The Development of Nanotechnology- based Detection Systems for the Diagnosis of Breast Cancer

UNIVERSITY *of the*
WESTERN CAPE

Mustafa Drah

A thesis submitted in fulfilment of the requirements for the degree of
Philosophiae Doctor in the Faculty of Science
Department of Biotechnology
University of the Western Cape

Supervisor: Prof Mervin Meyer

June 2015

Abstract

Breast cancer is one of the major causes of death in South Africa. About 1 in 29 South African women are at risk of developing this type of cancer in their lifetime. The global incidence of breast cancer also increases annually with over 1 million new cases diagnosed every year. Molecular diagnostic techniques such as qRT-PCR, Fluorescent In Situ Hybridization (FISH), Immunohistochemistry (IHC) and ELISA are used to diagnose breast cancer. Some of these diagnostic techniques make use organic fluorophores as fluorescent reporter molecules. The principle of all these diagnostic techniques is reliant on the detection of molecular biomarkers that are associated with the disease. In most cases these molecular biomarkers are DNA, RNA or proteins that are up-regulated in response to or as a result of the disease.

The first aim of this study was therefore to identify membrane proteins that are up-regulated in cancers that can potentially be used as biomarkers for the detection of breast cancer. The second aim of this study was to investigate the application of quantum dots in the development of a molecular diagnostic test that can detect a breast cancer biomarker.

The most commonly used method to identify molecular biomarkers for diseases have traditionally been gene expression analysis using technologies such as DNA microarray. These technologies have certain limitations and have therefore not been very successful in identifying useful disease biomarkers. Biomarker

discovery by proteomics can overcome some of these limitations and is potentially a more suitable method to identify molecular biomarkers for breast cancer.

In this study proteomics in combination with Stable Isotope Labelling with Amino Acids in Cell Culture SILAC was used to do a comparative analysis of the expression levels of membrane proteins present in a human breast cancer cell line (MCF-7) derived from a breast cancer patient and a human breast cell line (MCF-12A) derived from a healthy individual. This led to the identification of the transmembrane protein, GFRA1 as potential new biomarker for breast cancer. This study showed that this protein is over expressed in MCF-7 cells as compared to MCF-12A cells and that it is also highly expressed in the myoepthelial cells of the milk ducts of breast cancer patients.

This study also demonstrates the use of molecular beacon technology to develop a DNA probe for the detection of cDNA encoding the CK19 gene, which is a known biomarker for breast cancer. In the development of this probe, quantum dots were used as the fluorescence reporter. This molecular beacon probe was able to demonstrate the over expression of CK19 in MCF-7 cells. This study shows that this technology can potentially be used as a diagnostic test for breast cancer and since quantum dots are used in the development of these molecular beacon probes, this diagnostic test can potentially facilitate the development of multiplex detection systems for the diagnosis of breast cancer. Molecular beacon technology can potentially also be used to detect novel biomarkers such as GFRA1.

Key words

Breast cancer, Biomarkers, SILAC, Proteomics, bioinformatics, q RT-PCR, Nanotechnology, Molecular beacons, molecular diagnostic and Quantum Dots

UNIVERSITY of the
WESTERN CAPE

University of the western cape

Private Bag X 17, Bellville 7535, south Africa

Telephone: + 27-21-959 2255/959/2762

Fax: _ 27-21-959 1268/2266

FACULTY OF NATURAL SCIENCE

Plagiarism declaration

I declare that ***The Development of Nanotechnology- based Detection Systems for the Diagnosis of Breast Cancer*** is my own work, that it has not been submitted for any degree or examination in any other university and that all the sourcess I have used or quoted have been indicated and acknowledged by complete reference

Mustafa Drah **June 2015**

Signature:…….

Acknowledgments

I would like to express my sincere gratitude for the valuable assistance I received through my PhD journey from the following people:

Above all, I would like to acknowledge my Creator for giving me the faith, strength, good health, wisdom and giving me the strength and perseverance, without him I would not have finished this thesis.

- I declare my greatest Gratitude goes to my supervisor Prof. Mervin Meyer for allowing me the opportunity to carry out this study in his laboratory and for his helps, support, advices, encouragement and guidance from the beginning of my PhD study until now, and I acknowledge him for proofreading my thesis.

- I would like to thank Mr. Faghri February with his help with bioinformatics and q-PCR interference and analysis the results.

- I would like to acknowledge Dr. Zukile Mbita for his help with cell culture; protein extraction, 1 SDS PAGE and Western blot assays.

- I would like to thank Dr. Francois Taute with his help with conjugation of the nanoparticles.

- Also I would like to thank Dr. Ashley Pretorius for his help.

- I would like to thank the following people: Dr. Joyti Sharma, Dr. Nicole Sibuyi, Dr Ntevheleni Thovhogi, Miss Lauren Swartz, Miss Taahirah Boltman and Mrs Aml Jaballa for their support and motivation.

- I would like also to thank Prof. Mongi Benjeddou, Dr. Abram Madiehe and Prof. Martin Onani for their advices.

- My appreciation goes to Prof. Mohammed A. Elfagieh (National Cancer Institute Misurata, Libya) for providing the breast cancer tissues.

- I would like to thank Dr. Ruben Cloete and Mr. Mario Jonas with their help with bioinformatics.

- I would also like to acknowledge everyone in our lab (DST/Mintek NIC Lab) for his or her support during my study.

- I would like to thank Mrs. Bridget Daniels and Mrs. Elosie Braaf for their assistance in my project.

- My sincere gratitude goes to Libyan government and DST/Mintek NIC for funding and financial assistance for my project.

- I would also like to thank my parents for the support they provided me through my entire life and in particular, I must acknowledge my wife for her encouragement and editing assistance and I would to thank my brothers and my sister.

For my Parents and my family

Conferences and workshops contributions

1- **Mustafa Drah,** Paul Mushonga, Francois Taute, Martin Onani, Bongani Ndimba and Mervin Meyer The development of nanotechnology-based detection system for the diagnosis of breast cancer "at the 4th DST/Mintek NIC workshop 22-23 November 2011 at Minteck in Johannesburg, South Africa. **Oral Presentation.**

2- **Mustafa Drah,** Paul Mushonga, Francois Taute, Martin Onani, Bongani Ndimba and Mervin Meyer. The development of nanotechnology-based detection systems for the diagnosis of breast cancer, the 4th International Conference on Nanoscience and Nanotechnology 1-4 April 2012 at University of the free state, Bloemfontein, South Africa. **Poster Presentation.**

3- **Mustafa Drah,** Paul Mushonga, Francois Taute, Martin Onani and Mervin Meyer. The development of nanotechnology-based detection system for the diagnosis of breast cancer using Qdots, at the BROD-Biotechnology research day 4 September 2012 at University of the Western Cape. **Oral Presentation.**

4- **Mustafa Drah,** Paul Mushonga, Francois Taute, Martin Onani and Mervin Meyer. The application of Molecular beacon for the diagnosis of breast cancer, at the 5th DST/Mintek NIC workshop 21-22 September 2012 at MRC in Cape town, South Africa. **Oral Presentation (2nd Prize).**

