https://theses.gla.ac.uk/

Theses Digitisation:

https://www.gla.ac.uk/myglasgow/research/enlighten/theses/digitisation/

This is a digitised version of the original print thesis.

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses
https://theses.gla.ac.uk/
research-enlighten@glasgow.ac.uk

University of Glasgow

Division of Cancer Sciences

and Molecular Pathology

Methylation of the Heparan Sulphate D-Glycosaminyl-3-O-Sulphotransferase Gene Family in a Cohort of Breast Carcinomas

Craig Peter Charles Dick

Thesis submitted for the degree of Doctor of Philosophy

March 2006

ProQuest Number: 10391102

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.

ProQuest 10391102

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.
This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

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

GLASGOW
UNIVERSITY
LIBRARY:

Abstract

Background. Aberrant methylation of CpG islands is associated with down-regulation or silencing of the associated gene, and it is therefore an alternative mechanism to DNA mutation for silencing tumour suppressor genes in cancer. **Methods.** Using 'ICEAMP', a technique that utilises a methyl-binding-domain column to isolate methylated DNA, DNA pooled from 36 ductal invasive carcinomas was screened for aberrant CpG island methylation. This screen did not identify any novel CpG island methylation, however, previous screens of pooled DNA from 128 breast carcinomas had identified methylation at the CpG islands of the genes for heparan sulphate 3-O-sulphotransferase 3B (HS3ST3B) and heparan sulphate 3-O-sulphotransferase 1 (HS3ST1). Literature and database interrogation identified four other HS3ST genes, one of which does not have a CpG island (HS3ST5) and another (HS3ST2) that published data indicated was methylated in a high proportion of breast, pancreatic, and colonic cancers. Using MS-PCR a cohort of 80 breast tumours was analysed for methylation at six different CpG islands associated with HS3ST genes. HS3ST3B has two separate CpG islands, one associated with exon 1 (HS3ST3BE1) and one lying 3kb upstream from the promoter (HS3ST3B). **Results.** HS3ST1 and HS3ST3A showed no CpG island methylation. HS3ST2 was methylated in 45/80 cases, HS3ST4 in 34/80 cases, HS3ST3B in 30/80 cases and HS3ST3BE1 in 15/80 cases. MSPCR sequencing confirmed dense methylation of both HS3ST2 and HS3ST4 CpG islands in 10/10 cases. Using a cell culture system and RTPCR, transcriptional silence of HS3ST2 and HS3ST4 was found to correlate with methylation of the respective gene. Demethylation of the CpG island using 5-aza-2-deoxycytidine lead to re-expression of the associated gene. Statistical analysis showed there were significant correlations between HS3ST2 methylation status and grade ($p=0.015$), HS3ST3B methylation and lymphovascular invasion ($p=0.028$), and number of loci methylated and grade ($p=0.04$). **Conclusions.** Diverse patterns of promoter methylation involving the HS3ST genes are observed in breast cancer. Some genes in this family are frequently methylated and downregulated while two of the family, HS3ST1 and HS3ST3A, are never methylated. Individual HS3ST enzymes show substrate specificity for oligomeric sequences in heparan sulphate glycosaminoglycan chains, and 3-O-sulphation of heparan sulphate constitutes <5% of the sulfuryl groups in a heparan sulphate molecule, this means it may be suited to act as a modulator of cell surface biological processes. 3-O-sulphation is essential for antithrombin IIIa functioning, endocytosis of HSV-1, and potentiates bFGF signal transduction. The data presented suggest that 3-O-sulphation of heparan sulphate could play a role in human breast cancer. Further work is required to determine the biological significance of the dissimilarity in methylation of the different family members.

Preface

This thesis has been written with two people in mind. One is the person who will have time and the inclination to read the document as a whole. The other is someone without as much free time who requires an overview of the work. With the latter in mind a general discussion and overview section is included at the start of the thesis.

Acknowledgements

I would like to thank Breakthrough Breast Cancer for funding this PhD and also Barry Gusterson for his role in writing the grant and also for his help towards then end of the project when I needed further funding to pay part of my PhD fees. I am much obliged.

James and I became friends through the course of this PhD and I think that probably says more than enough. Thank you James.

Graeme Brock unfortunately had to leave for the USA and I think the UK lost a very enthusiastic, keen, dedicated and clever researcher. I hope he makes it back, with the family and a large grant. His ideas are all through this thesis and anything I learned in the lab I learned from him – his moto 'if you are going to cook it cook it right'. Thank you Graeme.

Strangely I'd also like to thank my examiners – the end result is better for their criticisms.

If I was only allowed to thank one person it would be Fiona, my wife, it would take a thesis to detail all that she has done and put up with, not to mention producing Peter who has made us happier than we thought possible.

Parents have a long lasting effect. At the end of this thesis I had corrections and the thought of the task at hand made me want to role into a ball and turn all the lights out. A few words of encouragement from my mum and dad and I had them completed in 2 weeks, despite a 4 month deadline. That is the effect that my mum and dad have always had on me, and I hope I pass it on to my own family because it is a debt I will never be able to repay.

I may be the first person to thank my mother in law, Mhari, who helped and supported even when we should have been helping and supporting her.

And so to the end.....

Author's Declaration

List of Abbreviations

°C	Degrees Celsius
[α-^{32}P]dCTP	α-^{32}P-labelled 2'-deoxycytidine-5'-triphosphate
λ	Wavelength
μ	Micro (10^{-6})
6-4PP	Pyrimidine-pyrimidone 6-4 photoproduct
8-oxoG	7,8-Dihydro-2'-deoxyguanosine
AP	Abasic
APS	Ammonium persulphate
AraC	1-ß-D-arabinofuranosyl-cytosine
b	Base
BER	Base excision repair
bp	Base pair
BrdU	5'-Bromo-3-deoxyuridine
BSA	Bovine serum albumin
CAG	Trinucleotide of cytosine, adenosine and guanine
CCTG	Tetranucleotide of cytosine, cytosine, thymine and guanine
CDK	Cyclin dependent kinase
cDNA	Complementary deoxyribonucleic acid
CGG	Trinucleotide of cytosine, guanine and guanine
Ci	Curie
CPD	Cyclobutane-pyrimidine dimer
cpm	Counts per minute
cps	Counts per second
CTG	Trinucleotide of cytosine, thymine and guanine
CUG	Trinucleotide of cytosine, uracil and guanine
CUG-BP	CUG-binding protein
dATP	2'-Deoxyadenosine-5'-triphosphate
dCTP	2'-Deoxycytidine-5'-triphosphate
DEPC	Diethylpyrocarbonate
dGTP	2'-deoxyguanosine-5'-triphosphate
DMEM	Dulbecco's modified Eagle medium
DMPK	Dystrophia myotonica protein kinase
DMSO	Dymethylsulphoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotidetriphosphate
DTT	Dithiotheitol
dTTP	2'-Deoxythymidine-5'-triphosphate
E. coli	*Escherichia coli*
EBV	Epstein-Barr virus
EDTA	Ethylenediaminetetra-acetic acid
EtBr	Ethidium bromide
EXP	Triplet repeat expansion proteins
FBS	Foetal bovine serum
FEN1	Flap structure-specific endonuclease 1
FITC	Fluoresceine isothiocyanate isomer I
g	Gravity acceleration
g	Gram
GAA	Trinucleotide of guanine, adenine and adenine
HD	Huntington disease
HNPCC	Hereditary non-polyposis colorectal cancer
HS3ST	Heparan sulphate D-glycosaminyl-3-O-sulphotransferase
IPTG	Isopropylthyo-ß-D-galactoside

k	Kilo (10^3)
KLHL1	Kelch-like 1
l	Litre
LB	Luria Bertani
LBCL	Lymphoblastoid cell lines
m	Milli (10^{-3})
M	Molar
MDR	Multidrug resistance
MED1	Methyl-CpG binding endonuclease 1
MLH	MutL homologue
MMR	Mismatch repair
MOPS	3-(*N*-Morpholino)-propanesulphonic
mRNA	Messenger ribonucleic acid
MSH	MutS homologue
MSPCR	Methylation Specific PCR
mUSF	Mouse upstream stimulatory factor
MW	Molecular weight
N	Nucleotide of adenine, cytosine, guanine or thymidine.
n	Nano (10^{-9})
neo	Neomycin
NER	Nucleotide excision repair
NMR	Nuclear magnetic resonance
NSAID	Nonsteroidal anti-inflammatory drug
OD	Optical density
OGG	7,8-Dihydro-2'-deoxiguanosine DNA glycosylase
p	Pico (10^{-12})
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCNA	Proliferating-cell nuclear antigen
PCR	Polymerase chain reaction
PIKK	Phosphoinositide 3-kinase related kinase
PKR	Double-stranded RNA-activated protein kinase
PMS	Post-meiotic segregation
PMSF	Phenylmethylsulphonyl fluoride
PPP2R2B	Protein phosphatase 2A, regulatory subunit ß
RNA	Ribonucleic acid
ROS	Reactive oxygen species
rRNA	Ribosomal ribonucleic acid
RSHL	Radial spokehead-like
S. cerevisiae	*Saccharomyces cerevisiae*
SBMA	Spinal and bulbar muscular atrophy
SCA	Spinocerebellar ataxia
S-DNA	Slipped-stranded deoxyribonucleic acid
SDS	Sodium dodecyl sulphate
SI-DNA	Slipped-stranded intermediate deoxyribonucleic acid
SIX5	Sine oculis related homeobox 5
SP-PCR	Small pool polymerase chain reaction
SV40	Simian virus 40
TCR	Transcription-coupled repair
TEMED	*NNN'N'*-Tetramethylethylenediamine
Tris	Tris(hydroxymethyl)amino methane
TRITC	Tetramethylrhodamine isothiocyanate
U	Unit
UTR	Untranslated region

UV	Ultraviolet
v	Volume
V	Volt
w	Weight
X-gal	5-Bromo-4-chloro-3-indolyl-ß-D-galactoside

1 General Discussion and Overview

This project was initiated in 2002. There was a considerable body of work at that time to support the idea that DNA methylation at CpG islands was critical for the silencing of unmutated tumour suppressor genes in cancer. This was of particular importance as therapeutic reversal of methylation, using DNA methyltransferase inhibitors such as 5-azacytidine, had been demonstrated in cell culture. It had also been documented that tumours differed in their methylation patterns, and the number of abnormally methylated genes in a tumour type was high, estimated to be 600 genes on average (Costello, Fruhwald et al. 2000). Although many tumour suppressor genes, identified through familial studies and loss of heterozygosity analysis, had been shown to have a role in familial forms of cancer it was not known if methylation events in sporadic cancers would frequently include these familial genes. This relatively novel field therefore offered considerable potential.

With this in mind, there were three aims for this project:

1. Isolate novel CpG islands methylated in breast cancer, but not in normal breast tissue.

2. Develop a 'high-throughput' technique to assess the methylation status of these CpG islands.

3. Use this technique to measure the frequency of methylation at the loci in a larger cohort of breast cancers and correlate this information with clinicopathological data.

As with all projects like this there were other ancillary aims. The idea that certain genes are 'turned off' by methylation at specific points in carcinogenesis was appealing and it was hoped that the identity of some of these genes would be discovered. In order to do this it was the intention to collect pre-invasive cancer tissue and also samples containing atypical

hyperplasias in order to compare their methylation profiles with known invasive cancers.

It was possible to tackle these three primary aims as there was access to a technique for isolating methylated DNA, and there was ethical approval to collect a cohort of breast tumours and access a tumour database located at Glasgow Royal Infirmary.

The technique for isolating methylated DNA uses the methyl binding domain of the Rat MeCP2 protein and had previously been used by Professor Adrian Bird and Dr Sally Cross in Edinburgh to isolate unmethylated CpG islands. Latterly Dr Graeme Brock developed a subtractive hybridisation step. This additional step isolates novel CpG islands methylated in breast cancer, but not in normal breast tissue. A CpG island isolated in this fashion, Heparan Sulphate 3-O-sulphotransferase *(HS3ST3B)*, acts as a molecular marker that differentiates between different grades of both ovarian and breast carcinomas (both in collaboration with Dr Tim Huang, then at the University of Missouri). The technique therefore appeared to be robust and at least one of the sequences isolated correlated with a clinicopathological marker

At the start of this project approximately 20 CpG islands had already been isolated in the collaborative studies undertaken by Dr Brock, using pooled samples of DNA from 155 different breast carcinomas contributed by Dr Huang. Pooled DNA was used as it was reasoned this would maximise the likelihood of identifying frequently methylated loci of potential importance in the genesis of breast cancer.

The first step in this project was to re-establish the methyl binding domain column facility. This was necessary as the MBD columns, previously constructed by Dr Brock, have a limited lifetime. Several columns were

constructed, using the same methods established by Professor Bird and Dr Cross, as subsequently modified by Dr Brock in his previous work.

Despite strict adherence to the established protocols in making the columns, very considerable difficulty was experienced isolating CpG islands. This phase of laboratory work continued over a period of about eight months, during which time approximately ten different MBD columns were constructed. Each column took not less than a week to manufacture. The failure of each column to isolate methylated CpG islands from the pooled breast tumour DNA lead to repeated cycles of discarding that column and building another one. The presumption was that the columns were somehow not functioning, but in retrospect this seems unlikely to have been the case. Control experiments using a plasmid containing an HMLH1 insert, which could be prepared in methylated, hemi-methylated and unmethylated forms did show appropriate binding to the columns.

Why, then, was it so difficult to isolate CpG islands from the pooled tumour DNA samples, which had been passed over the columns? One theoretical possibility is that the pooled DNA samples being worked on did not contain significant numbers of methylated CpG loci. As an absolute statement, this appears unlikely, given the known high frequency of epigenetic modification of DNA at many different loci in many different cancers, including breast cancers.

Another possibility, which may be more likely, is that although DNA methylated at CpG islands was indeed present in the pooled DNA samples, it was not present in sufficiently high proportion relative to unmethylated DNA. However, given that Dr Brock had been successful in identifying CpG islands using the same approach in pooled breast carcinoma DNA provided by Dr Huang, this explanation is not compelling. Another issue that only became apparent after the column phase of the project was abandoned was the nature of the pooled DNA that was being screened for aberrant CpG island methylation. Previously

the pooled tumour DNA had been from 155 different sources, and the pooled DNA used in this project was only from 36 different sources. This issue became apparent after discussing the nature of the samples that had been kindly gifted by Dr Huang. Unfortunately, there had been some miscommunication over the exact nature of this pooled breast cancer DNA. It may be that this had a bearing on the lack of success.

One other possible significant difference in the experimental protocol concerns the use of an HPLC system in previous work by Dr Brock, as opposed to the open-column (non-pressurised) system, which was adopted in the laboratory for this work. This change was adopted to avoid the very protracted protocols (lasting 2-3 days) necessitated by the HPLC system. In addition the lab did not own an HPLC system and the experiment would have been subject to another laboratory's assent. In order to create a library, it could occupy 2, or more weeks to complete one experiment. Although plasmid separation experiments showed that the open system did perform comparably to the HPLC system, the possibility that some degradation of performance may have been associated with this change cannot, ultimately, be excluded.

After isolating tumour DNA fractions in which there ought to have been enrichment for methylated DNA, including methylated CpG islands, and carrying out subtractive hybridisation with methylated DNA from normal breast tissue, the isolated sequences were cloned in *E. coli* and the colonies were hybridised with COT-1 placental DNA to remove repetitive elements. Remaining sequences were then digested with HpaII and HhaI to determine their GC content and size.

As CpG islands are in the size range 500-1200 base pairs, it was anticipated that successful isolations would be in this size range. The reality was that the inserts were almost never > 200 base pairs long, and could not, therefore, represent intact CpG islands.

About 90% of the isolates recovered were eliminated by COT-1 DNA hybridisation, COT-1 DNA hybridises repetitive CpG rich sequences that are not CpG islands, and these clones therefore likely to have been repetitive elements. This finding is was somewhat consistent with the expected ability of the MBD columns to isolate methylated DNA, as these repetitive elements are known to be densely methylated.

A question that has to be asked is whether the purification steps intended to isolate the CpG islands were in fact removing them. This seems unlikely because the screening process was both positively and negatively controlled for by the use of known clones containing repetitive elements and a known CpG island (BRCA1). These control experiments gave the expected results; if they did not (and this only happened once), that result was discarded. In total over 3,500 colonies were screened.

With hindsight, it is possible that greater success in the isolation of CpG islands from pooled carcinoma specimens might have been obtained with the HPLC-based column, but at the same time there was no reason to doubt the efficacy of the open system. It is also possible that there might have had more success by reusing some of the DNA samples provided by Dr Huang, but again there was no reason not to proceed with the samples to which there was access, and an independent data set would have had advantages.

After 8 months of relatively fruitless effort it was time to decide what to do next. Further effort could be invested in trying to get the column and screening protocols working, or change direction.

During the same period described above, there was work taking place in conjunction with Dr Graeme Brock on screening the subtractive hybridisation libraries previously cloned using the HPLC system and the DNA containing 155 pooled breast tumours.

In the course of this work, one of these libraries proved to be particularly rich in CpG islands. Two clones from this rich library were identified, by sequencing, to belong to the same gene family, viz. the Heparan Sulphate D-glycosaminyl-3-O-sulphotransferase (HS3ST) family. Review of the literature showed that frequent promoter methylation of a third member of this family, HS3ST2, had been described in a variety of tumours, including breast cancer.

At this point searches of the Human Genome database revealed that there were three other members of the HS3ST family (six in total), all possessing CpG islands, with the exception of HS3ST5. There was data pointing towards this family of extracellular matrix modifiers as being possibly involved in breast cancer. There was also published data showing that one of them, HS3ST2, was very frequently methylated, and silenced, in breast cancer.

These genes were also interesting in that the existence of six different genes, all encoding enzymes that carry out exactly the same chemical reaction, but have slightly different substrate specificities, suggests that there is a regulatory function associated with 3-O-sulphation. This idea is further supported by the fact that 3-O-sulphation of heparan sulphates amounts to <5% of all sulphation modifications of heparan sulphates. Additionally, it is known that 3-O-sulphation of heparan sulphate is required for dimerisation of FGF1 receptor on the cell surface and subsequent FGF1-R mediated signal transduction, suggesting a possible mechanistic link with human breast cancer. HS3ST1 activity is required for the activation of antithrombin IIIa and may therefore have an indirect role in tumour coagulation, metastasis and angiogenesis. Drosophila melanogaster Hs3st-B is a component of the Notch pathway. Reduction of Hs3st-B function by RNA interference has been shown to compromise Notch signalling, producing neurogenic phenotypes (Kamimura, Rhodes et al. 2004). It is also known that levels of Notch protein on the cell surface in this situation are markedly decreased suggesting Hs3st-B is involved in Notch signalling by affecting stability or intracellular trafficking of Notch

protein. In light of this information it was surmised that 3-O-sulphation of heparan sulphates may be of importance in breast cancer.

It was therefore decided to refocus by adopting the HS3ST gene family as the prime target of investigation, with the hypothesis that this was a group of putative Tumour Suppressor Genes, suggested by the high frequency of HS3ST2 methylation described in breast cancer. The aim was to look at them as a group and try to determine possible interactions between them.

All five of the HS3ST genes, HS3ST1, HS3ST2, HS3ST3A, HS3ST3B and HS3ST4, with CpG islands were examined in a series of 80 breast cancers for which there was fresh, unfixed, frozen DNA available. This was an essentially unselected, representative, series of breast cancers. The technique of methylation-specific PCR (MSPCR) was used in all cases.

Successful MSPCR analysis was achieved at all loci for all cases. All analyses were conducted in triplicate. In all but five of the experiments carried out the 3 analyses were completely concordant. In the 5 cases where there were discordant analyses experiments were repeated until a clear analysis was achieved.

It was found that there was indeed a high frequency of CpG island methylation at the HS3ST2 locus (45/80) and also at the HS3ST4 (34/80) locus in breast cancer, but methylation was never found at either locus in normal breast epithelial tissue. HS3ST1 and HS3ST3A were, in contrast, never methylated in breast cancer. This is of interest as HS3ST1 activity is an absolute requirement for antithrombin IIIa activity.

HS3ST3B has two separate CpG islands, one (HS3ST3B) upstream of the promoter region and one (HS3ST3B-E1) spanning the start of transcription and exon 1. The upstream promoter CpG island (HS3ST3B) was methylated in 30/80 carcinomas and the other (HS3ST3B-E1) in 15/80 carcinomas.

These data pointed to the possibility of HS3ST gene silencing as a significant event in breast cancer. It is believed that gene silencing by promoter region methylation is associated with dense methylation of CpG dinucleotides in the relevant CpG island. MSPCR alone did not allow this to be inferred, as the methylation specific primers only recognise a small number of CpG dinucleotides.

The next step therefore was to undertake sequencing following MSPCR in a representative subsets of cases. Where this was carried out, the intervening CpGs were shown to be completely (i.e. densely) methylated, as expected in the case of a gene silenced by methylation. A total of 20 cases were MSPCR sequenced for HS3ST2 and HS3ST4 (10 cases for each CpG island), all with the same result.

Although this was strong evidence for gene silencing, the relationship between promoter methylation and transcription for HS3ST2 and HS3ST4 was examined and using cell culture it was sought to determine if reversal of methylation *in-vitro* could lead to re-expression of these genes. The possibility of therapeutic reversal of abnormal methylation was one of the initial attractions of this area. These experiments were carried out in cell culture using the breast carcinoma cell line MCF7 and the ovarian cell line A2780 grown in the presence of 2-deoxy-5-azacytidine. These cell lines were chosen from a larger panel that was screened for the presence of HS3ST gene methylation.

It was shown that there is expression of the associated mRNA in those cell lines in which there is no HS3ST2 or HS3ST4 gene methylation, and that there is gene silencing in cell lines in which there is methylation of the respective HS3ST gene. Demethylation of a CpG island using 2-deoxy 5 azacytidine was associated with expression of the respective gene

It was also decided to look at associations between HS3ST gene methylation status and conventional pathological features of tumours examined. This was carried in light of published data suggesting an association between HS3ST2 methylation status and tumour grade and it was hoped that it might lead to clues regarding the pathology of tumours.

Statistical analysis was carried out on the MSPCR data: There is a positive correlation between methylation at the HS3ST3BE1 CpG island and methylation at HS3ST2 (p=0.009), HS3ST3B (p=0.04) and HS3ST4 (p=0.009). There is a correlation between HS3ST2 methylation and grade (p=0.015) and strangely there is a very strong correlation between lack of HS3ST2 methylation and nodal status unknown (p=0.009) and NPI status unknown (p=0.002). HS3ST3B methylation is correlated with lymphovascular invasion (p=0.028). HS3ST3BE1 methylation is correlated with age (p=0.04), NPI (p=0.047) and size (p=0.04). The number of loci methylated correlates with grade (p=0.04). The correlations that show a trend are as follows: there is a trend for number of loci methylated to be correlated with ER/PR negativity (p=0.054), NPI (p=0.06) and lymphovascular invasion (p=0.096).

The data presented in this thesis suggest a role for the HS3ST family genes in breast carcinogenesis.

However, these data do not, as yet, give any mechanistic insights into what the actual role of the HS3ST gene family might be in breast cancer, although what is known of their biochemical activities does allow for some speculation on possible roles. In particular, the ability of these gene products to modify the extracellular matrix in the vicinity of tumour cells could allow them a role in the emergence of invasive behaviour via interactions with the extracellular matrix. Another possibility is the potential for modification of cell signalling, either potentiating the effect of growth-stimulating signals or possibly by down regulating the effect of inhibitory or even cell-death signalling from the

extracellular matrix. Other possibilities include interactions with angiogenesis; the known interaction with antithrombin III may be relevant in this context.

It would be most helpful for further research if there were reagents to examine the expression of these gene products, either by in situ hybridisation or in the form of monoclonal antibodies which could be exploited through immunohistochemistry. This would allow the expression to be looked at on a larger scale, e.g. in tissue arrays, allowing a much more robust statistical analysis of their relevance in respect of classical pathological parameters. It would also be important to determine to what extent these genes are expressed by the epithelial, or stromal elements of the tumour; and in addition, whether and where they are expressed. Unfortunately there is no data indicating these antibodies have been raised and the high degree of homology between the enzymes may make this a very difficult task.

Clearly, it would be helpful if more data was available about substrate specificity and areas of overlap and redundancy between family members. This knowledge could enable a prediction of the ability of a complement of HS3ST enzymes to synthesise specific binding sites for growth factors or sites necessary for growth factor - receptor interactions. The context of this ability to produce binding sites is also important - a binding site on an ECM GAG may sequester a growth factor in the ECM and prevent it from interacting with the receptor on the cell surface and the same binding site on the correct cell surface proteoglycan may lead to growth factor receptor interaction and potentiation of growth factor signalling. HS3ST enzymes are the last modification to act on heparan sulphates. They are constructed and modified in the endoplasmic reticulum and they are therefore dependent on the previous modifications to generate a substrate for them. It would therefore be necessary to know the cell's complement of all the enzymes that modify heparan sulphate in order to build such a model.

The finding that methylation and down regulation of specific members of the HS3ST family takes place in breast cancer highlights the importance of the extracellular matrix in the carcinogenic process. Heparan sulphates and their modification have specific biological roles that are being modified by the cell through methylation and silencing of genes. The precise nature and the subtleties of the system are as yet not understood, however the advent of more powerful polysaccharide sequencing technology may offer the next advancement in this area.

2 Introduction

Breast cancer is the most common female malignancy and, after lung cancer, the second most common cause of female deaths due to a malignancy (McPherson, Steel et al. 2000). Based on current lifetime incidence of the disease, one in nine females in the United Kingdom will eventually develop breast cancer. Each year over 40,000 women in the UK are diagnosed with breast cancer, and one-third of them will die of it (CRUK 2005).

2.1 The Normal Breast

2.1.1 Breast Development

The mammary glands are often regarded as modified sweat glands; they appear in a rudimentary form around the fifth week of life when the mammary line thickens in the thoracic portion of the embryo to form the mammary ridge(Rosai 2004). Subsequently this ridge invaginates, buds and then grows down into the mesodermal tissue beneath (Rosai 2004). A small pit is left behind at the site of this event; just prior to birth this rises and forms the nipple. Each bud from the ridge develops into one lobe of the future breast, dividing to form arborising ducts and ultimately individual lobules (Ross 1984). At eight months gestation the stem of these buds canalises, under the influence of oestrogen, forming the lactiferous ducts and the surrounding mesoderm develops into the interlobar connective septa. Adipocytes infiltrate the breast around month five (Rosen 2004). The structure of the breast from this point on is not greatly altered until menarche when oestrogen and progesterone induce duct elongation, further branching morphogenesis, and ultimately proliferation of the terminal ducts to form the glandular lobules(Vogel, Georgiade et al. 1981).

2.1.2 The Adult Breast

The adult breast, although it varies in size, has certain constant features. It lies vertically from the second to the sixth ribs and extends transversely from the border of the sternum to the mid axillary line (J.S.P.Lumley). The muscles underlying the breast, pectoralis major, serratus anterior and the external oblique, are covered by the superficial fascia which forms the deep border of the breast (S.Chummy 2006). The breast extends into the axilla and the axillary tail may extend as far laterally as the posterior axillary line (S.Chummy 2006). The nipple is found in the centre of the breast and its position in relation to the underlying structures varies with the volume of breast tissue present(J.S.P.Lumley). The areola, a circular pigmented area surrounding the nipple, harbours the prominent sebaceous glands of Montgomery and a network of smooth muscle fibres beneath that contract during suckling (Ross 1984; Rosen 2004).

Each 'tree' of breast epithelium is surrounded by a basement membrane (Yoon, Herman et al. 2002). The nipple and the mouth of the ducts are lined by stratified squamous epithelium, this changes abruptly to a double layer of cuboidal epithelium/myoepithelium in the distal duct(Rosai 2004). Peripherally, the outer myoepithelial cells lose their cuboidal morphology, becoming flattened and elongated, but continue as an outer layer into the terminal duct lobular units(Rosen 2004).

2.1.3 Breast Changes of the Menstrual Cycle

There is a menstrual cycle of the breast, as well as the endometrium. It is a cycle dependent on the hypothalamic-pituitary axis: the hypothalamus releases gonadotrophin releasing hormone (GnRH), stimulating the anterior pituitary to produce follicle stimulating hormone (FSH) and leuteinising hormone (LH).

These hormones drive oocyte maturation and ovulation. This part of the cycle, up to the point of ovulation, is termed the follicular phase and breast lobules are relatively quiescent at this stage (Vogel, Georgiade et al. 1981), unlike the endometrium which proliferates during this part of the cycle (Ramakrishnan, Khan et al. 2002). During the luteal phase, following ovulation, FSH stimulates granulosa cells of the Graafian follicle to synthesise oestrogen and progesterone from precursors produced by adjacent thecal cells. The cuboidal lining cells of the terminal duct lobular unit (TDLU) respond to oestrogen by undergoing mitosis, but it is not clear that those lining the ducts proximal to the TDLU are affected in the same way(Vogel, Georgiade et al. 1981; Longacre and Bartow 1986; Ramakrishnan, Khan et al. 2002). The number of acini in the breast are said to increase as a result of cell proliferation (Nijhawan, Hemachandran et al. 2002); cells become vacuolated and the lobular stroma becomes oedematous(Rosai 2004). The sense of breast fullness experienced by some women in the premenstrual phase of the cycle may be explained by the increased stromal oedema. If fertilisation of the oocyte does not take place oestrogen and progesterone levels fall, mitotic stimulation to the breast epithelium is removed and there is a wave of subsequent apoptosis. Stromal oedema also subsides and there is infiltration of the intralobular stroma by lymphocytes (Longacre and Bartow 1986).

The total number of menstrual cycles a woman undergoes and her risk of developing breast cancer are directly related (MacMahon 2006). Exact reasons for this are not clear however, it may be that proliferation and apoptosis of the breast epithelium are dependent on the hormones of the menstrual cycle, with continuing rounds of cell proliferation promoting the development of precursor lesions from which invasive cancer ultimately emerges. Increased understanding of mammary stem cell niches and homeostatic control of differentiating epithelial cell populations in the breast are also likely to be important for understanding of these processes.

2.1.4 Lactation-Associated Breast Changes

During pregnancy the breast is prepared for lactation by a sustained period of cell proliferation resulting in a substantial increase in the bulk of the glandular tissue (Battersby and Anderson 1989). Overall gland size is increased with a relative reduction in the parenchymal elements (Joshi, Ellis et al. 1986). Following parturition the lactating state is established: this is the fully differentiated state of the human mammary gland. Subsequent suckling removes milk and an oxytocin-dependent feedback loop maintains the lactating state(Joshi, Ellis et al. 1986). Once suckling ceases a wave of involution follows, resulting in loss and remodelling of the glandular tissue with a return ultimately towards the 'resting' state(Joshi, Ellis et al. 1986).

2.1.5 The Postmenopausal Breast

Following the menopause, monthly cycles of stimulation and apoptosis associated with the menstrual cycle are a bsent and the breast undergoes gradual involution (TA 1989; Rosen 2004). The volume fraction of the breast composed of glandular elements decreases and there is a concomitant increase in connective tissue elements, mainly fat, in the postmenopausal breast (TA 1989) In old age the lobular elements of the breast almost disappear entirely ,however, hormone replacement therapy retards this process(TA 1989).

2.2 Breast Cancer

2.2.1 The History of Breast Cancer

The earliest known recording of breast disease comes from a 5000-year-old papyrus, bought in Luxor in 1820 by Edwin Smyth, which describes eight cases of tumours and ulcers of the breast (Weisman 1956). There are sporadic references in historical documents from then on, mainly descriptions of the disease(Rosen 2004). Hippocrates mentions breast cancer briefly in 400 BC, to note that he does not advocate any treatment or medication. However, at this time treatments were taking place: Herodotus states that a Persian physician living in Greece treated and cured King Darius' wife in 425 BC, but the exact surgery or medication is not revealed. Roman physicians performed extensive surgery, and the practice of surgery for breast cancer has continued from Roman times until present day, interspersed with periods where physicians used chemical or physical cautery for fungating lesions (Donegan).

Efforts to understand the pathogenesis of breast cancer were made as time and knowledge progressed. Henri LeDran's (1685-1770) postulate in 1757 that breast cancer originated in the gland itself then spread through lymphatics to the blood was a major advance (HF.LeDran 1757; Cotlar, Dubose et al. 2003). The idea is self evident in modern times, but prior to LeDran, René Descartes (1596-1650) had proposed 'impacted lymph' as the cause of breast cancer. It serves as a reminder that scientific facts are hard won and often run against the tide of popular thinking. Just after this Bernard Peyrilhe (1735-1804) went further and demonstrated the metastatic potential of breast cancer experimentally, aspirating material from a dog's mammary tumour and injecting it intraperitoneally (Co tlar, Dubose et al. 2003). On opening the abdominal cavity a few weeks later, numerous metastatic deposits were observed over the peritoneum.

Surgeons and patients after 1846 benefited greatly as general anaesthesia became available and experiences like that described by the author Fanny Burney (1752-1840), who underwent a mastectomy in 1811, undertaken by the Napoleon's surgeon, the celebrated Baron Larrey (1766-1842), and survived to write about it, were a thing of the past. If there is any doubt about the suffering involved in such an operation Burney depicts the procedure graphically,

"..When the wound was made, and the instrument was withdrawn, the pain seemed undiminished, for the air that suddenly rushed into those delicate parts felt like a mass of minute but sharp and forked poinards, that were tearing the edges of the wound but when again I felt the instrument - describing a curve - cutting against the grain, if I may so say, while the flesh resisted in a manner so forcible as to oppose and tire the hand of the operator, who was forced to change from the right to the left - then indeed I thought I must have expired. I attempted no more to open my eyes which remained hermetically shut...."

Antisepsis (1867) and subsequently asepsis (1871 onwards) became accepted practice, further reducing mortality and morbidity (Cotlar, Dubose et al. 2003). As anaesthesia improved surgical treatment of breast cancer became more aggressive, influenced by centrifugal theories of breast cancer spread. In 1867 Charles Moore at the Middlesex Hospital, and later William Halsted (1852-1922) at Johns Hopkins, described and championed radical mastectomy, advocating *en bloc* removal of pectoralis major and the axillary nodes along with the breast. In the middle of the twentieth century, as operations became even more radical, surgeons audited more thoroughly the effectiveness of a procedure; ultimately this saw the demise of such operations as the super-radical mastectomy, as a mortality rate of 12.5% and no increase in long-term survival sealed its fate (Robinson 1986).

Assessment of surgical effectiveness, combined with the understanding that breast cancer kills through metastasis, not local recurrence, has lead to today's more conservative approach. Surgery is tailored to ensure high-quality local control with very low local recurrence rates, achieved by either wide local excision and radiotherapy or simple mastectomy.

Surgical extirpation of a breast tumour remains the single most important element in breast cancer treatment (Levitt 1985). Radiotherapy is an important supplementary treatment; the effectiveness of wide local excision combined with radiotherapy was shown as long ago as the 1920s by the pioneering studies of Geoffrey Keynes (1887-1982), but its acceptance into routine practice took many years(Veronesi, Luini et al. 1993). Endocrine treatments (oophorectomy was first carried out by Beatson in 1896) and systemic chemotherapy reduce but do not entirely eliminate the risk of systemic relapse(Baum 1998; Rodger 2001).

Despite a large screening programme, access to rapid diagnosis and surgery, adjuvant treatments, and increased understanding of its biology, breast cancer continues to kill nearly 13,000 women every year in the UK (CRUK 2005). In the last 10-15 years there has, however, been a significant reduction in mortality from breast cancer(CRUK 2005), probably due to systematic application of available treatments, i.e. optimal primary surgery and radiotherapy, and appropriate use of chemotherapy and endocrine manipulations which reduce the risk of metastatic relapse. The impact of screening for breast cancer remains a contentious issue (Elmore, Armstrong et al. 2005; Gilbert 2005; Fryback, Stout et al. 2006; Maraqa and Lansdown 2006). It is largely accepted that postmenopausal women benefit, but assessment of the benefits to pre- and peri-menopausal women is difficult, and research is beset with confounding factors.

The realisation that the old fashioned, mutilating radical mastectomy is unnecessary is real progress. Sentinel node biopsy now offers better management of the axilla, adjuvant therapy and treatment of systemic disease prolong survival: these are major advances. Five year survival for breast cancer in Scotland is now 77%, compared to 18% for gastric cancer in women, colon and rectal carcinomas (48% and 50% respectively in females), prostate cancer (60%) and bladder cancer (60% in females). Also, when breast cancer 5 year survival is examined over time, there has been considerable improvement. For 1997-2001 five-year survival is 77%; for 1977-1981 it was 59.6% (NHSScotland 2003).

2.2.2 Clinical Presentation, Treatment and Prognosis of Breast Cancer

Breast cancer is subject to a screening program in most developed countries and as a result many breast cancers diagnosed are asymptomatic, picked up by mammography and confirmed by FNA or biopsy (Elmore, Armstrong et al. 2005). However, many cases also present with symptoms ranging from the skin changes of Paget's disease of the nipple, abnormal secretions, a simple lump, to neglected tumours that fungate and ulcerate (Rosai 2004). The presentation of the disease is therefore varied and depends as much on psychosocial factors and cultural awareness: breast cancer in Egypt often presents at an advanced stage because many patients there have little awareness of breast cancer(El-Monem 2003).

Treatment of breast cancer remains primarily surgical. The type of surgery depends on the site and size of the tumour, breast size and the wishes of the patient. Axillary clearance or sampling is required if there is invasion in order to stage disease. Due to the morbidity associated with axillary surgery, particularly clearance, there is a lot of research underway to determine the need for complete axillary clearance, the value of intraoperative frozen sections on sentinel lymph nodes, and how the lymph nodes should be processed. Sentinel node biopsy(Veronesi, Paganelli et al. 2003) has been shown to have a high specificity (100%) and sensitivity (>91%) for detecting axillary metastasis, and may prove to be the way forward(Fleissig, Fallowfield et al. 2005).

Adjuvant therapies for breast cancer include hormone receptor antagonists like Tamoxifen and the newer aromatase inhibitors, radiotherapy, chemotherapy and more recently anti-Her2 agonists (Cole, Gelber et al. 2001). These treatments are used subject to the hormone receptor status, grade and stage

of tumours. Herceptin/Trastuzumab is a monoclonal antibody specific for the Her2 protein, approved as a treatment for metastatic carcinoma in women with Her2 amplification (Lichtenstein, Berenson et al. 1990). It is often used as a third like treatment after others have failed. Recent results of Herceptin in the adjuvant setting, following standard chemotherapy regimes in patients with non-metastatic disease, show increased disease-free survival after one year of treatment (Piccart-Gebhart, Procter et al. 2005). Over the relatively short period of follow up there was, however, no significant difference in overall survival between those treated with Herceptin and those in the control group.

Nodal status remains the single most important factor in breast cancer prognosis: it is predictive of outcome and recurrence of the disease. However, there are patients with no detectable lymph node involvement, who should be at relatively low risk of recurrence and yet 20-33% of these women will develop recurrent disease after ten year follow up (Camphausen, Jones et al. 1986; Cole, Gelber et al. 2001). Therefore, pathological predictors of outcome in breast cancer remain imperfect.

In breast cancer, prognostic molecular profiles and predictive molecular signatures remain secondary to pathological observations at present. However, there are subgroups of patients, such as those with node negative tumours, in whom molecular profiling may be helpful in predicting clinical behaviour (van 't Veer, Dai et al. 2002). The van't Veer study in 2002 used DNA microarrays in the analysis of 117 primary breast tumours, and applied supervised classification to identify a gene expression signature strongly predictive of a short interval to distant metastases, in lymph node negative tumours. In addition, they established a signature that identified tumours occurring in BRCA1 carriers. The poor prognosis signature includes genes regulating the cell cycle, invasion, metastasis and angiogenesis. However, this study has come under some criticism lately for including patients in the study set that were

also used to build the algorithm (van 't Veer, Dai et al. 2002). The ultimate place therefore of these technologies remains undetermined.

2.2.3 Familial Breast Cancer

Familial cases account for around 5-10% of breast cancer, at most. BRCA1 and BRCA2 gene mutations are found in around 20-40% of these cases (Bishop 1999). Therefore, BRCA1 and BRCA2 germ line mutations are the cause of less than 5% of all breast cancer. Other familial syndromes that can manifest as breast cancer include Li-Fraumeni syndrome, linked to p53 mutations(Sidransky, Tokino et al. 1992); Ataxia telangectasia, associated with mutations in the ATM gene(Radford and Zehnbauer 1996) and Cowden's syndrome, in which 80% of cases have a mutation in the PTEN gene (Lynch, Ostermeyer et al. 1997). These together make up less than 5% of all familial breast cancer (Blamey, Wilson et al. 2000). From these figures two facts can be deduced 1): a large proportion of familial breast cancers are caused by unidentified mutations, probably disparate in character or of low penatrance (or else they would have been detected in the linkage studies that detected the BRCA1 and 2 mutations); and 2) most breast cancers are sporadic, and their aetiology must be explained in terms of the risk factors cited below.

2.2.4 Sporadic Breast Cancer

2.2.4.1 Epidemiology and Risk Factors in Breast Cancer

Epidemiological data, combined with data from animal and *in vitro* models of breast cancer, highlight three areas of importance in the pathogenesis of breast cancer: genetic factors, hormonal influences (such as increased levels or imbalance of endogenous oestrogens) and possible environmental influences. Many of the risk factors for breast cancer (e.g. nulliparity) suggest that an increased number of menstrual cycles confer a higher risk, but there are others, especially geographical variation, that require further explanation.

2.2.4.1.1 Age

The average age at diagnosis with breast cancer is sixty-five (MacMahon 2006). Breast cancer does not occur before the menarche and is rare under the age of twenty-five (Armstrong, Eisen et al. 2000). Incidence of breast cancer increases year on year (Armstrong, Eisen et al. 2000), but the rate at which it increases slows after the menopause and in some countries the incidence plateaus (MacMahon 2006). The change in the slope of the incidence curve at menarche and menopause exposes the hormonal nature of the condition. The incidence curve of colon cancer or gastric cancer when plotted against age does not have the same sigmoid appearance (Garfinkel, Boring et al. 1994; CRUK 2005).

2.2.4.1.2 Sex

One in two hundred cases of breast cancer occur in males in whom the disease tends to be diagnosed at a later stage. Interestingly, the prevalence of Klinefelter's syndrome in male breast cancers is 7.5%, implying a fifty-fold increased risk of breast cancer relative to normal males (Hultborn, Hanson et al. 1997; Weiss, Moysich et al. 2005). Exactly why is unclear.

2.2.4.1.3 Genetic Polymorphisms

As with drug and chemical metabolism, there is considerable inter-individual variability in the conjugation pathways for both oestrogens and catechol oestrogens. Polymorphismsin genes encoding relevant enzymes may define subpopulations of women with higher lifetime exposure to hormone-dependent growth promotion, or to cellular damage from particular oestrogens and/or oestrogen metabolites (Kristensen and Borresen-Dale 2000; Thompson and Ambrosone 2000; Haimov-Kochman, Lavy et al. 2002; Mitrunen and Hirvonen 2003) (Watts, Handel et al. 1992). Such variation could explain a proportion of the cancer susceptibility associated with reproductive effects and hormone exposure. Preliminary studies investigating the role of polymorphic genes encoding enzymes involved in oestrogen biosynthesis (CYP17 (Feigelson, Coetzee et al. 1997; Weston, Pan et al. 1998; Abrahamson, Tworoger et al. 2006; Hu, Xie et al. 2006; Silva, Cabral et al. 2006), CYP19 (Kristensen, Andersen et al. 1998; Olson, Ingle et al. 2006; Ribeiro, de Amorim et al. 2006), and 17-beta-HSD (Mannermaa, Peltoketo et al. 1994)) and conversion of oestrogen metabolites and their by-products (COMT (Bergman-Jungestrom and Wingren 2001), CYP1A1 (Huang, Chern et al. 1999; Hu, Xie et al. 2006), CYP1B1 (Watanabe, Shimada et al. 2000), GSTM1, GSTM3(Cheng, Chen et al. 2005), GSTP1 (Lavigne, Helzlsouer et al. 1997), GSTT1 and MnSOD (Cheng, Chen et al. 2005)) in modulating individual susceptibility to breast cancer have been carried out. Although some of these low-penetrance genes are good candidates for conferring an increased risk of breast cancer, better designed and considerably larger studies than most studies conducted so far will be needed before firm conclusions can be drawn (Mitrunen and Hirvonen 2003).

2.2.4.1.4 Geography

There are major differences in breast cancer incidence between geographically distinct populations; In North America the incidence is 86.3 per 100,000; in China the incidence is only 11.8 per 100,000 (McPherson, Steel et al. 2000). Environmental factors must also be important in breast cancer, as genetically

homogeneous populations living in different parts of the globe have such different rates. However, there is little definitive evidence as to what exactly these environmental factors are. Among immigrants from China to North America the incidence rises to match that of the host country after two generations. This appears to be strong evidence for an environmental effect, possibly due to assimilation of the host country's practices with regard to risk factors for breast cancer, including diet. The reverse population migration analysis has not been carried out, and may not even be possible. The search for single major environmental agents that causes human breast cancer has been unsuccessful, and perhaps strangely, smoking is not an important risk factor (McPherson, Steel et al. 2000). Factors including nutrition and body size are important in what is probably a multifactorial causation.

2.2.4.1.5 Dietary Factors and Obesity

The relationship between obesity and breast cancer risk is different depending on the menopausal status of the female. Premenopausal obesity reduces the chance of breast cancer, and postmenopausal obesity increases the risk of breast cancer (McTiernan, Rajan et al. 2003). This dichotomy might be resolved by analysing the effects of excess adipose tissue in these two different phases of reproductive life.

