Proteomic Analysis of Breast Cancer

By Kylie Roberts

Submitted in accordance with the University of Tasmania's PhD guidelines, June 2009.

Statement of Authenticity

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university, and to the best of my knowledge contains no copy or paraphrase of material previously published by another person except where due reference is made in the text.

Signed.....

Acknowledgements

Over the course of completing this PhD there are many people that I would like to thank.

Firstly to my supervisors, Professor Peter Stanton and Dr Roger Lord, thank you for you guidance and support, through you all I have evolved into a confident and self reliant researcher.

To Greg Woods, thank you for your support and guidance during what were some challenging times.

To the Immunology and Cancer research Group, thank you for your friendship and some great times over what has been a memorable few years.

To my family and friends, thank you for your support and help when it was needed.

Table of Contents

Chapter 1 – Literature Review
Pg…1

Chapter 2 - Project Aims
Pg…33

Chapter 3 – Materials and Methods
Pg…35

Chapter 4 - Differential analysis of the proteins in cancerous and healthy breast tissue
Pg…57

Chapter 5 – Differential analysis of Early and Late Onset Breast Cancer
Pg…95

Chapter 6 – Differential analysis of two different healthy cohorts
Pg…127

Chapter -7 Functional investigation into the involvement of b-casein, IGF-1 and STAT5 in breast cancer
Pg…135

Chapter 8 – final discussion
Pg…147

Chapter 9 – References
Pg…156

Abstract

Introduction: Breast cancer is a disease that will affect every woman in her life. Whether that affect will be direct or indirect is, at this stage almost impossible to determine. There have been several key findings in recent years with genes being found that are linked directly to breast cancer. With these being attributed to less than 10% of breast cancers diagnosed, the need for new biomarkers of breast cancer exists. This study aims to find new biomarkers though differential analysis of breast cancer tissue. Several different analysis have been conducted to provide insight in to potential new markers for diagnosing early onset breast cancers, looking for signs in otherwise healthy tissue and ascertaining differences between cancer and healthy tissue as a whole. This study also looks in to the functional areas of proteomic research, determining a need for further investigation in to IGF-1 and β-casein, which have been inconclusively linked to breast cancer in the past.

Methods: By employing selected proteomic technologies including iso-electric focusing and polyacrylamide gel electrophoresis and combining samples from populations from different locations across the globe, collections of two-dimensional gels from breast tissue were differentially compared to isolate proteins that have a high likelihood of being involved in breast cancer.

Results: This approach to breast cancer research has led to many proteins being isolated that have a role in breast cancer. The proteins found have a diverse range of roles from signaling proteins to structural proteins. Within those proteins successfully identified are proteins like HSP60, known to have a role in breast cancer, serum albumin which has limited information regarding a role in breast cancer and FLJ20309, one of several proteins that have been assumed from cDNA clones with little available information available regarding its function.

Conclusions: This study has led to the discovery of proteins that were previously thought to have limited, if no input in breast cancer. All of the proteins found require further experiments to fully elucidate their function in breast cancer however at least some of these are show promise as being of diagnostic and/or therapeutic value.

List of Figures

Fig1.1.1: Incidence and mortality rate of breast cancer in Taiwan, Greece and Australia. **Pg 2.**

Figure 1.1.2: The Age Standardised Rate (ASR) of breast cancer incidence per 100 000 women for separate age based cohorts for the Australian population. **Pg3.**

Figure 1.1.3: Incidence and ASR of breast cancer compared in the Australian population. *Incidence and ASR are measured differently. Incidence refers to the total number of cases reported in the stated year, whilst the ASR is an estimate of the number of women per 100 000 in each age group that is expected to report with breast cancer in each year. **Pg 4.**

Figure 1.1.4: Age standardised rates of mortality from breast cancer in Australia. A comparison between 1994 and 2004. **Pg 4.**

Figure 1.1.5: Incidence of Breast cancer on the Greek Island of Crete in 1994/1995. **Pg 5.**

Figure 1.1.6: Incidence of breast cancer in Taiwan based on the age standardised rate. The rate of incidence for 1997. **Pg6.**

Fig 1.2.1:vThe role of BRCA in DSB repair. (taken from Yoshida et al[9]). **Pg 7.**

Figure1.2.2: Roles of e2 and its receptor. (taken from Hayashi *et al* 2003[27].) **Pg 11.**

Figure 1.5.1: Key proteins operate and the cell cycle checkpoints. Noted are just a few proteins, that when damaged, de-regulate the cell cycle. **Pg 17.**

Figure 1.6.1: Sagittal section of a rat, and the comparative *in situ* MALDI-TOF image achieved by the selective highlighting of the peptides within the drug, 2hrs post treatment. This shows that the drug is directed strongly toward lung tissue, followed by traces in the excretory system of the kidney and bowel, but also highlights unexpected findings in the testis and an organ that was unidentified. (http://www.uab.edu/proteomics/pdf_files/2006/Mobley_Proteomics_Class.pdf.) **Pg 31.**

Fig. 3.1: Set up of a western blot cassette. The cassette would be folded and then the side corresponding to the right hand side of the the figure, would be aligned with the negative electrode. **Pg 45.**

Figure 5.1: Age distribution for each population cohort. Samples were placed in one of two categories, early inset = below 53 and late onset = above 53. **Pg 95.**

Figure 6.1: Total protein content of SSP 4402 in solution 2. NNH (Normal with no history) showed no expression of the isoform, whilst NWH (Normal with history) had very high expression levels **Pg 129.**

Figure 6.2: Staging analysis of pyruvate kinase 4402 in the cancerous cohorts form each population. **Pg 129.**

Figure 6.3: Analysis of the PP2A isoform shows an increase in the samples that have had a history of breast cancer, despite an obvious decrease in all of the cancer cohorts. **Pg 130.**

Figure 6.4: Staging analysis shows no clear trend for the expression of the PP2A isoform SSP 2402 in breast cancer for any of the populations **Pg 130.**

Figure 6.5: Analysis of hypothetical protein GI 51491284 shows that there are decreased levels of protein in samples without a history of breast cancer. **Pg 131.**

Figure 6.6: Staging analysis of isoform SSP 1202 shows that there is a peak in stage 1 for TCB and GCB, but this trend is not replicated in ACB. **Pg 131.**

Figure 6.7: Analysis of the alternate isoform 1204 of GI 51491284 shows more erratic expression, however presence in the controls with no history of breast cancer is absent **Pg 131.**

Figure 6.8: Staging analysis shows levels of 1204 are highest in stage one for ACB and TCB populations. GCB showed no expression oat all. **Pg 132.**

Figure 7.1a: Sequential extraction solution 2 from the extraction of breast tissue run in two dimensions on to a polyacrylamide gel TCB23 SOL2. Note the presence of artifacts on the image centre to right on the image, this is due to air bubble formation during imaging and is removed during the analysis and normalization process. **Pg138.**

Figure 7.1b: Western blot of IGF-1 using a monoclonal X terminal directed antibody indicating a dimmer located at 15kda **pg 139.**

Figure 7.1c: Western blot of β-casein using a monoclonal antibody directed against the y terminal. Results indicate a series of proteins of approximately 25kDa. **pg 139.**
Figure 7.1d: Western blot of STAT 5a . On the left is the gel obtained through separation and on the right an image of the western blot, indicating a breakdown product of STAT5a. **pg 140.**

Figure 7.1e: Using PMSF, the normalized gel on the left shows more distinct separation of the proteome. On the right, due to this additional serine protease inhibitor western blotting led to not only a full length version at 95kDa but also a dimer of STAT 5a at ~190kDa being revealed as present in the cytoplasm of breast tissue. These figures demonstrate normalized gel images. **Pg 140.**

Figure 7.2a: Trend analysis of healthy breast tissue shows that at the 6^{th} patient (N3) the βcasein level is steadily rising with the level of IGF-1. However from the 7^{th} patient (N1) the level of βcasein expression drops and is consistently much lower than that of IGF-1. **Pg 141.**

Figure 7.2b: Trend analysis of cancerous tissue shows that there are only several patients which are expressing higher levels of IGF-1 than β-casein. The lower levels of IGF-1, unlike in the healthy tissue fail to show a relationship between the level of IGF-1 and β-casein. **Pg 141.**

Figure 7.4.4a: Percentage of total protein expressed in each cell line treatment group, for both isoforms of β-casein and IGF-1 that were isolated. **Pg 143.**

List of Tables

Table 1.7.1: Comparative proteomic studies in to breast cancer conducted over the past 5 years. [84-87]. **Pg 25.**

Table 3.1: Chemicals and supplies used in this study. **Pg 36.**

Table 3.6.1: Treatment regime for T47D cell culture flasks. Appendix B contains the datasheets for each of the hormones administered. **Pg48.**

Table 4.1: Samples included in each of the analysis sets for the healthy versus cancer comparisons. There are two files for each of the extraction solutions. **Pg 56.**

Table 4.2: Summary of the proteins found to be of interest in the cancerous versus healthy analysis of breast tissue. # Significance was determined by a student t-test with a CI of 95%. **Pg 57.**

Table 4.3: Details from the MALDI-TOF analysis conducted at APAF for proteins differentially regulated between healthy and cancerous breast tissue. P**g 58.**

Table 5.1: Detailed list of which samples were in the Early onset and Late onset cohorts when division was based on the age on 53**. Pg 96.**

Table 5.2 : Summary of the proteins that were sent for sequencing by MALDI-TOF MS. 12 of the 26 proteins were identified. The data received from this sequencing and the relevant theoretical information is presented for the proteins Tat were identified. A MOWSE score of greater that 65 is a significant identification. **Pg97.**

Table 6.1: Results from MALDI-TOF MS analysis conducted at APAF in Sydney on proteins of interest in NNH versus NWH comparisons. **Pg 128.**

Table 7.3: The number of proteins that was up or down regulated in T47D breast cancer cells in response to different hormones was highly variable. **Pg 144.**

Table 7.4: The number of proteins commonly up and down regulated by each of the hormone treatments**. Pg 145.**

Table 8.1: Comparative proteomic studies in to breast cancer conducted over the past 5 years. **Pg 157.**

Abbreviations

%TPC	Percentage of total protein content
2D PAGE	Two dimensional poly acrylamide electrophoresis
2DE	Two dimensional electrophpresis
ACB	Australian Cancerous Breast
APAF	Australian Proteome Analysis Facility
ASR	Age Standardised Rate
BRCA	Breast cancer protein
DCIS	Ductal carcinoma in situ
DIRP5A	Didulfide-isomerise related protein 5a
DNA	Deoxyribonucleic acid
E_2	Estradiol
ER	Estrogen receptor
GCB	Greek Cancerous breast
GH	Growth Hormone
HRT	Hormone Replacement therapy
HSA	Human serum albumin
HSP	Heat Shock Protein
IDC	Invasive ductal carcinoma
IEF	Iso-electric focusing
IGF-1	Insulin like growth factor-1
IPG	Immobilised pH gradient
LCM	Laser capture microdissection
MALDI-TOF	Matrix assisted laser desorption ionization- time of flight mass
MS	spectrometry
Mw	Molecular weight
NB	Normal Breast
pI	Isoelectric point
PMF	Peptide mass fingerprint
PRL	Prolactin
RNA	Ribonucleic acid
SLRP	Small leucine rich protein
STAT5	Signal transducer and activator of transcription factor 5
TCB	Taiwanese Cancerous Breast

1 LITERATURE REVIEW

1.1 INCIDENCE OF BREAST CANCER 2
1.1.1 Breast cancer in Australia 3
1.1.2 Breast cancer in Greece 4
1.1.3 Breast cancer in Taiwan 5
1.2 RISK FACTORS 6
1.2.1 Genetic Impact 7
1.2.2 Lifestyle 8
1.2.3 Others lifestyle factors 12
1.3 THE GENETICS OF BREAST CANCER 13
1.3.1 Germ line mutations 14
1.3.2 Somatic Mutations 16
1.4 PROTEINS INVOLVED IN BREAST CANCER 18
1.4.1 Tumour Suppressor Genes 19
1.4.2 OncoGenes 19
1.5 PROTEINS IN BREAST CANCER 20
1.5.1 Proliferation and cell cycle control proteins 20
1.5.2 Adhesion Proteins 20
1.5.3 Apoptotic proteins 21
1.6 PROTEOMICS 21
1.6.1 History of Proteomics 21
1.6.2 Proteomics in Breast Cancer 24
1.6.3 Comparative Studies 25
1.6.4 Translational studies 29
1.6.5 Advancing technologies 30
1.6.6 Current Limitations 32

1.1 INCIDENCE OF BREAST CANCER

The incidence of breast cancer is increasing globally. Every year more and more women are being diagnosed with over 1.1 million women diagnosed in 2002 according to the International Agency for Research on Cancer (IARC) [1]. This is in contrast to the mortality rates due to breast cancer despite 400 000 women dying from the disease.

The rate at which incidence of breast cancer is increasing is not uniform for all countries. In western countries such as those in the EU, the USA and Australia the rates are rising at an average of 1.5% each year, whilst countries with traditionally lower incidence of breast cancer, such as many Asian countries, are increasing more rapidly, with an annual increase of between 2% for Japan and as much as 5% for China [1].

The mortality rates of breast cancer are on the decline however, with the rates of breast cancer associated death dropping by 3.6% annually [2]. This global drop in mortality has been attributed to better therapeutics and earlier diagnosis.

The incidence and mortality rates of breast cancer are related to many factors. These factors such as genetics, lifestyle and environment, which will be expanded on in section 1.2, all have the potential to alter the proteome of the cell and as this study will be based upon populations from Australia, Greece and Taiwan, it is important to understand the trends of breast cancer in these countries. The current incidence and mortality rates can be seen in figure 1.1.1.

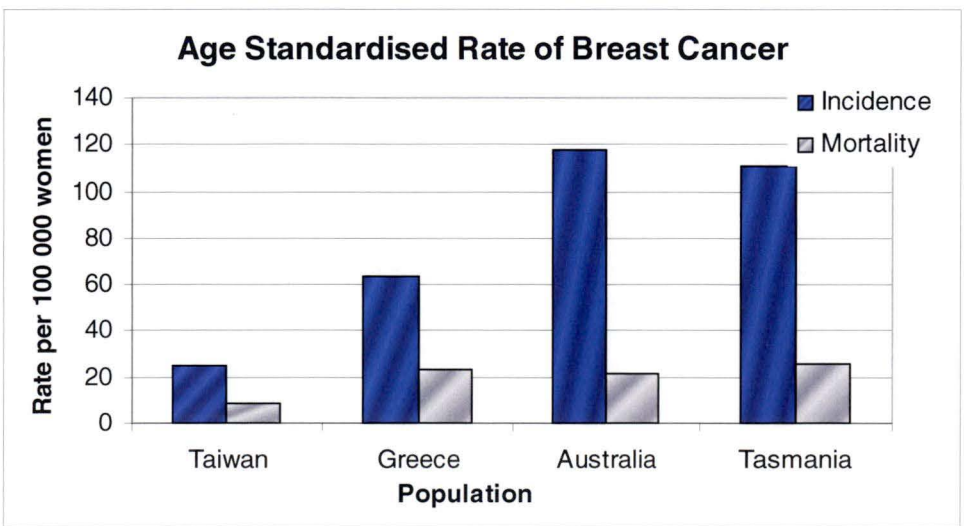


Figure 1.1.1: Incidence and mortality rate of breast cancer in Taiwan, Greece and Australia

1.1.1 BREAST CANCER IN AUSTRALIA

The rates of incidence and mortality of breast cancer in Australia have been tracked at a national level since 1983 by the Australian Institute for Health and Welfare. Data currently available extends to 2002. In this 19 year period the incidence has increased from an age standardised rate of 80 per 100 000 to 117 per 100 000 [2]. This 50% rise in incidence has been attributed, not to an actual increase in occurrence, but as being reflective of better screening and awareness of the disease within the community. Table 1.1.2 details the age breakdown of age at first diagnosis and the ASR per 100 000 women in that age group, whilst 1.1.3 looks at the actual rate of incidence of breast cancer in Australia with the data from 1.12 overlayed as a line graph, to demonstrate the distribution of incidence across the age range.

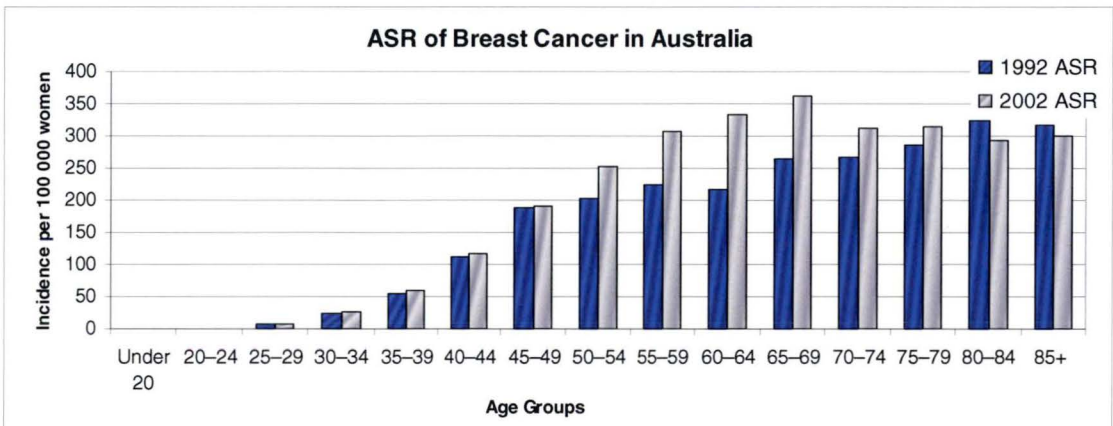


Figure 1.1.2: The Age Standardised Rate (ASR) of breast cancer incidence per 100 000 women for separate age based cohorts for the Australian population.

In contrast to incidence rates, the mortality rates due to breast cancer have improved for all age groups as can be seen in figure 1.1.4.

The decrease in mortality rates are a result of improved and earlier detection as well as advances in the treatment of the disease.

The age at which women are affected by the breast cancer in Australia, like around the world is an area of breast cancer research. The clinical outcome of breast cancer has been associated with younger women having poorer prognosis [3].

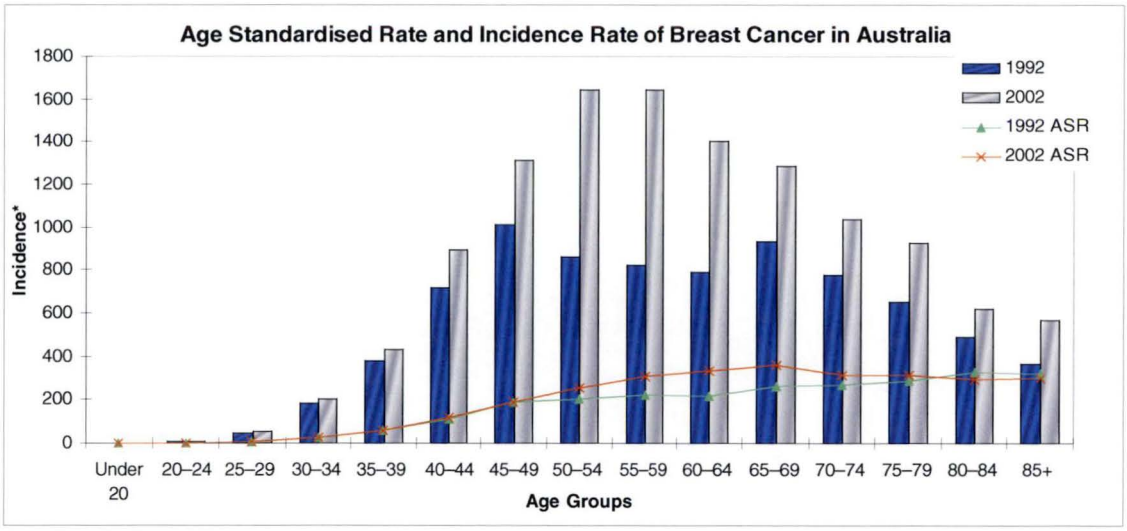


Figure 1.1.3: Incidence and ASR of breast cancer compared in the Australian population. *Incidence and ASR are measured differently. Incidence refers to the total number of cases reported in the stated year, whilst the ASR is an estimate of the number of women per 100 000 in each age group that is expected to report with breast cancer in each year.

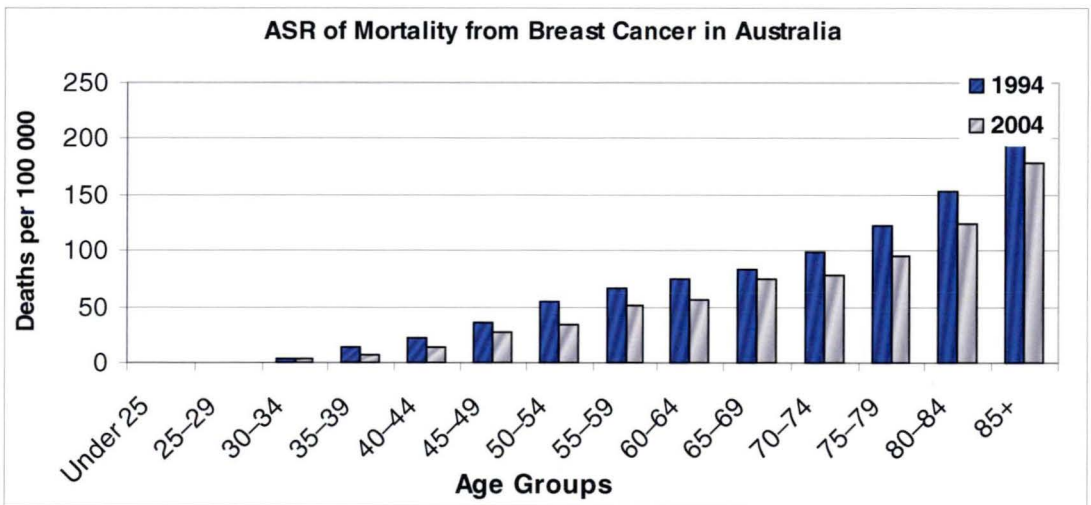


Figure 1.1.4: Age standardised rates of mortality from breast cancer in Australia. A comparison between 1994 and 2004.

In Australia 25% of women who were diagnosed with breast cancer are between the ages of 20 and 49 [2], this is a large percentage of women, for whom standard mammography is of limited use due to the higher density of breast tissue in this younger cohort of women [4].

1.1.2 Breast cancer in Greece

The incidence of breast cancer in the Greek women is lower than that of Australian women. The ASR of incidence for Greek women is 63 per 100 000 women per year with the ASR for the mortality from breast cancer is 23.0 [5].

The age distribution of incidence can be seen in Figure 1.1.4, however there are rates of incidence in the Greek province of Crete, as there is no national Greece cancer registry kept and the data has been taken form a report published by Vlachonikolis *et al.* in 2002 [6].

The differences in the rates of incidence between populations can be attributed to several factors including lifestyle, environment and genetics. All of these will have an effect on the cellular environment of the individual. This in turn will affect the proteome map that is generated from the tissue of women in each of the populations studied. The differences that are detected may not necessarily be the result of a difference in cancer tissue compared to healthy tissue, but in fact indicative of a cultural difference due to the genetic, lifestyle and environmental factors mentioned. Whilst these factors will be elaborated on in section 1.2, it is imperative that the differences between cancer alterations and population differences are distinguished between so as to find markers that are truly cancer specific.

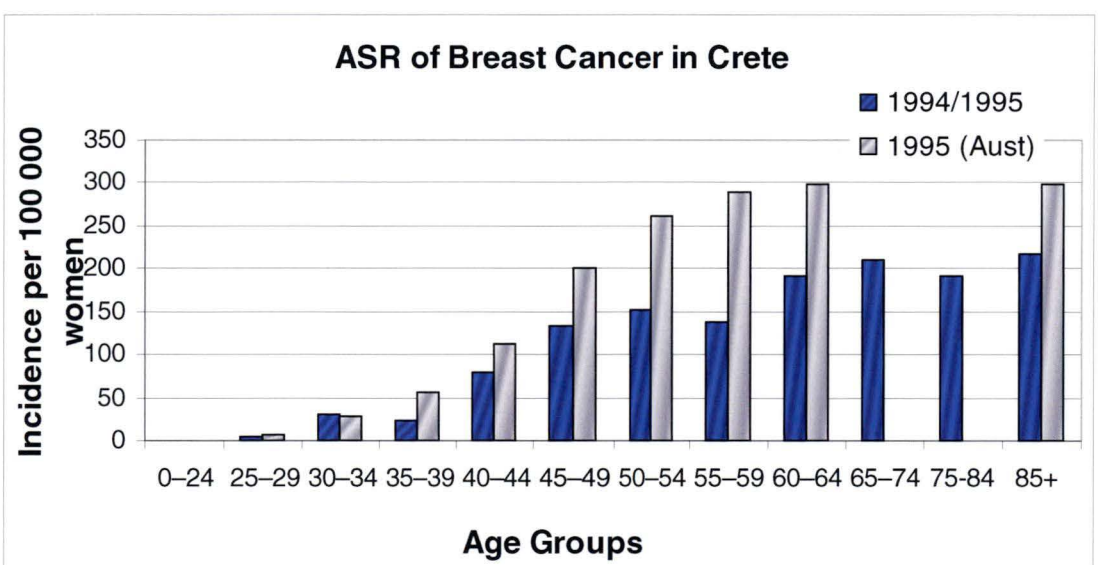


Figure 1.1.5: Incidence of Breast cancer on the Greek Island of Crete in 1994/1995.

1.1.3 BREAST CANCER IN TAIWAN

Breast cancer incidence in Taiwan has been following similar trends to those of many Asian countries. The rates of incidence are rising at twice the rate of western countries. This is anecdotally being attributed to the "westernisation of Asia", especially the influence of practices that increase the risk of getting breast cancer such as increased child bearing age and the use of the contraceptive pill. Other "westernisations" include a

changing diet and increased pollution. The current ASR of breast cancer in Taiwan is 25, with a very low mortality rate of 8.6 [7]. Figure 1.1.6 outlines the age standardised rate of breast cancer over the different age groups within the Taiwanese population.

Taiwan is unique compared to many countries as it has been reported to have a much higher frequency of early onset breast cancers. Incidence rates in women under the age of 50 are suggested to be upwards of 64% of newly diagnosed breast cancer cases [8]. The Taiwanese Cancer Registry has reported peaks of incidence in the age range of 45-49 years of age. All other cancers in Taiwanese women peak above the age of 70 [7].

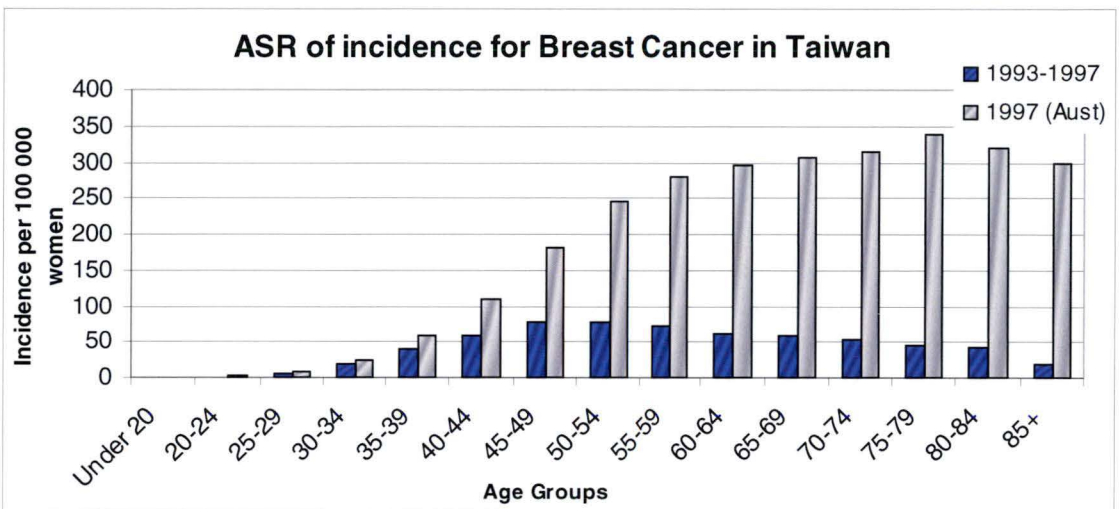


Figure 1.1.6: Incidence of breast cancer in Taiwan based on the age standardised rate. The rate of incidence for 1997 in Australia was used for comparison.

1.2 RISK FACTORS

Breast cancer is a multi-faceted disease that has many influencing factors. Whilst there has been no single cause found for breast cancer there are numerous factors that alter the likelihood of breast cancer, some more than others.

The risk factors that a woman is exposed to presumably alter the cellular environment and influence the course her breast cancer may take. The cellular environment will affect the proteome that will be observed at any given time. Understanding the variables involved is crucial to fully understanding the data that will emerge from a proteomic study such as this one.

Risk factors generally fall into two main types, genetic and epigenetic. Genetic risk factors can be further divided in two groups; germ line mutations and somatic

mutations. Epigenetic risk factors can also be further divided into lifestyle and environmental factors.

1.2.1 GENETIC IMPACT

Genetics of breast cancer, like most cancers, is complicated. There are many factors at play at one time and these will determine how a cancer will progress. This progression is determined by the genotype of the patient. Whilst up to 95% of genetic changes are due to a sporadic mutation there are several genes that have been linked directly to breast cancer.

Genetic mutations in the BRCA1 and BRCA2 gene on chromosomes 17 and 13 respectively, have been linked directly to increased breast cancer risk. The products of the BRCA genes, the BRCA1 and BRCA2 proteins in healthy patients are involved in the response to DNA damage and interact with each other as well as a myriad of other proteins involved in double stranded break repair (DSB repair) of DNA. A simplified overview of these interactions are outlined in Figure 1.2.1, taken from a review by Yoshida *et al* [9]. Mutations of key sites result in an inability of the BRCA proteins to perform their function which lead to an inability of the cell to repair DNA damage, hence replicating with faults in place, instead of damage being corrected, or apoptosis initiated.

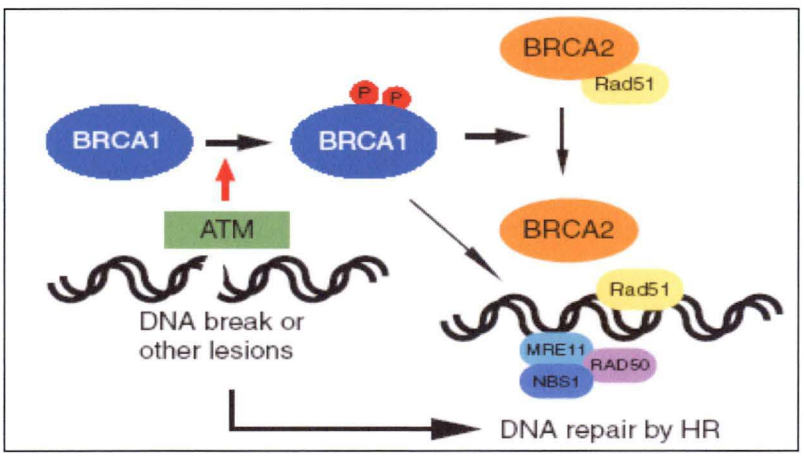


Figure 1.2.1: The role of BRCA in DSB repair. (taken from Yoshida et al[9]).

The presence of a germ line mutation in the BRCA genes increases the likelihood of breast cancer. BRCA1 gene mutations are reported to be associated with an 80% probability of breast cancer diagnosis before the age of 75. BRCA2 mutations are also highly indicative of a person's likelihood of being diagnosed with breast cancer

although not to the degree of BRCA1 mutations, with a lower probability of 30-50%. Mutations in the two BRCA genes are thought to contribute to up to 80% of familial inherited breast cancers, with approximate estimates of 20% of inherited breast cancers being due to undetermined genetic mutations. There have been several candidates over the last decade since the finding of the BRCA genes in 1993 but none have been found to be as prominent in breast cancer risk.

There may not be other highly associated genes to find, with many genetic mutations of low incidence, potentially making up the remainder of inherited breast cancers. Whilst the gene status of a person is important in elucidating risk, familial inheritance is only associated with 5-10% of all breast cancer cases, with sporadic mutations accounting for the remaining diagnoses.

Whilst the genetics of breast cancer will be dealt with in depth in section 1.3, sporadic mutations occur in genes that are responsible for the transcription of proteins involved in various functions including cell cycle check points, DNA repair, apoptosis, signal transduction, cellular adhesion and structural integrity. The causes behind the occurrence of DNA mutations in a majority of situations are stress on the replication and transcription processes. Somatic changes, those that are not inherited, may not directly cause cancer formation but they may allow for future mutations to go unnoticed or uncontrolled and may heighten the likelihood of a cancerous sporadic mutation.
Stress, a prominent cause of sporadic DNA damage can be inflicted in many ways and is often the result of epigenetic factors. Epigenetic factors that can influence the likelihood of DNA damage, and hence increase the risk of breast cancer are reviewed in the following section.

1.2.2 LIFESTYLE

The lifestyle or environment that a person is surrounded by can be almost as important as any genetic predisposition to breast cancer. It is highly probable that there is an intricate balance between the two. Reducing the environmental risk factors of breast cancer is a controllable aspect of breast cancer prevention, as opposed to the genetic aspects which are not controllable. Reducing the exposure to the following risk factors does not necessarily mean that you eliminate the development of breast cancer, but it

may help reduce the likelihood or prolong the time prior to onset of the disease, even in those with a genetic predisposition for the disease.

The best evidence for an environmental element of breast cancer is the increased rate following the migration of Japanese women, who traditionally have a very low incidence rate of breast cancer, to America, a nation of much higher incidence. The prevalence of breast cancer in these women and their descendents steadily increases through the generations to resemble the rates seen in American women rather than that of the Japanese women who remain living in Japan [10]. The reasons for these changes in incidence have been speculated to be associated with many factors including diet and exposure to hormones by way of the contraceptive pill and hormone replacement therapies.

1.2.2.1 DIET

Diet is one of the most controllable elements of a person's lifestyle. It can also have the biggest impact on an individual's health. It has been shown that a poor diet can lead to a plethora of diseases, cancer being alongside diabetes and heart disease as a major outcome of a poor diet. The problem however is in defining the specific elements of diet that result in the development and progression of these diseases.

The leptin receptor, which governs satiety is one of the major proteins that has been found to play a role in breast cancer through its role in activating transcription factors such as STAT 3 that up-regulate aromatase production and activate the Estrogen Receptor [11].

The intake of saturated fats is another identified risk factor which has been extensively researched and shown to increase the risk [12]. However, research has also shown that if saturated fats are eliminated an elevated likelihood of cancer onset is still present; this continued risk as being associated with the ratio of unsaturated fats. Studies from Greece have suggested that a balance between fatty acids is more important than total unsaturated fat consumption, with Omega 3 being "good" and Omega 6 being "bad" [13, 14]. This assumption is based on a surmise concerning the source of fats in the Palaeolithic era versus today and how the Greek diet has changed. This theory is supported when looking at the occurrence of breast cancers in multiple countries and the ratio of 3:6 Omega Fatty Acids. This difference can however also be tied to saturated

fats as omega 6 fatty acids tend to be associated with saturated fats, more often that omega 3 fatty acids. This has been supported by another recent study suggesting that an individual's "lipidome" is important, with the balance of all fats as opposed to the total levels of fats being what impacts cancer risk [15].

Other aspects of diet that have been brought into contention include the intake of antioxidants which mop up the DNA damaging free radicals. Antioxidants aid in protecting against stress at a cellular level. The greater the antioxidant level in the diet the more the body absorbs and the less DNA damage is done. Antioxidants are obtained through things like tea, red wine, and fruit [16] all of which have been studied to try and find the most protective foods. Soy, through the suspected actions of plant phyto-estrogens is postulated to be cancer protective, another possible but not yet well proven theory as is that of dairy consumption influencing the onset of cancer although this may be more reflective of total saturated fats in the diet, as dairy tends to be much higher in saturated fats [17]. Other aspects of dairy consumption is the potential chemical intake, with hormones such as Insulin-like Growth Factor -1 and bovine somatotropin absorbed by the milk from the food and environment of the cattle providing the milk for milk, cream, cheese and yoghurts.

The effects that diet can play in the development of breast cancer have been well documented, with many dietary elements already mentioned. There is one area of contention in the debate about diet and breast cancer however. The expressed proteome of the breast tissue is suspected to change in response to the environment, including dietary influences. This has been, until very recently, speculation with many dietary studies involving proteome changes focussing on the cells of the intestinal system and associated diseases [18-20] as well as the liver[21-25]. There has been evidence recently, though, to suggest diet can influence the proteome of the pectoralis major muscle in the chicken, indicating the ability of cells without a dietary function to be influenced at a protein expression level [26].

1.2.2.2 HORMONES

Hormones are an important factor when determining how to treat breast cancer and whether the treatment will succeed. Estrogen and Progesterone in particular are deeply embedded in the breast cancer puzzle. Estrogen and Progesterone levels, as well as their

receptor status are used to determine not only risk of breast cancer development but also the likely prognosis of the cancer and the treatment that will be administered.

In healthy tissue Estrogen binds to the Estrogen Receptor which through the actions of Estrogen Responsive Elements (ERE's) in promoter regions of DNA can potentially alter the transcription of nearly 300 different genes so far identified [27]. The products of these genes involve proteins that are involved in cell growth, differentiation and homeostasis. Figure 1.2.2 shows just some of the many interactions that the Estrogen receptor (ER) undergoes.

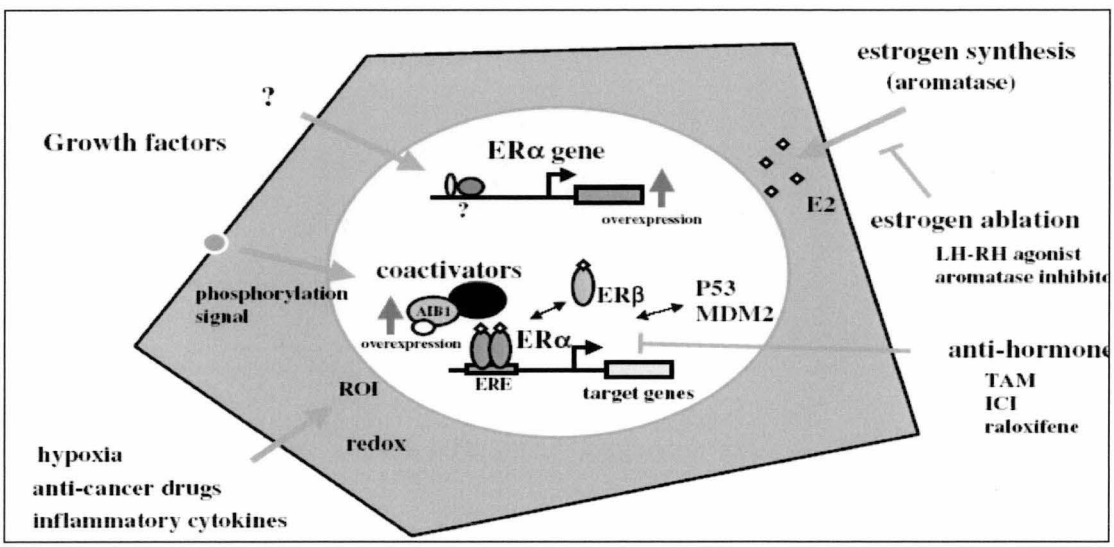


Figure1.2.2: Roles of e2 and its receptor. (taken from Hayashi *et al* 2003[27].)

Estradiol or E_2, the predominant form of the female hormone Estrogen, is the best characterised despite not yet being fully understood. E_2 is produced by the ovaries from menarche to menopause. The body maintains a constant cycle for 30-40 years, with the levels of E_2 fluctuating. It is known that a tumour will be either E_2 dependent or independent, with approximately 80% of breast cancers E_2 dependent [28].

This E_2 status is a crucial factor in treating breast cancer. For E_2 independent tumours therapies have revolved around indirect chemotherapies which act on the cell cycle, targeting rapidly dividing cells, such as those in the tumour which have little proliferative control. This however results in unwanted side effects in other rapidly dividing cells, such as hair follicles and intestinal lining cells, which lead to nausea, vomiting and hair loss. Adjuvant, targeted therapies have been the focus of much research. This led to the release of Herceptin, an antibody based targeted therapy which

targets HER2, a receptor that is not related to E_2 dependence, which is amplified in 20-25% of breast tumours[28, 29]. HER2 is involved in the control of cell growth, proliferation over expression results in over stimulated signalling of the cellular pathways in which it and the entire HER family are involved [30].

Reproductive history is a combination of risk factors which raise the levels of E_2 exposure over the course of a woman's life and increase the chances of developing breast cancer. The things that will alter the E_2 exposure include drugs such as the contraceptive pill [31]. Hormone Replacement Therapy (HRT) is another controversial field in breast cancer research. HRT was thought to be involved in increased breast cancer risk [32], however some research since has indicated that there is only a slight increase in the risk of breast cancer in women taking HRT [33-35] much of which can be lowered even more by careful selection of the HRT regime that is administered. Recent data on HRT suggests that the use of HRT results in lower production of E_2 and lower expression level of ER [36].

Conversely situations that decrease E_2 exposure such as breast feeding, a lower age at the birth of first child [37], the number of subsequent pregnancies, and higher age at menarche and menopause are all reported to be protective against breast cancer.

1.2.2.3 OTHERS LIFESTYLE FACTORS

There are several other factors of a person's lifestyle that may increase the risk of breast cancer. Socio-economic status has been tied to breast cancer with poorer women having higher incidence rates than those in higher socio economic classes. This has been linked to several things including, poorer education and worse access to screening programs [38-41].

Smoking has also been linked to breast cancer through the higher rates of exposure to carcinogenic chemicals. This has been debated, though, with some reports suggesting an anti-estrogenic effect implemented by the smoke, and hence conveying protection from breast cancer [42]. This protection would be limited to the duration of the persons being a "smoker" and may not negate the risks that are heightened by the other chemicals in the smoke.

1.2.2.4 ENVIRONMENTAL

The risks imposed by environmental factors are notoriously hard to prove. Many studies rely on retrospective studies and the reliability of the people participating.

Environmental factors encompass the exposure through day-to-day life in today's society that may be harmful and lead to the development of cancer. These factors include pollutants such as emissions from car exhausts and coal based chimney stacks that are heavily suspected to play a vital role in the presence of carcinogens in the local environment to lesser known risk factors that are emerging such as exposure to organo chlorides, bleach and ammonium based products used in day to day life in activities such as gardening and cleaning.

Exposure to radiation has also been shown to increase likelihood of breast cancer. Electromagnetic fields produced by transmission lines, TV's and household appliances; Ultra Violet light and sun exposure (UV radiation); and radiation from medical tests such as x-rays are also thought to increase overall cancer risk.

Studies linking environmental factors tend to be broad and aimed at cancer in general. One study by Johnson-Thompson and Guthrie (2000) showed that in the United States where African Americans and Hispanics have a higher rate of breast cancer than their Caucasian counterparts, the rate of breast cancer frequency has been linked to the number of hazardous waste facilities in that region [43]. It has been suggested that these two cohorts, African American and Hispanics, have a genetic pre-disposition for cancer occurrence. However in view of their low socio-economic status combined with these waste facilities, their lack of education and access to healthcare, when compared to that of the average caucasian American woman, it could be argued that environment not genetics is playing a more significant role.

1.3 THE GENETICS OF BREAST CANCER

The impact of genetics on human disease, especially cancer is undeniable. The exact role genes play is slowly being deciphered. Which genes are critical and which genes are altered as a consequence of other genetic changes are yet to be determined. The importance of the genes that are altered is a highly researched area and as more

information becomes available about their function and purpose, they become pieces of a very complicated puzzle.

Cancer is thought to have one of two starting points, a) an inherited or germ-line mutation, present in every cell in that individual, which over the individual's lifetime is highly likely to result in tumour formation or b) a sporadic, or somatic, mutation which is a spontaneous mutation causing a cascading effect of cellular changes. The latter will only be present in the daughter cells of the one that originally underwent mutation. Deleterious mutations of either type may end in cellular deregulation, and an uncontrollable cell cycle.

Mutations, both germ line and somatic, can be either silent, detrimental or beneficial and they can be dormant or unnoticed until future changes take place, which require the fully functional gene that has been mutated previously.

Cancer is a genetic disease but deciphering whether the alterations are common variations that occur in cancer, or susceptible genes, or whether the changes were present at birth is a difficult task.

1.3.1 GERM LINE MUTATIONS

Germ line mutations, also referred to as familial or hereditary mutations, had been suspected for many years before the isolation of the BRCA1 gene in 1993 at the locus 17q 12-21 [44]. Despite this discovery only 45% of familial breast cancers at the time could be attributed to mutations in this gene, leading to the discovery of BRCA2 at 13q 12-13 in 1995 [45]. More recent studies have suggested that ~ 50% of familial breast cancers can be attributed to mutations of BRCA1 and ~ 30% of mutations to BRCA2 leaving ~ 20% of inherited breast cancer based on familial pedigree without a specific associated gene or mutation at this time [46]. Others have suggested that combined, the two BRCA genes make up no more than 50% of incidence with 50% unaccounted for as yet [47]. Despite the identification of genes involved in breast cancer these only make up 5-10% of all breast cancers, with sporadic gene mutations composing the other 90-95% of breast cancer cases [2].

Despite the low incidence of familial breast cancers, known mutations are of high penetrance, estimates suggest upwards of 80% of patients with a known mutation in the

BRCA genes will have breast cancer in their lifetime [48]. Other genes such as CHEK2, PTEN and the CYP genes, which will be dealt with later in this section, are of low penetrance and hence result in cancer less often and require other mutational changes for them to affect the cellular regulation of the cell.

The BRCA1 and BRCA2 genes encode the transcription of the BRCA1 and BRCA2 proteins. These are involved in the regulation of double-stranded break (DSB) repair of DNA as well as the transcriptional response of other proteins such as p53 to DNA repair. They also maintain control of DNA damage at the cell cycle checkpoints. Their role in DNA surveillance, suggests in itself that mutations of this gene will result in critical cellular changes, allowing for the tumour to quickly grow in an uncontrolled manner. The specific mutation that a patient carries may determine the risk associated. Loss of the BRCA gene is embryonically lethal [49], therefore the gene must consist of at least one functional allele which will undergo deregulation at a later stage in development. BRCA associated tumours are more likely to be of high grade, with BRCA1 mutations being Estradiol (E_2) independent with normal HER2 expression, whilst BRCA2 is more likely to resemble spontaneously derived tumours and be E2 dependent [50, 51]. In a recent genetic sub typing of breast cancer, all BRCA1 mutations fell into a category that was distinguishable by basal cell similarities, lack of the Estrogen Receptor, and lack of HER2 expression and poor clinical prognosis. The BRCA2 mutations fell into sub type luminal A, a subtype resembling duct lining cells that give rise to a majority of breast tumour and have better clinical outlook [51].

The CHEK2 gene, also involved in cell cycle control, has recently been described as a possible familial gene mutation, estimated to cause up to 5 % of familial breast cancers [52]. It regulates checkpoint proteins and interacts closely with BRCA1. However the likelihood of the CHEK2 mutations found being a germ line mutation, as opposed to a spontaneous somatic change at this stage is contested. The role of CHEK2 as a familial breast cancer gene is in question with the publication of population based studies that show little difference between the normal population (1%) and the familial cancer patients (3%) in studies on multiple CHEK2 mutations from the UK, Australia [53] Finland[54] and the US [55]. It has also been shown in the Spanish population that the allelic mutation 1100delC, which has been focussed on most heavily in breast cancer, is present at an even lower rate, just 0.1% in over 400 families with familial non BRCA

related breast cancers studied [56]. This particular example shows how findings significant to one country may not necessarily be applicable to another population.

TP53, recently renamed and previously referred to as p53, is involved in many cell cycle pathways including apoptosis and it has various mutational forms [57]. TP53 is now referred to as a tumour suppressor and mutations can be inherited that will lead to Li Fraumeni Syndrome. A sporadic mutation is more likely to be a late event of tumourigenesis, and not an early change in the development of solid tumours[58].

Whilst the genes so far examined are cell cycle related, there are other genes that can be deleterious in other ways. If an inherited mutation is present in one of several metabolism related genes then breast cancer may arise. It has been shown in rodent models that mutations in genes CYP and NAT may induce cancer, by an inability to metabolise carcinogens properly, if the cells are exposed to them[59].

Other genetic polymorphisms have been suggested to lead to a predisposition to breast cancer, including the hormone related PTEN, CYP17 &19 [39]. Other publications have agreed with increased risk if mutations are present, but disagreed with the likelihood that such mutations are germ line mutations [60]. There are also carcinogen related CYP1A1, NAT 1 and NAT 2 genes, although while they promote tumorigenesis if it begins they do not directly initiate the process.

1.3.2 SOMATIC MUTATIONS

The search for germ line mutations began as an attempt to locate common genetic changes in breast cancer despite apparent familial inheritance being low. What was found was that the genes which are commonly mutated in familial cases of breast cancer are not commonly mutated in sporadic cases. Even sporadic BRCA1 mutations do not occur in the same location as those that are inherited [58]. Hence the search for genes independent of inheritance began.

Sporadic mutations are possible in any gene although some regions of some genes are more susceptible than others. Genetic mutations can be caused by many factors including UV exposure from the sun, radiation exposure from various radioactive sources and free radical exposure, mostly initiated through diet or stress, all of which cause DNA damage. DNA damage can be present in many forms. Single point

mutations in a critical codon of the gene may cause amplification of protein production or truncation of the transcribed protein. Mutations may also involve the deletion of an entire allele resulting in the complete absence of downstream proteins.

Once the sporadic mutations were located, determining whether the changes were a primary, causative event in tumourogeneis or a secondary, consequential change due to other mutations became essential.

Damage that results in cancerous changes is likely to be present in proliferative or DNA repair genes which will initiate the early stages of tumour formation. Figure 1.5.1 demonstrates where these proteins act within the cell cycle and demonstrates why they are so crucial.

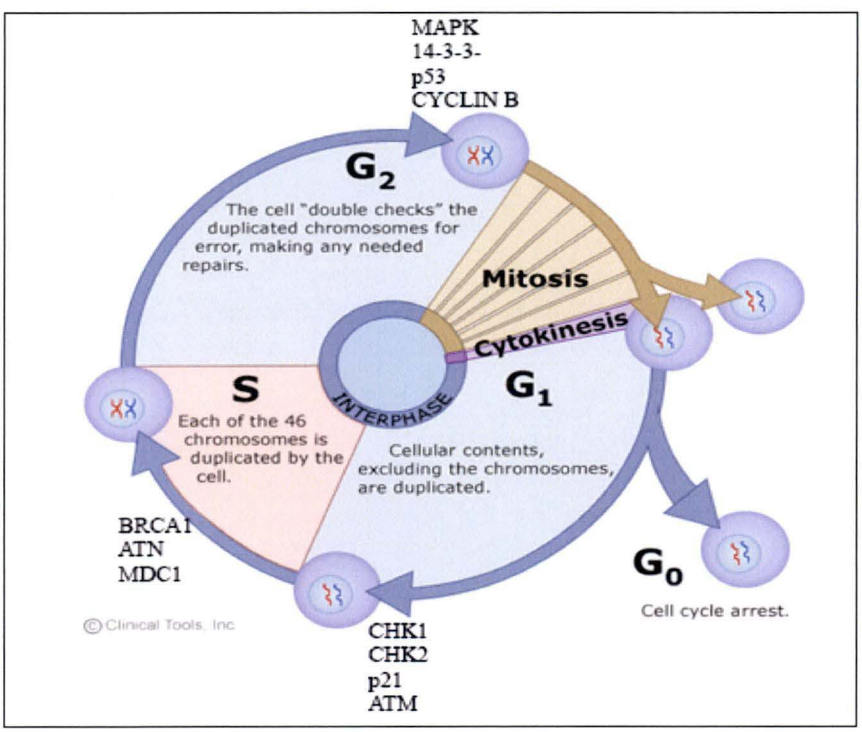


Figure 1.5.1: Key proteins operate and the cell cycle checkpoints. Noted are just a few proteins, that when damaged, de-regulate the cell cycle.

Whilst theoretically mutations can happen in any gene, common mutations, which are not germ line mutations that are found in breast cancers, have been found in the PTEN, TP53, ER, HER2, MYC and CCND1 genes.

The mutations that occur in these genes may cause aspects of the cell cycle to spiral out of control. Any mutation may cause a predisposition to cancer susceptibility due to

silent mutations in crucial genes. This may happen by the activation of a silently mutated oncogene such as HER2 or the loss of function of mutated tumour suppressor genes such as p53.

Other genes that are emerging as being important in breast cancer are the metabolic CYP and NAT genes which encode variants of Cytochrome p450 and N-acetyltransferases respectively. These genes produce proteins that are crucial in the metabolism of PAH's (Polycyclic Aromatic Hydrocarbons) and HAA's (Heterocyclic Aromatic Amines), both of which are common forms of carcinogens as well as hormones such as E_2. Mutations in CYP1A1 can alter the hydroxylation of E_2 to 2-hydroxyestradiol and favour CYP 1B1 and form 4-hydroxyestradiol, a carcinogen itself and up regulated in breast cancers [61]. There have been significant correlations between the presence of mutations in these genes increasing risk of breast cancer if the woman is a smoker and Caucasian [62]. Most CYP studies have involved Caucasian cohorts, however one study looks at a Japanese study and finds it to also be correlated to increased risk in Japanese women. Whether this can be extended to Asian women as a whole is yet to be seen. [63]. However one report suggests that there is a similar trend in Taiwanese women [64].

Gene modification of HER2 results in a constitutive up-regulation of the HER2 protein [65]. HER2 is a Tyrosine Kinase receptor that heterodimerises with other members of the HER family and stabilises them, causing constitutive activation of the target pathways. This constitutive activation results in deregulation of the cell cycle and is seen in 30% of breast cancers [66]. As outlined when discussing the risk factors of breast cancer, it has been possible to develop a targeted therapy towards the HER2 protein. This therapy is known as Herceptin, which binds to the receptor preventing binding and stabilisation of both heterodimers and homodimers and blocks downstream signalling cascades, negating the up-regulation of the gene and slowing down the process of tumourigenesis and progression.

1.4 PROTEINS INVOLVED IN BREAST CANCER

The study of a disease such as breast cancer involves a very broad range of fields. It includes better imaging in the form of advances in mammography, epidemiological studies, to find local trends and commonalities between the types of cancers and

potential risk factors, genetic studies to find linkage between occurrence and mutations, as well as proteins studies to examine how these genetic findings influence the cellular environment. These are just a few areas of research aimed at increasing the knowledge of breast cancer. In protein studies, it is important to first establish the role of a given protein in a healthy condition. Following this, it is then required that the alteration of this protein is found to know both how and why the regulation of this protein is dysfunctional and how to overcome or re-establish appropriate regulation. The data generated by the 'genomic era' has led to a huge number of hypothetical proteins which are poorly defined and little is known about their role within the cell. Now that there is data on the existence of these proteins and their amino acid sequence they are becoming more commonly recognised as being involved in disease through the ever expanding techniques of proteomics. The recognition of these proteins' involvement has led to a huge expansion in the amount of knowledge that can be obtained about these little known proteins and an increase in the functional studies occurring to determine their importance.

1.4.1 TUMOUR SUPPRESSOR GENES

Tumour suppressor genes are vital in the maintenance of normal cell cycle function. When these genes are mutated or lost then the protein that they encode, the tumour suppressor protein is mutated or no longer expressed by the cell. Effectively the "brakes" are taken off and the suppression that that protein performed is removed and the cell is able to proliferate in an uncontrolled manner. The genes that are recognised as commonly mutated in cancer include p53, Rb, and BRCA1.

1.4.2 ONCOGENES

Oncogenes, in contrast to tumour suppressor genes which cause a loss of function, when mutated cause the related oncoproteins to be over expressed, or expressed at inappropriate times. This results in an upset in the cells internal balance and effectively excessive production of a protein that may promote proliferation or be anti-apoptotic that is now present in very high levels constantly.

Oncoproteins include several proteins but most recently HER2 is the oncogene of interest. They also involve c-myc and Cyclin D1 and many members of the Heat Shock Protein Family.

1.5 PROTEINS IN BREAST CANCER

The complexity of breast cancer means that there are a number of pathways involved in the progression of breast cancer. The nature of cell biology results in most pathways being interconnected and at times interdependent. It is for these reasons that advances in the fight against cancer are slow and tedious. There are two key categories of proteins when discussing cancer, tumour suppressor proteins, where the normal expression is able to prohibit tumour formation by regulating key elements of cell cycle control and oncoproteins whose function, when mutated causes up regulation and hyperactivity of proliferative pathways. There are key functions of the cell that are commonly altered in the formation and progression of breast cancer and these key roles will be discussed.

1.5.1 PROLIFERATION AND CELL CYCLE CONTROL PROTEINS

Changes in a cells proliferation are one of the key functions that will determine the formation of tumours. Believed to be one of the first changes required for tumour formation, proliferation is the main function that mutations in tumour suppressor genes will influence. Tumour Suppressor Proteins such as p53 and BRCA1 are altered in such a way that allows the cells to proliferate without undergoing checks that involve correct DNA organisation and repair. Other key proteins in the regulation of the cell cycle include CHEK1and CHEK2 (cell cycle check point proteins) RADD 50 and RADD51 and GADD45, which are DNA repair proteins downstream of p53 and BRCA1. The proteins which would usually induce apoptosis under the right signal, such as the Tumour Suppressor Protein ATM may no longer be present or responding to the correct signals which results in uncontrolled proliferation of a cell.

1.5.2 ADHESION PROTEINS

Cellular adhesion proteins are important in the maintaining of a singular cancerous mass. When cellular adhesion begins to be deregulated, cells start to lose proteins such as cadherins and intergrins which allow the cells to detach from the tumour and begin

circulating in the blood stream. These cells may eventually become lodged in the smaller capillary networks which are often found in the brain or lung, where they become entrapped and continue to proliferate, resulting in metastases.

A group of proteins termed Metastasis Suppressor Proteins have been identified as key elements in the spread of breast cancer. These include proteins regulated by the CAD1 gene such as E-cadherin and the novel protein BRMS (breast metastasis suppressor 1). These are among the proteins attributed to the cells ability to dissociate from the tumour itself and induce the spread of disease [67].

1.5.3 APOPTOTIC PROTEINS

Activation of programmed cell death can occur through a variety of methods. Both mitochondrial stress and p53 can activate the caspase pathway via caspase 9. Fas and death receptors activate the caspase cascade through caspase 8 and 10. All lead to an array of changes within the nucleus that promote its DNA's breakdown and prevent cellular proliferation [68-70].

Apoptosis or cell death should occur if the cell is functioning irregularly and is properly interconnected with all other pathways, with, if you like fail safe modes, for when things go wrong the cell should self destruct. In cancer this doesn't happen. This can be due to either the over expression of anti-apoptotic factors or to the failure of pro-apoptotic factors. Either way the result is the same, the cell survives, only to continue proliferating.

1.6 PROTEOMICS

1.6.1 HISTORY OF PROTEOMICS

In 1995, a paper was published by Australian scientists in "Electrophoresis" that described the creation of a protein map, using techniques that would separate proteins and reveal what they referred to as the sample's "proteome" [71]. The term was quickly adopted and research articles since 1996 have referred to it as such ever since. In 1998 a review was published which referred to the study of the proteome as Proteomics [72]. Since that time, proteomics has grown from a novel approach to a field of science that is widely used and has captured the attention of researchers around the world.

This "new" science is the result of many decades of research and techniques coming together. Initially proteomics involved the combination of 2D SDS PAGE, a technique from the 1970's and Matrix Assisted Laser Desorption Ionisation – Time of Flight Mass Spectrometry (MALDI-TOF MS), founded in 1992 [73], but based on Mass Spectrometry techniques which have been used since 1955 [74]. These were coupled with the modern advances in image analysis and bioinformatics. Since then advances in all areas of proteins separation and analysis have been ongoing and have broadened the scope of proteomics in research.

The largest factor in the long term survival of this novel approach was the completion of the genome project. This allowed for the creation of searchable databases of all proteins based on the theoretical information derived from the human genome project that predicts the presence and sequences of all proteins from all genes. This almost absolute coverage of the human genome meant that regardless of what protein was isolated, it would be traceable to a gene on a chromosome, with the ability to predict function and downstream modifications.

The field of proteomics has emerged as a major science and as such there are now many techniques for separating proteins and examining a samples proteome. This review will deal primarily with the path of gel based proteomics and the advances in this field as this is the techniques that this study has adopted. Other avenues of proteome examination include mostly Liquid Chromatography (LC) based proteomic techniques however there is an expanding interest in *in situ* proteomics as the technology becomes available.

Advances that have influenced the progress of proteomics include many aspects of protein separation. Iso-electric focussing is the first dimension of the two dimensional separation process and the emergence of the IPG (Immobilised pH Gradient) enhanced the practicalities of this step. Whilst this development was well underway before the field of proteomics as we now know it began, it was first described in 1983 [75], it is a critical advancement of 2D protein separation This not only allowed the process to be simplified by replacing tube gels that had a pH buffer passed across it to allow for separation by pI, but also allowed for bulk manufacturing and quality control, leading to better reproducibility of the 2D image [76].

