Sporadic Breast Cancer in Young Women:

A Microarray Investigation

Thesis submitted for the Degree of
Doctor of Philosophy
at the University of Leicester

By

Gavin John Whyman B.Sc. (Leeds) M.Sc. (Leicester)
Breast Cancer Research Unit
Cancer Studies and Molecular Medicine
University of Leicester

December 2005

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

Abstract

The aim of this thesis was to test the hypothesis that sporadic breast cancers in young women arise because of a profile of genetic alterations specific to that age group. Previous studies by our research group had identified a higher frequency of Loss of Heterozygosity (LOH) in the BRCA1, BRCA2 and p53 intervals in younger breast cancer cases (< 35 years) when compared to matched older cases (> 55 years); BRCA1 and P53 data had been related to protein expression. Firstly BRCA2 expression was investigated in the same group using RT-PCR and immunohistochemistry. Secondly, methylation specific digestion and androgen receptor PCR were used to investigate skewed X-chromosome inactivation comparing younger and older cases. Finally, two separate microarray studies were performed to investigate both genomic copy number and gene expression changes. These were validated using semi-quantitative RT-PCR and qRT-PCR respectively.

Investigations of BRCA2 protein and mRNA expression were inconclusive. A higher frequency of skewed X chromosome inactivation was identified in young women (< 35 years) compared to older women (>55 years) with sporadic breast cancer in this study (P = 0.017), which might suggest the involvement of an X linked gene in the genesis of these tumours. Although it was not possible to relate cDNA expression levels to genomic copy number changes for individual tumours, the analysis of gene expression changes using Affymetrix Gene Chips identified a number of novel gene targets showing elevated expression in young breast cancers when compared to normal breast organoids. Several statistically significant candidate genes were investigated, and gene expression changes confirmed by qRT-PCR and for RBBP4 by immunohistochemistry. Of the significant gene targets that were ranked by SAM analysis DDB2 and RBBP4 are most noteworthy. Statistically significant differences in the level of expression of these two genes might suggest that they are novel markers of young breast cancers, meriting investigation in a larger series of cases.

Acknowledgements

First and foremost I'd like to thank my supervisors, Professor Rosemary Walker and Dr Jacqui Shaw. Without your support, expertise and encouragement this thesis wouldn't be close to completion. As this is the only page that you won't be proof reading may I apologise in advance for my poor English, bad grammar and changing tenses.

I am very lucky to have been working at the Breast Cancer Research Unit. This great unit is expertly managed by Karen Kulbicki. I'd obviously like to thank you for all your help at the "bucket chemistry" that is immunohistochemistry. But more importantly I'd like to thank you for your friendship over the years. You are the first person I turn to for a moan but the first to give me a good kick up the arse when I needed it. The sign of a true friend.

As I said, I was lucky to be working at BCRU, and this in small part to the folk that have worked there over the years. There have been far too many academic staff, post-docs, Ph.D., M.D., M.Sc. and B.Sc. students to name individually, but I would especially like to acknowledge; Louise Jones who showed me that being IT illiterate needn't slow your progress to professor; Jeni Luckett who proved that there are many different shades of black; and Rachael Alcock for introducing me to the wonderful world of 'Me to You Bears'.

One of the perks of working for BCRU was being able to tap in to vast expertise of the 'Old Pathology Lab'. Again, there are too many people to mention you all, so I'll shorten the list down to Leanne, Linda, Lindsey & Angie and I'd especially like to thank Howard Pringle for asking all those impossible questions forcing me to read far more than I would of liked.

No acknowledgement is complete without thanking your family and this is no different. Mum, Dad & our Sally – You've always been there for me, supporting me in everything that I do. So, while the support you've given me over the years isn't unexpected it's certainly not unappreciated (even though it may seem that way at times).

And last but certainly not least I'd like to thank my "Little Princess" Alicia Kulbicki. When ever I needed to relax with a pizza-picnic and a game of Guess Who or Marble Run you were there for me...Cheers lass.

Contents

Tables and Figures

Tables

Figures

Abbreviations

A	Adenine
AMV	Avian Myeloblastosis Virus
AP	Alkaline phosphatase
AR	Androgen Receptor
BAC	Bacterial artificial chromosome
BRCA1	Breast cancer associated gene 1
BRCA2	Breast cancer associated gene 2
C	Cytosine
cDNA	complementary DNA
CGH	Comparative genomic hybridisation
CORT	Cloning of receptor targets
CRE	cAMP response element
cRNA	complementary RNA
Ct	Cycle threshold
Cys	Cysteine
DAB	3, 3' - Diaminobenzidine
DCIS	Ductal carcinoma in situ
DMEM	Dulbecco/Vogt modified Eagle's minimal essential media
DNA	Deoxyribonucleic acid
dNTP	deoxyneucleotide triphosphate
ds	double stranded
DSB	Double strand break
ECL	Electochemiluminesance
ER	Oestrogen receptor
ES	Embryonic Stem Cell
FCS	Foetal calf serum
FFPE	Formalin fixed paraffin embedded

G	Guanine
H & E	Haematoxylin and Eosin
HRP	Horseradish peroxidase
IDC	Infiltrating ductal carcinoma
IEL	Isoelectic focusing
ILL	Infiltrating lobular carcinoma
IMS	Industrial methylated spirits
IVT	In vitro transcription
LCIS	Lubular carcinoma in situ
LN	Lymph node
LOH	Loss of heterozygosity
mRNA	Messenger RNA
NHEJ	Non-homologous end joining
NPA	No primary antibody
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCI	Phenol, chloroform, IAA
PCR	Polymerase chain reaction
PgR	Progesterone Receptor
PK	Proteinase K
qRT-PCR	Quantitative RT-PCR
RNA	Ribonucleic acid
RPMI	Roswell Park Memorial Institute media
rRNA	Ribosomal RNA
RT	Reverse transcription
SAM	Significance analysis of microarrays
SNP	Single nucleotide polymorphism
ss	Single stranded
St-ABC	Streptavidin binding complex

T	Thymine
TAE	Tris acetate EDTA
TBE	Tris borate EDTA
TBS	Tris buffered saline
TBST	TBS with Tween 20
tRNA	Transfer RNA
Tyr	Tyrosine
UTR	Untranslated region
XIST	X inactivation specific transcript

Chapter 1:

General Introduction to Breast Cancer

In North-West Europe, Australia and North America, breast cancer the most prevalent form of female cancer, accounting for 30% of malignancies in the UK female population (Office for National Statistics, 2003). There is approximately a 1 in 9 risk of any woman developing breast cancer in her lifetime and it is the most common cause of death in women aged 35-55 years. In 2000, there were 40,467 new breast cancer cases and 12,838 deaths attributed to breast cancer in women in the UK (Office for National Statistics, 2003).

1.1 Risk Factors for the development of Breast Cancer

1.1.1 Family history

The most significant risk factor in the development of breast cancer is having first degree relatives who have developed breast cancer whilst premenopausal (Clamp *et al*, 2002; Toms, 2004). Of all breast cancers, approximately 4% are due to inherited mutations in one of two breast cancer susceptibility genes, BRCA1 and BRCA2. Whilst these account for only a small fraction of all the breast cancer cases, pathogenic mutations in these genes have a high penetrance, conferring a 85% lifetime risk (Bertwistle and Ashworth, 2000; Kerr and Ashworth, 2001).

1.1.2 Sex and age

Breast cancer is almost exclusively a female disease, with greater that 99% of all sufferers being female (Office for National Statistics, 2003; Toms, 2004). Increasing age, as with many cancers, is also a significant risk factor, following a sigmoid-like relationship (Office for National Statistics, 2003, Toms, 2004).

1.1.3 Age at menarche

Women with an early age of menarche show a significantly increased risk of developing breast cancer. Average age of first menarche in developed countries has fallen from around 16-17 years in the mid-nineteenth century to 12-13 today (Tanner *et al*, 1973, Toms, 2004). The relative risk for premenopausal breast cancer is reduced by about 7% for each year that menarche is delayed after 12 years of age, and by about 3% for postmenopausal breast cancer (Clavel-Chapelon *et al*, 2002).

1.1.4 Age at menopause

Late menopause increases the risk of developing breast cancer (Kelsey *et al*, 1993). For each year menopause is delayed, there is an approximately 3% increase in breast cancer risk (Toms, 2004). This means that post menopausal women have a lower risk than premenopausal women of the same age and this is true whether the menopause is natural or induced through surgery (Collaborative Epidemiology Study, 1997).

1.1.5 Age at first full-term pregnancy

Women who do not have children (nulliparous) have an increased risk of developing breast cancer. However, the beneficial effects of having children are limited to women who have their first child before 30 years. Women who delay until they are in their mid to late thirties have a greater risk than nulliparous women (Walker, 1996).

1.1.6 Weight and diet

Obesity has been shown to moderately increase the risk of developing breast cancer, and is one of the few modifiable risk factors (van den Brant *et al*, 2000). About 8% of breast cancers may be attributable to being over weight and obesity (Bergstrom *et al*, 2001). In a pooled analysis, the risk of developing breast cancer increased by approximately 30% in

postmenopausal women with a BMI of over 28 kg/m^2 compared to those with a BMI of less than 20 kg/m^2 (van den Brant *et al*, 2000)

Whilst weight is an obvious consequence of diet, the nature of the diet is also significant. There is no good evidence as to why, but both in rodent studies and in epidemiological studies, a high fat diet correlates with an increased risk of developing breast cancer (Jasienska and Thune, 2001).

1.1.7 Geographic variation

Breast cancer is most prevalent in Northwest Europe, Australia, New Zealand and North America and is lowest in Southeast Asia and Africa (Ferlay *et al*, 2001; Ferlay *et al*, 2003). This is probably due to a combination of lifestyle and social-economic factors, including age at first full term pregnancy, post-menopausal weight and diet (Jasienska and Thune 2001). Moreover, migrants from low to high-risk countries acquire the risk of the host country within two generations (Shimizu *et al*, 1985; Tominaga *et al*; 1985, Ziegler *et al*, 1993).

1.1.8 Atypical hyperplasia

Benign breast diseases showing atypical epithelial hyperplasia impart a two to five fold increased risk of developing breast cancer in the future. Ordinary epithelial hyperplasia is also associated with a two fold increase (Dupont *et al*, 1985, Byrne *et al* 2000, Byrne *et al*, 2001).

1.2 Breast cancer histopathology

There are two main types of breast cancer, non-invasive breast carcinoma (carcinoma *in situ*) and invasive breast carcinoma.

1.2.1 Non-invasive breast carcinoma

In these lesions, the tumour cells are confined to the ductal or lobular units of the breast, with no penetration of the basement membrane. Depending on the cytological features these are divided into ductal or lobular carcinoma in situ.

1.2.1.1 Ductal carcinoma in situ (DCIS)

Until the advent of mammographic screening, DCIS was thought to account for only a very small proportion of breast cancers (5.0%). As the majority of women diagnosed with DCIS previously underwent mastectomy, there is very little evidence currently available to determine whether DCIS inevitably leads to invasive carcinoma (Trojani *et al*, 1991).

1.2.1.2 Lobular carcinoma in situ (LCIS)

Like DCIS, these lesions arise from the terminal ductules or acini. They are often multicentric, do not form a palpable tumour and have no characteristic radiographic signs. The incidence of LCIS is therefore difficult to determine. LCIS confers an increased risk of developing invasive carcinoma, but unlike DCIS can occur anywhere in the breast (Trojani *et al*, 1991).

1.2.2 Invasive carcinoma

1.2.2.1 Infiltrating ductal carcinoma (IDC)

These form the largest group of infiltrating carcinomas, accounting for approximately 73% of all infiltrating carcinomas. There is great histological variation within the type. The cells may be trabeculae or solid groups, and the cells range from regular to pleomorphic (Underwood *et al*, 1996).

1.2.2.2 Infiltrating lobular carcinoma (ILC)

Accounting for approximately 10% of all infiltrating breast carcinomas, these typically present as a palpable area of ill-defined thickness, in contrast to the prominent lump of ductal carcinoma. ILC is characterised by single cell infiltrations, often around pre-existing breast structures. This is more likely to be associated with bilateral breast cancer than IDC, and tends to be multicentric within the same breast (Underwood *et al*, 1996).

1.2.2.3 Special Types

There are several other types of breast carcinomas, however these are much rarer. They include mucoid or colloid (2.4%), tubular (1.2%), adenoid cystic (0.4%), cribriform (0.3%) and carcinosarcoma (0.1%). These percentages are the frequency of the tumour type of all invasive breast carcinomas (Joensuu and Toikkanen, 1995).

1.3 Tumour Stage

Irrespective of the tumour type, staging systems are used to assess the extent of spread of the carcinoma. There are essentially two internationally recognised staging systems. The first, a simple numbering system, is described in Table 1.1.

Stage	Description
I	Tumour up to 2 cm No lymph nodes affected No evidence of spread beyond the breast
II	Tumour between 2 and 5 cm and/or Lymph nodes in armpit affected No evidence of spread beyond armpit
III	Tumour more than 5 cm Lymph nodes in armpit affected No evidence of spread beyond armpit
IV	Tumour any size Lymph nodes in armpit often affected Cancer has spread to other parts of body

Table 1.1 Staging of breast cancers (Korkolis *et al*, 2004)

The second system, the TNM System, describes the tumour, node status and whether the tumour has metastasised (Table 1.2).

Stage	Description
T_1	Tumour 20 mm or less; no fixation or nipple retraction. Includes Paget's disease.
T_2	Tumour 20-50 mm or less than 20 mm but with tethering.
T_3	Tumour greater than 50 mm but less than 100 mm; or less than 50 mm but with infiltration, ulceration or fixation.
T_4	Any tumour with ulceration or infiltration wide of it, or chest wall fixation, or greater than 100 mm in diameter.
N_0	Node negative.
N_1	Axillary nodes mobile.
N_2	Axillary nodes fixed.
M_0	No distant metastases.
M_1	Distant metastases.

Table 1.2 The TNM system used to stage tumour progression (Singletary *et al*, 2002).

1.4 Tumour grade

In addition to the stage, tumours are classified according to their grade with a score from 1 to 3. The grade of the tumour is based on the level of differentiation within the tumour as evaluated based on tubule formation, nuclear pleomorphism and mitotic count. Both the stage and grade of a tumour are prognostic indicators. The higher the stage and grade, the poorer the prognosis (Elston and Ellis 1991).

1.5 The Genetics of Breast Cancer

Most of our knowledge of breast cancer genetics has come through studying breast cancer families. BRCA1 and BRCA2 are high penetrance cancer genes that offer a lifetime risk of developing breast cancer between 60-85% (Ford *et al*, 1998). However, only 4-6% of all breast cancers are believed to be attributable to mutation in either of these genes (de

Jong *et al*, 2002). Analysis of mutations has shown that these high penetrance genes account for less than 50% of all inherited breast cancer (Figure 1.1). A total of 33 genes in which sporadic mutations have been identified are given in Table 1.3.

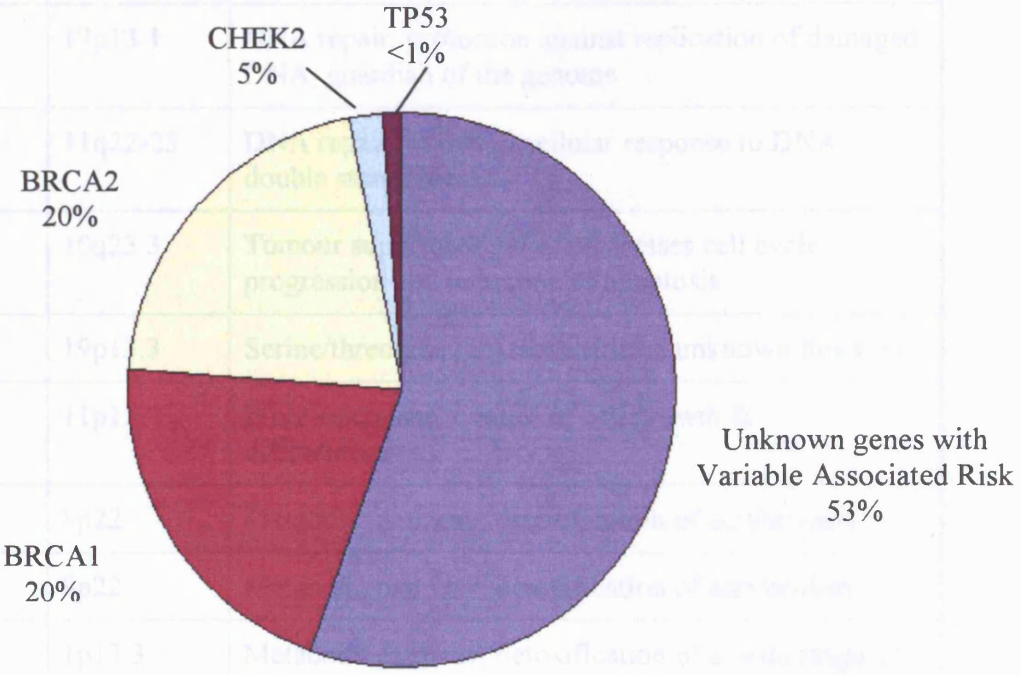

Figure 1.1 Inherited BRCA1 and BRCA2 mutations each occur in approximately 20% of breast cancer patients with a family history of breast cancer. TP53 mutations, which cause Li-Fraumeni syndrome, account for less than one percent of familial breast cancers, but women who survive childhood cancers have a 90% risk of developing breast cancer. Mutations in the cell cycle checkpoint kinase gene (CHEK2) account for 5% of cancers. The others are due to as yet undetermined factors such as additional susceptibility genes, exposure to hormonal and environmental factors and stochastic genetic events (Nathanson *et al*, 2001).

Gene	Location	Function
BRCA1	17q21	DNA repair
BRCA2	13q12-13	DNA repair
TP53	17p13.1	DNA repair, protection against replication of damaged DNA, guardian of the genome
ATM	11q22-23	DNA repair, sensor in cellular response to DNA double strand breaks
PTEN	10q23.3	Tumour suppressor gene, suppresses cell cycle progression and induction of apoptosis
LKB1	19p13.3	Serine/threonine kinase, otherwise unknown function
HRAS1	11p15	Proto-oncogene, control of cell growth & differentiation
NAT1	8p22	Metabolic pathway, detoxification of acrylamines
NAT2	8p22	Metabolic pathway, detoxification of acrylamines
GSTM1	1p13.3	Metabolic pathway, detoxification of a wide range of xenobiotics, including environmental carcinogens, chemotherapeutic agents and reactive oxygen species
GSTP1	11q13	Metabolic pathway, detoxification of numerous chemicals including chemotherapy agents and catechol oestrogens
GSTT1	11q	Metabolic pathway, detoxification of a wide range of xenobiotics, including environmental carcinogens, chemotherapeutic agents and reactive oxygen species
CYP1A1	15q	Metabolic pathway, oestrogen pathway, metabolism of oestrogens and PAHs
CYP1B1	2p21	Metabolic pathway, metabolism of PAHs
CYP2D6	22q11	Metabolic pathway, metabolism of many commonly prescribed drugs, including codeine
CYP17	10q24.3	Oestrogen pathway, balance of oestrogens, progesterones and androgens
CYP19	11q21.1	Oestrogen pathway, catalysing the conversion of androgens into oestrogens, determines local oestrogen

		level
ER	6q25.1	Oestrogen pathway, binding and transfer of oestrogens to the nuclei, ER modulates transcription of a number of growth factors including IGF-1 and TGFα
PR	11q24.1	Oestrogen pathway
AR	Xq11-12	Oestrogen pathway
COMP	22q11.2	Oestrogen pathway, conjunction and inactivation of catechol oestrogens
UGT1A1	2q37	Oestrogen pathway, phase II drug metabolism and maintains intracellular steady state levels of oestrogens
TNFα	6p21	Immuno pathway, central mediator in the inflammation response and immunological activities of tumour cells
HSP70	6p21	Molecular chaperone, regulate the structure, subcellular localisation, and turn over of cell proteins
HFE	6p21	Iron metabolism pathway
TFR	3q	Iron metabolism pathway
VDR	12q	Cell differentiation
APC	5q22	Inhibits the progression of cells from G1 to S phase, apoptosis, and cell-cell interactions
APOE	19q13.2	Lipid metabolism
CYP2E1	10q24.3	Metabolic pathway, metabolism of acetone, ethyl glycol and ethanol
EDH17B2	17q12-21	Oestrogen pathway, catalyses the reaction between oestrone and oestradiol
HER2	17q21	Proto-oncogene, control of cell growth and proliferation
TβR-1	9q33-34	Cell growth

Table 1.3 A summary of 33 genes most commonly associated with breast cancer, along with their genomic locus and known function (de Jong *et al*, 2002).

1.6 Breast cancer in young women

It has been known since the mid-70's that young women (< 40 years) who develop breast cancers have a poorer survival rate than older women (Wallgren *et al,* 1977). Indeed the 10 year disease-free survival rate for young women (< 35 years) is only 35% in comparison to 47% for older women (≥ 35 years) (Aebi *et al,* 2000). The 30-year survival rate falls to 19% in women under 30 years, and 10% if the patient has IDC (Feldman *et al,* 1998).

A number of studies have suggested that the poorer prognosis of young women with sporadic breast cancer is due to the differing biology and nature of the tumours from those of older women (Albain *et al,* 1994; Pillers 1992; Walker *et al,* 1996b; Wenger *et al,* 1993, Johnson *et al,* 2001).

1.6.1 Breast cancers in younger women are different

There are several markers that can provide information about poor outcome. These include the expression of p53 (Barnes *et al,* 1993; Thor *et al,* 1992), overexpression of the oncogenic protein HER-2 (Gullick *et al,* 1991; Walker *et al,* 1996b), lack of the oestrogen or progesterone receptors (Foekens *et al,* 1989; Reiner *et al,* 1990) and higher levels of Ki-67 antigen, a marker of proliferation (Railo *et al,* 1993). In order to determine whether breast cancers in young women were different from those in older women, these markers were used to assess the nature of 163 invasive breast carcinomas stratified by age by Walker *et al* (1996). These results are summarised in Table 1.4.

Age (years)	25-29	30-34	35-39	40-44	Control 50-67
n	18	30	40	75	70
ER (+)-ive	44%	57%	70%	49%	67%
PgR (+)-ive	33%	37%	60%	44%	48.5%
HER-2 (+)-ive	22%	20%	22.5%	17%	17%
p53 (+)-ive	67%	53%	45%	40%	37%
Proliferation					
low	6%	20%	42.5%	33%	50%
medium	22%	13%	17.5%	17%	10%
high	72%	67%	40%	50%	40%

Table 1.4 A comparison of prognostic markers of breast cancers stratified by age. (Walker *et al*, 1996).

Overall the carcinomas occurring in women less than 35 years had more aggressive features, with a higher incidence of grade III cases, p53 positivity, high proliferation and lacking the oestrogen and progesterone receptors.

Subsequent studies within our unit identified differences in the frequency of loss of heterozygosity (LOH) at the chromosomal intervals containing BRCA1 (17q21), BRCA2 (13q12-13) and TP53, in young patients (<35 years) compared to older cases (>55 years) matched for size, grade, and node status. At the BRCA1 locus, 64.5% of young breast cancers showed LOH for at the marker D17S855, whereas the 35.5% of older patients showed LOH at this marker ($p < 0.025$). At the BRCA2 D13S260 locus, LOH was detected in 74.2% and 30% of cases in young and older patients respectively ($p < 0.001$). Within the same study, LOH analysis was performed for three loci on chromosome 17p. This region contains the proto-oncogene *TP*53. At the most informative locus (D17S799)

the rate of LOH was 63.3% in younger women compared to 41.4% in older women (Johnson *et al*, 2002). This was not shown to be statistically significant. Gentile and colleagues (1999) investigated LOH at the p53 region of chromosome 17p and found 43% of informative young cases displayed LOH. In addition to this, it was also noted that 17% of young women had a mutation in the p53 gene and 46% showed protein accumulation. However, none of these factors were markers of poor prognosis.

Weber-Mangal *et al* (2003) also showed differences in sporadic breast cancer in young and older women using comparative genomic hybridisation (CGH). In cancers from young women genomic gains were observed clustered to chromosome arms 1q (65%), 8q (61%), 17q (50%), 20q (33%), 3q (21%), 1p (17%), 5p (17%) and 15q (17%). Genomic losses were commonly located on 8p (19%), 11q (11%), 16q (11%), 17p (11%) and 18q (10%). In comparison to the breast cancers of older women, genomic losses were more frequent in younger (<35 years) women. Losses on 8p22-p23 were more prevalent for patients with lymph node metastasis ($p = 0.002$), and grade III tumours were associated with gains in the long arm of chromosome 8 ($p = 0.01$).

Taking these data into consideration it is now generally accepted that early onset sporadic breast cancer is a discrete entity, with a novel aetiology. Different studies have considered various genetic alterations that may contribute to its development. One such investigation (Bergman-Jungestrom *et al*, 1999), led to the discovery of a single nucleotide polymorphism (SNP) in the cyp17 promoter, an enzyme involved in oestrogen biosynthesis. This T→C SNP produces a Sp1-type promoter site (CCACC box) and therefore increases the production of oestrogen. In this study, younger women were shown to be more likely to be either heterozygous or homozygous for the allele with the C mutation. As oestrogen exposure is a well-recognised risk factor, this suggests a possible role for oestrogen and the *CYP17* gene interaction in young breast cancer (Bergman-

Jungestrom *et al*, 1999). This may also explain why young women with ER+ tumours have a significantly poorer prognosis than ER- cases (Aebi *et al*, 2000).

Other studies of young women with breast cancer have shown a significant skewed X-chromosome inactivation pattern compared to older cases and normal populations. This might suggest that there is a low penetrance gene on the X-chromosome that not only skews the inactivation of the X-chromosome, but also may confer an increase risk of early onset breast cancer (Kristiansen *et al*, 2002).

1.7 BRCA1

In 1988, Newman *et al* were able to demonstrate that breast cancer could be inherited in and autosomal dominant manner in families (Newman *et al*, 1988). A further six years passed before the BRCA1 gene was identified and located at 17q21 using positional cloning techniques (Miki *et al*, 1994). BRCA1 is a large gene, spanning $\approx$ 84 kb and containing 24 exons (22 coding), with exon 11 being unusually large and comprising approximately 60% of the BRCA1 protein.

The BRCA1 protein comprises 1863 amino acid residues, with very little homology to other known proteins (Miki *et al*, 1994). The majority of functional information about BRCA1 has therefore come from studies of protein interactions. These interactions are summarised in Figure 1.2

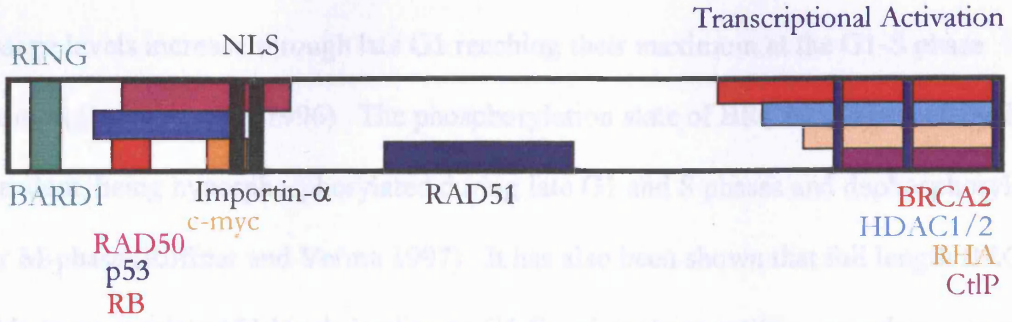

Figure 1.2 Functional domains of BRCA1, and a key sites of protein interactions. (Adapted from Welcsh *et al*, 2000)

1.7.1 BRCA1 in DNA repair

One of the primary functions of BRCA1 is in DNA repair. BRCA1 interacts directly with other proteins implicated in DNA repair, which include RAD50 (Zhong *et al,* 1999), RAD51 (Scully *et al,* 1997b; Scully *et al,* 1997a) and BRCA2 (Chen *et al,* 1999). In addition to the interaction of BRCA1 with RAD51, upon treatment with DNA damaging agents such as γ-radiation (Scully *et al,* 1997b), BRCA1 and RAD51 have been shown to co-localise at sites of DNA repair at intranuclear structures (Chen *et al,* 1999). More direct evidence connecting BRCA1 to DNA damage repair came from functional analysis using BRCA1-deficient embryonic stem (ES) cells that were found to be defective in transcription coupled repair (Gowen *et al,* 1998). In addition, the mutants showed impaired homology directed repair of chromosomal double strand breaks (DSBs) (Moynahan *et al,* 1999).

1.7.2 BRCA1 as a cell cycle regulator

BRCA1 may also have a role in the regulation of the cell cycle. BRCA1 associates with many cell-cycle proteins including E2F, CDC2 and various cyclins (Wang *et al,* 1997). BRCA1 mRNA levels have been shown to fluctuate throughout the cell cycle. The

message levels increase through late G1 reaching their maximum at the G1-S phase transition (Vaughn *et al,* 1996). The phosphorylation state of BRCA1 is also cell cycle dependent, being hyperphosphorylated during late G1 and S phases and dephosphorylated after M-phase (Ruffner and Verma 1997). It has also been shown that full length BRCA1 is able to up regulate p21 levels leading to G1-S cell-cycle arrest (Somasundaram *et al,* 1997).

1.7.3 A transcriptional role for BRCA1

The first of the c-terminal BRCT (BRCA c-terminal) domains interacts with activators and co-repressors. In order to act as a transcriptional regulator, the C-terminus of BRCA1 must interact with the RNA pol II holoenzyme (Anderson *et al,* 1998; Cui *et al,* 1998; Neish *et al,* 1998; Scully *et al,* 1997c). The second of the two domains interacts with p53 to stimulate p53-dependent transcription of p21 (Chai *et al,* 1999).

BRCA1 also contains a c-myc binding site between amino acid residues 175-303 and 433-511 (Wang *et al,* 1998). C-myc is a potent oncogene, which causes tumour formation when over expressed in mouse mammary glands (Stewart *et al,* 1984). Over expression of BRCA1 in these tumour cells was found to suppress the oncogenic potential of both c-myc and H-ras in a dose dependent manner (Wang *et al,* 1998).

From all these interacting proteins it is clear that BRCA1 forms a central hub, connecting DNA damage repair and cell cycle arrest using, at least in part, its ability to act as a transcriptional regulator.

1.7.4 The role of BRCA1 in sporadic breast cancers

BRCA1 has a clear role in inherited breast cancers, with approximately 20% of these having a BRCA1 mutation (Miki *et al,* 1994). However, inherited breast cancers only account for about 5% of all cases. Of the remaining 95%, there have been no reports of

mutations in the BRCA1 gene. There does seem to be a role for BRCA1 in sporadic breast carcinoma however, as many show loss of heterozygosity (LOH) at this locus and reduced BRCA1 protein levels (Yoshikawa *et al,* 1999).

One possible method of BRCA1 deactivation is through aberrant promoter hypermethylation. Alterations in promoter methylation patterns have been shown to silence gene expression, and have been noted in several tumour suppressor genes. This process has been shown to disrupt the binding transcriptional proteins to the promoter region. One such example is that of the CRE and RBF-1 regulatory elements in the RB1 gene, in which aberrant hypermethylation sufficiency disrupts trans- factor binding to dramatically reducing RB1 expression (Chen *et al,* 2000). Within the BRCA1 promoter region there is a cAMP response element (CRE) that has been shown to be hypermethylated in sporadic breast cancers (Magdinier *et al,* 1998; Mancini *et al,* 1998; Rice and Futscher 2000; Rice *et al,* 2000). The most compelling evidence comes from DiNardo *et al* (2001). Using a luciferase reporter assay, they were able to show that the CRE in BRCA1 is functional. Moreover, the function of the CRE is methylation sensitive with decreased activity when the CRE is hypermethylated. This is one possible method by which BRCA1 protein levels could be reduced without somatic mutation of the BRCA1 gene.

More recently, Wei *et al* (2005) described a subset of sporadic breast cancers which presented with a molecular and clinicopathological similar to that of inherited *BRCA1*-associated breast cancers. The mean copy numbers of *BRCA1* and *CEP17* per cell was significantly lower in cases with hypermethylation of the *BRCA1* promoter [1.87 vs. 2.3 ($P = 0.001$) and 1.85 vs. 2.29 ($P = 0.005$) respectively]. In addition, promoter hypermethylation was more frequent observed in young women ($P = 0.05$), with high-grade ($P = 0.001$), oestrogen receptor negative ($P = 0.04$)and progesterone receptor

negative (P = 0.01) tumours. This data supports the model of tumourigenesis in which *BRCA1* promotor hypermethylation may serve as the "first hit", in a similar manner to a germ line mutation.

1.8 BRCA2

1.8.1 Overview of BRCA2

BRCA2 was mapped to 13q12-q13; the *BRCA2* gene extends over 70 kb and contains 27 exons. The BRCA2 protein is one of the largest proteins to be identified, comprising 3,418 amino acid (a.a.) residues, and with the exception of the BRCT domain shows little homology to BRCA1 (Wooster *et al*, 1995). A schematic illustration of BRCA2 is shown in Figure 1.3, which demonstrates the functional domains identified in the BRCA2 protein (Welcsh *et al*, 2000).

Less is known about the function of BRCA2 than that of BRCA1. As with BRCA1, most of our knowledge comes from the protein interactions it forms.

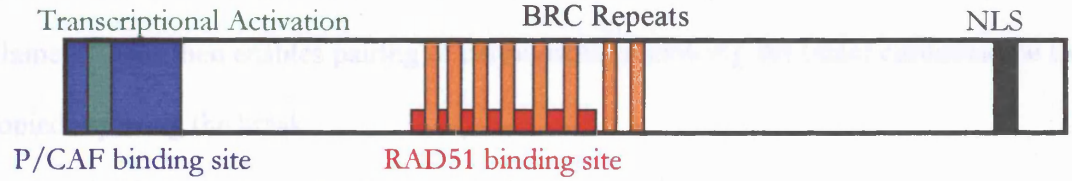

Figure 1.3 An illustration of the major BRCA2 functional domains. (Adapted from Welcsh *et al*, 2000).

1.8.2 BRCA2 and DNA repair

BRCA2 has been consistently linked to DNA repair, including the repair of double-strand breaks (DSB) by homologous recombination and the repair of oxidative damage by transcription coupled repair (Kanaar *et al*, 1998). There are two methods by which mammalian cells are able to repair DNA double strand breaks, non-homologous end

joining (NHEJ) and homologous recombination (Xia *et al,* 2001). There is substantial evidence to suggest that BRCA2 plays a vital role in homologous recombination (Kanaar *et al,* 1998). Within BRCA2, RAD51 interactions are mediated via six sequence specific repeats termed the BRC repeats. RAD51 catalyses strand exchange during directed repair of DNA double strand breaks by gene conversion. As RAD51 mediates its actions by forming a beaded string of RAD51 molecules the multiple RAD51 binding sites suggests a possible role for BRCA2. This is supported by the co-localisation of BRCA2 with BRCA1 and RAD51 at nuclear foci after cells have been exposed to ionising radiation (Yuan *et al,* 1999; Tarsounas *et al,* 2003). Moreover, both human and mouse BRCA2 null cells suffer from chromosomal instability (Xia *et al,* 2001).

As the first step in homologous recombination involves processing the DNA break to produce a single stranded region with a 3′ overhang, Yang *et al* (2005) proposed that BRCA2 then binds to the junction between the double and single stranded DNA with replication protein A (RPA) binding to the single stranded DNA. BRCA2 is then able to recruit Rad51, which displaces the RPA molecules forming a Rad51 nucleoprotein filament. This then enables pairing of chromosomes allowing the intact chromosome to be copied repairing the break.

Tutt *et al* (2001) performed a direct analysis of the ability of BRCA2 mutant cells to repair double-strand breaks created by endonucleases. Additionally, mouse embryonic stem (ES) cells with disruptions in *BRCA2* were found have compromised repair of these DSB by gene conversion (Moynahan *et al,* 1999). It appears that the ES cells used an alternative and lower fidelity, single strand annealing pathway (Tutt *et al,* 2001). Therefore BRCA2 probably acts as caretaker of the genome. By interacting with key enzymes that are involved in DNA repair, BRCA2 is able to maintain genomic integrity and stability. The use of an error-prone repair mechanism in BRCA2-compromised cells would account for

the genomic instability and may therefore contribute to the loss of tumour suppressor genes and accelerate tumour progression.

1.8.3 BRCA2 as a transcriptional regulator

The evidence to support BRCA2 as transcriptional regulator is less certain. When the product of *BRCA2* exon 3 (a. a. 22-105) is fused to the Lex A DNA-binding domain, it is capable of activating transcription in yeast (Milner *et al,* 1997). Within this region, there is a naturally occurring mis-sense mutation (Tyr42Cys) that reduces the transactivation activity. Whether this Tyr42Cys mutation predisposes to breast cancer or is a neutral polymorphism is unclear. The region of BRCA2 encoded by exon 3 has been shown to contain a functionally significant domain. This evidence comes from a Swedish family of breast and ovarian cancers, in which the co-inherited mutation in *BRCA2* is a deletion exon 3 (Nordling *et al,* 1998).

BRCA2 may also be able to activate transcription by modulating the acetylation of histones. BRCA2 interacts with P/CAF (p300/CBP-associated factor), a transcriptional co-activating protein that possesses histone acetyltransferase activity (Fuks *et al,* 1998). The acetylation of histones is able to facilitate translation by unwinding the chromatin allowing RNA pol II to bind to the DNA (Figure 1.4). Taken together, these functions suggest that BRCA2 could act as a modulator of transcription and the loss of this activity could facilitate tumour progression.

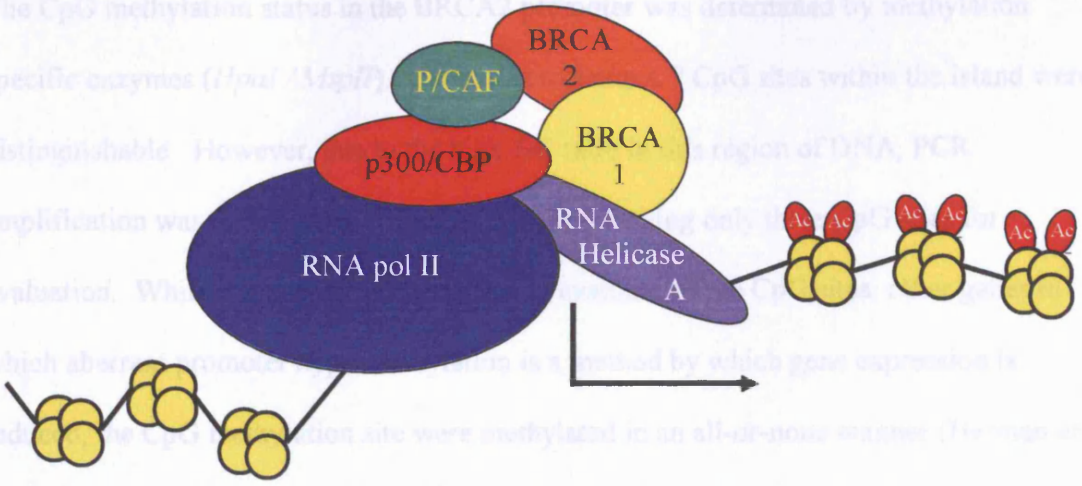

Figure 1.4 An illustration of how BRCA2 could interact with P/CAF to mediate the acetylation of histones to facilitate RNA pol II binding. (Taken from Welcsh *et al*, 2000).

1.8.4 Role of BRCA2 in sporadic breast cancer

BRCA2 mutations are thought to account for approximately 25% of familial breast cancer (Ford *et al*, 1998), however no mutations have been detected in sporadic breast cancer cases of (Nathanson *et al*, 2001). For certain tumour-suppressor genes alternative, epigenetic methods, of inactivation have been proposed. The methylation of normally unmethylated cytosine residues within the promoter region is one such method. This has been shown to be the case in the tumour-suppressor genes *RB1* (Ohtani-Fujita *et al*, 1993), *E-cadherin*, and *VHL* (Herman *et al*, 1994). To determine if BRCA2 was influenced by aberrant promoter hypermethylation Collins *et al* (1997) identified a CpG island close to the 5′ end of BRCA2 (Genbank accession number Z73360). This region conforms to the standards of a CpG island (Bird *et al* 1986) in that it has a combined G + C content greater than 60% and has an elevated CpG/GpC ratio. Within this CpG island, several transcription factor binding sites were identified (SP1, USF, AP2 and CP2), but there were no TATA or CAAT boxes observed.

The CpG methylation status in the BRCA2 promoter was determined by methylation specific enzymes (*HpaI / MspII*). Using this technique 7 CpG sites within the island were distinguishable. However, due to the high GC ratio in this region of DNA, PCR amplification was not possible at four of the sites, leaving only three CpG sites for evaluation. While it would have been ideal to evaluate all the CpG sites, other genes in which aberrant promoter hypermethylation is a method by which gene expression is reduced, the CpG methylation site were methylated in an all-or-none manner (Herman *et al,* 1994; Ohtani-Fujita *et al,* 1993; Yoshiura *et al,* 1995). Of the 64 sporadic breast cancers investigated, none showed methylation at any of the informative sites (Collins *et al,* 1997). These cases were unselected for age, in fact, no patient data was provided, and so it is possible that CpG island methylation is a factor in the sporadic breast cancers within a young cohort (<36 years of age) and warrants investigation.

1.9 Hypothesis & Aims

The aim of this thesis is to test the hypothesis that sporadic breast cancers in young women arise because of a profile of specific genetic alterations. A number of key questions will be investigated towards realising this aim:

1) What is the level of BRCA2 protein expression and localisation in young women (<35 yrs.) with sporadic breast cancer? Does this relate to LOH at BRCA2?
2) Can cDNA microarray be used to identify genomic copy number changes and specific gene expression profiles for sporadic breast cancers in young women (<38 yrs.)?
3) Do young women (<35 yrs) who develop sporadic breast cancer show skewed X-chromosome inactivation?

Chapter 2:

Materials & Methods

2.1 Materials

All chemical reagents were 'analar' grade purchased from Sigma, unless otherwise stated.

2.1.1 Breast Tissues

Breast tissue was available from breast cancer patients and women having elective breast reductive surgery (reduction mammoplasty) in the University Hospitals of Leicester NHS Trust. The breast cancer cases used are summarised in Table 2.1.

Consent was granted from all the patients in this study in accordance with the Ethical Policies of the University Hospitals Leicester NHS Trust LREC.

Younger Cases		
RW No.	Grade	Node Status
570	3	Neg
627	2	NK
734	2	Neg
784	3	Pos
810	3	Pos
862	3	Neg
939	3	Pos
1319	1	Neg
1376	2	Neg
1482	3	Pos
1496	3	Neg
1567	3	Neg
1579	2	Pos
1606	3	Neg

Older Cases		
RW No.	Grade	Node Status
1694	3	Neg
1936	3	Neg
2030	3	Neg
2062	3	Neg
2063	3	Neg
2074	3	Neg
2092	2	Neg
2093	3	Neg
2121	2	Pos
2123	3	Neg
2142	3	Pos
2146	2	Neg
2172	3	Neg
2182	3	Neg

1634	2	Pos
1639	3	Pos
1650	3	Pos
1686	3	Neg
1737	3	Neg
1752	2	Pos
1772	2	Neg
1820	3	Neg
1844	3	Neg
1857	3	Pos
1889	3	Pos
1932	3	Neg
1941	3	Neg
1962	2	Neg
2006	3	Pos
2019	3	Pos
2055	2	Pos
2157	3	Neg
2195	3	NK
2200	3	Pos
2204	3	Neg
2187	2	Neg
2188	3	Pos
2377	3	Pos
3044	3	Pos
3537	3	Neg

Table 2.1 A summary of the breast tumour samples used.

2.1.2 Cell Lines

All cell lines were originally purchased from the American Type of Culture Collection (ATCC, Rockville, Maryland). The cell lines used are shown in Table 2.2.

Cell Line	Description
MCF-7	Oestrogen receptor positive breast cancer cell line
T47D	Oestrogen receptor positive breast cancer cell line
ZR-75	Oestrogen receptor positive breast cancer cell line
MDA-MB-231	Oestrogen receptor negative breast cancer cell line
MDA-MB-486	Oestrogen receptor negative breast cancer cell line
MDA-MB-436	Oestrogen receptor negative breast cancer cell line
HBL-100	Oestrogen receptor negative, immortalised non-tumourigenic breast cell line

Table 2.2 A summary of the breast cell lines used.

Dulbecco/Vogt modified Eagle's minimal essential medium (DMEM) without phenol red: Sigma, D5921

Roswell Park Memorial Institute medium 1640 (RPMI) with phenol red: Sigma, R8758

L-glutamine: Sigma, G7513

2.1.3 Primary Antibodies

Three different primary anti-BRCA2 antibodies were used for both immunohistochemistry and western blot analysis. Two were purchased from BD PharMingen (catalogue numbers 69261a and 66066e), and the third was from Oncogene Research Products (catalogue

number PT146T). Whilst all had been previously used for western blot analysis, only antibody PT146T had been used for immunohistochemistry (Table 2.3).

Cat Number	66066e	69261a	PT146T
Manufacturer	BD PharMingen	BD PharMingen	Oncogene Research Products
Clone	Polyclonal	D13-2067	Polyclonal
Ag Binding Site	1324-1347	n/s	3245-3418
Positive Control	HBL-100	HBL-100	MCF-7
Negative control			CAPAN-1
Optimised Concentration			
Immunohistochemistry	1:100	1:100	1:100
Western Blot	1 µg/ml	1 µg/ml	2 µg/ml
Reference	Bernard-Gallon *et al*, 2000		Moll *et al*, 1999

Table 2.3 Primary anti-BRCA2 antibodies used in this study.

The Western blot analysis provided on the company data sheets are provided in Figure 2.1.

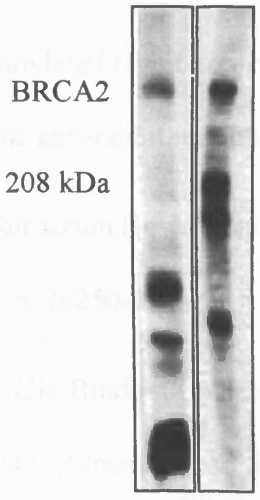

Figure 2.1 These are the western blots provided on the data sheets for the two anti-BRCA2 antibodies 66066e (left) and PT176T (right) showing BRCA2 and the smaller bands assumed to be BRCA2 degradation products.

