of fragments was confirmed by visualization on 1.5% agarose gels stained with ethidium bromide (section 3.3.1).

2.2.2.2 Immunoprecipitation - pDMR screen (chapter 5).

1. 4 µg of sheared genomic DNA resuspended in 240 µl of water were denatured for 10 min at 95-100°C and then placed on ice for 5 min.

2. 250 µl of 2 X IP buffer and 10 µl of 5MeC-mAb (10µg) were added to the DNA sample and incubated at 4ºC with slow rotation for 2 hours.

3. 30 µl of Dynabeads were washed twice with 700 µl of 1 X IP buffer.

4. Dynabeads were magnetically captured using a magnetic rack.

5. DNA-5MeC-Ab sample was added to the pre-washed beads and incubated at 4°C with slow rotation for two hours.

6. Dynabeads were magnetically captured and washed three times with 700µl of 1 x IP buffer.

7. Dynabeads were resuspended in 200 µl of Proteinase K buffer and 2 µl of proteinase K was added to the solution.

8. Solution was incubated at 50°C for 2 hours with rotation in a hybridization oven.

9. Dynabeads were magnetically captured and sample was removed using a P200 gilson pipette.

10. 700 µl of binding buffer (Zymo kit) was added to the sample.

11. Sample was applied to a filter column (Zymo kit) and centrifuged for 10 sec at maximum speed.

12. Filter columns were washed twice with wash buffer (Zymo kit)

13. Immunoprecipitated DNA was eluted twice with 15 µl water (1 min incubation at room temperature prior to centrifugation).

14. The DNA concentration was determined with a NanoDrop (using 1 OD_{260} = 33µg ssDNA)

2.2.2.3 Immunoprecipitation - tDMR screen (chapter 4).

For this screen (due to restricted DNA availability) MeDIP and input DNA was amplified by ligation-mediated PCR (LM-PCR) (Oberley et al., 2004) following the procedure below:

1. 2.5 μg sheared DNA was incubated with 1 X buffer 2, 10 X BSA, 1.2μl dNTP mix (10mM each), 3 Units of T4 DNA polymerase and distilled water to a final volume of 120μl for 20 minutes at 12°C.

2. The reaction was cleaned up using a Zymo-5 kit according to the manufacturer's instructions but the final elution was done in 30μl of TE buffer pH 8.5.

3. The adaptors JW102 (5'-gcggtgacccgggagatctgaattc–3') and JW103 (5'-gaattcagatc-3') were ligated to the cleaned-up DNA by incubation overnight at 16ºC in a reaction containing 40 μl adaptor mix (50μM each), 6 μl T4 DNA ligase 10 X buffer, 5 μl T4 DNA ligase (400U/μl) and distilled water to a final volume of 100μl.

4. DNA was cleaned up as described above.

5. To fill in the overhangs, the sample DNA was incubated at 72ºC for 10 min with 1μl dNTP mix (10mM each), 5μl 10 X AmpliTaq Gold PCR buffer, 3μl $MgCl_2$ (25mM), 5U AmpliTaq Polymerase and distilled water to a final volume of 50μl.

6. DNA was cleaned up as described above.

7. 50 ng of the ligated DNA sample was set aside as the input fraction.

8. 1.2 μg of the ligated DNA sample was denatured for 10 min at 100°C and then placed on ice for 5 min.

9. Immunoprecipitation was performed in 1 X IP buffer and 3 μl of 5-MeC-mAb with incubation at 4ºC with slow rotation for 2 hours.

10. 10 μl Dynabeads (6.7 x 10^8 beads/ml) were washed in 1 X IP buffer according to the manufacturer's instructions and, added to the DNA-antibody mixture and then incubated at 4ºC with slow rotation for 2 hours.

11. The Dynabead-Ab-DNA mixture was washed three times with 500 µl IP buffer and finally resuspended in 100 µl of Proteinase K buffer.

12. 1µl of proteinase K (50 U/ml) was added and incubated at 50ºC for 2 hours with rotation.