Reproducibility is one area that proteomics was heavily criticised for in its early stages, as the capturing of a dynamic system at one point in time by such sensitive measures as Mw and pI can be hard to reproduce. Hence many replicates of gels were needed to assure that what was seen was in fact a true reflection of the sample characteristics and not an artefact of gel-to-gel variations.

Given this need for reproducibility one of the next steps was the commercialisation of pre-cast polyacrylamide gels. Pre-cast gels have been referred to in the literature since 1977 [77], but the manufacturing and marketing of pre-cast gels by biotechnology companies made 2D proteomics more accessible and reliable. Since the pre-cast gels were placed on the market refinements have been made, offering different gradients densities to target particular ranges of Mw, increasing the size of the gel systems to increase the separation and resolution of the proteins and to stabilise the acrylamide, to lengthen shelf life of the gels and create less wastage of research resources.

Alongside these separation techniques were the advances in sample preparation. The treatment of samples to produce a sample suitable for the separation by IEF and 2D PAGE is constantly advancing. IEF posed many problems in the early stages as it required samples to have a charge of 0. Many uncharged proteins precipitate and hence solubilisation was a major issue to overcome. Breaking open cellular compartments with a non-ionic detergent also required intense research to assure membrane breakdown, without denaturation. Samples consist of very different make ups depending on their source. Some will be highly glycosylated, or have high albumin content, others will be rich in lipoproteins, others will have high salt content and all of these will affect the success of a 2D image.

Improvements in the stains available for 2D gels were the next steps to improve. The common older stains such as coomassie blue have lower sensitivities to protein and as such are only visible for protein concentration of ~10ng or higher. Silver stain, although it was more sensitive, visualising down to 1ng of protein, in it's classical formulation was not MALDI-TOF MS compatible. Fluorescent stains evolved, specifically for proteomics, including Sypro Ruby, and Flamingo Pink, and silver stains that were MS compatible were also devised and marketed accordingly. Whilst many researchers in 2D proteomics choose an IPG strip and ready cast gel to conduct their studies, the stain is

where the first major variation occurs and much of this has to do with the costs involved in obtaining fluorescent imagers over densitometers that will read silver stain and coomassie stains. As the usage becomes more common costs are dropping and the more sensitive stains like Sypro Ruby (down to 0.1ng) are now accessible to smaller institutions as well.

Advances since the inception of proteomics have also been happening in the field of Mass Spectrometry. MALDI-TOF based MS, was the original MS used for proteomic applications.

This process cleaves trypsinised proteins in to peptides. These peptides are then bound to a matrix which is presented to a laser which ionises the peptides. Upon ionisation/excitation of the peptides they leave the matrix and travel througha vaccum where at the opposite end is a sensor, measuring the time taken from excitation to reach the sensor. This produces a 'time of flight' for each peptide. The time of flight is correlated to the size of the peptide, with the larger peptides being slower to move when ionised. The size and the number of peptides will be unique for each protein. It is this combination of peptide information which has been termed a proteins "peptide mass fingerprint" or PMF.

This is where the genome project is applied to proteomics. The genome project has enabled the formation of peptide databases, such as MASCOT, that theoretically trypsinise all of the proteins in the genome. From the PMF these databases can provide identification of a protein, or they can postulate a gene location and function if it is previously unknown.

1.6.2 PROTEOMICS IN BREAST CANCER

Very early on in the field of proteomics it's potential application in breast cancer research was discussed [78]. Several publications, mostly reviews, addressed the theoretical practicality of such a technique in the world of medical research and more specifically cancer research. The possibility of finding bio-markers for breast cancer was addressed in several of these early publications [79-81]. Whilst the theory was in place, it took until 2001 for the first paper identifying a novel protein to be published by Le Naour *et al.* which detailed the protein RS/DJ-1 as a potential circulating marker [82].

1.6.3 COMPARATIVE STUDIES

In 2002 Pucci *et al* [83] published the first breast cancer protein map, and differential studies into changes between healthy and diseased tissues began. Several studies since have used a differential approach to ascertain novel proteins that may be involved in breast cancer [84-87]. All of these studies differ in quite significant ways, resulting care needing to be taken when the comparison of one study with the other occurs. The proteins that were isolated by each study can be found in Table 1.7.1.

The first list of potential biomarkers was published by Wulfkhele *et al.* in 2002 [87]. Wulfkhele *et al.* used 6 patients, obtaining both a tumour sample and a matched healthy tissue sample from each patient. The samples were DCIS (n=4), or DCIS associated with IDC (n=2). The tissue was either extracted whole or fixed, sectioned and then underwent Laser Capture Microdissection (LCM) to result in the final working sample. 2D protein separation was conducted using IEF and PAGE, using both Sypro Red, a sensitive fluorescence based stain and coomassie blue a comparatively insensitive stain. Proteins were determined to be of interest based on the intensity of staining. Wulfkhele *et al* found 57 differentially expressed proteins ranging from cytokeratins to chaperones which can be found listed in Table 1.7.1.

Proteins	Wulfkhele *et al*	Somiari *et al*	Luo *et al*	Deng *et al*
(LKEEYQSLIR)	✓			
14-3-3 ζ/δ	✓			
78kDa Glucose related protein (GRP78)/ Ig heavy chain binding protein (BIP)				✓
Actin, beta	✓	✓		
Actin-related protein 3 (Arp3)	✓			
Aldo-keto reductase family member 1, member a1			✓	
Alpha 1 antitrypsin				✓
Alpha-1-antitrypsin x2		✓		
Annexin 1			✓	
Annexin II	✓			
Annexin II	✓			
Annexin V	✓			
Annexin VII	✓			
Anti-thrombin III	✓			
Apolipoprotein - 1				✓
Biliverdin reductase B	✓		✓	
CapZ	✓			
Carbonic anhydrase I: carbonic dehydratase		✓		
Cathepsin B	✓			
Cathespin D				✓
Cellular retinoic acid-binding protein 2 (CRABP2)	✓			
Chain b of recombinant human fibrinogen fragment			✓	
Coactosin-like protein	✓			
Cofilin	✓			
Cofilin 1 non-muscle			✓	
Cycophilin a			✓	
Cycophilin b precursor			✓	
Cytokeratin 18(424 AA)		✓		

Proteins	Wulfkhele *et al*	Somiari *et al*	Luo *et al*	Deng *et al*
Cytoskeletal regulation	✓			
Destrin	✓		✓	
Dimethylarginine dimethylaminohydrolase (DDAH)	✓			
Elongation factor 1α	✓			
Elongation factor beta (ef-1-beta)				✓
Endoplasmic reticulum associated amyloid beta peptide binding protein			✓	
Eukaryotic translation initiation factor eIF3h	✓			
FABP-3 (HFABP)	✓			
FABP-5 (EFABP)	✓			
Fatty acid binding protein 4			✓	
Fibrinogen gamma		✓		
FK506 binding protein	✓			
FLJ10652 (KDVKPHPR)	✓			
GDP-mannose phosphorylase B	✓			
Gelsolin		✓		
Glycerol-3-phosphate dehydrogenase 1 (soluble)		✓		
GRP78	✓			
GRP94	✓			
HnRNPs A2	✓			
HnRNPs A2/B1	✓		✓	
HnRNPs A3	✓			
HnRNPs C	✓			
HnRNPs L	✓			
HnRNPs X	✓			
Hsp 27	✓			
Hsp 90	✓			
Human carbonic anhydrase 1			✓	
IG heavy Chain				✓
Ig Light Chain				✓
Intracellular chloride channel 1 (NCC27, CLIC1)	✓			
Lactate dehydrogenase a			✓	
Lactoferrin	✓			
L-Plastin	✓			
Lumican		✓		
Macrophage capping protein (CapG)	✓			
Manganese SOD			✓	
MGC: 22710			✓	
Nascent polypeptide-associated complex, a polypeptide	✓			
Peroxiredoxin 1			✓	
Peroxiredoxin 1 (Thioredoxin peroxidase)	✓			
Peroxiredoxin 2	✓			
Peroxiredoxin 5	✓			
Phosphoethanolamine binding protein (PEBP)	✓			
Phosphoethanolamine-binding protein (PEBP), prostatic-binding protein	✓			
Phosphoglcerate mutase 1			✓	
Phosphoglycerate kinase			✓	
Phosphoribosylaminomidazole coarboxylase			✓	
Polymeric immunoglobulin receptor (pIgR)	✓			
PRO2619		✓		
Procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase)		✓		
Profilin 1	✓		✓	
Prohibitin	✓			
Proteasome (prosome, macropain) activator subunit 1 (PA28 alpha)		✓		
Proteasome subunit alpha type 1 (PSMA1)				✓
Protein phosphatase 2A (PP2A)	✓			
Pyruvate kinase, M1 isozyme (EC 2.7.1 40)		✓		
Rab 11a	✓			
Rho GDI	✓			
Ribosomal protein s12 (RPS12)				✓
S100A11	✓			
S100A7	✓			
S100A8	✓			
Selenium binding protein	✓			
Serum Albumin				

Serum albumin x 2		✓		
Similar to Peroxiredoxin			✓	
Similar to trisephosphate isomerase			✓	
Similar to tubulin, beta 5		✓		
Small ubiquitin-related modifier 3 precursor (SMT3a) (SUMO3)				✓
Stathmin	✓			
Superoxide dismutase (SOD)				✓
Transgelin (Acidic)	✓			
Transgelin (Basic)	✓			
Transgelin 2				✓
Translationally controlled tumour protein (TCTP)				✓
Transthyretin				✓
Type VI collage n -1 chain	✓			
Type VI collagen_2 chain	✓			
Ubiquinol-cytochrome reductase complex core protein 2			✓	
Ubiquitin-conjugating enzyme 9 (Ubc9)	✓			
Uracil dna glycosylase			✓	
Voltage dependent anion channel 2			✓	
Voltage-dependent anion Channel protein 1 (VDAC)	✓			
α B-crystallin	✓			
α-Enolase	✓			

Table 1.7.1: Comparative proteomic studies in to breast cancer conducted over the past 5 years. [84-87].

Following the publication by Wulfkhele *et al* in 2002 was a publication in 2003 by Somiari *et al* [86]. Somiari used a different approach, using DIGE (Differential In-Gel Electrophoresis) as a method for comparative analysis between healthy and diseased tissue. Based on a Caucasian population of Americans, 4 samples of IDC were used, with reduction mammoplasty forming the basis of healthy tissue. Somiari *et al* found anywhere between 25 and139 proteins up regulated by >3 fold and 60 to 196 down regulated proteins in the samples examined. They listed 27 proteins, selected on their consistency across all samples, as being potentially significant, these can be found in Table 1.7.1.

Alongside the tissue based comparative studies were serum based studies, searching for circulating biomarkers. First proposed in 2001 by Kennedy *et al* [88], the first research in this area was published in 2002 by Li *et al.* [89], and followed by Rui *et al.* in 2003 [90]. Li *et al.* used SELDI–TOF MS (Surface Enhanced Laser Desorption Ionisation – Time of Flight MS) to determine three significant MS peaks that were different between the 41 healthy and the 103 cancerous patients that were recruited for the study. Rui *et al.* used a more traditional approach but was able to determine from 52 healthy samples and 76 cancerous samples, using 2D based electrophoresis that both the up regulation of HSP27 and the down regulation of 14-3-3σ sigma may have potential for determining malignancy, based on their combined expression pattern.

In 2004 there was a heavier emphasis in the literature on novel concepts using alternative body fluids as a source of potential biomarkers and protein profiling.

Alexander *et al* [91] looked at the potential of nipple aspirate to provide novel prognostic biomarkers for breast cancer. The benefits of this alternate approach are reflected in the number of samples recruited for the study, with 52 diseased samples and 53 healthy samples obtained. The lower complexity of sample resulted in lower numbers of proteins on the gels produced which allowed for determination of the entire proteome of nipple aspirate and observations of differential expression by eye in 10 different proteins to be adequate for analysis.

Hudelist *et al*, [92] published the first attempt to create a diagnostic and prognostic protein profile for breast cancer. There were some limitations in this approach, which was based on the protein chip and labelling slides with antibody targets, similar to the DNA micro array concept, but with proteins. Also the profile published was based on the findings from one patient, whom provided the healthy and the cancerous sample, although it did demonstrate the potential of the technology for future studies on a larger scale.

This was followed in 2005 by Pawlik *et al* [93] who did away with the use of gels and instead conducted direct MS using SELDI-TOF (surface enhanced laser desorption time of flight analysis). This is a process that results in large amounts of data due to the role that bioinformatics plays in this approach.

Also in 2005 Luo *et al.* published another comparative proteomics based paper comparing healthy to cancerous samples, and this was based on a Chinese population, 4 women were recruited, 2 healthy and 2 with ductal carcinoma [85].

2006 saw the publication of many follow up papers, Hudelist *et al* published a more extensive paper on the promise of protein profiling in breast cancer [94]. Moving away from the original protein chip technology towards direct analysis via a 2DE gel, Hudelist *et al* identified 32 differentially regulated proteins, 13 previously unconnected to breast cancer. The prognostic ability of these proteins, following extensive confirmation was used to identify an independent cohort of 50 breast cancers.

Pawlik *et al* in 2006 published a paper, again on the proteomics of nipple aspirate, but this time focussed on the importance of early diagnosis of disease through the nipple aspirate proteome [95], identifying Vitamin D binding protein as a potentially up regulated marker in nipple aspirates from women with breast cancer.

1.6.4 TRANSLATIONAL STUDIES

Whilst the comparative studies between healthy and cancerous have been broadly dealt with other applications of proteomics in breast cancer research has also progressed in the field of translational research. By using proteomics to analyse the effect that an agent, which may be anything from a therapeutic drug through to radiation through to growth factors, may have on breast cancer cells, it is possible to decipher reactions and pathway involvement that may have otherwise been missed. Whilst the proteome map of breast tissue was being developed in the early parts of this decade, similar studies on cell lines were being conducted with MCF-7 studies published in 2002 and T47D in 2004 [96, 97]. The 2002 paper by Chen *et al* saw the beginning of cell lines as tools for translational research in breast cancer, when Chen and colleagues described the protein profile change between Doxorubicin treated and untreated cells. Whilst Chen *et al* used the MCF-7 proteome to describe protein changes, the differences between non malignant breast epithelial cells and MCF-7 cells was not published until 2005 [98].

These cell lines, along with others have been used to develop the knowledge of known systems in breast cancer as well. Studies involving cell lines have revealed that new information about proteins such as IFG, HER2 [99] and 14-3-3σ.

14-3-3σ is a good example of the use of cell lines to determine function and significance in a given cell type. Originally, Laronga *et al.* discovered 14-3-3σ as a protein of interest in breast cancer by using genetic based over expression studies [100]. Moreiera in 2005 showed using a proteomic technique that 14-3-3σ is down regulated sporadically and is not a common event as was first thought to be the case in breast cancer [101]. In 2006, also via proteomics, 14-3-3σ was described by Liu *et al.* to be involved in drug resistance, with up regulation observed in the MCF-7 derived Adriamycin resistant cell line [102].

Drug resistance is another field where translation proteomic research has proven useful. Since 2004 there has been an increase in the literature published that revolves around translational studies on cell lines. Brown *et al* in 2004, also using MCF-7 derived cell lines for both susceptibility and resistance to Doxorubicin, found multiple proteins via LCMS to be potentially involved in drug resistance, with Metallothionein, Ubiquitin

related proteins and Cofilin based proteins, most frequently deregulated in multiple isoforms [103].

Anti-estrogen resistance has also been heavily looked at through functional proteomic studies using cell lines. Huber *et al* in 2004 detailed resistance in T47D cell lines, describing 44 proteins that were differentially regulated between resistant and sensitive cell lines[97].

Besada *et al* in 2006 looked at Tamoxifen resistance in breast cell line 3366, with 2Dgel analysis revealing 21 proteins that are differentially regulated between the resistant and sensitive cultures. Of these 21 proteins, Peroxiredoxin 1 and 4 were described, this family of proteins is one that is becoming the focus of much attention with Peroxiredoxin 2 found to be involved in the radiation resistance of breast cancer cell line MCF-7 by Wang *et al.* in a proteomics based study in 2005 [104].

The amount of translational data in recent years has increased rapidly. This explosion of publications has led to the need for a "bank of information". This has been addressed in small part by the online access of a breast cancer proteomics database maintained by the Danish Centre for Translational Breast Cancer Research (http://proteomics.cancer.dk/). Two protein separation techniques, Iso-electric Focussing and NEPHGE (non-Equilibrium pH Gradient Electrophoresis), for the 1st dimension have been used to present two protein maps for breast adipose tissue. The images published are dependant upon the sample preparation, and hence difficult to reproduce, and the identification of just 25% of the proteins in the protein map highlights the issues involved in effective separation and identification.

1.6.5 ADVANCING TECHNOLOGIES

Mass Spectrometry technologies are advancing at a rapid rate. The sensitivity of the machinery is improving beyond comprehension, with the biggest advances coming in the identification of the small peptides at low concentrations. Also of note in relation to MS is the development of techniques to enable MALDI-TOF MS on peptides digested *in-situ* (Figure 1.6.1). The continued development of which might lead to the need for sample separation to become a thing of the past. These are currently developmental however and tend to be out of reach of the financial budgets of most institutions.

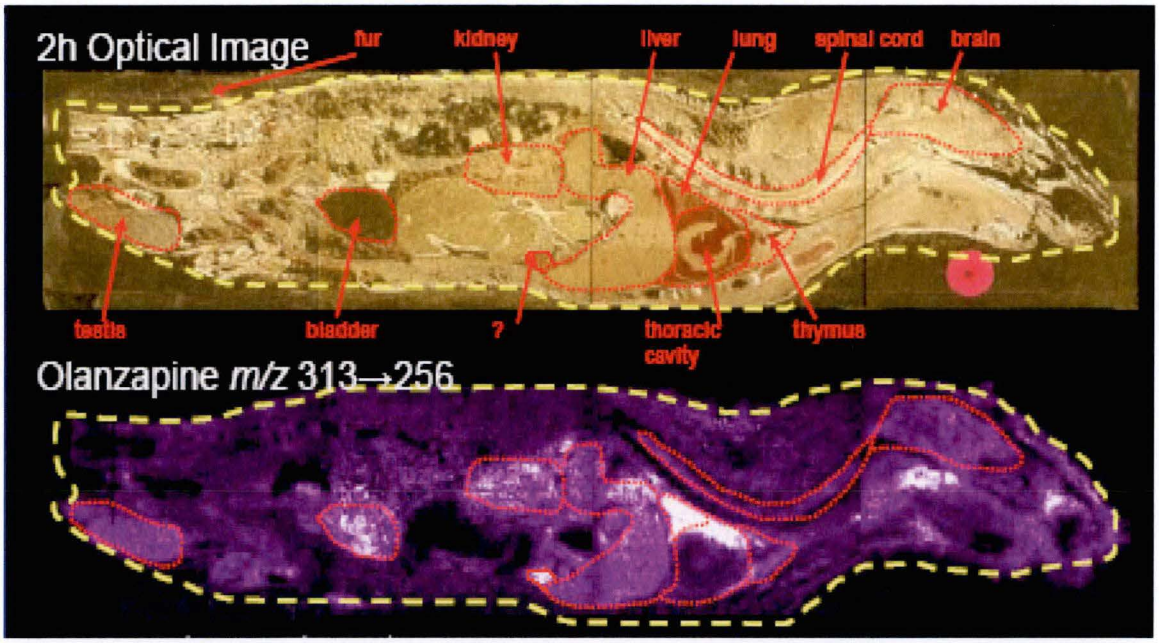


Figure 1.6.1: Sagittal section of a rat, and the comparative *in situ* MALDI-TOF image achieved by the selective highlighting of the peptides within the drug, 2hrs post treatment. This shows that the drug is directed strongly toward lung tissue, followed by traces in the excretory system of the kidney and bowel, but also highlights unexpected findings in the testis and an organ that was unidentified. http://www.uab.edu/proteomics/pdf_files/2006/Mobley_Proteomics_Class.pdf

Other improvements that are increasing options involve Laser Capture Microdissection (LCM) which has the ability to take individual cells from a sample and analyse these specifically. As described, early studies required 50 000 – 100 000 cells to obtain enough protein [87], but with the accuracy of the technique and the improved MS technologies, proteomics studies from just 3000 cells are emerging [105]. For the study of breast cancer improvements like this are very beneficial. It removes the need to use whole tissue sample, and instead enables analysis of cells on a very specific basis, for example the periphery v's the interior of the tumour. It also allows selection of cell type, epithelial cells from adipocytes, a cell type common in breast tissue. Limitations of this technique are concentrated on the low numbers and therefore low protein concentrations this limitation is however being aided by the more sensitive MS's on the market such as Fourier Transform Mass Spectrometers.

Finally the other major advances, without which proteomics would never have occurred, are the biotechnological advances, namely the handling and storage of data.

1.6.6 CURRENT LIMITATIONS

Proteomic research into breast cancer over the last decade has, despite a slow start, progressed and diversified rapidly. There are some key areas of the work to date that need to be addressed.

The sample size for studies that involve tissue specimens is very low, some have as few as two cancer specimens. All studies, except those out of the United States, are of single ethnicity, and whilst there are several attempts to utilise the techniques for early stage detection, few have applied it to early onset, in younger women.

Whilst there are many factors that limit the studies, the most prominent difficulty is the diversity of techniques that are applied in the field of proteomics, making determination of differences, and then confirming the reports, potentially very difficult to complete. Whilst the public access database by the centre for translation breast cancer research in Denmark is a step in the right direction for accessing information about the proteins involved in breast cancer, this is limited, and always needs to be considered in respect to the techniques used to obtain the image. If samples are whole tissue, LCM derived, LC – gel free analysed, or of cell line origin are all elements that may produce conflicting data between proteomic studies, stressing the requirement for reproducibility within each individual study.

2 Project aims

The first aim is to isolate and identify proteins that are significantly up or down regulated when comparing cancer tissue from each population to a healthy control set of breast tissue. Proteins resulting from this analysis will lead to the development of new targets for diagnostic, prognostic and therapeutic applications that are applicable to multiple populations from around the world.

The second aim is, using the same set of samples, to isolate and identify proteins that are significantly altered between early and late onset cancers. This is in order to identify key proteins that may be associated with the increased aggressiveness that is reported in breast cancer in younger women.

The healthy cohort comprises two groups of patients, those who have hade breast reduction surgery and those who have had prophylactic surgery due to previous diagnosis of breast cancer. The third aim is to conduct a pilot study into whether any early detection markers for breast cancer exist in macroscopically healthy tissue from patients that have a history of breast cancer.

IGF-1 and β-casein, both of which are regulated through the same transcription factor STAT5 are constantly contradicted in the literature with regard to breast cancer involvement. The final aim is to take a functional approach to breast cancer proteomics in order to identify whether there is a relationship between these two proteins that have been controversially linked to breast cancer in the past.

3 MATERIALS AND METHODS

3.1 CHEMICALS LIST 36
3.2 SOLUTIONS AND BUFFERS 37
3.3 SAMPLE COLLECTION 42
3.4 SAMPLE PROCESSING 42
3.5 WESTERN BLOTTING: 44
3.6 CELL CULTURE 46
3.7 STAINING 49
3.8 ANALYSIS 50

3.1 CHEMICALS LIST

Chemical	Company
Acetic Acid	BDH Chemicals, Australia
Acrylamide	Bio-Rad Laboratories, CA, USA
Agarose (Low melt electrophoresis grade)	Bio-Rad Laboratories, CA, USA
Ampholytes	Bio-Rad Laboratories, CA, USA
APS (Ammonium persulphate)	Sigma Chemicals Company, USA
β – mercaptoethanol	Serva, NY, USA
Bromophenol Blue (Tracer Dye)	Sigma Chemicals Company, USA
CHAPS (Non-ionic detergent)	Bio-Rad Laboratories, CA, USA
Criterion Gel (pre-cast, 11cm, 4-15%)	Bio-Rad Laboratories, CA, USA
DTT (Dithiothreitol)	Bio-Rad Laboratories, CA, USA
ECL (Enzyme Chemiluminescence) Kit	Amersham Pharmacia Biotech, UK Limited
EDTA (Ethylenediaminetetraacetic acid)	Univar, Ajax Chemicals, Australia
Glycerol	Sigma Chemicals Company, USA
Glycine	Sigma Chemicals Company, USA
Hybond, nitrocellulose membrane	Amersham Pharmacia Biotech, UK Limited
Hydrochloric acid	Rhone Poulenc Laboratory Products, France
Iodoacetamide	Bio-Rad Laboratories, CA, USA
IPG Ready Strips 11cm, pH 3-10	Bio-Rad Laboratories, CA, USA
Low fat milk powder	Diploma
Methanol	BDH Chemicals, Australia
Mineral oil	Bio-Rad Laboratories, CA, USA
NaCl	Sigma Chemicals Company, USA
SB 3-10 (Strong non-ionic detergent)	Bio-Rad Laboratories, CA, USA
SDS (Ionic detergent)	Bio-Rad Laboratories, CA, USA
Silver Stain Plus Kit - Development Accelerator solution - Fixative enhancer Concentrate - Image Development Reagent - Reduction Moderator Solution - Silver Complex Solution	Bio-Rad Laboratories, CA, USA
TBP (Tributyl phosphine)(strong protein reducer)	Bio-Rad Laboratories, CA, USA
TEMED	Sigma Chemicals Company, USA
Thiourea	Aldrich Chemical Company
Triton-X 100 (Ionic detergent)	Sigma Chemicals Company, USA
Trizma Base (Buffer)	Sigma Chemicals Company, USA
Urea	Bio-Rad Laboratories, CA, USA
Prolactin	Sigma Chemicals Company, USA
RPMI	Invitrogen Victoria, AUST
Insulin	Sigma Chemicals Company, USA
Growth Hormone	Sigma Chemicals Company, USA
Estradiol	Sigma Chemicals Company, USA
Gentamicin	Sigma Chemicals Company, USA
Glutamine	Sigma Chemicals Company, USA
Trypsin	Sigma Chemicals Company, USA
Nuclease Mix	Amersham Biosciences
Bio-Rad DC Protein Assay Kit I (#500-0111)	Bio-Rad Laboratories, CA, USA
PMSF	Sigma Chemicals Company, USA
FCS (Foetal Calf Serum)	Invitrogen Victoria, AUST.

Table 3.1: Chemicals and supplies used in this study.

3.2 SOLUTIONS AND BUFFERS

3.2.1 TANK BUFFER CONCENTRATE

30.28g	Tris
144.14g	Glycine
10g	SDS

Make to 1L with Milli Q water. Mix thoroughly until all is dissolved. To make 1 x Tank Buffer, 100mls of Tank Buffer Concentrate was combined with 900mls of Milli Q water.

3.2.2 TRIS BUFFERED SALINE CONCENTRATE

87.66g	NaCl
12.11g	Tris

Reagents were combined and made up to 1L with Milli Q water in a Schott bottle. The solution was placed on a magnetic stirrer and thoroughly mixed until the NaCl and Tris were dissolved. To make 1 x TBS, 100mls of TBS Concentrate was combined with 900mls Milli Q water.

3.2.3 0.1% AGAROSE SOLUTION

100ml	Tank Buffer
1g	Agarose
2µg	Bromophenol Blue

All reagents were combined in a Schott bottle. Prior to use these were heated on high in a microwave for 30 seconds, mixed and then heated for a further 30 seconds. The heated solution was then allowed to cool, but not set, so as not to denature the proteins. This was stored in a Schott bottle at room temperature and re-heated when required.

3.2.4 ELECTROBLOT BUFFER

7.5g	Tris
36g	Glycine
500mL	Methanol

Reagents were added to a large conical flask and made up to 2.5L with Milli Q water. The buffer was then transferred to a suitable container for storage at room temperature.

3.2.5 EQUILIBRATION BUFFER

36g	Urea
10mL	20% SDS

3.3mL	1.5M Tris
40mL	50% Glycerol

All reagents were combined. Due to the viscosity of the glycerol this was warmed for 30 seconds in a microwave on high to make measuring easier. It was made up to 100mL with Milli Q water and stored in Schott bottle at room temperature.

3.2.6 EQUILIBRATION BUFFER/DTT SOLUTION

200mg of DTT was dissolved into 10mls of equilibration buffer immediately prior to use.

3.2.7 EQUILIBRATION BUFFER/IAA SOLUTION

250mg of Iodoacetamide (IAA) was dissolved into 10mls of equilibration buffer immediately prior to use.

3.2.8 SEQUENTIAL EXTRACTION SOLUTIONS

3.2.9 SOLUTION 1

96.8mg	Tris
20mL	water

Tris and water was added to a 50mL falcon tube and divided into 1mL aliquots and stored in Eppendorf tubes at –20°C.

3.2.10 SOLUTION 2

24g	Urea (dried)
385µg	TBP
2g	4% CHAPS
0.2%	Carrier ampholytes
2µg	Bromophenol Blue

Reagents were added to a 50mL falcon tube and divided into 1mL aliquots and stored in eppendorf tubes at –20°C.

3.2.11 SOLUTION 3

1.5g (dissolved into 3mL water)	Urea (dried)
760mg	Thiourea
50µL of 200mM stock	TBP
100mg	2% CHAPS
100m g	SB 3-10
25µL/ 5mL sample	Carrier ampholytes

24.2mg	Tris
2µg	Bromophenol Blue

Reagents were added to a 50mL falcon tube and divided into 1mL aliquots and stored in Eppendorf tubes at –20°C.

3.2.12 FIXATIVE ENHANCER SOLUTION (FES)

500mL	Methanol
100mL	Acetic Acid
100mL	Fixative Enhancer Concentrate
300mL	Milli Q water

All reagents were added to a 1L Schott bottle and stored at room temperature.

3.2.13 LAEMMELI BUFFER

0.6g	SDS
1.9mL/1M	Tris
3mL	Glycerol
150µL	β-Mercaptoethanol (optional)
pH 6.8	

Reagents were combined and made up to 8.5mL with Milli Q water. A tiny amount of Bromophenol Blue (BPB) was added for tracking purposes and buffer was stored in 850µL aliquots in a -20°C freezer. 150µL of β-mercaptoethanol was added prior to use when a reducing buffer was required.

3.2.14 5% LOW FAT MILK POWDER (LFMP)

5g	LFMP
100mL	1 x TBS

LFMP and 1 x TBS were combined in a conical flask, sealed with parafilm and shaken vigorously until all LFMP was fully dissolved.

3.2.15 RPMI COMPLETE MEDIA

500ml	1640 RPMI
50ml	Foetal Calf Serum
5ml	L-Glutamine
500µl	Gentamicin
400µl	Bovine Insulin

All ingredients were added to the 1640 RPMI, using aseptic technique to avoid media contamination. Media was made up as required and discarded when pH was

compromised (indicated by the phenol red pH indicator dye in the 1640 RPMI). The complete media was stored at 4°C.

3.2.16 SEPARATION GEL: 8.5%

4mL	30% acrylamide / bis acrylamide 29:1 solution
3mL	5 x sep solution
8mL	Milli Q water
50µL	10% APS*
15µL	Temed

*10% APS (Ammonium persulphate) made up fresh each day, (0.1g APS/1mL Milli Q water)

Polyacrylamide, 5 x Sep solution and water were combined in a beaker. Immediately prior to pouring the gel APS and Temed were added to polymerise the solution to form a gel. The above volumes were doubled to pour two 7cm gels.

3.2.17 5 X SEP SOLUTION

112.5g	Tris
2.5g	SDS
pH 8.8	

Tris and SDS were combined and made up to 1L with Milli Q water and stored at room temperature until required.

3.2.18 SEPARATION GEL: 10% (1 X 7CM MINI GEL)

7mL	30% acrylamide / bis acrylamide 29:1 solution
3mL	5 x sep solution
5mL	Milli Q water
50µL	10 % APS
15µL	Temed

Polyacrylamide, 5 x sep solution and water were combined in a beaker. Immediately prior to pouring the gel APS and Temed were added to polymerise the solution to form a gel.

3.2.19 10 X STACK SOLUTION

75.5g	Tris
5g	SDS
6.8	pH

Tris and SDS were combined and made up to 1L with Milli Q water and stored at room temperature until required.

3.2.20 3% STACK GEL

1mL	30% polyacrylamide
1mL	10 x Stack Solution
8mL	Milli Q water
30µL	10% APS
12µL	Temed

Polyacrylamide, 10 x stack solution and water were combined in a beaker. This was done whilst preparing the separation gel and then covered with parafilm to avoid evaporation. Immediately prior to pouring the gel APS and Temed were added and mixed thoroughly to polymerise the solution to form a gel.

3.2.21 SILVER STAINING SOLUTION

35mL	Milli Q water
5mL	Silver Complex Solution
5mL	Reduction Moderator Solution
5mL	Image Development Reagent
50mL	Development Accelerator Solution

All reagents except Development Accelerator Solution were combined according to the instruction sheet (Bio-Rad Laboratories). Directly prior to use Development Accelerator Solution at room temperature was added and silver stain solution was used immediately.

3.2.22 STOP SOLUTION FOR SILVER STAIN

25mL	Acetic Acid
475mL	Milli Q water

Above reagents were combined and stored in a Schott bottle at room temperature and used when required.

3.2.23 T-NET EXTRACTION SOLUTION

100mls	Triton X-100
7.9g	NaCl
1.7g	EDTA
5.4g	Tris

The above reagents were added to 900mls of Milli Q water and stored at room temperature until required.

3.3 SAMPLE COLLECTION

Ethical approval was gained (RHH 2001.02) and Australian samples were obtained with informed consent from patients undergoing modified radical mastectomy, lumpectomy, or in the case of healthy tissue from patients undergoing reduction mammoplasty. Samples were retrieved from Anatomical pathology divisions at the RHH, following resection of required tissue for histological examination. Remaining sample was then removed of macroscopic fatty deposits and frozen at -70°C.

Greek and Taiwanese samples were obtained from the Prolypsis Breast cancer centre in Athens and the Chang Gung Memorial Hospital in Taiwan. These samples were then sent, on dry ice to the Discipline of Surgery at the University of Tasmania, where on arrival they were stored, alongside the Australian samples at -70°C.

3.4 SAMPLE PROCESSING

3.4.1 TISSUE PROCESSING

100mg of tissue was dissected from the tumour and then sonicated in 1ml of Sequential Extraction Kit (Bio-Rad Laboratories CA, USA) reagent 1 and 20µl of protease inhibitor (Sigma).

After the tissue was homogenised it was transferred from the vessel into a 1.5ml microcentrifuge tube (Eppendorf Centrifuge # 5415D) and centrifuged for 15 minutes at 13000rpm. The supernatant was removed and collected in a fresh tube with the sample name, date and "SOL1". The pellet was transferred back into the vessel and then sonicated with 1ml of extraction solution 2 and 20µl of Tributyl phosphine (TBP). Following sonication, centrifugation was repeated and then the same process for solution 3. Samples were frozen at -20° until further processing occurred.

The samples that were used in this study are detailed in Appendix A.

3.4.2 CELL LINE (T47D) PROCESSING

Sample processing for the cell line T47D was slightly altered to that of the tissue, due to advances in protocol since the beginning of the tissue work. Extraction was carried out by taking the pellet of cells (see section 3.6 for how the pellet was obtained) and making

the following modifications to the tissue processing protocol. At the end of homogenisation and before the centrifugation of extraction solution 2, 10µl of Nuclease Mix (Amersham Biosciences) was added and incubated at room temperature for 30 minutes. For Extraction solution 2 and extraction solution 3 only 500µl of buffer was used as opposed to the 1ml as was used for the tissue.

Each sample extraction then underwent a Bradford protein assay (Bio-Rad Laboratories) to determine protein concentration.

3.4.3 2D SDS-PAGE Tissue

3.4.4 Iso-Electric Focussing - Tissue

185µl of extracted tissue sample was pipetted into a 1.5ml microcentrifuge tube and a minute amount of bromophenol blue was added. This was then pipetted along a channel of an 11cm disposable rehydration tray. An 11cm IPG strips (Bio-Rad) of pH 3-10 were removed from the freezer, their protective coating removed and was then placed into each channel, acrylamide side down. Mineral oil was then layered on to the top of the sample/IPG strips and left to rehydrate overnight. 6 ml of water was placed on to dry paper wicks (Bio-Rad) which were then placed over the electrodes in the channels of the IEF focussing tray. Rehydrated IPG strips were blotted on damp filter paper and then transferred to an IEF focussing tray and focussed according to the following protocol:

Once focussed the strips were either frozen at -20°C until required or equilibrated ready for the 2nd dimension separation.

3.4.5 Equilibration for Tissue

Strips were placed acrylamide side up into the rehydration tray, which was cleaned during focussing, and then 4ml of equilibration buffer containing 2% DTT was added and then placed on a rocker to incubate for 10 minutes. The DTT was decanted and then 4ml of Equilibration buffer containing 2.5% Iodoacetamide was added and incubated for a further 10 minutes. The Buffer was decanted and the strips separated on to a 2nd dimension.

3.4.6 SDS-PAGE

The IPG strips were taken from the trays in which they were equilibrated and place on to the top of a 4-10% Tris-HCl Criterion Gel (Bio-rad) and set in place with a 1% agarose solution with Mw marker added to the pH 3 end. Once the agarose was set the tank and upper well was filled with tank buffer and then the samples were separated at 200V for 55 minutes. Once finished the gels were placed in a Fixative Enhancing Solution of Silver Staining.

3.4.7 2D SDS-PAGE FOR CELL LINE SAMPLES.

3.4.7.1 ISO-ELECTRIC FOCUSSING - CELL LINE.

For cell lines, 90µl of extracted samples was added to 90µl of the corresponding extraction buffer to make a total volume of 180µl loaded on to each strip for rehydration overnight. Remaining protocol is as per tissue protocol.

3.4.7.2 EQUILIBRATION FOR CELL LINE

Some variations to the above protocol were made for the equilibrating of cell line samples. Firstly IPG strips were place acrylamide side down, as this was found to produce better separation. Secondly 20µl of TBP was added to the Equilibration buffer with DTT to enable better reduction of the sample and ensure that any post extraction protein bonds have been dissociated.

3.4.7.3 SDS PAGE

Samples were run on to 10-20% Tris-acrylamide gels and underwent separation for 80 minutes.

3.5 WESTERN BLOTTING:

3.5.1 ELECTROTRANSFER

For both mini and the Criterion systems, the filter paper, fibre pads and Hybond nitrocellulose membrane (Amersham Pharmacia Biotech, UK Limited) were pre-wet in electroblot buffer. For the mini system the nitrocellulose was left in electroblot buffer, the central electrode stand removed from the tank and glass plates were carefully

separated. The stack gel was removed with a tissue and the gel was floated from the glass plate to the top of the nitrocellulose in the electroblot buffer. The gel and the nitrocellulose were removed from the buffer and placed on a chopping board and the nitrocellulose was trimmed to size with a sharp scalpel blade. Gel and nitrocellulose were then placed using forceps on to filter paper in orientation as seen in Fig 3.1.

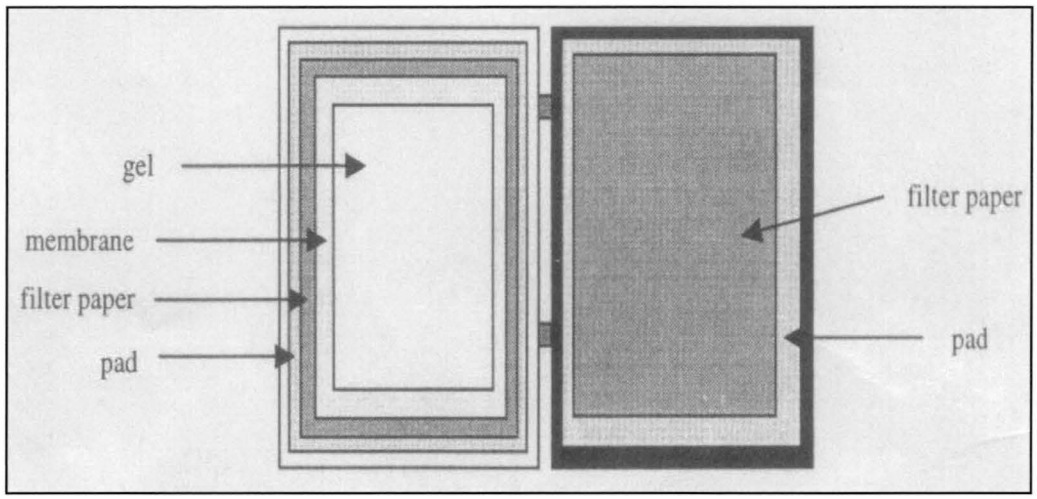


Figure. 3.1: Set up of a western blot cassette. The cassette would be folded and then the side corresponding to the right hand side of the the figure, would be aligned with the negative electrode.

The criterion system is similar, the gel was removed from plates and placed carefully on top of filter paper, instead of the nitrocellulose because of the fragility of the gel (7cm = 1.5mm thick, 11cm = 1.0mm thick). To avoid manipulating the gel too much an estimation of size for the nitrocellulose was made and then the nitrocellulose was placed, using forceps on top of the gel and the cassette closed, ensuring the correct orientation was achieved as seen in Fig 3.8. The Criterion tank was connected to the Power Pac 200 (Bio-Rad Laboratories) because of its higher amperage capacity (2 Amps) and set to operate at 100V for 30 minutes. The Mini Protean tank was connected to a Power Pac, 3000 or 200 and operated overnight at 30V.

3.5.2 IMMUNOBLOTTING

The transfer cassette was removed from the tank and the gel and filter paper were discarded. The membrane was placed into a flat bottomed dish and the membrane covered with 5% LFMP (50mls for 7cm mini gels and 100mls for 11cm criterion gels). The membranes were incubated for 30 minutes on a Ratek Platform Mixer (Crown

Scientific) with a set speed of 45. The membrane was rinsed with 1 x TBS twice and then the primary antibody was applied as per instructions on the corresponding antibody datasheets. These can be found in Appendix B.

The primary antibody was poured off into a labelled 50mL falcon tube and kept refrigerated for further use. Rinsed in 1 x TBS twice, the secondary antibody was applied and incubated for 1 hour with the platform mixer set at 45 and then the secondary antibody was discarded. The membrane was rinsed three times with 1 x TBS and then proceeded by three consecutive 5 minute washes in 1 x TBS with the first two 5 minute washes having the platform mixer set at 45 for the first two washes and then at 50 for the final wash. Ensuring that membranes are lying flat equal volumes of reagent one and two from an ECL kit (Amersham Pharmacia Biotech, UK Limited), (2ml of each reagent per membrane for mini's and 5 ml for Criterion blots) was pipetted on to the membrane and incubated for 60 seconds, blotted dry on paper towel and then covered with cling wrap. The membrane was placed into an auto-radiography cassette and taken to the dark room in the Royal Hobart Hospital's radiology department and then exposed to Hyperfilm (Amersham Pharmacia Biotech, UK Limited) for 30 seconds, 60 seconds and then longer durations as required. The film was developed using an automated Kodak RP X-OMAT Processor and then examined for protein labelling.

3.6 CELL CULTURE

T47D breast carcinoma cells were used to set up an in vitro culture model to look at changes under the influence of a set of hormones. Cells were obtained from ATCC in Manassas VA in the USA. On arrival cells were cultured, split, and frozen to ensure back up frozen stock.

3.6.1 ESTABLISHMENT OF CELL LINE

3.6.1.1 THAWING CELLS

T47D stock cells were removed from liquid Nitrogen, placed into a 37°C water bath and rapidly thawed until only small pellet of ice crystals remained.

The content of the cryovial was transferred into an empty 10mL centrifuge tube. Gradually RPMI media was added dropwise until 1mL of media had been added and then RPMI was added a little faster until all media was added, keeping the tube and cells on ice at all times.

Once the RPMI had been added the tube was centrifuged for 5 min @ 1000rpm. The supernatant was discarded and the pellet of cells resuspended in 5ml of RPMI. Two different T25 flasks of different dilutions were set up, one with 1ml of the concentrated cells and the other with 4ml of the concentrated cells. The T25 flasks were filled to 5ml total volume with complete RPMI. These flasks were then incubated at 37°C with 5% CO_2.

3.6.2 SPLITTING CELLS

Confluent flasks were removed from the incubator and the RPMI removed and discarded. The cells were then rinsed with 10ml of sterile PBS. Depending on the size of the flask 3ml, for T25 and 10ml for T75 flasks of sterile Trypsin EDTA solution was added to the flask. Once the cells had detached from the base of the flask, which would take 5 minutes in the 37°C incubator, 5-10ml of complete RPMI media was added and the trypsin EDTA effect on the cells was halted.

The cell suspension was transferred to a centrifuge tube and spun for 8 minutes @ 1200rpm to pellet the cells.

Cells were then resuspended in PBS and washed and then centrifuged again. The washed pellet was then resuspended in RPMI media and 100ml or 800ml of suspension was used to reseed T75 flasks containing 10ml RPMI depending on the intended purpose of the flask.

3.6.3 FREEZING CELLS

Cells were harvested by following the same initial steps as splitting the cells. After pelleting the cells they were washed in sterile PBS and then resuspended in chilled RPMI. If using a T75 flask 4ml was used and if using a T25 2ml was used. Placing the tube on ice, cold DMSO/FCS was added in a dropwise fashion until it was in equal proportion to the RPMI that was used. Cells were then placed in 1ml aliquots in to cryovial and frozen at -80°C insulated in a Styrofoam container to prevent sudden

freezing. After 24 hours the cells were transferred to liquid Nitrogen for long term storage.

3.6.4 TREATMENT OF CELLS

An 8 day time course was set up for each treatment of T47D cells as follows:

Day 1: A T75 stock flask was spilt, after pellet resuspension into 1ml of Complete RPMI, the suspension was then split into 3 flasks. Two of these flasks underwent treatment and were each seeded with 100µl of cells. The third flask was seeded with the remaining 800µl of cell suspension and was grown as the stock flask.

Day 4: Stock flask from Day1 reaches confluence and is therefore split as for Day 1 which gave rise to another two treatment flasks as well as a fresh stock.

Day 5: Half (5ml) of the media from the flasks to be treated from Day 1 is replaced with fresh complete media to ensure continued growth of cells to a point of confluence.

Day 7: Cells are treated with the desired hormones. The two flasks were treated one hour apart to allow for extraction period at 24 hour harvest. Table 3.6.1 outlines the treatments administered.

Day 8: Cells are harvested as per the protocol and the pellet is then extracted according to the sequential extraction protocol.

Treatment	GH	E2	PRL
1	**X**	-	-
2	-	**X**	-
3	-	-	**X**
4	**X**	**X**	-
5	**X**	-	**X**
6	-	**X**	**X**
7	**X**	**X**	**X**
8 (control)	-	-	-

Table 3.6.1: Treatment regime for T47D cell culture flasks. Appendix B contains the datasheets for each of the hormones administered.

3.6.5 EXTRACTION OF T47D CELLS

Following harvesting the washed cell pellet was resuspended in PBS again and then resuspended in 40 µl of protease inhibitor and 1ml of Extraction 1 Buffer and then transferred into a vial for homogenisation. The cells were homogenised and then spun and 13200 rpm for 15 minutes. Whilst Solution 1 was being centrifuged, 10 µl of TBP was added to each of the solution 2 and 3 tubes to ensure reduction of the remaining proteins and aid focussing at later stages. The pellet from Solution 1 was then placed back into the vial and homogenised again with 500µl of Extraction solution 2. The suspension was transferred into a 1.5ml Eppendorf tube and then 10 µl of Nuclease Mix (Amersham biosciences) was added and left to incubate at room temperature for 30 minutes. The solution was then spun at 13200 rpm for 15 minutes. The supernatant was removed and then the pellet was homogenised using Extraction buffer 3. The suspension was then spun again at 13200 rpm for 15 minutes and then the supernatant removed. All supernatants were stored at -20°C and labelled SOL1, 2 and 3 respectively. The pellet was also labelled and frozen at -20°C.

3.7 STAINING

3.7.1 SILVER STAINING

Gels in fixative were rocked for 20 minutes and then either left overnight or processed immediately. Fixative was discarded and the gels were washed in Milli Q twice for 10 minutes each time. Once washed, stain was added until desired intensity was reached and then staining was stopped with a 5% acetic acid solution for 15 minutes. After the silver reaction was stopped the gels were rinsed for 5 minutes in Milli Q and then the gels were transferred to zip lock bags and stored with 1ml of 0.05% NA Azide at 4°C after being imaged on a GS 800 Densitometer (Bio Rad).

3.7.2 SYPRO RUBY STAINING

Sypro Ruby staining was performed on samples where the intention was to send proteins from them for MALDI-TOF analysis.

After 2D separation the gel was placed in a fixative consisting of 40% Methanol and 10% acetic acid and then rocked for at least 30 minutes. Following fixation 50ml of

Sypro Ruby stain was added and the gel was placed in a Styrofoam esky to eliminate light and the rocked over night. Following staining the gel was rocked in destain solution consisting of 10% methanol and 7% acetic acid for 30 minutes and then rinsed in Milli Q for 15 minutes to limit background. Gel was then imaged on a ChemiDoc XRS (Bio Rad).

3.8 ANALYSIS

3.8.1 SCANNING THE GELS

For silver stain images PDQuest was opened from the desktop computer connected to the GS-800 densitometer. From the FILE menu GS-800 was selected, this opened up the "Acquire Image" options box which has three steps.

Step 1 was to select the type of stain, in this case silver, that was used on the gel being imaged. Step 2 was to conduct a preview scan to ensure that all of the gel was captured and that it was orientated correctly and then adjust the scan area as appropriate based on the preview scan to give the smallest area required to capture the whole gel, this was important to reduce the file size significantly. Step 3 was to set the resolution settings, there were left on the defaults of 63.5 microns for both the X and Y axis for all images scanned.

The image was then acquired, this starts the scanning process which when finished opens up the captured image in a new window and is now in an editable form.

For Sypro gels imaged on the ChemiDoc XRS, which uses the Bio Rad software Quantity One. The UV transluminator settings on the instrument were selected and the image capture time was set on 0.25 seconds. The image was then saved and then reopened in PDQuest as a 2D image and was ready for any required editing or analysis.

3.8.2 EDITING IMAGES

The raw scan images were rotated if necessary to place the pH 3 end of the gel on the left hand side of the screen. The images were then cropped according to advance crop setting "simple spot". Images were then appropriately named and saved, for example: extraction solution 2 of sample 4 from the Taiwanese cohort would be saved as: TCB 4 SOL2.

Once the raw images were resized, named and saved they were then edited through automatic spot detection system using preset parameters "simple spot." This process creates two new image types required before forming a Matchset. Firstly a filtered image where the file is altered to remove the background based on the parameters that are selected through the automated spot detection. The second file type is a Gaussian image that is based on a computer generated image of the "spot". This process uses Guassian modelling techniques to eliminate any streaking effect that may be present.

3.8.3 CREATING MATCHSETS

Once the three image types were created they were then added to one of two matchsets. TCB SOL2 gels were added to the matchset CBvNB.ms and the matchset was renamed CBvNBvTB.ms with copies of the original made and saved in a backup folder. TCB SOL3 gels were added to CBvNB sol3. Images were then edited to ensure two things: a) that all feasible spots were included, and b) that artefacts, such as air bubbles were removed. Although this is a very time consuming process it was necessary for the reliability of results and each gel was zoomed in on and checked individually. It should be noted that any artefacts that are removed from the analysis are still visible in the Raw image.

The images from each separate file are listed in Table 4.1 in the next chapter.

3.8.4 MATCHSET ANALYSIS

Once the gels are completed it is then possible to allocate certain gels to certain "Replicate Groups" this was done to create 3 groups; ACB (Australian Cancerous Breast), NB (Normal Breast) and TCB (Taiwanese Cancerous Breast). Along with these Replicate Groups, "Classes" can be formed to allow the input of other data as it concerns each gel, for example a samples histological status, Estrogen Receptor status, Progesterone Receptor status, and whether they fall into the early onset or late onset group, whether they received pre-surgery therapy and if so what type, their eye colour, postcode or favourite food can also be added if desired. Information can also be added here concerning the treatment of the gel, for example the 2D running time, the addition of reducing agents, the addition of nucleases etc. This ability to add such a diverse range of information makes this a very important and useful function for downstream analysis.

It needs to be mentioned here that the gels for the GCB (Greek Cancerous Breast) patients were originally placed into a separate Matchset and so they were in a separate file to the other samples. As this GCB containing Matchset also included ACB, the SSP data from CBvNB.ms was able to be imported and the same reference number for each protein was established.

Once replicate groups and classes have been set up, analysis can be conducted. For this project analysis was conducted between each population and Healthy tissue for 5 fold up regulated, 10 fold up regulated, 5 fold down regulated, 10 fold down regulate, statistically significant, and a combination of statistically significant and each up or down regulated in each group. Also the analysis of Early versus Late onset analysis was conducted for the above groups as was as analysis of the control cohort where a history of breast cancer was the differential factor thanks to the ability to add classes.

The data for each analysis was then exported in to Excel. Similarities between the 3 populations were then identified to provide a group of proteins of interest. These proteins were then sent to be analysed by MALDI-TOF MS at the Australian Proteomic Analysis Facility. The detailed analysis results for each protein that was successfully identified can be seen in Appendix C.

4 DIFFERENTIAL ANALYSIS OF THE PROTEINS IN CANCEROUS AND HEALTHY BREAST TISSUE.

4.1 SUMMARY 54

4.2 BACKGROUND 55

4.3 RESULTS 56

4.3.1 Results of MALDI-TOF MS 58

4.3.1.1 Serum albumin (1701) 58

4.3.1.2 Osteoglycin (2403 and 2406) 59

4.3.1.3 De-Isomerase Related Protein 5A (2501) 60

4.3.1.4 Lumican (2701) 61

4.3.1.5 Chain A of the Ga Molecule (4702) 62

4.3.1.6 Folliculin (3701) 62

4.3.1.7 Heat Shock Protein 60 (3702 and 4103) 63

4.3.1.8 UDP - Glycosyltransferase (7112) 64

4.3.1.9 FLJ 20309 (7204) 64

4.3.1.10 RP3-39312.2 (8203) 65

4.3.1.11 L-Plastin (7203) 65

4.4 DISCUSSION 80

4.4.1 Serum Albumin 80

4.4.2 Osteoglycin 81

4.4.3 Lumican 82

4.4.4 DIRP5a 83

4.4.5 Chain A of the GA module 84

4.4.6 Folliculin 86

4.4.7 Heat Shock Protein 60 86

4.4.8 UDP-Glycosyltransferase 87

4.4.9 FLJ 20309 89

4.4.10 RP3-393D12.2 90

4.4.11 L-Plastin 91

4.5 CONCLUSIONS 91

4.1 SUMMARY

The comparative analysis between diseased and healthy tissue is essential in the study of disease. Recognising the differences between the two and then being able to decipher the critical changes is an important part of any medical research.

In this chapter, not only has there been analysis of the local Tasmanian community and the breast cancers within it, but samples from Taiwan and Greece have also been included to enable the recognition of global biomarkers for breast cancer.

The limitations surrounding a lot of research are due to the assumption that one set of patients will behave the same as any other across the globe. Findings have been published to this extent where a protein or gene, believed to be of importance, is only important to a limited few in that population, for example the CHEK 2 gene discussed in 1.3.1. This problem has been overcome in this study by the inclusion of these other tissue samples from a variety of ethnic and cultural backgrounds.

Using a proteomic approach it has been possible to identify proteins that are up and down regulated in breast cancer tissue when compared to healthy tissue. The advantages of this are that it is possible to look at many proteins at once and identify them all simultaneously and then determine their importance and likely involvement.

Although not known for its ability to be quantified, in this study, with the use of the high powered software program PDQuest it has been possible by using normalisation as the concentration of loaded samples is known. This step has made these analyses very relevant, with densitometry reflecting directly what is happening in the tissue post transcription.

The results here show the findings of 13 proteins that were able to be identified using MALDI-TOF mass spectrometry. Each of these have been analysed at depth and discussed within this chapter. The proteins isolated involve a wide variety of cellular processes, from structural integrity of the extracellular matrix, in lumican and osteoglycin, to the possible carrier protein serum albumin and the metabolic protein UDP-glycosyltransferase. Hypothetical proteins, whose sequence has been determined from cDNA based extrapolation have also been isolated and may prove to be important in breast cancer development.

4.2 BACKGROUND

There were three cancerous cohorts from three separate populations from around the world included in this analysis. Appendix A shows specific data on the patients included in this study.

This and the following two chapters will detail proteins found in connection with different analyses conducted on the same samples. This chapter deals with the comparison of cancer tissue samples to a selection of healthy tissue from the local community. The following chapter, Chapter 5 deals with a further division of the samples cohorts, this time into early and late onset cancers. The chapter after that, Chapter 6 looks at a comparison within the healthy cohort, into those samples with a history of breast cancer and those without a history. Each chapter isolates and identifies proteins of interest. For the proteins that are successfully identified, each analysis type is conducted in order to gain an improved understanding in how that particular protein may be influencing the tissue and at what stage.

There were four matchset files created in PDQuest that were used to obtain the data for all of the following results chapters. The members of each are listed in table 4.1. Two are for solution two and two are for solution 3. The matchsets that included the Greek samples were compiled by Dr K.Bhatia for his Medical Masters degree. The image for each sample was retained; however image editing was updated with the use of an updated version of PDQuest (an upgrade from v6.0 to v7.1).

Data for the comparisons of the Greek samples to the other populations was exported and compared manually in Excel software.

The results presented in all Plates are measured on the percentage of total protein content (%TPC). This calculation was made possible by the loading of similar concentration of the gel and then normalisation of the gels conducted with in PDQuest. This normalisation process assumes that the valid proteins on the gel make up a total of 100% of the protein, this then give a value, in percentage of what the proteins of interest contribute to this total 100%.

CBvNBvTCB Sol2a			ACBvGCB_Sol2		CBvNB_Sol3a			ACBvGCB_Sol3	
AUST	TCB	NB	AUST	GCB	AUST	TCB	NB	AUST	GCB
ACB1	TCB1	NB1	ACB1	GCB1	ACB1	TCB1	NB1	ACB1	GCB1
ACB2	TCB1a	NB2	ACB2	GCB2	ACB2	TCB1a	NB2	ACB2	GCB2
ACB3	TCB2	NB3	ACB3	GCB3	ACB3	TCB2	NB3	ACB3	GCB3
ACB4	TCB3	NB4	ACB4	GCB4	ACB4	TCB4	NB4	ACB4	GCB4
ACB5	TCB4	NB5	ACB5	GCB5	ACB5	TCB4a	NB5	ACB5	GCB5
ACB6	TCB4a	NB6	ACB6	GCB6	ACB6	TCB5	NB6	ACB6	GCB6
ACB7	TCB5	NB7	ACB7	GCB7	ACB7	TCB6	NB7	ACB7	GCB7
ACB8	TCB6	NB8	ACB8	GCB8	ACB8	TCB7	NB8	ACB8	GCB8
ACB9	TCB7	NB9	ACB9	GCB9	ACB9	TCB8	NB9	ACB9	GCB9
ACB10	TCB8	NB10	ACB10	GCB10	ACB10	TCB9	NB10	ACB10	GCB10
ACB11	TCB9	NB11	ACB11	GCB11	ACB11	TCB10	NB11	ACB11	GCB11
ACB12	TCB10	NB12	ACB12	GCB12	ACB12	TCB11		ACB12	GCB12
ACB13	TCB11		ACB13	GCB13	ACB13	TCB12		ACB13	GCB13
ACB14	TCB12		ACB14	GCB14	ACB14	TCB13		ACB14	GCB14
ACB15	TCB13		ACB15	GCB15	ACB15	TCB14		ACB15	GCB15
ACB16	TCB14		ACB16	GCB16	ACB16	TCB15		ACB16	GCB16
ACB17	TCB15		ACB17		ACB17	TCB16		ACB17	
ACB18	TCB16		ACB18		ACB18	TCB17		ACB18	
ACB19	TCB17		ACB19		ACB19	TCB18		ACB19	
ACB20	TCB18		ACB20		ACB20	TCB19		ACB20	
ACB21	TCB19		ACB21		ACB21	TCB20		ACB21	
ACB22	TCB21		ACB22		ACB22	TCB21		ACB22	
ACB23	TCB22		ACB23		ACB23	TCB22		ACB23	
ACB24	TCB23		ACB24		ACB24	TCB23		ACB24	
	TCB24				CIS	TCB24			
	TCB25				ACB25	TCB25			
	TCB26				ACB26	TCB26			
	TCB27				ACB27	TCB27			
	TCB28								

Table 4.2: Samples included in each of the PDQuest Matchsets for the healthy versus cancer comparisons. There are two files for each of the extraction solutions.

4.3 RESULTS

The selection of proteins to undergo further investigation was based on analysis performed using PDQuest software (Bio-Rad, USA). As described in section 3.8 matchsets were created and image editing performed, followed by the formation of replicate groups and classes within the program. The collation of gels in this manner

allowed for the comparison of each population type to the control group. Comparisons were performed to show the up and down regulation of proteins displaying a 5 or 10 fold change.

Analysis was performed using an unpaired student t-test to show which proteins were detected as having significantly different concentrations. Following both the up/down regulation comparison and the statistical significance analysis a Boolean analysis was done to display proteins present in both analyses. A summary of these can be seen in table 4.2.