The primary anti-RBBP4 antibody (Abcam/Gentrex, clone ab86) was a mouse monoclonal antibody generated from a fusion protein with amino acid residues 1-425 of human RBBP4. It was an IgG1 class antibody with a kappa light chain.

2.1.4 Immunohistochemistry: Buffers and reagents

Haematoxylin: Haematoxylin (0.1% (w/v), BDH, 340 374T), Aluminium potassium sulphate (105.4 mM, Fisher, A/2440/53), Citric acid (4.8 mM, Fissons, C/6200), Chloral hydrate (302.2 mM, Fisher, C/4280/53), Sodium iodate (0.93 mM, DBH, 30171)

Eosin: Eosin Yellowish (14.4 mM, DBH, 341973), Calcium Chloride (5×10^{-4} %), formaldehyde (3.9×10^{-2} %)

Tris buffered saline (TBS): Tris (50 mM, Roche, 708 976), Sodium Chloride (300 mM, Fisher, S/3160/63), pH 7.65

Proteinase K: Roche, 3-115-836

Citrate buffer: Citric Acid (10 mM, Fisher, C/6200/53), pH to 6.0

Secondary antibodies: Biotinylated rabbit anti-mouse immunoglobulin (DAKO Cytomation, E03554) and swine anti-rabbit immunoglobulin (DAKO Cytomation, E0353).

Normal serum: Normal rabbit serum (Invitrogen, 1612-099 and Gibco, 16120-099) and normal swine serum (Invitrogen, 26250-084)

Tertiary Detection: Streptavidin Binding Complex (St-ABC) conjugated with horseradish peroxidase (DAKO Cytomation) and St-ABC conjugated with alkaline phosphatase (DAKO Cytomation).

DAKO Duet Kit: Includes biotinylated goat anti-mouse/rabbit immunoglobulin and streptavidin binding complex (St-ABC) conjugated with horseradish peroxidase (K0492)

Hydrogen peroxide: Fisher, H/1650/17

3, 3'-Diaminobenzidine (DAB): DAB (500µl, D5637), TBS (9.5ml), H_2O_2 (100µl, 3% (v/v)).

NBT/BCIP Tablets: Roche, 1 697 471.

Mountant: DPX mountant (VWR, 3 6029 414) and Aquamount (VWR, 362263H) as appropriate.

2.1.5 Western Blot: Buffers and reagents

Phosphate buffered saline (PBS): Sodium Chloride (130 mM, Fisher, S/3160/63), Di-sodium hydrogen orthophosphate (Na_2HPO_4) (3 mM, , Fisher, S/4480/53), Sodium di-hydrogen orthophosphate (NaH_2PO_4) (7 mM, VWR, 10245).

Gold lysis buffer: Octylphenoxyethanol (Triton X-100) (1% (v/v), Fisher, T/3751/08), Tris (pH 8.0, 30 mM, Roche, 708 976), sodium chloride (137 mM, Fisher, S/3160/63), glycerol, (15%, Sigma, G-6279), EDTA, (5 mM, , Fisher , D/0700/53)

BioRad Quantification Reagent: BioRad, 500-0006

Polyacrylamide gels: 37.5:1 acrylamide/bis-acrylamide (National Diagnostics, EC-890)

Tris/Glycine Western Running Buffer: Tris (44.6 mM, Roche, 708 976), Glycine (95.9 mM), Sodium dodecylsulphate (1.73 M, Sigma, L-5750).

4x loading buffer: Tris (pH 8.6) (0.26 M, Roche, 708 976), Glycerol (40% (v/v), Sigma, G-6279), β-Mercaptoethanol (20% (v/v), VWR, 44142W), Sodium dodecylsulphate (0.27 M, Sigma, L-5750), Bromophenol blue (0.01% (w/v), Electran 44305)

Gel saver pipette tips: Fisher, FB 56207

SeeBlue 2: Invitrogen, LC 5925

Nitrocellulose membrane: BioRad, 162-0112

Wet transfer buffer: Tris (44.6 mM, Roche, 708 976), Glycine (14.3 mM) ± Methanol (10%, Fisher, M/4000/17)

Semi-dry transfer buffer: Tris (25 mM, Roche, 708 976), Glycine (192 mM) ± Methanol (40%, Fisher, M/4000/17)

TBS with Tween 20 (TBST): Tris, (50 mM, Roche, 708 976), Sodium Chloride (300 mM, Fisher, S/3160/63), Polyoxyethylene sorbitan mono laurate (Tween 20) (0.1% (v/v), Sigma, P-1379)

Secondary antibodies: HRP conjugated rabbit anti-mouse (Invitrogen, 11182) was used for monoclonal primary antibodies and HRP conjugated goat anti-rabbit (Invitrogen, 10885) for polyclonal primary antibodies.

Amersham Western ECL Detection Solution: Amersham, RPN3103K

2.1.6 DNA Digestion

HpaII: New England Biolabs, R0171S

HpaII Digestion Buffer: Bis-Tris-Propane-HCl (10 mM), $MgCl_2$ (10 mM), dithiothreitol (1 mM), pH 7.0

2.1.7 RNA Extraction

Tri Reagent (Sigma, T9424), a mixture of guanidine thiocyanate and phenol was used for RNA extraction. The RNA was precipitated with the addition of glycogen as a carrier (Invitrogen, 10814-010).

2.1.8 Reverse Transcription

AMV reverse transcriptase: Promega, M5101

AMV buffer: Tris-HCl (250mM, pH 8.3), KCl (250mM), $MgCl_2$ (50mM), dithiothreitol (50mM), Spermidine (2.5mM)

dNTP mix: Roche, 1 969 064

RNasin: Promega, N2115

Random hexamers: Promega, C1181

2.1.9 Semi Quantitative Polymerase Chain Reaction

Taq Polymerase: Promega, M166B

AJ buffer: Tris-HCl (250mM, pH 8.3), NH_4SO_4 (11mM), $MgCl_2$ (4.5mM), dNTP (200μM each), Bovine Serum Albumin (110μg/ml), β-mercaptoethanol (6.7mM), EDTA (4.4μM, pH 8.0)

2.1.10 Quantitative RT-PCR

The primers and probes for AKAP1 (Applied Biosystems, 2268) and CEBP/A (Applied Biosystems, 1426) were inventoried primer sets and so were amplified using the standard master mix (Applied Bio Systems, 4331182).

DDB2, GRN, RBBP4 and TGFBI were amplified using Platinum® SYBR® Green qPCR SuperMix UDG (Invitrogen, 11733-038). Along with the PCR master mix this included the ROX reference dye. Primer sequences are given in Table 2.4

2.1.11 Agarose Gel Electrophoresis

Agarose gels: Roche, 1 388 991

Tris Acetate EDTA (TAE) Buffer: Tris (40mM), Glacial Acetic Acid (20mM), EDTA (0.1mM) pH 8.0

Tris Borate EDTA (TBE) Buffer: Tris (90mM), Boric acid (90mM) EDTA (2mM) pH 8.3

Ethidium bromide: Made up to 10mg/ml stock in UP water (Sigma, E-1385)

100bp DNA ladder: Invitrogen, 15628-050

2.1.12 PCR Primer sequences

All primers were designed using the Primer3 website (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) and check for any secondary structures with NetPrimer (http://www.premierbiosoft.com/) using sequences taken from the NCBI database (http://www.ncbi.nlm.nih.gov/), with the exception of the BRCA2 and AR primers which were taken from the papers cited (Table 2.4).

Gene Symbol	Forward Primer	Reverse Primer	Amplicon Size (bp)	Annealing Temp (°C)	Cycles	Accession number	Reference
AR	TCCAGAATCTGTTCCAGAGCGTGC	GCTGTGAAGGTTGCTGTTCCTCAT	variable (≈ 100)	61.5	29	NM_000044	Buller, 1991
B2M	TACTCACGTCATCCAGCAGA	GAAAGACCAGTCCTTGCTGA	159	60	34	NM_004048	
BA	AAAGACCTGTACGCCAACAC	GATCCACACGGAGTACTTGC	153	62	34	NM_001101	
BRCA2	GTTGTGAAAAAAACAGGACTTG	CAGTCTTTAGTTGGGGTGGA	353	61	34	NM_000059	Bieche, 1999
DDB2	GCAAGCAGAGGTGGTGATT	AAAGTGTCCCAGTCCCACAG	100	60	n/a	NM_00107	
GAPDH	CATCATCCCTGCCTCTACTG	TTGGCAGGTTTTTCTAGACG	157	60	34	BC_001601	
GAPDH	AGAACATCATCCCTGCCTC	GCCAAATTCGTTGTCATACC	347	60	34	BC_001601	
GRN	CTAGCACCTCCCCCTAACCA	CTGACAGGGAAGGCCTTAGA	96	62	n/a	NM_002087	
RBBP4	TGGTGGTCATACTGCCAAGA	ATTTGCCACACTTGCATGAT	104	62	n/a	X_71810	
TGFBI	TGGACAGACCCTGGAAACTC	TGAACAGGGTCCCGTACCT	125	60	n/a	NM_000358	

Table 2.4 This table summaries the (RT-) PCR primers used. Those marked "n/a" in the Cycles column did not require a sub-saturation end point to be determined as they were used for quantitative RT-PCR

2.1.13 BioAnalysis

RNA 6000 Nano Chip: Ambion, 5065

RNA 6000 Ladder: Ambion, 4476

2.1.14 Affymetrix GeneChip®

SuperScript TM reverse transcription kit: Invitrogen, 11917-010

Enzo® BioArray TM High Yield TM RNA Transcription Labelling Kit:

Affymetrix, 900182

2.2 Introduction to Methods

2.2.1 Immunohistochemistry

Albert H. Coons and his colleagues (1950) were the first to label antibodies with a fluorescent dye, and use these to identify antigens in tissue sections. With the expansion and development of immunohistochemistry, enzyme labels have been introduced such as peroxidase (Nakane and Pierce 1966; Avrameas and Uriel 1966) and alkaline phosphatase (Mason and Sammons 1978). Colloidal gold (Faulk and Taylor 1971) can be used for immunohistochemical reactions at both the light and electron microscopic level. Other labels include radioactive elements, and the immunoreaction can be visualized by autoradiography.

Most of the immunohistochemical staining methods in use today are based on the high affinity ($K = 10^{-15}$ M) that streptavidin (Streptomyces avidinii) and avidin (chicken egg) have for biotin. Both possess four binding sites for biotin, but due to the molecular orientation of biotin, fewer than four molecules will actually bind. Avidin is a

glycoprotein and has an isoelectric point of 10, it has a propensity to non-specific binding to lectin-like and negatively charged tissue components at physiological pH. It has largely been replaced by streptavidin.

The inherent amplification of sensitivity made the avidin- and streptavidin-biotin methods more desirable than previous methods (such as PAP and APAAP). The basic sequence of reagent applications consists of primary antibody, biotinylated secondary antibody, followed by the reformed streptavidin-biotin-enzyme complex. Horseradish peroxidase and alkaline phosphatase are the most commonly used enzyme labels. The basic structure of antigen-antibody-ABC binding is stylised in Figure 2.2.

Figure 2.2 This shows a stylised structure of the binding that takes place during an immunohistochemical reaction, also showing the degree of signal amplification possible. Where, the light blue square is the antigen, the red antibody is the primary, and the blue antibody is the secondary antibody. Green triangles represent biotin with HRP conjugated (red dots) and the blue cross represents (strept)avidin. The slide and tissue sample are represented by the black and red crossed bars respectively.

2.2.2 Formalin Fixed Paraffin Embedded Tissue

Formalin fixation is the most commonly used method of tissue fixation, using 40% (w/v) formaldehyde in water. This is usually a neutral salt tonicity with a buffering system to

maintain the pH. Formalin based fixatives are well tolerated by most tissues and have good penetration. Unfortunately there may be some shrinkage or distortion during the fixation or subsequent paraffin-embedding but generally these fixatives are appropriate for most immunohistochemistry (Boenisch *et al*, 2001).

Formaldehyde fixes by reacting with basic amino acids to form cross-linking methylene bridges. This results in a low permeability to macromolecules while not significantly altering the intracellular proteins. Small (10 x 10 x 3 mm) tissue pieces promptly fixed in neutral buffered formalin for 6-12 hours will generally show good cytological preservation and immunolocalisation, with only the minimal of antigen retrieval (Boenisch *et al*, 2001).

Although some antigens are not well demonstrated after formalin fixation, most can be demonstrated after the appropriate pre-treatment such as enzymic protein digestion and/or antigen retrieval. If antibodies are to be used on formalin-fixed paraffin embedded (FFPE) tissue sections, there are three considerations to be taken into account. Firstly, does the formaldehyde react with the epitope under investigation. Secondly, does it react with adjacent amino acids causing conformational changes, and thirdly does processing destroy the epitope under investigation. If there are any conformational changes as a result of the formaldehyde, these can often be reversed by antigen retrieval or proteolytic digestion, any conformational changes that result from the processing of the tissue are irreversible (Boenisch *et al*, 2001).

Conformational changes that destroy epitopes or alter them to reduce the reactivity with the antibody can occur for a number of reasons. These are most commonly due to chemical fixation or heat when processing. Many epitopes are heat sensitive and during the paraffin embedding process the tissue is heated to the melting point of the paraffin wax, typically 50-60°C. At 60°C some epitopes have a half life of only 10-15 minutes (Boenisch *et al*, 2001).

In order to achieve consistent and reproducible immunohistochemistry it is vital that the fixation and embedding processes are standardised in order to reduce epitope masking and loss variability. As there is no standard duration of formalin fixation, when comparing results between laboratories it is important that the methods for producing FFPE tissue sections are comparable (Boenisch *et al*, 2001).

Until the mid-1980s it was uncertain whether DNA could be extracted from formalin fixed paraffin embedded tissue. At this time, two independent studies demonstrated that DNA could be extracted from this tissue and whilst the DNA was partially degraded (with a size range of 100-1500 bp) the DNA was double stranded and suitable for cleavage by restriction enzymes (Dubeau *et al* 1986).

2.2.3 PAGE & Western Blotting

Western blotting is one of the most commonly used protein analysis assays used as it provides a direct method for identifying, monitoring and relative quantification specific proteins. The general technique for western blotting can be adapted for use with SDS-polyacrylamide Gel electrophoresis (PAGE) (which fractionates by mass), native PAGE (which fractionates by charge/mass), isoelectric focusing (IEF) (which fractionates by pI) or 2D-PAGE (which serially fractionates by these two methods) (Alberts *et al*, 1994).

Due to the variation in western blotting objectives there are slight variants to the method, for example whole cell lysates or subcellular fractions could be run on the gel. In general, a complex protein mixture (such as a cell lysate, cell extract or purified protein preparation) is fractionated by gel electrophoresis. After separation, the proteins are transferred onto a solid support, such as nitrocellulose, polyvinylidene fluoride (PCDF) or nylon. The protein under investigation is identified using a specific primary antibody. A band is visualised using an appropriate enzyme linked secondary antibody and a electro-

chemiluminescent substrate then exposing the membrane to X-ray film. This is summarised in Figure 2.3 (Mathews and van Holde, 1996).

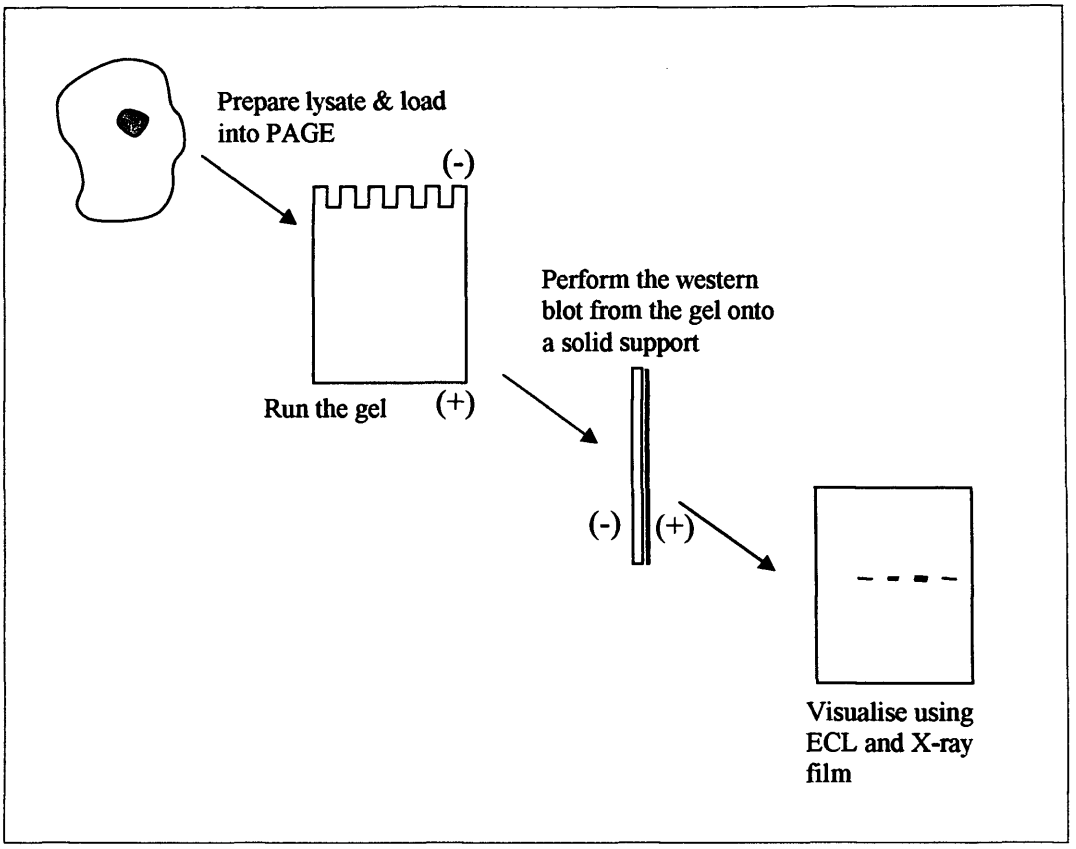

Figure 2.3 Summary of the process of PAGE and western blotting.

2.2.4 Polymerase Chain Reaction (PCR)

The PCR process amplifies specific DNA sequences from a longer DNA molecule. A typical reaction mixture includes the template DNA, a thermostable DNA polymerase, two oligonucleotide primers, deoxynucleotide triphosphates (dNTPs), reaction buffer, magnesium as wells as additional additives. The reaction process is facilitated by a thermalcycler, an automated instrument that takes the reaction mixture though a series of temperatures, then cycles through the temperature changes. Each PCR cycle should theoretically double the amount of the targeted template (amplicon) in the reaction (Saiki *et al*, 1985).

Each cycle of the PCR amplification consists of three steps. The first of these denatures the double stranded DNA template by heating to 94 °C. The second step reduces the temperature to prepare the DNA for amplification by allowing the specific annealing of the primers to the template. The temperature is dependent on the primer sequence, but usually falls between 50-70°C. Finally, the polymerisation reaction is catalysed by the thermostable DNA polymerase. This is achieved by increasing the temperature to 72°C. A schematic of this is shown in Figure 2.4 (Stryer, 1995).

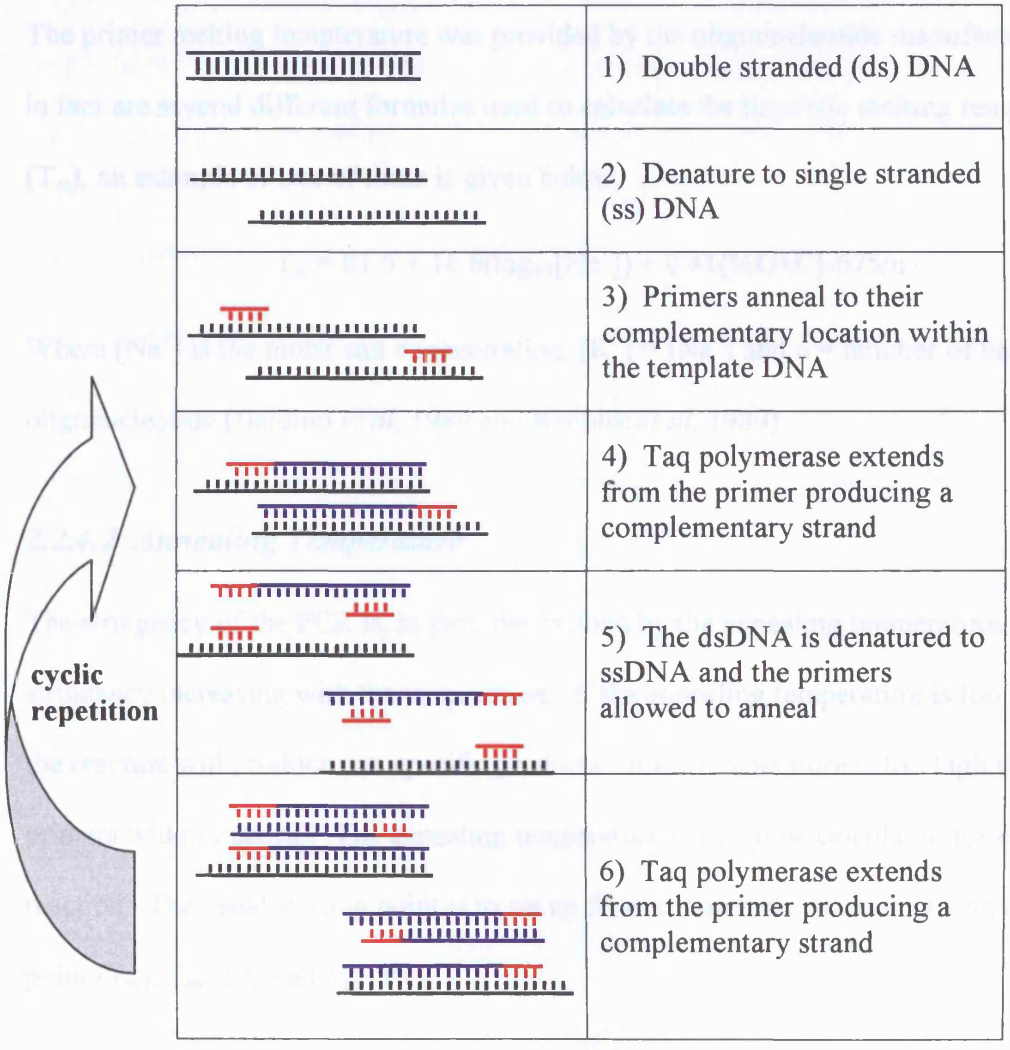

Figure 2.4 Summary of the PCR cyclic process

2.2.4.1 Primer Design

PCR primers generally range in length from 15-30 bases and are designed to flank the region of interest. Primers should have a 40-50% G/C content and care should be taken to avoid sequence which produce internal secondary structures. The 3′ end of the primers should not be complementary to avoid primer dimers. Also, runs of three G or C nucleotides should be avoided near the 3′. Ideally the primers should anneal at the same temperature, which is determined by the primer melting temperature.

The primer melting tempterature was provided by the oligonucleotide manufacturer. There in fact are several different formulae used to calculate the theoretic melting temperature (T_m), an example of one of these is given below:

$$T_m = 81.5 + 16.6(\log_{10}[Na^+]) + 0.41(\%G+C)-675/n$$

Where $[Na^+]$ is the molar salt concentration; $[K^+] = [Na^+]$ and n = number of bases in the oligonucleotide (Baldino *et al*, 1989 and Rychlik *et al*, 1989).

2.2.4.2 Annealing Temperature

The stringency of the PCR is, in part, determined by the annealing temperature, the stringency increasing with the temperature. If the annealing temperature is too low then the reaction will produce non specific products. If the temperature is too high then the primers will not anneal. The annealing temperature needs to be calculated for each reaction. The usual starting point is to set up three comparable reactions using the melting point (T_m), T_m -2°C and T_m -5°C.

2.2.4.3 Cycle Times

In order to prevent non-specific product formation (and to reduce the reaction time) the time for each stage of the reaction should be kept to a minimum whilst allowing enough

time for the reaction to take place. The usual standard is to allow 30 seconds for annealing, and 1 min/product kb for the denaturation and extension step.

2.2.4.4 Magnesium Concentration

The magnesium concentration is a crucial factor affecting the performance of Taq polymerase. Reaction components, including the template DNA, chelating agents present in the sample (e.g. EDTA or citrate), dNTPs and proteins, can affect the amount of free magnesium. Taq polymerase is inactive without magnesium, however excessive magnesium reduces enzyme fidelity and may increase the level of non-specific amplification (Williams *et al*, 1989; Eckert and Kunkel; 1990; Ellsworth *et al*, 1993)

2.2.4.5 Reverse Transcription (RT)- PCR

Of the numerous methods developed to measure gene expression in tissues and cells RT-PCR is by far the most sensitive and versatile. In its crudest form it can determine the presence or absence of mRNA transcripts. In its most refined, quantitative RT-PCR (qRT-PCR) can quantify the number of transcripts per cell.

As originally described, RT-PCR employed either avian myeloblastosis virus (AMV) or Moloney murine leukaemia virus (MMLV or MuLV) to reverse transcribe for first strand cDNA. The second strand synthesis and amplification is carried out using Taq polymerase in a standard PCR reaction (Murakawa *et al*; 1988, Wang *et al*, 1992; Aatsinki *et al*; 1994).

2.2.5 Quantitative RT-PCR (qRT-PCR)

Real time PCR allow the detection of PCR amplification products during the early phases of the PCR reaction, whereas traditional PCR detects the PCR products at the end of the PCR reaction. End point detection is limited by poor precision, low sensitivity, low

resolution and offers size based discrimination only (http://www.appliedbiosystems.com/support/).

The traditional PCR reaction follows a sigmoidal relationship between the number of copies and cycle number. During the exponential phase there is an exact doubling of PCR products per PCR cycle (assuming 100% efficiency). In the linear phase the reaction slows down as reaction components are being consumed. This phase shows high variability between reactions. The plateaux phase reflects that the reaction stops. No more products are made through a combination of reaction component depletion and product/product dimerisation out competing product/primer dimerisation (http://www.appliedbiosystems.com/support/).

Even the most sensitive qPCR system is not sensitive enough to detect products during the early cycles of a PCR reaction, the first detectable products being during the linear phase of the reaction. This is also the most variable and therefore the most informative phase. Most qPCR methodologies therefore quantify a reaction by its Ct value, which is the cycle number at which PCR products are first detectable. So reactions with a higher starting template concentration have a lower Ct value (Figure 2.5).

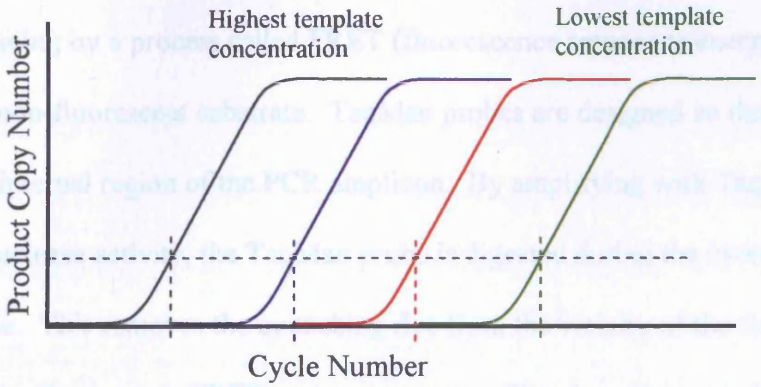

Figure 2.5 This shows how the linear phase of the PCR reaction is the most informative, i.e. a small change in the cycle number results in a large change in the copy number. With the horizontal dashed line representing the point at which PCR products are first detectable, the vertical coloured dash lines would give the Ct value for each of the reactions. So the reaction with the highest initial starting concentration of template is detectable first and so has the lowest Ct value (black line). Conversely the reaction with the lowest initial template concentration takes longer to be detectable and has the largest Ct value (green).

2.2.5.1 Quantitative PCR Chemistries

SYBR® Green

SYBR Green is the most inexpensive of all the chemistries used for real time studies but is also the least specific, binding to the minor groove of all double stranded DNA. Once bound to the double stranded DNA the dye is able to fluoresce when irradiated. As more double stranded DNA is produced the SYBR Green dye signal increases proportionally. However, as the binding is non-specific, it will bind and fluoresce to primer dimers and non-specific amplification products, which would skew Ct values (http://www.appliedbiosystems.com/support/).

TaqMan® Probes

TaqMan probes are oligonucleotides that contain a fluorescent dye (usually towards to the 5′ end) and a quenching dye (typically at the 3′ end). When the two are in close proximity

and the fluorescent dye is irradiated it transfers energy to the quencher dye rather than inducing fluorescing by a process called FRET (fluorescence resonance energy transfer), so producing a non-fluorescent substrate. TaqMan probes are designed so that they hybridise to an internal region of the PCR amplicon. By amplifying with Taq polymerase with a 5′ endonuclease activity, the TaqMan probe is digested during the extension phase of the PCR cycle. This removes the quenching dye from the vicinity of the fluorescent dye and so it is able to fluoresce as FRET no longer occurs. The rate of increased fluorescence is proportional to the rate of probe cleavage. As the probe binds to the amplicon, this adds an additional layer of product specificity to the quantification (http://www.appliedbiosystems.com/support/).

Molecular Beacons

Similar to TaqMan Probes, Molecular Beacons also utilise FRET. However, in addition to an intra-amplicon banding region they also contain a hairpin loop. This ensures that in free solution the quenching dye is in close proximity to the fluorescent dye. When the probes bind to the amplicons they unfold, moving the quencher from the fluorescent dye and producing a signal. As the reaction progresses the number of amplicons increase allowing more amplicon/probe hybridisation. The signal produced is therefore proportional to the total amount of product (Antony *et al*, 2001, Abravaya *et al*, 2003, http://www.molecular-beacons.org/).

2.2.6 Agilent BioAnalysis of RNA

In order to determine the quality of RNA extracted from cryo-preserved tissue, an aliquot was run on an Agilent RNA 6000 Nano LabChip using an Agilent BioAnalyser 2100. This has become a standard method for the determination of RNA quality. The system requires

as little as 5ng of total RNA to determine the quality of the RNA using capillary electrophoresis.

In order to ensure reproducibility the kit incorporates a marker in every sample and a marker and ladder well. The features of a successful ladder are 5 RNA peaks, one marker peak, and that all the peaks are well resolved. An example of this is shown in Figure 2.6.

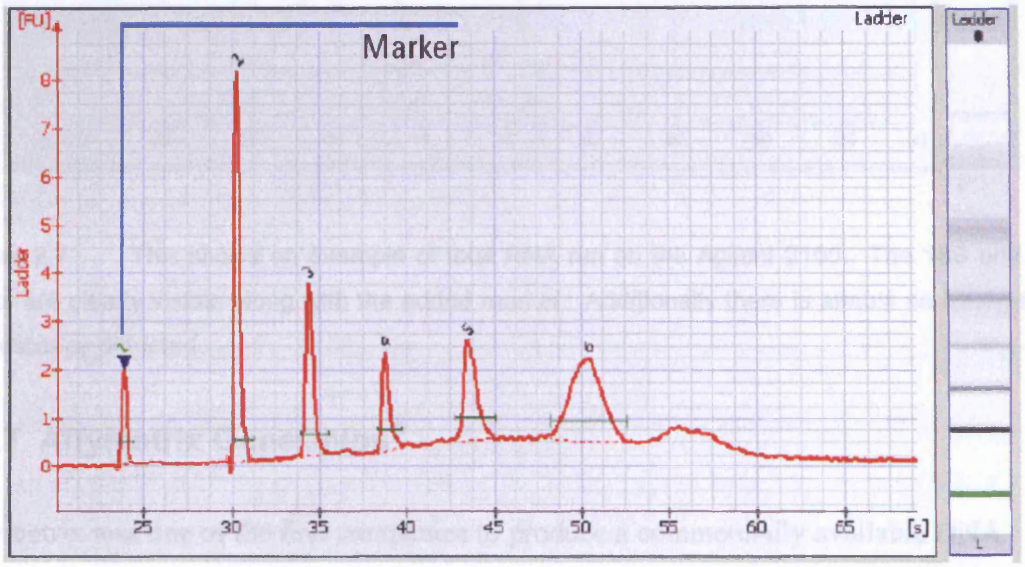

Figure 2.6. An example of how the Nano 6000 ladder runs on the Agilent BioAnalyser 2100. This is a perfect example of a ladder, there are 5 RNA peaks, the marker peak and all 6 peaks are well resolved.

A successful electrogram for a total RNA extraction (from eukaryotic cells) should show three clear peaks. These correspond to the marker and the 18S and 28S rRNA peaks, with low background fluorescence. Additionally the total florescence (that is the area under the curve) for the 28S peak should be double that of the 18S peak. See Figure 2.7.

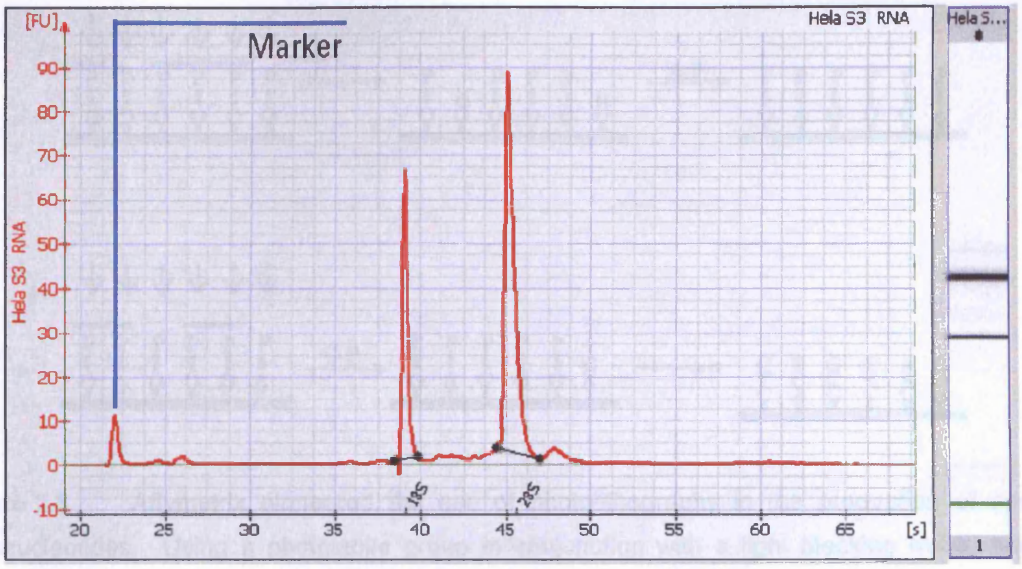

Figure 2.7 This shows an example of total RNA run on the Agilent 2100. The 18S and 28S peaks are clearly visible along with the added marker. Additionally there is almost no background fluorescence detected.

2.2.7 Affymetrix GeneChips®

Affymetrix was one of the first companies to produce a commercially available DNA microarray (http://www.affymetrix.com/). The arrays are produced in situ onto a flat solid support. Affymetrix pioneered the use of photolithographic masking and a photochemically removable group to enable a matrix of different oligonucleotide sequences to be produced. This cyclic process is illustrated in Figure 2.8.

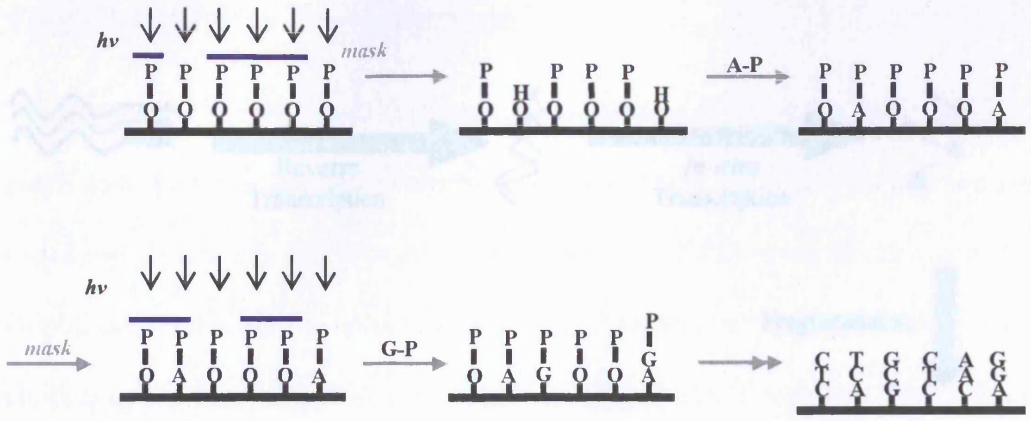

Figure 2.8 Affymetrix pioneered the use of photolithography in the production of specific oligonucleotides. Using a photolabile group in conjunction with a light blocking mask, specific dNTPs can be added to a growing oligonucleotide.

Total RNA is reverse transcribed using an oligo-dT primer containing a T7 polymerase site. Amplification and labelling of the sample is achieved by carrying out an in vitro transcription reaction in the presence of a biotinylated dNTP. This results in a cDNA population with amplification approximately 30-100 fold. The biotin labelled cRNA probes for the test and control samples are hybridised onto separate arrays. Binding of a streptavidin-conjugated fluorescent marker allows the array to be read using a scanning confocal laser. The different signals produced by the test and control array can then be compared (Harkin *et al*, 2000).

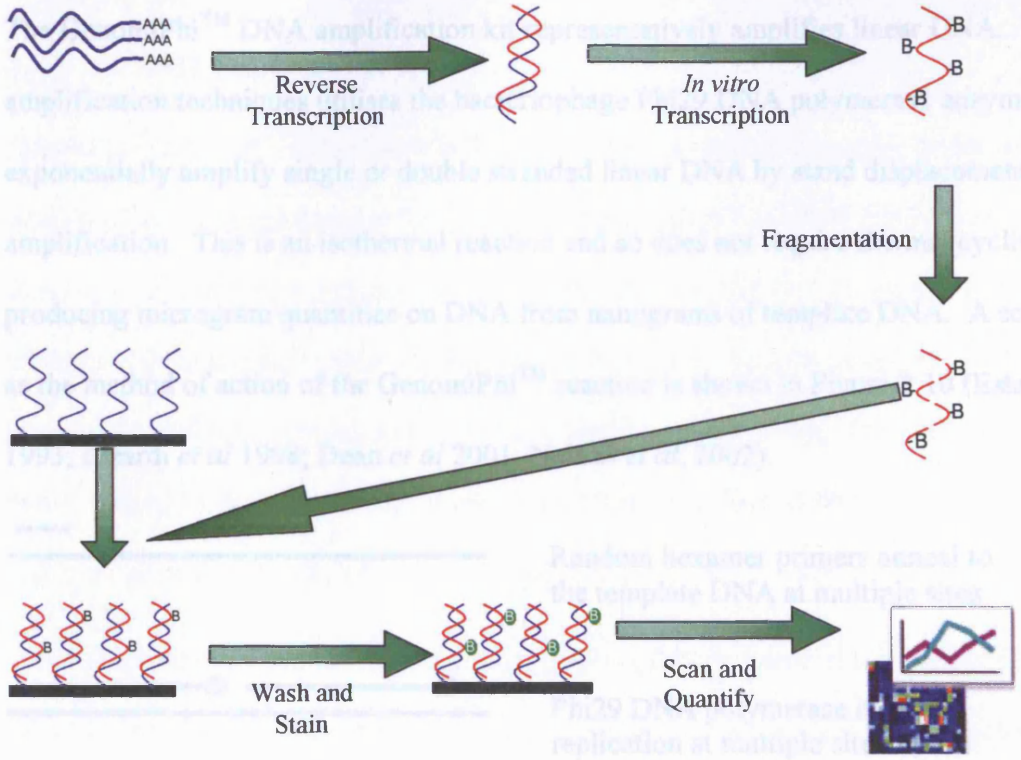

Figure 2.9. This shows a schematic for the labelling and hybridisation of RNA for use on Affymetrix GeneChips®.

There are many different applications for microarray technology. This can vary from a high throughput sequence analysis study as employed by researches on the human genome research project to RNA expression studies (Hedenfalk *et al*, 2001; Perou *et al*, 2000) . Tumour profiling studies aim to identify transcriptional changes as the cellular phenotype moves from that of a normal to tumour cell. An example of tumour profiling comes from Alizadeh *et al* (2000) who have been able to identify a distinct subclass of diffuse B-cell lymphoma that was not distinguishable by conventional methods.

2.2.8 GenomiPhi™ amplification

The GenomiPhi™ DNA amplification kit representatively amplifies linear DNA. The amplification techniques utilises the bacteriophage Phi29 DNA polymerase enzyme to exponentially amplify single or double stranded linear DNA by stand displacement amplification. This is an isothermal reaction and so does not require thermal cycling, producing microgram quantities on DNA from nanograms of template DNA. A schematic as the method of action of the GenomiPhi™ reaction is shown in Figure 2.10 (Estaban *et al* 1993; Lizardi *et al* 1998; Dean *et al* 2001; Nelson *et al*, 2002).

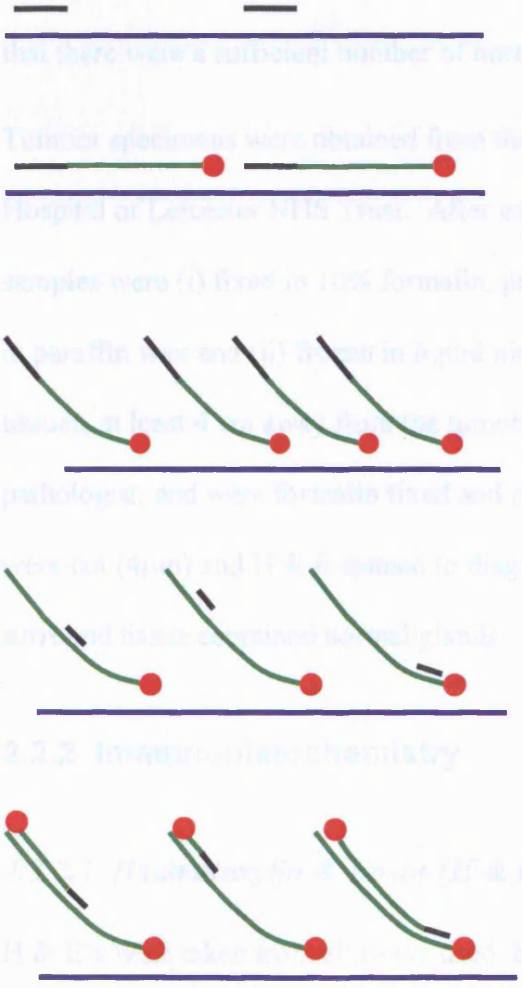

Figure 2.10. This schematic shows the method of action for GenomiPhi™ reaction.

2.3 Methods

2.3.1 Breast tissues

Normal breast tissue was available from elective breast reductive surgery (reduction mammoplasty) from patients in the University Hospitals of Leicester NHS Trust. The samples were taken by a pathologist and (i) fixed in 10% formalin, processed through graded alcohols and embedded in paraffin wax following standard protocols, and/or (ii) snap frozen in liquid nitrogen and stored in a liquid nitrogen fridge. Prior to use of the tissue, a 4μm section was cut and stained with Haematoxylin and Eosin (H & E) to ensure that there were a sufficient number of normal lobules within the section.

Tumour specimens were obtained from the pathology department of the University Hospital of Leicester NHS Trust. After examination by a pathologist, the dissected tumour samples were (i) fixed in 10% formalin, processed through graded alcohols and embedded in paraffin wax and (ii) frozen in liquid nitrogen as above. Uninvolved surrounding tissues, at least 4 cm away from the tumour, was also taken from the specimens by a pathologist, and were formalin fixed and paraffin embedded and snap frozen. Sections were cut (4μm) and H & E stained to diagnose the tumour type and to ensure that the surround tissue contained normal glands

2.2.2 Immunohistochemistry

2.3.2.1 Haematoxylin & Eosin (H & E) Staining

H & E's were taken from all tissue used, both frozen and FFPE. The FFPE sections were cut at 4μm, and then de-waxed in xylene (5min) and rehydrated by equilibrating the sections in decreasing concentrations of IMS (99%, 99%, 95%) for 2 min each followed by 5 min equilibration in running tap water. The sections were then stained in haematoxylin

for 3 min, after which they were washed in running tap water for 5 minutes. The sections were then immersed in eosin for 2 minutes, again washing in tap water afterwards. The sections were then dehydrated through increasing concentrations of IMS, and cleared in xylene before being mounted with DPX mountant.

After fixation in alcohol (95%, 5 seconds), frozen sections were H & E stained by immersing in haematoxylin for 20 seconds, after which they were washed in running tap water. The sections were then immersed in eosin for 2 seconds, and quickly dehydrated to xylene before being mounted with DPX.

2.3.2.2 Proteinase K digestion antigen retrieval with DAKO Duet detection

Sections were de-waxed in xylene (5 min), and then rehydrated by equilibrating the sections in decreasing concentrations of IMS (99%, 99%, 95%) for 2 min each followed by 5 min equilibration in running tap water. The sections were incubated in Proteinase K (5 or 10μg/ml) at 37 °C for 1 hour. Following this antigen retrieval, each section was rinsed with running tap water and equilibrated in TBS for 5 min Endogenous peroxidase activity was then blocked by immersing in 2% hydrogen peroxide (30 min), after which the sections were equilibrated in TBS. To block non-specific binding the sections were covered in normal goat serum (1:5 diluted in TBS, 30 min). The negative controls were left in goat serum, whereas the test slides had primary antibody applied and were incubated over night at 4 °C.

After this incubation, the slides were washed for 5 min in TBS. The Dako Duet kit was used, following the manufacturers instructions. In summary, the sections were incubated in a biotinylated goat anti-mouse/rabbit antibody (1:100, 30 min), after which were washed in TBS. The tertiary, prepared 30 min in advance was applied for 30 min at a dilution of 1:1000. This tertiary and final amplification step is labelled with a horseradish peroxidase,

allowing chromogenic detection with DAB. The DAB was applied to the sections for at least 5 min after which the DAB was washed off with water. The sections were counterstained in haematoxylin (20 sec) and washed in running tap water for at least 5 min. The sections could then be dehydrated in increasing concentrations of IMS through to xylene. The sections were mounted with a glass cover slip using DPX mountant.