Peripheral synthesis of oestrogen is dependent on aromatisation of the A ring of C19 androgens, a reaction catalysed by aromatase cytochrome P450 (coded for by the CYP19 gene). The product of this reaction is the phenolic A ring of C18 oestrogens (Lorincz and Sukumar 2006). In obese post menopausal women peripheral adipose tissue deposits are the main sites of oestrogen biosynthesis Grodin (Grodin, Siiteri et al. 1973). In premonopausal women peripheral oestrogen synthesis is not the primary source of the hormone, however it is still synthesised peripherally and excessive oestrogen can inhibit ovulation in premenopausal obese women (Barber, McCarthy et al. 2006).

Anovulatory cycles prevent the formation of the corpus luteum and hence may reduce subsequent proliferation and apoptosis of the breast epithelium (Vogel, Georgiade et al. 1981; Ramakrishnan, Khan et al. 2002). In the case of the premenopausal obese female the net effect is to reduce the number of menstrual cycles to which the breast is exposed. Hence, pre-menopausal obese female risk may be reduced in comparison to a non-obese female. Postmenopausally, peripheral synthesis of oestrogen may expose the breast epithelium to increased levels of this mitogen compared to non-obese females. This is to large extent a simplification of the effect of obesity as a risk factor for developing breast cancer and only covers the proposed effect of increased peripheral oestrogen synthesis in adipose tissue. Other hypotheses linking obesity to breast cancer place more emphasis on increased circulating insulin and insulin like growth factor in the obese; acting as mitogens (Lorincz and Sukumar 2006), and a newer proposal champions the autocrine, paracrine and endocrine functions of adipocytes as major influences on breast epithelial malignancies (Dizdar and Alyamac 2004).

The role of preadipocytes in tumor growth is complicated. A proposed definition of preadipocytes is an immature adipocyte positive for S-100 protein, a marker that is not found in fibroblasts or endothelial cells. Some studies suggest that pre-adipocytes support the growth of breast tumor cells (Chamras, Bagga et al. 1998), while other studies show that preadipocytes are either inhibitory (Johnston, Rondinone et al. 1992) or have no effect (Manabe, Toda et al. 2003), on breast cancer cell growth.

2.2.4.1.6 Infective Agents

Despite the existence of murine mammary transforming virus (MMTV), which causes mammary gland malignancies in mice (Zhou, Chen et al. 2001), no equivalent infective factor has been found in humans (McPherson, Steel et al. 2000).

2.2.4.1.7 Parity and the Menstrual Cycle

Age, early menarche, late menopause, nulliparity and a longer time between menarche and first child are all associated with increased risk of breast cancer (Hartge 2003). During the menstrual cycle there are large fluctuations in the levels of progestogens and oestrogens, which have an effect on gene expression in breast epithelial cells. All of the risk factors cited above, with the exception of age after the menopause, independently increase the number of menstrual cycles that a female undergoes. Therefore, it may be proposed that the increased exposure to this process, or to the hormones that are involved in the cycle, in some way drives carcinogenic processes in breast epithelium.

A complete molecular explanation of the link between the menstrual cycle and breast cancer is not fully established (Hulka, Liu et al. 1994; McTiernan, Rajan et al. 2003). It is known that breast epithelium and stroma express oestrogen and progesterone receptors and several growth factors are released in response to these hormones; some are stimulatory e.g. platelet derived growth factor and endothelium derived growth factor, and some are inhibitory e.g. transforming growth factor beta(Hartge 2003). There are also effects on the intralobular stroma in response to hormone synthesis and withdrawal. The effect during the reproductive years of these hormones, as well as increased epithelial cell proliferation, may create an evolutionary pressure that selects cells on the basis of ability to avoid apoptosis on withdrawal of female sex hormones. Each cycle then represents a round of evolutionary selection, explaining the link between the number of cycles and risk of breast cancer.

2.2.4.1.8 Benign Proliferative Breast Disease

Proliferative, but benign breast diseases are markers of increased risk of developing breast cancer. (Rosen 1993) . Hypotheses as to why this relationship exist include abnormal oestrogen and progesterone metabolism promoting formation of benign entities and cancers independently. In this model, benign

proliferations are markers of abnormal hormone metabolism rather than precursor lesions to cancer (Mitrunen and Hirvonen 2003).

2.2.5 The Pathology of Breast Cancer

The breast is composed of a variety of cell types and the relative proportions of the constituents change depending on the age and hormonal status of the person(Ross 1984; Rosai 2004). Therefore, the breast can be the site of tumours that are derived from fibroblasts, endothelial cells, smooth muscle, adipocytes etc. However, by far the most common tumours in the breast are carcinomas(Ross 1984; CRUK 2005).

Although there is marked variation in the histological phenotype of breast carcinomas from case to case it is thought that these tumours derive from epithelial cells lining the terminal duct lobular unit (Wellings 1980). These cells proliferate and undergo intracellular secretory changes in response to oestrogen and progesterone stimulation, then apoptosis in response to withdrawal of these hormones during the menstrual cycle. Increasingly it is being recognised that stem cell compartments can be identified within normal and neoplastic cell populations and markers characterising such cells are beginning to be identified (Liu, Dontu et al. 2005).

Invasive breast cancer is thought to arise in most cases from an intraepithelial proliferation of cells that have not yet acquired the necessary phenotype to invade the surrounding connective tissue(Buerger, Simon et al. 2000). Depending on the histological appearance of the process it may be labelled as ductal carcinoma in situ (DCIS), lobular carcinoma in situ (LCIS), or unusually one of the rarer subtypes. Less marked abnormalities may be characterised as atypical hyperplasias of ductal or lobular type. Retrospective pathology studies have shown that low-grade forms of DCIS left without treatment ultimately progress to invasive disease in 75% of cases, while 25% do not(Aubele, Werner et al. 2002). This begs the question what is the difference in these two identical histological entities at the molecular level? So far no unique molecular signatures predictive of the invasive phenotype have been identified.

Perhaps in some cases it is the host immune phenotype that is important (DCIS often attracts a brisk inflammatory response).

2.2.6 DNA Methylation and Breast Cancer

2.2.6.1 Introduction

Epigenetic cellular alterations establish heritable features that do not involve change to the DNA sequence. DNA methylation is the commonest epigenetic alteration in the mammalian genome and has a role in cancer biology (Pellise, Castells et al. 2004), viral infections (Zsiros, Jebbink et al. 1999), activity of mobile elements in the genome, somatic gene therapy, cloning, transgenic (Chai, Locke et al. 2001) technology, genomic imprinting, developmental abnormalities, mental health (Strunnikova, Schagdarsurengin et al. 2005) and X chromosome inactivation.

Although there are rare reports of CpA and CpT methylation in the mammalian genome, DNA methylation usually involves covalent attachment of a methyl group to the 5' position of a cytosine nucleotide, and almost always in the context of a 5'-CpG-3' palindrome. The observed frequency (1 in 160 to 1 in 320) of CpG dinucleotides in the human genome is only about 5-10% of the expected frequency of about 1 in 16 dinucleotides (Bird 1980). Uracil formed by deamination of unmethylated cytosine can be removed by uracil glycosylase(Lindahl 1974; Lindahl 1997.), but thymine generated by deamination of 5-methylcytosine, cannot be so removed, and the mutation of 5mC to T is up to 50-fold greater than other transitions. This susceptibility to mutation may have depleted CpG dinucleotides through time. Remaining CpGs are not randomly distributed in the genome, and nor is their methylation. Clusters of methylated CpGs e.g. in repetitive elements are found interspersed with CpG-rich regions known as CpG islands (Bird 1986). These GC rich (60-70%) islands are 200 bp-2,000 bp long, are usually found in association with the promoter or first exon of a gene, and their CpG content approaches the frequency predicted for the genome as a whole (observed:expected ratio >0.5). CpG islands are associated with both housekeeping genes and genes with

tissue-specifc patterns of expression, and of the predicted 35,000 genes in the human genome, approximately half have associated CpG islands.

The transfer of a methyl group to cytosine is catalysed by a DNA methyltransferase enzyme (DNMTs). To date eight DNMTs have been detected: DNMT1, DNMT1b, DNMT1o, DNMT1p, DNMT2, DNMT3a, DNMT3b (plus isoforms) and DNMT3l. DNMT3a and DNMT3b have *de novo* methylation activity, acting on sites where both CpGs in a palindrome are unmethylated. DNMT1, which can also carry out *de novo* methylation, is more efficient at maintenance methylation during DNA replication, transferring a methyl group to the non-methylated cytosine in a hemi-methylated CpG. Mice deficient in either DNMT1, DNMT3a or DNMT3b fail to develop, or die immediately after birth. (Bird and Wolffe 1999). DNA methylation patterns are established early in development. At the morula stage of *Xenopus* embryogenesis a wave of demethylation sweeps over the zygote reducing global CpG methylation by approximately 30%, with subsequent remethylation of the genome prior to implantation(Monk, Boubelik et al. 1987). The control of this process remains unclear at present.(Bird 2002)

In biological systems most reactions can be undone and DNA methylation is no exception. The enzyme 5-methylcytosine glycosylase has demethylating activity; it removes methylated cytosines from DNA leaving deoxyribose, and subsequently local DNA repair enzymes reconstitute the nucleotide by adding an unmethylated cytosine in place of the methylated one. An alternative translational form of the methyl binding domain protein MBD2, named MBD2b also possesses DNA demethylating activity. It is thought to act by hydrolysing 5-methylcytosine to cytosine and methanol. However this result has only been published by one group and other laboratories have failed to replicate these results in mammalian and Xenopus based studies(Ng, Zhang et al. 1999; Wade, Gegonne et al. 1999).

The potential importance of DNA methylation in cancer biology was recognised when CpG island methylation was associated with down-regulation of transcription, and began to be recognised as an alternative to DNA mutation as a means of satisfying the Knudson 'two hit' hypothesis (Jones and Laird 1999). However, the role of methylation is more far reaching than as a pure repressor of transcription.

Tumour DNA is generally globally hypomethylated (Bernardino, Roux et al. 1997; Johanning, Heimburger et al. 2002), and altering methylation enzymes to create mice with hypomethylated DNA can either promote or repress tumourogenesis. For example, multiple intestinal neoplasia (MIN) mice that carry a germline mutation in the mouse APC orthologue are predisposed to the development of intestinal tumours. MIN mice heterozygous for DNMT1 $^{-/+}$ treated with 2-deoxy-5-azacytidine show fewer polyps (Laird, Jackson-Grusby et al. 1995). On an MLH1-deficient background DNMT1 $^{-/+}$ mice again show a reduction in intestinal adenomas, however the risk of lymphoma in these mice is increased (Trinh, Long et al. 2002). Homozygote mice with a crippled form of DNMT1, that reduces DNMT1 activity to 10% of wild type, have substantial, genome-wide hypomethylation in all tissues and develop aggressive T-cell lymphomas with a high frequency of chromosome 15 trisomy(Trinh, Long et al. 2002). Therefore, the role of methylation in carcinogenesis may be more complicated than was at first supposed.

2.2.6.2 DNA Methylation and Gene Regulation

Eukaryotic gene expression is a complex process involving the integration of *cis*-acting elements within the DNA and *trans*-acting modulators in the form of transcription factors. DNA methylation of *cis* elements at the promoter of a gene is associated with downregulation of that gene. In contrast methylation of the coding region of the gene has a variable effect on transcription (Wajed, Laird et al. 2001).

Currently there is much debate over how promoter methylation integrates with downregulation of transcription. Several explanations have been offered by those who believe methylation is central to gene silencing.

Stereochemical interference with protein-DNA interaction is the simplest model to emerge. The transcription factors AP-2, NF-kappa B, E2F, CREB and c-Myc/Myn all recognise sequences that contain CpG residues and methylation of CpGs in these elements inhibits binding of these transcription factors (Tate and Bird 1993; Jones and Wolffe 1999).

DNA methylation also recruits specific repressor complexes that recognise and bind methylated DNA (Bader, Walker et al. 1999; Singal, Tu et al. 2000). MeCP1 and MeCP2 were first of these to be characterised. Using the common methyl binding domain motif to interrogate the genome database three others were found: MBD1, MBD2 and MBD4 (Hendrich, Abbott et al. 1999). More recently Kaiso, a repressor that binds methylated DNA through a zinc finger motif, has been characterised (Prokhortchouk, Hendrich et al. 2001). MBD4 is unusual amongst this group as it is associated with DNA repair. The others repress transcription in cell culture by interfacing with histone deacetylase complexes (Fuks, Hurd et al. 2003).

DNA methylation and its relationship with histone modification is complex. Three of the transcriptional repressors mentioned above, MeCP2, MBD2 and

MBD3 are found complexed with MeCP-2, MeCP-1 and Mi-2 which have histone deacetylase activity. Methylated reporter gene constructs with binding sites for these complexes have shown partial reversal of transcriptional silencing following treatment with the HDAC inhibitor Trichostatin A. This suggests that DNA methylation and histone deacetylation are both required to maintain complete silencing of a gene.

Methylation of histone H3 at lysine 9 (H3K9) appears to be a molecular marker of transcriptional silencing. Acetylation and methylation at this site appear to be mutually exclusive: acetylated H3 is commonly reported in the context of unmethylated, transcriptionally active euchromatin and H3K9 methylation in the setting of densely methylated, transcriptionally silent heterochromatin(Herman and Baylin 2003).

The temporal relationship between DNA methylation, histone H3 acetylation/methylation and transcriptional activity has been difficult to unravel experimentally. However, it has recently been shown, using a transfected GSTP1 gene in a prostate cancer cell line, that: (a) histone acetylation is independent of gene expression; (b) histone deacetylation is triggered by low levels of DNA methylation; (c) spread of DNA hypermethylation across the CpG island is associated with MBD2 but not MeCP2 binding; and (d), H3K9 methylation occurs after histone deacetylation and is associated with extensive DNA methylation of the CpG island (Stirzaker, Song et al. 2004).

With regard to which is more important in terms of gene silencing, DNA methylation or histone deacetylation and methylation: it appears that reversal of DNA methylation leads to gene re-expression and eventual reversal of the histone modifications, however reversal of the histone modifications does not lead to DNA demethylation (Fahrner, Eguchi et al. 2002).

2.2.6.3 DNA Methylation and Cancer

In the last decade or so DNA methylation in cancer has received intense study. In general, tumour DNA shows global hypomethylation, but aberrant hypermethylation of specific tumour suppressor CpG islands.

2.2.6.3.1 Hypermethylation in Cancer

Hypermethylation in cancer is technically easier to detect and therefore has been reported to a far greater degree than hypomethylation. Hypermethylation of a CpG island associated with a promoter region is generally accepted to imply transcriptional silencing of that gene. CpG island methylation has been shown to be an early event in cancer and may be related to aging (Jaenisch and Bird 2003). Methylation profiles in disparate cancers have been studied extensively and DNA methylation is reported in many solid and haematological tumours(Costello, Fruhwald et al. 2000). Some genes are methylated in many different tumour types, e.g. p16INK4a and RASSFIA, while other genes are specifically methylated in particular cancers e.g. GSTP1 methylation occurs in more than 90% of prostate tumours, but it is reported as unmethylated in AML (Melki, Vincent et al. 1999; Millar, Paul et al. 2000).

As methylation profiles were analysed, distinct patterns emerged in many malignancies, with known familial tumour suppressor genes methylated in sporadic forms of these tumours: e.g. VHL in clear cell carcinoma of the kidney (Herman, Latif et al. 1994), BRCA1 in breast cancer (Catteau, Harris et al. 1999), HMLH1 and APC in colon cancer (Kane, Loda et al. 1997). In familial cancers it was also shown that the non-mutated allele was methylated in a high proportion of cases. (Esteller, Fraga et al. 2001). As technology progresses the focus of hypermethylation is changing from analysis of a few loci at a time to multiple CpG island analysis using DNA microarrays. In breast tumours, CpG

island methylation patterns established using DNA microarrays correlate with ER status (Yan, Yang et al. 2001).

The role of DNA hypermethylation in carcinogenesis is proven in many tumour types (Table 2.1 page 67) there is also evidence that it can be involved in the modulation of transcription after an established breast cancer cell-line has been created. On disruption of estrogen receptor (ER) alpha signalling by small interfering RNA, polycomb repressors and histone deacetylases were recruited and initiated stable repression of the progesterone receptor (PR) gene, a known ER alpha target, in breast cancer cells. This event was accompanied by DNA methylation of the PR promoter. Re-establishing ER alpha signalling alone was not sufficient to reactivate the PR gene; DNA demethylation was also required. Methylation microarray analysis has further shown that progressive DNA methylation occurs in multiple ER alpha targets in breast cancer genomes. (Leu, Yan et al. 2004). Aberrations of DNA methylation arising after a malignant cell line is established suggest that DNA methylation control is abrogated during the carcinogenic process and remains so. There may also be a stochastic quality to DNA methylation, if the altered methylation pattern of a cell confers an advantage it may be selected as in a Darwinian model of evolution. However, other studies also suggest that DNA methylation does not take place until after transcriptional silencing has occurred and expression from a promoter will block DNA methylation (Song, Stirzaker et al. 2002). Therefore, DNA methylation and the alteration of the histone code may in some cases be secondary to the effects of transcription factor binding.

Hanahan and Weinberg formulated six novel capabilities a cell has to acquire to become a cancer cell: limitless replicative potential, self-sufficiency in growth signals, insensitivity to growth-inhibitory signals, evasion of programmed cell death, sustained angiogenesis and tissue invasion and metastasis (Hanahan and Weinberg 2000). This review not only highlighted the hurdles a cell has to overcome, but also made the observation that there was possibly more than

one pathway to gain each capability. Jones and Widschwendter examined the scope of DNA methylation as a means to overcome these barriers and were able to identify numerous targets in disparate cell types (Widschwendter and Jones 2002).

The observations of Jones and Widschwendter, in the context of Hanahan and Weinberg's formulation raises the point that, although there may be several different ways to overcome, e.g. cell cycle checkpoints, methylation of certain tumour suppressor genes, e.g. P16INK4A, are extremely common in disparate tumours (Lamy, Sesboue et al. 2002; Shao and Nguyen 2002; Esteller 2003; Silva, Silva et al. 2003; Lind, Thorstensen et al. 2004; Lund, Andersson et al. 2004). Genes that play such a central role are more likely to be methylated rather than the numerous ancillary genes required for the same phenotype.

Methylation of genes involved in DNA repair leads to increased mutation rates as the repair genes are transcriptionally silenced. This was first shown for the HMLH1 gene, mutated in patients with type II familial non-polyposis colon cancer (Herman, Umar et al. 1998). A precursor cell in this case develops a mismatch repair phenotype that results in instability of microsatellite sequences. Methylation of this gene is probably an early event in the carcinogenic process, as it has been observed in colonic epithelium that is normal histologically but has microsatellite instability (Nakagawa, Chadwick et al. 2001; Nakagawa, Nuovo et al. 2001).

MGMT removes DNA adducts and this gene is methylated in a number of tumours - lung, colon, lymphoid (Esteller, Catasus et al. 1999; Fleisher, Esteller et al. 1999; Sanchez-Cespedes, Esteller et al. 1999; Gallardo, Esteller et al. 2004; Lind, Thorstensen et al. 2004) and therefore is not expressed. Lack of MGMT expression leads to increased accumulation of O-6 methylguanine adducts, leading in turn to increased G to A transition mutations. Tumours with

MGMT silencing are more likely to have TP53 and K-RAS mutations (Esteller, Gonzalez et al. 2001) (Esteller, Hamilton et al. 1999).

2.2.6.3.2 Methylation and Increased Mutation Rates in the genome

CpG methylation increases mutation rates by hydrolytic deamination of cytosine to thymine (Schmutte and Jones 1998). Fifty percent of the mutations in TP53 in somatic cells occur at methylated sites (Fahrner, Eguchi et al. 2002). This increased mutagenic rate is observed in the germ line of all organisms that methylate their DNA.

Methylated CpGs are are also more likely to undergo UV induced deamination - this is due to the methyl group shifting the UV absorption spectrum for cytosine to a spectrum closer to that which is prevalent in sunlight. Not only does methylation affect the cytosine directly, but it also has an indirect effect on the adjacent guanidine residue, making it more susceptible to G to T transversions mutations, induced by the tobacco carcinogen metabolite benzo(a)pyrene diol epoxide (Pfeifer 2000; Pfeifer, Tang et al. 2000)

2.2.6.3.3 Hypomethylation in Cancer

Hypomethylation of the genome is a global phenomenon and tends to be associated with higher grades and by inference more poorly differentiated tumours (Ehrlich 2002; Ehrlich, Jiang et al. 2002). It is found in solid malignancies, such as breast and cervical (Kang, Kim et al. 2001), prostate (Prasad, Thraves et al. 1999) and hepatocellular carcinomas (Yoshikawa, Matsubara et al. 2001), and also in haematological malignancies such as chronic B-cell lymphocytic leukaemia (Feinberg and Vogelstein 1983).

Hypomethylation has a role in the genetic disease immunodeficiency, centromeric instability, and facial abnormalities (ICF). This condition is characterised by facial dysmorphism, mental retardation, and recurrent infections. The disease is linked to mutations in the DNMT3B gene (Hansen,

Wijmenga et al. 1999) that leads to chromatin instability as a result of abnormalities in the methylation of DNA. Hypomethylation of the pericentromeric regions of chromosomes 1,9 and 16 is found in both this condition and in breast cancers.(Narayan, Ji et al. 1998). All Children born with this condition die before adulthood due infections that cannot be combated by a crippled immune system. It is therefore not known whether abnormalities in DNMT3B lead to increased incidence of cancers.

The role of hypomethylation in carcinogenesis may be threefold. Initially there were suspicions that hypomethylation of CpG regions would lead to activation of oncogenes, previously repressed by methylation. It was initially reported that H-RAS hypomethylation was found in genes of cancer cells compared with their normal counterparts (Feinberg and Vogelstein 1983). This study was carried out on only 5 cases and larger studies have not yielded further evidence, so demethylation as a means of directly activating oncogenes remains only a theoretical possibility. The reactivation of transposable elements through demethylation, and their role in oncogene expression has, however, been established: L1 mutational inserts disrupt the APC gene and the CMYC gene in colon and breast cancer, respectively (Costello, Fruhwald et al. 2000). Hypomethylation has also been shown to promote chromosomal instability (Eden, Gaudet et al. 2003) and therefore cancer.

2.2.6.3.4Folic Acid, Gene Polymorphisms, DNA Methylation and Cancer

There are two main branches to the metabolism of methyl groups in a cell. One is concerned with purine and thymidine synthesis and the other involves methionine and s-adenosylmethionine for protein and polyamine synthesis and methylation reactions. Methylenetetrahydrofolate reductase (MTHFR) transfers methyl groups from the first of these branches to the second. Specifically, it converts 5,10-methylenetetrahydrofolate to 5-methyl-tetrahydrofolate which then transfers a methyl group to homocysteine to produce methionine.

Deficiency in this enzyme leads to increased homocysteine in the serum and urine and causes thrombo-occlusive vascular disease and mental retardation. Two less severe polymorphisms also exist, MTHFR C677T and MTHFR A1298C. The C677T variant is associated with higher serum homocysteine levels and increased risk of vascular disease and neural tube defects. It is also associated with an increased risk of some cancers: a 677TT genotype confers 6-fold increased risk of developing oesophageal squamous carcinoma [adjusted odds ratio (OR), 6.18; 95% confidence interval (CI), 3.32-11.51] compared with those who had the 677CC genotype (Baiqiu, Songbin et al. 2000; Hong, Song et al. 2001; Song, Xing et al. 2001; Song, Song et al. 2001). Breast cancer has also been investigated for an association between polymorphisms in MTHFR. A significant excess of the A677T polymorphisms was observed among the breast cancer cases (odds ratio 1.43, 95% confidence interval 1.02-2.00). The association was more pronounced in those with breast cancer diagnosis under the age of 40 years (odds ratio 1.66, 95% confidence interval1.12-2.41) (Campbell, Baxter et al. 2002) . The MTHFR 677TT variant genotype conferred a 2-fold increased risk of cancer of the gastric cardia (odds ratio, 2.04; 95% confidence interval, 1.28-3.26), and a significantly elevated risk was also seen in MTHFR 677CT heterozygotes (odds ratio, 1.56; 95% confidence interval, 1.03-2.36). The MTHFR A1298C polymorphism had no effect on risk, which is in keeping with other studies (Miao, Xing et al. 2002).

The mechanism by which a 677TT genotype confers increased risk of cancer is still open to speculation. It has been suggested by some authors that inadequate thymidine pools, caused by the altered MTHFR enzyme, lead to increased incorporation of uracil into DNA. This results in increased strand breaks and creates precursors for chromosome translocations and deletions. Confusingly the 677TT genotype seems to be associated with a decreased risk of leukaemia (Wiemels, Smith et al. 2001).

Hypomethylation and hypermethylation of genes is associated with folate deficiency. In animals, folate deficiency initially results in exon-specific hypomethylation of *TP53* and increased DNA methyltransferase activity (Kim, Pogribny et al. 1996); as the deficiency continues there is an increase in both *TP53* and genome wide methylation (Pogribny, Miller et al. 1997).

Interpretation of these results is difficult as there are confounding factors. The deficiency in the MTHFR gene would cause decreased folate, increased uracil incorporation and in turn lead to DNA strand breaks. p53 would be upregulated in response to this DNA damage. The methylation changes noted in the p53 gene might be secondary to expression, not a direct consequence of folate metabolism.

However, population studies do indicate that low serum folate and homozygosity for C677T confer an increased risk of colorectal adenomas when compared to wild type and heterozygotes, but only in individuals with low folate levels. Further work in this field is necessary to provide more detailed understanding of the mechanisms at work, and the effect of intervention to increase folate levels.

2.2.6.4 DNA Methylation as a Marker for Tumour Diagnosis and Prognosis

With exception of the lymphomas and some other haematological malignancies, most molecular markers of cancer still have less prognostic significance than the histological and clinical staging of the disease. Methylation aberrations of the genome occur at an early stage in tumours and there are tumour-specific forms of methylation. Relatively recently breast tumours have been sub-classified by their molecular profiles, and it has been claimed that there are gene expression signatures predictive of a short interval to distant metastases in node-negative patients (without tumour cells in local lymph nodes at diagnosis). The estimated hazard ratio for distant metastases in the group with a poor-prognosis signature, in comparison to the group with the good-prognosis signature, was 5.1 (95 percent confidence interval, 2.9 to 9.0; $p<0.001$). These molecular categories cannot be differentiated by microscopy alone (van 't Veer, Dai et al. 2002; van de Vijver, He et al. 2002).

Earlier diagnosis of breast cancer is an important goal. Detection of sub-clinical tumours through mammographic breast screening may not have exhausted the possibilities for increased survival, and smaller yet may need to be the goal. Early detection of tumours by picking up methylated DNA in the serum is a possibility. As tumours grow and invade they shed DNA into the circulation and methylation changes in this DNA can be detected through MSPCR of the serum. There have been no reports of DCIS DNA detection in serum and at present there is no clear evidence to indicate at what stage in a tumour's evolution tumour DNA appears in the serum. There is also the question as to which methylation changes should be looked at in relation to which tumours, and the issue of methylation changes with increased age is also an important consideration. Large population-based studies are required to assess these issues and at present the methylation profiles of cancers and the ageing genome have not been characterised to the degree required to move into such a large study. The production of CpG island DNA methylation chips that cover

the entire genome may lead to profiling of methylation in the aging genome and a profile of breast tumour methylation. As yet molecular cancer screening is in its infancy.

2.2.6.5 DNA Methylation and Novel Therapeutic Strategies

That epigenetic modifications to the genome are potentially reversible has been shown many times in cell culture (Plumb, Strathdee et al. 2000). Therapeutic reversal of DNA methylation with re-expression of critically affected genes is a plausible aim and presents an attractive option for the treatment of some malignancies. Many clinical trials are being carried out at present using azacytidine, phenylbutyrates and histone deacetylase inhibitors to reactivate genes in haematological(Ruter, Wijermans et al. 2004) and solid tumours(Strathdee, MacKean et al. 1999).

The most well trialled demethylating drugs in use at present are azacytidine (5-azacytidine), decitabine (5-2-deoxyazacytidine), fazarabine (1-B-D-arabinofurasonyl-5-azacytosine), and dihydro-5-azacytidine. They all share a common structural motif in the form deoxycytidine, but they differ in the functional group located at the fifth position in the pyrimidine ring. Reversal of epigenetic changes using histone deacetylase inhibitors (HDACIs) is also in trial (Thiagalingam, Cheng et al. 2003).

The mechanism of action of the different deoxycytidine analogues varies, with the functional group they possess. Decitabine itself is phosphorylated by the cell (Lubbert 2000) and is subsequently incorporated into DNA causing disruption of DNA synthesis, structural instability and DNA damage(D'Incalci, Covey et al. 1985; Juttermann, Li et al. 1994). It also binds to DNMT, which is clearly the pharmacological aim. Azacytidine is preferentially incorporated into RNA (Santini, Kantarjian et al. 2001). The direct DNA damage caused by these agents may account for their myelotoxicity. Pre-clinical trials with decitabine

confirm that they can reverse methylation changes in cell lines and cells from human leukaemia patients(Momparler, Bouchard et al. 1984; Steuber, Krischer et al. 1996). The evidence of effectiveness in clinical trials differs between solid tumours, where the response rate is reported as less than 10% (Goffin and Eisenhauer 2002), and haematological malignancies, where the results are more encouraging. In childhood induction-resistant AML the combination of etoposide, amsacrine and azacytidine had a higher complete response rate than etoposide and amsacrine by themselves (Kantarjian, O'Brien et al. 2003). CML and myelodysplastic syndrome have shown a significant response to decytibine (Daskalakis, Nguyen et al. 2002; Kantarjian, O'Brien et al. 2003). In the responding myelodysplastic syndromes the p15-cyclin dependent kinase inhibitor is demethylated and re-establishment of normal gene expression is reported (Galm, Wilop et al. 2004). This is supportive evidence for demethylation as a therapeutic mechanism, rather than DNA damaging pathways. Si-RNA techniques directed against DNMT1, DNMT3A and DNMT3B have also been used to demethylate genes, but only in cell culture (Espada, Ballestar et al. 2004). One of the most significant conclusions from this study was that DNMT1 interference was required for demethylation, but not DNMT3A or DNMT3B.

DNA methylation and histone acetylation co-ordinate to control gene expression and HDACIs activate gene expression from methylated genes. HDACIs alone, however, are not adequate to activate gene expression from densely methylated genes, but when HDACIs are used in combination with a DNA demethylating agent they have a synergistic effect (Cameron, Bachman et al. 1999). Three patients with cutaneous T-cell lymphoma involved in a phase I clinical trial showed a partial response to treatment with depsipeptide, an HDACI, and one patient with peripheral T-cell lymphoma had a complete response (Piekarz, Robey et al. 2001). HDACIs also induce remission in transgenic models of promyelocytic leukaemia(He, Tolentino et al. 2001).

Combination therapies using an HDACI and DNA demethylating drug are in clinical trials at present.

Surprisingly, sodium valproate, a drug traditionally used for treatment of epilepsy and bipolar disorder, has an effect on methylation-mediated gene expression, as do the anti-arrhythmic drugs hydralazine and procainamide. (Segura-Pacheco, Trejo-Becerril et al. 2003). The sodium valproate cell culture studies show direct inhibition of HDACs, with an associated increase in 5-lipoxygenase mRNA in both proliferating and differentiated cells. Demethylation of ER, RARB and p16 in cell culture experiments using procainamide and hydralazine have been demonstrated in breast cancer derived cell lines (Segura-Pacheco, Trejo-Becerril et al. 2003). It would be interesting to analyse breast cancer patients who have been treated for cardiovascular disease with these drugs in co-ordination with normal chemotherapy for breast cancer. The model of drug resistance in ovarian tumours suggests that the HMLH1 gene is down-regulated through methylation (Plumb, Strathdee et al. 2000). The end-result of this is to increase the chance of mutations in the cell, favouring the emergence of drug resistance. This exact model may not hold in breast cancer, however there is evidence that tumour methylation is altered as cancers progress and therefore reversal of methylation may revert a tumour to a drug susceptible phenotype. Women with breast cancer who have taken procainamide or hydralazine during chemotherapy might have longer disease free survival if methylation aberrations are prevented from developing both during the chemotherapy and after treatment. Clinical evaluation of these drugs in a case of cervical carcinoma and a head and neck cancer case resulted in re-expression of p16 and RARB (Segura-Pacheco, Trejo-Becerril et al. 2003). This is evidence that they are viable demethylating agents and further work is needed to establish their role, if any.

Side effects of demethylation include, not only myelotoxicity (a dose limiting side effect), but also activation of dormant oncogenes. This was reported in the case of a pancreatic cancer which was thought to metastasize more rapidly as a consequence of invasion promoting genes being released from methylation (Sato, Maehara et al. 2003). There are also issues with regard to dosage. Monitoring methylation of genes in the peripheral circulation is used in some clinical trials as a means of assaying the effect of the demethylating agent. Unfortunately, this surrogate may not give an accurate picture of what is going on in solid tumours.

Combinations of drugs may be the way forward, as synergy between HDACIs and demethylating drugs has been shown. Scheduling is also important - if chemoresistance is encountered due to methylation then chemotherapy should be planned for the 'window' produced by demethylating drugs. Monitoring of demethylation treatment may become more advanced as CpG island chips become available, allowing profiling of a large number of CpG islands in one sample. Individualised treatment may then allow more effective treatment than standardised regimes.

2.3 Project Aims

It is thought that specific DNA methylation events occur early in the evolution of breast cancer. This project aspired to screen a sample of breast cancers for novel, aberrantly methylated CpG islands (CpG islands). Following this initial gathering of 'targets' it was proposed to characterise target methylation changes in a larger cohort of new breast tumours. In addition, methylation profiles of the larger cohort were to be related to available clinico-pathological data.

Screening for methylation changes in cancer has evolved as technology has become more advanced. At the inception of this project there were few screening techniques available: restriction landmark genomic scanning (RLGS) (Costello, Fruhwald et al. 2000) was difficult to interpret and had only been used successfully in one laboratory. Differential methylation hybridization (DMH) (Huang, Perry et al. 1999) is limited to a candidate gene approach and does not screen the whole genome. RLGS and DMH are also dependent on methylation-sensitive restriction enzyme digests as a means of differentiating a methylated locus from an unmethylated one. Restriction digests have been shown to leave up to 10% of restriction sites undigested despite extended digestion with more than one set of isoschizomers, and incomplete digests may lead to false positive signals.

A technique was selected which had been used successfully in this laboratory (Brock, Huang et al. 2001), that relied on the methyl binding domain of the rat MeCP2 protein. It is a technique first developed by Professor Adrian Bird and Dr. Sally Cross of Edinburgh University to isolate CpG islands in the human genome. It was later adapted by Dr Graham Brock, with the addition of a subtractive hybridisation step, to isolate differentially methylated sequences from tumour DNA. This approach has yielded novel CpG islands in the past (Yan, Chen et al. 2001), and the methylation status of one of these, HS3ST3B,

in combination with the methylation status of a few other genes, can segregate breast tumours into ER positive or negative tumours. (Yan, Chen et al. 2001). The methylation status of HS3ST3B also has a similar role in ovarian tumours (Professor Tim Huang, Ohio University, personal communication). It was therefore regarded it as a tested method that was functioning robustly.

Following screening for CpG islands the targets were to be sequenced to assess their CpG and GC content and provide DNA code that could be used to interrogate the human genome database. If the genes met the criteria for CpG islands they were to be included in a DNA array that was being developed in the laboratory at this time.

It is generally accepted that dense methylation of a CpG island is associated with transcriptional silencing of that gene (Esteller 2005). Despite the robustness of this association correlation of the relationship between transcription and methylation was also sought.

These were the aims at the start of this project - some were realised and some were not.

Gene	Role in Tumor Development	Site of Tumour
APC	Deranged regulation of cell proliferation, cell migration, cell adhesion, cytoskeletal reorganization, and chromosomal stability	Breast Lung Esophageal
BRCA1	Implicated in DNA repair and transcription activation	Breast Ovarian
CDKN2A/p16	Cyclin-dependent kinase inhibitor	GIT Head and neck NHL Lung
DAPK1	Calcium/calmodulin-dependent enzyme that phosphorylates serine/threonine residues on proteins; Suppression of apoptosis	Lung
E-cadherin	Increasing proliferation, invasion, and/or metastasis	Breast Thyroid Gastric
ER	Hormone resistance	Breast Prostate
GSTP1	Loss of detoxification of active metabolites of several carcinogens	Prostate Breast Renal
hMLH1	Defective DNA mismatch repair and gene mutations	Colon Gastric Endometrium Ovarian
MGMT	p53-related gene involved in DNA repair and drug resistance	Lung Brain
p15	Unrestrained entry of cells into activation and proliferation	Leukemia Lymphoma Squamous cell carcinoma, lung
RASSF1A	Loss of negative regulator control of cell proliferation through inhibition of G_1/S-phase progression	Lung, Breast, Ovarian, Nasopharyngeal, Kidney
Rb	Failure to repress the transcription of cellular genes required for DNA replication and cell division	Oligodendroglioma Retinoblastoma
VHL	Altered RNA stability through and erroneous degradation of RNA-bound proteins	Renal cell cancer

Table 2.1 Table of genes that are methylated in cancer.

Abbreviations: APC, adenomatous polyposis coli; BRCA1, breast cancer 1; CDKN2A/p16, cyclin-dependent kinase 2A; DAPK1, death-associated protein kinase 1; ER, estrogen receptor; GSTP1, glutathione S-transferase Pi 1; hMLH1, Mut L homologue 1; MGMT, O-6 methylguanine-DNA methyltransferase; RASSF1A, Ras association domain family member 1; Rb, retinoblastoma; VHL, von Hippel-Lindau; GIT, gastrointestinal tract; NHL, non-Hodgkin's lymphoma.

3.2.1 Protein Purification

Figure 3-3 shows a 15% polyacrylamide gel protein gel which was run at 200V for 1hour. The lanes contain progressive steps in the extraction process starting from the crude bacterial extract, ending with the purified form of the protein which was eluted from the nickel agarose matrix. The specific contents of each lane are given in the figure legend. The gel proves that a band weighing between 10 and 15kDa was purified from the nickel agarose slurry. The cloned Rat MBD domain and associated polyhistidine tail weighs 11.2KDa. This protein purification process was carried out for every new column that was manufactured; this amounted to eight columns in total, although only six of these were used to construct libraries.

3.2.2 Column Testing

Figure 3-4 is a 2% agarose gel, which was run at 100V for 2hours and the associated radiograph of the gel. The sites of the restriction enzymes HpaII and HhaI are given in Table 3-2 and methyltransferase recognition sites and density of methylation are given in Table 3-3. This set of experiment is designed to test for efficacy of an HMBD column to bind densely methylated plasmid DNA at 0.4M NaCl and 0.6M NaCl and also to test that densely methylated plasmid DNA is eluted at 0.85M NaCl. An HMBD column was loaded with a linearised, end-labelled radioactive pTEST plasmid and then washed with 0.4M, 0.6M and 0.85M NaCl. Following the collection of each fraction the DNA was precipitated resuspended, split into three digests and 1ug of unmethylated unlabelled pTEST plasmid was added to each digest. The 3 samples from each fraction were digested with HpaII or HhaI or were undigested (as a negative control). HpaII and HhaI cut the unmethylated plasmid, HpaII cuts partially methylated plasmid, and neither cut the densely methylated plasmid. Only the unmethylated, unlabelled pTEST plasmid can be seen on the agarose gel; the

end-labeled pTEST plasmid, which was fractionated on the HMBD column, falls below the mass required to be visible on an agarose gel. It can be seen from the agarose gel that the restriction digests proceeded to completion and none were inhibited by the process of fractionation. The lower picture is an autoradiograph of the agarose gel after it was dried and incubated at -80oC with the radiographic film for 48hours. This autoradiograph demonstrates that there is unmethylated plasmid (digested by both HpaII and HhaI) in the low salt fraction, partially methylated plasmid (digested by HpaII, but not HhaI) in the medium salt fraction and densely methylated plasmid (resistant to digestion by either enzyme) in the high salt fraction. It can be concluded that this HMBD column binds and separates pTEST plasmid DNA on the basis of methylation density. This assay of a column's ability to separate pTEST plasmid DNA on the basis of methylation density was carried out for each new column that was constructed prior to fractionation of tumour and normal DNA. All columns tested were able to fractionate plasmid DNA on the basis of methylation.

3.2.3 Subtractive Hybridisation Library Generation

Figure 3-6 A), B) and C) shows the successive steps in the construction of CPD 3 and CPD 4 subtractive hybridization libraries. These are representative results. The first gel A) is an agarose gel of CPD3 and CPD4 following PCR with the catchlinker 2.4/2.44. This PCR step is carried out to enrich the library for tumour specific methylated DNA. In the case of these two particular libraries the portion of gel indicated by the white box was cut out and gel purified, and the DNA was then cloned into pGem Teasy. This step was carried out as previous libraries had given only fragments of DNA less than 200bp in size. This was thought to have been a result of the cloning process favouring smaller fragment sizes and perhaps the bacteria favouring smaller inserts. In order to bias the cloning process in favour of larger inserts the gel purification approach

outlined above was carried out. This gel also proves that the 2.4/2.44 ligation step was working.

Figure 3-6B) shows an agarose gel with PCR products obtained from cloned products of library CPD3. These clones were selected for PCR on the basis of the COT-1 DIG labeled screen. Colonies that did not label with the COT-1 probe were included for PCR. PCR was carried out using T4 and SP6 primers which are located in the pGem Teasy vector. Again the decision to use a clone following PCR analysis was based on the size of the product. If the product was greater than 350bp in size it was included for further analysis (i.e. subsequent digestion with HpaII and HhaI). 350bp was chosen as a cut off point because the vector that the insert was cloned into added 120bp to the insert size. So in the case of this set of clones only 2 (numbers 4 and 7) were included for HpaII and HhaI digestion. This gel does, however prove, that the cloning process was to some extent working, although there was still a bias towards small inserts being cloned despite the gel purification step.

Figure 3-6C) 92 shows the result of digesting the two clones from the CPD3 library with the restriction enzyme HpaII. The restriction site for this enzyme is frequently found in CpG islands multiple times and is an indirect assay of the CG content of a sequence of DNA. Restricting the cloned inserts with HpaII resulted in four fragments. There is one restriction site in the vector sequence, this means that there are only 2 sites within the cloned insert. Figure 3-7 95, shows the HMLH1 CpG island and it can be seen that there is a discrepancy in the number of HpaII sites in the HMLH1 CpG island and these cloned inserts. It was therefore interpreted that these inserts did not contain many CG dinucleotides, were not CpG islands and subsequently they were not submitted for sequencing.

3.2.4 Summary of Results of Libraries Constructed

Subtractive Hybridisation Library	Clones examined	Clones excluded by DIG-COT-I screening	Clones excluded after PCR analysis	Clones excluded after HpaII.	Candidate CpG Islands
CPD1	490	480	8	2	0
CPD2	1200	1079	96	25	0
CPD3	490	490	0	0	0
CPD4	490	478	5	7	0
CPD5	490	474	8	8	0
CPD6	490	470	8	12	0

Table 3-1 Results of screening the subtractive hybridisation libraries generated from six different HMBD columns.

This is a summary of results from six subtractive hybridisation libraries that were constructed using the six different HMBD columns and tumour DNA donated by Professor Tim Huang. Each library was constructed using a new HMBD column. The table shows the original number of clones picked and examined in the first column and then goes on to give the number of colonies excluded in the subsequent screening steps.

Table 3-1 above shows the number of clones examined in the cases of the hybridisation libraries CPD1-CPD6. Each library was constructed using a new HMBD column and subsequent fractionation, subtraction and screening steps. In total 3,650 clones were picked from these libraries (six libraries in total), 3,471 (95.9%) were excluded on the basis of COT-1 DIG labelled screening of libraries, 125 (3.4%) were ruled out on the basis of size and 54 (1.7%) were rejected by the HpaII screening step.

3.2.5 PCR for the HS3ST2 CpG Island in Subtractive Hybridisation Libraries

Figure 3-5 shows a PCR for the HS3ST2 CpG island which was found in a subtractive hybridization library constructed by Dr. Brock. PCR was carried out on the subtractive hybridization libraries CPD1-CPD5 for the HS3ST2 CpG island and also on a library constructed by Dr. Brock that was known to contain the HS3ST2 CpG island. This result shows that HS3ST2 is not present in the CPD libraries and it infers that the libraries are not enriched for densely methylated CpG islands in the same way as those that were constructed in the lab previously. It cannot be stated that the CPD libraries contained no CpG islands from this gel, however, it may be stated that they do not contain enrichment for a CpG island that has been shown to be methylated in 75/88 breast tumours in a previous study (Miyamoto, Asada et al. 2003). This is quite compelling evidence that there may be a problem with either the tumour samples being used for the fractionation process or the subtractive hybridization process. The column appears to work correctly in that it can fractionate densely methylated pTEST plasmid. However, there may be a difference in the functionality of the column to fractionate densely methylated genomic DNA.