Population	Up regulated	Up regulated and significant#	Down regulated	Down regulated and significant	Significant
Solution 2					
Australian					
5x	24	5	8	4	37
10x	8	0	3	0	
Taiwanese					
5x	13	0	13	6	32
10x	5	0	5	1	
Greek					
5x	18		8		61
10x	9		0		
Solution 3					
Australian					
5x	25	3	13	5	30
10x	7	0	0	0	
Taiwanese					
5x	15	0	15	10	45
10x	4	0	1	1	
Greek					
5x	18	6	7	1	74
10x	9	4	1	0	

Table 4.2: Summary of the proteins found to be of interest in the cancerous versus healthy analysis of breast tissue. # Significance was determined by a student t-test with a CI of 95%.

From this selection of proteins, 22 were chosen to be sent for sequence analysis. This selection process was based on the determination of significance in combination with up or down regulation. These proteins were excised from the gel and then sent to the Australian Proteome Analysis Facility (APAF) for sequence analysis by Matrix Assisted Lased Desorption Ionisation – Time Of Flight mass spectrometry (MALDI-TOF MS.

Sequencing results were mixed, with 13 of the 22 being identified and the other 9 remaining unknown. Details of each of the identified proteins being listed below in table 4.3.

SSP	ID	Mowse Score	Exp. Mw/pI	Th. Mw/pI	Accession number
Solution 2					
1701	Serum albumin	30	57/6	69 3/6.13	gi 23307793
1703	-	-	58/5 8	-	-
2403	Osteoglycin OG	56	29/6.2	33.9/5.33	gi·33150528
2406	Osteoglycin OG	54	30/6.2	33 9/5 33	gi 33150528
2501	Disulfide Isomerase-Related Protein 5	108	43/6.1	46.2/4 95	gi 1710248
2701	Lumican	52	58/6.1	38.4/6 16	gi.30582253
3702	HSP60	70	53/6.4	61/5.7	gi:6996447
4702	Chain A, crystal structure .	446	60/6.7	65 2/5.57	gi:55669910
6303	-	-	26/7 6	-	-
9302	-	-	27/9 1	-	-
Solution 3					
2403	-	-	38/6 4	-	-
2906	-	-	259/6.4	-	-
3701	Folliculin	30	67/6.2	64 4/5 83	gi 22255880
4103	HSP60	42	27/7	61.0/5 7	gi 6996447
7101	-	-	26/8		-
7112	UDP glycosyltransferase 1 family, polypeptide A6 isoform 2	35	32/8	29.6/9 27	gi·45827767
7203	L-Plastin	42	32/8	70/5.2	gi:4504965
7204	FLJ20309	41	34/7 9	98 1/8.48	gi·38488718
8101	-	-	27/8.1		-
8108	-	-	31/8.1		-
8203	RP3-39312 2	38	33/8 2	32.7/7 77	gi·7529595
9305	-	-	37/>9.6		-

Table 4.3: Details from the MALDI-TOF analysis conducted at APAF for proteins differentially regulated between healthy and cancerous breast tissue.

4.3.1 Results of MALDI-TOF MS

From the sequencing of proteins that were of interest in the cancer versus healthy analysis, several findings specific to each protein were seen.

For each protein that was chosen to be sequenced, the results for the early versus late analysis were also examined. This was to add strength to the results already found and to ensure that a protein had not been missed by the analysis of the early versus late onset patients.

The following outlines the results gained for each of the identified proteins for both the cancer versus healthy and the early versus late onset analyses.

The referred to plates can be found, in order of appearance in the text, from page 67-79. There is a fold out legend corresponding to the information in these plates at the rear of this thesis.

4.3.1.1 Serum albumin (1701)

A serum albumin variant was isolated as a protein of interest for the cancer v healthy analysis. Plate 4.1a demonstrates that for serum albumin 1701, there was a significant

level of down regulation, approximately 65%, in the cancer samples when compared to the healthy tissue, the exact values of significance can be seen in Appendix D. As explained in section 4.2 regardless of the analysis that prompted selection for identification, cancer versus healthy, early onset versus late onset and, in the control group, history of breast cancer versus no history of breast cancer.

The serum albumin 1701 variant occurred in more than 90% of the control samples this frequency fell to 55% or less in the each of the cancerous populations. There were significant difference between all cancerous cohorts and the healthy tissue (Plate 4.1a).

Early versus late onset analysis revealed no change in the concentration of serum albumin 1701 based on the age of onset (Plate 4.1b). All breast cancer populations remained significantly down regulated compared to the healthy tissue. There was no change in the pattern of expression either, with high levels in the healthy and low levels in the cancerous samples, indicating that the level of expression of serum albumin 1701 is independent of age.

When the concentration of serum albumin in the samples were divided into the stage classification of the tumour, there was a trend observed, showing serum albumin 1701 levels decreasing with increasing staging (Plate 4.1c).

Analysis was conducted to enable observation of the differences between the healthy patients with a history of breast cancer and those without a history. These results showed that in the case of serum albumin there was no significant change, with the levels in both control groups remaining significantly higher than the cancerous samples (Plate 4.1d).

4.3.1.2 Osteoglycin (2403 and 2406)

Two proteins, SSP 2403 and 2406, that were sent for MALDI-TOF MS for the cancer versus healthy analysis were identified as osteoglycin, a small leucine rich proteoglycan (SLRP) that binds to collagen during fibril formation.

The concentration of osteoglycin in the control group was consistently higher, with the levels in the diseased tissue of all patients being down regulated by 34% and 50% respectively for each isoform when compared to the healthy cohorts (Plate 4.2a and 4.3a).

Osteoglycin isoform 2403 was expressed in 66% of healthy tissue samples. This rate of expression drops in the diseased tissue to 33% of cancer patients.

In the early versus late analysis the late onset patients remained significantly altered between the healthy and diseased patients (p=0.043 and p=0.036 for ACB and TCB respectively). In the early onset the significance was lost despite similar expression patterns being observed (Plate 4.2b).

This change in expression level was also present when the healthy cohort was split into those with and without a history of breast cancer. The level of osteoglycin in the two separated healthy cohorts was altered little, with no significant change in either isoform (Plate 4.2d).

Osteoglycin isoform 2406 was found to be in a higher percentage of normal tissue, 83% of all patients recruited, with the frequency dropping to 38% or less for the diseased samples. The concentration and significance for this isoform can be seen in Plate 4.3a. This too is unaffected by the age of onset of the patient as seen in Plate 4.3b. This is despite the absence of significance in the late onset patients that was seen in the early onset patients. The trend is however present with considerably higher levels in the healthy group for both the early and late onset cohorts.

Staging analysis suggests that there is a decreasing concentration with an increase in the stage of the tumour. This was observed in all populations however statistical analysis showed this decline to be insignificant (Plate 4.3c).

There was no significant difference between tissue samples form women with a history and those without a history of breast cancer (Plate 4.3d).

4.3.1.3 DE-ISOMERASE RELATED PROTEIN 5A (2501)

The comparison between the cancer and healthy tissue showed that this protein was only present in the Taiwanese patients. De-isomerase related protein 5a (DIRP5a) was observed in 17 out of the 28 TCB patients.

The early versus late analysis was found to be higher in the late onset patients however this was not significant (p=0.12) (Plate 4.4a and Plate 4.4b). The expression of DIRP 5a was more prevalent in late onset patients.

The staging analysis that was conducted on this protein found that there was a decrease in the expression level as the stage increased, i.e. got worse, however these changes were not significant from one stage to the next (Plate 4.4c).

4.3.1.4 LUMICAN (2701)

Lumican, like osteoglycin is also a member of the SLRP family. It too is associated with the structural integrity of tissue by way of collagen fibril formation.

In this study lumican was isolated as one of the proteins of interest for the cancer versus healthy analysis. The level of expression seen in cancerous tissue dropped to below 30% of that seen in the control group (Plate 4.5a). This difference was only significant in the TCB v NB comparison however ($p=0.028$), with the ACB comparison not reaching significance ($p=0.065$). This isoform of Lumican (SSP 2701) was seen to be in 66% of the control group and below 50% in each of the cancerous cohorts.

The significance of these results has been influenced however by the large variation in expression levels. The early versus late analysis shows that the levels in the early onset patients were not significant but this was not the case in the late onset patients with the concentration reaching a significant difference between the control group and the Taiwanese cohort ($p=0.04$) (Plate 4.5b).

Staging analysis for Lumican 2701 showed that there were significant changes between the concentration in the healthy controls and each tumour stage group, however there was no significance between the stages, offering no clear indication at which point this protein may become important. The levels were very low in all stages with the exception of GCB in stage 2, this was a result that was skewed by 1 very high reading from one individual patient (Plate 4.5c).

When looking at the comparison of the two healthy cohorts for this protein it was seen that there was a slight increase in patients with a history of breast cancer but this was not significant ($p=0.65$) with an even distribution of those who were and were not expressing the Lumican 2701.(Plate 4.5d).

4.3.1.5 CHAIN A OF THE GA MOLECULE (4702)

The levels of protein expression between populations of diseased tissue and the control group are similar. The results did not reach significance due to the variability in expression that was observed despite an average 2 fold increase (Plate 4.6a).

When early onset v late onset analysis was performed, no significance was reached. As is shown in Plate 4.6b, there is an increase in the levels of expression between the ages of onset in the healthy tissue that is not seen in the diseased tissue. This rise in expression in the older patients is not significant (p=0.32), again due to a high variation.

The levels of Chain A of the Ga Molecule decreases 2 fold compared to the healthy tissue, but unlike the other sequenced proteins in this study the percentage of patients expressing it remains unchanged with 50-70% of each cohort expressing the Chain A protein.

4.3.1.6 FOLLICULIN (3701)

Folliculin, unlike many of the proteins sequenced, is increased in the cancer samples. This concentration difference is only significant for the comparison between Australian cancerous and healthy tissues (Plate 4.7a). The differences in the other populations are not significant (p= 0.45 for the Taiwanese and p=0.27 for the Greek). The frequency of expression was seen as 85% of healthy samples, 91% of ACB samples and only 48% and 8% of TCB and GCB respectively.

Significance is only seen in the early onset analysis between ACB and NB. This significance it lost in the late onset samples. There is also an increase evident in the older, healthy samples but this is also not significant (p=0.16) (Plate 4.7b).

The staging analysis of this protein has showed no pattern of expression between stages, with each population showing varying concentrations at stages one, two and three. Several stages were significantly different to the healthy control however in the ACB and TCB comparisons (Plate 4.7c).

Analysis for the controls with a history compared to the controls without history of breast cancer shows a lower concentration of Folliculin 3701 in the samples with no history. This decreased concentration is almost significant (p=0.08), suggesting a larger cohort of healthy controls may result in significant difference. (Plate 4.7d).

4.3.1.7 HEAT SHOCK PROTEIN 60 (3702 AND 4103)

Two isoforms of Heat shock protein 60 (HSP60) were identified. One isoform is a cytoplasmic form HSP 60, this has been assigned the SSP of 3702. The other is nucleic form, identified in solution 3 and has been assigned an SSP of 4103.

The cytoplasmic isoform of HSP 60 is significantly up regulated in the cancerous state of all populations (Plate 4.8a) and is independent of age (Plate 4.8b). This isoform was found to be much more common in cancer patients with 95%, 71% and 93% of ACB, TCB and GCB patients expressing the protein, with just 58% of healthy patients expressing the cytoplasmic form of HSP 60.

In the staging analysis that was conducted it was seen that there is a significant change in the levels seen in each stage and the healthy tissue for all comparisons except GCB stage 3 v NB, ACB stage 3 v NB and TCB stage 1 v NB. No comparative analysis between stages gave any significant results (Plate 4.8c) (See Appendix D).

For the history versus no history analysis it can be seen that the cytoplasmic HSP is more frequent in patients with a history, 5/7 patients, compared to 2/5, but there is little change in the concentration level (Plate 4.8d).

The nucleic HSP 60 (SSP 4103) expression levels, which in this case has been identified in the highly hydrophobic portion of the cellular extracts (ie: Solution 3), appears to have aberrant expression levels between the diseased populations (Plate 4.9a). The concentration differences were not significant due to the low numbers of patients expressing the protein, only 27% in the control group and no samples in the Australian cancer cohort.

This hydrophobic HSP 60 protein was seen in 37% and 50% of Taiwanese and Greek patients respectively with little difference in concentration level being seen when the early versus late onset analysis was conducted however there was a skew of most patients to be in the older onset populations (Plate 4.9b).

The levels are higher in Stage 2 for the Greek and the TCB patients than in the other stages, however these results are not significant compared to healthy tissue. The change between TCB stage one and two is significant (P=0.02) (Plate 4.9c), suggestive that HSP 60 is involved in the progression of breast cancer, whether this is a causative or

subsequent action of progression however, is yet to be determined. A larger local cohort of both healthy and cancerous tissue would be advantageous is deciphering a link between HSP 60 and staging.

Control group analysis between those with and without a history of breast cancer showed a decrease in patients with a history of breast cancer however this is only in 3 patients in total and no significance was reached in this analysis (Plate 4.9d).

4.3.1.8 UDP - GLYCOSYLTRANSFERASE (7112)

UDP-glycosyltransferase was found to be entirely absent in all patients from Australia both control and cancerous. The Taiwanese however had expression in 5 patients (18.5% of samples) whilst the Greek showed expression in 9 patients (56% of samples) (Plate 4.10a). This suggests the possibility of population based or genetic differences rather than breast cancer specific changes.

There appeared to be no differences that were dependant upon the age of onset for the protein (Plate 4.10b).

Staging analysis suggests a peak in stage one, however with the low numbers of patients involved and no population matched controls to compare it to no significance was achieved (Plate 4.10c).

4.3.1.9 FLJ 20309 (7204)

FLJ 20309 is a hypothetical protein, meaning that during database searching for a match to the MALDI-TOF derived PMF no known previously described proteins matched. There was however a match to a hypothetical protein, whose existence is assumed from the human genome project.

This hypothetical protein FLJ 20309 was found only in Australian and Taiwanese cancer samples. Due to no expression in the healthy samples they were both determined to be significant results. Fifty eight percent of Australian patients expressed FLJ 20309 however the frequency of expression in the Taiwanese cancer samples was much lower at 18.5% (Plate 4.11a).

All of the TCB patients were early onset patients. But this was not the case in the ACB cohort where the expression of FLJ 20309 was split between the two age based cohorts,

with 2 patients in the early onset group, whilst the other 13 were late onset patients (Plate 4.11b).

Staging analysis for this protein suggests that there may be a role for FLJ 20309 in the early stages of cancer as there is no observed expression in Stage 3 tumours for ACB or TCB cancers (Plate 4.11c).

4.3.1.10 RP3-39312.2 (8203)

RP3-39312.2 is also a hypothetical protein that was found almost entirely in Australian samples, except for one Taiwanese sample. The differences between the Australian cancerous samples and the healthy breast tissue were significant (p=0.02) with expression in 42% and 9% of the samples respectively (Plate 4.12a).

The analysis for early onset versus late onset breast cancers for RP3-39312.2 suggests that it is more likely to have a role in late onset conditions as the concentration of the protein is highly decreased in the early onset patient that expressed it. The sole healthy sample was in the older age group and all but one ACB sample was a late onset patient. The sole TCB sample was also a late onset patient (Plate 4.12b).

The staging analysis suggests a role in stage 2 for RP3-39312.2 with the ACB cancers having higher concentrations during this stage, although significance wasn't reached for comparisons between stage 2 and stage 1 or stage 2 and stage 3 (Plate 4.12c).

The sole healthy sample that expressed RP3-393D12.2 is a patient that doesn't have a history of breast cancer and it is in very low levels. None of the patients with a history of breast cancer had any expression of RP3-393D12.2 (Plate 4.12d).

4.3.1.11 L-PLASTIN (7203)

L-plastin was found to be in much higher concentrations in the ACB tissue samples when compared to the healthy tissue. This was seen to a lesser extent in the GCB samples where an increased concentration was not as distinct. Very low levels were seen in the sole TCB patient that expressed L-Plastin (Plate 4.13a).

The frequency of L-Plastin was high with 22 out of 23 patients from the ACB cohort expressing it which dropped to 7 out of 12 in the healthy tissue. The occurrence in TCB

was very low with just 1 patient, whilst GCB had 7 out of 16 patients expressing L-Plastin in this form.

There was a significant difference present between the ACB samples and the healthy tissue (P=0.000002). Concentrations were also much higher in the ACB samples in the early onset group when compared to the late onset patients (Plate 4.13b), with the healthy samples changing little in concentration between age groups.

Staging showed an erratic level of expression with no trend, except for retaining significant differences to the healthy tissue (Plate 4.13c).

Patients in both healthy cohorts, with and without a history of breast cancer expressed L-Plastin and no difference in concentration was observed between the two groups of patients (Plate 4.13d).

Plate 4.1: Serum Albumin 1701

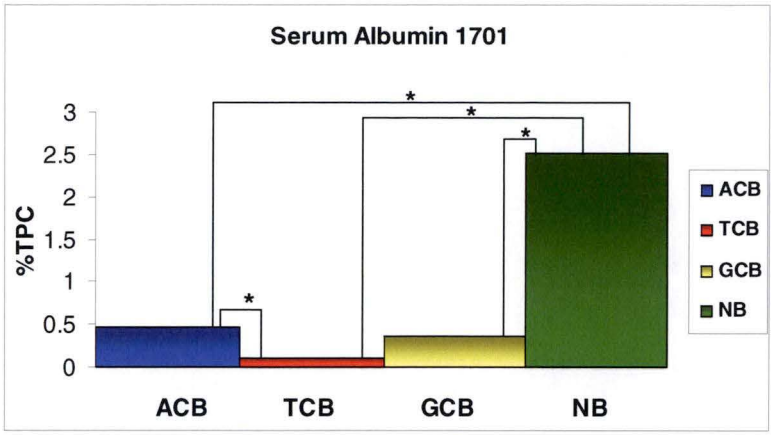


Plate 4.1a:

Serum albumin when compared to healthy tissue is down regulated in the cancerous state.

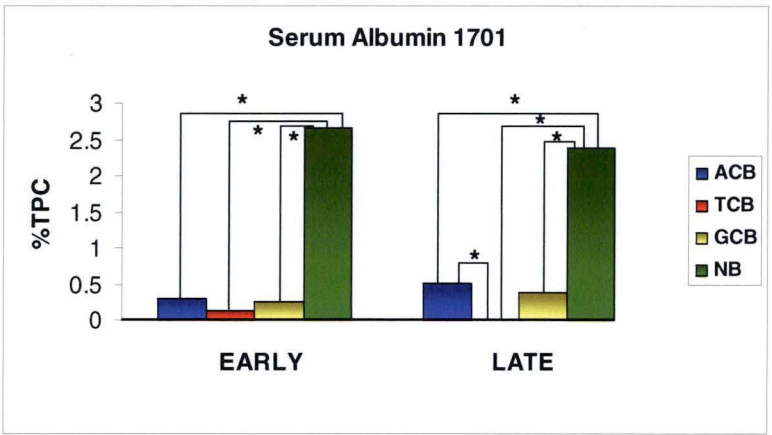


Plate 4.1b:

The expression of serum albumin in 1701 shows no difference in pattern between the early and the late onset cancers.

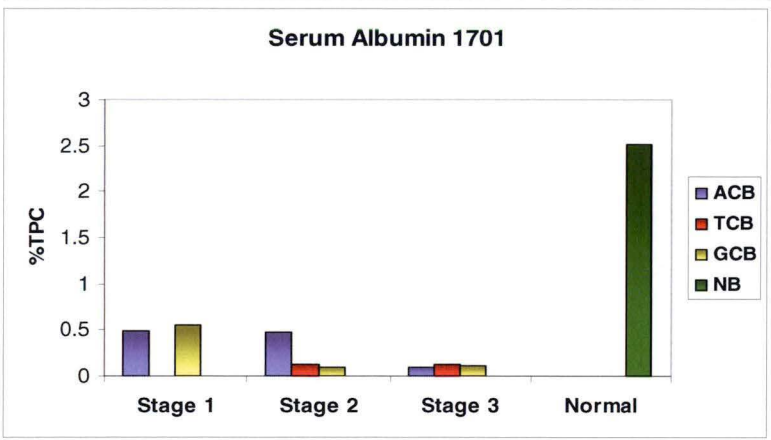


Plate 4.1c:

Staging analysis shows that there is little difference between the stages, although all are significantly down regulated compared to the healthy controls.

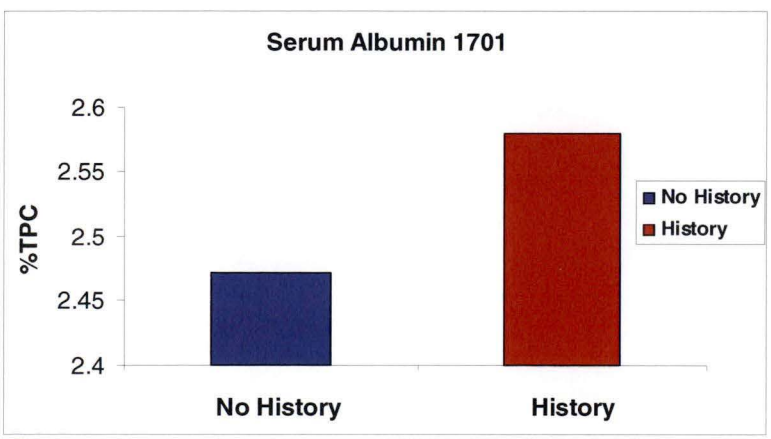


Plate 4.1d:

When comparing the two control cohorts those with a history are slightly up regulated.

Plate 4.2: Osteoglycin 2403

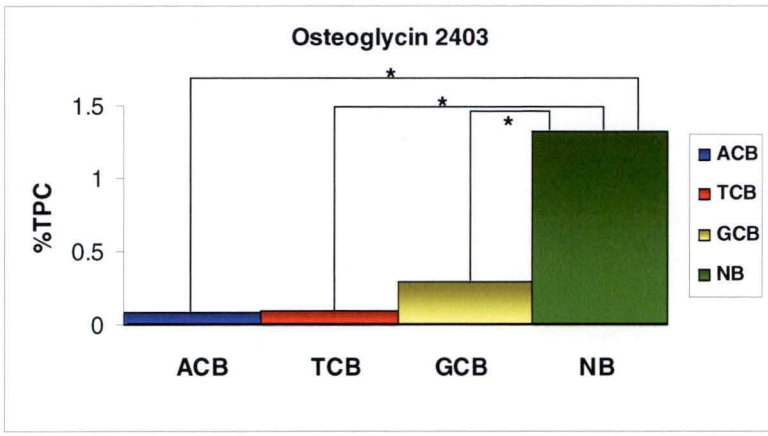


Plate 4.2a:

Comparisons of cancer samples to healthy tissue show a significant down regulation in osteoglycin 2403.

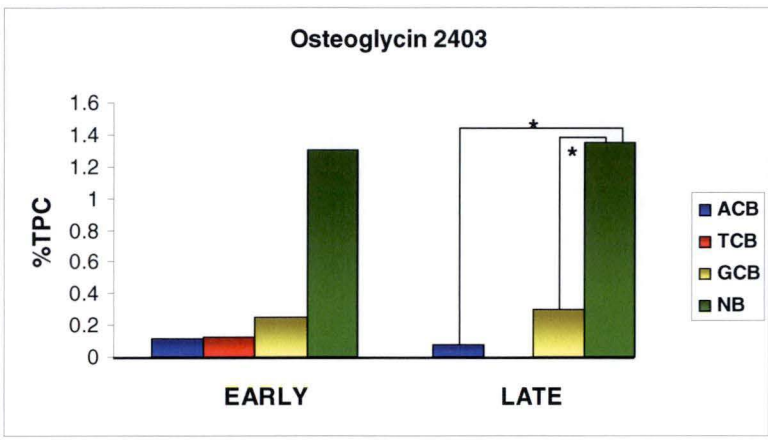


Plate 4.2b:

The comparison of early and late onset patients shows that there is no dependence upon age at onset of disease. The significance between healthy and cancerous tissue is maintained in the Late onset cohorts.

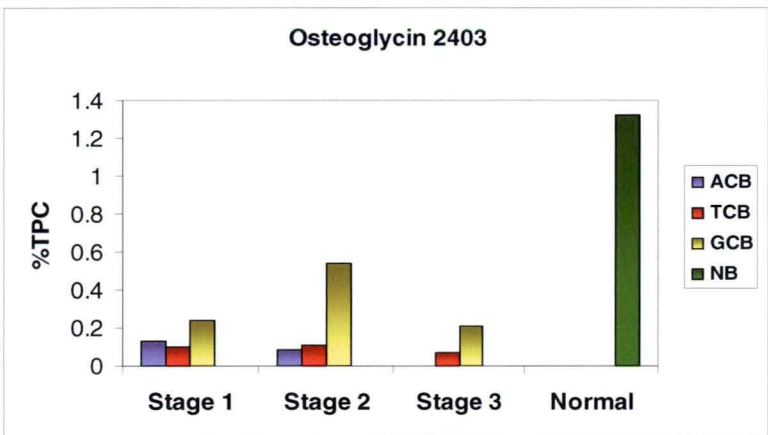


Plate 4.2c:

Staging analysis shows a continual trend of decreasing concentration in progressive stages of cancer. These concentrations were not changed significantly between stages, only between the healthy and diseased tissue.

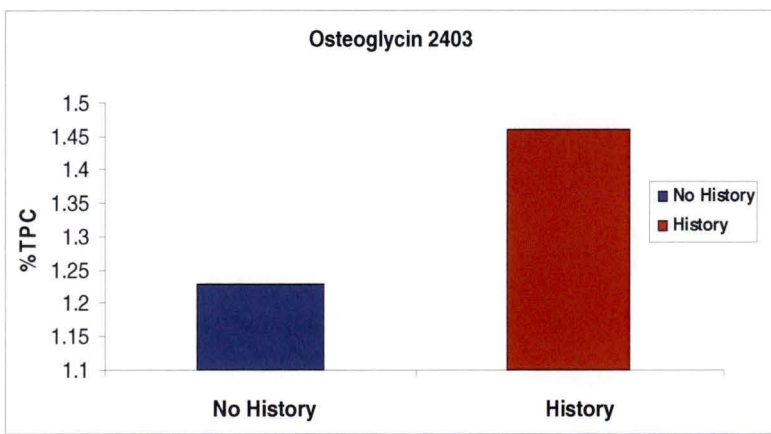


Plate4.2d:

Analysis shows that the concentration of osteoglycin 2403 in healthy tissue is lower in those without a history of breast cancer ($p=0.825$).

Plate 4.3: Osteoglycin 2406

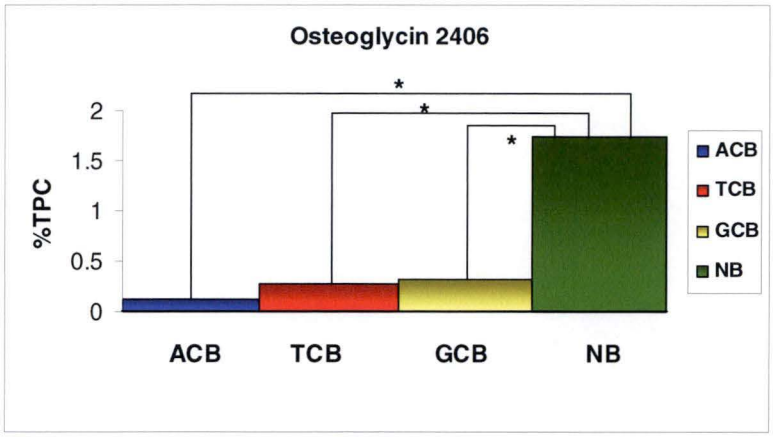


Plate 4.3a:

Analysis shows that there is a significant down regulation in all populations of osteoglycin.

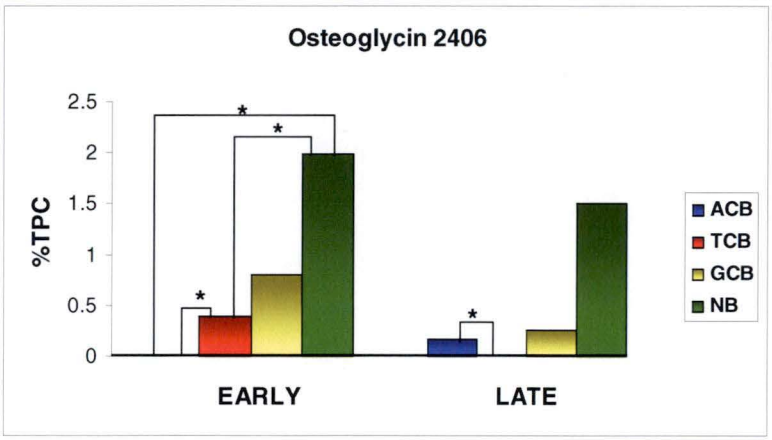


Plate 4.3b:

Comparisons show that there is little change between early and late onset rates of breast cancer. Significance however is only present in the early cancers.

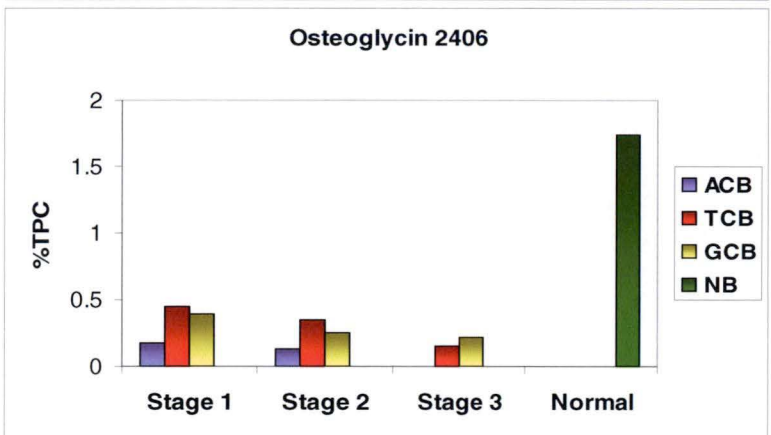


Plate 4.3c:

This figure indicates that the level of osteoglycin is consistent in all stages of breast cancer. A small downward trend may be seen however this was not significant.

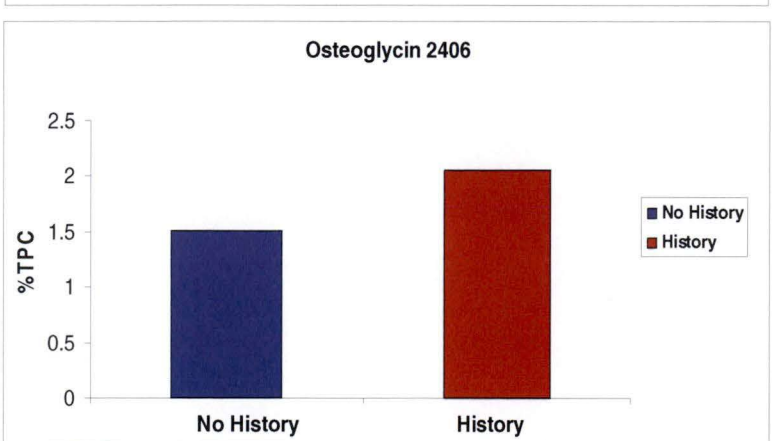


Plate 4.3d:

Analysis showed that there was no correlation of expression to a past history of breast cancer.

Plate 4.4: Lumican 2701

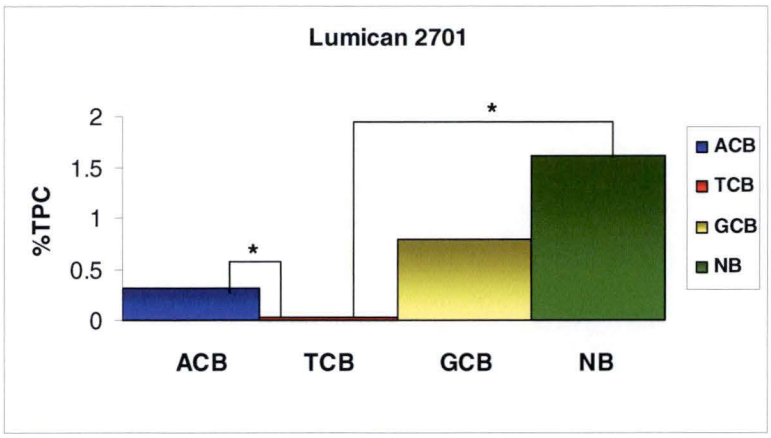


Plate 4.4a:

Lumican is shown to be present in higher levels in healthy tissue. This comparison only reached significance in the Taiwanese population.

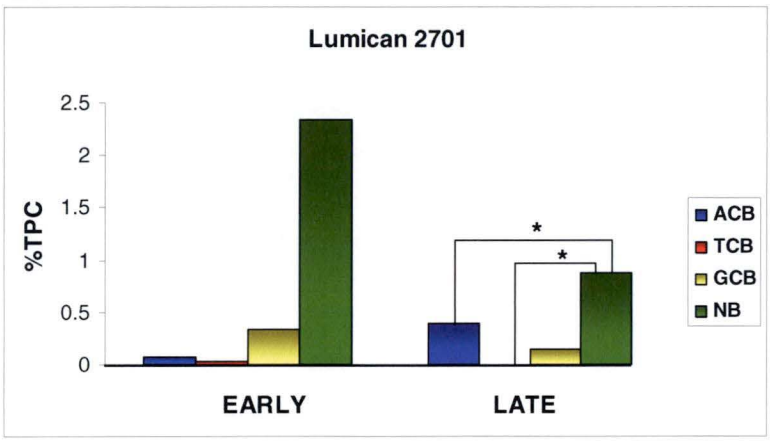


Plate 4.4b:

Significant regulation differences were seen in the late onset cancers when compared to the age matched healthy tissue. Whilst a decrease in lumican 2701 can be seen in healthy tissue this is not a significant change ($p=0.28$).

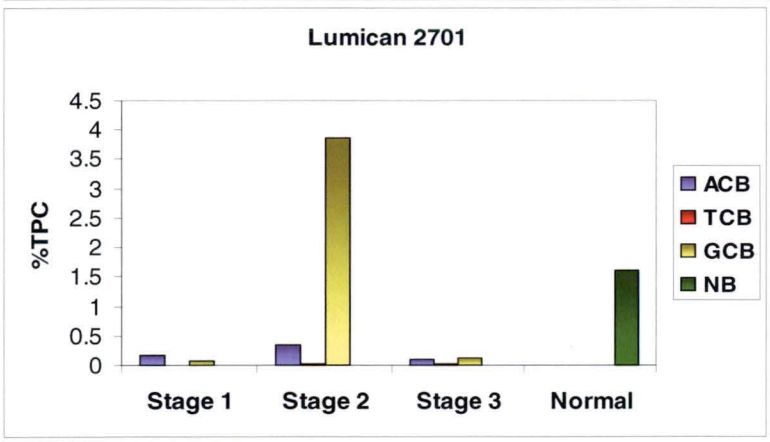


Plate 4.4c:

Lumican is consistently higher in healthy tissue. The only exception is in the stage 2 Greek cohort where the level is higher.

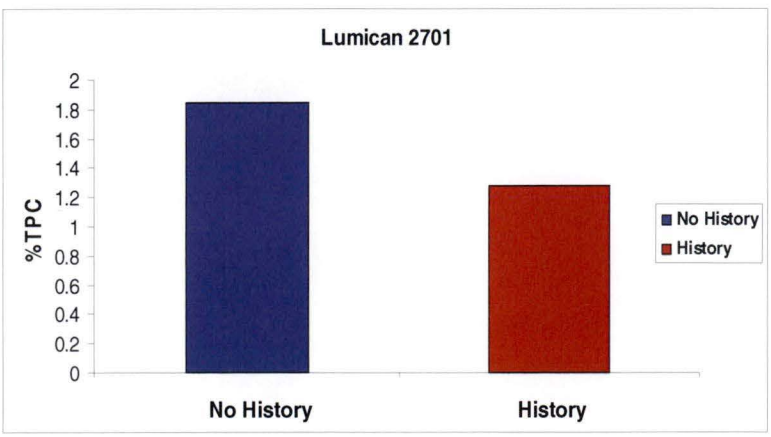


Plate 4.4d:

Comparison of lumican 2701 between the healthy samples with a history and the healthy samples without history, showed little difference in concentration.

Plate 4.5: DIRP5a 2501

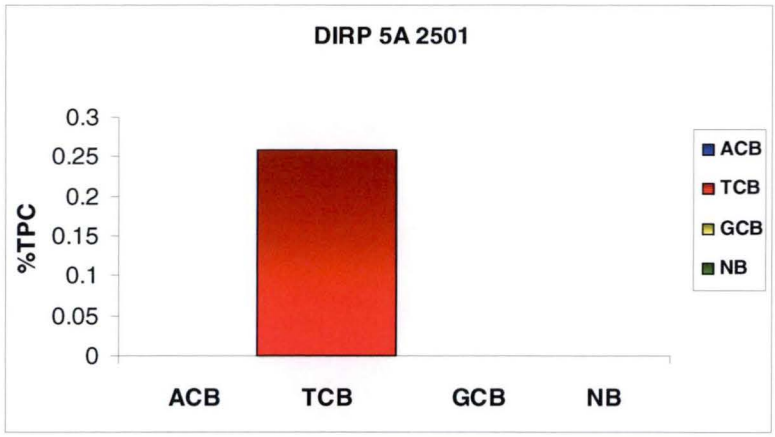


Plate 4.5a:

DIRP 5a was a protein found only in the Taiwanese cancers.

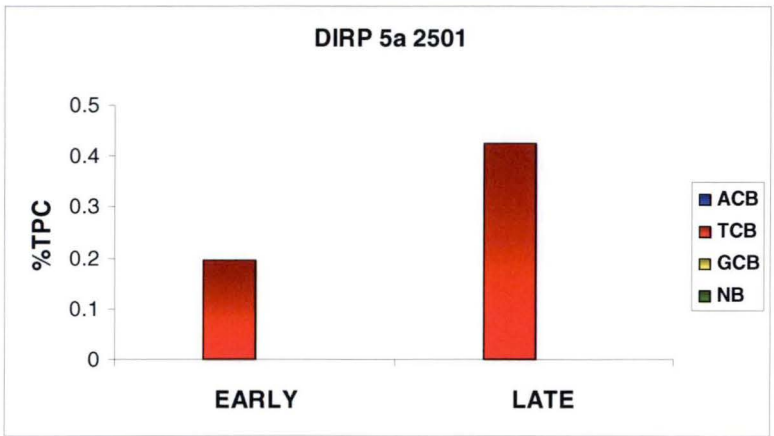


Plate 4.5b:

DIRP5a was seen to be up regulated in the Late onset samples, however this is was not a significant change.

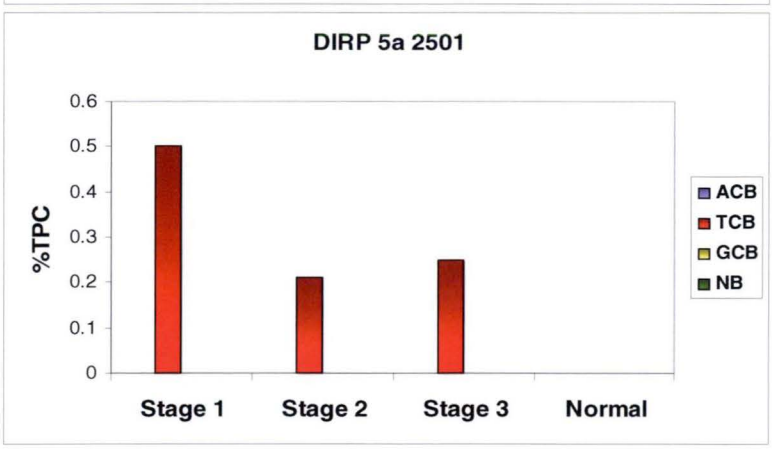


Plate 4.5c:

Normal comparison is not possible for this protein, however levels appear to be higher in the early stages, although this is not significant.

Plate 4.6: Chain A of the Ga Module 4702

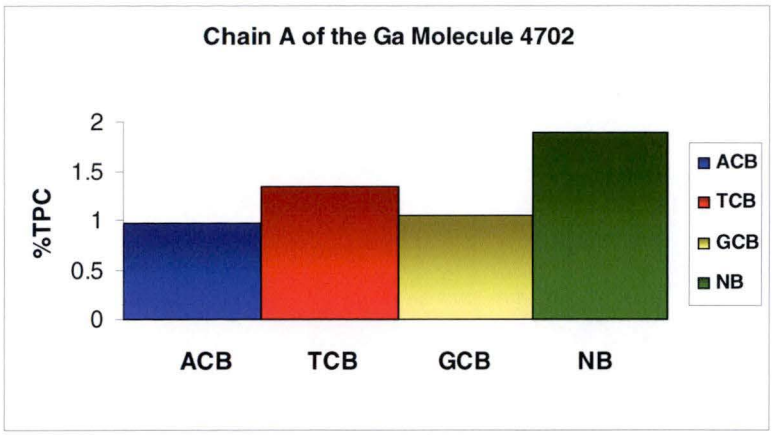


Plate 4.6a:

Chain A of the Ga Module was seen to be in higher concentrations in healthy tissue, however this was not significant.

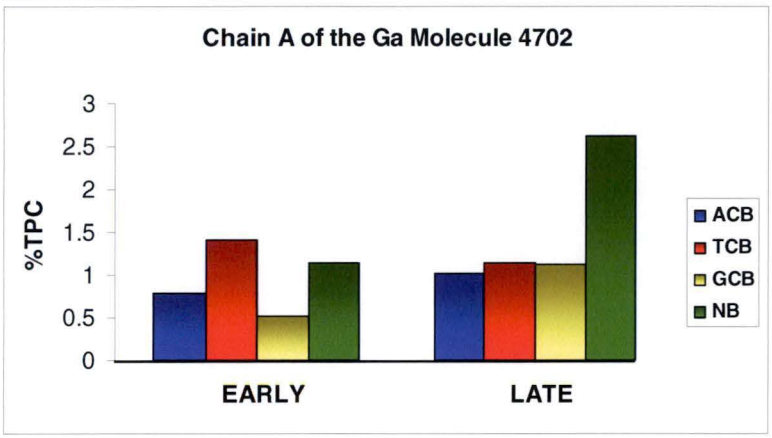


Plate 4.6b:

Whilst significance was not reached, data suggests that there is a possible difference between the concentrations in the early and late cohorts of healthy tissue (p=0.32). The differences seen between healthy and cancerous tissue are not seen in Early onset tissue.

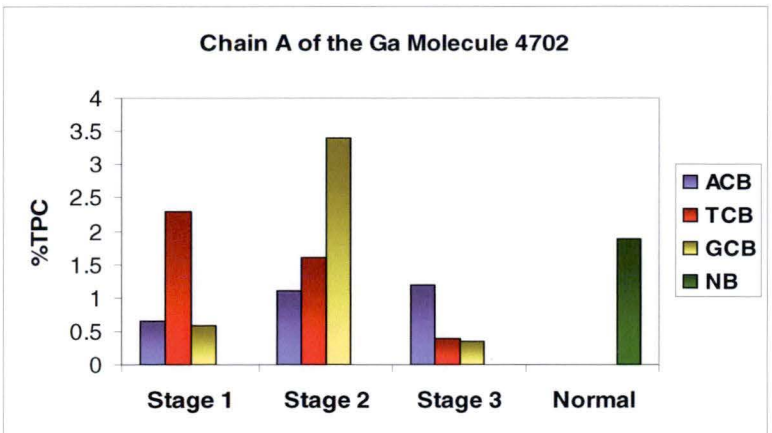


Plate 4.6c:

Staging analysis of the Chain A of the Ga Module protein, shows no regulation differences between the stages of the cancers.

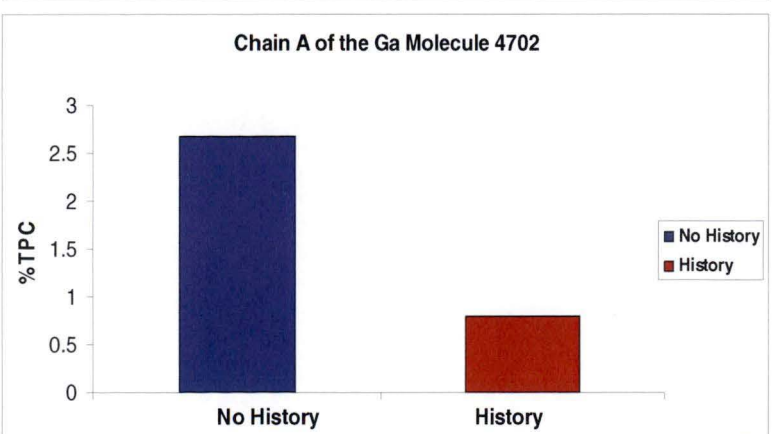


Plate 4.6d:

Analysis of normal tissue and normal with a history indicate that there is a higher concentration in patients with no history of breast cancer. This difference is not significant.

Plate 4.7: Folliculin 3701

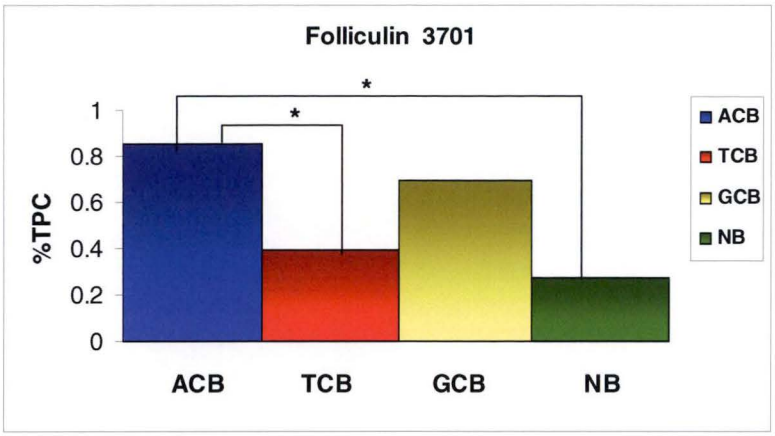


Plate 4.7a:

Folliculin is significantly up regulated in the ACB population compared to the control group. No significance was reached in the comparison of healthy tissue to the TCB or GCB populations

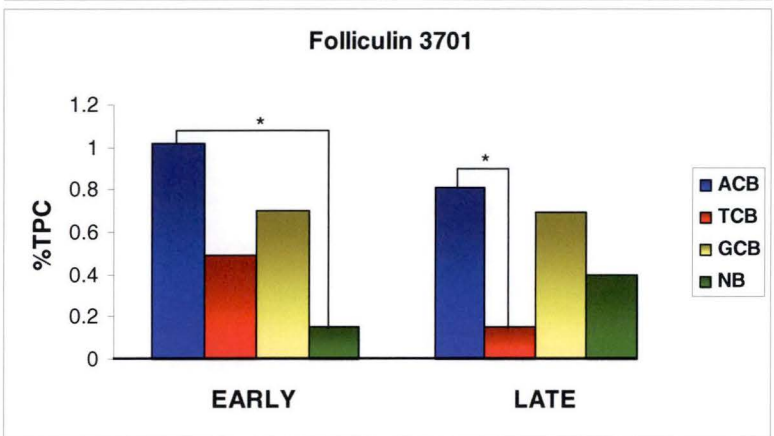


Plate 4.7b:

Significant results were seen in the early onset cohort between the Australian cohorts, this was lost in the late onset patients, where there is evidence of an increasing concentration with age in the control group, although this increase is not significant (p=0.16).

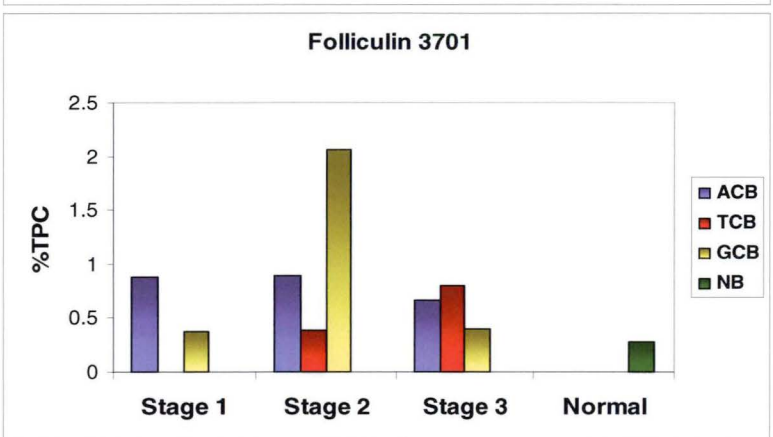


Plate 4.7c:

Staging analysis shows that there is a higher concentration in all stages of cancer compared to healthy tissue. Significance between stages was not present, except for TCB stage 1/2 comparison. ACB samples were significantly different to the control in all stages except 1/Normal comparison.

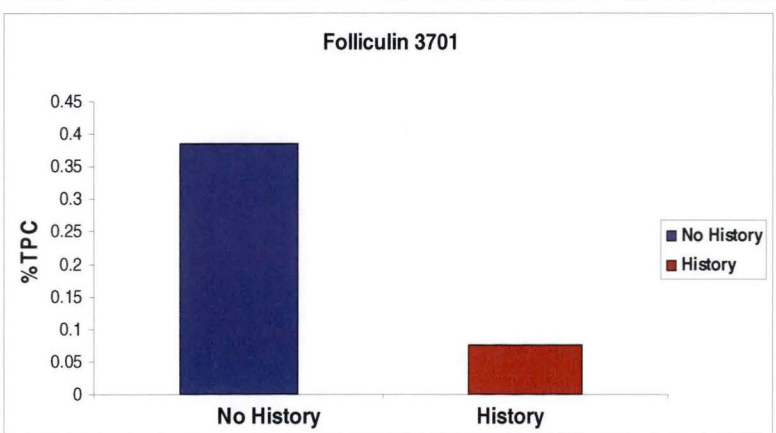


Plate 4.7d:

Healthy cohort comparison shows a higher concentration level in those without a history of breast cancer. This is not significant p=0.08.

Plate 4.8: Heat Shock Protein 3702

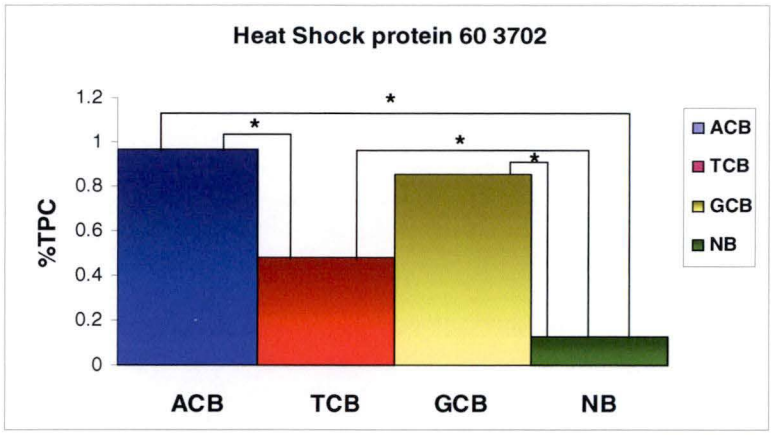


Plate 4.8a:

HSP 3702 is significantly up-regulated compared to controls in all cancer populations.

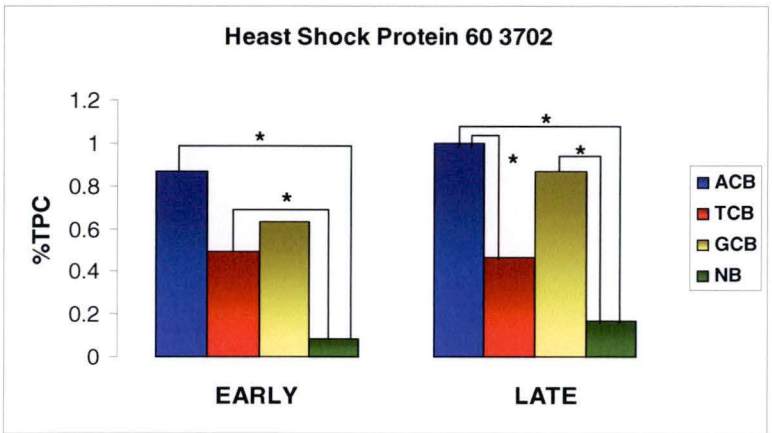


Plate 4.8b:

Early versus late analysis showed no change in the expression pattern of HSP 60 with increased concentration seen in all cancerous cohorts in both age sets.

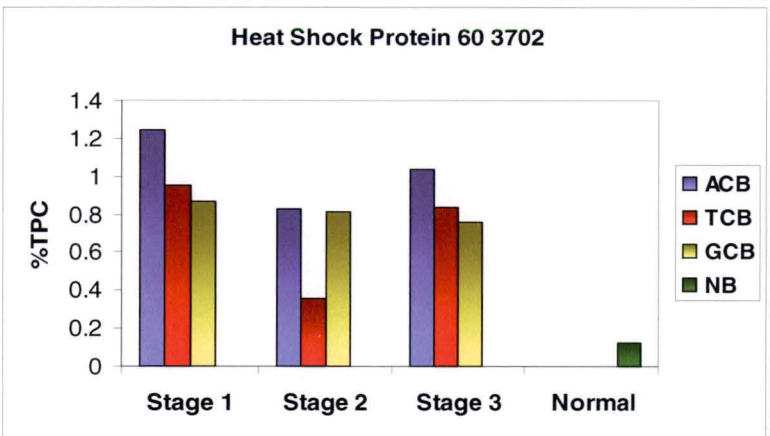


Plate 4.8c:

Staging analysis shows that the cytoplasmic concentration of HSP60 is consistent in all tumour stages.

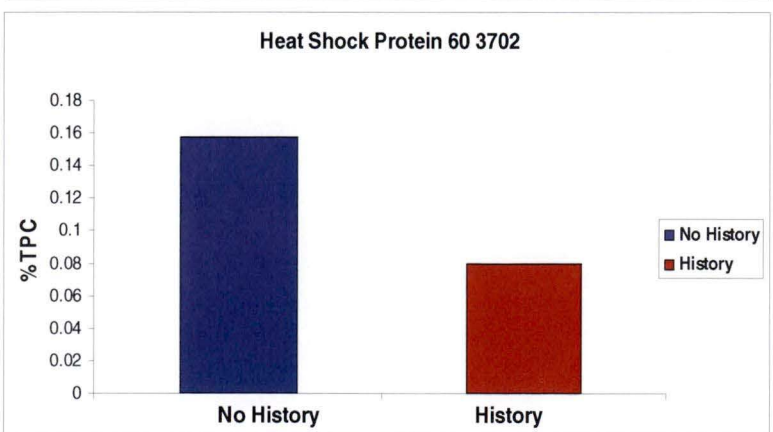


Plate 4.8d:

Controls with a history of breast cancer have a slightly lower concentration of HSP 60 however this is not significant.

Plate 4.9: Heat Shock Protein 4103

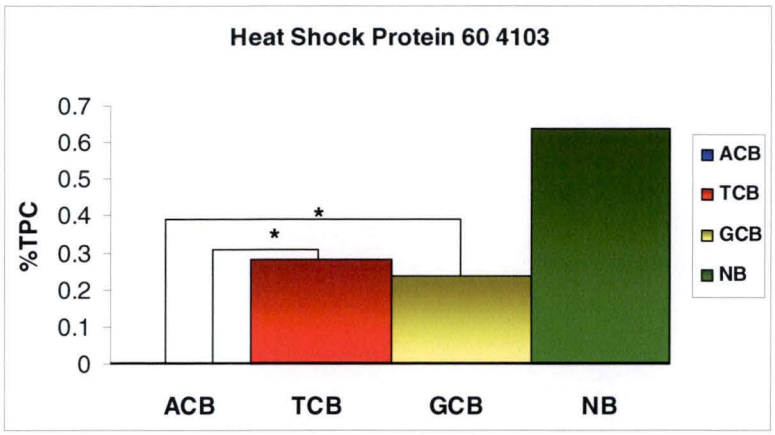


Plate 4.9a:

Analysis shows a down regulation in cancerous tissue, and HSP 60 is entirely absent in ACB samples.

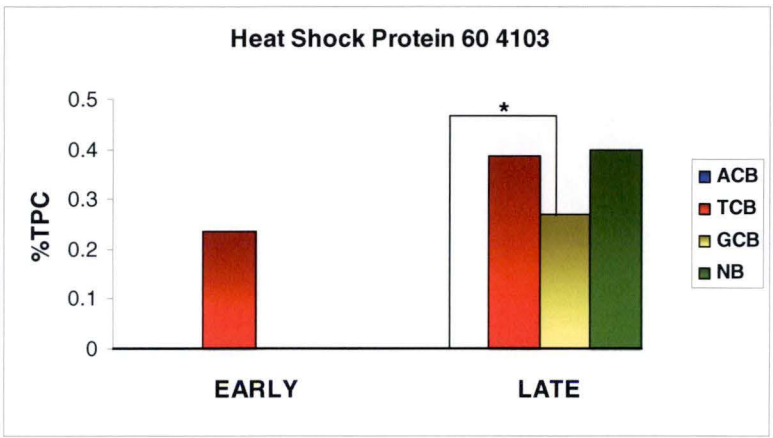


Plate 4.9b:

Trends indicate that late onset breast cancer has higher levels of HSP 60 4103. The healthy samples only express HSP 60 4103 in older patients.

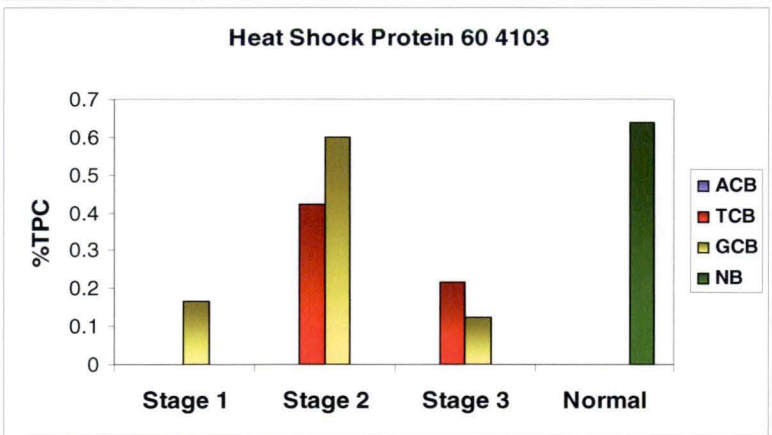


Plate 4.9c

Staging analysis shows a peak in Stage 2 in the TCB and the GCB populations. The concentrations seen however are not as high as those in the healthy patients.

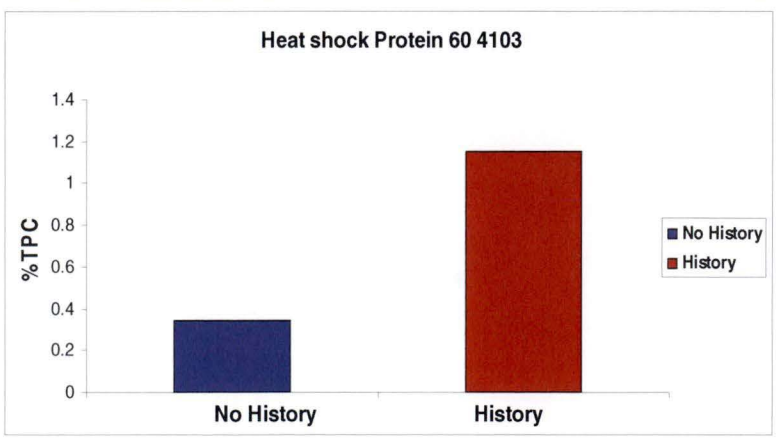


Plate 4.9d:

Concentration is much higher in the controls with a history of breast cancer. This level is not significant.

Plate 4.10: UDP Glycosyltransferase 7112

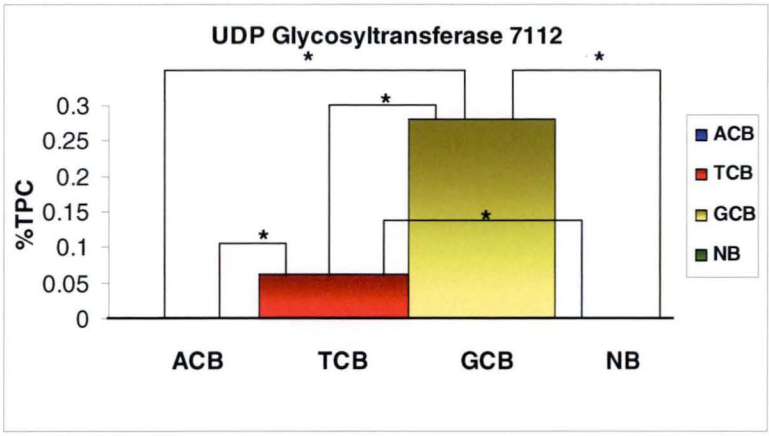


Plate 4.10a:

Cancer versus healthy analysis shows that there is no expression in Australian samples.