13. The sample was cleaned up using a Zymo kit-5 (using 700 µl binding buffer).

14. The DNA concentration was determined with a NanoDrop (using 1 OD_{260} = 33µg ssDNA) and diluted to 1 ng/µl.

15. Two separate LM-PCRs were performed for IP and input fraction respectively. LM-PCR was performed in a final volume of 50µl containing 10 µl distilled water, 10 µl Advantage-GC buffer, 10 µl GC- melt, 3.1 µl 25 mM $Mg(OAc)_2$, 5 µl JW-102 primer (10 µM), 1.4 µl dNTPs, 1 µl Advantage-GC polymerase and 10 µl DNA (1ng/µl). Reaction conditions were as follows: 1 cycle at 95ºC for 2 min for initial denaturation, 20 cycles at 94ºC for 30 sec, 68ºC for 3 min and 1 cycle at 68ºC for 10 min.

16. After LM-PCR, the reactions were cleaned up using a QIAquick PCR Purification kit and eluted with 50µl of water (pre-heated to 50ºC).

2.2.3 Bisulphite Sequencing

2.2.3.1 Primer Design

1. Primers were designed to complement bisulphite treated DNA

2. Primers were designed using primer 3 (Rozen and Skaletsky, 2000)

3. Primers were designed to be 22 bp long (where possible), to have about 30-40% GC content and a melting temperature around 58°C. All primers were designed to contain at least two C to U transitions as a marker of successful bisulphite treatment and to exclude CpG sites, where methylation can vary. Primer sets were further designed to yield amplicons of 300 to 400 bp in size.

The complete list of all primer pair sequences is provided in appendix table 2.1.

2.2.3.2 Bisulphite treatment

Genomic DNA (500 ng) was subjected to sodium bisulphite conversion using the EZ DNA methylation Kit according to the manufacturer's instructions. Elution step was performed with 20 μl elution buffer. The basis of bisulphite treatment is described in chapter 1 (section 1.3.6.1).

2.2.3.3 PCR amplification of bisulphite treated DNA

20 ng of bisulphite converted DNA was used for each PCR. Reactions (25 μl) were set up in 96-well plates (Applied Biosystems) and contained 17.5 μl water, 1 x AmpliTaq Gold buffer, 2 μl primer mix (10 mM each), 2mM $MgCl_2$, 0.5 μl dNTP mix (10 mM each) and 1U of AmpliTaq Gold. The thermal cycling conditions were as follows:

i. 95°C for 5 min

ii. 94°C for 30 sec

iii. 57°C for 1 min

iv. 72°C for 1 min

v. steps ii to iv were repeated 40 times

vi. 72°C for 5 min

PCR products were confirmed by visualization on 1.5% agarose gels using ethidium bromide staining. PCR products were cleaned up using Millipore PCR filter plates as follows. After adding 40 μl of water to the PCR reactions, they were loaded into the filter plate. Plate was placed on a vacuum manifold (10 mmHg for 12 min). Subsequently, 25 μl of water were loaded on the filter plate. After vortexing the plate for 10 min, purified PCR products were retrieved by aspiration.

2.2.3.4 Sequencing

Cleaned PCR fragments were sequenced from both ends using the dideoxy chain terminator method (Sanger et al., 1977), with V3.1 Bigdye terminator chemistry. Sequencing reactions (10 μl) were set up in 96 well plates and contained 0.5 μl Big Dye, 2 μl sequencing reaction buffer, 3 μl primer (3μM), 4μl of cleaned PCR product (section 2.2.3.3) and 2.5 μl water. Thermal cycling conditions were as follows: 1min at 96°C and 45 cycles at 96°C for 10 sec; 50°C for 10 sec; 60°C for 2 min.