2.3.2.3 Pressure Cooking antigen retrieval with DAKO Duet detection

The sections were de-waxed, rehydrated, and equilibrated in TBS, as above. Using a pressure cooker (Prestige, 6189), the slides were boiled in citrate buffer (≈126 °C) for 2, 3 or 5 min at full pressure. The sections were quickly cooled in running tap water and equilibrated in TBS. Endogenous peroxidase activity was eliminated by immersing in hydrogen peroxide (2%, 30 min). After washing in TBS (5 min) non-specific binding was blocked by incubating the sections in normal goat serum (1:5, diluted in TBS) for 30 min then the sections were incubated in the primary antibody (4 °C, over night).

The subsequent detection was performed using the Dako duet kit following the manufacturers guidelines, see section 2.3.2.2.

2.3.2.4 Pressure Cooking Antigen Retrieval with Strep ABC/Alkaline Phosphatase Detection

As previously described in section 2.3.2.2, the sections were de-waxed, rehydrated, and equilibrated in TBS. Using a pressure cooker, the slides were boiled in citrate buffer (126 °C) for 5 min at full pressure. The sections were quickly cooled in running tap water and equilibrated in TBS. After washing in TBS (5 min) non-specific binding was blocked by incubating the sections in normal serum (1:5, diluted in TBS) for at least 30 min. When the primary antibody was monoclonal, normal rabbit serum was used; when the primary

was polyclonal, normal swine serum was used. After which the sections were incubated in the primary antibody (4 °C, over night).

After this incubation, the slides were washed for 5 min in TBS. The secondary antibody was then be applied to the sections. The secondary used depended on the primary antibody used. If the primary was monoclonal (therefore raised in mouse) a biotinylated rabbit anti-mouse (1:400) was used, however if the primary was a polyclonal antibody (raised in a rabbit) a biotinylated swine anti-rabbit (1:600) secondary was used. The secondary was incubated for 30 min at room temperature, after which each section was washed in TBS (5 min). During this incubation the tertiary amplification step was prepared. This was an ABC (steptavadin binding complex) with a conjugated alkaline phosphatase, allowing the complex to form. After the 30 min incubation the secondary was washed off with TBS and the tertiary (ABC/AP) was applied for 30 min. After this period the sections were washed with TBS and the chromogenic reagent NBT/BCIP was applied. The time taken for the NBT/BCIP to develop is quite variable, ranging from 10 – 40 min. The sections should therefore be inspected microscopically in order to determine when to stop the reaction. As the product produced is soluble in organic solutions, the alkaline phosphatase reaction is stopped by washing in water and the sections were mounted under coverslips using aquamount.

2.3.2.5 Immunohistochemistry for Retinoblastoma Binding Protein 4 (RBBP4)

As previously described in section 2.3.2.2, the sections were de-waxed, rehydrated, and equilibrated in TBS. Antigen retrieval was achieved by microwaving the sections, fully submerged in citrate buffer for 20 minutes (750W), after which the sections were allowed to stand in the buffer for 30 minutes. The sections were then washed in running tap water for 10 minutes before blocking endogenous peroxidase activity by immersing in 3% (v/v)

hydrogen peroxide for 10 minutes. After washing in running water, and incubating in normal rabbit serum at a 1:5 dilution (10 minutes, room temperature). The primary RBBP4 antibody was applied to the sections for 1 hour at room temperature (1:2000).

The primary was washed off with 2 x 5 minute washes in TBS and the secondary antibody applied (Rabbit anti-mouse biotinylated) for 30 min at room temperature. This was removed with 2 x 5 minute washes in TBS, after which HRP conjugated ABC was incubated for 30 minutes at a dilution of 1:1000. This was washed in TBS, and visualised with DAB (9.5 ml TBS, 500 µl DAB 100µl 3% (v/v) H_2O_2). Excess DAB was removed by washing in running water. The sections were counterstained with Mayers Haematoxylin, before washing in running water. The sections were then dehydrated, cleared and mounted in DPX.

2.3.3 Cell Culture

The cell lines HBL-100, MCF-7, T47D, MDA-MB-231 and MDA-MB-468 were cultured in Dulbecco/Vogt modified Eagle's minimal essential medium (DMEM) without phenol red (D5921). This was supplemented 5mM L-glutamine (G7513) and 10% foetal calf serum (FCS). ZR-75-1 cells were cultured in Roswell Park Memorial Institute medium 1640 (RPMI) with phenol red (R8758) supplemented with 5mM L-glutamine and 10% FCS. All the cell culture was performed at 37°C in air supplemented with 5% CO_2.

2.3.4 Western Blot

2.3.4.1 Protein Extraction and Quantification

Cells were grown to subconfluence in 125 cm^2 tissue culture flasks. The media was removed and cells washed in sterile PBS. The flasks were then placed on ice. Gold lysis buffer (500 µl) was added to the cells, and using a sterile cell scraper the cells were scraped

from the flask. Once transferred to an eppendorf they were lysed by passing through a 25G syringe. This was then sedimented by centrifugation (7,500g, 5 min). The supernatant contained the soluble proteins and the pellet contained the membrane bound proteins. To avoid the effects of repeat freeze/thawing the protein solutions were aliquoted and snap frozen in liquid nitrogen. One aliquot was retained for quantification. This was achieved using BioRad's Protein Quantification Reagent, a modification of the Bradford Assay, which was used following the manufacturers instructions. This ensured that exactly 90 μg of total protein lysate was loaded into each lane.

2.3.4.2 Separating Proteins by Polyacrylamide Gel Electrophoresis (PAGE)

Polyacrylamide gels 1.5 mm in thickness were made with a concentration ranging from 5 to 20% in the separation gel and 3 to 10% in the stacking gel. Before loading the proteins were denatured by boiling in loading buffer (95-99°C, 10 min), then quickly chilled on ice. Using a gel saver pipette tip the protein solution was loaded into the polyacrylamide gel. A pre-stained protein ladder was used. The gels were run at 100V for between 2 and 5 hours or until the coloured 148 kDa marker had almost run off the end of the gel, depending on the density of the acrylamide gel.

2.3.4.3 Transfer to a Nitrocellulose Membrane

Prior to transfer the gel, membrane and blotting paper were washed in transfer buffer for 15 min. The transfer apparatus used was a BioRad semi-dry and wet blotter. Both methods use an electric current to transfer the proteins from the gel, on the cathode side, to the nitrocellulose membrane, on the anode side, of the apparatus.

When using the Semi-Dry blotter, 3MM blotting paper soaked in semi-dry transfer buffer was placed either side of the gel and membrane. To fix the proteins on the membrane, the

anode transfer buffer included methanol. To ensure even transfer care was taken to remove all air bubbles, and current was applied at 5mA/cm^2. This process took up to 3 hours.

The wet blotter was assembled in a similar manner but was fully submerged in the transfer buffer. The blotter was run overnight at room temperature at 30V.

After transfer, the membranes were wrapped stained in Ponceau's solution to confirm protein transfer then wrapped in Saran barrier wrap and stored at 4°C.

2.3.4.4 Detection of Proteins

Non-specific binding was blocked by washing the membrane in a solution of milk proteins (30 min, 5% Marvel dissolved in TBS with Tween 20 (TBST)). Following blocking, the membrane was rolled into a 15 ml centrifuge tube containing 1 ml of the primary antibody diluted in TBST. This was then allowed to incubate overnight at 4°C. Following incubation, the membrane was washed three times in TBST (10 min and then the membrane was incubated in the appropriate secondary antibody, either HRP conjugated rabbit anti-mouse for monoclonal antibodies or HRP conjugated goat anti-rabbit for polyclonal antibodies. Incubation was at room temperature for 1 hour, following which the membrane was washed three times in TBST (10 min). The membrane was covered in Amersham Western ECL Detection Solution for 5 min at room temperature. Excess liquid was removed and the membrane was wrapped in Saran barrier wrap. The membrane was then exposed to X-ray film for 3 min to 3 hours, and the film was developed using AGFA Curix 60 film developer.

2.3.5 DNA extraction

DNA was extracted from formalin fixed, paraffin embedded tissue. Sections were cut (10x 4 μm), de-waxed and re-hydrated by washing in xylene twice and then decreasing concentrations of IMS (99% x 2 and 95%). The tissue was suspended in 300 μl of a Tris/SDS solution (0.05 M Tris, 0.1% SDS) containing Proteinase K (1 mg/ml) at 60°C overnight.

To this an equal volume of phenol/chloroform/IAA (25:24:1) was added mixed, and centrifuged at 10,000g for 3 minutes. The top aqueous layer was removed and added to an equal volume of chloroform/IAA (24:1). This was mixed well, centrifuged and the top aqueous layer again removed.

To precipitate the DNA, 3 volumes of cold (-20°C) ethanol and 1/10 volume sodium chloride (1 M) was added and stored at -20°C for at least 1 hour. After centrifugation at high speed the precipitated DNA was pelleted, and washed in 70% ethanol.

Once the DNA pellet had been dried, it was resuspended in 50 μl DNase/RNase free water. The DNA solution was quantified using spectrophotometry (see Section 2.sss) and stored at 4°C.

2.3.6 Assessing DNA quality and quantity

In order to qualitatively assess the quality, i.e. the presence or not of fragmentation of the extracted DNA it was run on a low percentage (0.8%) TAE agarose gel. High quality DNA has a high molecular weight. This is demonstrated on the TAE gel by a low mobility. DNA that is of lower quality (or had been previously digested) shows a smear on a TAE gel because DNA fragments have a variety of sizes and therefore varying mobility.

DNA purity was assessed using standard photometric techniques, where the A260/A280 ratio should be greater than 1.8.

DNA was also quantified using spectrophotometry, using the formula

$$\frac{A260 \times dist \times dilution \times k}{1000} = conc(\mu g / ml)$$

Where: $A260$ is the optical density at 260 nm,

dist is the light path distance in cm

dilution is the dilution factor of the solution

k is a constant (50 for double stranded DNA and 40 for single stranded DNA and RNA)

2.3.7 DNA Restriction

Following the manufactures guidelines, 5µg of high molecular weight DNA was diluted with the manufacturers reaction buffer, sterile water and 1U of HpaII to a volume of 50µl. In a negative control, the HpaII was replaced with sterile water. This was incubated at 37°C for 16 hours to ensure complete digestion of the DNA. To inactivate the enzyme the temperature was raised to 94°C for 5 minutes before cooling and storing at 4°C. To ensure complete digestion an aliquot was run on a TAE agarose gel as described in Section 2.3.6

2.3.8 RNA Extraction

Total RNA was extracted using Sigma TRI Reagent™. One millilitre of TRI Reagent was used for either 10^7 cells from culture or for between 15-20 snap frozen tissue sections, cut to 7 µm, depending on the tissue size. Tissue was dispersed and cells lysed by passing the suspension through gradually increasing grade needles until it would flow freely through a 25G needle. If the tissue was particularly fatty the suspension was centrifuged (12,000g,

10 minutes, 4°C) so excess fat which formed a layer on top could be removed and discarded. Chloroform (200µl) was added, and the tube vigorously shaken for 15 seconds to form an emulsion. After standing at room temperature for 15 minutes the partially separated suspension was fully separated by centrifugation (12,000g, 10 minutes, 4°C). A red organic phase contained the protein; the cloudy interface contained DNA; and a colourless aqueous phase contained the RNA.

The aqueous layer was removed, and 500µl isopropanol and 1 µl (20 µg) glycogen added, followed by incubation for 10 minutes at room temperature, then centrifugation for 10 minutes (12,000g, 4°C). The RNA pellets were washed twice in 75% ethanol (1 ml) and centrifuged for 5 minutes (7,500g, 4°C). Ethanol was removed, the RNA pellet allowed to air dry with inversion for 10 minutes (room temperature), and then dissolved in an appropriate volume of RNase/DNase free water; this was usually 20 µl. RNA concentrations were determined using spectrophotometry, samples were then split into small aliquots and snap frozen on dry ice, then stored at -80°C until required.

2.3.9 Reverse Transcription

All reverse transcription (RT) reactions were carried out using AMV reverse transcriptase on 1 µg total RNA using random priming (random hexamers).

The RT reaction mixture was produced as described in Table 2.5.

Component	Concentration
Magnesium chloride ($MgCl_2$)	5 mM
RT - Buffer	1x
dNTP mix	1 mM
RNasin®	1 U/µl
AMV RT	0.75 U/µl
Random primers	25 ng/µl
Total RNA	50 ng/µl
RNase/DNase free water to final volume	20 µl

Table 2.5 RT reaction mix composition.

Once each reaction was prepared the RT reactions were incubated at room temperature for 10 minutes, then at 42°C for 15 minutes. After this the reaction mix was incubated at 95°C for 5 minutes to deactivate the AMV RT. The resulting cDNA was stored at 4°C.

2.3.10 Polymerase Chain Reaction

To minimise imprecision, a master mix was created when performing PCR reactions. This would be specific for each primer set used, requiring only the addition of a template, be that DNA or cDNA, and the Taq polymerase. The composition of the final PCR reaction mixture is given in Table 2.6.

Once the master mix had been created it was aliquoted into a PCR reaction tube. To this the template DNA was then added.

The thermal cycling was then started; at the first annealing point the reaction was paused to allow the addition of Taq polymerase. Once the PCR reaction had finished the product was stored at 4°C.

Component	Amount
AJ buffer (10x)	5 µl (1x)
Right primer (Sigma Genesis)	10 pmol (0.2 pmol/µl)
Left primer (Sigma Genesis)	10 pmol (0.2 pmol/µl)
Taq polymerase (Promega, M166B)	1 U
Template	1-5 µl
Total Volume	50µl

Table 2.6 The composition of a PCR reaction.

Most (RT-)PCR reactions have a very similar cycle structure with an initial denaturing step, followed by the cyclic denaturing, annealing and extension phases and finishing with a final extension step (illustrated in Figure 2.11). The annealing temperature of the reaction is optimised by running parallel reactions at a number of different temperatures and using gel electrophoresis to determine which gave the most specific product. The temperatures assessed are based on the melting point of the primers (Tm), with the temperatures (Tm)°C, (Tm-1)°C and (Tm-2)°C being used.

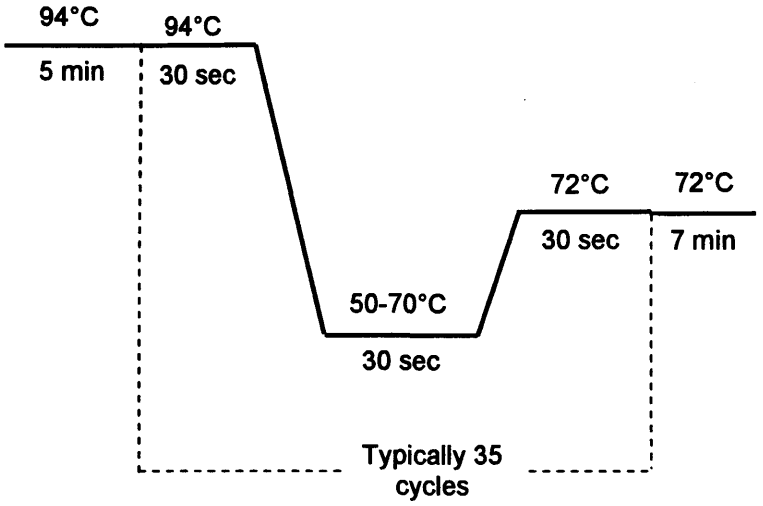

Figure 2.11 This shows a typical (RT-)PCR cycle structure. There is an initial denature step (red) at 94°C for 5 minutes. The (black) cyclic phase composes of a denaturing (94°C), annealing (50-70°C) and extending (72°C) step all for 30 seconds. The reaction is then finished with a final 7 minute extension at 72°C (blue).

2.3.11 Agarose Gel Electrophoresis

PCR products were run on agarose gels to determine the size of the products; it was also possible to semi-quantify levels of expression by band intensity. The gels used were between 2-3% agarose dissolved in 1x TAE buffer and run at 100V in a bath of 1x TAE. Ethidium bromide was added to visualise the double stranded DNA (2 µl/100 ml). Twelve microlitres of PCR product were added to 3 µl of 10x PCR loading buffer prior to loading, and for each row a DNA size ladder was also run (0.5 µg/lane).

Once the gel had been run the PCR products were visualised on a UV transluminator and digitally captured using FlashPoint 128 capture card and Vision Works 2.0.

2.3.12 Quantitative RT-PCR

All the qRT-PCR amplifications were performed on a Stratagene Mx4000, and the resulting data analysed using the stand alone software (v. 4.00) for the Mx4000.

2.3.12.1 TaqMan® Assay-on-Demand

The qRT-PCR primers and probes for AKAP1 and CEBP/A were supplied by Applied Biosystems. Both were inventoried primer sets, and so use a standard master mix. Each qRT-PCR reaction was produced to a total volume of 20µl, using 10µl of the 2x master mix, 1µl of the 20x primer/probe set, 1µl cDNA and made up to 20µl with RNase free water.

Using the standard techniques described above, cDNA was generated. The Assay-on-Demand primer and probe sets were all pre-optimised to operate using a 2-phase cycle that combines dimerisation and elongation. Both of the primer/probe sets required identical conditions with an initial 5 min denaturation at 94°C, followed by 40 cycles of 15 seconds denaturation at 94°C and 30 seconds dimerisation/elongation at 60°C. As it has become industry standard, ROX was used as the baseline control and FAM as the variable fluorophore.

2.3.12.2 SYBR Green II qRT-PCR

The SYBR Green reactions were performed using Platinum® SYBR® Green qPCR SuperMix UDG. The reaction mixture for each qRT-PCR reaction was composed of 25µl Platinum® SYBR® Green qPCR SuperMix UDG master mix, 1 µl ROX reference dye, 10 µM the forward and reverse primer, 1 µl cDNA and made to a total volume of 40 µl with RNase free waster. The PCR reaction cycle was similar to that of a traditional PCR 3 phase reaction. There was an initial denature step of 4 minutes at 95°C which was

followed by 45 cycles consisting of denaturation (95°C, 15 seconds), annealing (30 seconds) and extension (72°C, 30 seconds) phases. As described in section 2.2.3.3 the annealing temperature was determined by running the parallel reactions at (Tm)°C, (Tm-1)°C and (Tm-2)°C.

The specificity of the reaction must also be determined, as SYBR Green binds to all double strand DNA, including non-specific PCR products and primer dimers. The specificity of the reaction was determined by running a melting curve at the end of the reaction. This was produced by measuring fluorescence levels at 1°C increasing increments and plotting the amplitude changes. An example of a specific amplification is shown in Figure 2.12.

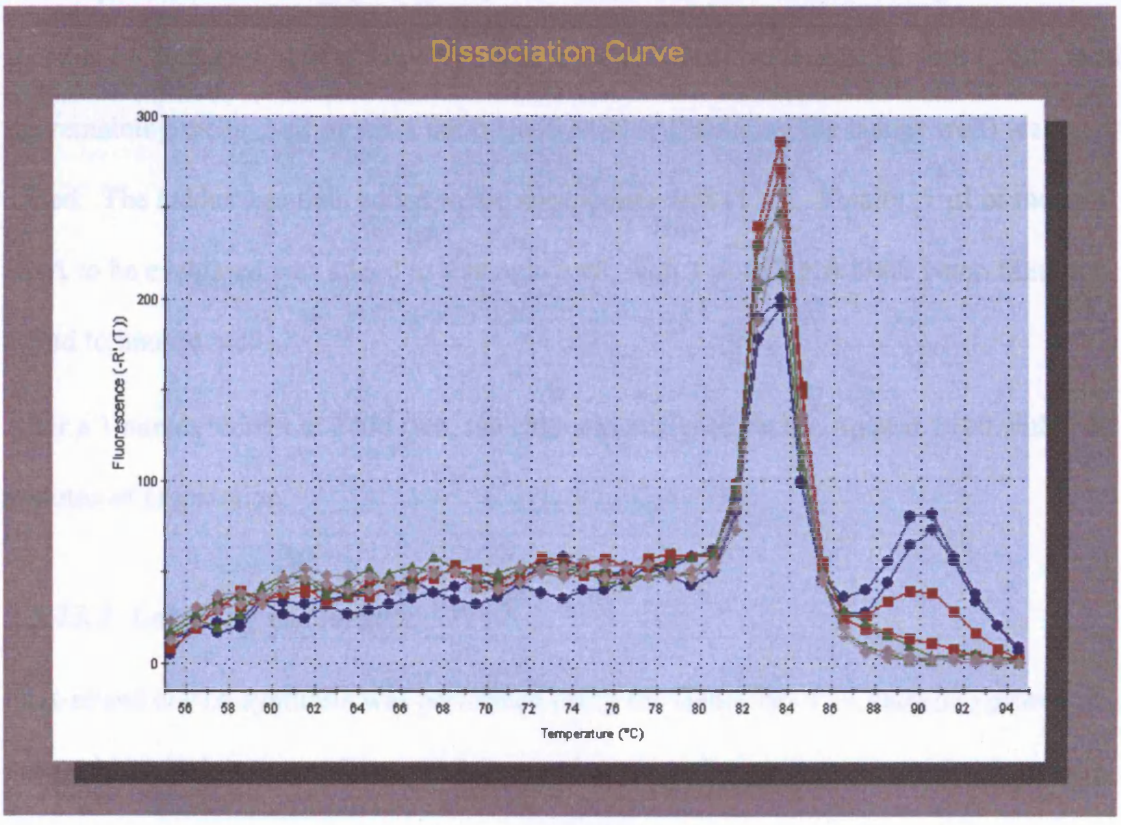

Figure 2.12 An example of a SYBR Green dissociation curve. At the end of the reaction a thermal dissociation curve can be plotted to determine the melting point of the products. This will give an indication as to the size of the product and the confirm that there were no non-specific PCR products produced. Only one peak of any significance can be seen, with all the different reactions represented. This was for the GRN reaction.

2.3.13 Affymetrix GeneChips®

RNA was isolated using TriReagent™, the quality was determined using the Agilent BioAnalyser and quantified using spectrophotometry.

2.3.13.1 BioAnalysis

The Agilent BioAnalyser was used as per the manufacturers' directions. These are summarised below, assuming all reagents have been reconstituted; only requiring loading into the RNA 6000 Nano Chip.

The chip was loaded with 9 µl of gel-dye mix and pressurised in the Priming Station for 30 seconds. A further 9 µl of gel-dye mix was loaded into the wells marked with a 'G'. In all the remaining wells, 5 µl of RNA 6000 Nano Marker (including the ladder well) was added. The ladder was then added to the appropriate well (1 µl). Finally, 1 µl of the total RNA to be evaluated was added to a sample well, with 1 µl of RNA 6000 Nano Marker added to unused wells.

After a 1 minute vortex at 2400 rpm, the chip was analysed on the Agilent 2100 within 5 minutes of preparation.

2.3.13.2 Labelling the targets

First-strand cDNA synthesis was performed using the GeneChip T7-Oligo(dT) promoter primer kit, which is a modification of the standard SuperScript™ reverse transcription kit produced by Invitrogen Life Technologies. This kit includes the T7-Oligo(dT) promoter primer:

5'- GGCCAGTGAATTGTAATACGACTCACTATAGGGAGGCGG-$(dT)_{24}$-3'

A summary of the reaction can be found in Table 2.7.

	Reagents in reaction	Volume	Final concentration or amount
1. Primer Hybridisation Incubate at 70°C for 10 min. Place on ice.	T7-$(dT)_{24}$ primer	2 µl	100 pmol
	RNA	(variable)	5 µg
	RNase free water	(variable)	
2. Temp Adjustment Add to the above tube, mix and incubate at 42°C for 2 min	5x buffer	4 µl	1x
	0.1M DTT	2 µl	10 mM
	10mM dNTP mix	1 µl	500 µM
3. First strand synthesis Add to the above tube and incubate at 42°C for 1 hour	SuperScript RT	1 µl	200 U
Total Volume		20 µl	

Table 2.7 Protocol for first strand cDNA synthesis

To this 150 µl of second strand synthesis reaction mixture was added. The composition of which is shown in Table 2.8.

Component	Final Concentration or amount
5x Reaction buffer	1x
dNTP mix	200 µM each
E. coli DNA Ligase	10 U
E. coli DNA polymerase I	40 U
E. coli RNase H	2U
RNase/DNase free water	(91 µl)
(Final Volume)	(150 µl)

Table 2.8 Composition of the second strand cDNA synthesis mix.

This was incubated at 16°C for 2 hours when T4 DNA Polymerase (10 U) was added. The mixture was incubated at 16°C for a further 5 minutes, after which 10 µl EDTA (0.5M) was added. Prior to labelling the double strand cDNA was cleaned-up using a spin-kit.

The synthesis of biotin-labelled cRNA was performed using the Enzo® BioArray™ HighYield™ RNA Transcription Labelling Kit. The protocol was performed as described in the product literature, but is summarised as follows.

At room temperature, and in the order indicated in Table 2.9 a reaction mix was created for each IVT reaction.

Reagent	Volume
Template ds-cDNA	10 µl
RNase/DNase free water	12 µl
10x HY Reaction Buffer (vial 1)	4 µl
10x Biotin-Labelled Ribonucleotides (vial 2)	4 µl
10x DTT (vial 3)	4 µl
10x RNase Inhibitor Mix (vial 4)	4 µl
20x T7 RNA Polymerase (vial 5)	2 µl
Total Volume	40 µl

Table 2.9 IVT reaction mixture and order.

This reaction was incubated at 37°C for 5 hours with gentle mixing every 30-45 minutes. Once the IVT had completed, the biotin-labelled cRNA was cleaned up using the spin-column kit provided and the cRNA was quantified using spectrophotometry. An aliquot was run on the Agilent BioAnalyser to ensure that full length biotin labelled cRNA has been produced.

The cRNA was then fragmented by incubating at 94°C in Fragmentation buffer for 35 minutes.

2.3.13.3 Hybridisation & Detection

A hybridisation cocktail was produced for each array. The composition of this is shown in Table 2.10.

Component	Concentration /Amount
Fragmented cRNA	10 µg
Control oligonucleotide B2	50 pM
20x Eukaryotic hybridisation controls (bioB, bioC, bioD, cre)	1.5, 5, 25 & 100 pM respectively
Herring sperm DNA	0.1 mg/ml
Acetylated BSA	0.5 mg/ml
2x Hybridisation buffer	1x
Water (RNase/DNase free)	--
(Final Volume)	(200 µl)

Table 2.10 Hybridisation cocktail for each probe array.

Prior to incubation with the hybridisation cocktail the array was incubated with 1x hybridisation buffer at 45°C for ten minutes, and the hybridisation cocktail was heated to 99°C for five minutes and then at 45°C for a further 5 minutes. After these incubations the hybridisation cocktail was added to the soaked array and incubated at 45°C for 16 hours. The hybridisation oven was set to rotate at a speed of 60 rpm.

Post hybridisation washes were performed on a GeneChip Fluidics Station 400. The array was then analysed using the Agilent Gene Array Scanner G2500A using standard techniques.

2.3.13.4 Analysis

The Affymetrix GeneChips were analysed using the open source programmes dCHIP v.1.3 (Cheng Li and Wing Hung Wong) and SAM v.1.21(Significance Analysis of Microarrays) (Gil Chu, Balasubramanian Narasimhan, Robert Tibshirani and Virginia Tusher) with a tolerance set to filter out changes less than 2 fold in either direction.. User guides for both these programmes can be found in Appendix 1.

Chapter 3:

BRCA2 Immunohistochemistry, Western Blot and RT-PCR Analysis

3.1 BRCA2

BRCA2 has a known role in familial breast cancer with mutations in the gene accounting for approximately 25% of all familial cases (Nathanson *et al*, 2001). Although no mutations have been found in sporadic breast cancer patients, evidence of BRCA2 involvement has come from LOH studies with significantly higher levels of LOH found in breast cancers from younger women compared to older women (Johnson *et al*, 2002).

As the investigations into mutations of *BRCA2* have proved unfruitful the focus of research has begun to shift towards investigating epigenetic modifications of *BRCA2*. Such modifications include promoter methylation status (Chan *et al*, 2002; Hilton *et al*, 2002; Alvarez *et al*, 2005) and UTR mutation analysis (Schubert *et al*, 1997; Hedau *et al*, 2004).

In order to have a cellular response, the downstream effects of any potential genetic or epigenetic modification must manifest itself as a change within the protein structure, or in the levels of the translated protein. Therefore, changes detected by immunohistochemistry or western blot analysis may originate from as of yet undiscovered genetic or epigenetic modifications.

3.2 Aims

The aim of this chapter was to examine BRCA2 mRNA and proteins levels in sporadic breast cancers in young women, and where possible relate this to previous LOH data from the research group

The specific objectives were:

1) to optimise commercially available antibodies and use these for immunohistochemistry studies in breast cancers in from young women and control normal breast tissue.
2) To perform western blots on breast cancer cells lines to confirm antibody specificity.
3) To carry out semi-quantitative RT-PCR analysis to determine BRCA2 mRNA levels in breast cancers in young women.

3.3 Methods

3.3.1 Immunohistochemistry

Three different anti-BRCA2 antibodies were available to evaluate for immunohistochemistry. However, only one of these was endorsed by the manufacturer for immunohistochemical use (66066e). The other two were only recommended for western blot analysis and immunoprecipitation.

3.3.1.1 66066e

The optimisation of antibody 66066e involved a range of antigen retrieval techniques including; (i) pressure cooking in citrate buffer at maximum pressure for either 2 or 5 minutes, (ii) Proteinase K digestion at 5 or 10 μg/ml for 1 hour and (iii) a combination of pressure cooking and Proteinase K digestion. In addition both HRP (DAKO Duet Kit) and alkaline phosphatase (using NBT/BCIP) detection systems were used. Tumour from 7 breast cancer cases, normal surrounding tissue from 3 breast cancer cases, 6 reduction mammoplasties

and paraffin embedded MCF-7 cells were used for the optimisation of this antibody.

3.3.1.2 PT146T

Pressure cooking in citrate acid buffer is one of the most effective forms of antigen retrieval. Therefore it was sufficient to use this method alone for this antibody. Pressure cooking at full pressure for 2 or 3 minutes was performed on MCF-7 cytoblocks, normal breast (RM224, RM225 and RM 227) and tumour tissue (RW1686T).

Horseradish peroxidase conjugated St-ABC was used at the tertiary detection system with DAB as the chromogen.

3.3.1.3 69261a

As with PT146T pressure cooking in citrate buffer for either 2 or 3 minutes was used for antigen retrieval. In addition, tissue sections were also pre-treated by microwaving in citrate buffer. Horseradish peroxidase conjugated St-ABC was used at the tertiary detection system with DAB as the chromogen. This immunoglobulin was optimised on normal surrounding tissue (RW2130S)

3.3.2 Western Blot Analysis

Western blot analysis was performed on seven breast cell lines, representing a range of breast cell lines, using four different antibodies. Two of these were the anti-BRCA2 antibodies used previously for immunohistochemistry (66066e and PT146T); an anti-BRCA1 antibody and an anti-MUC1 (EMA) antibody were used as positive controls.

As BRCA2 is such a large protein with a predicted mass of 400kDa (Tavtigian *et al*, 1997) two controls were required. BRCA1, with a mass of 220kDa (Chen *et al*, 1995; Chen *et al*, 1996), was used to confirm transfer of a high molecular weight protein. MUC1 was used to show protein transfer of a very high molecular weight protein. The mass of MUC1 ranges from approximately 250kDa up to 400kDa. This range is because MUC1 is a highly glycosylated protein, and like most glycoproteins the carbohydrate composition is less rigidly determined than the protein element.

Initial experiments utilised semi-dry transfer of the proteins from the polyacrylamide gel to the nitrocellulose membrane. These gave very little protein transfer, most probably due to the long transfer times generating excessive heat and so degrading the proteins. Subsequent transfers using a wet transfer overnight, were performed at lower ampere and so generated less heat. These gave excellent transfers, all results shown are the result of wet transfers.

3.3.3 RT-PCR

Using the RT-PCR protocol (see Sections 2.3.9 and 2.3.10), the relative expression of BRCA2 mRNA was determined using agarose gel electrophoresis and ethidium bromide detection, with all reaction run in duplicate. β-actin mRNA was used as a positive control and reference marker.

3.4 Results

3.4.1 Immunohistochemistry

3.4.1.1 66066e

This was the only primary antibody of the three characterised that was reported as being suitable for immunohistochemistry.

Under all conditions the no primary antibody (NPA) control free from any background staining when DAB was used as the chromogen (Figure 3.1a). When using an alkaline phosphatase detection system coupled with NBT/BCIP, only the typical non-specific, evenly distributed background staining was present.

For the majority of the breast tissues examined, using the NBT/BCIP detection system, no staining above the background was detectable. However in a small number of normal breast sections the cytoplasm of the luminal epithelial cells showed slightly stronger staining (Figure 3.1b). Using the DAB detection system, there was either no staining above the control, or there was excessive non-specific background staining (Figure 3.1c) which was not eliminated by increasing the blocking conditions. In a very limited number of cases it was possible to isolate areas of possible 'true' staining, however these patterns altered depending on the antigen retrieval methods that were used. For example, RM306 showed staining of the luminal contents and of the apical poles of the luminal epithelial cells, but only after 5 minutes pressure cooking pre-treatment (Figure 3.1d). Using 1 hour PK-10 digestion antigen retrieval, in a different normal sample (RM305) some of the luminal epithelial cells showed positive staining in the cell nuclei, but not in the luminal contents (Figure 3.1e).

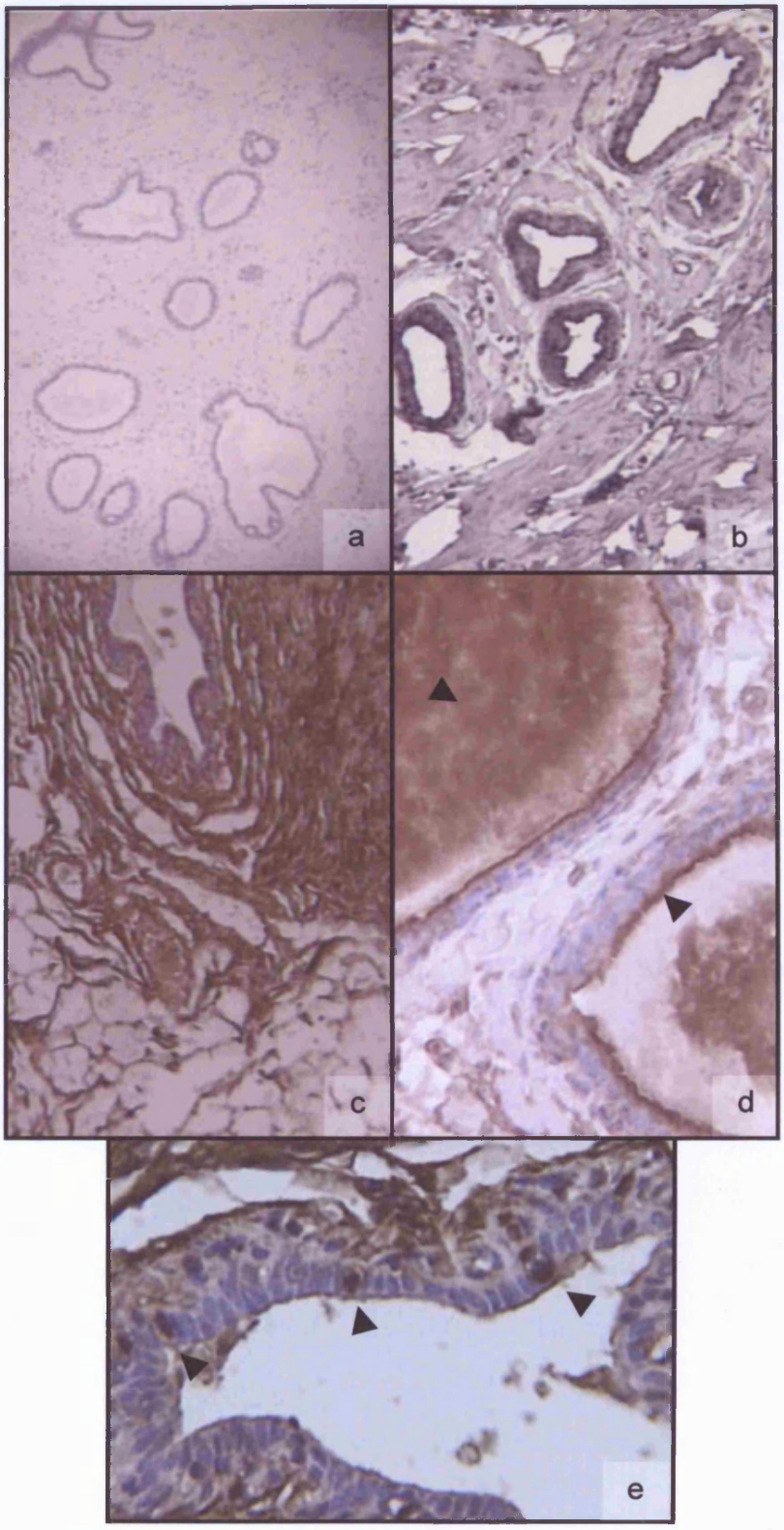

Figure 3.1. Images of BRCA2 immunohistochemistry. The various panels show; (a) The no primary antibody control on RM306 after 5 minutes pressure cooking using HRP/DAB based detection. (b) The antibody 66066e on RM301 after 1 hour PK-10 pre-treatment using alkaline phosphatase and NBT/BCIP detection. Here the cytoplasm of the luminal epithelial cells appear to be staining positively. (c) RM305 after 1 hour PK-10

pre-treatment, using the primary antibody 66066e and HRP/DAB detection. This shows the extent of the background staining observed. (d) This shows RM306 after 5 minutes pressure cooking, using the antibody 66066e and HRP/DAB detection. Under these conditions both the luminal contents and apical poles of the epithelial cells are stained positive. (e) This is the same section as (c) but under higher magnification showing nuclear staining.

3.4.1.2 PT146T

Using the antibody PT146T none of the NPA negative controls showed any staining, whereas the normal breast showed slightly positive staining of the cytoplasm in the luminal epithelial cells (Figure 3.2a and b) but was completely negative in the tumour tissue tested.

This antibody was also tested on FFPE MCF-7 cells. Neither the NPA (Figure 3.2c and d) nor positive with 2 minutes pressure cooking showed any staining. After 3 minutes pressure cooking however the whole cells stained positive (Figure 3.2e). As cells receive less fixation than tissues and so require less antigen retrieval it would suggest that this was all non-specific staining rather than specific immuno-detection.

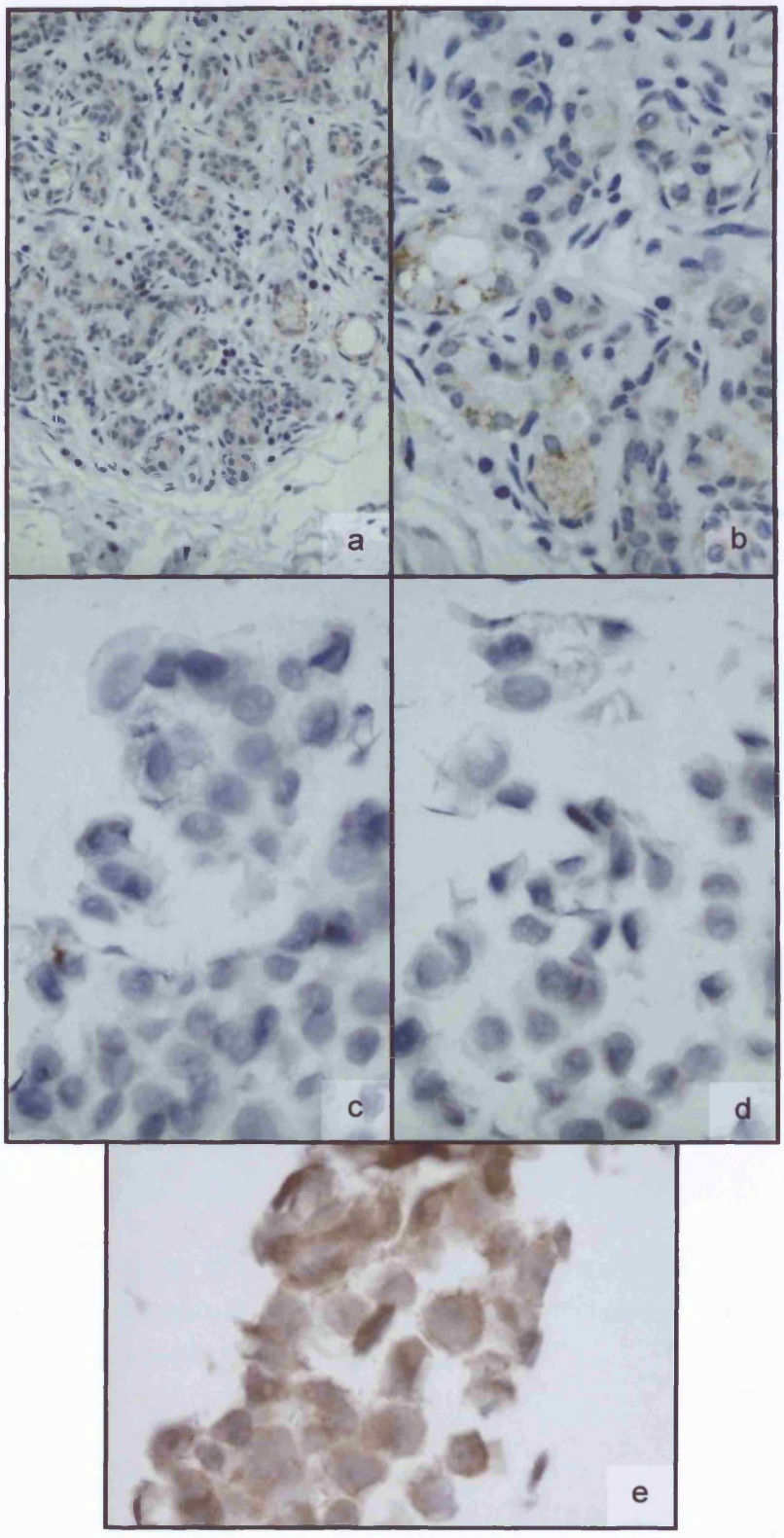

Figure 3.2 Images of BRCA2 immunohistochemistry. The various panels show; (a) RM225 is shown after 2 minutes pressure cooking pre-treatment using the antibody PT146T and HRP/DAB based detection. Here the cytoplasm of luminal epithelial cells appears to be stained positive. (b) This is the same section as (a) but under higher

magnification showing the cytoplasmic staining more clearly. (c) This shows the NPA negative control for FFPE MCF-7 cells after 2 minutes pressure cooking with HRP/DAB detection. Under the same conditions but using the primary antibody PT146T is shown in (d) and with 3 minutes pressure cooking in (e). There is no observable difference in (d) compared to the NPA suggesting that the staining observed in (e) is all non-specific

3.4.1.3 69261a

All antigen retrieval techniques failed to produce any staining, specific or non-specific. All the test sections were identical to the NPA negative controls (Figure 3.3a) although there were a few speckles of random precipitate in the microwave NPA control (Figure 3.3a)

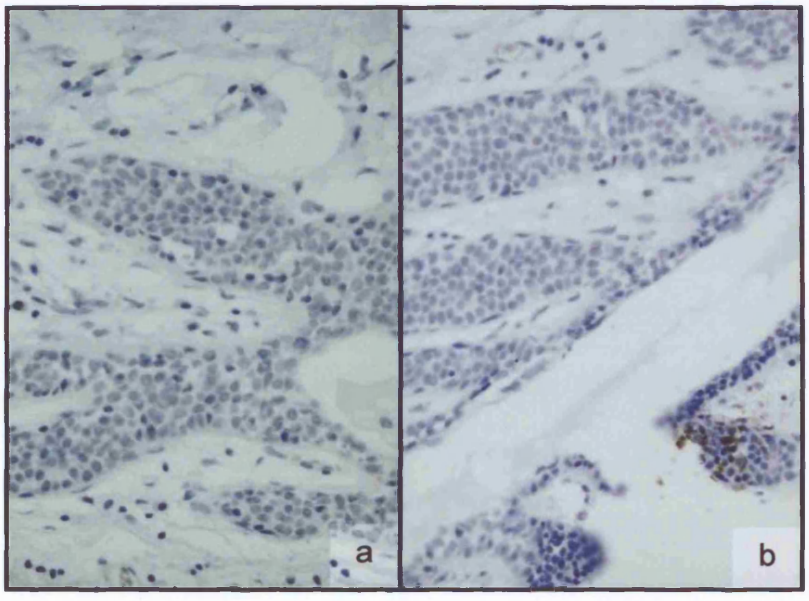

Figure 3.3 (a) shows the NPA control for RW2130s after 2 minutes pressure cooking and HRP/DAB based detection, with (b) showing the same case after microwave pre-treatment. Here non-specific DAB precipitate can be seen.

3.4.2 Western Blot Analysis

Having failed to perform successful immunohistochemistry, Western blots were carried out in order to confirm the specificity, or otherwise, of the three anti-BRCA2 primary antibodies under investigation. Representative western blot

analysis of all four primary antibodies used are shown in Figure 3.5 and summarised below:

A. When the Western blot was performed using the anti-BRCA2 antibody 66066e the strongest band detected was also the largest, at approximately 200kDa in size. There were also several smaller bands detected. Its also important to note that not every lane was positive, nor were the positive lanes equal in intensity. There was no band detected at 400kDa, the predicted size of BRCA2.

B. Using the anti-BRCA2 antibody PT146T, two bands of almost equal intensity were detected. These were approximately 200 & 300 kDa in size, and almost equally detected for each cell line studied. Whilst several smaller products were also detected, there was no band at 400 kDa as predicted for BRCA2.

C. The first of the positive controls was BRCA1. This antibody detected two bands, one approximately 250kDa, representing the BRCA1 protein. The second was smaller, approximately 200kDa in size. The staining was not equal with HBL-100 being stronger, and only six of the seven cell lines were positive for BRCA1.

D. The second positive control was MUC1/EMA. Only four of the cell lines were positive for MUC1, and as expected for a glycoprotein few distinct bands were visible, but bands were visualised up to 400kDa. Staining smears were also seen down to 250 kDa, with additional staining at smaller sizes than this.

