Analysis of the Breast Cancer Proteome

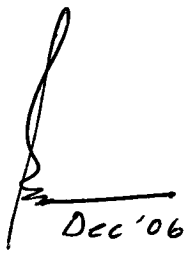
Dr Kiron D Bhatia M.B., B.S., submitted in fulfilment of the requirements towards the award of a Master in Medical Science (Surgery) degree

Discipline of Surgery

University of Tasmania

December
2006

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution except by way of background information and duly acknowledged in the thesis, and to the best of my knowledge and belief no material previously published or written by another person except where due acknowledgement is made in the text of the thesis.



Dr Kiron D Bhatia

ABSTRACT

ANALYSIS OF THE BREAST CANCER PROTEOME

BHATIA KD, Lord RJ, Stanton PD

Discipline of Surgery, University of Tasmania, Hobart

PURPOSE

The discovery of the oestrogen and *her-2* receptor proteins paved the way for the subsequent development of adjunctive therapeutic options for the treatment of breast cancer. The advancement of proteomic technology has enabled the identification of such unique proteins differentially expressed in a range of disease states including breast cancer. It is our aim to use this technique to identify potential target molecules for future therapy.

METHODS

Breast cancer samples weighing 100mg were obtained from patients in Tasmania and Greece. Normal breast tissue was obtained as control from patients undergoing reduction mammoplasty. Sample preparation was performed using a sequential extraction protocol to yield two solutions per tissue sample. This was followed by two-dimensional gel electrophoresis and resultant gels analyzed using PDQuest software.

RESULTS

Quantitative analysis revealed a 5-fold increased expression of 54 spots and decreased expression of 20 spots in breast cancer versus normal tissue using solution 2 and 3 extracted gel images. All such identified spots were located within pI (Iso-electric point) 6.0-8.2 and molecular weight 25-75kDa. Further detailed analysis enabled the identification of a number of spots showing differential expression by clinical stage including metastatic cancer, histological subtype, racial subgroup and amongst familial cancers.

CONCLUSION

The further characterization of identified spots will involve the exclusion of known markers using immunoblotting techniques with unknown spots subjected to mass spectrometric analysis or amino acid sequence determination. This approach may allow the identification of new diagnostic and/or therapeutic targets.

ACKNOWLEDGEMENT

Firstly, I would like to thank Dr R Lord, Prof P Stanton and Miss Kylie Roberts for their guidance and assistance in the laboratory. Due acknowledgement is given to the following individuals and organizations for their contribution and assistance towards the completion of this project. I would like to thank the patients who participated in this study by the provision of necessary breast tissue, both in Australia and overseas. The RHH research foundation and the Paul McKay Bolton foundation are thanked for providing the necessary funding toward the purchase of equipment and reagents. The Bio-Rad Corporation is thanked for their technical assistance in establishing a proteomic laboratory. The various surgeons who contributed towards the procurement of breast tissue samples including Prof P Stanton, Mr R Brodribb, Mr R Linacre, Mr S Sinha, Mr S Wilkinson, Mr F Kimble, Mr C Edwards and Prof S Vassilaros (Athens) are thanked along with the Department of Pathology, RHH and Hobart Pathology for their co-operation and assistance in provision of adequate tissue samples.

LIST OF ABBREVIATIONS

ATM	Ataxia Telangiectasia Mutation
BRCA-1	Breast Cancer-1 gene
BRCA-2	Breast Cancer-2 gene
DCC	Deleted in Colon Cancer gene
DCIS	Ductal Carcinoma in Situ
DNA	Deoxyribonucleic Acid
DTT	Dithiothreitol
EGFR	Epidermal Growth Factor Receptor
ESI-MS	Nanoelectrospray Ionization Mass Spectrometry
FGF	Fibroblast Growth Factor
FISH	Fluorescent In Situ Hybridization
GAP	GTPase Activating Protein
GDP	Guanosine Diphosphate
GTP	Guanosine Triphosphate
HER	Human Epidermal Growth Factor Receptor
HRT	Hormone replacement therapy
IARC	International Agency for Research on Cancer
IEF	Iso-Electric Focus
IR-MALDI-MS	Infra-Red Matrix Adsorbed Laser Desorption Ionization Mass Spectrometry
IPG	Immobilized pH Gradient
LCM	Laser Capture Micro-dissection
LCIS	Lobular Carcinoma in Situ
MALDI-TOF	Matrix Adsorbed Laser Desorption Ionization Time Of Flight
NEPHGE	Non Equilibrium pH dependant Gradient Electrophoresis
RNA	Ribonucleic Acid
mRNA	messenger RNA
RHH	Royal Hobart Hospital
SDS-PAGE	Sodium Dodecyl Sulphate- Polyacrylamide Gel Electrophoresis
TBP	Tri-Butyl Phosphine

TABLE OF CONTENTS

AIMS

1. To elucidate the differential expression of proteins in cancerous breast tissue.

2. To study racial differences in breast cancer, by the analysis of tumour samples from Tasmania and Greece.

3. To identify protein spots that are associated with a metastatic phenotype of breast cancer, by the analysis of progressive clinical stage and axillary node tissue.

FOREWORD

In *Eugene Delacroix*'s painting c1830 titled "Liberty Leading the People", Liberty is represented as an inspiring, bare-chested woman at the front of battle, stirring the men around her to fight.

It is our endeavour to study nascent breast tissue, malignant and benign, including those from different populations using protein biochemistry and proteomic technology. This subject has not been studied extensively, to date, using the above-mentioned means.

1.0 LITERATURE REVIEW

1.1 PATHOGENESIS

There are a multitude of aetiopathological factors that need to be taken into account in order to establish a causal basis for this study.

Geographical considerations

Australia is one of the few countries in which cancer registration occurs on a national basis and data are obtained from multiple sources. The Menzies Centre for Population Health Research (Hobart, Tasmania) data indicate that the risk of developing any cancer by the age of 75 years is approximately 1 in 3 according to figures derived in 1998 [1]. Cancer incidence generally increased with age and the highest incidences were found in the >60 year age groups. Of these prostate and lung cancers were most frequent in males and breast cancer in females [1]. Breast cancer, mentioned in the context of this study shall imply female beast cancer and not the much rarer condition of male breast cancer.

In 1998, breast cancer accounted for 27% of all cancers diagnosed in the Tasmanian population [1].

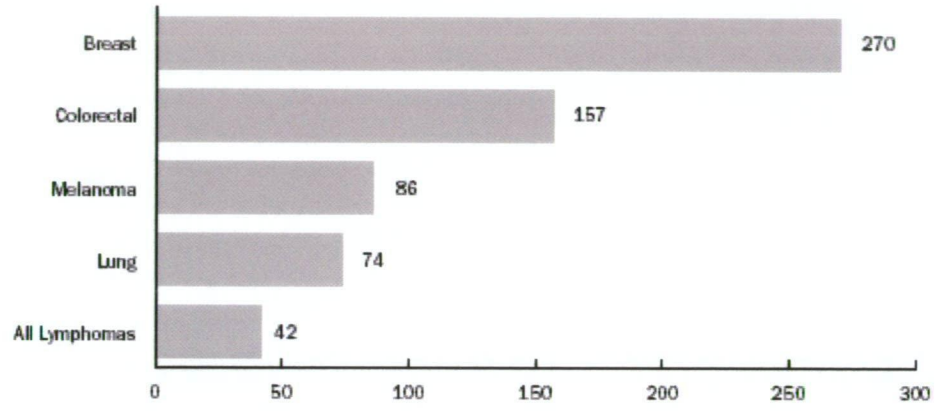


Figure 1: Common cancers, 1998: number of cancers diagnosed in all females in Tasmania [1].

Deaths due to Breast Cancer featured prominently among cancer related deaths in Tasmania in 1998 (61 cases or 14% of all cancer related deaths) [1].

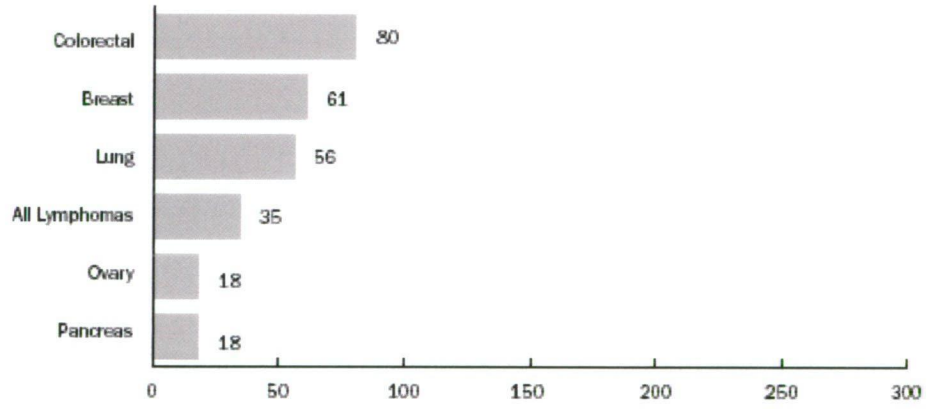


Figure 2: Common causes of cancer related deaths, 1998: total number in all females [1].

At a national level, the incidence and mortality of breast cancer can be compared. Data from the Australian Institute of Health and Welfare suggests that even on a national level breast cancer is the most commonly occurring cancer in females with a total national incidence of 10,096 cases of breast cancer cases in 1997 (27.8% of all new cancer cases) with a total mortality of 2,596 cases (17.4% of all cancer deaths) in the same year [2]. These figures correlate with those for Tasmania from the Menzies

Centre for Population Health Research. Interestingly, data also suggests a gradual increase of 2.1% per annum in incidence rates between the years 1992 and 1997 whereas mortality rates have declined slightly by 1.6% per annum [2].

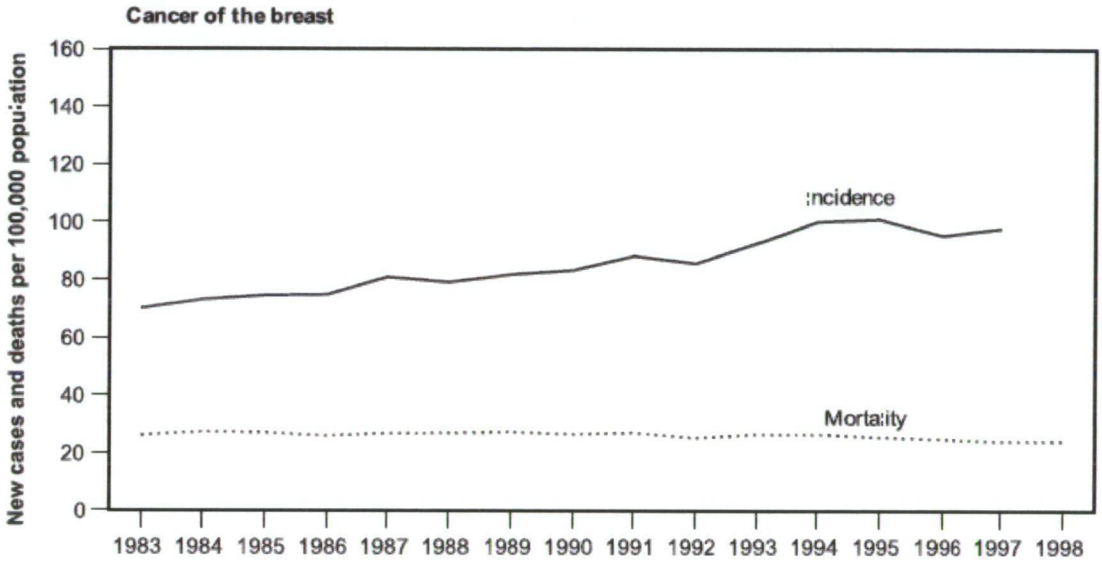


Figure 3: Trends in age-standardized incidences and mortality rate for cancer of the breast, Australia, 1983-1998 [2].

Incidence and mortality figures gathered worldwide vary for a large number of reasons. These may be caused by differential risk factor exposure, genetic susceptibility, detection and treatment practices, and even coding and registration procedures. It is clear however, that Australia features prominently on the world stage in terms of breast cancer incidence. The second population group involved in this study is from Athens, Greece whose incidence figures are intermediate (84.4 per 100,000 population in females of all ages). Breast cancer is largely a disease of affluent western societies such as the United States, Netherlands and Australia, and less common in countries such as Greece, Japan and Taiwan [2]. Some of these factors could be genetic in nature, but they are thought more likely to be environmental as shown by the fact that when migrants from low incidence locales

migrate to an area where the incidence is higher, they take on the higher incidence rates of their adopted nations. These factors are as yet undetermined but diet, reproductive patterns and nursing habits are thought to be involved [3-5]. We had chosen a differential analysis of tumours obtained from Greece to attempt to answer these questions.

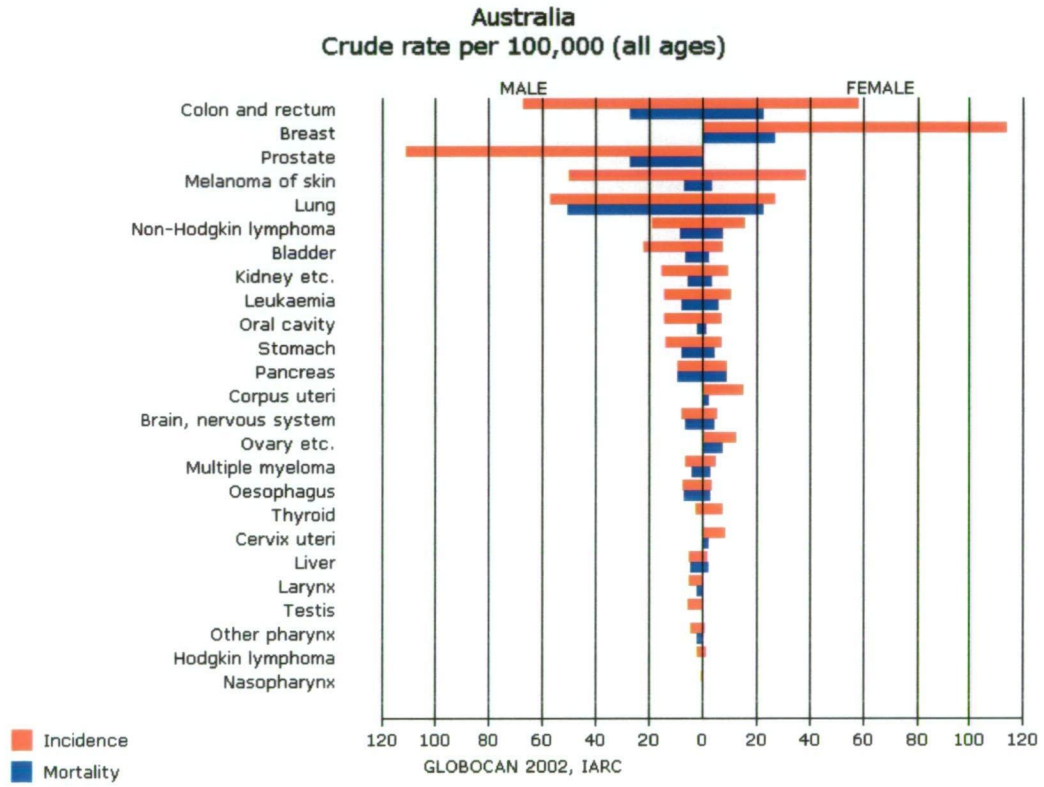


Figure 4: Cancer incidence and mortality in Australia according to IARC 2002 [6].

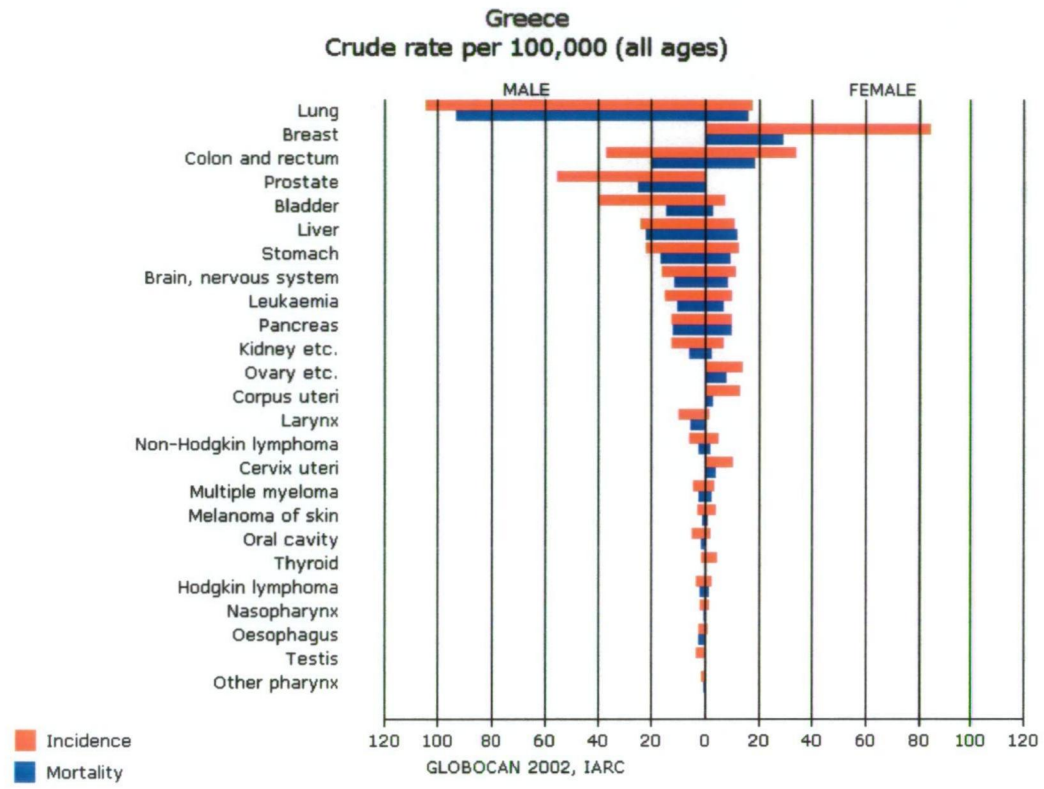


Figure 5: Cancer incidence and mortality in Greece according to IARC 2002 [6].

Genetics and family history

There has been a great deal of interest in the genetic links predisposing an individual towards the development of cancer in the modern era. The risk of acquiring a genetically linked breast cancer is said to be higher if one or more of the below mentioned variables are present.

- A diagnosis of breast cancer before the age of 50 years.
- A first degree relative who has been previously diagnosed with breast cancer.
- A known inherited genetic defect in the family such as BRCA-1, BRCA-2 or ATM.

Not surprisingly, most patients in our study group were in fact post menopausal. As such, in those who did display a family history, the significance is uncertain and hence did not warrant genetic testing. The specific inherited mutations account for only 5-10% of all breast cancers and of these the most pertinent are *BRCA-1* and *BRCA-2*.

- The *BRCA-1* gene is located on chromosome 17q21 and predisposes to breast and ovarian cancer [7]. Like the p53 gene it is also a tumour suppressor gene and its product encodes a protein secreted from breast epithelium that inhibits cell growth after binding to cell surface receptors. It has been calculated that women with a *BRCA-1* mutation have a 50% chance of developing breast cancer by the age of 50 years and an 80% chance by the age of 65 years.
- The *BRCA-2* gene is located on chromosome 13q12-13 and is also thought to be a tumour suppressor gene linked to breast but not ovarian cancer. It carries with it a similar lifetime risk although these tumours tend to have a better prognosis due to oestrogen receptor positive status. At this stage it is not entirely clear how mutations in these genes manifest in an increased predisposition to breast cancer, but there is evidence indicating that a ring protein (*BARD-1*) is required to interact with the *BRCA-1* segment and that the *BARD1-BRCA1* interaction is disrupted by mis-sense mutations indicating that *BARD-1* may be involved in mediating tumour suppression by *BRCA-1* [8].
- The Li-Fraumeni syndrome delineates a condition in which individuals inherit a germ line mutation in the *p53* tumour suppressor gene located on chromosome 17p13.1. Individuals with this inherited mutation show an increased susceptibility to develop a wide range of tumours including carcinomas, sarcomas, lymphomas and brain tumours. The condition accounts

for < 1% of breast cancers and is much less common than other inherited mutations [9]. In this syndrome, breast cancer appears in women under 40 years of age and is more often bilateral than in the general population.

- The syndrome of Ataxia Telangiectasia involves a gene located on chromosome 11q22 encoding a mutated protein kinase. The ATM gene plays a critical role in the repair of DNA damage. Patients with this syndrome display an increased susceptibility towards breast cancer directly and indirectly via the targeting of tumour suppressor genes such as BRCA1 and p53. However, while women with Ataxia Telangiectasia display an increased predisposition to breast cancer, an increased frequency of ATM mutations is not found in breast cancer [10].

Aside these, there are other factors that have a significant relationship with breast cancer and these include growth promoting factors such as the proto-oncogene *c-erb B-2* (also called *her-2, c-neu*) and sporadic mutations in growth inhibiting factors such as the tumour suppressor genes, *Rb* (retinoblastoma) and *p53* [9].

Environmental factors

- Obesity – The relationship of obesity to the incidence of breast cancer is ill defined, however Trentham-Dietz *et al* have documented the greater incidence of breast cancer in the top quintile groups for height at age 20 years, recent weight and body mass index. In women who reached their highest adult weight at younger ages (< 45 years), increasing weight loss since that age showed a reduced risk of post-menopausal breast cancer. However in women

whose highest adult weight was seen after the age of 45 years, this reduced risk after weight loss was not seen [11]. In addition researchers have found differences between cohorts of increasing adiposity for rates of recall, biopsy and cancer detection. Hunt KA *et al* found an increase in the rates of recall ($p < 0.0001$), biopsy ($p < 0.0002$) and screen detected cancers ($p < 0.015$) as well as an increased median cancer size ($p < 0.02$) and advanced stage at diagnosis ($p = 0.046$) [12]. However, these results are inferential at best.

- Alcohol – The effects of alcohol in the causation of breast cancer are also ill defined. Some authors conclude that alcohol should be considered to be an established risk factor given the greater proportion of positive studies versus studies in which no causal link was identified [13].

- Smoking – The effects of smoking have been well documented to include a tumour initiating, mutagenic effect as well as, in relation to breast cancer a protective anti-oestrogenic effect. A confounding study documented by Manjer J *et al* has actually concluded that the incidence of breast cancer was higher in ex-smokers than it was found to be in current smokers or those that had never smoked. The age adjusted relative risk for all smokers was said to be 1.31 and in pre-menopausal smokers it was 1.57. The incidence of breast cancer was inversely related to the time period since quitting and was found to be highest in those that had given up 12 months prior to screening (RR=2.76) [14]. The relationship between breast cancer, smoking and the oral contraceptive pill would also need to be taken into account, although data to this effect is lacking. Clearly more studies need to be performed with this

relationship in mind even though the above study included a cohort of 10,902 women.

- Dietary fat intake – The relationship of dietary fat intake has been studied previously and results are difficult to interpret. Zemla B *et al* found no significant association with dietary fat intake when studying differences in incidence between native and migrant populations in Poland [15]. La Vecchia C *et al* studied various groups in Italy and found a significant inverse relationship of breast cancer with the intake of olive oil and other vegetable oils, but not with butter or margarine [3]. In China where the incidence of breast cancer is a fifth of that of the United States, modest associations were found by Yuan JM *et al* between the intake of unsaturated, mono- and polyunsaturated fats and breast cancer among two population groups (Shanghai and Tianjin) of varying incidence [4]. This group also displayed a strong inverse association with breast cancer and the intake of carotene, crude fibre and vitamin C. Thus, data on a relationship of breast cancer with dietary fat intake is conflicting and the results difficult to interpret.

In addition to the above-mentioned variables, breast cancer incidence has tended to fluctuate in certain areas along with public awareness of the subject. This is of course, an indirect association and reflects increased participation in screening programmes. In 1974 nationwide public attention was created when the wives of the President and Vice-President of the United States were diagnosed with breast cancer and set out to undergo surgery. During this period there was a documented sharp rise in attendance at breast screening clinics throughout all demographic groups. This can also be

reflected in more recent times with the development of the disease in celebrities. Their impact today perhaps, is not as significant given the increased general awareness of the subject.

Endocrine factors

There are a number of endocrine factors that are said to have a proven causal link and others that have less well established influences. Known positive influences towards the development of breast cancer include an age at menarche < 12 years and an age at menopause > 55 years, which are said to have a relative risk of 1.3 and 1.5-2 respectively [16]. The early onset of menstrual cycles, late menopause, delayed childbearing and a lower cumulative period of lactation reflect a prolonged exposure to endogenous oestrogens and render women with an increased propensity towards the development of breast cancer. Several reports have cited the protection from breast cancer afforded by prolonged breast-feeding. However, it has been suggested that racial differences in endocrine function do not explain the geographical variation in disease prevalence and a direct causal link will only be established with risk factor standardised studies [17].

Less well established influences in the causation of breast cancer include hormone replacement therapy (HRT) and the oral contraceptive pill. Data on HRT is conflicting but overall consensus appears to be that a short duration < 5 years appeared to be safe whereas a duration > 5 years was associated with a relative risk of > 1.4 [18]. The oral contraceptive pill is not reported to have an increased risk for breast cancer by numerous authors, but there is controversy here as well. A large

review of 54 epidemiological studies published in the Lancet in 1996 including 150,000 women has shown a very small increase, but an increase nonetheless of 0.05 excess cases per 100,000 women aged 16-19 years and 0.47 excess cases per 100,000 women aged 25-29 on the pill. It is interesting in this context, that these tumours were found to be less advanced [19].

1.2 TUMOUR MARKERS

In order to assist the surgical management of breast cancer patients, it has become common practice to perform serum and/or immunohistochemical testing for tumour markers. These may include steroid hormone receptors, *CEA,* epithelial mucins (e.g. *CA15.3*), tumour suppressor genes (e.g. *p53, retinoblastoma gene*), oncogenes (e.g. *c-myc, c-erb B2(HER2), ras*), proteases (e.g. *cathepsin D*), growth factor receptors (e.g. *EGFR*), proliferation markers (e.g. *Ki-67*), adhesion molecules (e.g. *E-cadherin, integrins*) and tumour DNA ploidy and S phase fraction [20]. A detailed discussion of these is beyond the scope of this thesis and only the salient features of proteinaceous markers that are clinically relevant are mentioned below.