Sequencing clean up was performed as follows:

1. 10μl water was added to the samples (total 20 μl).
2. 50 μl of precipitation mix was added, plate was sealed and agitated briefly
3. Plate was centrifuged at 4000 rpm for 20 min at 4°C.
4. Supernatant was tipped off and plate was drained by placing it upside down on tissue paper.
5. 100 μl of chilled 70% ethanol was added to each well.
6. Plate was centrifuged at 4000 rpm for 3 min at 4°C.
7. Steps 4-6 were repeated.
8. Plate was centrifuged upside down on a tissue at 250 rpm for 1 min.
9. Plate was left unsealed in the dark for an hour to evaporate any residual ethanol.
10. Plate was given to the Wellcome Trust Sanger Institute sequencing facilities. Sequencing reactions were analyzed on 3730 ABI sequencing machines (Applied Biosystems, USA).

2.2.3.5 ESME analysis

Quantitative methylation rates were estimated from bisulphite sequence traces using the ESME software (Lewin et al., 2004). ESME estimates, at any given CpG site, the average methylation level from all the copies of DNA amplicons generated during PCR and is therefore, compared to sub-cloning, a more accurate representation of

methylation levels. ESME calculates quantitative methylation values from signal proportions represented by different dyes in four-dye sequence trace files after correcting for imbalanced and over-scaled signals, incomplete bisulphite conversion, quality problems and base-call artefacts. ESME is useful for estimating methylation levels in samples with heterogeneous methylation levels (e.g. human tissues) and it has been used extensively by the HEP (see also section 1.3.6.1) (Eckhardt et al., 2006; Rakyan et al., 2004).

2.2.4 Quantitative real-time PCR

2.2.4.1 Primer Design

Primer pairs for all qRT-PCR assays were designed with Primer 3. The amplicons generated by these primers were 100 to 200 bp long. Primer pairs used for expression studies were designed for regions across intron-exon boundaries to avoid false positives arising from amplification of contaminating genomic DNA. The complete list of primer sequences is provided in appendix table 2.1.

2.2.4.2 qRT-PCR amplification

qRT-PCR was performed using an ABI Prism 7300 Sequence Detection System, using Optical MicroAmp 96-well plates and optical adhesive covers (Applied Biosystems). For each qRT-PCR reaction (total volume of 13.5 μl), 6.5 μl SYBR Green PCR master mix and 2.5 μl primer mix (1.5 μM each.) were used. 15 ng of cDNA and 30 ng of DNA were used for expression and MeDIP validation assays respectively.

Reaction conditions were as follows: 1 cycle at 50°C for 2 minutes, 1 cycle at 95°C for 10 minutes, 40 cycles at 95°C for 15 sec and at 60°C for 1 minute. Reactions were done in triplicates.

2.2.4.3 qRT-PCR assay data analysis

The efficiency, reproducibility and dynamic range of the assay was determined by constructing a standard curve using serial dilutions of a known template every time a new primer pair was used. Primer pairs were used only if efficiency of the assay was 90 to 100%, the slope of the curve around 3.0 and the Ct values for all technical replicates were similar. The presence of non-specific products was identified by constructing melting curves for each primer pair at 0.1°C intervals between 60°C and 95°C.

MeDIP validation assay

To evaluate the relative enrichment of target sequences after MeDIP, the C_t of the MeDIP fraction was normalized (for each sequence tested) to the C_t of the input (ΔC_t). Subsequently I normalised the ΔC_t of each target sequence to the ΔC_t of an unmethylated control sequence ($\Delta\Delta C_t$). Finally I calculated the enrichment as follows:

$$E = 2^{-\Delta\Delta Ct}$$

Expression assay

Gene expression in the cancer cell lines used in this study was determined in relation to a reference gene. As reference, the *ubiquitin C* gene (*UBC*) was used. (Vandesompele et al., 2002). Additional housekeeping genes (*RPL13A* and *GAPDH1*) were also tested and based on validation experiments UBC was assessed to be an adequate representative (figure 2.1). The EBV-transformed cell lines GM10851 and GM15510 were used as controls for normal *UBC* expression.