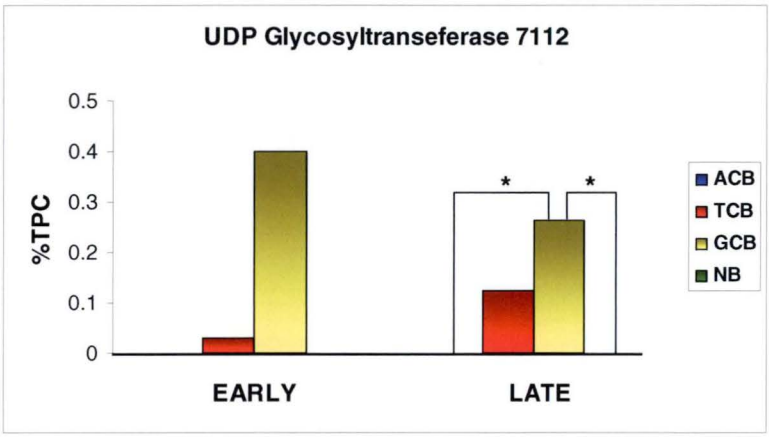


Plate 4.10b:

Age based analysis proved to be inconclusive with no significant changes between early and late onset cancers for the TCB or GCB cohort.

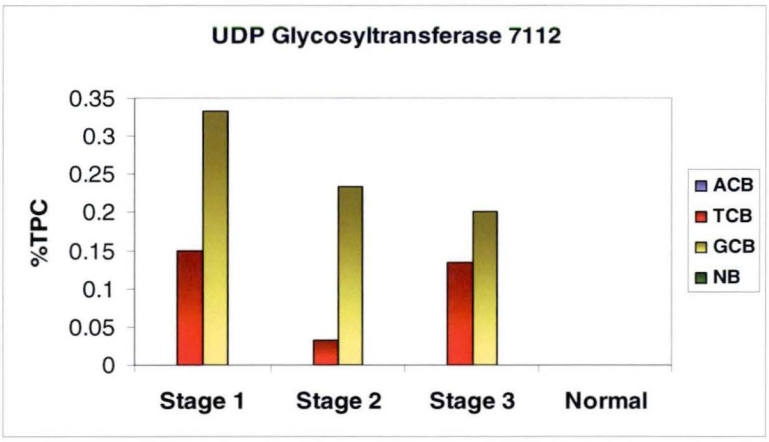


Plate 4.10c:

Staging analysis shows no conclusive results, however average concentration was higher in stage 1 than any other stage for both the GCB and TCB populations.

Plate 4.11: FLJ 20309 7204

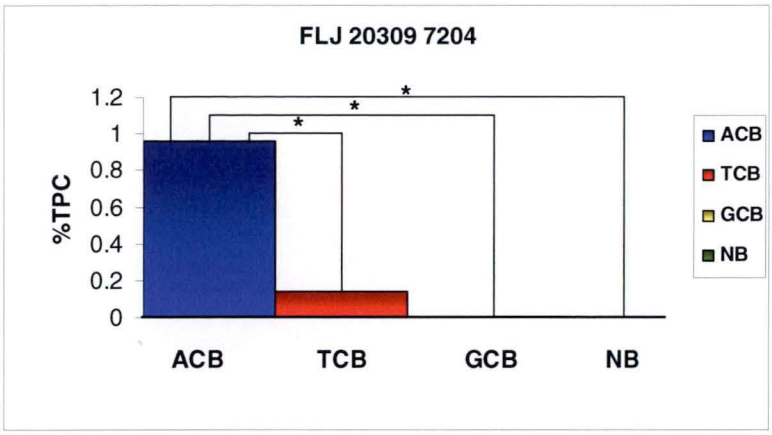


Plate 4.11a:

This protein was strongly expressed in ACB, with some TCB patients expressing it. There was no expression in healthy tissue.

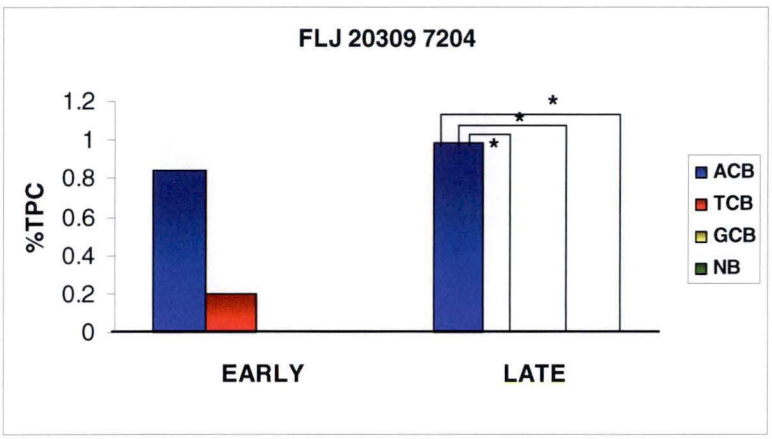


Plate 4.11b:

Analysis shows that there is no age dependence of this protein, with both age groups showing equal concentrations.

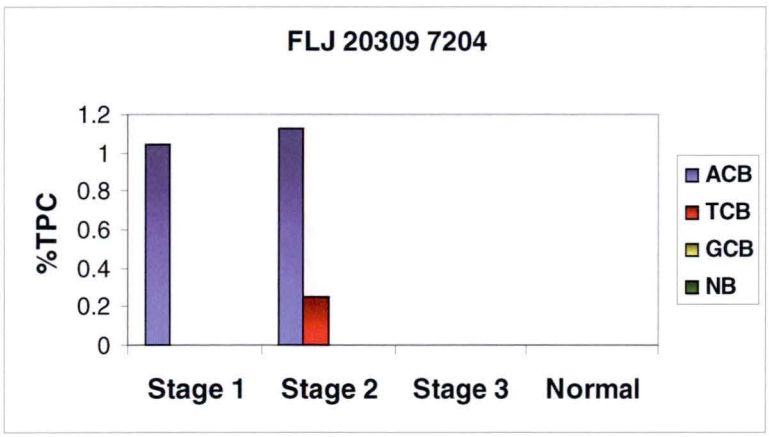


Plate 4.11c:

Staging analysis shows that this protein is expressed in the earlier stages of cancer.

Plate 4.12: Rp3-393D12.2 8203

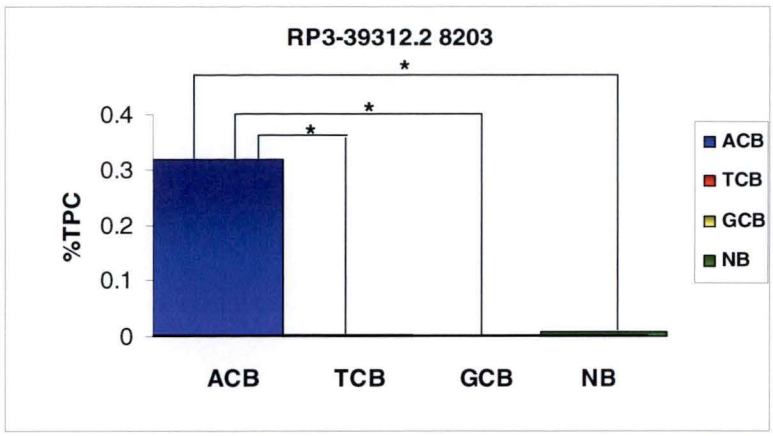


Plate 4.12a:

Concentration levels indicate that there is a large increase in concentration in ACB compared to NB.

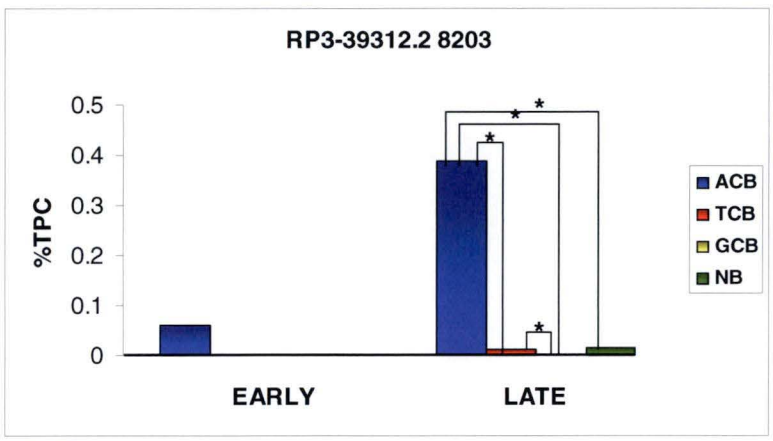


Plate 4.12b:

Onset based analysis indicated that this protein may be more involved in the late onset patients.

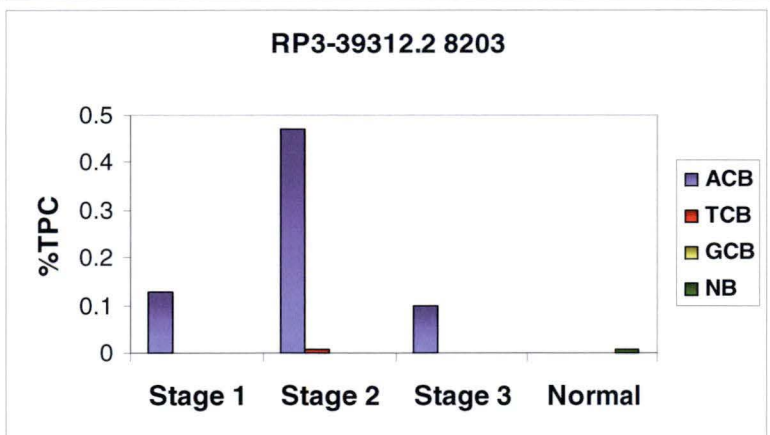


Plate 4.12c:

Staging analysis indicates a peak in level two however the concentration in all stages is above that in the control cohort.

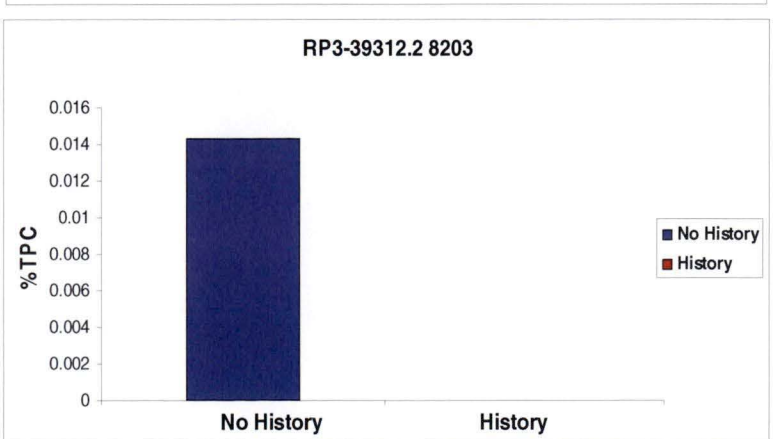


Plate 4.12d:

The low concentration seen in the healthy tissue is only in those patients that do not have a history of breast cancer.

Please note the change in axis for this graph. As the NB concentration was low, axes were varied for graphical purposes.

Plate 4.13: L-Plastin 7203

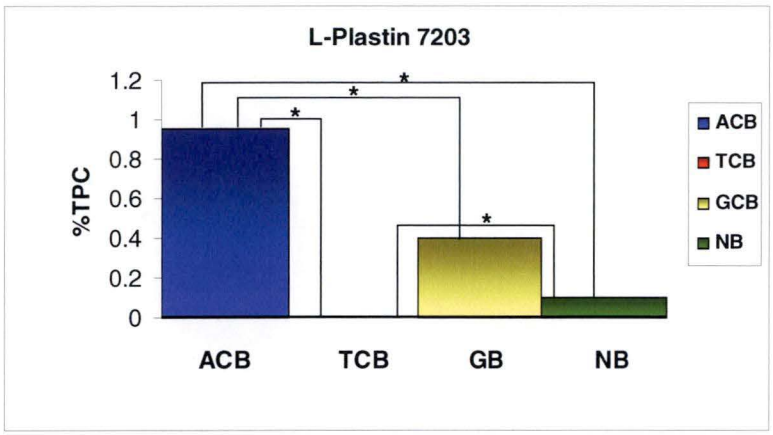


Plate 4.13a:

The cancer cohorts were up regulated compared to the healthy cohort. This difference in regulation was only significant for the ACB comparison to healthy tissue.

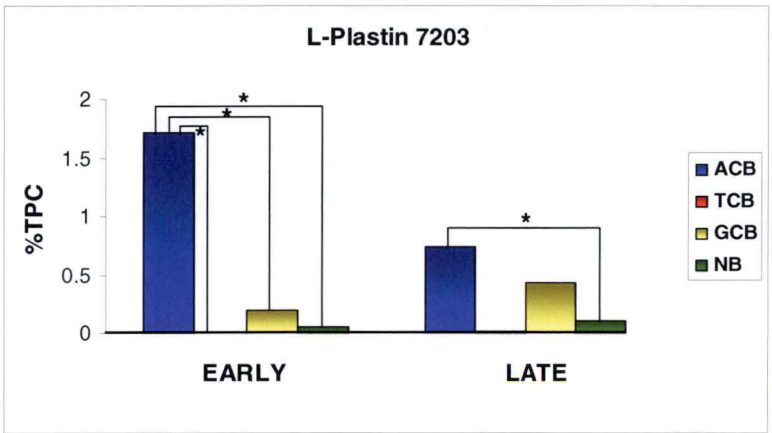


Plate 4.13b:

Early versus late comparisons indicate that there is a difference between the levels in the onset cohorts of the ACB patients. This change is not present in the healthy tissue.

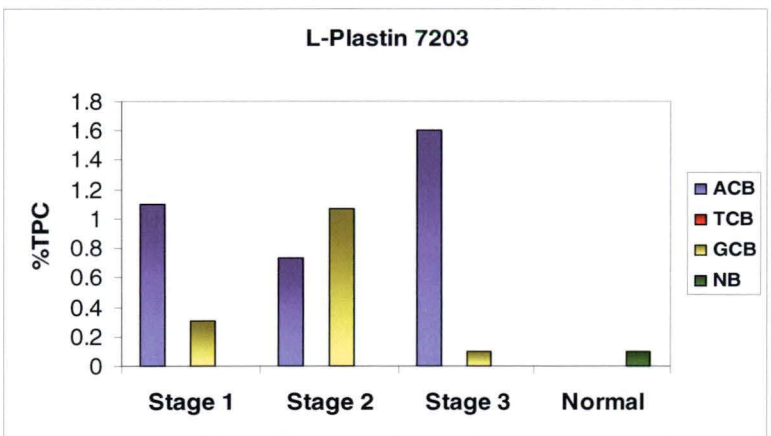


Plate 4.13c:

Staging analysis was inconclusive with concentrations levels varying between stage and population.

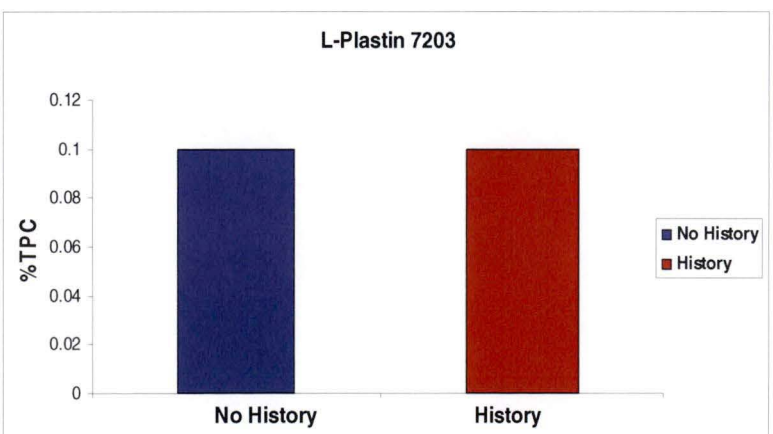


Plate 4.14d:

Comparison of healthy tissue showed no difference between those that did and those that didn't have a history of breast cancer.

4.4 DISCUSSION

4.4.1 SERUM ALBUMIN

In the cancer versus healthy analysis one human serum albumin protein (HSA) was found to be significantly altered between healthy and cancerous tissue. The identification was defined in the MS data obtained from the Australian Proteome Analysis Facility (APAF) as a tentative identification due to a low MOWSE score of 30. The MOWSE score is the product of an algorithm that takes in to account parameters such as the likelihood of the derived peptides matching other proteins, the experimental Mw and pI. A value of 65 or higher is defined as significant identification. Confirmation of the proteins identity would require a lengthy set of further experiments as antibodies to specific isoforms of HSA are unable to be obtained commercially.

The concentration of this particular form of HSA was lower in the cancerous tissue which indicates several possibilities. This particular HSA sequence was identified as several of the proteins sent for sequencing, suggesting several isoforms of HSA within the cell (see Chapter 5 for the results concerning these alternate isoforms). This particular isoform was cytoplasmic and was 10kDa below the theoretical molecular weight which may contribute heavily to the low MOWSE score that was achieved. This lower than expected Mw may be due to cleavage of the protein, this is unlikely though as the MS analysis aligned peptides to within 30 amino acids of both the C and N termini.

Post translational modification is possible although also not likely due to the reduced nature of the sample, which was intended to eliminate post translational modifications. The reduced nature of the protein however may not be assumed, as was seen in preliminary data on the volatility of TBP, the primary reducing agent employed in this study[106].

Also possible and most probable is the possibility of splicing variants from alternate gene transcription. Splicing of the HSA gene has been reported in Watkins *et al* [107] and Madison *et al* [108] which has shown several forms of circulating HSA. In HSA 1701 there were no sequence alignments between amino acid 300 and 569 of the HSA

sequence which adds strength to the potential of a splicing variant, eliminating that section of the protein.

The function of this isoform of HSA in breast cancer is unknown however it is possible that it is a transporter protein. HSA involvement in the movement of proteins from the extracellular environment to the intracellular environment has been reported in the literature for the breast cancer cell lines MCF 7 and MDA-MB 231 [109, 110]. An increase in circulating HSA has been seen in a previous report [111], where it was been found to be inversely proportional to survival. This study however has found in the case of this variant that it is decreased in the tissue. The lower concentration of HSA 1701 in the diseased tissue may be due to several possibilities. It may be a direct effect of DNA damage to Chromosome 4, where the HSA gene lies. It may also be due to a reduced expression of the protein that it carries, and hence through feedback mechanisms that are yet to be determined, less of this form of HSA is produced as less is required for transportation purposes.

HSA 1701 was shown not to differentially be expressed when the age of onset analysis was conducted. This constant concentration of HSA 1701 irrespective of age suggests that there is role that HSA 1701 plays that is required in the normal functioning of healthy breast tissue. What can't be known without further investigation is the role of HSA 1701 in healthy tissue and how it is disrupted in the cancerous tissue.

Staging analysis of HSA 1701 showed a trend to decrease as tumour stage increased. This raises some important questions about the role of HSA in the intracellular environment in the early stages of breast cancer. If this variant of HSA has a role in protein transport, the protein that it is transporting may be very important in elucidating the role of HSA in breast cancer. Whether HSA is decreased and hence the transporting variant has decreased or whether the protein being transported is decreasing, resulting in little or no synthesis and therefore a decrease of HSA 1701 is yet to be determined.

4.4.2 OSTEOGLYCIN

Osteoglycin is a member of the small leucine-rich proteoglycan family (SLRP). Both isoforms of osteoglycin that were isolated were found in much higher concentrations and more frequently in the healthy patients. The decrease that was seen in the cancerous patients was not altered between the three populations.

The early versus late analysis indicates that there is not an age related factor at play in the decrease of osteoglycin in breast tissue. This is the same with the staging analysis, as no difference in concentration was seen between stages.

Osteoglycin is a protein that was originally located in bone and has since been studied extensively in the cornea as well. Also referred to as osteoinductive factor 1 and mimecan, osteoglycin has been accredited with many potential functions. Kukita *et al* [112] report its inhibitory function on the formation of osteoclast like cells whilst Tasheva *et al* [113] associate osteoglycin to interferon γ regulation. Fernandez *et al* [114] located the protein in atherosclerotic plaques and there has also been an association of osteoglycin to vascularisation by Shanahan *et al* [115].

These very varied reports of osteoglycin function make it difficult to decipher at this stage what the proteins involvement in breast cancer might be. One study by Hu *et al* [116] suggest a role for osteoglycin as an effective marker for the diagnosis of pituitary tumours. Hu *et al* also showed osteoglycin expression in healthy tissue with a drop in the level of frequency in the tumour tissue. They however do not allude to the role which osteoglycin may play in tumour development or progression. Lee *et al* [117] has also shown a down regulation of the gene expression of osteoglycin in gastric cancers.

The best evidence to date of osteoglycin involvement in cancer is in a report by Tasheva *et al* [118]. They showed that that there is a p53 binding site on the osteoglycin protein, which activates the transcription of the osteoglycin protein. p53 was not investigated in this study and therefore may indicate where the intracellular role of osteoglycin may lie.

4.4.3 LUMICAN

Like osteoglycin, lumican is a small leucine-rich proteoglycan (SLRP). In this study we found lumican 2701 to be down regulated in breast cancer tissue compared to healthy controls. This down regulation is not fully supported by the literature on lumican function in breast tissue. Some studies have found that lumican is up regulated in the tumour compared to the healthy surrounding tissue[119], others have found it not to be up regulated in the tumour cells, but in the surrounding fibroblasts [120]. Some have found that there is a relationship between low levels of lumican expression, poor prognosis and tumour histology [119, 121, 122].

There was no link between tumour stage and lumican 2701 in this study. There was no indication of lumican 2701 being involved in early onset tumour progression, as seen by the results from the early versus late analysis suggesting no sign of being age related, as was suggested by one study [123].

Of most interest in regard to lumican 2701 in breast cancer is the study by Leygue *et al* [119], where they describe some anomalies of lumican expression. Lumican in the protein form was found not to necessarily correlate with lumican mRNA *in situ.* This supports the possibility of lumican having alternate forms, which supports what has been found in this study, with other detected isoforms occurring in the following results chapter, Chapter 5.

Understanding the role that Lumican plays in the tissue is difficult with contradicting accounts of its behaviour within the tissue and in *in vitro* studies. One microarray study shows up regulation of the LUM gene, which encodes lumican, in all glioblastoma multiform brain tumours [124]. This, along with the up-regulation of mRNA seen in human breast cancers indicates that it is possible that not all forms of lumican have been fully defined.

4.4.4 DIRP5A

Only expressed in Taiwanese patients, this protein disulfide-isomerase related protein 5a (DIRP5a), is most likely a splicing variant that is specific to the Taiwanese population (ref splicing paper.). In a previous study DIRP5a has been demonstrated to be up regulated in prostate cancer cell lines [125]. The results from this study as well as these other two studies suggest a need to look further at the role of this protein in breast cancer.

DIRP5a is located in the endoplasmic reticulum. Where it was located in this study is a little harder to determine as it was found in the mildly hydrophobic fraction which may incorporate proteins from the endoplasmic reticulum. With the sequential extraction techniques used it is difficult to say for sure whether the detergents would have broken down the membrane of the endoplasmic reticulum, or dissociated the protein from it.

There has been evidence to suggest that DIRP5a co-locates on a 2D gel with a heavy DIRP5a isoform [126]. As the theoretical and experimental Mw's are 46kDa and 43kDa

respectively, the results here appear to have isolated an unbound form of DIRP5a. Also the co-location of any protein would affect the Mass Spectrometry that was conducted and as a result lead to a failure to identify the protein. In the Maniratanachote *et al* study the co-location is possibly due to different extraction methods and in particular differences in the reduction of the protein-protein interactions within the samples processed [126].

Hayano and Kikuchi [127] have shown that DIRP5a is stress inducible in cancer cell lines. Up regulation of this protein is not determinable is this study as this isoform of DIRP5a was not seen in any other population. It would be necessary to gain healthy Taiwanese samples to look at this isoform specifically to determine whether it is up regulated or different in the populations at a genetic level, leading to many potential studies to clarify this isoforms role as a marker of breast cancer.

4.4.5 CHAIN A OF THE GA MODULE

The presence of this protein is a peculiar one. Chain A of the GA module is a bacterial protein related to PAB (peptostrepotcoccal albumin binding protein)[128]. As it is found in several bacterial species endemic to the human body, in particular the skin, GI tract, urogenital tract and the oral cavity, the presence in breast tissue is interesting.

The MS results for this protein are very strong, with a MOWSE score of 446 and sequence coverage of 33%. It is therefore unlikely to be any other protein. It may be a human homologue of the Chain A of the GA module protein, however the existence of such a protein has not been reported in any other literature previously.

The presence of this protein having bacterial origins that have the potential to be skin borne automatically requires the discussion of potential sample contamination. There are several reasons why this was deemed unlikely to be the case. The samples were obtained originally from three separate locations around the world, using different doctors, and surgeons, so the scenario of one contamination source at the point of collection is not plausible.

The second possibility is that its' presence is due to post collection contamination, most likely at the time of processing, this too is highly unlikely. Due to the continuing nature

of this study the ACB, GCB, TCB and NB samples were handled and extracted by different people at different times and hence it is unlikely to be the case.

Thirdly the consistency of the samples from healthy tissue having higher concentrations than the diseased tissue, despite not reaching significance, all favours the result being a reflection of the samples as opposed to sample contamination.

Having accepted the results as a true reflection of the proteome, the presence of a bacterial protein in a human soft tissue sample is interesting. Reports have shown that Chain A of the GA module is present in strains of *Finegoldi magna* (previously known as Peptostreptococcus magnus) and some Streptococci. It has also been reported that the presence of a GA module increases the bacteria's virulence. These bacteria have also been directly associated with wound infections and soft tissue abscess [129]. This however confuses the results seen as there is a decrease in the number of patients expressing the protein in the diseased tissue, which wouldn't be expected of bacteria that cause infections. There are several possibilities as to why this is so.

When looking at the healthy with history analysis it appears that there is an increase in those patients that have a history of breast cancer, with 6 out of 7 healthy samples having a high expression level. When looking closer at the no history patients it appears that there is only 1 out of 4 samples and this level is quite high, skewing the results between the two subsets of healthy patients. The frequency is higher in patients with a history, and may therefore be the result of breast cancer and associated treatments; lowering the immune system and allowing the establishment of bacteria within the tissue or it could be an early indicator of breast cancer and have a primary role in its initiation.

Also the presence of a human analogue of this protein has not been confirmed nor deemed not to exist. It is possible therefore that it exists and this is what has been identified here. In the bacterial system the Ga module is a protein that conjugates with HSA, hence there may be a human protein that under normal healthy conditions binds with HSA but functions differently when bound by the bacterial Ga module by either resulting in new signalling cascades, or truncating those that would normally occur. Another explanation may be that the potentially suppressed immune system of the patient is susceptible to the bacteria which is endemic to a large majority of people, with

Finegoldi magna being the most common pathogenic gram negative bacteria [130]. This may explain the increase in the older, healthy population where the levels are increased, potentially due to an ailing immune system. Without further conformational studies into this protein the reason for its presence is speculation at this stage.

4.4.6 FOLLICULIN

The gene encoding folliculin has been associated with several diseases. Inherited mutations have been associated with a condition known as Birt-Hogg Dube syndrome, whilst somatic mutations have been associated with lung disorders, kidney tumours and benign skin lesions[131]. It is thought that folliculin in healthy tissue may act as a tumour suppressor, although results from this study do not support that claim.

Despite it being reported to have role in kidney and lung tumours, folliculin has not to date been associated with breast cancer. In this study significance was only reached when comparing the Australian cancerous and healthy tissue cohorts so it needs to be considered that this may be a difference that is specific to the Australian population, maybe even the Tasmanian population given the nature of sample selection. The concentration levels indicated a trend of up regulation in the other populations. A healthy control group for each may be required to determine the exact function that this protein may play in the future as a prognostic or diagnostic marker with a global reach.

In the early versus late analysis that was conducted there is a rise in the older healthy patients however as this group contains all of those patients with a history of the disease it is difficult to pinpoint age, or previous history as the reason for this increase in folliculin concentration.

4.4.7 HEAT SHOCK PROTEIN 60

Two isoforms of HSP 60 were identified. The cytoplasmic HSP 60 that was identified was present in much higher concentrations in the cancerous patients and this is likely to be due to the stress response, the number of healthy patients expressing this protein was 58% of healthy patients, however 5 of these had previous exposure to breast cancer, which may provide a reason as to it's expression here. This finding is supported by previous studies that HSP 60 is up regulated in malignancies [132]. The difference

between the patients with and without a history of breast cancer is not a significant one however.

This study supports the findings that HSP 60 is increased in cancer. The increase in malignant samples is not supported by the staging data of this study which showed no increase in HSP 60 with increase in staging.

The second of the HSP 60 isoforms (4103) was isolated from Solution 3, indicating a nuclear/mitochondrial location due to the high level of hydrophobicity. The higher concentration that has been determined in the healthy patients is not unusual as the constitutively expressed protein is stored in the mitochondria until times of cellular stress when it is then translocated to the cytoplasm where it is involved in apoptosis [133] and cell signalling, via a chaperoning function [134]. This cytoplasmic relocation is demonstrated nicely by the second isoform which was found in the cytoplasmic, mildly hydrophobic fraction, and was found to be up regulated in the cancerous patients compared to the healthy patients.

Whilst HSP 60 is not a novel finding when in the search for a marker for breast cancer it demonstrates quite clearly the effectiveness of targeting a search for new markers using the techniques employed for this study by replicating the results that have been previously observed for this protein when using alternate techniques.

This is overshadowed by the low numbers expressing the protein in nuclear form. The number of patients expressing the protein in the healthy samples was very low for the nuclear isoform, this suggest that there is potentially another isoform of HSP 60 in the nucleus, or that its reported constitutive expression is not present in the Australian healthy population. The existence of an alternate form is supported by the low Mw for the HSP 60 protein, suggesting that it may be in the process of being either synthesised or degraded.

4.4.8 UDP-GLYCOSYLTRANSFERASE

This protein was sequenced due to its high level of difference between the early and late onset of healthy to Taiwanese patients as there was a high level of importance placed on this cohort due to its high level in early onset patients compared to the other populations. Later analysis showed no presence in the ACB cohort. This may be due to

several reasons, but the expression of and the concentration at which it is expressed has been shown to be highly dependent upon many factors including genetics [135]. This may be what is seen here, a strongly influenced genetic isoform of UDP-Glycosyltransferase, which is also known as UDP - Glucuronosyltransferase. It is likely that this protein would be found in an altered form if a pan UDP-glycosyltransferase antibody was used to conduct a western blot of a 2D separation, provided that the epitope was on the appropriate polymorphisms of the protein.

There was a lack of significance when there was separation of the cohorts into early and late onsets. The trend of a decrease that was seen in the in the late onset Taiwanese group which the proteins selection for sequencing was based on, is questioned by the up-regulation between the age of onset that was seen in the GCB samples. So whilst the age of onset is unlikely to be of importance in this proteins expression there is still a possibility that there is a role for this protein in cancer occurrence.

The concentration levels of this protein have been reported to be highly variable. It is usually found in the liver where it catalyses glucuronidation [136]. It has also been found to bind to paracetamol and other similar phenolic drugs [137]. There has also been a suggestion that it may be altered in relation to stress via Nrf2, a transcription factor in stress mediated responses. It is this interaction with Nrf2 that may play a key role in carcinogenesis, with studies showing that there is an increase in carcinogen induced urinary tumours in Nrf2 null mice, with decreases in tumour incidence in mice that are treated with Oltipraz a chemotherapy drug that re-establishes the pathways of all UDP-glycosyltransferase 1 enzymes [138].

This study has managed to show that there is expression of the UDP-glycosyltransferase protein in breast tissue. This concentration level appears to be quite variable in this isoform. It also opens the way to look at the genetic variability of the UGT1 gene having strong population specific expression.

There have been studies that indicate that in rats there is an increase of UDP-glycosyltransferase in the early stages of hepatocarcinoma [139, 140], this was shown to be the case in human breast cancers, with expression higher in early stage cancers, in patients that showed expression.

4.4.9 FLJ 20309

FLJ 20309 is a protein that was originally found in a cDNA collection derived from the Japanese population. It has been determined that the gene encoding FLJ 20309 lies on chromosome 2 and there have been 3 variants found, two of 51 kDa and one of 98 kDa. None of these correspond to the 34kDa variant that was found in this study, suggesting that this form may be a variant that is separately synthesised or it is a degraded version of one of the previously observed variants. The origin of the isoform in this study warrants further investigation.

Expression based on genomic data shows the potential presence of FLJ 20309 in a variety of human tissues, including the mammary gland. (http://www.genesniffer.org/examples/T1D/chr2/genes/FLJ20309.htm).

Sequence analysis has suggested that it has a role in chromatin remodelling and is DNA associated, binding to the hiNo8o containing complex. It has been shown to occur in many human tissues and has homologues in many other species [141]. Cytogenetic studies have connected this part of the chromosome to several diseases including IDDM, cataracts polymorphic congenital and retinitis pigmentosa (http://www.jbirc.aist.go.jp/hinv/orphan/pathology.cgi?id=HIX0002768). None of these are specific to tumours, in fact they are heavily linked to the ocular region, a known tumour free area. Whether FLJ 20309 acts as a potential tumour suppressor or oncoprotein is yet to be determined.

Here it can be seen that FLJ 20309 is up regulated in cancerous tissue and is absent in the healthy tissue. This also suggests that it may have an oncogenic effect. Staging data shows a high presence and frequency in both stages one and two. There is no evidence of FLJ 20309 in stage three cancers.

At this stage it is not known whether other isoforms of FLJ 20309 are present in the breast cancer tissue or whether there is just one isoform.

Whilst functional studies are yet to be carried out on the FLJ 20309 protein, once they are done there may be a much clearer reason as to why it is expressed in over 50% of breast cancers examined and not in any healthy tissue.

4.4.10RP3-393D12.2

This was originally an hypothetical protein that was determined from the mRNA of cDNA constructs. Since its isolation, RP3-393D12.2 has been given several names, including FHL5, LIM protein ACT and ACT (ACT=activator of CREM in Testis). As the name ACT suggests it has been determined to be an activator of cAMP-responsive element modulator in testis [142].

Its existence in other human tissue is yet to be determined; however it has appeared in this study of breast tissue. It has also been shown to have a role in spermatogenesis and better understanding of its specific role is required to understand it's presence in breast tissue. It has also been found to be up regulated in numerous tumour cell lines [143]. There has also been evidence to show that it has an involvement in actin binding in the pillar cells of eels, with conclusions in that study suggesting a role in maintenance of cellular integrity [144].

Results here show that RP3-393D12.2 is upregulated significantly in breast cancers, it is possible however that this particular variant is a predominantly Australian specific isoform as it was rarely seen in the Taiwanese population and not at all in the Greek cancers.

The early versus late analysis suggests it has a role in late onset cancers as opposed to the early onset cancers. This finding when put together with the fact that it has been suggested to be involved in cell integrity in time of stress may imply that it is, in fact, a stress response protein and is functioning in a protective manner. Its absence in early onset cancer may be a clue addressing the aggressive nature of early onset tumours.

Looking purely at the cancer sample data it is possible that RP3-393D12.2 is a protein involved in the aging process, but this does not explain the presence in only one healthy sample. Tumour suppression proteins by definition have roles in healthy tissue, suggesting RP3-393D12.2 has either a secondary role in tumour progression as an up regulated protein in response to transcription activation or there is a second isoform of the protein in healthy tissue, yet to be located that changes in response to activation, indicating a true tumour suppressor protein.

Preliminary investigation has located potential epitopes for antibody targeting for further functional examination, with long term sites on therapeutic use of these targets for therapy delivery or function alteration.

4.4.11 L-PLASTIN

L-plastin was originally isolated as a leucocyte plastin protein. But L-Plastin has since shown to be involved in many other functions. Of interest is its up regulation in human tumours of most origins[145].

L-Plastin has been reported to be involved in cellular processes such as proliferation, invasion and loss of E-cadherin in colon cancer metastasis[146].

The exact role of L-Plastin in breast cancer is yet to be elucidated but there has been one study in particular that showed a lack of L-plastin in healthy tissue [147]. This is not the case in this study where 7 healthy samples express the protein, two of these NB2 and NB7 have had no previous association with breast cancer, despite expression though, the concentration is very low and is most likely attributed to the expression mediated by the estrogen and progesterone receptors.

The roles that were elucidated in a colon cancer study suggest a potential metastatic predictive quality for the protein [146], however the presence and concentration was seen to be the same in all stages of breast cancer with no significant variance between stage 1, 2 or 3, only between each stage and the healthy tissue.

There was no expression of this isoform of L-Plastin in the TCB samples. This may be due to potential genetic variants.

4.5 CONCLUSIONS

This chapter has identified many proteins that have potential involvement in breast cancer. Many of the proteins identified and discussed here, in combination with further research will offer new information in terms of how proteins are altered in the development and progression of breast cancer.

Some of these proteins may not offer any further use to breast cancer treatment, diagnosis, or prognosis. Others like FLJ 20309 however, offer huge potential as

candidates in targeted therapy once more information has been established in regard to its structure and specific function.

5 DIFFERENTIAL ANALYSIS OF EARLY AND LATE ONSET BREAST CANCER

5.1 SUMMARY 94
5.2 BACKGROUND 95
5.2.1 Age of Onset Cohorts 95
5.2.2 Analysis Approach 96
5.3 RESULTS 96
5.3.1 α-1 collagen VI (2804, 1806 and 1821) 97
5.3.2 Crystallin α-B (8102) 98
5.3.3 Ig L-chain, V-region (5202) 99
5.3.4 Cofilin (7003 and 7004) 100
5.3.5 Lumican (1805 and 2802) 100
5.3.6 Proapolipoprotein (2101) 102
5.3.7 Serum Albumin (501 and 3804) 102
5.4 DISCUSSION 116
5.4.1 α-1 collagen VI – 2804. 116
5.4.2 Cofilin 117
5.4.3 Crystallin α-B 119
5.4.4 Ig L-Chain, V-Region 120
5.4.5 Lumican 121
5.4.6 Proapolipoprotein 123
5.4.7 Serum Albumin 123
5.5 CONCLUSIONS 125

5.1 SUMMARY

Breast cancer is a disease that presents differently in every woman. There are some commonalities that are recurring and it is these similarities, like E_2 dependence and HER2 over expression that have been targeted in the past for treatment development and determining prognosis. Some similarities have been recognised through the documentation of anecdotal evidence and have only been confirmed by retrospective studies. This is the case for the importance of age at which a woman is diagnosed with breast cancer. A comprehensive literature review by Sidoni *et al* [148] shows that the age at which a woman is first diagnosed with breast cancer is associated with more aggressive tumours. The reasons behind this are speculative however younger women appear to have a higher likelihood of a more aggressive tumour phenotype which can directly influence the survival rate of women in this cohort.

The number of younger women that are diagnosed with breast cancer is relatively high, with 25% being under the age of 50 [2]. This is below the age which standard screening programs are targeted. Sidoni *et al* [148] state that 48% of women under 40 who are diagnosed with breast cancer have genetic defects in the BRCA genes. So whilst familial inheritance plays a large role in the early onset of breast cancer this is only the case for less than half of these younger women. What makes these younger women more susceptible to tumours and at higher risk of aggressive tumours is the focus of this chapter.

By comparing the proteome of tissue samples from women who had late onset cancer to tissue from women who had early onset cancers it is possible to recognise proteins that vary between one age cohort and the other. The distribution of the ages of the samples collected is slightly limited due to the lower incidence of early onset cancers. However with the number of Taiwanese patients that are included in this study it is possible to make preliminary findings and locate proteins that under more intense scrutiny may prove to be important in the aggressive natures of the early onset tumours.

From this analysis 26 proteins were sent for sequencing of which 12 were identified by MALDI-TOF Mass Spectrometry. These proteins varied in their concentrations and frequencies and reveal a collection of proteins that may provide potential new markers for prognosis and potential therapies, given more intensive studies in the future.

5.2 BACKGROUND

5.2.1 AGE OF ONSET COHORTS

On examination of the data available for the patients, combined with the knowledge that the tumours of women in lower age groups have a poorer clinical outcome due to more phenotypically aggressive tumours it was decided to investigate further.

The ages of each cohort were plotted (Fig 5.1) and then an arbitrary division of the samples in to groups that contained the patients that were over 53 and under 53 based on an observed separation of the patients collected was performed. These classes were termed "early onset" and "late onset" within each population. The healthy controls were also separated into the under 53 and above 53 age groups and also termed "early" and "late" onset for consistency, despite their lack of disease onset. The patients which were included in each group are listed in Table 5.1.

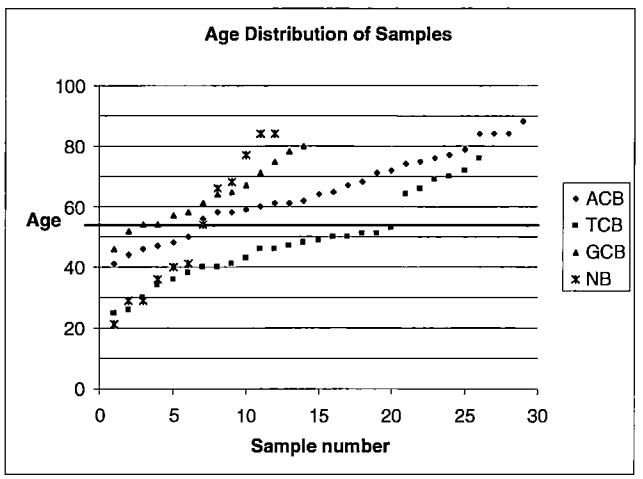


Figure 5.1: Age distribution for each population cohort. Samples were placed in one of two categories, early onset = below 53 and late onset = above 53.

The numbers of patients in the lower age group was relatively low for the Australian (ACB) and the Greek (GCB) population, the number of healthy samples (NB) in each age group provided a 50:50 split of the available samples. From the TCB population two thirds were in the younger age group, making it feasible to look closer at this issue.

This separation of the populations into these classes provided 28 patients in the early onset and 40 patients in the late onset when the cancer populations were combined and 6 healthy samples in each of the age based cohorts.

Healthy		Australian		Taiwanese		Greek	
E	L	E	L	E	L	E	L
NB2	NB1	ACB2	ACB1	TCB2	TCB1	GCB3	GCB1
NB5	NB3	ACB6	ACB3	TCB3	TCB1A	GCB15	GCB2
NB6	NB4	ACB9	ACB4	TCB4	TCB13		GCB4
NB8	NB7	ACB20	ACB5	TCB4a	TCB15		GCB5
NB11	NB9	ACB18	ACB7	TCB5	TCB18		GCB6
NB12	NB10	ACB19	ACB 8	TCB6	TCB24		GCB7
			ACB10	TCB7	TCB25		GCB8
			ACB11	TCB8	TCB27		GCB9
			ACB12	TCB9			GCB10
			ACB13	TCB10			GCB11
			ACB14	TCB11			GCB12
			ACB15	TCB12			GCB13
			ACB16	TCB14			GCB14
			ACB17	TCB16			GCB16
			ACB21	TCB17			
			ACB22	TCB19			
			ACB23	TCB21			
			ACB24	TCB22			
				TCB23			
				TCB26			

Table 5.1: Detailed list of which samples were in the early onset and late onset cohorts when division was based on the age on 53.

5.2.2 ANALYSIS APPROACH

The analysis of the data was very similar to the process adopted for the previous chapter. There is however one difference worth noting. Whilst the proteins of interest for the cancer versus healthy analysis were devised purely from the concentration level between samples, in the early versus late analysis the results from the Taiwanese patients were more heavily favoured due to the higher percentage of early onset individuals present. This was based on the assumption that the because of the larger number of samples in this cohort, protein targets that were found were more likely to be of clinical significance.

5.3 RESULTS

MALDI-TOF MS conducted at the Australian Proteomic Analysis Facility identified 12 of the 26 proteins isolated as of interest in this analysis. Table 5.2 outlines each of the proteins that were sequenced. For each of the proteins successfully identified detailed analysis was conducted.

The plates referred to in the text can be found between pages, 103 and 114. A fold out guide to the legends can be found at the rear of this thesis.

SSP	ID	Mowse Score	Exp. Mw/pI	Th. Mw/pI	Accession number
Solution 2					
1805	Lumican	126	64/5.8	38.3/6.16	gi:642534
2802	Lumican	43	64/6.1	38.4/6.16	gi:30582253
501	Serum Albumin	33	36/5.7	69.3/6.13	gi:23307793
2804	α-1 collagen VI	61	130/6.1	47.8/5.52	gi:30030
2101	proapolipoprotein	464	21/6.1	28.9/5.45	gi:178775
5202	Ig Light chain	39	22/6.9	22.8/6.7	gi:27552515
6603	-	-	38/7.2		
5803	-	-	80/6.7		
7003	Cofilin 1	49	18/7.8	18.5/8.22	gi:57099669
7004	Cofilin 2	42	18/7.9	18.7/7.66	gi:15030332
7201			21/7.8	-	-
7304			24/7.8	-	-
8501			38/8.4	-	-
Solution 3					
1607	-		57/6.1		-
1806	α-1 collagen VI	89	151/6.3	47.8/5.52	gi:30030
1821	α-1 collagen VI	31	93/6.2	47.7/5.63	gi:1915902
2905			149/6.4		-
3804	Serum albumin	199	69/6.9	69.3/6.13	gi:23307793
4302	-	-	36/7		-

4703			58/7.0		
5105	-	-	26/7.4		-
6201	-	-	39/7.6		-
6402	-	-	39/7.6		-
6203			31/7.7		-
7305	-	-	35/7.8		-
8102	Crystallin α B	55	26/8	20.1/6.76	gi:30582379

Table 5.2: Summary of the proteins that were sent for sequencing by MALDI-TOF MS. 12 of the 26 proteins were identified. The data received from this sequencing and the relevant theoretical information is presented for the proteins Tat were identified. A MOWSE score of greater that 65 is a significant identification.

5.3.1 α-1 COLLAGEN VI (2804, 1806 AND 1821)

Several isoforms of α-1 collagen VI were identified through the MALDI-TOF MS analysis of proteins observed to be of interest in the early versus late analysis (See Plates 5.1-5.3).

α-1 collagen VI Isoform SSP 2804 in the cytoplasmic extraction fraction was identified as one of the proteins, suggesting a mildly hydrophobic, cytoplasmic form of α-1 collagen VI. This form was not observed in the healthy tissue (See Plate 5.1a-c). It was seen in a total of 5 cancer patients, 1 ACB and 4 TCB. The ACB sample was from a late onset patient whilst the TCB patients were split evenly with 2 patients in each age of onset based cohort. No statistical significance was found for any of the comparisons conducted, due most likely to the infrequency of which it was found. Also most likely due to this low frequency, there was no significance found in the staging analysis. There was no observed expression of this isoform in any of the healthy samples.

A nuclear form of α-1 collagen VI was isolated and the concentration was higher and expression more frequent. This isoform, SSP 1806 was isolated from solution 3 and was seen in only 1 healthy sample, but was observed in 2 ACB, 13 TCB and 4 GCB patients (See Plate 5.2a-d). A general trend was for there to be an increased expression of SSP 1806 in the older patients but again with the low number of patients exhibiting expression this failed to reach significance in the early versus late analysis or the cancer versus healthy analysis.

SSP 1806 showed no clear pattern in the staging analysis however there were only four stage three samples, all were from the TCB cohort, that expressed α-1 collagen VI. All ACB and GCB samples that expressed SSP 1806 were either stage one or stage two.

The healthy sample that expressed SSP 1806 has a history of breast cancer, and is the only sample to express any of the three α-1 collagen VI isoforms detected in this study.

Plate 5.3a-c shows the analysis results for a third isoform of α-1 collagen VI, SSP 1821 from solution 3 which was seen in 5 TCB patients only. Four of the five patients expressing SSP 1821 were in the early onset group with only 1 late onset patient expressing it. This protein was expressed in much higher concentrations in the early onset patients than in the late onset patient, but when looking at the raw data it can be seen that one of the four patients is skewing the result as it is much higher than the other 3 (Appendix C). No significance is achieved for this protein (Appendix D). Cancer versus healthy analysis for this isoform was deemed unnecessary due to the lack of expression in any other cohorts.

5.3.2 CRYSTALLIN α-B (8102)

SSP 8102 was identified as crystallin α-B. This isoform of crystallin α-B was isolated in the solution 3 extract. Crystallin α-B was present in 100% of healthy tissue, this level of frequency drops to 11% and 37% in the TCB and GCB cohorts respectively whilst the ACB patients had a higher rate of expression at 83% of patients.

In both the early and late cohorts of healthy tissue there were significantly higher levels of crystallin α-B than any of the cancerous populations in either the late or early analysis groups except for the ACB early versus NB early comparison, where significance was not reached (p= 0.44). There were no significant changes in expression for any intra-population comparisons ie: early GCB versus late GCB comparison (See Plate 5.4a-d).

Cancer versus healthy analysis for crystallin α-B showed significance between all populations and the healthy tissue (P=0.004, 0.00015 and 0.004 for ACB, TCB and GCB comparison to NB respectively).

Staging analysis demonstrates no statistically significant difference between the tumour stages as the concentration levels were highly variable.

Analysis of the control samples with and without a history of breast cancer showed that there was no significant difference between the two cohorts.

5.3.3 IG L-CHAIN, V-REGION (5202)

Ig l-chain was identified as SSP 5202 and found to be inconsistently expressed between the populations, with 45% of ACB, 34% of TCB, 62% of GCB and 33% of healthy tissue expressing the protein. The healthy patients showed a decreasing trend in the Ig l-chain expression in the older cohort compared to the early cohort, however this difference was insignificant (See Plate 5.5a-d).

The decreasing trend of Ig light chain from early to late onset was replicated in the GCB samples and is statistically significant (P=0.01). The changes between the early and the late onset cohort when compared to the respective healthy early and late onset cohort were also significant.

The cancer patients in the TCB population showed this trend of decreasing levels in the late onset cohort but the level of change was not significant. The ACB patients maintained levels of SSP 5202 at a constant level.

The staging analysis shows no clear pattern to be consistent between all the populations and no comparisons came back as statistically significant. There was however an observed trend in the ACB samples that showed a decrease in concentration with an increase in tumour severity.

The history versus no history of breast cancer analysis in the control group shows that there was also no pattern. There was a higher average in the samples with no history of breast cancer, but this was not significant (P=0.27).

5.3.4 COFILIN (7003 AND 7004)

Two isoforms of cofilin were identified through MALDI-TOF MS. One of these isoforms, isoform SSP 7003 was not observed in healthy tissue, whilst a low number of patients from all cancer cohorts expressed this isoform of cofilin (See Plate 5.6a-d). The concentration of cofilin SSP 7003 is higher in the older cohort of each population however the low frequency of patients expressing cofilin 7003 has resulted in the findings failing to be statistically significant.

Staging analysis showed that there was no pattern of expression dependant upon the stage of the tumour. The GCB results were consistently higher than the other populations however this was the only consistency in findings.

Cofilin SSP 7004 was found to be up regulated in the cancers by 6.7, 3.4 and 17.6 fold for the ACB, TCB and GCB patients respectively (See Plate 5.7a-d). This isoform was found to be in very low concentrations in the healthy tissue. Like isoform SSP 7003, the concentration is higher in the late onset patients in the GCB and TCB samples but not the ACB or control samples. These results were not significant, which may again be due to the low percentage of the population expressing SSP 7004. Only 16% of healthy patients and 16.7%, 13.8% and 31% of the ACB, TCB and GCB patients respectively expressed SSP 7004. Two healthy patients expressed the isoform, and there was no apparent link to whether or not the patient had a history of breast cancer.

There were no clear conclusions from the staging analysis where the expression level of SSP 7004 appears to have specific association with stage of the tumour.

5.3.5 LUMICAN (1805 AND 2802)

Several forms of lumican were isolated in this study. SSP 1805 was one of the cytoplasmic isoforms that was identified (See Plate 5.8a-d). The frequency of expression of this SSP 1805 was similar in both the healthy and diseased samples with approximately 50% of people expressing this protein in both the healthy and cancer samples.

Statistical significance was only reached when the late onset TCB patients were compared to the late onset NB tissue. The early onset patients in all populations had erratic expression and significance was not reached. Concentration levels dropped much more significantly in the older onset cancers than they did in the early onset patients.

The change in expression level was more evident when the cancer versus healthy analysis that was conducted with statistically significant results in the ACB v NB comparison.

Staging for this isoform showed no significant association between concentration and staging however there does appear to be a slight trend of decreased concentration with increase in staging in all populations but this is not statistically significant.

Control sample analysis of patients with and without a history of breast cancer shows that there is no significant change in concentration between the two healthy cohorts.

SSP 2802 is another isolated protein that was identified as lumican (See Plate 5.9a-d). Lumican SSP 2802 was expressed in 66% of patients whilst the frequency of occurrence dropped to 38% in the Australian and Taiwanese and 56% in the Greek. This, like SSP 1805 was also found to be an isoform in the cytoplasm, however the trends of this protein differ in that they are highly expressed in healthy late onset individuals. The concentration is considerably lower in the cancers of both early and late onset cohorts.

The concentration of SSP 2802 is much higher in the older age group of healthy patients than it is in the younger age group with a comparison between the two healthy cohorts yielding a significance level of P=0.04. When the older cohort is directly compared to the younger cohort in the cancer samples the change in concentration is no longer observed and is reflected in the loss of significance for all populations.

Staging analysis shows that there is a decrease with staging in all but the GCB population. The GCB populations' average is skewed by a high concentration from one patient. Significance is reached for comparisons of stage 2 to NB and stage 3 to NB and between stage 2 and stage 3 in the ACB tissue.

Analysis of the control groups for SSP 2802 showed that there is a higher concentration in those patients with a history of breast cancer than those without, however this does not reach significance with this low number of patients (p=0.09).

5.3.6 PROAPOLIPOPROTEIN (2101)

Proapolipoprotein was expressed in a majority of samples, 91% of healthy samples, 62 % of ACB samples, 93% of TCB samples and 68 % of GCB samples.

The analysis of this protein showed that proapolipoprotein was expressed in higher concentrations in the healthy patients (See Plate 5.10a-d). In both the early versus late analysis and when the cohorts are combined to create the cancer versus healthy analysis, the differences are not statistically significant. In the late onset cohort it can be seen that the concentration of proapolipoprotein is lower in healthy tissue, but remains almost constant in the cancerous tissue, with only a slight decrease in concentration. These late onset results, also fail to reach significance.

When the cohorts are combined to conduct the cancer versus healthy analysis the only significant result is ACB comparison to TCB, suggesting a population, perhaps genetic,

based difference. Significance is almost reached however for the ACB/NB comparison (p=0.054) suggesting that the number of samples, if increased may result in statistical significance being achieved.

Staging analysis shows that there is no clear pattern of expression correlating the concentration to the stage of breast cancer. A peak in stage two in all cancer populations is present but no significance is reached.

Control cohort analysis shows that there is only one patient with a history that fails to express SSP 2101 and that there is a lower expression in those with a history than those without a history although this change is insignificant.

5.3.7 SERUM ALBUMIN (501 AND 3804)

Serum albumin was identified as two proteins of interest in the early versus late analysis. SSP 501 was isolated from solution 2 suggesting it is a cytoplasmic form, whilst 3804 was located in solution 3 suggesting it is a nuclear form.

The frequency of the isoform SSP501 was much lower in the healthy tissue with just 25% of patients expressing the protein. This frequency of occurrence was much higher in the cancer samples with rates of 91%, 38% and 87% for ACB, TCB and GCB respectively.

The concentration of SSP 501 was observed to be higher in the early onset ACB patients when compared to the NB patients (See Plate 5.11a-d). This was consistent in the late onset comparison of ACB versus NB as well as for TCB versus NB, where both comparisons achieved significance. There was an observed change in the concentration level of early and late NB samples. This however was not a significant change.

Staging analysis shows that there is a peak in SSP 501 concentrations in stage 2 for both the ACB and the GCB cohorts whilst no increase is observed in the TCB samples.

Comparative analysis shows that there is high expression in patients without a history of breast cancer compared to those with a history. This higher concentration is not statistically significant.

SSP 3804, also identified as serum albumin was seen in 36% of NB samples, whilst it was seen in 12%, 22% and 50% of ACB, TCB and GCB patients respectively.

SSP 3804 was not present in the early onset ACB samples, however it was observed in the late onset cohort. This was not replicated in the healthy cohorts where a similar concentration of SSP 3804 was seen in both age based cohorts. There were no common trends between populations or cohorts for the early versus late analysis or the cancer versus healthy analysis.

Staging analysis shows that no discernable pattern emerged for the staging when compared to concentration level.

Analysis of the control cohorts that do and don't have a history of breast cancer revealed no significant results.

Plate 5.1 - α-1collagen VI (2804)

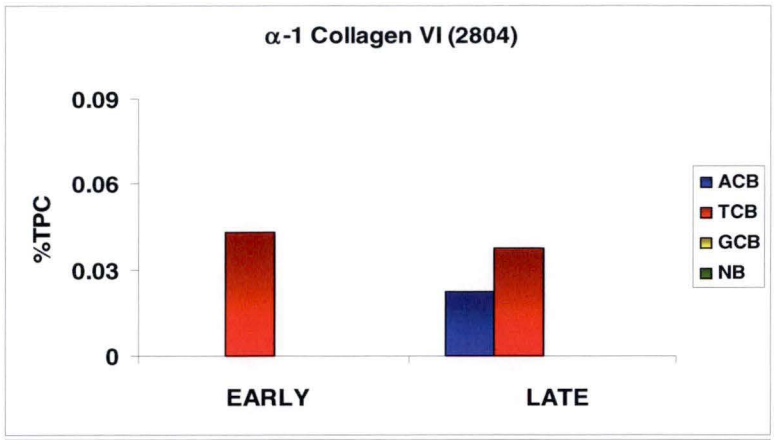


Plate 5.1a:
Early versus late analysis shows that there is no significant relationship between the expression level and age at disease onset.

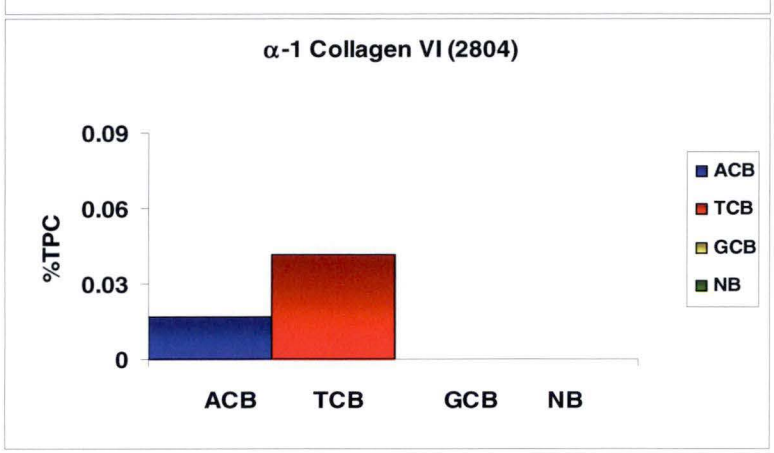


Plate 5.1b:
Cancer versus healthy analysis shows no expression in the GCB or NB cohorts.

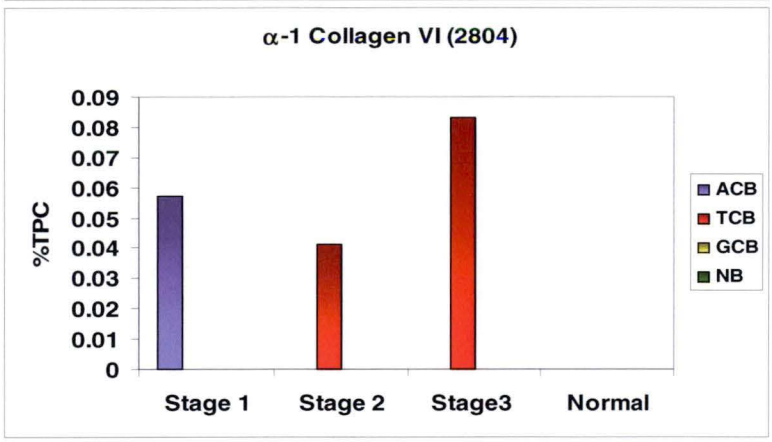


Plate 5.1c:
Staging results were inconsistent, with no trend seen for either ACB or TCB in relation to the amount of expression and the stage of breast cancer.

Plate 5.2 – α-1collagen VI (1806)

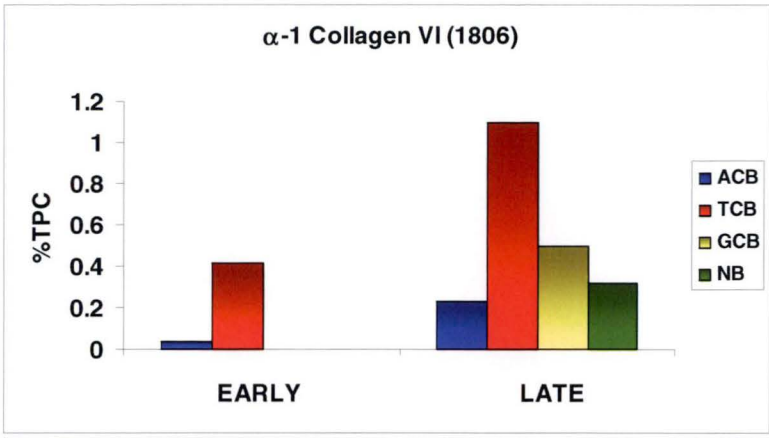


Plate 5.2a:
Early versus late analysis shows a trend to up regulation in the late onset cohort for all of the populations examined but no significant differences were observed.