1.2.1 Serum tumour markers

These include *CEA* and epithelial mucins (e.g. *CA15.3*). *CEA* is a family of high molecular weight glycoproteins produced in embryonic tissues of the gastrointestinal tract and its level is seen to be raised in between 25-50% of breast cancers. However, their levels can also be seen to be raised in a variety of conditions including hepatitis, inflammatory bowel disease, pancreatitis, chronic lung disease and even smoking.

The epithelial mucins are members of the *MUC-1* gene family, and these along with *CEA* have not been found to be useful for diagnostic purposes as their sensitivity is low (15-35% in early breast cancer) [21]. However, serum levels of these have been found to be related to tumour stage with significantly higher values obtained from patients with nodal involvement than in those without and in patients with larger tumours. This would implicate the use of these as independent prognostic factors and several studies have shown a shorter disease free survival and overall survival with rising levels, but others report conflicting data. Their main use appears to be in the monitoring of treatment and follow up for those with distant metastases and used in combination, their sensitivity increased to 95%, with patients responding to treatment showing decreasing levels of these markers [21].

1.2.2 Tissue tumour markers

Steroid hormone receptors

Steroid receptors are currently the only markers accepted in standard clinical practice, with immunohistochemistry for them being performed routinely as part of the histological diagnosis. Although, not useful in diagnosis, they have prognostic value in primary and metastatic breast cancer. The progesterone receptor is the product of a gene located on chromosome 11 and though its presence or absence in breast cancer is important, it appears to have less impact on prognosis.

In 1958, Elwood Jensen showed that by using a radioactive marker, only tissues that were able to respond to oestrogen, such as those of the female reproductive tract were

able to concentrate it when it was injected in their blood. By 1968, he had devised a test that could detect the presence of oestrogen receptors on breast cancer cells. Since then, Szego and Pietras in the 70's have identified the membrane bound oestrogen receptor which when activated by oestradiol, increased intracellular second messenger cAMP levels. This appears to be produced by a gene complex located on chromosome 14. Steroid hormone receptor positive status has been considered favourable for some time now and subsequent to its identification has become the most significant prognostic marker for breast cancer. Only recently, a second isoform of *OR* termed *ORβ* was discovered in 1996 [22] and the two isoforms exhibit different tissue distribution and overlapping but distinct actions. This has been validated by *OR* knockout studies in mice [23].

OR is a member of a large family of nuclear receptor transcription factors with characteristic structural domains. The two isoforms share 47% of their sequence identity with heterologous N-terminus transactivation domains (AF-1) and conserved DNA- and ligand-binding domains [24]. Hormone binding to the C-terminus ligand-binding domain induces a receptor dimerization and conformational change that triggers the stimulation or inhibition of genes. The exact mechanism of *OR* induced gene transcription is still poorly understood.

It has been stated that the *β* unit of *OR* is an important modulator of proliferation and invasion of breast cancer cells and loss of its expression may predispose to cancerous change [25]. Indeed, Palmieri *et al* have shown that *ORβ* variants and not *ORα* are expressed in human adult mammary fibroblasts. These results reveal that the proliferative signals derived from adult breast tissue in response to oestrogen are not

mediated by $OR\alpha$ as is currently thought [26]. Although it shall be some time yet before this controversy is clarified, proteomic sciences may help in elucidating this further.

The next step in the evolution of what we know about the oestrogen receptor led to its being used as a target for modulation in the treatment of breast cancer and other diseases in which it was involved. An anti-oestrogen drug, Tamoxifen was developed in 1962 for use as the morning after pill, although its precise mechanism of action is still unknown. A decade later its use in the treatment of breast cancer began to be investigated and it has now been the mainstay of hormonal therapy in breast cancer for the last 25 years. In relation to this, the Early Breast Cancer Trialists' Group had conclusively showed the improved 10 year survival for patients with early breast cancer (resectable) treated with adjuvant Tamoxifen [27]. Since then, further trials (Breast Cancer Prevention Trial) have been conducted by the NCI (National Cancer Institute, USA) and concluded in 1998 have shown that high risk women, when treated with Tamoxifen, reduced their chances of contracting the disease by 44% [28].

Sequence data for $OR\alpha$ has been listed on the ExPASy (Expert Protein Analysis System) server of the Swiss Institute of Bioinformatics (with corresponding mirror sites at the Australian Proteome Analysis Facility, Sydney and six other nations around the world) as a UniProtKB/Swiss-Prot entry since 1986 (ESR1_HUMAN, primary accession number P03372) and is made up of 595 amino acids with a molecular weight of 66.216kDa (http://au.expasy.org/uniprot/P03372). In addition, all eight described isoforms of $OR\beta$ have also been listed on the same server as UniProtKB/Swiss-Prot entries (ESR2_HUMAN, primary accession number Q92731)

since 1999, the prototype of these *ORβ-1* is made up of 530 amino acids with a molecular weight of 59.216kDa (http://au.expasy.org/uniprot/Q92731). Similarly, the progesterone receptor has its own corresponding UniProtKB/Swiss-Prot entry (PRGR_HUMAN, primary accession number P06401) since 1988, being defined in two isoforms A and B, consisting of 933 amino acids and a molecular weight of 98.995kDa (http://au.expasy.org/uniprot/P06401).

Proteases

Cathepsin D is a lysosomal aspartyl protease that can be rapidly assayed by conventional immunohistochemical techniques although routine use of this as a marker has not been supported due to the lack of significant data from multivariate analysis. However some authors have displayed strong associations between stromal *Cathepsin D* expression and high tumour tumour grade (p=0.003), increased tendency to local recurrence (p=0.009), regional recurrence (p=0.006), poorer disease free survival (p=0.0001) and poorer overall patient survival (p=0.0086) [29, 30]. Results are thus conflicting and results of further research are awaited. *Cathepsin D* has been listed on the ExPASy server as a UniProt/TrEMBLE entry since 2004 (CTSD protein, lysosomal aspartyl protease, Q6IB57_HUMAN, primary accession number Q6IB57) and is made up of 412 amino acids with a molecular weight of 44.552kDa (http://au.expasy.org/uniprot/Q6IB57).

Growth factor receptors

EGFR or ***Human Epidermal Growth Factor Receptor*** (*HER*) is part of the tyrosine kinase superfamily of receptors and has been shown to play an important role in the development of the normal breast as well as the progression into breast cancer. It has been found to be present in four isoforms (*HER1-4*). Recent evidence suggests that *HER1-3* are associated with a poor prognosis, whereas *HER-4* may be associated with good prognosis tumours [31].

Of these, *HER-2* is the most clinically relevant and is a 185kDa cell surface receptor, discovered in the 1970's which is a product of the *HER-2* or *c-erbB2* proto-oncogene and in the 1980's was shown by Slamon *et al* to be linked to breast cancer as an independent prognostic marker , associated with more aggressive tumours [32] [33]. *HER-2* alone has no known ligand, but forms heterodimers with *HER-1* (*EGFR*), *HER-3* and *HER-4*. This results in the activated receptor transmitting growth signals to the nucleus. Overexpression of this glycoprotein has been found to be associated with invasiveness, a lack of response to standard chemotherapy and is seen in up to 25% of breast cancers [34]. This can be correlated with several negative prognostic variables including ***OR*** negative status, high S-phase fraction, positive nodal disease, mutated *p53* and high nuclear grade.

HER-2 can be rapidly assayed using conventional immunohistochemical means (HercepTest), FISH (Fluorescent in situ hybridization, PathVysion) and CISH (Chromogenic in situ hybridization, Invitrogen). Immunohistochemical testing for this has now become routine in the assessment of breast cancer.

In parallel with the development in oestrogen receptor modulation, a recently developed monoclonal antibody to the *HER-2* receptor protein, Herceptin (Trastuzumab) is now available in some countries. Its mechanism of action is synergistic with chemotherapeutic agents and Tamoxifen [35], *HER-2* receptor down-regulation, limiting the signal for further growth [36] and the potentiation of antibody mediated cellular cytotoxicity (ADCC) by facilitating the interaction between NK (Natural Killer) cells and *HER-2* receptor/antibody complex causing lysis [37]. The landmark multicentric, randomized, controlled clinical Herceptin trial was published in 2001 encompassing 469 patients. In this report, Slamon *et al* reported that patients randomized to receive Herceptin and chemotherapy, versus those that received chemotherapy alone, showed a significantly longer median time to disease progression, higher overall response rate, longer median duration of response and longer median survival [38]. This protein has been listed on the ExPASy server as a UniProtKB/Swiss-Prot entry since 1987 (Receptor tyrosine protein kinase erbB-2, ERBB2_HUMAN, primary accession number P04626) (http://au.expasy.org/uniprot/P04626).

Proliferation markers

The expression of the human *Ki-67* protein has been found to be a marker of cellular proliferation as, being present in the nucleus, it is relocated to the chromosomal surface during mitosis and is present in all active phases of the cell cycle barring G0 (resting phase). The growth fraction of a tumour is represented by the *Ki-67* labelling index and this has been found to correlate with survival and tumour recurrence based on uni- and multivariate analyses. It does not appear however, to be related to tumour

metastases [39]. Pietilanen *et al* have shown by multivariate analysis in axillary node negative cases that tumour size and *Ki-67* labelling were independent prognostic factors [40]. The *Ki-67* molecule is a large 359kDa ubiquitous nuclear protein made of 3256 amino acids. It has been listed on the ExPASy server as a UniProtKB/Swiss-Prot entry since 1995 (Antigen KI-67, KI67_HUMAN, primary accession number P46013) (**http://au.expasy.org/uniprot/P46013**).

1.3 INTRODUCTION TO PROTEOMIC SCIENCES

All cells, eukaryotic or prokaryotic are made of a conglomeration of heterogeneous molecules, namely lipids, polysaccharides, proteins, nucleic acids, minerals and water. There can be no question as to the significance of DNA in the carriage of genetic code that allows propagation of the species. However, it has been at times quite difficult to link gene sequence data to the onset of disease. This disparity between the two is filled by secondary messengers and effector molecules such as RNA and proteins.

We know that the synthesis of effector molecules progresses through the transcription of a DNA template into mRNA and its subsequent translation to protein. The study of mRNA has arisen as a means of studying the translation of genetic code and certainly numerous studies have been performed with this idea in mind [41]. It is however, important to appreciate that while mRNA based approaches measure message abundance they do not measure the actual effector molecules, the proteins, themselves. A protein cannot be synthesized without its mRNA template, but there can be an abundance of protein in a cell when its mRNA is sparse and conversely,

there can be an abundance of mRNA in a cell and no protein synthesis. In addition, once a particular protein is synthesized from a single strand of mRNA, it can be spliced, truncated at its C- or N-termini or various substituent groups added. Hence, a particular protein can be post-translationally altered in numerous ways. These facts give rise to the analogy as stated by Wilkins, MR *et al* that the protein make-up of a cell at a particular time varies with its state of development, the pH and the environmental conditions. A hepatocyte for example, may have a drastically different protein make-up after as little as a glass of wine [42].

It was not until the 1970's that the separation of proteins was found to be possible in two dimensions. One of the first published reports of this technique by Kenrick KG *et al,* combined native Iso-electric focusing (IEF) with pore gradient SDS-PAGE (Sodium Dodecyl Sulphate – Polyacrylamide Gel electrophoresis) to separate serum proteins [43]. There was, however, a handicap in the form of a lack of capability in data analysis. With the development of information technology and specifically 'bio-informatics', software was developed that enabled the analysis of large amounts of experimental data. With this came the advent of proteomics, when, at the Siena 2D Electrophoresis meeting in late 1994, two Australian Scientists, Keith Williams and Marc Wilkins coined the word "proteome", denoting a set of PROTEins encoded by the genOME [42]. Indeed, the rise in popularity of proteomics as a science could be said to mirror software development.

With the mapping of the human genome, emphasis has been placed on the search for loci of cancer development. There have been numerous reports of identification of regions of loss of heterozygosity (LOH) and their implication in the pathogenesis of

breast cancer with the aid of the human genome[44-48]. In addition, Osborne *et al* have integrated data from 143 studies on allelic imbalance, including identification of BRCA1 and 2 loci as well 24 other regions on 16 chromosomes [44]. Other researchers have utilized genomic studies to display linkage analyses in 14 Finnish breast cancer families that did not display germline mutations in BRCA1 and 2 genes. Of most interest was region 2q32 and this would certainly warrant further study [49]. There is thus, a great potential for further progress.

The fields of genomics and proteomics are not necessarily conflicting but rather complementary, as mentioned above, proteins are the effector molecules of normal or disordered genetic loci. In other words, the proteomic profile of breast cancer is a reflection of the genetic make up of the tumour. However, proteins being effector molecules, are a potential target for modulation, implying that a disordered gene is already manifest. This has clinical implications for the treatment of breast cancer rather than prophylaxis. This conceptual difference was one of the main considerations behind the choice of proteomic study.

1.4 TWO DIMENSION GEL ELECTROPHORESIS

The human breast is a highly specialized, complex and tenacious tissue, consisting of a heterogeneous population of cells with varying function. Sample acquisition for the purpose of proteomic analyses generally requires bulk tissue collection. For general use purposes this may be adequate. In highly specialized laboratories, however, the advent of laser capture microdissection (LCM) techniques whereby a focused laser beam targets specific groups of cells, holds promise for the future. Retrieval of the

concerned group of cells is conducted by the concomitant activation of a transfer film placed adjacent to the tissue section [50].

A Medline search conducted in the year 2001 revealed little in the way of breast tissue sample preparation. However, there is evidence of the quality of sample preparation being the determinant of a reliable measurement of protein expression in patient cells or tissue biopsies [51]. As a product of the work of Franzen in the early 90's, guidelines are available for the preparation of tissue biopsies. Yet, information on the preparation of intact breast cancer/ normal breast tissue specifically is limited. Various publications exist on the application of 2D gel electrophoresis towards the separation of polypeptides amongst cultured human breast cell lines, the majority involving the use of MCF7 (oestrogen receptor positive) and MDA-MB231 (oestrogen receptor negative) and other less well described lines such as Hs578t, ZR75-1 and MB468. However, as highlighted by Giometti *et al*, these approaches failed to demonstrate the differences in terms of gene expression, between these lines and the tumours in situ [52].

The preparation of intact tissue samples has progressed greatly over the years. In the early years, solubilization consisted of a single stage extraction. It was only later that sequential extraction parameters evolved to aid in the separation of cytosolic, membrane and microsomal fractions. Weiss *et al* described a sequential extraction protocol for the separation of barley seed proteins in 1992 [53]. Molloy MP *et al* have spoken about the identification of membrane proteins in *E.coli* whole cell lysates using this method that were previously thought to be open reading frames [54]. Various biomedical laboratories have since formulated reagents that are available in

an easy to use 'kit' form along with *E.coli* whole cell lysates as a training tool. This has greatly facilitated the adequate preparation of samples.

The presence of nucleic acids has a detrimental effect on protein separation. DNA complexes dissociate under the denaturing conditions of sample preparation. This inhibits protein entry and migration into the gel. In addition, DNA binds to proteins and leads to artefact and streaking in gels. In 1986, Rabilloud described performing the extraction at high pH, at which proteins behave as anions and complex formation with anionic nucleic acids is minimized [55]. However, addition of endonuclease to the sample after solubilization has been found to be satisfactory and can be performed as a single step.

Despite the sensitivity of currently available technology including infra-red matrix assisted laser desorption ionisation mass spectrometry (IR-MALDI-MS), analysis of low abundance proteins remains challenging [42] and traditional biochemical techniques are employed to help visualize minor proteins and reduce protein sample complexity such as differential extraction [54], subcellular fractionation [56] and hydroxy-apatite chromatography [57]. The separation of basic proteins such as ribosomal and nuclear proteins had also long been held problematical and the system most widely used for the IEF of these proteins was non-equilibrium pH gradient electrophoresis (NEPHGE) and its attendant problems. A detailed description of these techniques is beyond the scope of this thesis. The advent of 'zoom in' IPG gel strips however, specifically over the alkaline range has been a major development in the resolution of basic proteins into highly reproducible 2D patterns [58].

Sample loading can be performed by various means and each has had its advocates. Traditional methods involve cup loading and in-gel re-hydration of Immobilized pH Gradient strips (IPG). Some authors also describe a paper-bridge loading method [59]. These have been reported to allow the application of a much larger protein load and an improved spot resolution. In addition, 2D gel electrophoresis techniques have become simplified of late by integrated systems where in-gel re-hydration and iso-electric focusing can be conducted as a simple, one step overnight procedure.

Immobilized pH gradient (IPG) strips were developed for the ease of predicting the position of unknown bands against proteins whose amino acid structure and iso-electric point (pI) were known. This approach had been experimentally verified by the work of Bjellqvist B *et al* [60]. Salt and buffer ions in the sample as well as in IPG strips can interact with the protein load and determine the ultimate resolution in the first dimension. Current practice dictates the application of sample along the entire length of the IPG strip. There have been reports by Strahler JR *et al* in 1988, however of more efficient sample entry occurring by its application at the anode end for samples re-hydrated in 1-2mM Tris base and at the cathode end for samples re-hydrated in 1-2mM Acetic acid [61]. This avoided in addition, the phenomenon of lateral band spread and increased spot resolution in the first dimension. This explanation is not tenable today, as the above description was applicable to the older tube gel systems that are being phased out in favour of IPG strips for ease of reproducibility. Current commercially available IPG strips are retailed in a dehydrated form (Bio-Rad).

Reducing agents such as Tributylphosphine (TBP) and Dithiothreitol (DTT) improve protein solubility by their action on disulfide bonds. Proteins cannot be viewed as independent units without bond disruption. DTT is a thiol reducing agent that forces equilibrium towards reduced cysteine groups and to a certain extent minimizes oxidation from the atmosphere. TBP is non-charged and hence, some authors believe that the use of TBP in place of DTT results in increased protein solubility and increased transfer to the second dimension [62]. However, as noted by Altland K *et al* in 1988, even in the presence of these strong reducing agents, polyacrylamide gels function as electron acceptors for sulfhydryl groups in proteins [63].

In order to achieve reproducibility between first dimension runs and permit full-scale proteome analysis, it is vital to maintain a set group of parameters. There are a number of variables that come into play, including total sample amount, voltage and time parameters. The temperature at which iso-electric focusing takes place also needs to be taken into account. Gorg *et al* described, in a landmark article published in 1991, the effects of temperature on spot position and pattern quality. Increased temperatures were found to reveal altered 2D images in terms of sample entry, resolution and background staining. First dimension separation at 20° C was found to be superior compared to 10 or 15° C in terms of resolution. At 30° C however, spots were discovered to have migrated, as evident on computer based analysis. Given this, they would have been even more surprised to find that the direction in which the spots migrated was not constant for all [64].

It has been said that broad range IPG strips are used for analysis of simple prokaryotic proteomes. It is clear though, that to enable the detailed analysis of more complex

eukaryotic organisms and tissues, a greater degree of focus is required. With this in mind, narrow range and very narrow range (gradient 3.5-4.5, 4-5, etc.) 'zoom-in' IPG strips have been developed. Mass spectrometry has identified the superior ability of very narrow range IPG strips to separate different proteins species and isoforms compared with broader range strips [65].

Following the second dimension separation on a polyacrylamide gel, visualization is undertaken by means of numerous staining techniques. In the context of proteomics, the most suitable stain available is silver and various commercially available 'kits' are available along with their protocols [66]. Spots containing between 10-100nanograms of protein can be easily detected and gels do not require specialized image acquisition tools. In addition, excised spots can be subjected to mass spectrometric analysis (although standard silver stain is not compatible with this, as an additional oxidative step in the process changes protein mass) [67]. In specialized laboratories however, SYPRO fluorescent stains are preferred for an even greater sensitivity (<10nanograms), although specialized and costly image acquisition tools are required for gel analysis. Lopez MF *et al* have reported the advantages of fluorescent stains over silver in detail and include a broader dynamic range and enhanced recovery of peptides from in-gel digests for the purpose of matrix assisted laser desorption ionisation time of flight (MALDI-TOF) spectrometry [68]. As reported by the work of Lauber WM *et al* this phenomenon also correlates with the number of protein database matches and percent sequence coverage [69].

1.5 PROTEIN IDENTIFICATION AND PROTEOME DATABASES

In the early years of 2D gel electrophoresis, although protein sequencing had been defined and automated [70], much data in gels was relegated to spot description and classification, as software capable of handling such data was not yet developed. Today however, advances in proteomic technology and software enable the precise protein spot identification on 2D gels. The aim of any proteome study is to allow the accurate identification of proteins preferentially expressed in a given tissue and newer third generation software packages such as Melanie 4 (Geneva Bioinformatics), ImageMaster 2D Platinum 5.0 (Amersham Biosciences) and PDQuest version 7.0 (Bio-Rad) have been developed to aid in this regard. The ultimate responsibility however, lies with the 'human' researcher being able to identify the protein spots of interest.

Spots of interest once identified need to be isolated. The Bio-Rad Corporation and others have developed automated spot excision tools (e.g. the Proteome Works spot cutter from Bio-Rad). This tool is integrated with PDQuest and allows precise localization of a particular spot of interest on a gel and concomitant isolation of spot containing gel material. Using such an approach, the working time, is claimed to be reduced by 50% with a corresponding reduction in errors [67].

The further characterization of proteome data can be performed by means of a peptide mass fingerprint obtained through matrix assisted laser desorption ionisation – time of flight (MALDI-TOF) mass spectrometry, liquid chromatography-tandem mass spectrometry [69], nano-electrospray ionisation mass spectrometry (ESI-MS) [71], or

by means of amino acid sequence analysis from blotted membranes subjected to automated Edman degradation [72]. These indirect and direct methods of protein sequencing are too complex to be discussed at any great length in the current context.

In addition to the above sophisticated techniques, there are simpler methods available for the characterization and identification of proteome data, though not necessarily as accurate. The matching of proteins can be performed with theoretical pI and molecular weight data (computed with the use of experimental standards) against those of known proteins in proteome databases such as SWISS-Prot/TrEMBL, PUMA, Prosite, and SWISS-2DPAGE. SWISS-Prot also allows cross-species matching against all proteins from a chosen genus, family, phylum or kingdom [42]. However, it needs to be mentioned that though SWISS-Prot contains a large number of entries, they are mainly from 30 or so species on which a protein sequence focus has been maintained. Clearly, this is not very accurate and mainly serves as a means of exclusion of other proteins. In relation to this, human breast proteome databases have been posted on the world wide web (http://proteomeweb.anl.gov/) since 1995 by the Argonne National Laboratory, Illinois [52], as well as at the Siena 2DPAGE database for malignant and normal breast proteins (http://www.bio-mol.unisi.it/2d/2d.htm). This will facilitate the identification of preferentially expressed polypeptides in breast cancer as well as generate interest in this subspecialty of proteomic investigation.

Our understanding of the pathogenesis of breast cancer has developed over the years along with the list of genetic and environmental factors involved. To this regard, there have been key scientific developments that have translated into a clinical application.

The oestrogen and her-2 receptors are prototypes that have led to therapy being developed to modulate their response. The differences displayed by these factors in varied population groups however, have not been addressed to date. This is the rationale behind our experiments.

2.0 MATERIAL AND METHODS

2.1 SAMPLE ACQUISITION AND ETHICS APPROVAL

1. Ethics approval

This was obtained from the Royal Hobart Hospital Ethics committee (*project no. 2001.02*) with the submission of a patient information sheet and consent form. A copy of the mentioned documents is included in this thesis. Similar arrangements were made for ethical approval from the Hobart Private Hospital and the Calvary Hospital.

2. Patient recruitment and sample acquisition

Surgeons performing the relevant surgery were contacted and their consent obtained for access to patients in the public hospital as well as the private sector. Patients themselves were subsequently approached and interviewed with regard to a clinical history including current medications, hormone replacement, previous treatment received and family history. Data was recorded in tabular form (Table 11, pg 72). This session included a discussion of consent issues and was most conveniently performed in a preadmission clinic setting, 1-2 weeks prior to the date of surgery. A potential complicating factor in sample procurement was the requirement for fresh breast and nodal tissue rather than tissue fixed in formalin.

Due to this, special arrangements were made with Hobart Pathology, CentrePath (private pathology diagnostics) and the Anatomical Pathology department at the Royal Hobart Hospital. This included the direct collection of suitable specimens by the author at the site and time of excision, to be transported to the relevant anatomical pathology laboratory where tissue was obtained if deemed feasible. Nodal tissue samples deemed suitable for collection were bisected and half of the node was retained by the relevant pathology lab for histological confirmation. Subsequently, the breast tissue specimen was fixed in formalin and the tissue sample for the purpose of proteomic investigation was transported to the surgical laboratory and stored at $-70^{o}C$ in an appropriately sized container without any required preparation.

2.2 SAMPLE PREPARATION FOR PROTEOME ANALYSIS

The following tables list the disposable goods, hardware and reagents used in sample preparation.

Instrument	Manufacturer/Distributor
Scalpel handle no.4	Bard-Parker (U.S.A.)
Sterile surgical blade no.25	Swann-Morton (England)
Airpure biological safety cabinet class II	Email Westinghouse
Disposable plastic petri dish	Disposable Products Pty Ltd (Australia)
Plastic weighing boats (cat no. W3001)	Sigma-Aldrich
Non-sterile gloves (powder free)	Ansell
Weighing machine (SC2020)	Ohaus (U.S.A.)
Pyrex 2ml homogenizer (cat no. 7727-02)	Crown-Scientific Pty Ltd
Ultra-Turrax T8 electric homogenizer	Inka Labortechnik
Gilson pipetman (20,200,1000microlitre)	John Morris Scientific
Tips pipetter, polypropylene (200,1000microlitre)	Sorenson
Micro test tubes 1.5ml	Eppendorf
Centrifuge 5415D	Eppendorf
Plastic test tube rack	Nalgene
Freezer F230	Fisher-Paykel
Refrigerator C270	Fisher-Paykel
Polystyrene 96 well cell, flat bottom	Corning
Microplate reader model 3550	Bio-Rad
Duran Jena-glas laboratory bottles 100ml	Schott-Mainz
Protean IEF system, including basic unit, 11 cm focusing trays with lids, 2 pairs of stainless steel forceps, pack of electrode wicks, mineral oil and cleaning brush (cat no. 165-4000)	Bio-Rad
Electrode wicks pre-cut (cat no. 165-4071)	Bio-Rad
Disposable re-hydration/equilibration tray 11cm (cat no. 165-4025)	Bio-Rad
ReadyStrip IPG strip, 11cm, pH 3-10 (cat no. 163-2014)	Bio-Rad
Criterion precast Tris-HCl gel, 4-15% (cat no. 345-0031)	Bio-Rad
Criterion 2D cell (cat no. 165-6001)	Bio-Rad
GS-800 Calibrated Imaging Densitometer, PC (cat no. 170-7980)	Bio-Rad
PDQuest 2-D Image Analysis Software, PC (cat no. 170-8603)	Bio-Rad

Table 1: Disposable goods and Hardware in 'Analysis of the breast cancer proteome'.