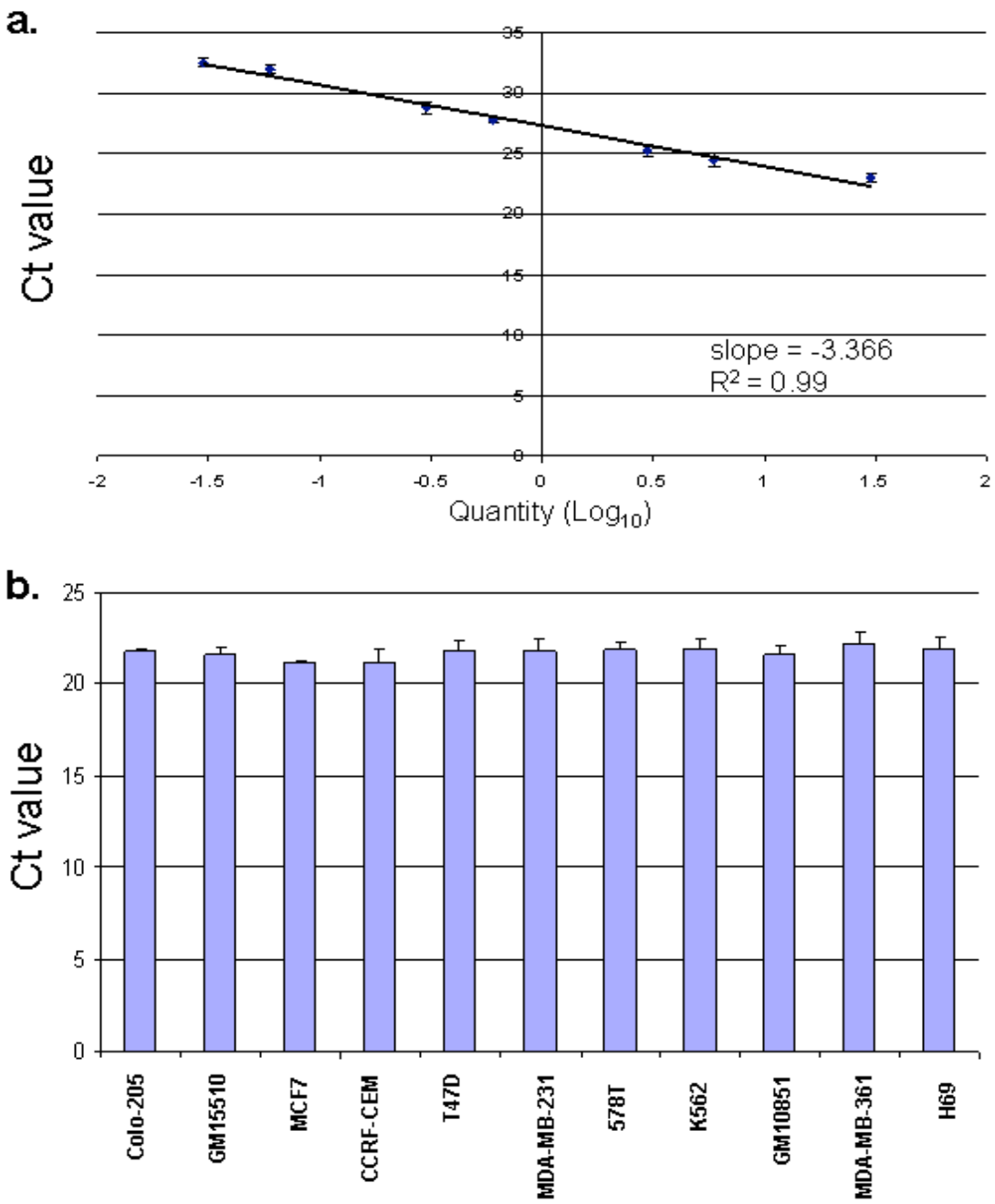

Figure 2.1 **Validation of UBC primers**. a). UBC detection by RT-PCR. UBC was detected by serial dilutions of human genomic DNA. The figure shows a standard curve generated from the real time amplification plot. Figure shows the average of three independent experiments. The efficiency of the reaction was 99% (R^2 =0.99) b). UBC detection in various cell lines. UBC detection by RT-PCR was performed on 15ng cDNA originated by the 11 cell lines tested for the pDMR screen (chapter 5). In all cases UBC detected within the same range of PCR cycles (Ct value close to 22 in all cases). Figure shows the average of three independent experiments for each cell line.