5- **Mustafa Drah**, Paul Mushonga, Francois Taute, Martin Onani, Faghri February, Abram Madiehe, and Mervin Meyer. The development of nanotechnology-based detection systems for the diagnosis of breast

cancer, the 2nd Biotechnology world congress 18-21 February 2013 at Dubai Women's College. Dubai**. Poster Presentation (2nd Prize).**

6- **Mustafa Drah**, Paul Mushonga, Francois Taute, Martin Onani, Faghri February, Abram Madiehe, and Mervin Meyer. The application of quantum dots in the diagnosis of breast cancer " at the BROD-Biotechnology research open day 30 October 2013 at University of the Western Cape. **Poster Presentation (1st Prize).**

7- **Mustafa Drah**, Paul Mushonga, Francois Taute, Martin Onani, Faghri February, Abram Madiehe, and Mervin Meyer. The development of nanotechnology-based detection systems for the diagnosis of breast cancer" at the ninth National cancer research institute NCRI 3-6th November 2013 at the BT convention centre, Liverpool, UK. **Poster Presentation.**

8- **Mustafa Drah**, Martin Onani, Faghri February and Mervin Meyer. The development of quantum dot-based in situ hybridization techniques, MAM-14 international symposium on macro-and supramolecular architectures and materials" 23-27 November 2014, Peermont hotels casino resorts, Johannesburg, South Africa. **Poster Presentation.**

TABLE OF CONTENTS

List of figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

List of tables

Chapter 1

Chapter 2

Chapter 3

List of abbreviation

Abbreviation	Full Name
%	Percentage
°C	Degree Celsius or degree centigrade
~	Approximately
∞	Infinity
≥	Greater than or equal to
3′	3 primer
5′	5 primer
A	Adenosine – Adenine
ACS	American Cancer Society
APS	Ammonium persulphate
AT	Annealing temperature
ATCC	American Type Culture Collection
AuNP	Gold Nanoparticle
BLAST	Basic sequence alignment search tool
bp	Base pair
BSA	Bovine serum albumin

C	Cytosine
cDNA	Complementary deoxyribonucleic acid
CdTe/ZnS	Cadmium telluride zinc sulphide
cm^2	Square centimeter
CO_2	Carbon dioxide
C_p	Crossing point
C_t	Threshold cycle
dH_2O	Distilled water
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
DTT	Dithiothreitol
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride
EDTA	Ethylene diamine tetra acetic acid
EGF	Epidermal growth factor
F P	Forward primer
FBS	Fetal Bovine serum
FDA	Food and Drug administration

FFPE	Formalin-fixed, paraffin-embedded
FISH	Fluorescence *In situ* hybridization
FISH	Fluorescent in situ hybridization
G	Guanosine-Guanine
GSH	L-Glutathione
h	Hour
IHC	Immunohistochemistry
InP/ZnSe	Indium phosphide zinc selenide
Kb	Kilo-base pair
KCL	Potassium chloride
kDa	Kilo Dalton
LC-MS/MS	Liquid chromatography-tandem mass spectrometry
M	Molar
MB	Molecular beacon
mg	Milligram
min	Minute
ml	Millilitre
mm	Millimeter

mM	Millimolar
MNPs	Magnetic nanoparticles
MOPS	4-Morpholine propanesulphonic acid
MRI	Magnetic resonance imaging
mRNA	Messenger ribonucleic acid
MS	Mass spectrometry
NaOH	Sodium Hydroxide
NCBI	National Centre for Biotechnology Information
ng	Nanogram
nm	Nanometer
nM	Nanomolar
PBS	Phosphate Buffer Slain
PFA	Paraformaldehyde
PVDF	polyvinylidene difluoride
Qdot	Quantum dots
qRT-PCR	Quantitative Real time polymerase chain reaction
RCF	Relative Centrifugal Force or gravitational
RNA	Ribonucleic acid
RNase	Ribonuclease

RP	Reverse primer
RPM	Revolutions per minute
RT	Room temperature
SAGE	Serial Analysis of Gene Expression
SCC	Saline sodium citrate
SDS	Sodium dodecyl sulphate
SDS PAGE	Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
Sec	Second
SILAC	Stable isotope labelling with amino acids in cell culture
T	Thymidine – Thymine
TBETween 20	Tris Buffered Saline containing Tween20
TBS	Tris Buffered Saline
TCEP	tris(2-carboxyethyl)phosphine
TEME	*N,N,N',N'*-Tetra methylethylene-diamine
Tris	Tris (hydroxymethyl) aminoethane
U	Unit
US	Ultrasound
UV	Ultraviolet

V	Volt
v:v	Volume to volume
W:v	Weight to volume
x	Times
µ	Micro
µg	Microgram
µL	Microliter

A list of biomarkers Abbreviations

Abbreviation	Full name
BRCA2	Breast cancer 2, early onset
c-Myc	V-Myc Avian Myelocytomatosis Viral Oncogene Homolog
CK19	Cytokeratin 19
FOXA1	Fork head box A1
ER	Estrogen receptor
PR	Progesterone receptor
PIK3CA	Hosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit ALPHA
PARP	Poly (ADP-ribose) polymerase
CA15-3	CANCER ANTIGEN 15-3
CA27-29	CANCER ANTIGEN 27-29

CHAPTER 1

1.1 Cancer

Cancer is a disease caused by the uncontrollable multiplication of groups of cells that arise from genetic alterations that affect tightly controlled regulatory systems required for the control of cell growth (Pietras *et al.,* 2010). All cancers arise as a result of changes that have occurred in the DNA sequence which results from a stepwise accumulation of genetic alterations within a cell (Stratton *et al.,* 2009). These alterations lead to abnormal proliferation and expansion, and ultimately invasion of surrounding tissues (Nass *et al.,* 1997, Antoniou *et al.,* 2003).

Six key features or hallmarks characterize cancer cells (Hanahan *et al.,* 2000). These hallmarks are 1) cancer cells can stimulate their own growth, 2) cancer cells can bypass growth suppressors, 3) cancer cells can evade apoptosis, 4) cancer cells can multiply indefinitely, 5) cancer cells develop excessive angiogenesis systems and 6) cancer cells can invade surrounding tissues (Hanahan *et al.,* 2000). Cancer pathophysiology can either be hereditary or triggered by environmental factors such as smoking, diet and exposure to ultraviolet radiation (Croce 2008). Genetic mutations in proto-oncogenes and tumour suppressor genes can lead to cancer (Croce 2008). Proto-oncogenes are involved in cell signalling processes under normal physiological conditions. However, these genes can be activated by genetic mutations to become oncogenes, which contributes to the development of cancer. Oncogenes can also be activated by the DNA de-methylation, which could lead to a dosage effect, i.e. producing an additional functional copy of the gene,

which in turn can also lead to cancer growth (Stratton *et al.,* 2009, Muthusamy *et al.,* 2010). Several oncogenes have been identified, such as growth factor receptors common in breast cancer, for example human epidermal growth factor 2 (*HER-2/neu*), Myelocytomatosis Viral Oncogene Homolog (*c-Myc*), Rat sarcoma (*RAS),* Mitogen-activated protein kinase *(ERK)* and (B-cell lymphoma 2) *Bcl-2* (Croce 2008).