Reagent/Chemical	Catalogue no.	Manufacturer
ReadyPrep Sequential Extraction Kit, includes 1 vial of reagent 1 (to make 50 ml), 2 vials of reagent 2 (to make 10 ml each), 2 vials of reagent 3 (to make 10 ml each), 1 vial containing 0.6 ml of 200mM TBP	163-2100	Bio-Rad
Protease inhibitor cocktail	P-8340	Sigma
TRIZMA base	T-1503	Sigma
Urea	161-0731	Bio-Rad
CHAPS	161-0460	Bio-Rad
Bio-Lyte 3/10 Ampholyte	163-1112	Bio-Rad
Thiourea	24025-7	Aldrich
Caprylyl sulfo-betaine SB3-10	D4266	Sigma
Tri-butyl phosphine (TBP)	163-2101	Bio-Rad
Protein assay dye concentrate	500-0006	Bio-Rad
Protein assay standard (bovine gamma globulin)	500-0005	Bio-Rad
Deoxyribonuclease	D-0876	Sigma
Iso-electric focus (IEF) standard	161-0310	Bio-Rad
Bromophenol blue	B-8026	Sigma
Mineral Oil	163-2129	Bio-Rad
Sodium dodecyl sulphate (SDS)	161-0301	Bio-Rad
Glycerol	G-7893	Sigma
Dithiothreitol (DTT)	161-0611	Bio-Rad
Iodoacetamide	163-2109	Bio-Rad
Glycine	G-7126	Sigma
Precision protein molecular weight standards	161-0372	Bio-Rad
Low Melting Point (LMP) overlay agarose	162-0019	Bio-Rad
Fixative enhancer concentrate	161-0461	Bio-Rad
Methanol	10158.6B	BDH
Acetic acid	100015N	BDH
Coomassie brilliant blue R-250	161-0400	Bio-Rad
Silver Stain Plus Kit	161-0449	Bio-Rad
Silver Complex Solution	161-0462	Bio-Rad
Reduction Moderator Solution	161-0463	Bio-Rad
Image Development Reagent	161-0464	Bio-Rad
Development Accelerator Concentrate	161-0448	Bio-Rad
Sodium Azide	30111	BDH

Table 2: Reagents and chemicals used in 'Analysis of the breast cancer proteome'.

Sample preparation required the thawing of earlier obtained breast tissue. This was performed at room temperature. It was determined that samples weighing 100mg, obtained after the dissection of macroscopically evident fat tissue in a disposable

plastic petri dish would be adequate for the purpose of breast cancer proteome investigation. Factors such as amount of tumour tissue available, adequate centrifugation using appropriate Eppendorf vials and the isolation of a satisfactory number of gel spots were taken into account when coming to this conclusion. This was performed within a biological safety cabinet and any unused sample was returned to the –70°C refrigerator.

Although a standardized sample weight was used, a modified Bradford protein assay was performed on randomly selected samples. Bradford dye concentrate (Bio-Rad) was used in a 1:5 dilution with bovine gamma globulin as control in a concentration of 1.43mg/ml of de-ionized distilled water (milli-Q).

Figure 6: Dissection of breast tissue.

Sequential extraction was performed according to the protocol described by Bio-Rad Laboratories.

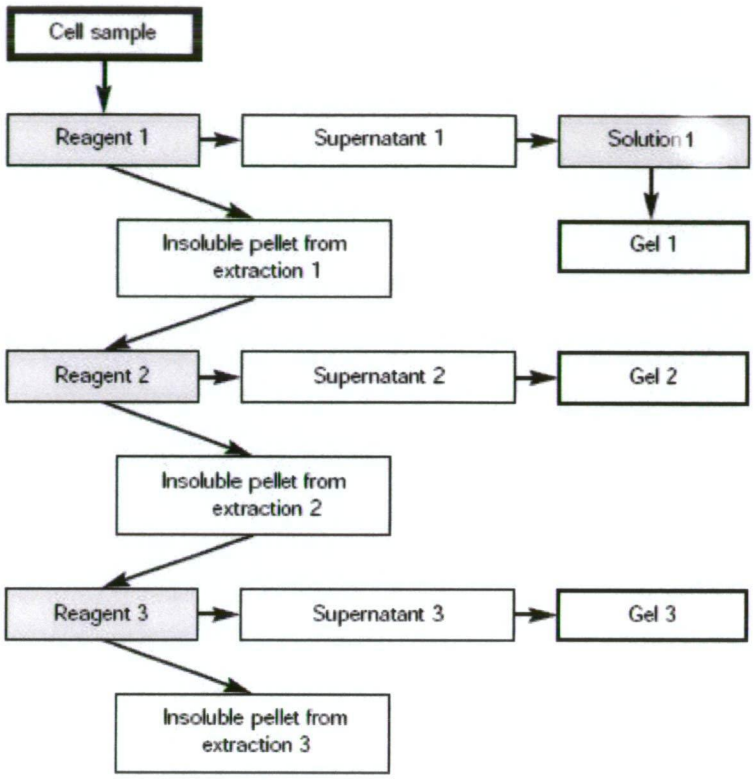


Figure 7: Flowchart for sequential extraction.

Extraction 1

- Homogenization step – To the tissue sample was added 20microlitres of protease inhibitor cocktail (mixture of protease inhibitors with broad specificity for inhibition of serine, cysteine, and aspartic proteases, and aminopeptidases) and 1 ml of sequential extraction reagent 1 (40mM Tris base). Homogenization was carried out manually in a hand held Pyrex 2ml homogenizer and aided by an electrical homogenizer. The contents were then transferred to a 1.5ml micro test tube (Eppendorf).
- Centrifugation step – The sample containing micro test tube was micro-centrifuged (5415D, Eppendorf) using an appropriate counter balance at a rate of 13,200 rpm for a period of 15 mins. Supernatant was then extracted and

transferred to a new micro test tube, labelled Solution 1 (with appropriate identifying data and date of manufacture) and stored at –20°C. The insoluble pellet was then extracted by forceps (swabbed with 70% ethanol) and transferred to the homogenizer for further processing.

Extraction 2

- Homogenization step – A 1:50 dilution (20microlitres) of reducing agent TBP in 1ml of sequential extraction Reagent 2 (Table 5, pg 68) was added to the insoluble pellet which was homogenized and transferred to a micro test tube.
- Centrifugation step – This was performed at 13,200 rpm for 13 mins. The obtained supernatant was labelled Solution 2.

Extraction 3

- Homogenisation step – a 1:50 dilution (20microlitres) of TBP reducing agent in 1ml of sequential extraction Reagent 3 (Table 6, pg 68) was added to the insoluble pellet from extraction 2. The contents were then transferred to a micro test tube as described above.
- Centrifugation step – performed as described above at 13,200 rpm for a period of 13 mins following which the supernatant was extracted, transferred to a new micro test tube and labelled Solution 3. The remaining insoluble pellet was discarded.

This protocol took into account the variable solubilization properties of the three reagents, with reagent 1 extracting hydrophilic cytosolic proteins and reagents 2 and 3 extracting increasingly hydrophobic proteins. The resultant solutions at times showed increased viscosity, reflecting the presence of nucleic acids as was evident by the poor separation during the first dimension focus. These were then subjected to enzymatic digestion by the addition of endonuclease or placement in a warm water bath. This was found to be more of a concern during the preparation of nodal samples. The sample solutions were then stored in standard Eppendorf vials at –20°C pending further processing.

2.3 SAMPLE LOADING

Due to the presence of potentially interfering abundant cytosolic proteins, Solution 1 was not processed further and only Solutions 2 and 3 were loaded on to IPG strips. To 185microlitres of thawed sample solution was added a speck of bromophenol blue to provide colour contrast.

Figure 8: Preparation of sample solution for re-hydration.

The solutions were then laid within the rows of plastic disposable re-hydration/ equilibration trays using standard micro-pipetting techniques. IPG strips of chosen pH gradient (3-10) and length (11cm) were then carefully placed on the solution taking great care to allow for even distribution and minimize the presence of air bubbles that could interfere with sample absorption. The IPG strips were then overlaid with 2ml of mineral oil and left for overnight passive re-hydration. Even absorption of sample could be estimated by the presence of a uniform thickness throughout the Acrylamide gel layer of the IPG strips.

A standard sample containing 15microlitres of IEF standard mixed with 170microlitres of reagent 3 provided for pI estimation in the first dimension focus. This contained nine natural proteins with a pI ranging from 4.45-9.6. The IEF standard was similarly incubated onto an IPG strip that was re-hydrated overnight.

2.4 FIRST DIMENSION SEPARATION

An Iso-electric focusing (IEF) tray was suitably prepared by placing filter paper wicks over the electrode wires and moistened with 8microlitres of de-ionized water (milliQ). Re-hydrated IPG strips could now be subjected to a first dimension focus. Mineral oil was discarded from the re-hydration trays and IPG strips retrieved, placed in the focusing tray and overlaid with a further 2ml of mineral oil per row. First dimension focusing then took place at room temperature (approximately 20°C) according to the protocol set forth Bio-Rad for first dimension separation using the Protean IEF cell [67].

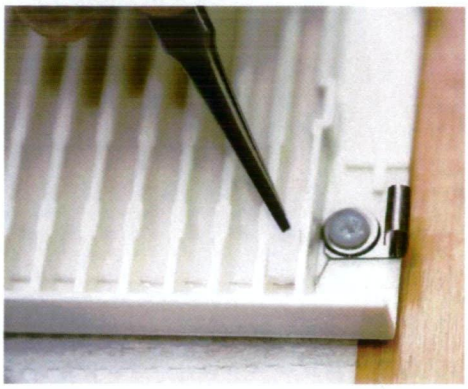


Figure 9: Placement of electrode wicks.

Figure 10: Placement of IPG strip in focusing tray.

Figure 11: The Protean IEF cell (Bio-Rad).

Step	Voltage (v)	Time (mins)	Volt-hours	Ramp
1	250	20	-	Linear
2	8000	150	-	Linear
3	8000	-	20,000	Rapid

Table 3: Set parameters for 11cm, pH gradient 3-10 IPG strips for the protean IEF cell (Bio-Rad).

Adequate focusing was ascertained by the gradual migration of bromophenol blue contrast to the opposing electrode.

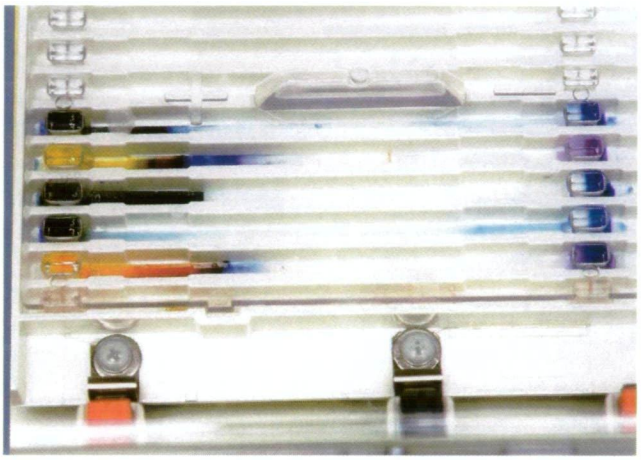


Figure 12: First dimension focus in progress.

2.5 EQUILIBRATION OF IPG STRIPS

In order to allow SDS binding in preparation for the second dimension separation, it was a pre-requisite to equilibrate focused IPG strips in SDS containing buffers. An equilibration base buffer was prepared according to the protocol set forth by Bio-Rad [67].

Equilibration took place in two steps. Firstly, according to Bio-Rad protocol (for 11cm strips), 4ml of buffer 1 (base buffer and Dithiothreitol to a concentration of 2.5%, that reduces sulfhydryl groups) was placed in a disposable plastic re-hydration/ equilibration tray and incubated with gentle agitation for a period of 10min. This was followed by placement in 4ml of buffer 2 (base buffer and dry Iodoacetamide to a concentration of 2.5% that alkylates sulfhydryl groups), and similarly incubated.

Following equilibration, the IPG strips were embedded onto the second dimension gels as described.

2.6 SECOND DIMENSION SEPARATION

In order to provide savings in time and labour, as well as enable reproducibility, precision poured pre-cast gradient gels were utilized (Criterion 4-15% Tris-HCl, Bio-Rad for 11cm IPG strips). In addition to equilibration, the transition from first dimension to second dimension focus involved the embedding of the IPG strip in the well of a pre-cast gel cassette. The gel cassette was removed from its sealed plastic container and its well was washed gently in de-ionized water (milliQ). An appropriate Criterion second dimension focusing cell (Bio-Rad) was then filled to the set mark with tank buffer to be followed by the placement of gel cassettes in their respective trays. The IPG strip was then carefully dipped in de-ionized water to rid it of excess buffer chemicals and air bubbles, placed in the cassette well and overlaid with low melting point Agarose (1% prepared solution in milliQ with bromophenol blue). Standard proteins were added to the gel in the form of an appropriately sized filter paper wick, containing 10microlitres of Precision Protein Molecular Weight Standards placed in the well beside the IPG strip. The stabilized IPG strip then underwent second dimension separation once its well was further topped with tank buffer.

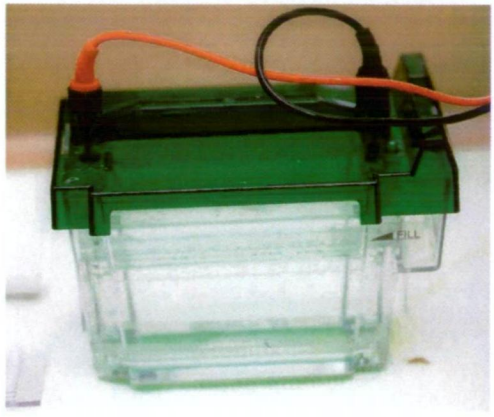


Figure 13: Criterion 2D focusing cell (Bio-Rad).

Separation was performed according to Bio-Rad protocol at 200 volts for 65 mins [67]. The Migrating front was identified with the help of the constituent bromophenol blue. Following the completion of second dimension separation, gels were extracted by breaking open the containing cassette, marked for orientation and stored. Gel cassette trays are shipped, packaged in plastic containers that were utilized as vessels for gel storage. To the containers were added 100ml of Fixative Enhancer Solution per gel followed by the gels themselves. These vessels were then placed in the refrigerator for overnight incubation.

2.7 STAINING TECHNIQUES

The silver stain plus kit (catalogue 161-0449, Bio-Rad) was utilized for gel staining as they were deemed to be sufficiently sensitive for our purpose. The components of the kit are as follows.

- Fixative Enhancer Concentrate
- Silver Complex Solution (containing NH_4NO_3 and $AgNO_3$)
- Reduction Moderator Solution (containing Tungstosilicic acid)

- Image Development Reagent (containing Formaldehyde)
- Development Accelerator Reagent

To enable the adequate staining of gels, the use of clean utensils and high quality de-ionized distilled water was absolutely essential. Staining took place using the following steps.

a. Preparation of Development Accelerator Solution

950ml of milliQ distilled water was placed in a 1L cylinder containing a Teflon stir bar. To the stirring water was added the entire 50g content of the Development Accelerator Reagent slowly. In order to adjust volume to 1 litre, further water could be added. Once fully dissolved, the solution was poured into the provided 1 litre bottle (provided in kit), labelled Development Accelerator Solution and stored at 4°C.

b. Rinse step

The staining vessel was retrieved from the refrigerator and Fixative Enhancer Solution was decanted. 100ml of milliQ water per gel was then added and the vessel gently agitated for a period of 10 mins. This cycle was repeated four times, decanting milliQ and replacing it with fresh milliQ after each cycle.

c. Stain step

A staining solution was prepared to a volume of 100ml per gel, to be used within a period of 5 minutes. A large beaker containing 35 ml milliQ water and a Teflon stir bar was placed on a stirrer. To this was added the following reagents in sequence,

- 5ml Silver Complex Solution
- 5ml Reduction Moderator Solution
- 5ml Image Development Reagent

Immediately before use, 50ml of room temperature Development Accelerator Solution was added to the beaker and stirred well. The contents of the beaker were then added to the staining vessel. Spots started to become visible within 5mins and staining was dependant on the quality of sample prepared. The average time to adequate staining was approximately 15mins.

d. Stop step

A 5% Acetic Acid solution was prepared, to be used as a Stop solution. Staining solution was decanted once the gels were adequately stained and 100ml of the Stop solution was added to the staining vessel. It was noted that a better result was obtained when gels were removed from the staining solution when slightly under-stained as the stop solution would increase spot definition and intensity. This was gently agitated for a further 15mins when stop solution was discarded and gels were ready for storage.

e. Storage step

Having discarded the stop solution, gels were rinsed in 100ml milliQ water for a further 5mins. This was then decanted and gels placed and sealed carefully in appropriately sized zip-lock clear plastic bags along with 2ml of 5% Sodium Azide.

The gels were now ready for image acquisition and long-term storage. However, in order to account for tumour heterogeneity as well as theoretical differences in the preparation of solutions, loading and running of gels on occasion, multiple gel images were created from selected samples (normal and cancerous) in order to confirm reproducibility. These individual images however were not differentially analysed or indeed kept on record.

2.8 IMAGE ACQUISITION

In order for 2D gels to be analysed with an image evaluation system they needed to be digitised. This was performed by way of Densitometry using the GS-800 calibrated imaging densitometer.

Figure 14: The GS-800 Calibrated Imaging Densitometer (Bio-Rad).

The GS-800 densitometer is integrated with PDQuest software (Bio-Rad), a programme that enables analysis of 2D gels. On the File menu of PDQuest, the GS-800 densitometer option was selected. The acquisition window of the densitometer was thus opened, displaying a control panel and a scanning window. The four basic steps to acquiring an image were then undertaken.

1. Selecting the application (gel)
2. Selecting the scan area by way of a preview scan. Using this as a guide the scan area was selected by dragging the mouse within the scan window. The border of the scan area was then marked with a frame.
3. Selecting the scan resolution. This was performed by the select scan resolution dialog box. For the purposes of this study a medium resolution (63.5_x X 63.5_y microns) was used as the default value.
4. Image acquisition. Once calibrated the densitometer could acquire images by selecting the acquire image button. As the scan was completed the image

opened with a default file name, date and username and automatically saved as such once enabled.

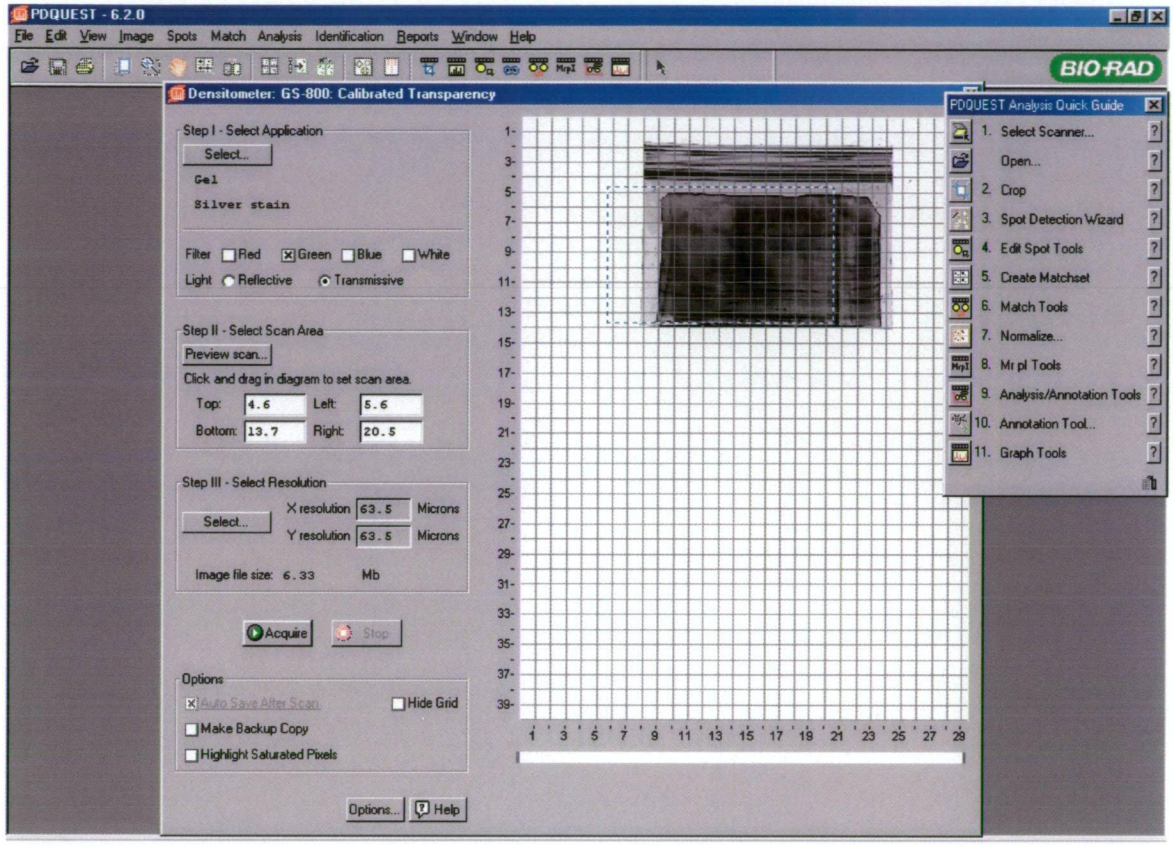


Figure 15: The GS-800 Calibrated Imaging Densitometer image acquisition sub-window.

2.9 IMAGE EDITING

This entailed the transformation of an image, adjusting brightness and contrast in order to optimise image display and did not alter the underlying data. Selecting Transform from the image menu opened a dialog box showing a smaller preview pane of the image to be transformed. Any changes were only reflected in the preview pane.

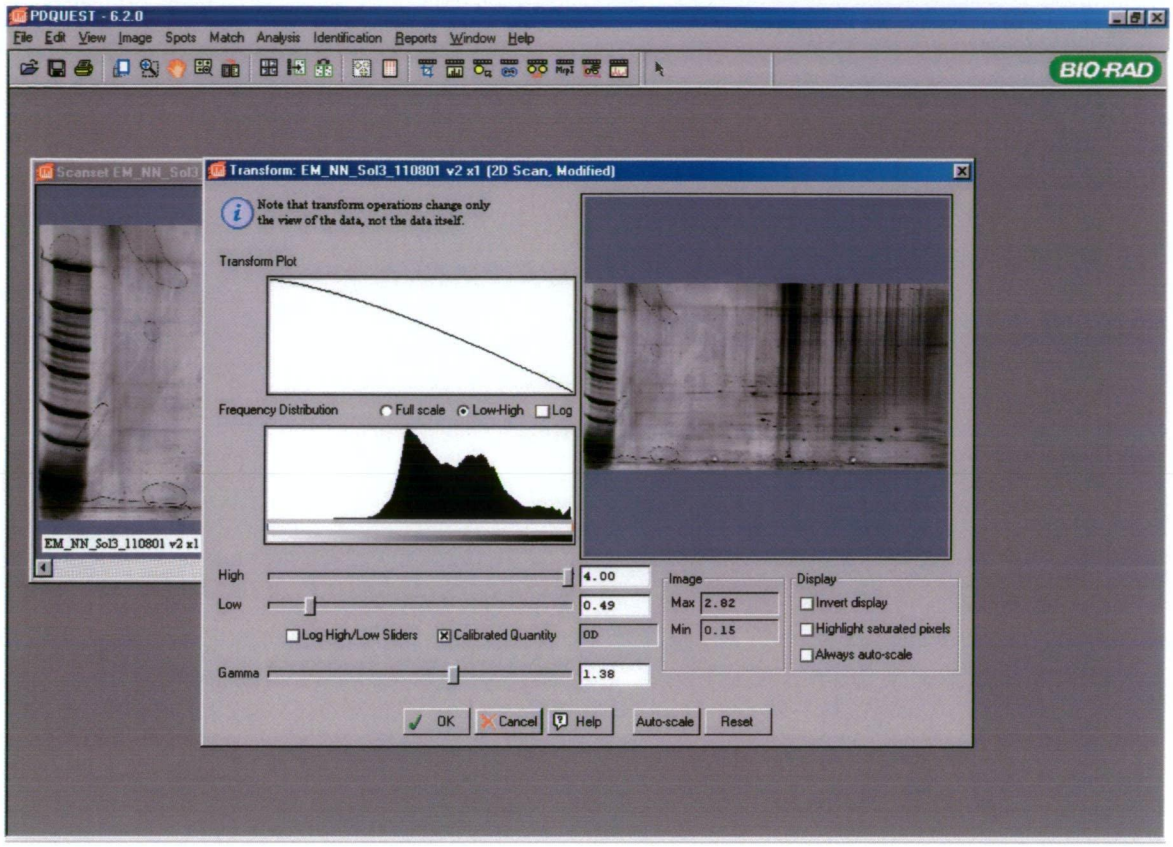


Figure 16: The Transform image dialog box.

Using the high/low and gamma (non-linear) sliders the image was optimised manually. The log checkbox enabled an easier discernment of subtle changes in signal intensity and better distinguish peaks. Once satisfactory, the transform key was activated and the new image appeared in a window labelled as version 2.

It was a pre-requisite when matching spots that all the Matchset images are of the same size and shape. In order for this to proceed, all images were cropped using the crop commands on the Image menu. Crop parameters were established using defined crop settings in terms of size. Accuracy of the crop was aided by the placement of a cross-hair on a landmark. Once the crop settings were selected, the cropped image appeared in a new window labelled as version 3.