Figure 3-3 HMBD Protein 15% polyacrylamide gel. This is a 15% polyacrylamide gel run in MOPS buffer for 1hour at 200V. The HMBD protein is 11.2KDa in size. The ladder used has markers at 15kDa and 10KDa which are marked with arrows. Lanes 1-12 are progressive stages in the purification of the protein. Lane 1 is a sample of the preinduction broth, lane 2 is the postinduction broth at 2hours, lane 3 is the flow through from loading the protein extract onto a sepharose anion exchange column, subsequent washes of this column with 0.4M, 0.85M and 1M NaCl are shown in lanes 4-6. Lane 7 is the remaining protein extract that did not bind to the nickel agarose matrix Lanes 8-12 are washes of a 200ul sample of the nickel agarose matrix-HMBD protein column: 200ul of matrix slurry was extracted and washed six times with 200ul of 250uM imidazole, an analogue of the histidine side chain. Imidazole competitively inhibits chelation of the nickel ion of the agarose matrix by the polyhistidine tail of the HMBD protein. At a concentration of 250uM all the protein bound to the nickel agarose matrix is eluted. These washes were pooled and the protein concentration calculated as described in Materials and Methods, the protein concentration ranged between 14-16mg/ml of matrix slurry.

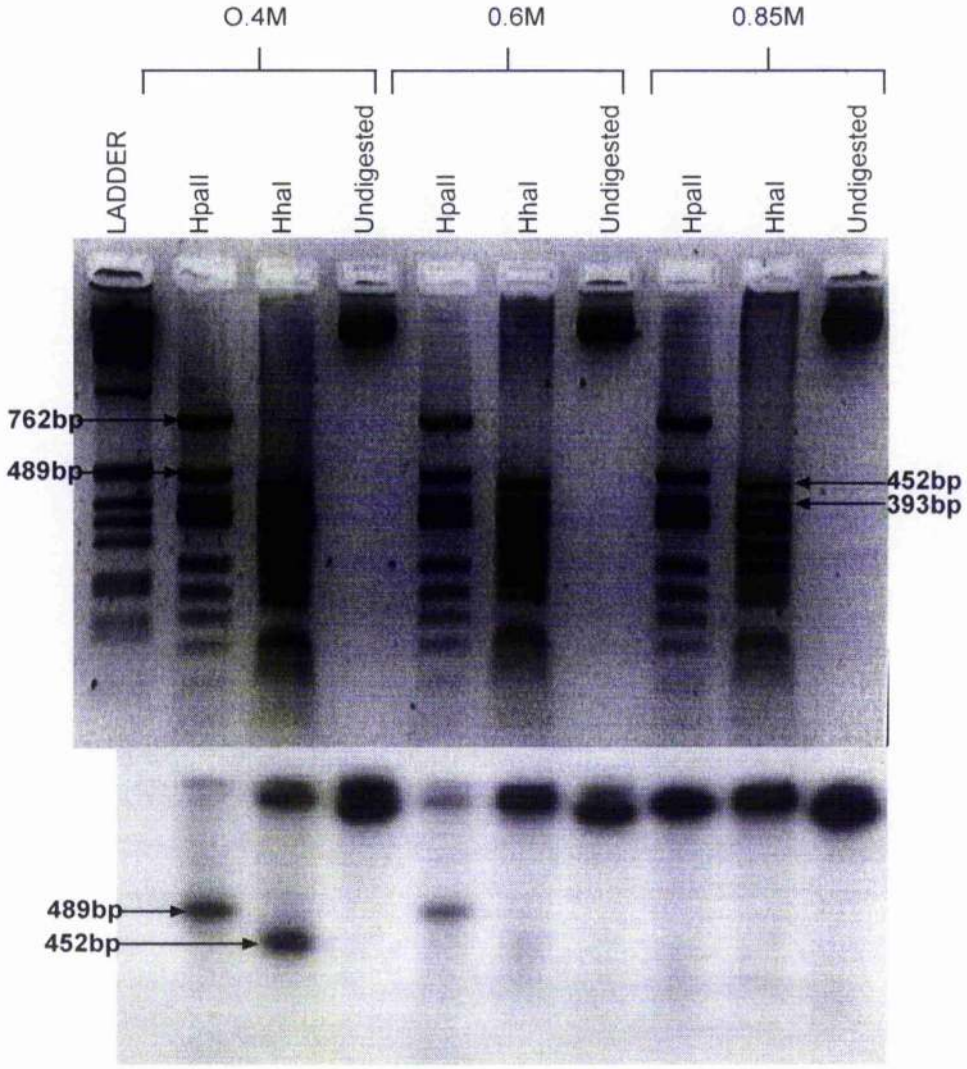


Figure 3-4 Separation of Unmethylated, Partially methylated and Densely Methylated pTEST Plasmid on the HMBD Column Plasmids were modified as described in the main text, endlabelled with dCTP mixed, loaded on to the HMBD column and fractions were collected over a NaCl gradient. The salt fractions are indicated above each lane. Following the collection of each fraction the DNA was precipitated, resuspended and 1ug of unmethylated unlabelled pGem-test plasmid was added to each fraction. The fractions were then digested with HpaII and HhaI: both enzymes will cut the unmethylated plasmid, HpaII will cut partially methylated plasmid, and neither will cut the densely methylated plasmid. The upper gel is the agarose gel digest of the fractions. Only the unmethylated plasmid spike can be seen; the labelled plasmids fall below the mass required to be visible on an agarose gel. It can be seen from this gel that the restriction digests proceeded to completion and none were inhibited. The lower picture is an autoradiograph of the the agarose gel after it was dried and incubated at -80oC with the radiographic film for 48hours. This autoraiograph demonstratres that there is unmethylated plasmid (digested by both HpaII and HhaI) in the low salt fraction, partially methylated plasmid (digested by HpaII, but not HhaI) in the medium salt fraction and densely methylated plasmid (resistant to digestion by either enzyme) in the high salt fraction. It can be concluded that this HMBD column binds and separates DNA on the basis of methylation density.

			PRODUCT SIZE(IN BASE PAIRS)		
	Recognition Site	Methylation Sensitive	Pgem-Test Unmmethylated	Pgem-Test Partially Methylated	Pgem-Test Densely Methylated
HpaII	C/CGG GGC/C	Yes	762, 489, 404, 351, 347, 309	762, 489, 404, 351, 347, 309	3543
HhaI	GCG/C C/GCG	Yes	452, 393, 337, 332, 270, 263	3543	3543
MspI	C/CGG GGC/C	No	762, 489, 404, 351, 347, 309	762, 489, 404, 351, 347, 309	762, 489, 404, 351,347, 309

Table 3-2 pGem-test product sizes following digestion with the restriction enzymes HpaII, MspI and HhaI.

Product sizes are given for the six largest products produced by cleavage of each enzyme. Product sizes are also given following methylation of the plasmid with HhaI methyltransferase, and Sss methyltransferase

.

	Recognition Site	No of sites in pGem - Test	Density of methylation (No of sites/ size of plasmid)
HhaI Methyltransferase	GCGC CGCG	28	0.0079
Sss Methyltransferase	CG GC	490	0.014

Table 3-3 Methyltransferase recognition sites and density of methylation in pGem-test.

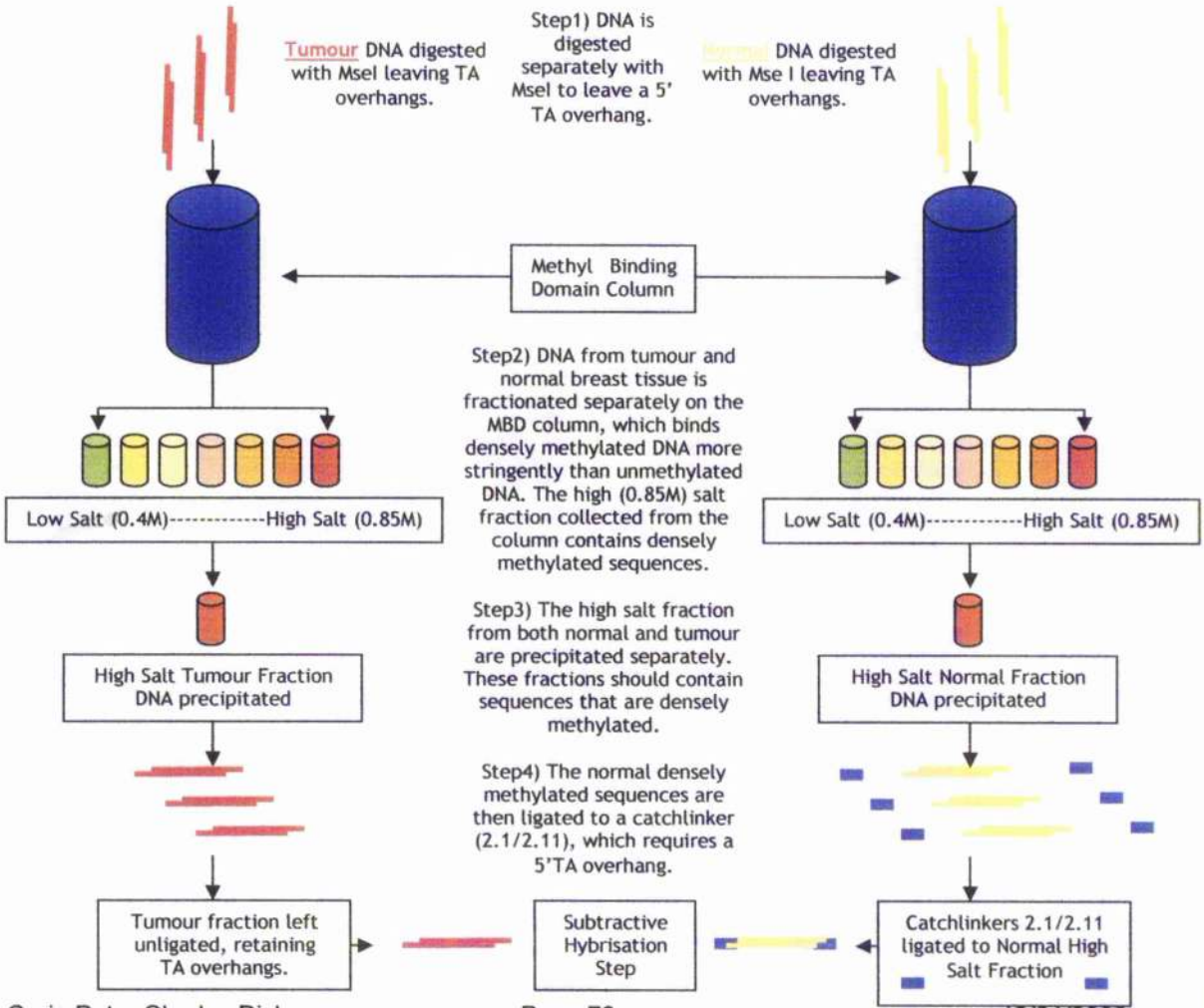
The site of modification is marked in bold.

3 Isolation of Methylated CpG Islands Using a Methyl Binding Domain Column

3.1 Introduction

Aberrant methylation of CpG islands in breast cancer is thought to be a significant, early event in breast carcinogenesis (Costello, Fruhwald et al. 2000). One of the principal aims of this project was to isolate, through a screening process called ICEAMP, novel CpG islands methylated in breast tumours. ICEAMP is an acronym derived from the title of an article, "Isolation of **C**p**G** **I**SLAND**S** Exhibiting **A**bnormal **M**ethylation **P**atterns" (Brock, Huang et al. 2001) and it is a technique developed by Dr. Graeme Brock. The steps of this process are outlined below.

3.1.1 Overview of Methodology



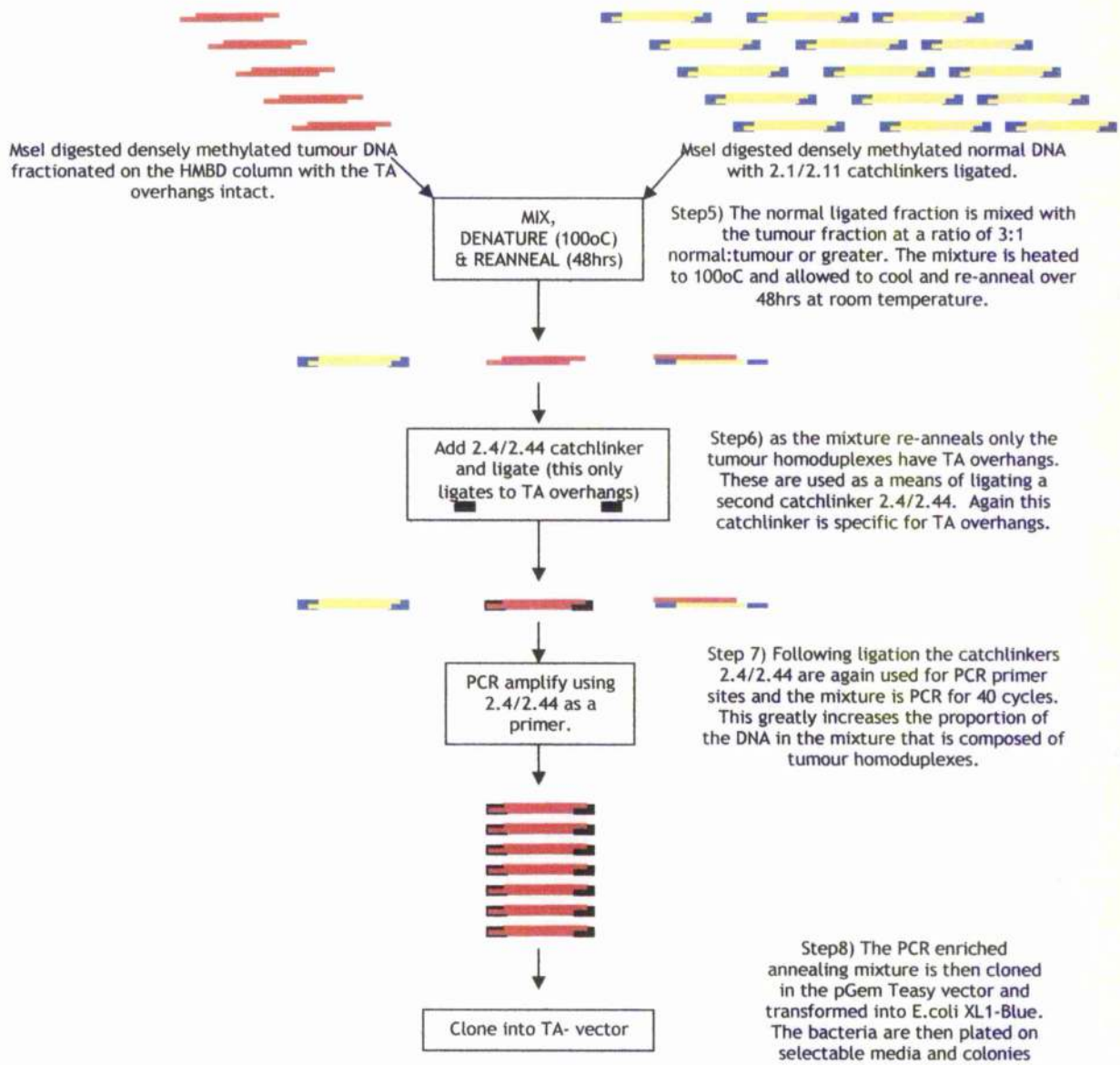


Figure 3-1 Flow diagram of DNA fractionation and subtractive hybridisation technique.

This is an outline of the DNA fractionation and subtractive hybridisation process. Normal genomic breast DNA is coloured yellow and breast tumour genomic DNA is coloured red. The catchlinker 2.4/2.44 is coloured black and the catchlinker 2.1/2.11 is coloured blue.

Figure 3-1 above is an outline of the DNA fractionation and subtractive hybridisation steps of the ICEAMP technique. The diagram shows separate samples of tumour DNA and normal DNA, which have been digested with MseI to leave 5' TA overhangs, being fractionated on a methyl binding domain column. This column is constructed as described in materials and methods and contains the methyl binding domain of the rat MeCP2 protein. The MBD domain is attached to a nickel-agarose matrix via a polyhistidine tail that was cloned into the protein. The column binds DNA, but has a greater affinity for densely methylated DNA compared to unmethylated DNA and partially methylated DNA. One protein molecule binds one methylated CpG and hence the greater the density of methylated CpG's the greater the affinity the DNA has to the column (Nan, Meehan et al. 1993; Cross, Charlton et al. 1994) (Brock, Huang et al. 2001). When functioning correctly the column will bind densely methylated DNA up to a salt concentration of 0.85M (Cross, Charlton et al. 1994) at which point the column no longer retains densely methylated DNA and it is released. This capability is utilised to fractionate the tumour and normal DNA separately and retrieve the densely methylated sequences from each. The 0.85M salt fraction which is shown in the diagram for both tumour and normal is then ethanol precipitated and the normal densely methylated sequences, which have TA overhangs, are ligated with the catchlinker 2.1/2.11, which will only ligate to DNA with a 5' TA overhang. Densely methylated tumour DNA is left with the TA overhangs intact at this point. It is worth stressing that up until this step tumour and normal DNA have not been mixed and the previous fractionation steps have all been carried out separately.

The next step is the subtractive hybridisation step. The normal methylated and now blunt ended DNA is mixed with the methylated tumour DNA, which has TA overhangs. This is then heated to 100oC and allowed to re-anneal at room temperature over a period of 48hours. As the mixture re-anneals there are three possibilities as seen in Figure 3-1 steps 5-6 above, the tumour will re-anneal to tumour (hopefully this will only occur in the case of densely

methylated CpG islands), the tumour will re-anneal to normal or normal will re-anneal to normal. The purpose of this step is to remove sequences, such as Alus Sines and Lines and other repetitive elements or non-CpG islands that are densely methylated and common to both tumour and normal. It is expected that densely methylated CpG islands in the tumour DNA will re-anneal to tumour. Assuming this is the case the penultimate step in the fractionation process is the ligation of a further catchlinker (2.4/2.44) to the reannealed mixture. Again this catchlinker will only anneal to strands of DNA with a 5' TA overhang and therefore should only anneal to tumour homoduplexes. If the fractionation process has worked correctly these should be densely methylated sequences that are found in tumour DNA but not normal DNA.

The last step in the subtractive hybridisation step is a PCR step, which uses the 2.4/2.44 catchlinkers as a primer site. 40 cycles of PCR are carried out. The purpose of this step is to greatly increase the number of tumour homoduplexes in the re-annealed mixture. The PCR mixture is then ligated into a selectable vector, in this case pGem Teasy, and transformed into E. coli XL-1 blue strain and plated on a selectable media. Clones with inserts are picked and seeded into the well of a 96 well plate. The subsequent screening process is outlined below:

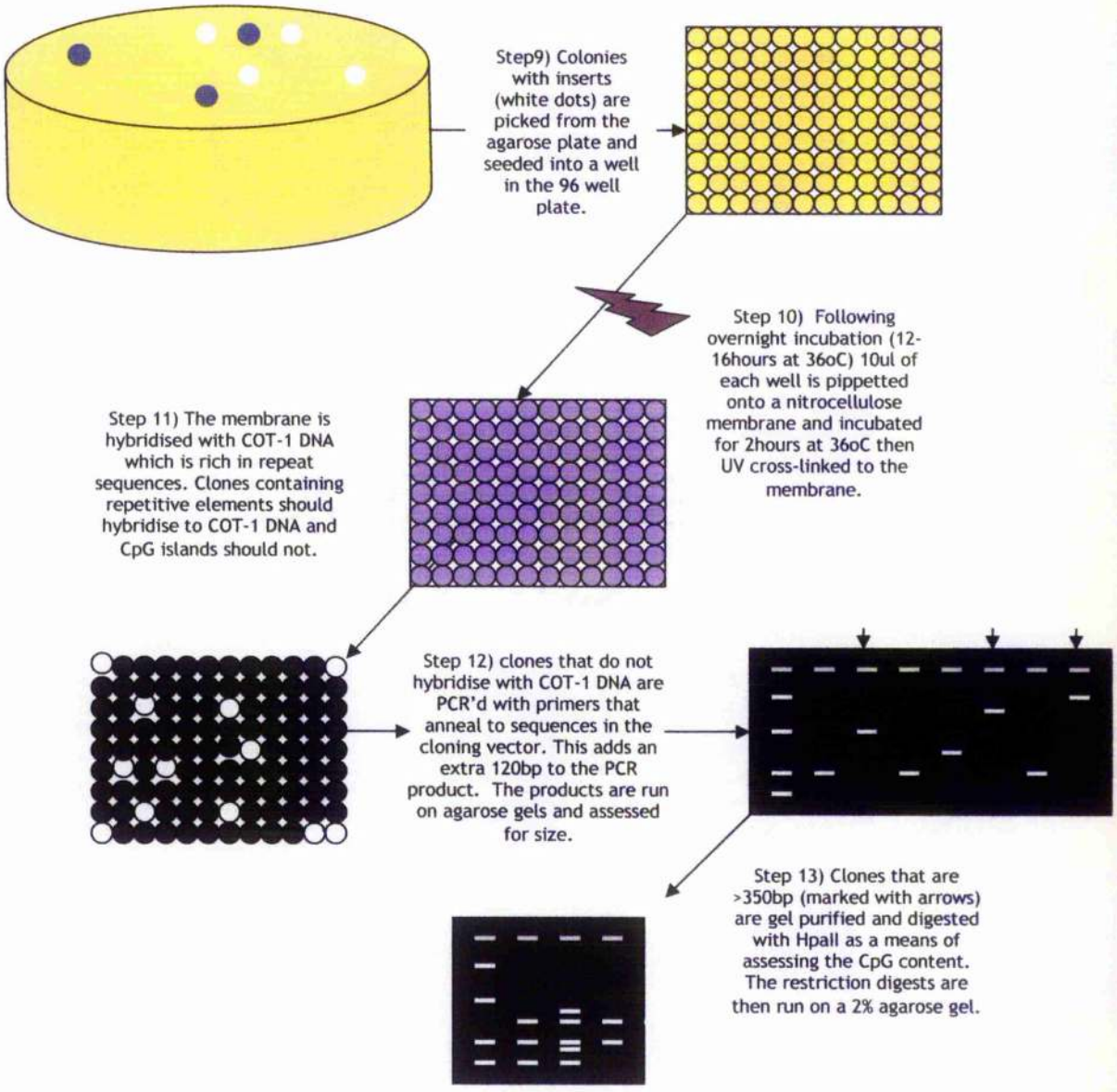


Figure 3-2 Schematic of the screening process for DNA libraries created using MBD columns and subtractive hybridisation.

3.1.1.1 Outline of Screening Process

The schematic above shows the screening process for the sequences that are cloned following subtractive hybridisation. In brief colonies with inserts are picked and grown in 100ul of L-broth containing ampicillin in a 96 well format. The colonies are then pipetted onto a nitrocellulose membrane, grown for a further 2 hours and then cross-linked to the membrane using UV radiation. The membrane is then hybridised with a COT-1 DIG labelled probe, which fluoresces and activates photographic paper. COT-1 DNA is rich in DNA from repetitive elements. These elements are usually densely methylated and constitute a high proportion of sequences in most of the libraries produced in the past (Brock, Charlton et al. 1999), it is therefore necessary or helpful to screen them out. Following hybridisation the membrane is dried and exposed to radiographic paper for 72hours at -80oC. In addition to the colonies that are pipetted onto the membrane controls are also cross linked to the membrane; CpG islands for BRCA1 and HMLH1 (which should not hybridise COT-1 DNA) and Alu repeat sequences which should.

Clones that did not hybridise the COT-1 probe were PCR amplified using primers located in the plasmid vector; this added 120bp to an insert's size. A slightly arbitrary figure of 350bp was used as a cut off point to select clones that were likely to be CpG islands. This figure was based on the literature at that time and also on previous experience of screening libraries. Clones larger than 350bp were gel purified and digested with HpaII, a restriction enzyme with the recognition sequence 'CCGG' whose site is generally dense in CpG islands (Bird, Taggart et al. 1985; Antequera, Macleod et al. 1989; Antequera and Bird 1993). HpaII digestion was used as a pseudo marker for CpG density. No specific density of HpaII sites was decided on as definitive of a CpG island and as discussed later on in this chapter this may have been an erroneous step in screening for CpG islands.

3.1.2 ICEAMP

There are two central features of the ICEAMP technique: 1) Isolation of methylated DNA from malignant and normal breast tissue using a poly histidine methyl binding domain (HMBD) column. 2) Subtraction of methylated normal DNA from methylated tumour DNA to generate sequences exclusively methylated in breast tumours.

3.1.2.1 HMBD Protein Expression and Purification

An HMBD column is constructed from a 79-amino acid methyl binding domain, cloned from the *Rattus norvegicus* MeCP2 protein, modified to include six histidine residues at the N-terminus. The HMBD protein is linked to a nickel-agarose matrix, which forms the solid support of the column, via chelation of the nickel ion by the poly-histidine sequence. Expression, purification and construction of the HMBD column are described below.

3.1.2.1.1 HMBD Expression

The plasmid pET6-HMBD, coding for the MBD-polyhistidine protein, was kindly donated by Dr Graeme Brock. This was transformed, expressed and purified, as described in Materials and Methods. Protein purification is demonstrated in Figure 3-3 below.

3.1.2.2 Column Function and Ability to Bind Methylated DNA

Following construction of an HMBD column it was tested for its ability to separate DNA on the basis of methylation density. A pGEM®-T Easy plasmid vector containing the HMLH1 CpG island (referred to as pTEST) was used to test each column. Three separate pTEST preparations were made, and linearised using NcoI. One was left unmethylated, the second was methylated using HpaII

methyltransferase and the third was methylated using Sss methyltransferase. This gave three different densities of pTEST methylation; unmethylated, partially methylated and densely methylated (table 2).

500ng of each pTEST plasmid was digested with the enzymes HpaII, HhaI (which are methylation sensitive) and MspI (which is not sensitive to methylation) to ensure the methylation reactions had modified pTEST appropriately. All three plasmids were end-labelled with ^{32}P gamma CTP, as described in Materials and Methods. All fractionations were carried out in the cold room at 4°C. 50ng of each pTEST was mixed in buffer A and loaded onto the HMBD column. The column was washed with three increasing NaCl concentrations of buffer A: Low Salt 0.4M (LS), Medium Salt 0.6M (MS) and High Salt 0.85M (HS). The three fractions were collected for further purification. To remove any remaining DNA the column was washed with 50mls of 1M NaCl buffer A and equilibrated with 0.4M NaCl buffer A. The HS fraction was diluted to 0.4M NaCl and loaded onto the HMBD column again. The fractionation process was repeated and the second HS fraction was collected (HS2). Repeating the fractionation step increased the purity of densely methylated DNA. The LS, MS and HS2 fractions were pooled and precipitated using isopropanol/ethanol, as described in Materials and Methods.

Each fraction was 'spiked' with 1ug of unmethylated, non-radioactive pTEST plasmid prior to digestion with HpaII and HhaI restriction enzymes. Addition of the unmethylated, non-radioactive plasmid is an internal control for the restriction digests. If the unmethylated plasmids are digested then the DNA that is in the three fractions must also have been subject to digestion and any inhibition of restriction is therefore due to methylation at their recognition sites. The three fractions: LS, MS and HS2 were digested and run on a 2% agarose gel. This was photographed to detect the non-radioactive pTEST digests. The gel was dried and radiographic film was used to detect

radioactive digest products. Exposure was carried out for 48hrs at -80°C. The results for column testing are shown in Figure **3-4** page 90.

3.1.2.3 Fractionation of Tumour and Normal Genomic DNA

The HMBD column binds double stranded methylated CpG dinucleotides, with one MBD molecule binding one methylated CpG dinucleotide; this characteristic was used to fractionate genomic DNA into unmethylated DNA (which elutes between 0.1-0.4M NaCl), partially methylated DNA (which elutes between 0.4-0.6M NaCl) and densely methylated DNA (which elutes between 0.6-0.85M NaCl).

100ug of normal breast carcinoma genomic DNA was digested with the restriction enzyme MseI (recognition site TTAA) to leave TA overhangs. This digest cleaves genomic DNA into smaller fragments whilst maintaining the integrity of the majority of CpG islands. MseI was heat inactivated, an aliquot of the digest was run on a 0.8% agarose gel, the remainder was diluted into 15mls of 0.1M NaCl buffer A and loaded onto an HMBD column. The normal HS2 (NHS2) eluate was collected, precipitated and ligated to catchlinker 2.1/2.11 (referred to as 2.1/2.11-NHS2 from now on). T4 DNA polymerase and dNTPs were then added to the ligation reaction to convert any remaining TA overhangs that had not ligated to 2.1/2.11, to blunt ended DNA.

100ug of breast carcinoma DNA was also digested with MseI, fractionated on the HMBD column and the tumour HS2 (THS2) fraction was collected. The source of breast carcinoma DNA in this case was pooled from of 36 ductal carcinomas (provided by Professor Tim Huang, Ohio University).

3.1.2.4 Subtractive Hybridisation

2.1/2.11-NHS2 was mixed, at a ratio of 3:1, with the THS2 fraction of the tumour DNA. The mixture was heated to 95°C for 5mins. and held at 68°C for 48hrs in hybridisation buffer.

As the mixture anneals over 48hours there are three possibilities:

Tumour DNA will reanneal to tumour DNA- theoretically this should only happen in the case of a sequence that is exclusive to THS2. This homoduplex will have TA overhangs, as a result of the MseI digest.

Tumour DNA will anneal to normal DNA - ideally this will occur when a sequence is common to THS2 and NHS2. The likelihood of common sequences forming heteroduplexes was increased by altering the ratio of NHS2:THS2 (range 3:1-10:1) in the subtractive hybridisation mixture. Heteroduplexes will not have TA overhangs, the NHS2 strand has a 2.1/2.11 catchlinker ligated to it.

Normal DNA will anneal to normal DNA. This homoduplex will not have TA overhangs.

Following hybridisation the mixture was cooled rapidly on ice and the catchlinker 2.4/2.44 was ligated to the subtractive hybridisation mixture. 2.4/2.44 catchlinker requires TA overhangs for ligation, therefore it will only ligate to THS2:THS2 homoduplexes. Again, any overhangs remaining were removed by T4 DNA polymerase and the addition of dNTPs.

The hybridisation mixture was amplified using 2.4 and 2.44 as primers, thus only amplifying THS2:THS2 homoduplexes, and products were cloned into

PGEM®-T Easy. All clones were transformed into XLI-Blue E.Coli and plated on selective media. Prior to cloning the PCR products were run on a gel, to check that the PCR step had worked. If the PCR step was successful a smear would be created on the gel. Cloning was not yielding sufficiently sized products, therefore the region of the gel that contained products greater than 500 base pairs was gel purified and used for cloning purposes. (See Figure 3A)

3.1.2.5 Subtractive Hybridisation Screening

Previous experience by Dr. Graeme Brock with the ICEAMP technique showed that 26% of 173 sequences isolated following library generation met the criteria of CpG islands (Brock, Huang et al. 2001).

To increase the efficiency of screening for CpG islands three additional steps were added to the original protocol:

Screening for repetitive elements using DIG labelled COT-1 DNA; repetitive elements made up 19% of the 173 clones isolated.

Screening for sizes >350bp; the average size of a clone in the original ICEAMP libraries was 350bp and the smaller clones were low in CpG islands, they tended to be bulk genomic DNA. (See Figure 3B)

Screening for HpaII restriction enzyme sites; these are frequent in CpG islands, but not in genomic DNA (see Figure 3C).

3.1.2.5.1 Screening Using a DIG Labelled COT-1 DNA Probe

A subtractive hybridisation library was plated on selective media. Each colony was used to seed a single well of a 96 well plate containing 100ul of specified medium. Colonies were grown overnight at 36°C then 10ul of each well was spotted onto a nitrocellulose membrane, which was placed on a selective agar medium, and again grown overnight at 36°C. The colonies were lysed, as

described in Materials and Methods, and DNA was cross-linked to the membrane using UV radiation.

Colonies were then screened with COT-1 DNA to identify repetitive elements: COT-1 DNA was labelled using the DIG-labelling system, as described in Materials and Methods, and hybridised to the nitrocellulose membrane. The membrane was dried and incubated with radiographic paper for 24 hours at -80°C. The autoradiograph was developed and assessed for COT-1 positive clones.

The COT-1 probe is rich in the repetitive element sequences that constituted 19% of previous subtractive hybridisation libraries. If a clone hybridised with the DIG labelled probe it was presumed to be due to the presence of a repetitive element in that clone and hence it was discarded as a potential CpG island.

3.1.2.5.2 Screening Using PCR and Restriction Digest.

Clones that made it through the COT-1 DIG labelled screen were seeded into 100ul volume of ddH2O, heated to 100°C, to lyse bacteria, and 1ul of the lysate was used as a PCR template. PCR using vector specific primers that spanned the cloned insert was carried out as described in Materials and Methods. Products were run on an agarose gel and assessed for size. If a clone was greater than 350bp it was thought to have potential as a CpG island and was selected for restriction enzyme digest with HpaII. The restriction enzyme digest gives an idea of the CpG and GC density of a clone. After this third and final level of selection any clones that met the criteria of a CpG island, and were specifically methylated in tumour but not normal DNA were sequenced.

3.2 Results

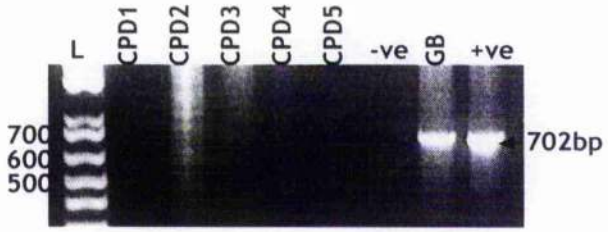


Figure 3-5 PCR for the HS3ST2 CpG Islands using an HMBD Subtractive Hybridisation Libraries

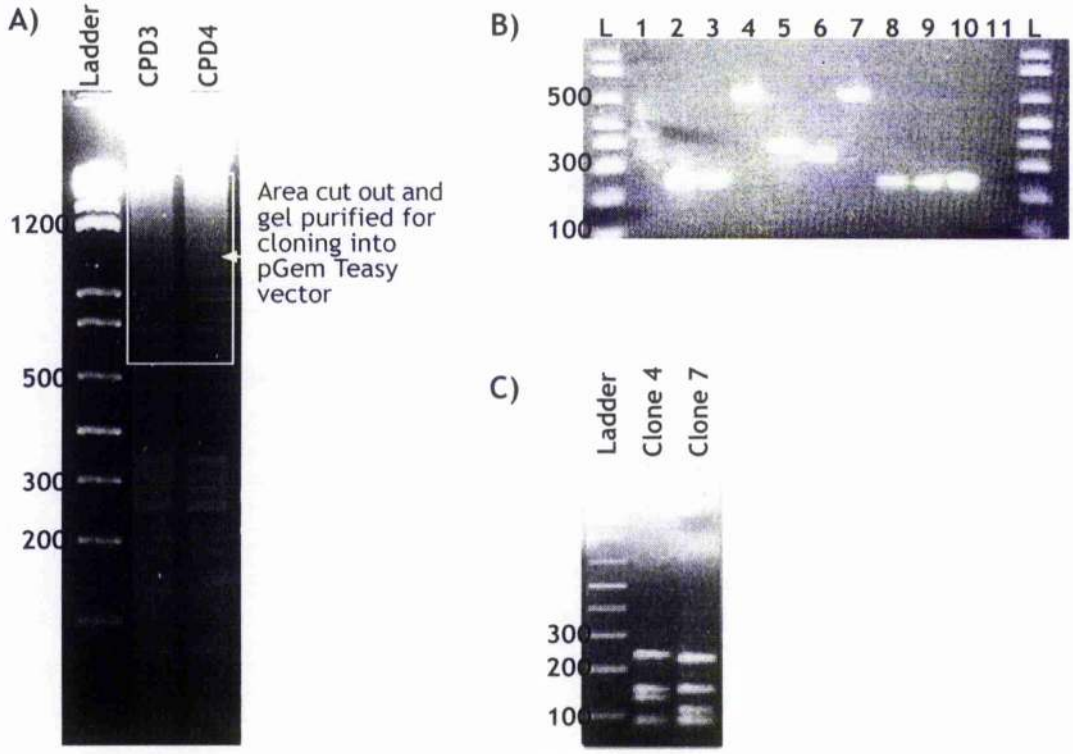


Figure 3-6 Preparation and Screening of Subtractive Hybridisation Libraries. A) PCR of the subtractive hybridisation libraries CPD3 and CPD4 with the catchlinker 2.4/2.44. In this case the area within the white box was cut out and gel purified to select larger fragments for cloning. Following screening of 490 clones for repetitive elements with DIG labelled COT-I DNA only 10 clones were selected for PCR analysis. B) PCR analysis of 10 clones from the CPD3 and CPD4 libraries. Of the 10 clones selected only two, clones 4 and 7 in this case, were large enough to merit further screening. It should be noted when interpreting the size of the PCR products in gel B that the primers used to amplify clones integrated 120 base pairs of the plasmid vector into the product, this was taken into account when assessing the size of cloned inserts. C) Restriction enzyme digests of clones 4 and 7 with HpaII. Each clone has three HpaII sites. There is one HpaII site in the 120bp vector sequence amplified by the PCR reaction, thus there are only 2 HpaII sites in both sequences. Therefore, the clones were not rich enough in HpaII sites to consider them as CGIs.

3.3 Discussion

Interpretation of what went wrong is difficult. In Dr. Brock's original ICEAMP paper, 173 sequences were selected and sequenced to identify their nature. The same approach was not used in this case. Rather than seeking to characterise all sequences generated by ICEAMP the purpose of the exercise was to screen for CpG islands. No samples made it through the screening process and therefore the identity of clones in the libraries was never revealed. It may be asked why the clones isolated were never sequenced and the answer to this is simply that none of them at that time appeared to merit sequencing - none passed the criteria set down at the inception of the project: they were either screened out at the COT1-DNA hybridisation stage, or they were of insufficient size to be a CpG island. In retrospect, for the sake of this thesis, it might have been useful to have sequence data for a few of these clones to prove the assumption that they were repetitive elements or indeed non-CpG island elements.

Review of this screening process is important. Firstly, could the COT-1-DIG screen have been overly sensitive, resulting in false positives? This could explain the lack of CpG islands in the libraries. To control for such an event each clone was grown in a single well of a 96 well plate. In each plate there were four clones known to have a repetitive element, an Alu, as positive controls. There were also four clones that had the HMLH1 CpG island, as negative controls. On two occasions only the screening process did result in either the CpG island clone hybridising with COT1 DNA, or the Alu clone failing to hybridise. In this situation selection of the clones based on that nitrocellulose (NC) membrane was not made, a new NC membrane was spotted with the colonies and a further hybridisation carried out. Without sequencing all clones from the libraries it would be hard to say that the COT-1-DIG screening process was producing false positives, but they were being controlled for.

More than 90% of clones from each library were excluded on the basis of the COT-1-DIG signal. This suggests that the libraries were extremely rich in repetitive elements. Why this was so is difficult to say, either the methylated DNA from the normal was not removing the repetitive elements from the tumour DNA in the subtractive hybridisation stage or the tumour DNA had an abundance of repetitive methylated elements in comparison to the normal DNA.

Efficacy and reproducibility of the fractionation process were issues that may have affected the content of the HS2 fraction when genomic DNA was fractionated. Using plasmid DNA as an assay of the column's functionality may not have been sensitive enough to trouble shoot a column's efficacy to fractionate genomic DNA. Fractionation of DNA on the column could have been assayed using female genomic DNA, assessing the ability to separate an imprinted genomic sequence such as the monoamine oxidase gene from the rest of the genome, rather than a plasmid DNA sequence from other plasmids. This would have been a more specific test of a column's ability to fractionate genomic DNA based on methylation at a CpG island.

The mechanism used to produce a salt gradient may have affected the reproducibility of the fractionation process. The columns produced were not used on an HPLC system, they were in an open system at atmospheric pressure and the salt gradient was stepped rather than linear. An open system and a stepped salt gradient had been used previously in the laboratory to produce subtractive hybridisation libraries, however when problems arise the more standardised a technique is the easier it is to track down the source of a problem. The deviations from the original technique may have made a difference to the ratio of methylated DNA:bulk genomic DNA in HS2 fractions.

Several HMBD columns were manufactured successfully and all fractionated pTEST plasmid on the basis of methylation density Figure **3-4** page 90.

However, this set of experiments aimed to isolate CpG islands that were exclusively methylated in breast tumours.

Enzyme	Cut Position	Fragment Size
HpaII	205	236
	207	205
	261	54
	497	32
	520	23
		2
HhaI	330	330
	370	182
		40

Table 3-4 HMLH1 CpG Island HpaII and HhaI Restriction Sites and Product Sizes.

The table above shows the position of the HhaI and HpaII restriction sites within the HMLH1 CpG island which is a 552bp sequence spanning the start of transcription (-60bp from the TATA box) and extends into the first exon of the gene. The figure below shows the restriction sites in a diagrammatic form.

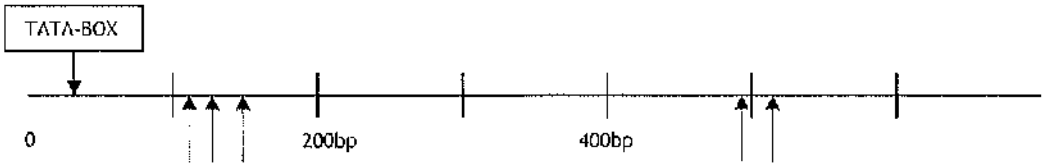


Figure 3-7 HMLH1 CpG Island HpaII and HhaI Restriction Sites

The figure above shows a line diagram of the CpG island of the HMLH1 gene. The TATA box is shown at position 60 and the sites of the restriction sites of the enzymes HpaII are marked by the red arrows and the position is given in Table 3-4.

In retrospect it cannot be stated whether or not the process failed due to the lack of sequencing data. Not sequencing some of the clones, especially the ones that were above 350bp in size was incorrect. HpaII digestion of these clones was only an indirect measure of the CG content, it did not give any indication of the nature, or the genomic location of the DNA and was an incomplete assessment of the identity of these clones. Another criticism of the last step in the assessment of the clones by HpaII digestion is apparent if the

size of the HMLH1 fragments following restriction are assessed, Table 3-4. Four are smaller than 100bp in size, one is only 2bp in size, one is 23bp, one 32bp and one 53bp. The 2bp fragment would definitely not be seen on an agarose gel and it is doubtful if the 23bp and 32bp fragments would be identified. This would leave three fragments visible, this is only one more than the number seen in clones 4 and 7 from the CPD3 library and these were only representative results from one library. The decision to exclude clones as non-CpG island sequences on the basis of an HpaII restriction digest may have been correct; however it is not proven on the basis of the analysis carried out.

3.4 Conclusions

The HMBD columns were capable of binding and fractionating pGem-test plasmids dependent on methylation status.

The subtractive hybridisation libraries were thought to be composed of repetitive elements, and non-CpG island sequences. This judgement was based on the restriction of clones with an HpaII digest. It has been shown that for the HMLH1 CpG island, at least, HpaII digestion may not be a good assay of CG content of a DNA sequence. Lack of sequence data makes it impossible to say whether there were any CpG islands in the clones isolated and failure to sequence the clones was an error.

Further work was required to assess the efficacy of the subtractive hybridisation process and screening steps.

4 Methylation Status of the HS3ST Family of Genes in a Cohort of 80 Breast Carcinomas

4.1 Introduction

The subtractive hybridisation libraries CPD1-CPD6 produced using the HMBD columns did not reveal any CpG island methylation in the group of breast tumours analysed. Whilst examining these libraries for CpG islands other subtractive hybridisation libraries previously produced in the laboratory by Dr Graham Brock were also being screened. CpG islands from two genes of the same family were found in one of these subtractive hybridisation libraries: Heparan sulphate D-glycosaminyl-3-O-sulphotransferase 3B (HS3ST3B) and Heparan sulphate D-glycosaminyl-3-O-sulphotransferase 1 (HS3ST1). This result prompted interest in the methylation status of this gene family.

A search of the literature and genome databases found one publication on Heparan sulphate D-glycosaminyl-3-O-sulphotransferase 2 (HS3ST2) methylation and breast cancer (Miyamoto, Asada et al. 2003). Using MSPCR, the authors of this study reported methylation of the HS3ST2 CpG island in seven breast cancer cell lines, 77/85 breast tumours, 8/10 colon tumours, 7/10 lung tumours and 10/10 pancreatic tumours. They also reported decreased transcription of HS3ST2 in 37 breast tumours that were methylated at the CpG island. This association was also found in the 7 breast cancer cell lines, and re-expression of HS3ST2 in cell lines was achieved by demethylation of the CpG island.

HS3ST3B CpG island methylation, following isolation using ICEAMP(Brock, Huang et al. 2001), was also previously reported in 30/93 breast tumours (Chen, Duncan et al. 2003). Furthermore, hypermethylation of this locus, in combination with 5 to 7 other CpG islands, was significantly correlated with hormone receptor status, clinical stages, and age at diagnosis of the patients analysed (Chen, Chen et al. 2003; Chen, Duncan et al. 2003).

The human genome database and Ensembl databases contain six entries to date for the heparan sulphate D-glycosaminyl-3-O-sulphotransferase gene family: HS3ST1, HS3ST2, HS3ST3A, HS3ST3B, HS3ST4, and HS3ST5 (Shworak, Liu et al. 1999; Xia, Chen et al. 2002). All have a CpG island apart from HS3ST5.

These six enzymes differ in substrate specificity but all catalyse the same reaction (Yabe, Shukla et al. 2001), namely the transfer of a sulphuryl group to

the 3' position of glucosamine within the polysaccharide heparan sulphate(Liu, Shworak et al. 1999). This modification is the last in a series of alterations that take place in the Golgi apparatus during heparan sulphate synthesis (David 1991). Biologically, heparan sulphates are important for various functions (Sasisekharan, Shriver et al. 2002) Table 4-1102and HS3ST1 activity is specifically required to produce the active site in antithrombin IIIa(Zhang, Lawrence et al. 2001).