Tumour suppressor genes suppress tumour formation. These genes inhibit cell proliferation and cell division during the cell cycle allowing time for the correction of mistakes that may occur during DNA replication (Chen *et al.,* 1998). However, mutations can inactivate the tumour suppressor genes, which can lead to the development of cancer. Tumour suppressor genes can also be inactivated by DNA methylation and histone deacetylation to render them non-functional (Deng *et al.,* 2000). Examples of tumour suppressor genes are Retinoblastoma gene (*RB*), Protein of 53 kilodaltons (kDa) (*p53*) and *INK4a* is a protein of 16 kDa ($p16^{INK4a}$) (Sherr 2004), breast cancer 1(*BRCA1*) and breast cancer 2 (*BRCA2*) (Hall *et al.,* 1990).

Cancer is considered to be an important global health problem because it affects millions of people all over the world (Xue *et al.,* 2008). It is predicted that cancer will become the main cause of morbidity and mortality in the coming decades all over the world (Bratu *et al.*, 2011). Cancer accounts for more than six million deaths a year worldwide, with ten million new cases are diagnosed each year (Xue

et al., 2008). Studies showed that in 2008, an estimated 196,3 million healthy life-years were lost worldwide because of cancer (Soerjomataram *et al.,* 2012).

In the United States of America (USA) alone, a quarter of the mortality tolls are attributed to cancer (Peng *et al.,* 2010). Cancer is recognised as the second leading cause of death in USA after heart diseases (American Cancer Society. 2014). In 2014, approximately 585,720 American people died from cancer (American Cancer Society. 2014). One in three women and one in two men in the USA are expected to develop cancer in their lifetime (Medley 2007). The total cancer incidence in the USA recorded between 2004 and 2008 were 1,638,910 cases (Siegel *et al.,* 2012). This included 226,870 cases for breast cancer, 12,170 cases for uterine cancer, 143,460 cases for colon and rectum cancer and 241,740 cases for prostate cancer (Siegel *et al.,* 2012). The most common cancer among women world wide is breast cancer (Agarwal *et al.,* 2009), which is also the second leading cause of cancer deaths in women (Gauger *et al.,* 2014). Therefore breast cancer is considered one of the most important diseases amongst women worldwide (Gast *et al.,* 2009).

Cancer is also a major problem in developing countries and the second most common cause of death (Xue *et al.,* 2008). Although some statistics are available for South Africa, the incidence of cancer in African countries is not very well studied and documented. The 2008 South African National Cancer Registry reported that the top cancers in South Africa was cancers of the breast, cervical,

bladder, stomach, oesophagus and colorectal (Cancer Association of South Africa. 2014).

1.2. Breast cancer

The female breast consists of lobes, lobules and milk ducts (Figure 1.1). The breast consists 15 to 20 lobes, each one made up of many smaller lobules, which are the milk producing mammary glands. The lobules are connected through milk ducts. The breast ducts are composed of the basement membrane and a layer of luminal epithelial and myoepithelial cells, while the surrounding stroma include leukocytes, fibroblasts, myofibroblasts, and endothelial cells (Polyak 2007). Breast cancer is a malignant disease that originates from cells in the breast tissue. One model for breast cancer suggests that the myoepithelial cells are genetically altered causing a decrease in the number of these cells, while the number of stromal fibroblasts, myofibroblasts, lymphocytes, and endothelial cells increases.

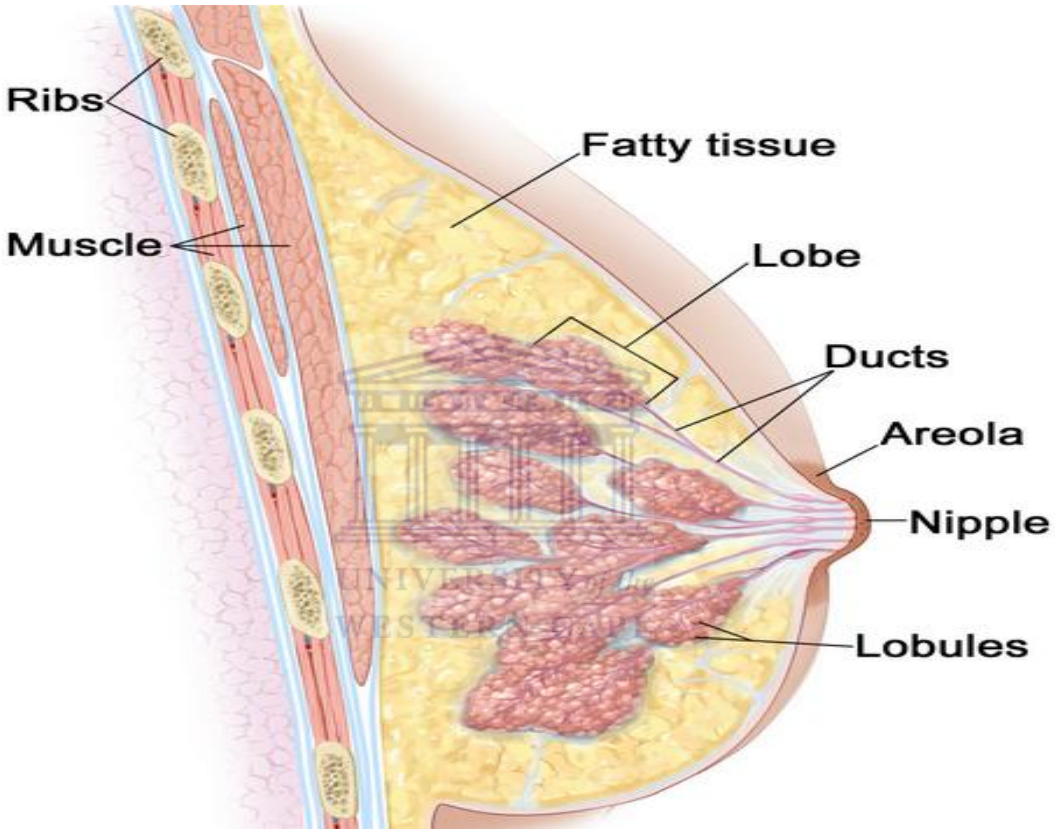


Figure 1. 1: The anatomy of the female breast. The lobes, lobules and ducts, are indicated (Winslow. 2011).

In terms of histology, therapeutic response and patient outcomes, breast cancer is considered a highly heterogeneous disease (Prat *et al.,* 2011, Network 2012). Breast cancer is an adenocarcinoma that is classified as either a ductal or lobular carcinoma (Hammer *et al.,* 2008) since these two structures are the primary origins of the malignancy. Most breast cancer cases arises from luminal epithelial cells (Fang *et al.,* 2009) which makes up the milk production glands and ducts that connect to the nipple (Prat *et al.,* 2011, Network 2012). Ductal and lobular carcinoma can be further classified into other pathological subgroups such as in situ or invasive (Hammer *et al.,* 2008).