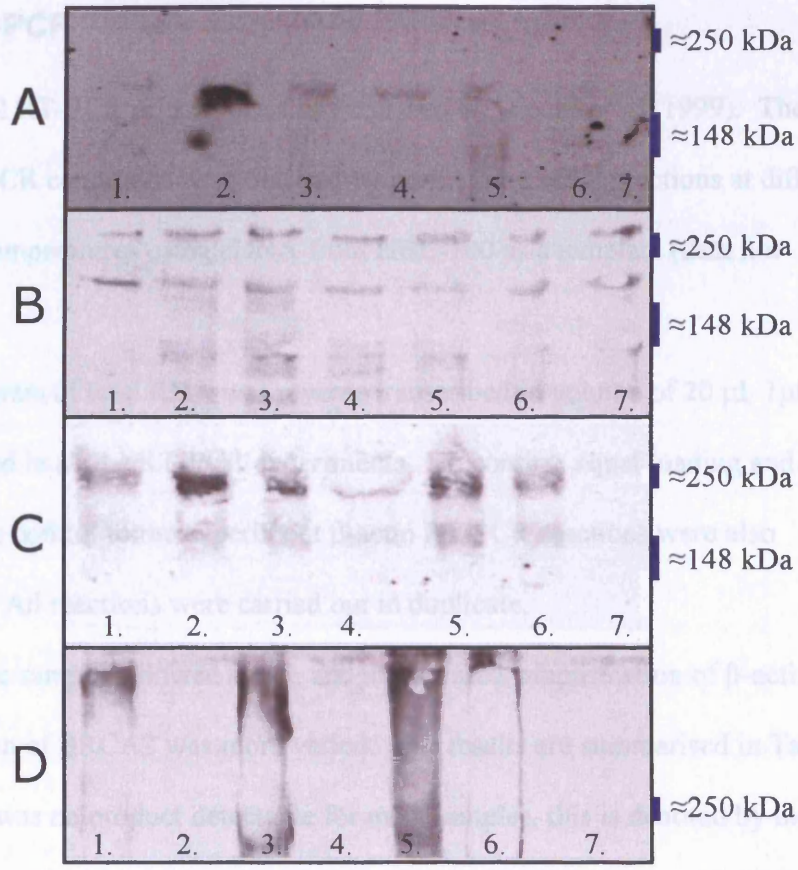

Figure 3.4 Western Blot analysis of two BRCA2 antibodies. Panel A is BRCA2 (66066e), panel B is BRCA2 (PT176T), panel C is BRCA1 and panel D is MUC1. From left to right the cell lines are; MCF-7 (1), HBL-100 (2), MBA-MD-486 (3), MBA-MD-231 (4), ZR-75 (5), T747D (6) & MBD-MD-436 (7).

The results from (D) show that proteins approximately 400kDa could be transferred from the polyacrylamide gel to a nitrocellulose and detected. The two anti-BRCA2 antibodies assessed were unable to identify a protein of approximately 380 kDa. However, they did consistently detect several smaller proteins, which could be BRCA2 degradation products. It is also important to note that using both anti-BRCA2 primary antibodies the strongest band was given by the non-tumourigenic breast cell line HBL-100 (Lane 2, Figure 3.4).

3.4.3 RT-PCR

The BRCA2 RT-PCR primers were as described by Bieche *et al* (1999). The published PCR conditions were verified by performing serial reactions at different annealing temperatures using cDNA from HBL-100 as a template (data not shown).

One microgram of total RNA was reverse transcribed in volume of 20 µl, 1µl of this was used in all the RT-PCR experiments. To confirm equal loading and to act as a positive control to the experiment β-actin RT-PCR reactions were also performed. All reactions were carried out in duplicate.

While all the samples showed equal, and unsaturated, amplification of β-actin the amplification of BRCA2 was more varied. The results are summarised in Table 3.1. There was no product detectable for most samples, this is denoted by the number "0". There were, however a few samples that did amplify with BRCA2, these samples were scored from 1 to 3 according to their band intensity. A score of 1 was given to bands of a low intensity, products that were of moderate intensity were scored as 2 and bands of high intensity were scored a 3.

Unfortunately the sample size is too small for any meaningful statistics but it is clear that a positive amplification did not shown any bias towards either the normal tissue (denoted by the prefix RM) or the tumour tissues (denoted by the prefix RW).

Twenty breast cancer samples, and 21 normal breast samples were evaluated. Of these 5 and 4 respectively were positive for BRCA2 by RT-PCR. Three breast organoid samples taken from reduction mammoplasty material were also assessed, all of which were negative. Interestingly, there was one paired primary tumour and lymph node metastasis, both of these samples were weakly positive for

BRCA2 by RT-PCR and had a score of 1. Two separate cultures of the non-tumourigenic breast cell line HBL-100 were also examined and both were positive as would be predicted from the western blotting data.

Sample	Density	Sample	Density	Sample	Density
RM166	0	RW1634	0	HBL-100	1
RM166	0	RW1639	3	HBL-100	2
RM167	0	RW1650	0	O1	0
RM170	0	RW1686	0	O2	0
RM170	0	RW1737	0	O3	0
RM170	0	RW1752	0	LN(T)	1
RM172	0	RW1772	0		
RM173	1	RW1820	0		
RM174	0	RW1844	0		
RM174	0	RW1857	1		
RM174	0	RW1857	0		
RM176	0	RW1889	0		
RM178	0	RW1941	1		
RM1820	0	RW2137	0		
RM2019	2	RW2195	0		
RM2195	1	RW2204	0		
RM2200	1	RW2215	0		
RM260	0	RW810A	0		
RM263	1	RW810A	0		
RM265	0	T(T)	1		
RM784	0	T(G)	0		

Table 3.1. A summary of the BRCA2 RT-PCR results. RM denotes normal tissue obtained from reduction mammoplasties, samples prefixed RW are tumour tissue. Samples O1, O2 & O3 are from organoid preparations. Sample LN(T) is lymph node

tissue matching tumour tissue from T(T). Band intensities were ranked from 0 (absent) through 1 (weak), 2 (moderate) to 3 (strong).

These results were reproducible, in separate amplifications, but using the same cDNA, and were produced after 35 PCR cycles.

3.5 Discussion

Despite using 3 different primary anti-BRCA2 antibodies under a range of different antigen retrieval and detection conditions it was not possible to satisfactorily optimise any of these to perform successful BRCA2 immunohistochemistry. While the manufacturers of these antibodies did not endorse their use for immunohistochemistry, two had been used for this purpose by other research groups (Table 2.2). The third antibody was the only other anti-BCRA2 antibody commercially available at the time and it was felt that its use in immunohistochemistry was worth exploring. Only few publications have described BRCA2 immunohistochemistry. Bernard-Gallon *et al*, (2002) reported that BRCA2 staining was found in the luminal secretions, in the nucleus and within the cytosol of normal tissue. In addition to this it was also demonstrated in the Golgi apparatus, endoplasmic reticulum, secretory vesicles and intranuclear hernia in invasive breast cancer.

By using a variety of immunohistochemistry conditions some similar staining patters were observed in this study. For example with the antibody 66066e, using 1 hour pre-treatment with PK-10 and NBT/BCIP detection, there was some staining of the cytoplasm of luminal epithelial cells. However, using the same antibody but with 5 minutes pressure cooking and DAB detection, the apical poles of the luminal epithelial cells, some nuclei of the luminal epithelial cells and the

luminal contents stained positively. However, as the background non-specific staining was quite strong this was not considered to be true positive staining. The paper by Bernard-Gallon does not comment on either background non-specific staining or if they experienced any difficulties reproducing their work. It is therefore not possible to accurately compare our data with the published study.

It is quite apparent that BRCA2 immunohistochemistry has been difficult to achieve for many research groups. Despite being discovered in 1995 (Wooster *et al*) a recent literature search for BRCA2 immunohistochemistry gave only 77 hits but only 14 of these had actually performed immunohistochemistry for BRCA2. Moreover, of these 14 papers 10 were from the Bernard-Gallon research group and one of the other research groups (Moll *et al*, 1999) is no longer using PT146T as the excessive antigen retrieval required gave very inconsistent results (Professor Ute Moll, personal communication).

In order to explore the reasons for the difficulties in optimising these antibodies for immunohistochemistry Western blots were performed to assess the specificity of the antibodies under investigation. The manufacturers data sheets provided with two antibodies available for Western blot analysis showed that there was cross reactivity with several smaller proteins (Figure 2.1). With no supporting evidence provided, the manufacturers suggest that these are BRCA2 degradation products.

Using specific antibodies to BRCA1 and MUC1 it was possible to detect very large proteins, of a similar size to BRCA2, with Western blotting. However, in the cell line lysates analysed no full length BRCA2 protein was detected. The largest molecular weight peptides that were detected were approximately 230 kDa and 300 kDa using primary antibodies 66066e and PT146T respectively, with

several smaller peptides also detected. As with BRCA2 immunohistochemistry there is very little published BRCA2 Western blot data. Again, of the 22 hits for "BRCA2 and Western blot" only 6 had performed BRCA2 Western Blots. One of these was by Bernard-Gallon (2003) in which they also discuss antibody cross reactivity. Yuan *et al* (1999) were also unable to detect high molecular weight BRCA2, but published findings based on a 230 kDa peptide.

Since the completion of this study there have been very few commercial anti-BRCA2 antibodies made available, nor have there been many BRCA2 immunohistochemistry or Western blot publications. This illustrates how complex antibody production and optimisation is for this huge protein.

The majority of previously published papers suggest that BRCA2 mRNA is ubiquitously expressed in breast tissue (Wooster *et al*, 1995; Tavtigian *et al*, 1996), in contrast to the findings of this study. While this is an admittedly small cohort, the data would suggest that BRCA2 mRNA is not ubiquitous within breast tissue. It is possible that the mRNA levels are linked to other factors such as the cell cycle when the rate of DNA duplication and repair vary. The PCR primers were taken from Beiche (1999) who were able to show that while BRCA2 levels showed wide variation in expression levels, there was a correlation between over expression and Grade III tumours. Due mainly to nuclear polymorphism and mitotic index, the authors suggest that *BRCA2* contributes to the proliferation rate of cells. It may therefore be the case that the stromal cells were diluting the epithelial cell RNA masking the true BRCA2 amplification. This could be resolved by extracting the RNA from epithelial cells isolated using laser capture microdissection. This technique was attempted but did not produce sufficient quantities of high quality cDNA for RT-PCR analysis. This would not however

explain the negative BRCA2 RT-PCR results obtained in the 3 organoids samples examined, which comprised an enriched population of both luminal epithelial and myoepithelial cells and warrants investigation in a larger series of samples.

Taking these data into consideration it is now generally accepted that early onset sporadic breast cancer is a discrete entity, with a novel aetiology. Different studies have considered various genetic alterations that may contribute to its development. One such investigation (Bergman-Jungestrom, 1999), led to the discovery of a single nucleotide polymorphism (SNP) in the cyp17 promoter, an enzyme involved in oestrogen biosynthesis. This T→C SNP produces a Sp1-type promoter site (CCACC box) and therefore increases the production of oestrogen. In this study, younger women were shown to be more likely to be either heterozygous or homozygous for the allele with the C mutation. As oestrogen exposure is a well-recognised risk factor, this suggests a possible role for oestrogen and the *CYP17* gene interaction in young breast cancer (Bergman-Jungestrom, 1999)

In conclusion, it was not possible to make any firm conclusions about BRCA2 protein levels in this study, despite extensive and repeated investigations. Moreover the data presented here suggest that any immunohistochemistry studies carried out with the specific BRCA2 antibodies used should be interpreted with caution, due to the poor antibody specificity and variation in staining seen. RT-PCR analysis was more successful, but as only few tissue samples, both tumour and normal, were positive for BRCA2 by RT-PCR it was not possible to relate these findings to previous LOH data in the young breast cancer cases. It would be important to use a more sensitive real-time PCR approach to accurately measure BRCA2 mRNA levels in any future studies.

Chapter 4:

Investigation of Skewed X-Chromosome Inactivation in Sporadic Breast Cancer in Young Women

4.1 Introduction

4.1.1 Introduction to X chromosome inactivation

X-chromosome inactivation in somatic cells was first demonstrated by Mary Lyon (1961), and is often referred to as Lyonisation. X-chromosome inactivation occurs at an early stage in the female embryo as a stochastic event, that is independent of any cells around it, and with equal probability of either X-chromosome being inactivated. Inactivation of the X-chromosome is most likely to be initiated from the X-inactivation centre by *XIST* (X inactivation specific transcript), which unlike most genes does not encode a protein and is expressed only from the inactive X-chromosome in post-inactivation somatic cells (Brown 1999). At the time of X-chromosome inactivation, *XIST* RNA functions in *cis* to spread an inactivating signal up and down the chromosome on which it resides (Willard *et al*, 1995). X-chromosome inactivation is then maintained throughout subsequent cell division and differentiation by continued *XIST* expression and by induction of methylation of DNA cytosine residues.

Therefore, due to X-chromosome inactivation, female tissue are normally comprised of a mosaic of two cell populations, expressing either a maternal or paternal X-chromosome. Deviations from random X-chromosome inactivation have been used to demonstrate clonal proliferation of malignant cells, positive or negative selection in carriers of adrenoleukodystrophy or X-linked immune disorders and primary disturbances of the X-inactivation process (Diaz-Cano *et al*, 2000; Indsto *et al*, 2001; Dobyns *et al*, 2004) . Such "skewed" X-inactivation can result as a chance occurrence, but has also been shown to be a consequence of specific factors affecting the X-inactivation process. Examples of these include female carriers of X-linked diseases such as Wiskott-Aldrich syndrome, Lesch-Nyhan syndrome, Barth syndrome and some immunodeficiency syndromes (Nyhan *et al*,

1970; Feron *et al*, 1988; Ørstavik *et al*, 1998; Christensen *et al*, 2000). Additionally, 33-50% of women 60 years or older also show an increased frequency of skewed X-chromosome inactivation, presumed to be the result of a slow but persistent selection process (Christensen *et al*, 2000; Brown, 1999).

4.1.2 Molecular analysis of X-Chromosome inactivation

There are essentially two strategies to study X-chromosome inactivation. Direct methods exploit phenotypic differences in cells that are exclusively dependent on whether the maternal or paternal X-chromosome is active. One such method was employed by Novelli *et al* (2003) to evaluate breast patch size. This method relied on a heterozygous C563T mutation in the X-linked glucose-6-phosphate dehydrogenase gene (*G6PD*), the so called Mediterranean mutation. This mutation causes a reduction in enzyme activity that is detectable by enzyme histochemistry. Cells with an active enzyme are visualised by a chromagenic reaction whereas the mutated *G6PD* cells do not catalyse the chromagenic reaction. Direct methods such as this are however limited in application to the study of specific cases that are mutation carriers. Indirect methods are therefore the more widely applicable, to determine X-chromosome inactivation patterns.

Since X-chromosome inactivation is predominantly mediated through epigenetic modifications such as gene hypermethylation, indirect methods of determining the X-chromosome inactivation pattern exploit this differential methylation status, typically using PCR based techniques (Lyon, 1972; Gartler and Riggs, 1983; Razin *et al* 1980; Grant *et al* 1992). The commonest technique takes advantage of a highly polymorphic trinucleotide repeat within exon 1 of the androgen receptor gene (*AR*), which is located on the X-chromosome. This locus is particularly useful as it is not only highly polymorphic but also is in close proximity to a CpG methylation site. The X-chromosome inactivation pattern is determined by pre-digesting DNA with the methylation sensitive restriction enzyme *Hpa*II,

before *AR* PCR. The enzyme will only cut the recognition sequence (CCGG) when the cytosines are unmethylated. Therefore, by pre-digesting DNA with *Hpa*II only the methylated inactive allele will be available for amplification. As the amplicon spans a highly polymorphic trinucleotide repeat, the two different alleles can be distinguished on the basis of their size and the ratio determined (Figure 4.1).

Traditionally, the ratio of the two alleles has been determined by separation using either agarose or polyacrylamide gel electrophoreses, with the bands detected using either ethidium bromide, silver staining or using radio-labelled primers that can be detected using X-ray film (Lyon, 1972; Busque *et al,* 1996). More recently, researchers have used to autosequencers that can quickly and accurately determine the ratio of the two alleles. Irrespective of the method used, an arbitrary value is taken to represent a skewed X-inactivation pattern. There is as yet no accepted standard value that represents a skewing of the X-inactivation pattern. For example a skewed X-chromosome inactivation pattern is considered to be when the ratio is less than 0.9 (Kristiensen, *et al*, 2002; Hatakeyama *et al*, 2004) or when there is a 3 x difference in band intensity (Buller *et al*, 1999).

Figure 4.1 (next page) Detection of skewed X-inactivation using *Hpa*II sensitivity. Skewed X-inactivation can be detected using methylation sensitive DNA restriction and amplification across a highly polymorphic region of the *AR*. Amplification of unrestricted DNA produces two equal bands when separated by gel electrophoresis. After *Hpa*II digestion, the unmethylated, active allele (blue) is digested and therefore fails to amplify. So, in a population of cells there are three possible outcomes. If the X-chromosome inactivation was random, then there will be equal amplification product for each allele. If however the X-inactivation was non-random, one of the alleles will amplify preferentially.

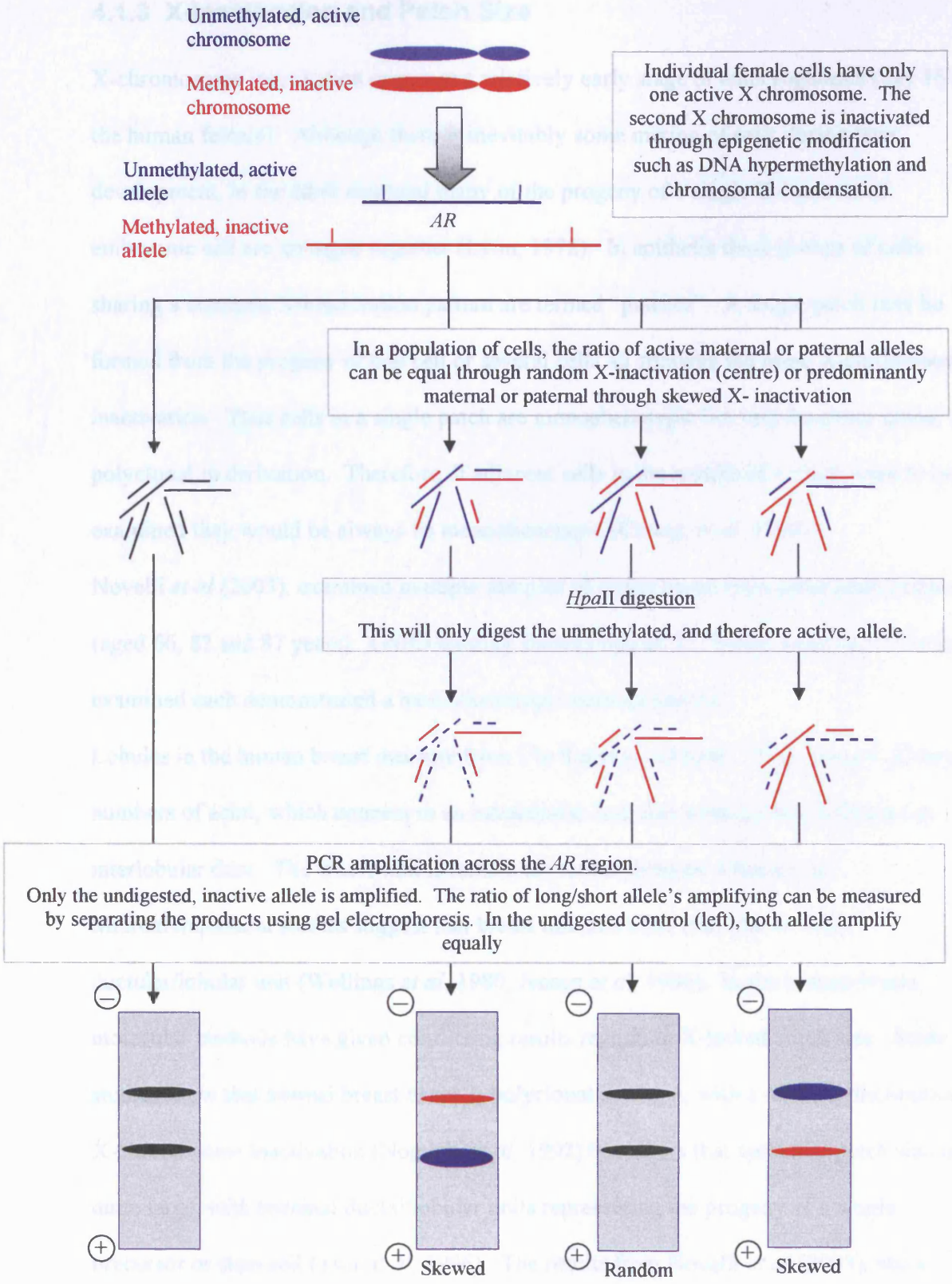
Unmethylated, active chromosome
Methylated, inactive chromosome
Unmethylated, active allele
Methylated, inactive allele
AR
Individual female cells have only one active X chromosome. The second X chromosome is inactivated through epigenetic modification such as DNA hypermethylation and chromosomal condensation.
In a population of cells, the ratio of active maternal or paternal alleles can be equal through random X-inactivation (centre) or predominantly maternal or paternal through skewed X- inactivation
HpaII digestion
This will only digest the unmethylated, and therefore active, allele.
PCR amplification across the AR region.
Only the undigested, inactive allele is amplified. The ratio of long/short allele's amplifying can be measured by separating the products using gel electrophoresis. In the undigested control (left), both allele amplify equally
Skewed
Random
Skewed

Figure 4.1

4.1.3 X-Inactivation and Patch Size

X-chromosome inactivation occurs at a relatively early stage of embryogenesis (day 16 in the human female). Although there is inevitably some mixing of cells during later development, in the adult mammal many of the progeny of a single X-inactivated embryonic cell are arranged together (Lyon, 1972). In epithelia these groups of cells sharing a common X-inactivation pattern are termed "patches". A single patch may be formed from the progeny of one cell or several cells all showing the same X-chromosome inactivation. Thus cells in a single patch are monophenotypic but may be either clonal or polyclonal in derivation. Therefore, if adjacent cells in the middle of a patch were to be examined they would be always be monophenotypic (Chung, *et al*, 1998).

Novelli *et al* (2003), examined multiple samples of breast tissue from three adult females (aged 66, 82 and 87 years). G6PD staining showed that all 111 small ducts and 57 lobules examined each demonstrated a monophenotypic staining pattern.

Lobules in the human breast measure from 1 to 8 mm in diameter. They consist of varying numbers of acini, which connect to an intralobular duct that subsequently becomes an interlobular duct. The whole unit is termed the terminal-ductal lobular unit.

Microarchitectural studies suggest that breast tumours arise from the terminal ductular/lobular unit (Wellings *et al*, 1980; Jensen *et al*, 1986). In the human breast, molecular methods have given conflicting results regarding X-linked patch size. Some studies show that normal breast tissue is polyclonal in origin, with a random distribution of X-chromosome inactivation (Noguchi *et al*, 1992) but others that epithelial patch size is quite large, with terminal ductal/lobular units representing the progeny of a single precursor or stem cell (Tsai *at al*, 1996). The results from Novelli *et al* (2003), show directly that the terminal lobulo-ductular unit often lies all within one patch/clone so that in

studies using X-chromosome inactivation it may be very difficult to ever demonstrate a polyclonal origin of breast tumours.

Figure 4.2 shows the positive and negative G6PD staining in breast tissue. Interestingly, both myoepithelial and luminal cells showed large patches of G6PD positivity and negativity. Based largely on tissue culture studies, it is believed that epithelial and myoepithelial cells have a common stem cell origin (Lakhani, *et al*, 1999, Gudjonsson *et al*, 2002). These data would support this hypothesis.

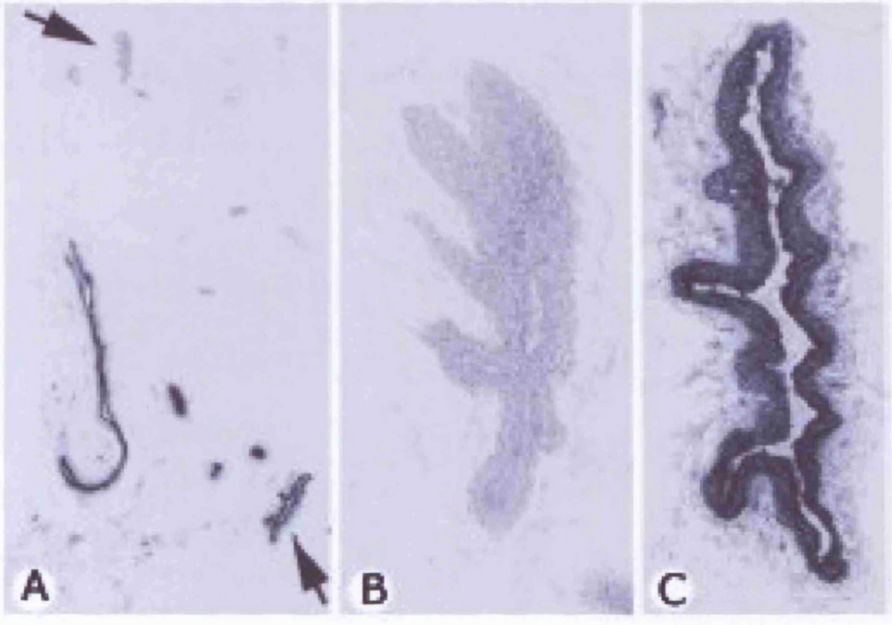

Figure 4.2 G6PD staining in human breast tissue. (A) A low-power view of benign breast tissue stained for G6PD activity. Arrowheads highlight positive and negative staining ducts shown in higher magnification in B and C. (B) A high-power view of a G6PD negative staining duct. (C) A high-power view of a positive-stained duct. Both luminal and myoepithelial cells stain in a similar fashion. (Magnifications: A, x 14; B and C, x 94). (Taken from Novelli *et al*, 2003)

4.1.4 Female cancers & skewed X-chromosome inactivation

Recently a correlation between skewed X-chromosome inactivation and ovarian cancer was demonstrated (Buller *et al*, 1999). A higher frequency of X-inactivation was demonstrated in women with invasive cancer when compared to borderline cancers and

healthy controls. Buller *et al* (1999) also reported an association between *BRCA1* germ line mutation carriers and the incidence of skewed X-inactivation.

In 2002, Kristiansen *et al*, investigated the X-inactivation patterns of Norwegian breast cancer patients, collected between 1984 and 1994. The median age of diagnosis was 60 years, ranging from 27 to 90. The normal controls were aged 20-40 years (144 cases), 19-65 years (138 cases), 73-93 years (202 cases) and 55-72 years (91 cases). The DNA was extracted from peripheral blood. PCR amplification across exon 1 of the *AR* with and without prior *Hpa*II digestion was performed to distinguish the active from inactive alleles. The PCR products were detected using an ABI autosequencer. The X-inactivation pattern was deemed skewed when the ratio of products was ≥10:1, and the results analysed using Pearson's Chi-squared (χ^2).

The X-inactivation pattern was found to be skewed in the younger patients (≤48 years) when compared to the young control group (20-40 years) (22% and 2 % respectively, p=0.003), and in the patients aged 27-65 years compared to blood donors aged 19-65 years (7% and 0.7% respectively, p=0.005).

Kristiansen *et al* also investigated the length of the CAG repeat length but failed to find a correlation between the repeat length and age of diagnosis or between breast cancer patients and controls.

Whilst acknowledging that familial breast cancer patients show earlier onset of the disease, they were unable to verify whether the patients were sporadic or familial breast cancer patients. One of the aims of this study was therefore to investigate sporadic breast cancer patients only.

4.2 Aims

This study investigated sporadic breast cancer in young women (<36 years) compared to matched post menopausal women (> 55 years), testing the null hypothesis that there is no

significant difference in the frequency of skewed X-chromosome inactivation in these two cohorts.

4.3 Method

As discussed above, the large patch sizes in breast can cause distorting effects and so it is difficult to use breast tissue to determine where X-inactivation is skewed. Therefore DNA from uninvolved lymph nodes removed at the time of wide local excision/mastectomy was used to extract DNA from each case. All tissues were formalin fixed and paraffin embedded.

4.4 Results

An example of the X-Inactivation gels produced and the subsequent analysis by Scion densitometry are shown in Figure 4.3.

The bands produced in lane a were analysed using Scion Densitometry and the graphs produced are shown, also labelled a, with the top band on the left and the bottom band on the right. This was continued for lanes b through d. Lanes a and b are the result of PCR amplification across the CpG island in the *AR* gene without and with *Hpa*II digestion respectively for case RW2195. Lanes c and d show the same but for case RW2137. From the gel image alone it is clear that both alleles are amplified approximately equally in lanes a and b, suggesting that this patient showed random (50:50) X inactivation. In lane c there were two PCR products detected, confirmed by densitometry, but only one product in lane d, indicating skewed X-inactivation.

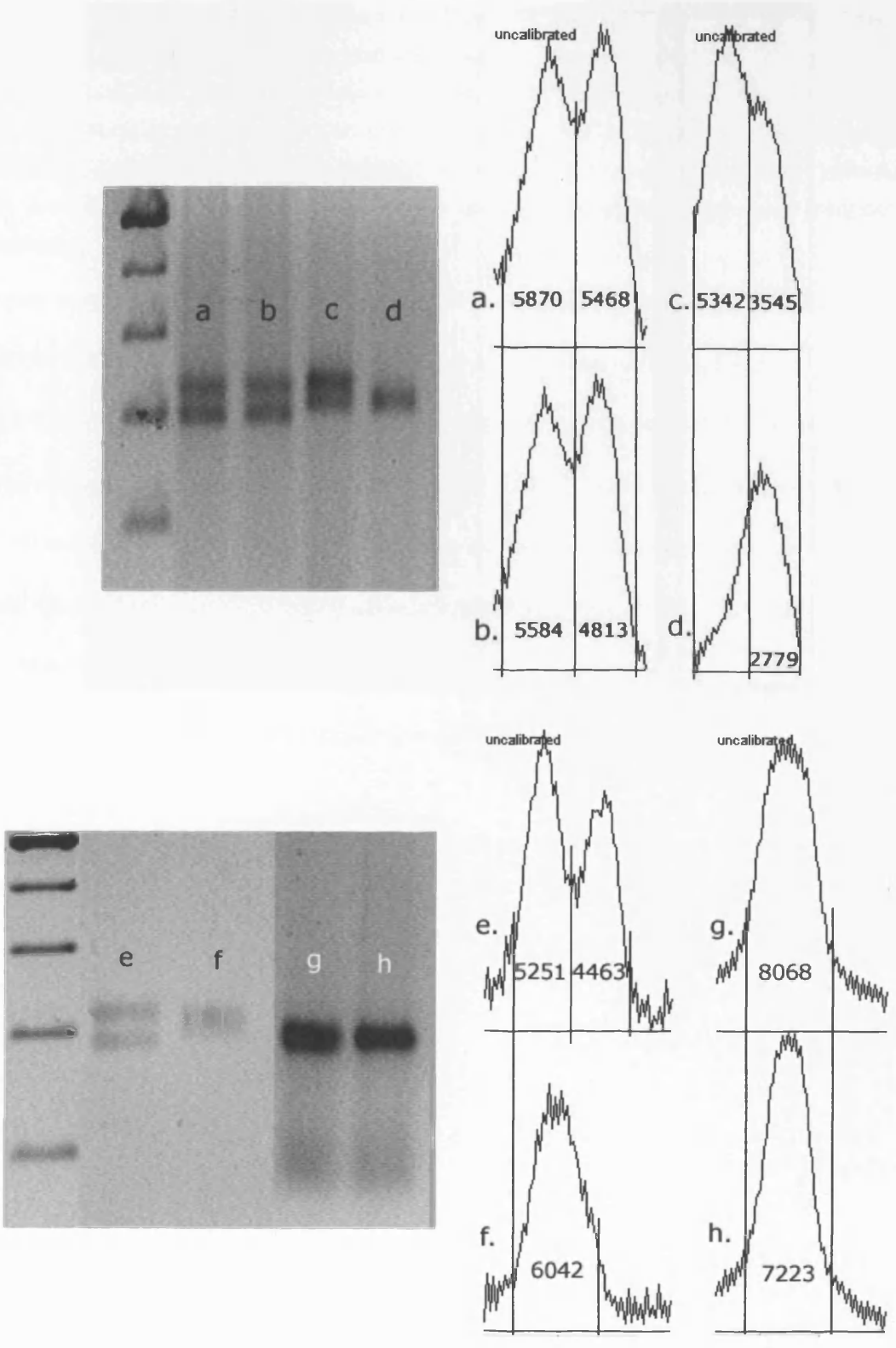

Figure 4.3 Examples of skewed, random and homozygous (non-informative) X-chromosome inactivation. Lanes a and b represent case RW2195 undigested and *Hpa*II digested respectively. Lanes c and d represent case RW2137 without and with *Hpa*II digestion respectively. The Scion densitometry (top right panel) confirms that in lanes a and b there were two alleles approximately equally expressed suggesting a random X-inactivation pattern. Whereas, the Scion densitometry

for lanes c and d shows that there were two alleles without *Hpa*II digestion but only one with, suggesting a skewed X-chromosome inactivation pattern. Lanes e and f represent case RW2107 undigested and *Hpa*II digested respectively. Lanes g and h represent case RW2113 without and with *Hpa*II digestion respectively. Both the gel image (bottom left panel) and Scion densitometry (bottom right panel) show that case RW2107 has a skewed X-inactivation pattern. Whereas as only one product is detectable with and without *Hpa*II digestion making case RW2107 non informative.

Sixteen cases aged <35 years and 21 cases ≥55 years were examined (Table 4.1). 11/14 (78.6%) informative cases from young women showed skewed X-inactivation, 7/19 (36.8%) informative cases from older women showed X-inactivation. These data were compared using Chi-squared(χ^2) analysis (SPSS, v. 12.0.1) and there was a significant difference (P=0.017). This suggests that there was a significant increase in skewed X-inactivation of younger (<35 years) cases compared to older women (>55 years) with sporadic breast cancers.

Case	Cohort	No. of alleles detected at AR		Skewed X-inactivation
		- *Hpa*II	+ *Hpa*II	
RW2196	Young breast cancer cases (≤35 years)	2	1	✓
RW2019		2	1	✓
RW2105		2	1	✓
RW2195		2	2	✗
RW2137		2	1	✓
RW2215		2	1	✓
RW2055		2	1	✓
RW2220		2	1	✓
RW2201		2	2	✗
RW2157		2	1	✓
RW2204		2	1	✓
RW1962		2	2	✗
RW1944		1	1	NI
RW2200		2	1	✓
RW2151		1	1	NI
RW2006		2	1	✓
Number (and %) informative cases				11/14
				78.60%

Case	Cohort	No. of alleles detected at AR		Skewed X-inactivation
		- *Hpa*II	+ *Hpa*II	
RW1930	Older breast cancer cases (≥55 years)	2	1	✓
RW2030		2	2	✗
RW2062		2	1	✓
RW2063		2	2	✗
RW2074		2	2	✗
RW2092		1	1	NI
RW2187		2	2	✗
RW2121		2	2	✗
RW2078		2	2	✗
RW2377		2	1	✓
RW3044		2	1	✓
RW2107		2	1	✓
RW2113		1	1	NI
RW2114		2	1	✓
RW2123		2	1	✓
RW2182		2	2	✗
RW2146		2	2	✗
RW2172		2	2	✗
RW1936		2	2	✗
RW1694		2	2	✗
RW2142		2	2	✗
Number (and %) informative cases				7/19
				36.80%

Table 4.1 Skewed X-chromosome inactivation in young and older breast cancer cases. When there were two alleles without *Hpa*II digestion and only one with, the distribution was skewed (✓). When there were two alleles in each reaction the distribution was random (50:50) (✗). When only one allele was detectable with undigested DNA the case was homozygous for the *AR* polymorphism and therefore non-informative (NI).

4.5 Discussion

A higher frequency of skewed X-chromosome inactivation has been identified in young women (≤35 years) with sporadic breast cancer compared to older women (≥55 years) with sporadic breast cancer in this study. The comparison between young and older women with breast cancer is valid, and offers a more powerful comparison than a cancer vs. normal breast comparison. As there is no recognised consensus at which ratio of alleles is designated "skewed" there is no reason to presume that the on/off methodology used in this study is any less valid than the 3x intensity or 90:10 ratio reported by other groups. Kristiansen *et al* (2002) found a higher frequency of skewed X- inactivation in young (27-40 years & 27-45 years) breast cancer patients than in controls (20-40 years & 19-45 years). Neither middle aged women with breast cancer (55-72 years) nor older women with breast cancer (73-90 years) showed a higher frequency of X chromosome inactivation with respect to their controls (55-72 years and 73-93 years respectively).

As previously discussed, women show an increased tendency towards a skewed X-chromosome inactivation pattern as they age. It would therefore be ideal to compare younger women with breast cancer to those without and likewise for older women. Unfortunately it was not possible to obtain age matched lymph node tissue from normal controls and as breast tissue develops in patches it would be unsuitable for this purpose. However, given that women develop a skewed X-inactivation pattern as they age, the comparison of young women with sporadic breast cancer to older women with sporadic breast cancer is a robust one as probability would predict a skewing in the favour of the older cohort.

These data show a much greater probability of skewed X inactivation in young women with breast cancer than shown in Kristiansen *et al* (78.6 % compared to 22%). Other than the age of the patients, Kristiansen does not make any other comments to the patients or

the tumours. All the patients in this study had sporadic breast cancer, and were from cohorts with narrower age restraints than those studies by Kristiansen. Additionally, whereas Kristiansen extracted DNA from fresh peripheral blood, this study used FFPE lymph nodes. DNA from FFPE tissue is of a lower quality, is highly fragmented and often looses epigenetic modifications such as DNA methylation. As these post-fixation events take place independently of biological function it would not affect the pre-fixation X-Chromosome inactivation status, but may make it more pronounced. This would explain why in this study either one allele or two equal alleles were detectable, resulting in the frequency of skewing being greater than previously reported. The method used was the simplest, and probably the least sensitive of all discussed in Section 5.1.2. It could also explain why the weakly amplifying allele was undetectable, again leading to a increase in the frequency of skewing.

Buller *et al* demonstrated that women with invasive ovarian cancer had a higher frequency of skewed X-inactivation than either borderline cancer patients or normal controls. In addition, 9 of the 11 patients with a mutated *BRCA1* gene showed skewed X-chromosome inactivation. Buller *et al* did not stratify the cases by age, however these results are in general agreement to those obtained in this study.

Both authors discuss two possible models involving the mutation of a X-linked tumour suppressor gene or oncogene. In the first model, X-inactivation is a chance event resulting in random X-inactivation, with inactivation of the second X chromosome occurring as a subsequent event resulting in inactivation of both alleles. The second model suggests that the mutated gene provides a proliferative advantage, thus results in a skewing of the inactivation pattern. This would also explain why older women with breast cancer do not have a high frequency of skewed X inactivation, as they women with a skewed X-

inactivation profile would be predicted to develop breast cancer at a younger age. The data from this investigation support this hypothesis.

In conclusion, this investigation has shown a high frequency of skewed X-chromosome inactivation in a cohort of breast cancers in young women (<36 years), in comparison to older women (≥55 years). Taken together these data provide evidence for the involvement of an X-linked gene in sporadic breast cancers in young women and warrants further study.

Chapter 5:

Microarray Profiling of Sporadic Breast Cancer in Young Women (≤35 years)

5.1 Introduction

Gene expression profiling using microarray technology allows for the simultaneous analysis of tens of thousands of markers across the entire genome. The real challenge of microarray techniques is in the extraction of useful information from a potentially large data set. This can be by either comparing intensities on a gene by gene basis or by analysing biologically related genes. The most commonly used methods are hierarchical clustering (Weigelt *et al*, 2005, Stec *et al*, 2005), *K*-means clustering (Callagy *et al*, 2005), self-organising maps (Pantazi *et al*, 2002, Nattkemper *et al*, 2005) and support vector machines (Listgarten *et al*, 2004). All of these evaluate the data without respecting the physical location of the genes within the genome. This can therefore obscure gene expression changes that result from know (or unknown) changes at the genomic level, such as DNA amplification, deletion, rearrangement and epigenetic modification such as DNA methylation.

5.1.1 CGH and Array CGH

Comparative Genomic Hybridisation (CGH) detects and maps DNA sequence copy number changes throughout the whole genome onto a cytogenetic map provided by metaphase chromosomes (Kallioniemi *et al*, 1992). This method limits the resolution to 10-20 Mb, restricting the detection of closely spaced aberrations and is limited by the accuracy of the available cytogenetic information. Array CGH however, provides the capacity to map genome copy number at a resolution determined by the size and spacing of the clones on the array. This is facilitated as array CGH uses genomic BAC, P1, cosmid or cDNA clones as the hybridisation target rather than metaphase chromosomes (Solinas-Toldo *et al*, 1997, Pinkel *et al*, 1998, Pollack *et al*, 1999 and Snijders *et al*, 2001). The relative copy number can be measured at these loci by labelling the test and reference

DNA with different fluorescent dyes and determining the ratio as with conventional CGH (Pinkel *et al* 1998).

5.1.2 Gene Amplification in Breast Cancer

Gene amplification effectively leads to an increase in gene dosage and has been seen in both scheduled genomic changes and as an unscheduled event such as that associated with resistance to cytotoxic drugs (Turton *et al*, 2001). Through systematic surveys, many proto-oncogenes have been identified within regions of amplification, of which many are the homologues of retroviral transforming genes, such as *MYC*, *ABL*, RAS^{K}, RAS^{H}, *MYB*, *HER2*, *ERBB2*, *MYCN* and *MYCL* (Schwab and Amler, 1990). The copy number of the amplified genes has been found to range from 5 to over 500 (Schwab, 1998). Although the amplified region of DNA is usually much larger than the transcriptional unit of a proto-oncogene propagating the amplification, continuous amplified regions of DNA, typically several hundred kilobases may represent the amplification of a single gene, such as the *MYCN* gene in neuroblastoma (Corvi *et al*, 1995). In breast cancer, however, it is far more common for complex discontinuous regions of up to 20 Mb to be amplified. This means that genes adjacent to the dominant proto-oncogene may also be amplified (Szeptowski *et al*, 1993; Guan *at al*, 1994).

Within breast cancer, such examples include the genes *Her-2*, *MYC*, *CCND1* and *ZNF-217* (Ross and Fletcher, 1999), which are known to be co-amplified in a significant proportion of tumours and their amplification is linked to high grade carcinoma. In addition *MYC* is associated with higher rates of proliferation and an unfavourable prognosis (Mizukami *et al*, 1995, Berns *et al*, 1996). Analysis has shown that some of these changes correlate to clinical behaviour. Given that there are often multiple amplification regions, which tend to be quite large in breast cancer, it would suggest that breast cancer cells are susceptible to gene amplification.

5.1.3 The MRC Toxicology in-house cDNA Array

5.1.3.1 Array Methodology

The in house cDNA microarrays were produced using a purpose built DNA microarrayer, based on the Stanford model, using a 16 and 32-tip head which produce 100μm cDNA spots. The arrays produced contain approximately 6500 cDNA clones based on the known human gene set from Research Genetics (Invitrogen). Further details about the cDNA arrays can be found at http://www.le.ac.uk/mrctox/microarray_lab.

5.1.3.2 Amplification of Chromosome 17q.

A major region that frequently shows amplification in breast cancers is 17q11-q21, which contains the oncogene *Her-2* (Slamon *et al*, 1987). Amplification and overexpression of *Her-2* is found in 10-34% of primary breast cancers (Venter *et al* 1987, Gusterson *et al* 1988). This amplification is not unique to breast cancer with ovarian, endometrial, prostate, gastric, oesophageal and head & neck cancers also showing amplification of the region (Slamon *et al*, 1989; Press *et al*, 1994; Ishikawa *et al*, 1997; Ross *et al*, 1997; Brien *et al*, 2000).

As chromosomal amplifications are usually attributed to a single target gene, the phenotypic changes observed in chromosome 17q11-q21 amplifications were traditionally attributed to the overexpression of *Her-2*. It has been observed, however, that *Her-2* is not always over expressed in cells with an amplification of chromosome 17q11-q21, suggesting that this is not the only gene capable of driving the amplification of the region. Using a combination of CGH and mRNA expression profiling very few genes have been shown to be both amplified and upregulated in this region. Using a chromosome 17 specific cDNA microarray Kauraniemi *et al* (2001) were able to identify 12 transcripts that showed a constantly increased copy number as well as being over expressed in at least 3 of

the 17 breast cancer cell lines investigated. These included the *HER2* gene as well as *GRB7*, *MLN64*, *PPARBP*, *ZNF-144* along with 7 ESTs. This data would suggest that genes other than *HER2* may also be responsible for propagating the chromosome 17q11-q21 amplification. Moreover, patients with both a amplification of this region and overexpression of *Her-2* develop cancers with a more aggressive phenotype than those that only over express *Her-2* (Farabegoli *et al*, 1999). *GRB7*, *PPARBP*, and *ZNF-144* were investigated together with *HER2* in this study.

Chromosomal amplifications are also well reported at the 17q23 locus, first identified by Kallioniemi *et al* (1994). Amplification of this locus is seen in approximately 20% of primary breast cancers and seems to be more common in high-grade tumours and either *BRCA1* or *BRCA2* mutation carriers (Roylance *et al*, 1999). It has also been shown that amplification of this locus is associated with a poorer prognosis (Tirkkonen *et al* 1997). Using CGH and expression profiling of both breast cell lines and primary breast cancer several potential oncogenes have been identified from within the commonly amplified region, *RPS6KB1*, *APPBP2*, *TRAP240*, *TBX2*, and *PPM1D* (Monni *et al*, 2001, Clark *et al*, 2002 and Sinclair *et al*, 2003). This could account for the poorer prognosis for patients with 17q23 amplification.

5.1.4 Genes analysed to validate the MRC Toxicology in-house cDNA data

5.1.4.1 HER2

In 1985, HER2 (NEU, ERB-β2) was independently mapped by two groups (Yang-Feng *et al*, 1985; Coussens *et al*, 1985) to 17q12-q22. This was later narrowed down to 17q21.1 by Muleris *et al* (1997) using fluorescence in situ hybridisation (FISH).

Slamon *et al* (1989) first described the role of HER2 in breast and ovarian cancer, later noting that HER2 was over expressed in 25-30% of breast cancers, increasing the aggressiveness of the cancers (Slamon *et al*, 2001). Numerous other studies have also linked HER2 amplification and/or overexpression to poorer prognosis (Menard *et al* 2001, Nabholtz *et al* 2002).