Heparan sulphates are large polysaccharides constructed from a disaccharide unit: there are up to 200 of these disaccharide units in a heparan sulphate molecule. The monomer is constituted from three sugars: iduronic acid or glucuronic acid which is bound to glucosamine. The backbone of the heparan sulphate chain is constructed by heparan sulphate polymerase, the sugar is then modified by heparan sulphate N-deacetylase and N-sulphotransferase, then uronosyl C5-epimerase. At the end of this sequence of reactions the sugar is sulphated, firstly at the 2' position, then the 6' position and finally HS3ST enzymes transfer a sulphuryl group to the 3' position. These latter modifications split the heparan sulphate into different zones; sulphated, non-sulphated and transition zones. The sulphated zones are thought to harbour important domains for biological functions such as growth factor binding and enhancement of enzymes e.g. antithrombin IIIa (Liu, Shriver et al. 2002; Sasisekharan, Shriver et al. 2002; Shriver, Liu et al. 2002).

Previous modifications to heparan sulphate molecules determine where 3-O-sulphation by HS3ST enzymes occurs (Yabe, Shukla et al. 2001). 3-O-sulphation was first described in 1969 (Danishefsky, Steiner et al. 1969) and accounts for <5% of the sulphuryl groups found on a heparan sulphate molecule. This rare sulphation reaction is performed by six different enzymes, all of which have different substrate specificities. This suggests 3-O-sulphation creates an important functional group that has specific biological activities.

The importance of heparan sulphates in cancer has been well documented (Liu, Shriver et al. 2002; Liu, Shriver et al. 2002; Sasisekharan, Shriver et al. 2002; Shriver, Liu et al. 2002) but the relationship of specific heparan sulphate modifications to cancer biology is not as clear or well researched. There is evidence that heparan sulphate sulphation patterns are important, even

necessary for certain biological functions: the two paradigms are HS3ST1 and antithrombin IIIa function (HajMohammadi, Enjyoji et al. 2003): in short heparin sulphate dramatically enhances the neutralization of coagulation proteases by antithrombin IIIa, provided it has been sulphated at appropriate sites by HS3ST1. Another paradigm is the well studied relationship between heparan sulphate, FGF receptor dimerisation and FGF (Ishihara, Guo et al. 1993) (Kinnunen, Huang et al. 2005). The sulphation pattern of a heparan sulphate covalently linked to the FGF monomer is important for dimerisation of the bFGF receptor, lack of sulphation inhibits dimerisation and prevents signal transduction. In addition it has been shown that Hs3st-B is involved in Notch signalling in *Drosophila melanogaster* by affecting stability or intracellular trafficking of Notch protein. Reduction of Hs3st-B function by transgenic RNA interference compromised Notch signaling, producing neurogenic phenotypes (Kamimura, Rhodes et al. 2004).

From this evidence it appears that sulphation patterns of heparan sulphates are important and in some cases necessary for their function. The FGF family is large and if heparan sulphate sulphation patterns are important in bFGF signalling then they may also be important in signaling of receptors in this family.

The biological significance of heparan sulphates, the importance of heparan sulphation patterns, the rarity of heparan sulphate 3-O-sulphation and the relatively high number of enzymes required for this modification indicate that this is an important family of enzymes which may have a role to play in carcinogenesis. We therefore sought to define the methylation status at each HS3ST CpG island in a cohort of 80 breast tumours.

Table 4-1 Proteins of the cellular microenvironment bound by heparin/heparan sulphate (partial list)

Morphogenesis and Tissue Repair	Host Defence	
Morphogens	Coagulation	Anti-angiogenic factors
Activin	Antithrombin III	Angiostatin
BMP-2, -4	Factor Xa	Endostatin
Chordin	Leuserpin	Cell adhesion molecules
Frizzled-related peptides	Tissue factor pathway inhibitor	L-selectin
Sonic hedgehog		MAC-1
Sprouty peptides	Thrombin	N-CAM
Wnts (1-13)	Growth factors	PECAM-1
ECM components	EGF family	Chemokines
Fibrin	Amphiregulin	C-C
Fibronectin	Betacellulin	CXC
Interstitial collagens	Heparin-binding EGF	Cytokines
Laminins	Neuregulin	IL-2, -3, -4, -5, -7, -12
Pleiotropin (HB-GAM)	FGFs (1-15)	GM-CSF
Tenascin	IGF-II	Interferon
Thrombospondin	PDGF-AA	TNF-
Vitronectin	TGF-1,2	EC superoxide dismutase
Tissue remodelling factors	VEGF-165, 189	pro-arg-rich antimicrobial peptides
Tissue plasminogen activator	Growth factor binding proteins (BP)	Bac-5, -7
		PR-39
Plasminogen activator inhibitor	Follistatin	Energy metabolism
	IGF BP-3, -5	Agouti-related protein
Protease nexin I	TGF- BP	ApoB, ApoE
	Proteinases	Lipoprotein lipase
	Neutrophil elastase	Triglyceride lipases
	Cathepsin G	

Abbreviations: BMP, bone morphogenetic protein; FGF, fibroblast growth; EGF, epidermal growth factor; IGF, insulin-like growth factor; PDGF, platelet-derived growth factor; TGF, transforming growth factor; IL, interleukin; GM-CSF, granulocyte-monocyte colony stimulating factor; TNF, tumor necrosis factor; N-CAM, neural cell adhesion molecule; PECAM, platelet-endothelial cell adhesion molecule; Bac, bactinecin.bCell adhesion molecules are also involved in morphogenesis.

4.1.1 PCR for HS3ST2 and HS3ST4 CpG Islands in Subtractive Hybridisation Libraries

Prior to embarking on a full investigation of methylation status of the HS3ST genes in a large cohort of breast tumours a screen of a successful hybridization library (constructed by Dr Graeme Brock) for the presence of the HS3ST2 and HS3ST4 CpG islands was made. A successful library was determined by assessing the number of methylated CpG islands that had been isolated from it. A PCR product was identified in this library for both the HS3ST2 and HS3ST4 products.

Figure 4.3 contains a 1% agarose gel and shows a PCR for the HS3ST4 CpG island using a subtractive hybridization library, constructed using a protein column manufactured by Dr. Graham Brock. This figure demonstrates that the CpG island for the HS3ST4 gene is present in this particular library. This infers, assuming that the fractionation process was 100% successful, that this particular CpG island was densely methylated in at least one of the tumours in the 128 which were pooled and used to construct this library. This result is of more significance as it is used as evidence for pursuing the HS3ST family as a target for methylation in breast carcinogenesis. This result is also consistent with the hypothesis that expression of these genes could be modulated through methylation at the CpG islands in breast cancer.

4.1.2 *Tumour Collection and Harvesting of DNA*

Tumour samples were collected in line with the ethics approval given for this project (Appendix 1). In total ninety tumour samples were collected either by signed consent prior to surgery (39 cases), or from tissue bank archives in the Royal Infirmary, Glasgow (51 cases). The tissue in both cases was removed from the specimen, frozen in liquid nitrogen and stored at -70°C. Naked eye assessment of the tissue was made to determine if tumour was present or not and this was confirmed by frozen section histology in all cases. Of the ninety samples collected 80 were analysable. Clinico-pathological data was derived from the original pathology report where possible, or from the entry in the tumour data bank if this was not possible. Prior to MSPCR analysis histology reports were collected where possible for all samples and relevant data was added to a database. The samples were then linked to the clinical information and anonymised to prevent identification, in line with the ethical consent. The identifiers used for the tumours throughout this thesis relate to the database created for this analysis alone and cannot be linked to patient identities.

Table 4-6 gives a summary of the tumour details including the number with unknown clinicopathological variables. The tumours were categorised; by age into pre-menopausal (<50), peri-menopausal or early menopausal (51-65) and late menopausal (>65); by tumour size, as shown in the table; into grades 1,2,or 3; by nodal status, which in this sense is the same as stage; by NPI into low risk (<3.4), medium risk (3.4-5.4) and high risk (>5.4); by ER/PR status and by the presence or absence of lymphovascular space invasion.

ER/PR status was determined positive if either ER or PR was greater than 40%. The reason for grouping ER and PR together was somewhat practical and also biological. The practicality was older tumours from the tumour bank in Glasgow Royal Infirmary only had ER status and not PR status and the newer tumours, which were prospectively collected had both ER and PR status. In either case the presence of a positive result indicated that the tumour would be treated with either tamoxifen or latterly with an aromatase inhibitor. The biological argument is ER or PR positivity indicates a tumour that may be responsive to steroid antagonists. The problem with this approach is there is a group of ER negative tumours in samples gleaned from the tumour bank that are PR positive,

but have been called negative, and this must be remembered when interpreting results.

4.1.2.1 DNA Extraction

DNA was extracted as described in Materials and Methods and Figure 4-4 2% TBE agarose gel of DNA extracted from tumour samples, shows representative extractions from a number of tumours. In the case of any tumour which produced a band that was largely limited to the well and had not progressed or smeared down the gel it was deemed to be of good quality, non - degraded DNA. In comparison tumour CDTD123 was borderline, but acceptable, tumour CDTD100 was of poor quality and was not used for MSPCR and tumour CDTDXX1 was non-existent. The tumours which did not produce good quality DNA were re-extracted a second time and where necessary a third time. In the case of CDTDXX1 the tumour was extracted a second time, unsuccessfully and then eventually yielded DNA on the third attempt. This is why CDTD100 and CDTDXX1 appear in subsequent figures for MSPCR. The complete history of the extractions for all tumours is given in the table in the appendix. However, the table below shows the ten tumours that did not produce any DNA in the course of three extraction attempts. The corresponding year of collection can be found in the appendix (range 1982-1988), they were all from GRI tumour bank and significantly they were from a freezer reported to have had the power supply inadvertently turned off, for an unspecified period of time, on more than one occasion, in the last ten years. With such storage conditions it is likely that samples will degrade.

Tumour ID	Result of DNA Extractions		
CDTDBXX	(-)	(-)	(-)
CDTD4Y4	(-)	(-)	(-)
CDTD2X2	(-)	(-)	(-)
CDTD7A6	(-)	(-)	(-)
CDTD6JQ	(-)	(-)	(-)
CDTD0J6	(-)	(-)	(-)
CDTDCID	(-)	(-)	(-)
CDTD1DC	(-)	(-)	(-)
CDTDD2C	(-)	(-)	(-)
CDTD333	(-)	(-)	(-)

Table 4-2 Summary of tumours that failed to produce DNA.

4.1.3 Methylation Specific PCR for the HS3ST CpG Islands

Methylation specific PCR (MSPCR) was carried out on each tumour sample, according to protocols (Materials and Methods.)

4.1.3.1 MSPCR Controls

In order to control for the reaction itself, each CpG island was cloned using primers given in Materials and Methods. This was then transformed into XLI-Blue E.Coli, grown on selective media and harvested using the Quiagen plasmid extraction kit. Plasmids containing each CpG island were then methylated with SssI Methyltransferase, creating a stock of methylated CpG islands for each gene. These methylated plasmids were modified with sodium bisulphite and used as positive controls for the methylated form of the CpG island.

4.1.3.2 Primer Design

Primers were designed for each region examined with the aim of creating a product that lay within the CpG island of each gene. CpG island status was determined using the web-based program GRAIL, found in the human genome database search engine when performing BLAST searches. The sequence derived from this was then analysed further using the program gene-jockey to determine the GC content, CpG content and presence of HpaII and HhaI sites. The criteria used to determine a CpG island have been stated earlier and we verified that each proposed set of primers spanned a region that met these criteria.

Target sequences were then manipulated to determine the base sequence that would result from treatment with sodium bisulphite for a methylated target sequence and also for an unmethylated target sequence. The web-based program 'Primer 3' was then used to design primers for each sequence (details of primer sequences are given in Materials and Methods.) We aimed to produce primers that had CpG sequences at the 3' end, to reduce the possibility of mis-priming, and in addition we aimed for a high annealing temperature, also to avoid this problem.

4.1.3.3 Determination of MSPCR Results

The determination of what constitutes a positive result and a negative one is a somewhat contentious issue when using MSPCR. Faint bands appearing in some MSPCR reactions, were ruled to be 'negatives' in this study. The term 'faint' in this laboratory was defined as the PCR product being less intense than the 200 base pair band of the DNA ladder when 10ul of the ladder was loaded per well. The 200bp band in the DNA ladder corresponds to approximately 100ng of DNA. The visual limit on an agarose gel is thought to be between 20-50ng. Admittedly this is a somewhat arbitrary method, however it was practical, easily applicable and the use of an internal standard is likely to improve specificity.

Consistency in declaring a sample to be 'negative' or 'positive' overall by MSPCR requires a decision tree of some sort. To tackle this two sets of primers for each CpG island were designed, one set was specific for the methylated form of the CpG island, the other for the unmethylated form. When conducting each MSPCR there was a positive control for the methylated form of each CpG island and two negative controls: unmodified genomic DNA and a PCR mixture without any DNA template. The unmethylated reactions also used the same negative controls.

MSPCR for a tumour sample therefore has the following possible outcomes:

Table 4-3 Possible outcomes in MSPR analysis of a locus

Methylated	Unmethylated	Outcome	Determination
+	+	1	Methylated
+	-	2	Methylated
-	+	3	Unmethylated
-	-	4	Void

Methylated and unmethylated PCRs were set up for each CpG island and the results interpreted as follows: Outcome 4 was determined as void and the reactions were repeated. If the result was the same the DNA was remodified and MSPCR was repeated. Outcomes 1-3 is were considered valid results providing the positive and negative MSPCR controls were appropriate. Using these rules false negatives, resulting from DNA degradation or from failure of reactions were controlled for. False positives were also controlled for through the two negative controls used in methylated MSPCRs.

A tumour sample was tested at least three times for methylation until two identical, concordant valid outcomes were obtained. In all cases, apart from 5 tumours, concordant results were obtained for all three reactions. The tumours and the associated results that did not produce concordant results are given in the Table 4-4 below.

Tumour Database ID.	DNA Extractions			Successive MSPCR's for HS3ST2M				HS3ST2M Final Result	Year	Source
CDTD100	(-)	(-)	(+)	(+)	(-)	(+)	(+)	(+)	2002	F
CDTDBMX	(+)			(+)	(+)	(-)	(+)	(+)	2002	F
CDTDPLY	(+)			(+)	(+)	(-)	(+)	(+)	1989	TB
CDTDC9H	(+)			(+)	(-)	(+)	(+)	(+)	2002	F
CDTDXX1	(-)	(-)	(+)	(+)	(+)	(-)	(+)	(+)	2000	TB

Table 4-4 Tumours which did not give concordant results for MSPCR for the HS3ST2 CpG Island.

Source (TB=Tumour Bank, F=Fresh)

As can be seen from the table all the non-concordant results came from the HS3ST2 gene, and all of the non-concordant results are false negatives. Two of the samples showed difficulty in extraction of DNA, but significantly they are not the oldest tumour samples to be analysed.

Reasons for false negatives for methylation in this context are most probably technical and relate to the possible degradation of samples by repeated freeze thawing, this is a well recognised draw-back of the sodium bisulphite treatment of DNA(Clark, Harrison et al. 1994). Other specific reasons for false negative results for methylation status include degradation of DNA during sodium bisulphite modification causing breakage of the sugar phosphate backbone of the molecule - this may result in degradation of the methylated form of the gene in a sample that has a low proportion of methylated DNA; failure of specific PCR reactions; failure of a well in a PCR block; and also poor technique when setting up numerous reactions at one time. In the context of the number of PCR reactions carried out and the results generated this relates 5 false negative reactions out of 240 (2%).

The experience gained from doing such a large number of PCRs lead to the realisation that the number of times a sample was freeze thawed was related to the likelihood of failed PCRs with that sample. This likelihood could be reduced

by the use of a buffered solution, rather than ddH20 alone for storage, however, this was not 100% effective.

Representative results for all genes, including 2 loci of the HS3ST3BE1 gene are given in the results section. MSPCR for methylated and unmethylated forms of the gene are shown in each figure and tumours are identified by their database identification code. This code is also used for the gels showing DNA extractions. The code can be used to identify the tumour in the appendix.

The figures are accompanied by a table that describes the result for each sample's MSPCR, methylated and unmethylated, for the given gene, and where possible the overall result of this specific round of MSPCR. Again the tumour database identification code is used in the table and can be related to the information in the appendix. It should be noted that some results cannot be interpreted from these gels alone as either a methylated or unmethylated result is absent. These results are indicated as such in the table

4.1.4 Methylation Status of HS3ST2 and HS3ST4 in Normal Breast Tissue

Methylation of the HS3ST gene family in normal breast tissue was assessed using Methylation Specific PCR (MSPCR). This was to determine if CpG island methylation was tumour specific or if it also occurs in histologically normal breast. Methylation of the HS3ST2 and HS3ST4 CpG islands in normal breast tissue was not detected in any of the 20 normal breast samples analysed. Normal breast in this case included normal adjacent breast tissue from mastectomy specimens for breast cancer (15 cases) and normal breast tissue from reduction mammoplasty specimens (5 cases). All samples were assessed histologically to ensure they contained normal breast epithelial tissue only and did not contain carcinoma or benign proliferative disease. Representative results are shown in Figure 4-5 MSPCR for HS3ST2 in normal breast tissue taken from mastectomy specimens, and Figure 4-6 MSPCR for HS3ST4 in normal breast tissue taken from mastectomy specimens

4.1.5 Structure of the HS3ST3B Locus

Table 4-5 below shows the HS3ST gene loci. HS3ST3B and HS3ST3A are homologues, located on chromosome 17. They are more than 90% identical at the nucleotide transcript level and overlap in their substrate specificities (Shworak, Liu et al. 1999). However, HS3ST3B is different from HS3ST3A in having two possible CpG islands. One is located in exon 1 (HS3ST3BE1) and the other is 3kb upstream of this, proximal to the HS3ST3B promoter. It is this proximal CpG island that has been analysed in previous reports (Chen, Chen et al. 2003). Interestingly the human genome database software predicts a transcript from the anti-sense strand at this upstream CpG island, suggesting the CpG island at this locus may be related to HS3ST3B and possibly the putative gene coded on the antisense strand. It is unusual for a CpG island to be separated into two parts and therefore both GC rich regions were analysed separately. For the purposes of identification they were named as HS3ST3B (the upstream CpG island) and HS3ST3BE1 (exon 1 CpG island).

Table 4-5 Chromosome location of the HS3ST Genes

Gene	Location
HS3ST1	4p: 11.08m
HS3ST2	16q:22.73m
HS3ST3A	17p:13.3
HS3ST3B	17p:14.15m
HS3ST4	16q:25.61m
HS3ST5	6q:114.49m

4.1.6 MSPCR Sequencing of the HS3ST2 and HS3ST4 Loci

MSPCR sequencing was carried out on both the HS3ST2 and HS3ST4 loci to determine if the CpG dinucleotides between the primer annealing sites were also methylated. It is expected that gene silencing will be associated with dense methylation of the relevant CpG island [Laird, 2005 #2190.

Ten tumours positive by MSPCR for HS3ST2 methylation and ten tumours positive by MSPCR for HS3ST4 methylation were again subjected to MSPCR and the products were cloned into pGEM Teasy as described in Materials and Methods. These clones were transformed into E.Coli pLys and grown on selective media. Clones were then sequenced as described in Materials and Methods and compared to the non-modified sequence.

4.2 Results

Table 4-6. Summary of clinicopathological variables.

	No. of patients (n=80)
Age	
<50	17 (21%)
51-65	25 (31%)
>65	30 (38%)
Unknown	8 (10%)
Tumour size (mm)	
0-20	16 (20%)
21-50	41 (51%)
>50	7 (9%)
Unknown	16 (20%)
Grade	
1	9 (11%)
2	24 (30%)
3	32 (40%)
Unknown	15 (19%)
Nodal status	
Negative	18 (23%)
1-3 Positive	21 (26%)
>3 Positive	19 (24%)
Unknown*	22 (28%)
NPI	
<3.4	7 (9%)
3.4-5.4	28 (35%)
>5.4	23 (29%)
Unknown	22 (28%)
ER status	
Negative	20 (25%)
Positive	39 (49%)
Unknown	21 (26%)
LV invasion	
Negative	33 (41%)
Positive	26 (33%)
Unknown	21 (26%)

*Includes one patient with inadequate negative sample (only one node sampled).

Table 4-6 above summarises the clinicopathological data associated with the tumour bank that was used for this analysis. Samples that do not have clinicopathological data available for a parameter are labelled as unknown for that variable. The reason for lack of information for that parameter varied with each case and a specific cause is given for each sample in the appendix. However, there were a few primary reasons: 1) The sample did not have a case number attached to it and no clinical data could be retrieved other than what was available in the tumour bank data book 2) The patient had limited surgery, either a very small biopsy for diagnostic purposes or a 'toilet' procedure without axillary sampling for lymph node status.

Some of the variables can be compared to local and national statistics to give an idea if the tumour bank is representative of breast cancers in the UK and locally. The ratio of age groups <50, 51-64 and >64 in the tumour bank is 1:1.5:1.8. Using the online resource from the Scottish Health Statistics the ratio of these age groups for the west of Scotland in 2003 was 1:2:2.4. This discrepancy in proportions may be explained by the tumour bank's probable skew towards non-screen detected younger patients. Sampling screen detected lesions without a definitive diagnosis of invasive malignancy compromises patient safety, this category of lumpectomy specimen was therefore not sampled and probably lead to the change in ratios described above.

The ratio of grades in the tumour bank was 2:5.4:7.2 for grades 1,2 and 3, this is slightly out of step with the grading ratios suggested by the Royal College of Pathologists (Elston and Ellis 1991) which are 2:3:5 for symptomatic breast cancers. The increased ratio of higher-grade cancers may be explained by the sampling of larger tumours and non-sampling of smaller tumours. Larger tumours tend to have higher grades and therefore the tumour bank has a greater proportion of these tumours(Elston and Ellis 1991).

4.2.1 Methylation at CpG Islands and Correlation with Clinicopathological Data.

Table 4-7 Number of Tumours Positive for Methylation at each Locus

Gene	Number Methylated
HS3ST1	0/80
HS3ST2	45/80
HS3ST3A	0/80
HS3ST3B	30/80
HS3ST3BE1	15/80
HS3ST4	34/80

Methylation at each single locus was correlated with the clinicopathological parameters listed in Table 4-6. Fisher's exact test was used to calculate significance of methylation at a locus with respect to ER/PR positivity and lymphovascular invasion. The chi-squared test for trend was used to calculate significance for NPI, nodal status, grade, tumour and age. Where the continuous variables; age, tumour size and NPI were concerned a Mann Whitney test was also performed to test for significance. The association between methylation at different loci was assessed using Fisher's exact test. Methylation of more than one locus and relationship to clinicopathological variables was assessed using the chi-squared test for trend.

The full results of all the statistical analyses carried out are presented in Table 4-14, Table 4-15, Table 4-16, Table 4-17, Table 4-18, Table 4-19, Table 4-20, Table 4-21, Table 4-22 and Table 4-23 at the end of the results section. Correlations with a p value of 0.05 or less are termed significant and those that have a p value between 0.05 and 1 are termed as showing a trend.

The correlations that were significant are as follows:

There is a positive correlation between methylation at the HS3ST3BE1 CpG island and methylation at HS3ST2 (p=0.009) and HS3ST4 (p=0.009).

There is a negative correlation between methylation at the HS3ST3BE1 CpG island and methylation HS3ST3B (p=0.04).

There is a correlation between HS3ST2 methylation and grade (p=0.015) and strangely there is a very strong correlation between lack of HS3ST2 methylation and nodal status unknown (p=0.009) and NPI status unknown (p=0.002).

HS3ST3B methylation is correlated with lymphovascular invasion (p=0.028).

HS3ST3BE1 methylation is correlated with age (p=0.04), NPI (p=0.047) and size (p=0.04).

The number of loci methylated correlates with grade (p=0.04)

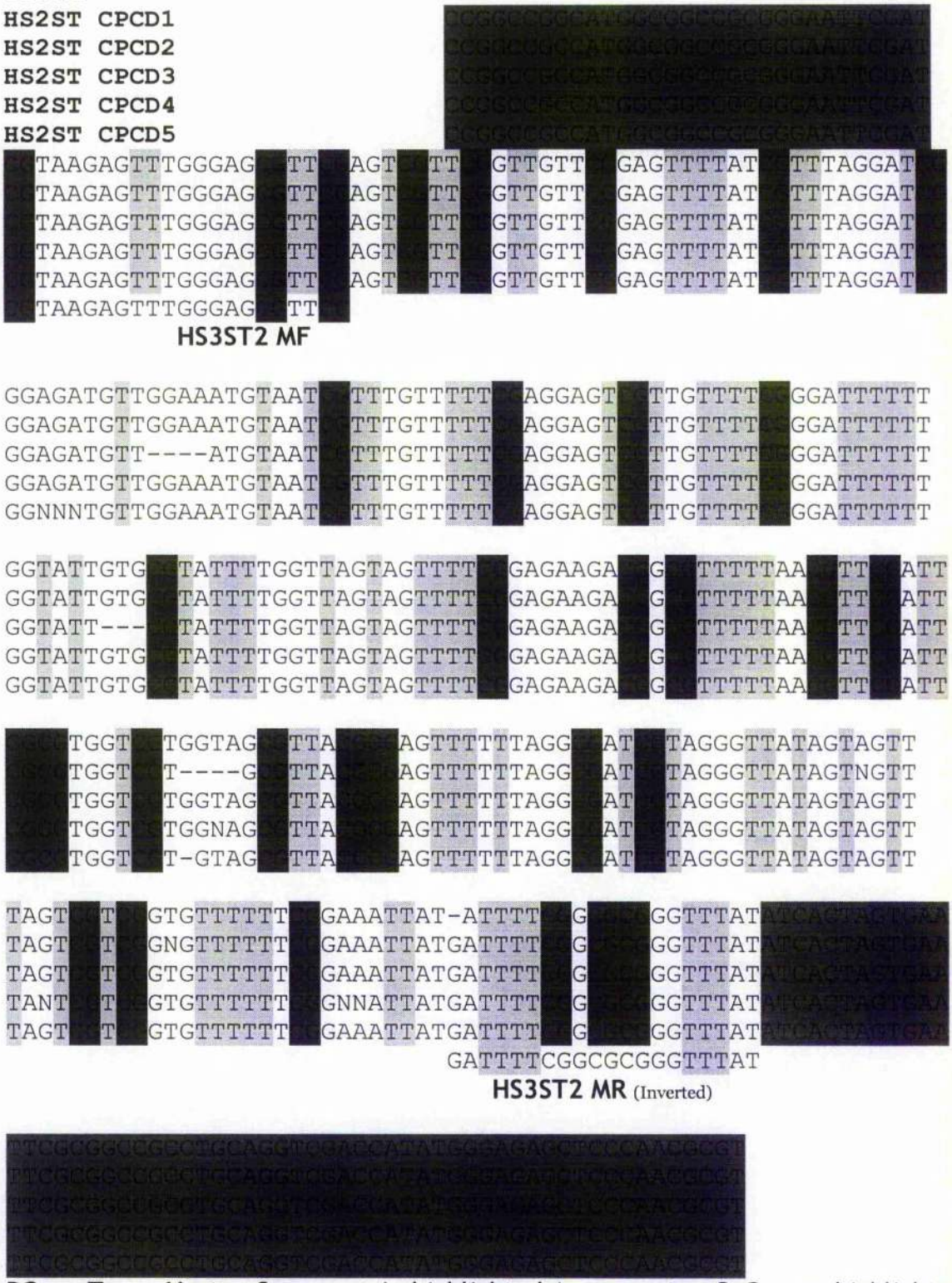
The correlations that show a trend are as follows:

There is a trend for number of loci methylated to be correlated with ER/PR negativity (p=0.054), NPI (p=0.06) and lymphovascular invasion (p=0.096).

4.2.2 MSPCR Sequencing

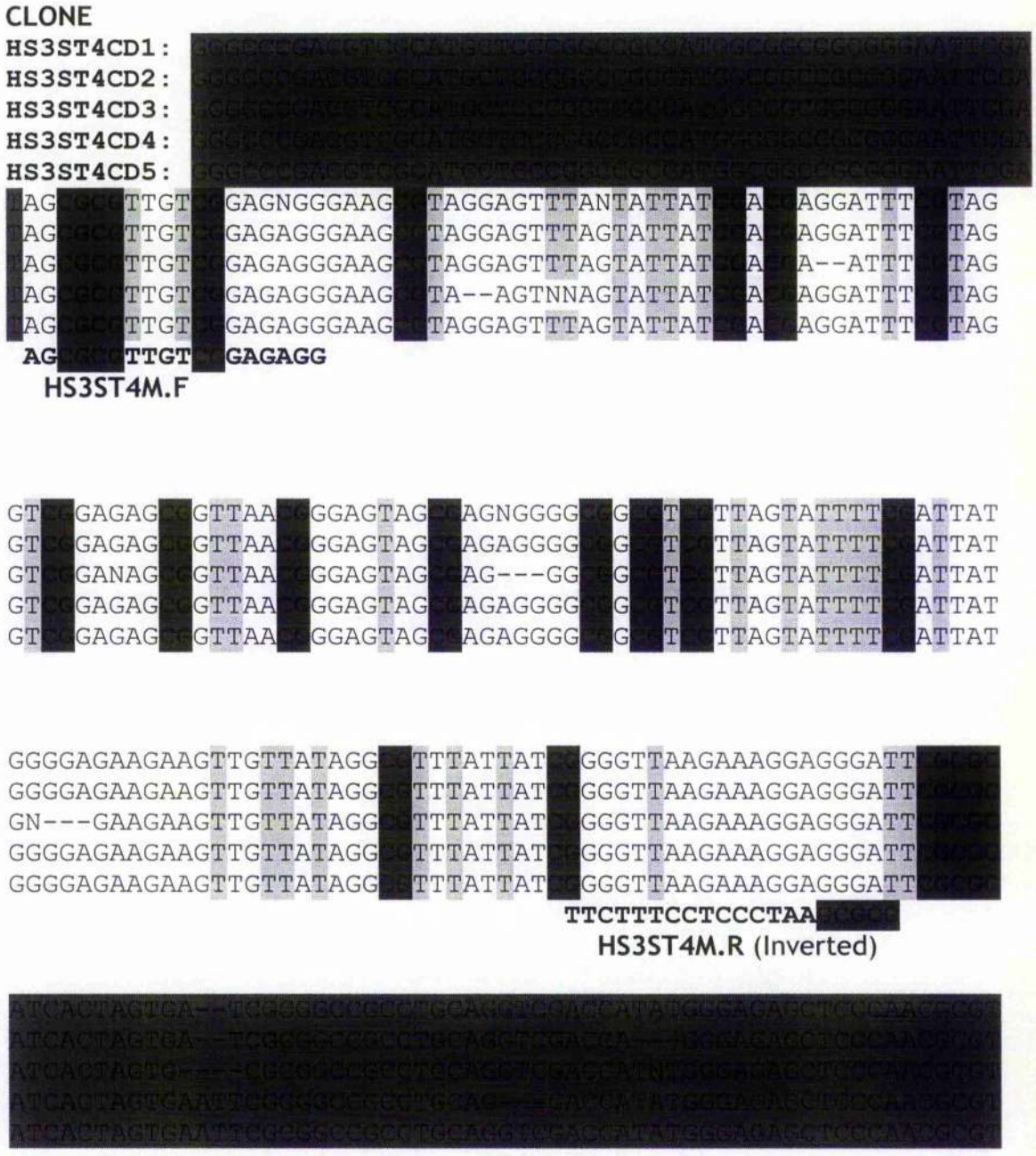
Representative results of MSPCR sequencing for HS3ST2 and HS3ST4 are given in Figure 4-1 and Figure 4-2 respectively. In all cases the CpG dinucleotides between primer annealing sites were methylated, confirming that methylation at the HS3ST2 and HS3ST4 CpG islands is not restricted to the primer annealing sites. In addition all of the intervening cytosines, not found in the context of a CpG dinucleotide, were converted to uracils, proving that this finding is not an artefact of incomplete sodium bisulphate conversion.

Figure 4-1 HS3ST2 MSPCR Sequencing Data



PGem Teasy Vector Sequence is highlighted in magenta. CpGs are highlighted red. Cytosines converted to thymidines are highlighted yellow. Primer sequences are also given and annotated. Clones HS3ST2 CPD1- HS3ST2 CPD5 are examples of sequencing results obtained from MSPCR of tumours followed by cloning into PGem Teasy and sequencing as described in Materials and Methods. Sequences were aligned using the Gene Jockey program. N delineates a nucleotide that was unreadable and a dash indicates an artificial gap created by the program to allow best fit alignment

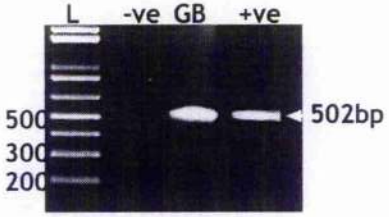
Figure 4-2 HS3ST4 MSPCR Sequencing Data



PGem Teasy Vector Sequence is highlighted in magenta. CpGs are highlighted red. Cytosines converted to thymidines are highlighted yellow. Primer sequences are also given and annotated. Clones CD1-CD5 are examples of sequencing results obtained from MSPCR of tumours followed by cloning into PGem Teasy and sequencing as described in Materials and Methods. Sequences were aligned using the Gene Jockey program. N delineates a nucleotide that was unreadable and a dash indicates an artificial gap created by the program to allow best fit alignment.

Figure 4-3 PCR for HS3ST4 in GB Subtractive Hybridisation Library.

1% Agarose gel run at 100V for 1hr. +ve control (+) is normal genomic DNA in this case and -ve control is PCR mixture with 1ul of ddH20 instead of normal genomic DNA.



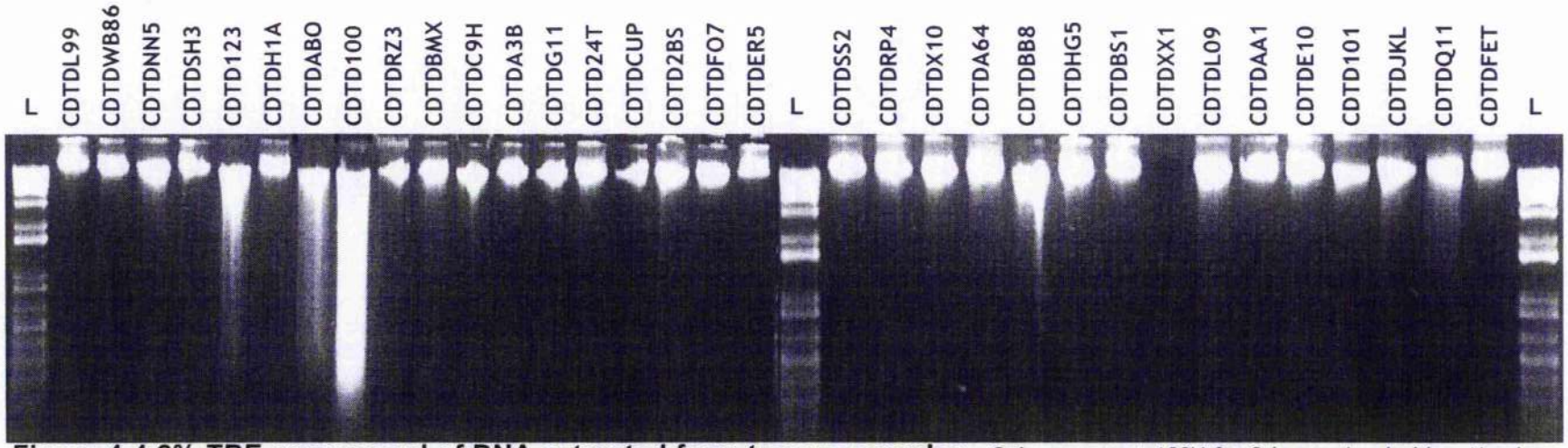


Figure 4-4 2% TBE agarose gel of DNA extracted from tumour samples. Gel was run at 100V for 2 hours. L = ladder. Numbers relate to tumour samples. Molecular marker sizes are given in base pairs. This gel was essentially a quality control step to ensure the DNA extracted for modification with sodium bisulphite was of a high quality. For example, tumour CDTDL99 was deemed to be of good quality, non - degraded DNA. In comparison tumour CDTD123 was borderline, but acceptable, tumour CDTD100 was of poor quality and was not used for MSPCR and tumour CDTDXX1 was non-existent. All samples were quantified on a Beckman UV-spectrophotometer as described in Materials and Methods and 500ng of each sample was added to each lane. This gave an additional level of quality control, ensuring the spectrophotometer was giving accurate readings for each sample. The tumour samples are given their database number and a table showing the result of DNA quality assesment is given below; the position of the tumour on the table in the appendix is given for ease of cross-referencing results

Craig Peter Charles Dick

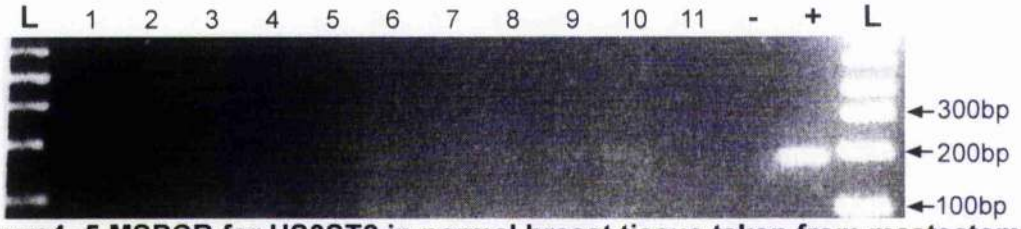


Figure4 -5 MSPCR for HS3ST2 in normal breast tissue taken from mastectomy specimens L=ladder. Marker sizes are given in base-pairs. All tissue was taken from either normal meastectomy specimens for reduction mamoplasty or from normal breast tissue in a mastectomy specimen for tumour.

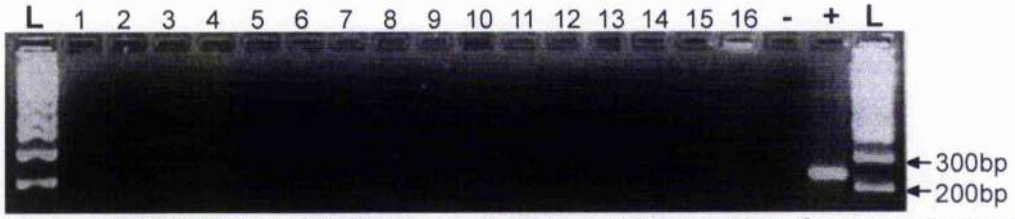


Figure 4 -6 MSPCR for HS3ST4 in normal breast tissue taken from mastectomy specimens L=ladder. Marker sizes are given in base-pairs. All tissue was taken from either normal meastectomy specimens for reduction mamoplasty or from normal breast tissue in a mastectomy specimen for tumour.

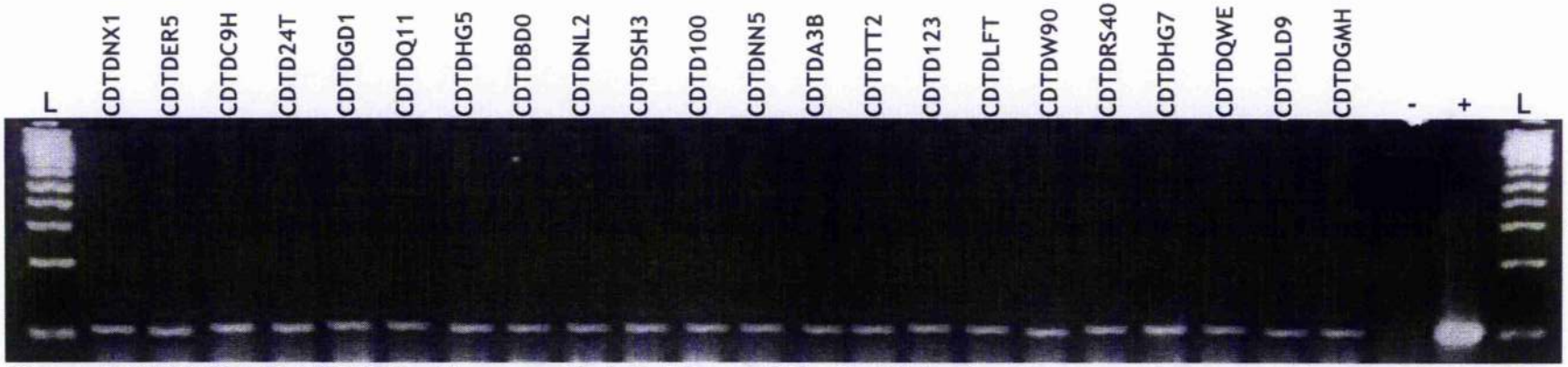


Figure 4.7 HS3ST1UM MSPCR for the unmethylated form of the CpG island. PCR products were run on a 2% agarose gel for 2hrs at 100V. (L)=ladder. Tumours are identified by their tumour database ID code(-) is the negative control, the reaction mixture plus unmodified normal genomic DNA. (+) is a positive control ,the cloned CpG island, then modified by sodium bisulphite. Marker sizes are given in base pairs.

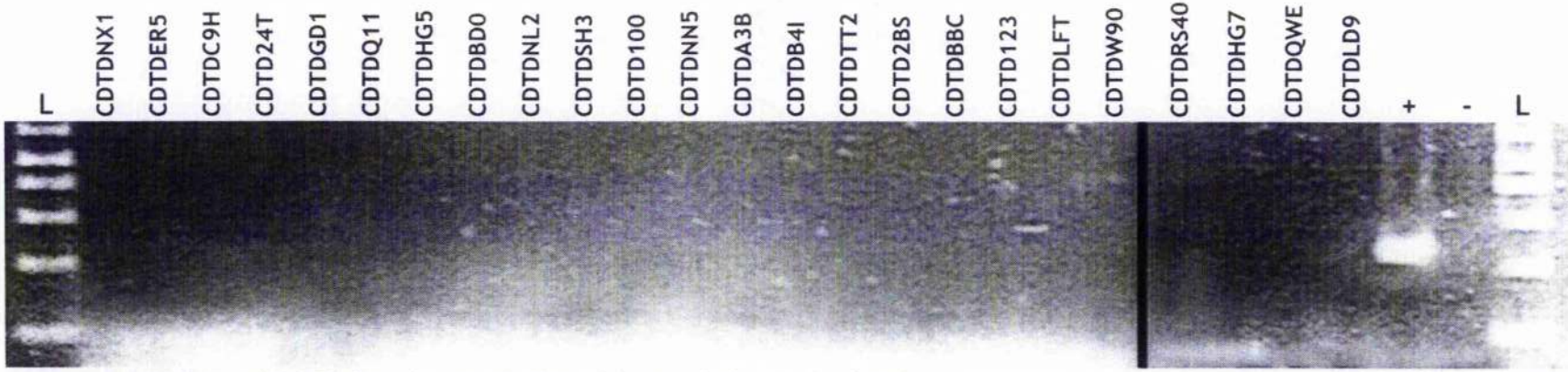


Figure 4.8 HS3ST1 MSPCR for the methylated form of the CpG island. PCR products were run on a 2% agarose gel for 2hrs at 100V. (L)=ladder. Tumours are identified by their tumour database ID code(-) is the negative control, the reaction mixture plus unmodified normal genomic DNA. (+) is a positive control ,the cloned CpG island, then modified by sodium bisulphite. Marker sizes are given in base pairs

21.12.06
Page122
Craig Peter Charles Dick

Table 4-8 Methylation Status of HS3ST1 in Relation to the Tumours Described In the gels on the previous page.

HS3ST1M	HS3ST1UM	Tumour Database ID	Outcome
-	+	CDTDNX1	Unmethylated
-	+	CDTDER5	Unmethylated
-	+	CDTDC9H	Unmethylated
-	+	CDTD24T	Unmethylated
-	+	CDTDGD1	Unmethylated
-	+	CDTDQ11	Unmethylated
-	+	CDTDHG5	Unmethylated
-	+	CDTDBD0	Unmethylated
-	+	CDTDNL2	Unmethylated
-	+	CDTDSH3	Unmethylated
-	+	CDTD100	Unmethylated
-	+	CDTDNN5	Unmethylated
-	+	CDTDA3B	Unmethylated
-	Absent	CDTDB4I	Requires UM
-	+	CDTDTT2	Unmethylated
-	Absent	CDTD2B5	Requires UM
-	Absent	CDTDBBC	Requires UM
-	+	CDTD123	Unmethylated
-	+	CDTDLFT	Unmethylated
-	+	CDTDW90	Unmethylated
-	+	CDTDRS40	Unmethylated
-	+	CDTDHG7	Unmethylated
-	+	CDTDQWE	Unmethylated
-	+	CDTDLD9	Unmethylated
Absent	+	CDTDGMH	Requires M

(-) indicates a negative result for that form of the gene, (+) indicates a positive result for that form of the gene and the outcome is given in the right hand column.