The relative difference in expression level of a target gene in cancer cell lines (test sample) compared to the EBV-transformed cell lines (controls) was determined as follows:

First the C_T of the target gene was normalized to the C_T of the reference gene for both test sample and control sample as follows:

$$\Delta C_{T(test)} = C_{T(target, test)} - C_{T(ref, test)}$$

$$\Delta C_{T(control)} = C_{T(target, control)} - C_{T(ref, control)}$$

Second, the ΔC_T of the test sample was normalised to the ΔC_T of the control as follows:

$$\Delta\Delta C_T = \Delta C_{T(test)} - \Delta C_{T(control)}$$

Finally the expression ratio was calculated as follws:

$$2^{-\Delta\Delta C_T} = \text{Normalised expression ratio}$$

2.2.5 Bacterial Cloning

Cloning of PCR fragments was performed using the TOPO TA Cloning Kit following the manufacture's protocol. Briefly, 1μl of PCR product was added to a mixture containing 4μl water, 1μl salt solution (1.2 M NaCl, 0.06 M $MgCl_2$) and 0.083μl TOPO vector. After 30 min incubation at room temperature 25μl of TOPO chemically competent cells were added to the mixture. Cells were heat-shocked for 45 sec at 42°C and 150μl of SOC medium was added immediately to the cells. Cells were grown for 1h (37°C, 200 rpm) before plating onto LB-amp plates.

2.2.6 Mini-preps of plasmid DNA

A single colony was inoculated into 1ml of LB broth containing ampicillin and grown overnight at 37°C at 320 rpm. For this purpose sterile 96 deep well blocks (2ml capacity) were used. On the following day the cells were pelleted for 2 min at 4000 rpm and resuspended in 120μl GTE buffer on ice. After ensuring complete resuspension of the pellet 120μl NaOH/SDS and 120μl KoAc were added to the cells.

140μl of the cell lysates were removed from the bottom of each deep well and dispensed into a Costar filter plate, which was placed on top of a Costar 3365 storage plate containing 140μl of 100% isopropanol, per well. Plates were centrifuged for 15 min at 400rpm and 4°C. Filter plate was discarded and isopropanol tipped-off. After addition of 100μl of 70% ethanol to the wells and 5 min centrifugation (4000rpm, 4°C) the plate was dried. Finally plasmids were dissolved by adding 60μl of water in each well.

2.2.7 Restriction Digests

Restriction digests of plasmid DNA (up to 10μg) were carried out using 1 x EcoRI buffer and 20 units of EcoRI enzyme. Samples (20μl reaction) were incubated at 37°C for 2 hours, and the resulting digest was confirmed by agorose gel electrophoresis.

2.2.8 Colony PCR

Following bacterial transformation (section 2.2.5) individual colonies were picked using sterile toothpicks and resuspended in 25μl PCR reaction buffer. For this purpose 96 well plates kept on ice were used. M13 forward and reverse primers were used for colony PCR. PCR conditions were as follows:

i. 95°C for 8 min
ii. 94°C for 30 sec
iii. 55°C for 30 sec
iv. 72°C for 1 min
v. Steps ii to iv repeated 30 times
vi. 72°C for 10 min

PCR products were confirmed by agarose gel electrophoresis and staining with ethidium bromide.

2.2.9 MHC tile path array

2.2.9.1 Generation of amino-linked probes

A total of 1747 overlapping plasmid clones were used to generate the array. Of those, 1662 clones (average insert size 2 kb) were picked from the HapMap chromosome 6 library (The International HapMap Project, 2003) and 85 clones were generated by cloning gap-spanning PCR amplicons (average insert size 332 bp). In addition, I generated and included 43 PCR-derived clones as controls. Generation of gap and control clones is described in the following section (2.2.9.2).