1.2.1 Prevalence of breast cancer

Breast cancer is the second leading cause of death after lung cancer in the USA (Jemal *et al.,* 2007) and remains one of the leading causes of cancer-related deaths in the Western world (Kuhl *et al.,* 2007). It is responsible for over 500 000 deaths per year worldwide (Königsberg *et al.,* 2011). The incidences of breast cancer is also increasing in developing countries such as South Africa (Agarwal *et al.*, 2009). One in nine women are at risk of developing breast cancer (Callesen *et al.*, 2008).

In the USA, breast cancer affected about 232,620 women and men in 2011 and was responsible for 39,970 deaths in the same year (Siegel *et al.*, 2011). In 2014, it was estimated that about 232,670 new cases of breast cancer were diagnosed among the women in the USA, making it the most frequently diagnosed cancer in women (Siegel *et al.*, 2014). In Australia, about 2 % of the women between the

ages of 20 and 34 years and 11 % of women between the ages of 35 and 44 years are affected by breast cancer (Hickey *et al.*, 2009). It was reported by Ferlay *et al.* (2007) that breast cancer accounted for 429,900 cases of cancer in European women in 2006 (Ferlay *et al.*, 2007). In South Africa, breast cancer is the most common type of cancer amongst women, affecting about 16.6 % of women (Vorobiof *et al.*, 2001). According to the 2008 South African National Cancer Registry, the number of the female breast cancer patients in the different population groups were as follows: 2,884 cases amongst blacks, 1,976 cases amongst whites, 834 cases amongst coloureds and 330 cases amongst Asians (Cancer Association of South Africa. 2015).

1.2.2 Types of breast cancer

Upon diagnosis, breast cancer is classified primarily by the histological appearance of the cancer cells. There are two main types of breast cancer: carcinoma in situ and invasive (infiltrating) carcinoma. These two types of breast cancer will be briefly described below.

1.2.2.1 Breast carcinoma in situ

Breast carcinoma in situ (CIS) is further classified into either ductal or lobular CIS (Figure 1.2) based on cytological features (Malhotra *et al.*, 2010). Ductal carcinoma in situ (DCIS) is one of the most common types of non-invasive cancer. This type of cancer is defined when the cancer cells are present within the ducts, but have not yet spread and infiltrate the surrounding breast tissues (Bravaccini *et al.,* 2013). DCIS affects the basement membrane of the breast

myoepithelium. The neoplastic cells proliferate and accumulate within the milk ducts. In the healthy breast, the ducts have a single epithelial layer, however, during development of DCIS neoplastic cells grow uncontrollably into a lesion resulting in multiple layers that accumulate inside the ducts. The breast stroma including the extracellular matrix, lymphatics, blood vessels, stromal cells, immune cells and fat cells can either promote or suppress the carcinogenic process by responding to secretory signals due to oxidative stress and nutrient deprivation that arise from the tumour cells accumulating within the duct (Espina *et al.*, 2011). If not treated, DCIS may progress at a later stage to invasive ductal carcinoma (IDC), but the mechanism of this transition is not yet well understood (Sue *et al.*, 2013).

1.2.2.2 Invasive (or infiltrating) carcinoma

Like DCIS, IDC originates from the milk ducts then spreads through the walls of the ducts to surrounding breast tissues (Johnson *et al.*, 2012). IDC is the most common type of invasive breast cancer, accounting for 72 to 80 % of all invasive breast cancers (Arps *et al.*, 2013) and accounts for 8 to 14 % of all breast cancer cases (Malhotra *et al.*, 2010). Invasive carcinoma is divided into six categories, that includes tubular, ductal lobular, infiltrating ductal, mucinous, medullary and invasive lobular carcinoma (ILC) (Figure 1.2).

ILC also known as infiltrating lobular carcinoma is a major invasive tumour type that originates from the lobules and compared to IDC is more likely to be positive for hormone receptors (Arpino *et al.,* 2004). ILC is characterized by a general

thickening of an area of the breast, usually the section above the nipple and toward the arm. This type of cancer is difficult to visualize by mammography (Cao *et al.,* 2012), and if not treated within 3 years after disease diagnosis, the diseased cells can spread to different parts of the body such as the bones, lungs, liver and brain (Weigelt *et al.*, 2005, Fernandez *et al.*, 2013, Switzer *et al.,* 2014).

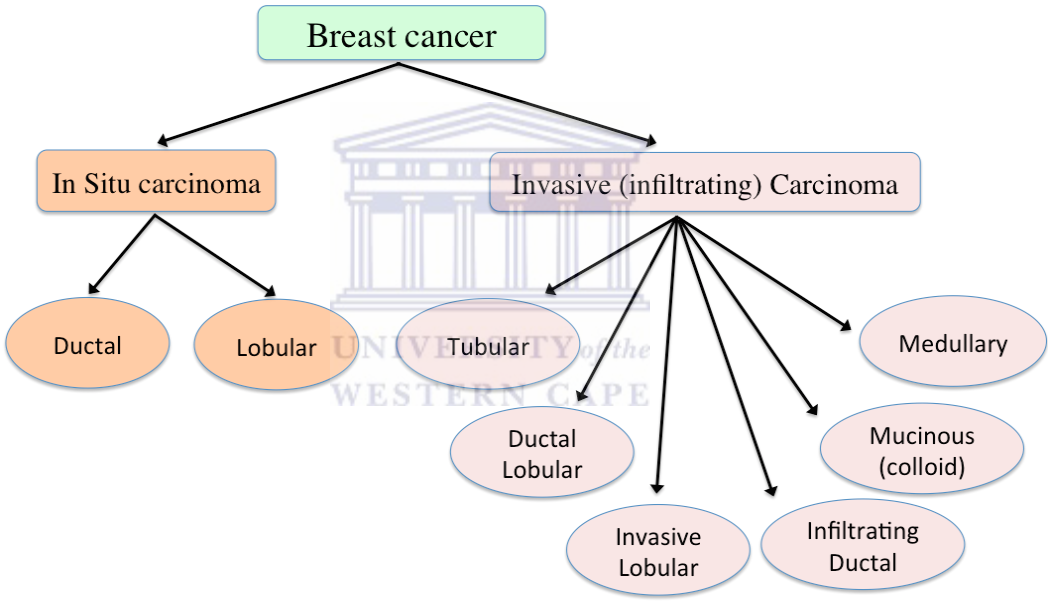


Figure 1. 2: The classification of breast cancers. This classification based on Nuclear Pleomorphism, Glandular/Tubule Formation and Mitotic (Malhotra *et al.,* 2010).