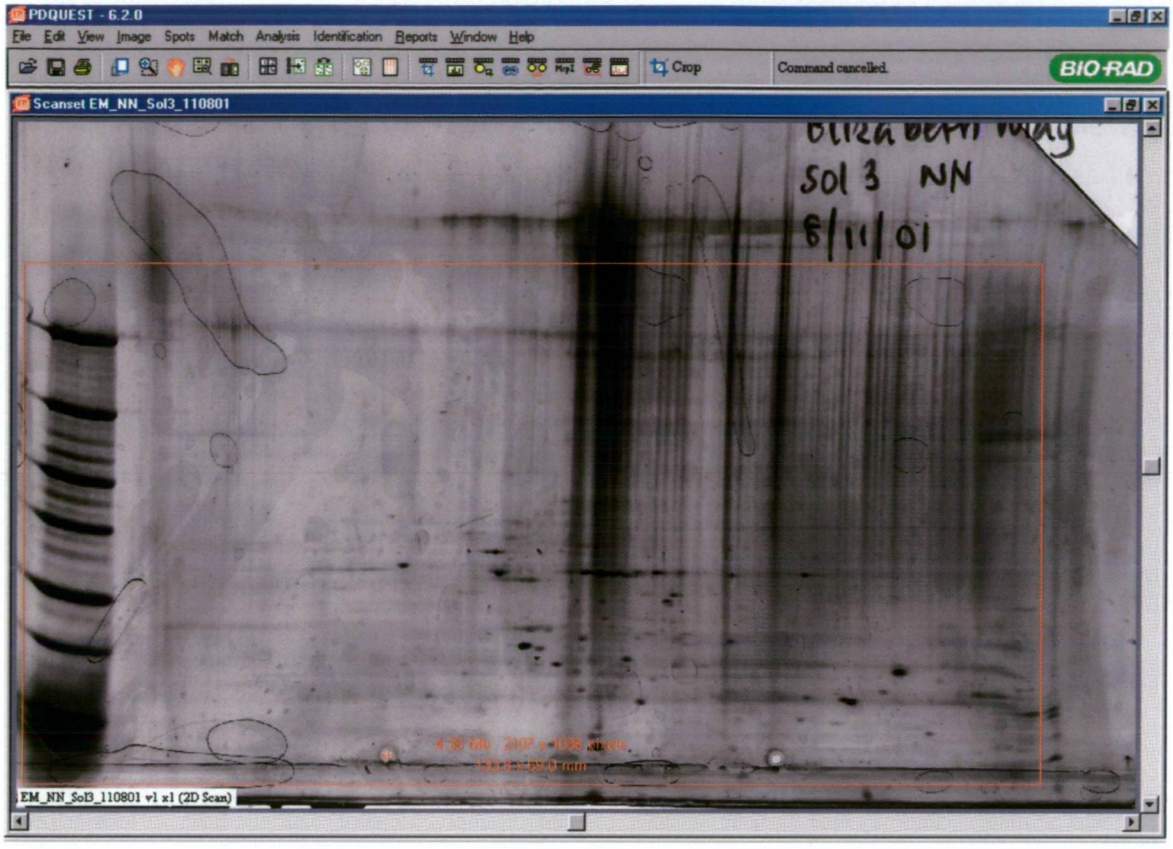


Figure 17: Placement of established crop parameters following acquisition and transformation of gel images.

2.10 SPOT DETECTION

The spot detection wizard from the spots menu on the PDQuest interface was used to select parameters for detecting spots in the gel scans.

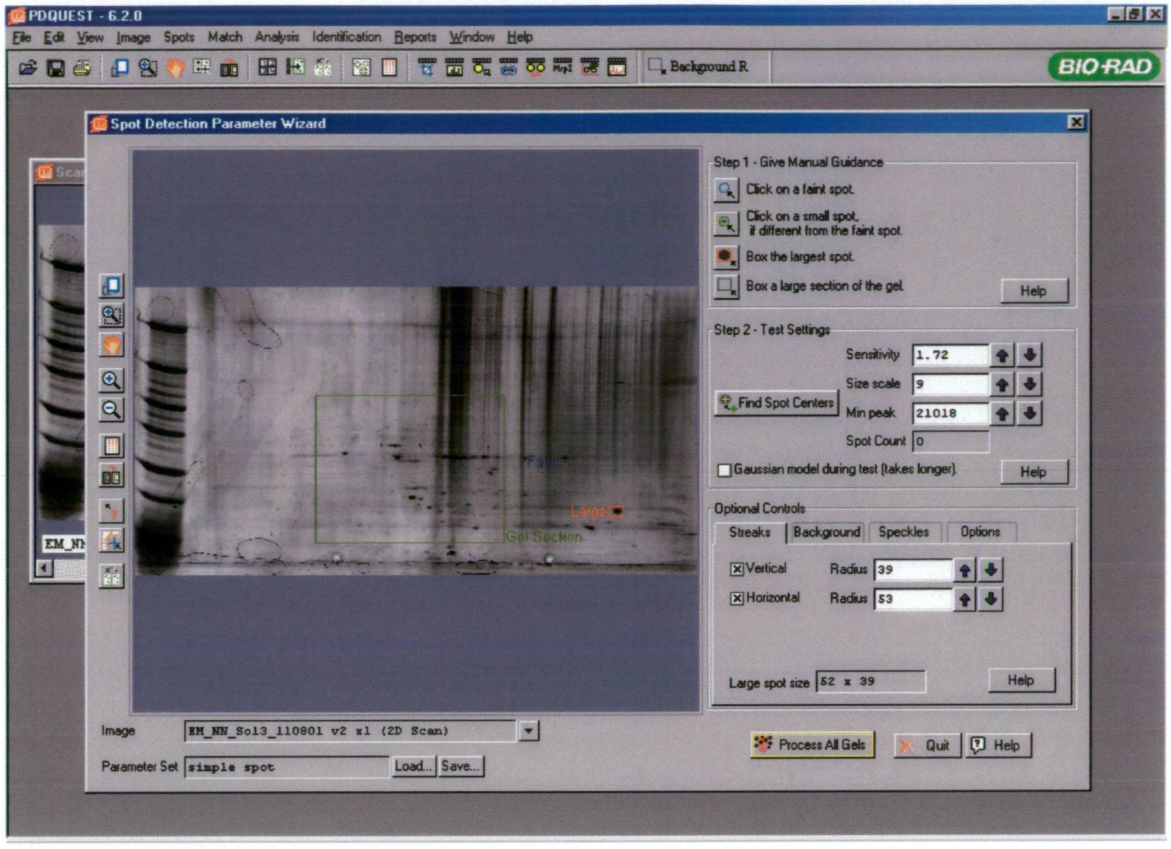


Figure 18: PDQuest spot detection Wizard.

The wizard requests manual guidance by marking features on the gel scan that were used to automatically select correct detection parameters, such as a faint spot, smallest spot, large spot and a large representative section of the gel image as background. The sensitivity of detection could be adjusted by changing the above-mentioned parameters. Test settings enabled the display of spots detected in the wizard window. In addition optional controls could be adjusted to remove gel streaking and background.

However, in most instances, in order to avoid the indiscriminate detection of artefact as spots, a minimum sensitivity was utilized for spot detection using the wizard interface. The gel was then scanned manually by scrolling through the image and

spots added or removed using the edit spot tool. This method though somewhat laborious was found to be more accurate in terms of spot detection. Legitimate protein spots presented themselves as regular, round to ovoid entities, often in chains or clusters that could be easily differentiated from linear, irregular entities that were artefact.

Once spot detection was complete, PDQuest filtered the original gel scan containing the defined number of spots and created a synthetic image containing the spots displayed as Gaussian intensity distributions. There are thus four images per gel, the original 2D gel scan, a cropped version of the scan, a filtered version and a synthetic Gaussian image. These images are all part of the same scan set and share the same root file name.

2.11 MATCHSETS

In order to compare spots across a large number of gels, a Matchset was created. This contained all the gels of a given experiment. A standard image (synthetic) was chosen from the Matchset gels, which was used as a reference gel in order to compare data. This was usually the gel image that contained the maximum number or most clearly defined spots. A particular Matchset could contain about a hundred gel images. The greater the number of images generally equates with a more difficult and arduous analysis, as the sub-window for viewing a particular spot in all the gel images of a particular Matchset becomes progressively smaller.

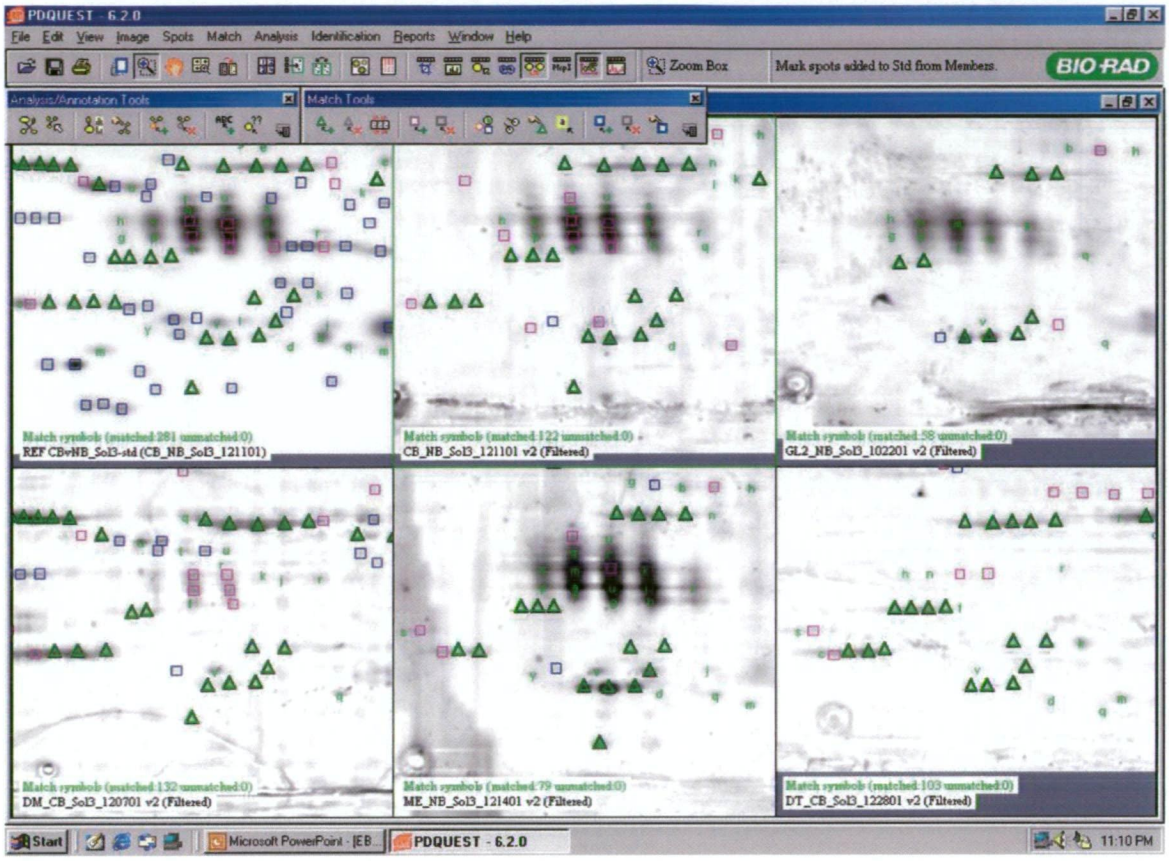


Figure 19: An example of a created Matchset with the top left hand image representing the reference image.

In our experiment, a total of nine Matchsets were created, as determined by the needs of the study (see Table 7). The last three mentioned Matchsets were based on the CBvNB_Sol3 template. Once a Matchset had been created and a standard template selected, matching of spots was conducted.

a. Land-marking and auto-matching

This was performed firstly by annotating well-resolved spots that were present consistently in all gel images as landmark spots and confirmed by studying the region at different levels of magnification. Spots were land-marked at all corners of the gel

images by selecting 'landmark' from the match menu, match toolbar or by right clicking and selection. The cursor moved along each gel image placed at the corresponding point and land-marked spots are displayed as a green triangle. The programme then by default performed an auto-match once at least two spots in each gel image had been land-marked. Auto-matched spots were displayed as green letters in lower case and unmatched spots as red ellipses.

b. Manual matching

Unmatched spots, could be manually matched by a similar process by selecting 'manual match' from the match toolbar. This did not alter the configuration of the match that was determined by the previous placement of landmark spots. These spots represent those that the computer had failed to recognize in the previous auto-match due to minor shifts in position.

c. Adding unmatched spots to the standard

Spots that failed to be recognized in the auto-match and those that could not be accounted for by manual matching were generally those not present in the gel image chosen as the standard. These spots were added to the standard image by similarly selecting 'add spot to standard' from the match toolbar and the spots could then be manually matched if possible. Spots that could not be manually matched were then seen as unique spots.

d. Ascribing molecular weight and pI values

During the run of experiments, standard proteins were required with known molecular weight and pI values in order to assign these parameters to the respective protein spots of interest. In the first dimension run this was done by the simultaneous focusing of a group of standard proteins with known pI values on an IPG strip. This enabled us to calculate the pI values of unknown proteins in our experiment by comparison. In the second dimension run however, a slightly different approach was utilized. To the left of the focused IPG strip placed in its well in the pre-cast gel cassette was placed a filter paper wick approximately 10mm in length soaked in 10microlitres of precision protein molecular weight standards. The standard proteins would separate alongside the experimental proteins in horizontal bands running down the left border of the gel. These standards were pre-stained with Coomassie blue and hence were visible even prior to staining with silver.

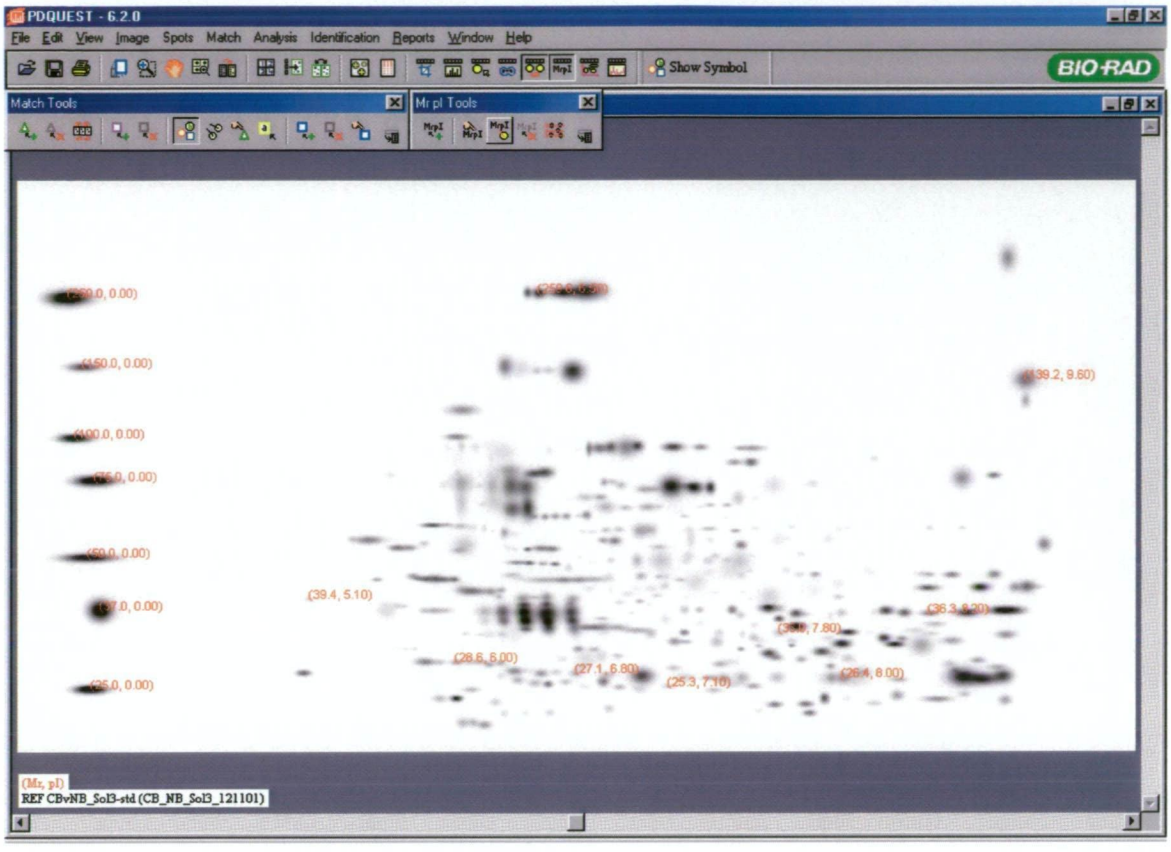


Figure 20: Ascription of Molecular weight and pI data onto a reference image (with the first number denoting Molecular weight followed by pI).

2.12 REPLICATE GROUPS AND ANALYSIS SETS

Analysis sets were then created in order to seek a particular group of proteins of interest. However, for our purposes it was not enough to elucidate the differential expression of proteins in one gel versus another in the Matchset. Our objective was to determine the differential expression of proteins in breast cancer as a whole. PDQuest provides for these requirements in the form of Replicate group analysis whereby all the gel images from a particular group of patients are treated as similar. In each of our Matchsets, Replicate groups were created depending upon the needs of the exercise as displayed in the table below.

Match set	Replicate groups
CBvNB_Sol3	Cancer breast, normal breast, CIS
CBvNB_Sol2	Cancer breast, normal breast
CBvCN_Sol3	Cancer breast, cancer node, normal node
CBvCN_Sol2	Cancer breast, cancer node, normal node
ACBvGCB_Sol3	Australian cancer breast, Greek cancer breast
Familial_Sol3	Familial cancer breast (first/second degree), cancer breast
Histosubtype_Sol3	Ductal cancer, lobular cancer, mixed
Clinicalstage_Sol3	Normal breast, CIS, clinical stages I, IIA, IIB, IIIA, IIIB (AJC)

Table 4: Required Matchsets and Replicate groups.

Once Replicate groups were created, the average spot quantities in the members of the group were sought by selecting the quantitation mode as Replicate group quantitation. This was located under the Edit menu. With replicate group quantitation each individual spot's quantitation was displayed by selecting the quantity table function on the spots menu. The average value of that spot's quantitation was then displayed along with the co-efficient of variation and the number of gels in the replicate group.

In order to accurately compare spot quantities between different gels, compensation must be made for variation in spot density that was not due to differential expression. This is referred to as Normalization. With a Match set open, the normalize button was selected from the Analysis menu. The particular method chosen was 'Total quantity in valid spots' (as sample load was consistent) with a scaling factor of PPM (parts per million).

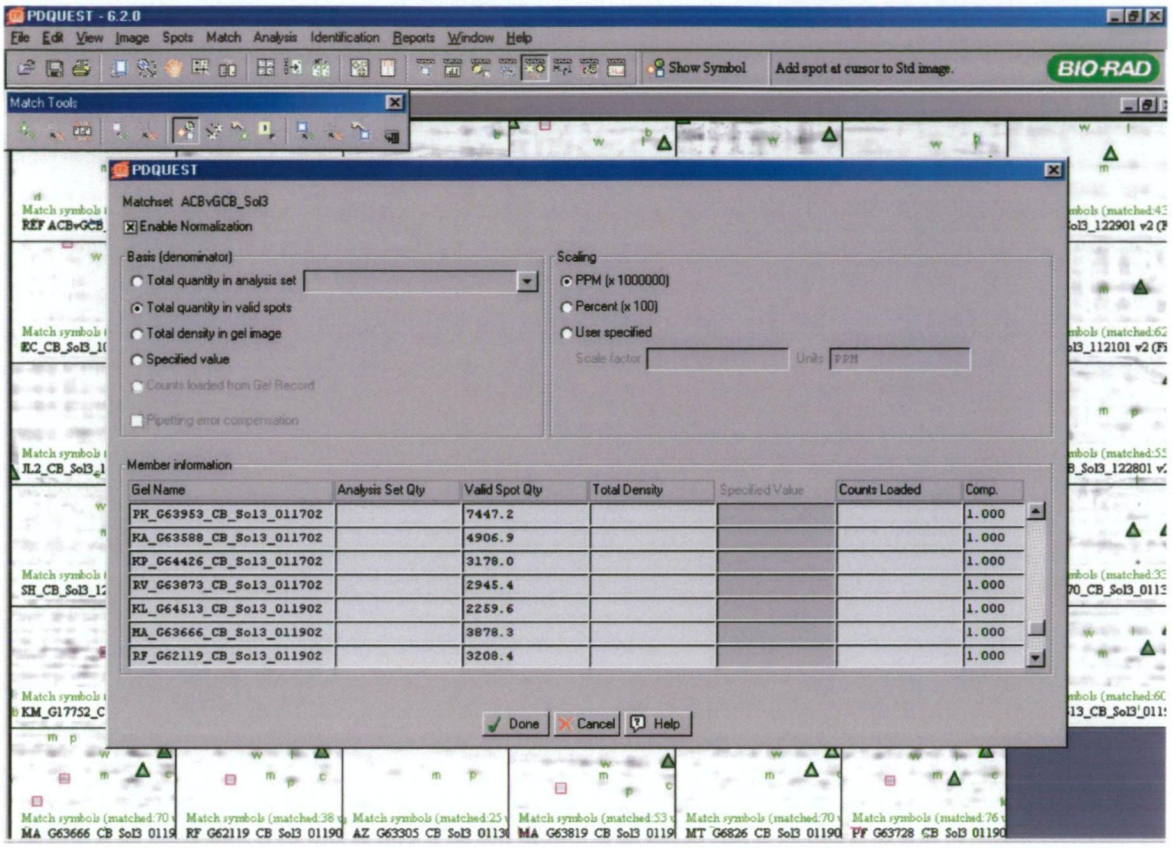


Figure 21: The Normalization sub-window using a scaling factor of PPM (parts per million).

Having normalized gel data, we were now ready to begin analysis. The 'create Analysis set' button was selected from the Analysis menu or toolbar. To begin with, a 'Quantitative Analysis set' was created by clicking on the 'Type' button and selecting the required Analysis set type in order to perform fold change analysis. The next step was to select Replicate groups (A and B) for comparison using the compare command and selecting the required group from the drop down menu. We could now dictate the fold change requirements of the experiment. In order to obtain spot data that was significant a fold change factor was selected depending on the number of spots in the set and the need for a fold change that was statistically significant. This was usually of the order of 5 fold. By then clicking on the 'Go' button, the 'Create Analysis set' dialog box was replaced by the Match set standard image containing the highlighted

spots that met the above criteria. For example, if the 'Increased in A' button was clicked and a factor of 5.00 was entered, this would include spots in the Analysis set whose quantitation was at least five times that of the corresponding spot in B. Similarly, clicking on the 'Decreased in A' button and entering a factor of 5.00 would include spots in the Analysis set whose quantitation was less than a fifth of that of the corresponding spot in B.

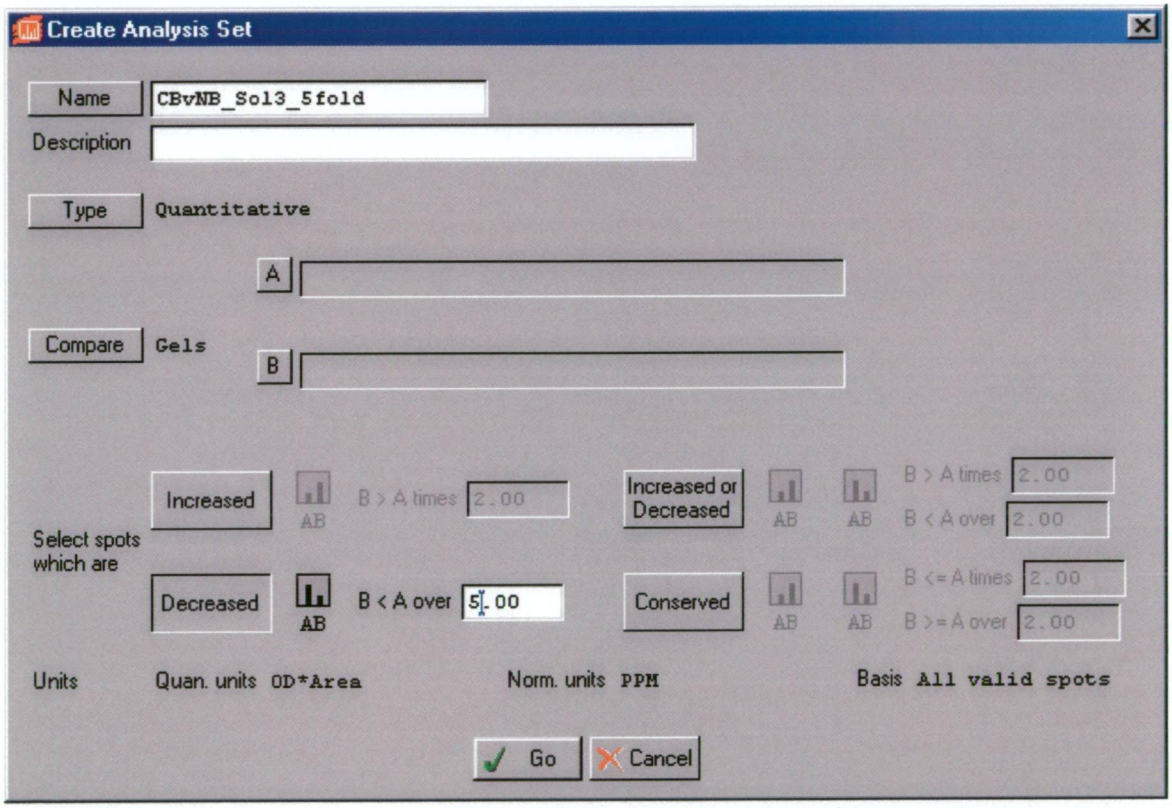


Figure 22: The Create Analysis Set dialog box denoting a Quantitative Analysis Set with a five-fold sensitivity.

For the purposes of analysing spot data that was present in already defined Analysis sets, Boolean Analysis sets were created using Boolean operators. This was performed by clicking on the 'Type' button on the 'Create Analysis set' dialog box and selecting Boolean. The wanted Analysis sets for comparison were selected along

with the Boolean operator required. By then clicking on the 'Go' button, the 'Create Analysis set' dialog box disappears, to be replaced by the Match set standard image containing the highlighted spots that meet the above criteria.