Moreover, there is good evidence that a val-655-ile variant is associated with breast cancer. A population-based case-control study of the val-655-ile polymorphism found that the valine allele was associate with an increased risk of breast cancer, particularly in younger women (Xie *et al*, 2000). Due to the significant ethnic differences of breast cancer incidence compared to other solid tumours, a study of allele frequencies was undertaken across 3 continents and over 7 ethnic groups. The frequency of the valine-allele varied between 1 to 24%, with the continental African populations having the lowest frequency corresponding to the lower incidence and lower risk of developing breast cancer (Ameyaw *et al* 2002).

5.1.4.2 ZNF144

The functions of ZNF144 are poorly characterised but it is known that its primary function is as a transcriptional regulator, more specifically a transcriptional repressor for sequences, which contain GACTNGACT (Kanno *et al* 1995). It may also have some tumour suppressor activity (Ishiwatari *et al* 1997).

Tagawa *et al* (1990) originally cloned the mouse homolog (MEL18), and determined that the protein contained a novel cystine-rich zinc finger motif. It was later shown that the protein structure also contained a RING-finger motif and a proline/serine rich region. The RING finger proteins may represent a mammalian homolog of the Drosophila polycomb gene group (Ishida *et al* 1993).

5.1.4.3 PSMB3

Proteasome subunit, beta type 3 (*PSMB3*) encodes one of the beta subunits of the 20S core structure of the proteosome, the most important component of the ATP-dependent proteolytic pathway in eukaryotes (Nandi *et al* 1997). The 20S core can assemble with other protein subunits to form the 26S proteosome complex that is responsible for proteolysis within the ubiquitin pathway. Additional assemblies have been suggested to have other cellular regulatory functions (Coux *et al* 1996).

5.1.4.4 PPARBP

The peroxisome proliferator activated receptor binding protein (PPARBP) is a transcriptional co-activator for the PPAR (Peroxisome proliferator activated receptors) class of steroid receptor mediated transcription factors and was mapped to 17q12-q21.1 (Zhu *et al*, 1999, Frade *et al*, 2000). There are three isoforms, termed PPARα, PPARβ and PPARγ, all of which are involved in lipid metabolism (Reddy *et al* 1986, Issemann and Green *et al* 1990).

PPARBP has many functions, including roles within the multi-protein thyroid hormone receptor-associated protein (TRAP), the vitamin D_3 receptor interacting protein (DRIP), and activator recruitment cofactor (ARC) complex (Zhu *et al* 2000). The primary role, however, is to increase the transcriptional activity of PPARγ (Zhu *et al* 1999).

5.1.4.5 CrkRS

Cdc2-related kinase with an arginine/serine-rich domain (CrkRS) is a member of the Crk family. These are a group of proteins that show varying homology (42-55%) to CDC2 (cell division cycle 2), a tyrosine kinase that has a major role in the control of the cell cycle (Meyerson *et al* 1992). CrkRS is ubiquitously expressed in all tissue types and is thought to localise to SC35 nuclear speckles. Whilst the protein is poorly characterised, it has been

postulated to act as a link between the transcriptional and splicing machinery of the cell (Ko *et al* 2001).

5.1.4.6 GRB7

Growth factor receptor-bound 7 (*GRB7*) was mapped to the 17q21-q22 in 1997 by Dong *et al.* Prior to this the protein was cloned using the CORT method (cloning of receptor targets) and identified as a 535 residue peptide, containing an SH2 (src homology) domain. The GRB7 family of proteins are known to interact with a variety of cell surface receptors and other signalling molecules (Han *et al* 2001). Using a yeast 2-hybrid system it was shown that RND1 interacts with GRB7. More specifically that it in the switch II loop of RND1 (a region responsible for guanine nucleotide exchange in all GTPases) and the SH2 domain of GRB7 that interact (Vayssiere *et al*, 2000).

These interactions have been mapped, for example, between the switch II loop of RNDI, , and the SH2 domain of GRB7. GRB7 also interacts with the cytoplasmic region of EGF receptors, such as Her2, via its SH2 domain (Stein *et al* 1994).

GRB7 is frequently amplified and over expressed, along with *HER2* in breast cancer (Stein *et al* 1994), and it has been suggested that it may have a role in cellular migration (Han and Guan *et al* 1999, Han *et al* 2001).

5.1.5 Expression Profiling of Breast Cancer

The use of microarrays to assess genome wide expression profiles has only been part of the scientific mainstream since the late 1990's but since then there have been several hundred publications on breast cancer alone. Despite the shear number of publications, the majority fall into one of three broad topics (1) to group tumours by their pathology, (2) to elucidate signalling pathways and discover novel therapeutic targets, and (3) to predict tumour behaviour.

5.1.5.1 Grouping tumours by pathology and molecular pathology

This category of publications can be sub-divided in to those which pre-selected cases into distinct cohorts (i.e. were "supervised") and those which did not.

Those studies that defined distinct cohorts can be considered to follow a more traditional scientific approach. Given that the cases were classified by well defined and highly potent pathological and molecular pathological markers it is no surprise that the hierarchical clusters that form are based around these features. Many of the clusters identified were, for example, formed on the basis of steroid receptor status (Nagai *et al*, 2004) outcome (van 't Veer *et al*, 2002) and BRAC1 or BRCA2 mutation (Hendenfalk *et al*, 2001). This has subsequently been refined to characterise different cell populations with a single tumour (Man *et al*, 2005).

For studies in which cases were not pre-selected into specific cohorts, the same strong molecular pathological and pathological markers are still the most dominant clustering influence. Examples include nodal status (Ahr *et al*, 2002) and ER status (Sørtie *et al*, 2001).

Quite recently these two approaches have been combined. West *et al* (2001) for example used pre-selected cohorts (based on ER and LN status) and hierarchical clustering. From this Bayesian regression models were produced with predictive capabilities based on gene expression profiles. Using a second unselected set of tumour cases, the genes expression profiles were validated. This method proved to be highly predictive for ER status but less so for LN status, although the authors were optimistic that this could be refined.

Looking for tumour specific novel therapeutic targets, cDNA subtraction and microarray were utilised (Amatschek *et al*, 2004). By using pooled RNA from normal tissues, genes that were differentially expressed in breast cancer, lung squamous cell cancer, lung adenocarcinomas and renal cell carcinoma were identified. Thirty nine genes were found

to be selectively over expressed in breast cancers (Table 5.1). Using this technique is was also possible to account for genes that were the consequence of inflammatory infiltration.

Gene name	Accession no
Secreted protein, acidic, cysteine-rich (osteonectin) (SPARC)	NM_003118
Collagen, type I, α1 (COL1A1)	NM_000088
Collagen, type I, α2 (COL1A2)	NM_000089
Collagen, type III, α1 (COL3A1)	NM_000090
Fibronectin 1 (FN1), transcript variant 1	NM_002026
N-acetyltransferase 1 (NAT1)	NM_000662
Osteoblast specific factor 2 (fasciclin I-like) (OSF-2)	NM_006475
5T4 oncofetal trophoblast glycoprotein (5T4)	NM_006670
Thrombospondin 2 (THBS2)	NM_003247
KIAA0225 protein (KIAA0225)	D86978
Pre-B-cell leukemia transcription factor 1 (PBX1)	NM_002585
Collagen, type VI, α3 (COL6A3)	NM_057167
Platelet-derived growth factor receptor, β polypeptide (PDGFRB)	NM_002609
Similar to glucosamine-6-sulfatases (SULF2)	NM_018837
Matrix metalloproteinase 11 (stromelysin 3) (MMP11)	NM_005940
Interferon, α -inducible protein (G1P3), transcript variant 3	NM_022873
Transducin (β)-like 1 (TBL1)	NM_005647
Fer-1 (*C. elegans*)-like 3 (myoferlin) (FER1L3)	NM_013451
Matrix metalloproteinase 13 (collagenase 3) (MMP13)	NM_002427
Cyclin D1 (PRAD1: parathyroid adenomatosis 1) (CCND1)	NM_053056
Melanophilin (MLPH)	NM_024101
Nonmetastatic cells 1, protein (NM23A)	NM_000269

Chondroitin sulfate proteoglycan 2 (versican) (CSPG2)	NM_004385
Prolactin receptor (PRLR)	NM_000949
Small inducible cytokine subfamily A (Cys-Cys), member 19 (SCYA19)	NM_006274
Homo sapiens H3 histone, family 3B (H3F3B)	NM_005324
Stanniocalcin 2 (STC2)	NM_003714
Transcription factor AP-2 β (TFAP2B)	NM_003221
X-box binding protein 1 (XBP1)	NM_005080
Cathepsin K (pycnodysostosis) (CTSK)	NM_000396
Protease, serine, 11 (IGF binding) (PRSS11)	NM_002775
Activated RNA polymerase II transcription cofactor 4 (PC4)	NM_006713
Chromosome 1 open reading frame 29 (C1orf29)	NM_006820
Collagen, type V, α1 (COL5A1)	NM_000093
Hypothetical protein IMPACT (IMPACT)	NM_018439
Melanoma differentiation-associated protein-5 (MDA5)	NM_022168
Nonmetastatic cells 2, protein (NM23B)	NM_002512
Plasminogen activator, urokinase (PLAU)	NM_002658
Bone marrow stromal cell antigen 2 (BST2)	NM_004335

Table 5.1. Genes found to be amplified in Breast Cancer using cDNA subtraction. (Taken from Amatschek et al, 2004)

5.1.5.2 Signal pathway elucidation and novel target discovery

Harkin *et al* (2000) were one of the first groups to publish gene expression profiles following cellular manipulations. Using inducible cells, the effects of BRCA1 induction was observed at the mRNA expression level. The authors identified that GADD45 was induced by BRCA1, and that this correlated with BRCA1-mediated activation of the c-jun *N*-terminal kinase/stress-activated protein kinase JNK/SPK pathway.

The effects of oestrogen on human breast cancer ZR-75.1 cells has also been investigated (Cictiello *et al*, 2004) by comparing the expression profiles of cells incubated with and without 17β-oestradiol. Hierarchical clustering identified 344 genes, which had altered gene expression. These genes were shown to be involved in the cell cycle, cell survival, resistance to stress and chemotherapy as well as RNA and protein synthesis. Interestingly, there is some evidence of oestrogen response elements within chromosome 17, as this region showed several clusters of up regulation with oestrogen treatment. Moreover, when MCF-7 cells were cultured in the presence of exogenous BRCA1, a known inhibitor of oestrogen transcriptional activity, the number of genes differentially expressed in the presence of oestrogen was decreased. This reduction meant that only about 10% of the genes showed up or down regulation with oestrogen treatment (Xu *et al*, 2005).

5.1.5.3 Predicting tumour behaviour

About 60-70% of patients with lymph-node negative breast cancer are successfully treated with local or regional treatment alone (Early Breast Cancer Trialists' Collaborative Group, 1998). However, currently there are few diagnostic tools available to identify high-risk patients. Wang *et al* (2005) identified a 76-gene signature, that was highly informative at identifying patients who developed distant metastasis within 5 years (hazard ratio 5.67 [95% CI 2.59-12.4]). This was achieved using hierarchical clustering of 115 node negative breast cancer cases to determine the gene signature and was validated using a further 171 node negative breast cancer cases. With further development, it is hoped that this could prove to be a useful diagnostic tool within the pathologist armoury in predicting tumour behaviour.

The gene expression profile of 42 genes were also able to differentiate long term survivors (>9 years) from short-term survivors (<3years) with node positive tumours (Amatschek *et al*, 2004). While some of the genes found to be associated with short-term survival had

been reported previously (*TGF-β3*, *VCAM-1*, *CD44*, thyroid hormone receptor and Cyclin B1) many had not (*ERG2*, *B-Myb*, *MTH1* and *NET-1*) (Herrera-Gayol *et al*, 1999, Megha *et al*, 1999, Byrne *et al*, 2000, Ghellal *et al*, 2000 and Barrett *et al*, 2002). Previous to this, a set of 70 genes were found to be predictive at stratifying patients with a good or poor prognosis (van 't Veer *et al*, 2002). While the 42 and 70 gene sets did not correlate, many have closely related cellular functions, such as in cell cycle regulation, invasion and metastasis, angiogenesis and signal transduction and may reflect inter-tumoural heterogeneity between the different cohorts investigated.

5.2 Aims

The aim of this chapter was to use genomic array approaches to identify novel gene changes in sporadic breast cancers in young women. The specific objectives were:

(i) to carry out array CGH analysis in collaboration with the MRC Toxicology Unit and to validate any significant copy number changes identified,

(ii) to perform cDNA microarray with Affymetrix GeneChips® comparing tumour, normal breast tissue an HBL-100 cell line RNA,

(iii) to validate novel gene targets identified in (ii) using qRT-PCR and immunohistochemistry.

5.3 Methods

5.3.1 Samples analysed

5.3.1.1 Gant Array & Chromosome 17q RT-PCR

For the in house Gant Array, DNA was extracted from snap frozen material of 6 different young breast cancer cases. To act as the normal control, DNA was also extracted from

matching peripheral blood. The DNA was quantified using spectrophotometry and the quality assessed by running an aliquot on a 0.8% agarose gel (Section 2.3.6).
In order to investigate gene expression on chromosome 17q, cDNA was generated from 26 breast cancer cases one with a lymph node metastasis along with 16 normal reduction mammoplasties. In addition, breast organoids from a reduction mammoplasty, and the breast cell line HBL-100 were also processed, following standard protocols (Section 2.3.3). The total RNA extracted for the reverse transcription was quantified using spectrophotometry and the quality of the RNA assessed using an Agilent BioAnalyser.

5.3.1.2 Gene expression array

Total RNA was extracted from frozen tissue for over 35 cases from women with breast cancer occurring ≤35 years, a proportion of who had been studied previously (Walker *et* al, 1996, Johnson *et al* 2002). A further two cases were obtained prospectively. The quality of the extracted RNA was assessed using an Agilent BioAnalyser.

For the normal control four different sources were investigated (i) RNA extracted from matching uninvolved surrounding tissue taken at least 4 cm away from the tumour, (ii) RNA extraction from laser capture microdissected reduction mammoplasties, (iii) normal breast material from reduction mammoplasties, enzymatically digested with organoid populations removed for RNA extraction (provided by Kellie Mulligan) and (iv) the non-tumourigenic breast cell line HBL-100.

Eight Affymetrix HG-U133A GeneChips were available for this study, with the labelling of the target RNA, hybridisation and analysis of the chip as described in Sections 2.1.9 and 2.2.4. Samples analysed on the GeneChips comprised duplicate arrays for each of the two prospectively collected tumour samples from young women, one normal breast organoid sample and the breast cell line HBL-100.

5.3.1.3 Q-RT-PCR Verification

In order to verify the microarray data obtained q-RT-PCR was performed for a selection of genes that showed variation across the samples. Two different methodologies were employed; 4 genes were analysed using SYBR Green analysis, and two using commercially available TaqMan probes. The TaqMan probes were purchased from the ABI Assays on-Demand™ range of pre-optimised primers and probes. For full primer details see Section 2.1.12.

The qRT-PCR was performed on the two young age breast cancers (≤ 37 years), prospectively collected, that were also analysed on the Affymetrix GeneChip and a lymph node metastasis from one of these patients. In addition, for some of the qRT-PCR reactions a selection of breast cell line cDNA was also available for analysis. The cell lines available were MCF-7, ZR-75, MB-MDA-486, T47D and MB-MDA-231 which were cultured as described in Section 2.3.3. The RNA was extracted and reverse transcribed as described previously Section 2.3.8 and 2.3.9.

5.3.1.4 RBBP4 Immunohistochemistry

Formalin fixed paraffin embedded material from 12 older (>55 years) and 16 younger women (<36 years) were available for this study. Immunohistochemistry was performed as described in Section 2.3.2.5. Two factors were considered when scoring the staining observed, the intensity of the staining (Score_I), and the percent of cells staining positively (Score_C). The final score (Score_T) was achieved by adding these values (Table 5.2)

Score_I	Intensity of staining
1	Weak
2	Weak/moderate
3	Moderate
4	Moderate/Strong
5	Strong

Score_C	Percentage of cells staining
1	≤50%
2	51-70%
3	>70%

Score_T =
Score_I + Score_C

Table 5.2 The scoring method used to semi-quantify the RBBP4 immunohistochemical staining. The total score (Score_T) was determined by adding the intensity score (Score_I) to the percent of cells staining positively (Score_C)

5.4 Results

5.4.1 Array CGH analysis of breast tumour DNA

DNA was extracted from 6 snap frozen breast cancer cases using standard PCI (phenol, chloroform, IAA) methods. All cases were grade III, infiltrating ductal carcinomas from women with early onset disease (<36 years). Matching constitutional DNA was extracted from peripheral blood as the normal control DNA for the hybridisation. This was extracted from frozen blood using a Qiagen DNA extraction column following the manufactures instructions. DNA quality was assessed by agarose gel electrophoresis and quantity determined by spectrophotometry.

The labelling and hybridisations were carried out by another PhD student, Nicola Turton. All 6 of the young breast cancer samples analysed showed chromosomal amplification. While some of the amplifications were unique to a particular case, several were common amongst 2 or more of the cases. The most frequent amplifications were at 1q21, 8q24, 12q23, 13q14, 17q11-12, 17q21, 20q12 and 20q13. Areas containing more that one

amplified gene were seen in several tumours, most noticeably the long arm of chromosome 17 (Figure 5.1).

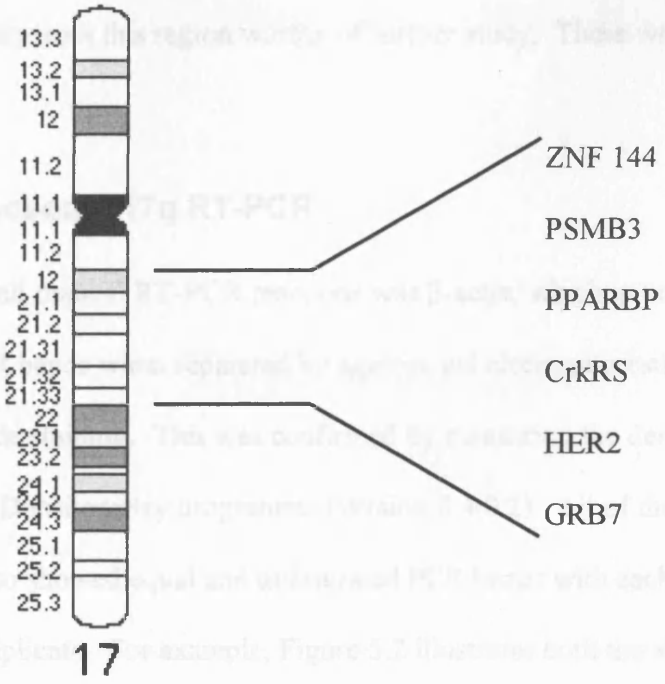

Figure 5.1 Ideogram of Chromosome 17. This illustrates the region of Chromosome 17q found to be amplified along with the genes selected for PCR analysis.

Initially a number of different chromosomal gene targets were selected for verification in tumour RNA and tissue sections. However, DNA sequencing of the clones revealed a number of errors in the identity of the clones, which was shown to be approximately 30% for the array used, so that it was not sensible to follow up these targets. Unfortunately, these sequencing errors took some time to identify, which hampered the progress of the project. As a result, none of the actual CGH data was available to guide the analysis of the later Affymetrix GeneChip gene expression data. Ideally the CGH study should have been repeated with new spotted arrays.

Therefore it was decided to focus on a more detailed analysis of the chromosome 17q amplification, firstly due to the known importance of *HER2* amplification in breast cancers and secondly as the array CGH had identified 3 known genes from chromosome 17q

HER2, *ZNF144*, and *GRB7* as being amplified. Analysis of GeneBank sequence data available at the time, and the consensus of other published literature lead to the selection of three other targets from this region worthy of further study. These were, *PSMB3*, *PPARBP* and *CrkRS*.

5.4.2 Chromosome 17q RT-PCR

The control for all manual RT-PCR reactions was β-actin, which gave equal and unsaturated PCR bands when separated by agarose gel electrophoresis and detected using ethidium bromide staining. This was confirmed by measuring the density of the bands using the Scion Densitometry programme (version β.4.0.2). All of the genes under investigation also showed equal and unsaturated PCR bands with each RT-PCR reaction carried out in triplicate. For example, Figure 5.2 illustrates both the absolute and relative density for the bands for β-actin and GRB7. The relative density was determined by dividing by the absolute density of the weakest band.

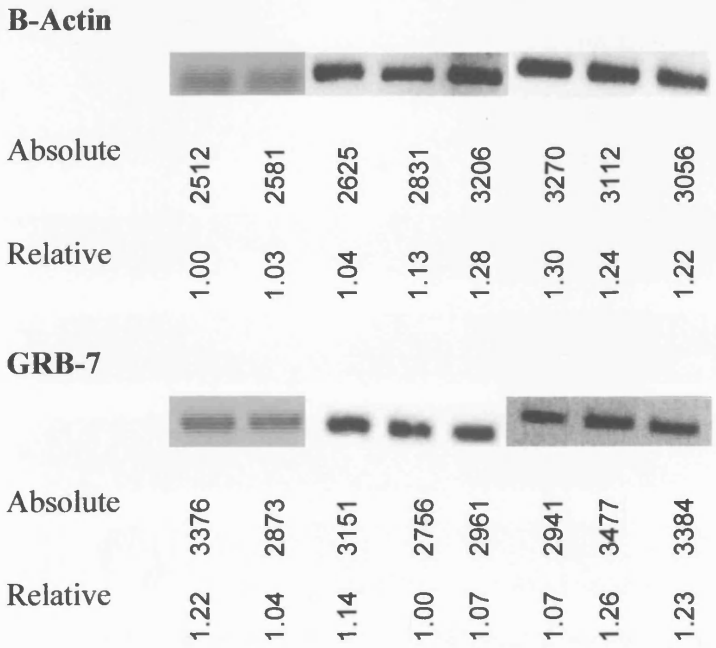

Figure 5.2 RT-PCR analysis of GRB-7 in breast tumours and controls showing the absolute and relative density of the PCR bands for 8 representative reactions (β-actin top, GRB-7 bottom). From left to right; T1, T2, O2 (tumour tissue and breast organoids also analysed by Affymetrix microarray), RM263, RM173 (normal breast), RW1772, RW1844 (tumour tissue) and HBL-100 (breast cell line). These are representative examples of the 23 tumour samples and 16 reduction mammoplasties that were investigated.

Examples of RT-PCR for the 7 other genes investigated are shown in Figure 5.3. Although there were small ratio changes between samples there was no obvious evidence of tumour specific amplification for any of the targets identified. Given the limitations of accurately quantifying manual RT-PCR results attempts were made to more accurately quantify this data by calibrating the Scion Densitometry programme to produce standard curves using known controls. Attempts to improve the quantification involved measuring PCR products generated from a serial dilution of template DNA analysed on several different gels and repeat separation of a fixed amount (1 μg) of 100 bp DNA ladder (Invitrogen). However, the calibration curves generated using the standards did not improve the analysis as the

scion densitometry programme was not accurate enough to distinguish bands of similar intensity (data not shown).

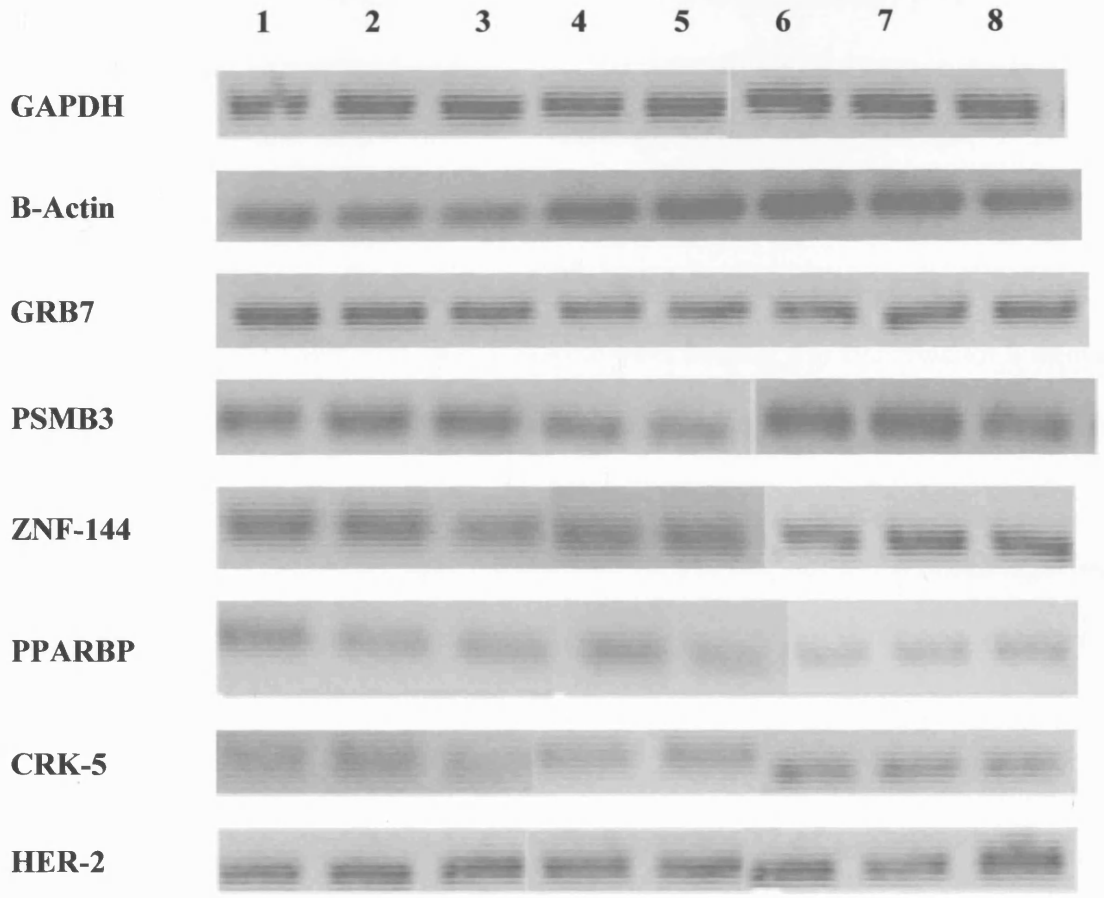

Figure 5.3 RT-PCR analysis of 6 gene targets from chromosome 17q. This shows representative PCR products for all the genes investigated using cDNA obtained from T1 (1), T2 (2), Organoids (3) (tumour tissue and breast organoids also analysed by Affymetrix microarray), RM263 (4), RM173 (5) (normal breast), RW1772 (6), RW1844 (7) (tumourous tissue) and HBL-100 (8) (breast cell line). These are representative examples of the 23 tumour samples and 16 reduction mammoplasties that were investigated.

5.4.3 Affymetrix GeneChip Analysis

5.4.3.1 An Imperfect Array?

One of the most complicated aspects of GeneChip microarray technology is the bewildering number of strategies available. Whether investigators choose to pre-select their cohorts, use unsupervised clustering or one of the many analysis packages, what is common is that the more cases studied, and the more repeats the better. While there is no doubt that array prices have fallen in the last few years, for the average research lab they are still very expensive. For this research project grant funding was available for 8 arrays, and as the tests needed to be run at least in duplicate this enabled only 4 different samples to be analysed.

Before embarking on the project advice was sought how to best utilise the limited number of chips and to obtain the most information. There were only two options to choose from. The first utilised individual cases per chip, the second would be to pool RNA from several cases. Both approaches have their advantages and disadvantages. Using one case per chip a full profile for each case can be obtained. However, with such a small number of chips available only a small number of cases can be evaluated and so a true profile of breast cancer could not be obtained.

By pooling samples a whole breast cancer profile would be obtained. However, it would not be possible to determine if the expression changes were the result of a cumulative effect of all the tumours or the dominate effect of a single case.

There was also debate as the most suitable control. Ideally this should be matched normal tissue, but as the majority of grossly dissected normal breast is composed of fat, it would be difficult to extract sufficient epithelial cell RNA. The alternative would be to use microdissected surrounding normal breast. Again this has several limitations. Firstly, it is technically difficult to isolate normal ducts using only a light haematoxylin background

stain. This is also compounded by the number of cells required, with a recent publication suggesting that approximately 10,000 cells are needed (Glanzer and Eberwine, 2004) Another option is to use breast organoids. These are isolated after enzymatic digestion of reduction mammoplasties and cell separation to produce an enriched population of ductal epithelial and myoepithelial cells. Given that breast cancers have an epithelial origin it would be advantageous to further separate the mixture, it is however not yet technically possible without separation and culture, which might change the expression profile. The final control option was to use the non-tumourigenic breast cell line HBL-100, which has a phenotype more akin to normal breast. This cell line was used with the caveat that it is far from "normal", indeed our own studies have shown that it is highly invasive and myoepithelial-like (Gordon *et al*, 2003).

While discussions were underway it became apparent that it was not going to be a difficult decision after all. Although the frozen tumour samples had been stored appropriately in liquid nitrogen vapour they had been used previously, leaving them open to the effects of freeze/thawing. A total of 76 breast tissue samples were extracted (41 normal, 35 tumour aliquots), and analysed using an Agilent BioAnalyser however, the 28S and 18S rRNA bands were very variable when present and the 28S/18S ratios were poor (Figure 5.4). As a result none of the RNA samples extracted from stored tissue samples were of a suitable quality for labelling for the Affymetrix GeneChip Array. Therefore it was necessary to recruit new patients to obtain fresh tissue for RNA extraction.

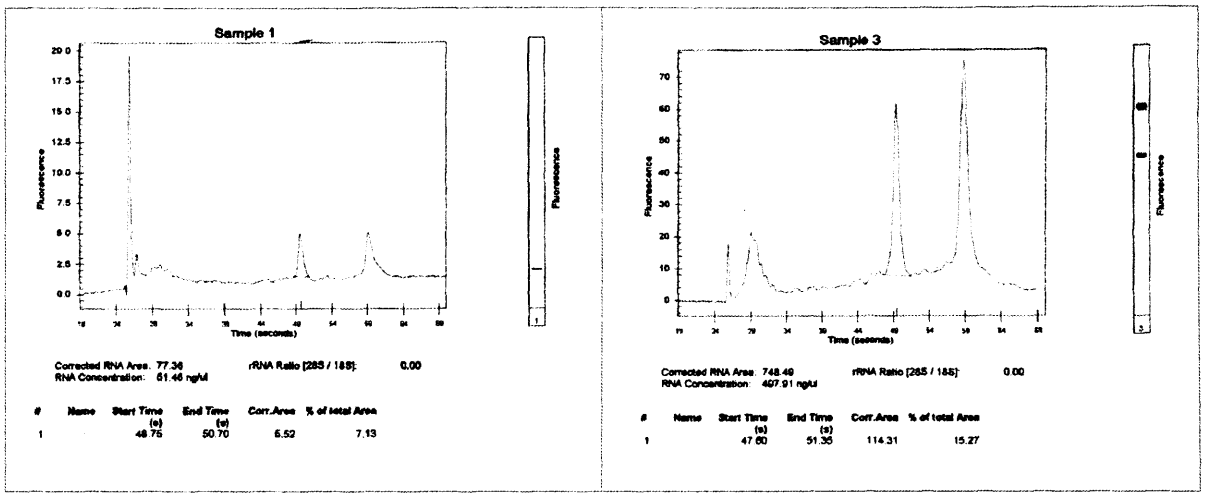

Figure 5.4 Agilent 2100 BioAnalyser analysis of total RNA extracted from frozen tissue sections. Capillary electrophoresis analysis and a gel representation of separated RNA.

Due to time limitations only two young breast cancer cases could be recruited in the time available. RNA was extracted from 5 mm^3 of tissue, selected by a consultant Histopathologist, and dropped freshly into 1 ml of TriReagent. H & E analysis of a frozen section confirmed that the tissue was greater than 90% tumour. Both of these tumours gave good quality RNA with clear 18S and 28S rRNA peaks on the capillary electrophoresis trace (Figure 5.5). With the RNA available, the best uses of the chip was to run 4 samples, each in duplicate, evaluating the expression profile of the two tumours prospectively collected (Tumour 1 & Tumour 2), with a normal breast organoid isolated from reduction mammoplasty and HBL-100s as controls.

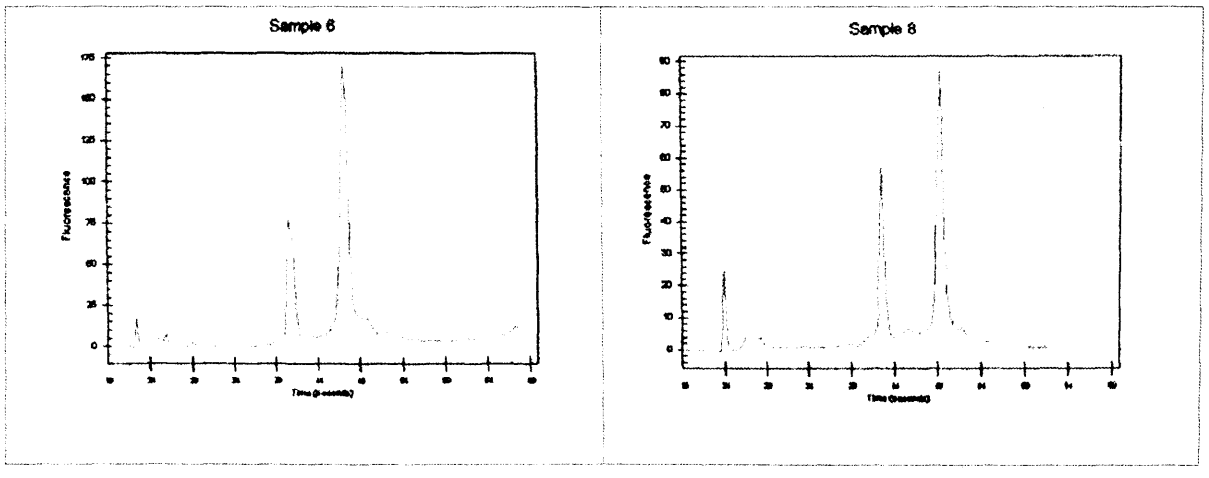

Figure 5.5 Agilent 2100 BioAnalyser analysis of total RNA extracted from two young breast cancer cases. The two large peaks show the 18S and 28S rRNA

5.4.3.2 Array Analysis

RNA was labelled and hybridised to the 8 available GeneChips as described in section 2.3.13. The Affymetrix GeneChips were analysed using the open source programmes dCHIP (Cheng Li and Wing Hung Wong [www.dchip.org]) and SAM (Significance Analysis of Microarrays) (Gil Chu, Balasubramanian Narasimhan, Robert Tibshirani and Virginia Tusher [www-stat.stanford.edu/~tibs/SAM/]) with a tolerance set to filter out changes less than 2 fold in either direction. User guides for both these programmes can be found in Appendix 1.

5.4.3.3 Tumour 1 vs. Tumour 2

There are 15,000 genes on the Affymetrix GeneChips, SAM analysis identified 147 genes which were differentially expressed between the two tumours investigated (Figure 5.6). Fifty of these showed higher expression in Tumour 1, and 97 showed higher expression in Tumour 2. For the full list of differentially expressed genes see Appendix 2. Of the 147 differentially expressed genes about 75% do not have a suspected role in breast cancer or are hypothetical proteins. Since there was less than 1% difference between the two tumour cases it was deemed acceptable to use a joint Tumour 1 and Tumour 2 comparison. One caveat of note is that *HER2* expression was 5.7-fold higher in Tumour 1 compared to Tumour 2, suggesting that Tumour 1 only was *HER2* amplified.

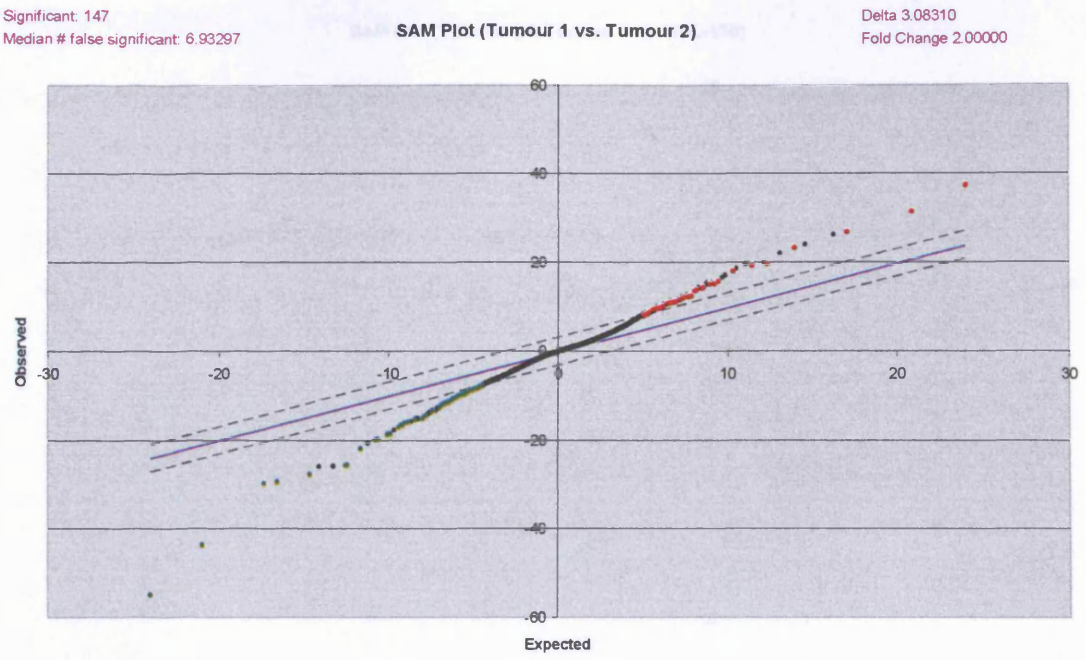

Figure 5.6 SAM analysis comparing Tumour 1 with Tumour 2. The blue line demonstrates y = x, i.e. a perfect match for Tumour 1 and Tumour 2 expression. The dashed black lines show y = x ± 2. This demonstrates the 2 fold error rate for the comparison. The black points therefore represent genes that are not differentially expressed in the two tumours, the red points show increased expression in Tumour 1 and the green points increased expression in Tumour 2.

5.4.3.4 Tumour 1 & Tumour 2 vs. HBL-100

SAM analysis was used to compare Tumour 1 and 2 with the HBL-100 cell line. This identified 470 genes that were up regulated and 285 genes that were down regulated in the two tumours (Figure 5.7).

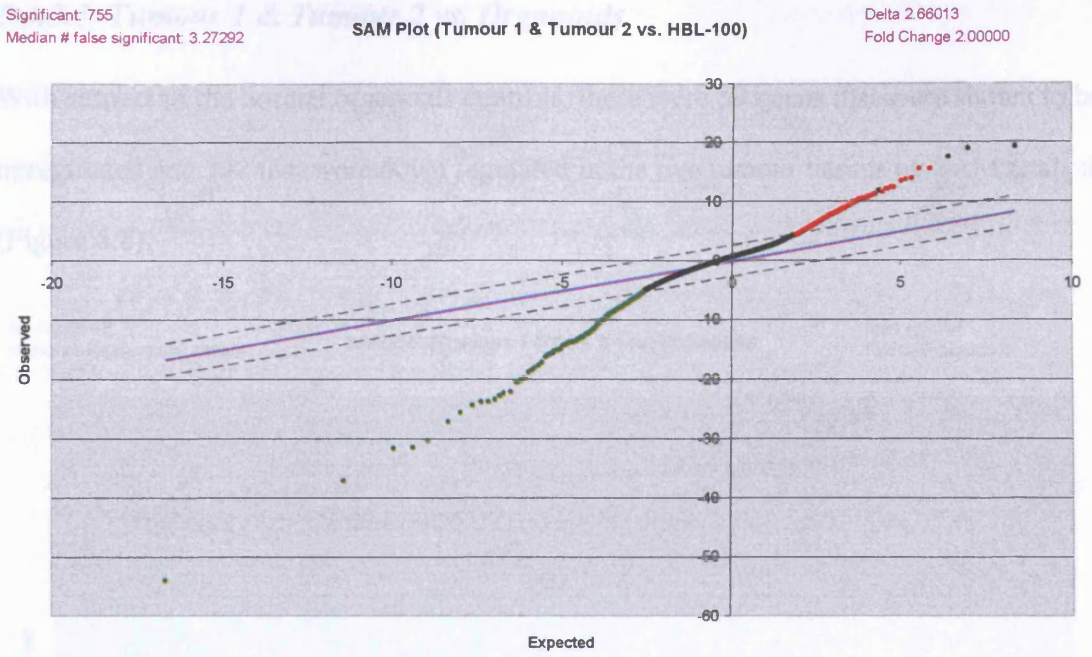

Figure 5.7 SAM analysis comparing Tumours 1 & 2 with. HBL-100. The blue line demonstrates y = x, i.e. a perfect match for Tumours 1 & 2 and the HBL-100 expression. The dashed black lines show y = x ± 2. This demonstrates the 2 fold error rate for the comparison. The black points therefore represent genes that are not differentially expressed, the red points show increased expression in the tumours and the green points decreased expression in the tumours.

Of the 755 genes that show differential gene expression between the two tumours and HBL-100, 78 were also differentially regulated in the tumours when compared to organoid RNA (see below). Of these 78 genes, 3 showed an inverted relationship. Both spermine N1-acetyltransferase and B-cell CLL/lymphoma 6 show a decreased expression in the Tumours compared to the organoids but a relative increase compared to HBL-100 and a hypothetical protein (FLJ10604) showed the reverse.

In addition, three of the genes selected for investigation by qRT-PCR from the tumour vs. organoid comparison were also differentially expressed (upregulated) when compared to HBL-100. These were the genes GRN, AKAP1 and C/EBPA. A complete data set generated by SAM analysis and table of the genes altered compared to both the HBL-100 and breast organoids can be found in Appendix 3

5.4.3.5 Tumour 1 & Tumour 2 vs. Organoids

With respect to the normal organoids controls, there were 69 genes that were shown to be upregulated and 372 that were down regulated in the two tumour tissues by SAM analysis (Figure 5.8).

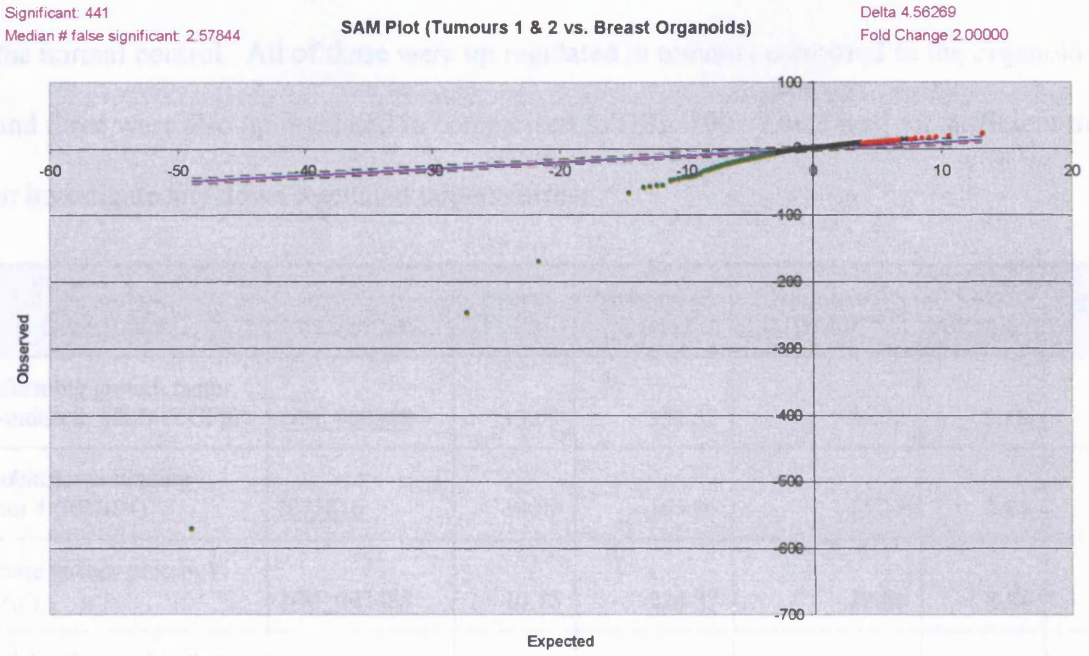

Figure 5.8 SAM analysis comparing tumour 1 and tumour 2 with a normal breast organoid RNA. The blue line demonstrates y = x, i.e. a perfect match for Tumours 1 & 2 and the breast organoid expression. The dashed black lines show y = x ± 2. This demonstrates the 2 fold error rate for the comparison. The black points therefore represent genes that are not differentially expressed, the red points show increased expression in the tumours and the green points decreased expression in the tumours.

Since the comparison between tumours and organoid RNA is likely to reveal true tumour specific gene changes it was decided to concentrate on validation of candidate gene targets identified by this study. The most significant fold changes were ranked and then bioinformatic analysis was used to identify suitable candidates for further study. This involved investigating each gene in cancer using gene sequence related references

(UniGene) and literature (PubMed). This analysis was discussed in detail with both supervisors and with Professor J. Lunec, the then head of the in house Affymetrix microarray facility. Based on their potential significance in breast cancer and evidence of differential gene expression by the SAM analysis, six genes were selected for further study (Table 5.3). All of these genes showed at least a 2 fold change between the tumours and the normal control. All of these were up regulated in tumours compared to the organoid and three were also up regulated in comparison to HBL-100. There was not sufficient time to investigate any down regulated targets further.

Gene Name	Gene ID	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	q-value (%)
transforming growth factor, beta-induced, 68kD (TGFβI)	NM_000358	13.09	352.52	26.92	3.53	0.38
retinoblastoma binding protein 4 (RBBP4)	X71810	10.89	165.96	15.23	2.04	0.38
A kinase anchor protein 1 (AKAP)	NM_003488	10.75	224.37	20.86	3.98	0.38
CCAAT/enhancer binding protein, alpha (CEBPA)	NM_004364	9.04	112.25	12.41	2.96	0.58
damage-specific DNA binding protein 2 (48kD) (DDB2)	NM_000107	9.78	51.67	5.27	2.30	0.58
Granulin (GRN)	BC000324	8.77	245.93	28.03	5.90	0.58

Table 5.3 SAM analysis data for the 6 candidate genes selected for further study. the full SAM Analysis (Appendix 4) showing the genes of interest, to be analysed further.