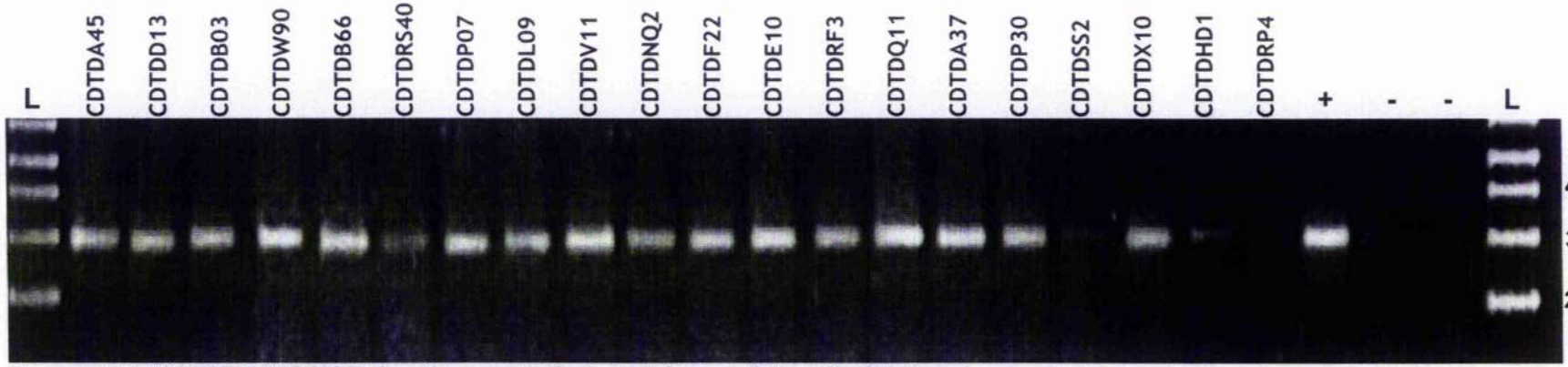


Figure 4.9 HS3ST2 MSPCR for the unmethylated form of the CpG island. PCR products were run on a 2% agarose gel for 2hrs at 100V. (L)=ladder. Tumours are identified by their tumour database ID code(-) is the negative control, the reaction mixture plus unmodified normal genomic DNA. (+) is a positive control ,the cloned CpG island, then modified by sodium bisulphite. Marker sizes are given in base pairs

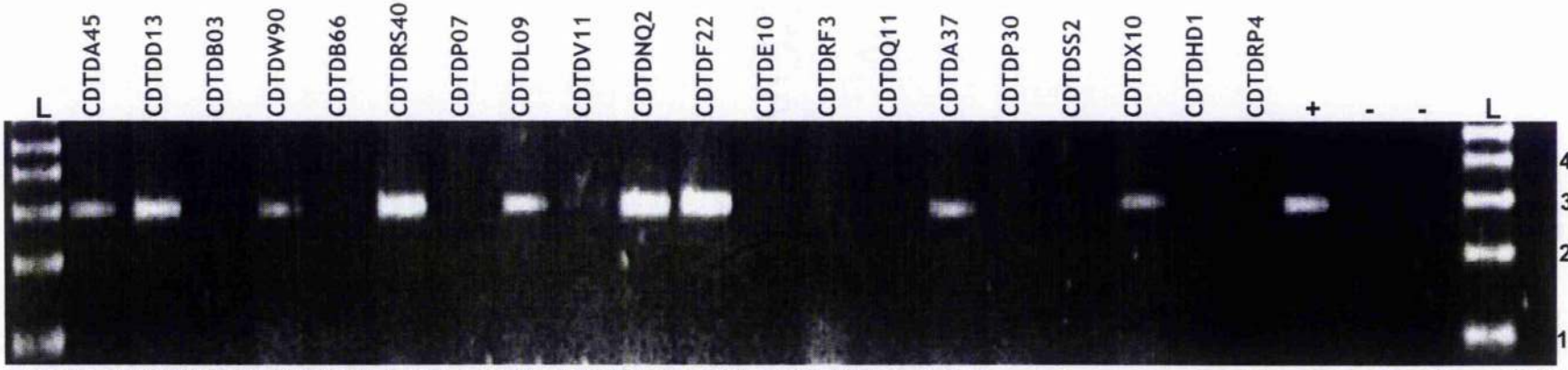


Figure 4.10 HS3ST2 MSPCR for the methylated form of the CpG island. PCR products were run on a 2% agarose gel for 2hrs at 100V. (L)=ladder. Tumours are identified by their tumour database ID code(-) is the negative control, the reaction mixture plus unmodified normal genomic DNA. (+) is a positive control ,the cloned CpG island, then modified by sodium bisulphite. Marker sizes are given in base pairs

21.12.06
Page 124
Craig Peter Charles Dick

Table 4-9 Methylation status of HS3ST2 in relation to the tumours described in the gels in the previous page

HS3ST2M	HS3ST2UM	Tumour Database ID	Outcome
+	+	CDTDA45	Methylated
+	+	CDTDD13	Methylated
-	+	CDTDB03	Unmethylated
+	+	CDTDW90	Methylated
-	+	CDTDB66	Unmethylated
+	+	CDTDRS40	Methylated
-	+	CDTDP07	Unmethylated
+	+	CDTDL09	Methylated
-	+	CDTDV11	Unmethylated
+	+	CDTDNQ2	Methylated
+	+	CDTDF22	Methylated
-	+	CDTDE10	Unmethylated
-	+	CDTDRF3	Unmethylated
-	+	CDTDQ11	Unmethylated
+	+	CDTDA37	Methylated
-	+	CDTDP30	Unmethylated
-	-	CDTDSS2	Unmethylated
+	+	CDTDX10	Methylated
-	-	CDTDHD1	Repeat/Void
-	-	CDTDRP4	Unmethylated

(-) indicates a negative result for that form of the gene, (+) indicates a positive result for that form of the gene, and the outcome is given in the right hand column.

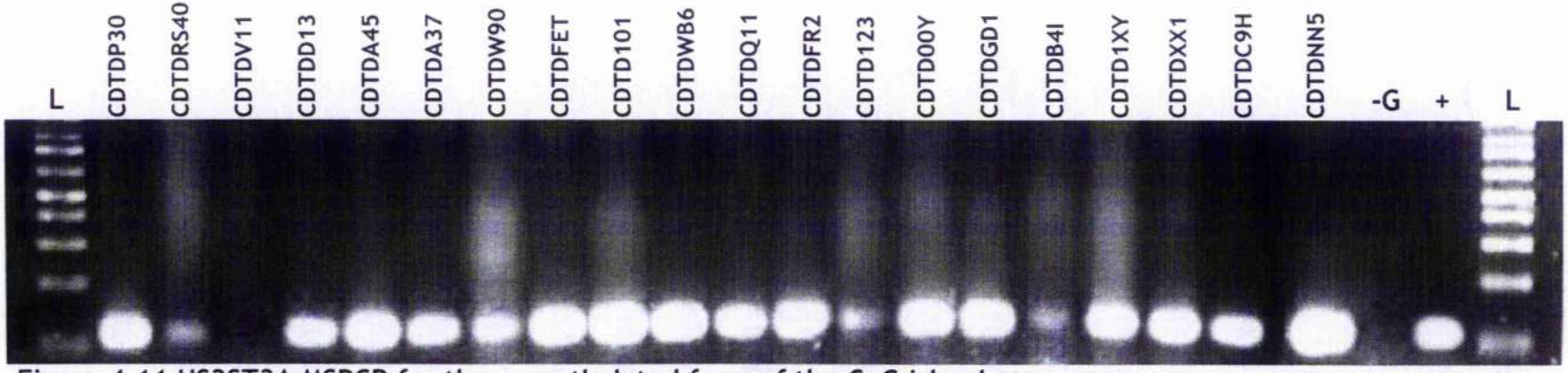


Figure 4.11 HS3ST3A MSPCR for the unmethylated form of the CpG island. PCR products were run on a 2% agarose gel for 2hrs at 100V. (L)=ladder. Tumours are identified by their tumour database ID code(-) is the negative control, the reaction mixture plus unmodified normal genomic DNA. (+) is a positive control ,the cloned CpG island, then modified by sodium bisulphite. Marker sizes are given in base pairs

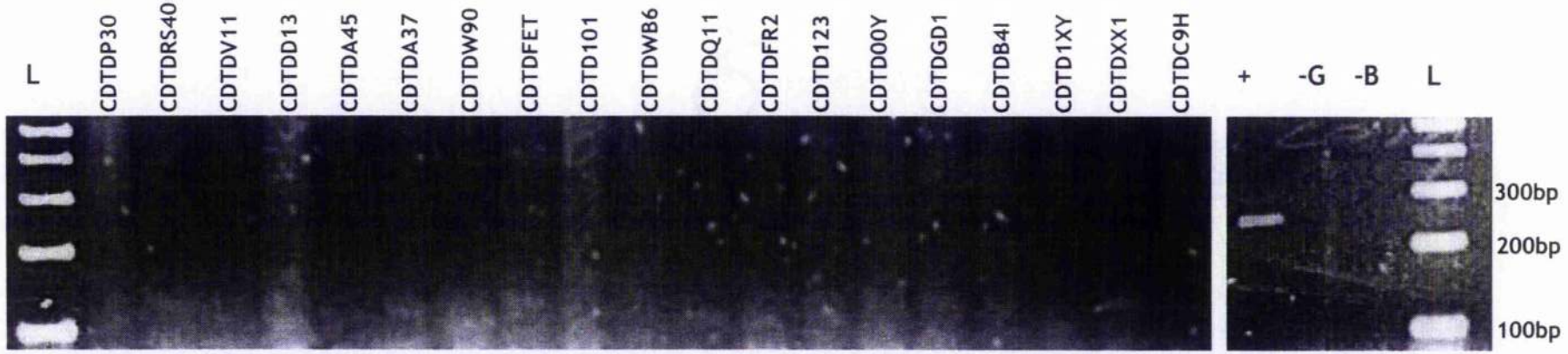


Figure 4.12 HS3ST3A MSPCR for the methylated form of the CpG island. PCR products were run on a 2% agarose gel for 2hrs at 100V. (L)=ladder. Tumours are identified by their tumour database ID code(-) is the negative control, the reaction mixture plus unmodified normal genomic DNA. (+) is a positive control ,the cloned CpG island, then modified by sodium bisulphite. Marker sizes are given in base pairs

Craig Peter Charles Dick

Table 4-10 Methylation status of HS3ST3A in relation to the tumours described in the gels in the previous page

HS3ST3AUM	HS3ST3AM	Tumour Database ID	Outcome
+	-	CDTDP30	Unmethylated
+	-	CDTDRS40	Unmethylated
-	-	CDTDV11	Unmethylated
+	-	CDTD13	Unmethylated
+	-	CDTDA45	Unmethylated
+	-	CDTDA37	Unmethylated
+	-	CDTDW90	Unmethylated
+	-	CDTDFET	Unmethylated
+	-	CDTD101	Unmethylated
+	-	CDTDWB6	Unmethylated
+	-	CDTDQ11	Unmethylated
+	-	CDTDFR2	Unmethylated
+	-	CDTD123	Unmethylated
+	-	CDTD00Y	Unmethylated
+	-	CDTDGD1	Unmethylated
+	-	CDTDB4I	Unmethylated
+	-	CDTD1XY	Unmethylated
+	-	CDTDXX1	Unmethylated
+	-	CDTDC9H	Unmethylated
+	-	CDTDNN5	Unmethylated

(-) indicates a negative result for that form of the gene, (+) indicates a positive result for that form of the gene, and the outcome is given in the right hand column.

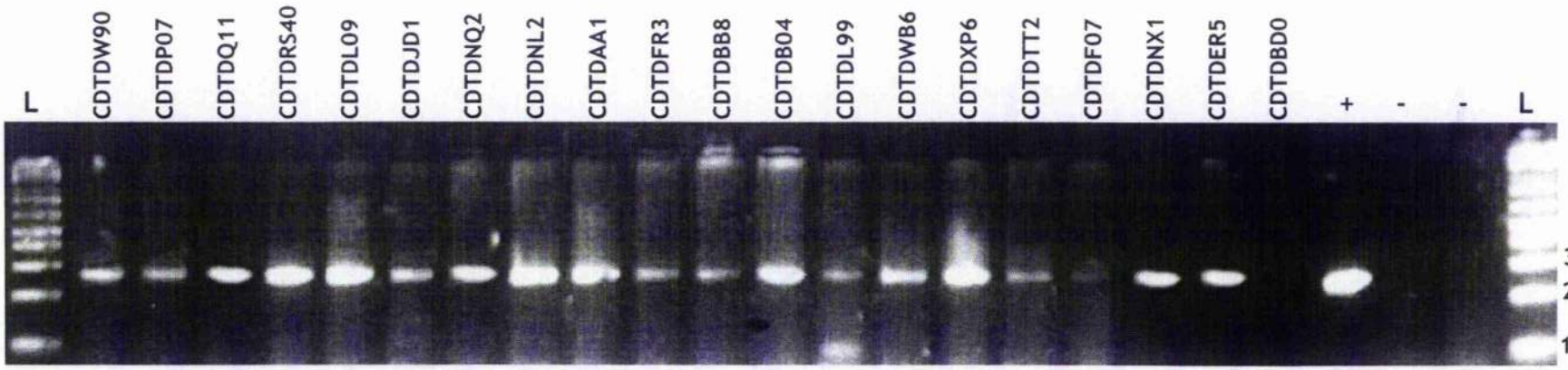


Figure 4.13 HS3ST3B MSPCR for the unmethylated form of the CpG island. PCR products were run on a 2% agarose gel for 2hrs at 100V. (L)=ladder. Tumours are identified by their tumour database ID code(-) is the negative control, the reaction mixture plus unmodified normal genomic DNA. (+) is a positive control ,the cloned CpG island, then modified by sodium bisulphite. Marker sizes are given in base pairs

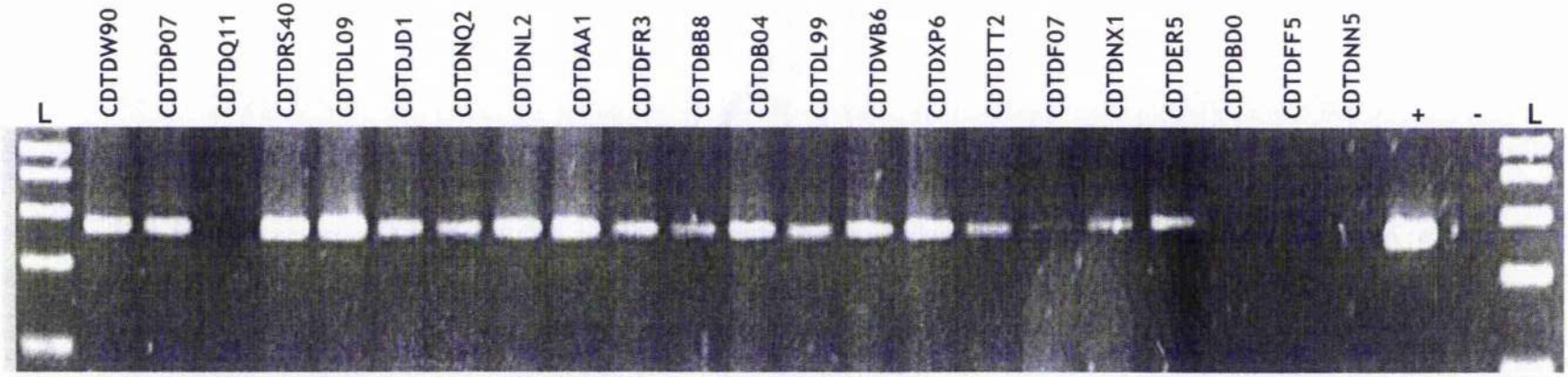


Figure 4.14 HS3ST3B MSPCR for the methylated form of the CpG island. PCR products were run on a 2% agarose gel for 2hrs at 100V. (L)=ladder. Tumours are identified by their tumour database ID code(-) is the negative control, the reaction mixture plus unmodified normal genomic DNA. (+) is a positive control ,the cloned CpG island, then modified by sodium bisulphite. Marker sizes are given in base pairs

23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44

Table 4-11 Methylation status of HS3ST3B in relation to the tumours described in the gels in the previous page

HS3ST3B M	HS3ST3B UM	Tumour Database ID	Outcome
+	+	CDTDW90	Methylated
+	+	CDTDP07	Methylated
-	+	CDTDQ11	Unmethylated
+	+	CDTDRS40	Methylated
+	+	CDTDL09	Methylated
+	+	CDTDJD1	Methylated
+	+	CDTDNQ2	Methylated
+	+	CDTDNL2	Methylated
+	+	CDTDAA1	Methylated
+	+	CDTDFR3	Methylated
+	+	CDTDBB8	Methylated
+	+	CDTDB04	Methylated
+	+	CDTDL99	Methylated
+	+	CDTDWB6	Methylated
+	+	CDTDXP6	Methylated
+	+	CDTDTT2	Methylated
E	E	CDTDF07	Void
+	+	CDTDNX1	Methylated
+	+	CDTDER5	Methylated
-	-	CDTDBD0	Void

(-) indicates a negative result for that form of the gene, (+) indicates a positive result for that form of the gene, and the outcome is given in the right hand column.

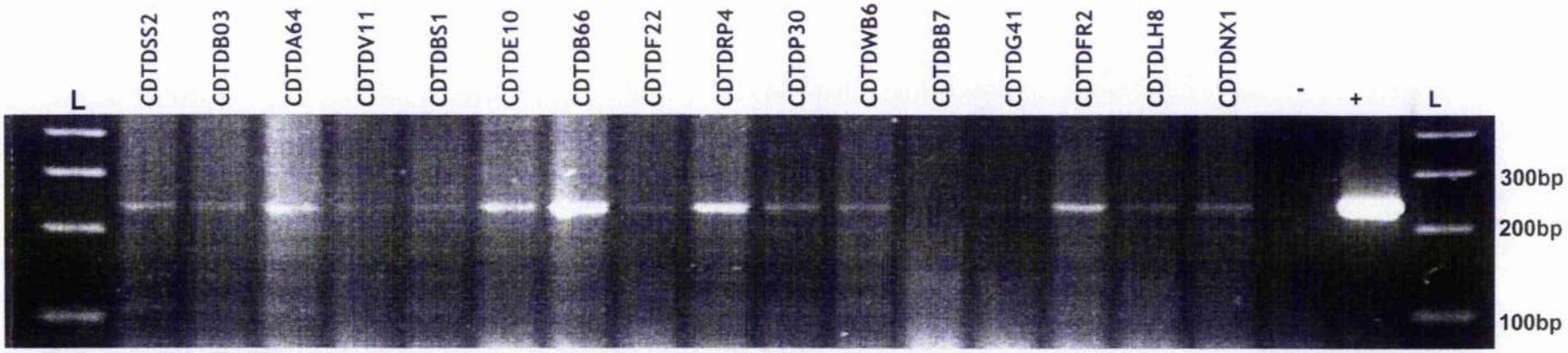


Figure 4.15 HS3ST3BE1 MSPCR for the unmethylated form of the CpG island. PCR products were run on a 2% agarose gel for 2hrs at 100V. (L)=ladder. Tumours are identified by their tumour database ID code(-) is the negative control, the reaction mixture plus unmodified normal genomic DNA. (+) is a positive control ,the cloned CpG island, then modified by sodium bisulphite. Marker sizes are given in base pairs

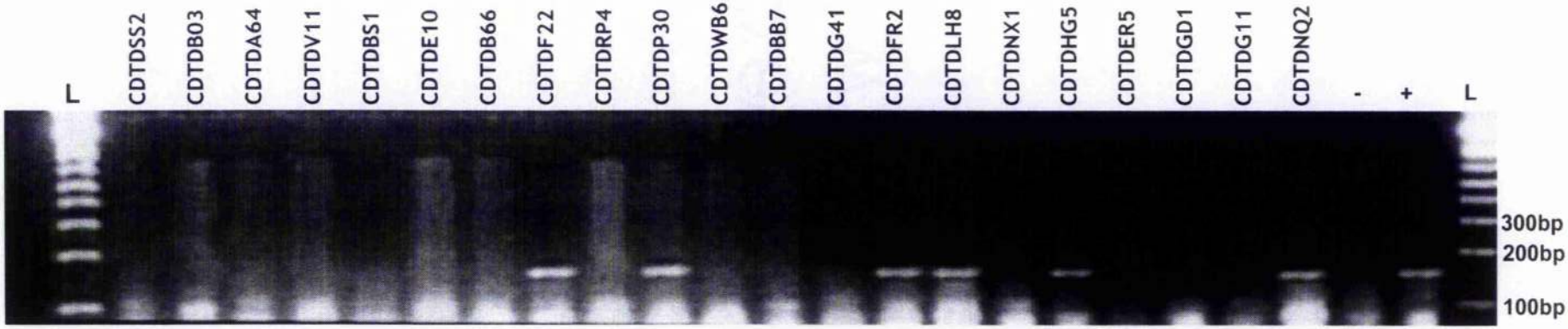


Figure 4.16 HS3ST3BE1 MSPCR for the methylated form of the CpG island. PCR products were run on a 2% agarose gel for 2hrs at 100V. (L)=ladder. Tumours are identified by their tumour database ID code(-) is the negative control, the reaction mixture plus unmodified normal genomic DNA. (+) is a positive control ,the cloned CpG island, then modified by sodium bisulphite. Marker sizes are given in base pairs

Craig Peter Charles Dick

Table 4-12 Methylation status of HS3ST3BE1 in relation to the tumours described in the gels in the previous page

HS3ST3BE1 M	HS3ST3BE1 UM	Tumour Database ID	Outcome
-	+	CDTDSS2	Unmethylated
-	+	CDTDB03	Unmethylated
-	+	CDTDA64	Unmethylated
-	-	CDTDV11	Void
-	-	CDTDBS1	Void
-	-	CDTDE10	Void
-	+	CDTDB66	Unmethylated
+	-	CDTDF22	Methylated
-	+	CDTDRP4	Unmethylated
+	+	CDTDP30	Methylated
-	+	CDTDWB6	Unmethylated
-	-	CDTDBB7	Void
-	-	CDTDG41	Void
+	+	CDTDFR2	Methylated
+	+	CDTDLH8	Methylated
-	+	CDTDNX1	Unmethylated
+		CDTDHG5	Methylated
-		CDTDER5	Requires UM
-		CDTDGD1	Requires UM
-		CDTDG11	Requires UM
+		CDTDNQ2	Methylated

(-) indicates a negative result for that form of the gene, (+) indicates a positive result for that form of the gene, and the outcome is given in the right hand column

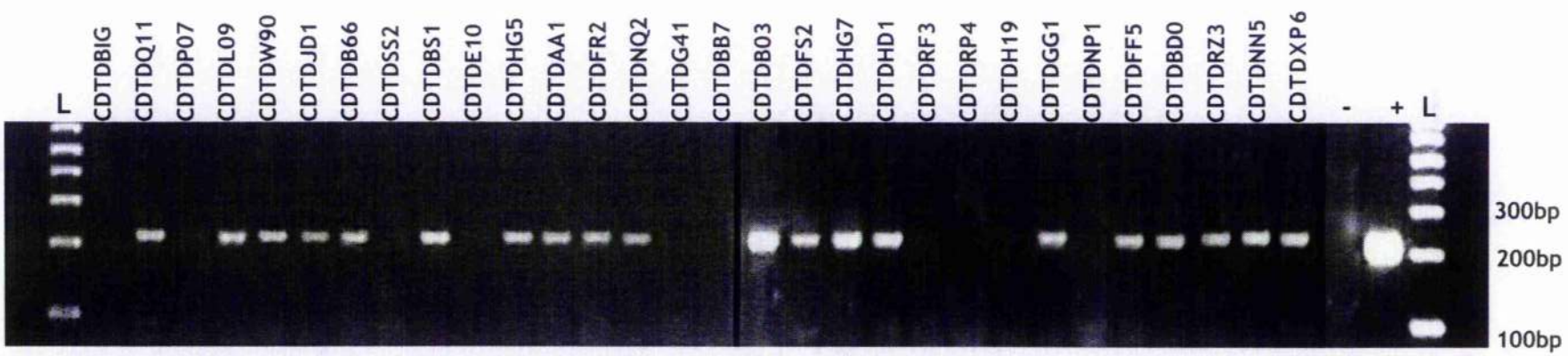


Figure 4.17 HS3ST4 MSPCR for the unmethylated form of the CpG island. PCR products were run on a 2% agarose gel for 2hrs at 100V. (L)=ladder. Tumours are identified by their tumour database ID code(-) is the negative control, the reaction mixture plus unmodified normal genomic DNA. (+) is a positive control ,the cloned CpG island, then modified by sodium bisulphite. Marker sizes are given in base pairs

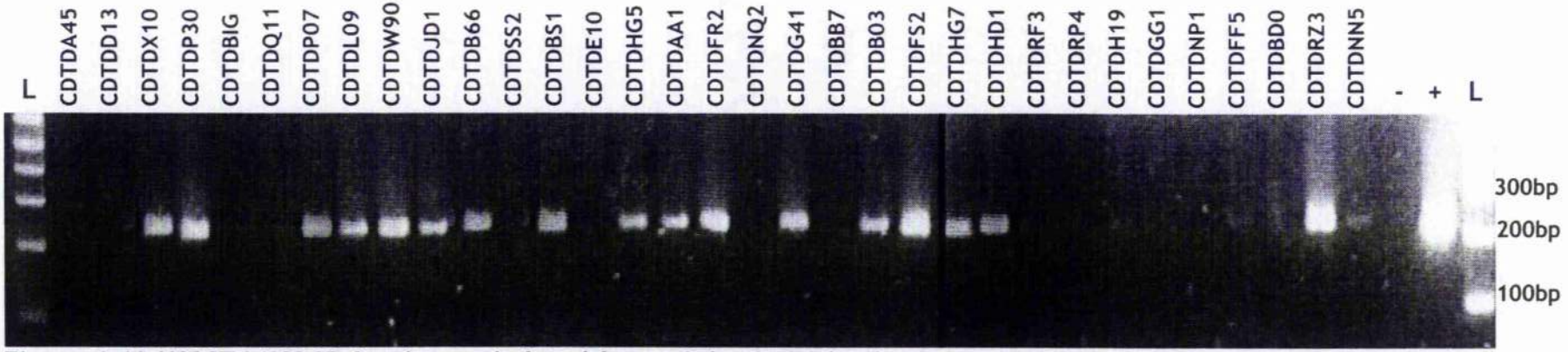


Figure 4.18 HS3ST4 MSPCR for the methylated form of the CpG island. PCR products were run on a 2% agarose gel for 2hrs at 100V. (L)=ladder. Tumours are identified by their tumour database ID code(-) is the negative control, the reaction mixture plus unmodified normal genomic DNA. (+) is a positive control ,the cloned CpG island, then modified by sodium bisulphite. Marker sizes are given in base pairs

Table 13 Methylation status of HS3ST4 in relation to the tumours described in the gels in the previous page

HS3ST4M	HS3ST4UM	Tumour Database ID	Outcome
-	Absent	CDTDA45	Requires UM
-	Absent	CDTDD13	Requires UM
+	Absent	CDTDX10	Methylated
+	Absent	CDTDP30	Methylated
-	-	CDTDBIG	Void
-	+	CDTDQ11	Unmethylated
+	-	CDTDP07	Methylated
+	+	CDTDL09	Methylated
+	+	CDTDW90	Methylated
+	+	CDTDJD1	Methylated
+	+	CDTDB66	Methylated
-	-	CDTDSS2	Void
+	+	CDTDBS1	Methylated
-	-	CDTDE10	Void
+	+	CDTDHG5	Methylated
+	+	CDTDAA1	Methylated
+	+	CDTDFR2	Methylated
-	+	CDTDNQ2	Unmethylated
+	-	CDTDG41	Methylated
-	+	CDTDBB7	Unmethylated
+	+	CDTDB03	Methylated
+	+	CDTDFS2	Methylated
+	+	CDTDHG7	Methylated
+	-	CDTDHD1	Methylated
-	-	CDTDRF3	Void
-	-	CDTDRP4	Void
-	-	CDTDH19	Void
-	+	CDTDGG1	Unmethylated
-	-	CDTDNP1	Void
-	+	CDTDFF5	Unmethylated
-	+	CDTDBD0	Unmethylated
+	+	CDTDRZ3	Methylated
E	+	CDTDNN5	Repeat
Absent	+	CDTDXP6	Requires M

(-) indicates a negative result for that form of the gene, (+) indicates a positive result for that form of the gene, and the outcome is given in the right hand column, in the case of the tumour CDTDNN5 the band was evaluated as equivocal as the sample had floated out of the well whilst it was being loaded and the PCR was repeated in this case.

Table 4-14 Association between methylation status at different loci.

	HS3ST2		HS3ST3B		HS3ST3BE1	
	Negative	**Positive**	**Negative**	**Positive**	**Negative**	**Positive**
HS3ST3B						
Negative	21 (60%)	29 (64%)				
Positive	14 (40%)	16 (36%)				
	P=0.82					
HS3ST3BE1						
Negative	33 (94%)	32 (71%)	37 (74%)	28 (93%)		
Positive	2 (6%)	13 (29%)	13 (26%)	2 (7%)		
	P=0.009		P=0.040			
HS3ST4						
Negative	20 (57%)	26 (58%)	32 (64%)	14 (47%)	40 (62%)	6 (40%)
Positive	15 (43%)	19 (42%)	18 (36%)	16 (53%)	25 (38%)	9 (60%)
	P=1.00		P=0.16		P=0.009	

Significance assessed by Fisher's exact test

Table 4-15 Association between HS3ST2 methylation and clinicopathological variables.

	HS3ST2 Methylation		
	Negative n=35	Positive n=45	Significance*
Age			
<50	8 (23%)	9 (20%)	χ^2=0.02, df=1, P=0.90
51-65	9 (26%)	16 (36%)	
>65	13 (37%)	17 (38%)	Including unknown:
Unknown	5 (14%)	3 (7%)	χ^2=1.83, df=3, P=0.61
Tumour size (mm)			
0-20	7 (20%)	9 (20%)	χ^2=0.51, df=1, P=0.48
21-50	15 (43%)	26 (58%)	
>50	2 (6%)	5 (11%)	Including unknown:
Unknown	11 (31%)	5 (11%)	χ^2=5.57, df=3, P=0.13
Grade			
1	5 (14%)	4 (9%)	χ^2=5.93, df=1, P=0.015
2	13 (37%)	11 (24%)	
3	7 (20%)	25 (56%)	Including unknown:
Unknown	10 (29%)	5 (11%)	χ^2=10.99, df=3, P=0.012
Stage			
1	4 (11%)	14 (31%)	χ^2=0.86, df=1, P=0.35
2	8 (23%)	13 (29%)	
3	7 (20%)	12 (27%)	Including unknown:
Unknown	16 (46%)	6 (13%)	χ^2=11.54, df=3, P=0.009
NPI			
<3.4	2 (6%)	5 (11%)	χ^2=1.83, df=1, P=0.18
3.4-5.4	13 (37%)	15 (33%)	
>5.4	4 (11%)	19 (42%)	Including unknown:
Unknown	16 (46%)	6 (13%)	χ^2=14.74, df=3, P=0.002
ER status			
Negative	5 (14%)	15 (33%)	P=0.16
Positive	18 (51%)	21 (47%)	Including unknown:
Unknown	12 (34%)	9 (20%)	χ^2=4.48, df=2, P=0.11
LV invasion			
Negative	12 (34%)	21 (47%)	P=0.79
Positive	11 (31%)	15 (33%)	Including unknown:
Unknown	12 (34%)	9 (20%)	χ^2=2.28, df=2, P=0.32

* The first significance test for each variable is the chi squared test for trend or Fisher's exact test (for ER status and LV invasion) excluding the "unknown" category. The second test is the chi squared test for association including the "unknown" category

Table 4-16 Association between HS3ST3B methylation and clinicopathological variables.

	HS3ST3B Methylation		
	Negative n=50	Positive n=30	Significance*
Age			
<50	9 (18%)	8 (27%)	χ^2=0.78, df=1, P=0.64
51-65	16 (32%)	9 (30%)	
>65	20 (40%)	10 (33%)	Including unknown:
Unknown	5 (10%)	3 (10%)	χ^2=0.91, df=3, P=82
Tumour size (mm)			
0-20	7 (14%)	9 (30%)	χ^2=2.79, df=1, P=0.10
21-50	29 (58%)	12 (40%)	
>50	5 (10%)	2 (7%)	Including unknown:
Unknown	9 (18%)	7 (23%)	χ^2=4.09, df=3, P=0.25
Grade			
1	6 (12%)	3 10%)	χ^2=0.81, df=1, P=0.37
2	17 (34%)	7 (23%)	
3	18 (36%)	14 (47%)	Including unknown:
Unknown	9 (18%)	6 (20%)	χ^2=1.35, df=3, P=0.72
Stage			
1	12 (24%)	6 (20%)	χ^2=0.43, df=1, P=0.58
2	14 (28%)	7 (23%)	
3	11 (22%)	8 (27%)	Including unknown:
Unknown	13 (26%)	9 (30%)	χ^2=0.57, df=3, P=0.90
NPI			
<3.4	4 (8%)	3 (10%)	χ^2=0.01, df=1, P=0.93
3.4-5.4	19 (38%)	9 (30%)	
>5.4	14 (28%)	9 (30%)	Including unknown:
Unknown	13 (26%)	9 (30%)	χ^2=0.56, df=3, P=0.91
ER status			
Negative	10 (20%)	10 (33%)	P=0.09
Positive	28 (56%)	11 (37%)	Including unknown:
Unknown	12 (24%)	9 (30%)	χ^2=3.03, df=2, P=0.22
LV invasion			
Negative	26 (52%)	7 (23%)	P=0.03
Positive	13 (26%)	13 (43%)	Including unknown:
Unknown	11 (22%)	10 (33%)	χ^2=6.39, df=2, P=0.04

* The first significance test for each variable is the chi squared test for trend or Fisher's exact test (for ER status and LV invasion) excluding the "unknown" category. The second test is the chi squared test for association including the "unknown" category.

Table 4-18Association between HS3ST4 methylation and clinicopathological variables.

	HS3ST34 Methylation Negative n=50	Positive n=30	Significance*
Age			
<50	8 (17%)	9 (27%)	χ^2=0.03, df=1, P=0.89
51-65	7 (37%)	8 (24%)	
>65	16 (35%)	14 (41%)	Including unknown:
Unknown	5 (11%)	3 (9%)	χ^2=2.18, df=3, P=0.54
Tumour size (mm)			
0-20	9 (20%)	7 (21%)	χ^2=0.001, df=1, P=0.98
21-50	23 (50%)	18 (53%)	
>50	4 (9%)	3 (9%)	Including unknown:
Unknown	10 (22%)	6 (18%)	χ^2=0.21, df=3, P=0.98
Grade			
1	4 (9%)	5(15%)	χ^2=0.01, df=1, P=0.93
2	15 (33%)	9 (27%)	
3	17 (37%)	15(44%)	Including unknown:
Unknown	10 (22%)	5 (15%)	χ^2=1.64, df=3, P=0.65
Stage			
1	10 (22%)	8 (27%)	χ^2=0.04, df=1, P=0.85
2	13 (28%)	8 (20%)	
3	10 (23%)	9 (27%)	Including unknown:
Unknown	13 (28%)	9 (27%)	χ^2=0.40 df=3, P=0.94
NPI			
<3.4	4 (9%)	3 (9%)	χ^2=1.40, df=1, P=0.50
3.4-5.4	18 (39%)	10 (29%)	
>5.4	11 (24%)	12 (35%)	Including unknown:
Unknown	13 (28%)	9 (27%)	χ^2=1.43, df=3, P=0.70
ER status			
Negative	9 (20%)	3 (32%)	P=0.27
Positive	24 (52%)	8 (44%)	Including unknown:
Unknown	13 (28%)	4 (24%)	χ^2=1.71, df=2, P=0.43
LV invasion			
Negative	21 (42%)	12 (40%)	P=0.20
Positive	12 (31%)	14 (40%)	Including unknown:
Unknown	13 (28%)	8 (20%)	χ^2=2.05, df=2, P=0.36

* The first significance test for each variable is the chi squared test for trend or Fisher's exact test (for ER status and LV invasion) excluding the "unknown" category. The second test is the chi squared test for association including the "unknown" category.

Table 4-17 Association between HS3ST3BE1 methylation and clinicopathological variables.

	HS3ST3BE1 Methylation Negative n=50	Positive n=30	Significance*
Age			
<50	16 (18%)	1 (7%)	χ^2=4.32, df=1, P=0.04
51-65	21 (32%)	4 (27%)	
>65	21 (32%)	9 (60%)	Including unknown:
Unknown	7 (11%)	1 (7%)	χ^2=4.67, df=3, P=0.20
Tumour size (mm)			
0-20	16 (25%)	0 (0%)	χ^2=4.99, df=1, P=0.04
21-50	31 (48%)	10 (67%)	
>50	5 (8%)	2 (13%)	Including unknown:
Unknown	13 (20%)	3 (20%)	χ^2=4.99, df=3, P=0.17
Grade			
1	9 (14%)	0 (0%)	χ^2=2.38, df=1, P=0.31
2	19 (29%)	5 (33%)	
3	25 (39%)	7 (47%)	Including unknown:
Unknown	12 (19%)	3 (20%)	χ^2=2.37, df=3, P=0.50
Stage			
1	14 (22%)	4 (27%)	χ^2=0.01, df=1, P=0.94
2	18 (28%)	3 (20%)	
3	15 (23%)	4 (27%)	Including unknown:
Unknown	18 (28%)	4 (27%)	χ^2=0.49 df=3, P=0.92
NPI			
<3.4	7 (11%)	0 (10%)	χ^2=3.93, df=1, P=0.047
3.4-5.4	24 (37%)	4 (27%)	
>5.4	16 (25%)	7 (47%)	Including unknown:
Unknown	18 (28%)	4 (27%)	χ^2=4.05, df=3, P=0.26
ER status			
Negative	17 (26%)	3 (20%)	P=0.73
Positive	31 (48%)	8 (53%)	Including unknown:
Unknown	17 (26%)	4 (27%)	χ^2=0.27, df=2, P=0.088
LV invasion			
Negative	27 (42%)	6 (40%)	P=0.75
Positive	20 (31%)	6 (40%)	Including unknown:
Unknown	18 (28%)	3 (20%)	χ^2=0.60, df=2, P=0.74

* The first significance test for each variable is the chi squared test for trend or Fisher's exact test (for ER status and LV invasion) excluding the "unknown" category. The second test is the chi squared test for association including the "unknown" category.

Table 4-19 Association between HS3ST2 methylation and continuous clinicopathological variables.

	HS3ST2 Methylation		
	Negative	Positive	Significance
Age	61 (47 - 75)	62 (52 - 70)	P=0.66
Tumour size (mm)	25 (17-34)	30 (22-40)	P=0.16
NPI	4.6 (4.2-5.4)	5.4 (4.4-6.4)	P=0.09

Results shown as median (interquartile range). Significance assessed by Mann-Whitney test.

Table 4-20 Association between HS3ST3B methylation and continuous clinicopathological variables.

	HS3ST3B Methylation		
	Negative	Positive	Significance
Age	62 (53 - 74)	60 (46 - 67)	P=0.15
Tumour size (mm)	30 (25-40)	24 (18-36)	P=0.18
NPI	4.7 (4.2-5.7)	5.4 (4.4-6.6)	P=0.46

Results shown as median (interquartile range). Significance assessed by Mann-Whitney test.

Table 4-21 Association between HS3ST3BE1 methylation and continuous clinicopathological variables.

	HS3ST3BE1 Methylation		
	Negative	Positive	Significance
Age	60 (48 - 72)	68 (60 - 72)	P=0.13
Tumour size (mm)	28 (18-38)	32 (26-47)	P=0.11
NPI	4.7 (4.2-5.9)	5.6 (4.5-5.9)	P=0.15

Results shown as median (interquartile range). Significance assessed by Mann-Whitney test.

Table 4-22 Association between HS3ST4 methylation and continuous clinicopathological variables

	HS3ST4 Methylation		
	Negative	Positive	Significance
Age	61 (53 -72)	62 (46 - 72)	P=1.00
Tumour size (mm)	30 (21-40)	29 (21-36)	P=0.88
NPI	4.9 (4.2-5.6)	5.4 (4.3-6.5)	P=0.62

Results shown as median (interquartile range). Significance assessed by Mann-Whitney test.

Table 4-23 Association between number of methylated loci and clinicopathological variables.

	Number of methylated loci					
	0 n=13	1 n=27	2 n=24	3 n=15	4 n=1	Significance*
Age						
<50	1 (8%)	7 (26%)	7 (29%)	2 (13%)	0 (0%)	χ^2=0.13,
51-65	4 (31%)	9 (33%)	8 (33%)	4 (27%)	0 (0%)	df=1,
>65	5 (39%)	9 (33%)	8 (33%)	7 (47%)	1(100%)	P=0.72
Unknown	3 (23%)	2 (7%)	1 (4%)	2 (13%)	0 (0%)	
Tumour size (mm)						
0-20	3 (23%)	6 (22%)	2 (8%)	5 (33%)	0 (0%)	χ^2=0.10,
21-50	5 (39%)	14 (52%)	15 (63%)	6 (40%)	1(100%)	df=1,
>50	1 (8%)	2 (7%)	2 (8%)	2 (13%)	0 (0%)	P=0.75
Unknown	4 (31%)	5 (19%)	5 (21%)	2 (13%)	0 (0%)	
Grade						
1	1 (8%)	5 (19%)	2 (8%)	1 (7%)	0 (0%)	χ^2=4.18,
2	6 (46%)	8 (27%)	6 (25%)	4 (27%)	0 (0%)	df=1,
3	2 (15%)	9 (33%)	12 (50%)	8 (53%)	1(100%)	P=0.04
Unknown	4 (30%)	5 (19%)	4 (17%)	2 (13%)	0 (0%)	
Nodal status						
Negative	1 (8%)	7 (26%)	5 (29%)	5 (13%)	0 (0%)	χ^2=0.12,
1-3 Positive	4 (31%)	6 (33%)	8 (33%)	3 (27%)	0 (0%)	df=1,
>3 Positive	2 (39%)	7 (33%)	5 (33%)	4 (47%)	1(100%)	P=0.94
Unknown	6 (23%)	7 (7%)	6 (4%)	3 (13%)	0 (0%)	
NPI						
<3.4	1 (8%)	3 (11%)	1 (4%)	2 (13%)	0 (0%)	χ^2=3.53,
3.4-5.4	5 (39%)	12 (44%)	7 (30%)	4 (27%)	0 (0%)	df=1,
>5.4	1 (8%)	5 (19%)	10 (42%)	6 (40%)	1(100%)	P=0.06
Unknown	6 (46%)	7 (26%)	6 (25%)	3 (20%)	0 (0%)	
ER status						
Negative	1 (8%)	6 (22%)	7 (29%)	5 (33%)	1(100%)	χ^2=3.71,
Positive	8 (62%)	13 (48%)	12 (50%)	6 (40%)	0 (0%)	df=1,
Unknown	4 (31%)	8 (30%)	5 (21%)	4 (27%)	0 (0%)	P=0.05
LV invasion						
Negative	6 (46%)	13 (48%)	9 (38%)	5 (33%)	0 (0%)	χ^2=2.77,
Positive	3 (23%)	7 (26%)	8 (33%)	7 (47%)	1(100%)	df=1,
Unknown	4 (31%)	7 (26%)	7 (30%)	3 (20%)	0 (0%)	P=0.01

* Chi squared test for trend excluding the "unknown" category.

Discussion

Methylation within the HS3ST gene family is variable in breast tumours, HS3ST3A and HS3ST1 are never methylated, whereas the others HS3ST2, HS3ST3B and HS3ST4 are all found to be methylated. The high number of tumours methylated at the HS3ST2, HS3ST3B and HS3ST4 loci suggests that methylation of these genes may be significant in the carcinogenic process. However, it must be borne in mind that this argument is based on the premise that methylation has a functional consequence that is selected for during tumourogenesis. The alternative hypotheses are that HS3ST1 and 3A loss of function are non-viable for a cell and loss of these enzymes through methylation is never selected for: Loss of HS3ST1 expression, may be selected against as it is required for enhanced activity of anti-thrombin III. Silencing of such an enzyme may lead to greater activation of the clotting system during invasive growth and metastasis, a disadvantage for a cancer cell. The role of HS3ST3A is not as well defined, but the substrate specificity of the enzyme is similar to and overlaps that of HS3ST3B. Again, it may be speculated that HS3ST3A and HS3ST3B function is of great significance as there has been gene duplication in this case and preservation of this arrangement during the evolution of humans. Methylation of the other loci may be a purely stochastic event and have no functional significance for a cell. However, it has been shown that tumour DNA is globally hypomethylated (Bernardino, Roux et al. 1997) and there is specific hypermethylation of areas in the genome. Hypomethylation is the default position of a tumour cell DNA as a whole and therefore any hypermethylation, it could be argued, is significant.

Methylation of HS3ST2 has been shown before to be a high frequency event in breast cancers [Miyamoto, 2003 #43] and other tumours, so this result is in keeping with previous analysis. In the study by Dr. Miyamoto (Miyamoto, Asada et al. 2003) 77 out of 85 breast tumours were methylated at the HS3ST2 CpG island, this amounts to 88% of their sample number. In contrast, we found 45/80 tumours to be methylated at this locus, 56% of the sample number. This difference in findings may be due to disparity in breast tumours or it could be due to sensitivity, or specificity of MSPCR.

In order to validate the results of MSPCR at the HS3ST2 and HS3ST4 loci and also to assess the density of methylation at these loci methylation specific sequencing was carried out on a total of 20 clones. The intervening CpG dinucleotides, located between each set of primers, were all found to be methylated. This suggests that methylation detected using the MSPCR primers was also a surrogate marker for dense methylation and therefore should correspond to transcriptional silencing of that gene. In the previously reported study of Dr. Miyamoto methylation at the HS3ST2 locus was also correlated with down-regulation of transcription.