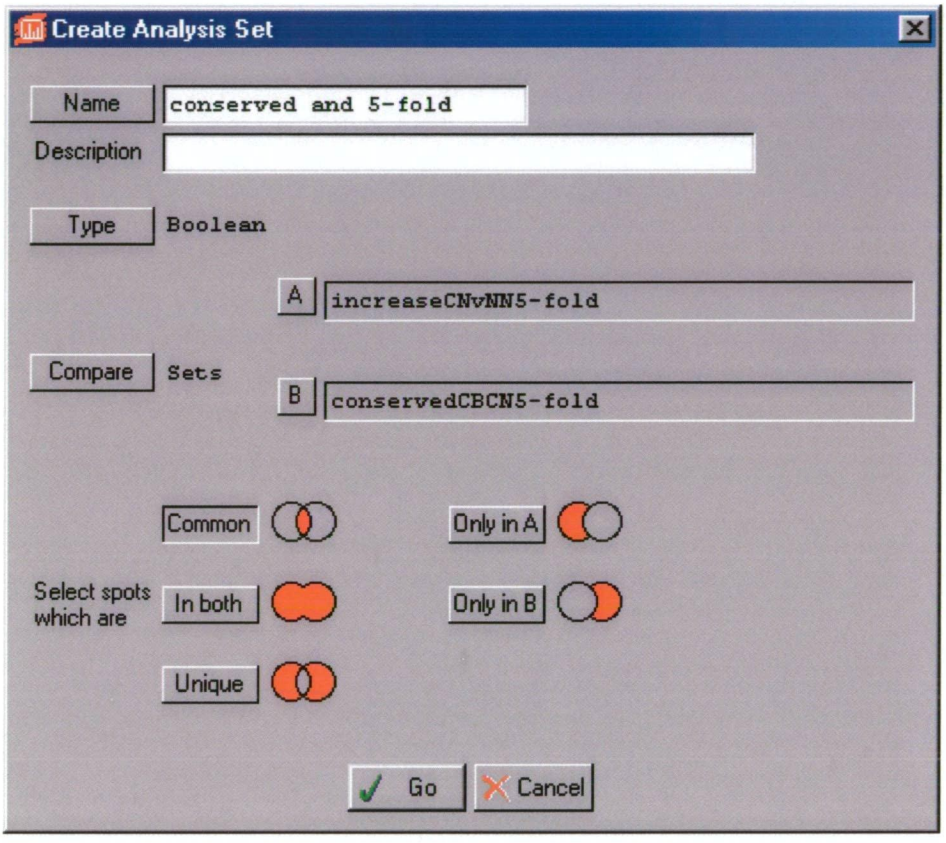


Figure 23: The Create Analysis Set dialog box denoting Boolean operators.

2.13 LIST OF REAGENTS

Reagent	Amount (final concentration)
Urea	9.61gm (8M)
Trizma base	0.1gm (40mM)
CHAPS	0.8gm (4% w/v)
Bio-Lyte 3/10 Ampholyte	40microlitres (0.2% w/v)
MilliQ	Adjust to 20ml

Table 5: Reagent 2.

Reagent	Amount (final concentration)
Urea	6.01 gm (5M)
Trizma base	0.1 gm (40mM)
Thiourea	3.04 gm (2M)
CHAPS	0.4 gm (2% w/v)
SB 3-10	0.4 gm (2% w/v)
Bio-Lyte 3/10 Ampholyte	40microlitres (0.2% w/v)
MilliQ	Adjust to 20ml

Table 6: Reagent 3.

Reagent	Amount (final concentration)
Urea	36gm (6M)
20% SDS	10ml (2%)
1.5M Trizma base	3.3ml (0.05M)
50% Glycerol	40ml (20%)
MilliQ	Adjust to 100ml

Table 7: Equilibration base buffer.

Reagent	Amount
Trizma base	30.28gm
Glycine	144.14gm
SDS	10gm
MilliQ	Adjust to 1litre

Table 8: Tank Buffer.

Reagent	Volume
Methanol	200ml
Acetic Acid	40ml
Fixative Enhancer Concentrate	40ml
MilliQ	Adjust to 400ml

Table 9: Fixative Enhancer Solution.

3.0 RESULTS

3.1 PATIENT DATA

The following Table details patient and sample data from the Tasmanian study group comprising patients of largely Anglo-Celtic background. In total, 28 breast cancer samples were obtained, of which 26 were invasive cancer samples. In some cases, as mentioned below, patients provided nodal cancer, normal breast and nodal tissue samples. All patients in the group were female and no samples of malignant or normal, breast or nodal tissue were obtained from male patients.

Sample ID	Age	Histology	TNM stage	Receptor status			Pre-operative therapy
				Oest	Prog	*c-erbB2*	
1	80	NOS	T2N0M0(IIA)$^{+}$	+	+	-	XRT
2	87	Ductal	T2N0M0(IIA)	+	+	-	
3	75	Ductal	T2N0M0(IIA)	+	+	-	
4	53	Mixed	T2N0M0(II)	+	-	-	
5	50	Ductal	T3N2M0(IIIA)*	-	-	+++	
6	77	Ductal	T2N0M0(IIA)	++	+	+++	
7	49	Ductal	T3N2M0(IIIA)	+	-	-	Chemotherapy
8	64	Lobular	T2N0M0(IIA)	NP	NP	NP	
9	47	Mixed	T2N1M0(IIB)*	+	+	-	
10	78	Ductal	T1N0M0(I)	+	+/-	-	
11	65	Mixed	T1N0M0(I)	+	+	NP	
12	59	Ductal	T1N0M0(I)	++	-	NP	
13	87	Lobular	T2N0M0(IIA)	+	+	NP	
14	91	Ductal	T2N0M0(IIA)	+	+	NP	
15	70	Ductal	T1N0M0(I)$^{+}$	+	+	NP	
16	67	Ductal	T_{is}N0M0(O)$^{+}$	+	+	NP	
17	80	Ductal	T1N0M0(I)$^{+}$	+	-	NP	
18	68	Ductal	T1N0M0(I)	+	+	NP	
19	51	Ductal	T1N0M0(I)$^{+}$	-	-	-	
20	61	Lobular	T2N1M0(IIB)$^{+}$	+	+/-	-	
21	62	Ductal	T1N0M0(I)$^{+}$	+	+	NP	
22	44	Lobular	T2N1M0(IIB)	+	+	-	
23	82	Ductal	T2N2M0(IIIA)*	+	+	-	
24	79	Lobular	T3N0M0(IIB)$^{+}$	+	+	NP	
25	61	Ductal	T_{is}N0M0(O)	+	-	NP	
26	74	Tubular	T1N1M0(IIA)$^{+}$	+	+	NP	
27	64	Ductal	T4N2M0(IIIB)*	-	-	NP	
28	87	Ductal	T2N1M0(IIB)*	NP	NP	NP	

Table 10: Tumour characteristics of Tasmanian group with dominant histology, clinical stage (TNM and AJC), receptor status and relevant pre-operative treatment. XRT=radiation treatment, NOS=Not otherwise specified, NP=Not performed, *denotes nodal cancer tissue obtained, $^{+}$denotes normal node tissue obtained.

The table below details salient family history and pre-operative treatment characteristics.

Patient ID	Family history		Pre-operative therapy	Menopause	HRT
	1°	2°			
1	-	-	XRT	Post	-
2	-	-	-	Post	+
3	-	-	-	Post	-
4	-	+	-	Post	+
5	-	+	-	Peri	-
6	+	-	-	Post	-
7	-	-	Chemotherapy	Pre	-
8	-	-	-	Post	+
9	+	-	-	Peri	-
10	-	-	-	Post	-
11	-	-	-	Post	-
12	-	-	-	Peri	-
13	-	-	-	Post	-
14	-	+	-	Post	+
15	-	-	-	Post	-
16	-	-	-	Post	-
17	+	-	-	Post	+
18	-	-	-	Pre	-
19	-	-	-	Pre	-
20	-	-	-	Peri	-
21	-	-	-	Peri	-
22	-	-	-	Pre	-
23	-	-	-	Post	-
24	-	-	-	Post	-
25	-	-	-	Peri	-
26	-	-	-	Post	+
27	+	-		Post	+
28	-	-	-	Post	-

Table 11: Further patient data including family history and relevant menopausal and pre-operative treatment characteristics.

Sample ID	Age	Concurrent tumour in ipsilateral breast
1	32	-
2	80	+
3	87	+
4	69	-
5	87	+
6	57	-
7	39	-
8	24	-
9	32	-
10	44	-
11	71	-
12	43	-

Table 12: Normal breast tissue samples obtained for this study.

As mentioned above, a total of 12 samples of normal breast tissue were obtained from the Tasmanian population. Of these, 9 were from patients having undergone reduction mammoplasty and the remaining were normal breast tissue samples from patients who had undergone mastectomy for cancer.

For the purpose of accurate spot quantitation throughout Matchset member gels, it was important to load a standard amount of protein. This was performed mainly by using a standardized sample weight of 100mg, after removing any macroscopically evident fat tissue as alluded to in the methods section. However, another way of testing this was to perform a modified Bradford assay and this was performed using randomly chosen cancerous and normal breast samples, showing grossly similar spectral absorbance.

3.2 BREAST CANCER

As detailed above in the methods section, a number of Matchsets, with their contained replicate groups were created in order to answer a particular question. Within these Matchsets were contained analysis sets that displayed the requisite spot data. As displayed in Figure 24 below the Matchset CBvNB_Sol3 standard Gaussian image contains identified spots from cancerous breast and normal tissue of solution 3 extraction. A total of 403 protein spots (not including molecular weight markers) were identified. The highlighted spots indicate quantitative analysis sets increased and decreased 5-fold respectively. By these means, 30 spots were found to show a minimum of a 5-fold increase (including 8 spots showing a 10-fold increase) in expression for cancerous breast tissue versus normal breast tissue. Conversely 15 spots showed a minimum of a 5-fold decrease (including 1 spot showing a 10-fold decrease) in expression in cancerous breast tissue versus normal breast tissue. The majority of proteins identified were found to have a molecular weight of 25-75kDa (kilodaltons) and lie within a pI range 6.0-8.2 respectively. In this context it was noted that gel images obtained from a more malignant phenotype displayed spots showing increased expression located towards the basic end of the gel image and spots showing decreased expression located towards the acidic end of the gel image, whereas the converse was true with gel images from a normal phenotype.

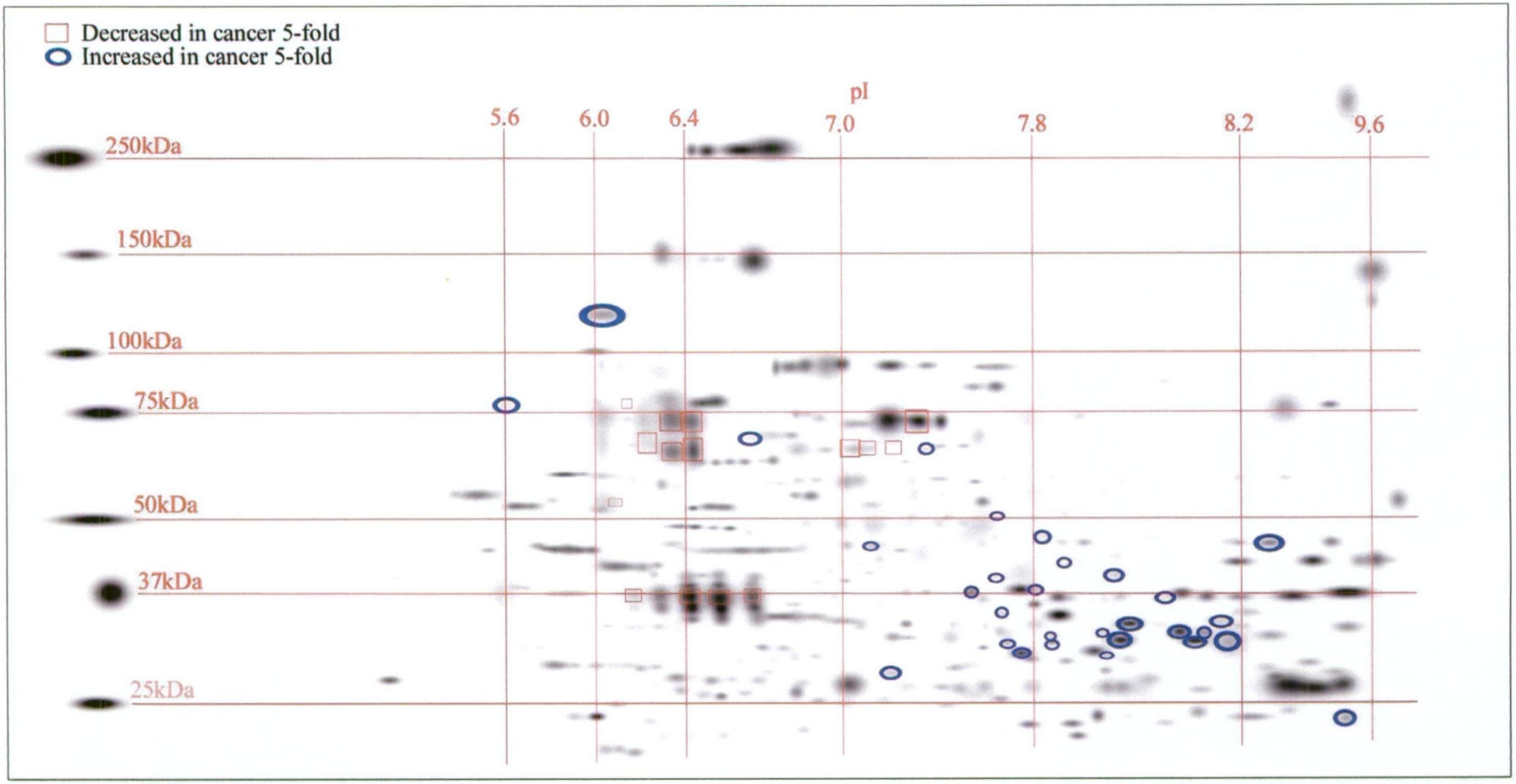


Figure 24: Identified spots from solution 3 extraction showing increased and decreased quantitative expression in breast cancer versus normal breast tissue.

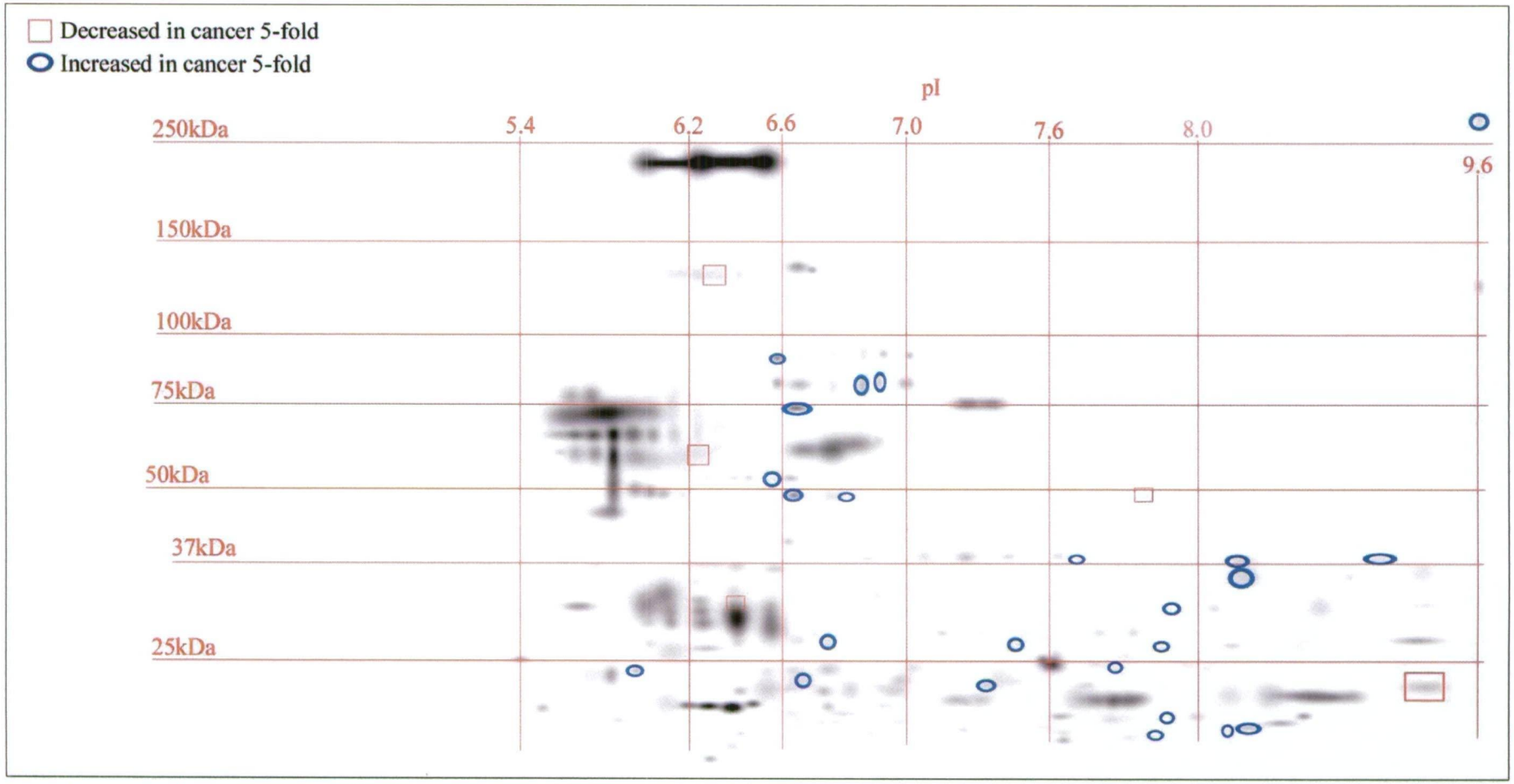


Figure 28: Identified spots from solution 2 extraction showing increased and decreased quantitative expression in breast cancer versus normal breast tissue.

The second extraction step in sample preparation yielded the Matchset CBvNB_Sol2. This revealed a further 24 spots showing a minimum 5-fold increase (including 5 spots showing a 10-fold increase) in expression in cancerous breast tissue versus normal breast tissue. Similarly, 5 spots showed a minimum 5-fold decrease (including 1 spot showing a 10-fold decrease) in expression in cancerous breast versus normal breast tissue.

Although not a major point of focus in this study, two patients provided non-invasive cancer samples and as mentioned earlier, were studied as a distinct replicate group in Matchset CBvNB_Sol3. Although limited by sample numbers, their analysis revealed a quantitative 5-fold increase in expression of 8 spots and a 5-fold decrease in expression of 6 spots in DCIS (ductal carcinoma in situ) versus normal breast tissue. This correlates with data obtained from analysis of invasive cancerous breast tissue with some spots in common between the two sets, particularly the cluster between 25-37kDa and pI 7.8-8.2, suggesting that these proteins were expressed early in the course of tumour biology.

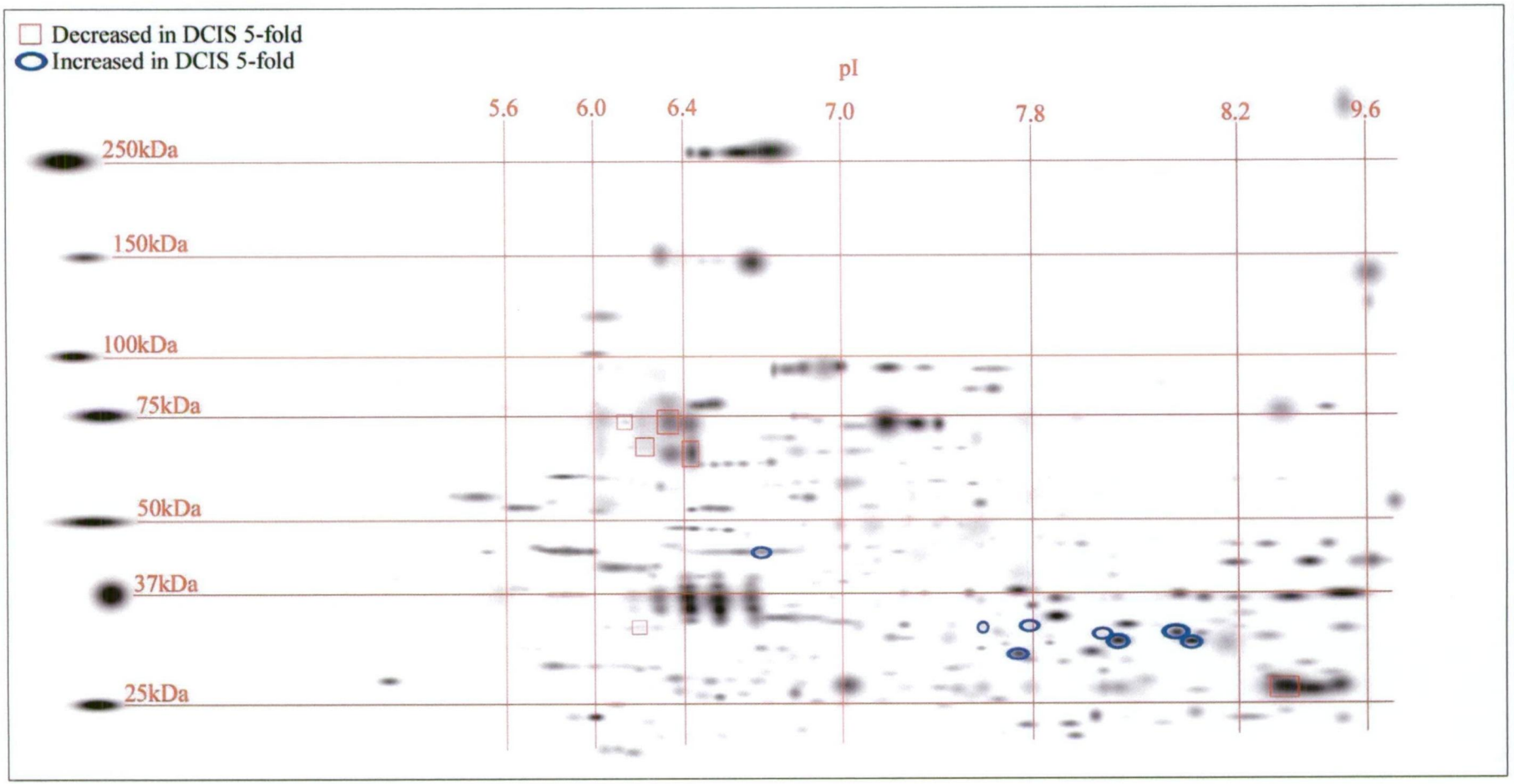


Figure 29: Identified spots from solution 3 extraction showing increased and decreased quantitative expression in DCIS versus normal breast tissue.

3.3 FAMILIAL BREAST CANCER

As mentioned in Table 11 (pg 72), 7 patients were identified in the database who gave a family history of breast cancer, including 4 with a first degree family history. Their analysis was conducted by the creation of Matchset familial_Sol3 containing the replicate groups of familial breast cancer (first/ second degree) and cancer breast as mentioned in the methods section. A quantitative analysis set was then created using a factor of 2-fold differential expression between familial breast cancer and non familial breast cancer. Figure 25 displays 27 spots showing a minimum 2-fold increase in expression in familial breast cancer versus non-familial breast cancer, whereas 54 spots showed a minimum 2-fold decrease in expression in familial breast cancer versus non-familial breast cancer respectively.

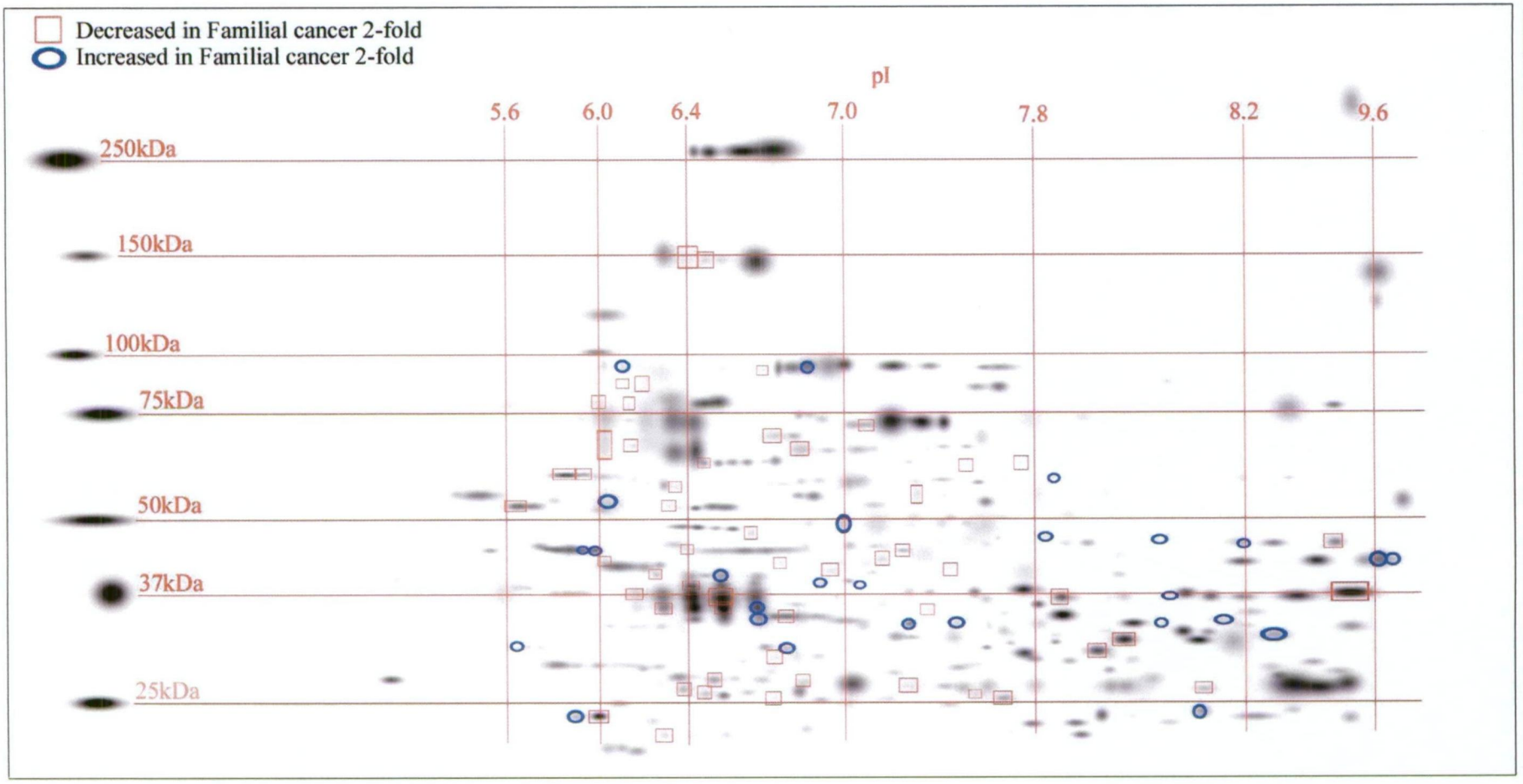


Figure 25: Identified spots from solution 3 extraction showing increased and decreased quantitative expression in familial breast cancer versus non-familial breast cancer.