5.4.4 qRT-PCR Verification with organoids

5.4.4.1 Primer Optimisation and House Keeper Validation

For four of the gene targets primer sets were designed and manual PCR conditions were first optimised by varying the annealing temperature in AJ Buffer (see Materials and Methods). PCR conditions were considered optimal when a single specific product was

visible with no evidence of non-specific products or primer artefacts (Figure 5.9). This served to define a starting annealing temperature for the qRT-PCR runs. Primer sets and annealing temperatures were again optimised for qRT-PCR confirmed by a single specific product melt temperature (see Materials and Methods).

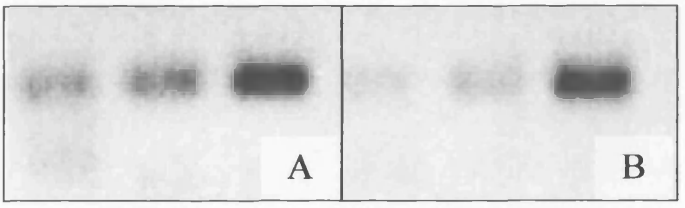

Figure 5.9 Optimisation of primer annealing temperature by manual PCR. (A) GRN, (B) RBBP4

β-actin was selected as a reference or "house keeping gene" for this study. A standard curve was generated from a serial dilution of HBL-100 cDNA which confirmed PCR efficiency (Figure 5.10A). The level of expression of each candidate gene was compared relative to β-actin in a number of cDNA samples including the two primary tumours, organoid and HBL-100. For one of the primary tumours there was also cDNA from a lymph node metastasis. For each sample 1μg of RNA was converted to cDNA and triple repeat qRT-PCR reactions were carried out using $^{1}/_{25}$ of this cDNA. Comparison of cycle thresholds showed consistent values for β-actin reactions (Figure 5.10B). The results presented in the following section give the Ct values (mean ± SD) for at least duplicate qRT-PCR reactions, the fold change compared to the organoids reaction is also given. Where this was significant ($p > 0.05$, 1-tailed t-test) this is marked with an asterisk (*).

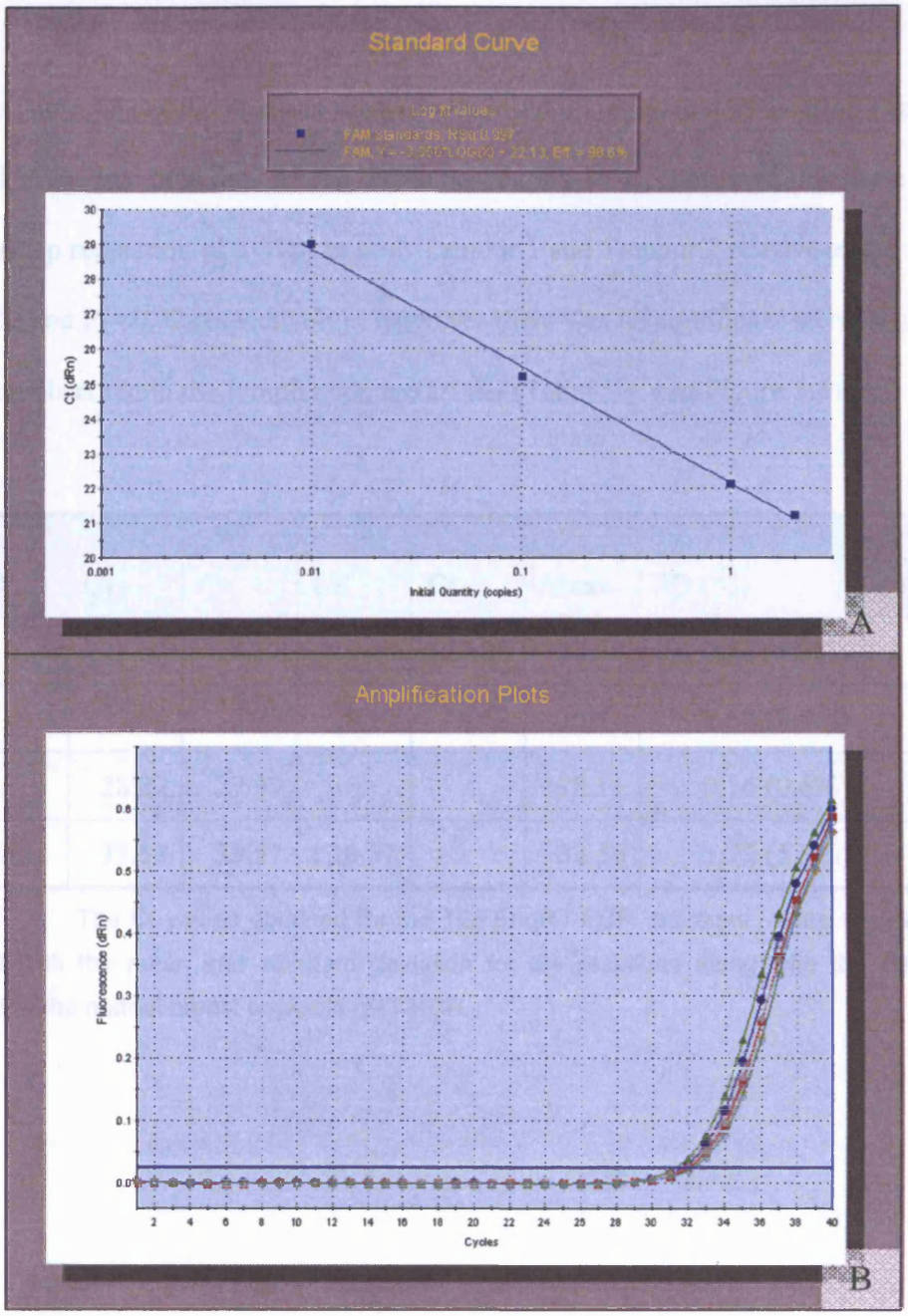

Figure 5.10 Validation of β-actin qRT-PCR between different cDNA samples. (A) Amplification plots for HBL-100 cDNA (2µl, 1µl, 0.1µl and 0.01µl). (B) Ct values for β-actin qRT-PCR using different cDNA templates.

5.4.4.2 TGFβI

The Affymetrix GeneChip analysis showed a 3.5 fold up regulation in the gene expression of TGFβI in the two tumours. SYBR Green based qRT-PCR, confirmed that there was a significant up regulation of TGFβI in both Tumour 1 and Tumour 2 relative to the organoid ($P = 0.002$ and $P = 0.02$ respectively). However there was no significant increase in expression observed in the lymph node metastasis (Table 5.4. and Figure 5.11).

TGFβI	Ct	Ct	Ct	Ct	Mean	SD (%)	Fold Change
T1	28.64	28.37	28.46	27.88	28.34	0.33 (1.2%)	18.62
LN1	32.64	32.71	28.18	28.10	30.41	2.62 (8.6%)	4.44
T2	28.22	27.99			28.11	0.16 (0.6%)	21.88
Organoids	33.53	33.57	30.57		32.56	1.72 (5.3%)	1.00

Table 5.4 The Ct values obtained for the TGFβI qRT-PCR reactions. Here the raw data is presented with the mean and standard deviation for the reactions along with the fold change compared to the normal breast organoid qRT-PCR.

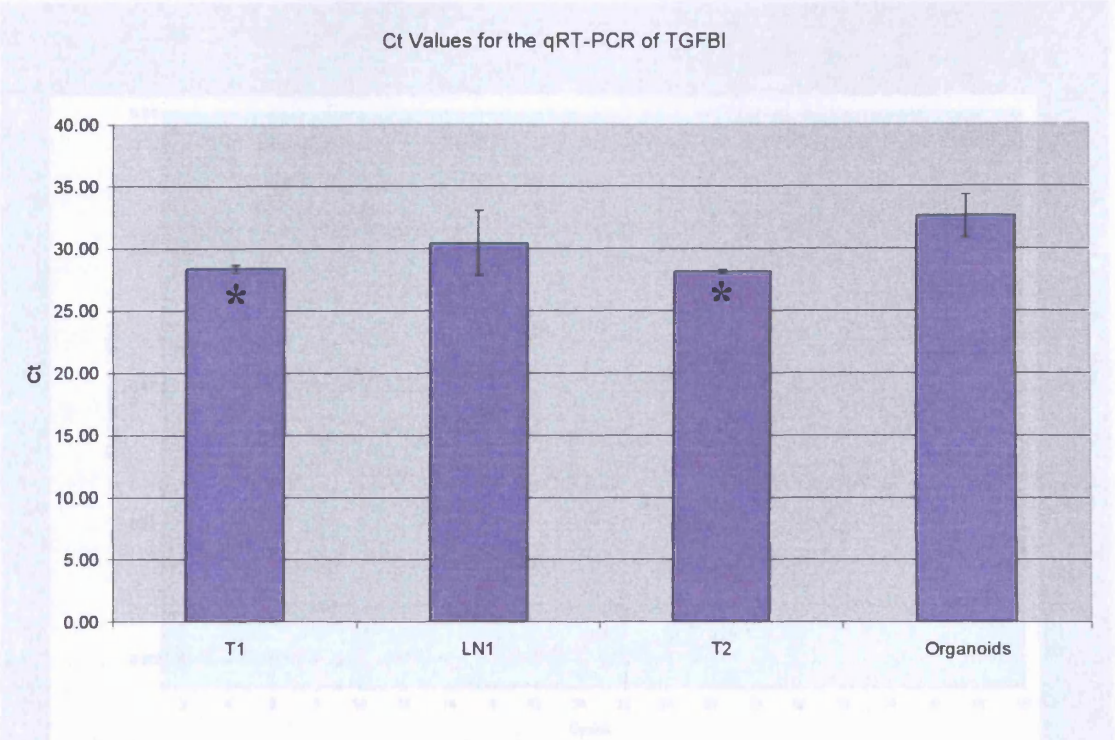

Figure 5.11 (previous) Comparison of TGFβI expression between tumours and normal organoid by qRT-PCR. The graph shows the mean Ct value (± standard deviation) of the reaction. Significant differences ($p \leq 0.05$) between the test reactions and the organoid control are marked with an asterisk (*).

5.4.4.3 RBBP4

For RBBP4, the GeneChip analysis predicted a 2 fold up regulation in gene expression in the tumour samples compared to the breast organoid. SYBR Green chemistry was used for RBBP4 qRT-PCR. The primary data are shown for duplicate reactions in Figure 5.12 and Table 5.5. These results show that there was good reproducibility between replicate reactions. The qRT-PCR data showed comparable results to the array for T2 (2.24, P = 0.026), however T1 showed a much larger fold change (approximately 16 fold, P = 0.013). Interestingly, the lymph node metastasis from T1 showed a huge up regulation, with a 192 fold increase relative to the organoid (P = 0.002) (Figure 5.13).

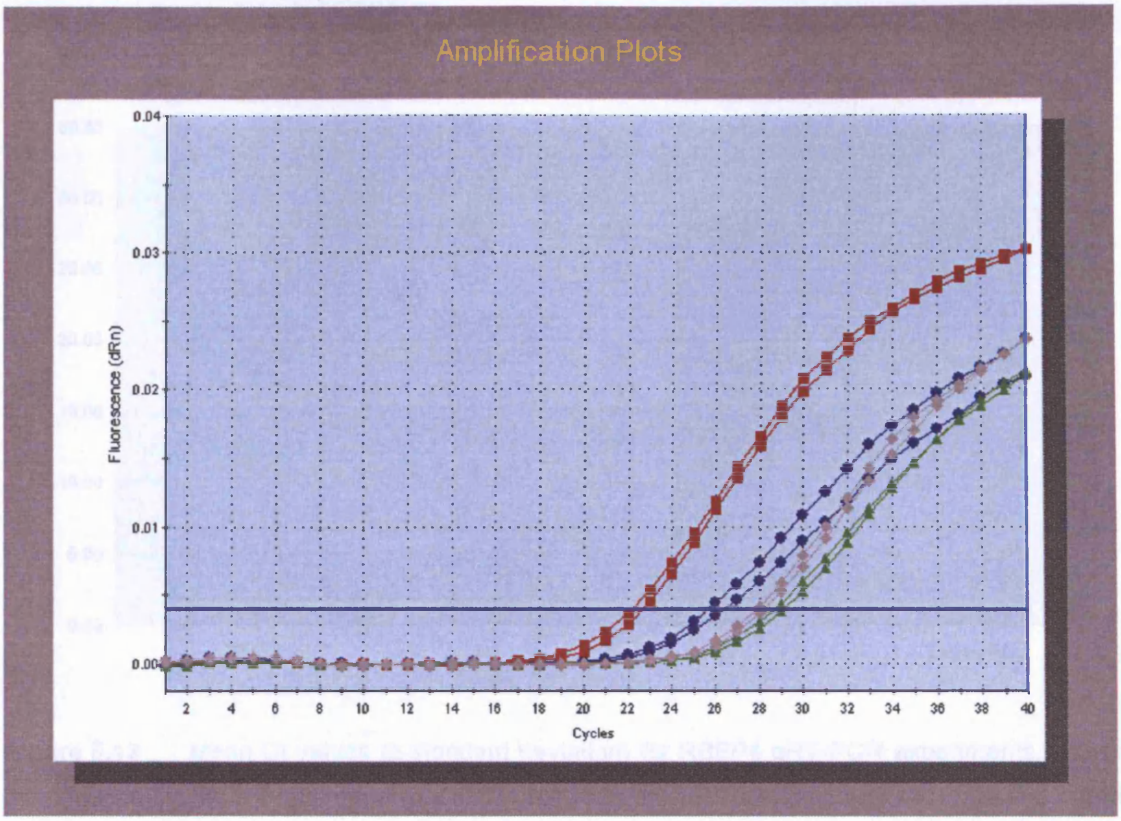

Figure 5.12 (Previous) PCR amplification plot for the RBBP4 qRT-PCR reactions. The horizontal blue line shows the point at which PCR products are first accurately detectable above the background. The cycle number at which the curve crosses this line defines the Ct value for that reaction. (Red is LN1, blue is T1, grey is T2, green is Breast Organoid)

RBBP4	Ct	Ct	Mean	SD (%)	Fold Change
T1	25.24	25.93	25.59	0.49 (1.9%)	16.51
LN1	22.29	21.79	22.04	0.35 (1.6%)	192.67
T2	28.67	28.26	28.47	0.29 (1.0%)	2.24
Organoids	29.47	29.79	29.63	0.23 (0.8%)	1.00

Table 5.5 Ct values for the RBBP4 qRT-PCR reactions. Here the raw data is presented with the mean and standard deviation for the reactions along with the fold change compared to the normal breast organoid qRT-PCR.

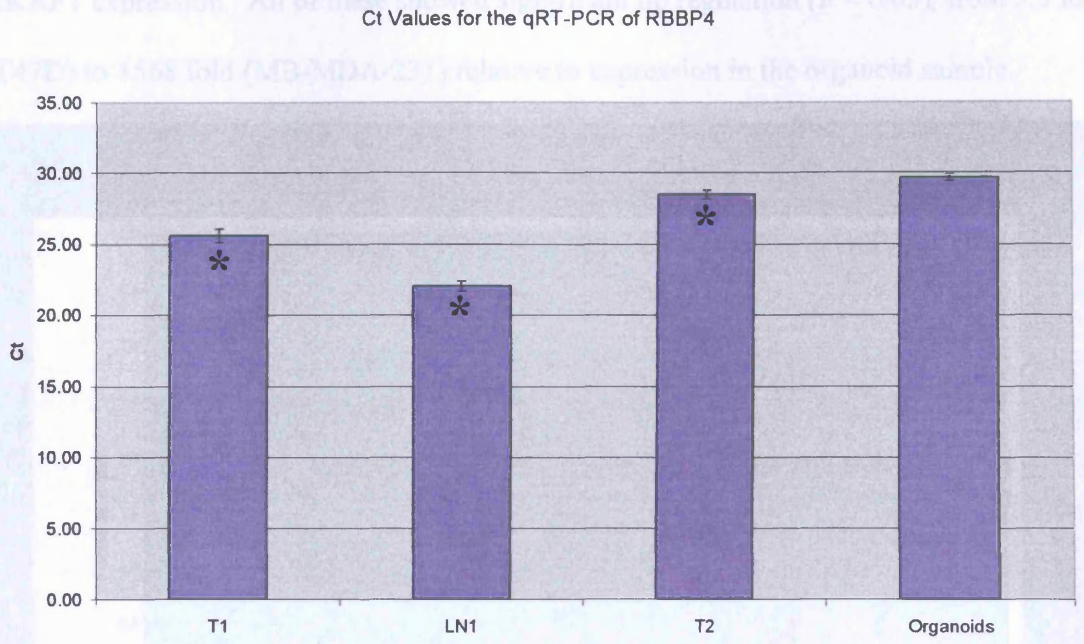

Figure 5.13 Mean Ct values (± standard deviation) for RBBP4 qRT-PCR experiments. Where there was a significant difference ($p \leq 0.05$) between the test reactions and the organoid control this is marked with an asterisk (*).

5.4.4.4 AKAP1

A validated commercially available TaqMan primer and probe set was used to investigate AKAP1 gene expression by qRT-PCR. Duplicate reactions showed good correlation for Ct values (Figure5.14). The primary microarray data suggested a 4 fold increase in expression of AKAP1, in the two tumours compared to the organoid RNA. However, analysing each tumour individually using TaqMan® qPR-PCR showed different results. Tumour 2 did not express AKAP1 at all and Tumour 1 was upregulated by 41 fold ($P = 0.015$). The lymph node metastasis showed a slight, but significant ($P = 0.039$), decrease in the level of expression relative to the organoid (Table 5.6 and Figure 5.15). There is no obvious reason for these significant differences and it is difficult to investigate this further using two different commercial systems. Five breast cell lines were also examined for

AKAP1 expression. All of these showed significant up regulation ($P < 0.05$), from 5.5 fold (T47D) to 1568 fold (MB-MDA-231) relative to expression in the organoid sample.

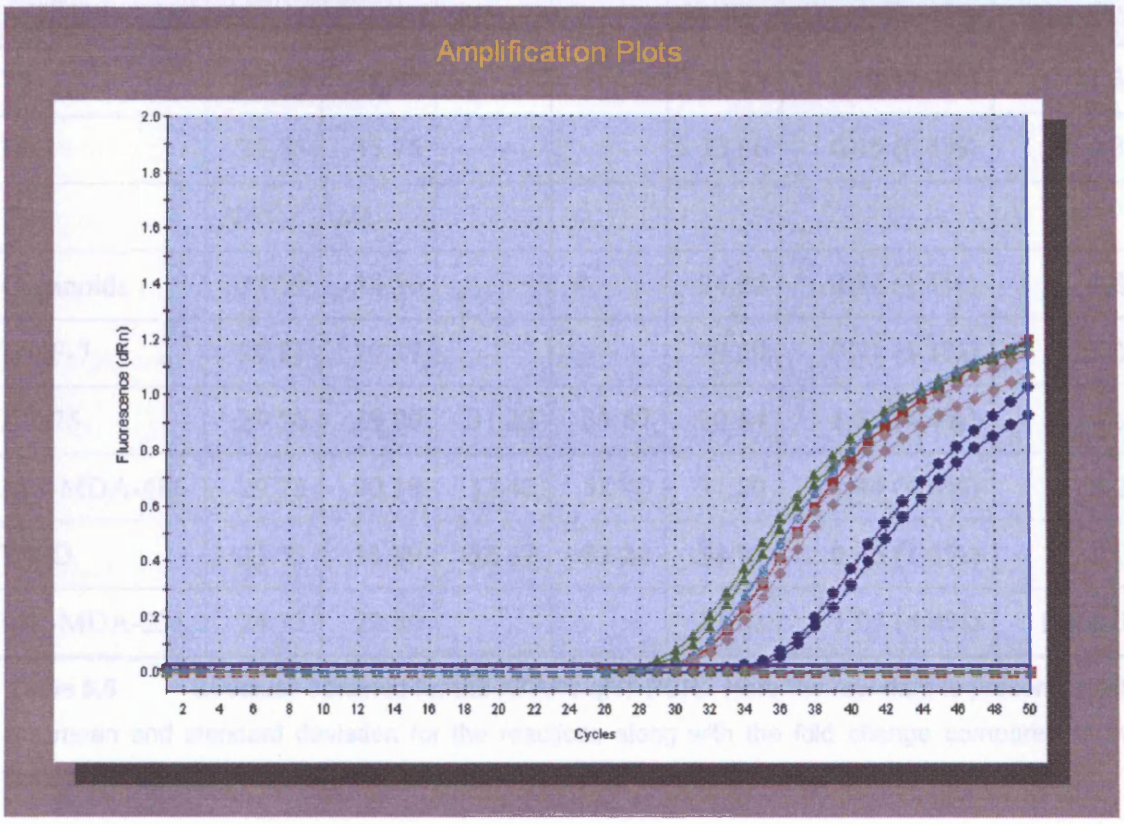

Figure 5.14 PCR amplification plot for the AKAP1 qRT-PCR experiments. The horizontal blue line shows the point at which PCR products are first accurately detectable above the background. The cycle number at which the curve crosses this line defines the Ct value for that reaction. (Green is MCF-7,cyan is ZR-75, red is MB-MDA-486, grey is T47D and blue is LN1)

AKAP	Ct	Ct	Ct	Ct	Mean	SD (%)	Fold Change
T1	28.88	28.87			28.88	0.01 (0.0%)	41.36
LN1	35.55	35.76			35.66	0.15 (0.4%)	0.38
T2	*NA*	*NA*					
Organoids	33.99	34.50			34.25	0.36 (1.1%)	1.00
MCF-7	29.81	29.37			29.59	0.31 (1.1%)	25.19
ZR-75	29.56	29.09	31.22	31.87	30.44	1.32 (4.4%)	14.03
MB-MDA-486	29.73	30.19	32.46	32.40	31.20	1.44 (4.6%)	8.28
T47D	31.53	31.49	32.37	32.24	31.91	0.46 (1.5%)	5.05
MB-MDA-231	24.37	22.89			23.63	1.05 (4.4%)	1,568.32

Table 5.6 Ct values obtained for the AKAP1 qRT-PCR. Here the raw data is presented with the mean and standard deviation for the reactions along with the fold change compared to the normal breast organoid qRT-PCR (where *NA* indicates no amplification).

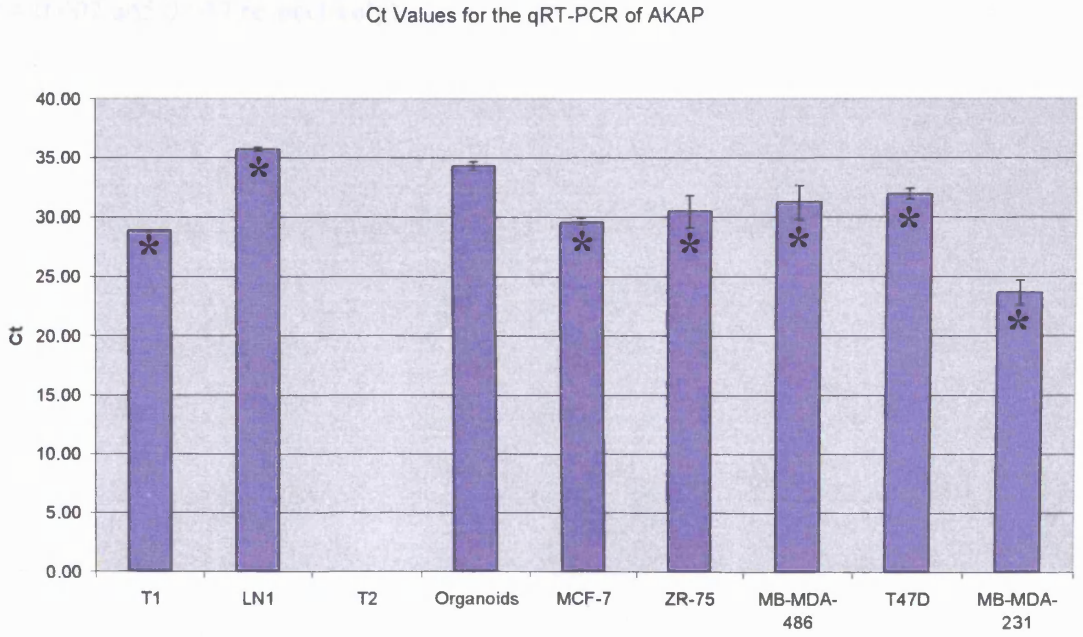

Figure 5.15 (previous) A graph showing the mean Ct value (± standard deviation) for the AKAP1 qRT-PCR. Where there is a significant difference ($p \leq 0.05$) between the test reactions and the organoid control it is marked with an asterisk (*).

5.4.4.5 C/EBPA

Gene expression of C/EBPA was also determined using TaqMan® qRT-PCR. Amplification plots of Ct values for duplicate reactions using T1, T2, LN1 and the organoid samples showed good reproducibility (Figure 5.16). The primary microarray data had suggested approximately a 3 fold up regulation in expression in the two tumours. TaqMan® qRT-PCR confirmed this but the level of expression in the two tumours was significantly higher (51 fold and 170 fold (mean) up regulation). However due to wide variability in the replicates for tumour 1 this only proved to be significant for Tumour 2 ($P = 0.002$) (Table 5.7 and Figure 5.17). The lymph node metastasis also showed significant up regulation (125 fold, $P = 0.002$). Of the five cell lines investigated, MB-MDA-486, T47D and MB-MDA-231 did not show a significantly different level of expression to the organoid. Expression in MCF-7 and ZR-75 was significantly higher than in the organoid ($P = 0.002$ and 0.037 respectively).

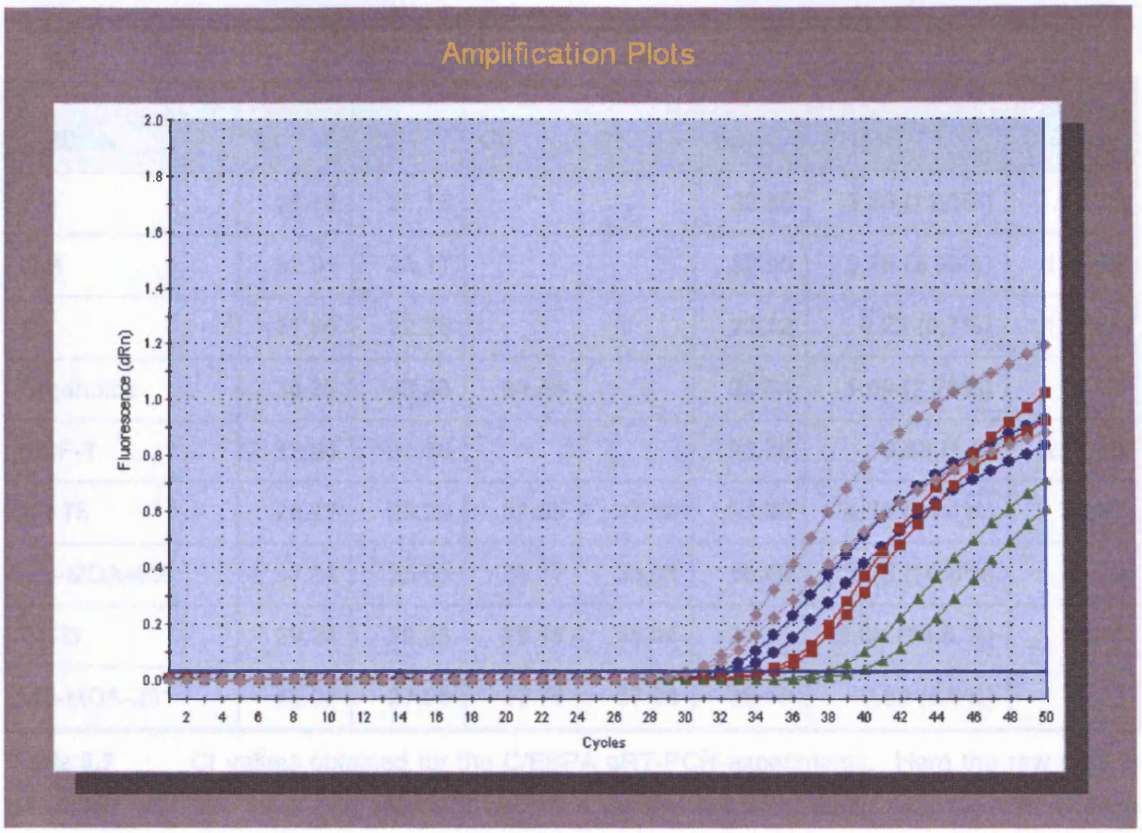

Figure 5.16 PCR amplification plot for the C/EBPA qRT-PCR reaction. The horizontal blue line shows the point at which PCR products are first accurately detectable above the background. The cycle number at which the curve crosses this line defines the Ct value for that reaction. (Grey is ZR-75, blue isMCF-7, red is MB-MDA-486 and green is T47D)

C/EBPA	Ct	Ct	Ct	Ct	Mean	SD (%)	Fold Change
T1	36.56	31.15			33.86	3.83 (11.3%)	51.33
LN1	32.01	33.11			32.56	0.78 (2.39%)	125.95
T2	31.96	32.28			32.12	0.23 (0.7%)	170.86
Organoids	38.39	40.56	39.66		39.54	1.09 (2.76%)	1.00
MCF-7	32.04	31.36			31.70	0.48 (1.5)	228.60
ZR-75	29.27	29.29	37.65	37.18	33.35	4.70 (14.1%)	72.97
MB-MDA-486	33.34	33.06	39.77	39.63	36.45	3.75 (10.3%)	8.50
T47D	29.24	29.33	39.18	38.94	34.17	5.64 (16.5 %)	41.19
MB-MDA-231	43.08	37.98	37.75	37.95	39.19	2.60 (6.6%)	1.27

Table 5.7 Ct values obtained for the C/EBPA qRT-PCR experiments. Here the raw data is presented with the mean and standard deviation for the reactions along with the fold change compared to the normal breast organoid qRT-PCR

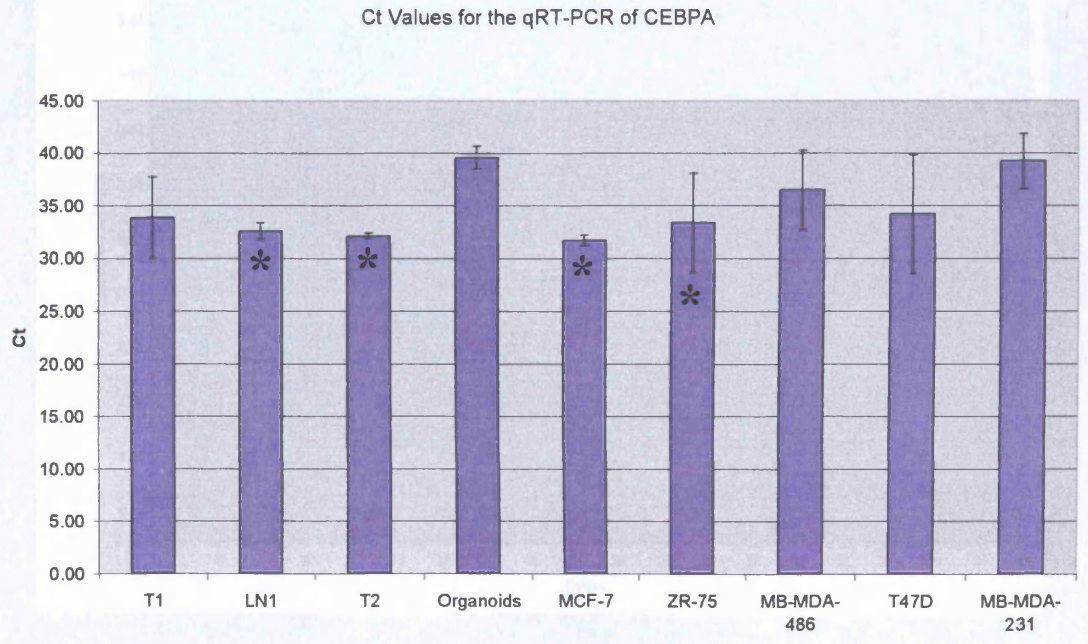

Figure 5.17 (previous) A graph showing the mean Ct value (± standard deviation) for the C/EBPA qRT-PCR. Where there is a significant difference ($p \leq 0.05$) between the test reactions and the organoid control it is marked with an asterisk (✻).

5.4.4.6 DDB2

GeneChip analysis predicted that DDB2 was up regulated by approximately 2.3 fold in the two tumours relative to the organoid RNA. However, using SYBR Green qRT-PCR a more substantial change was seen in both tumours (18 ($P < 0.001$) and 7-fold ($P < 0.008$) for T1 and T2 respectively). In addition the matching lymph node metastasis for Tumour 1, showed a 11-fold up regulation of the mRNA expression ($P < 0.001$). The Ct values for duplicate reactions are presented in Figure 5.18 and Table 5.8. The results are presented graphically in Figure 5.19.

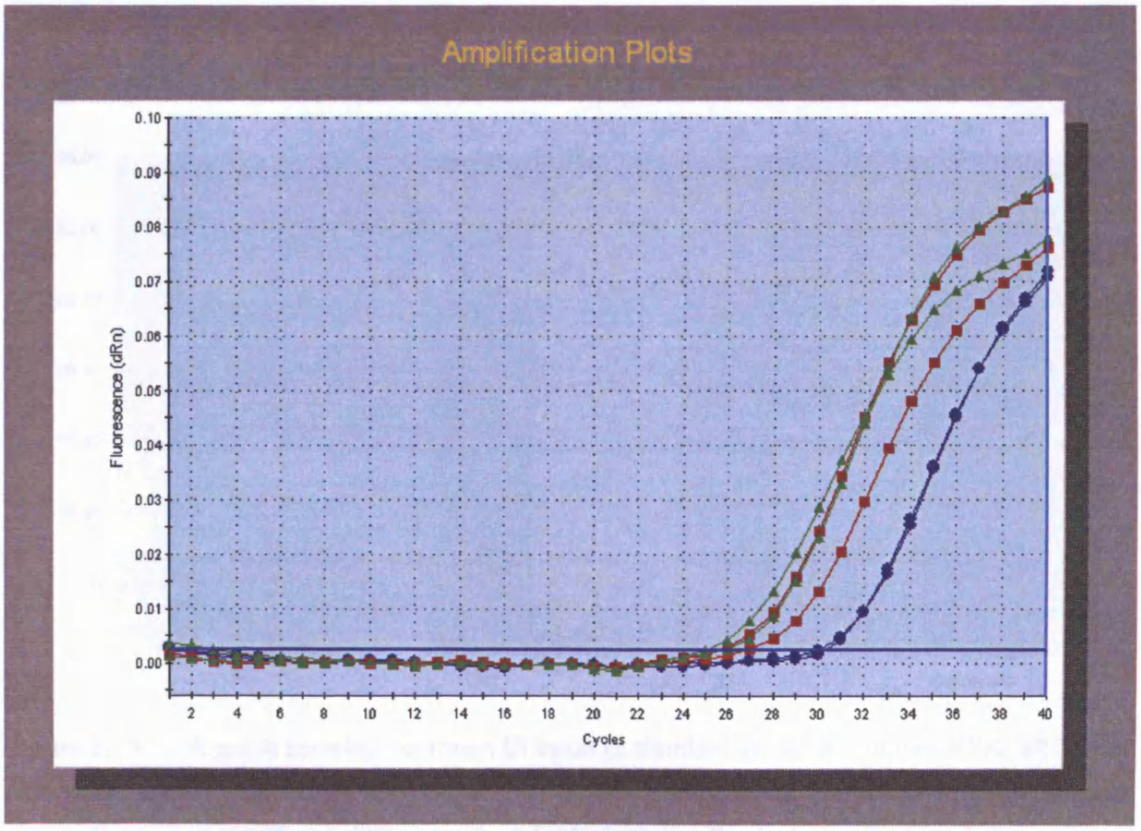

Figure 5.18 (previous) PCR amplification plot for the DDB2 qRT-PCR reaction. The horizontal blue line shows the point at which PCR products are first accurately detectable above the background. The cycle number at which the curve crosses this line defines the Ct value for that reaction. (Green is T1, red is LN1 and blue is breast organoids)

DDB2	Ct	Ct	Ct	Ct	Ct	Ct	Mean	SD (%)	Fold Change
T1	25.02	25.53	25.18	25.13	23.66	24.19	24.79	0.71 (2.9%)	18.17
LN1	27.34	25.08	23.89	25.88	25.26	25.46	25.49	1.13 (4.4%)	11.18
T2	25.27	25.41	27.23	27.22			26.28	1.09 (4.2%)	6.43
Organoids	30.68	30.97	28.06	28.24	28.50	27.36	28.97	1.49 (5.1%)	1.00

Table 5.8 Ct values obtained for the DDB2 qRT-PCR. Here the raw data is presented with the mean and standard deviation for the reactions along with the fold change compared to the normal breast organoid qRT-PCR

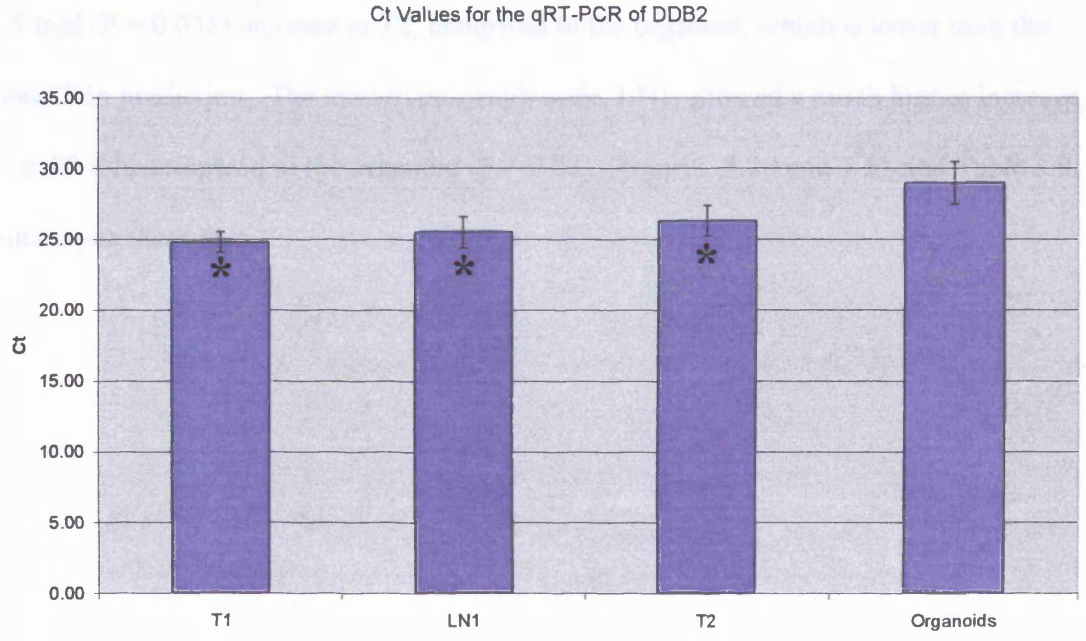

Figure 5.19 A graph showing the mean Ct value (± standard deviation) for the DDB2 qRT-PCR. Also noted above each bar is the relative fold change compared to the breast organoid control. Where there is a significant difference ($p \leq 0.05$) between the test reactions and the organoid control it is marked with an asterisk (*).

Analysis of DDB2 gene expression has recently been investigated by another student in the laboratory (Mr Oliver Ng). He was able to investigate expression in 8 other tumour cases. Increased expression was found in all cases but with wide variability in the level of expression. There were 3 cases less than 35 years in age. The fold increase ranged from 11.10 to 24.06 with a mean of 18.97. In 5 cases from women over 50 years of age the up regulation was significantly lower, ranging from 1.54 to 9.01 with a mean of 3.54. These differences were statistically significant (P = 0.0043). These preliminary data support my findings and may suggest DDB2 as a novel marker for breast cancers in young women and warrants further investigation.

5.4.4.7 GRN

The predicted up regulation of GRN mRNA was 5.9 fold from the Affymetrix GeneChip experiments. SYBR Green qRT-PCR showed a 16.0 fold increase (P = 0.023) in T1 and a 3.5 fold (P.= 0.035) increase in T2, compared to the organoid, which is lower than the GeneChip prediction. The metastatic lymph node, LN1, showed a much higher increase, over 80 fold compared to the organoid (P = 0.01). Figures, 5.20 and 5.21 and Table 5.9 summarise these data.

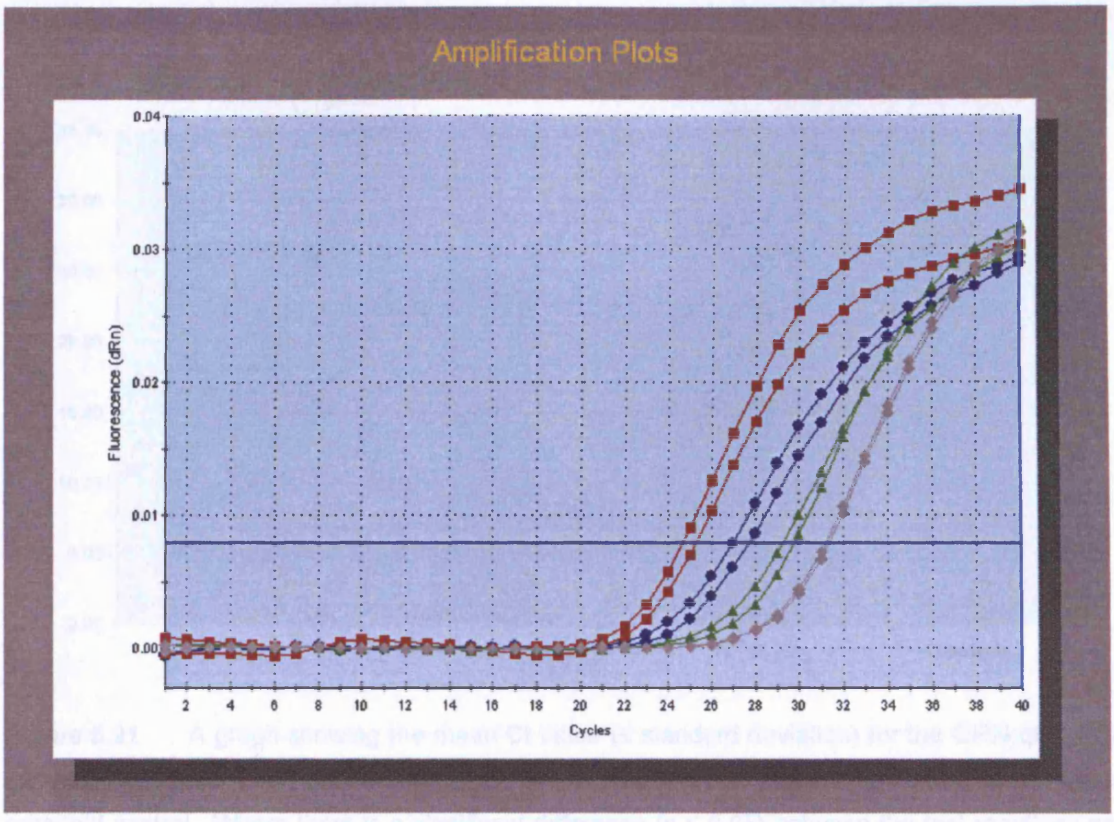

Figure 5.20 PCR amplification plot for the GRN qRT-PCR experiments. The horizontal blue line shows the point at which PCR products are first accurately detectable above the background. The cycle number at which the curve crosses this line defines the Ct value for that reaction. (Red is LN1, blue is T1,green is T2 and grey is breast organoid)

GRN	Ct	Ct	Mean	SD (%)	Fold Change
T1	27.39	26.67	27.03	0.51 (0.5%)	15.94
LN1	24.40	25.00	24.70	0.42 (0.4%)	80.17
T2	29.50	28.98	29.24	0.37 (0.4%)	3.45
Organoids	31.11	30.94	31.03	0.12 (0.1%)	1.00

Table 5.9 Ct values obtained for the GRN qRT-PCR. Here the raw data is presented with the mean and standard deviation for the reactions along with the fold change compared to the normal breast organoid qRT-PCR

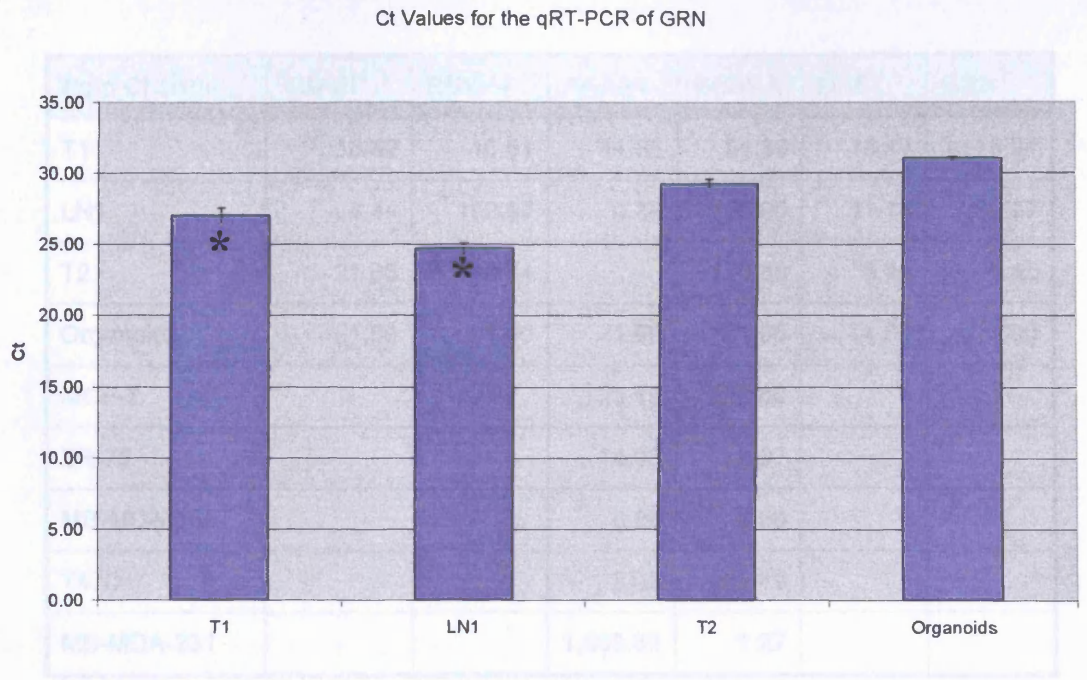

Figure 5.21 A graph showing the mean Ct value (± standard deviation) for the GRN qRT-PCR experiments. Also noted above each bar is the appropriate fold change compared to the breast organoid control. Where there is a significant difference ($p \leq 0.05$) between the test reactions and the organoid control it is marked with an asterisk (*).