Distant metastasis of the cancer is common in breast cancer patients during the late stages of the disease (Weigelt *et al.,* 2005). Distant organ invasion begins when the primary subpopulations of breast tumour cells acquire metastatic ability through acquired somatic mutations. These mutations determine the site where the tumour cells are likely to invade. Several prognostic markers have been studied to identify breast cancer patients at risk of disease metastasis; the markers include the tumour size, angioinvasion, gene and protein profiling. For instance, patients with tumours that are 2-5 cm in size and have tumour emboli in more than three blood vessels are at a higher risk of metastasis (Weigelt *et al.*, 2005).

Metastasis of cancer into other tissues or organs occurs when the cancer cells migrate into the circulatory blood system and once arrested by capillaries of the distant tissues the cells invades the tissue and starts to proliferate and grow into a new tumour. The cancer cells are depended on angiogenesis to spread from its original location to other sites in the body and for a continuous supply of oxygen and nutrients to grow and multiply (Weigelt *et al.*, 2005).

The bones and the lungs are the primary targets in breast cancer metastasis (Weigelt *et al.*, 2005). Once spread to the bones, the patients suffer from extreme pain and have increased risk of fractures (Jimenez-Andrade *et al.*, 2010). When the cancer cells get into the lungs, these patients can have difficulty with breathing, chest tightness, chronic cough, pleural effusion, loss of appetite and weight loss (Temel *et al.*, 2007).

1.2.3 Breast cancer treatment

1.2.3.1 Surgery

Surgery is usually the first line of treatment, however, mandatory radiation or chemotherapy and hormonal therapy is required thereafter (Collins *et al.*, 2011, Li *et al.*, 2011). The complete resectioning of the cancerous tissue is a standard approach to treat breast cancer (Ziogas *et al.*, 2009). A large number of breast cancer patients undergo mastectomy, a surgical procedure that involves the removal of the whole breast (Coopey *et al.*, 2013). In the event that the patient is diagnosed at an early stage, removal of the cancerous tissue as opposed to the whole breast is preferred (Osman *et al.*, 2013). This may be achieved using breast-conserving surgery such as lumpectomy, partial mastectomy and modified radical mastectomy. Lumpectomy is a type of surgery that entails the removal of a tumour (lump) and some of the normal tissue surrounding the tumour (Sabel *et al.*, 2009). Partial mastectomy is a procedure used in patients with invasive breast cancers. It entails the removal of a section of the breast that is cancerous together with some of the normal tissue surrounding it (McCahill *et al.*, 2012). Whereas modified radical mastectomy entails removing the entire breast and its tissues (Loukas *et al.*, 2011). Current strategies employed in the treatment of breast cancer are associated with adverse health effects.

Sentinel lymphatic drainage and axillary node dissection

Lymphatic drainage and dissection is often performed in breast cancer patients when the cancer has metastasized into sentinel lymph node (SLN) and axillary lymph nodes (ALN) (Mansel *et al.*, 2006, Del Bianco *et al.*, 2008). Lymphatic

drainage follows post surgery to relieve chronic lymphatic oedema in patients; the technique uses manual massage to redirect the blood flow towards healthy areas (Martín *et al.,* 2011).

The SLN is the first lymph node to receive lymphatic drainage from a tumour and is the most likely route a primary tumour will migrate from (Mansel *et al.*, 2006, Del Bianco *et al.*, 2008). The ALN dissection is usually performed in patients with invasive breast cancer. It is sometimes used as a prognostic factor to determine the stage of the breast cancer (Smeets *et al.*, 2013). This procedure however, has been associated with short and long term side effects, such as in lymphedema, nerve injury, and shoulder dysfunction (Krag *et al.*, 2010). SLN biopsy is considered an accurate method for the detection of axillary metastasis in patients diagnosed with breast cancer at an early stage (Jonjić *et al.*, 2012).

1.2.3.2 Radiation therapy

Irradiation therapy can reduce the risk of cancer recurrence in patients who had breast-conserving surgery followed by radiotherapy (Buchholz 2009). This in turn can reduce the mortality of breast cancer after breast-conservation surgery. A drawback of irradiation therapy is the toxic effects of radiation, but this is depended on the course and the dose of radiation (Whelan *et al.*, 2010).

Radiation therapy can also increase the rate of cardiovascular disease in ageing women (Wood *et al.*, 2001). This therapy should also not be given to patients suffering from systemic lupus erythematosus, scleroderma and pregnant women,

especially during the first trimester of their pregnancy (Buchholz 2009, Whelan *et al.*, 2010).

1.2.3.3 Chemotherapy

Chemotherapy is a treatment administered to cancer patients, which is aimed at killing rapidly multiplying cells, which include cancer cells (Smith *et al.*, 2002). Chemotherapy is considered to be useful treatment for older women, however, the side effects associated with chemotherapy, such as vomiting, hair loss and nausea has a negative impact on the patients' quality of life (Niikura *et al.*, 2013). Therefore, chemotherapy is beneficial to premenopausal women suffering from breast cancer (Dellapasqua *et al.*, 2005). Niikura *et al.* (2013) reported that previous studies found chemotherapy, as an adjuvant therapy, in women with breast cancer as the preferred treatment. However, younger women preferred chemotherapy as compared to older women (Niikura *et al.*, 2013). Recently neoadjuvant therapy has been considered as an approach for the treatment of breast cancer in the early stages of the disease in order to increase the rate of breast-conserving therapy and also as a means to reduce the extent of surgery (Gampenrieder *et al.*, 2013).

1.2.3.4 Endocrine therapy

Endocrine therapy (hormone therapy) reduces the disease–related mortality and improves disease-free survival in patients with an early stage of breast cancer (Romond *et al.*, 2005). Niikura *et al.* (2013) reported that endocrine therapy as opposed to chemotherapy was the preferred treatment option (Niikura *et al.*,

2013). For women with hormone receptor-positive breast cancer, adjuvant endocrine therapy is often advised, such as tamoxifen treatment which reduces breast cancer recurrence and mortality in patients with hormone receptor-positive cancer (Connor *et al.*, 2013).

1.2.3.5 Targeted therapy

Targeted therapy refers to a drug or other substance that is directed to cancer-specific markers using molecules that will recognize the cancer cells; and in so doing block their growth and spread without harming the healthy cells (Nahta *et al.*, 2006). Some of the targeting moieties used in the treatment of breast cancer include monoclonal antibodies, tyrosine kinase inhibitors and poly (ADP-ribose) polymerase (PARP) inhibitors. Several targeted therapies approved by the Food and Drug administration (FDA) include monoclonal antibodies such as trastuzumab (Herclon, Herceptin) and pertuzumab patients (Mohamed *et al.*, 2013). Both therapies block the action of the *HER-2*/neu growth factor, which is known to promote the growth of breast cancer cells. However, pertuzumab is used as a combinatorial therapy in metastasized breast cancer patients (Mohamed *et al.*, 2013). PARP is a group of enzymes encoded by the *PARP1* gene that detects and interferes with DNA repair, and is involved in base excision repair and repairing of double stranded DNA breaks, which is also one of the most important aspects in the treatment of cancer (Weil *et al.*, 2011, Davar *et al.*, 2012). Tyrosine kinase inhibitors act by blocking tyrosine kinases, which plays a critical role in growth factor signalling (Arora *et al.*, 2005).