Double-stranded amino-linked amplicons were generated from each clone using vector-specific PCR in 50 mM KCl, 5 mM Tris pH 8.5 and 2.5 mM $MgCl_2$ (10 min at 95°C; followed by 35 cycles of 95°C for 1 min, 60°C for 1.5 min, 72°C for 7 min; and a final extension of 72˚C for 10 min - Forward primer 5'-CCCAGTCACGACGTTGTAAAACG-3'; Reverse primer 5'-AGCGGATAACAATTTCACACAGG-3'). In order to generate strand-specific array probes, two separate PCR reactions were performed for each clone, in one case using a 5'-aminolinked primer for the forward strand, and in the other case, for the reverse strand. After quality assessment of the products by gel electrophoresis, spotting buffer was added directly to a final concentration of 250 mM sodium phosphate pH 8.5, 0.00025% w/v sarkosyl, 0.1% sodium azide, and the products were filtered (Multiscreen-GV filter plates, Millipore).

2.2.9.2 Gap closure and control clones

As mentioned in the previous section (2.2.9.1) 85 gap and 45 control clones were generated. Using the appropriate primer pair (appendix table 2.1.1.1), amplicons corresponding to gap regions and control regions were generated by PCR. For this purpose commercially available human genomic DNA was used. Generation of clones for the gaps and controls was performed as described in section 2.2.5 and

2.2.6. Successful cloning amplicons confirmed by restriction digestions (2.2.7) and colony PCR (2.2.8)

2.2.9.3 Array printing and processing

Array printing and processing were performed at the Wellcome Trust Sanger Institute Microarray Facility as follows:

1. Array probes were printed onto amino binding slides (Motorola) at 20-25ºC, 40-50% relative humidity using a MicroGrid II arrayer (Biorobotics/Apogent Discoveries).

2. The array probes were printed in a 24 block format with spots in duplicates.

3. The slides were transferred into a microscope slide rack and placed in a humid chamber (NaCl saturated with water in an air-tight container) and incubated for 24-72 hours at room temperature.

4. The slides were removed from the humidity chamber and immersed in a 1% (w/v) solution of ammonium hydroxide and incubated for 5 minutes with gentle shaking.

5. The slides were then transferred to a solution of 0.1% (w/v) sodium dodecyl sulphate and incubated for 5 min with gentle shaking.

6. The slides were briefly rinsed in Milli-Q ddH2O (Miili-Q plus 185 purification system) at room temperature and then placed in 95ºC Milli-Q ddH2O for 2 minutes to completely denature the bound DNA elements resulting to single-stranded strand-specific array probes.

7. The slides were transferred to ice-cold Milli-Q ddH2O and then briefly rinsed two times in Milli-Q ddH2O at room temperature.

8. The slides were dried by spinning at 180 g for 5 min.

9. The slides were stored in a slide box and kept at room temperature in a cool dry place until used.

10. The final slide consists of 24 blocks (19 x 20) and a total of 7832 probes.

2.2.10 Microarray hybridization

Fluorescent labelling was performed using a modified Bioprime labelling kit in a 130.5 μl reaction (topped up with distilled water) containing 100 ng DNA, 15 μl dNTP mix (2 mM dATP, 2 mM dTTP, 2 mM dGTP, and 0.5 mM dCTP), and 1.5 μl Cy5/Cy3 dCTP (1mM). The reactions were purified using Micro-spin G50 columns (Pharmacia-Amersham) in accordance with the manufacturer's instructions. Reference and test samples were combined and precipitated with 55 μl of 3 M sodium acetate (pH 5.2) in 2.5 volumes of ethanol with 90 μg human C_ot1 DNA. The DNA pellet was resuspended in hybridization buffer (see Materials) containing 300 μg yeast tRNA. Hybridization was performed for 24 hours at 37°C on a MAUI hybridization platform. Finally, the arrays were washed serially in wash solution 1 for 5 min at room temperature, in wash solution 1 for 5 min at 60°C, four times in wash solution 2 for 20 min at room temperature, in wash solution 3 for 10 min at room temperature and finally in HPLC water for 10 min at room temperature. Subsequently the arrays were dried (by centrifugation – 3 min at 800 rpm) and stored in the dark.