4.2.3 Statistical Correlation of Methylation and Clinicopathological Variables

Prior to discussing any of these findings it must be stated that the likelihood of finding significant correlations in any form of statistical analysis becomes more probable the greater the number of tests performed. It is with this caveat in mind that the results for these analyses are interpreted.

The negative correlation between methylation at HS3ST3BE1 and HS3ST3B, and the positive correlation between methylation at HS3ST3BE1 and HS3ST2 and HS3ST4 is interesting. These observations are open to some speculation. With respect to 3B and 3BE1 methylation it could be argued that methylation at either of these two loci causes the same effect and therefore methylation may be mutually exclusive. There is however, evidence to support the idea that methylation at one site leads to infiltration of methylation at adjacent sites (Millar, Ow et al. 1999) and this does n ot appear to h appen in this case. Therefore it may be argued that methylation at only one of these sites confers a closed chromatin structure that prevents further methylation at the adjacent site, certainly it would be interesting to know the functional effect of methylation at wach of the loci - it is not known if they both have the same effect.

In the case of HS3ST2 and HS3ST4 the seeding hypothesis does not hold as these genes are located on different chromosomes and therefore it may be (assuming methylation has a functional significance in the case of these genes) there is an advantage conferred from having lack of HS3ST3B and HS3ST2 or HS3ST3B and

HS3ST4. This dual loss of expression would have the functional effect of reducing the complement of HS3ST genes in the endoplasmic reticulum and therefore some substrates would escape sulphation. If this confers an advantage to a cell in the form of increased growth factor binding, or potentiation then it may be selected for. Alternatively it may be that the promoter regions of these genes have similar features and are likely to be methylated concordantly. There is evidence to suggest that methylation of CpG islands occurs following down-regulation of transcription and this is dependent on lack of expression from the gene(Song, Stirzaker et al. 2002; Stirzaker, Song et al. 2004). If these genes are silenced due to lack of a common upstream factor then methylation at these loci may occur concordantly and this may explain the positive association.

Increasing tumour grade is correlated with methylation at the HS3ST2, and number of loci methylated correlates with grade. Tumour grade is calculated by counting mitotic activity, assessing nuclear pleomorphism and lastly by evaluation of tumour differentiation by formation of acinar structures. It is thought to indicate tumour aggressiveness, larger, more advanced tumours tend to have a higher grade and smaller, earlier detected lesions tend to be of lower grade. It is of interest that HS3ST2 is more likely to be methylated the higher the grade of the tumour, and suggests that silencing of this gene is required in later stages of tumourogenesis, not in the transformation of benign to DCIS, or DCIS into an invasive phenotype. The finding that number of loci methylated correlates with grade is in keeping with other studies that have shown increased CpG island methylation with increasing grades of tumour.(Herman, Civin et al. 1997; Soares, Pinto et al. 1999; Salem, Liang et al. 2000; Dong, Pang et al. 2001; Uhlmann, Rohde et al. 2003). The significance of this latter finding is not readily interpretable as the study does not have the power to say which with confidence which specific genes are methylated as tumour grade increases. It would be of greater interest to know if two specific genes tended to increase in frequency of methylation in association with increasing tumour grade as this may allow more specific questions to be asked over substrate overlaps in the associated transferases, or perhaps promoter similarities between genes.

There is a very strong correlation between lack of methylation at HS3ST2 and NPI unknown and nodal status unknown. In fact this is the strongest correlation of all he statistical analyses. The tumours that are negative for HS3ST2

methylation and have missing NPI or nodal status are all from the GRI tumour bank with the exception of 2 tumours which were collected fresh. The reason for this correlation may be due to a common factor in the storage of these specimens causing degradation of the methyl group, they may have been in a batch that was analysed together and there was a technical problem with that particular batch of PCRs or modifications. However, all reactions were carried out in triplicate, therefore it would have had to be a recurring problem. The age at operation dates from 1989 to 2003 for this group, which is not disparate from the remainder of the tumours in the bank. To some extent this finding remains an anomaly, however it suggests that there is a common non-pathological factor linking the tumours and the status of the HS3ST2 locus.

HS3ST3B methylation is correlated with lymphovascular invasion. This finding can be related in a positive way to previous work by Tim Huang and Graeme Brock (Brock, Huang et al. 2001; Huang, Cheng et al. 2003) that showed HS3ST3B was correlated with stage and grade of breast tumours. The previous studies did not provide a functional explanation of why this correlation existed. The data in this study would suggest that LVI may be increased in tumours that lack HS3ST3B expression thereby leading to increased nodal metastases and increased grade.

HS3ST3BE1 methylation is correlated with age, NPI and size. It is difficult to speculate on the significance of these particular findings in relation to the HS3ST3B gene product, as it is not known what the functional significance of methylation at this locus is. It may be related to the product, but equally it may not and previous studies (Brock, Huang et al. 2001) have concentrated on the downstream locus rather than this one.

The trend for number of loci methylated to be correlated with ER/PR negativity, NPI and lymphovascular invasion may be consistent with the finding in previous studies that methylation of CpG islands as a whole is increased with grade, stage and degree of differentiation of a tumour (Costello, Fruhwald et al. 2000). NPI is calculated from grade and stage, and LVI is associated with lymph node metastasis and thereby with increased stage.

There are a few thorny issues that have to be addressed in the discussion section that this analysis has thrown up. The first is the lack of any HS3ST1 methylation, despite this gene being found in a subtractive hybridisation library and being

pivotal in the decision to investigate the methylation status of these genes as a whole family. MSPCR only analysis the methylation status of the few CpGs which are located in the primer sites and as such it is a limited examination of a CpG island. There have been studies that demonstrate downregulation of a gene by methylation at only a few CpG dinucleotides (Clark, Harrison et al. 1997; Graff, Herman et al. 1997; Millar, Ow et al. 1999; Stirzaker, Song et al. 2004) and it may be that this is the case in HS3ST1, however, the method used to isolated this CpG island would suggest it is densely methylated. Therefore it is either the case that the CpG island was not methylated in the initial tumours used to create the library and represents an error at that stage, or the number of tumours in this study was not significant enough to pick up methylation at that locus, or the HS3ST1 primers were targeted at the wrong section of the CpG island. Further work is required to resolve this issue.

With regard to the statistical tests performed on the data for HS3ST3BE1 methylation a slight aberration, or discontinuity is found between the Mann-Whitney test for the continuous variables and Fisher's exact test for these variables when there are examined as categorical statistics. The p values for Age, Tumour and NPI in the Mann-Whitney test are 0.13, 0.11 and 0.15 respectively and the p values in Fisher's exact test are 0.04. 0.04 and 0.047. The discrepancy is more apparent than real when the p values are looked at individually, however, the Mann-Whitney test should be more sensitive than Fisher's exact test at detecting a significant association.. The reason for this is unclear however the groups for age and tumour size and NPI have a number of samples that fall into the middle category in each case, but are at the extreme ends of the categories. These tumours do not have as much effect on the correlation when they are viewed as continuous variables. This issue may be more fully resolved with increased numbers of specimens.

An important question that arise out these findings is functional and relates to the relationship that methylation at these loci has with transcription of the associated gene. In most of the studies published to date it has been shown that dense CpG island methylation correlates with silencing of the associated gene(Esteller 2005). Indeed, this relationship has been demonstrated already for HS3ST2 by Dr Miyamoto in Tokyo. However, the proportion of tumours, in this study, methylated at HS3ST2 is different to Miyamoto's study. Therefore, it

was necessary to establish the relationship in this particular setting, as transcription may be dependent on which region of a CpG island is methylated. It was also sought to establish this relationship with regard to HS3ST4, and HS3ST3B, as these genes had never been investigated with regard to methylation and transcription. This was possible in the case of HS3ST4, but technical difficulties were encountered in the case of HS3ST3B as the gene is 96% homologous to HS3ST3A, making analysis of transcription at an RNA level impossible.

The biological effect of turning the genes off is not fully understood, specifically what this means in terms of substrates in heparan sulphate molecules that would still have sulphation at the 3'OH and substrates that would no longer be sulphated. From a critics point of view this is important, it may be that silencing at these loci has no biological effect and there is sufficient substrate overlap between HS3ST3A and HS3ST1 to catalyse the necessary modifications to all the heparan sulphates a tumour cell requires and turning the others off is efficient, but not advantageous.

The issue of density of methylation at the HS3ST2 and HS3ST4 loci is addressed above and it appears that for these two genes at least MSPCR of the CpG island shows an association with dense methylation at that sight. The issue of the biological effect of turning these genes off has not been broached experimentally in this particular piece of work, however, an examination of the literature suggests that each HS3ST gene is required for specific substrates. There is some substrate overlap between enzymes, but this is not broad and certain substrates can only be sulphated by one specific HS3ST enzyme (Liu, Shworak et al. 1999). The relationship between methylation at the HS3ST2 and HS3ST4 genes and transcription is covered in the next chapter.

5 The Relationship of CpG Island Methylation to RNA Expression at the HS3ST2 and HS3ST4 Loci

5.1 Expression of HS3ST2 and HS3ST4 in Normal Breast Tissue

As demonstrated by MSPCR HS3ST2 and HS3ST4 genes are frequently methylated in breast cancers. The significance of this observation is dependent on the normal expression pattern of these genes, specifically if they are expressed in normal breast tissue. Therefore, analysis of expression in normal breast tissue was undertaken.

During the collection of breast tumours from mastectomy specimens normal breast tissue was collected from sites more than 4cm away from the tumour. The tissue was snap-frozen in liquid nitrogen, sectioned and RNA was extracted using Tri-reagent (Sigma) as described in Materials and Methods. Histological normality was confirmed on frozen section in every case.

DNA contamination was assessed using GAPDH trans-intronic primers and RNA concentrations were determined by UV spectroscopy, as described in Materials and Methods.

RTPCR for HS3ST2 and HS3ST4 was carried out on the extracted RNA using the invitrogen superscript II system. In brief: first strand synthesis used a poly T primer, the cDNA derived from this step was used as a template for PCR. In order to control for DNA contamination of PCRs two first strand synthesis reactions were set up for every sample, identical apart from the omission of Superscript II (SSII) reverse transcriptase from one of the reactions. Two PCRs were then set up for each sample: one used the reaction that contained SSII reverse transcriptase as a template, the other used the reaction that did not contain SSII as a negative control. If a PCR product appeared in the negative control there was DNA contamination of that sample and interpretation was null and void.

GAPDH RTPCR was also carried out on all RNA samples using primers which span intron 2 within the GAPDH gene. The forward primer anneals to a exon 1 and the reverse primer anneals to exon 2. The size of the product is therefore an indicator of the template that the primers anneal to. If a sample is pure processed mRNA the primers produce a band 320bp in size and if a sample is

contaminated by DNA or non-processed mRNA a band 350bp larger (the size of intron 1) is visible on a gel. If contamination following RTPCR with GAPDH trans-intronic primers was detected the sample was not used for RNA analysis.

5.2 Correlation of Methylation at the HS3ST2 and HS3ST4 CpG Islands with Transcription In Cell Culture and In Vivo

A screen for methylation at the HS3ST2 and HS3ST4 loci was carried out on 7 different cell lines: MCF7, A2780, SVCT, HeLa, MCF-8, SUM149PT and SUM 102PT, using MSPCR. MCF7, the breast derived tumour cell line was found to be methylated at the HS3ST2 CpG island and A2780, an ovarian cell-line (gifted by Prof. Bob Brown) was found to be methylated at the HS3ST4 CpG island. These two cell lines were used as in vitro tools, to study the relationship between methylation at the CpG island and RNA expression from the respective gene.

Methylation status of the CpG islands of both genes was determined using MSPCR. RNA expression was detected using RTPCR, as described previously. We also treated cells with 5-aza-2-deoxycytidine to reverse the methylation of the CpG islands.

A2780 cells and MCF 7 cells were seeded at a density of $6x10^5$ cells/10 cm plate on day 0, and exposed to 0, 0.1, 0.2, 0.5, 1 or 2mM 5-aza-2-deoxycytidine (5-aza-dC; Sigma) for 24 hrs on days 1 and 3. The cells were cultured in fresh media after each treatment, and harvested on day 4 using Tri-reagent, as described in Materials and Methods.

On harvesting the cells both DNA and RNA was extracted, methylation status was determined by MSPCR. The RNA was harvested and used for RTPCR as described previously.

5.3 Correlation of Methylation at the HS3ST2 and HS3ST4 Loci with Transcription in Breast Tumour Samples

Methylation and transcriptional correlation is apparent in the cell lines analysed. However, cell culture is an artificial environment and the observations made in these models do not always hold true in vivo.

Carrying out RTPCR in a non-quantatitive fashion is a reasonable approach in a cell culture system as it gives a positive or negative result in a sample that is believed to contain a homogeneous population of cells. Tumour extracts, from any site, contain an admixture of marrow-derived cells, fibroblasts, endothelial cells, and in the case of breast tissue adipocytes, myoepithelial cells and possibly normal adjacent epithelial cells. Carrying out non-quantitative RTPCR on breast tumours requires that the sample is almost exclusively breast tumour, containing no, or as little as possible 'contamination' from normal cells.

Two approaches were used to overcome this difficulty: in samples that there was enough tissue tumours were examined by naked eye and if breast tumour material was clearly identified it was dissected using a scalpel from the surrounding normal tissue. Margins were taken within the tumour mass and normal tissue was never present at the margins. In other samples micro-dissection was carried out using the Arcturus Laser Capture Microscope and the associated RNA Extraction kit, as described in Materials and Methods.

The two approaches were not used side by side, use of the gross dissection process was made first of all as a means of deciding whether micro-dissection needed to be used in this instance.

5.3.1 Micro-dissection of Tumour Samples

Breast tumour samples were collected and stored as described previously. Samples were prepared as described in Materials and Methods. Micro-dissection of samples was undertaken using the Arcturus Laser Capture System and the associated RNA extraction kit was used for harvesting RNA.

5.4 Results

5.4.1 Expression of HS3ST2 and HS3ST4 in Normal Breast Tissue

RNA expression of HS3ST2 and HS3ST4 was demonstrated in 14/14 cases (100%) of cases of normal tissue. Figure 5-1 and Figure 5-4 show representative results for RTPCR carried out on normal breast tissue specimens and these results are summarised in Table 5-3 RTPCR for HS3ST2 and HS3ST4 in Normal Tissue. The figures show two reactions for each sample, one contains a cDNA template that was constructed using the reverse transcriptase enzyme superscript II (+) and the other lane contains template constructed without the addition of the SSII reverse transcriptase (-). The latter is a negative control for the RTPCR reaction and would only be positive if there was contamination by DNA. In all cases examined no such contamination was seen.

5.4.2 Expression of HS3ST2 and HS3ST4 in Cell Culture

5.4.2.1 HS3ST2 and Cell Line MCF7

Figure 5-2 shows a number of MSPCR reactions for HS3ST2 that were carried out on the cell line MCF7. The figure shows six groups of two cell culture samples. Cell cultures for increasing concentrations of 2-deoxy-5-azacytidine were set up (Untreated-2.0uM) and a parallel culture was set up for each dose, but this culture was not treated with the demethylating agent. This parallel culture should retain the methylated state of the HS3ST2 gene as it is not treated with 2-deoxy-5-azacytidine. The figure shows that the HS3ST2 locus remains methylated when untreated and there is complete loss of methylation at a concentration of 0.2uM and above.

Figure 5-3 shows the corresponding HS3ST2 RTPCR results for the cell cultures that are detailed in the associated MSPCR reactions. The negative control for each RTPCR reaction is a template which was constructed without addition of the superscript II enzyme, as detailed above. In this figure it can be seen that there is a product just visible at the 0.1uM concentration, however, complete demethylation of the MSPCR primer sites does not occur until 0.2uM. This

indicates there may be partial demethylation at this locus, either between cells, or within the CpG island and expression takes place. In order to address this question HS3ST2 MSPCR for the unmethylated form of the island would have to be carried out.

The association of HS3ST2 methylation status and expression in MCF7 is summarised in Table 5-4 Correlation of Methylation and Transcription of the HS3ST2 Locus in the Cell Line MCF7.

5.4.2.2 HS3ST4 and Cell Line A2780

Figure 5-5 and Figure 5-6 show the equivalent reactions detailed in 5.4.2.1 for HS3ST4 in the setting of the cell line A2780. In this case re-expression of HS3ST4 does not take place until the 0.2uM concentration, which corresponds with the loss of HS3ST4 methylation.

The relationship of HS3ST2 methylation status and expression in A2780 is summarised in Table 5-5.

5.4.2.3 Expression of HS3ST2 and HS3T4 in Grossly Dissected Tumour Samples

In total ten samples were used from the tumour bank to assess methylation and associated RNA expression in HS3ST2 and HS3ST4. The details of the tumours used and their results are given in Table 5-6 and the statistical analysis of these results is shown in the tables below.

Table 5-1 HS3ST2 Methylation and Transcription in Grossly Dissected Tumour Samples

HS3ST2 RTPCR	HS3ST2 Methylation		*Significance
	Methylated	Unmethylated	
RTPCR (+ve)	4 (57%)	2 (67%)	P=1.00
RTPCR (-ve)	3 (43%)	1 (33%)	
Total	7 (100%)	3 (100%)	

*Significance calculated Fisher's exact test.

Table 5-2HS3ST4 Methylation and Transcription in Grossly Dissected Tumour Samples

HS3ST4 RTPCR	HS3ST4 Methylation		*Significance
	Methylated	Unmethylated	
RTPCR (+ve)	2 (40%)	4 80%)	P=0.52
RTPCR (-ve)	3 (60%)	1 (20%)	
Total	5 (100%)	5 (100%)	

*Significance calculated Fisher's exact test.

The tables above indicate that there is no significant correlation between methylation status of these two genes and expression in the case of the grossly dissected tumour specimens. Reasons for this are most probably due to the admixture of cell types that constitute a breast tumour sample. While methylation can be found in tumour samples the quantification of methylation was not assessed in these samples. It may be that methylation does not occur in all tumour cells, or also that a sample is not a homogeneous sample of tumour alone but includes normal breast epithelial cells. Samples will undoubtedly contain normal cells of some type be they endothelial or adjacent fibroblasts. Previously in this chapter it has been shown that the two genes HS3ST2 and HS3ST4 were expressed in normal breast tissue. It is expected it is unlikely that a single breast tumour sample will contain a pure tumour cell sample and therefore at least some RNA from normal cells will be present.

5.4.2.4 Micro-dissection

Micro-dissection and RNA extraction was largely unsuccessful in contrast to micro-dissection and DNA extraction. Samples did not produce the quality of RNA that was required to amplify GAPDH, our control mRNA. Various attempts were made at modifying the cutting, staining, handling and extraction process to prevent RNA degradation.

In the first instance the sample was frozen immediately in liquid nitrogen, took a section and fixed this immediately in alcohol. The time from patient to fixation was less than an hour in some circumstances. This method in itself was not enough to prevent degradation. Time from fixation, through staining and eventual micro-dissection and harvesting was further reduced. The machinery was cleaned and prepared with RNase Zap as was all the glassware and DEPC

treated water was used to dilute staining medium and ethanol. In summary, this technique was hampered by technical difficulties.

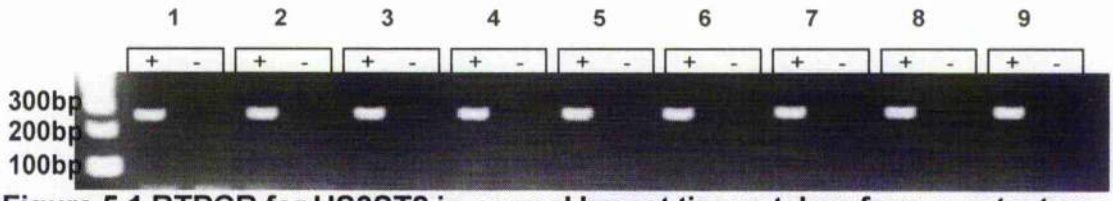


Figure 5.1 RTPCR for HS3ST2 in normal breast tissue taken from mastectomy specimens.

L=ladder, numbers relate to the specimen that the tissue was taken from, - is the negative control (the RNA extraction without the addition of superscriptII at the reverse transcriptase stage), + indicates that reverse transcriptase was added to the RNA extraction. Marker sizes are given in base pairs.

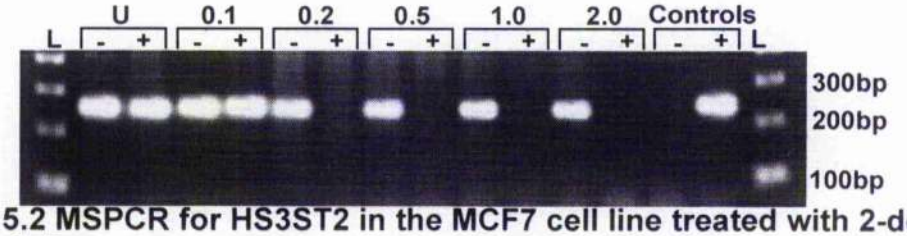


Figure 5.2 MSPCR for HS3ST2 in the MCF7 cell line treated with 2-deoxy-5-azacytidine

Following either no treatment (U) with the demethylating drug 2-deoxy-5-azacytidine or treatment with 0.1uM, 0.2uM, 0.5uM, 1uM or 2uM. (-) is the negative control for each treatment; cells were grown in parallel, but not treated with 2-dexy-5-azacytidine. (+) indicates that the cell line was treated with the associated dose of 2-deoxy-5-azacytidine. L is the molecular ladder and sizes are given in base pairs. This gel demonstrates demethylation of HS3ST2 in the MCF7 cell line following treatment with 0.2uM 2-deoxy-5-azacytidine, or concentrations above this. The postive control reaction is uses a plasmid template that is artificially methylated and treated with sodium bisulphite and the negative control contains normal unmodified genomic DNA.

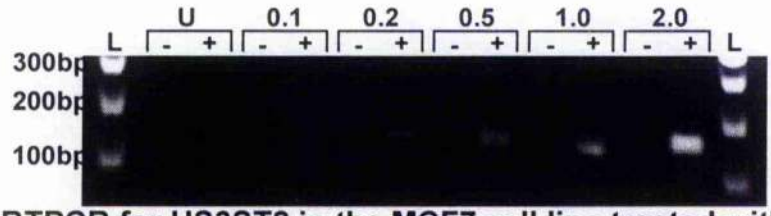


Figure 5.3 RTPCR for HS3ST2 in the MCF7 cell line treated with 2-deoxy-5-azacytidine

Following either no treatment (U) with the demethylating drug 2-deoxy-5-azacytidine or treatment with 0.1uM, 0.2uM, 0.5uM, 1uM or 2uM. (-) is the negative control for each treatment; reverse transcriptase was not added to the RNA extraction mixture. (+) indicates that reverse transcriptase was added to the RNA extraction mixture. L is the molecular ladder and sizes are given in base pairs. This gel demonstrates re expression of HS3ST2 in the MCF7 cell line following treatment with 0.1uM 2-deoxy-5-azacytidine, or concentrations above this. The negative controls remain blank demonstrating that any signal is due to the presence of RNA and DNA contamination is not an issue.

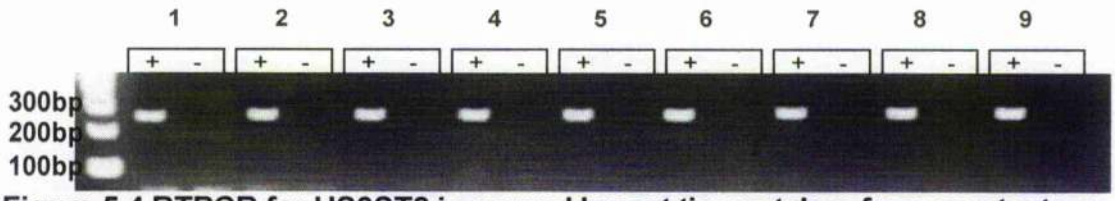


Figure 5.4 RTPCR for HS3ST2 in normal breast tissue taken from mastectomy specimens.

L=ladder, numbers relate to the specimen that the tissue was taken from, - is the negative control (the RNA extraction without the addition of superscriptII at the reverse transcriptase stage), + indicates that reverse transcriptase was added to the RNA extraction. Marker sizes are given in base pairs.

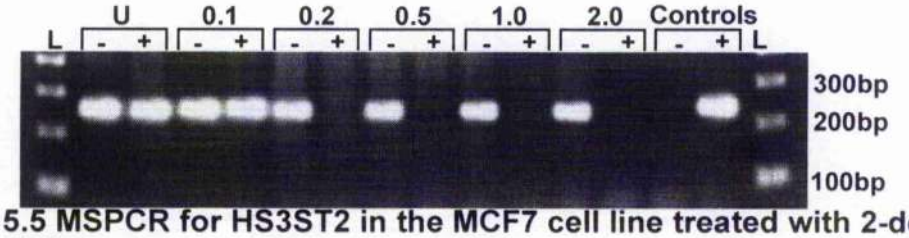


Figure 5.5 MSPCR for HS3ST2 in the MCF7 cell line treated with 2-deoxy-5-azacytidine

Following either no treatment (U) with the demethylating drug 2-deoxy-5-azacytidine or treatment with 0.1uM, 0.2uM, 0.5uM, 1uM or 2uM. (-) is the negative control for each treatment; cells were grown in parallel, but not treated with 2-dexy-5-azacytidine. (+) indicates that the cell line was treated with the associated dose of 2-deoxy-5-azacytidine. L is the molecular ladder and sizes are given in base pairs. This gel demonstrates demethylation of HS3ST2 in the MCF7 cell line following treatment with 0.2uM 2-deoxy-5-azacytidine, or concentrations above this. The postive control reaction is uses a plasmid template that is artificially methylated and treated with sodium bisulphite and the negative control contains normal unmodified genomic DNA.

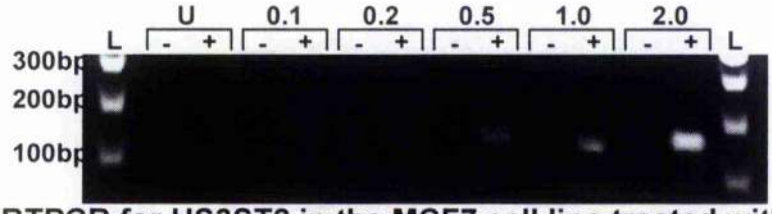


Figure 5.6 RTPCR for HS3ST2 in the MCF7 cell line treated with 2-deoxy-5-azacytidine

Following either no treatment (U) with the demethylating drug 2-deoxy-5-azacytidine or treatment with 0.1uM, 0.2uM, 0.5uM, 1uM or 2uM. (-) is the negative control for each treatment; reverse transcriptase was not added to the RNA extraction mixture. (+) indicates that reverse transcriptase was added to the RNA extraction mixture. L is the molecular ladder and sizes are given in base pairs. This gel demonstrates re expression of HS3ST2 in the MCF7 cell line following treatment with 0.1uM 2-deoxy-5-azacytidine, or concentrations above this. The negative controls remain blank demonstrating that any signal is due to the presence of RNA and DNA contamination is not an issue.

Table 5-3 RTPCR for HS3ST2 and HS3ST4 in Normal Tissue.

Tissue Number	GAPDH	HS3ST2	HS3ST4
1	++	++	++
2	++	++	++
3	++	++	++
4	++	++	++
5	++	++	++
6	++	++	++
7	++	++	++
8	++	++	++
9	++	++	++
10	++	++	++
11	++	++	++
12	++	++	++
13	++	++	++
14	++	++	++

(+)=positive, (-)=negative. 14/14 samples were positive for HS3ST2 expression and 14/14 samples were positive for HS3ST4 expression in normal breast tissue (100%) in the case of each gene.

Table 5-4 Correlation of Methylation and Transcription of the HS3ST2 Locus in the Cell Line MCF7

[5-aza-2-deoxycytidine]	HS3ST2 Methylation Status	HS3ST2 RTPCR
0	Methylated	Negative
0.1uM	Methylated	Positive
0.2uM	Unmethylated	Positive
0.5uM	Unmethylated	Positive
1uM	Unmethylated	Positive
2uM	Unmethylated	Positive

Table 5-5 Methylation and Transcription of the HS3ST4 Locus in the Cell Line A2780

[5-aza-2-deoxycytidine]	HS3ST4 Methylation Status	HS3ST4 RTPCR
0	Methylated	Negative
0.1uM	Unmethylated	Positive
0.2uM	Unmethylated	Positive
0.5uM	Unmethylated	Positive
1uM	Unmethylated	Positive
2uM	Unmethylated	Positive

Table 5-6 Grossly Dissected Tumours and HS3ST2 and HS3ST4 Methylation Status and RTPCR Results

Tumour Database ID.	HS3ST2M	RTPCR HS3ST2	HS3ST4M	RTPCR HS3ST4	Source
CDTDRP4	(+)	(+)	(-)	(+)	F
CDTDA64	(+)	(+)	(-)	(+)	F
CDTD101	(+)	(+)	(-)	(-)	F
CDTDW90	(+)	(+)	(+)	(-)	TB
CDTDFS2	(+)	(-)	(+)	(+)	F
CDTDHD1	(-)	(+)	(+)	(-)	F
CDTD4FT	(-)	(-)	(+)	(+)	TB
CDTDWB6	(+)	(-)	(+)	(-)	F
CDTDFF5	(-)	(+)	(-)	(+)	F
CDTD123	(+)	(-)	(-)	(+)	TB

(+) indicates a positive result, (-) indicates a negative result. TB=tumour bank, F=fresh

5.5 Discussion

The association of methylation at the HS3ST2 and HS3ST4 CpG islands with silencing of transcription is in keeping with the general consensus in the literature - dense methylation of a CpG island correlates with transcriptional silencing of that gene (Esteller 2005; Esteller 2005; Esteller 2005).

Removing methylation from these CpG islands restored transcription from the respective genes *in vitro* and this supports the idea that methylation directly inhibits transcription. However, as with all demethylating agents, the effect is not specific to the gene that is being studied. It is always possible that demethylation and re-expression of other genes, such as transcription factors, is responsible for the re-expression of HS3ST2 and HS3ST4. This is a universal confounding factor in this type of experiment and is not particular to the HS3ST loci.

This chapter, and the investigation of the expression profile of this family of genes is limited. Essentially a relationship between methylation at the HS3ST2 and HS3ST4 loci and transcriptional silencing of the respective gene is supported by these results. Of more importance is the effect that non-expression has - does lack of 3-O-sulphation in specific molecules offer a means of acquiring one of the six novel capabilities proposed by Weinberg and Hanahan?: limitless replicative potential, self-sufficiency in growth signals, insensitivity to growth-inhibitory signals, evasion of programmed cell death, sustained angiogenesis and tissue invasion and metastasis(Hanahan and Weinberg 2000). Certainly heparan sulphate molecules are involved in almost all of these processes, as can be seen from the table at the end of this chapter.

3-O-Sulphation is the rarest modification of a heparan sulphate molecule(Danishefsky, Steiner et al. 1969) and therefore it is ideally suited to selectively regulate biological activities. Individual isoforms of heparan sulphate D-glucosaminyl 3-O-sulphotransferase (HS3ST) exhibit sequence-specific action(Raman, Myette et al. 2002), which creates heparan sulphate structures with distinct biological functions. For example, HS3ST1 preferentially generates binding sites for anti-thrombin III (Zhang, Lawrence et al. 2001), whereas HS3ST3

isoforms create binding sites for the gD envelope protein of herpes simplex virus 1 (HSV-1), which enables viral entry to cells(Liu, Shriver et al. 2002).

HS3ST enzymes comprise a presumptive sulphotransferase domain and a divergent N-terminal region. Domain swaps between cDNA species were carried out in an attempt to localize the domain or motif that determines a family member's substrate specificity(Yabe, Shukla et al. 2001). In this experiment the N-terminal region of HS3ST1 was fused with the sulphotransferase domain of HS3ST3A to generate N1-ST3(A). Similarly, the N-terminal region of HS3ST3A was fused to the sulphotransferase domain of HS3ST1 to generate N3(A)-ST1. Wild-type and chimaeric enzymes were transiently expressed in COS-7 cells and extracts were analysed for selective generation of binding sites for anti-thrombin. HS3ST1 was 270-fold more efficient at forming anti-thrombin-binding sites than HS3ST3A, indicating its significantly greater selectivity for substrates that can be 3-O-sulphated to yield such sites. N3(A)-ST1 was as active as HS3ST1, whereas the activity of N1-ST3(A) was as low as that of HS3ST3A. Analysis of Chinese hamster ovary cell transfectants revealed that only HS3ST3A and N1-ST3(A) generated gD-binding sites and conveyed susceptibility to infection by HSV-1.

This experiment demonstrates that the substrate-specific properties of HS3STs are defined by the self-contained sulphotransferase domain and are not directly influenced by the divergent N-terminal region. This finding re-emphasises the specificity of each family member of the HS3ST family and shows that they have definite functions that are not redundant.

Heparan sulphate molecules can be sequenced. The technique is akin to the Maxam and Gilbert technique for sequencing DNA molecules - the terminal sugar is covalently linked to a chemical reagent, cleaved from the end of the heparan sulphate and sent through a mass spectrometer. The mass-profile of this terminal monomer reveals the sugar's identity and its sulphation pattern. Unfortunately, this technique is reliant on a pure sample of a heparan sulphate and, as can be imagined, sequencing an *in vitro* sample that is composed of many differentially modified heparan sulphates may prove to be impossible. The reason for mentioning this issue is that currently the function of HS3ST enzymes is assayed using tertiary assessments of their enzymatic activity e.g. the ability

to produce functioning anti-thrombin IIIa, or the internalisation of HSV-1 particles into cells in cell culture. There is no technique to directly detect 3-O-sulphation of heparan sulphate molecules and this is a major hurdle in understanding how each family member confers a different functional domain.

Heparan-sulphate glycosaminoglycans (HSGAGs) act at the cell-extracellular-matrix (ECM) interface to modulate cell signaling, thereby regulating how a cell perceives its environment(Perrimon and Bernfield 2000). HSGAGs also interact with various extracellular signaling molecules: growth factors, morphogens, enzymes and chemokines(DiGabriele, Lax et al. 1998; Sasisekharan, Shriver et al. 2002). The specificity of these interactions is dependent on HSGAG sequence, spacing of binding sites and the three-dimensional structure of the HSGAG chain (Gallagher 1995; Vives, Pye et al. 1999). HSGAGs, depending on location and sequence, impinge on tumour onset and progression in various ways, some of which are pro-tumourogenic and others of which are anti-tumourogenic(Berry, Kwan et al. 2001; Liu, Shriver et al. 2002; Liu, Shriver et al. 2002; Liu, Shriver et al. 2002; Shriver, Liu et al. 2002). HSGAGs at the tumour-cell surface actively modulate the tumourogenic process by regulating autocrine signaling loops that lead to unregulated cell growth(Sasisekharan and Venkataraman 2000; Liu, Shriver et al. 2002; Sasisekharan, Shriver et al. 2002). HSGAGs impinge on how an organism responds to a growing tumour, including the recruitment of cells of the immune system to the tumour site(Esko, Rostand et al. 1988), formation of a fibrin shell around the tumour that acts as a protective barrier(Stringer and Gallagher 1997; Rao and Pendurthi 1998; HajMohammadi, Enjyoji et al. 2003) and development of new blood vessels to the site of the growing tumour(Vlodavsky, Goldshmidt et al. 2002; Blackhall, Merry et al. 2003; Chen, Adenekan et al. 2004; Reiland, Sanderson et al. 2004). Compelling clinical evidence indicates that pharmacological doses of heparin, a highly sulphated HSGAG, can have a marked effect on tumour metastasis. Clinical trials have been designed to determine the exact benefits of heparin therapy in cancer(Borsig, Wong et al. 2001; Borsig 2003; Borsig 2004). With all this interest centered around the very substrate that the HS3ST enzymes act on it would be surprising if 3-O-sulphation, or lack of it, did not in some way modulate the function of heparan sulphate.

5.6 Conclusions

The findings in this chapter can be summarised as follows:

Methylation of the CpG islands of HS3ST2 and HS3ST4 is correlated with gene silencing.

Demethylation of the CpG islands at these loci is correlated with re-expression of the associated gene

The analysis of expression in the gross tumour samples is uninformative as is the analysis using micro-dissected material.

6 Materials and Methods

6.1 General Materials

6.1.1 Chemicals and Reagents

AnalaR and molecular biology grade chemicals were obtained from Fisher Scientific, Merck Ltd. (BDH Laboratory Supplies), Riedel-de Haën or Sigma Chemical Company Ltd. Other reagents obtained from alternative sources are mentioned in Table 6.1.

Table 6.1 Chemicals and reagents

Chemicals and reagents	Supplier
[α-^{32}P]dCTP (3000 Ci/nmol)	Amersham Pharmacia Biotech
30% (w/v) acrylamide/bis 29:1 (3.3% C)	BioRad
Acetic acid	Fisher Scientific U.K. Ltd.
Agar	Difco Laboratories
Agarose	Boehringer Mannheim
Bacto® tryptone	Difco Laboratories
Bacto® yeast extract (without amino acids)	Difco Laboratories

Boric acid	Fisher Scientific U.K. Ltd.
Chloroform	Fisher Scientific U.K. Ltd.
Ethanol, absolute 100%, analytical reagent	Bamford Laboratories
ExpressHyb hybridization solution	Clontech
Hydrochloric acid	Fisher Scientific U.K. Ltd.
IPTG	Boehringer Mannheim
PMSF	Boehringer Mannheim
Sephadex® G-50	Amersham Pharmacia Biotech
Sepharose Cation Exchange Matrix	Amersham Pharmacia Biotech
Sodium dodecyl sulphate	Anachem
X-gal	Boehringer Mannheim

6.2 General Disposable Plasticware

General disposable plasticware materials used during the course of this project are listed in Table 6.2 below.

Table 6.2. Plastic materials and suppliers.

Plastic Materials	Supplier
0.2 ml micro-tubes	ABgene®
5ml Chromatography Columns	Novagen
15 ml centrifuge tubes, gamma irradiated	Sterilin
20 µl, 200 µl and 1 ml filter pipette tips	Rainin Instrument Co. Inc. and Greiner Labortechnik
200 µl and 1 ml pipette tips	Sarstedt
50 ml centrifuge tubes	Sterilin
90 mm Petri dishes	Philip Harris Scientific
Thin-wall polycarbonate 96-well plates and thermosealers	Costar

6.2.1 Tissue Culture Disposable Materials

Tissue culture disposable plastic materials are listed in Table 6.3 below.

Table 6.3. Tissue culture materials and suppliers.

Disposable Materials	Supplier
100 x 20 mm polystyrene tissue culture dishes	Corning
25 cm^2 tissue culture flasks	Costar and Iwaki
5, 10 and 25 ml pipettes	Corning
6-, 12- and 96-well tissue culture microplates	Iwaki
60 x 15 mm polystyrene tissue culture dishes	Corning
75 and 150 cm^2 tissue culture flasks	Iwaki
Cell scrapers	Sigma
Cryo 1°C freezing container	Nalgene

6.2.2 Enzymes

The enzymes used for DNA analysis are listed in Table 6.4, together with their suppliers, from where their reaction buffers were also obtained.

Table 6.4. Enzymes and suppliers.

Enzyme	Supplier
Taq DNA polymerase	Bioline
T4 DNA ligase	GibcoBRL Life Technologies
EcoRI	GibcoBRL Life Technologies
MseI	New England Biolabs
MspII	New England Biolabs
HhaI	New England Biolabs
HpaII	New England Biolabs
*Sss*Methyltransferase	New England Biolabs
*Hha*Methyltransferase	New England Biolabs
Proteinase K	Boehringer Mannheim
Superscript II	Invitrogen

6.3 Nucleic Acids Size markers and Mass Ladders

Following gel electrophoresis, DNA and RNA samples were sized and/or quantified using the following size markers or mass ladders listed in Table 2.5.

Table 6.5. Nucleic acid size markers and/or mass ladders and their suppliers.

Size marker and/or mass ladder	Supplier
1 kb^+ ruler	GibcoBRL Life Technologies
2.5 kb molecular ruler	BioRad
*Hae*III digested ϕX174 DNA	New England Biolabs
High MW mass ladder	GibcoBRL Life Technologies
*Hind*III digested λ DNA	New England Biolabs
Low DNA mass ladder	GibcoBRL Life Technologies
RNA ladder	GibcoBRL Life Technologies

6.3.1 Kits

The kits used during the course of this project are listed in Table 6.6.

Table 6.6. Kits and suppliers.

Kit	Supplier
CpG Modification Kit (S7820)	Chemicon International Serologicals
Nucleon® BACC kit	Nucleon
PGEM®-Teasy vector system	Promega
QIAprep® miniprep kit	Qiagen
QIAquick® spin kit	Qiagen
Ready-to-go DNA labelling beads (-dCTP)	Amersham Pharmacia Biotech

6.3.2 Membranes and Paper

Hybond™-N⁺ and Hybond™-N nylon membranes were obtained from Amersham Pharmacia Biotech.

Schleicher & Schuell supplied gel blotting paper. Saran wrap was obtained from Dow.

6.3.3 Photography and Autoradiography

Agarose gels were photographed using a UVP gel documentation system 7500. Konica X-ray film and autoradiography cassettes were obtained from Genetic Research Instrumentation Ltd. Films were developed using an X-Ograph Compact X2 system (X-Ograph Ltd.). Both fixing and developer solutions were supplied by Kodak.

Micro-dissection photographs were taken with a Nikon coolpix digital camera.

6.3.4 Microscopes

An Arcturus laser dissection microscope was used during the course of this project for micro-dissection of patient samples. Cells were collected on an CapSure® LCM Caps.

6.3.5 Spectrophotometer

A Shimadzu UV-1201 UV-Vis spectrophotometer was used to determine purity and concentration of nucleic acids in aqueous solution.

6.3.6 DNA Crosslinker

A Stratagene® UV crosslinker 2400 was used during the course of this project, to fix DNA onto nylon membranes.

6.4 Experimental Materials

6.4.1 Bacterial Host Strains

Escherichia coli TOP10 strain was obtained from Invitrogen, Dh 5 alpha strain from Invitrogen, XL1-Blue Supercompetent and XL2-Blue Ultracompetent strains were obtained from Stratagene. E.Coli strain BL21 (DE3) pLys was obtained from a communal laboratory stock.

6.4.2 Vectors

pGEM® T-Easy bacterial vector was obtained from Promega and pCR2.1 vector was purchased from Invitrogen.

6.4.3 DNA Sources

pET6HMB, a modified form of pET11D (Novagen), containing the Rat MecP2 HMB domain coding region, was gifted from Dr R. Brock.

6.4.4 Oligonucleotides

The oligonucleotide primers used in this project were obtained from Genosys. The primer sequences are listed in Table 2.7.

Table 6.7. Oligonucleotides.

Name, 5'-3' sequence, melting temperature (Tm) and target sequence.

Oligonucleotide	Sequence 5'-3'	Tm (°C)	Product	CpGs	C's
HS3ST1UMF	GGTTGTGTTGTTTTTGGGTGTGGT	64 °C	102bp	4	9
HS3ST1UMR	TCCCCACTTTCCACAAAAACTCCT	64 °C		2	4
HS3ST1MF	CGTTTATGTGTTATTGATTAGGAGGT	65 °C	283bp	1	9
HS3ST1MR	CTTACGCACGCCGATAATAA	65 °C		3	6
HS3ST1F	ATTGATCTGC GGCGACAGCT	64 °C	723bp	na	na
HS3ST1R	CCCTAGTAAT TGAAGGCCTG	64 °C		na	na
HS3ST2UMF	TGGGAGTGTTTGAGTTGTTTG	59 °C	278bp	4	9
HS3ST2UMR	TCCATAAACCCACACCAAAA	59 °C		3	9
HS3ST2MF	CGTAAGAGTTTGGGAGCGTTCG	65 °C	285bp	3	7
HS3ST2MR	ATAAACCCGCGCCGAAAATC	65 °C		3	10
HS3ST2F	GGCATTTCCCGAAGAGCCAGA	64 °C	713bp	na	na
HS3ST2F	CTTGGGTGAGCCGGAGTGGTT	64 °C		na	na

HS3T3AUMF	TTTGAGTGGGAGATAGGAGGTTTGTG	66 °C	114bp	3	7
HS3ST3AUMR	CCAACCCAACCCAACCACAC	66 °C		3	4
HS3ST3AMF	TTCGTTTGGGCGCGTTTT	62 °C	252bp	3	8
HS3ST3AMR	AAAACAACCGATAAACGCCTACGA	62 °C		1	10
HS3ST3AF	ATGGGGGACA CTGGGGAAGG	59 °C	663bp	na	na
HS3ST3AR	GGTCTTGGCA GCGGTGTCG	59 °C		na	na
HS3ST3BUMF	AGTTTTGGGTTGGATTTATTTGG	60 °C	264bp	3	10
HS3ST3BUMR-	ACACATAAAAATTTCACAACACAACC	60 °C		4	9
HS3ST3BMF	AGTTAAGTTTCGGGTCGGATTT	60 °C	222bp	2	8
HS3ST3BMR	TTCACGACACAACCTACGACTC	60 °C		2	4
HS3ST3BF	Cloned from GB library		875bp	na	na
HS3ST3BR	Cloned from GB library			na	na
HS3ST3BE1.1UMF.1	TGGTTGTGGGTATATGGGGGTAA	65 °C	242bp	4	8

HS3ST3BE1.1UMR.1	ACAACACAAACTTCCTCCTCACCA	65 °C		3	6
HS3ST3BE1MF	TGTTCGGCGGGATTTACGTT	65 °C	149bp	3	10
HS3ST3BE1MR	CTCACCGACGACGAAAACGA	65 °C		4	11
HS3ST3BE1F	TTCCTGCGCA GTTCGCCTCT	59 °C	654bp	na	na
HS3ST3BE1R	CGGCCATCTC CTTGCCTGAA	59 °C		na	na
HS3ST4UMF	TTTATTTAAGGAGGTGTTGTTTGA	57 °C	223bp	0	6
HS3ST4UMR	TAAATTCCAACCACACAAATACA	57 °C		2	5
HS3ST4MF	TTTTGTTTTGTTTTTCGCGTTT	59 °C	240bp	2	14
HS3ST4MR	CCAACCACGCAAATACGTTTA	59 °C		2	5
HS3ST4F	TTCCACGCCCTTCGAGCATC	62 °C	778bp	na	na
HS3ST4R	CGGAGGTGGGAGGCATCCTT	62 °C		na	na

6.5 Solutions

6.5.1 General Solutions

λ DNA x *Hin*dIII and ϕx174 DNA x *Hae*III size ladders

25 ng/µl λ DNA digested with *Hin*dIII and 25 ng/µl ϕx174 DNA digested with *Hae*III, 1X TE, 1X DNA loading dye.