1.2.3.6 Nanotechnology-based treatments for breast cancer

Nanoparticles ranging in the sizes between 1 and 100nm have physicochemical properties that renders them useful in medicine (Yang *et al.*, 2013). Nanotechnology has shown a great promise in the treatment and diagnosis of cancer. Nanomaterials capable of delivering chemotherapeutic agents directly to the breast cancer tissues bring hope in the fight against breast cancer. Due to their small size, nanoparticles have large surface area. Smaller nanoparticles have an increased loading capacity that can be manipulated for biological use. The nanoparticles can be used in multi-modal systems by attaching targeting, diagnostic and therapeutic agents. By targeting disease associated markers it is possible to minimize bystander effects to healthy tissues and increase drug efficacy at lower dosages (Tanaka *et al.*, 2009, Tharkar *et al.*, 2014).

Nanotechnology-based strategies for breast cancer theranostics are under extensive study and development to overcome toxicity issues associated with conventional chemotherapeutic agents and to design early diagnostics assays. Nanoparticles as drug delivery agents can be personalised by targeting specific or diseased tissues, thereby reducing the dosage and frequency of drug treatment (Saadeh *et al.*, 2014).

In breast cancer diagnostics, magnetic nanoparticles (MNPs) and semiconductor nanoparticles are employed as contrast agents for the detection of primary and metastatic tumours at the molecular level and assess the disease stage (Ahmed *et al.*, 2013). Detection of tumours that has accumulated the nanoparticles can be

achieved by traditional techniques such as MRI. Multiplexing is possible as different sized nanoparticles targeting different markers can be used at once to determine the position of the cancer (Ahmed *et al.*, 2013)

Furthermore, gold nanoparticles (AuNPs) have been widely studied for development of clinical diagnostic applications. The properties of AuNPs has shown great promise to be used in medical field to develop Point-of-care (POC) diagnostic systems for the specific detection of DNA and RNA sequences and immunoassays (Baptista *et al.*, 2008).

Another successful application of nanoparticles is the use of liposomes as cancer drug delivery vehicles. Liposomes are the first drug-carrying nano-carriers to reach cancer clinics (Markovsky *et al.*, 2012). They are self-assembled closed colloidal structures composed of concentric phospholipid bilayers surrounding a central aqueous core (Parveen *et al.*, 2012). They offer simultaneous loading of hydrophobic (non-polar molecules/drugs) into the lipid bilayer while hydrophilic (polar molecules/drugs) can be encapsulated in the aqueous core allowing a variety of therapeutic cargo (e.g. anti-cancer drugs, DNA, peptides, vaccines, enzymes, and imaging agents) to be loaded into this assembly (Perche *et al.*, 2013). They have been reported to be biocompatible, biodegradable, have low immunogenicity and excellent safety profiles in humans. It also offers unlimited therapeutic cargo loading, increased pharmokinetic and pharmodynamic abilities and it is relatively inexpensive for mass production. These factors make liposomes superior as nanoparticle carrier system for therapeutics.

The first liposomal drug to gain approval by the FDA in 1995 was Doxil or Caelyx (Perche *et al.*, 2013), and has been used to treat a wide variety of advanced stage cancers such as ovarian cancer, metastatic breast cancer and AIDS-related Kaposi's sarcoma. Doxil is a PEGylated liposomal formulation of doxorubicin. Other commercialized phase III liposomal doxorubicin liposomes include: Myocet, LipoDox, Thermodox and DaunoXome. Drugs such as Vincristine (Marqibo), Paclitaxel (Lipusu and LEP-ETU), Cisplatin (Lipoplatin) have also been formulated in liposomes and are undergoing phase III clinical trials.

1.2.4. Diagnostic methods for breast cancer

1.2.4.1 Clinical diagnosis

1.2.4.1.1 Breast self-examination

Breast self-examination (BSE) is a screening method used by women to check for any irregularities in the breast tissue (Petro-Nustas *et al.*, 2013). BSE is a physical examination of the breast using fingers to feel or detect any possible lumps in the breast tissue. This technique is simple and allows women to take charge of their own health (Harris *et al.*, 2002). The American Cancer Society (ACS) recommended that women should familiarise themselves with what the normal state of the breast tissue and frequently check for any abnormalities by feeling the breast using the pads of their fingers (Evans, 2012).

1.2.4.1.2 Mammography

Mammography is a screening method that has shown to reduce mortality rates associated with breast cancer when followed up with more superior diagnostic

methods and treatment. With this screening approach it is especially women aged between 39 and 69 years that can benefit (Berg 2009, Brodersen *et al.*, 2010, Nelson *et al.*, 2009). In developed countries, mammography form part of general health services (Weedon-Fekjær *et al.*, 2008). Screening rates vary between woman and appears to be linked to the level of the women's education (Evans, 2012). Screening rates in women without a tertiary education is lower as compared to women with tertiary education, which suggest that this is be linked to the women's socioeconomic circumstances.

Mammography is limited in that it isn't sensitive enough to detect cancer in dense breast tissues (Evans, 2012). A small margin of false-negative and false-positive results has also been associated with mammography which might cause a delay in the diagnosis and treatment of breast cancer (Brodersen *et al.*, 2010). It was also shown that mammography cannot detect all the breast cancer foci (Kuhl *et al.*, 2007). Pataky *et al*, (2013) carried out a comparative analysis on the cost effectiveness of Magnetic Resonance Imaging (MRI) versus mammography for breast cancer diagnosis in BRCA1/2 mutation carriers (Pataky *et al.*, 2013). This study found that the cost of mammography is lower compared to MRI but that it is also less sensitive than MRI.

1.2.4.1.3 Ultrasound

Ultrasound (US) is widely available, inexpensive and generally well received by patients (Kelly *et al.*, 2010). Breast US is primarily used to distinguish between solid masses and cysts in the breast (Nothacker *et al.*, 2009). In developing

countries, US has become popular among lower level health centres (Gonzaga 2010) and is more sensitive than mammogram especially in patients with dense breast tissue (Madjar 2010, Luparia *et al.*, 2013). US are more superior to mammography when evaluating the tumour size (Keune *et al.*, 2010) and is an low cost method for evaluating palpable breast cancer (Gonzaga 2010).

1.2.4.1.4 Magnetic Resonance Imaging (MRI)

MRI is a technique mainly used for the assessment of complex lesions found in women with a high risk of developing breast cancer (Warner 2011). MRI as a screening tool for breast cancer and has shown a higher degree of sensitivity as compared to mammography (Bryce *et al.*, 2000, Saslow *et al.*, 2007). It provides more information about the breast tissue vascularity that cannot be obtained from mammography (Liberman 2004). False-negative and false-positive results after MRI screening can be attributed to inherent technology limitations of MRI, human error and/or the patient's characteristics (Saslow *et al.*, 2007).