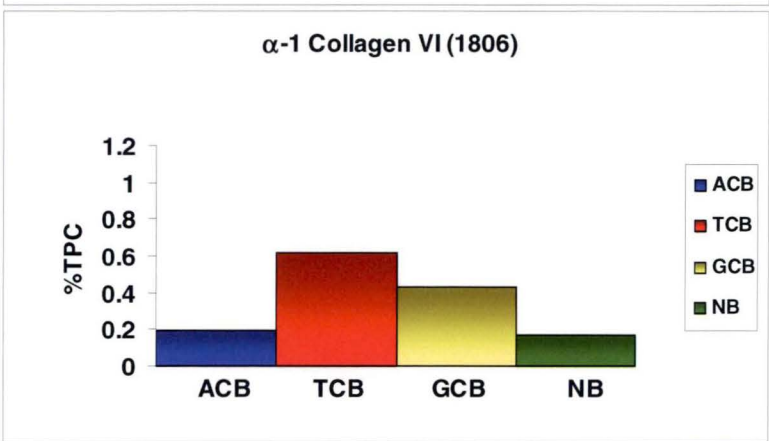


Plate 5.2b:
There is an up regulation in both TCB and GCB when comparing them to the NB cohort. This difference is not observed in the ACB population.

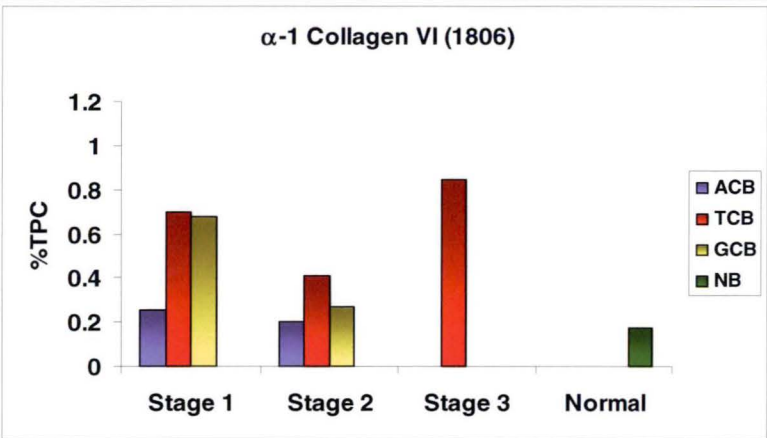


Plate 5.2c:
TCB was the only population to express SSP1806 in stage 3, with ACB and GCB expression limited to stage 1 and 2 tumours.

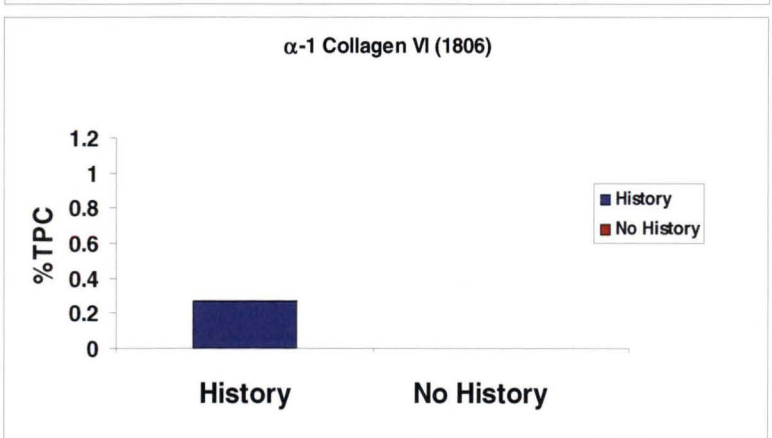


Plate 5.2d:
Analysis shows that there is only expression in the patients that have had a history of breast cancer

Plate 5.3– α-1collagen VI (1821)

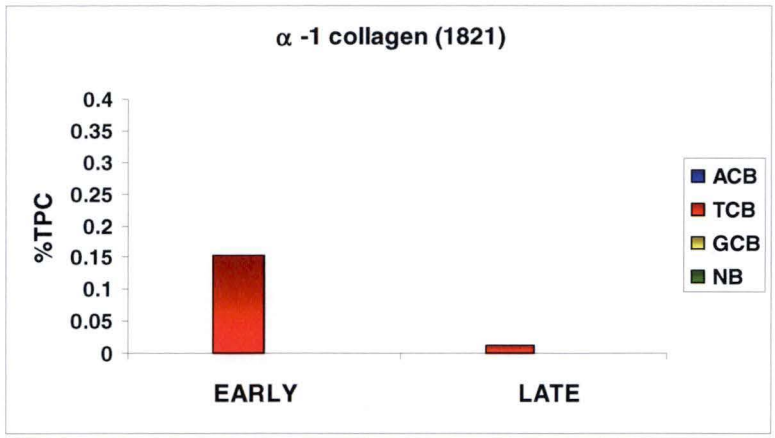


Plate 5.3a:
Early versus late analysis shows a trend towards a large decrease in the late onset patient group in the TCB population.

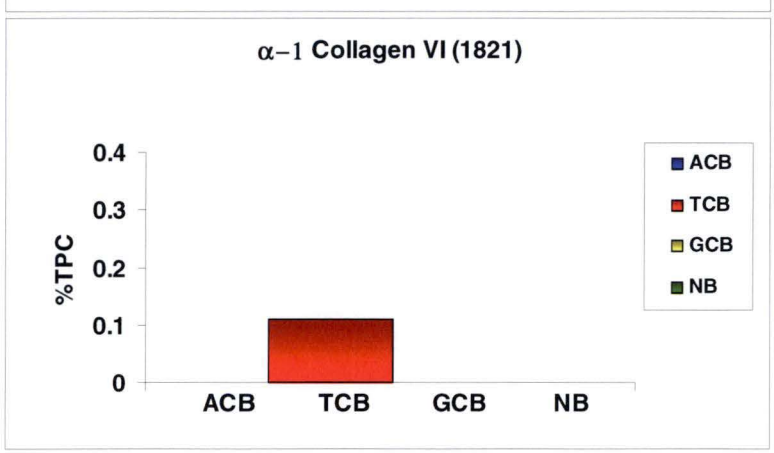


Plate 5.3b:
TCB was the only patient group to express this isoform of α-1collagen VI.

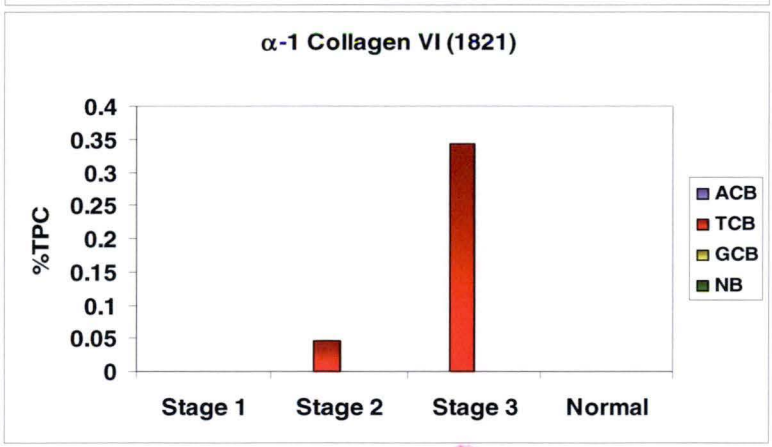


Plate 5.3c:
Expression was higher in the stage 3 tumours in the TCB population. This trend is insignificant.

Plate 5.4 Crystallin α-B

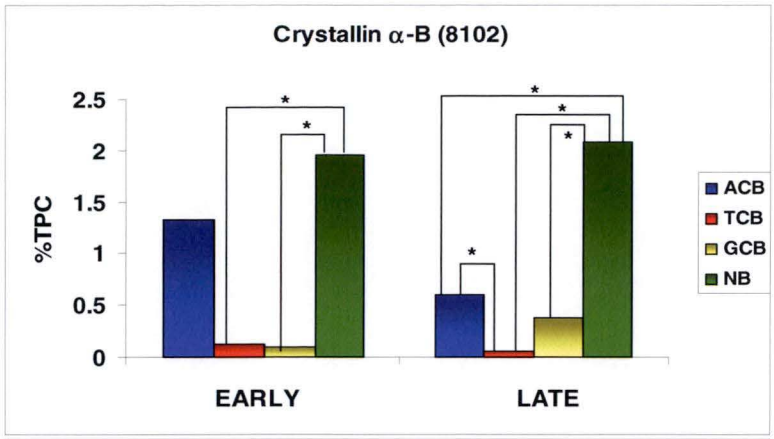


Plate 5.4a:
Early versus late analysis shows that there is no variance in the expression level for either of the NB cohorts, whist there is a change in the levels for all cancerous cohorts.

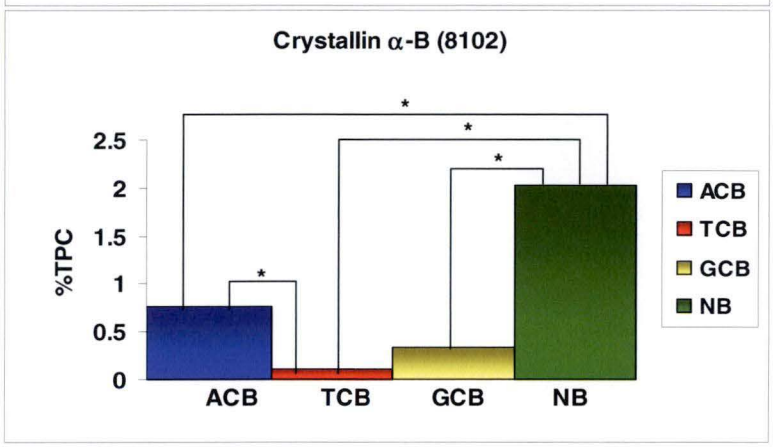


Plate 5.4b:
Cancer versus healthy analysis shows that there is a significantly lower level of crystallin in all of the cancer populations than in the control samples.

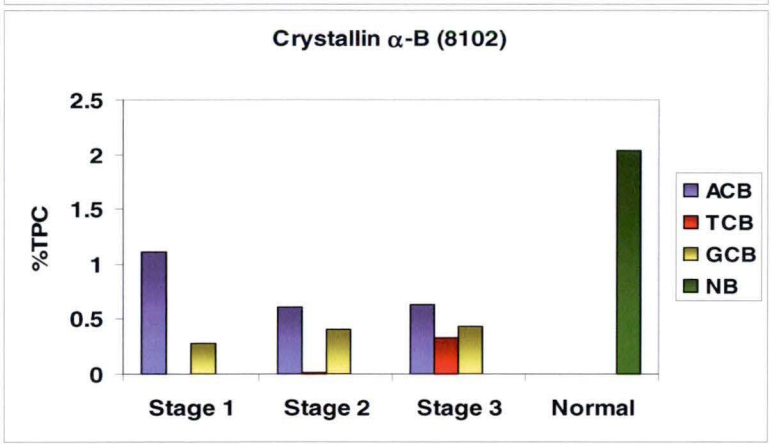


Plate 5.4c:
Staging result were inconsistent, with high variations between staging and populations.

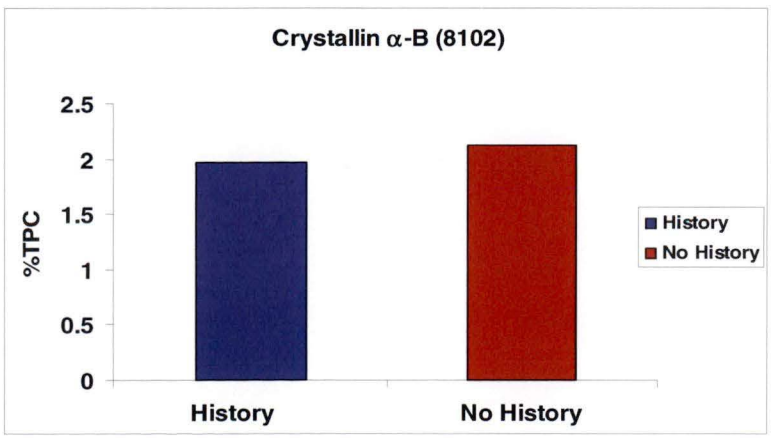


Plate 5.4d:
There was no difference between the healthy patients that have a history of breast cancer and those that do not have a history.

Plate 5.5: Ig L-Chain, V region

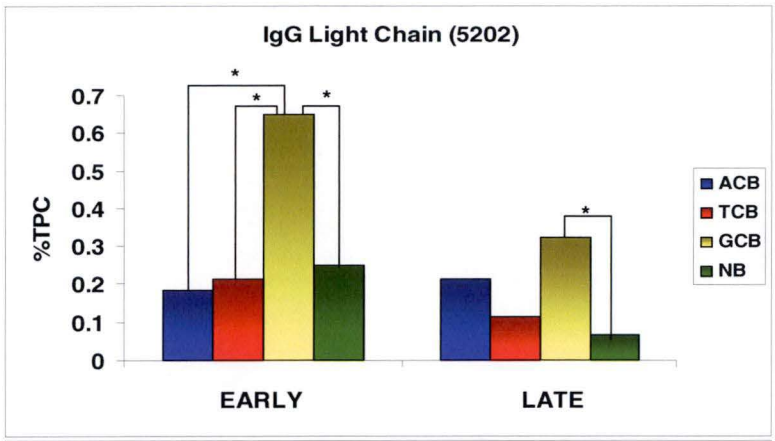


Plate 5.5a:
Control samples show a decreasing trend in late onset samples that is replicated in the TCB and the GCB populations but not in the ACB population. The GCB cohort is significantly different to all early onset cohorts as well as the healthy late onset cohort.

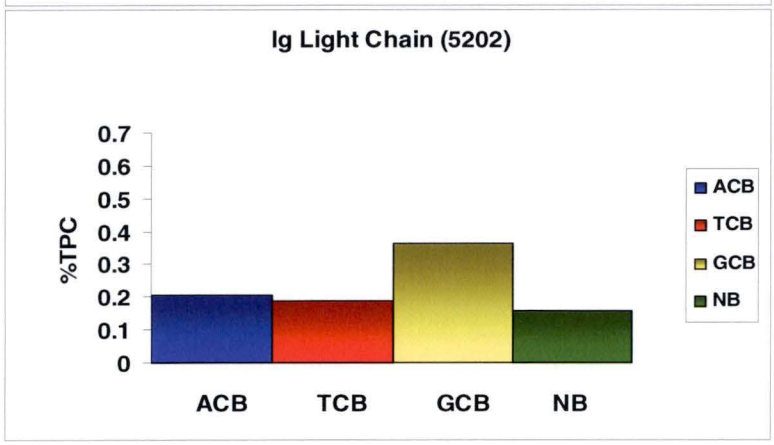


Plate 5.5b:
Cancer versus healthy analysis showed no significant findings between any of the populations.

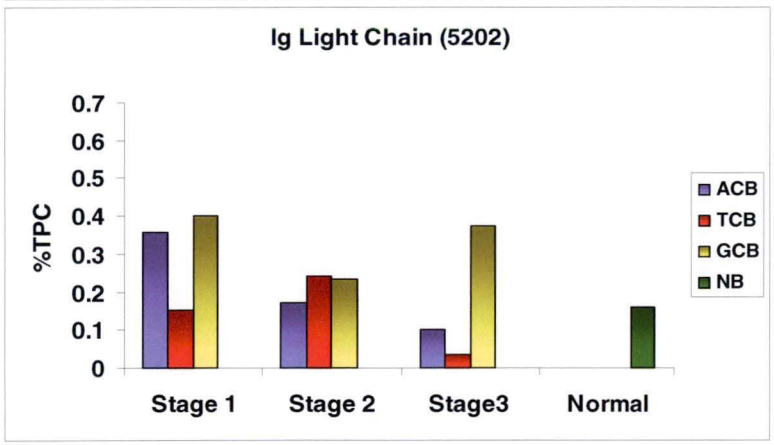


Plate 5.5c:
Staging analysis shows that there is a decreasing trend in the ACB population as the stage increases. Other populations are more inconsistent. No observed changes were significant.

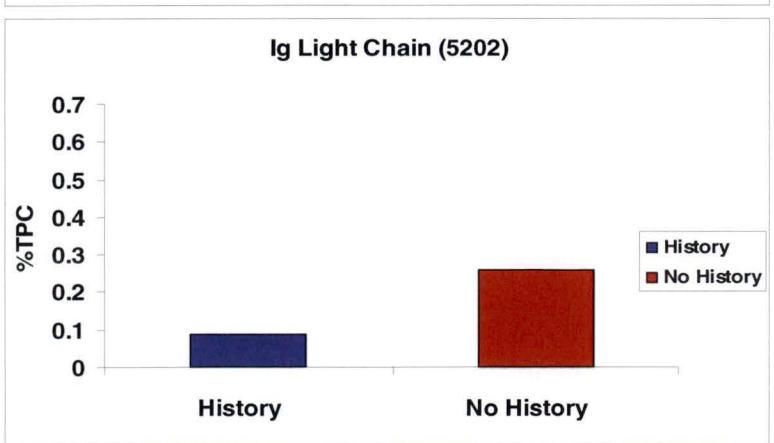


Plate 5.5d:
There is an observed increase in the expression level in patients without a history of breast cancer. This trend is not significant.

Plate 5.6: Cofilin 7003

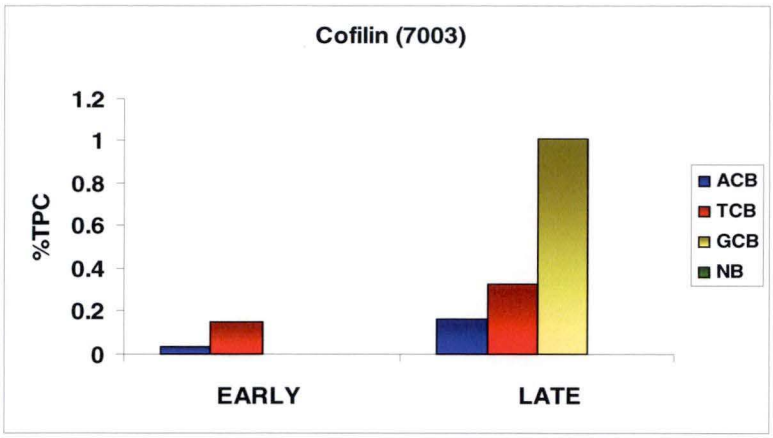


Plate 5.6a:
Early versus late analysis shows an up regulation in all late onset tumour cohorts.

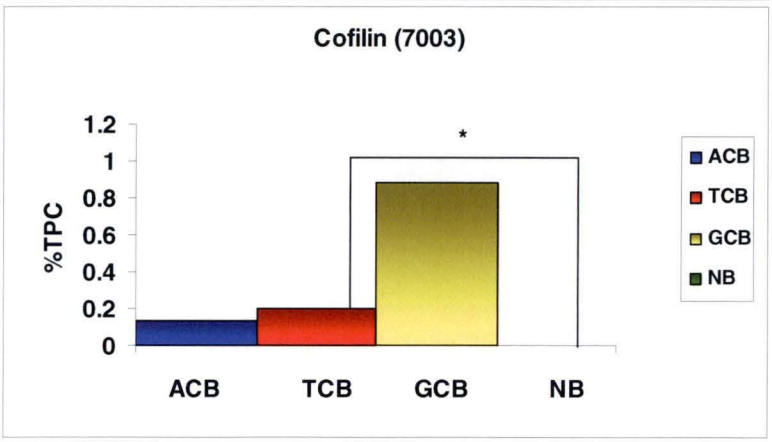


Plate 5.6b:
There was no expression in NB. Results were only significant for TBC v NB. GCB and ACB expression was too inconsistent to yield significant results.

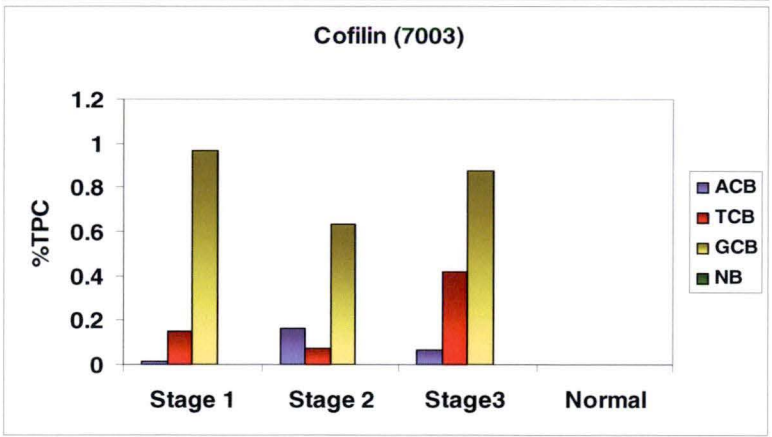


Plate 5.6c:
Staging analysis shows there was no trend observed between expression level and stage of tumour for and of the populations.

Plate 5.7: Cofilin 7004

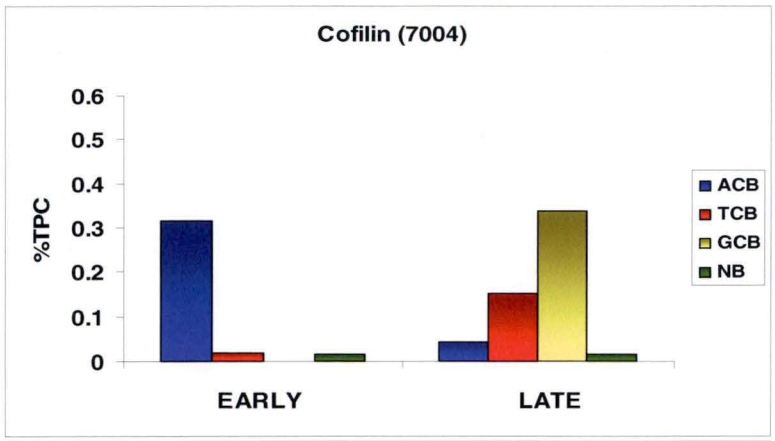


Plate 5.7a:
Inconsistent results were seen between the cohorts, with no clear pattern of up or down regulation emerging. Control sample expression remained unaltered between age based cohorts.

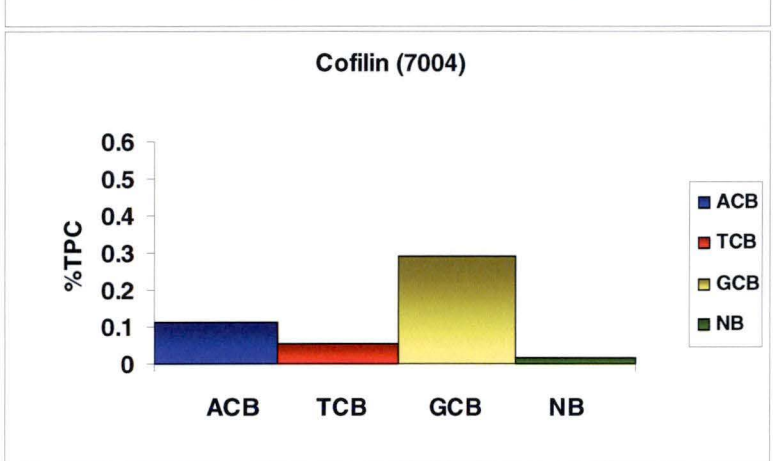


Plate 5.7b:
No significance was reached between the cohorts, despite there being a trend for the level of SSP 7004 to be higher in tumour cohorts.

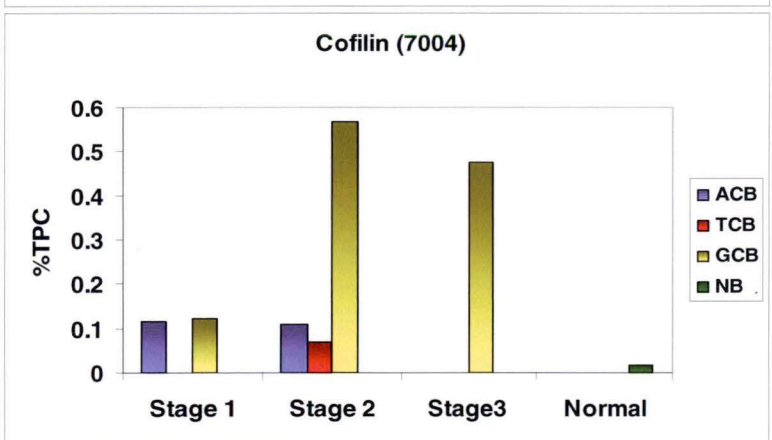


Plate 5.7c:
Staging analysis shows that apart from 2 GCB patients, expression of SSP 7004 is confined to stage1 and 2 tumours.

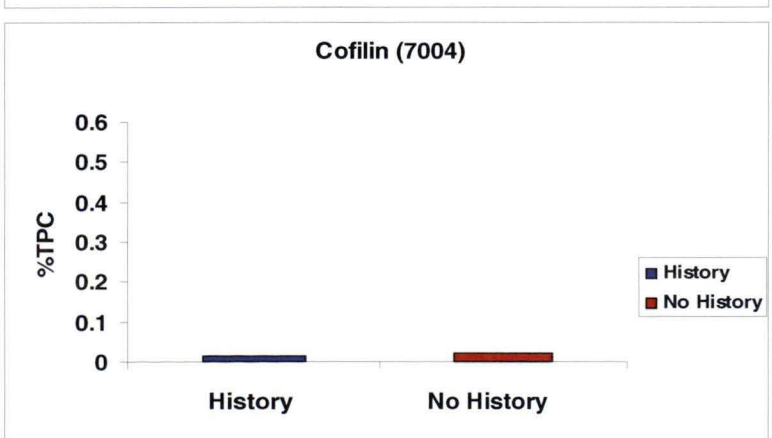


Plate 5.7d:
There was no difference between healthy samples that had a history of breast cancer and those that didn't have a history of breast cancer.

Plate 5.8: Lumican (1805)

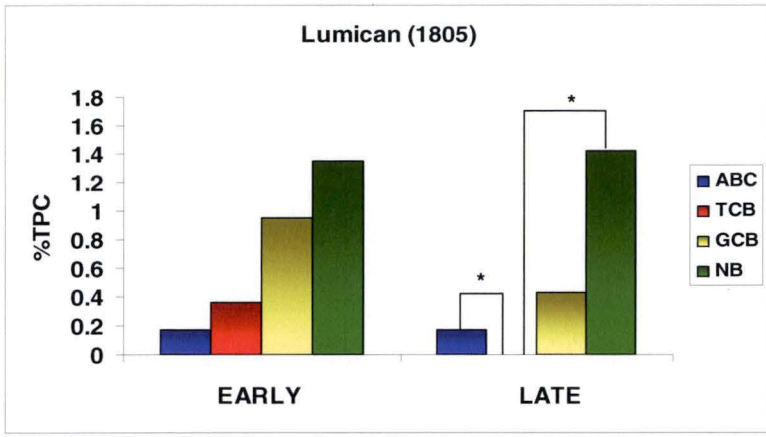


Plate 5.8a:
There was no change in the level of expression for ACB or NB. A decreasing trend in expression was seen in the TCB and GCB samples between the early and late onset cohorts, these decreases are not significant.

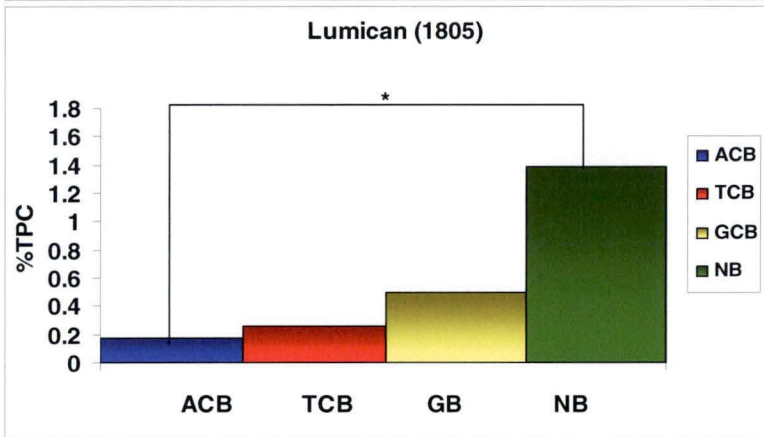


Plate 5.8b:
Levels in Control tissue are significantly higher than the ACB tissue samples. Although levels showed a higher trend in NB than in TCB or GCB, significance was not reached.

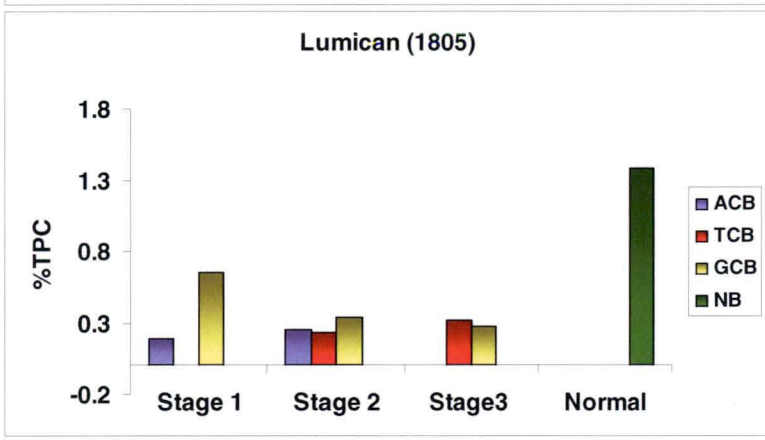


Plate 5.8c:
There was no staging pattern determined with constant levels of expression seen in all cohorts in all stages.

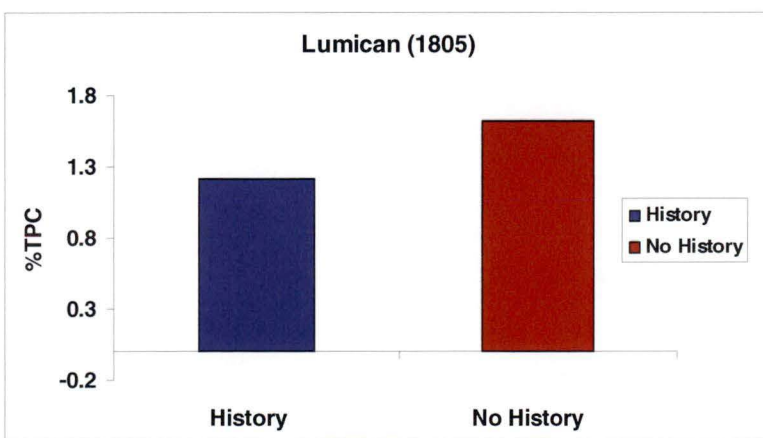


Plate 5.8d:
There was no significant difference between the healthy samples that do and don't have a history of breast cancer.

Plate 5.9: Lumican (2802)

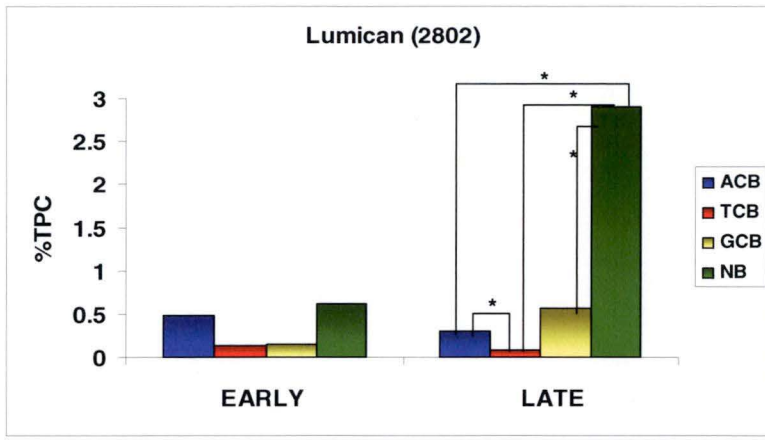


Plate 5.9a:
There is increased expression in late onset controls. This increase is not seen in late onset cancer patients in any of the populations.

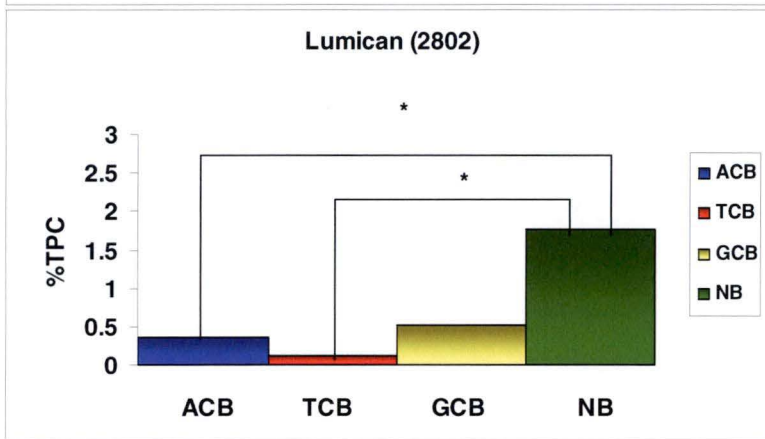


Plate 5.9b:
Decreased expression is seen in significant levels in the TCB and the ACB cohorts compared to the controls. GCB levels show a decreased trend but the level is not significant.

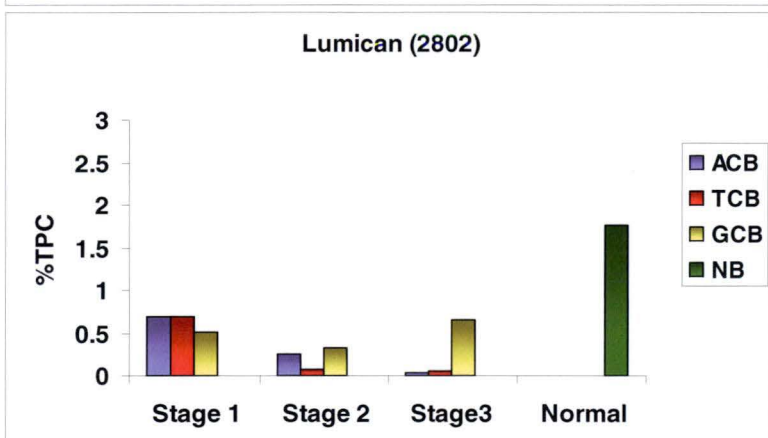


Plate 5.9c:
A decreasing trend is seen as the staging in each population increases. This difference is not significant.

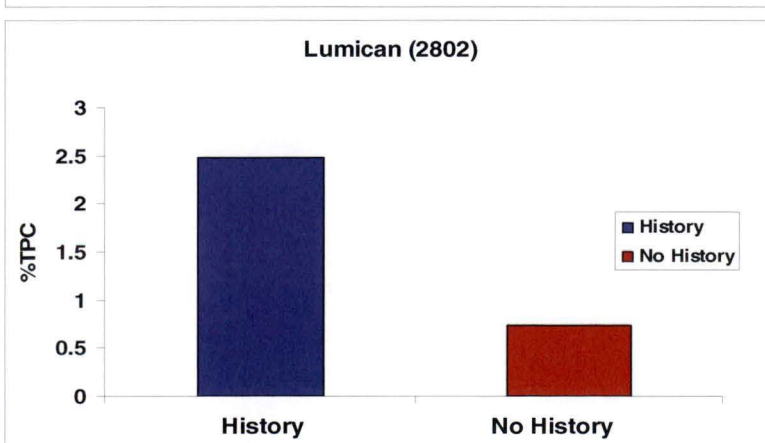


Plate 5.9d:
Patients with a history trend toward a higher level of expression that those with no history of breast cancer. This level was not quite significant (P=0.09).

Plate 5.10: Proapolipoprotein (2101)

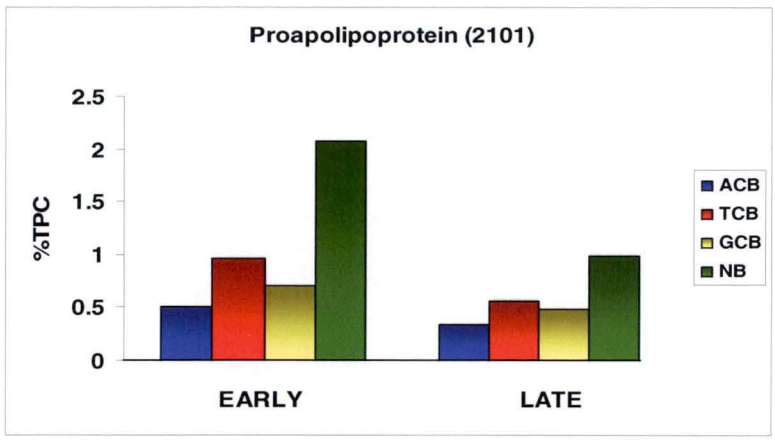


Plate 5.10a:
There is a trend toward a drop in the late onset patients compared to the early onset patients. There is a greater change in the NB population, although no comparisons are statistically significant

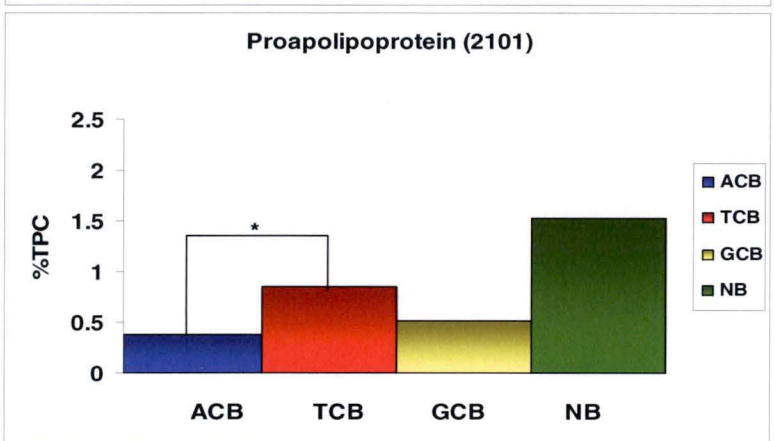


Plate 5.10b:
Levels of SSP 2101 indicate a trend higher in the control population, but this change is not significant.

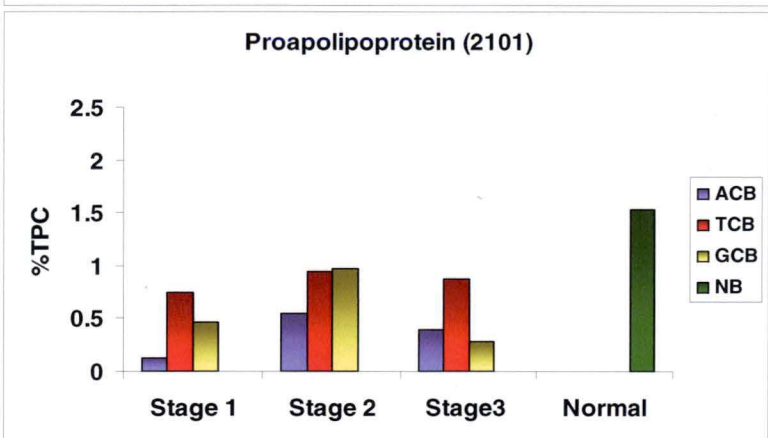


Plate 5.10c:
Staging analysis shows that there is a trend toward a peak concentration of SSP 2101 in stage 2 in all populations, but this is not significant.

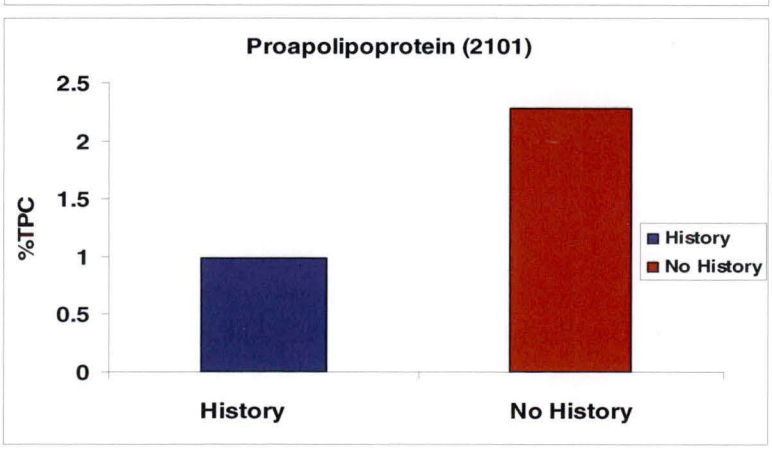


Plate 5.10d:
The level of concentration trends higher in controls with no history of breast cancer. The difference between the two is almost significant. (P=0.09).

Plate 5.11: Serum albumin

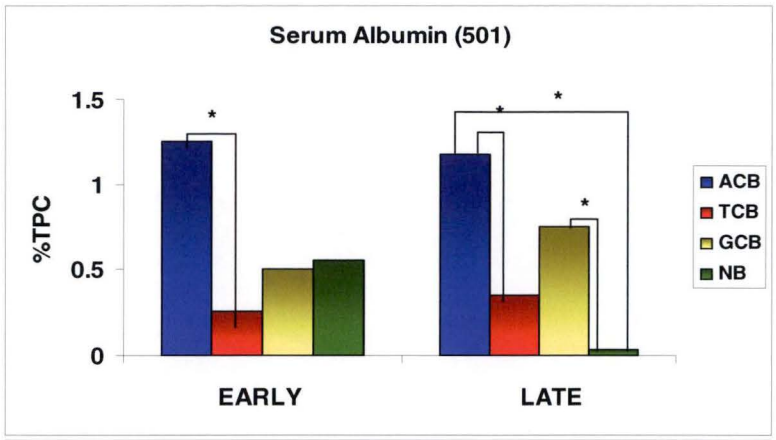


Plate 5.11a:
Early versus late analysis shows that whilst the levels drop between ages in the NB cohort they are increased or stay the same in the cancer populations.

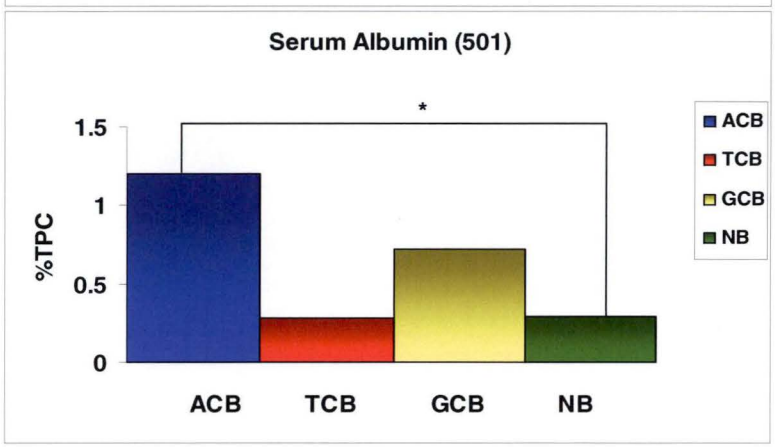


Plate 5.11b:
Levels are significantly higher in the ACB population compared to the controls.

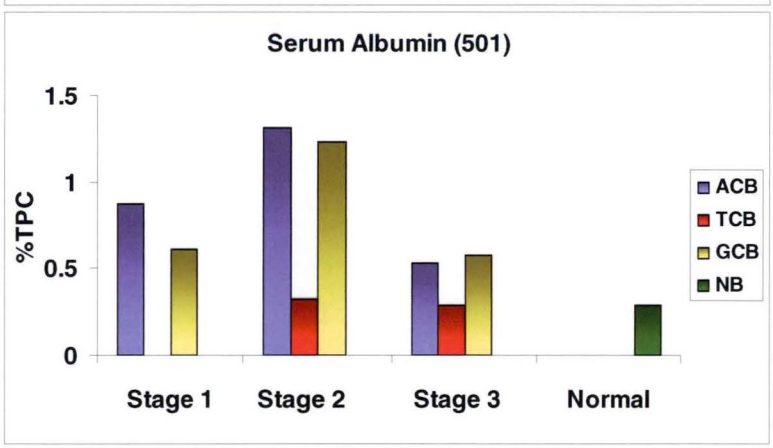


Plate 5.11c:
Staging levels are highly inconsistent, but there does appear to be a trend suggesting a concentration peak in stage two tumours in all populations.

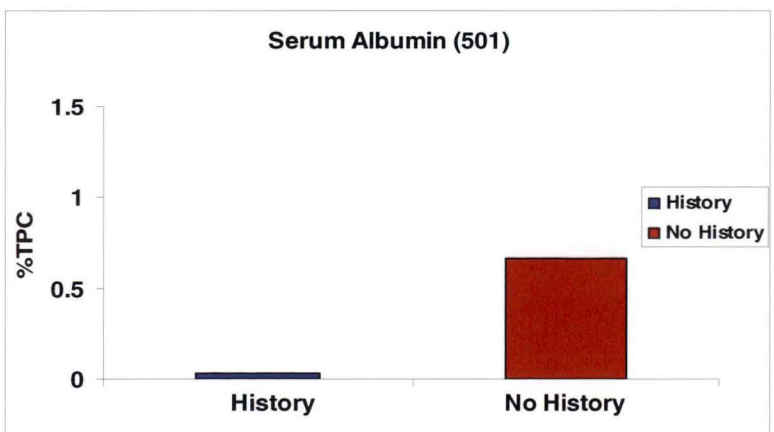


Plate 5.11d:
Levels are considerably, but not statistically significantly higher in patients that do not have a history of breast cancer.

Plate 5.12 – Serum albumin (3804)

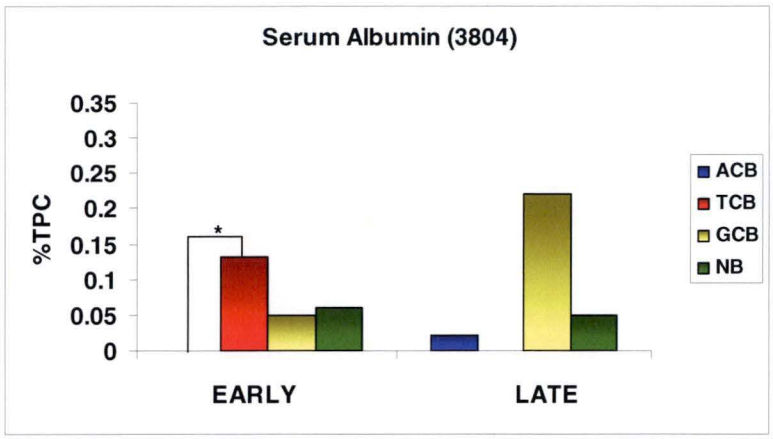


Plate 5.12a:
Results for the early versus late analysis are inconsistent with different patterns of expression seen in each age of onset comparison for each population.

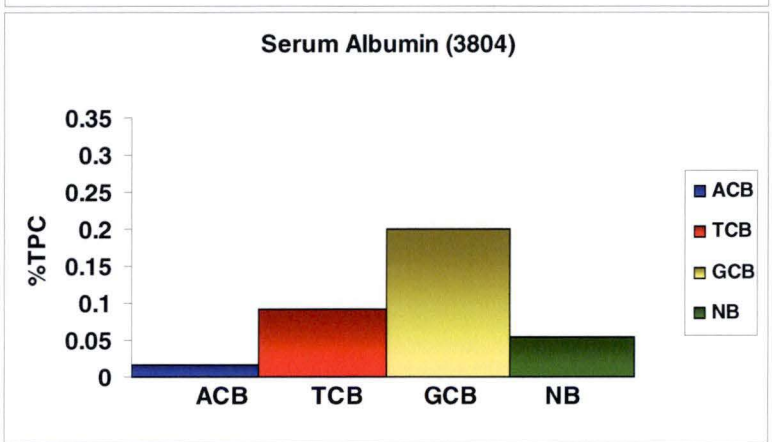


Plate 5.12b:
Cancer versus healthy expression was inconsistent with no significant results found.

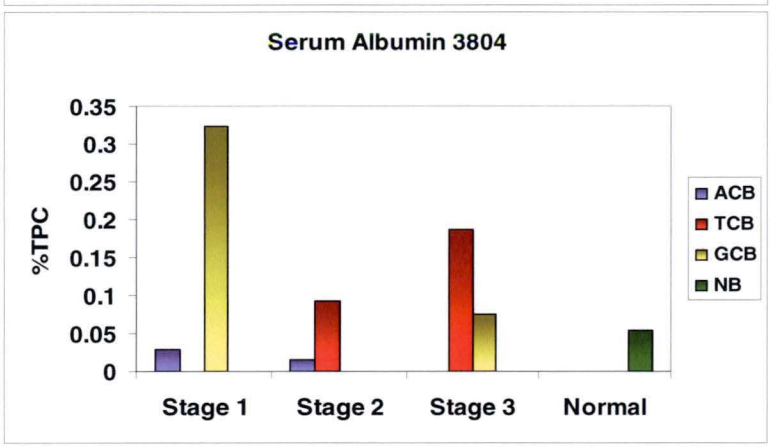


Plate 5.12c:
Staging analysis shows erratic expression levels and no significance between expression levels and the stage of tumour.

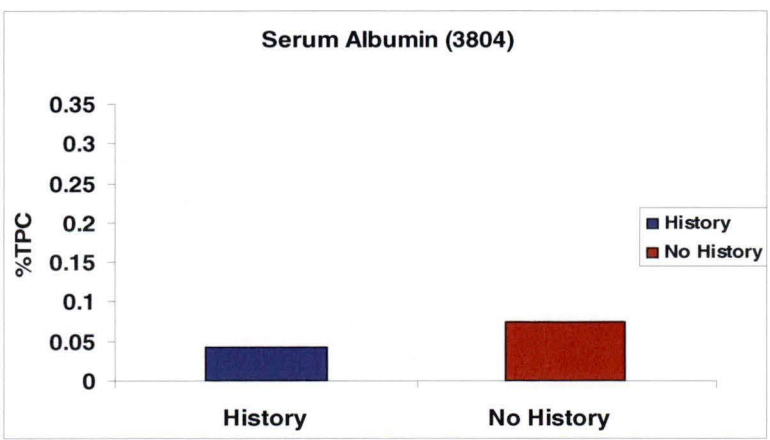


Plate 5.12d:
Analysis shows that there is little difference in expression between patients with and without a history of breast cancer.

5.4 DISCUSSION

5.4.1 α-1 COLLAGEN VI – 2804.

This protein has been previously associated with breast cancer but the exact influence that it has on breast cancer is unknown.

The change in concentration of this protein has to be dealt with carefully as it is a structural protein. The breast is an organ that over the life time of a woman changes. Some proteins may be influenced by the age of the woman, whether or not the woman had ever been pregnant and if so, how many times she has been pregnant, whether she breast fed and how long for if she did. These biological functions will change the structural make up of the tissue dramatically. As a woman ages the breast tissue becomes less dense, and this is a healthy normal occurrence. The early versus late analysis helps to demonstrate that for isoform SSP 1806, the change in concentration levels may simply be due to an involvement in the aging process. This being said, when taking a closer look at the concentration in the healthy tissue samples, this isoform is only seen in one sample, one that has a history of breast cancer, suggesting that this isoform may in fact be a potential marker of cancer induced changes within the breast tissue.

The information gained from the identification of α-1 collagen VI may have some interesting implications for breast cancer. SSP 2804, a cytoplasmic variant, had previously been identified and has been associated with inflammatory diseases such as myopathy and osteoarthritis [149].

Collagen VI functions as an important protein in the extracellular matrix of cells. Whilst collagen VI can be secreted by many types of cells, adipocytes are one of the major secretors of collagen VI. As breast tissue is rich with adipocytes, it is not unusual to detect collagen VI forms in breast tissue. One study that has been conducted shows that collagen VI is highly up regulated in breast tissue of mice [150], and DNA micro arrays, along with H&E staining of human breast tissue show a consistent upregulation of α 3 collagen VI [150]. In the same study treatment of MCF-7 cells with α–1 collagen caused the upregulation of β-catentin and cyclin D1, known markers in breast cancer.

Whilst this study did not look directly at the α–1 collagen VI form *in situ*, it suggests that collagen VI plays more than a structural role in breast cancer development.

The theoretical characteristics of α-1 collagen VI infer a 47kDa molecular weight, however it was isolated at 130kDa in this study. A previous 2D SDS-PAGE study showed that there is a 140kDa form [151], which it is suspected to be the form that has been isolated in this study. Initially qualifying for this analysis by the large difference in concentration between the early and the late onset in the Taiwanese cohort the result was not replicated in other populations, suggesting a Taiwanese specific variation that is related to cancer in younger women.

The suspected nuclear form SSP1806 was located in only 1 healthy patient and this is a patient that has had a history of breast cancer. No other healthy tissue expressed the protein, implying a potential role for the α-1 collagen isoform in breast cancer. Analysis of the staging data for this protein shows that it is much more common in the early stages of breast cancer. This observation supports the theory that α-1 collagen has a role in the early stages of breast cancer. Only TCB had expression of SSP 1806 in stage 3 tumours suggesting the possibility of a genetic variant in the ACB and GCB samples that isn't present in the TCB during these latter stages.

The low frequency of these isoforms in the cancer population that was studied is an aspect that means further screening for this and other specific isoforms of α-1collagen VI in both cancer and healthy tissue is required and makes it difficult to draw solid conclusions at this stage.

5.4.2 COFILIN

Cofilin is a protein that is involved in actin binding and the formation and polymerisation of intracellular actin rods [152]. Cofilin has been found to be both intranuclear and cytoplasmic, with localisation to the nucleus in times of induced stress reported [153, 154].

In this study cofilin was isolated in two different forms, a muscle (7003) and non-muscle (7004) isoform in the cytoplasmic fraction of the extraction.

Cofilin has been isolated in several studies relating to cancers [85, 155-158] and has been shown in previous studies to have altered expression in cancerous tissue compared

to the healthy tissue [85], and to be present in all forms of astrocytoma of the brain [157]. This latter study provided no healthy control for comparisons but it does support the staging data that was obtained in this study with the protein found in all stages at equal levels.

In direct contrast to the findings in this study is the breast cancer study by Luo *et al* [85]. Through proteome analysis they found a decrease of cofilin 1 non-muscle form in breast cancer tissue when compared to healthy tissue. Luo *et al* used only 5 breast cancer samples with matched normal tissue from the same patients suggesting the number of patients may be a limiting factor in the Luo *et al* study. The extraction process that was carried out in Luo *et al* was a simple whole cell extraction. This places all proteins in the same solution. This difference in extraction may begin to explain the contradicting results between the studies as in their study there is no discrimination between cytoplasmic and nuclear proteins as has been done in this study. This study has been able to distinguish between different isoforms of Cofilin in different cellular locations. It is here where the importance of one isoform may be hidden by looking at a whole cell extraction as opposed to cellular components individually.

It has also been reported that a failure to phosphorylate SER 23 in the cofilin sequence can result in an inability to localise to the nucleus. It may be that in these few patients that have shown expression in the cytoplasm that they are unable to phosphorylate at this site due to a mutation of the sequence, hence there is no expression in healthy tissue due to successful nuclear localisation. This needs to be looked at more closely to decipher whether cofilin in either isoform may have a role in breast cancer prognosis.

One other explanation, arising from a more recent study describes a relationship between LIMK1 a Serine Threonine Protein Kinase and the activation of the cofilin pathway as opposed to the concentration of cofilin itself. The level of activation is reported to be proportional to the degree of migration, invasion and metastatic potential of mammary tumours [158]. In this study the activation status of the pathway is currently undetermined and hence it can only be postulated to be causing the variation in results that were observed.

In an even more recent study cofilin is reported to be involved in barb formation of the actin fibrils, acting in a chemotactic manner, directing the growth of the cancer cells

[156]. Whilst another study reports that in cell membrane fractions cofilin is upregulated in a super invasive breast tumour cell line [155].

The results for this analysis in the early versus late analysis failed to reach statistical significance. This is attributed to the high variation in the concentration of cofilin in both isoforms. The absence of statistical significance at this stage suggests a requirement to understand better the role of cofilin in breast tissue and then in breast cancer and the pathways in which it could be involved as well as locate other forms of the protein in other locations so as to gain a bigger picture. Isoform 7003 specifically, which was not present in the healthy tissue certainly may offer some potential as a drug target in the future.

5.4.3 CRYSTALLIN α-B

Crystallin α-B is a member of the small heat shock protein family (sHSP). Sequenced due to its difference between early and late onset patients originally, this difference was not significant but there is a visible trend between the early and late onset proteins when comparing the healthy to the cancerous in each age group. The levels in the healthy tissue were higher, particularly in the late onset group. When analysis was done for cancer versus healthy however the level of significance increased. This suggests crystallin α-B may have a role in breast cancer.

Isolated in this study was a nuclear isoform of crystallin α-B. In this study no corresponding cytoplasmic protein was isolated. The involvement of crystallin α-B in breast cancer recently has been postulated to be as a novel oncogene, indicating an upregulation of crystallin α-B[159, 160].

Studies have shown that an increasing presence of crystallin α-B may indicate lymph node involvement [161]. It may have a role in metastasis[162], and may be an indicator of poor prognosis for both breast cancer [159] and head and neck cancers [163].

Studies have shown crystallin α-B to inhibit activation of caspase-3 via TRAIL signalling[164], as well as cause the inhibition of RAS activation[165], both actions result in the inhibition of apoptosis. Crystallin α-B is also reported to be involved in filament bundle formation during mitosis and that it is involved in focal adhesion [166]. More specific to breast cancer are studies showing chemotherapy induced expression of

crystallin α-B [167] and the presence of crystallin α-B being linked to Tamoxifen resistance [168].

This study contradicts current literature, showing a decrease in expression within the hydrophobic compartments of the cell however, all of the studies mentioned fail to discriminate between the concentration in the cytoplasm and the nucleus. It is feasible therefore that there is a cytoplasmic isoform that is elevated, obscuring observation of the proteins decrease in the nucleus when expression is evaluated using a whole cell extraction based method.

As there was no discrimination between patients and their treatment during the patient selection process whether or not the concentration of crystallin α-B is effected by treatments has not been examined in this study due to the difficulty in obtaining medical records for the international cohorts.

There is a need to identify the presence or absence, and the concentration level of crystallin α-B in the cytoplasm before any conclusion can be drawn from the results in this study. It could be postulated however that the observed nuclear depletion is due to cytoplasmic increase and relocation of crystallin α-B. Variation in the concentration levels, or an alternative genetic form for TCB (which were heavily weighted in the selection of these proteins) was the reason that it was not seen, or it may have been isolated but not yet identified due to the MAALDI-TOF MS analysis.

The question that would then need to be asked is whether all of the roles of crystallin α-B in the cell are equally important or whether the roles within the nucleus are jeopardised by the mass exodus of the protein into the cytoplasm.

Despite being originally identified for its differential regulation in the age of onset analysis, it appears that crystallin α-B, and it's down regulation may have more of a general significance in cancer as a whole, as opposed to being an age specific indicator at this point.

5.4.4 IG L-CHAIN, V-REGION

Ig L-Chain, is a protein that has been shown to have variants in a variety of cancers [169-173]. The types of cancer that are noted to have altered Ig L-chain expression are, not surprisingly, lymphomas and blood cancers such as CLL (Chronic Lymphoid

Leukemia) and B-cell associated malignancies [170, 172, 174]. There has been limited documented evidence for a role in solid tumours, despite variations in Ig chains being originally located in colon cancer [169, 171]. Involvement in breast cancer of alternate Ig forms has until recently been even less frequent [173, 175, 176].

These studies in to the involvement in breast cancers are still being validated. Kotlan *et al* in 1999 found variable v-regions chains in the plasma cells from around a single medullary carcinoma [173]. In that study Kotlan *et al* suggested that infiltration by B-cells is an indicator of good prognosis. This is supported by a 1992 paper by Kawata *et al* [177], however more recent literature, particularly in renal and ovarian solid tumours suggests that infiltration is an indicator of poor prognosis [178-180]. In 2003 Kotlan *et al* then demonstrated that there may be potential in using the antibody binding regions on the B-cells that surround the tumours as potential targets for antibody engineering [175]. The study however was still specific to medullary carcinomas. Wang *et al* recently expanded the types of carcinomas in which these variations occur to include NOS tumours (Not otherwise specified), however this was from a sample number of three[176]. So whilst pilot studies have been performed and published, their use is still limited, although Wang *et al* was able to show that whilst the V-region in L-chains were altered, it was not the same sequence altered in each patient.

The results from this study are highly variable between populations, age groups and staging and for this reason Ig L-Chain is unlikely to be found to have a role in the progression or ability to diagnose or predict outcome using the analysis performed in this study.

The results seen appear to reflect a protein that is highly dependent upon the individual and may not reflect the state of their tumour, but more the state of their health and immune system.

More studies to determine the importance of the Ig L-chain v-region would need to be conducted to ascertain the role of Ig in the breast, and the results may in no way be related to the patient's breast cancer status.

5.4.5 LUMICAN

Lumican is a protein that was located in several isoforms as discussed in the previous chapter. Chapter 4 discussed how lumican is mostly reported to be up regulated in

breast tissue [86, 119]. Although one study by Troup *et al* supports a decrease in the concentration of lumican and associates it with better survival, there is no control comparisons made [121]. These studies, although they take in to account potential variation in cell population types, fail to recognise the possibility of different forms of lumican. This study has showed a total of 3 variants of lumican, whether these are due to post translational modification or genetic variation is still to be determined.