3.4 HISTOLOGICAL TYPING

Using the Matchset CBvNB_Sol3 as a template, a further Matchset was created (Histosubtype_Sol3) containing the replicate groups ductal cancer and lobular cancer. These were quantitatively analysed using a minimum of 5-fold differential expression versus normal breast tissue. Figure 27 below details the quantitative identification of 26 spots that showed a 5-fold increase in expression in ductal cancer samples versus normal breast tissue and 19 spots that showed a 5-fold increase in expression in lobular cancer samples versus normal breast tissue. By analysing this data using Boolean parameters, it was determined that 15 of the 26 identified spots were unique to ductal cancer, whereas 8 of the 19 identified spots were unique to lobular cancer.

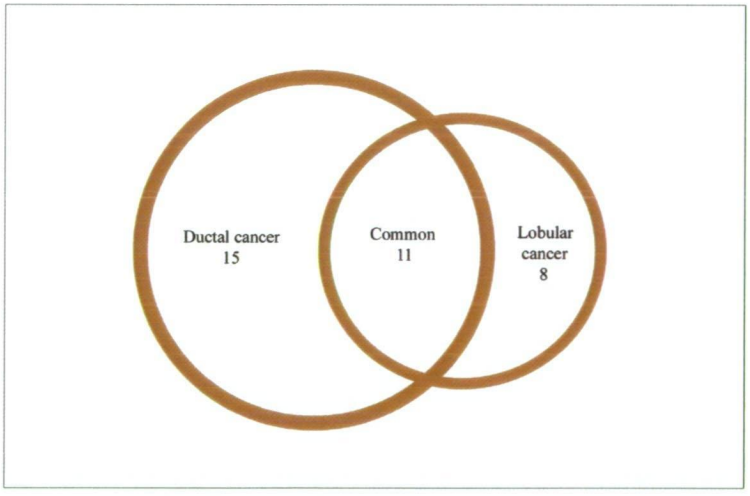


Figure 26: Boolean analysis of quantitative sets Increased in ductal cancer 5-fold and Increased in Lobular cancer 5-fold showing 15 unique spots in ductal cancer and 8 unique spots in lobular cancer respectively.

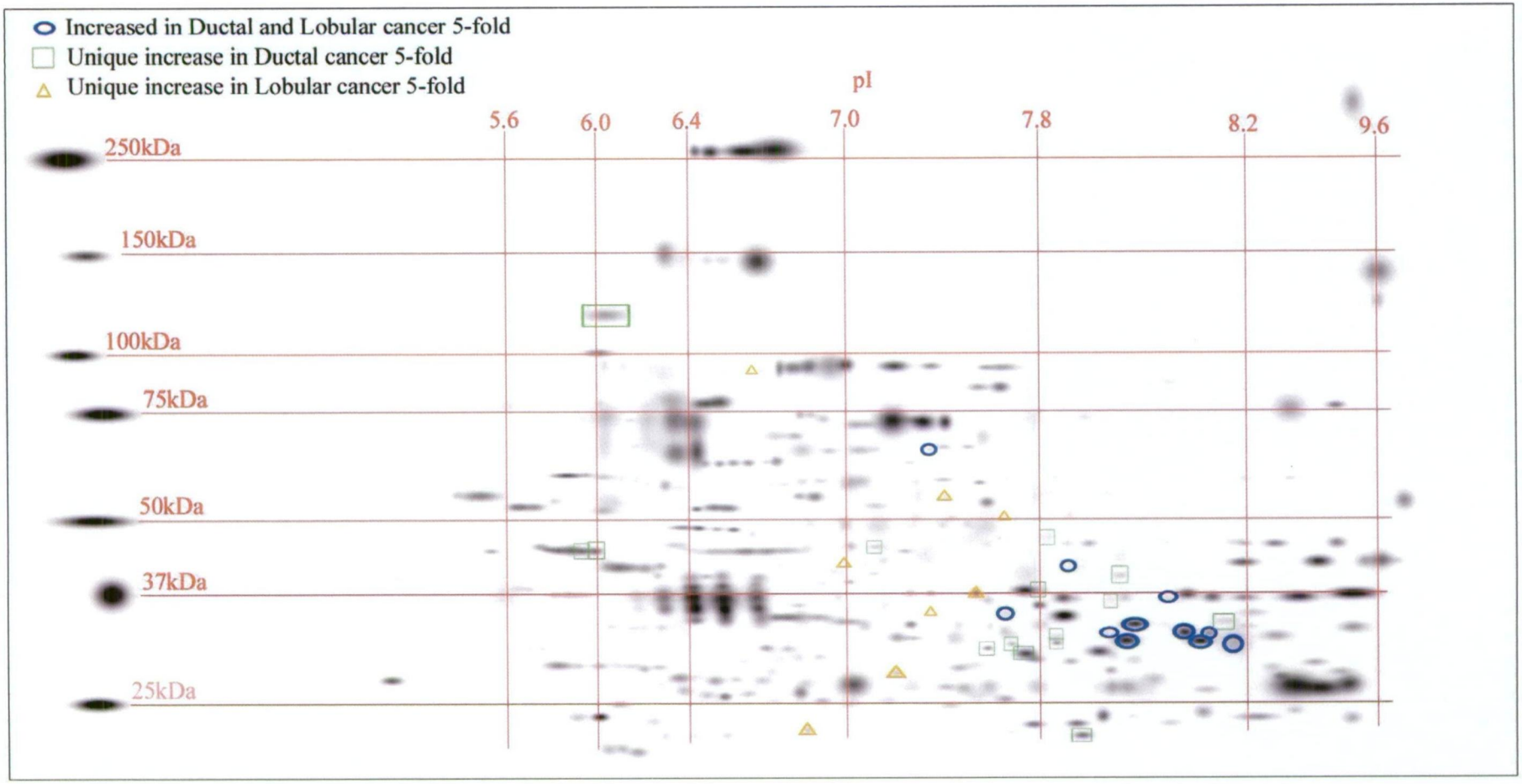


Figure 27: Identified spots from solution 3 extraction showing increased quantitative expression in ductal and lobular breast cancer versus normal breast tissue.

3.5 CLINICAL STAGE

In order to study the differential expression of protein spots in sequence with tumour biological progression, Matchset Clinicalstage_Sol3 was established based on the CBvNB_Sol3 template. A 5-fold quantitative analysis was performed between the various clinical stages of breast cancer (based on the AJC classification) samples obtained by the creation of appropriate replicate groups within the Matchset as mentioned in the methods section. The figure below accounts for a logical stepwise change in the spot expression profile with clinical stage progression. A total of 96 protein spots have shown stage differentiated quantitation.

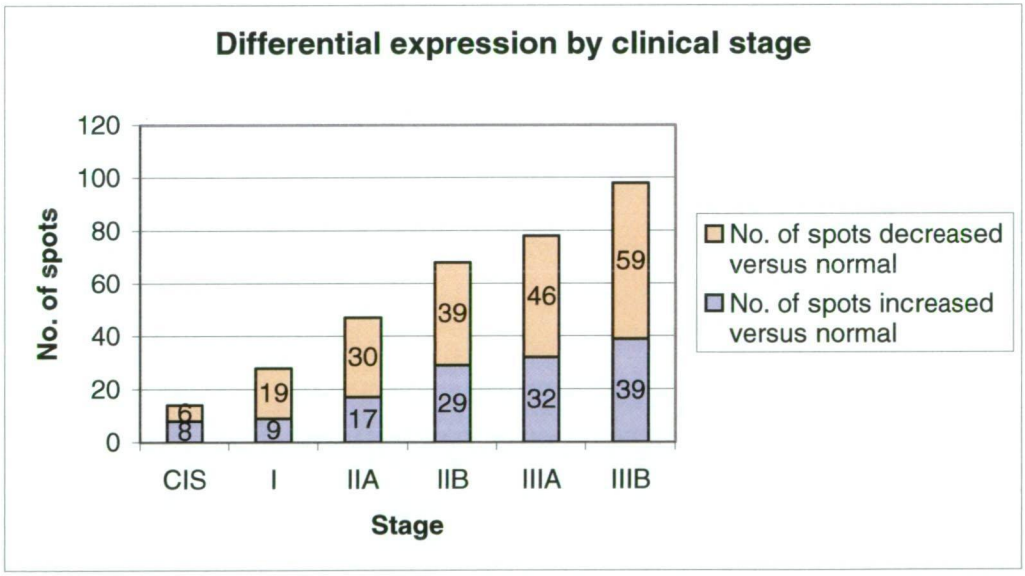


Figure 28: The differential expression of spots by clinical stage as studied in Tasmanian breast cancer group (Matchset Clinical stage_Sol3).

3.6 RACIAL DIFFERENTIATION

A total of 16 breast cancer samples were obtained from Athens and their salient features are mentioned below. No samples of nodal cancer, or normal breast/ nodal tissue were obtained.

Sample ID	Age	Histology	TNM stage	Receptor status			Pre-operative therapy
				Oest	Prog	*c-erbB2*	
1	54	Mucinous	T1N0M0(I)	+++	+++	-	
2	71	Ductal	T1N1M0(IIA)$^{+}$	+++	++	++	
3	61	Ductal	T1N0M0(I)	+++	+++	+	
4	51	Lobular	T1N0M0(I)*	++++	++	+	
5	61	Ductal	T3N2M1(IV)	++	+	+++	
6	64	Ductal	T2N1M0(IIB)	++	-	+++	
7	58	Lobular	T2N2M0(IIIA)	++	+	+	
8	46	Ductal	T1N1M0(IIA)*	+++	+	+	
9	75	Ductal	T2N1M0(IIB)	++	+	+	
10	57	Ductal	T1N1M0(IIA)	+++	++	+++	
11	54	Ductal	T3N2M0(IIIA)	+++	++	++	Chemotherapy
12	81	Ductal	T3N0M0(IIB)	-	-	-	
13	67	Ductal	T2N2M0(IIIA)	+	+	+	
14	80	Ductal	T1N0M0(I)	+++	++	++	
15	65	Ductal	T1N0M0(I)	++	++	-	
16	78	Ductal	T1M0(I)	-	++	-	Tamoxifen for 6yrs, Re-operation for local recurrence

Table 13: Tumour characteristics of Athens group of patients with dominant histology, clinical stage (TNM and AJC), receptor status and relevant pre-operative treatment. *denotes first degree family history of breast cancer, $^{+}$denotes second degree family history of breast cancer.

Matchset ACBvGCB_Sol3 contains differential spot data between gels processed from Tasmanian and Athens group breast cancer samples. A 2-fold quantitative analysis was conducted of these replicate groups to reveal the increased expression of 43 spots and decreased expression of 37 spots in Tasmanian breast cancer. Further analysis revealed 5 of 43 spots showing a minimum 5-fold increase and 3 of 37 spots

showing a minimum 5-fold decrease in expression in Tasmanian breast cancer, respectively.

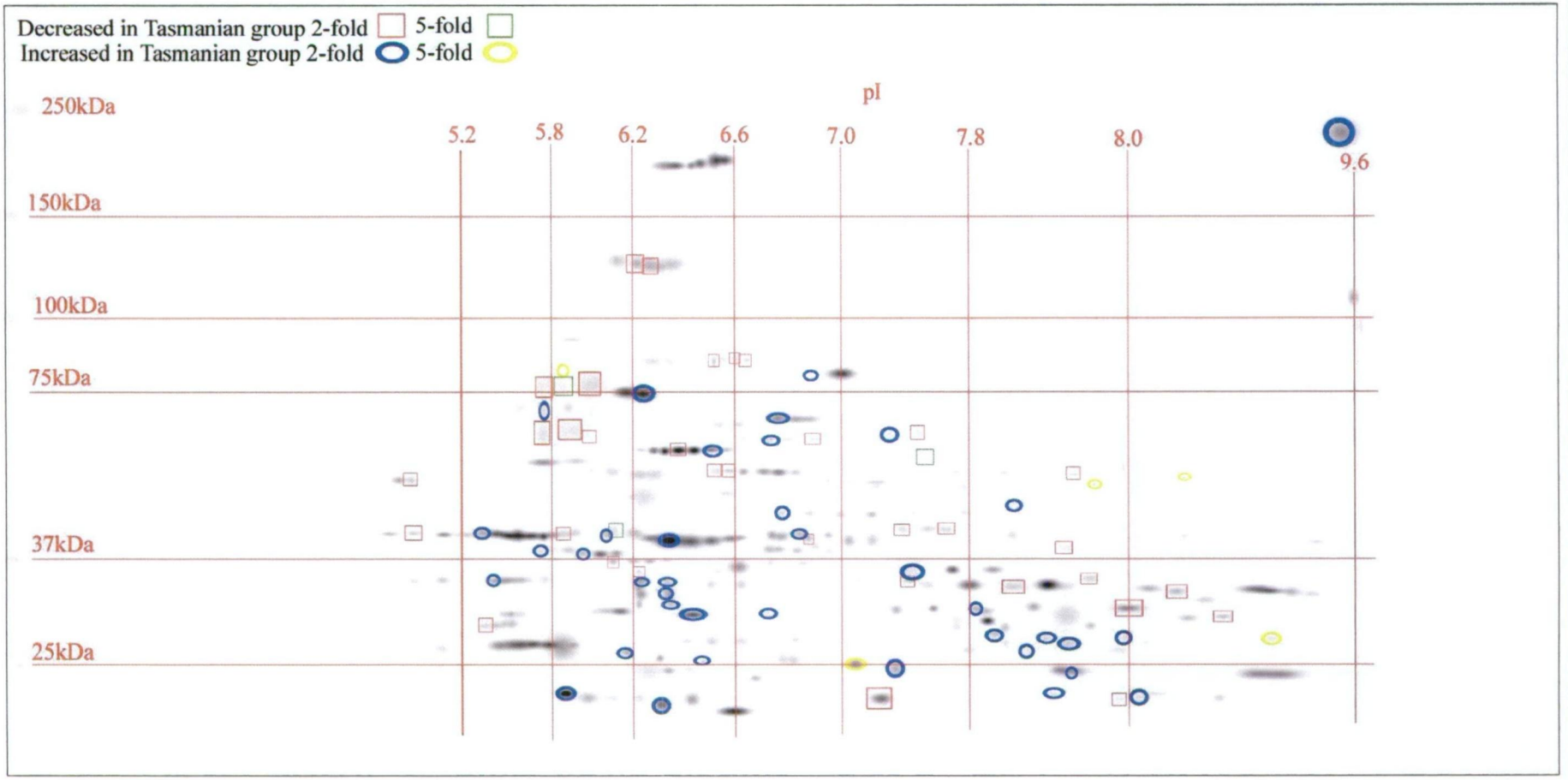


Figure 29: Identified spots from solution 3 extraction showing increased and decreased quantitative expression in Tasmanian group versus Athens group.

Extraction 2 contributed towards the production of gel Gaussian images contained in Matchset ACBvGCB_Sol2. The analysis of this reveals a quantitative 2-fold increase in the expression of 29 spots, of which 3 displayed a minimum 5-fold increase in Tasmanian breast cancer samples versus Athens group breast cancer samples. 23 spots were found to display a minimum 2-fold decrease in expression in Tasmanian breast cancer samples.

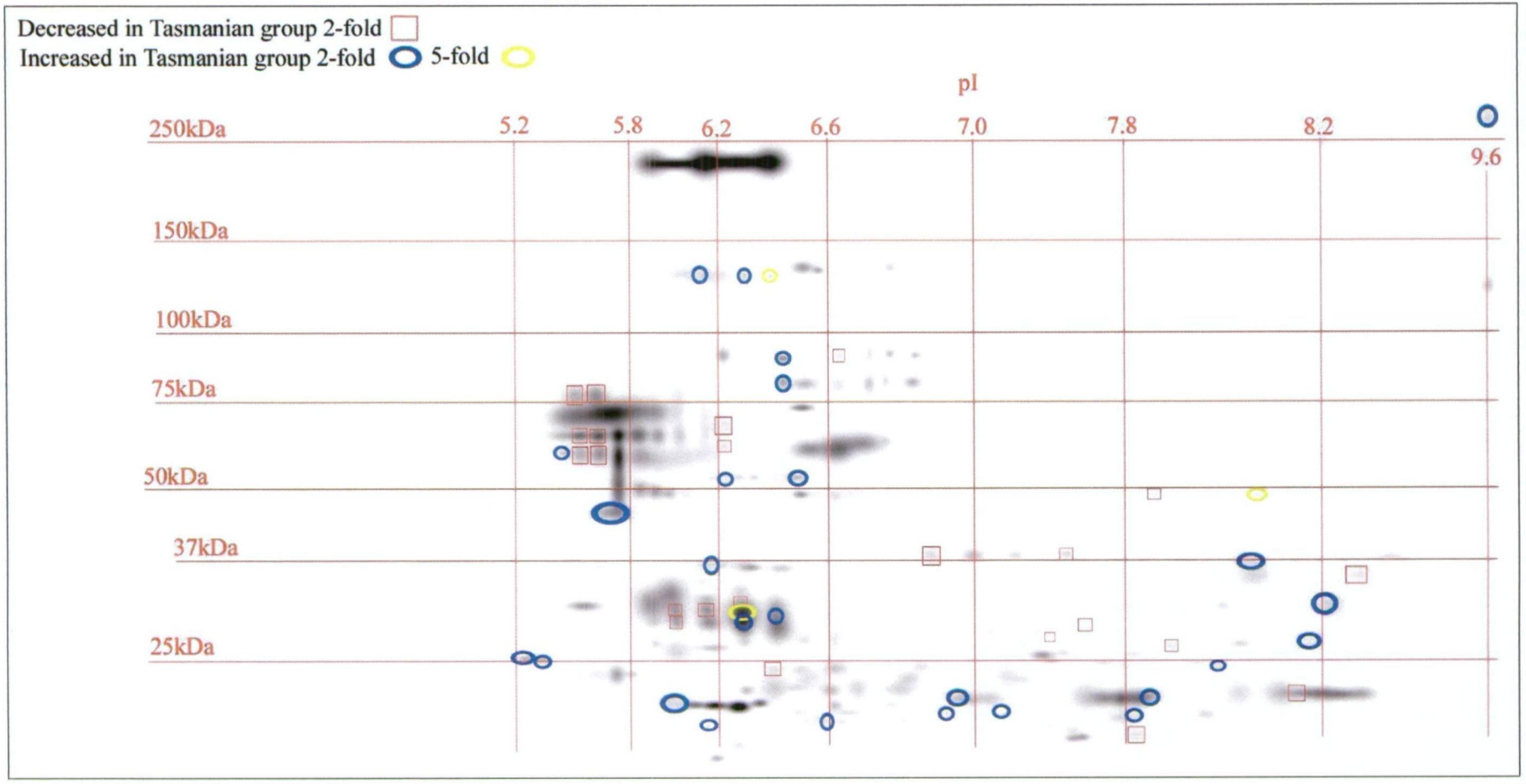


Figure 30: Identified spots from solution 2 extraction showing increased and decreased quantitative expression in Tasmanian group versus Athens group.

3.7 METASTATIC PHENOTYPE

Node positive disease enabled the study of the metastatic potential of breast cancer. A total of five patients only, provided nodal cancer tissue samples, whereas another nine patients provided normal node tissue (Table 10, pg 71). Matchset CBvCN_Sol3 was created containing the replicate groups of Tasmanian breast cancer, nodal cancer and normal nodes as described in the methods section. A quantitative analysis set was then created looking at all the spots that were conserved between Tasmanian breast and nodal cancer within a 2-fold sensitivity. These amounted to 95 spots in total. These spots could not be present to any significant degree in normal nodes and hence, further quantitative analysis sets were created to reveal a minimum 2-fold increase in expression of 31 spots, of which 11 spots displayed a minimum 5-fold increase in expression in nodal cancer versus normal nodes. By Boolean analysis of these sets, 11 spots were found to be conserved in nodal cancer and showed a minimum 2-fold increase when compared to normal nodes. Of these spots, 2 were found to show a 5-fold increase in expression in nodal cancer.

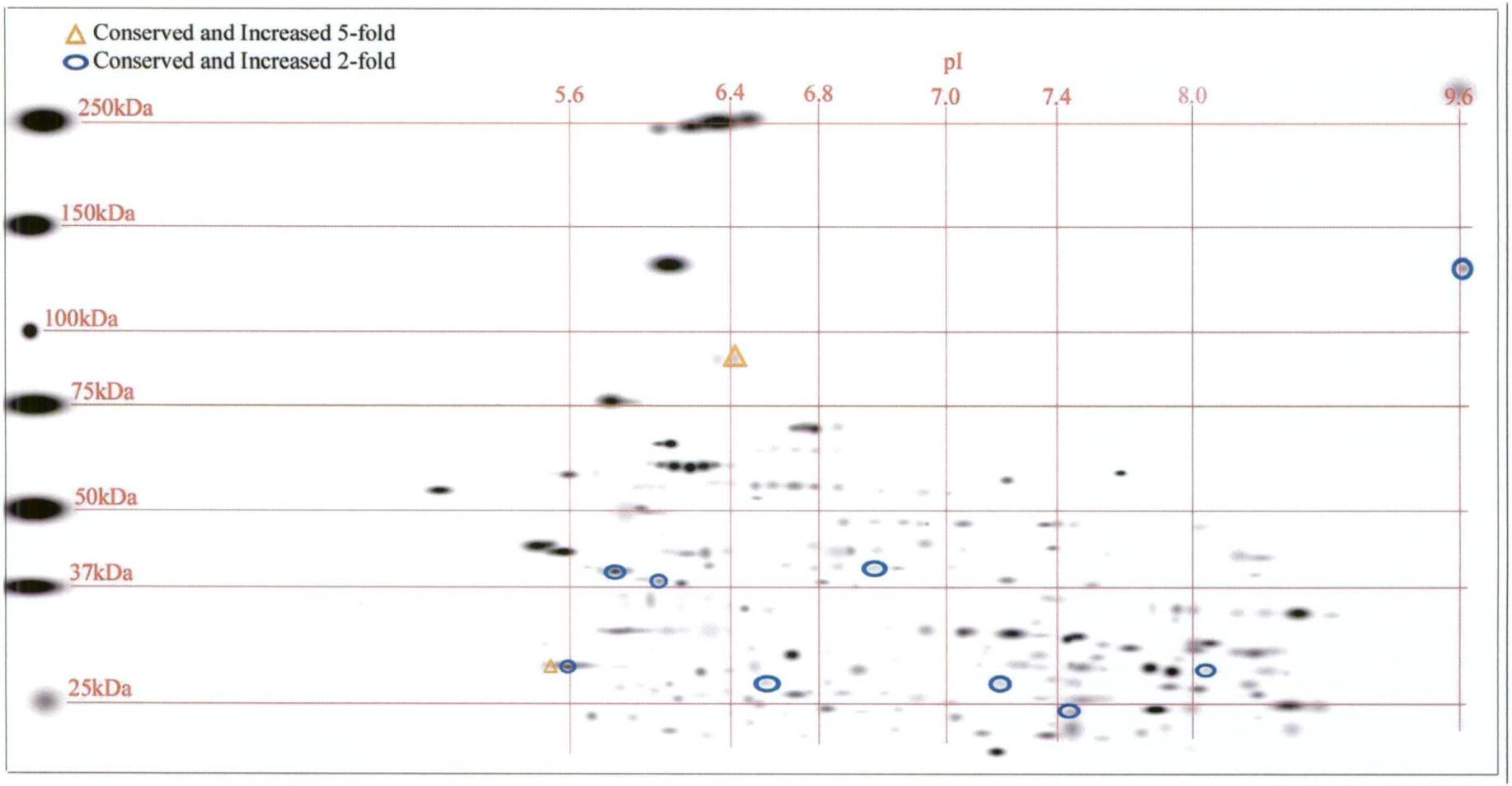


Figure 31: Identified spots from solution 3 extraction showing conservation between breast cancer and nodal cancer. Further analysis of these spots revealed a set showing increased quantitative expression versus normal nodal tissue.

The quantitative and Boolean analysis of the corresponding Solution 2 Matchset CBvCN_Sol2 revealed 40 spots that were found to be conserved between Tasmanian breast and nodal cancer within a 2-fold factor. Of these, 8 spots were found to show a 2-fold increase in expression when compared to normal nodes, with 4 of these showing a 5-fold increase.

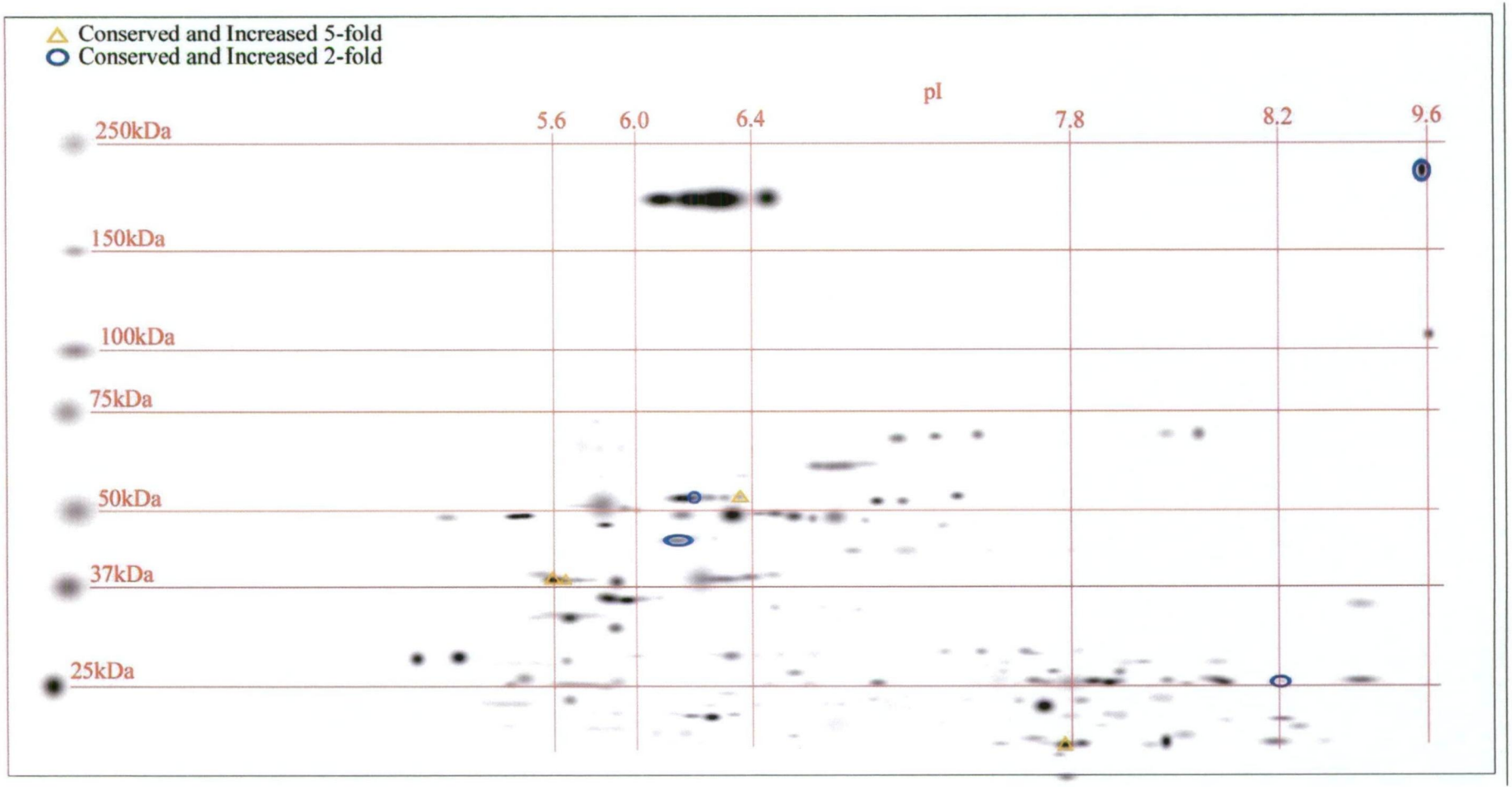


Figure 32: Identified spots from solution 2 extraction showing conservation between breast cancer and nodal cancer. Further analysis of these spots revealed a set showing increased quantitative expression versus normal nodal tissue.