1.2.4.2 Molecular diagnostic methods

The molecular diagnosis of cancer is based on the detection of molecular changes that is associated with the disease and depends on the identification of biomarkers that is associated with these molecular changes. Biomarkers refer to molecules that are used as an indicator of the biological state, behaviour and function of the cells (Strimbu *et al.*, 2010). Pathologically significant biomarkers include the expression or altered expression of genes and gene products, proteins and lipids (Nie *et al.*, 2007). A number of different molecular diagnostic methods can be

used to evaluate the presence of such biomarkers in a patient sample. This includes techniques such as quantitative real time Polymerase Chain Reaction (qRT-PCR), enzyme-linked immunosorbent assay (ELISA), Fluorescent in situ hybridization (FISH) and Immunohistochemistry (IHC). For the purpose of this discussion, only FISH, IHC and qRT-PCR will be highlighted.

1.2.4.2.1 Fluorescent *in situ* hybridization

Fluorescent in situ hybridization (FISH) is a cytogenetic technique used to detect the physical location of specific gene or specific nucleic acid sequences (DNA or RNA) in intact chromosome and to assess if multiple copies of a specific gene or nucleic acid sequence (e.g. *HER-2* gene for breast cancer) is present (Tanke *et al.*, 2005). Using fluorescence microscopy, FISH allows for the identification of chromosomal abnormalities for a given condition. This is based on the use of fluorescent probes having a complementary sequence to that of the target gene. Three probes are generally used in FISH, to detect either the whole-chromosome painting or repetitive and locus-specific sequences of a particular gene (Tanke *et al.*, 2005, Bishop 2010). The presence or absence of the disease is assessed based on the fluorescent reporter molecule attached to the complementary nucleic acid in the tissue section.

In breast cancer patients, this technique can distinguish between malignant and benign pigment lesions (Nijhawan *et al.*, 2012). FISH has been approved by the FDA to evaluate the gene expression of *HER-2*. This is achieved based on scoring systems set by the FDA and American Society Clinical Oncology/College of

American Pathologists (ASCO/CAP). *HER-2* gene expression in breast cancer patients is confirmed by comparing the gene copy number or the ratio between gene and the centomeric region of chromosome 17 (CEP17). Patients with *HER-2* copy number > 4 or *HER-2*/CEP 17 > 2 are considered to have breast cancer. If the ratio between *HER-2*/CEP17 is < 1.8 the result is negative, whereas the score of 1.8 to 2.2 suggest a borderline case (Brunelli *et al.,* 2008, Sapino *et al.*, 2013). In primary breast cancer tissue, FISH can also be used to evaluate the expression of the estrogen receptor (ER) and the progesterone receptor (PR) (Ma *et al.*, 2013).

The fluorescent reporter molecules used in these FISH is usually an organic fluorophores such as texas red, fluorescein, etc. Herein, also lie the limitations of FISH in that organic fluorophores are sensitive to photobleaching and limits multiplexing possibilities. Photo-instability of organic fluorophores causes photobleaching, which is the fading of the fluorescent signal over time and can negatively affect the FISH result. Organic fluorophores typically also have large emission spectra, which limits the application of multiple organic fluorophores to detect multiple targets/biomarkers, since the potential for spectral overlap is high for organic fluorophores.

1.2.4.2.2 Immunohistochemistry (IHC)

Immunohistochemistry (IHC) has an important role in the medical field, especially in pathology for the diagnosis of diseases (Liu *et al.*, 2011). IHC is a standard technique used to detect the expression of proteins using an antibody that

binds to specific molecules present in the diseased cells (Mahmoud *et al.*, 2011). IHC is based on the same principle as FISH; however, it is used to determine protein expression instead of gene expression as is the case in FISH. Monoclonal and polyclonal antibodies, which targets specific antigens/biomarkers within cancerous breast tissue, has been identified (Alam *et al.*, 2013), and are used to determine the protein expression in formalin-fixed, paraffin-embedded (FFPE) tissue samples (Byers *et al.*, 2007).

IHC is commonly used to evaluate the HER-2 status in breast cancer patients (Garrison *et al.*, 2013). The FDA has approved IHC for the assessment of HER-2 protein expression in breast cancer patients, based on the scoring system established by the FDA and ASCO/CAP. The results are based on the ratio of HER-*2* copy number to that of CEP 17. A score between 0 and 1+ indicate that the patient is negative, a 2+ score is weakly positive and considered and considered as borderline case which needs further tests to confirm the disease status. A positive result will have a 3+ score (Brunelli *et al.*, 2008, Sapino *et al.,* 2013). IHC has also been used to analyze the expression of PR and ER in locally recurrent and primary tissue of breast cancer patients (Ma *et al.*, 2013).

However, IHC is the preferred technique because smaller amounts of tissue is required to carry out the test; frozen or fixed tumour tissue samples may be used, the technique is widely available, and the assessment is based on the use of light or fluorescence microscopy. Detection of the antigen can either be done using chromogenic and fluorescence methods. Chromogenic reporters yield an intensely

colored product that can be analyzed with a light microscope. Fluorescence detection methods make use of fluorescent tags such as texas red, fluorescein, etc. Consequently, IHC suffers from similar drawbacks as FISH.

1.2.4.2.3 Quantitative Real Time Polymerase Chain Reaction (qRT-PCR)

Reverse transcriptase assays have become the most commonly used method for characterising gene expression patterns in different sample populations. It is a technique that collects and generates data, in real time, with progressive PCR cycles (Derveaux *et al.*, 2010). It provides quantitative information as each amplicon is doubled by emitting a fluorescent signal, thus combining amplification and detection in a single step (Wong *et al.*, 2005, Guénin *et al.*, 2009). It has become the preferred method due to its high sensitivity, good reproducibility and wide dynamic range (Derveaux *et al.*, 2010). The reliability of qRT-PCR lies in its sensitivity and ability to detect a single copy of a specific transcript. However, some limitations associated with this technique include non-specific amplification, variations in amplification efficiencies and hetero-duplex formation (Pfaffl 2001). Successful qRT-PCR amplification relies on a number of factors, termed indicators of a good qRT-PCR. These include the integrity of RNA, cDNA and the absence of DNase (Fleige *et al.*, 2006). These steps are crucial as it is most likely to introduce variations in the sample due to the presence of salts, phenol and other inhibitors that may be carried over from RNA extraction (Pfaffl 2001). However, there are many ways to circumvent such variations, if encountered. Amongst these are: the type of fluorescent probe used for detection

of the DNA as well as the quantification strategy which is very crucial since it determines the type of output desired for a particular experiment.

A PCR reaction can be categorized by four major phases. Initially there is linear and the early exponential phase. This is followed by the exponential phase (log-linear) and the plateau phase as shown in Figure 1.3. The most important phases in quantification are the early exponential and log-linear phases as this is where the actual copy number and the amplification efficiency for a particular transcript is calculated from. Data gathered from these phases are important for calculating background signal and amplification efficiency. Rn is the intensity of the fluorescent emission of the reporter dye divided by the intensity of fluorescent emission of the passive dye (a reference dye incorporated into the PCR master mix to control for differences in master mix volume). ΔRn is calculated as the difference in Rn values of a sample and either no template control or background, and thus represents the magnitude of signal generated during PCR. (Wong *et al.*, 2005).