There were 2 isoforms of lumican for the early versus late analysis and they showed in healthy tissue an upregulation in the late onset for SSP 2802 in healthy tissue and no change for SSP 1805.

These patterns of expression were not replicated in the cancerous cohorts for either isoform. SSP 1805 was clearly depleted in concentration in the both the early and late onset patients, indicating its' expression was irrespective of age. This was confirmed with the significant results in the cancer versus healthy analysis for this isoform.

The staging data suggests that there was a decrease in concentration with an increasing severity for SSP 2802. It is therefore possible that SSP 2802 may be a useful isoform to look at in future studies as it appears the higher the concentration the lower the histological stage of the tumour.

The history versus no history analysis is intriguing because for SSP 1805 there is a decreased level in the patients with a history. This is reversed for SSP 2802, where the concentration is higher in the patients who have a history of breast cancer, almost reaching statistical significance. One study has indicated a role for lumican in apoptosis via the Fas mediated pathway and suppression of cell proliferation [181]. This suggests that it is possible that SSP 2802 peaks in expression just prior to tumour formation, also indicating a higher level lumican being expressed at a genetic level. This suggests an attempt to prevent the progression of tumour formation, essentially acting in a tumour suppressive fashion. This hypothesis however, needs further investigation in to areas such as the isolation of multiple variants of lumican in this study, as well as comparative studies which use whole tissue and cell lines as well as LCM acquired tissue to ascertain the function, and location of the majority of the Lumican protein.

5.4.6 PROAPOLIPOPROTEIN

Proapolipoprotein was seen in this study to follow a down regulated trend in cancerous tissue. Proapolipoprotein is the precursor protein for apolipoprotein of which there are several members. The gene bank entry for this particular sequence implies proapolipoprotein A-1 as the protein identified. Apolipoprotein A-1 is the major protein component of HDL.

The results found here are in contrast to the concentration observed in serum proapolipoprotein in Huang *et al* [182]. Huang *et al* suggested proapolipoprotein A-1 to have a role as an upregulated serum marker for breast cancer. Proapolipoprotein in that study was found to be upregulated by 38%, however they also showed that Apolipoprotein in all post translational isoforms was down-regulated in the serum. The fact that there is a down-regulation in the actual tumour itself of proapolipoprotein, which is cleaved to give the protein circulating Apolipoprotein A-1, suggests that there is validity in both results.

The decreased concentration in tumours that was observed in this study would explain a resulting decrease in the serum levels of the downstream cleavage products as seen by Huang *et al.* It does not explain the increased level of the precursor in the serum though.

The trend toward lower levels of proapolipoprotein were not significant due the high variation of expression, suggesting a high level of individual affect on the levels of this protein as opposed to a genetic based difference.

Not only has proapolipoprotein been associated with breast cancer but it has also been shown to be de-regulated in colorectal cancer [183] and in Alzheimer's [184], suggesting a fundamental role in disease progression as opposed to being limited to cancer only.

5.4.7 SERUM ALBUMIN

Most of the discussion as to serum albumin's involvement was considered in the previous chapter. There is the need to further discuss serum albumin given the results from this analysis.

Serum albumin in the hydrophobic samples (SSP 3804) suggests that it is either translocating proteins to the mitochondria or to the nucleus. There has been evidence as

mentioned for the role of serum albumin as protein transporter [110]. Serum albumin has also been shown to co localise in colon cancer cell lines with a chemotherapy drug Suramin. This co-localisation inhibits the effectiveness, by decreasing the levels transported to the nucleus [185].

SSP 3804 showed erratic concentration levels and no firm conclusions can be drawn. Sequenced for its potential role in differing between early and late onset cancers it has become evident that this is probably not the case with the levels in each population varying too much to make any clear conclusions.

Results for the cytoplasmic form however, are more conclusive. For SSP 501 in the early versus late analysis there is a decreased concentration in the healthy samples in the late onset cohort that is not seen in any of the cancer populations. This suggests a role in younger patients for serum albumin SSP 501 that is not required in older subjects in healthy individuals. Whatever this role for SSP 501 may be, it appears it is still required in the cancerous tissue, or it is indirectly associated through an up-regulation of this isoform. The reason for up-regulation of serum albumin needs to be determined to find out whether it is due to changes in protein or gene regulation. The expression of serum albumin SSP 501 is also higher in the earlier stages of cancer, ie: stages 1 and 2 as opposed to stage 3. The exception here is in the TCB samples where there is no expression in stage one cancers. There are only two stage 1 TCB samples though, which is not enough to base a conclusive result on for this population.

Serum albumin has been shown to bind to Estradiol and aid signalling in the breast cancer cell line MCF-7 [110]. Serum albumin is also internalised by MCF-7 cells as opposed to Sex Hormone Binding Globulin (SHBG) which functions similarly at the cell membrane to serum albumin during E_2 delivery [109].

Whether serum albumin is involved in the E_2 dependence and delivery for tumour cells is yet to be described, but may prove to be a potential explanation for what is observed. In breast cancer the E_2 requirements of the cell do not necessarily diminish with age, unlike in healthy patients where the older age group who may have gone through menopause and no longer have an E_2 requirement in the breast tissue. If this E_2 requirement is what is causing the sustained and increased concentration of serum

albumin SSP 501, it may provide further ways of treating E_2 dependent tumours given further validation and future investigations.

5.5 CONCLUSIONS

This chapter has demonstrated that there are many proteins that are altered in breast cancer that vary accordingly with the age of the patient. This analysis has shown that there is a need for population matched controls to determine the significance of many of the findings related to age of onset and breast cancer in this study. There are many proteins identified that require further investigation, but of particular promise at this stage is αB-crystallin, a protein that is increasingly appearing to be an oncoprotein.

6 DIFFERENTIAL ANALYSIS OF TWO DIFFERENT HEALTHY COHORTS

6.1 SUMMARY 127

6.2 RESULTS 128

6.2.1 MALDI- TOF MS 128

6.2.2 Pyruvate Kinase – SSP 4402 128

6.2.3 Phosphatase 2A – SSP 2402 129

6.2.4 Gi 51491284 – SSP 1201 and SSP 1204 130

6.3 DISCUSSION 132

6.3.1 Pyruvate Kinase 132

6.3.2 Phosphatase 2A – SSP 2402 133

6.3.3 Gi 51491284 134

6.4 CONCLUSIONS 135

6.1 SUMMARY

The results of any study rely heavily on the validity of the controls. When the controls for this study were being collected, two types of healthy tissue were collated to form the healthy cohort. The first of these control types are true controls. They are comprised of breast tissue from women that have had mammoplasty surgery for reasons unassociated to breast cancer. The second set of tissue samples that form the healthy tissue cohort are samples excised from the breast tissue that was associated with breast cancer by being in the same or opposing breast when a total or partial mastectomy was the treatment of choice. The healthy sample was taken from macroscopically normal breast tissue following the removal of tissue that also contained tumour cells.

This gave the benefit of a higher number of patients in the normal cohort, but it also allowed for there to be potentially undetected differences between the two healthy cohorts. These differences may reflect changes that are due to circulating factors that may cause affects distal to the primary tumour site. The ability to locate an early diagnostic from the comparison of these tissues was investigated further.

Using the same software as the other analysis, PDQuest, two classes were formed normal with history (NWH) and normal no history (NNH). The two groups were analysed for proteins significantly different and up and down regulated by at least 5 fold. These comparisons led to 8 proteins being sequenced. MALDI-TOF MS analysis was able to gain a positive identification for 4 of the proteins. Whilst the role of these proteins are yet to be fully established, there is a basis for further investigation as they may hold some interesting insight into potential early diagnostics for breast cancer.

6.2 RESULTS

6.2.1 MALDI- TOF MS

Following analysis using PDQuest software 8 proteins were found to be differentially regulated to a significant degree. Sequencing at APAF in Sydney revealed the identity of 4 of the 8 proteins. The results can be seen in Table 6.1. The proteins that were successfully identified have had the results outlined in more detail as follows.

SSP	ID	Mowse Score	Exp. Mw/pI	Th. Mw/pI	Accession number
Solution 2					
4402	pyruvate kinase	46	29/6.5	57.8/7.58	gi:35505
8307	-	-	24/8	-	-
Solution 3					
703	-	-	74/5.9		
1201	gi/51491284	71	33/6.2	96.7/9.59	gi:51491284
1204	gi/51491284	82	33/6.1	96.7/9.59	gi:51491284
2402	phosphatase 2A	27	37/6.4	56.1/6.6	gi:18490282
7303	-	-	36/8.7	-	-
8106	-	-	31/8	-	-

Table 6.1: Results from MALDI-TOF MS analysis conducted at APAF in Sydney on proteins of interest in NNH versus NWH comparisons.

6.2.2 PYRUVATE KINASE – SSP 4402

Pyruvate kinase was isolated as a cytoplasmic protein that was found to be up regulated in healthy tissue that had had previous association with breast cancer (Figure 6.1). There was no presence of this form found in the healthy, unassociated tissue with the expression exhibited in all patients with a history except one. Statistical analysis showed that when comparing the cancerous cohorts to the healthy controls cohort significance was not reached. When the healthy cohort was split in to patients with (NWH) and with out a history (NNH) of breast cancer the analysis then reached significant levels. All comparisons made to the cohort with history were significant, whilst the comparisons to the cohort with no history were not significant. Statistical significance was also reached when comparing the two healthy groups (P=0.03).

This proteins presence in the cancerous cohorts, specifically the levels in different stages of breast cancer was also analysed with little association to cancer staging evident. These staging results can be seen in Figure 6.2.

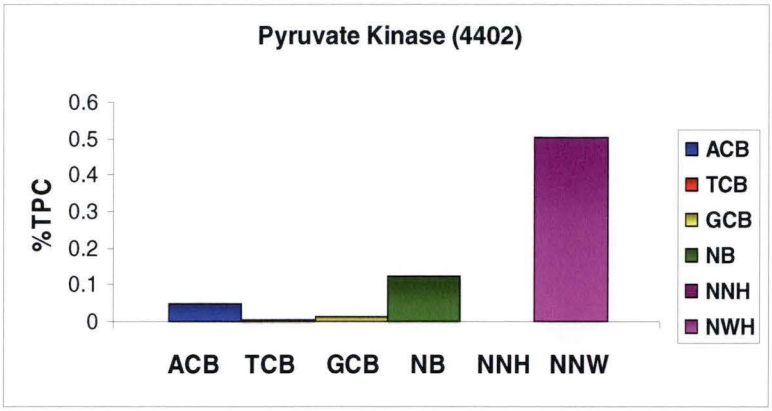


Figure 6.1: Total protein content of SSP 4402 in solution 2. NNH (Normal with no history) showed no expression of the isoform, whilst NWH (Normal with history) had very high expression levels.

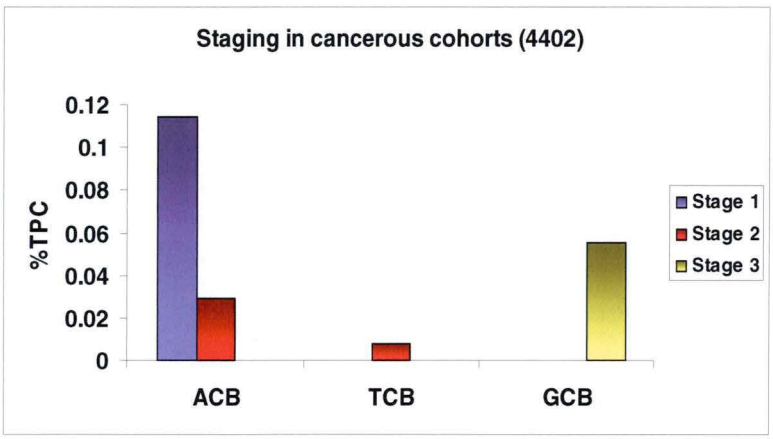


Figure 6.2: Staging analysis of pyruvate kinase 4402 in the cancerous cohorts form each population.

6.2.3 PHOSPHATASE 2A – SSP 2402

This protein, Phosphatase 2A – an isoform of the regulatory subunit B56, was found to be in higher concentrations in the healthy tissue, but this difference was not found to be significant when the entire control cohort was combined for comparisons. When the healthy cohort was separated in to the two healthy tissue cohorts the concentration of the patients with a history was seen to be higher than the concentration seen in the patients without a history (Figure 6.3). Despite the average of both normal groups being higher than the cancerous cohorts, there was a lack of significance due to high variability and low numbers when those without a history of breast cancer were compared to each of the populations. The concentration difference between the cancerous group and the NWH samples however was significant for each of the three cancerous groups that were compared. Staging analysis for this protein show little association of SSP 2402 with the stage of breast cancer (Figure 6.4).

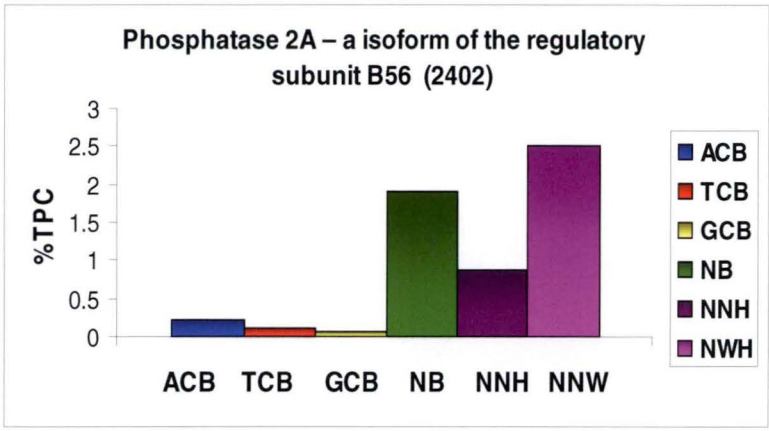


Figure 6.3: Analysis of the PP2A isoform shows an increase in the samples that have had a history of breast cancer, despite an obvious decrease in all of the cancer cohorts.

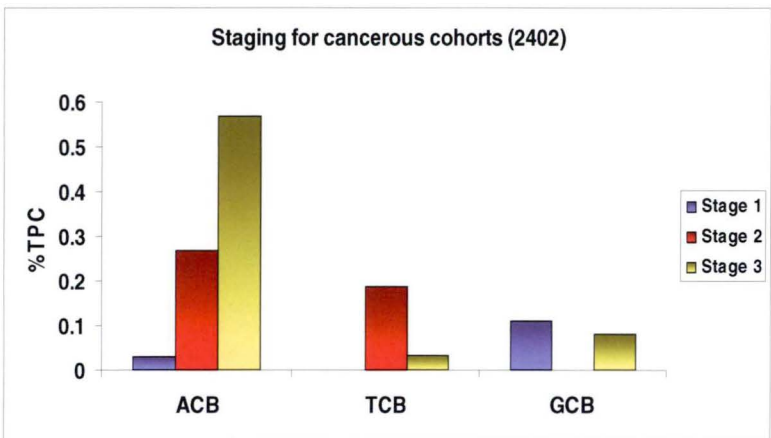


Figure 6.4: Staging analysis shows no clear trend for the expression of the PP2A isoform SSP 2402 in breast cancer for any of the populations

6.2.4 Gi 51491284 – SSP 1201 and SSP 1204

Gi 51491284, a hypothetical protein extrapolated from cDNA clones, was identified as 2 of the 8 proteins sent for identification. Both isoforms are presumed to be hydrophobic as they were isolated from solution 3.

For the first isoform SSP 1201, Figure 6.5 demonstrates that in all of the cohorts the concentration of SSP 1201 appears constant, including in the healthy patients, with no significant comparisons between any of the populations. However, when the healthy cohort is split in to the NWH and NNH groups it becomes apparent that there is a significant difference in concentration between the two (P=0.03) and that the NNH is significantly different to the cancers whilst the NWH cohort is not. This protein was present in all 7 patients with history and only 2 of the 4 patients without history. In the cancer cohorts the protein was present in 78%, 52% and 93% for ACB, TCB and GCB respectively. Staging analysis for SSP 1201 showed no clear association with increased staging in any of the cancerous cohorts (Figure 6.6).

Isoform SSP 1204, shown in Figure 6.7, is less frequent than its 1201 counterpart, with presence in 39%, 52% and 0% of ACB, TCB and GCB samples. In the normal cohort, 4 out of the 7 patients that had a history of breast cancer expressed this protein whilst none of the proteins that have had no contact with breast cancer expressed the protein. SSP 1204 in the healthy tissue with no history is significantly different to all the populations based cohorts of breast cancer and is almost significantly different to the NWH cohort (P=0.055).

Staging analysis, as seen in figure 6.8 indicates that the level of this isoform may peak in stage 1, with concentration decreasing in latter stages of breast cancer.

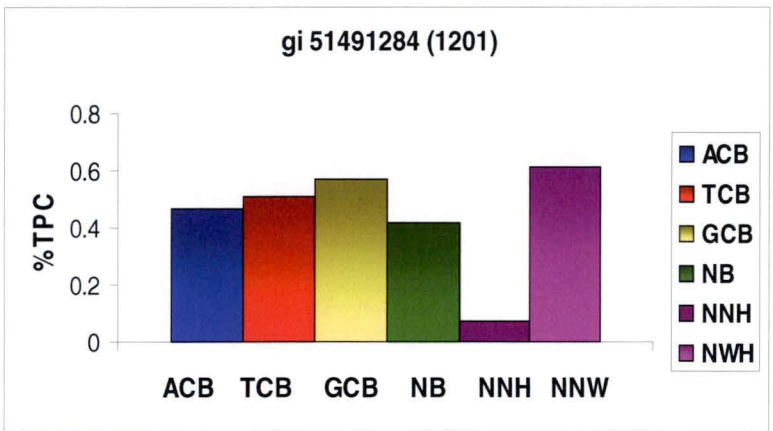


Figure 6.5: Analysis of hypothetical protein GI 51491284 shows that there are decreasing trend of protein concentration in samples without a history of breast cancer

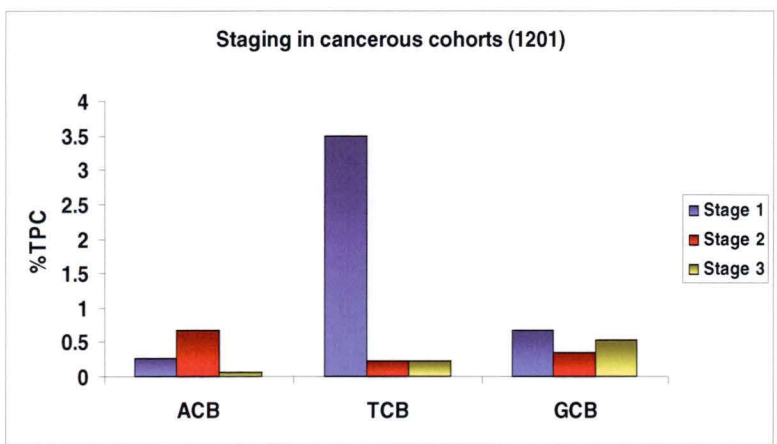


Figure 6.6: Staging analysis of isoform SSP 1202 shows that there is a peak in stage 1 for TCB and GCB, but this trend is not replicated in ACB.

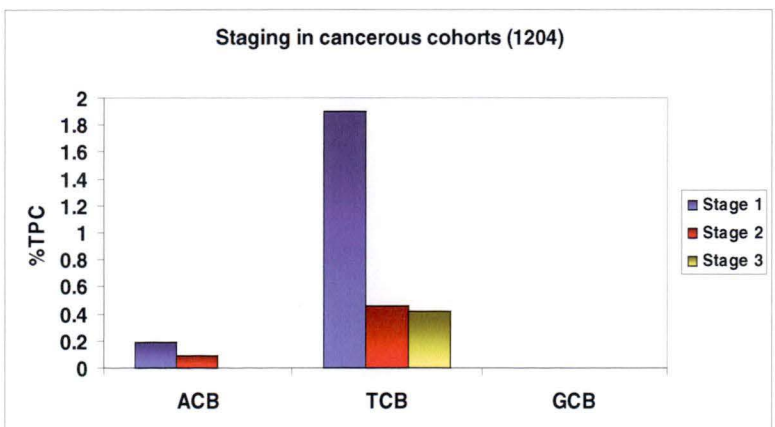


Figure 6.7: Analysis of the alternate isoform 1204 of GI 51491284 shows more erratic expression, however presence in the controls with no history of breast cancer is absent

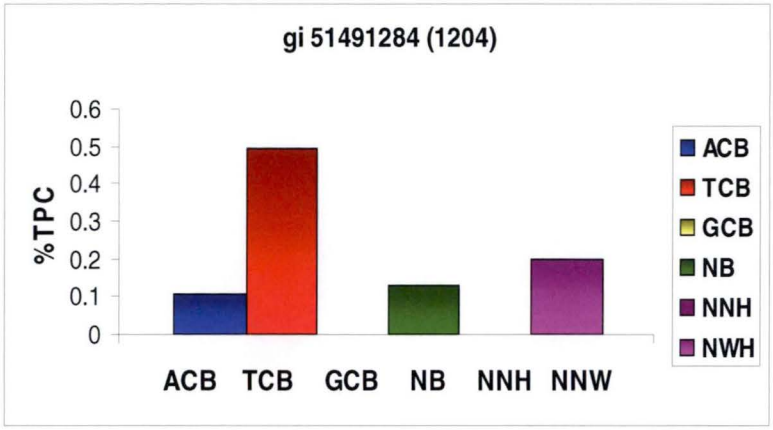


Figure 6.8: Staging analysis shows levels of 1204 are highest in stage one for ACB and TCB populations. GCB showed no expression at all. There is a trend toward expression only in healthy tissue with a history of BC.

6.3 Discussion

6.3.1 Pyruvate Kinase

Pyruvate kinase was seen in this study to be down regulated in cancerous tissue compared to healthy with a history, but was not seen in healthy tissue with no history of breast cancer. This suggests that there may be a role for pyruvate kinase in the early stages of breast cancer, it may also be involved in signaling or as a circulating factor. Isidoro *et al* [186] found pyruvate kinase to be involved in breast cancer in the past in a proteomic study to find diagnostic targets, as has been done here.

Pyruvate kinase is a glycolytic enzyme that is present in many forms. Pyruvate kinase's inferred importance in cancer became clearer when tumour specific M2 was isolated in the late 1970's. Before that however, pyruvate kinase had been implicated in several cancers including breast cancer in 1970 [187]. Consistent research over the last 3 decades has seen the physical and enzymatic properties of pyruvate kinase heavily researched. Most recently studies relating pyruvate kinase to breast cancer have indicated that the levels of the M2 isozyme are altered by the use of Herceptin [188].

This finding may explain the result that indicates the lower levels of pyruvate kinase in the cancer tissue. The tumour specific nature of trazbutamab (Herceptin), and the high levels of pyruvate kinase in the healthy tissue of women with a history of breast cancer suggest that the effect of the drug on pyruvate kinase levels may be localised. This is

speculation as the treatment regimes for the patients involved have not been examined in this study.

6.3.2 PHOSPHATASE 2A – SSP 2402

Phosphatase 2a (PP2A) has been shown to be involved in many aspects of breast cancer, including the regulation of Estrogen Receptor expression, and telomerase activity in breast cancer cells [189-191].

The results for the expression of this regulatory subunit of PP2A are difficult to interpret. There is a concentration increase in the healthy tissue associated with breast cancer compared to the unassociated healthy tissue. The concentration in the cancerous tissue is even lower though, suggesting that there may be an early, perhaps tumour establishing role of PP2a in breast cancer. Alternatively, given that all of the samples except one in the healthy, breast cancer associated tissue are in an older age group it may be an age related difference that fails to be exhibited in cancers such as the need for ER in many cases. If comparison of ER status to the concentration of PP2a was to be conducted with these tissues it may be possible to look at this more conclusively.

Another possibility for the differential regulation between the cancer tissue and the tissue that was associated with breast cancer may lie within a paper by Rossini *et al* in 1997 [192]. Rossini *et al* showed that a loss of PP2A function through inhibition led to apoptosis via an unknown 53kda kinase.

In 2000 Ruediger *et al* isolated two mutations of the A subunit of PP2A that were specific to breast carcinoma. One of these is thought to prevent B subunit binding, whilst the other is thought to be unable to bind either B or C subunits [193]. Ruediger *et al* went on then implicate PP2A as a tumour suppressor.

These results offer explanation for the results found here. Absence of PP2A isoforms in the true healthy controls may indicate a normally functioning system, the excess seen in the normal tissue with a history of breast cancer association, indicates that PP2A deregulation may be a precursor for cancer development, whilst the level of the isoform has gone up the protein as a whole is potentially dysfunctional, and the lack of function results in apoptosis. The low level changes in the cancer tissue may reflect an ongoing deregulation of the PP2A protein, but secondary changes to oncoproteins involved in

the apoptotic pathway have resulted in a loss of apoptosis, despite abnormal PP2A functioning.

The question must be raised as to PP2A's use as a diagnostic as the Australian tissue showed expression of this isoform in all 3 late stage cancers included in this study.

Whether there is a population based element to the isoforms is yet to be determined, however preliminary results here and previous studies indicate that the role of PP2a in breast cancer prognosis and this isoform specifically in Australian women are in need of further investigation.

6.3.3 GI 51491284

Gi 51491284, also known as DKFZp781, is a hypothetical protein whose existence was inferred from the cDNA sequencing of tissue that originated from the retina according to the NCBI database [194]. This study now confirms its presence in breast tissue.

Sequence data and structural modeling of the protein suggests it has a role as a zinc binding protein. It has been found in this study to be a nuclear protein, which supports this zinc finger (ZNF) binding hypothesis. Other structural implications in the NCBI entry suggest the presence of a helicase which implies a role in DNA replication, repair and/or remodeling.

There have been no previous functional studies on this protein but the results found here presents strong evidence for a ZNF binding function in breast cancer tumours and that this may be present at a very early stage and may elude to outcome or drug response, although these possibilities are yet to be confirmed.

The potential of isoforms, other that the two identified here is high and both normal and cancerous tissues need to be examined to determine the existence of other forms of this previously hypothetical form. The biggest indicator that there may be other isoforms is visible in Table 6.1 when looking at the theoretical and experimental Mw's of the two isoforms sequenced. They are approximately one third the size of what is expected from the sequence. The MALDI-TOF MS based peptide sequencing revealed matches along the entirety of the sequence, suggesting alternate splicing, post translational cleavage or modifications as opposed to a C or N terminal truncation. The possibility of the protein being in the process of synthesis or degradation is also ruled out by this observation.

Whilst isoform SSP 1204 is not particularly common, SSP 1201 may provide crucial clues to the prognosis and detection of breast cancer and the likelihood of continued risk of the disease. Two patients that have no history of breast cancer expressed the SSP 1201 isoform and whether they go on to develop breast cancer is of high interest. It may take many years of follow up for these patients to present with breast cancer, if they present at all.

6.4 CONCLUSIONS

This chapter has demonstrated the importance of true healthy controls. This pilot investigation has revealed several proteins that may be involved in either the early stages of breast tumour formation, or in the signaling of tumour cells once a tumour has formed.

Whether these changes are exclusive to breast tissue remains to be seen, however the broad potential of these proteins in cancer treatment and diagnosis, if further investigations replicate these initial investigations findings would offer new hope in the push towards earlier diagnosis and treatment of cancer.

7 FUNCTIONAL INVESTIGATION INTO THE INVOLVEMENT OF β-CASEIN, IGF-1 AND STAT5 IN BREAST CANCER.

7.1 SUMMARY 137
7.2 RESULTS 138
7.2.1 Western blotting of target proteins 138
7.2.2 Investigation in to a relationship between IGF-1 and β-casein 140
7.2.3 T47D cultures 142
7.2.4 Proteomic Analysis and Western Blots of cell cultures 143
7.2.5 Differentially regulated proteins in vitro 144
7.3 DISCUSSION 145
7.3.1 Western blotting of IGF-1, β-casein and STAT5a 145
7.3.2 A relationship between IGF-1 and β-casein 146
7.3.3 In vitro investigations into β-casein, IGF-1 and STAT5 147
7.4 CONCLUSION 148

7 FUNCTIONAL INVESTIGATION INTO THE INVOLVEMENT OF β-CASEIN, IGF-1 AND STAT5 IN BREAST CANCER.

7.1 SUMMARY 137
7.2 RESULTS 138
7.2.1 Western blotting of target proteins 138
7.2.2 Investigation in to a relationship between IGF-1 and β-casein 140
7.2.3 T47D cultures 142
7.2.4 Proteomic Analysis and Western Blots of cell cultures 142
7.2.5 Differentially regulated proteins in vitro 144
7.3 DISCUSSION 145
7.3.1 Western blotting of IGF-1, β-casein and STAT5a 145
7.3.2 A relationship between IGF-1 and β-casein 146
7.3.3 In vitro investigations into β-casein, IGF-1 and STAT5 147
7.4 CONCLUSION 148

7.1 SUMMARY

In previous studies we have shown β-casein to be up regulated in breast cancer [195]. In this study β-casein was investigated further, adding a functional component to this proteomic analysis of breast cancer.

β-casein is a protein that is expressed at terminal differentiation of mammary epithelial cells in response to prolactin (PRL), typically during lactation [196]. β-casein is then secreted as one of several milk proteins into the ducts of the breast. PRL induces β-casein transcription through a set of receptor kinase mediators, specifically the JAK 2/ STAT5 phosphorylation pathway [197].

β-casein has often been used in breast cancer studies as an indicator of terminal differentiation, but reports correlating β-casein specifically to breast cancer prognosis have been contradictory [198-202].

Another protein that is often speculated to be involved in breast cancer but is still inconclusive is insulin-like growth factor-1 (IGF-1). IGF-1, as its name suggests acts as a growth promoter, its transcription is activated by growth hormone (GH) and it is secreted. IGF-1 like β-casein is regulated through the JAK2/ STAT5 pathway [203]. It is this commonality that lead to the investigation into whether there is a relationship between β-casein and IGF-1 in breast cancer.

Analysis of the relationship began by comparing the data for cancer and healthy controls in the Australian population using the samples described in previous chapters, based on protein identification from western blotting.

The analysis of the trend between IGF-1 and β-casein suggested that there is a relationship in healthy tissue. The apparent trend was not present in cancerous samples and this led to further questions. A mammary carcinoma cell line T47D was cultured to look at the affect of different hormone combinations on the concentration of IGF-1 and β-casein *in vitro*. Prolactin (PRL), growth hormone (GH) and estradiol (E_2) were all used in varying combinations. The results from the treatment of the cells were inconclusive. IGF-1 and β-casein levels varied with no relationship evident between the two.

Differential analysis of the proteins that were altered by these hormone applications are still to be fully elucidated, however they offer promise for finding the important regulators in the JAK2/STAT5 activation pathway.

7.2 RESULTS

7.2.1 WESTERN BLOTTING OF TARGET PROTEINS

Both IGF-1 and β-casein were located by western blotting. The location of these proteins can be seen in figure 7.1a, 7.1b and 7.1c. Confirmation of their identities was attempted by excising the proteins and sending them to the Australian Proteomic Analysis Facility (APAF) in Sydney for MALDI-TOF MS analysis. The results from the MALDI-TOF MS were inconclusive for the identity of both proteins.

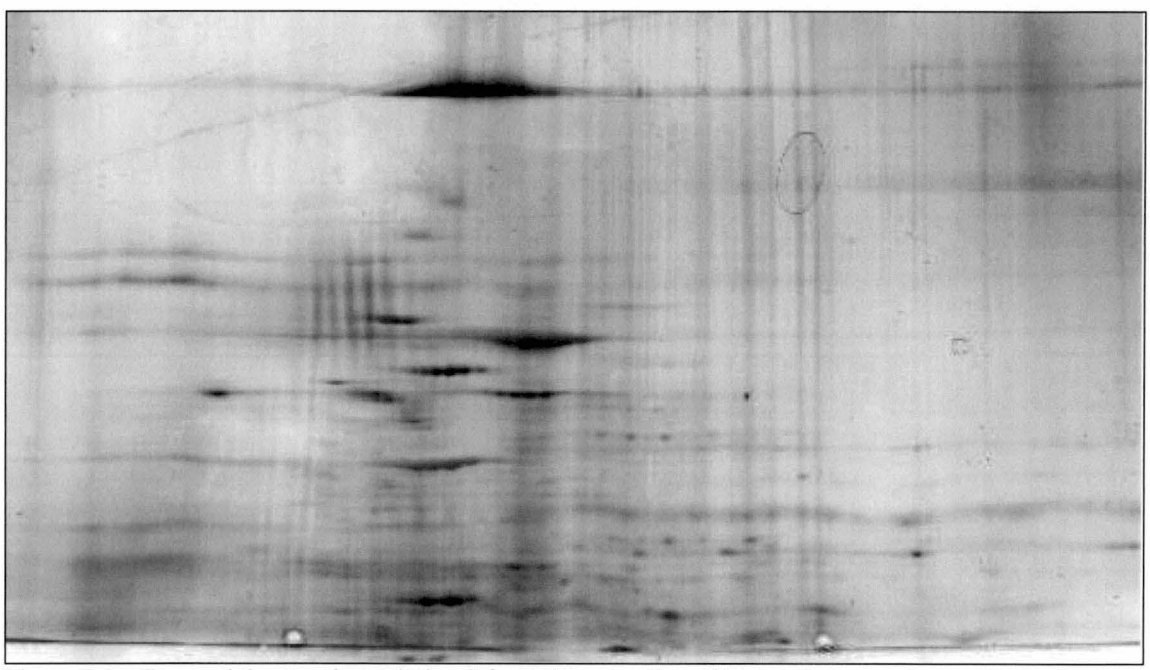


Figure 7.1a: Sequential extraction solution 2 from the extraction of breast tissue run in two dimensions on to a polyacrylamide gel TCB23 SOL2. Note the presence of artifacts on the image centre to right on the image, this is due to air bubble formation during imaging and is removed during the analysis and normalization process.

Western blotting of STAT5a and STAT5b initially revealed a lack of full length STAT5 present in the samples (Figure 7.1d). Literature has indicated the existence of a STAT5 specific protease [204]. To investigate whether the protease inhibitor cocktail that was employed in this study was sufficient led to the separating of freshly extracted samples, resulting in the development of an altered protocol.

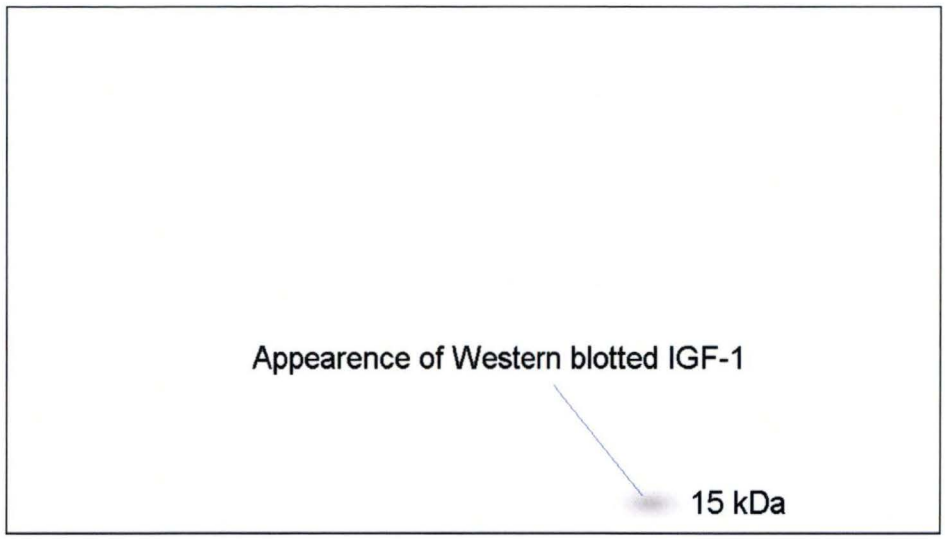


Figure 7.1b: Western blot of IGF-1 using a monoclonal X terminal directed antibody indicating a dimer located at 15kda.

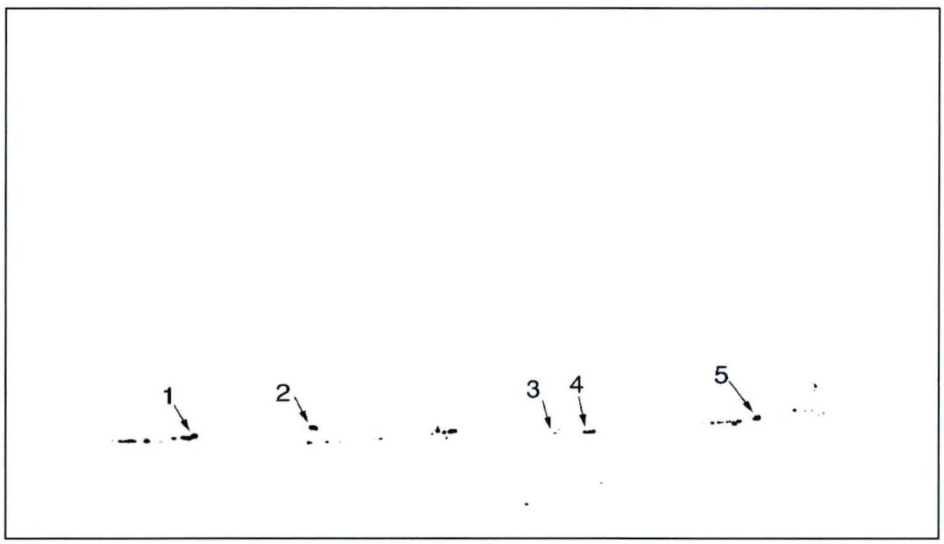


Figure 7.1c: Western blot of β-casein using a monoclonal antibody directed against the y terminal. Results indicate a series of proteins of approximately 25kDa.

PMSF (phenylmethylsulfonyl fluoride), a serine protease inhibitor was added to the extraction of the samples, the details of these changes are outlined in Chapter 3. The newly extracted tissue was separated and the gels were re-run. The western blotting process was then repeated. The improvement in the gels and then subsequent determination of the location of STAT5a can be seen in Figure 7.1e.

Attempts were made to locate STAT5b on both the pre- and post- PMSF treated gels but they were unsuccessful.

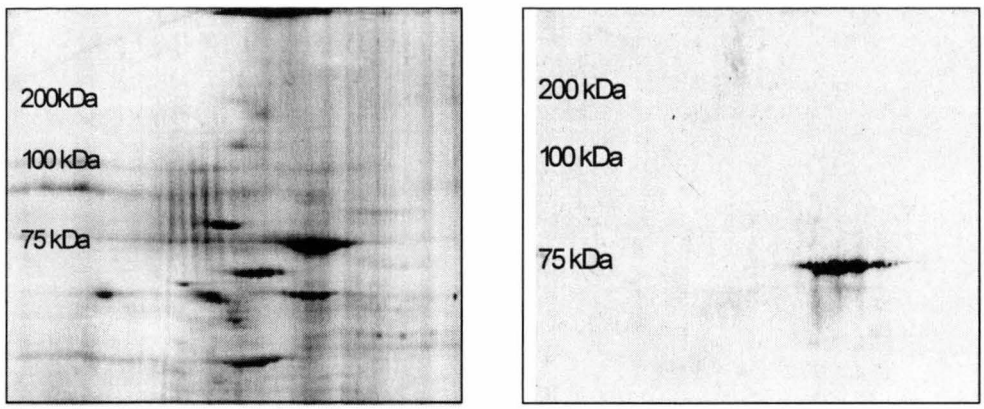


Figure 7.1d: Western blot of STAT 5a. On the left is the gel obtained through separation and on the right an image of the western blot, indicating a breakdown product of STAT5a.

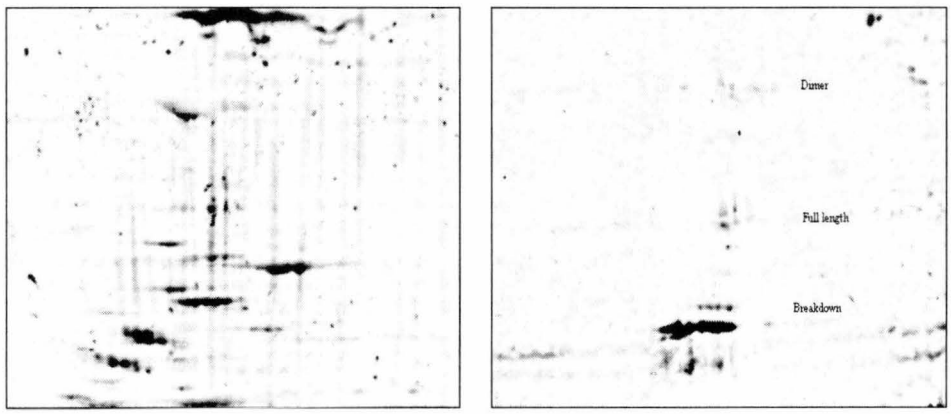


Figure 7.1e: Using PMSF, the normalized gel on the left shows more distinct separation of the proteome. On the right, due to this additional serine protease inhibitor western blotting led to not only a full length version at 95kDa but also a dimer of STAT 5a at ~190kDa being revealed as present in the cytoplasm of breast tissue. These figures demonstrate normalized gel images

7.2.2 A RELATIONSHIP BETWEEN IGF-1 AND β-CASEIN?

To investigate any potential relationship between the expression levels of IGF-1 and β-casein in human breast tissue the concentration levels for the Australian tissue samples, healthy and cancerous, were intentionally ordered from highest concentration to lowest concentration of IGF-1. For each patient the corresponding, collated, β–casein concentration data for the isoforms detected in Roberts *et al* in 2004 [1], were included alongside the IGF-1 data in a graphical format. The results from this analysis can be seen in Figure 7.2a and 7.2b.

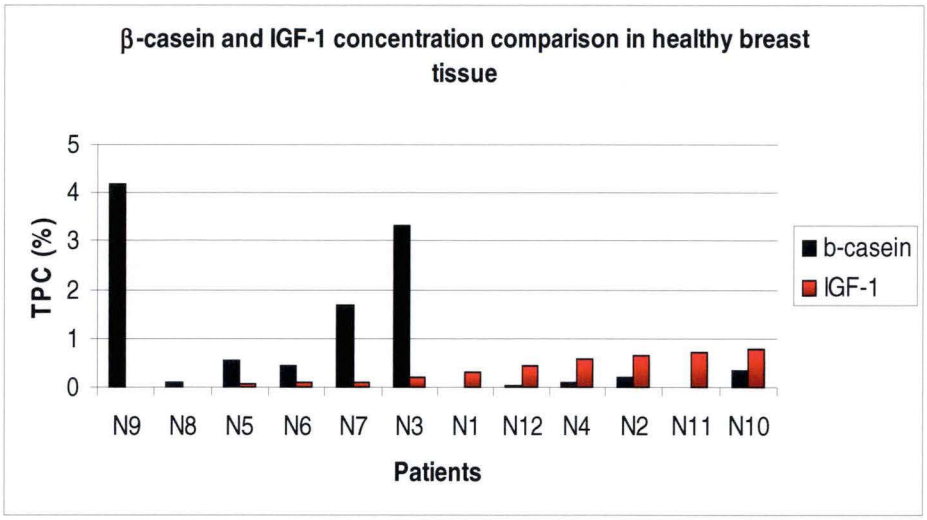


Figure 7.2a: Trend analysis of healthy breast tissue shows that at the 6th patient (N3) the βcasein level is steadily rising with the level of IGF-1. However from the 7th patient (N1) the level of βcasein expression drops and is consistently much lower than that of IGF-1.

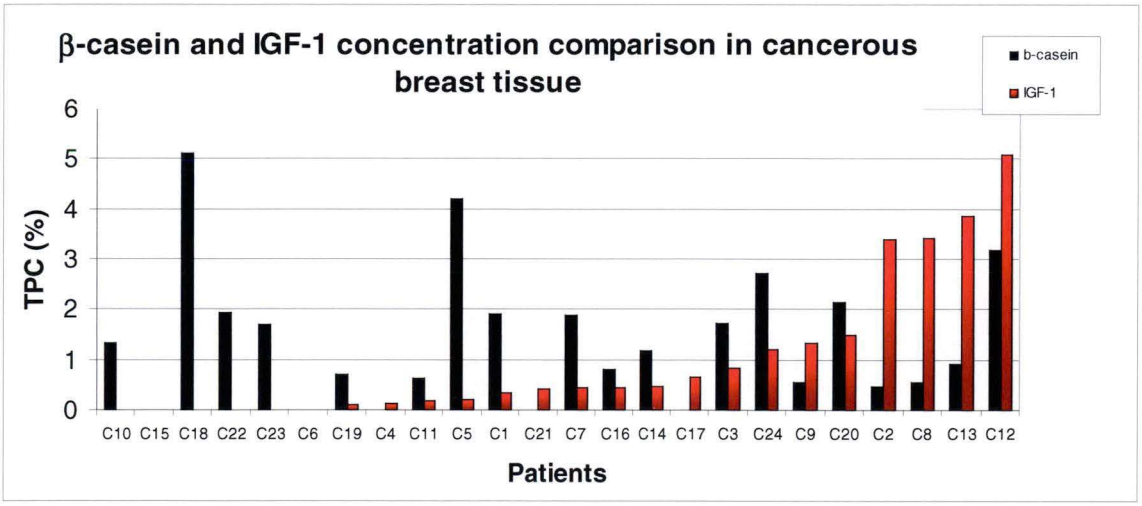


Figure 7.2b: Trend analysis of cancerous tissue shows that there are only several patients which are expressing higher levels of IGF-1 than β-casein. The lower levels of IGF-1, unlike in the healthy tissue fail to show a relationship between the level of IGF-1 and β-casein.

Healthy tissue analysis in this fashion observed an inversely proportional relationship between IGF and β–casein. In the cancerous tissue it was found that there was no such relationship and expression levels of both proteins are more erratic.

Correlation analysis confirms the lack of a correlative relationship in the cancerous tissue with correlation coefficient of r = 0.03. The same analysis style suggested that there is also no relationship in the healthy tissue. Further observation of the healthy samples shows that the concentration level of β-casein drops and then begins to increase again. Analysis done again, taking into account this pattern found, a correlation coefficient of r = 0.95 and 0.63 for the first 6 samples and the final 5 samples respectively is obtained in the healthy tissue.

The pattern of β-casein expression that was observed could not be attributed to either the age of the patients or any history of breast cancer for the patients. It is possible however that the concentration of IGF-1 is acting as an inhibitor of β-casein expression after it reaches a threshold level. In the cancer tissue, 4 samples have higher IGF-1 levels than β-casein levels. Three are stage 2 tumours one is a stage 3 tumour. As of 2006, 3 out of these four women have died (75%) compared with 39% of all of the Australian cancer patients that participated in the study.

7.2.3 T47D CULTURES

T47D, a breast carcinoma cell line, was set up to investigate the STAT5 pathway, a signaling pathway known to regulate IGF-1 and β-casein through two of its members; STAT5a and STAT5b, *in vitro*. The following preliminary experiments were conducted to establish the viability of investigation of the STAT pathway using this cell line.

To establish the amount of protein that is needed for these analyses, cells were harvested from both T25 and T75 cell culture flasks when the culture reached confluence. The total weight of cells collected for the T25 flask was 40mg and 150mg for T75 flask. This difference determined that all experimental treatments were to be conducted in T75 flasks that were at confluence. The protein concentration was also much higher for Solution 3 extracts with the greater amount of starting material which is worth considering for later experiments on these samples if nuclear proteins become of interest.

7.2.4 Proteomic Analysis and Western Blots of Cell Cultures

The protocol for cell harvesting and treatment can be seen in Chapter 3. Post harvesting the cells underwent sequential extraction, as was also outlined in Chapter 3. The cells also underwent Bradford protein analysis to determine how much protein had been extracted and determine the dilution, if any to load on to the gels to achieve 1μg/μl of protein to load onto the 2 dimensional gels. This was to ensure as much consistency as possible.

Western blotting of β-casein and IGF-1 was conducted on gels from each of the treatment groups. There were, however two isoforms detected in some of the gels that underwent comparison. Only the cells that had received treatment from all three hormones showed signs of both IGF-1 and β-casein, whilst GH alone treated cells were the only other treatment group to show IGF-1 expression.

The concentration levels of IGF-1 and β-casein were low and there were no patterns from the presence or absence of either observed from any of the hormone combinations that were applied to the cell cultures, this can be seen in Figure 7.2.4a.

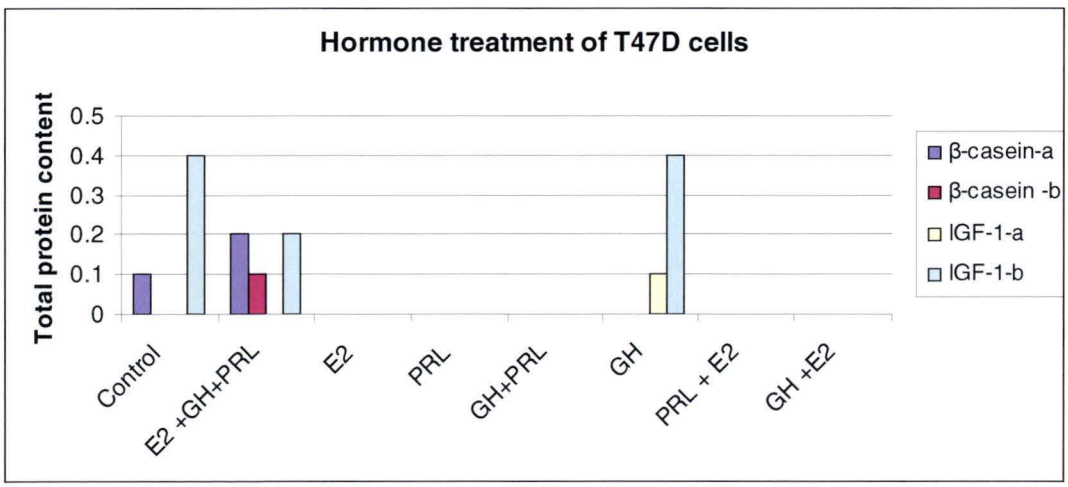


Figure 7.4.4a: Percentage of total protein expressed in each cell line treatment group, for both isoforms of β-casein and IGF-1 that were isolated.

Where only two hormones were applied there appeared to be no expression of either protein, in either isoform. This suggested that the 24 hour time point may have not been optimal, despite previous reports to the contrary or that the application of only two of the chosen hormones is not optimal for transcribing either IGF-1 or β-casein.

7.2.5 DIFFERENTIALLY REGULATED PROTEINS IN VITRO

Differential analysis was conducted on the gels obtained from the *in-vitro* study of IGF-1 and β-casein. Differentially regulated proteins were found when comparing the treatment groups. Of particular interest are the control versus hormone alone and the combinations of each hormone with E_2.

The numbers of proteins up and down regulated can be seen in Table 7.3. Interestingly it can be seen that treatments with more than one hormone limit the number of proteins that are down regulated dramatically. Conversely however the introduction of two or more hormones increases the number of proteins that are up regulated in all cases but GH alone.

Control versus...	Up regulated Proteins	Down regulated proteins
Estradiol (E_2)	19	58
Prolactin (PRL)	19	59
Growth Hormone (GH)	39	12
GH + E_2	32	7
PRL + E_2	28	4
PRL + GH	27	5
PRL + GH + E_2	24	5

Table 7.3: The number of proteins that was up or down regulated in T47D breast cancer cells in response to different hormones was highly variable.

The number of proteins that are altered by each treatment yields some clear observations. The application of E_2 or PRL result in more proteins being down regulated than up regulated, and this is the opposite for GH treated cells. The total number of proteins that are altered declines dramatically when all three hormones are combined. The proteins that are commonly up or down regulated are the ones that are of most interest. There numbers are much smaller and although they are yet to be identified, the comparison of common proteins can be seen in table 7.4.

	E_2		PRL		GH		GH/E_2		PRL/E2		PRL/GH		$PRL/E_2/GH$	
	↑	↓	↑	↓	↑	↓	↑	↓	↑	↓	↑	↓	↑	↓
E_2	-	-	5	15	2	4	6	5	3	0	5	2	0	1
PRL	5	15	-	-	4	6	8	3	5	0	4	1	4	1
GH	2	4	4	6	-	-	9	2	6	1	6	1	11	2
GH/E_2	6	5	8	3	9	2	-	-	7	0	6	0	5	1
PRL/E_2	3	0	5	0	6	1	7	0	-	-	6	0	5	0
PRL/GH	5	2	4	1	6	1	6	0	6	0	-	-	5	0
$PRL/E_2/GH$	0	1	4	1	11	2	5	1	5	0	5	0	-	-

Table 7.4: The number of proteins commonly up and down regulated by each of the hormone treatments.

7.3 DISCUSSION

7.3.1 WESTERN BLOTTING OF IGF-1, β-CASEIN AND STAT5A

The western blotting of IGF-1 and β-casein showed their location on a 2D gel. The inability of MALDI-TOF MS to confirm their identities is an area that needs to be addressed. Both proteins are very small, with β-casein being 27kDa and IGF-1 just 7kDa. These small sizes provide limited trypsinisation targets, namely lysine and arginine residues. This results in a limited number of peptides, combined with this is a small protein concentration and the sample becomes unstable under the MALDI-TOF MS process. Whilst MS instrumentation is becoming more sensitive, these proteins are at the lower limits of success for the basic MALDI-TOF procedure that was employed.

STAT5 was also western blotted and located on a 2D gel. A failure of the original methodology to show full length STAT5 on 2D gel was determined and whilst the reason behind this was found, it raises more questions surrounding the effectiveness of the protease inhibitor cocktail and the possibility other proteins may not have been protected from proteolytic degradation.

STAT5b was elusive in all western blots regardless of PMSF addition. This has been attributed to the traditional lack of STAT5b in mammary tissue.

7.3.2 A RELATIONSHIP BETWEEN IGF-1 AND β-CASEIN

The presence of a unique trend between IGF-1 and β-casein protein concentrations in healthy tissue was observed. The most curious finding here was the drop in β-casein once IGF-1 reached a threshold level followed by a continued rise in β-casein again. What is causing this "resetting" of β-casein expression is of interest. Prior to the decrease, the level of β-casein are consistently higher than that of the corresponding IGF-1. After the drop, the β-casein concentration level fails to reach that of IGF-1. The suggestion would be that there is a threshold IGF-1 concentration, once the IGF-1 concentration rises above this level it inhibits the expression of β-casein. The mechanisms behind this are, at this stage unclear. It was determined from that neither β-casein or IGF-1 concentration was dependent on the age of the patients or whether they had a personal history of breast cancer.

The activation of β-casein and IGF-1 transcription are dependant upon the JAK2/STAT5 pathway, in breast tissue. The inhibition of β-casein expression has been reported previously through the phosphorylation of Y779 in the Transcriptional Activation Domain (TAD) of STAT5a due to a lack of co-stimulation [205]. So it is possible that recruitment of STAT5a to transcribe IGF-1 may be due to IGF-1 concentrations influencing the phosphorylation at this residue and therefore inhibiting β-casein formation.

Autocrine production of IGF-1 in lung cancer tissue has recently been reported [206]. It is possible that a similar mechanism is being seen here in healthy breast tissue, with the autocrine pathway becoming self sustaining at a minimum concentration level of IGF-1. The signaling cascade of this autocrine system is yet to be elucidated, however given the lack of STAT5b in breast tissue along with the high level of reported redundancy between the STAT5 proteins in mammary tissue [207], the recruitment of STAT5a to aid IGF-1 production, which has been reported in 2003 [208] is suspected, resulting in a decreased rate of β-casein production.

Whilst it is widely recognized that β-casein is an endpoint indicator of mammary differentiation, it may prove to be important in future studies to investigate the IGF-1 status alongside it. If an autocrine system of IGF-1 exists, indicated by a higher

concentration of IGF-1 over β-casein, then it suggests that there may be a higher likelihood of deregulation of other STAT5a related pathways. Studies have suggested that poor prognosis is linked to a depletion of STAT5a, and others have linked it to an increase in IGF-1. This study supports the involvement of IGF-1 in the prognosis of breast cancer. It is possible that these two factors may be directly linked.

Having said this, the presence of higher IGF-1 in healthy tissue needs to be further investigated. Whether there is a genetic reasoning or the elevated IGF-1 levels seen are a result of signaling, needs to be determined. The three women with high levels of IGF-1 in the healthy cohort, all have no history of breast cancer, whether it is predictive of likelihood to develop breast cancer is an important element in determining the importance of IGF-1 in the early detection of breast cancer.

Understanding this mechanism will offer insight into prognosis of breast cancer patients in the future. The inconclusive results seen previously for IGF-1 and β-casein individually may be resolved by looking at these proteins as a collective.

7.3.3 IN VITRO INVESTIGATIONS INTO β-CASEIN, IGF-1 AND STAT5

The *in vitro* studies using T47D cell lines showed inconclusive results for the presence of a link between β-casein and IGF-1 using the hormones known to activate the pathways. It is highly likely that there are more elements at play than just the presence of an activator, and a co-activator, which was present in the form of E_2 in this study. Whilst co-activation has been shown to be an important aspect of STAT5 pathways activation [205, 209], it has also proven itself to be a much more complicated signaling pathway, activated and sustained by many other protein cascades including the MAPK and PI3K pathways [210].

The differential analysis confirmed the complexity of this system by demonstrating the importance of looking at pathways in a true context. In this study with the treatment of just one hormone many proteins were altered, the more hormones that were used to treat the cells the lower the number of proteins that were altered. Rarely in the human body would a cell be influenced by one individual hormone at a time.

The lowered number of proteins in the multi-treatment approach may indicate the proteins that are functionally altered significantly in the cell and hence shed light on the susceptible targets of this JAK2/STAT5 pathway.

The differential analysis revealed proteins that are consistently altered between treatment groups. It is these proteins that may offer an important insight in to the proteins that are more prone to change *in vivo*. Whilst there is more work to do in this area, the identity of these proteins may reveal targets that are crucial points in the deregulation of a cell and the initiation of tumour formation.

7.4 CONCLUSION

The results from the functional approach to proteomics in this study were not as definitive as originally hypothesized. This pathway is a very complicated pathway and the differential analysis that followed the functional treatments allude to the complexity of this system and how much there is to still understand. Whilst a distinct relationship between β-casein and IGF-1 was not in this case proven, the possibility still warrants further investigation and the footing for this research has been laid in the differential analysis that was performed.

8 Final Discussion

8.1 Introduction 150
8.2 Key findings from this study 153
8.2.1 Confirmation of proteins known to be involved in breast cancer 153
8.2.2 New candidates in the search for breast cancer biomarkers 153
8.2.3 Novel proteins in the search for breast cancer targets 153
8.2.4 IGF-1 – new potential in the prognosis of breast cancer 154
8.3 Differential to functional experiments – new proteins to find. 154
8.4 Future experiments 155
8.4.1 Confirmation of the proteins isolated 155
8.4.2 Functional and differential Cell culture studies 155
8.4.3 Antibody production to the unknown proteins 156
8.5 Final Conclusions 156

8.1 INTRODUCTION

This study set out to isolate proteins that are involved in breast cancer that may be used in the future as a prognostic, diagnostic or therapeutic target. Proteomics was chosen as the primary technique to achieve the aims and during the time that this study has occurred there have been many advances in the technology associated with proteomics. It is important to acknowledge these improvements, as whilst many could not be incorporated in this study due to a desire to ensure reproducibility of the gels between the samples extracted in 2000, namely the Greek and Australian samples, with those extracted specifically for this thesis, further studies involving the findings presented here demand that these changes be incorporated to provide the best chance at fully characterising the chosen target proteins. Many of these changes lie in two areas, sample extraction and mass spectrometry based identification.

Despite the limitations surrounding this project, there were several key findings that may promote better treatments prognosis of breast cancer in the future. The differential analysis approach used in this study has successfully identified proteins likely to be involved in cancer establishment of progression, it has successfully pinpointed several proteins that may result, with further research in establishing why breast cancer is more aggressive in younger women and it has shown proteins that may be involved in a more systemic change in response to breast cancer, indicated by protein changes in apparently healthy tissue.

The functional study into the interactions of IGF-1 and β-casein was less conclusive in determining the exact way in which the proteins may affect each other. It has raised several questions though and implicates that there is a relationship between the two proteins, but through what signaling mechanism remains unclear. This in vitro approach has also enabled a differential study to take place, which with further analysis will aid in determining just how hormones affect breast cancer cell lines and through what signaling cascades.

By combining approaches both comparative and functional proteomic approaches this study has several key findings, in relation to new markers for breast cancer as well as the use of proteomics in the attempt to find breast cancer target proteins,

8.2 LIMITATIONS

8.2.1 SAMPLE EXTRACTION TECHNIQUES

As previously stated there have been many developments in proteomic sample preparation including the reduction and solubilisation of whole tissue samples. Techniques today allow for separation of specific sub cellular components or specific post translational modifications, allowing for the study of proteins in a given sub cellular environment, or in their active or non active forms. In 2003 when this study began it was desired that the Greek and Australian gels from previous phenotype comparisons between populations be included. This required the sample extraction methods remain the same to ensure reproducible and comparable gels. It should be mentions that whilst the gel images were used for this study, complete reanalysis of the gel occurred with an updated version of the gel analysis software PDQuest 7.1. Whilst these slightly imperfect gels limited some of the lines of investigation, it served its purpose sufficiently in that it resulted in reproducible images of breast cancer from all three populations and extended the reach of this comparative study to incorporate populations from around the world.