4.0 DISCUSSION

4.1 STUDY DESIGN

With the burgeoning rise of proteomic technology in the last decade, it became clear that this science could be applied to the study of a large variety of disease conditions including cancer. As such, breast cancer as a disease state had not been extensively studied using this technology with reports in the literature being relatively few in number [51, 73]. Secondly, it was felt that the analysis of intact tumour tissue would be more representative of the disease state in the community as alluded to by Giometti *et al*, as opposed to the previously studied cultured breast epithelial cell lines [52].

The incidence of breast cancer in Tasmania as detailed in the literature review was consistent with the national average. Background knowledge into the aetiological significance of breast cancer was lacking, with little objective clinical evidence to corroborate. However, the purpose of the study was to establish a baseline proteomic profile of the disease that could then be analysed to find potential diagnostic markers and therapeutic targets not unlike the history of the development in *OR* and *HER-2* receptor modulation. In order to do this, a larger patient population translating to more tissue samples would have been ideal but not entirely necessary and we were able to obtain, what was thought to be, a sufficient number. Further, in order to expand our sphere of study, we chose firstly, to include patients undergoing axillary dissection and obtain nodal tissue for comparison. This would attempt to identify a metastatic phenotype of breast cancer. However, the proportion of breast cancer patients providing for malignant nodal samples was under-represented. This was not

unexpected and is likely to be an ongoing concern in any experiment of this nature. The reasons for this are multifactorial. Firstly, with an earlier stage at diagnosis, fewer patients are likely to present with malignant axillary nodal disease translating to a trend towards breast conservation surgery and nodal sampling techniques. Secondly, it was found that this also meant that with the smaller number of nodes provided for pathological examination, it was less likely that a node would be obtained for our experiments.

In order to consolidate our existing patient database and to highlight possible differences in expression between racially diverse population groups, we sought assistance from overseas colleagues. The opportunity thus arose to obtain breast cancer samples from the Prolipsis clinic in Athens, Greece. The lack of a control group from within this population was seen as a limiting factor but the logistical difficulties in solving this were not deemed practicable and any significant findings obtained would need to take this into account. We had earlier approached colleagues from Japan, but due to a change in legislature in that country and an effective ban on the export of human tissue for experimental purposes, this avenue did not materialize. We then sought other colleagues from Taiwan, from whom we were successful in obtaining tissue. However, due to time constraints, this group has been excluded from this thesis. The preparation and analysis of these samples is ongoing and with time this shall aid in the production of a more comprehensive proteome map of the disease. To the author's knowledge, the study of breast cancer from epidemiologically distinct population groups has not been investigated previously.

4.2 SAMPLE PREPARATION

It was realized that the most convenient method of sample preparation, given the time constraints of the study, was to use previously established protein separation protocols. The Bio-Rad corporation has conveniently formulated the 2D starter kit (catalogue no. 163-2105) including a protein sample (*E.coli* whole cell lysate) and sufficient reagents to re-hydrate, focus and transfer to the second dimension gel, ten 11 cm ReadyPrep IPG strips [67]. Although the use of this kit was not included in our protocol, it did however serve as a training tool and provided for a backup source of reagents. However, the fact that reagents were individually prepared in the laboratory in batches does also raise the possibility of a theoretical difference in chemical concentrations amongst reagents.

It has been well known that reproducibility of sample preparation is the key to a successful proteomics experiment [74]. However, as mentioned previously, there has been little work in the past on intact breast tissue. As stated by Bio-Rad, most protein mixtures will require some experimentation to determine optimum conditions for 2D-PAGE [67]. The 2D starter kit and ReadyPrep sequential extraction kits (see methods section) provided readymade reagent solutions for training. For this purpose, tissue from our already established breast tissue bank was not suitable. As an alternative to the *E.coli* cell lysate provided with the starter kit, we chose to use an animal protein source and the most suitable, was commercially available chicken breast without the exclusion of subcutaneous fat. Samples were prepared according to the protocols established by Bio-Rad and gels produced by their two dimensional electrophoresis showed adequate spot identification, although a large section of the gel was occupied

by myosin chains. Analysis of these images was not conducted. The next step was to determine how normal breast tissue would behave under these circumstances. Homogenization of sample, which was quite easy with chicken breast proved to be a challenge with intact normal breast tissue. The UltraTurrax (see methods section) electric homogeniser was not suitable for, in particular, normal breast tissue that would continually entwine itself in the instruments blade apparatus. The only suitable alternative was to transfer the contents to a Pyrex manual homogeniser (see methods section), which although laborious, provided for a satisfactory homogenisation. This process needed to be interrupted momentarily every 2-3 minutes because generated heat in the homogeniser necessitated placement of sample in the freezer for cooling. Malignant breast tissue on the other hand, when subjected to homogenisation behaved unexpectedly. This tissue would fragment quite easily in the manual homogeniser and hence this was the preparation method of choice for all obtained cancer samples. Admittedly, it would not be entirely implausible for this variation in preparation technique to have accounted for some of the differences in gel pattern between normal and cancerous breast tissue.

Sequential extraction is a means to sample solubilization and reducing sample complexity. As stated by Molloy MP *et al*, increasingly powerful solubilizing solutions used in sequence have been shown to increase the total number of spots detected [54]. These solutions extract an overlapping set of proteins. Reagent 1 (containing 40mM Tris) extracts soluble proteins such as cytosolic proteins, whereas Reagents 2 and 3 (containing differing concentrations of chaotropic agents, detergents, carrier ampholytes and reducing agents) extract proteins of intermediate and low solubility. Urea and Thiourea are commonly used chaotropic agents that

disrupt hydrogen bonds and prevent the formation of protein aggregates. Detergents such as CHAPS disrupt hydrophobic interactions and increase the solubility of proteins at their respective pI [67]. Newer detergents have emerged such as SB 3-10 (see methods section) and ASB-14 that have helped isolate previously undetected membrane proteins and have found increasing use in proteomics experiments [75]. Each of the three reagents extract an overlapping set of proteins and greatly increase the number of protein spots identified, however as mentioned in the methods section, Solution 1 was not processed beyond the preparation stage in order to exclude potentially interfering abundant cytosolic proteins .

The role of carrier ampholytes lies in the supplementation of salt concentration in a sample solution. Certain proteins require salt for adequate solubility but the concentration of these salts is usually restricted to 40mM as they limit the applied voltage and hence increase the time required for focusing. Their concentration is therefore limited to <0.2% w/v and in our experiment, were used as a mixture over a large pH range (3-10). We experimented with the idea of using a mixture of carrier ampholytes and these would have been over a pH range of 5-8 and 8-10 in a ratio of 2:1. However, to avoid repeating all the experiments in our limited given time frame and the delay in transport of ampholyte solutions from the Bio-Rad Corporation in the U.S.A., this was not able to be performed.

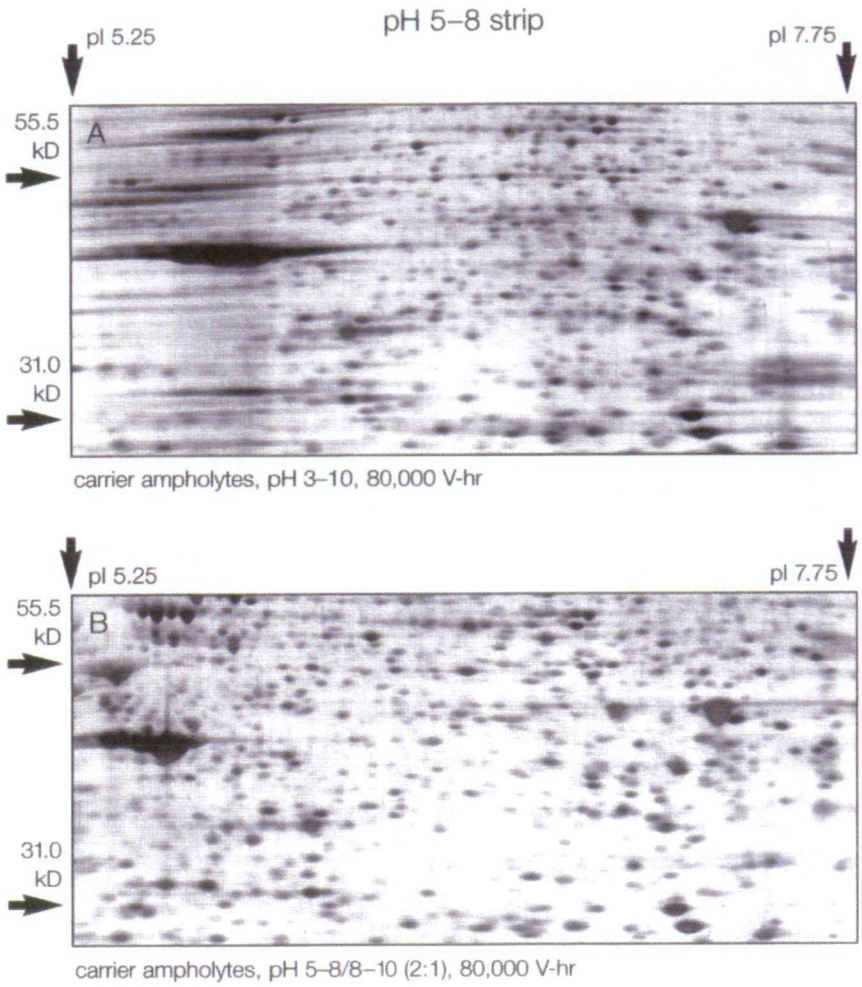


Figure 33: Matching sections of 2D images showing separation of 110 micrograms of a cytosolic extract of human lymphoblastoid cell line loaded on a 17 cm pH 5-8 ReadyStrip IPG strip and second dimension separation on a 10-24% gradient gel. Image B shows greatly improved focusing with the use of pH 5-8/8-10 ampholyte mixture in a ratio of 2:1 [67].

The use of reducing agents such as DTT (Dithiothreitol, used in IPG strip equilibration, see methods section) and TBP (Tributylphosphine, used in Reagents 2 and 3, see methods section) reduce disulfide bonds and increase spot resolution [62]. We found that increasing the concentration of TBP to 20microlitres or a 1:50 dilution from the prescribed 1:100, significantly increased spot resolution and that increasing it further to a 1:25 dilution did not aid greatly, but substantially increased the utilization and cost of the volatile reagent [67].

The new and more powerful detergents, the use of a variable mixture of carrier ampholytes and reducing agents shall aid greatly in the production of more clearly resolved gel images.

The presence of nucleic acids, particularly DNA, was found to interfere with adequate sample preparation. The DNA complexes would markedly increase the viscosity of the sample solution and hinder protein entry and migration on the IPG strip. As mentioned earlier, this was particularly the case with lymph node samples, with protein spots at times showing poor separation on 2D gels. The use of endonuclease helped overcome this problem.

4.3 TWO DIMENSIONAL PAGE

Resolution in the first dimension run is dependent on the pH gradient, strip length and the applied electrical field. As stated by Garfin *et al*, the difference in pI (Iso-electric point) between two adjacent IEF resolved protein bands is directly proportional to the square root of the pH gradient and inversely proportional to the square root of the voltage gradient at the position of the bands [76]. In simpler terms this theory states that, band resolution will be maximal in narrow pH ranges when high voltages are applied. One of the drawbacks of this statement is that, with high applied voltages, heat is generated. Hence, for the sake of not compromising resolution, thin IPG gel strips are preferred as they dissipate heat more readily [67]. Thin IPG strips have been the benchmark tool for the first dimension focus during the last 15 years.

The advent of 'zoom-in' narrow range (pH 4-7, 5-8) and very narrow range (pH 4.5-5.5, 5-6) IPG strips have added greatly to the resolution of a first dimension run. Of late, it has become increasingly clear that, additional resolving power is needed to separate complex eukaryotic proteomes [65, 77]. Our laboratory has recently acquired narrow range IPG strips (pH 5-8) and these shall help increase spot detection.

A pre-requisite to the second dimension run is the equilibration of IPG strips and this ensures that proteins are coated with Dodecyl Sulfate and Cysteines are reduced and alkylated. We found it more convenient to use 2.5% DTT in our protocol (see discussion of sample preparation) both for ease of base buffer preparation and increased resolution compared to the prescribed 2% concentration [67]. The loading of IPG strips onto its well in the Criterion 2D gel cassette was also complicated by the delayed binding of low melting point Agarose overlay solution when used at the prescribed concentration of 0.5%. We found that a 1% solution avoided migration of the IPG strip and marker proteins as well as the unnecessary waiting period while the Agarose set.

Our decision to use gradient gels for the second dimension run was based on the need to analyse proteins over a large molecular weight range and perform a global comparison of gels. The next logical step in our experiments would have been to perform the second dimension run on narrower gradient or single percentage precast Criterion gels e.g. 12.5% resolving, 4% stacking (total percentage of Acrylamide plus cross-linker) (Bio-Rad, catalogue no. 345-0018) that would resolve proteins within a mass range of 20-100kDa. This would help the further resolution and detection of spots within the molecular weight range of virtually all our currently detected spots.

Another feature that would aid high-throughput experiments would have been the use of Criterion Dodeca cells (Bio-Rad, catalogue no. 165-4130) capable of accommodating 12 gels per run, although for our purposes this would have equated with a substantial increase in cost and a more tedious experiment given limitations in other associated equipment and manpower, and hence not seriously considered.

Gel staining techniques also needed to be adapted to suit laboratory conditions. The Criterion precast gel cassettes are packaged in appropriately sized plastic containers and these proved to be ideal staining vessels, the only limiting factor being the size of the laboratory platform mixer. This restricted the number of gels that could be stained at any one time to six. The protocol we followed for Silver Stain Plus was largely, as described by Bio-Rad. This was a modification of the standard silver stain that was developed in the late 70's by Merrill *et al*, and various commercially available kits are available, each with their own protocols [78, 79]. Essentially, the use of Silver Stain Plus enables the mass spectrometric analysis of excised spots that standard stain does not. This is due to an additional oxidative step in standard silver staining technique that changes protein mass. There is a drawback however, in that the visualization of heavily glycosylated proteins and lipoproteins can be less sensitive with Silver Stain Plus. Staining results were variable, depending on the quality of sample prepared. The quality of gel images obtained was also dependant on the period of time incubated in stop solution (5% Acetic Acid) and the immediate placement of gels in the stop solution would highlight spots and potentially give rise to over-stained gels. Overnight incubation of stained gels in the refrigerator was also a practice we learnt to avoid, as this caused them to enlarge and distort, needing an extended wash with milliQ (see storage step, methods section).

The sensitivity of Silver Stain Plus has been said to be 30 to 50-fold times that of Coomassie Brilliant Blue R-250, the standard stain for protein detection. The fluorescent SYPRO Ruby and Orange stains, and radio-isotope labelling have features that are desirable in high throughput laboratories. They are sensitive to 1-10nanograms of protein but require costly image acquisition tools. They do however, allow detection of glycoproteins, lipoproteins, low molecular weight proteins and metalloproteins that are not well detected by other stains [67]. This is also reflected by the fact that these stains have shown a greater number of protein database matches and percent sequence coverage, as stated by Lauber WM *et al* [69].

4.4 GEL ANALYSIS

The acquisition of gel images was conducted by densitometry which measures the intensity of a light beam before and after being attenuated by a protein spot. Wet gel images could also be acquired by the water-proof scan platen. However, for reasons of practicality, we chose to store prepared gels in appropriately sized zip lock clear plastic bags containing 2 ml of 5 % Sodium Azide as a preservative. This made gels extremely portable and could be handled without a relative fear of breakage. Images could then be scanned directly with the gels contained in their zip lock bags and their acquisition was not compromised using this method.

In order to compensate for sample variation, errors in the preparation, loading and running of two dimensional gels, it was useful to treat the gels as being similar [67], particularly as the biological question in mind was the essential difference between malignant and normal breast tissue. This similar treatment was referred to as replicate

group quantitation and was necessary to avoid large numbers of analysis sets between Matchset member Gaussian images. Spot quantitation was thus reflected as an average within a particular group and made for a more accurate quantitation. The disadvantage of using this approach was that, as mentioned above, spots that were qualitatively expressed in one group versus another could not be determined as the number of analysis sets required would be too great. A further disadvantage of replicate group analysis in the currently used version of PDQuest (v6.2.0) was that it failed to recognize a member gel image to be present in more than one replicate group, necessitating the creation of multiple copies of a Matchset template.

We found PDQuest to be a program that required considerable tuition. An enclosed self help tutorial CD-ROM was extremely useful. Following a lengthy period of familiarization and hands on experience, scan and crop settings, spot detection and matching could be performed with relative ease. There are more recent programs such as PDQuest version 7.0 (Upgrade catalogue no. 170-8626), Melanie IV and ImageMaster 2D Platinum 5.0 that are now available. The latest version of PDQuest incorporates the newly developed SPoT (Streamlined PrOteomic Tools) technology that entails an auto-matching algorithm, significantly reducing the need for manual intervention during matching. In addition, exported TIFF image files can include display of analysis overlays and annotations as they appear on screen. These and other features including a match summary table and integration with Worksbase™ (Bio-Rad) bioinformatics database shall greatly facilitate spot detection, analysis and identification.

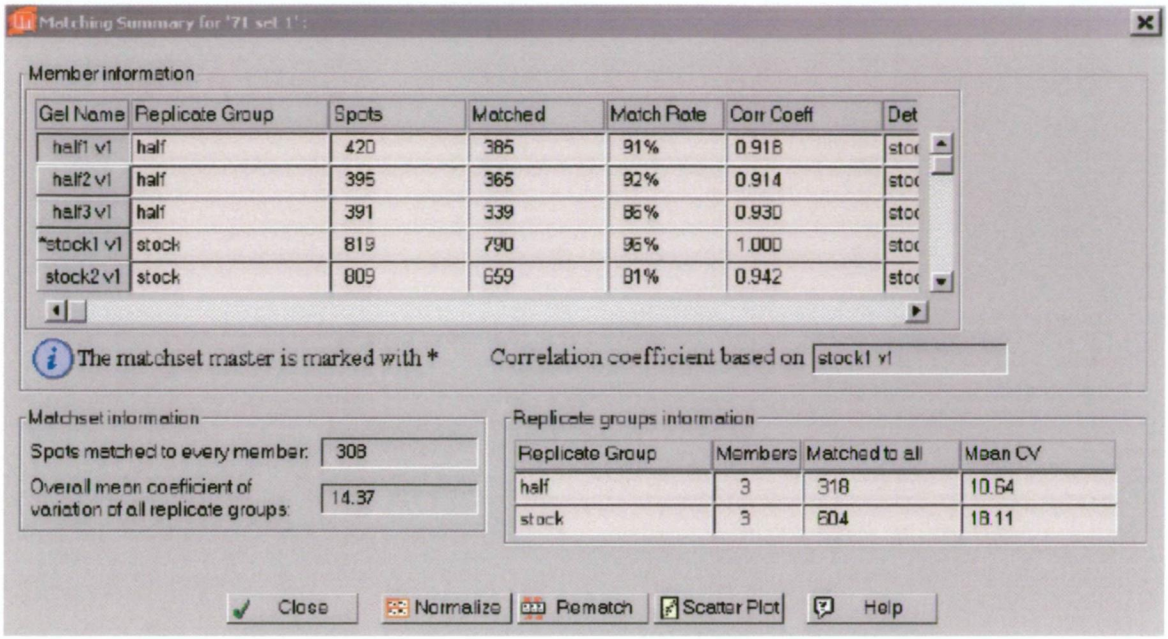


Figure 34: PDQuest version 7.0 Match Summary table that details number of spots detected, percentage match and correlation co-efficient between Matchset member gels.

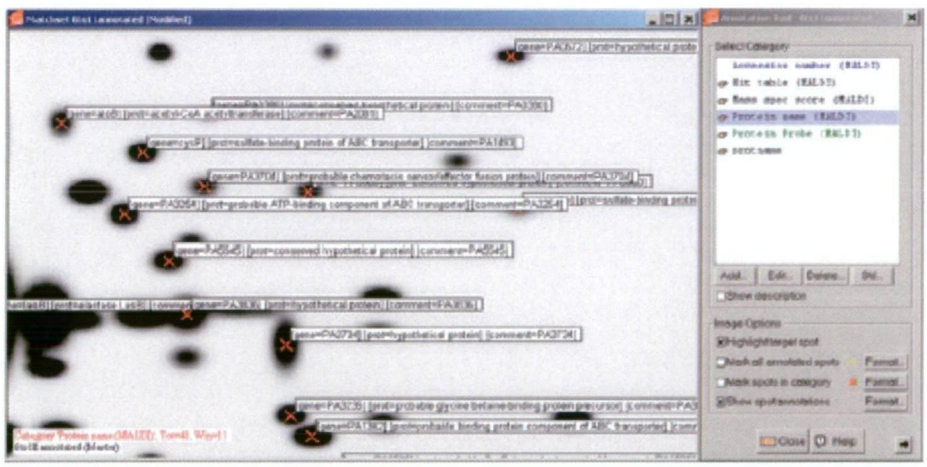


Figure 35: PDQuest version 7.0.1 integrated with Worksbase™ software allows direct import of mass spectrometric data from Micromass and MASCOT (Matrix Sciences).

4.5 RESULTS

Standard maps of breast cancer including protein spots from Australian and Greek breast cancer samples, normal breast tissue, cancerous and normal node tissue, and DCIS were created. A composite map was created, containing a total of 403 protein spots when looking at Australian breast cancer and normal breast tissue. It could be argued that sample size may have been limited particularly in the case of nodal tissue and DCIS. However, the purpose of the study was to primarily isolate the differential expression of protein spots in Australian breast cancer samples compared to normal breast tissue and for this reason emphasis was placed on such.

A recognizable pattern of spot expression could be seen in gels between cancerous and normal breast tissue and particular clusters of spots could be seen to be associated with a more malignant or normal phenotype. A total of 74 spots showed a minimum 5-fold change in expression between Australian breast cancer and normal breast tissue samples. Of these, 54 spots have shown a minimum 5-fold increase in expression in breast cancer versus normal breast tissue. Although this number is only likely to drop with the exclusion of some spots, this is significant in that other authors such as Bini *et al* have reported finding only 32 spots that showed a differential increase in expression in breast cancer versus normal breast tissue [80]. Further, Bini *et al* performed their separation using 18cm Immobiline strips of pH3-10 and 9-16% SDS gradient 2D gels [80]. Amongst those 54 spots that we have found however, is a cluster of spots, 8 in number between 25-37kDa and pI 7.8-8.2 that were found to be up-regulated early in the course of the disease, given that they showed a quantitative

increase in expression in gels obtained from DCIS samples. These would be a good starting point for further characterization.

The spot expression profile of familial breast cancer samples may herald the identification of a marker for these tumours. The sample size is however, limited with only 7 patients in the group displaying a family history of the disease, of which only 4 displayed a first degree history. Thus, further samples may be required along with a more sensitive analysis (2-fold change at present). It is also not known if any of these familial breast cancer samples were linked to any of the known genetic breast cancer syndromes as none have been tested specifically for such. Nonetheless, a search for their translated proteins (including *BRCA-1, BRCA-2, p53, Rb, bax, bcl-2*) shall be conducted and a differential expression sought.

Similarly, histological sub-typing is associated with prognosis and may have a bearing on overall treatment and outcome. As mentioned in the results section, 15 spots were found that were unique to ductal cancer and 8 spots that were unique to lobular cancer. This has not been previously reported in the literature and further characterization may herald the identification of an early marker for histological differentiation, although the value of this compared to existing parameters is debatable.

The primary reason for studying these tumours from a view point of clinical stage was to attempt to identify an early marker for local invasion and nodal metastases. In addition, as Figure 28 in the results section shows, with tumour biological progression there are 96 spots that we have found, to have shown a 5-fold change in expression.

These spots may represent loss of tumour suppressor factors or conversely, of amplification of tumour promoting factors. In which case, tumours would be able to be further sub-classified based on the presence or absence of these factors and their likelihood of progression to the next clinical stage. An appropriate starting point for the investigation of these would be amongst the spots showing a fold change in expression between normal, DCIS and early breast cancer. There are in addition, other tumour parameters that could be studied and include histological grade that, akin to clinical stage differentiation would seek to identify proteins that are differentially expressed with biological progression. This has not been looked at previously in the reported literature but had to be excluded due to time constraints.

The analysis of breast cancer from racially diverse origins has shown a minimum 2-fold differential expression of 132 spots using Solution 2 and 3 extractions (including 11 showing a minimum 5-fold change) between the Tasmanian and Athens groups. The further characterization of these 11 spots may reveal the results of genetic, environmental or hormonal influences that determine the differences in incidence seen between the two groups. As per the authors' knowledge, this has not been studied previously. Ideally, as mentioned earlier, this ought to have included a cohort of normal breast tissue samples from Athens but the logistical difficulties of arranging this, persuaded us from including this group. The future addition of breast cancer samples from other regions such as Taiwan will serve to complete the full spectrum of incidence.