2.2.11 Microarray Scanning

Microarrays were scanned using a ScanArray Express HT scanner (PerkinElmer) as follows:

1. Cy3 and Cy5 images at 5 μm resolution were acquired using the ScanArray 4000 confocal laser-based scanner (Perkin-Elmer) at laser power of 100% and a photo multiplier tube (PMT) value of 75% and 70% respectively. All arrays were scanned using the same parameters to avoid introducing another variable as part of the scanning process.

2. The software ScanArray Express (Perkin Elmer) was used to quantify the fluorescent intensities of the spots using the fixed circle quantisation and the TOTAL normalization method. This software can automatically locate the spot position on the

scanned image of the array to obtain the signal intensity values. Mean intensity ratios (intensity-background) were reported for each spot representing an array element.

2.2.12 Microarray Data Analysis

For each sample we analysed two biological replicates. All hybridizations were performed with fluorochrome-reversed pairs of two-colour labelled probes (two dye swaps as technical replicates). For the purpose of this analysis I treated the forward and reverse probes as replicates. Hence, for each sample tested, I obtained 16 measurements derived from quadruplicate spots on 4 array hybridizations (two biological replicates plus dye swaps). Fluorescence intensities were determined using the ScanArray Express software (PerkinElmer). Fusion of dye-swap and biological replicate results and subsequent analyses were performed using R packages from Bioconductor (Gentleman et al., 2004). For each probe, log-ratios were normalised within arrays using a Local Linear Regression (loess) which is efficient in removing dye effects (Smyth and Speed, 2003) and average intensities were normalised between arrays (Yang, 2003) leaving previously normalised ratios unchanged. Dye-swapped samples and biological replicates were defined in a design matrix where rows represent samples (observations) and columns represent effects of interests (parameters). Subsequent analyses were performed according to the design matrix by fitting a gene-wise generalised linear model to log-ratios with the generalised least squares method. This analysis takes advantage of the correlation structure arising from the four duplicated spots (Smyth, 2005) which is expected to be constant. Finally, ranking of the features according to their evidence of discrepancy between effects as defined in the design matrix, was performed by using empirical Bayes methods (Smyth, 2004) where moderated t-statistics test each individual effect equal to zero. Estimated p-values were subjected to multiple testing by using the False Discovery Rate (FDR) method (Benjamini, 1995). A threshold of p-value < 0.001 was used.

This analysis was performed by Dr. Gregory Lebfevre at the Wellcome Trust Sanger Institute.

2.2.13 Identification of genomic features of DMRs

I used the Application Programme Interface (API) to extract the features that are in the Ensembl functional genomics dataset. The whole chromosome 6 was scanned using a 2 kb sliding window in 1 kb steps. For each window, I counted the number of each type of feature within the bounds of the window. This way, a discrete probability distribution was generated, which, for a randomly selected window, determines how likely is to observe a certain number of features.

Windows that overlap an assembly gap were ignored, as this would bias the results.

For each DMR and for feature type, I used the feature count and the probability distribution to calculate:

1. The probability that a random window of that size would have exactly that number of features.

2. The probability that a random window of that size would have more or the same number of features (the right-hand tail of the distribution: if this value is small, it would suggest that the feature is enriched); 95% confidence interval was used.

3. The probability that a random window of that size would have less or the same number of the feature (the left-hand tail of the distribution: if this value is small, it would suggest that the feature is depleted); 95% confidence interval was used.

It should be noted that the probability distribution was generated for a 2 kb window, but DMRs were not exactly 2 kb, so some scaling was done to allow for this difference, i.e. if a DMR was 4 kb and had 6 features, then for a 2 kb window it would be scaled to 3 features.

This analysis was performed by Dr. Stephen Rice at the Wellcome Trust Sanger Institute.