Future studies are now able to optimize the procedures used, depending on the source of the samples to be included. This was seen in Chapter 7 where the T47D cells were extracted, using an optimized protocol, allowing for a more optimal protein load on the gel, higher concentrations of protein inhibitors, altered reduction techniques to ensure a lack of post extraction dimerisation of unrelated proteins and the treatment with nuclease to prevent nucleic acid-protein interaction which causes a lack of focusing during IEF.

Chapter 7 also highlighted problems with the protease inhibitor in the initial protocol not being adequate to isolate STAT5a. With a little adjustment to the proteases used the image achieved through the 2 dimensional separation process was enhanced significantly. Whilst this limitation does not undermine the findings of this study, it may suggest that there are more proteins to find in relation to breast cancer, and it may also explain why the MALDI-TOF analysis done at the beginning of this study provided so many unidentifiable proteins.

Experiments yet to be conducted can now choose to optimise samples even further, using extraction techniques more suited to specific sub cellular regions as opposed to the whole cell approach that was used in these studies.

8.2.2 TECHNOLOGICAL ADVANCES

Mass spectrometry sensitivity, whilst it has improved at exponential rates is an area of the research that this study had limited control over, due to the outsourcing of these experiments to APAF. Since this study began MALDI-TOF mass spectrometry has seen many improvements. There are now mass spectrometers that are potentially able to successfully identify proteins such as IGF-1 of β-casein, or proteins in lower concentrations. These highly developed Mass spectrometers were not however owned by the venue the proteins were sent to for processing, nor were they marketed at the time these proteins were sequenced.

Future applications of mass spectrometry in regards to the proteins found to be differentially regulated in this study that are yet to be identified may yield a better success rate at identification with a more up to date mass spectrometer, or with more efficient extraction techniques.

8.2.3 CONTROL SAMPLE LIMITATIONS

Throughout the first three chapters involving multiple populations there was constant limiting factor. This is the lack of control samples from each of the three nationalities, as opposed to just the Australian healthy samples. The first step to confirming the multiregional application of these results is to obtain healthy control samples from both the Greek and Taiwanese populations. Failing this, a much larger cohort of Australian samples, both healthy and malignant tissue, needs to be examined with separate studies conducted in both Greece and Taiwan.

Despite the advances and limitations, this project has successfully achieved its aims. It has found a number of proteins that offer potential for new targets for further research in breast cancer. This study has made a large step towards identifying global markers of breast cancer that will be significant to more than just the local population, however there are still are still many questions to be answered to put these new markers to their optimal use.

8.3 KEY FINDINGS

8.3.1 CONFIRMATION OF PROTEINS KNOWN TO BE INVOLVED IN BREAST CANCER

There were several proteins that have been found to be involved in breast cancer previously, that were isolated in this study. The most significant of these was the finding of heat shock protein 60 (HSP 60, or chaperonin). The expression levels of HSP 60 seen in the cytoplasm and the nucleus reiterated what is already known about HSP 60. The expression of HSP 60 in higher concentrations in the nucleus/mitochondria (extraction solution 3) of healthy tissue samples was complemented by the higher level of HSP 60 in the cytoplasm (extraction solution 2) of the cancer samples. This pattern indicates that the known behavior of HSP 60 to migrate to the cytoplasm in times of stress is being observed by the proteomics techniques employed, indicated that the extraction process and the analysis process were working as expected.

8.3.2 NEW CANDIDATES IN THE SEARCH FOR BREAST CANCER MARKERS

Several candidate proteins were found in the hunt for new markers. These included some proteins that had not been associated with breast cancer in the past or had had limited association with breast cancer. Isolated proteins, such as lumican, osteoglycin, and folliculin, with further study, offer new possibilities for breast cancer as potential therapeutic targets and diagnostic factors. Some proteins such as α-B crystallin, when this study begun had not been associated with breast cancer are now suspected to be an oncoprotein. This change in perception of this protein demonstrates the ever changing field of knowledge surrounding proteins involved in cancer and also demonstrates the need to make a change from focusing on the genes to focusing on the state of the functional proteins within the tissues of interest.

8.3.3 NOVEL PROTEINS IN THE SEARCH FOR BREAST CANCER TARGETS

FLJ 20309, RP3-393D12.2 and Gi 51491284 all offer new potential for the clinical approach to breast cancer. They were isolated in this study and have not been previously identified in their protein form. They are hypothetical proteins that, through various

studies are assumed to exist, based on cDNA constructs. Whilst the proteins that already have known and verified functions and structures are important in furthering breast cancer research and understanding the disease it is these new proteins that offer the most potential for new approaches to breast cancer. Specifically, finding out their role in breast tissue, how they may regulating pathways that lead to malignancies, and the ways in which to best take advantage of them are all of extreme interest.

8.3.4 IGF-1 – NEW POTENTIAL IN THE PROGNOSIS OF BREAST CANCER

IGF-1 was looked at in the context of trying to elucidate the reasons behind it continually being questioned in the literature as an indicative protein in breast cancer.

Constantly, IGF-1 has been reported to be up-regulated and then a series of papers will be published contradicting this. Here, it is shown that there is an up-regulation of IGF-1 in healthy tissue and in cancerous tissue in some but not all cases. These cases are not linked by age or previous involvement in breast cancer. It is suggested here though with preliminary data, based on the survival of the women in this trial that the level of IGF-1 in the tumour may indicate a poor prognosis. This may only be true however when taking in to account the concentration of β-casein and their transcription factor STAT5A. Whilst there is still a lot of research to be done to determine the exact predictive ability of IGF-1 in breast cancer patients it does offer a potential marker of prognosis.

8.4 DIFFERENTIAL TO FUNCTIONAL EXPERIMENTS

The proteins here that are differentially regulated by both PRL and by GH indicate proteins that are common in the signaling cascade. Whilst the functional component of this study yielded inconclusive results it results in more questions than it answered. The differential analysis has led to a large number of proteins that are highly likely to be involved in the JAK2/STAT5 pathway. Identification of these proteins will shed light on to the crucial points at which STAT regulation is directly controlled by the hormone that activates it. These are many proteins here that may provide a lot of promise for breast cancer in particular for the development of future therapy targets. The earlier finding of IGF-1 being a potential indicator of prognosis highlights why it is so important to understand, in intricate detail the workings of this pathway.

8.5 FUTURE EXPERIMENTS

Every chapter in this study has led to many questions being posed. There are several key areas though which offer a lot of potential for future research.

8.5.1 CONFIRMATION OF THE PROTEINS ISOLATED

Completing the identification process of these proteins is crucial to determining the best direction to take further experiments. PCR based studies to firstly confirm the specific protein and then determine whether the changes in protein concentration are due to transcription at a genetic level, or post translation changes is a key priority. Immunohistochemical studies may also be of use for determining the exact sub cellular location of proteins that have only one isoform. For the proteins with ore that one isoform highly specific antibodies would be needed to differentiate between the multiple isoforms. Although this is a difficult task it is possible with improvements in phosphorylation and glycosylation directed antibodies that have been developed in recent years to investigate post translational modifications and the activation states of proteins.

8.5.2 FUNCTIONAL AND DIFFERENTIAL CELL CULTURE STUDIES

The brief investigation into the role different hormone activators of the JAK2/STAT5 pathway highlighted some key issues with cell culture studies. The greater the number of hormones that was added the fewer the number of proteins were differentially regulated, indicated by the large drop in the number of proteins up or down regulated when 2 or all three of the proteins were applied. With the differential results that were presented, the proteins that are commonly altered in the multi-treatments groups need to be identified and their role deciphered. These proteins are the ones that change in a system that is more like the environment with in the breast tissue and therefore offer the proteins that are more likely to respond to the treatments that are developed in the future as others that are limited by multiple hormone application are assumed to be highly controlled and possible less likely to be deregulated.

Further cell culture systems need to be developed, cells that are both estrogen dependent and estrogen independent need to be looked at in order to establish the precise role of

these differentially regulated proteins. Cell culture systems other than breast cancer cell lines also need to be investigated to determine how widely applicable the changes are.

8.5.3 ANTIBODY PRODUCTION TO THE UNKNOWN PROTEINS.

The third area in which to follow up is in the area of the newly recognised proteins. These proteins offer many exciting possibilities. The first step to be made is the development of synthetic peptides and antibodies against key regions of the proteins thought to play a role in breast cancer. From this study 3 proteins have warranted such an investigation. RP3-393D12.2 and FLJ 20309, identified in chapter 4 indicated that they both have a potential use in the future treatment of breast cancer. At this stage however the functions of these proteins, if any, in healthy breast tissue is unknown. The third protein of interest was found in Chapter 6 with the isolation of a purely hypothetical protein Gi 51491284 in two separate isoforms. This like those from chapter 4, these proteins offer many new avenues in which to direct breast cancer research.

Preliminary investigations have shown that there are promising outlooks for further investigation into these proteins. The theoretical peptide generation algorithm by Jameson and Wolf has indicated that there are several areas on each protein that are highly antigenic and should allow for the generation of the appropriate peptides against these proteins. With these antibodies generated it is then possible to delve further in to the functions that these proteins have. They also offer potential new targeted therapy options. With these antibodies is possible that they will either offer a delivery system for new drugs, or be able to alter the proteins function and improve the course of breast cancer in millions of women all over the world.

8.6 FINAL CONCLUSIONS

This study provides one of few investigations that look at whole tissue from a global population. It has highlighted the variation of one population to the other despite all other aspects being equal. There were many proteins that were seen in just one or two populations. This doesn't rule out the proteins' involvement, it just requires further investigation in to why the differences exist and whether or not they will impair the proteins' ability to be used as a maker.

In Chapter 1, Table 1.1 listed the proteins that had been found to date in several comparative proteomic studies in to the proteins involved in breast cancer. That table

has now been amended to include the proteins identified in this study as seen in Table 8.1.

Proteins	Wulfkhele *et al*	Somiari *et al*	Luo *et al*	Deng *et al*	PhD study
(LKEEYQSLIR)	✓				
14-3-3 ζ/δ	✓				
78kDa Glucose related protein (GRP78)/ Ig heavy chain binding protein (BIP)				✓	
Actin, beta	✓	✓			
Actin-related protein 3 (Arp3)	✓				
Aldo-keto reductse family member 1, member a1			✓		
Alpha 1 antitrypsin				✓	
Alpha-1-antitrypsin x2		✓			
Annexin 1			✓		
Annexin II	✓				
Annexin II	✓				
Annexin V	✓				
Annexin VII	✓				
Anti-thrombin III	✓				
Apolipoprotein A1				✓	
Biliverdin reductase B	✓		✓		
CapZ	✓				
Carbonic anhydrase I: carbonic dehydratase		✓			
Cathepsin B	✓				
Cathespin D				✓	
Cellular retinoic acid-binding protein 2 (CRABP2)	✓				
Chain A of the Ga module complex					✓
Chain b of recombinant human fibrinogen fragment			✓		
Coactosin-like protein	✓				
Cofilin	✓				✓
Cofilin 1 non-muscle			✓		✓
Cycophilin a			✓		
Cycophilin b precursor			✓		
Cytokeratin 18(424 AA)		✓			
Destrin	✓		✓		
Dimethylarginine dimethylaminohydrolase (DDAH)	✓				
Disulfide-isomerase related protein 5a					✓
Elongation factor 1α	✓				
Elongation factor beta (ef-1-beta)				✓	
Endoplasmic reticulum associated amyloid beta peptide binding protein			✓		
Eukaryotic translation initiation factor eIF3h	✓				
FABP-3 (HFABP)	✓				
FABP-5 (EFABP)	✓				
Fatty acid binding protein 4			✓		
Fibrinogen gamma		✓			
FK506 binding protein	✓				
FLJ 20309					✓
FLJ10652 (KDVKPHPR)	✓				
Folliculin					✓
GDP-mannose phosphorylase B	✓				
Gelsolin		✓			
Gi 51491284					✓
Glycerol-3-phosphate dehydrogenase 1 (soluble)		✓			
GRP78	✓				
GRP94	✓				
HnRNPs A2	✓				
HnRNPs A2/B1	✓		✓		
HnRNPs A3	✓				
HnRNPs C	✓				
HnRNPs L*	✓				
HnRNPs X	✓				
Hsp 27	✓				
Hsp 60					✓

Proteins	Wulfkhele *et al*	Somiari *et al*	Luo *et al*	Deng *et al*	PhD study
Hsp 90	✓				
Human carbonic anhydrase 1			✓		
IG heavy Chain				✓	
Ig Light Chain				✓	
Intracellular chloride channel 1 (NCC27, CLIC1)	✓				
Lactate dehydrogenase a			✓		
Lactoferrin	✓				
L-Plastin	✓				✓
Lumican		✓			✓
Macrophage capping protein (CapG)	✓				
Manganese SOD			✓		
MGC: 22710			✓		
Nascent polypeptide-associated complex, a polypeptide	✓				
Osteoglycin					✓
Peroxiredoxin 1			✓		
Peroxiredoxin 1 (Thioredoxin peroxidase)	✓				
Peroxiredoxin 2	✓				
Peroxiredoxin 5	✓				
Phosphoethanolamine binding protein (PEBP)	✓				
Phosphoethanolamine-binding protein (PEBP), prostatic-binding protein	✓				
Phosphoglcerate mutase 1			✓		
Phosphoglycerate kinase			✓		
Phosphoribosylaminomidazole coarboxylase			✓		
Polymeric immunoglobulin receptor (pIgR)	✓				
PRO2619		✓			
Proapolipoprotein					✓
Procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase)		✓			
Profilin 1	✓		✓		
Prohibitin	✓				
Proteasome (prosome, macropain) activator subunit 1 (PA28 alpha)		✓			
Proteasome subunit alpha type 1 (PSMA1)				✓	
Protein phosphatase 2A (PP2A)	✓				✓
Pyruvate Kinase					✓
Pyruvate kinase, M1 isozyme (EC 2.7.1.40)		✓			
Rab 11a	✓				
Rho GDI	✓				
Ribosomal protein s12 (RPS12)				✓	
RP3-39312.2					✓
S100A11	✓				
S100A7	✓				
S100A8	✓				
Selenium binding protein	✓				
Serum Albumin		✓			✓
Similar to Peroxiredoxin			✓		
Similar to trisephosphate isomerase			✓		
Similar to tubulin, beta 5		✓			
Small ubiquitin-related modifier 3 precursor (SMT3a) (SUMO3)				✓	
Stathmin	✓				
Superoxide dismutase (SOD)				✓	
Transgelin (Acidic)	✓				
Transgelin (Basic)	✓				
Transgelin 2				✓	
Translationally controlled tumour protein (TCTP)				✓	
Transthyretin				✓	
Type VI collagen α1 chain	✓				✓
Type VI collagen α2 chain	✓				
Ubiquinol-cytochrome reductase complex core protein 2			✓		
Ubiquitin-conjugating enzyme 9 (Ubc9)	✓				
UDP Glycosyltransferase					✓

Proteins	Wulfkhele *et al*	Somiari *et al*	Luo *et al*	Deng *et al*	PhD study
Uracil dna glycosylase			✓		
Voltage dependent anion channel 2			✓		
Voltage-dependent anion Channel protein 1 (VDAC)	✓				
α B-crystallin	✓				✓
α-Enolase	✓				

Table 8.1: Comparative proteomic studies in to breast cancer conducted over the past 5 years.

This proteomic approach to breast cancer has shown been effective in the search for new targets for breast cancer research. Proteomics has proven itself to be an effective tool in recognising differentially regulated proteins in both whole tissue and cell culture models. This study has led to the several future areas of research that had previously been looked at in a very limited fashion in regards to breast cancer. The revelation of entirely new proteins with highly limited information, but very strong evidence linking them to breast cancer in this study are the proteins that show the most promise. Understanding why these proteins are expressed the way they are and utilising that will offer new options in the future treatment of breast cancer.

9 REFERENCES

1. Parkin, D.M., et al., *Global cancer statistics, 2002.* CA Cancer J Clin, 2005. **55**(2): p. 74-108.
2. AIHW. *Breast cancer in Australia: an overview, 2006.* Cancerseries no. 34. cat. no. CAN 29 2006 [cited.
3. Xiong, Q., et al., *Female patients with breast carcinoma age 30 years and younger have a poor prognosis: the M.D. Anderson Cancer Center experience.* Cancer, 2001. **92**(10): p. 2523-8.
4. Lannin, D.R., et al., *Difficulties in diagnosis of carcinoma of the breast in patients less than fifty years of age.* Surg Gynecol Obstet, 1993. **177**(5): p. 457-62.
5. http://info.cancerresearchuk.org/cancerstats/geographic/cancerineu/incidenceandmortality/. [cited.
6. Vlachonikolis, I.G., T.J. Aletra, and V. Georgoulias, *Incidence of breast cancer on Crete, 1994-1995.* Eur J Cancer, 2002. **38**(4): p. 574-7.
7. http://crs.cph.ntu.edu.tw/crs_e/estimate.htm. [cited.
8. Cheng, S.H., et al., *Unique features of breast cancer in Taiwan.* Breast Cancer Res Treat, 2000. **63**(3): p. 213-23.
9. Yoshida, K. and Y. Miki, *Role of BRCA1 and BRCA2 as regulators of DNA repair, transcription, and cell cycle in response to DNA damage.* Cancer Sci, 2004. **95**(11): p. 866-71.
10. Ziegler, R.G., et al., *Migration patterns and breast cancer risk in Asian-American women.* J Natl Cancer Inst, 1993. **85**(22): p. 1819-27.
11. Sulkowska, M., et al., *Leptin--from regulation of fat metabolism to stimulation of breast cancer growth.* Pathol Oncol Res, 2006. **12**(2): p. 69-72.
12. Lee, H.P., et al., *Dietary effects on breast-cancer risk in Singapore.* Lancet, 1991. **337**(8751): p. 1197-200.
13. Simopoulos, A.P., *The Mediterranean diets: What is so special about the diet of Greece? The scientific evidence.* J Nutr, 2001. **131**(11 Suppl): p. 3065S-73S.
14. Simopoulos, A.P., *The traditional diet of Greece and cancer.* Eur J Cancer Prev, 2004. **13**(3): p. 219-30.
15. Bougnoux, P., B. Giraudeau, and C. Couet, *Diet, cancer, and the lipidome.* Cancer Epidemiol Biomarkers Prev, 2006. **15**(3): p. 416-21.
16. La Vecchia, C., A. Altieri, and A. Tavani, *Vegetables, fruit, antioxidants and cancer: a review of Italian studies.* Eur J Nutr, 2001. **40**(6): p. 261-7.
17. Moorman, P.G. and P.D. Terry, *Consumption of dairy products and the risk of breast cancer: a review of the literature.* Am J Clin Nutr, 2004. **80**(1): p. 5-14.
18. Breikers, G., et al., *Potential protein markers for nutritional health effects on colorectal cancer in the mouse as revealed by proteomics analysis.* Proteomics, 2006. **6**(9): p. 2844-52.
19. Drew, J.E., et al., *Salicylate modulates oxidative stress in the rat colon: a proteomic approach.* Biochem Pharmacol, 2006. **72**(2): p. 204-16.

20. Lenaerts, K., et al., *Glutamine regulates the expression of proteins with a potential health-promoting effect in human intestinal Caco-2 cells.* Proteomics, 2006. **6**(8): p. 2454-64.
21. Chanson, A., et al., *Proteomic analysis reveals changes in the liver protein pattern of rats exposed to dietary folate deficiency.* J Nutr, 2005. **135**(11): p. 2524-9.
22. tom Dieck, H., et al., *Transcriptome and proteome analysis identifies the pathways that increase hepatic lipid accumulation in zinc-deficient rats.* J Nutr, 2005. **135**(2): p. 199-205.
23. Park, J.Y., J.K. Seong, and Y.K. Paik, *Proteomic analysis of diet-induced hypercholesterolemic mice.* Proteomics, 2004. **4**(2): p. 514-23.
24. Junghans, P., et al., *Dietary protein-related changes in hepatic transcription correspond to modifications in hepatic protein expression in growing pigs.* J Nutr, 2004. **134**(1): p. 43-7.
25. Beattie, J.H., et al., *Hepatic responses to dietary stress in zinc- and metallothionein-deficient mice.* Exp Biol Med (Maywood), 2006. **231**(9): p. 1542-7.
26. Corzo, A., et al., *Protein expression of pectoralis major muscle in chickens in response to dietary methionine status.* Br J Nutr, 2006. **95**(4): p. 703-8.
27. Hayashi, S., et al., *Estrogen and growth factor signaling pathway: basic approaches for clinical application.* J Steroid Biochem Mol Biol, 2003. **86**(3-5): p. 433-42.
28. Yamashita, H., et al., *Immunohistochemical evaluation of hormone receptor status for predicting response to endocrine therapy in metastatic breast cancer.* Breast Cancer, 2006. **13**(1): p. 74-83.
29. Engel, R.H. and V.G. Kaklamani, *HER2-Positive Breast Cancer: Current and Future Treatment Strategies.* Drugs, 2007. **67**(9): p. 1329-41.
30. Chan, S.K., M.E. Hill, and W.J. Gullick, *The role of the epidermal growth factor receptor in breast cancer.* J Mammary Gland Biol Neoplasia, 2006. **11**(1): p. 3-11.
31. La Vecchia, C., et al., *Oral contraceptives and breast cancer: a cooperative Italian study.* Int J Cancer, 1995. **60**(2): p. 163-7.
32. Rossouw, J.E., et al., *Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial.* Jama, 2002. **288**(3): p. 321-33.
33. Coombs, N.J., et al., *HRT and breast cancer: impact on population risk and incidence.* Eur J Cancer, 2005. **41**(12): p. 1775-81.
34. Dinger, J.C., et al., *Breast cancer risk associated with different HRT formulations: a register-based case-control study.* BMC Womens Health, 2006. **6**: p. 13.
35. Espie, M., et al., *Breast cancer in postmenopausal women with and without hormone replacement therapy: preliminary results of the MISSION study.* Gynecol Endocrinol, 2006. **22**(8): p. 423-31.
36. Hall, P., et al., *Hormone-replacement therapy influences gene expression profiles and is associated with breast-cancer prognosis: a cohort study.* BMC Med, 2006. **4**: p. 16.
37. Kroman, N., et al., *Time since childbirth and prognosis in primary breast cancer: population based study.* Bmj, 1997. **315**(7112): p. 851-5.

38. Lagerlund, M., et al., *Socio-economic factors and breast cancer survival--a population-based cohort study (Sweden).* Cancer Causes Control, 2005. **16**(4): p. 419-30.
39. Dumitrescu, R.G. and I. Cotarla, *Understanding breast cancer risk -- where do we stand in 2005?* J Cell Mol Med, 2005. **9**(1): p. 208-21.
40. Baquet, C.R. and P. Commiskey, *Socioeconomic factors and breast carcinoma in multicultural women.* Cancer, 2000. **88**(5 Suppl): p. 1256-64.
41. Smigal, C., et al., *Trends in breast cancer by race and ethnicity: update 2006.* CA Cancer J Clin, 2006. **56**(3): p. 168-83.
42. Tansavatdi, K., B. McClain, and D.M. Herrington, *The effects of smoking on estradiol metabolism.* Minerva Ginecol, 2004. **56**(1): p. 105-14.
43. Johnson-Thompson, M.C. and J. Guthrie, *Ongoing research to identify environmental risk factors in breast carcinoma.* Cancer, 2000. **88**(5 Suppl): p. 1224-9.
44. Hall, J.M., et al., *Linkage of early-onset familial breast cancer to chromosome 17q21.* Science, 1990. **250**(4988): p. 1684-9.
45. Wooster, R., et al., *Identification of the breast cancer susceptibility gene BRCA2.* Nature, 1995. **378**(6559): p. 789-92.
46. Ford, D., et al., *Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium.* Am J Hum Genet, 1998. **62**(3): p. 676-89.
47. Easton, D.F., D. Ford, and D.T. Bishop, *Breast and ovarian cancer incidence in BRCA1-mutation carriers. Breast Cancer Linkage Consortium.* Am J Hum Genet, 1995. **56**(1): p. 265-71.
48. Goldberg, J.I. and P.I. Borgen, *Breast cancer susceptibility testing: past, present and future.* Expert Rev Anticancer Ther, 2006. **6**(8): p. 1205-14.
49. Hakem, R., J.L. de la Pompa, and T.W. Mak, *Developmental studies of Brca1 and Brca2 knock-out mice.* J Mammary Gland Biol Neoplasia, 1998. **3**(4): p. 431-45.
50. Moynahan, M.E., *The cancer connection: BRCA1 and BRCA2 tumor suppression in mice and humans.* Oncogene, 2002. **21**(58): p. 8994-9007.
51. Sorlie, T., et al., *Repeated observation of breast tumor subtypes in independent gene expression data sets.* Proc Natl Acad Sci U S A, 2003. **100**(14): p. 8418-23.
52. Meijers-Heijboer, H., et al., *Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations.* Nat Genet, 2002. **31**(1): p. 55-9.
53. Jekimovs, C.R., et al., *Low frequency of CHEK2 1100delC allele in Australian multiple-case breast cancer families: functional analysis in heterozygous individuals.* Br J Cancer, 2005. **92**(4): p. 784-90.
54. Kilpivaara, O., et al., *CHEK2 variant I157T may be associated with increased breast cancer risk.* Int J Cancer, 2004. **111**(4): p. 543-7.
55. Friedrichsen, D.M., et al., *Frequency of CHEK2 mutations in a population based, case-control study of breast cancer in young women.* Breast Cancer Res, 2004. **6**(6): p. R629-35.
56. Osorio, A., et al., *The breast cancer low-penetrance allele 1100delC in the CHEK2 gene is not present in Spanish familial breast cancer population.* Int J Cancer, 2004. **108**(1): p. 54-6.

57. Kumar, S., et al., *p53 in breast cancer: mutation and countermeasures.* Front Biosci, 2007. **12**: p. 4168-78.
58. Kenemans, P., R.A. Verstraeten, and R.H. Verheijen, *Oncogenic pathways in hereditary and sporadic breast cancer.* Maturitas, 2004. **49**(1): p. 34-43.
59. Raunio, H., et al., *Diagnosis of polymorphisms in carcinogen-activating and inactivating enzymes and cancer susceptibility--a review.* Gene, 1995. **159**(1): p. 113-21.
60. FitzGerald, M.G., et al., *Germline mutations in PTEN are an infrequent cause of genetic predisposition to breast cancer.* Oncogene, 1998. **17**(6): p. 727-31.
61. Miyoshi, Y., et al., *Association of genetic polymorphisms in CYP19 and CYP1A1 with the oestrogen receptor-positive breast cancer risk.* Eur J Cancer, 2003. **39**(17): p. 2531-7.
62. Sillanpaa, P., et al., *CYP1A1 and CYP1B1 genetic polymorphisms, smoking and breast cancer risk in a Finnish Caucasian population.* Breast Cancer Res Treat, 2006.
63. Miyoshi, Y., et al., *Breast cancer risk associated with CYP1A1 genetic polymorphisms in Japanese women.* Breast J, 2002. **8**(4): p. 209-15.
64. Huang, C.S., *Genetic polymorphism of estrogen- and carcinogen-metabolizing genes in association with breast cancer risk in Taiwanese women.* Gan To Kagaku Ryoho, 2002. **29 Suppl 1**: p. 104-5.
65. Wang, S.E., et al., *HER2 kinase domain mutation results in constitutive phosphorylation and activation of HER2 and EGFR and resistance to EGFR tyrosine kinase inhibitors.* Cancer Cell, 2006. **10**(1): p. 25-38.
66. Casalini, P., et al., *Role of HER receptors family in development and differentiation.* J Cell Physiol, 2004. **200**(3): p. 343-50.
67. Welch, D.R., P.S. Steeg, and C.W. Rinker-Schaeffer, *Molecular biology of breast cancer metastasis. Genetic regulation of human breast carcinoma metastasis.* Breast Cancer Res, 2000. **2**(6): p. 408-16.
68. Bours, V., et al., *Nuclear factor-kappa B, cancer, and apoptosis.* Biochem Pharmacol, 2000. **60**(8): p. 1085-9.
69. Song, R.X. and R.J. Santen, *Apoptotic action of estrogen.* Apoptosis, 2003. **8**(1): p. 55-60.
70. Green, A.M. and N.D. Steinmetz, *Monitoring apoptosis in real time.* Cancer J, 2002. **8**(2): p. 82-92.
71. Wasinger, V.C., et al., *Progress with gene-product mapping of the Mollicutes: Mycoplasma genitalium.* Electrophoresis, 1995. **16**(7): p. 1090-4.
72. James, P., *Protein identification in the post-genome era: the rapid rise of proteomics.* Q Rev Biophys, 1997. **30**(4): p. 279-331.
73. Yip, T.T. and T.W. Hutchens, *Mapping and sequence-specific identification of phosphopeptides in unfractionated protein digest mixtures by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry.* FEBS Lett, 1992. **308**(2): p. 149-53.
74. Bailey, H.S., *Mass spectrometry as a clinical research tool.* SD J Med Pharm, 1955. **8**(5): p. 168-71.
75. Gelfi, C. and P.G. Righetti, *Preparative isoelectric focusing in immobilized pH gradients. II. A case report.* J Biochem Biophys Methods, 1983. **8**(2): p. 157-72.
76. Corbett, J.M., et al., *Positional reproducibility of protein spots in two-dimensional polyacrylamide gel electrophoresis using immobilised pH gradient*

isoelectric focusing in the first dimension: an interlaboratory comparison. Electrophoresis, 1994. **15**(8-9): p. 1205-11.
77. G. raorg, A., W. Postel, and R. Westermeier, *[A simple method for preparing pre-cast polyacrylamid gel plates for isoelectric focusing].* Z Lebensm Unters Forsch, 1977. **164**(3): p. 160-2.
78. Osin, P., et al., *Experimental pathology and breast cancer genetics: new technologies.* Recent Results Cancer Res, 1998. **152**: p. 35-48.
79. Hondermarck, H., et al., *Proteomics of breast cancer for marker discovery and signal pathway profiling.* Proteomics, 2001. **1**(10): p. 1216-32.
80. Wulfkuhle, J.D., et al., *New approaches to proteomic analysis of breast cancer.* Proteomics, 2001. **1**(10): p. 1205-15.
81. Kobayashi, R., *A proteomics approach to find a new breast cancer-specific antigenic marker.* Clin Cancer Res, 2001. **7**(11): p. 3325-7.
82. Le Naour, F., et al., *Proteomics-based identification of RS/DJ-1 as a novel circulating tumor antigen in breast cancer.* Clin Cancer Res, 2001. **7**(11): p. 3328-35.
83. Pucci-Minafra, I., et al., *Proteomic patterns of cultured breast cancer cells and epithelial mammary cells.* Ann N Y Acad Sci, 2002. **963**: p. 122-39.
84. Deng, S.S., et al., *Comparative proteome analysis of breast cancer and adjacent normal breast tissues in human.* Genomics Proteomics Bioinformatics, 2006. **4**(3): p. 165-72.
85. Luo, Y., et al., *Comparative proteome analysis of breast cancer and normal breast.* Mol Biotechnol, 2005. **29**(3): p. 233-44.
86. Somiari, R.I., et al., *High-throughput proteomic analysis of human infiltrating ductal carcinoma of the breast.* Proteomics, 2003. **3**(10): p. 1863-73.
87. Wulfkuhle, J.D., et al., *Proteomics of human breast ductal carcinoma in situ.* Cancer Res, 2002. **62**(22): p. 6740-9.
88. Kennedy, S., *Proteomic profiling from human samples: the body fluid alternative.* Toxicol Lett, 2001. **120**(1-3): p. 379-84.
89. Li, J., et al., *Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer.* Clin Chem, 2002. **48**(8): p. 1296-304.
90. Rui, Z., et al., *Use of serological proteomic methods to find biomarkers associated with breast cancer.* Proteomics, 2003. **3**(4): p. 433-9.
91. Alexander, H., et al., *Proteomic analysis to identify breast cancer biomarkers in nipple aspirate fluid.* Clin Cancer Res, 2004. **10**(22): p. 7500-10.
92. Hudelist, G., et al., *Use of high-throughput protein array for profiling of differentially expressed proteins in normal and malignant breast tissue.* Breast Cancer Res Treat, 2004. **86**(3): p. 281-91.
93. Pawlik, T.M., et al., *Significant differences in nipple aspirate fluid protein expression between healthy women and those with breast cancer demonstrated by time-of-flight mass spectrometry.* Breast Cancer Res Treat, 2005. **89**(2): p. 149-57.
94. Hudelist, G., et al., *Proteomic analysis in human breast cancer: identification of a characteristic protein expression profile of malignant breast epithelium.* Proteomics, 2006. **6**(6): p. 1989-2002.
95. Pawlik, T.M., et al., *Proteomic analysis of nipple aspirate fluid from women with early-stage breast cancer using isotope-coded affinity tags and tandem*

mass spectrometry reveals differential expression of vitamin D binding protein. BMC Cancer, 2006. **6**: p. 68.

96. Chen, S.T., et al., *Proteomics reveals protein profile changes in doxorubicin--treated MCF-7 human breast cancer cells.* Cancer Lett, 2002. **181**(1): p. 95-107.
97. Huber, M., et al., *Comparison of proteomic and genomic analyses of the human breast cancer cell line T47D and the antiestrogen-resistant derivative T47D-r.* Mol Cell Proteomics, 2004. **3**(1): p. 43-55.
98. Sandhu, C., et al., *Global protein shotgun expression profiling of proliferating mcf-7 breast cancer cells.* J Proteome Res, 2005. **4**(3): p. 674-89.
99. Zhang, D., et al., *Proteomic study reveals that proteins involved in metabolic and detoxification pathways are highly expressed in HER-2/neu-positive breast cancer.* Mol Cell Proteomics, 2005. **4**(11): p. 1686-96.
100. Laronga, C., et al., *Association of the cyclin-dependent kinases and 14-3-3 sigma negatively regulates cell cycle progression.* J Biol Chem, 2000. **275**(30): p. 23106-12.
101. Moreira, J.M., et al., *Down-regulation of the tumor suppressor protein 14-3-3sigma is a sporadic event in cancer of the breast.* Mol Cell Proteomics, 2005. **4**(4): p. 555-69.
102. Liu, Y., et al., *Identification of 14-3-3sigma as a contributor to drug resistance in human breast cancer cells using functional proteomic analysis.* Cancer Res, 2006. **66**(6): p. 3248-55.
103. Brown, K.J. and C. Fenselau, *Investigation of doxorubicin resistance in MCF-7 breast cancer cells using shot-gun comparative proteomics with proteolytic 18O labeling.* J Proteome Res, 2004. **3**(3): p. 455-62.
104. Wang, T., et al., *The role of peroxiredoxin II in radiation-resistant MCF-7 breast cancer cells.* Cancer Res, 2005. **65**(22): p. 10338-46.
105. Umar, A., et al., *NanoLC-FT-ICR MS improves proteome coverage attainable for approximately 3000 laser-microdissected breast carcinoma cells.* Proteomics, 2007. **7**(2): p. 323-9.
106. 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.
107. Watkins, S., et al., *A donor splice mutation and a single-base deletion produce two carboxyl-terminal variants of human serum albumin.* Proc Natl Acad Sci U S A, 1991. **88**(14): p. 5959-63.
108. Madison, J., et al., *Genetic variants of serum albumin in Americans and Japanese.* Proc Natl Acad Sci U S A, 1991. **88**(21): p. 9853-7.
109. Germain, P., et al., *Use of confocal microscopy to localize the SHBG interaction with human breast cancer cell lines--a comparison with serum albumin interaction.* Cell Mol Biol (Noisy-le-grand), 1997. **43**(4): p. 501-8.
110. Germain, P., et al., *Internalization and biological effects of serum albumin in the breast cancer MCF-7 and MDA-MB 231 cells.* Cell Mol Biol (Noisy-le-grand), 1995. **41**(8): p. 1119-29.
111. Lis, C.G., et al., *Is serum albumin an independent predictor of survival in patients with breast cancer?* JPEN J Parenter Enteral Nutr, 2003. **27**(1): p. 10-5.
112. Kukita, A., et al., *Osteoinductive factor inhibits formation of human osteoclast-like cells.* Proc Natl Acad Sci U S A, 1990. **87**(8): p. 3023-6.

113. Tasheva, E.S. and G.W. Conrad, *Interferon-gamma regulation of the human mimecan promoter.* Mol Vis, 2003. **9**: p. 277-87.
114. Fernandez, B., et al., *Osteoglycin expression and localization in rabbit tissues and atherosclerotic plaques.* Mol Cell Biochem, 2003. **246**(1-2): p. 3-11.
115. Shanahan, C.M., et al., *Identification of osteoglycin as a component of the vascular matrix. Differential expression by vascular smooth muscle cells during neointima formation and in atherosclerotic plaques.* Arterioscler Thromb Vasc Biol, 1997. **17**(11): p. 2437-47.
116. Hu, S.M., et al., *The mimecan gene expressed in human pituitary and regulated by pituitary transcription factor-1 as a marker for diagnosing pituitary tumors.* J Clin Endocrinol Metab, 2005. **90**(12): p. 6657-64.
117. Lee, J.Y., et al., *Analysis of gene expression profiles of gastric normal and cancer tissues by SAGE.* Genomics, 2003. **82**(1): p. 78-85.
118. Tasheva, E.S., et al., *Transcriptional activation of bovine mimecan by p53 through an intronic DNA-binding site.* Biochim Biophys Acta, 2001. **1517**(3): p. 333-8.
119. Leygue, E., et al., *Lumican and decorin are differentially expressed in human breast carcinoma.* J Pathol, 2000. **192**(3): p. 313-20.
120. Naito, Z., *Role of the small leucine-rich proteoglycan (SLRP) family in pathological lesions and cancer cell growth.* J Nippon Med Sch, 2005. **72**(3): p. 137-45.
121. Troup, S., et al., *Reduced expression of the small leucine-rich proteoglycans, lumican, and decorin is associated with poor outcome in node-negative invasive breast cancer.* Clin Cancer Res, 2003. **9**(1): p. 207-14.
122. Vuillermoz, B., et al., *Influence of aging on glycosaminoglycans and small leucine-rich proteoglycans production by skin fibroblasts.* Mol Cell Biochem, 2005. **277**(1-2): p. 63-72.
123. Vuillermoz, B., et al., *The small leucine-rich proteoglycan lumican inhibits melanoma progression.* Exp Cell Res, 2004. **296**(2): p. 294-306.
124. Hagedorn, M., et al., *Accessing key steps of human tumor progression in vivo by using an avian embryo model.* Proc Natl Acad Sci U S A, 2005. **102**(5): p. 1643-8.
125. Segawa, T., et al., *Androgen-induced expression of endoplasmic reticulum (ER) stress response genes in prostate cancer cells.* Oncogene, 2002. **21**(57): p. 8749-58.
126. Maniratanachote, R., et al., *Chaperone proteins involved in troglitazone-induced toxicity in human hepatoma cell lines.* Toxicol Sci, 2005. **83**(2): p. 293-302.
127. Hayano, T. and M. Kikuchi, *Molecular cloning of the cDNA encoding a novel protein disulfide isomerase-related protein (PDIR).* FEBS Lett, 1995. **372**(2-3): p. 210-4.
128. de Chateau, M., E. Holst, and L. Bjorck, *Protein PAB, an albumin-binding bacterial surface protein promoting growth and virulence.* J Biol Chem, 1996. **271**(43): p. 26609-15.
129. Lejon, S., et al., *Crystal structure and biological implications of a bacterial albumin binding module in complex with human serum albumin.* J Biol Chem, 2004. **279**(41): p. 42924-8.
130. Riggio, M.P. and A. Lennon, *Specific PCR detection of Peptostreptococcus magnus.* J Med Microbiol, 2003. **52**(Pt 4): p. 309-13.

131. Nickerson, M.L., et al., *Mutations in a novel gene lead to kidney tumors, lung wall defects, and benign tumors of the hair follicle in patients with the Birt-Hogg-Dube syndrome.* Cancer Cell, 2002. **2**(2): p. 157-64.
132. Franzen, B., et al., *Analysis of polypeptide expression in benign and malignant human breast lesions.* Electrophoresis, 1997. **18**(3-4): p. 582-7.
133. Garrido, C., et al., *Heat shock proteins: endogenous modulators of apoptotic cell death.* Biochem Biophys Res Commun, 2001. **286**(3): p. 433-42.
134. Maguire, M., et al., *Comparative cell signalling activity of ultrapure recombinant chaperonin 60 proteins from prokaryotes and eukaryotes.* Immunology, 2005. **115**(2): p. 231-238.
135. Cheng, T.C., et al., *Breast cancer risk associated with genotype polymorphism of the catechol estrogen-metabolizing genes: a multigenic study on cancer susceptibility.* Int J Cancer, 2005. **113**(3): p. 345-53.
136. Bock, K.W. and C. Kohle, *UDP-glucuronosyltransferase 1A6: structural, functional, and regulatory aspects.* Methods Enzymol, 2005. **400**: p. 57-75.
137. Miners, J.O., et al., *Characterization of paracetamol UDP-glucuronosyltransferase activity in human liver microsomes.* Biochem Pharmacol, 1990. **40**(3): p. 595-600.
138. Iida, K., et al., *Nrf2 is essential for the chemopreventive efficacy of oltipraz against urinary bladder carcinogenesis.* Cancer Res, 2004. **64**(18): p. 6424-31.
139. Bock, K.W., et al., *Persistently increased expression of a 3-methylcholanthrene-inducible phenol uridine diphosphate-glucuronosyltransferase in rat hepatocyte nodules and hepatocellular carcinomas.* Cancer Res, 1990. **50**(12): p. 3569-73.
140. Fischer, G., D. Ulrich, and K.W. Bock, *Effects of N-nitrosomorpholine and phenobarbital on UDP-glucuronyltransferase in putative preneoplastic foci of rat liver.* Carcinogenesis, 1985. **6**(4): p. 605-9.
141. Jin, J., et al., *A mammalian chromatin remodeling complex with similarities to the yeast INO80 complex.* J Biol Chem, 2005. **280**(50): p. 41207-12.
142. Kotaja, N., et al., *Abnormal sperm in mice with targeted deletion of the act (activator of cAMP-responsive element modulator in testis) gene.* Proc Natl Acad Sci U S A, 2004. **101**(29): p. 10620-5.
143. Morgan, M.J. and S.A. Whawell, *The structure of the human LIM protein ACT gene and its expression in tumor cell lines.* Biochem Biophys Res Commun, 2000. **273**(2): p. 776-83.
144. Mistry, A.C., et al., *FHL5, a novel actin-binding protein, is highly expressed in eel gill pillar cells and responds to wall tension.* Am J Physiol Regul Integr Comp Physiol, 2004. **287**(5): p. R1141-54.
145. Park, T., Z.P. Chen, and J. Leavitt, *Activation of the leukocyte plastin gene occurs in most human cancer cells.* Cancer Res, 1994. **54**(7): p. 1775-81.
146. Foran, E., et al., *The leukocyte protein L-plastin induces proliferation, invasion and loss of E-cadherin expression in colon cancer cells.* Int J Cancer, 2006. **118**(8): p. 2098-104.
147. Lapillonne, A., et al., *Expression patterns of L-plastin isoform in normal and carcinomatous breast tissues.* Anticancer Res, 2000. **20**(5A): p. 3177-82.
148. Sidoni, A., et al., *Breast cancer in young women: clinicopathological features and biological specificity.* Breast, 2003. **12**(4): p. 247-50.

149. Swoboda, B., et al., *Increased content of type-VI collagen epitopes in human osteoarthritic cartilage: quantitation by inhibition ELISA.* J Orthop Res, 1998. **16**(1): p. 96-9.
150. Iyengar, P., et al., *Adipocyte-derived collagen VI affects early mammary tumor progression in vivo, demonstrating a critical interaction in the tumor/stroma microenvironment.* J Clin Invest, 2005. **115**(5): p. 1163-76.
151. Wegrowski, J., G. Bellon, and A. Randoux, *Human type VI collagen: isolation and characterization of alpha 1 and alpha 2 chains by two-dimensional preparative gel electrophoresis.* Biochim Biophys Acta, 1990. **1039**(2): p. 189-96.
152. Nishida, E., S. Maekawa, and H. Sakai, *Cofilin, a protein in porcine brain that binds to actin filaments and inhibits their interactions with myosin and tropomyosin.* Biochemistry, 1984. **23**(22): p. 5307-13.
153. Ohta, Y., et al., *Dephosphorylation of cofilin accompanies heat shock-induced nuclear accumulation of cofilin.* J Biol Chem, 1989. **264**(27): p. 16143-8.
154. Abe, H., R. Nagaoka, and T. Obinata, *Cytoplasmic localization and nuclear transport of cofilin in cultured myotubes.* Exp Cell Res, 1993. **206**(1): p. 1-10.
155. Dowling, P., et al., *Proteomic analysis of isolated membrane fractions from superinvasive cancer cells.* Biochim Biophys Acta, 2007. **1774**(1): p. 93-101.
156. Hitchcock-Degregori, S.E., *Chemotaxis: Cofilin in the driver's seat.* Curr Biol, 2006. **16**(24): p. R1030-2.
157. Odreman, F., et al., *Proteomic studies on low- and high-grade human brain astrocytomas.* J Proteome Res, 2005. **4**(3): p. 698-708.
158. Wang, W., et al., *The activity status of cofilin is directly related to invasion, intravasation, and metastasis of mammary tumors.* J Cell Biol, 2006. **173**(3): p. 395-404.
159. Moyano, J.V., et al., *alphaB-Crystallin is a novel oncoprotein that predicts poor clinical outcome in breast cancer.* J Clin Invest, 2006. **116**(1): p. 261-270.
160. Gruvberger-Saal, S.K. and R. Parsons, *Is the small heat shock protein alphaB-crystallin an oncogene?* J Clin Invest, 2006. **116**(1): p. 30-32.
161. Chelouche-Lev, D., et al., *alphaB-crystallin as a marker of lymph node involvement in breast carcinoma.* Cancer, 2004. **100**(12): p. 2543-8.
162. Cicek, M., et al., *Identification of metastasis-associated proteins through protein analysis of metastatic MDA-MB-435 and metastasis-suppressed BRMS1 transfected-MDA-MB-435 cells.* Clin Exp Metastasis, 2004. **21**(2): p. 149-57.
163. Chin, D., et al., *Alpha B-crystallin, a new independent marker for poor prognosis in head and neck cancer.* Laryngoscope, 2005. **115**(7): p. 1239-42.
164. Kamradt, M.C., F. Chen, and V.L. Cryns, *The small heat shock protein alpha B-crystallin negatively regulates cytochrome c- and caspase-8-dependent activation of caspase-3 by inhibiting its autoproteolytic maturation.* J Biol Chem, 2001. **276**(19): p. 16059-63.
165. Li, D.W., et al., *Calcium-activated RAF/MEK/ERK signaling pathway mediates p53-dependent apoptosis and is abrogated by alpha B-crystallin through inhibition of RAS activation.* Mol Biol Cell, 2005. **16**(9): p. 4437-53.
166. Djabali, K., et al., *alphaB-crystallin interacts with cytoplasmic intermediate filament bundles during mitosis.* Exp Cell Res, 1999. **253**(2): p. 649-62.
167. Ishiguro, Y., et al., *Chemotherapy-induced expression of alpha B-crystallin in neuroblastoma.* Med Pediatr Oncol, 1997. **29**(1): p. 11-5.

168. Besada, V., et al., *Proteomics of xenografted human breast cancer indicates novel targets related to tamoxifen resistance.* Proteomics, 2005.
169. Chen, P.F., et al., *Amplification of immunoglobulin transcripts by the non-palindromic adaptor polymerase chain reaction (NPA-PCR). Nucleotide sequence analysis of two human monoclonal antibodies recognizing two cell surface antigens expressed in ovarian, cervix, breast, colon and other carcinomas.* Hum Antibodies Hybridomas, 1994. **5**(3-4): p. 131-42.
170. Anderson, M.L., et al., *Cloning and sequence analysis of an Ig lambda light chain mRNA expressed in the Burkitt's lymphoma cell line EB4.* Nucleic Acids Res, 1985. **13**(8): p. 2931-41.
171. Wang, A.C. and H.H. Fudenberg, *Chemical studies on a monoclonal immunoglobulin from a patient with carcinoma of the colon.* Immunol Commun, 1975. **4**(5): p. 483-97.
172. Solomon, A., et al., *Immunocytochemical detection of kappa and lambda light chain V region subgroups in human B-cell malignancies.* Am J Pathol, 1990. **137**(4): p. 855-62.
173. Kotlan, B., et al., *Immunoglobulin variable regions usage by B-lymphocytes infiltrating a human breast medullary carcinoma.* Immunol Lett, 1999. **65**(3): p. 143-51.
174. Ebeling, S.B., M.E. Schutte, and T. Logtenberg, *Molecular analysis of VH and VL regions expressed in IgG-bearing chronic lymphocytic leukemia (CLL): further evidence that CLL is a heterogeneous group of tumors.* Blood, 1993. **82**(5): p. 1626-31.
175. Kotlan, B., et al., *Immunoglobulin repertoire of B lymphocytes infiltrating breast medullary carcinoma.* Hum Antibodies, 2003. **12**(4): p. 113-21.
176. Wang, Y., et al., *Focused antibody response in plasma cell-infiltrated non-medullary (NOS) breast cancers.* Breast Cancer Res Treat, 2007.
177. Kawata, A., et al., *Tumor-infiltrating lymphocytes and prognosis of hepatocellular carcinoma.* Jpn J Clin Oncol, 1992. **22**(4): p. 256-63.
178. Dong, H.P., et al., *NK- and B-cell infiltration correlates with worse outcome in metastatic ovarian carcinoma.* Am J Clin Pathol, 2006. **125**(3): p. 451-8.
179. Sellin, L., et al., *Acute renal failure due to a malignant lymphoma infiltration uncovered by renal biopsy.* Nephrol Dial Transplant, 2004. **19**(10): p. 2657-60.
180. Tokar, M., et al., *Rituximab in a patient with acute renal failure due to B-cell lymphomatous infiltration of the kidneys.* Leuk Lymphoma, 2004. **45**(4): p. 819-20.
181. Vij, N., et al., *Lumican suppresses cell proliferation and aids Fas-Fas ligand mediated apoptosis: implications in the cornea.* Exp Eye Res, 2004. **78**(5): p. 957-71.
182. Huang, H.L., et al., *Biomarker discovery in breast cancer serum using 2-D differential gel electrophoresis/ MALDI-TOF/TOF and data validation by routine clinical assays.* Electrophoresis, 2006. **27**(8): p. 1641-50.
183. Yu, B., et al., *Comparative study of proteome between primary cancer and hepatic metastatic tumor in colorectal cancer.* World J Gastroenterol, 2004. **10**(18): p. 2652-6.
184. Davidsson, P., et al., *Proteome analysis of cerebrospinal fluid proteins in Alzheimer patients.* Neuroreport, 2002. **13**(5): p. 611-5.

185. Gilad, L.A., et al., *Regulation of vitamin D receptor expression via estrogen-induced activation of the ERK 1/2 signaling pathway in colon and breast cancer cells.* J Endocrinol, 2005. **185**(3): p. 577-92.
186. Isidoro, A., et al., *Breast carcinomas fulfill the Warburg hypothesis and provide metabolic markers of cancer prognosis.* Carcinogenesis, 2005.
187. Hilf, R., et al., *Enzymes, nucleic acids, and lipids in human breast cancer and normal breast tissue.* Cancer Res, 1970. **30**(6): p. 1874-82.
188. Hoopmann, M., et al., *Tumor M2 pyruvate kinase--determination in breast cancer patients receiving trastuzumab therapy.* Cancer Lett, 2002. **187**(1-2): p. 223-8.
189. Keen, J.C., et al., *Epigenetic regulation of protein phosphatase 2A (PP2A), lymphotactin (XCL1) and estrogen receptor alpha (ER) expression in human breast cancer cells.* Cancer Biol Ther, 2004. **3**(12): p. 1304-12.
190. Keen, J.C., et al., *Protein phosphatase 2A regulates estrogen receptor alpha (ER) expression through modulation of ER mRNA stability.* J Biol Chem, 2005. **280**(33): p. 29519-24.
191. Li, H., et al., *Protein phosphatase 2A inhibits nuclear telomerase activity in human breast cancer cells.* J Biol Chem, 1997. **272**(27): p. 16729-32.
192. Rossini, G.P., C. Pinna, and R. Viviani, *Inhibitors of phosphoprotein phosphatases 1 and 2A cause activation of a 53 kDa protein kinase accompanying the apoptotic response of breast cancer cells.* FEBS Lett, 1997. **410**(2-3): p. 347-50.
193. Ruediger, R., H.T. Pham, and G. Walter, *Disruption of protein phosphatase 2A subunit interaction in human cancers with mutations in the A alpha subunit gene.* Oncogene, 2001. **20**(1): p. 10-5.
194. http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?db=protein&val=51491284. [cited.
195. Roberts, K., et al., *Proteomic analysis of selected prognostic factors of breast cancer.* Proteomics, 2004. **4**(3): p. 784-92.
196. Chughtai, N., et al., *Prolactin induces SHP-2 association with Stat5, nuclear translocation, and binding to the beta-casein gene promoter in mammary cells.* J Biol Chem, 2002. **277**(34): p. 31107-14.
197. Chida, D., et al., *Transcriptional regulation of the beta-casein gene by cytokines: cross-talk between STAT5 and other signaling molecules.* Mol Endocrinol, 1998. **12**(11): p. 1792-806.
198. Hendrick, J.C. and P. Franchimont, *Radio-immunoassay of casein in the serum of normal subjects and of patients with various malignancies.* Eur J Cancer, 1974. **10**(11): p. 725-30.
199. Schedin, P., T. Mitrenga, and M. Kaeck, *Estrous cycle regulation of mammary epithelial cell proliferation, differentiation, and death in the Sprague-Dawley rat: a model for investigating the role of estrous cycling in mammary carcinogenesis.* J Mammary Gland Biol Neoplasia, 2000. **5**(2): p. 211-25.
200. Cohen, C., et al., *Tumor-associated antigens in breast carcinomas. Prognostic significance.* Cancer, 1987. **60**(6): p. 1294-8.
201. Earl, H.M., et al., *Immunohistochemical study of beta- and kappa-casein in the human breast and breast carcinomas, using monoclonal antibodies.* Cancer Res, 1989. **49**(21): p. 6070-6.

202. Bartkova, J., et al., *Lack of beta-casein production by human breast tumours revealed by monoclonal antibodies.* Eur J Cancer Clin Oncol, 1987. **23**(10): p. 1557-63.
203. Frost, R.A., G.J. Nystrom, and C.H. Lang, *Regulation of IGF-I mRNA and signal transducers and activators of transcription-3 and -5 (Stat-3 and -5) by GH in C2C12 myoblasts.* Endocrinology, 2002. **143**(2): p. 492-503.
204. Lee, C., et al., *Characterization of the Stat5 protease.* J Biol Chem, 1999. **274**(38): p. 26767-75.
205. Yamashita, H., et al., *Role of serine phosphorylation of Stat5a in prolactin-stimulated beta-casein gene expression.* Mol Cell Endocrinol, 2001. **183**(1-2): p. 151-63.
206. Velcheti, V. and R. Govindan, *Insulin-like growth factor and lung cancer.* J Thorac Oncol, 2006. **1**(7): p. 607-10.
207. Bromberg, J., *Signal transducers and activators of transcription as regulators of growth, apoptosis and breast development.* Breast Cancer Res, 2000. **2**(2): p. 86-90.
208. Hwa, V., et al., *Transcriptional regulation of insulin-like growth factor-I (IGF-I) by interferon-gamma (IFN-gamma) requires STAT-5b.* J Biol Chem, 2003.
209. Neilson, L.M., et al., *Coactivation of Jak1 Positively Modulates Prolactin-Jak2 Signaling in Breast Cancer: Recruitment of ERK and Stat3 and Enhancement of Akt and Stat5a/b Pathways.* Mol Endocrinol, 2007.
210. Nyga, R., et al., *Activated STAT5 proteins induce activation of the PI 3-kinase/Akt and Ras/MAPK pathways via the Gab2 scaffolding adapter.* Biochem J, 2005. **390**(Pt 1): p. 359-66.

1 Appendix A

Sample ID	Age	Histology	TNM stage	1.0.1		Receptor status
				Oest	Prog	*c-erbB2*
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)	++	+	+++
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

Appendix A1: Australian samples included in this study (ACB).

Sample No.		Age at collection	Pathology Findings	Stage
T1	64	64	IDC, PR, H2N	
T1a	64	64		
T2	36	36	IDC,ER,PR	4
T3	38	38	IDC,ER,PR	2B
T4	43	43	IC, H2N	2B
T4a	43	43		
T5	41	41	IDC, ER	3A
T6	30	30	MuC,ER,PR,H2N	3A
T7	51	51	IDC,ER,PR,H2N	3A
T8	47	47	IDC,H2N	2A
T9	49	05-01-1952	IDC,ER,PR,H2N	2A
T10	46	01-05-1955	IDC,ER,PR,H2N	2B
T11	50	25-05-1951	IDC, ER, PR	3A
T12	50	09-11-1951	IC, ER,PR,H2N	2A
T13	72	05-03-1929	IDC, H2N	2B
T14	51	02-02-1950	IDC, PR	1
T15	69	28-02-1932	IDC,ER,PR,H2N	3B
T16	25	12-12-1976	IDC,ER,PR	2B
T17	40	01-06-1961	IDC,ER,PR,H2N	1
T18	40	15-10-1961	IDC,ER,PR,H2N	2A
T19	53	48-11-1948	IDC,ER,PR,H2N	2A
T20	61	18-07-1940	IDC,ER,PR,H2N	3A
T21	46	30-10-1955	IDC,ER,PR,H2N	2A
T22	49	04-07-1952	IDC, H2N	2A
T23	48	30-12-1953	IDC, H2N	3
T24	76	16-01-1925	IDC	2
T25	68	03-01-1933	MC, H2N	2A
T26	26	01-05-1975	IDC,ER,PR,H2N	2A
T27	70	08-01-1931	IDC, H2N	2A
T28	34	05-08-1967	FC	

Appendix Aa: Taiwanese samples included in this study (TCB).

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

Appendix A3: Greek cancer samples included in this study (GCB).

Sample ID	Age	Any History of Breast cancer
1	32	Previous radiotherapy for a large tumour
2	80	
3	87	Concurrent tumour in ipsilateral breast
4	69	Benign breast disease
5	87	-
6	57	-
7	39	Breast had previous breast cancer
8	24	Previous lumpectomy
9	32	Breast had previous breast cancer -
10	44	Two tumours previously
11	71	-
12	43	-

Appendix A4: Healthy samples that were included in this study (NB).

2 Appendix B

Casein Ab-1 (Clone F20.14)

Mouse Monoclonal Antibody

Cat. #MS-935-P0, -P1, or -P (0.1ml, 0.5ml, or 1.0ml at 200µg/ml) (Purified Ab with BSA and Azide)
Cat. #MS-935-P1ABX or -PABX (0.1ml or 0.2ml at 1.0mg/ml) (Purified Ab without BSA and Azide)

Description: Casein is synthesized by mammary epithelium only in the presence of lactogenic hormone combination of insulin, hydrocortisone, and prolactin. Enhancement of the rate of secretion of newly-synthesised caseins occurs when mammary explants are challenged in vitro with agents that activate protein kinase A (PKA). The PKA-sensitive step(s) in casein secretion is early in the exocytosis pathway but inhibition of PKA does not impair casein maturation. PKA is located on membranes of vesicles situated in the Golgi region. Casein positive cells are present in all breast carcinomas with a variable distribution and degree of staining. Casein serves as an in vitro marker of terminally differentiated breast luminal epithelial cells.

Mol. Wt. of Antigen: 28kDa

Epitope: Not determined

Species Reactivity: Human. Others-not tested

Clone Designation: F20.14

Ig Isotype: IgG_1

Immunogen: Purified casein and alpha-lactalbumin.

Applications and Suggested Dilutions:

- Western Blotting (Not verified)
- Immunohistology (Acetone-fixed frozen only)

The optimal dilution for a specific application should be determined by the investigator.

Positive Control: Breast cancer cells and breast carcinomas.

Cellular Localization: Cytoplasmic

Supplied As:

200µg/ml of antibody purified from ascites fluid by Protein G chromatography. Prepared in 10mM PBS, pH 7.4, with 0.2% BSA and 0.09% sodium azide. Also available without BSA and azide at 1mg/ml.

Storage and Stability:

Ab with sodium azide is stable for 24 months when stored at 2-8°C. Antibody WITHOUT sodium azide is stable for 36 months when stored at below 0°C.

Key References:

1. Burchell *et. al.* Hybridoma, 1985; 4:341-350.

Limitations and Warranty:

Our products are intended FOR RESEARCH USE ONLY and are not approved for clinical diagnosis, drug use or therapeutic procedures. No products are to be construed as a recommendation for use in violation of any patents. We make no representations, warranties or assurances as to the accuracy or completeness of information provided on our data sheets and website. Our warranty is limited to the actual price paid for the product. NeoMarkers is not liable for any property damage, personal injury, time or effort or economic loss caused by our products.