As mentioned earlier, we found 19 protein spots that showed a minimum 2-fold increase in expression in cancerous nodal tissue compared to normal nodal tissue, that

were conserved between cancerous breast and nodal tissue, including 6 spots that showed a 5-fold change. The significance of this lies in the fact that this approach has not been previously reported in the literature. The number of studies looking at the metastatic potential of cancer in general has been limited. To this end, Schwalke MA *et al* and Osada T *et al* have used a 2D PAGE approach to study pancreatic and hepatocellular carcinoma respectively, but breast cancer has not been previously evaluated [81, 82]. Our attempt was to try and identify protein spots that would be differentially expressed in a tumour that was seen as being, more capable of axillary metastases. Firstly, this would have benefits in terms of, identification of tumours with a greater metastatic potential and thereby identify patients who would benefit from a more radical axillary clearance or chemotherapy. Secondly, as mentioned earlier, this may allow the identification of a marker of axillary metastases and its study as a potential target for the development of an adjunctive therapeutic agent. However, there are limitations to the study of nodal disease as also discussed previously.

4.6 FUTURE IMPLICATIONS

Our experience with proteomic sciences has led us to believe that the refinement of spot set data is an ongoing exercise, taking into account the inclusion of further breast cancer samples from different sources and with varied phenotypic expression. To this effect, it would have been ideal to procure further cancer samples from other population groups at the lower end of the spectrum of incidence such as Taiwan to add to our current dataset. Although also not feasible for the purpose of this thesis, it would have been ideal to have obtained more familial breast cancer and nodal tissue

samples given the small numbers obtained for the purpose of this study. It would also be fair to say that the analysis of our current gel images using the later versions of gel analysis software would yield more accurate match and analysis set data.

In order to find new potential markers and/or therapeutic targets it was deemed vital to exclude already known markers of the disease. Although details of this are not included in this thesis, we have identified five known prognostic factors and include *OR-α, β-casein, cytokeratin 7, calponin* and *bax* using a 2DPAGE and Western blot approach. It has been interesting to note in this regard that *ORα*, the most commonly studied prognostic factor of breast cancer, was found to be present in two isoforms. A 63kDa isoform that is detected by currently used immunohistochemical means, was up-regulated in cancerous breast tissue samples. The second 80kDa isoform was qualitatively expressed in cancerous breast tissue samples and not in normal breast tissue samples. This finding presents significant advantages in using a proteomics approach rather than conventional techniques [83].

Keratin		Bax		β-Casein		p130cas		Calponin		ER	
M_r	pI	M_r	pI	M_r	pI	M_r	pI	M_r	pI	M_r	pI
70	6.7	16	–	22	–	140	7.9	50	6.9	63	7.3
56	6.6	16	7.9	22	–	140	–	50	7.0	80	7.9
56	6.7	16	–	22	7.1	280	–	35	–		
56	6.8			22	7.6						
56	6.65			22	–						
56	6.75										
56	6.85										

Figure 36: Silver Stain Plus and overlay Western blot obtained molecular weight and pI data for selected prognostic factors. ER=estrogen receptor [83].

Using this approach it is hoped that other known markers of the disease shall also be identified and excluded.

The refinement of our current data shall be an ongoing process. Protein spots, as they are excluded by comparison, identified by mass spectrometry and/or sequenced, will be determined to be, more or less significant to the pathogenesis of breast cancer from our knowledge of the disease today. With the advancement of existing software available for proteomic analysis, this process is likely to become more precise. Enhanced spot identification, editing and matching is likely to produce an easier analysis and a more refined spot set. The development of newer integrated tools such as the spot excision tool (ProteomeWorks, Bio-Rad) would greatly aid the accurate recovery of gel spots (though we will need to resort to sending intact gels to APAF, Australian Proteome Analysis Facility for further processing). Bioinformatics databases will be the way of the future, but security issues may need to be addressed even though many of these require a multi-level secure log-in for entry. These and other advancements in the field address the needs of a growing science with great potential for the discovery of new potential diagnostic markers/ therapeutic targets, not just in breast cancer but any disease.

As previously discussed in the literature review, the evolution of our knowledge of the oestrogen receptor began with its discovery, leading to its significance as a prognostic marker and finally becoming a target for modulation. In later years, the exact process could be said to have occurred in the case of Herceptin, largely due to the work of Slamon *et al* [33, 38, 84]. Only through the discovery of the *OR* and *HER-2* proteins have they evolved into significant prognostic markers and eventual targets for modulation. It is hoped that with further investigation and characterization, a similar protein be identified that may follow the examples above. We believe that proteomic sciences will be the key to this discovery.

REFERENCES

1. UTAS, *Cancer in Tasmania, Incidence and Mortality 1998.* 2001, Menzies Centre, Tasmanian Cancer Registry.
2. AIHW, *Cancer in Australia 1997.* 2000, Australian Institute of Health and Welfare: Canberra.
3. la Vecchia, C., et al., *Olive oil, other dietary fats, and the risk of breast cancer (Italy).* Cancer Causes Control, 1995. 6(6): p. 545-50.
4. Yuan, J.M., et al., *Diet and breast cancer in Shanghai and Tianjin, China.* Br J Cancer, 1995. 71(6): p. 1353-8.
5. Stoppard, D.M., *The Breast Book. The essential guide to breast care and breast health for women of all ages.* 1996, London: Dorling Kindersley Limited.
6. IARC, *GLOBOCAN 2002.*
7. Wooster, R., et al., *Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13.* Science, 1994. 265(5181): p. 2088-90.
8. Wu, L.C., et al., *Identification of a RING protein that can interact in vivo with the BRCA1 gene product.* Nat Genet, 1996. 14(4): p. 430-40.
9. Cotran, K., Robbins, *Pathological basis of disease.* 2001.
10. Khanna, K.K. and G. Chenevix-Trench, *ATM and genome maintenance: defining its role in breast cancer susceptibility.* J Mammary Gland Biol Neoplasia, 2004. 9(3): p. 247-62.
11. Trentham-Dietz, A., et al., *Weight change and risk of postmenopausal breast cancer (United States).* Cancer Causes Control, 2000. 11(6): p. 533-42.
12. Hunt, K.A. and E.A. Sickles, *Effect of obesity on screening mammography: outcomes analysis of 88,346 consecutive examinations.* AJR Am J Roentgenol, 2000. 174(5): p. 1251-5.
13. Kjaerheim, K., *[Alcohol, breast cancer and causal inference in epidemiology].* Tidsskr Nor Laegeforen, 1997. 117(26): p. 3771-6.
14. Manjer, J., et al., *Breast cancer incidence in relation to smoking cessation.* Breast Cancer Res Treat, 2000. 61(2): p. 121-9.
15. Zemla, B., *The role of selected dietary elements in breast cancer risk among native and migrant populations in Poland.* Nutr Cancer, 1984. 6(3): p. 187-95.
16. Bilimoria MM, M.M., *The women at increased risk for breast cancer: evaluation and management strategies.* Cancer J Clin, 1995. 46: p. 263.
17. Bulbrook, R.D., *Geographical variation in endocrine function and its relation to breast cancer incidence: some general considerations.* Breast Cancer Res Treat, 1991. 18 Suppl 1: p. S37-40.
18. Bieber, E.J. and R.B. Barnes, *Breast cancer and HRT--what are the data?* Int J Fertil Womens Med, 2001. 46(2): p. 73-8.
19. *Breast cancer and hormonal contraceptives: collaborative reanalysis of individual data on 53 297 women with breast cancer and 100 239 women without breast cancer from 54 epidemiological studies. Collaborative Group on Hormonal Factors in Breast Cancer.* Lancet, 1996. 347(9017): p. 1713-27.
20. Denley H, P.S., Elston CW, Lee ASH, Ellis IO, *Preoperative assesment of prognostic factors in breast cancer.* J Clin Pathol, 2001. 54: p. 20-24.

21. Molina R, J.J., Filella X, Zanon G, Pahisa J, Munoz M, *c-erB2 oncoprotein, CEA and CA15.3 in patients with breast cancer: prognostic value.* Breast Cancer Res Treat, 1998. 51: p. 109-119.
22. Kuiper GGJM, E.E., Pelto-Huiko M, Nilsson S and Gustafsson JA, *Cloning of a natural estrogen receptor expressed in rat prostate and ovary.* Proc Natl Acad Sci USA, 1996. 93: p. 5925-5930.
23. Krege JH, e.a., *Generation and reprodcutive phenotypes of mice lacking estrogen receptor beta.* Proc Natl Acad Sci USA, 1998. 95: p. 15677-15682.
24. Pike, A.C.W., et al., *Structure of the ligand-binding domain of oestrogen receptor beta in the presence of a partial agonist and a full antagonist.* EMBO J., 1999. 18(17): p. 4608-4618.
25. Lazennec, G., et al., *ER beta inhibits proliferation and invasion of breast cancer cells.* Endocrinology, 2001. 142(9): p. 4120-30.
26. Palmieri, C., et al., *The expression of oestrogen receptor (ER)-beta and its variants, but not ERalpha, in adult human mammary fibroblasts.* J Mol Endocrinol, 2004. 33(1): p. 35-50.
27. Group, E.B.C.T.C., *Effects of adjuvant tamoxifen and of cytotoxic therapy on mortality in early breast cancer: an overview of 61 randomized trials among 28, 896 women.* N Engl J Med, 1988. 319: p. 1681-92.
28. Fisher, B. and J. Costantino, *Highlights of the NSABP Breast Cancer Prevention Trial.* Cancer Control, 1997. 4(1): p. 78-86.
29. O'Donoghue, A.E., et al., *Cathepsin D in primary breast carcinoma: adverse prognosis is associated with expression of cathepsin D in stromal cells.* Breast Cancer Res Treat, 1995. 33(2): p. 137-45.
30. Joensuu, H., S. Toikkanen, and J. Isola, *Stromal cell cathepsin D expression and long-term survival in breast cancer.* Br J Cancer, 1995. 71(1): p. 155-9.
31. Tovey, S.M., et al., *Outcome and human epidermal growth factor receptor (HER) 1-4 status in invasive breast carcinomas with proliferation indices evaluated by bromodeoxyuridine labelling.* Breast Cancer Res, 2004. 6(3): p. R246-51.
32. Slamon, D.J., et al., *Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene.* Science, 1987. 235(4785): p. 177-82.
33. Slamon, D.J. and G.M. Clark, *Amplification of c-erbB-2 and aggressive human breast tumors?* Science, 1988. 240(4860): p. 1795-8.
34. Slamon, D. and M. Pegram, *Rationale for trastuzumab (Herceptin) in adjuvant breast cancer trials.* Semin Oncol, 2001. 28(1 Suppl 3): p. 13-9.
35. Pegram M, H.S., Lewis G, et al, *Inhibitory effects of combinations of HER-2/neu antibody and chemotherapeutic agents used for treatment of human breast cancers.* Oncogene, 1999. 18: p. 2241-2251.
36. Baselga J, N.L., Albanell J, et al, *Recombinant humanized anti-HER-2 antibody (Herceptin) enhances the antitumour activity of paclitaxel and doxorubicin against HER-2/neu overexpressing human breast cancer xenografts.* Cancer Res, 1998. 58: p. 2825-2831.
37. Pegram MD, B.D., Wirth C, et al. *Antibody dependant cell mediated cytotoxicity in breast cancer patients in Phase III clinical trials of a humanized anti-HER-2 antibody.* in *Proc Am Assoc Cancer Res.* 1997.

38. Slamon DJ, L.-J.B., Shak S, Fuchs H, Paton V, Bajamonde A, et al, *Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2.* N Engl J Med, 2001. 344: p. 783-792.
39. Scholzen T, G.J., *The Ki-67 protein: from the known and the unknown.* J Cell Physiol, 2000. 182(3): p. 311-22.
40. Pietilainen T, L.P., Aaltomaa S, Eskelinen M, Kosma VM, Syrjanen K, *The important prognostic value of Ki-67 expression as determined by image analysis in breast cancer.* J Cancer Res Clin Oncol, 1996. 122(11): p. 687-92.
41. Wan, *Cloning differentially expressed mRNAs.* Nat Biotechnol, 1996. 14: p. 1685-1691.
42. Wilkins MR, W.K., Appel RD, Hochstrasser DF (Eds.), *Proteome Research: New Frontiers in Functional Genomics.* 1997: Springer.
43. Kenrick KG, M.J., *Iso-electric focusing and gradient gel electrophoresis: a two-dimensional technique.* Anal Biochem, 1970. 33: p. 204-207.
44. Osborne, R.J. and M.G. Hamshere, *A genome-wide map showing common regions of loss of heterozygosity/allelic imbalance in breast cancer.* Cancer Res, 2000. 60(14): p. 3706-12.
45. Richard, S.M., et al., *Nuclear and mitochondrial genome instability in human breast cancer.* Cancer Res, 2000. 60(15): p. 4231-7.
46. Shen, C.Y., et al., *Genome-wide search for loss of heterozygosity using laser capture microdissected tissue of breast carcinoma: an implication for mutator phenotype and breast cancer pathogenesis.* Cancer Res, 2000. 60(14): p. 3884-92.
47. Fukino, K., et al., *Combined total genome loss of heterozygosity scan of breast cancer stroma and epithelium reveals multiplicity of stromal targets.* Cancer Res, 2004. 64(20): p. 7231-6.
48. Ellis, N.A., et al., *Localization of breast cancer susceptibility loci by genome-wide SNP linkage disequilibrium mapping.* Genet Epidemiol, 2006. 30(1): p. 48-61.
49. Huusko, P., et al., *Genome-wide scanning for linkage in Finnish breast cancer families.* Eur J Hum Genet, 2004. 12(2): p. 98-104.
50. Banks, R.E., et al., *The potential use of laser capture microdissection to selectively obtain distinct populations of cells for proteomic analysis--preliminary findings.* Electrophoresis, 1999. 20(4-5): p. 689-700.
51. Franzen, B., et al., *Sample preparation of human tumors prior to two-dimensional electrophoresis of proteins.* Electrophoresis, 1995. 16(7): p. 1087-9.
52. Giometti, C.S., K. Williams, and S.L. Tollaksen, *A two-dimensional electrophoresis database of human breast epithelial cell proteins.* Electrophoresis, 1997. 18(3-4): p. 573-81.
53. Weiss, W., W. Postel, and A. Gorg, *Application of sequential extraction procedures and glycoprotein blotting for the characterization of the 2-D polypeptide patterns of barley seed proteins.* Electrophoresis, 1992. 13(9-10): p. 770-3.
54. Molloy, M.P., et al., *Extraction of membrane proteins by differential solubilization for separation using two-dimensional gel electrophoresis.* Electrophoresis, 1998. 19(5): p. 837-44.
55. Rabilloud, *Procedures for two-dimensional electrophoretic analysis of nuclear proteins.* J Chromatogr, 1986. 351: p. 77-89.

56. Taylor, R.S., et al., *Proteomics of rat liver Golgi complex: minor proteins are identified through sequential fractionation.* Electrophoresis, 2000. 21(16): p. 3441-59.
57. Fountoulakis, M., et al., *Enrichment of low abundance proteins of Escherichia coli by hydroxyapatite chromatography.* Electrophoresis, 1999. 20(11): p. 2181-95.
58. Gorg, A., et al., *Very alkaline immobilized pH gradients for two-dimensional electrophoresis of ribosomal and nuclear proteins.* Electrophoresis, 1997. 18(3-4): p. 328-37.
59. Sabounchi-Schutt, F., et al., *An immobiline DryStrip application method enabling high-capacity two-dimensional gel electrophoresis.* Electrophoresis, 2000. 21(17): p. 3649-56.
60. Bjellqvist, B., et al., *The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences.* Electrophoresis, 1993. 14(10): p. 1023-31.
61. Strahler, J.R., et al., *Effect of salt on the performance of immobilized pH gradient isoelectric focusing gels.* Electrophoresis, 1988. 9(2): p. 74-80.
62. Herbert, B.R., et al., *Improved protein solubility in two-dimensional electrophoresis using tributyl phosphine as reducing agent.* Electrophoresis, 1998. 19(5): p. 845-51.
63. Altland, K., et al., *Isoelectric focusing of basic proteins: the problem of oxidation of cysteines.* Electrophoresis, 1988. 9(9): p. 474-85.
64. Gorg, A., et al., *Temperature-dependent spot positional variability in two-dimensional polypeptide patterns.* Electrophoresis, 1991. 12(9): p. 653-8.
65. Westbrook, J.A., et al., *Zooming-in on the proteome: very narrow-range immobilised pH gradients reveal more protein species and isoforms.* Electrophoresis, 2001. 22(14): p. 2865-71.
66. Yan, J.X., et al., *A modified silver staining protocol for visualization of proteins compatible with matrix-assisted laser desorption/ionization and electrospray ionization-mass spectrometry.* Electrophoresis, 2000. 21(17): p. 3666-72.
67. Bio-Rad, *2-D Electrophoresis for Proteomics, A Methods and Products Manual*, ed. L.H. D Garfin.
68. Lopez, M.F., et al., *A comparison of silver stain and SYPRO Ruby Protein Gel Stain with respect to protein detection in two-dimensional gels and identification by peptide mass profiling.* Electrophoresis, 2000. 21(17): p. 3673-83.
69. Lauber, W.M., et al., *Mass spectrometry compatibility of two-dimensional gel protein stains.* Electrophoresis, 2001. 22(5): p. 906-18.
70. Edman, P. and G. Begg, *A protein sequenator.* Eur J Biochem, 1967. 1(1): p. 80-91.
71. Wilm, M., et al., *Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry.* Nature, 1996. 379(6564): p. 466-9.
72. Wilkins, M.R., et al., *Rapid protein identification using N-terminal "sequence tag" and amino acid analysis.* Biochem Biophys Res Commun, 1996. 221(3): p. 609-13.
73. Franzen, B., et al., *Analysis of polypeptide expression in benign and malignant human breast lesions: down-regulation of cytokeratins.* Br J Cancer, 1996. 74(10): p. 1632-8.

74. Rabilloud, T., *Solubilization of proteins for electrophoretic analyses.* Electrophoresis, 1996. 17(5): p. 813-29.
75. Chevallet, M., et al., *New zwitterionic detergents improve the analysis of membrane proteins by two-dimensional electrophoresis.* Electrophoresis, 1998. 19(11): p. 1901-9.
76. Garfin, *Isoelectric focusing*, in *Separation Science and Technology*, A. S, Editor. 2000, Academic Press: San Diego. p. 263-298.
77. Hoving, S., H. Voshol, and J. van Oostrum, *Towards high performance two-dimensional gel electrophoresis using ultrazoom gels.* Electrophoresis, 2000. 21(13): p. 2617-21.
78. Merrill CR, G.D., Sedman SA and Ebert MH, *Ultrasensitive stain for proteins in polyacrylamide gels shows regional variation in cerebrospinal fluid proteins.* Science, 1981. 211: p. 1437-1438.
79. Gottlieb M, C.M., *Silver staining of native and denatured eucaryotic DNA in agarose gels.* Anal Biochem, 1987. 165: p. 33-37.
80. Bini, L., et al., *Protein expression profiles in human breast ductal carcinoma and histologically normal tissue.* Electrophoresis, 1997. 18(15): p. 2832-41.
81. Schwalke MA, D.C., Bleday R et al, *Protein differences in human pancreatic cancer cell lines with diverse metastaic potential.* Arch Surg, 1990. 125: p. 469-471.
82. Osada T, S.M., Ino Y, et al, *E-cadherin is involved in the intrahepatic metastasis of hepatocellular carcinoma.* Hepatology, 1996. 24: p. 1460-1467.
83. Roberts K, B.K., Stanton P, Lord R, *Proteomic analysis of selected prognostic factors of breast cancer.* Proteomics, 2004. 4: p. 784-792.
84. Slamon, D.J., et al., *Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer.* Science, 1989. 244(4905): p. 707-12.

LIST OF FIGURES

LIST OF TABLES

APPENDICES

CONSENT FORM

"Proteome Analysis of Breast Cancer" RHH ethics 2001.02

Statement by the patient:

I have read and understood the "Information Sheet" for this research study. I understand that breast tissue removed, as a course of my surgical management will be provided to researchers involved in this study. I have been informed that the results of the study may not be of any direct benefit to my medical management. Any questions that I have asked have been answered to my satisfaction. I agree to participate in this investigation and understand that I may withdraw at any time without prejudice to my future medical care. I agree that research data gathered for the study may be published provided that I cannot be identified as a subject.

Name of Patient:

Signature of Patient:

Date:

Statement by the researcher:

I have explained this research study and the implications of participation in it, to the volunteer patient. I believe that the consent is informed and that she understands the implications of participation.

Name of Researcher:

Signature of Researcher:

Date:

INFORMATION SHEET

for patients diagnosed with breast cancer

"Proteome Analysis of Breast Cancer"
Dr Roger Lord, Dr Kiron Bhatia & Professor Peter Stanton (Chief Investigators)

Study procedure and possible risks:

As you have been diagnosed with breast cancer and have been scheduled for surgery to remove the tumour, you are invited to participate in the study. There are no additional risks in being involved in this study nor are there any additional procedures that will take place. Your consent is sought purely for the scientific use of the tissue being removed.

Purpose of Study:

The aim of this project is to produce a basic protein map of breast tissue, be it cancerous or otherwise normal. The researchers of this proposed study seek firstly, to look at differences in known biological markers of breast cancer between two populations of women, one group from Japan and the other from Australia. Women from Asian regions have a much lower mortality rate for breast cancer when compared to those in Australia. It has been suggested that environmental and/or dietary factors may account for this observation but as yet there is no data to clearly indicate that this is the case. If biological differences between these groups do exist, then these might present as higher or lower levels of known biological markers of the disease.

The second part of the study uses a state of the art separation method that allows complex mixtures of molecules to be resolved. Using this method a biological map of all the molecules that make up a piece of tissue, such as a breast cancer, can be separated to produce a map, which is unique for that tissue. Normal breast tissue (e.g. from women under going breast reduction) would thus have a different map to breast cancer or for that matter advanced cancers that have broken away from the main breast tumour and begun to grow in another region of the body. The differences between maps for the normal breast tissue compared with cancerous tissue may represent molecules that have been generated as a result of the cancer. These molecules could be markers of the disease or targets for future vaccines or pharmaceutical drugs.

Inclusion or exclusion criteria for this research project:

This study includes patients who are either having a breast cancer removed, including any advanced cancers that have spread to the lymph nodes and women under going breast reduction surgery. There are no exclusion criteria for this project.

Participation and withdrawal from the research project:

Participation in this research project is entirely voluntary and no financial payment will be given. If you decide to take part in this study, there will be no prejudice against your future care if you decide to withdraw prior to surgery.

Confidentiality:

Information and materials obtained during this project will remain strictly confidential and shall only be used for the purposes of this study. Only the investigators directly involved in this project will have access to identifying data. The Privacy Act, 1988 gives this research project the guidelines for conduct. Your consent to participate in this research project will be an informed one and the investigator will answer any questions you have. Any unused tissue may be kept for future analysis in relation to this study alone, and will be disposed of in a professional manner when no longer required.

Information Sheet and Consent Form:

Copies of the information Sheet and signed consent forms will be given to you to keep.

For more information about this research project, you can contact:

Dr Roger Lord
Discipline of Surgery
GPO Box 252-28 Hobart
Tasmania 7001 Australia
Tel (03) 6226 4896
Email: John.Lord@utas.edu.au

Dr Kiron Bhatia
Discipline of Surgery
GPO Box 252C-28 Hobart
Tasmania 7000 Australia
Tel (03) 6226 4875
Email: kdbhatia@postoffice.utas.edu.au

If you have any concerns of an ethical nature or complaints about the manner in which the project is conducted, you may contact the chair of the Ethics Committee:

Dr Rosalie Parton
6222 8226
or
Chris Hooper
Secretary, RHH ethics committee
6226 2763

INFORMATION SHEET

for patients undergoing breast reduction surgery

"Proteome Analysis of Breast Cancer"
Dr Roger Lord, Dr Kiron Bhatia & Professor Peter Stanton (Chief Investigators)

Study procedure and possible risks:

As you are to undergo breast reduction surgery, you are invited to participate in this study so as to provide normal tissue for comparison with breast cancer tissue. There are no additional risks in being involved in this study nor are there any additional procedures that will take place. Your consent is sought purely for the scientific use of the tissue being removed.

Purpose of Study:

The aim of this project is to produce a basic protein map of breast tissue, be it cancerous or otherwise normal. The researchers of this proposed study seek firstly, to look at differences in known biological markers of breast cancer between two populations of women, one group from Greece and the other from Australia. Women from Mediterranean regions have a much lower mortality rate for breast cancer when compared to those in Australia. It has been suggested that environmental and/or dietary factors may account for this observation but as yet there is no data to clearly indicate that this is the case. If biological differences between these groups do exist, then these might present as higher or lower levels of known biological markers of the disease.

The second part of the study uses a state of the art separation method that allows complex mixtures of molecules to be resolved. Using this method a biological map of all the molecules that make up a piece of tissue, such as a breast cancer, can be separated to produce a map, which is unique for that tissue. Normal breast tissue (e.g. from women undergoing breast reduction) would thus have a different map to breast cancer or for that matter advanced cancers that have broken away from the main breast tumour and begun to grow in another region of the body. The differences between maps for the normal breast tissue compared with cancerous tissue may represent molecules that have been generated as a result of the cancer. These molecules could be markers of the disease or targets for future vaccines or pharmaceutical drugs.

Inclusion or exclusion criteria for this research project:

This study includes patients who are either having a breast cancer removed, including any advanced cancers that have spread to the lymph nodes and women under going breast reduction surgery. There are no exclusion criteria for this project.

Participation and withdrawal from the research project:

Participation in this research project is entirely voluntary and no financial payment will be given. If you decide to take part in this study, there will be no prejudice against your future care if you decide to withdraw prior to surgery.

Confidentiality:

Information and materials obtained during this project will remain strictly confidential and shall only be used for the purposes of this study. Only the investigators directly involved in this project will have access to identifying data. The Privacy Act, 1988 gives this research project the guidelines for conduct. Your consent to participate in this research project will be an informed one and the investigator will answer any questions you have. Any unused tissue may be kept for future analysis in relation to this study alone, and will be disposed of in a professional manner when no longer required.

Information Sheet and Consent Form:

Copies of the information Sheet and signed consent forms will be given to you to keep.

For more information about this research project, you can contact:

Dr Roger Lord
Discipline of Surgery
GPO Box 252-28 Hobart
Tasmania 7001 Australia

Tel (03) 6226 4896
Email: John.Lord@utas.edu.au

Dr Kiron Bhatia
Discipline of Surgery
GPO Box 252C-28 Hobart
Tasmania 7000 Australia

Tel (03) 6226 4875
Email: kdbhatia@postoffice.utas.edu.au

If you have any concerns of an ethical nature or complaints about the manner in which the project is conducted, you may contact the chair of the Ethics Committee:

Dr Rosalie Parton
6222 8226
or
Chris Hooper
Secretary, RHH ethics committee
6226 2763