0.5M EDTA pH 8.0

0.5M EDTA with NaOH to pH 8.0.

1 kb+ ladder

60 ng/µl 1 kb ladder, 1X DNA loading dye in 1X TBE.

10% (w/v) SDS

10 % (w/v) SDS in H_2O.

100X Denhardt's solution

2% (w/v) Ficoll®400, 2% (w/v) polyvinylpyrolidone, 2% (w/v) BSA.

100X Ethidium bromide-acridine orange stock solution

2.75 mM ethidium bromide, 1.25 mM acridine orange, 2% (v/v) ethanol in H_2O.

1M Tris•HCl pH 8.0

1 M Trizma base and HCl to pH 8.0.

1X TBE

90 mM Trizma base, 90 mM orthoboric acid, 2 mM EDTA.

2.5 kb molecular ruler

333 ng/µl of 2.5 kb molecular ruler in 1X TE, 1X Orange G loading dye.

20X SSC

3.0 M NaCl, 0.3 M sodium citrate.

20X SSPE

3.0 M NaCl, 0.2 M NaH_2PO_4, 20 mM EDTA, pH 8.0.

3 M Sodium acetate pH 7.5

3 M sodium acetate with NH_3 to pH 7.5.

5X DNA loading dye

0.5% (w/v) SDS, 0.25% (w/v) xylene cyanol, 0.25% (w/v) bromophenol blue, 1.5% (w/v) Ficoll®400, in 3X TBE.

5X Orange G loading dye

0.06 % (w/v) Orange G, 50 % (v/v) glycerol in H_2O.

75% (v/v) Ethanol

75% (v/v) absolute ethanol in H_2O.

80% (v/v) Ethanol

80 % (v/v) absolute ethanol in H_2O.

Amplisize molecular ruler

333 ng/µl of amplisize molecular ruler in 1X TE, 1X Orange G loading dye.

Denaturing solution

0.5 M NaOH, 1.5 M NaCl in H_2O.

Ethidium bromide

Stock solution: 25 mM in H_2O.

Working concentration: 500 nM.

Lysis buffer

50 mM Tris•HCl pH 8.0, 100 mM EDTA pH 8.0, 0.5% (w/v) SDS in H_2O.

Neutralising solution

1.5 M NaCl, 0.5M Trizma base and HCl to pH 7.5.

PCR buffer

45 mM Tris•HCl pH 8.8, 11 mM ammonium sulphate, 4.5 mM $MgCl_2$, 6.7 mM 2-mercaptoethanol, 4.4 µM EDTA, 1 mM dATP, 1 mM dCTP, 1 mM dGTP, 1 mM dTTP and 113 µg/ml BSA.

Phenol

Phenol saturated in 10 mM Tris pH 8.0, 1 mM EDTA.

Phenol:chloroform:isomayl alcohol, 25:24:1

Phenol:chloroform:isomayl alcohol, 25:24:1, saturated in 10 mM Tris pH 8.0, 1 mM EDTA.

Proteinase K

Stock solution: 20 mg/ml proteinase K in filter sterilised 50 mM Tris•HCl pH 8.0.

TAE

40 mM Tris•acetate, 1 mM EDTA in H_2O.

TE

10 mM Tris•HCl pH 8.0, 1 mM EDTA pH 8.0 in H_2O.

6.5.2 Bacterial Solutions, Media and Antibiotics

Ampicillin

Stock solution: 50 mg/ml in H_2O (stored at -20°C).

Working concentration: 100 µg/ml.

Chloramphenicol

Stock solution: 50 mg/ml in H_2O (stored at -20°C).

Working concentration: 20ug/ml.

IPTG

Stock solution: 100 mg/ml in H_2O (stored at -20°C).

Working concentration: 10 µg/ml.

Luria-Bertani (LB) medium

3% (w/v) Bacto® tryptone, 0.5% (w/v) Bacto® yeast extract, 1% (w/v) NaCl.

LB agar contained 1.5% (w/v) agar.

SOB medium

2% (w/v) Bacto® tryptone, 0.5% (w/v) Bacto® yeast extract, 0.85 mM NaCl, 0.25 mM KCl pH 7.0 with NaOH. Sterilized 10 mM $MgSO_4$ added prior to use.

SOC medium

0.04% (w/v) glucose in SOB medium.

X-gal

Stock solution: 50 mg/ml in dimethylformamide (stored at -20°C).

Working concentration: 50 µg/ml.

6.5.3 *Tissue Culture Solutions, Media and Antibiotics*

Tissue culture media, serum, antibiotics and solutions were obtained from GibcoBRL Life Technologies or Sigma.

Dulbecco's modified Eagle medium (DMEM)

DMEM with 862 mg/l L-alanyl-L-glutamine (GlutaMax^{l}), 4 mg/l pyrodoxine•HCl, 4500 mg/l glucose, 110 mg/l sodium pyruvate.

Trypsin/EDTA solution

Stock solution: 5.0 g/l trypsin, 2.0 g/l EDTA, 8.5 g/l NaCl (stored at -20°C).

Working concentration: 0.5 g trypsin, 0.2 g EDTA•4Na/l in PBS.

Penicillin and streptomycin solution

Stock solution: 10,000 U/ml penicillin, 10,000 µg/ml streptomycin. Utilising penicillin G (sodium salt) and streptomycin sulphate: prepared in normal saline (stored at -20°C).

Working concentration: 100 U/ml penicillin, 100 µg/ml streptomycin.

10X Dulbecco's phosphate buffered saline (PBS)

8 g/l NaCl, 0.2 g/l KCl, 2 g/l KH_2PO_4, 2.16 g/l $Na_2HPO_4 \bullet 7H_2O$.

Foetal bovine serum (FBS)

Origin E.C. Virus and mycoplasma tested.

6.6 Tissue Culture Methods

6.6.1 Feeding Cultured Cells

Cells were grown in a 5% CO_2 incubator at 37°C. Initially cells were cultured in 25 cm^2 flasks during the first 5-10 passages. All solutions (cell growth medium, PBS and trypsin/EDTA) were warmed to 37°C. The old medium (5 ml) was removed carefully, avoiding scratching the surface of the cells. Cells were washed gently with 3 ml of PBS, and 5 ml of fresh culture medium was added. The same standard protocol was adapted for cells growing in 6-well clusters: each well was washed with 2 ml of PBS and the cells grown in 3 ml of cultured medium.

6.6.2 Subculturing Cultured Cells

Cells growing in 25 cm^2 flasks were washed in PBS as described before: 2 ml of trypsin/EDTA solution added and the cells incubated at 37°C, 5% CO_2 for 5 minutes. The cells were observed under the microscope to check if they had rounded up and lifted off the surface of the culture flask. The flask was tapped gently against the bench to help the cells dissipate from the surface of the flask. If necessary, the cells were incubated for an extra two minutes and a cell scraper used to make sure that nearly every cell was floating. Trypsin digestion was stopped by the addition of 3 ml of fresh culture medium. The cell suspension was pippetted up and down repeatedly to dissociate cell clumps, and transferred into a 15 ml falcon tube. The cells were collected by centrifugation at 200 *g* for 5 minutes, the supernatant removed and the cell pellet resuspended in 5 ml of fresh culture medium. A variable amount of the cell suspension, dependent on the split ratio (1:5 or 1:10), was collected and mixed with fresh growth medium, up to a final volume of 5 ml. For the higher split ratios cells were not collected by centrifugation after trypsin digestion

prior to seeding. Instead, a small aliquot of the cell suspension (usually 100 µl or 125 µl for split ratios of 1:50 or 1:40, respectively) was mixed with 5 ml of fresh medium and plated in a fresh 25 cm^2 tissue culture flask. Cells were finally returned to the 37°C, 5% CO_2 incubator. At every passage the number of population doublings was determined based on the cell number.

6.6.3 Measuring Cell Counts and Determining Population Doubling Times

Following trypsin digestion and neutralisation with standard growth medium, cells were counted on a haemocytometer, using a phase contrast microscope. The haemocytometer was covered with the coverslip and a drop of cell suspension was dropped at the edge of the coverslip on both sides of the chamber. At least 100 cells were counted, and the number of cells/ml calculated. The number of population doublings (PD) was determined, based on the cell number for two consecutive passages as follows:

Equation 1

$$PD = \log_2 \left[\frac{\text{cell number at passage n+1}}{\text{cell number at passage n}} \times \text{split ratio} \right]$$

Finally, the population doubling times were calculated by dividing the number of population doublings, by the time over which they had occurred:

$$PDT = PD / \text{time}$$

6.6.4 Freezing Cultured Human Cells in Liquid Nitrogen

A single cell suspension was obtained following trypsin digestion as described before (section 6.6.2). Cells were precipitated by centrifugation at 200 *g* for 5 minutes, resuspended in fresh culture medium at a concentration of ~1.3-3 x 10^6 cells/ml and transferred into 2 ml cryovials. Dymethylsulphoxide (DMSO) was added to a final concentration of 10% (v/v), and then mixed well by gently inverting the tubes.

The vials were transferred into a freezing container with isopropanol at room temperature, and cooled to -70°C for at least 4 hours. Finally the frozen samples were moved to liquid nitrogen (-190°C) until needed.

6.6.5 Thawing Cultured Human Cells

Frozen cell samples stored in liquid nitrogen were quickly thawed by immersion of the cryovials in a water bath at 37°C. Five ml of fresh warm medium were added to the cells, and mixed gently by repeated pipetting. The cells were then collected by centrifugation at 200 *g* for 5 minutes, resuspended in fresh culture medium, and finally plated on a tissue culture flask or plate.

6.7 Preparation, purification and analysis of DNA

6.7.1 Preparation of Plasmid DNA

Qiagen kits were used in order to obtain high yields of good quality plasmid DNA. DNA was purified from 5-ml bacterial cultures by alkaline lysis followed by anion-exchange columns, according to the manufacturer's protocol.

6.7.2 DNA Extraction from Human Tissues

A 10mg to 20 mg piece of breast tissue was milled and placed in a 1.5 ml screw top Eppendorf tube with 550 µl of lysis buffer and 15 µl of 20 mg/ml of proteinase K to a final concentration of 530 µg/ml of proteinase K. The tissue was incubated overnight at 60°C. The lysate was briefly centrifuged at 21,000 *g* to precipitate the debris. Two hundred and fifty µl of supernatant was transferred into a fresh tube, an equal volume of phenol was added and mixed vigorously by vortexing for 5 seconds to emulsify the two phases. The mixture was centrifuged for 5 minutes at 21,000 *g* to separate the two phases. The upper phase was placed in a clean tube, 250 µl of phenol:chloroform:isoamyl alcohol (25:24:1) added and mixed vigorously by vortexing for 5 seconds. The upper phase was removed and transferred to a clean tube. To remove any remaining traces of phenol, 250 µl of chloroform was emulsified with the aqueous phase by vortexing for 5 seconds.

The upper aqueous phase was transferred to a clean tube along with 25 µl of 3M sodium acetate, pH 7.5 and mixed briefly by inverting the tube 5 to 10 times. The DNA was finally precipitated by the addition of 500 µl of ice-cold absolute ethanol, incubated at -20°C for at least 1 hour, and spun at 21,000 *g* in a bench top centrifuge. Supernatant was decanted, the pellet rinsed with 1 ml of ice-cold 80% (v/v) ethanol and precipitated by centrifugation at 21,000 *g*

in a top bench centrifuge. The DNA pellet was either air dried at room temperature for 30 to 60 minutes, or at 4°C overnight. Once dried the DNA pellet was dissolved in 100 to 200 µl of TE and heated to 60°C for 30 minutes, or at 37°C overnight.

6.7.3 DNA Extraction from Cultured Human Cells

Cell culture DNA samples were extracted using a Nucleon DNA extraction kit for blood and tissue culture, following the manufacturer's protocol. Briefly, cells were collected by centrifugation at 800 *g* for 5 minutes, resuspended in washing "Buffer MA"[1] and left on ice for 5 minutes. Cells were collected by centrifugation at 1,300 *g* and subsequently lysed in "Reagent B"[2] by gently pipetting up and down, followed by a 30-minute incubation at 37°C. Proteins were precipitated by the addition of 5M sodium perchlorate and subsequently extracted by the addition of chloroform, Nucleon resin and centrifugation at 370 *g* for 2 minutes. The upper aqueous phase was removed and transferred into a fresh 1.5 ml Eppendorf tube, without disturbing the resin layer or the chloroform phase. An extra centrifugation step at 5,000 *g* for 2 minutes was performed to pellet down any resin that might have been carried over. The supernatant was collected and transferred into a fresh 1.5 ml Eppendorf tube. DNA was precipitated by the addition of ice-cold absolute ethanol and centrifugation at 21,000 *g* for 10 minutes. The supernatant was discarded and the DNA pellet washed in ice-cold 80% (v/v) ethanol and either air-dried at room temperature for 30 minutes, or at 4°C overnight. Once dried, the DNA pellet was dissolved in 100 µl of TE at 60°C for 30-60 minutes, or at 37°C overnight, and finally stored at -20°C.

6.7.4 Determination of DNA Concentration

When necessary, the concentration and purity of DNA in aqueous solution was established by measuring the UV absorbance of the solution at wavelengths ranging from 200-300 nm. Pure double-stranded DNA solutions have an absorbance maximum at 260 nm, at which an optical density of 1 corresponds to 50 mg/ml of DNA in the solution. The purity of DNA samples was calculated by comparing the ratio of the OD at 260 nm (OD_{260}) to 280 nm (OD_{280}), at which wavelength proteins have an absorbance maximum. An OD_{260}/OD_{280} ratio of 1.8 or greater is taken as an acceptable level of purity. DNA samples were diluted 1:100 in H_2O and the spectrophotometer baseline corrected with H_2O. The OD was measured and the concentration calculated as 100 x 50 x OD_{260} mg/ml.

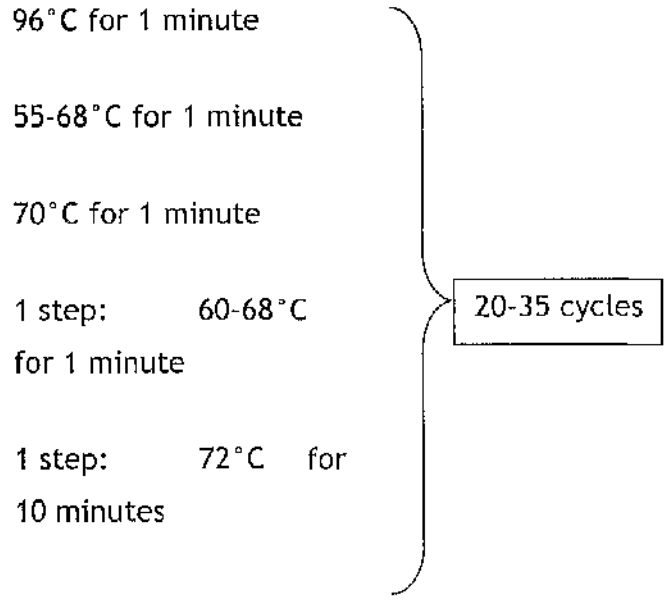
Alternatively, DNA concentrations were determined by electrophoresis on agarose gels, followed by densitometry analysis against low molecular weight DNA mass ladder (GibcoBRL, Life Technologies), using Kodak Digital Science 1D software (Kodak).

6.7.5 Polymerase Chain Reaction (PCR)

All PCR amplifications were carried out in either a Biometra Uno thermal cycler or a Biometra T3 thermal cycler, in either ABgene 96 well plates or ABGene thin walled 200 µl tubes. Each reaction was overlaid with white light mineral oil and then sealed. The thermal cycler lid was preheated to 105°C.

PCR amplifications were performed with 1 µM of each primer, 1X PCR buffer and 0.5 U *Taq* DNA polymerase in a total volume of 10 µl. The annealing temperature varied according to the primers selected to carry out the amplification, but in general fell within a range from 55°C to 68°C. Reactions were thermal cycled under the following conditions:

Table 6.8 Generic PCR Cycle



6.7.6 DNA Cloning Techniques

6.7.6.1 Restriction Endonuclease Digestion of Plasmid DNA

Restriction digests were carried out with 1-5 µg of purified plasmid DNA, 5-10 units of restriction endonuclease and 1X manufacturer's restriction endonuclease reaction buffer in a total of 10 µl with H_2O were incubated at 37°C for 90-120 minutes.

6.7.6.2 DNA Ligation

Bacteriophage T4 DNA ligase was used to catalyse the formation of phosphodiester bonds between the 3'-hydroxyl groups and the 5'-phosphate groups of the DNA inserts and vectors. A recombinant plasmid with 2 single-stranded nicks results owing to the removal of the 5'-phosphate groups from

the digested vector. The nicks are repaired after the recombinant plasmids have been introduced into competent bacteria.

Ligations of vector and insert DNA were incubated at 4°C overnight with 1 U of bacteriophage T4 DNA ligase and 1X T4 DNA ligase buffer in a total volume of 10 µl. Generally, DNA molecules were ligated at vector:insert molar ratio of 1:3. Control ligations, with one component missing, were also incubated at 4°C overnight.

6.7.6.3 Transformation of Competent Bacterial Cells

Fifty µl aliquots of competent cells were thawed on ice for 10 minutes. One µl of the ligation mixture or 10 ng of control plasmid were added and after gentle mixing, the bacteria were incubated on ice for 30 minutes. The cells were heat-shocked at 42°C for 45 seconds and cooled on ice for 2 minutes. Pre-warmed SOC medium (450 µl) was added to the cells and the samples incubated, with shaking (225rpm), at 37°C for 1 hour. The transformed cells were split into two aliquots of 50 µl and 450 µl, which where subsequently spread on a Luria-Bertani (LB) plate containing 100 µg/ml ampicillin, 50 µg/ml 5-bromo-4-chloro-3-indolyl-ß-D-galactoside (X-gal) and 10 µg/ml isopropylthyo-ß-D-galactoside (IPTG), dried for up to 30 minutes, and incubated at 37°C overnight. Ampicillin, X-gal and IPTG are used to positively select transformed bacteria, which are resistant to the antibiotic ampicillin, and are white because of the insertional inactivation of the *lacZ* gene of the recombinant plasmid. As a result β-galactosidase is not synthesised and the bacterial colonies are white rather than blue.

6.7.6.4 Generation of Plasmid Stocks

Stocks of bacteria transformed with recombinant plasmids were prepared for long-term storage. A colony from an agar plate was used to inoculate 5 ml of

LB medium, containing 100 µg/ml of ampicillin and incubated at 37°C overnight. An aliquot of overnight culture (usually 500 µl) was mixed with an equal volume of 2% (w/v) peptone, 40% (v/v) glycerol, by repeated inversion. The stocks were stored at -70°C until needed.

6.7.6.5 Gel Electrophoresis

6.7.6.5.1 Agarose gel electrophoresis

DNA molecules were separated according to their size by agarose gel electrophoresis. Solutions of 0.8-2% (w/v) of agarose in 0.5X TBE were prepared in a microwave, the solution was cooled to approximately 50°C, ethidium bromide was added to a final concentration of 500 nM and the gel was cast in a horizontal tray. DNA samples were mixed with 1X DNA loading dye and electrophoresed in 0.5X TBE for periods of time varying from 2-16 hours. An aliquot of a DNA size ladder (usually 100-200 ng) was also electrophoresed and used as a size marker. Separated DNA samples were visualised using an UV transilluminator (wavelength 254 nm) and photographed.

6.7.6.6 Southern Blotting

DNA transfer from agarose gels onto a nylon membrane by Southern "squash" blotting

Once the gels had run sufficient distance, a scalpel and a stainless steel ruler were used to cut off excess gel not required. Gels were rinsed in deionised water, incubated in depurinating solution for 10 minutes, denaturing solution for 30 minutes and neutralising solution for 30 minutes and rinsed in deionised water between incubations. incubations were performed with gentle shaking.

A piece of Magna nylon membrane the same size as the gel was wet in deionised water first, and then in neutralising solution. A piece of Saran Wrap, somewhat larger in size than the gel, was placed on the bench, with a piece of

gel blotting paper, the same size as the gel and wet in neutralising solution, layered on top. The gel was placed on the Saran Wrap, and the membrane layered on top of the gel. Any air bubbles trapped under the membrane were carefully removed. Two sheets of gel blotting paper were rinsed in neutralising solution and layered onto the membrane. The blot was topped with a thick layer of paper towels, a glass plate and a weight (500 g for a small gel, 1 kg for a large gel). The DNA was transferred from the gel onto the membrane by capillary action for 3 up to 16 hours. The blot was dismantled in reverse order and the membrane placed on a piece of dry gel blotting paper, the DNA side up. The membrane was baked for at least 20 minutes in an 80°C oven and the DNA fixed to the membrane by exposure to 1,200 J/m^2 in a UV crosslinker.

6.7.7 Preparation of Radiolabelled Double-Stranded Probes

Double-stranded DNA probes were radiolabelled using α-^{32}P-labelled 2'-deoxycitidine-5'-triphosphate ([α-^{32}P]dCTP) and the Ready-to-go kit, following the manufacturer's instructions. Briefly, 30 ng of DNA double-stranded probe, 5 ng of DNA size marker and H_2O to a final volume of 45 µl were denatured at 100°C for 5 minutes and quenched on ice for 3 minutes. The lyophilised reaction mix was resuspended in the DNA solution. Five µl of [α-^{32}P]dCTP (3,000 Ci/nmol) were added and the mixture incubated at 37°C.After 1 hour the reaction was terminated by the addition of 250ul of ddH20

Double-stranded DNA probes were purified using a 1.0 ml Sephadex® G-50 column, prepared in a 1.0 ml plastic syringe. The Sephadex® was compacted by centrifugation at 1,800 *g* for 2 minutes. The probe was added to the column and centrifuged at 2,000 *g* for 3 minutes. The eluate was collected and the percentage of incorporation calculated by taking counts per minute (cpm) readings before and after purification.

6.7.8 Bisulphite Modification of DNA

Depending on the sample used the DNA was modified with a Chemicon CpG modification kit, according to the manufacturer's protocol, or using the following protocol.

6.7.8.1 Denaturation.

1ug of DNA was digested using either MseI or BglII in a total volume of 25ul, the reaction was incubated at 36°C for 3 hours. Denaturation of the DNA was carried out by adding 2.77ul of 3M NaOH to give a final concentration of 0.3M NaOH. This solution was incubated for 15minutes at 37°C.

6.7.8.2 Deamination.

A solution of 4M sodium bisulphite and 6.25M urea was made up freshly for each reaction, pH was adjusted to 5 using concentrated NaOH. 22ul of the digestion reaction was added to 208ul of the sodium bisulphite reaction and hydroquinone was added to a final concentration of 0.5mM. The final concentrations of urea and sodium bisulphite were 5.36M and 3.44M respectively. This reaction was overlaid with mineral oil and subjected to the following program, using the Biometra thermal cycler: 55°C for 15minutes followed by 30 seconds at 95°C for 25 cycles - this reaction takes approximately 6-8hours depending on the machine.

6.7.8.3 Desulphonation.

The reaction was then cleaned using a Quiagen PCR purification column and eluted in 50ul of TE (Tris-EDTA). 5.5ul of 3M NaOH was added to the eluent and incubated for 20minutes at 36°C.

6.7.8.4 Neutralisation and Desalting.

65ul of 5M NH_4OAc, 2ul of 20mg.ml^{-1} glycogen solution, and 360ul of 100% ethanol were added to each reaction. The mixture was stored at -20°C overnight and centrifuged at 16Krpm on a benchtop centrifuge at 4°C and the supernatant removed. The pellet was then washed twice with 70% ethanol and allowed to air dry for 20mins before being dissolved in 30ul of TE pH8 and aliquoted into 10ul volumes, before being stored at -20°C.

1ul of each reaction was used as a template for MSPCR.

6.8 Preparation, Purification and Analysis of RNA

6.8.1 RNA Extraction from Cultured Human Cells and Human Tissue

RNA is vulnerable to degradation from nucleases, to optimise the successful handling of RNA, the following precautions were taken 1) RNA work was restricted to a dedicated bench. 2) Dedicated equipment and stock reagents were used wherever possible and gloves were worn at all times.3) All solutions required for RNA related techniques were made using diethylpyrocarbonate (DEPC) treated water. 4) DEPC was added to a concentration 0.1% (v/v) to deionised water, mixed and allowed to stand overnight before the inactivation of the DEPC by autoclaving. 5) All plasticware was soaked in active DEPC-treated water before autoclaving, all glassware was soaked in active DEPC water before baking.

Cells were grown on a 75cm^2 tissue culture flask until confluence was achieved. Approximately 6 x 10^6 cells were collected after trypsin digestion (see section 6.6.2), washed once in 1X PBS and lysed in 1 ml of Tri Reagent™ by repeated pipetting. The lysate was allowed to stand for 5 minutes, to ensure complete dissociation of nucleoprotein complexes. Following cell lysis, 200 µl of chloroform were added, the mixture was shaken vigorously for 15 seconds and allowed to stand for 15 minutes at room temperature. The mixture was then separated into three phases by centrifugation at 12,000 *g* for 15 minutes at 4°C. The aqueous phase was transferred into a fresh tube, 500 µl of isopropanol were added and the mixture was allowed to stand for 10 minutes at room temperature. The RNA was precipitated by centrifugation at 12,000 *g* for 10 minutes at 4°C. The supernatant was removed, the RNA pellet washed in 1 ml of 75% (v/v) ethanol and collected by centrifugation at 7,500 *g* for 5 minutes at 4°C. The RNA pellet was let to air-dried for 10-30 minutes and

finally resuspended in 40 µl of DEPC treated ddH_2O. Repeated pipetting at 60°C for 15 minutes facilitated RNA solubilisation.

6.8.2 Determination of RNA Concentration

The concentration and purity of RNA in aqueous solution was quantified by measuring the UV absorbance of the solution at wavelengths ranging from 200-300 nm. Pure RNA solutions have an absorbance maximum at 260 nm, at which a solution containing 40 mg/ml of RNA gives rise to an absorbance of 1. The purity of the DNA samples was estimated by comparing the ratio of the OD at 260 nm to 280 nm, at which wavelength proteins have an absorbance maximum. An OD_{260}/OD_{280} ratio of 1.65 or greater is taken as an acceptable level of purity. RNA samples were diluted 1:80 in H_2O and the spectrophotometer baseline corrected with H_2O. The OD was measured and the concentration calculated as 80 x 40 x OD_{260} mg/ml.

6.8.3 RNA Visualisation by Agarose Gel Electrophoresis

In order to assess RNA integrity and quality, RNA samples were electrophoresed through a 1% (w/v) agarose gel containing 0.1% (w/v) SDS, 500 nM ethidium bromide, and prepared in 1X TAE. Five µg of RNA were diluted in DNA loading buffer to a final volume of 10 µl, heated at 65°C for 5 minutes and immediately cooled on ice. The samples were electrophoresed at 100 V for 1 hour. 100-200 ng of 1 kb^+ DNA size marker was also electrophoresed. Separated RNA molecules were visualised using a UV transilluminator (wavelength 254 nm) and photographed.

6.9 Slide Preparation for Laser Capture Micro-dissection

The sections were placed in the following solutions for the annotated times.

1.	100% ethanol 15 seconds
2.	95% ethanol - 15 seconds
3.	70% ethanol - 15 seconds
4.	Deionized water - 15 seconds
5.	Mayer's Hematoxylin - 30 seconds
6.	Deionized water - rinse (x 2) - 15 seconds
7.	70% ethanol - 15 seconds
8.	Eosin Y - 5 seconds
9.	95% ethanol - 15 seconds
10.	95% ethanol - 15 seconds
11.	100% ethanol - 15 seconds
12.	100% ethanol - 15 seconds
13.	Xylenes (to ensure dehydration of the section) - 60 seconds
14.	Air-dry for approximately 2 minutes.

Table 6.9 Slide Preparation for LCM

This protocol was adapted from the NCI protocols database. The critical issue in micro-dissection appears to be the elimination of all water from the slides that are to be harvested. We only used tissue that had been frozen in liquid

nitrogen. This was then fixed in 100% ethanol following cryosectioning. Myer's haemotoxylin solution (Sigma), in varying concentrations was used for staining, along with Eosin Y solution (Sigma). The tissue is now ready for LCM.

The sections were then microdissected on an Arcturus LCM using the Capsure™ lids. DNA was harvested from the material by overnight digestion with proteinase K solution. RNA was collected using the Picopure™ RNA kit from Arcturus, as per the manufacturer's instructions.

6.10 Patient Samples

Breast tumour samples used in this study, obtained from patients undergoing mastectomies or lumpectomies, were classified as infiltrating ductal carcinomas or lobular carcinomas. Adjacent normal parenchyma was obtained to serve as a control. DNA (5 µg) was extracted from a matched pair of samples (36 were derived from tumour and 36 from the adjacent normal tissue) using standard techniques then digested with *Mse*I, 10 U (NEB) for 3-4 h. The digested DNA from the normal sample was fractionated using an MBD column as described. This was followed by washing with 20 ml of buffer containing 1.0 M NaCl to remove any remaining bound DNA. The column was then equilibrated by washing with 20 ml of buffer containing 0.1 M NaCl followed by the fractionation of samples derived from tumour DNA. Samples eluting from the MBD column between 0.75 and 0.85 M NaCl were precipitated and used in this analysis.

6.11 ICEAMP

6.11.1 *Manufacture of HMBD Column*

The recombinant, His-tagged MBD (HMBD) from *Rattus norvegicus* MeCP2 was expressed in *Escherichia coli* BL21(DE3)LysS using the plasmid pET6HMBD, kindly donated by Dr. Graham Brock and used previously(Brock, Huang et al. 2001) . The HMBD sequence is based on a previous study that defined the minimal protein fragment required to recognize doulble stranded methylated-CpG. The poly-His tag does not interfere with binding. An overnight culture was diluted 100 times and incubated at 37 °C to *A*600 = 0.3, before induction with isopropyl -thiogalactoside. Cells were grown in defined medium enriched with 4.g.l.-1 glucose and 5g.l.-1 (NH4)2SO4. Harvested cells were resuspended in buffer A (50 mM sodium phosphate (pH 7.0), 50 mM NaCl), combined with five volumes of Buffer MA plus 5 M urea (due to the protein being expressed in an insoluble form), sonicated and pelleted. The supernatant was loaded onto a sepharose cation-exchange column and was washed sequentially with Buffer MA plus 5 M urea, Buffer MA plus 0.25 M NaCl and 3.75 M urea, and Buffer MA plus 0.4 M NaCl and 3 M urea. The protein was eluted in Buffer MA plus 1 M NaCl. The material was purified further on a nickel-agarose column. The purity of the protein was ~98 % (as judged by SDS-polyacrylamide gel electrophoresis) and the yield was 10 - 40 mg l−1 of culture. Prior to using the HMBD column it was washed in 100mls of BufferA (All operations were performed at 4 °C.

6.11.2 *DIG Labelling of COT-1DNA*

COT-1 DNA was purchased from Roche, as was the High Prime DIG labelling Kit. COT-1 DNA was used as a template for the DIG labelling reaction. Labelling was carried out as per the manufacturer's instruction.

In brief 10ug of COT-1 DNA was used as a template, to which one vial with 20 µl 5x conc. labeling mixture containing: random hexamer primer mixture, 1 unit/µl Klenow enzyme, labeling grade, 1 mM dATP, dCTP, dGTP, each, 0.65 M dTTP, 0.35 mM DIG-11-dUTP, alkali-labile 5x reaction buffer with 50% glycerol (v/v) was added plus ddH20 to a final volume of 100ul. The reaction was incubated overnight at 36oC and stored at -20oC as per the manufacturer's instructions.

6.11.3 Subtractive Hybridisation

Linker-adaptors (10 mM) LA2.1 5'-TAGTTAACGCGCTGCATGAGTA-3' and LA2.11 5'-TACTCATGCAGCGCGTTAAC)-3' were added to the driver DNA (~150 ng) along with T4 ligase (400 U) and 10x ligase buffer (NEB) in a total volume of 20 µl, then incubated overnight at 4°C. This was followed by incubation for 30 min at 37°C with 1 U of T4 DNA polymerase (NEB) and dNTPs (100 µM). Driver and tester, mixed in a ratio of 3:1-10:1, were cleaned using a Qiagen PCR Purification Kit column (Qiagen Ltd, UK) and resuspended in 20 µl of distilled H20 with 50 mM HEPES (pH 7.5), 100 mM NaCl, 50 mM EDTA (pH 8) and 0.1% w/v SDS, total volume 50 µl, based on a method described previously . Following a 5 min incubation at 95°C the solution was held at 68°C for 48 h, placed immediately on ice for 10 min and cleaned using a Qiagen PCR clean-up column. After resuspending in 20 µl of Tris-HCl pH 8.0, linker adaptors (20 mM) LA2.4 5'-TACTTCTTGCGCCAAGACGTT-3' and LA2.44 5'-AACGTCTTGGCGCAAGAAG-3', T4 ligase (800 U) and 10x ligase buffer (NEB) were added, total volume 30 µl, followed by overnight incubation at 4°C. A Qiagen column was used to clean the mixture, and an aliquot (1 µl) used in a PCR reaction with linker LA2.44 as primer.

6.12 Statistical Methods

Methylation at each single locus was correlated with the clinicopathological parameters listed in Table 4-6. Fisher's exact test was used to calculate significance of methylation at a locus with respect to ER/PR positivity and lymphovascular invasion. The chi-squared test for trend was used to calculate significance for NPI, nodal status, grade, tumour and age. Where the continuous variables; age, tumour size and NPI were concerned a Mann Whitney test was also performed to test for significance. The association between methylation at different loci was assessed using Fisher's exact test. Methylation of more than one locus and relationship to clinicopathological variables was assessed using the chi-squared test for trend.

7 Appendices

7.1 Ethics Approval

7.2 Tumour Database

Position in Table	Tumour Database ID.	Result of 1st DNA Extraction	Result of 2nd DNA Extraction	Result of 3rd DNA Extraction	HS3ST1M	HS3ST2M	HS3ST3AM	HS3ST3BM	HS3ST3BE1M	HS3ST4M		SIZE		GRADE		NODES +VE		NODES FOUND		ER,PR		AGE AT DIAGNOSIS
1	CDTDSS2	(+)			0.00	1.00	0.00	0.00	0.00	0.00		13.00		2.00		0.00		14.00		1.00		54.00
2	CDTDRP4	(+)			0.00	1.00	0.00	0.00	0.00	0.00		80.00		3.00		4.00		14.00	X			35.00
3	CDTDH19	(+)			0.00	1.00	0.00	0.00	1.00	0.00		29.00		2.00		13.00		18.00		1.00		48.00
4	CDTDB03	(+)			0.00	0.00	0.00	0.00	0.00	1.00		30.00		2.00		2.00		10.00		1.00		89.00
5	CDTDP30	(+)			0.00	0.00	0.00	0.00	1.00	1.00		25.00		2.00		0.00		15.00		1.00		75.00
6	CDTDGG1	(+)			0.00	0.00	0.00	0.00	0.00	0.00		10.00		2.00		1.00		15.00		1.00		60.00
7	CDTDRS40	(+)			0.00	1.00	0.00	1.00	0.00	0.00		22.00		2.00		0.00		12.00		1.00		67.00
8	CDTDA64	(+)			0.00	1.00	0.00	0.00	0.00	0.00		40.00		3.00		4.00		21.00		1.00		48.00
9	CDTDB66	(+)			0.00	0.00	0.00	0.00	0.00	1.00	X		X		X		X		X		X	
10	CDTD101	(+)			0.00	1.00	0.00	0.00	0.00	0.00		30.00		2.00		0.00		10.00		1.00		88.00
11	CDTDVIP	(-)	(+)		0.00	1.00	0.00	0.00	0.00	0.00		30.00		3.00		0.00		27.00		0.00		51.00
12	CDTDV11	(+)			0.00	0.00	0.00	0.00	0.00	0.00	X		X		X		X		X		X	
13	CDTDFET	(+)			0.00	1.00	0.00	0.00	0.00	0.00		35.00		3.00		0.00		16.00		1.00		62.00
14	CDTDP07	(+)			0.00	0.00	0.00	1.00	0.00	1.00		24.00		3.00		0.00		24.00		0.00		36.00
15	CDTDW90	(+)			0.00	1.00	0.00	1.00	0.00	1.00	X		X		X		X		X		X	
16	CDTDF22	(+)			0.00	1.00	0.00	0.00	1.00	0.00		35.00		3.00		1.00		11.00		1.00		61.00
17	CDTDA37	(+)			0.00	1.00	0.00	0.00	0.00	0.00		28.00		3.00		3.00		11.00		0.00		56.00
18	CDTDX10	(+)			0.00	1.00	0.00	0.00	0.00	1.00		26.00		3.00		9.00		14.00		1.00		61.00
19	CDTDD13	(+)			0.00	1.00	0.00	0.00	0.00	0.00		46.00		3.00		2.00		12.00		1.00		33.00
20	CDTDA45	(+)			0.00	1.00	0.00	0.00	0.00	0.00		35.00		3.00		0.00		15.00	X			57.00
21	CDTDFS2	(+)			0.00	1.00	0.00	0.00	1.00	1.00		27.00		2.00	X			X		1.00		68.00
22	CDTDHG7	(+)			0.00	1.00	0.00	0.00	1.00	1.00		34.00		3.00		3.00		11.00		1.00		72.00
23	CDTDQ11	(+)			0.00	0.00	0.00	0.00	0.00	0.00		25.00		2.00		2.00		11.00		1.00		83.00
24	CDTDLH8	(+)			0.00	1.00	0.00	0.00	1.00	1.00		30.00		2.00		4.00		11.00		1.00		68.00
25	CDTDL09	(+)			0.00	1.00	0.00	1.00	0.00	1.00		17.00		2.00		0.00		4.00	X			72.00
26	CDTDBS1	(+)			0.00	1.00	0.00	0.00	0.00	1.00		47.00		3.00		3.00		13.00		0.00		62.00
27	CDTDE10	(+)			0.00	0.00	0.00	0.00	0.00	0.00		40.00		3.00		0.00		1.00		1.00		59.00
28	CDTDHG5	(+)			0.00	1.00	0.00	0.00	1.00	1.00		95.00		3.00		0.00		26.00	X			54.00
29	CDTDAA1	(+)			0.00	1.00	0.00	1.00	1.00	1.00		36.00		3.00		9.00		13.00		0.00		72.00
30	CDTDFR2	(+)			0.00	1.00	0.00	0.00	1.00	1.00		26.00		3.00		0.00		11.00		1.00		65.00
31	CDTDRF3	(+)			0.00	0.00	0.00	1.00	0.00	0.00		12.00		2.00		5.00		14.00		1.00		37.00
32	CDTDBB8	(+)			0.00	1.00	0.00	1.00	0.00	0.00		20.00		3.00		1.00		21.00		0.00		57.00
33	CDTDJKL	(+)			0.00	1.00	0.00	0.00	0.00	0.00		11.00		2.00		1.00		11.00		0.00		72.00
34	CDTDNQ2	(+)			0.00	1.00	0.00	1.00	1.00	0.00		25.00		3.00		0.00		19.00		0.00		70.00
35	CDTDQWE	(-)	(+)		0.00	1.00	0.00	0.00	0.00	0.00		22.00		1.00		0.00		17.00		1.00		74.00
36	CDTDHD1	(+)			0.00	0.00	0.00	0.00	0.00	1.00		15.00		1.00		7.00		13.00		1.00		78.00

37	CDTDLOQ	(+)		0.00	1.00	0.00	0.00	0.00	0.00		25.00	3.00	0.00	21.00		0.00	76.00
38	CDTDJD1	(+)		0.00	0.00	0.00	1.00	0.00	1.00		24.00	3.00	3.00	14.00		1.00	66.00
39	CDTDFRD	(+)		0.00	1.00	0.00	0.00	1.00	0.00		50.00	2.00	3.00	10.00		1.00	68.00
40	CDTD6MH	(+)		0.00	1.00	0.00	1.00	0.00	0.00		80.00	3.00	2.00	16.00		1.00	60.00
41	CDTDB04	(+)		0.00	1.00	0.00	1.00	0.00	0.00		47.00	3.00	5.00	23.00		1.00	71.00
42	CDTDG41	(+)		0.00	1.00	0.00	0.00	0.00	1.00		52.00	2.00	0.00	14.00		1.00	80.00
43	CDTD4FT	(+)		0.00	0.00	0.00	1.00	0.00	1.00		27.00	1.00	0.00	20.00		1.00	40.00
44	CDTDBB7	(-)	(+)	0.00	0.00	0.00	0.00	0.00	0.00		23.00	2.00	4.00	10.00		1.00	76.00
45	CDTDL99	(+)		0.00	0.00	0.00	1.00	0.00	1.00		17.00	2.00	1.00	18.00		1.00	76.00
46	CDTDWB6	(+)		0.00	1.00	0.00	1.00	0.00	1.00		38.00	3.00	11.00	15.00		0.00	46.00
47	CDTDNL2	(+)		0.00	0.00	0.00	1.00	0.00	1.00		30.00	3.00	4.00	4.00		0.00	67.00
48	CDTDLFT	(+)		0.00	0.00	0.00	0.00	0.00	0.00		140.00	3.00	X		X		0.00	77.00
49	CDTDFF5	(+)		0.00	0.00	0.00	0.00	0.00	0.00		25.00	1.00	3.00	17.00		1.00	81.00
50	CDTDBD0	(+)		0.00	0.00	0.00	0.00	0.00	0.00		16.00	2.00	0.00	5.00		1.00	64.00
51	CDTDBBC	(+)		0.00	0.00	0.00	0.00	0.00	0.00		45.00	2.00	3.00	25.00		1.00	39.00
52	CDTDABO	(+)		0.00	1.00	0.00	0.00	0.00	0.00		38.00	1.00	5.00	17.00		0.00	43.00
53	CDTDXP6	(+)		0.00	0.00	0.00	1.00	0.00	0.00		18.00	2.00	4.00	16.00		1.00	53.00
54	CDTD123	(-)	(+)	0.00	1.00	0.00	1.00	0.00	0.00	X		X	X	X	X		X	
55	CDTDRZ3	(+)		0.00	1.00	0.00	1.00	0.00	1.00		18.00	1.00	0.00	7.00		1.00	62.00
56	CDTDSH3	(+)		0.00	1.00	0.00	1.00	0.00	1.00		20.00	3.00	13.00	15.00		0.00	58.00
57	CDTDNP1	(+)		0.00	0.00	0.00	0.00	0.00	1.00		40.00	1.00	0.00	0.00		0.00	33.00
58	CDTD100	(-)	(-) (+)	0.00	1.00	0.00	0.00	0.00	0.00		13.00	1.00	1.00	11.00		1.00	70.00
59	CDTDNN5	(+)		0.00	1.00	0.00	0.00	1.00	1.00		70.00	3.00	4.00	0.00		0.00	73.00
60	CDTDBMX	(+)		0.00	1.00	0.00	1.00	0.00	0.00		38.00	3.00	1.00	1.00		0.00	40.00
61	CDTDC9H	(+)		0.00	1.00	0.00	1.00	0.00	1.00		19.00	3.00	2.00	12.00		0.00	46.00
62	CDTDA3B	(+)		0.00	0.00	0.00	1.00	0.00	1.00	X		1.00	X	X	X		57.00	
63	CDTDH1A	(+)		0.00	1.00	0.00	1.00	0.00	0.00	X		X	X	X	X		54.00	
64	CDTD00Y	(+)		0.00	0.00	0.00	0.00	1.00	0.00	X		X	X	X	X		72.00	
65	CDTD24T	(+)		0.00	0.00	0.00	0.00	0.00	1.00	X		X	X	X	X		44.00	
66	CDTDG11	(+)		0.00	1.00	0.00	1.00	0.00	1.00		18.00	2.00	1.00	9.00		100.00	67.00	
67	CDTDCUP	(+)		0.00	0.00	0.00	0.00	0.00	0.00		15.00	2.00	6.00	16.00		100.00	72.00	
68	CDTDB4I	(+)		0.00	0.00	0.00	0.00	0.00	0.00	X		X	X	X	X		59.00	
69	CDTDTT2	(+)		0.00	0.00	0.00	1.00	0.00	0.00	X		X	X	X	X		73.00	
70	CDTD2BS	(+)		0.00	0.00	0.00	1.00	0.00	0.00	X		X	X	X	X		X	
71	CDTDGD1	(+)		0.00	0.00	0.00	0.00	0.00	1.00		30.00	2.00	1.00	4.00		90.00	55.00	
72	CDTDF07	(+)		0.00	0.00	0.00	1.00	0.00	0.00		35.00	3.00	7.00	7.00		50.00	61.00	
73	CDTD1XY	(+)		0.00	0.00	0.00	1.00	0.00	0.00		85.00	2.00	X	X	X		61.00	
74	CDTDER5	(+)		0.00	0.00	0.00	1.00	0.00	1.00	X		X		X	X		43.00	
75	CDTD42Y	(+)		0.00	1.00	0.00	0.00	1.00	0.00	X		X		X	X		52.00	
76	CDTDXX1	(-)	(-) (+)	0.00	1.00	0.00	0.00	0.00	1.00		30.00	3.00	5.00	5.00		0.00	39.00	
77	CDTDNX1	(-)	(+)	0.00	0.00	0.00	1.00	0.00	1.00		30.00	3.00	X	X		0.00	48.00	
78	CDTDPLY	(+)		0.00	1.00	0.00	0.00	1.00	1.00	X		X	X	X	X		X	
79	CDTDFTA	(-)	(+)	0.00	0.00	0.00	0.00	0.00	0.00	X		X	X	X	X		X	
80	CDTDBIG	(+)		0.00	0.00	0.00	0.00	0.00	0.00	X		X	X	X	X		X	
81	CDTDBXX	(-)	(-) (-)	X	X	X	X	X	X	X		X	X	X	X		X	
82	CDTD4Y4	(-)	(-) (-)	X	X	X	X	X	X	X		X	X	X	X		X	
83	CDTD2X2	(-)	(-) (-)	X	X	X	X	X	X	X		X	X	X	X		X	
84	CDTD7A6	(-)	(-) (-)	X	X	X	X	X	X	X		X	X	X	X		X	