Material Safety Data:

This product is not licensed or approved for administration to humans or to animals other than the experimental animals. Standard Laboratory Practices should be followed when handling this material. The chemical, physical, and toxicological properties of this material have not been thoroughly investigated. Appropriate measures should be taken to avoid skin and eye contact, inhalation, and ingestion. The material contains 0.09% sodium azide as a preservative. Although the quantity of azide is very small, appropriate care should be taken when handling this material as indicated above. The National Institute of Occupational Safety and Health has issued a bulletin citing the potential explosion hazard due to the reaction of sodium azide with copper, lead, brass, or solder in the plumbing systems. Sodium azide forms hydrazoic acid in acidic conditions and should be discarded in a large volume of running water to avoid deposits forming in metal drainage pipes.

For Research Use Only

47790 Westinghouse Dr., Fremont, CA 94539, USA • Tel.: (510) 991-2850 • Fax: (510) 991-2860
E-Mail: neomark@ix.netcom.com • Website: www.neomarkers.com

IGF-I Ab-1 (Clone M23)

Mouse Monoclonal Antibody

Cat. #MS-1508-P0, -P1, or -P (0.1ml, 0.5ml, or 1.0ml at 200µg/ml) (Purified Ab with BSA and Azide)
Cat. #MS-1508-P1ABX or -PABX (0.1ml or 0.2ml at 1.0mg/ml) (Purified Ab without BSA and Azide)
Cat. #MS-1508-R7 (7.0ml) (Ready-to-Use for Immunohistochemical Staining)
Cat. #MS-1508-PCS (5 Slides) (Positive Control for Histology)

Description: Insulin-like growth factor 1 (IGF-1) is involved in regulation of neuronal growth and development in central and peripheral nervous system. It is known to protect neurons against cell death induced by amyloidogenic derivatives, glucose or serum deprivation through pathways involving AKT kinase and transcription factor FKHRL1 phosphorylation. Activation of the insulin–like growth factor system has emerged as a key factor for tumor progression and resistance to apoptosis in many cancers like breast and thyroid cancers.

Mol. Wt. of Antigen: ~21kDa

Species Reactivity: Human, Mouse, and Rat. Others-not tested

Clone Designation: M23

Ig Isotype: IgG_1

Immunogen: Purified human IGF-1.

Applications and Suggested Dilutions:

- Immunohistology (formalin/paraffin)
  (Ab 10-20µg/ml for 60 min at RT)
- [Staining of formalin-fixed tissues REQUIRES boiling tissue sections in 10mM citrate buffer, pH 6.0, (*NEOMARKERS'* Cat. #AP-9003), for 10-20 min followed by cooling at RT for 20 min.]

The optimal dilution for a specific application should be determined by the investigator.

Positive Control: Pancreas

Cellular Localization: Cytoplasmic

Supplied As:

200µg/ml of antibody purified from tissue culture supernetant by Protein G chromatography. Prepared in 10mM PBS, pH 7.4, with 0.2% BSA and 0.09% sodium azide. Also available without BSA and azide at 1mg/ml.

or

Prediluted antibody which is ready-to-use for staining of formalin-fixed, paraffin-embedded tissues.

Storage and Stability:

Ab with sodium azide is stable for 24 months when stored at 2-8°C. Antibody WITHOUT sodium azide is stable for 36 months when stored at below 0°C.

Suggested References:

1. Rotwein p, et al. (1986) J. Biol. Chem 261: 4828-4832.
2. Sandberg-Nordqvist A C, et al. (1993) Cancer Res. 53: 2475-2478.
3. Zheng W H, et al. (2000) J Neural.Transm. Suppl 2000: 261-272.

Limitations and Warranty:

Our products are intended FOR RESEARCH USE ONLY and are not approved for clinical diagnosis, drug use or therapeutic procedures. No products are to be construed as a recommendation for use in violation of any patents. We make no representations, warranties or assurances as to the accuracy or completeness of information provided on our data sheets and website. Our warranty is limited to the actual price paid for the product. NeoMarkers is not liable for any property damage, personal injury, time or effort or economic loss caused by our products.

Material Safety Data:

This product is not licensed or approved for administration to humans or to animals other than the experimental animals. Standard Laboratory Practices should be followed when handling this material. The chemical, physical, and toxicological properties of this material have not been thoroughly investigated. Appropriate measures should be taken to avoid skin and eye contact, inhalation, and ingestion. The material contains 0.09% sodium azide as a preservative. Although the quantity of azide is very small, appropriate care should be taken when handling this material as indicated above. The National Institute of Occupational Safety and Health has issued a bulletin citing the potential explosion hazard due to the reaction of sodium azide with copper, lead, brass, or solder in the plumbing systems. Sodium azide forms hydrazoic acid in acidic conditions and should be discarded in a large volume of running water to avoid deposits forming in metal drainage pipes.

For Research Use Only

Manufactured by:
NEOMARKERS, INC.
47790 Westinghouse Dr.,
Fremont, CA 94539

Distributed by:
LAB VISION CORPORATION
Tel.: (800) 828-1628 • Fax. (510) 991-2826
labvision@labvision.com • www.labvision.com

Thermo SCIENTIFIC

DATA SHEET
Rev 092402E

Stat5a Ab-1

Rabbit Polyclonal Antibody

Cat. #RB-095-P0, -P1, or -P (0.1ml, 0.5ml, or 1.0ml at 1.0mg/ml) (Purified Ab with BSA and Azide)
Cat. #RB-095-P1ABX or -PABX (0.5ml or 1.0ml at 1.0mg/ml) (Purified Ab without BSA and Azide)
Cat. #RB-095-B0, -B1, or -B (0.1ml, 0.5ml, or 1.0ml at 1.0mg/ml) (Biotin-Labeled Ab with BSA and Azide)
Cat. #RB-095-R7 (7.0ml) (Ready-to-Use for Immunohistochemical Staining)
Cat. #RB-095-PCS (5 Slides) (Positive Control for Histology)
Cat. #RB-095-PCL (0.1ml) (Positive Control for Western Blot)
Cat. #PP-095 (1.0ml at 200µg/ml) (Blocking Peptide with BSA and Azide)

Description: Stat5a protein is a transcription factor activated by hormone and cytokines. Two highly related, but distinct Stat5 genes (Stat5a and Stat5b) were identified in mouse. The amino acid sequences of Stat5a and Stat5b show ~96% sequence similarity, and both proteins are co-expressed in most tissues of both virgin and lactating mice. However, differential accumulation of Stat5a and Stat5b mRNA has been reported for both muscle and mammary tissue. Stat5a is critically involved in a variety of physiological functions, including reproduction, lactation, immune function and somatic growth.

Mol. Wt. of Antigen: 95kDa

Epitope: C-terminal

Species Reactivity: Human, Mouse, and Rat. Others-not known.

Immunogen: A synthetic peptide corresponding to C-terminal sequence of human Stat5a.

Applications and Suggested Dilutions:

- Gel Supershift
  Immunoprecipitation (Not suitable)
- Western Blotting (5-10µg/ml for 2hrs at RT)
- Immunohistology (Formalin/paraffin)
  (Ab 2.5-5µg/ml for 30 min at RT)

* [Staining of formalin-fixed tissues REQUIRES boiling tissue sections in 10mM citrate buffer, pH 6.0, (*NEOMARKERS'* Cat. #AP-9003), for 10-20 min followed by cooling at RT for 20 min.]

The optimal dilution for a specific application should be determined by the investigator.

Positive Control: A431 cells. Tonsil.

Cellular Localization: Nuclear

Supplied As: Total IgG purified from rabbit antiserum by Protein A chromatography. Prepared at 1mg/ml in 10mM PBS, pH 7.4, with 0.2% BSA & 0.09% sodium azide. Also available without BSA and azide at 1mg/ ml, or Prediluted antibody which is ready-to-use for staining of formalin-fixed, paraffin-embedded tissues.

Storage and Stability:

Ab with sodium azide is stable for 24 months when stored at 2-8°C. Antibody WITHOUT sodium azide is stable for 36 months when stored at below 0°C.

Suggested References:

1. Frank DA. Mol Med 1999 Jul; 5(7):432-56.

Limitations and Warranty:

Material Safety Data:

This product is not licensed or approved for administration to humans or to animals other than the experimental ani mals. Standard Laboratory Practices should be followed when handling this material. The chemical, physical, and toxicological properties of this material have not been thoroughly investigated. Appropriate measures should be taken to avoid skin and eye contact, inhalation, and ingestion. The material contains 0.09% sodium azide as a preservative. Although the quantity of azide is very small, appropriate care should be taken when handling this material as indicated above. The National Institute of Occupational Safety and Health has issued a bulletin citing the potential explosion hazard due to the reaction of sodium azide with copper, lead, brass, or solder in the plumbing systems. Sodium azide forms hydrazoic acid in acidic conditions and should be discarded in a large volume of running water to avoid deposits forming in metal drainage pipes.

For Research Use Only

Thermo Fisher Scientific
Anatomical Pathology
47777 Warm Springs Blvd.
Fremont, CA 94539, USA
Tel: 1-510-991-2800
Fax: 1-510-991-2826
http://www.thermo.com/labvision

Manufactured by:
NeoMarkers
For
Lab Vision Corporation

EC REP

Thermo Fisher Scientific
Anatomical Pathology
93-96 Chadwick Road, Astmoor
Runcorn, Cheshire WA7 1PR, UK
Tel: 44-1928-562600
Fax: 44-1928-562627
Labvision.uk@thermofisher.com

NEOMARKERS
YOUR COLLEAGUE IN CANCER RESEARCH

DATA SHEET
Rev 030702D

Stat5b Ab-2

Rabbit Polyclonal Antibody

Cat. #RB-096-P0, -P1, or -P (0.1ml, 0.5ml, or 1.0ml at 1.0mg/ml) (Purified Ab with BSA and Azide)
Cat. #RB-096-P1ABX or -PABX (0.5ml or 1.0ml at 1.0mg/ml) (Purified Ab without BSA and Azide)
Cat. #RB-096-B0, -B1, or -B (0.1ml, 0.5ml, or 1.0ml at 1.0mg/ml) (Biotin-Labeled Ab with BSA and Azide)
Cat. #RB-096-PCL (0.1ml) (Positive Control for Western Blot)
Cat. #PP-096 (1.0ml at 200µg/ml) (Blocking Peptide with BSA and Azide)

Description: Stat5b protein is a transcription factor activated by hormone and cytokines. Two highly related, but distinct Stat5 genes (Stat5a and Stat5b) were identified in mouse. The amino acid sequences of Stat5a and Stat5b show ~96% sequence similarity, and both proteins are co-expressed in most tissues of both virgin and lactating mice. However, differential accumulation of Stat5a and Stat5b mRNA has been reported for both muscle and mammary tissue. Stat5b is critically involved in a variety of physiological functions, including reproduction, lactation, immune function and somatic growth.

Mol. Wt. of Antigen: 90kDa

Epitope: C-terminal

Species Reactivity: Human, Mouse, and Rat. Others-not known.

Immunogen: A synthetic peptide from C-terminal sequence of human Stat5b.

Applications and Suggested Dilutions:

- Gel Supershift
- Immunoprecipitation (Use Protein A)
  (Ab at 10µg/mg protein lysate)
- Western Blotting (5-10µg/ml for 2hrs at RT)

The optimal dilution for a specific application should be determined by the investigator.

Positive Control: Jurkat cells.

Cellular Localization: Nuclear

Supplied As:

Total IgG purified from rabbit anti-serum by Protein A chromatography. Prepared at 1mg/ml in 10mM PBS, pH 7.4, with 0.2% BSA & 0.09% sodium azide. Also available without BSA and azide at 1mg/ml.

Storage and Stability:

Ab with sodium azide is stable for 24 months when stored at 2-8°C. Antibody WITHOUT sodium azide is stable for 36 months when stored at below 0°C.

Suggested References:

1. Daino H, et al. Blood. 2000 Apr 15; 95(8):2577.
2. Frank DA. Mol Med 1999 Jul; 5(7):432-56.
3. Stephanou A, Latchman DS. Gene Expr 1999; 7(4-6):311-9.

Limitations and Warranty:

Our products are intended FOR RESEARCH USE ONLY and are not approved for clinical diagnosis, drug use or therapeutic procedures. No products are to be construed as a recommendation for use in violation of any patents. NeoMarkers makes no representations, warranties or assurances as to the accuracy or completeness of information provided on our data sheets and website. Our warranty is limited to the price paid for the product. NeoMarkers is not liable for any property damage, personal injury, time or effort or economic loss caused by our products.

Material Safety Data:

This product is not licensed or approved for administration to humans or to animals other than the experimental ani mals. Standard Laboratory Practices should be followed when handling this material. The chemical, physical, and toxicological properties of this material have not been thoroughly investigated. Appropriate measures should be taken to avoid skin and eye contact, inhalation, and ingestion. The material contains 0.09% sodium azide as a preservative. Although the quantity of azide is very small, appropriate care should be taken when handling this material as indicated above. The National Institute of Occupational Safety and Health has issued a bulletin citing the potential explosion hazard due to the reaction of sodium azide with copper, lead, brass, or solder in the plumbing systems. Sodium azide forms hydrazoic acid in acidic conditions and should be discarded in a large volume of running water to avoid deposits forming in metal drainage pipes.

For Research Use Only

47790 Westinghouse Dr., Fremont, CA 94539, USA • Tel.: (510) 991-2850 • Fax: (510) 991-2860
E-Mail: neomark@ix.netcom.com • Website: www.neomarkers.com

Cell Line Designation: T-47D
ATCC® Catalog No. HTB-133

Table of Contents:

- Cell Line Description
- Biosafety Level
- Use Restrictions
- Handling Procedure for Frozen Cells
- Handling Procedure for Flask Cultures
- Medium Renewal
- Complete Growth Medium
- Cryoprotectant Medium
- References
- Replacement Policy
- Specific Batch Information

Cell Line Description

Organism: *Homo sapiens* (human)
Tissue: mammary gland; breast; from metastatic site: pleural effusion; ductal carcinoma
Age: 54 years
Gender: female
DNA profile (STR analysis):
Amelogenin: X
CSF1PO: 11,13
D13S317: 12
D16S539: 10
D5S818: 12
D7S820: 11
TH01: 6
TPOX: 11
vWA: 14
Morphology: epithelial
Growth properties: adherent
DoublingTime: 32 hrs
Receptors: calcitonin; androgen receptor, positive; progesterone receptor, positive; glucocorticoid; prolactin; estrogen receptor, positive.
Isoenzymes: AK-1, 1; ES-D, 2; G6PD, B; GLO-I, 1-2; PGM1, 1; PGM3, 1
Tumorigenic: forms colonies in soft agar
Depositors: I. Keydar
Comments: The T-47 line was isolated by I. Keydar from a pleural effusion obtained from a 54 year old female patient with an infiltrating ductal carcinoma of the breast.
This differentiated epithelial substrain (T-47D) was found to contain cytoplasmic junctions and receptors to 17 beta estradiol, other steroids and calcitonin.It will form colonies in soft agar.
The cells express the WNT7B oncogene [PubMed: 8168088].
Cytogenetic analysis: This is a hypotriploid human cell line. The modal chromosome number is 65 occurring at 50% and polyploidy at 0.8%. 18 marker chromosomes are common to most cells, of which 7 are paired and 11 are single-copied. The t(8q14q), t(9q17q), t(10q17p) are among 7 paired markers common to most cells. N7, N9, and N10 are absent and N11 is generally present in 4 copies. DM's occurred, but infrequently. Q-band examination did not show the presence of a Y chromosome.

Note: Cytogenetic information is based on initial seed stock at ATCC. Cytogenetic instability has been reported in the literature for some cell lines.
Purified DNA: from this line is available as ATCC HTB-133D (10µg).
Total RNA from this line is available as ATCC HTB-133R (100µg)

Biosafety Level: 1

This cell line is not known to harbor an agent known to cause disease in healthy adult humans. Handle as a potentially biohazardous material under at least Biosafety Level 1 containment. This cell line has **NOT** been screened for Hepatitis B, human immunodeficiency viruses or other adventitious agents. Cell lines derived from primate lymphoid tissue may fall under the regulations of 29 CFR 1910.1030 Bloodborne Pathogens. ATCC recommends that appropriate safety procedures be used when handling all cell lines, especially those derived from human or other primate material. Detailed discussions of laboratory safety procedures are provided in **Laboratory Safety: Principles and Practice** (Fleming et al., 1995) the ATCC manual on quality control (Hay et al., 1992), the Journal of Tissue Culture Methods (Caputo, 1988), and in the U.S. Government Publication, ***Biosafety in Microbiological and Biomedical Laboratories*** , 4th ed. HHS Publication No. (CDC) 93-8395. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. Washington DC: U.S. Government Printing Office; 1999. The entire text is available online at www.cdc.gov/od/ohs/biosfty/bmbl4/bmbl4toc.htm.

Use Restrictions

These cells are distributed for research purposes only. ATCC recommends that individuals contemplating commercial use of any cell line first contact the originating investigator to negotiate an agreement. Third party distribution of this cell line is discouraged, since this practice has resulted in the unintentional spreading of cell lines contaminated with inappropriate animal cells or microbes.

Handling Procedure for Frozen Cells

To insure the highest level of viability, thaw the vial and initiate the culture as soon as possible upon receipt. If upon arrival, continued storage of the frozen culture is necessary, it should be stored in liquid nitrogen vapor phase and not at –70°C. Storage at –70°C will result in loss of viability.

SAFETY PRECAUTION: ATCC highly recommends that protective gloves and clothing always be used and a full face mask always be worn when handling frozen vials. *It is important to note that some vials leak when submersed in liquid nitrogen and will slowly fill with liquid nitrogen. Upon thawing, the conversion of the liquid nitrogen back to its gas phase may result in the vessel exploding or blowing off its cap with dangerous force creating flying debris.*

1. Thaw the vial by gentle agitation in a 37°C water bath. To reduce the possibility of contamination, keep the O-ring and

cap out of the water. Thawing should be rapid (approximately 2 minutes).

2. Remove the vial from the water bath as soon as the contents are thawed, and decontaminate by dipping in or spraying with 70% ethanol. *All of the operations from this point on should be carried out under strict aseptic conditions.*

3. Transfer the vial contents to a centrifuge tube containing 9.0 ml complete culture medium. and spin at approximately 125 xg for 5 to7 minutes.

4. Resuspend cell pellet with the recommended complete medium (see the specific batch information for the culture recommended dilution ratio). and dispense into a 25 cm^2 or a 75 cm^2 culture flask. *It is important to avoid excessive alkalinity of the medium during recovery of the cells. It is suggested that, prior to the addition of the vial contents, the culture vessel containing the complete growth medium be placed into the incubator for at least 15 minutes to allow the medium to reach its normal pH (7.0 to 7.6).*

5. Incubate the culture at 37°C in a suitable incubator. A 5% CO_2 in air atmosphere is recommended if using the medium described on this product.

Handling Procedure for Flask Cultures

The flask was seeded with cells (see specific batch information) grown and completely filled with medium at ATCC to prevent loss of cells during shipping.

1. Upon receipt visually examine the culture for macroscopic evidence of any microbial contamination. Using an inverted microscope (preferably equipped with phase-contrast optics), carefully check for any evidence of microbial contamination. Also check to determine if the majority of cells are still attached to the bottom of the flask; during shipping the cultures are sometimes handled roughly and many of the cells often detach and become suspended in the culture medium (but are still viable).

2. **If the cells are still attached,** aseptically remove all but 5 to 10 ml of the shipping medium. The shipping medium can be saved for reuse. Incubate the cells at 37°C in a 5% CO_2 in air atmosphere until they are ready to be subcultured.

3. **If the cells are not attached,** aseptically remove the entire contents of the flask and centrifuge at 125 xg for 5 to 10 minutes. Remove shipping medium and save. Resuspend the pelleted cells in 10 ml of this medium and add to 25 cm^2 flask. Incubate at 37°C in a 5% CO_2 in air atmosphere until cells are ready to be subcultured.

Subculturing Procedure

Volumes used in this protocol are for 75 cm^2 flask; proportionally reduce or increase amount of dissociation medium for culture vessels of other sizes.

1. Remove and discard culture medium.

2. Briefly rinse the cell layer with 0.25% (w/v) Trypsin- 0.53 mM EDTA solution to remove all traces of serum which contains trypsin inhibitor.

3. Add 2.0 to 3.0 ml of Trypsin-EDTA solution to flask and observe cells under an inverted microscope until cell layer is dispersed (usually within 5 to 15 minutes).

Note: To avoid clumping do not agitate the cells by hitting or shaking the flask while waiting for the cells to detach. Cells that are difficult to detach may be placed at 37°C to facilitate dispersal.

4. Add 6.0 to 8.0 ml of complete growth medium and aspirate cells by gently pipetting.

5. Add appropriate aliquots of the cell suspension to new culture vessels.
Subcultivation Ratio: 1: 3 to 1: 5

6. Incubate cultures at 37°C.

Note: For more information on enzymatic dissociation and subculturing of cell lines consult Chapter 10 in ***Culture of Animal Cells, a manual of Basic Technique*** by R. Ian Freshney, 3rd edition, published by Alan R. Liss, N.Y., 1994.

Medium Renewal

Two to three times weekly

Complete Growth Medium

These cells are grown in RPMI 1640 medium with 2 mM L-glutamine that is modified by ATCC to contain:

10 mM HEPES
1 mM sodium pyruvate
4.5 g/L glucose
1.5 g/L sodium bicarbonate

Supplemented with:

10 % fetal bovine serum.
0.2 units/ ml bovine insulin

This medium is formulated for use with a 5% CO_2 in air atmosphere.

This ATCC modified and tested medium formulation (without the additional supplements or serum described above) is available as ATCC Catalog No. 30-2001.

ATCC tested fetal bovine serum is available as ATCC Catalog No. 30-2020.

Cryoprotectant Medium

Complete growth medium described above supplemented with 5% (v/v) DMSO.

Cell culture tested DMSO is available as ATCC Catalog No. 4-X.

Additional Information

Additional product and technical information can be obtained from the catalog references and the ATCC Web site at www.atcc.org, or by e-mail at tech@atcc.org.

S4776 Somatotropin (STH) human pituitary

Sigma

Synonym
Growth Hormone

CAS Number
9002-72-6
MDL number
MFCD00081960

Expand/Collapse All

Price and Availability

Product Number	List Price AUD	Your Price AUD	Available to Ship	Quantity	Actions
S4776-1VL		612.80	06.06.2005 details...		

Descriptions

Biochem/physiol Actions
Increases mass of most tissues by increasing cell number rather than cell size, mobilizes fat stores

Packaging
Vial contains approx. 4 I.U. hGH, with 0.1 mg ammonium bicarbonate and 3.5 mg mannitol.

Properties

mol wt
mol wt 22,125 Da by calculation

storage temp.
2-8°C

Safety

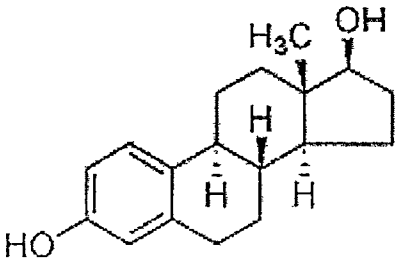
Hazard Codes
B

Related Categories
... Mass Spectrometry > **Peptide and protein standards for mass spectrometry analysis**

... Somatotropin > **Somatotropin (STH)**

E1024 Estradiol

Sigma meets USP testing specifications for



Synonym
β-Estradiol

1,3,5-Estratriene-3,17β-diol

17β-Estradiol

3,17β-Dihydroxy-1,3,5(10)-estratriene

Dihydrofolliculin

Molecular Formula
$C_{18}H_{24}O_2$

Molecular Weight
272.38

CAS Number
50-28-2
Beilstein Registry Number
1914275

EG/EC Number
2000238

MDL number
MFCD00003693

Expand/Collapse All

Price and Availability

Product Number	List Price AUD	Your Price AUD	Available to Ship	Quantity	Actions
E1024-1G		41.70	06.06.2005 details...		
E1024-25G		586.90	06.06.2005 details...		
E1024-100G		2,270.00	06.06.2005 details...		

Descriptions

L7009 Prolactin from human pituitary

Sigma ~30 IU/mg

Synonym
Lactogenic Hormone

LTH

Luteotropic Hormone

CAS Number
9002-62-4
MDL number
MFCD00131553

Expand/Collapse All

Price and Availability

Product Number	List Price AUD	Your Price AUD	Available to Ship	Quantity	Actions
L7009-1UG		112.70	06.06.2005 details...		
L7009-5UG		371.10	06.06.2005 details...		

Descriptions

Biochem/physiol Actions
Prolactin is a neuroendocrine hormone. The prolactin receptor is a transmembrane glycoprotein that belongs to the cytokine hematopoietic receptor family. A large number of cells and organs express the receptor, including B cells, T cells, macrophages, monocytes and neutorphils. Prolactin signal transduction involves the JAK/STAT families and the src kinase family. Induces lactation; inhibits secretion of gonadotropins; release is inhibited by dopamine.

Other Notes
RIA active LTH

Unit Definition
Specific activity by RIA per an international reference preparation of LTH for immunoassay

Properties

storage temp.
2-8°C

Related Products

Related product
L4021, Prolactin human

Related Categories
... Luteotropic Hormone (Prolactin) > **Prolactin**

3 Appendix C

ACB	Healthy v Cancer Analysis							Early v Late Analysis							
Solution 2	Chain A-4702	DIRP 5a - 2501	HSP 60-3702	Lumican - 2701	Osteoglycin (2403)	Osteoglycin (2406)	Serum albumin-1701	Serum albumin-501	Lumican 1805	Proapolipo - 2101	Lumican 2802	α1-collagen IV -2804	IgG Light chain 5202	Cofilin-7003	Cofilin- 7004
ACB 1	0.9	0	1.3	1 5	0.9	0	0	0 5	0.4	0	0	0	0	1.8	0
ACB 2	1.5	0	0.5	0.3	0	0	0.3	0 3	0	0	0.1	0	0	0.2	0
ACB 3	2.1	0	1.5	0.2	0	0	0.9	1.7	0 2	2.3	0.9	0	0.3	0	0
ACB 4	1.1	0	0 5	0	0	0	0	0	0	0	0	0	0.5	0	0
ACB 5	0.5	0	0 5	0.5	0	0	0	0.6	0.2	0.1	0.9	0	1	0	0
ACB 6	1.1	0	1.7	0	0	0	0	0 7	0	0.2	0	0	0	0	0
ACB 7	1	0	0 9	0	0	0	0	0.6	0	1	0	0	0.3	0	0
ACB 8	0	0	0.9	2.1	0.4	0	1.2	1.2	0.7	0.4	0.6	0	0	0	0
ACB 9	2.1	0	1.2	0.1	0	0	1	1.3	0	0.2	2.5	0	0	0	0
ACB 10	4.5	0	1.5	0	0	0	0	2.8	0	0.2	0	0	0	0	0
ACB 11	1.4	0	0.1	0.5	0.8	0.9	0.7	0.1	1 1	0.3	0.3	0 4	0 6	0.1	0 1
ACB 12	4	0	0	0	0	0.4	0.5	1.7	0.3	0	0	0	0.5	0	0
ACB 13	0.3	0	0 6	0.5	0	0	1.1	1.6	0	0	0	0	0	0	0
ACB 24	0.7	0	0 6	1.3	0	0	1.1	0.9	0	0	0.3	0	0.1	0	0
ACB 14	0.2	0	1 3	0.1	0	0	0	0.5	0	0.4	0.5	0	0	0	0
ACB 15	0 5	0	0 3	0.3	0	0	1.9	0.5	0 2	0.3	0.7	0	0 3	0	0
ACB 16	0	0	0.5	0	0.4	0	0	1 3	0	0	0 1	0	0	0	0
ACB 17	0.8	0	0.7	0	0	0	0.3	3	0	2.3	0	0	0	0	0
ACB 18	0	0	1	0	0	0	0	1	0	0.1	0.3	0	0.8	0	0.7
ACB 19	0	0	0.4	0	0	0	0.2	1.2	1	0.2	0	0	0.3	0	1.2
ACB 20	0	0	0.4	0	0	0.7	0.3	2.2	0	0	0	0	0	0	0
ACB 21	0.4	0	2.5	0	0.4	0	1.5	4.6	0	0	0.6	0	0	0	0.7
ACB 22	0	0	1.6	0	0	0	0	0	0	0.2	0 6	0	0	0	0
ACB 23	0	0	2 7	0	0	0	0	0.4	0	0.8	0	0	0.2	1	0
Averages	0.96	0	0.966	0.308	0.120	0 083	0.458	1.19	0 17	0.37	0.35	0.01	0.20	0.12	0.11

Appendix C1:Protein concentrations (based on the % of total protein content) for Solution 2 proteins of interest identified in the ACB population.

Solution 3	FLJ20309-7204	Folliculin-3701	HSP-60-4103	I-plastin-7203	RP-3-393D12.2-8203	Upd-glyc -7112	α-1 collagen VI 1806	α-1 collagen VI 1821	Serum albumin-3804	Crystallin αB -8102
ACB 1	0.4	2.2	0	0.2	0	0	0	0	0	0.3
ACB 2	0	0.7	0	1.5	0.3	0	0	0	0	0.4
ACB 3	0	1.1	0	0.8	0	0	0	0	0	1.3
ACB 4	0	1.4	0	0.4	0.7	0	0	0	0	1.1
ACB 5	0	0.4	0	0.3	0.4	0	0	0	0	0.4
ACB 6	0	0.7	0	2.3	0	0	0	0	0	1.2
ACB 7	0	0.6	0	1	0	0	0	0	0	0.3
ACB 8	0.4	0.3	0	0.2	0	0	2.6	0	0	1.1
ACB 9	3.9	2.1	0	2.7	0	0	0	0	0	3.9
ACB 10	1.3	0.4	0	0.5	0	0	0	0	0	0
ACB 11	0	0.2	0	0.6	0.4	0	0	0	0.2	0.3
ACB 12	1.4	0.9	0	2.2	0.8	0	0	0	0	0.5
ACB 13	5	1.5	0	0.3	2.6	0	0	0	0	0.2
ACB 14	1.8	0.5	0	1	0.1	0	0	0	0	2.6
ACB 15	0	0	0	1.9	0	0	0	0	0	0
ACB 16	1.9	0.3	0	1	0	0	0	0	0	0.4
ACB 17	1.3	0.3	0	0.3	0.7	0	0	0	0.1	0.6
ACB 18	0.3	0.8	0	0	0	0	0	0	0	1
ACB 19	0	0.8	0	2.1	0	0	0.2	0	0	0.1
ACB 20	1.5	0.4	0	0.8	1.2	0	0	0	0.1	0.2
ACB 21	0.8	1.6	0	0.5	0	0	1.6	0	0	0
ACB 22	1.2	2	0	0.9	0.1	0	0	0	0	1.1
ACB 23	0.8	0.5	0	0.5	0	0	0	0	0	0.5
Averages	0.95	0.85	0	0.95	0.31	0	0.19	0	0.01	0.76

Appendix C2:Protein concentrations (based on the % of total protein content) for Solution 3 proteins of interest identified in the ACB population.

	Serum albumin-1701	Osteoglycin (2403)	Osteoglycin (2406)	DIRP 5a - 2501	Lumican - 2701	HSP 60-3702	Chain A - 4702	Serum albumin-501	Lumican 1805	Proapolipo - 2101	Lumican 2802	α1-collagen IV -2804	IgG Light chain 5202	Cofilin-7003	Cofilin- 7004
TCB 1A	0	0	0	0.5	0	0.2	1.8	0.5	0	0.4	0	0	0.6	0.2	0.2
TCB 1	0	0	0	0.3	0	0	0	0	0	0.8	0	0	0.2	1 5	0
TCB 2	0	0	0	0.4	0	0.6	0.4	0	0.4	0.5	0	0	0	0	0
TCB 3	0	0	0	0	0	0	3.5	0	0	0.6	0	0	0.4	0.3	0
TCB 4a	0	0	0	0	0	0.7	1	1.9	0	0.9	0	0	0.2	0	0
TCB 4	0	0	1.6	0	0	1	0.5	0	0	0.5	0	0	0	0	0
TCB 5	0	0.1	0.7	0.2	0	1	0.3	0	1	0.9	0	0	0	1.4	0
TCB 6	0	0	0	0.7	0	1.1	0	0	0	1.6	0	0	0.1	0	0
TCB 7	0	0	0	0	0	0.7	0	0.6	0	0.1	0	0	0	1.1	0
TCB 8	0.6	0	0	0.7	0.4	0.6	2 5	0	0	1.7	0	0	2.2	0	0
TCB 9	0	1.2	0	0	0	0	2.3	0	0	2.1	0	0	0	0	0
TCB 10	0	0	0.5	0.4	0	0.2	0	0	1.3	1.3	0	0	1.2	0	0
TCB 11	0.8	0 1	0.1	0	0.1	0	0.4	0 2	0.8	0 4	0	0.5	0 1	0	0
TCB 12	1.5	0	0 5	0	0.1	0	1.4	0	0	1 3	0	0	0	0	0
TCB 13	0	0	0	1.1	0	0.5	0	1.6	0	0.2	0	0.1	0.1	0	0
TCB 14	0	0	0	0 9	0	1 2	3.9	0	0	0 5	0	0	0	0.3	0
TCB 15	0	0	0	0 5	0	0 8	0	0.3	0	0.5	0.3	0	0	0	0
TCB 17	0	0.2	0.9	0.1	0	0.7	0.7	0	0	0.6	0	0	0	0	0.2
TCB 16	0	0	0	0.2	0.1	0	2.1	0	0	1	1.4	0	0.3	0	0
TCB 19	0	0.4	1 2	0	0	0	1.9	0.4	0	0.8	0	0	0	0.4	0
TCB 18	0	0	0	0.6	0	0.4	0.5	0	0	1.2	0	0	0	0	0
TCB 21	0	0.2	0.3	0.2	0	0.3	3 2	0.2	1.1	1.1	0.1	0	0	0	0
TCB 22	0	0	1.9	0	0	0	0.6	0	0	1.3	0	0	0	0	0
TCB 23	0	0.2	0.1	0.1	0	1.4	1.7	0.6	0.1	1 7	0	0	0	0	0
TCB 24	0	0	0	0.4	0	0.7	1	0	0	0.3	0.3	0.2	0	0	1
TCB 25	0	0	0	0	0	1.1	0.1	0	0	0	0	0	0	0	0
TCB 26	0	0	0	0	0	0.6	0.8	1.4	1.6	0.7	0 7	0.4	0	0	0
TCB 27	0	0	0	0	0	0	5.7	0	0	1.5	0	0	0	0.5	0
TCB 28	0	0.2	0.3	0.2	0	0 2	2.4	0 5	1.3	0	0.6	0	0	0	0 2
Avg.	0.1	0.09	0.28	0.26	0.02	0.48	1.33	0.275	0.22	0.87	0.1	0.04	0.19	0 20	0.0

Appendix C3:Protein concentrations (based on the % of total protein content) for Solution 2 proteins of interest identified in the TCB population.

Solution 3	α-1 collagen VI 1806	Folliculin-3701	HSP-60-4103	Upd-glyc -7112	I-plastin - 7203	FLJ20309 -7204	RP-3-393D12.2-8203	α-1 collagen VI-1821	serum albumin-3804	Crystallin αB -8102
TCB 1A	0.00	0.00	0.00	0.00	0.10	0.00	0.00	0.00	0.00	0.40
TCB 1	2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TCB 2	1.90	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TCB 4a	0.00	0.40	0.00	0.00	0.00	0.70	0.00	0.30	0.00	0.00
TCB 4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TCB 5	0.00	0.00	0.40	0.00	0.00	0.00	0.00	2.40	0.50	0.00
TCB 6	0.20	1.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TCB 7	3.30	2.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.30
TCB 8	0.00	0.30	0.90	0.00	0.00	0.20	0.00	0.10	0.00	0.10
TCB 9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.40	0.00
TCB 10	0.00	0.60	0.60	0.20	0.00	0.00	0.00	0.00	0.00	0.00
TCB 11	0.00	0.40	0.20	0.00	0.00	0.00	0.00	0.00	0.80	0.00
TCB 12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.00	0.00
TCB 13								0.10	0.00	0.00
TCB 14	1.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TCB 15	0.40	0.40	0.30	0.50	0.00	0.00	0.00	0.00	0.00	0.00
TCB 16	0.40	0.00	2.20	0.00	0.00	0.80	0.00	0.00	0.20	0.00
TCB 17	0.00	0.00	0.00	0.30	0.00	0.00	0.00	0.00	0.00	0.00
TCB 18	0.80	0.00	1.30	0.50	0.00	0.00	0.00	0.00	0.00	0.00
TCB 19	0.00	0.80	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TCB 20	2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TCB 21	0.00	0.00	0.00	0.00	0.00	1.80	0.00	0.00	0.30	0.00
TCB 22	0.50	0.70	0.00	0.00	0.00	0.00	0.00	0.00	0.30	0.00
TCB 23	0.00	2.10	0.00	0.10	0.00	0.00	0.00	0.00	0.00	0.00
TCB 24	0.60	0.50	0.90	0.00	0.00	0.00	0.10	0.00	0.00	0.00
TCB 26	0.20	0.00	0.00	0.00	0.00	0.20	0.00	0.00	0.00	0.00
TCB 27	0.00	0.30	0.60	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Avg.	0.62	0.39	0.28	0.06	0.00	0.14	0.00	0.11	0.09	0.10

Appendix C4:Protein concentrations (based on the % of total protein content) for Solution 3 proteins of interest identified in the TCB population.

Solution 2	Cancer v Healthy proteins							Early v Late onset proteins							
Data	Serum albumin-1701	Osteoglycin (2403)	Osteoglycin (2406)	DIRP 5a - 2501	Lumican - 2701	HSP 60-3702	Chain A - 4702	Serum albumin-501	Lumican 1805	Proapolipo - 2101	Lumican 2802	α1-collagen IV -2804	IgG Light chain 5202	Cofilin-7003	Cofilin- 7004
GCB 1	0.00	0 00	0.00	0.00	0.00	0.56	0 00	0.80	0.00	0.50	0.00	0.00	0 70	0 00	0.20
GCB 2	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0 50	0.70	0.00	0.00	0.00	0.10	0.00	0.00
GCB 3	0.00	0 00	0.00	0.00	0.00	0.54	0.00	0.80	0.00	0.00	0.00	0.00	0.70	0.00	0.00
GCB 4	0.00	0.00	0.00	0.00	0.00	2.91	0.67	0.80	3.00	0.40	0.00	0.00	1.10	0.00	0.70
GCB 5	0.00	0 00	0.00	0.00	0.00	0.00	7.37	3 70	1.00	0.00	0.50	0.00	0.00	0.00	0.00
GCB 6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0 50	0 00	0.00	0.00	0.00	0.60	0.00	0.00
GCB 7	1.25	0.00	0.51	0.00	0.00	0.54	2.46	0.70	0 00	0.30	2.80	0.00	0.00	0.50	0.00
GCB 8	0.00	0.00	0.00	0.00	0.00	1.88	0.54	0.60	0.00	0.30	0.00	0.00	0 00	0.00	0.00
GCB 9	0.16	1.62	0 75	0.00	11.14	0.99	1.46	0.00	0.00	2 40	0.50	0 00	0 00	1.80	0.00
GCB 10	2.08	0.00	0.00	0.00	0 00	0.29	0.00	0.40	0.20	0.10	0.60	0.00	0.60	8 20	0 40
GCB 11	0 46	0.84	0.89	0.00	0 50	1.19	1.38	0.60	1.00	0.50	2 30	0.00	0 20	0 00	0.00
GCB 12	0.00	0.00	0 00	0 00	0.00	0.53	0.00	0.40	0.10	0 10	0 30	0.00	0 00	3.50	1.70
GCB 13	0.00	0.91	0.00	0.00	0.00	0.22	0.00	1.40	0.00	0.00	0.80	0.00	0.50	0.00	0.00
GCB 14	0.10	0.00	0.00	0.00	0 42	1.44	1.36	0.00	0.00	0.50	0.00	0.00	0.70	0 10	1.70
GCB 15	1.18	0.74	1 41	0 00	0.00	0.64	0.58	0.20	1.90	1 40	0.30	0.00	0.60	0.00	0.00
GCB 16	0.50	0.50	1.60	0 00	0.68	0.73	1.04	0.10	0.00	1.70	0.10	0.00	0.00	0.00	0.00
Average	0.36	0.29	0.32	0.00	0 80	0.85	1.05	0.72	0.49	0.51	0.51	0.00	0.36	0.88	0.29

Appendix C5:Protein concentrations (based on the % of total protein content) for Solution 2 proteins of interest identified in the GCB population.

Solution 3	α-1 collagen VI 1821	α-1 collagen VI 1806	Crystallin αB -8102	FLJ20309 -7204	Follicullin-3701	HSP-60-4103	I-plastin -7203	RP-3-393D12.2	serum albumin-3804	Upd-glyc -7112
GCB 1	0.0	0.0	1.7	0.0	0.6	0.0	0.0	0.0	0.3	0.4
GCB 2	0.0	5.3	0.9	0.0	0.2	0.5	0.0	0.0	0.2	0.6
GCB 3	0.0	0.0	0.0	0.0	1.1	0.0	0.0	0.0	0.0	0.7
GCB 4	0.0	0.0	1.2	0.0	0.0	0.6	1.1	0.0	0.0	0.9
GCB 5	0.0	0.0	0.0	0.0	0.0	0.0	3.2	0.0	0.0	0.7
GCB 6	0.0	0.0	0.0	0.0	0.5	0.0	0.4	0.0	0.0	0.4
GCB 7	0.0	0.0	0.2	0.0	0.4	0.0	0.5	0.0	0.2	0.0
GCB 8	0.0	0.2	0.0	0.0	0.5	0.0	0.0	0.0	0.3	0.6
GCB 9	0.0	0.8	1.2	0.0	6.0	1.7	0.0	0.0	0.0	0.0
GCB 10	0.0	0.0	0.0	0.0	0.4	0.3	0.6	0.0	0.2	0.0
GCB 11	0.0	0.0	0.0	0.0	0.5	0.3	0.0	0.0	0.0	0.0
GCB 12	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0
GCB 13	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	1.7	0.0
GCB 14	0.0	0.0	0.0	0.0	0.2	0.1	0.0	0.0	0.0	0.0
GCB 15	0.0	0.0	0.2	0.0	0.3	0.0	0.4	0.0	0.1	0.1
GCB 16	0.0	0.6	0.0	0.0	0.3	0.1	0.2	0.0	0.2	0.1
Average	0.0	0.4	0.3	0.0	0.7	0.2	0.4	0.0	0.2	0.3

Appendix C6:Protein concentrations (based on the % of total protein content) for Solution 3 proteins of interest identified in the GCB population.

Solution 2	α-1 collagen VI-2804	Chain A - 4702	Cofilin-7003	DIRP 5a-2501	HSP 60-3702	IgG light Chain-5702	Lumican-2701	Osteloglycin 2403	Osteoglycin 2406	Proapolipoprotein 2101	Serum albumin 3804
NB 1	0.0	0.0	0.0	0.0	0.1	0.0	6.2	2.8	3.1	1.6	0.8
NB 2	0.0	0.0	0.0	0.0	0.2	0.0	3.2	3.7	4.8	1.7	2.8
NB 3	0.0	0.9	0.0	0.0	0.0	0.0	5.9	0.0	1.6	1.8	2.6
NB 4	0.0	3.1	0.0	0.0	0.2	0.4	3.0	1.0	1.9	0.7	4.7
NB 5	0.0	0.0	0.0	0.0	0.0	0.8	0.5	2.6	0.0	1.3	4.0
NB 6	0.0	0.0	0.0	0.0	0.0	0.5	0.0	1.0	1.8	0.8	0.0
NB 7	0.0	8.3	0.0	0.0	0.4	0.0	6.6	3.5	0.1	1.6	4.0
NB 8	0.0	2.9	0.0	0.0	0.1	0.2	7.6	1.6	0.5	1.0	3.0
NB 9	0.0	1.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.2
NB 10	0.0	2.5	0.1	0.0	0.0	0.0	9.5	1.7	1.4	0.2	2.2
NB 11	0.0	0.0	0.0	0.0	0.0	0.0	9.5	0.4	0.2	7.0	7.7
NB 12	0.0	4.0	0.1	0.0	0.2	0.0	5.0	2.6	0.5	0.6	1.7
Averages	0.0	1.9	0.0	0.0	0.1	0.2	4.8	1.7	1.3	1.5	2.8

Appendix C7:Protein concentrations (based on the % of total protein content) for Solution 2 proteins of interest identified in the NB population.

Solution3	α-1 collagen VI 1821	α-1 collagen VI 1806	Crystallin aB-8102	FLJ20309 -7204	Folliculin-3701	HSP-60-4103	I-plastin - 7203	RP-3-393D12.2	serum albumin-3804	Upd-glyc -7112
NB 1	0.00	1.90	4.70	0.00	0.00	0.00	0.20	0.00	0.00	0.00
NB 2	0.00	0.00	0.90	0.00	0.30	0.00	0.10	0.00	0.30	0.00
NB 3	0.00	0.00	1.70	0.00	0.10	0.00	0.10	0.00	0.10	0.00
NB 4	0.00	0.00	1.70	0.00	0.60	0.00	0.00	0.00	0.00	0.00
NB 5	0.00	0.00	3.20	0.00	0.00	2.00	0.30	0.00	0.00	0.00
NB 6	0.00	0.00	2.40	0.00	0.00	2.60	0.00	0.00	0.00	0.00
NB 7	0.00	0.00	2.10	0.00	1.10	2.40	0.20	0.00	0.10	0.00
NB 8	0.00	0.00	1.30	0.00	0.30	0.00	0.10	0.00	0.00	0.00
NB 9	0.00	0.00	1.00	0.00	0.20	0.00	0.10	0.10	0.00	0.00
NB 10	0.00	0.00	1.30	0.00	0.40	0.00	0.00	0.00	0.10	0.00
NB 11	0.00	0.00	2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Averages	0.00	0.17	2.03	0.00	0.27	0.64	0.10	0.01	0.05	0.00

Appendix C8:Protein concentrations (based on the % of total protein content) for Solution 3 proteins of interest identified in the NB population.

4 Appendix D

Comparison	HSA 1701	Osteoglycin 2403	Osteoglycin 2406	DIRP5a 2501	Lumican 2701	Chain A 4702	HSP 60 3702
Cancer v Healthy							
ACB/NB	0.001	0.013	0.001		0.065	0.24	0.0000004
TCB/NB	0.0003	0.013	0.002	0.0001	0.028	0.47	0.0003
GCB/NB	0.0006	0.034	0.003		0.039	0.033	0.003
ACB/GCB	0.609	0.137	0.185		0.496	0.862	0.590
ACB/TCB	0.010	0.923	0.185	0.0001	0.020	0.306	0.005
TCB/GCB	0.134	0.111	0.795	0.0001	0.282	0.600	0.119
Early v Late							
TCB-E/NB-E	0.012	0.175	0.021	0.003	0.112	0.752	0.0009
TCB-E/ACB-E	0.372	0.957	0.005	0.003	0.552	0.199	0.155
TCB-E/GCB-E	0.753	0.703	0.694	0.003	0.532	0.295	0.231
NB-E/ACB-E	0.014	0.174	0.010		0.116	0.671	0.014
NB-E/GCB-E	0.014	0.230	0.345		0.162	0.515	0.067
ACB-E/GCB-E	0.875	0.692	0.499		0.568	0.715	0.356
TCB-L/NB-L	0.021	0.0360	0.052	0.011	0.042	0.317	0.084
TCB-L/ACB-L	0.002	0.194	0.032	0.003	0.016	0.881	0.021
TCB-L/GCB-L	0.109	0.054	0.057	0.011	0.567	0.993	0.150
ACB-L/NB-L	0.048	0.043	0.072		0.209	0.251	0.0002
NB-L/GCB-L	0.038	0.078	0.089		0.979	0.296	0.008
ACB-L/GCB-L	0.554	0.155	0.512		0.567	0.860	0.644
NB-E/NB-L	0.790	0.956	0.544		0.288	0.329	0.314
ACB-E/ACB-L	0.332	0.738	0.032		0.049	0.629	0.641
TCB-E/TCB-L	0.109	0.048	0.005	0.126	0.109	0.721	0.872
GCB-E/GCB-L	0.714	0.892	0.617		0.554	0.454	0.366

Appendix D1: Student T-Test results for the proteins identified in Cancer v Healthy comparisons in Solution 2. CI % of 0.05.

Comparison	Folliculin	HSP-4103	UPD-Glyc.	L-Plastin	FLJ	RP3
TE/NE	0.051440289	0.302112723	0.110523938	0.141927448	0.07859573	-
TE/AE	0.140434404	0.070357747	0.110523938	0.021825117	0.449922313	0.373900966
TE/GE	0.69995354	0.070357747	0.434461753	0.499999999	0.07859573	-
NE/AE	0.026940046	0.182615634	-	0.025870619	0.389972553	0.373900966
NE/GE	0.378260963	0.182615634	0.409665529	0.705101589	-	-
AE/GE	0.575216113	-	0.409665529	0.031799493	0.335069721	0.373900966
TL/NL	0.210956735	0.977998906	0.17047066	0.062650513	-	0.845506322
TL/AL	0.000895546	0.064410619	0.17047066	0.35081531	0.002939946	0.027600738
TL/GL	0.216119031	0.597567128	0.257643254	0.094851092	-	7.26838E-05
NL/AL	0.091239157	0.363217468	-	0.000127565	0.002939946	0.029360704
NL/GL	0.518579458	0.7689731	0.009412398	0.181930664	-	0.363217468
AL/GL	0.791625996	0.044411857	0.009412398	0.246430518	0.002939946	0.023212646
NE/NL	0.166651395	0.478767118	-	1	-	0.363217468
AE/AL	0.525646109	-	-	0.107313987	0.862586849	0.062648632
TE/TL	0.071623014	0.49557017	0.299193125	0.350616663	0.07859573	0.350616663
GE/GL	0.990684033	-	0.730231249	0.489274409	-	-
AC/TC	0.013768075	0.009359681	0.043122514	6.68543E-06	0.007066924	0.020167746
AC/GC	0.676902416	0.045234036	0.00348826	0.03997715	0.001795301	0.018871346
AC/N	0.001531431	0.083607853	-	2.78759E-05	0.001795301	0.022331959
TC/N	0.456235281	0.325533358	0.043122514	0.009636413	0.07896976	0.561836555
TC/GC	0.438942999	0.767913142	0.018387373	0.06948555	0.07896976	0.326527022
N/GC	0.27653257	0.274514036	0.00348826	0.162983596	-	0.340893132

Appendix D2: Student T-Test results for the proteins identified in Cancer v Healthy comparisons in Solution 3. CI % of 0.05.

	ACB					TCB					GCB				
	1/N	2/N	3/N	1/2	2/3	1/N	2/N	3/N	1/2	2/3	1/N	2/N	3/N	1/2	2/3
Solution 2															
HSA-1701	0.001	0.001	0.000	0.976	0.080	0.000	0.000	0.000	0.203	0.953	0.002	0.001	0.000	0.108	0.838
Osteoglycin-2403	0.018	0.014	0.009	0.764	0.182	0.015	0.015	0.013	0.966	0.630	0.028	0.014	0.009	0.637	0.614
Osteoglycin-2406	0.002	0.001	0.001	0.779	0.112	0.123	0.004	0.001	0.866	0.290	0.006	0.001	0.001	0.687	0.937
DIRP5-2501						0.430	0.016	0.087	0.601	0.793					
Lumican-2701	0.044	0.079	0.037	0.387	0.251	0.027	0.030	0.028	0.163	0.533	0.034	0.079	0.037	0.409	0.414
Chain A-4702	0.134	0.351	0.365	0.414	0.815	0.845	0.717	0.073	0.736	0.017	0.111	0.351	0.365	0.292	0.262
HSP 60-3702	0.028	0.000	0.122	0.348	0.635	0.180	0.028	0.014	0.223	0.061	0.044	0.000	0.122	0.918	0.925
Solution 3														1v2	2v3
Folliculin-3701	0.070	0.012	0.003	0.951	0.276	0.023	0.555	0.171	0.030	0.287	0.530	0.458	0.478	0.479	0.486
HSP60-4103	0.084	0.084	0.084			0.084	0.569	0.250	0.026	0.318	0.195	0.958	0.160	0.515	0.480
UDP-Glyc.- 7112						0.500	0.334	0.158	0.577	0.289	0.024	0.423	0.182	0.727	0.906
L-Plastin-7203	0.029	0.005	0.057	0.375	0.130	0.008	0.008	0.008			0.139	0.460	1.000	0.553	0.461
FLJ 20309-7204	0.100	0.010		0.895	0.010		0.078		0.078	0.078					
Rp3-393D12.2- 8203	0.147	0.050	0.459	0.148	0.137	0.341	0.832	0.341	0.334	0.334	0.341	0.341	0.341		

Appendix D3: Student T-Test results for the staging based data on proteins identified in Early v Late comparisons. CI % = 0.05.

Comparison	HSA 501	Lumican 1805	Proapolipoprotein 2101	Lumican 2802	α-1 collagen 2804	Ig l chain 5202	Cofilin 7003	Cofilin 7004
Cancer v Healthy								
ACB/NB	0.016	0.042	0.054	0.32	0.032	0.641	0.137	0.141
TCB/NB	0.975	0.058	0.229	0.015	0.076	0.810	0.0176	0.310
GCB/NB	0.234	0.139	0.090	0.059		0.097	0.124	0.077
ACB/GCB	0.133	0.170	0.537	0.502	0.327	0.153	0.189	0.265
ACB/TCB	0.0006	0.425	0.007	0.074	0.381	0.836	0.558	0.428
TCB/GCB	0.080	0.336	0.116	0.087	0.076	0.165	0.229	0.130
Early v Late								
TCB-E/NB-E	0.622	0.360	0.317	0.345	0.165	0.844	0.090	0.879
TCB-E/ACB-E	0.046	0.364	0.279	0.433	0.165	0.863	0.213	0.218
TCB-E/GCB-E	0.561	0.646	0.780	0.932	0.165	0.003	0.090	0.162
NB-E/ACB-E	0.322	0.283	0.188	0.832		0.733	0.363	0.213
NB-E/GCB-E	0.939	0.785	0.313	0.373		0.033		0.363
ACB-E/GCB-E	0.188	0.559	0.826	0.471		0.016	0.363	0.192
TCB-L/NB-L	0.145	0.040	0.279	0.019	0.197	0.654	0.117	0.320
TCB-L/ACB-L	0.020	0.026	0.298	0.021	0.662	0.337	0.456	0.438
TCB-L/GCB-L	0.216	0.075	0.762	0.063	0.197	0.105	0.304	0.376
ACB-L/NB-L	0.0006	0.059	0.104	0.026	0.331	0.146	0.165	0.519
NB-L/GCB-L	0.012	0.121	0.215	0.037		0.046	0.125	0.074
ACB-L/GCB-L	0.025	0.287	0.516	0.317	0.331	0.363	0.197	0.105
NB-E/NB-L	0.390	0.953	0.341	0.044		0.264		1
ACB-E/ACB-L	0.880	0.976	0.679	0.684	0.331	0.856	0.284	0.255
TCB-E/TCB-L	0.684	0.007	0.077	0.525	0.893	0.465	0.395	0.330
GCB-E/GCB-L	0.570	0.680	0.811	0.177		0.012	0.125	0.061

Appendix D4: Student T-Test results for the proteins identified in Early v Late comparisons in Solution 2. CI % = 0.05.

Comparison	α1-collagen 1806	α1-collagen 1821	HSA-3804	αB-crystallin 8102
Cancer v Health				
ACB/NB	0.932		0.237	0.004
TCB/NB	0.090	0.222	0.430	0.0001
GCB/NB	0.495		0.194	0.0003
ACB/GCB	0.506		0.100	0.081
ACB/TCB	0.067	0.222	0.068	0.003
TCB/GCB	0.626	0.222	0.345	0.170
Early v Late				
TCB-E/NB-E	0.052	0.241	0.388	0.008
TCB-E/ACB-E	0.074	0.241	0.049	0.146
TCB-E/GCB-E	0.052	0.241	0.321	0.872
NB-E/ACB-E	0.363		0.519	0.275
NB-E/GCB-E			0.904	0.008
ACB-E/GCB-E	0.363		0.624	0.137
TCB-L/NB-L	0.144	0.350	0.075	0.013
TCB-L/ACB-L	0.073	0.350	0.187	0.002
TCB-L/GCB-L	0.276	0.350	0.083	0.074
ACB-L/NB-L	0.851		0.242	0.043
NB-L/GCB-L	0.724		0.176	0.024
ACB-L/GCB-L	0.560		0.109	0.262
NB-E/NB-L	0.363		0.884	0.860
ACB-E/ACB-L	0.245		0.963	0.436
TCB-E/TCB-L	0.146	0.282	0.021	0.565
GCB-E/GCB-L	0.212		0.205	0.187

Appendix D5: Student T-Test results for the proteins identified in Early v Late comparisons in Solution 3. CI % = 0.05.

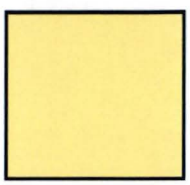
			ACB					TCB					GCB		
Solution 2	**1/N**	**2/N**	**3/N**	**1v2**	**2v3**	**1/N**	**2/N**	**3/N**	**1v2**	**2v3**	**1/N**	**2/N**	**3/N**	**1v2**	**2v3**
HSA - 501	0.162	**0.018**	0.433	0.296	**0.028**	0.309	0.920	0.978	0.053	0.835	0.307	0.527	0.341	0.665	0.647
Lumican - 1805	**0.049**	0.058	**0.024**	0.716	**0.030**	**0.024**	0.055	0.078	0.087	0.727	0.262	0.119	0.077	0.547	0.894
Proapolip. -2101	**0.023**	0.119	0.089	0.141	0.716	0.211	0.308	0.283	0.572	0.797	0.083	0.566	**0.039**	0.570	0.445
Lumican - 2802	0.125	**0.024**	**0.012**	0.224	0.068	0.335	**0.013**	**0.012**	0.530	0.828	0.071	**0.033**	0.195	0.618	0.617
α1-coll. VI - 2804	0.356			0.356			0.130	0.363	0.130	0.646					
Ig-l Chain - 5202	0.300	0.887	0.665	0.323	0.574	0.966	0.614	0.144	0.685	0.167	0.136	0.784	0.295	0.573	0.647
Cofilin - 7003	0.356	0.341	0.423	0.384	0.593	0.500	0.090	0.179	0.690	0.253	0.317	0.392	0.391	0.764	0.828
Cofilin - 7004	0.362	0.419	0.166	0.972	0.341	0.166	0.384	0.166	0.251	0.251	0.250	0.434	0.346	0.516	0.902
Solution 3															
α1-coll. VI - 1806	0.771	0.919	0.341	0.852	0.337	0.587	0.423	0.236	0.750	0.443	0.426	0.783	0.341	0.535	0.423
α1-coll. VI - 1821							0.082	0.356	0.082	0.421					
HSA-3804	0.527	0.215	0.082	0.677	0.165	0.082	0.460	0.339	0.046	0.498	0.168	0.082	0.811	0.104	0.391
αB-crystallin- 8102	0.198	0.002	0.013	0.412	0.922	0.000	0.000	0.002	0.337	0.367	0.000	0.025	0.021	0.796	0.968

Appendix D6: Student T-Test results for the staging based data for proteins identified in Early v Late comparisons. CI % = 0.05.

	ACB v TCB	ACB v GCB	ACB v NB	TCB v GCB	TCB v NB	GCB v NB	ACBv NNH	ACB v NWH	TCB v NNH	TCB v NWH	GCB v NNH	GCB v NWH	NNH v NWH
Solution 2													
4402	0.214	0.351	0.275	0.536	0.059	0.082	0.168	**0.047**	0.326	**0.034**	0.333	**0.036**	**0.033**
8307	0.100	0.114	0.052	0.736	0.793	0.459	0.893	**0.006**	0.265	0.182	0.309	**0.012**	0.133
Solution 3													
703	0.840	0.228	0.298	0.209	0.273	0.893	0.219	0.144	0.212	0.327	0.408	0.125	0.188
1201	0.854	0.651	0.791	0.828	0.698	0.516	**0.007**	0.542	**0.032**	0.706	**0.014**	0.856	**0.032**
1204	**0.043**	**0.003**	0.741	**0.011**	0.063	0.061	**0.003**	0.321	**0.011**	0.151	N/A	0.056	0.056
2402	0.170	0.103	**0.025**	0.690	**0.018**	**0.016**	0.518	**0.033**	0.447	**0.027**	0.430	**0.026**	0.214
7303	**0.032**	0.279	**0.024**	0.158	0.646	0.116	**0.009**	0.055	0.193	0.867	**0.039**	0.273	0.244
8106	**0.007**	**0.004**	**0.005**	**0.021**	0.378	0.104	**0.004**	**0.006**	**0.021**	0.816	N/A	0.103	0.103

Appendix D7: Statistical analysis results from the Healthy with a history (NWH) comparison to the Healthy without a history of breast cancer (NNH). CI % = 0.05.

A guide to plates and figures in this thesis

ACB	Australian Cancerous Breast	
TCB	Taiwanese Cancerous Breast	
GCB	Greek Cancerous Breast	
NB	Normal Breast	
*	Indicates a statistically significant finding	
%TPC	The percentage of total protein content that that protein constitutes.	
NWH	Normal breast tissue with a history of breast cancer	
NNH	Normal breast tissue with no history of breast cancer	