5.4.4.8 qRT-PCR Summary

In summary qRT-PCR analysis confirmed that the 6 selected genes showed elevated expression in the two tumours relative to the organoid RNA samples However the Affymetrix GeneChip analysis appeared to underestimate this change as the qRT-PCR fold changes were usually greater . Table 5.10 shows the fold changes obtained for the two tumours, metastatic lymph node and 5 breast cell lines.

Fold Change	TGFBI	RBBP4	AKAP1	CEBPA	DDB2	GRN
T1	18.62	16.51	41.36	51.33	18.17	15.94
LN1	4.44	192.67	0.38	125.95	11.18	80.17
T2	21.88	2.24		170.86	6.43	3.45
Organoids	1.00	1.00	1.00	1.00	1.00	1.00
MCF-7			25.19	228.60		
ZR-75			14.03	72.97		
MB-MDA-486			8.28	8.50		
T47D			5.05	41.19		
MB-MDA-231			1,568.32	1.27		

Table 5.10 Summary of fold change identified for 6 selected genes in two tumours, metastatic lymph node and 5 breast cell lines identified by qRT-PCR.

5.4.5 RBBP4 Immunohistochemistry

In order to further validate the gene expression changes it was desirable to examine protein expression as well as mRNA levels. This work was carried out towards the end of the research project and there was only sufficient time to begin protein analysis. A product search was carried out and commercial antibodies were only available for RBBP4 and DDB2. The DDB2 antibody was suggested to be appropriate for Western blotting but not tissue studies. As it was not possible to obtain protein from fresh tissue it was therefore decided to perform a limited immunohistochemistry analysis in tumour tissue sections using the RBBP4 antibody only.

A series of experiments were carried out with the support of Mrs Karen Kulbicki, a research BMS2 in the group, to optimise the conditions for immunohistochemistry. This involved varying the antibody dilution and antigen retrieval with different times and pressure cooking (Section 2.3.2.5) (Data not shown).

A total of 28 tumour cases were examined, which comprised 16 cases <36 years of age and 12 cases >55 years. The RBBP4 staining was independently semi-quantified using a modified version of the quick scoring method. The intensity of the staining (Score_I) and the percentage of cells that stained positively (Score_C) were observed with the final score (Score_T) being the sum of these (Table 5.11). Figure 5.22 shows representative staining from tissue with weak (A), moderate (B) and strong (C) nuclear staining in greater than 70% of the cells. A common feature was that malignant cells were more likely to be positive than normal breast within the same section as shown in Figure 5.22D. The summary of staining patterns is presented in Table 5.11. These data were analysed using the Mann-Whitney statistical test (SPSS v.12), but there were no significant differences identified between young and older onset cases (P = 0.330, 0.208 and 0.151 for Score_I, Score_C and Score_T respectively)

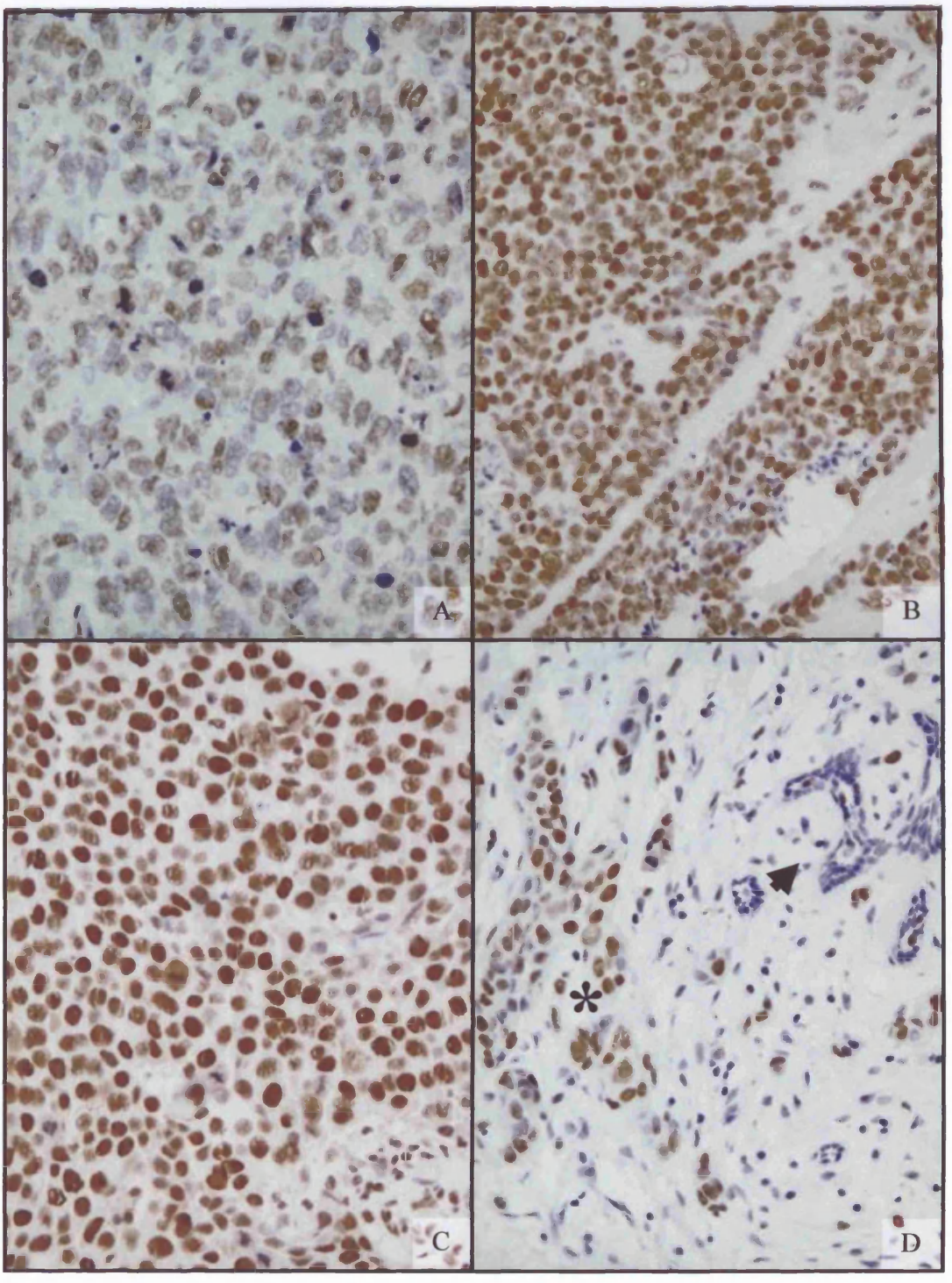

Figure 5.22 RBBP4 immunohistochemistry in breast tumour tissue sections. Nuclear staining was observed with varying intensity. Panel (A) shows weak staining, panel (B) shows moderate staining and panel (C) shows strong staining. Panel (D) illustrates cancer (*) staining positive with the normal (◀) negative for RBBP4.

Young onset cases (<36 years)			
Case	Score_C	Score_I	Score_T
RW2200	1	3	4
RW2796	2	2	4
RW2151	2	2	4
RW2019	2	2	4
RW2215	3	2	5
RW2055	3	2	5
RW2204	3	3	6
RW2201	3	3	6
RW2144	3	3	6
RW2137	3	3	6
RW2006	3	3	6
RW1941	3	3	6
RW2200	5	2	7
RW2157	4	3	7
RW2195	5	3	8
RW2105	5	3	8

Older onset cases (>55 years)			
Case	Score_C	Score_I	Score_T
B30	1	1	2
RW1930	1	1	2
B81	3	1	4
178	2	2	4
53	3	1	4
RW2114	1	3	4
B115	2	3	5
177	3	3	6
6	3	3	6
38	4	2	6
61	4	3	7
3	4	3	7

Table 5.11 Summary of the staining pattern observed for the RBBP4 immunohistochemistry in breast tumour cases.

5.5 Discussion

5.5.1 Aims and Limitations

The initial aim of this chapter was to perform array CGH and array expression analysis, both of which were to be confirmed using qRT-PCR. Any interesting candidate genes were then to be examined at the protein level using immunohistochemistry.

Due to several technical problems it was not possible to fully achieve these aims. The first of these resulted from a production error in the in house cDNA arrays used for CGH analysis. Unfortunately it took several months of clone validation before it was apparent that approximately 30% of the sequences tested were either incorrect or contaminated by other clones. This meant that there was no accurate array CGH data to guide the project. As discussed in Section 5.4.3.1, the Affymetrix GeneChip analysis was also less than ideal. Financial constraints limited the number of GeneChips available, the frozen archived samples were of a poor quality insufficient for use to extract RNA and unfortunately time restrictions limited the prospective collection to 2 young breast cancer cases. The time constraints also limited the number of samples and genes that could be verified by qRT-PCR. Since completing my experimental work a number of the targets have been followed by other members of the group and comment will be included where appropriate in this discussion.

5.5.2 Chromosome 17q Amplification

Since there were a number of problems with the original in house array it was difficult to follow up the CGH data. Therefore it was decided to investigate a number of genes in the 17q12-21 interval for co-amplification using semi-quantitative PCR. However, there were no significant differences in the level of expression observed in the cDNA samples analysed. This may result from inter-tumoural heterogeneity or may reflect errors in the semi-qualitative measurement using β-actin as the reference gene. From this investigation there is no strong evidence to suggest that expression of the genes analysed from chromosome 17q12-21 (genes HER-2, Crk-5, PSMB3, PPARBP, ZNF-144 and GRB7) differs in young tumours compared to either normal tissue or breast cell lines. Obviously, due to the array contamination errors discussed previously it has not been possible to compare the gene copy number changes from these young women.

Whilst semi-quantitative RT-PCR is not the most sensitive technique available to show the relative differences in gene expression it is sensitive enough to detect when there are any differences to more accurately quantify.

From reviewing the published literature it is clear that amplification of the chromosomal region 17q12-21 often plays an important role in the initiation and/or propagation of breast cancer (Winston *et al*, 2004; Nessling *et al*, 2005; Reyal *et al*, 2005; Mohsin *et al*, 2005; Rudlowski *et al*, 2005). However there is a less clear consensus of opinion as to the important genes within this region, which maintain the amplification. The most effective way to investigate such a large area of the genome would be through comparative genomic hybridisation on a custom array containing all the genes in the 1.5Mb region that was initially found to be amplified on the MRC cDNA array. The gene expression analysis would also be best performed on a similar custom array. Should the cohort be sufficiently large, and composed of tumour tissue and matching microdissected surrounding normal breast it would be possible to determine which genes are consistently amplified and which are consistently up regulated. This would obviously be best linked to clinical outcome data to fully investigate the significance of the gene amplifications and overexpression.

5.5.3 Affymetrix GeneChip Analysis

The 4 selected RNA samples were successfully labelled and subsequently hybridisation to the eight Affymetrix GeneChips. Of the 441 genes found to be differentially regulated in Tumour 1 and Tumour 2 compared to breast organoids, there were many that showed potential as tumour specific markers. The six genes that were investigated further show some potential and will be discussed separately. There were, however, a number of genes that appeared to show differences that were probably the consequence of sample variation when comparing whole tumour tissue and isolated breast organoids which predominantly comprise myoepithelial cells and luminal epithelial cells. Examples include the leukocyte

immunoglobulin-like receptor gene, which appeared to be upregulated in the tumour samples (by SAM analysis), but this was most probably due to an inflammatory infiltrate into the tumour. A second example would be SMA5. The apparent decrease in the expression of this gene is actually more likely to be due to the fact that there was a more diverse cell population in the tissue compared to the organoids.

Due to the structural complexity and variability of normal breast, with its naturally infiltrating ductal structure, whole breast samples were not a viable option for extraction of "normal" breast cell RNA. Enzymatically extracted breast organoids are an enriched population of breast epithelial and myoepithelial cells. While it would be ideal to isolate exclusively the epithelial cells this was not possible without culture and so the samples selected were the most suitable controls available at the time.

The only protocols currently approved by Affymetrix, all require large amounts of full length mRNA with an intact poly A tail. While this high quality mRNA is available for extraction from cell lines, this is usually not the case from frozen archival material. The cost of reagents and GeneChips make it unrealistic for publicly funded research labs to optimise an alternative. This said this should be achievable for Affymetrix to provide an alternative protocol for mRNA from tissues. The problems of poly A tail loss could be over come by using random primers rather than oligo dT primers. The addition of dideoxy-terminators could limit the labelled target to approximately 100 bp, which would also reduce the effect of mRNA fragmentation. The use of techniques such as this also make the use of mRNA from laser capture dissected material possible.

5.5.4 Gene expression profiling by qRT-PCR

Almost uniformly, across all the genes investigated and on both tumour samples investigated there was quite a large discrepancy in the mRNA up regulation determined by the Affymetrix GeneChip and the qRT-PCR. This is most probably due to the nature of

the RNA extracted. While identical methodologies were followed (using TriReagent) the origin of the cells differed. The organoids were taken from a cell suspension in tissue culture media and digestive enzymes but produced very high quality RNA. Tumour RNA however, was extracted from grossly dissected material, which was manually agitated and disrupted in the TriReagent. While the RNA was high quality, with little degradation and so suitable for the GeneChip, it was not of the same quality as the organoid RNA. While the Affymetrix labelling protocol does try to compensate for inefficient labelling of the target RNA, it is primarily designed for cell line RNA, which is more likely to be of a very high quality and full length. This difference could therefore explain why the GeneChip under estimated the changes in mRNA expression compared to qRT-PCR, as there was potentially slightly less labelled tumour RNA than organoid RNA. As the sample quality requirements for qRT-PCR are less stringent this was not an issue for the qRT-PCR and more accurate fold changes were observed.

It would be premature to presume that the genes identified play a significant role in breast cancer initiation or progression. However, from this study and previous investigations of these genes and their products it is clear that there is a possible role for these genes in the aetiology of breast cancer in young women.

5.5.5 Target gene analyses

5.5.5.1 TGFβI

Mapped to 5q31, as the name suggests the expression of this gene is induced by TGFβ1 activation. First identified in an adenocarcinoma cell line, TGFβI is a 693 amino acid protein, that contains a fas1 domain. The terminal Arg-Gly-Asp of this domain shows an $\alpha_3\beta_1$ integrin binding motif (Skonier *et al*, 1992; Nam *et al*, 2003).

It has been well documented that $\alpha_3\beta_1$ integrin, through interactions with other proteins in the cell membrane and cytosol, mediates adhesion to the extracellular matrix (reviewed by Chrenek *et al*, 2001). TGFβI is actively synthesized during vessel remodelling, with tube forming endothelial cells showing strong staining at the contact zone between the endothelial cells and the extracellular matrix (Rawe *et al*, 1997). It is therefore conceivable that TGFβI could play a role in tumour mediated angiogenesis and metastasis through these interactions. While there have not been any previous reports of TGFβI involvement in breast cancer, it has been reported to be over expressed in colorectal cancers (Buckhaults *et al*, 2001).

5.5.5.2 RBBP4

The retinoblastoma protein is thought to suppress cell growth by binding to intracellular proteins. RBBP4 was identified by its ability to bind to the carboxyl-terminus of RB1. RBBP4 has a predicted size of 425 amino acid residues, and shows 30% homology to the *S. cerevisiae* MSI1 protein, thought to be a negative regulator to the ras signal transduction pathway. Both immunohistochemistry and subcellular fractionation studies have shown RBBP4 to be localised within the nucleus (Qian *et al*, 1993; Qian and Lee, 1995)

RBBP4 was found to be one of the three subunits of chromatin assembly factor 1 (Tyler *et al*, 1996; Zang *et al*, 1997). Moreover, RBBP4 was found to bind not only to the BRCT domain of BRCA1, but that this binding facilitates an association with the histone deacetylaters HDAC1 and HDAC2 (Taunton *et al*, 1996; Verreault *et al*, 1998; Yarden and Brody *et al*, 1999; Zhang *et al*, 2000).

The signalling mechanism responsible for activating genes through the cAMP regulated enhancer (CRE) is one of the most studies pathways (Montminy, 1997). Following the activation of activation of some G protein coupled receptors, the catalytic subunit of

protein kinase A (PKA) is released from the regulatory subunit and is transported to the nucleus. Here it phosphorylates the CRE binding transcription factor CREB. As CREB is phosphorylated at the same position by several protein kinases, including calcium-calmodulin, it has been proposed that CREB serves as a general activator of CRE containing sequences. Once phosphorylated, CREB is able to interact with the co-activator CREB binding protein (CBP). Significantly in this context, the association of phosphorylated CREB with CBP promotes the binding of RBBP4, providing a link between transcriptional activation and histone (de)acetylation (Shaywitz and Greenberg, 1999; Zhang *et al*, 2000).

Despite this, the mechanisms by which RBBP4 gene products modulate cell growth are still as yet unclear. RBBP4 mRNA has shown to be significantly increased in hepatocellular carcinoma (Song *et al*, 2004), and decreased in mucoepidermoid carcinoma, of the salivary glands (Leivo *et al*, 2005). No studies of breast cancer have been published to date.

It would therefore be plausible that increased RBBP4 gene expression (and increased protein expression as suggested by the immunohistochemistry) could provide a direct link between BRCA1, cAMP based transcriptional activation, and histone (de)acetylation, which promote or propagate tumourigenesis.

5.5.5.3 AKAP1

In eukaryotic cells, cytoplasmic cAMP activates several isoforms of cAMP-dependent protein kinases (PKAs) involved in signal transduction. PKAs are composed of 2 regulatory and two catalytic subunits. Of these regulatory subunits there are two classes: type I are found in cytoplasmic PKAs, and type II found to associate with the particulate fraction of cell lysates. The effects of individual PKA isoforms is determined by their cellular location, specified by binding to distinct A-kinase anchor proteins, which tether the

kinase to the regulatory subunit. Therefore altering the intracellular AKAP protein composition though altered gene expression could result in unintentional cAMP dependent effects such as enhanced proliferation (Huang *et al* 1997; Houslay and Adams, 2003).

5.5.5.4 C/EBPA

Mapped to the 19q13.1, CEBPA is has been characterised mostly in relation to the haematopoietic system (Szpirer *et al* 1992; Fröhling *et al* 2004). Here, CEBPA is exclusively expressed by myelomonocytic cells, and is upregulated during granulocytic differentiation. In fact, in CEBPA null mice, no mature granulocytes are observed, whereas all other cell types are unaffected (Pabst *et al*, 2001).

CEBPA is poorly characterised in solid tumours but has been under investigation in relation to acute myeloid leukaemia (AML) for many years. Of AML-M2 patients without (8, 21) translocation, 16% showed a mutation in the CEBPA gene. In patients with an (8, 21) translocation this mutation rate was halved to about 8% (Helbing., 2004). Whilst the relevance of the translocation is unclear, further studies have shown that CEBPA directly interacts with, and inhibits, the kinases CDK2 and CDK4 resulting in an arrest of cell proliferation. This action is mediated by blocking the interaction of CDK2/4 with associated cyclins. Indeed, both Cdk2 and Cdk4 activity is increased in mouse Cebpa knockout livers, leaving to increased proliferation.

This block of cell proliferation by CEBPA would seem contradictory to facilitating cancer progression. One step to fully elucidate the role of CEBPA in breast cancer would be to sequence the whole gene to ensure that it is in fact wild type CEBPA gene product that is being over expressed (Wang *et al*, 2001).

5.5.5.5 DDB2

The DDB2 gene, mapped to 11p12.11, encodes a 48 kDa protein that localises to the cell nucleus. DDB2 is involved in both global genomic repair and nucleotide excision repair, especially involving repair of UV-induced DNA damage (Hwang et al. 1999, Tang et al, 2000) Within the 5′ UTR of the gene sequence, there is a p53 response element, with DDB2 mRNA levels being dependent on p53 basal levels, which also increases in response to DNA damage (Hwang et al, 1999). BRCA1 is able to up regulate DDB2 in response to both UV light and chemotherapy agents independently of p53 (Hartman and Ford, 2002). Moreover, in cells without DNA damage, DDB2 initiates cell cycle progression from G_1 to S phase by stimulating E2F1. Significantly increased levels of DDB2 could therefore be a result of p53 expression and/or DNA damage. In this study, qRT-PCR showed over expression of DDB2 in both primary tumours. Other studies in the research group have shown that this gene is also expressed at high levels in MCF-7, T47-D and ZR-75-1 cell lines. More interestingly a preliminary study has recently showed statistically significant differences in mRNA expression between breast cancer cases from younger (<35 years) and older (>50 years) women, with higher levels of expression seen in younger breast cancer cases ($P = 0.0043$). This is preliminary evidence to support the identification of novel gene targets specific to sporadic breast cancer in young women by this microarray study and the investigation should be followed up in a larger series of cases.

As with CEBPA, it would be advantageous to sequence the DDB2 gene in breast cancer cases to determine if a functional wild type gene is over expressed, or if the effects of DDB2 are related to a gene mutation.

5.5.5.6 GRN

Mapped to 17q21.32, the GRN protein is a member of a family of cys-rich proteins (Bhandari *et al*, 1992; Thornton *et al*, 1999). GRN is a growth factor that regulates the

dynamics of tissue development, growth and wound repair. This action is mediated by the phosphorylation of Shc and $_{p44/42}$MAPK in the ERK signal pathway as well as phosphorylating B/AKT protein kinase and p70 in the PI3K pathway (Zanocco-Marani *et al*, 1999).

GRN has previously been linked to breast cancer (Lu and Serrero, 1999; Lu and Serrero, 2000). It has been found that GRN is able to lead to oestrogen independence in ER positive cell lines, and as such GRN is highly expressed in ER negative cell lines such as MDA-MB-468 (Lu and Serrero, 2000). In ER positive cells it is thought that the ER is responsible for regulating GRN expression (Lu and Serrero, 1999). While the authors note that it could be a consequence rather than cause, GRN levels were twice that in unanchored cells compared to anchored cells, suggesting a potential role in cell invasion.

5.5.6 RBBP4

From the few sections examined, which contained both normal and tumour material, such as that in Figure 5.22(D), it is clear that RBBP4 protein was over expressed in tumour tissue compared to normal tissue. There has only been one other immunohistochemistry publication to date that has studied this protein. Fukuoka *et al* (2004) found a significant correlation in the up-regulation of RBBP4 and HDAC2 protein and mRNA expression in adenocarcinoma and squamous cell carcinoma of the lung, although this did not prove to be predictive. This is therefore a potential target in several cancer types that requires further investigation.

While it would have been ideal to assess the other genes investigated by qRT-PCR at the protein level, at the time of this research project there was only a commercially available antibody for RBBP4 that was suitable for use in immunohistochemistry. As other

candidates are investigated in ongoing studies it should hopefully be possible to investigate other targets at the protein level.

Chapter 6:

Conclusions and future perspectives

It is nearly 30 years since the first report focussing on the specific nature of breast cancer in young women (< 40 years) (Wallgren *et al,* 1977). Familial breast cancers are significant in this age group, however, the differing biology and nature of sporadic breast cancers in young women has also been noted and related to poor prognosis (Albain *et al,* 1994; Pillers 1992; Walker *et al,* 1996b; Wenger *et al,* 1993, Johnson *et al,* 2001). Indeed, there are several markers that can provide information about poor outcome. These include the expression of p53 (Barnes *et al,* 1993; Thor *et al,* 1992), overexpression of HER-2 (Gullick *et al,* 1991; Walker *et al,* 1996b), lack of the oestrogen or progesterone receptors (Foekens *et al,* 1989; Reiner *et al,* 1990) and higher levels of Ki-67 (Railo *et al,* 1993).

Previous data generated within the Breast Cancer Research Unit identified a high level of LOH at the BRCA1, BRCA2 and TP53 loci in breast cancers in young women (Johnson *et al.* 2002) and for BRCA1 and p53 this was related to protein expression (S. Johnson, PhD Thesis). In order to follow up the BRCA2 LOH, attempts were made to relate these data to BRCA2 mRNA and protein expression levels as part of this thesis. Unfortunately, this was not very successful. RT-PCR analysis only identified a small number of BRCA2 positive cases and the three commercial antibodies tested were shown to be inappropriate for use on tissue sections by immunohistochemistry, or for western blotting. In the future, it would be useful to revisit analysis of BRCA2 mRNA expression using qRT-PCR, which is more sensitive and accurate than semi-quantitative manual PCR methods. It should be possible to compare BRCA2 expression in cohorts of sporadic breast cancers, stratified by age and matched for type grade and node status. If new anti-BRCA2 antibodies become available, which are shown to be more reliable, these could also be used for tissue studies. Other studies, which

focus on epigenetic silencing of BRCA2 could also be initiated to investigate whether this mechanism is important for gene silencing in sporadic breast cancers in young women.

Analysis of skewed X-chromosome inactivation identified a statistically significant difference in the level of skewing in young breast cancer cases when compared to a similar cohort of older cases. These data support a previous published study (Kristiansen *et al*, 2002), which suggest a high level of skewing in young breast cancer cases. Taken together these data suggest that an X-linked gene may be responsible for this skewing, and this could be investigated further. Other researchers have previously reported allelic loss on the X-chromosome, but no significant linkage has been identified. Moreover, the microarray analysis carried out in this thesis did not identify any significant targets from the X-chromosome and so it did not seem sensible to pursue these studies further at this stage.

The cDNA microarray study was most successful. A number of novel gene targets have been identified that show high levels of expression in the two breast cancers from young women when compared to normal breast organoids using Affymetrix arrays. Six genes were selected for further study and these were followed up by qRT-PCR in the original RNA samples, and a number of other samples. These studies largely confirmed the primary microarray data and showed significant up regulation in tumours compared to organoid controls, at a level higher than shown in the primary microarray. Most notably, the results for DDB2 generated as part of this thesis, and in subsequent studies in the research group suggest that DDB2 expression is greater in sporadic breast cancers in young women when compared to older women. Very recently, a second target RARRES3 (retinoic acid receptor responder 3) has also been shown to be more highly expressed in young breast cancer cases (Sinead Lambe, personal communication). This finding is novel, since no cancers previously investigated have shown

over expression of RARRES3 (Jiang et al, 2005). However, analysis of a larger cohort of sporadic breast cancers is needed to validate these findings. The research group are currently collaborating with colleagues in London to enable the analysis of a large series of breast cancer cases from young women, with known BRCA gene status.

If these two genes or any other novel targets prove to be significant, then it would be important to investigate their functional significance in breast cell line models. Recently, Sun and Chao (2005) have shown that over expression of DDB2 has anti-apoptotic effects. It would be interesting to study DDB2 and RARRES3 expression in breast cancer cell lines, and use gene transfection and siRNA technologies to investigate the functional significance of these two genes in these cells on measurable hallmarks of cancer such as proliferation, apoptosis and invasion. It will also be important to investigate the underlying mechanism leading to changes in the level of gene expression. If shown to be related to gene copy number, then FISH and array CGH methods should reveal these changes. These novel gene targets might have future potential as biomarkers for diagnosis and or prognosis.

In conclusion, this study has extended previous findings by the research group and has generated novel data to support the hypothesis that sporadic breast cancers in young women arise because of specific genetic alterations. The most significant findings of this thesis have resulted from the successful completion of pilot Affymetrix cDNA microarray studies using freshly collected tumour tissue from two sporadic breast cancers in young women. A number of novel gene targets have been identified including DDB2 and RARAES3 which are worthy of further investigation.

References

Aatsinki, J.T. *et al.* (1994) *BioTechniques* **16**, 282.

Abravaya K, Huff J, Marshall R, Merchant B, Mullen C, Schneider G, and Robinson J (2003) Molecular beacons as diagnostic tools: technology and applications. *Clin Chem Lab Med* **41**, 468-474

Aebi, S., Gelber, S., Castiglione-Gertsch, M., Gelber, R. D., Collins, J., Thurlimann, B., Rudenstam, C. M., Lindtner, J., Crivellari, D., Cortes-Funes, H., Simoncini, E., Werner, I. D., Coates, A. S. & Goldhirsch, A. (2000) Is chemotherapy alone adequate for young women with oestrogen-receptor-positive breast cancer? *Lancet,* **355,** 1869-74.

Ahr, A., Holtrich, U., Solbach, C., Scharl, A., Strebhardt, K., Karn, T. & Kaufmann, M. (2001) Molecular classification of breast cancer patients by gene expression profiling. *J Pathol,* **195,** 312-20.

Albain, K. S., Allred, D. C. & Clark, G. M. (1994) Breast cancer outcome and predictors of outcome: are there age differentials? *J Natl Cancer Inst Monogr,* 35-42.

Alberts B, Bray D, Lewis L, Raff M, Roberts K, Watson J. D. (1994) *Molecular Biology of the Cell, 3rd Edition.* Garland Publishing, London.

Alvarez, S., Diaz-Uriarte, R., Osorio, A., Barroso, A., Melchor, L., Paz, M. F., Honrado, E., Rodriguez, R., Urioste, M., Valle, L., Diez, O., Cigudosa, J. C., Dopazo, J., Esteller, M. & Benitez, J. (2005) A predictor based on the somatic genomic changes of the BRCA1/BRCA2 breast cancer tumors identifies the non-BRCA1/BRCA2 tumors with BRCA1 promoter hypermethylation. *Clin Cancer Res,* **11,** 1146-53.

Amatschek, S., Koenig, U., Auer, H., Steinlein, P., Pacher, M., Gruenfelder, A., Dekan, G., Vogl, S., Kubista, E., Heider, K. H., Stratowa, C., Schreiber, M. & Sommergruber, W. (2004) Tissue-wide expression profiling using cDNA subtraction and microarrays to identify tumor-specific genes. *Cancer Res,* **64,** 844-56.

Ameyaw M. M, Tayeb M, Thornton N, Folayan G, Tariq M, Mobarek A, Evans D. A, Ofori-Adjei D, McLead H. L. (2002) Ethnic variation in the HER-2 codon 655 genetic polymorphism previously associated with breast cancer. *J HUm Genet* **47**(4), 172-5

Anderson, P. a. (1987) Diagnostic Histopathology of the Breast. Edinburgh, Churchill Livingstone.

Anderson, S. F., Schlegel, B. P., Nakajima, T., Wolpin, E. S. & Parvin, J. D. (1998) BRCA1 protein is linked to the RNA polymerase II holoenzyme complex via RNA helicase A. *Nat Genet,* **19,** 254-6.

Antony T and Subramaniam V (2001) Molecular beacons: nucleic acid hybridization and emerging applications. *J Biomol Struct Dyn* **19,** 497-504

Avrameas S, Uriel J. (1966) Method of antigen and antibody labelling with enzymes and its immunodiffusion application *C R Acad Sci Hebd Seances Acad Sci D.* **262**(24):2543-5

Baldino, F. *et al.* (1989) *Meth. Enzymol.* **168**, 761

Barnes, D. M., Dublin, E. A., Fisher, C. J., Levison, D. A. & Millis, R. R. (1993) Immunohistochemical detection of p53 protein in mammary carcinoma: an important new independent indicator of prognosis? *Hum Pathol,* **24,** 469-76.

Barrett, K. L., Demiranda, D. & Katula, K. S. (2002) Cyclin b1 promoter activity and functional cdk1 complex formation in G1 phase of human breast cancer cells. *Cell Biol Int,* **26,** 19-28.

Bergman-Jungestrom, M., Gentile, M., Lundin, A. C. & Wingren, S. (1999) Association between CYP17 gene polymorphism and risk of breast cancer in young women. *Int J Cancer,* **84,** 350-3.

Bergstrom, A., Pisani, P., Tenet, V., Wolk, A. & Adami, H. O. (2001) Overweight as an avoidable cause of cancer in Europe. *Int J Cancer,* **91,** 421-30.

Berns EM, Klijn JG, Smid M, van Staveren IL, Look MP, van Putten WL, Foekens JA. (1996) TP53 and MYC gene alterations independently predict poor prognosis in breast cancer patients. *Genes Chromosomes Cancer* **16**(3), 170-9

Bernard-Gallon, D., Peffault De Latour, M., Rio, P., Favy, D., Hizel, C., Vissac, C. & Bignon, Y. J. (2000) BRCA2 protein expression in sporadic breast carcinoma with or without allelic loss of BRCA2. *Int J Cancer,* **86,** 453-6.

Bertwistle, D. & Ashworth, A. (2000) BRCA1 and BRCA2. *Curr Biol,* **10,** R582.

Bhandari, V.; Palfree, R. G. E.; Bateman, A (1992) Isolation and sequence of the granulin precursor cDNA from human bone marrow reveals tandem cysteine-rich granulin domains *Proc. Nat. Acad. Sci U S A* **89**, 1715-1719

Bieche, I., Nogues, C. & Lidereau, R. (1999) Overexpression of BRCA2 gene in sporadic breast tumours. *Oncogene,* **18,** 5232-8.

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

Boenisch M. S. (2001) Handbook: Immunochemical Straining Methods 3rd Edition. (www.DAKO.com)

Brien T. P, Odze R. D, Sheehan C. E, McKenna B. J, Ross J. S. (2000) HER-2/neu gene amplification by FISH predicts poor survival in Barrett's esophagus-associated adenocarcinoma. *Hum Pathol* **31**(1), 35-9

Brown C. J. (1999) Skewed X-Chromosome Inactivation: Cause or Consequence? *J Nat Cancer Inst.* **91,** 301-303

Buckhaults P, Rago C, St Croix B, Romans KE, Saha S, Zhang L, Vogelstein B, Kinzler KW (2001) Secreted and cell surface genes expressed in benign and malignant colorectal tumors *Cancer Res* **61**(19):6996-7001

Buller R.E, Sood A.K, Lallas T, Buekers T, Skilling J.S. (1999) Association between nonrandom X-chromosome inactivation and BRCA1 mutation in germline DNA of patients with ovarian cancer. *J Natl Cancer Inst* **91**(4), 339-46

Busque L, Mio R, Mattioli J, Brais E, Blais N, Lalonde Y, Maragh M, Gilliland D. G (1996) Nonrandom X-inactivation patterns in normal females: lyonization ratios vary with age. *Blood* **88**(1), 59-65

Byrne, C., Connolly, J. L., Colditz, G. A. & Schnitt, S. J. (2000) Biopsy confirmed benign breast disease, postmenopausal use of exogenous female hormones, and breast carcinoma risk. *Cancer,* **89,** 2046-52.

Byrne, C., Schairer, C., Brinton, L. A., Wolfe, J., Parekh, N., Salane, M., Carter, C. & Hoover, R. (2001) Effects of mammographic density and benign breast disease on breast cancer risk (United States). *Cancer Causes Control,* **12,** 103-10.

Callagy, G., Pharoah, P., Chin, S. F., Sangan, T., Daigo, Y., Jackson, L. & Caldas, C. (2005) Identification and validation of prognostic markers in breast cancer with the complementary use of array-CGH and tissue microarrays. *J Pathol,* **205,** 388-96.

Cancer, C. G. o. H. F. f. B. (2001) Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. *Lancet,* **358,** 1389-99.

Chai, Y. L., Cui, J., Shao, N., Shyam, E., Reddy, P. & Rao, V. N. (1999) The second BRCT domain of BRCA1 proteins interacts with p53 and stimulates transcription from the p21WAF1/CIP1 promoter. *Oncogene,* **18,** 263-8.

Chan, K. Y., Ozcelik, H., Cheung, A. N., Ngan, H. Y. & Khoo, U. S. (2002) Epigenetic factors controlling the BRCA1 and BRCA2 genes in sporadic ovarian cancer. *Cancer Res,* **62,** 4151-6.

Chen, J. J., Silver, D., Cantor, S., Livingston, D. M. & Scully, R. (1999) BRCA1, BRCA2, and Rad51 operate in a common DNA damage response pathway. *Cancer Res,* **59,** 1752s-1756s.

Chen, Y., Chen, C. F., Riley, D. J., Allred, D. C., Chen, P. L., Von Hoff, D., Osborne, C. K. & Lee, W. H. (1995) Aberrant subcellular localization of BRCA1 in breast cancer. *Science,* **270,** 789-91.

Chen, Y., Farmer, A. A., Chen, C. F., Jones, D. C., Chen, P. L. & Lee, W. H. (1996) BRCA1 is a 220-kDa nuclear phosphoprotein that is expressed and phosphorylated in a cell cycle-dependent manner. *Cancer Res,* **56,** 3168-72.

Chen Y, Lee W. H, Chew H. K. (1999) Emerging roles of BRCA1 in transcriptional regulation and DNA repair. *J Cell Physol* **181**(3):385-92

Chrenek MA, Wong P, Weaver VM. (2001) Tumour-stromal interactions. Integrins and cell adhesions as modulators of mammary cell survival and transformation *Breast Cancer Res* **3**(4):224-9

Cicatiello, L., Scafoglio, C., Altucci, L., Cancemi, M., Natoli, G., Facchiano, A., Iazzetti, G., Calogero, R., Biglia, N., De Bortoli, M., Sfiligoi, C., Sismondi, P., Bresciani, F. & Weisz, A. (2004) A genomic view of estrogen actions in human breast cancer cells by expression profiling of the hormone-responsive transcriptome. *J Mol Endocrinol,* **32,** 719-75.

Chung I.M., Schwartz S.M., Murry C.E. (1998) Clonal architecture of normal and atherosclerotic aorta: implications for atherogenesis and vascular development. *Am J Pathol* **152**(4), 913-23

Clamp, A., Danson, S. & Clemons, M. (2002) Hormonal risk factors for breast cancer: identification, chemoprevention, and other intervention strategies. *Lancet Oncol,* **3,** 611-9.

Clark, J., Edwards, S., John, M., Flohr, P., Gordon, T., Maillard, K., Giddings, I., Brown, C., Bagherzadeh, A., Campbell, C., Shipley, J., Wooster, R. & Cooper, C. S. (2002) Identification of amplified and expressed genes in breast cancer by comparative hybridization onto microarrays of randomly selected cDNA clones. *Genes Chromosomes Cancer,* **34,** 104-14.

Clavel-Chapelon, F. (2002) Differential effects of reproductive factors on the risk of pre- and postmenopausal breast cancer. Results from a large cohort of French women. *Br J Cancer,* **86,** 723-7.

Coons A. H., Kaplan M. H., (1950) Localization of antigen in tissue cells; improvements in a method for the detection of antigen by means of fluorescent antibody. *J Exp Med* **91**(1):1-13

Corvi R, Savelyeva L, Schwab M (1995) Duplication of N-MYC at its resident site 2p24 may be a mechanism of activation alternative to amplification in human neuroblastoma cells. **55**(16), 3471-4

Coussens, L., Yang-Feng, T. L., Liao, Y. C., Chen, E., Gray, A., McGrath, J., Seeburg, P. H., Libermann, T. A., Schlessinger, J., Francke, U. & et al. (1985) Tyrosine kinase receptor

with extensive homology to EGF receptor shares chromosomal location with neu oncogene. *Science,* **230,** 1132-9.

Coux, O., Tanaka, K. & Goldberg, A. L. (1996) Structure and functions of the 20S and 26S proteasomes. *Annu Rev Biochem,* **65,** 801-47.

Cui, J. Q., Shao, N., Chai, Y., Wang, H., Reddy, E. S. & Rao, V. N. (1998) BRCA1 splice variants BRCA1a and BRCA1b associate with CBP co-activator. *Oncol Rep,* **5,** 591-5.

de Jong, M. M., Nolte, I. M., te Meerman, G. J., van der Graaf, W. T., Oosterwijk, J. C., Kleibeuker, J. H., Schaapveld, M. & de Vries, E. G. (2002) Genes other than BRCA1 and BRCA2 involved in breast cancer susceptibility. *J Med Genet,* **39,** 225-42.

Dean, F. *et al,* (2001) *Genome Research* **11**, 1095-1099

DiNardo, D. N., Butcher, D. T., Robinson, D. P., Archer, T. K. & Rodenhiser, D. I. (2001) Functional analysis of CpG methylation in the BRCA1 promoter region. *Oncogene,* **20,** 5331-40.

Dong, L. Q., Du, H., Porter, S. G., Kolakowski, L. F., Jr., Lee, A. V., Mandarino, L. J., Fan, J., Yee, D. & Liu, F. (1997) Cloning, chromosome localization, expression, and characterization of an Src homology 2 and pleckstrin homology domain-containing insulin receptor binding protein hGrb10gamma. *J Biol Chem,* **272,** 29104-12.

Dubeau L, Chandler L. A, Gralow J. R, Nichols P. W, Jones P. A. (1986) Southern blot analysis of DNA extracted from formalin-fixed pathology specimens. *Cancer Res* **46**(6):2964-9.

Dupont, W. D. & Page, D. L. (1985) Risk factors for breast cancer in women with proliferative breast disease. *N Engl J Med,* **312,** 146-51.

Eckert, K.A. and Kunkel, T.A. (1990) *Nucl. Acids Res.* **18**, 3739.

Ellsworth, D.L., Rittenhouse, K.D. and Honeycutt, R.L. (1993) *BioTechniques* **14**, 214

Elston, C. W. & Ellis, I. O. (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,** 403-10.

Estaban, J. A., *et al* (1993) *J Biol Chem* **11**, 2719-2726

Farabegoli F, Ceccarelli C, Santini D, Baldini N, Taffurelli M, Marrano D, Trere D, Derenzini M (1999) c-erbB-2 over-expression in amplified and non-amplified breast carcinoma samples *Int J Cancer* **84**(3), 273-7

Faulk W. P, Taylor G. M. (1971) An immunocolloid method for the electron microscope. *Immunochemistry.* **8**(11):1081-3

Feldman, A. L. & Welch, J. P. (1998) Long-term outcome in women less than 30 years of age with breast cancer. *J Surg Oncol,* **68,** 193-8.

Ferlay, J., Bray, F., Pisani, P. & al, e. (2001) Globocan 2000: Cancer Incidance, Mortality and Prevalence Worldwide, version 1.0: IARC CancerBase no.5 Lyon, IARC Press.

Ferlay, J., Bray, F., Sankila, R. & al, e. (2003) EUCAN: Cancer Incidance, Mortality and Prevalance in the European Union 1998, version 5.0: IARCPress.

Foekens, J. A., Portengen, H., van Putten, W. L., Peters, H. A., Krijnen, H. L., Alexieva-Figusch, J. & Klijn, J. G. (1989) Prognostic value of estrogen and progesterone receptors measured by enzyme immunoassays in human breast tumor cytosols. *Cancer Res,* **49,** 5823-8.

Ford, D., Easton, D. F., Stratton, M., Narod, S., Goldgar, D., Devilee, P., Bishop, D. T., Weber, B., Lenoir, G., Chang-Claude, J., Sobol, H., Teare, M. D., Struewing, J., Arason, A., Scherneck, S., Peto, J., Rebbeck, T. R., Tonin, P., Neuhausen, S., Barkardottir, R., Eyfjord, J., Lynch, H., Ponder, B. A., Gayther, S. A., Zelada-Hedman, M. & et al. (1998) Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. *Am J Hum Genet,* **62,** 676-89.

Frade R, Balbo M, Barel M (2000) RB18A, whose gene is localized on chromosome 17q12-q21.1, regulates in vivo p53 transactivating activity *Cancer Res* **60**(23), 6585-9

Frohling S, Schlenk RF, Stolze I, Bihlmayr J, Benner A, Kreitmeier S, Tobis K, Dohner H, Dohner K (2004) CEBPA mutations in younger adults with acute myeloid leukemia and normal cytogenetics: prognostic relevance and analysis of cooperating mutations *J Clin Oncol* **22**(4), 624-33

Fukuoka J, Fujii T, Shih JH, Dracheva T, Meerzaman D, Player A, Hong K, Settnek S, Gupta A, Buetow K, Hewitt S, Travis WD, Jen J. (2004) Chromatin remodeling factors and BRM/BRG1 expression as prognostic indicators in non-small cell lung cancer *Clin Cancer Res* **10**(13), 4314-24

Fuks, F., Milner, J. & Kouzarides, T. (1998) BRCA2 associates with acetyltransferase activity when bound to P/CAF. *Oncogene,* **17,** 2531-4.

Gartler S.M, Riggs A.D. (1983) Mammalian X-chromosome inactivation. *Annu Rev Genet.* **17,** 155-90.

Gentile, M., Olsen, K., Dufmats, M. & Wingren, S. (1999) Frequent allelic losses at 11q24.1-q25 in young women with breast cancer: association with poor survival. *Br J Cancer,* **80,** 843-9.

Ghellal, A., Li, C., Hayes, M., Byrne, G., Bundred, N. & Kumar, S. (2000) Prognostic significance of TGF beta 1 and TGF beta 3 in human breast carcinoma. *Anticancer Res,* **20,** 4413-8.

Glanzer JG, Eberwine JH (2004) Expression profiling of small cellular samples in cancer: less is more. *Br J Cancer* **90**(6), 1111-4

Gordon LA, Mulligan KT, Maxwell-Jones H, Adams M, Walker RA, Jones JL (2003) Breast cell invasive potential relates to the myoepithelial phenotype *Int J Cancer* **106**(1), 8-16

Gowen, L. C., Avrutskaya, A. V., Latour, A. M., Koller, B. H. & Leadon, S. A. (1998) BRCA1 required for transcription-coupled repair of oxidative DNA damage. *Science,* **281,** 1009-12.

Grant, M., Zuccotti, M. & Monk, M. (1992) Methylation of CpG sites of two X-linked genes coincides with X-inactivation in the female mouse embryo but not in the germ line. *Nat Genet,* **2,** 161-6.

Guan X. Y., Meltzer P. S., Dalton W. S., Trent J. M. (1994) Identification of cryptic sites of DNA sequence amplification in human breast cancer by chromosome microdissection. *Nat Genet* **8**(2), 155-61

Gudjonsson, T., Ronnov-Jessen, L., Villadsen, R., Rank, F., Bissell, M. J. & Petersen, O. W. (2002) Normal and tumor-derived myoepithelial cells differ in their ability to interact with luminal breast epithelial cells for polarity and basement membrane deposition. *J Cell Sci,* **115,** 39-50.

Gullick, W. J., Love, S. B., Wright, C., Barnes, D. M., Gusterson, B., Harris, A. L. & Altman, D. G. (1991) c-erbB-2 protein overexpression in breast cancer is a risk factor in patients with involved and uninvolved lymph nodes. *Br J Cancer,* **63,** 434-8.