85	CDTD6JQ	(-)	(-)	(-)	X	X	X	X	X	X	X	X	X	X	X	X
86	CDTD0J6	(-)	(-)	(-)	X	X	X	X	X	X	X	X	X	X	X	X
87	CDTDCID	(-)	(-)	(-)	X	X	X	X	X	X	X	X	X	X	X	X
88	CDTD1DC	(-)	(-)	(-)	X	X	X	X	X	X	X	X	X	X	X	X
89	CDTDD2C	(-)	(-)	(-)	X	X	X	X	X	X	X	X	X	X	X	X
90	CDTD333	(-)	(-)	(-)	X	X	X	X	X	X	X	X	X	X	X	X
					0.00	45.00	0.00	30.00	15.00	34.00						

1 Positives	4.00	3.00	0.00	5.00
2 Positives	11.00	7.00	5.00	9.00
3 Positives	25.00	14.00	7.00	15.00

1 Negatives	4.00
2 Negatives	30.00
3 Negatives	51.00

VASCINV	LYMPHINV	STAGE	NPI = G+S+(D. CM/5)	Site of Operation	Year of Operation	Source (TB=Tumour Bank, F=Fresh)
0.00	0.00	1	3.26	GRI	2002	F
1.00	1.00	3	7.60	GRI	2003	F
0.00	1.00	3	5.58	WIG	2003	F
0.00	0.00	2	4.60	WIG	2002	F
1.00	1.00	1	3.50	WIG	2003	F
0.00	0.00	2	4.20	WIG	2003	F
X	X	1	3.44	WIG	2002	F
1.00	1.00	3	6.80	WIG	2002	F
X	X	X	X	GRI	1989	TB
0.00	0.00	1	3.60	WIG	2003	F
0.00	0.00	1	4.60	GRI	1999	TB
X	X	X	X	GRI	1995	TB
0.00	0.00	1	4.70	WIG	2003	F
0.00	0.00	1	4.48	WIG	2003	F
X	X	X	X	GRI	1999	TB
0.00	0.00	2	5.70	GRI	1989	TB
0.00	0.00	2	5.56	GRI	1992	TB
1.00	1.00	3	6.52	GRI	1996	TB
0.00	0.00	2	5.92	GRI	1999	TB
X	X	1	4.70	GRI	1997	TB
0.00	0.00	X	X	WIG	2003	F
1.00	1.00	2	5.68	GRI	1998	TB
0.00	1.00	2	4.50	GRI	1998	TB
1.00	1.00	3	5.60	GRI	1998	TB
X	X	1	3.34	GRI	1999	TB
X	X	2	5.94	GRI	1999	TB
0.00	0.00	1	4.80	GRI	1999	TB
0.00	0.00	1	5.90	GRI	1998	TB
1.00	1.00	3	6.72	GRI	1998	TB
0.00	0.00	1	4.52	GRI	1998	TB
0.00	1.00	3	5.24	GRI	1999	TB
1.00	1.00	2	5.40	GRI	1999	TB
0.00	0.00	2	4.22	GRI	1999	TB
0.00	0.00	1	4.50	GRI	1999	TB
0.00	0.00	1	2.44	WIG	2003	F
0.00	1.00	3	4.30	WIG	2003	F

0.00	0.00	1	4.50	WIG	2003	F
0.00	1.00	2	5.48	WIG	2003	F
0.00	0.00	2	5.00	WIG	2003	F
0.00	0.00	2	6.60	GRI	1998	TB
0.00	1.00	3	6.94	GRI	1999	TB
0.00	0.00	1	4.04	GRI	1999	TB
0.00	0.00	1	2.54	GRI	1998	TB
0.00	0.00	3	5.46	WIG	2002	F
0.00	0.00	2	4.34	WIG	2002	F
1.00	1.00	3	6.76	WIG	2002	F
1.00	1.00	3	6.60	WIG	2003	F
0.00	0.00	X	X	WIG	2003	F
0.00	1.00	2	3.50	WIG	2003	F
0.00	0.00	1	3.32	WIG	2002	F
0.00	0.00	2	4.90	WIG	2002	F
1.00	1.00	3	4.76	WIG	2002	F
1.00	1.00	3	5.36	WIG	2002	F
X	X	X	X	GRI	1993	TB
0.00	0.00	1	2.36	WIG	2003	F
0.00	1.00	3	6.40	WIG	2003	F
0.00	0.00	X	X	WIG	2002	F
0.00	0.00	2	3.26	GRI	2002	F
0.00	1.00	3	7.40	WIG	2002	F
0.00	0.00	2	5.76	WIG	2002	F
1.00	1.00	2	5.38	WIG	2002	F
X	X	X	X	GRI	1998	TB
X	X	X	X	GRI	1998	TB
X	X	X	X	GRI	1998	TB
X	X	X	X	GRI	1998	TB
0.00	1.00	2	4.36	WIG	2002	F
1.00	1.00	3	5.30	GRI	2003	F
X	X	X	X	GRI	2000	TB
X	X	X	X	GRI	2002	TB
X	X	X	X	GRI	2000	TB
0.00	0.00	2	4.60	WIG	2002	F
1.00	1.00	3	6.70	WIG	2002	F
X	X	X	X	GRI	1999	TB
X	X	X	X	GRI	1999	TB
X	X	X	X	GRI	1999	TB
0.00	0.00	3	6.60	GRI	2000	TB
0.00	1.00	X	X	WIG	2002	F
X	X	X	X	GRI	1989	TB
X	X	X	X	GRI	1989	TB
X	X	X	X	GRI	1988	TB
X	X	X	X	GRI	1984	TB
X	X	X	X	GRI	1989	TB
X	X	X	X	GRI	1988	TB
X	X	X	X	GRI	1988	TB

X	X	X	X	GRI	1987	TB
X	X	X	X	GRI	1988	TB
X	X	X	X	GRI	1989	TB
X	X	X	X	GRI	1982	TB
X	X	X	X	GRI	1984	TB
X	X	X	X	GRI	1986	TB

GRI	55		
WIG	35		
		F	39
		TB	51

Tumour Database ID.	Reason for missing data
CDTDSS2	
CDTDRP4	
CDTDH19	
CDTDB03	
CDTDP30	
CDTDGG1	
CDTDRS40	
CDTDA64	
CDTDB66	
CDTD101	
CDTDVIP	
CDTDV11	
CDTDFET	
CDTDP07	
CDTDW90	
CDTDF22	
CDTDA37	
CDTDX10	
CDTDD13	
CDTDHG7	
CDTDQ11	
CDTDLH8	
CDTDL09	Tumour could not be linked back to a patient.
CDTDBS1	Tumour could not be linked back to a patient.
CDTDE10	
CDTDHG5	ER and PR immuno not reported.
CDTDAA1	
CDTDFR2	
CDTDRF3	
CDTDBB8	
CDTDJKL	
CDTDNQ2	
CDTDQWE	
CDTDHD1	

	CDTDLOQ
	CDTDJD1
	CDTDFRD
	CDTD6MH
	CDTDB04
	CDTDG41
	CDTD4FT
	CDTDBB7
	CDTDL99
	CDTDWB6
	CDTDNL2
Axillary clearance not carried out	CDTDLFT
	CDTDFF5
	CDTDBD0
	CDTDBBC
	CDTDABO
	CDTDXP6
No clinical details were entered in the tumour database.	CDTD123
	CDTDRZ3
	CDTDSH3
	CDTDNP1
	CDTD100
	CDTDNN5
	CDTDBMX
	CDTDC9H
No clinical details were entered in the tumour database.	CDTDA3B
No clinical details were entered in the tumour database.	CDTDH1A
No clinical details were entered in the tumour database.	CDTD00Y
No clinical details were entered in the tumour database.	CDTD24T
	CDTDG11
	CDTDCUP
Age only entered in the tumour database	CDTDB4I
	CDTDTT2
No clinical details were entered in the tumour database.	CDTD2BS
	CDTDGD1
	CDTDF07
Grade, size and age only entered in database.	CDTD1XY
Age only entered in the tumour database	CDTDER5
Age only entered in the tumour database	CDTD42Y
	CDTDXX1
Axillary clearance not carried out	CDTDNX1
No clinical details were entered in the tumour database.	CDTDPLY
No clinical details were entered in the tumour database.	CDTDFTA
No clinical details were entered in the tumour database.	CDTDBIG
DNA degraded during storeage.	CDTDBXX
DNA degraded during storeage.	CDTD4Y4
DNA degraded during storeage.	CDTD2X2
DNA degraded during storeage.	CDTD7A6

DNA degraded during storeage.	CDTD6JQ
DNA degraded during storeage.	CDTD0J6
DNA degraded during storeage.	CDTDCID
DNA degraded during storeage.	CDTD1DC
DNA degraded during storeage.	CDTDD2C
DNA degraded during storeage.	CDTD333

8 Bibliography

Abrahamson, P. E., S. S. Tworoger, et al. (2006). "Associations between the CYP17, CYPIB1, COMT and SHBG polymorphisms and serum sex hormones in post-menopausal breast cancer survivors." Breast Cancer Res Treat.

Antequera, F. and A. Bird (1993). "Number of CpG islands and genes in human and mouse." Proc Natl Acad Sci U S A **90**(24): 11995-9.

Antequera, F., D. Macleod, et al. (1989). "Specific protection of methylated CpGs in mammalian nuclei." Cell **58**(3): 509-17.

Armstrong, K., A. Eisen, et al. (2000). "Assessing the risk of breast cancer." N Engl J Med **342**(8): 564-71.

Aubele, M., M. Werner, et al. (2002). "Genetic alterations in presumptive precursor lesions of breast carcinomas." Anal Cell Pathol **24**(2-3): 69-76.

Bader, S., M. Walker, et al. (1999). "Somatic frameshift mutations in the MBD4 gene of sporadic colon cancers with mismatch repair deficiency." Oncogene **18**(56): 8044-7.

Baiqiu, W., F. Songbin, et al. (2000). "Study of the relationship between psoriasis and the polymorphic site C677T of methylenetetrahydrofolate reductase." Chin Med Sci J **15**(2): 119-20.

Barber, T. M., M. I. McCarthy, et al. (2006). "Obesity and polycystic ovary syndrome." Clin Endocrinol (Oxf) **65**(2): 137-45.

Battersby, S. and T. J. Anderson (1989). "Histological changes in breast tissue that characterize recent pregnancy." Histopathology **15**(4): 415-9.

Baum, M. (1998). "Polychemotherapy for early breast cancer." Lancet **352**(9139): 1554.

Bergman-Jungestrom, M. and S. Wingren (2001). "Catechol-O-Methyltransferase (COMT) gene polymorphism and breast cancer risk in young women." Br J Cancer **85**(6): 859-62.

Bernardino, J., C. Roux, et al. (1997). "DNA hypomethylation in breast cancer: an independent parameter of tumor progression?" Cancer Genet Cytogenet **97**(2): 83-9.

Berry, D., C. P. Kwan, et al. (2001). "Distinct heparan sulfate glycosaminoglycans are responsible for mediating fibroblast growth factor-2 biological activity through different fibroblast growth factor receptors." Faseb J **15**(8): 1422-4.

Bird, A. (2002). "DNA methylation patterns and epigenetic memory." Genes Dev **16**(1): 6-21.

Bird, A., M. Taggart, et al. (1985). "A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA." Cell **40**(1): 91-9.

Bird, A. P. (1980). "DNA methylation and the frequency of CpG in animal DNA." Nucleic Acids Res **8**(7): 1499-504.

Bird, A. P. (1986). "CpG-rich islands and the function of DNA methylation." Nature **321**(6067): 209-13.

Bird, A. P. and A. P. Wolffe (1999). "Methylation-induced repression--belts, braces, and chromatin." Cell **99**(5): 451-4.

Bishop, D. T. (1999). "BRCA1 and BRCA2 and breast cancer incidence: a review." Ann Oncol **10 Suppl 6**: 113-9.

Blackhall, F. H., C. L. Merry, et al. (2003). "Binding of endostatin to endothelial heparan sulphate shows a differential requirement for specific sulphates." Biochem J **375**(Pt 1): 131-9.

Blamey, R. W., A. R. Wilson, et al. (2000). "ABC of breast diseases: screening for breast cancer." Bmj **321**(7262): 689-93.

Borsig, L. (2003). "Non-anticoagulant effects of heparin in carcinoma metastasis and Trousseau's syndrome." Pathophysiol Haemost Thromb **33 Suppl 1**: 64-6.

Borsig, L. (2004). "Selectins facilitate carcinoma metastasis and heparin can prevent them." News Physiol Sci **19**: 16-21.

Borsig, L., R. Wong, et al. (2001). "Heparin and cancer revisited: mechanistic connections involving platelets, P-selectin, carcinoma mucins, and tumor metastasis." Proc Natl Acad Sci U S A **98**(6): 3352-7.

Brock, G. J., J. Charlton, et al. (1999). "Densely methylated sequences that are preferentially localized at telomere-proximal regions of human chromosomes." Gene **240**(2): 269-77.

Brock, G. J., T. H. Huang, et al. (2001). "A novel technique for the identification of CpG islands exhibiting altered methylation patterns (ICEAMP)." Nucleic Acids Res **29**(24): E123.

Buerger, H., R. Simon, et al. (2000). "Genetic relation of lobular carcinoma in situ, ductal carcinoma in situ, and associated invasive carcinoma of the breast." Mol Pathol **53**(3): 118-21.

Cameron, E. E., K. E. Bachman, et al. (1999). "Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer." Nat Genet **21**(1): 103-7.

Campbell, I. G., S. W. Baxter, et al. (2002). "Methylenetetrahydrofolate reductase polymorphism and susceptibility to breast cancer." Breast Cancer Res **4**(6): R14.

Camphausen, R. T., R. L. Jones, et al. (1986). "Structure and relevance of the oligosaccharide hapten of Mycobacterium avium serotype 2." J Bacteriol **168**(2): 660-7.

Catteau, A., W. H. Harris, et al. (1999). "Methylation of the BRCA1 promoter region in sporadic breast and ovarian cancer: correlation with disease characteristics." Oncogene **18**(11): 1957-65.

Chai, J. H., D. P. Locke, et al. (2001). "Retrotransposed genes such as Frat3 in the mouse Chromosome 7C Prader-Willi syndrome region acquire the imprinted status of their insertion site." Mamm Genome **12**(11): 813-21.

Chamras, H., D. Bagga, et al. (1998). "Preadipocytes stimulate breast cancer cell growth." Nutr Cancer **32**(2): 59-63.

Chen, C. M., H. L. Chen, et al. (2003). "Methylation target array for rapid analysis of CpG island hypermethylation in multiple tissue genomes." Am J Pathol **163**(1): 37-45.

Chen, D., B. Adenekan, et al. (2004). "Syndecan-1 expression in locally invasive and metastatic prostate cancer." Urology **63**(2): 402-7.

Chen, J., M. B. Duncan, et al. (2003). "Biosynthesis of 3-O-sulfated heparan sulfate: unique substrate specificity of heparan sulfate 3-O-sulfotransferase isoform 5." Glycobiology **13**(11): 785-94.

Cheng, T. C., S. T. Chen, et al. (2005). "Breast cancer risk associated with genotype polymorphism of the catechol estrogen-metabolizing genes: a multigenic study on cancer susceptibility." Int J Cancer **113**(3): 345-53.

Clark, S. J., J. Harrison, et al. (1997). "Sp1 binding is inhibited by (m)Cp(m)CpG methylation." Gene **195**(1): 67-71.

Clark, S. J., J. Harrison, et al. (1994). "High sensitivity mapping of methylated cytosines." Nucleic Acids Res **22**(15): 2990-7.

Cole, B. F., R. D. Gelber, et al. (2001). "Polychemotherapy for early breast cancer: an overview of the randomised clinical trials with quality-adjusted survival analysis." Lancet **358**(9278): 277-86.

Costello, J. F., M. C. Fruhwald, et al. (2000). "Aberrant CpG-island methylation has non-random and tumour-type-specific patterns." Nat Genet **24**(2): 132-8.

Cotlar, A. M., J. J. Dubose, et al. (2003). "History of surgery for breast cancer: radical to the sublime." Curr Surg **60**(3): 329-37.

Cross, S. H., J. A. Charlton, et al. (1994). "Purification of CpG islands using a methylated DNA binding column." Nat Genet **6**(3): 236-44.

CRUK (2005). Cancer Incidence and Mortality Statistics.

D'Incalci, M., J. M. Covey, et al. (1985). "DNA alkali-labile sites induced by incorporation of 5-aza-2'-deoxycytidine into DNA of mouse leukemia L1210 cells." Cancer Res **45**(7): 3197-202.

Danishefsky, I., H. Steiner, et al. (1969). "Investigations on the chemistry of heparin. VI. Position of the sulfate ester groups." J Biol Chem **244**(7): 1741-5.

Daskalakis, M., T. T. Nguyen, et al. (2002). "Demethylation of a hypermethylated P15/INK4B gene in patients with myelodysplastic syndrome by 5-Aza-2'-deoxycytidine (decitabine) treatment." Blood **100**(8): 2957-64.

David, G. (1991). "Biology and pathology of the pericellular heparan sulphate proteoglycans." Biochem Soc Trans **19**(4): 816-20.

DiGabriele, A. D., I. Lax, et al. (1998). "Structure of a heparin-linked biologically active dimer of fibroblast growth factor." Nature **393**(6687): 812-7.

Dizdar, O. and E. Alyamac (2004). "Obesity: an endocrine tumor?" Med Hypotheses **63**(5): 790-2.

Donegan, W. L. History of Breast Cancer.

Dong, S. M., J. C. Pang, et al. (2001). "Concurrent hypermethylation of multiple genes is associated with grade of oligodendroglial tumors." J Neuropathol Exp Neurol **60**(8): 808-16.

Eden, A., F. Gaudet, et al. (2003). "Chromosomal instability and tumors promoted by DNA hypomethylation." Science **300**(5618): 455.

Ehrlich, M. (2002). "DNA methylation in cancer: too much, but also too little." Oncogene **21**(35): 5400-13.

Ehrlich, M., G. Jiang, et al. (2002). "Hypomethylation and hypermethylation of DNA in Wilms tumors." Oncogene **21**(43): 6694-702.

El-Monem, D. H. A. (2003). D. J. J. Going.

Elmore, J. G., K. Armstrong, et al. (2005). "Screening for breast cancer." Jama **293**(10): 1245-56.

Elston, C. W. and I. O. Ellis (1991). "Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up." Histopathology **19**(5): 403-10.

Esko, J. D., K. S. Rostand, et al. (1988). "Tumor formation dependent on proteoglycan biosynthesis." Science **241**(4869): 1092-6.

Espada, J., E. Ballestar, et al. (2004). "Human DNA methyltransferase 1 is required for maintenance of the histone H3 modification pattern." J Biol Chem **279**(35): 37175-84.

Esteller, M. (2003). "Cancer epigenetics: DNA methylation and chromatin alterations in human cancer." Adv Exp Med Biol **532**: 39-49.

Esteller, M. (2005). "Aberrant DNA methylation as a cancer-inducing mechanism." Annu Rev Pharmacol Toxicol **45**: 629-56.

Esteller, M. (2005). "DNA methylation and cancer therapy: new developments and expectations." Curr Opin Oncol **17**(1): 55-60.

Esteller, M. (2005). "Dormant hypermethylated tumour suppressor genes: questions and answers." J Pathol **205**(2): 172-80.

Esteller, M., L. Catasus, et al. (1999). "hMLH1 promoter hypermethylation is an early event in human endometrial tumorigenesis." Am J Pathol **155**(5): 1767-72.

Esteller, M., M. F. Fraga, et al. (2001). "DNA methylation patterns in hereditary human cancers mimic sporadic tumorigenesis." Hum Mol Genet **10**(26): 3001-7.

Esteller, M., S. Gonzalez, et al. (2001). "K-ras and p16 aberrations confer poor prognosis in human colorectal cancer." J Clin Oncol **19**(2): 299-304.

Esteller, M., S. R. Hamilton, et al. (1999). "Inactivation of the DNA repair gene O6-methylguanine-DNA methyltransferase by promoter hypermethylation is a common event in primary human neoplasia." Cancer Res **59**(4): 793-7.

Fahrner, J. A., S. Eguchi, et al. (2002). "Dependence of histone modifications and gene expression on DNA hypermethylation in cancer." Cancer Res **62**(24): 7213-8.

Feigelson, H. S., G. A. Coetzee, et al. (1997). "A polymorphism in the CYP17 gene increases the risk of breast cancer." Cancer Res **57**(6): 1063-5.

Feinberg, A. P. and B. Vogelstein (1983). "Hypomethylation distinguishes genes of some human cancers from their normal counterparts." Nature **301**(5895): 89-92.

Fleisher, A. S., M. Esteller, et al. (1999). "Hypermethylation of the hMLH1 gene promoter in human gastric cancers with microsatellite instability." Cancer Res **59**(5): 1090-5.

Fleissig, A., L. J. Fallowfield, et al. (2005). "Post-operative arm morbidity and quality of life. Results of the ALMANAC randomised trial comparing sentinel node biopsy with standard axillary treatment in the management of patients with early breast cancer." Breast Cancer Res Treat: 1-15.

Fryback, D. G., N. K. Stout, et al. (2006). "The Wisconsin Breast Cancer Epidemiology Simulation Model." J Natl Cancer Inst Monogr(36): 37-47.

Fuks, F., P. J. Hurd, et al. (2003). "The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation." J Biol Chem **278**(6): 4035-40.

Gallagher, J. T. (1995). "Heparan sulphate and protein recognition. Binding specificities and activation mechanisms." Adv Exp Med Biol **376**: 125-34.

Gallardo, F., M. Esteller, et al. (2004). "Methylation status of the p15, p16 and MGMT promoter genes in primary cutaneous T-cell lymphomas." Haematologica **89**(11): 1401-3.

Galm, O., S. Wilop, et al. (2004). "DNA methylation changes in multiple myeloma." Leukemia **18**(10): 1687-92.

Garfinkel, L., C. C. Boring, et al. (1994). "Changing trends. An overview of breast cancer incidence and mortality." Cancer **74**(1 Suppl): 222-7.

Gilbert, F. J. (2005). "Screening for breast cancer in women at moderate and high risk." Clin Oncol (R Coll Radiol) **17**(4): 240-3.

Goffin, J. and E. Eisenhauer (2002). "DNA methyltransferase inhibitors-state of the art." Ann Oncol **13**(11): 1699-716.

Graff, J. R., J. G. Herman, et al. (1997). "Mapping patterns of CpG island methylation in normal and neoplastic cells implicates both upstream and downstream regions in de novo methylation." J Biol Chem **272**(35): 22322-9.

Grodin, J. M., P. K. Siiteri, et al. (1973). "Source of estrogen production in postmenopausal women." J Clin Endocrinol Metab **36**(2): 207-14.

Haimov-Kochman, R., Y. Lavy, et al. (2002). "[Review of risk factors for breast cancer--what's new?]." Harefuah **141**(8): 702-8, 761.

HajMohammadi, S., K. Enjyoji, et al. (2003). "Normal levels of anticoagulant heparan sulfate are not essential for normal hemostasis." J Clin Invest **111**(7): 989-99.

Hanahan, D. and R. A. Weinberg (2000). "The hallmarks of cancer." Cell **100**(1): 57-70.

Hansen, R. S., C. Wijmenga, et al. (1999). "The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome." Proc Natl Acad Sci U S A **96**(25): 14412-7.

Hartge, P. (2003). "Genes, hormones, and pathways to breast cancer." N Engl J Med **348**(23): 2352-4.

He, L. Z., T. Tolentino, et al. (2001). "Histone deacetylase inhibitors induce remission in transgenic models of therapy-resistant acute promyelocytic leukemia." J Clin Invest **108**(9): 1321-30.

Hendrich, B., C. Abbott, et al. (1999). "Genomic structure and chromosomal mapping of the murine and human Mbd1, Mbd2, Mbd3, and Mbd4 genes." Mamm Genome **10**(9): 906-12.

Herman, J. G. and S. B. Baylin (2003). "Gene silencing in cancer in association with promoter hypermethylation." N Engl J Med **349**(21): 2042-54.

Herman, J. G., C. I. Civin, et al. (1997). "Distinct patterns of inactivation of p15INK4B and p16INK4A characterize the major types of hematological malignancies." Cancer Res **57**(5): 837-41.

Herman, J. G., F. Latif, et al. (1994). "Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma." Proc Natl Acad Sci U S A **91**(21): 9700-4.

Herman, J. G., A. Umar, et al. (1998). "Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma." Proc Natl Acad Sci U S A **95**(12): 6870-5.

HF.LeDran (1757). Memoires avec un précis de plusieurs observations

sur le cancer.

. Memories de l'academie royale de

chirurgie.

Hong, S. H., J. Song, et al. (2001). "Genetic variation of the methylenetetrahydrofolate reductase and cystathionine beta-synthase genes in Korean patients with coronary artery disease and a new polymorphism in intron 7." Mol Cell Probes **15**(2): 119-23.

Hu, M. B., W. Xie, et al. (2006). "[Study on the relationship between polymorphisms of genes (CYP17, CYP19 and SULT1A1) and susceptibility to breast cancer in Chinese women]." Zhonghua Liu Xing Bing Xue Za Zhi **27**(4): 351-5.

Huang, C. S., H. D. Chern, et al. (1999). "Breast cancer risk associated with genotype polymorphism of the estrogen-metabolizing genes CYP17, CYP1A1, and COMT: a multigenic study on cancer susceptibility." Cancer Res **59**(19): 4870-5.

Huang, E., S. H. Cheng, et al. (2003). "Gene expression predictors of breast cancer outcomes." Lancet **361**(9369): 1590-6.

Huang, T. H., M. R. Perry, et al. (1999). "Methylation profiling of CpG islands in human breast cancer cells." Hum Mol Genet **8**(3): 459-70.

Hulka, B. S., E. T. Liu, et al. (1994). "Steroid hormones and risk of breast cancer." Cancer **74**(3 Suppl): 1111-24.

Hultborn, R., C. Hanson, et al. (1997). "Prevalence of Klinefelter's syndrome in male breast cancer patients." Anticancer Res **17**(6D): 4293-7.

Ishihara, M., Y. Guo, et al. (1993). "Regulation of biosynthesis of the basic fibroblast growth factor binding domains of heparan sulfate by heparan sulfate-N-deacetylase/N-sulfotransferase expression." J Biol Chem **268**(27): 20091-5.

J.S.P.Lumley Surface Anatomy: The Anatomical Basis of Clinical Examination

Jaenisch, R. and A. Bird (2003). "Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals." Nat Genet **33 Suppl**: 245-54.

Johanning, G. L., D. C. Heimburger, et al. (2002). "DNA methylation and diet in cancer." J Nutr **132**(12): 3814S-3818S.

Johnston, P. G., C. M. Rondinone, et al. (1992). "Identification of a protein factor secreted by 3T3-L1 preadipocytes inhibitory for the human MCF-7 breast cancer cell line." Cancer Res **52**(24): 6860-5.

Jones, P. A. and P. W. Laird (1999). "Cancer epigenetics comes of age." Nat Genet **21**(2): 163-7.

Jones, P. L. and A. P. Wolffe (1999). "Relationships between chromatin organization and DNA methylation in determining gene expression." Semin Cancer Biol **9**(5): 339-47.

Joshi, K., J. T. Ellis, et al. (1986). "Cellular proliferation in the rat mammary gland during pregnancy and lactation." Lab Invest **54**(1): 52-61.

Juttermann, R., E. Li, et al. (1994). "Toxicity of 5-aza-2'-deoxycytidine to mammalian cells is mediated primarily by covalent trapping of DNA methyltransferase rather than DNA demethylation." Proc Natl Acad Sci U S A **91**(25): 11797-801.

Kamimura, K., J. M. Rhodes, et al. (2004). "Regulation of Notch signaling by Drosophila heparan sulfate 3-O sulfotransferase." J Cell Biol **166**(7): 1069-79.

Kane, M. F., M. Loda, et al. (1997). "Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repair-defective human tumor cell lines." Cancer Res **57**(5): 808-11.

Kang, J. H., S. J. Kim, et al. (2001). "Methylation in the p53 promoter is a supplementary route to breast carcinogenesis: correlation between CpG methylation in the p53 promoter and the mutation of the p53 gene in the progression from ductal carcinoma in situ to invasive ductal carcinoma." Lab Invest **81**(4): 573-9.

Kantarjian, H. M., S. O'Brien, et al. (2003). "Results of decitabine (5-aza-2'deoxycytidine) therapy in 130 patients with chronic myelogenous leukemia." Cancer **98**(3): 522-8.

Kim, Y. I., I. P. Pogribny, et al. (1996). "Exon-specific DNA hypomethylation of the p53 gene of rat colon induced by dimethylhydrazine. Modulation by dietary folate." Am J Pathol **149**(4): 1129-37.

Kinnunen, T., Z. Huang, et al. (2005). "Heparan 2-O-sulfotransferase, hst-2, is essential for normal cell migration in Caenorhabditis elegans." Proc Natl Acad Sci U S A **102**(5): 1507-12.

Kristensen, V. N., T. I. Andersen, et al. (1998). "A rare CYP19 (aromatase) variant may increase the risk of breast cancer." Pharmacogenetics **8**(1): 43-8.

Kristensen, V. N. and A. L. Borresen-Dale (2000). "Molecular epidemiology of breast cancer: genetic variation in steroid hormone metabolism." Mutat Res **462**(2-3): 323-33.

Laird, P. W., L. Jackson-Grusby, et al. (1995). "Suppression of intestinal neoplasia by DNA hypomethylation." Cell **81**(2): 197-205.

Lamy, A., R. Sesboue, et al. (2002). "Aberrant methylation of the CDKN2a/p16INK4a gene promoter region in preinvasive bronchial lesions: a prospective study in high-risk patients without invasive cancer." Int J Cancer **100**(2): 189-93.

Lavigne, J. A., K. J. Helzlsouer, et al. (1997). "An association between the allele coding for a low activity variant of catechol-O-methyltransferase and the risk for breast cancer." Cancer Res **57**(24): 5493-7.

Leu, Y. W., P. S. Yan, et al. (2004). "Loss of estrogen receptor signaling triggers epigenetic silencing of downstream targets in breast cancer." Cancer Res **64**(22): 8184-92.

Levitt, S. H. (1985). "Primary treatment of early breast cancer with conservation surgery and radiation therapy. The effect of adjuvant chemotherapy." Cancer **55**(9 Suppl): 2140-8.

Lichtenstein, A., J. Berenson, et al. (1990). "Resistance of human ovarian cancer cells to tumor necrosis factor and lymphokine-activated killer cells: correlation with expression of HER2/neu oncogenes." Cancer Res **50**(22): 7364-70.

Lind, G. E., L. Thorstensen, et al. (2004). "A CpG island hypermethylation profile of primary colorectal carcinomas and colon cancer cell lines." Mol Cancer 3(1): 28.

Lindahl, T. (1974). "An N-glycosidase from Escherichia coli that releases free uracil from DNA containing deaminated cytosine residues." PNAS USA 71: 3649-3653.

Lindahl, T., P. Karran, and R. D. Wood. (1997.). "DNA excision repair pathways. ." Curr. Opin. Genet. Dev. 7: 158-169.

Liu, D., Z. Shriver, et al. (2002). "Dynamic regulation of tumor growth and metastasis by heparan sulfate glycosaminoglycans." Semin Thromb Hemost 28(1): 67-78.

Liu, D., Z. Shriver, et al. (2002). "Tumor cell surface heparan sulfate as cryptic promoters or inhibitors of tumor growth and metastasis." Proc Natl Acad Sci U S A 99(2): 568-73.

Liu, J., Z. Shriver, et al. (2002). "Characterization of a heparan sulfate octasaccharide that binds to herpes simplex virus type 1 glycoprotein D." J Biol Chem 277(36): 33456-67.

Liu, J., N. W. Shworak, et al. (1999). "Expression of heparan sulfate D-glucosaminyl 3-O-sulfotransferase isoforms reveals novel substrate specificities." J Biol Chem 274(8): 5185-92.

Liu, S., G. Dontu, et al. (2005). "Mammary stem cells, self-renewal pathways, and carcinogenesis." Breast Cancer Res 7(3): 86-95.

Longacre, T. A. and S. A. Bartow (1986). "A correlative morphologic study of human breast and endometrium in the menstrual cycle." Am J Surg Pathol 10(6): 382-93.

Lorincz, A. M. and S. Sukumar (2006). "Molecular links between obesity and breast cancer." Endocr Relat Cancer 13(2): 279-92.

Lubbert, M. (2000). "DNA methylation inhibitors in the treatment of leukemias, myelodysplastic syndromes and hemoglobinopathies: clinical results and possible mechanisms of action." Curr Top Microbiol Immunol **249**: 135-64.

Lund, G., L. Andersson, et al. (2004). "DNA methylation polymorphisms precede any histological sign of atherosclerosis in mice lacking apolipoprotein E." J Biol Chem **279**(28): 29147-54.

Lynch, E. D., E. A. Ostermeyer, et al. (1997). "Inherited mutations in PTEN that are associated with breast cancer, cowden disease, and juvenile polyposis." Am J Hum Genet **61**(6): 1254-60.

MacMahon, B. (2006). "Epidemiology and the causes of breast cancer." Int J Cancer **118**(10): 2373-8.

Manabe, Y., S. Toda, et al. (2003). "Mature adipocytes, but not preadipocytes, promote the growth of breast carcinoma cells in collagen gel matrix culture through cancer-stromal cell interactions." J Pathol **201**(2): 221-8.

Mannermaa, A., H. Peltoketo, et al. (1994). "Human familial and sporadic breast cancer: analysis of the coding regions of the 17 beta-hydroxysteroid dehydrogenase 2 gene (EDH17B2) using a single-strand conformation polymorphism assay." Hum Genet **93**(3): 319-24.

Maraqa, L. and M. Lansdown (2006). "Study design and statistics in the epidemiology of breast cancer." Br J Cancer **95**(9): 1301; author reply 1302-3.

McPherson, K., C. M. Steel, et al. (2000). "ABC of breast diseases. Breast cancer-epidemiology, risk factors, and genetics." Bmj **321**(7261): 624-8.

McTiernan, A., K. B. Rajan, et al. (2003). "Adiposity and sex hormones in postmenopausal breast cancer survivors." J Clin Oncol **21**(10): 1961-6.

Melki, J. R., P. C. Vincent, et al. (1999). "Concurrent DNA hypermethylation of multiple genes in acute myeloid leukemia." Cancer Res **59**(15): 3730-40.

Miao, X., D. Xing, et al. (2002). "Susceptibility to gastric cardia adenocarcinoma and genetic polymorphisms in methylenetetrahydrofolate reductase in an at-risk Chinese population." Cancer Epidemiol Biomarkers Prev **11**(11): 1454-8.

Millar, D. S., K. K. Ow, et al. (1999). "Detailed methylation analysis of the glutathione S-transferase pi (GSTP1) gene in prostate cancer." Oncogene **18**(6): 1313-24.

Millar, D. S., C. L. Paul, et al. (2000). "A distinct sequence (ATAAA)n separates methylated and unmethylated domains at the 5'-end of the GSTP1 CpG island." J Biol Chem **275**(32): 24893-9.

Mitrunen, K. and A. Hirvonen (2003). "Molecular epidemiology of sporadic breast cancer. The role of polymorphic genes involved in oestrogen biosynthesis and metabolism." Mutat Res **544**(1): 9-41.

Miyamoto, K., K. Asada, et al. (2003). "Methylation-associated silencing of heparan sulfate D-glucosaminyl 3-O-sulfotransferase-2 (3-OST-2) in human breast, colon, lung and pancreatic cancers." Oncogene **22**(2): 274-80.

Momparler, R. L., J. Bouchard, et al. (1984). "5-aza-2'-deoxycytidine therapy in patients with acute leukemia inhibits DNA methylation." Leuk Res **8**(2): 181-5.

Monk, M., M. Boubelik, et al. (1987). "Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development." Development **99**(3): 371-82.

Nakagawa, H., R. B. Chadwick, et al. (2001). "Loss of imprinting of the insulin-like growth factor II gene occurs by biallelic methylation in a core region of H19-associated CTCF-binding sites in colorectal cancer." Proc Natl Acad Sci U S A **98**(2): 591-6.

Nakagawa, H., G. J. Nuovo, et al. (2001). "Age-related hypermethylation of the 5' region of MLH1 in normal colonic mucosa is associated with microsatellite-unstable colorectal cancer development." Cancer Res **61**(19): 6991-5.

Nan, X., R. R. Meehan, et al. (1993). "Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2." Nucleic Acids Res **21**(21): 4886-92.

Narayan, A., W. Ji, et al. (1998). "Hypomethylation of pericentromeric DNA in breast adenocarcinomas." Int J Cancer **77**(6): 833-8.

Ng, H. H., Y. Zhang, et al. (1999). "MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex." Nat Genet **23**(1): 58-61.

NHSScotland (2003). NHSScotland Cancer Incidence (Up to 1999) and Mortality (Up to 2001) statistic.

Nijhawan, R., M. Hemachandran, et al. (2002). "Proliferative breast disease revisited." Acta Cytol **46**(5): 1027-8.

Olson, J. E., J. N. Ingle, et al. (2006). "A comprehensive examination of CYP19 variation and risk of breast cancer using two haplotype-tagging approaches." Breast Cancer Res Treat.

Pellise, M., A. Castells, et al. (2004). "Detection of lymph node micrometastases by gene promoter hypermethylation in samples obtained by endosonography-guided fine-needle aspiration biopsy." Clin Cancer Res **10**(13): 4444-9.

Perrimon, N. and M. Bernfield (2000). "Specificities of heparan sulphate proteoglycans in developmental processes." Nature **404**(6779): 725-8.

Pfeifer, G. P. (2000). "p53 mutational spectra and the role of methylated CpG sequences." Mutat Res **450**(1-2): 155-66.

Pfeifer, G. P., M. Tang, et al. (2000). "Mutation hotspots and DNA methylation." Curr Top Microbiol Immunol **249**: 1-19.

Piccart-Gebhart, M. J., M. Procter, et al. (2005). "Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer." N Engl J Med **353**(16): 1659-72.

Piekarz, R. L., R. Robey, et al. (2001). "Inhibitor of histone deacetylation, depsipeptide (FR901228), in the treatment of peripheral and cutaneous T-cell lymphoma: a case report." Blood **98**(9): 2865-8.

Plumb, J. A., G. Strathdee, et al. (2000). "Reversal of drug resistance in human tumor xenografts by 2'-deoxy-5-azacytidine-induced demethylation of the hMLH1 gene promoter." Cancer Res **60**(21): 6039-44.

Pogribny, I. P., B. J. Miller, et al. (1997). "Alterations in hepatic p53 gene methylation patterns during tumor progression with folate/methyl deficiency in the rat." Cancer Lett **115**(1): 31-8.

Prasad, S. C., P. J. Thraves, et al. (1999). "Differential expression of stathmin during neoplastic conversion of human prostate epithelial cells is reversed by hypomethylating agent, 5-azacytidine." Int J Oncol **14**(3): 529-34.

Prokhortchouk, A., B. Hendrich, et al. (2001). "The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor." Genes Dev **15**(13): 1613-8.

Radford, D. M. and B. A. Zehnbauer (1996). "Inherited breast cancer." Surg Clin North Am **76**(2): 205-20.

Ramakrishnan, R., S. A. Khan, et al. (2002). "Morphological changes in breast tissue with menstrual cycle." Mod Pathol **15**(12): 1348-56.

Raman, R., J. Myette, et al. (2002). "Identification of structural motifs and amino acids within the structure of human heparan sulfate 3-O-sulfotransferase that mediate enzymatic function." Biochem Biophys Res Commun **290**(4): 1214-9.

Rao, L. V. and U. R. Pendurthi (1998). "Tissue factor on cells." Blood Coagul Fibrinolysis **9 Suppl 1**: S27-35.

Reiland, J., R. D. Sanderson, et al. (2004). "Heparanase Degrades Syndecan-1 and Perlecan Heparan Sulfate: FUNCTIONAL IMPLICATIONS FOR TUMOR CELL INVASION." J Biol Chem **279**(9): 8047-55.

Ribeiro, F. S., L. M. de Amorim, et al. (2006). "CYP19 (TTTA)n polymorphism and breast cancer risk in Brazilian women." Toxicol Lett **164**(1): 90-5.

Robinson, J. (1986). "Treatment of Breast Cancer through the ages." Am J Surg. **1986 Mar;151(3):317-33**.

Rodger, A. (2001). "Polychemotherapy for early breast cancer." Lancet **358**(9299): 2170.

Rosai, J. (2004). Rosai and Ackerman's Surgical Pathology.

Rosen, P. P. (1993). "Proliferative breast "disease". An unresolved diagnostic dilemma." Cancer **71**(12): 3798-807.

Rosen, P. P. (2004). Rosen's Breast Pathology.

Ross, j. a. R., L (1984). Histology.

Ruter, B., P. W. Wijermans, et al. (2004). "DNA methylation as a therapeutic target in hematologic disorders: recent results in older patients with myelodysplasia and acute myeloid leukemia." Int J Hematol **80**(2): 128-35.

S.Chummy, S. (2006). Last's Anatomy: Regional and Applied.

Salem, C., G. Liang, et al. (2000). "Progressive increases in de novo methylation of CpG islands in bladder cancer." Cancer Res **60**(9): 2473-6.

Sanchez-Cespedes, M., M. Esteller, et al. (1999). "Molecular detection of neoplastic cells in lymph nodes of metastatic colorectal cancer patients predicts recurrence." Clin Cancer Res **5**(9): 2450-4.

Santini, V., H. M. Kantarjian, et al. (2001). "Changes in DNA methylation in neoplasia: pathophysiology and therapeutic implications." Ann Intern Med **134**(7): 573-86.

Sasisekharan, R., Z. Shriver, et al. (2002). "Roles of heparan-sulphate glycosaminoglycans in cancer." Nat Rev Cancer **2**(7): 521-8.

Sasisekharan, R. and G. Venkataraman (2000). "Heparin and heparan sulfate: biosynthesis, structure and function." Curr Opin Chem Biol **4**(6): 626-31.

Sato, N., N. Maehara, et al. (2003). "Effects of 5-aza-2'-deoxycytidine on matrix metalloproteinase expression and pancreatic cancer cell invasiveness." J Natl Cancer Inst **95**(4): 327-30.

Schmutte, C. and P. A. Jones (1998). "Involvement of DNA methylation in human carcinogenesis." Biol Chem **379**(4-5): 377-88.

Segura-Pacheco, B., C. Trejo-Becerril, et al. (2003). "Reactivation of tumor suppressor genes by the cardiovascular drugs hydralazine and procainamide and their potential use in cancer therapy." Clin Cancer Res **9**(5): 1596-603.

Shao, Z. M. and M. Nguyen (2002). "Tumor-specific DNA in plasma of breast cancer patients." Anticancer Drugs **13**(4): 353-7.

Shriver, Z., D. Liu, et al. (2002). "Emerging views of heparan sulfate glycosaminoglycan structure/activity relationships modulating dynamic biological functions." Trends Cardiovasc Med **12**(2): 71-7.

Shworak, N. W., J. Liu, et al. (1999). "Multiple isoforms of heparan sulfate D-glucosaminyl 3-O-sulfotransferase. Isolation, characterization, and expression of human cdnas and identification of distinct genomic loci." J Biol Chem **274**(8): 5170-84.

Sidransky, D., T. Tokino, et al. (1992). "Inherited p53 gene mutations in breast cancer." Cancer Res **52**(10): 2984-6.

Silva, J., J. M. Silva, et al. (2003). "Concomitant expression of p16INK4a and p14ARF in primary breast cancer and analysis of inactivation mechanisms." J Pathol **199**(3): 289-97.

Silva, S. N., M. N. Cabral, et al. (2006). "Breast cancer risk and polymorphisms in genes involved in metabolism of estrogens (CYP17, HSD17beta1, COMT and MnSOD): possible protective role of MnSOD gene polymorphism Val/Ala and Ala/Ala in women that never breast fed." Oncol Rep **16**(4): 781-8.

GLASGOW UNIVERSITY LIBRARY