Gusterson BA, Machin LG, Gullick WJ, Gibbs NM, Powles TJ, Elliott C, Ashley S, Monaghan P, Harrison S (1988) c-erbB-2 expression in benign and malignant breast disease. *Br J Cancer.* **58**(4), 453-7

Han DC, Guan JL. (1999) Association of focal adhesion kinase with Grb7 and its role in cell migration *J Biol Chem* **274**(34), 24425-30

Han DC, Shen TL, Guan JL. (2001) The Grb7 family proteins: structure, interactions with other signaling molecules and potential cellular functions. *Oncogene* **20**(44), 6315-21

Harkin, D. P. (2000) Uncovering functionally relevant signaling pathways using microarray-based expression profiling. *Oncologist* **5,** 501-7.

Hartman AR, Ford JM (2002) BRCA1 induces DNA damage recognition factors and enhances nucleotide excision repair *Nat Genet* **32**(1), 180-4

Hatakeyama, C., Anderson, C. L., Beever, C. L., Penaherrera, M. S., Brown, C. J. & Robinson, W. P. (2004) The dynamics of X-inactivation skewing as women age. *Clin Genet,* **66,** 327-32.

Hedau, S., Jain, N., Husain, S. A., Mandal, A. K., Ray, G., Shahid, M., Kant, R., Gupta, V., Shukla, N. K., Deo, S. S. & Das, B. C. (2004) Novel germline mutations in breast cancer susceptibility genes BRCA1, BRCA2 and p53 gene in breast cancer patients from India. *Breast Cancer Res Treat,* **88,** 177-86.

Hedenfalk, I., Duggan, D., Chen, Y., Radmacher, M., Bittner, M., Simon, R., Meltzer, P., Gusterson, B., Esteller, M., Kallioniemi, O. P., Wilfond, B., Borg, A. & Trent, J. (2001) Gene-expression profiles in hereditary breast cancer. *N Engl J Med,* **344,** 539-48.

Helbling D, Mueller BU, Timchenko NA, Hagemeijer A, Jotterand M, Meyer-Monard S, Lister A, Rowley JD, Huegli B, Fey MF, Pabst T (2004) The leukemic fusion gene AML1-MDS1-EVI1 suppresses CEBPA in acute myeloid leukemia by activation of Calreticulin *Proc Natl Acad Sci U S A* **101**(36):13312-7

Herman, J. G., Latif, F., Weng, Y., Lerman, M. I., Zbar, B., Liu, S., Samid, D., Duan, D. S., Gnarra, J. R., Linehan, W. M. & et al. (1994) Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma. *Proc Natl Acad Sci U S A,* **91,** 9700-4.

Herrera-Gayol, A. & Jothy, S. (1999) Adhesion proteins in the biology of breast cancer: contribution of CD44. *Exp Mol Pathol,* **66,** 149-56.

Hilton, J. L., Geisler, J. P., Rathe, J. A., Hattermann-Zogg, M. A., DeYoung, B. & Buller, R. E. (2002) Inactivation of BRCA1 and BRCA2 in ovarian cancer. *J Natl Cancer Inst,* **94,** 1396-406.

Houslay MD, Adams DR (2003) PDE4 cAMP phosphodiesterases: modular enzymes that orchestrate signalling cross-talk, desensitization and compartmentalization *Biochem J* **15,** 370(Pt 1):1-18

Huang, L. J.; Durick, K.; Weiner, J. A.; Chun, J.; Taylor, S. S (1997) Identification of a novel protein kinase A anchoring protein that binds both type I and type II regulatory subunits *J Biol Chem* **272,** 8057-8064

Hwang, B. J.; Toering, S.; Francke, U.; Chu, G (1999) p48 activates a UV-damaged-DNA binding factor and is defective in xeroderma pigmentosum group E cells that lack binding activity *Molec Cell Biol* **18**, 4391-4399

Indsto, J. O., Cachia, A. R., Kefford, R. F. & Mann, G. J. (2001) X inactivation, DNA deletion, and microsatellite instability in common acquired melanocytic nevi. *Clin Cancer Res,* **7,** 4054-9.

Ishida, A., Asano, H., Hasegawa, M., Koseki, H., Ono, T., Yoshida, M. C., Taniguchi, M. & Kanno, M. (1993) Cloning and chromosome mapping of the human Mel-18 gene which encodes a DNA-binding protein with a new 'RING-finger' motif. *Gene,* **129,** 249-55.

Ishiwatari, H., Nakanishi, K., Kondoh, G., Hayasaka, N., Li, Q., Yamashita, A., Inoue, H. & Hakura, A. (1997) Suppression of tumor growth by the 3' untranslated region of mel-18 in 3Y1 cells transformed by the E6 and E7 genes of human papillomavirus type 18. *Cancer Lett,* **117,** 57-65.

Issemann, I. & Green, S. (1990) Activation of a member of the steroid hormone receptor superfamily by peroxisome proliferators. *Nature,* **347,** 645-50.

Jasienska, G. & Thune, I. (2001) Lifestyle, hormones, and risk of breast cancer. *Bmj,* **322,** 586-7.

Jensen, D. E., Proctor, M., Marquis, S. T., Gardner, H. P., Ha, S. I., Chodosh, L. A., Ishov, A. M., Tommerup, N., Vissing, H., Sekido, Y., Minna, J., Borodovsky, A., Schultz, D. C., Wilkinson, K. D., Maul, G. G., Barlev, N., Berger, S. L., Prendergast, G. C. & Rauscher, F. J., 3rd (1998) BAP1: a novel ubiquitin hydrolase which binds to the BRCA1 RING finger and enhances BRCA1-mediated cell growth suppression. *Oncogene,* **16,** 1097-112.

Jiang S. Y, Chou J. M, Leu F. J, Hsu Y. Y, Shih Y. L, Yu J. C, Lee M. S, Shyu R. Y (2005) Decreased expression of type II tumor suppressor gene RARRES3 in tissues of hepatocellular carcinoma and cholangiocarcinoma. *World J Gastroenterol* **11**(7), 948-53

Joazeiro, C. A., Wing, S. S., Huang, H., Leverson, J. D., Hunter, T. & Liu, Y. C. (1999) The tyrosine kinase negative regulator c-Cbl as a RING-type, E2-dependent ubiquitin-protein ligase. *Science,* **286,** 309-12.

Joensuu, H. & Toikkanen, S. (1995) Cured of breast cancer? *J Clin Oncol,* **13,** 62-9.

Johnson, S. M., Shaw, J. A. & Walker, R. A. (2002) Sporadic breast cancer in young women: prevalence of loss of heterozygosity at p53, BRCA1 and BRCA2. *Int J Cancer,* **98,** 205-9.

Kallioniemi, A., Kallioniemi, O. P., Sudar, D., Rutovitz, D., Gray, J. W., Waldman, F. & Pinkel, D. (1992) Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science,* **258,** 818-21.

Kallioniemi, O. P., Kallioniemi, A., Mascio, L., Sudar, D., Pinkel, D., Deaven, L. & Gray, J. (1994) Physical mapping of chromosome 17 cosmids by fluorescence in situ hybridization and digital image analysis. *Genomics,* **20,** 125-8.

Kanaar, R., Hoeijmakers, J. H. & van Gent, D. C. (1998) Molecular mechanisms of DNA double strand break repair. *Trends Cell Biol,* **8,** 483-9.

Kanno, M., Hasegawa, M., Ishida, A., Isono, K. & Taniguchi, M. (1995) mel-18, a Polycomb group-related mammalian gene, encodes a transcriptional negative regulator with tumor suppressive activity. *Embo J,* **14,** 5672-8.

Kauraniemi, P., Barlund, M., Monni, O. & Kallioniemi, A. (2001) New amplified and highly expressed genes discovered in the ERBB2 amplicon in breast cancer by cDNA microarrays. *Cancer Res,* **61,** 8235-40.

Kelsey, J. L., Gammon, M. D. & John, E. M. (1993) Reproductive factors and breast cancer. *Epidemiol Rev,* **15,** 36-47.

Kerr, P. & Ashworth, A. (2001) New complexities for BRCA1 and BRCA2. *Curr Biol,* **11,** R668-76.

Ko, T. K., Kelly, E. & Pines, J. (2001) CrkRS: a novel conserved Cdc2-related protein kinase that colocalises with SC35 speckles. *J Cell Sci,* **114,** 2591-603.

Korkolis D.P, Tsoli E, Fouskakis D, Yiotis J, Koullias G.J, Giannopoulos D, Papalambros E, Nikiteas N.I, Spiliopoulou C.A, Patsouris E, Asimacopoulos P, Gorgoulis V.G (2004) Tumor histology and stage but not p53, Her2-neu or cathepsin-D expression are independent prognostic factors in breast cancer patients. *Anticancer Res* **24**(3b):2061-8

Kristiansen, M., Langerod, A., Knudsen, G. P., Weber, B. L., Borresen-Dale, A. L. & Orstavik, K. H. (2002) High frequency of skewed X inactivation in young breast cancer patients. *J Med Genet,* **39,** 30-3.

Lakhani, S. R., Chaggar, R., Davies, S., Jones, C., Collins, N., Odel, C., Stratton, M. R. & O'Hare, M. J. (1999) Genetic alterations in 'normal' luminal and myoepithelial cells of the breast. *J Pathol,* **189,** 496-503.

Leivo I, Jee KJ, Heikinheimo K, Laine M, Ollila J, Nagy B, Knuutila S (2005) Characterization of gene expression in major types of salivary gland carcinomas with epithelial differentiation *Cancer Genet Cytogenet* **156**(2):104-13

Listgarten, J., Damaraju, S., Poulin, B., Cook, L., Dufour, J., Driga, A., Mackey, J., Wishart, D., Greiner, R. & Zanke, B. (2004) Predictive models for breast cancer susceptibility from multiple single nucleotide polymorphisms. *Clin Cancer Res,* **10,** 2725-37.

Lizardi, P. *et al,* (1998) *Nat Genet* **19,** 225-323

Lu R, Serrero G (1999) Stimulation of PC cell-derived growth factor (epithelin/granulin precursor) expression by estradiol in human breast cancer cells. *Biochem Biophys Res Commun* **256**(1):204-7

Lu R, Serrero G (2000) Inhibition of PC cell-derived growth factor (PCDGF, epithelin/granulin precursor) expression by antisense PCDGF cDNA transfection inhibits tumorigenicity of the human breast carcinoma cell line MDA-MB-468 *Proc Natl Acad Sci USA* **97**(8), 3993-8

Lyon, M. F. (1961) Gene action in the X-chromosome of the mouse (Mus musculus L.). *Nature,* **190,** 372-3.

Lyon, M. F. (1972) X-chromosome inactivation and developmental patterns in mammals. *Biol Rev Camb Philos Soc,* **47,** 1-35.

Magdinier, F., Ribieras, S., Lenoir, G. M., Frappart, L. & Dante, R. (1998) Down-regulation of BRCA1 in human sporadic breast cancer; analysis of DNA methylation patterns of the putative promoter region. *Oncogene,* **17,** 3169-76.

Man, Y. G., Zhang, Y., Shen, T., Zeng, X., Tauler, J., Mulshine, J. L. & Strauss, B. L. (2005) cDNA expression profiling reveals elevated gene expression in cell clusters overlying focally disrupted myoepithelial cell layers: implications for breast tumor invasion. *Breast Cancer Res Treat,* **89,** 199-208.

Mancini, D. N., Rodenhiser, D. I., Ainsworth, P. J., O'Malley, F. P., Singh, S. M., Xing, W. & Archer, T. K. (1998) CpG methylation within the 5' regulatory region of the BRCA1 gene is tumor specific and includes a putative CREB binding site. *Oncogene,* **16,** 1161-9.

Mason D. Y, Sammons R. (1978) Alkaline phosphatase and peroxidase for double immunoenzymatic labelling of cellular constituents. *J Clin Pathol.* **31**(5):454-60

Mathews C. K. and van Holde K. E. (1996) *Biochemistry 2nd Edition.* The Benjamin/Cummings Publishing Company, Wokingham, UK

Megha, T., Lazzi, S., Ferrari, F., Vatti, R., Howard, C. M., Cevenini, G., Leoncini, L., Luzi, P., Giordano, A. & Tosi, P. (1999) Expression of the G2-M checkpoint regulators cyclin B1 and P34CDC2 in breast cancer: a correlation with cellular kinetics. *Anticancer Res,* **19,** 163-9.

Menard S, Fortis S, Castiglioni F, Agresti R, Balsari A (2001) HER2 as a prognostic factor in breast cancer *Oncology* **61** Suppl 2:67-72

Meyerson, M., Enders, G. H., Wu, C. L., Su, L. K., Gorka, C., Nelson, C., Harlow, E. & Tsai, L. H. (1992) A family of human cdc2-related protein kinases. *Embo J,* **11,** 2909-17.

Miki, Y., Swensen, J., Shattuck-Eidens, D., Futreal, P. A., Harshman, K., Tavtigian, S., Liu, Q., Cochran, C., Bennett, L. M., Ding, W. & et al. (1994) A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science,* **266,** 66-71.

Milner, J., Ponder, B., Hughes-Davies, L., Seltmann, M. & Kouzarides, T. (1997) Transcriptional activation functions in BRCA2. *Nature,* **386,** 772-3.

Mizukami, Y., Nonomura, A., Takizawa, T., Noguchi, M., Michigishi, T., Nakamura, S. & Ishizaki, T. (1995) N-myc protein expression in human breast carcinoma: prognostic implications. *Anticancer Res,* **15,** 2899-905.

Mohsin SK, O'Connell P, Allred DC, Libby AL. (2005) Biomarker profile and genetic abnormalities in lobular carcinoma in situ. *Breast Cancer Res Treat* **90**(3):249-56

Moll U, Lau R, Sypes M. A, Gupta M. M, Anderson C. W (1999) DNA-PK, the DNA-activated protein kinase, is differentially expressed in normal and malignant human tissues. *Oncogene* **18**(20), 3114-26

Monni, O., Barlund, M., Mousses, S., Kononen, J., Sauter, G., Heiskanen, M., Paavola, P., Avela, K., Chen, Y., Bittner, M. L. & Kallioniemi, A. (2001) Comprehensive copy number and gene expression profiling of the 17q23 amplicon in human breast cancer. *Proc Natl Acad Sci U S A,* **98,** 5711-6.

Montminy M (1997) Transcriptional regulation by cyclic AMP *Annu Rev Biochem* **66,** 807-22

Moynahan, M. E., Chiu, J. W., Koller, B. H. & Jasin, M. (1999) Brca1 controls homology-directed DNA repair. *Mol Cell,* **4,** 511-8.

Muleris, M., Almeida, A., Malfoy, B. & Dutrillaux, B. (1997) Assignment of v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 2 (ERBB2) to human chromosome band 17q21.1 by in situ hybridization. *Cytogenet Cell Genet,* **76,** 34-5.

Murakawa, G.J. *et al.* (1988) *DNA* 7, 287.

Nabholtz JM, Reese DM, Lindsay MA, Riva A. (2002) HER2-positive breast cancer: update on Breast Cancer International Research Group trials. *Clin Breast Cancer* **3** Suppl 2:S75-9

Nagai, M. A., Da Ros, N., Neto, M. M., de Faria Junior, S. R., Brentani, M. M., Hirata, R., Jr. & Neves, E. J. (2004) Gene expression profiles in breast tumors regarding the presence or absence of estrogen and progesterone receptors. *Int J Cancer,* **111,** 892-9.

Nandi D, Woodward E, Ginsburg D. B, Monaco J. J (1997) Intermediates in the formation of mouse 20S proteasomes: implications for the assembly of precursor beta subunits. *EMBO J* **16**(17), 5363-75

Nam JO, Kim JE, Jeong HW, Lee SJ, Lee BH, Choi JY, Park RW, Park JY, Kim IS (2003) Identification of the αvβ3 integrin-interacting motif of betaig-h3 and its anti-angiogenic effect *J Biol Chem* **278**(28), 25902-9

Nathanson, K. L., Wooster, R. & Weber, B. L. (2001) Breast cancer genetics: what we know and what we need. *Nat Med,* **7,** 552-6.

Nattkemper, T. W., Arnrich, B., Lichte, O., Timm, W., Degenhard, A., Pointon, L., Hayes, C. & Leach, M. O. (2005) Evaluation of radiological features for breast tumour

classification in clinical screening with machine learning methods. *Artif Intell Med,* **34,** 129-39.

Neish, A. S., Anderson, S. F., Schlegel, B. P., Wei, W. & Parvin, J. D. (1998) Factors associated with the mammalian RNA polymerase II holoenzyme. *Nucleic Acids Res,* **26,** 847-53.

Nelson, J. R., *et al* (2002) *Biotechniques* **32,** S44-S47

Nessling M, Richter K, Schwaenen C, Roerig P, Wrobel G, Wessendorf S, Fritz B, Bentz M, Sinn HP, Radlwimmer B, Lichter P. (2005) Candidate genes in breast cancer revealed by microarray-based comparative genomic hybridization of archived tissue *Cancer Res* **65**(2), 439-47

Newman, B., Austin, M. A., Lee, M. & King, M. C. (1988) Inheritance of human breast cancer: evidence for autosomal dominant transmission in high-risk families. *Proc Natl Acad Sci U S A,* **85,** 3044-8.

Noguchi S, Motomura K, Inaji H, Imaoka S, Koyama H (1992) Clonal analysis of human breast cancer by means of the polymerase chain reaction *Cancer Res* **52**(23), 6594-7

Novelli, M., Cossu, A., Oukrif, D., Quaglia, A., Lakhani, S., Poulsom, R., Sasieni, P., Carta, P., Contini, M., Pasca, A., Palmieri, G., Bodmer, W., Tanda, F. & Wright, N. (2003) X-inactivation patch size in human female tissue confounds the assessment of tumor clonality. *Proc Natl Acad Sci U S A,* **100,** 3311-4.

Nyhan, W. L., Bakay, B., Connor, J. D., Marks, J. F. & Keele, D. K. (1970) Hemizygous expression of glucose-6-phosphate dehydrogenase in erythrocytes of heterozygotes for the Lesch-Nyhan syndrome. *Proc Natl Acad Sci U S A,* **65,** 214-8.

Ohtani-Fujita, N., Fujita, T., Aoike, A., Osifchin, N. E., Robbins, P. D. & Sakai, T. (1993) CpG methylation inactivates the promoter activity of the human retinoblastoma tumor-suppressor gene. *Oncogene,* **8,** 1063-7.

Orstavik, K. H., Orstavik, R. E., Naumova, A. K., D'Adamo, P., Gedeon, A., Bolhuis, P. A., Barth, P. G. & Toniolo, D. (1998) X chromosome inactivation in carriers of Barth syndrome. *Am J Hum Genet,* **63,** 1457-63.

Pabst T, Mueller BU, Harakawa N, Schoch C, Haferlach T, Behre G, Hiddemann W, Zhang DE, Tenen DG (2001) AML1-ETO downregulates the granulocytic differentiation factor C/EBPalpha in t(8;21) myeloid leukemia *Nat Med* 7(4), 444-51

Pantazi, S., Kagolovsky, Y. & Moehr, J. R. (2002) Cluster analysis of Wisconsin Breast Cancer dataset using self-organizing maps. *Stud Health Technol Inform,* **90,** 431-6.

Pillers, E. M. (1992) Histological grade of breast cancer in younger women. *Lancet,* **339,** 1483.

Pinkel, D., Segraves, R., Sudar, D., Clark, S., Poole, I., Kowbel, D., Collins, C., Kuo, W. L., Chen, C., Zhai, Y., Dairkee, S. H., Ljung, B. M., Gray, J. W. & Albertson, D. G. (1998) High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet,* **20,** 207-11.

Pollack, J. R., Perou, C. M., Alizadeh, A. A., Eisen, M. B., Pergamenschikov, A., Williams, C. F., Jeffrey, S. S., Botstein, D. & Brown, P. O. (1999) Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat Genet,* **23,** 41-6.

Press MF, Pike MC, Hung G, Zhou JY, Ma Y, George J, Dietz-Band J, James W, Slamon DJ, Batsakis JG, *et al* (1994) Amplification and overexpression of HER-2/neu in carcinomas of the salivary gland: correlation with poor prognosis. *Cancer Res* **54**(21):5675-82

Puck, J. M. (1998) The timing of twinning: more insights from X inactivation. *Am J Hum Genet,* **63,** 327-8.

Qian, Y.-W.; Lee, E. Y.-H. P (1995) Dual retinoblastoma-binding proteins with properties related to a negative regulator of Ras in yeast *J Biol Chem* **270,** 25507-25513

Qian, Y.-W.; Wang, Y.-C. J.; Hollingsworth, R. E., Jr.; Jones, D.; Ling, N.; Lee, E. Y.-H. P (1993) A retinoblastoma-binding protein related to a negative regulator of Ras in yeast *Nature* **364**, 648-652

Railo, M., Nordling, S., von Boguslawsky, K., Leivonen, M., Kyllonen, L. & von Smitten, K. (1993) Prognostic value of Ki-67 immunolabelling in primary operable breast cancer. *Br J Cancer,* **68,** 579-83.

Rawe IM, Zhan Q, Burrows R, Bennett K, Cintron C (1997) Beta-ig. Molecular cloning and in situ hybridization in corneal tissues *Invest Ophthalmol Vis Sci* **38**(5):893-900

Razin, A. & Riggs, A. D. (1980) DNA methylation and gene function. *Science,* **210,** 604-10.

Reddy J. K, Reddy M. K, Usman M. I, Lalwani N. D, Rao M. S (1986) Comparison of hepatic peroxisome proliferative effect and its implication for hepatocarcinogenicity of phthalate esters, di(2-ethylhexyl) phthalate, and di(2-ethylhexyl) adipate with a hypolipidemic drug. *Environ Health Perspect* **65**:317-27

Reiner, A., Neumeister, B., Spona, J., Reiner, G., Schemper, M. & Jakesz, R. (1990) Immunocytochemical localization of estrogen and progesterone receptor and prognosis in human primary breast cancer. *Cancer Res,* **50,** 7057-61.

Reyal F, Stransky N, Bernard-Pierrot I, Vincent-Salomon A, de Rycke Y, Elvin P, Cassidy A, Graham A, Spraggon C, Desille Y, Fourquet A, Nos C, Pouillart P, Magdelenat H, Stoppa-Lyonnet D, Couturier J, Sigal-Zafrani B, Asselain B, Sastre-Garau X, Delattre O, Thiery JP, Radvanyi F. (2005) Visualizing chromosomes as transcriptome correlation maps: evidence of chromosomal domains containing co-expressed genes-- a study of 130 invasive ductal breast carcinomas. *Cancer Res* **65**(4), 1376-83

Rice, J. C. & Futscher, B. W. (2000) Transcriptional repression of BRCA1 by aberrant cytosine methylation, histone hypoacetylation and chromatin condensation of the BRCA1 promoter. *Nucleic Acids Res,* **28,** 3233-9.

Rice, J. C., Ozcelik, H., Maxeiner, P., Andrulis, I. & Futscher, B. W. (2000) Methylation of the BRCA1 promoter is associated with decreased BRCA1 mRNA levels in clinical breast cancer specimens. *Carcinogenesis,* **21,** 1761-5.

Ross, J. S. & Fletcher, J. A. (1999a) HER-2/neu (c-erb-B2) gene and protein in breast cancer. *Am J Clin Pathol,* **112,** S53-67.

Ross, J. S. & Fletcher, J. A. (1999b) The HER-2/neu oncogene: prognostic factor, predictive factor and target for therapy. *Semin Cancer Biol,* **9,** 125-38.

Ross D. D., Yang W, Abruzzo L. V., Dalton W. S., Schneider E, Lage H, Dietel M, Greenberger L, Cole S. P., Doyle L. A. (1997) Atypical multidrug resistance: breast cancer resistance protein messenger RNA expression in mitoxantrone-selected cell lines. *J Natl Cancer Inst* **91**(5), 429-33

Roylance, R., Gorman, P., Harris, W., Liebmann, R., Barnes, D., Hanby, A. & Sheer, D. (1999) Comparative genomic hybridization of breast tumors stratified by histological grade reveals new insights into the biological progression of breast cancer. *Cancer Res,* **59,** 1433-6.

Rudlowski C, Schulten HJ, Golas MM, Sander B, Barwing R, Palandt JE, Schlehe B, Lindenfelser R, Moll R, Liersch T, Schumpelick V, Gunawan B, Fuzesi L. (2005) Comparative genomic hybridization analysis on male breast cancer. *Int J Cancer* Dec 13; [Epub ahead of print]

Ruffner, H. & Verma, I. M. (1997) BRCA1 is a cell cycle-regulated nuclear phosphoprotein. *Proc Natl Acad Sci U S A,* **94,** 7138-43.

Rychlik, W., Spencer, W.J. and Rhoads, R.E. (1989) *Nucl. Acids Res.* **18**, 6409.

Saiki R. K, Scharf S, Faloona F, Mullis K. B, Horn G. T, Erlich H.A, Arnheim N. (1985) Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**(4732):1350-4.

Schubert, E. L., Lee, M. K., Mefford, H. C., Argonza, R. H., Morrow, J. E., Hull, J., Dann, J. L. & King, M. C. (1997) BRCA2 in American families with four or more cases of breast or ovarian cancer: recurrent and novel mutations, variable expression, penetrance, and the possibility of families whose cancer is not attributable to BRCA1 or BRCA2. *Am J Hum Genet,* **60,** 1031-40.

Schwab M (1998) Amplification of oncogenes in human cancer cells *Bioessays* **20**(6), 473-9

Schwab M, Amler L.C., (1990) Amplification of cellular oncogenes: a predictor of clinical outcome in human cancer. *Genes Chromosome Cancer.* **1**(3), 181-93

Scully, R., Anderson, S. F., Chao, D. M., Wei, W., Ye, L., Young, R. A., Livingston, D. M. & Parvin, J. D. (1997a) BRCA1 is a component of the RNA polymerase II holoenzyme. *Proc Natl Acad Sci U S A,* **94,** 5605-10.

Scully, R., Chen, J., Ochs, R. L., Keegan, K., Hoekstra, M., Feunteun, J. & Livingston, D. M. (1997b) Dynamic changes of BRCA1 subnuclear location and phosphorylation state are initiated by DNA damage. *Cell,* **90,** 425-35.

Scully, R., Chen, J., Plug, A., Xiao, Y., Weaver, D., Feunteun, J., Ashley, T. & Livingston, D. M. (1997c) Association of BRCA1 with Rad51 in mitotic and meiotic cells. *Cell,* **88,** 265-75.

Shaywitz AJ, Greenberg ME (1999) CREB: a stimulus-induced transcription factor activated by a diverse array of extracellular signals *Annu Rev Biochem* **68,** 821-61

Shimizu, H., Ross, R. K., Bernstein, L., Yatani, R., Henderson, B. E. & Mack, T. M. (1991) Cancers of the prostate and breast among Japanese and white immigrants in Los Angeles County. *Br J Cancer,* **63,** 963-6.

Sinclair, C. S., Rowley, M., Naderi, A. & Couch, F. J. (2003) The 17q23 amplicon and breast cancer. *Breast Cancer Res Treat,* **78,** 313-22.

Singletary S. E., Allred C, Ashley P, Bassett L. W., Berry D, Bland K. I., Borgen P. I., Clark G, Edge S. B., Hayes D. F., Hughes L. L., Hutter R. V., Morrow M, Page D. L., Recht A, Theriault R. L., Thor A, Weaver D. L., Wieand H. S., Greene F. L. (2002) Revision of the American Joint Committee on Cancer staging system for breast cancer. *J Clin Oncol.* **20**(17):3628-36

Skonier J, Neubauer M, Madisen L, Bennett K, Plowman GD, Purchio AF. (1992) cDNA cloning and sequence analysis of beta ig-h3, a novel gene induced in a human adenocarcinoma cell line after treatment with transforming growth factor-beta. *DNA Cell Biol* **11**(7), 511-22

Slamon, D. J. (1987) Proto-oncogenes and human cancers. *N Engl J Med,* **317,** 955-7.

Slamon, D. J., Godolphin, W., Jones, L. A., Holt, J. A., Wong, S. G., Keith, D. E., Levin, W. J., Stuart, S. G., Udove, J., Ullrich, A. & et al. (1989) Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. *Science,* **244,** 707-12.

Slamon, D. J., Leyland-Jones, B., Shak, S., Fuchs, H., Paton, V., Bajamonde, A., Fleming, T., Eiermann, W., Wolter, J., Pegram, M., Baselga, J. & Norton, L. (2001) Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N Engl J Med,* **344,** 783-92.

Snijders, A. M., Nowak, N., Segraves, R., Blackwood, S., Brown, N., Conroy, J., Hamilton, G., Hindle, A. K., Huey, B., Kimura, K., Law, S., Myambo, K., Palmer, J., Ylstra, B., Yue, J. P., Gray, J. W., Jain, A. N., Pinkel, D. & Albertson, D. G. (2001) Assembly of microarrays for genome-wide measurement of DNA copy number. *Nat Genet,* **29,** 263-4.

Solinas-Toldo, S., Lampel, S., Stilgenbauer, S., Nickolenko, J., Benner, A., Dohner, H., Cremer, T. & Lichter, P. (1997) Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. *Genes Chromosomes Cancer,* **20,** 399-407.

Somasundaram, K., Zhang, H., Zeng, Y. X., Houvras, Y., Peng, Y., Zhang, H., Wu, G. S., Licht, J. D., Weber, B. L. & El-Deiry, W. S. (1997) Arrest of the cell cycle by the tumour-suppressor BRCA1 requires the CDK-inhibitor p21WAF1/CiP1. *Nature,* **389,** 187-90.

Song H, Xia SL, Liao C, Li YL, Wang YF, Li TP, Zhao MJ (2004) Genes encoding Pir51, Beclin 1, RbAp48 and aldolase b are up or down-regulated in human primary hepatocellular carcinoma *World J Gastroentol* **10**(4), 509-13

Sorlie, T., Perou, C. M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Thorsen, T., Quist, H., Matese, J. C., Brown, P. O., Botstein, D., Eystein Lonning, P. & Borresen-Dale, A. L. (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A,* **98,** 10869-74.

Stec, J., Wang, J., Coombes, K., Ayers, M., Hoersch, S., Gold, D. L., Ross, J. S., Hess, K. R., Tirrell, S., Linette, G., Hortobagyi, G. N., Fraser Symmans, W. & Pusztai, L. (2005) Comparison of the predictive accuracy of DNA array-based multigene classifiers across cDNA arrays and Affymetrix GeneChips. *J Mol Diagn,* **7,** 357-67.

Stein D, Wu J, Fuqua SA, Roonprapunt C, Yajnik V, D'Eustachio P, Moskow JJ, Buchberg AM, Osborne CK, Margolis B (1994) The SH2 domain protein GRB-7 is co-amplified, overexpressed and in a tight complex with HER2 in breast cancer *EMBO J* **13**(6):1331-40

Stewart, T. A., Pattengale, P. K. & Leder, P. (1984) Spontaneous mammary adenocarcinomas in transgenic mice that carry and express MTV/myc fusion genes. *Cell,* **38,** 627-37.

Stryer L. (1995) *Biochemistry 4th Edition.* W. H. Freeman and Company, New York

Su, L. K., Wang, S. C., Qi, Y., Luo, W., Hung, M. C. & Lin, S. H. (1998) Characterization of BRCA2: temperature sensitivity of detection and cell-cycle regulated expression. *Oncogene,* **17,** 2377-81.

Sun CL, Chao CC (2005) Cross-resistance to death ligand-induced apoptosis in cisplatin-selected HeLa cells associated with overexpression of DDB2 and subsequent induction of cFLIP. *Mol Pharmacol.* **67**(4), 1307-14

Szpirer C, Riviere M, Cortese R, Nakamura T, Islam MQ, Levan G, Szpirer J (1992) Chromosomal localization in man and rat of the genes encoding the liver-enriched transcription factors C/EBP, DBP, and HNF1/LFB-1 (CEBP, DBP, and transcription factor 1, TCF1, respectively) and of the hepatocyte growth factor/scatter factor gene (HGF) *Genomics* **13**(2), 293-300

Tagawa, M., Sakamoto, T., Shigemoto, K., Matsubara, H., Tamura, Y., Ito, T., Nakamura, I., Okitsu, A., Imai, K. & Taniguchi, M. (1990) Expression of novel DNA-binding protein with zinc finger structure in various tumor cells. *J Biol Chem,* **265,** 20021-6.

Tang, J. Y.; Hwang, B. J.; Ford, J. M.; Hanawalt, P. C.; Chu, G. (2000) Xeroderma pigmentosum p48 gene enhances global genomic repair and suppresses UV-induced mutagenesis *Molec Cell* **5,** 737-744

Tanner, J. M. (1973) Growing up. *Sci Am,* **229,** 34-43.

Tanner, J. M., (1973) Trends towards earlier menarche in London, Oslo, Copenhagen, the Netherlands and Hungary. *Nature* **243**: 95-96

Tarsounas, M., Davies, D. & West, S. C. (2003) BRCA2-dependent and independent formation of RAD51 nuclear foci. *Oncogene,* **22,** 1115-23.

Taunton J, Hassig CA, Schreiber SL (1996) A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p *Science* **272**(5260), 408-11

Thor, A. D., Moore, D. H., II, Edgerton, S. M., Kawasaki, E. S., Reihsaus, E., Lynch, H. T., Marcus, J. N., Schwartz, L., Chen, L. C., Mayall, B. H. & et al. (1992) Accumulation of p53 tumor suppressor gene protein: an independent marker of prognosis in breast cancers. *J Natl Cancer Inst,* **84,** 845-55.

Thornton, M. A.; Poncz, M.; Korostishevsky, M.; Yakobson, E.; Usher, S.; Seligsohn, U.; Peretz, H (1999) The human platelet alpha-IIb gene is not closely linked to its integrin partner beta-3 *Blood* **94**, 2039-2047

Tirkkonen, M., Johannsson, O., Agnarsson, B. A., Olsson, H., Ingvarsson, S., Karhu, R., Tanner, M., Isola, J., Barkardottir, R. B., Borg, A. & Kallioniemi, O. P. (1997) Distinct somatic genetic changes associated with tumor progression in carriers of BRCA1 and BRCA2 germ-line mutations. *Cancer Res,* **57,** 1222-7.

Tominaga, S. (1985) Cancer incidence in Japanese in Japan, Hawaii, and western United States. *Natl Cancer Inst Monogr,* **69,** 83-92.

Toms J. R., (2004) Cancer Research UK CancerStats Monograph. London. Cancer Research UK

Trojani, M. (1991) *A Colour Atlas of Breast Histopathology*, Chapman and Hall, UK.

Tsai, Y. C., Lu, Y., Nichols, P. W., Zlotnikov, G., Jones, P. A. & Smith, H. S. (1996) Contiguous patches of normal human mammary epithelium derived from a single stem cell: implications for breast carcinogenesis. *Cancer Res,* **56,** 402-4.

Turton, N. J., Judah, D. J., Riley, J., Davies, R., Lipson, D., Styles, J. A., Smith, A. G. & Gant, T. W. (2001) Gene expression and amplification in breast carcinoma cells with intrinsic and acquired doxorubicin resistance. *Oncogene,* **20,** 1300-6.

Tutt A, Bertwistle D, Valentine J, Gabriel A, Swift S, Ross G, Griffin C, Thacker J, Ashworth A. (2001) Mutation in Brca2 stimulates error-prone homology-directed repair

of DNA double-strand breaks occurring between repeated sequences. *EMBP J* **20**(17):4704-16

Tyler JK, Bulger M, Kamakaka RT, Kobayashi R, Kadonaga JT (1996) The p55 subunit of Drosophila chromatin assembly factor 1 is homologous to a histone deacetylase-associated protein *Mol Cell Biol* **16**(11), 6149-59

Underwood, J. C. E. (1996) *General and Systemic Pathology, 2nd Edition,* London, Churchill Livingstone.

van 't Veer, L. J., Dai, H., van de Vijver, M. J., He, Y. D., Hart, A. A., Mao, M., Peterse, H. L., van der Kooy, K., Marton, M. J., Witteveen, A. T., Schreiber, G. J., Kerkhoven, R. M., Roberts, C., Linsley, P. S., Bernards, R. & Friend, S. H. (2002) Gene expression profiling predicts clinical outcome of breast cancer. *Nature,* **415,** 530-6.

van den Brandt, P. A., Spiegelman, D., Yaun, S. S., Adami, H. O., Beeson, L., Folsom, A. R., Fraser, G., Goldbohm, R. A., Graham, S., Kushi, L., Marshall, J. R., Miller, A. B., Rohan, T., Smith-Warner, S. A., Speizer, F. E., Willett, W. C., Wolk, A. & Hunter, D. J. (2000) Pooled analysis of prospective cohort studies on height, weight, and breast cancer risk. *Am J Epidemiol,* **152,** 514-27.

Vaughn, J. P., Davis, P. L., Jarboe, M. D., Huper, G., Evans, A. C., Wiseman, R. W., Berchuck, A., Iglehart, J. D., Futreal, P. A. & Marks, J. R. (1996) BRCA1 expression is induced before DNA synthesis in both normal and tumor-derived breast cells. *Cell Growth Differ,* **7,** 711-5.

Vayssiere B, Zalcman G, Mahe Y, Mirey G, Ligensa T, Weidner KM, Chardin P, Camonis J. (2000) Interaction of the Grb7 adapter protein with Rnd1, a new member of the Rho family. *FEBS Lett* **467**(1), 91-6

Venter, D. J., Tuzi, N. L., Kumar, S. & Gullick, W. J. (1987) Overexpression of the c-erbB-2 oncoprotein in human breast carcinomas: immunohistological assessment correlates with gene amplification. *Lancet,* **2,** 69-72.

Verreault A, Kaufman PD, Kobayashi R, Stillman B (1998) Nucleosomal DNA regulates the core-histone-binding subunit of the human Hat1 acetyltransferase *Curr Biol* **8**(2), 96-108

Walker, R. A. (1996) General and Systemic Pathology. Edinburgh, Churchill Livingstone.

Walker, R. A., Lees, E., Webb, M. B. & Dearing, S. J. (1996) Breast carcinomas occurring in young women (< 35 years) are different. *Br J Cancer,* **74,** 1796-800.

Wallgren, A., Silfversward, C. & Hultborn, A. (1977) Carcinoma of the breast in women under 30 years of age: a clinical and histopathological study of all cases reported as carcinoma to the Swedish Cancer Registry, 1958-1968. *Cancer,* **40,** 916-23.

Wang, H., Shao, N., Ding, Q. M., Cui, J., Reddy, E. S. & Rao, V. N. (1997) BRCA1 proteins are transported to the nucleus in the absence of serum and splice variants BRCA1a, BRCA1b are tyrosine phosphoproteins that associate with E2F, cyclins and cyclin dependent kinases. *Oncogene,* **15,** 143-57.

Wang, R.-F., Cao, W.W. and Johnson, M.G. (1992) *BioTechniques* **12,** 702.

Wang, Q., Zhang, H., Kajino, K. & Greene, M. I. (1998) BRCA1 binds c-Myc and inhibits its transcriptional and transforming activity in cells. *Oncogene,* **17,** 1939-48.

Wang, H.; Iakova, P.; Wilde, M.; Welm, A.; Goode, T.; Roesler, W. J.; Timchenko, N. A (2001) C/EBP-alpha arrests cell proliferation through direct inhibition of Cdk2 and Cdk4 *Molec Cell* **8**, 817-828

Wang, Y., Klijn, J. G., Zhang, Y., Sieuwerts, A. M., Look, M. P., Yang, F., Talantov, D., Timmermans, M., Meijer-van Gelder, M. E., Yu, J., Jatkoe, T., Berns, E. M., Atkins, D. & Foekens, J. A. (2005) Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet,* **365,** 671-9.

Weber-Mangal, S., Sinn, H. P., Popp, S., Klaes, R., Emig, R., Bentz, M., Mansmann, U., Bastert, G., Bartram, C. R. & Jauch, A. (2003) Breast cancer in young women (< or = 35 years): Genomic aberrations detected by comparative genomic hybridization. *Int J Cancer,* **107,** 583-92.

Wei M, Grushko T. A, Dignam J, Hagos F, Nanda R, Sveen L, Xu J, Fackenthal J, Tretiakova M, Das S, Olopade O. I. (2005) BRCA1 promoter methylation in sporadic breast cancer is associated with reduced BRCA1 copy number and chromosome 17 aneusomy. *Cancer Res* **65**(23):10692-9.

Weigelt, B., Hu, Z., He, X., Livasy, C., Carey, L. A., Ewend, M. G., Glas, A. M., Perou, C. M. & Van't Veer, L. J. (2005) Molecular portraits and 70-gene prognosis signature are preserved throughout the metastatic process of breast cancer. *Cancer Res,* **65,** 9155-8.

Welcsh, P. L., Owens, K. N. & King, M. C. (2000) Insights into the functions of BRCA1 and BRCA2. *Trends Genet,* **16,** 69-74.

Wellings, S. R. (1980) A hypothesis of the origin of human breast cancer from the terminal ductal lobular unit. *Pathol Res Pract,* **166,** 515-35.

Wenger, C. R., Beardslee, S., Owens, M. A., Pounds, G., Oldaker, T., Vendely, P., Pandian, M. R., Harrington, D., Clark, G. M. & McGuire, W. L. (1993) DNA ploidy, S-phase, and steroid receptors in more than 127,000 breast cancer patients. *Breast Cancer Res Treat,* **28,** 9-20.

West M, Blanchette C, Dressman H, Huang E, Ishida S, Spang R, Zuzan H, Olson JA Jr, Marks JR, Nevins JR (2001) Predicting the clinical status of human breast cancer by using gene expression profiles *Proc Natl Acad Sci U S A* **98**(20), 11462-7

Willard, H. F. (1995) The Sex Chromosomes and X-Chromosome inactivation. *The metabolic and molecular bases of inherited disease.* New York, McGraw-Hill.

Williams, J.F. (1989) *BioTechniques* 7, 762.

Winston J.S., Ramanaryanan J, Levine E (2004) HER-2/neu evaluation in breast cancer are we there yet? *Am J Clin Pathol* **121**, Suppl:S33-49.

Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, Collins N, Gregory S, Gumbs C, Micklem G. (1995) Identification of the breast cancer susceptibility gene BRCA2. *Nature* **378**(6559):789-92.

Xia, F., Taghian, D. G., DeFrank, J. S., Zeng, Z. C., Willers, H., Iliakis, G. & Powell, S. N. (2001) Deficiency of human BRCA2 leads to impaired homologous recombination but maintains normal nonhomologous end joining. *Proc Natl Acad Sci U S A,* **98,** 8644-9.

Xie D, Shu XO, Deng Z, Wen WQ, Creek KE, Dai Q, Gao YT, Jin F, Zheng W (2000) Population-based, case-control study of HER2 genetic polymorphism and breast cancer risk. *J Natl Cancer Inst.* **92**(5), 412-7

Xu, J., Fan, S. & Rosen, E. M. (2005) Regulation of the estrogen-inducible gene expression profile by the breast cancer susceptibility gene BRCA1. *Endocrinology,* **146,** 2031-47.

Yang-Feng, T. L. S., A. L.; Weinberg, R. A.; Francke, U (1985) Oncogene from rat neuro/glioblastomas (human gene symbol NGL) is located on the proximal long arm of human chromosome 17 and EGFR is confirmed at 7p13-q11.2. *Cytogenet. Cell Genet,* **40,** 784.

Yang, H., Li, Q., Fan, J., Holloman, W. K. & Pavletich, N. P. (2005) The BRCA2 homologue Brh2 nucleates RAD51 filament formation at a dsDNA-ssDNA junction. *Nature,* **433,** 653-7.

Yarden RI, Brody LC (1999) BRCA1 interacts with components of the histone deacetylase complex. *Proc Natl Acad Sci U S A* **96**(9), 4983-8

Yoshikawa, K., Honda, K., Inamoto, T., Shinohara, H., Yamauchi, A., Suga, K., Okuyama, T., Shimada, T., Kodama, H., Noguchi, S., Gazdar, A. F., Yamaoka, Y. & Takahashi, R. (1999) Reduction of BRCA1 protein expression in Japanese sporadic breast carcinomas and its frequent loss in BRCA1-associated cases. *Clin Cancer Res,* **5,** 1249-61.

Yoshiura, K., Kanai, Y., Ochiai, A., Shimoyama, Y., Sugimura, T. & Hirohashi, S. (1995) Silencing of the E-cadherin invasion-suppressor gene by CpG methylation in human carcinomas. *Proc Natl Acad Sci U S A,* **92,** 7416-9.

Yuan, S. S., Lee, S. Y., Chen, G., Song, M., Tomlinson, G. E. & Lee, E. Y. (1999) BRCA2 is required for ionizing radiation-induced assembly of Rad51 complex in vivo. *Cancer Res,* **59,** 3547-51.

Zanocco-Marani T, Bateman A, Romano G, Valentinis B, He ZH, Baserga R (1999) Biological activities and signaling pathways of the granulin/epithelin precursor *Cancer Res* **59**(20), 5331-40

Zhang HS, Gavin M, Dahiya A, Postigo AA, Ma D, Luo RX, Harbour JW, Dean DC (2000) Exit from G1 and S phase of the cell cycle is regulated by repressor complexes containing HDAC-Rb-hSWI/SNF and Rb-hSWI/SNF. *Cell* **101**(1):79-89

Zhong, Q., Chen, C. F., Li, S., Chen, Y., Wang, C. C., Xiao, J., Chen, P. L., Sharp, Z. D. & Lee, W. H. (1999) Association of BRCA1 with the hRad50-hMre11-p95 complex and the DNA damage response. *Science,* **285,** 747-50.

Zhu, Y., Qi, C., Jain, S., Le Beau, M. M., Espinosa, R., 3rd, Atkins, G. B., Lazar, M. A., Yeldandi, A. V., Rao, M. S. & Reddy, J. K. (1999) Amplification and overexpression of peroxisome proliferator-activated receptor binding protein (PBP/PPARBP) gene in breast cancer. *Proc Natl Acad Sci U S A,* **96,** 10848-53.

Zhu, Y., Qi, C., Jia, Y., Nye, J. S., Rao, M. S. & Reddy, J. K. (2000) Deletion of PBP/PPARBP, the gene for nuclear receptor coactivator peroxisome proliferator-activated receptor-binding protein, results in embryonic lethality. *J Biol Chem,* **275,** 14779-82.

Ziegler, R. G., Hoover, R. N., Pike, M. C., Hildesheim, A., Nomura, A. M., West, D. W., Wu-Williams, A. H., Kolonel, L. N., Horn-Ross, P. L., Rosenthal, J. F. & Hyer, M. B. (1993) Migration patterns and breast cancer risk in Asian-American women. *J Natl Cancer Inst,* **85,** 1819-27.