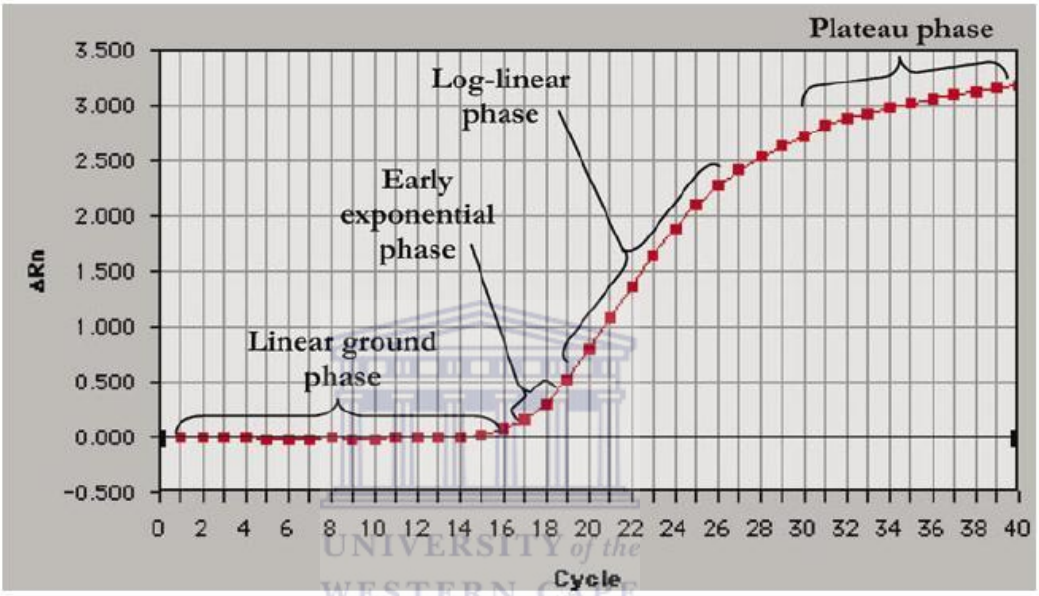


Figure 1. 3: Phases of the PCR amplification curve. The PCR amplification curve charts the accumulation of fluorescent emission at each reaction cycle. The curve can be broken into four different phases: the linear ground, early exponential, log-linear, and plateau phases. This graph was generated with ABI Prism SDS version 1.9 software (Applied Biosystems). (Wong *et al.*, 2005).

The linear ground phase represents the initial phase of the reaction and accounts for very little or no fluorescence other than that ascribed to background fluorescence, whereas the early exponential phase accounts for fluorescence above that of the background fluorescence. At this stage, the amplicon is quantified at each cycle since this is the point of the cycle in which detection occurs and is referred to as the crossing point (C_p). The value obtained at the Cp, represents the starting copy number of the template and is the value used to calculate the output. During the log linear phase, the reaction reaches its optimal amplification with each product being doubled in each PCR cycle (Wong *et al.*, 2005). Linearity can be detected at this point, and the slope can be used to calculate the efficiency of the reaction with a slope of 3.32 equating to 100 % efficiency. The steeper the slope the lower the efficiency and that could be due to factors like handling errors or inferior cDNA (Pfaffl 2001).

The last phase of the PCR is where the PCR components are starting to get depleted and this affects the amplification of the PCR product. Consequently the amplification plateau as a result of a decrease in PCR efficiency as the product is not doubled at every subsequent cycle (Pfaffl 2001). This phase is not used for quantification purposes as the target input is depleted.

1.2.4.2.3.1 Quantitation strategies in q-Real-Time-PCR

1.2.4.2.3.2 Absolute quantification

Absolute quantification essentially relies on known standard housekeeping genes to generate a standard curve of known concentration by using a serial dilution of the input target DNA. The aim of using a standard curve is to generate a linear relationship between the C_p and initial amounts of template, whether it is RNA or cDNA. This method permits the accurate quantification of unknown concentrations of template by extrapolation based on the C_p values. The main advantage of a standard curve is that it generates highly reproducible and stable data, and allows for the accurate calculation of PCR efficiency for each template reaction. However, this approach is laborious and is highly reliable on whether good laboratory practices were employed during experimentation (Pfaffl 2001, Schmittgen *et al.*, 2008).

1.2.4.2.3.3 Relative Quantification

The relative quantitation strategy, measures changes in gene expression based on either an external standard or a reference sample, known as a calibrator (Wong *et al.*, 2005).

This approach takes into consideration changes in mRNA levels across multiple samples and then quantifies their expression levels by the comparing it to the expression levels of an internal control usually in a control sample. The expression levels are usually expressed in terms of a fold increase or decrease

relative to the expression of the internal in the control sample (Pfaffl 2004, Dhanasekaran *et al.*, 2010).

Comparison to absolute quantification the approach of relative quantification does not require standards with known concentrations and the reference can be any transcript (Pfaffl 2001). Currently there are numerous models available to calculate the gene expression ratios based on relative quantitation, which include the comparative C_t (threshold cycle) method, the Livack method, the Pfaffl method and several others (Wong *et al.*, 2005).

The comparative Ct method can be highly inaccurate as it assumes an equal PCR amplification for all template samples, and most of the time this is not the case. However, this approach is currently used extensively in custom designed and pre-packed assays plates for pathway analysis and diagnostics research. Although this is now an accepted approach in gene transcript expression analysis, it is still not the most accurate approach to use. For the purpose of this study, only one method will be elaborated on based on a relative quantification approach using standard curves.

1.2.4.2.3.4 Pfaffl Method

The Pfaffl model is the model of choice in most research environments. This model combines gene quantification and normalisation into a single calculation. The most crucial aspect of this model is that uses the amplification efficiencies of both the target and reference genes to correct for the differences in amplification

efficiencies between the two assays in both the control and sample assays performed. The Pfaffl method employs an excel-based software, known as REST®, which automates the data analysis. There is an in-build pairwise Fixed Reallocation Randomisation test to determine whether results is significant and is indicates whether the reference gene is suitable for normalisation (Wong *et al.*, 2005). Expression ratios are calculated based on imported C_p from the experiments performed and expresses the target gene ratio differences between the control and test sample and generates a plot of the ratios for all samples tested (Pfaffl 2001).

1.2.4.3 Nanotechnology-based solutions for the diagnosis of cancer

1.2.4.3.1The use of quantum dots as fluorescent tags in molecular probes

Quantum dots (Qdot) are nanometre-sized clusters that are composed of a few hundred to several thousands of atoms, usually from groups II-VI, III-V and IV-VI (Rosenthal *et al.*, 2011). Qdot are often synthesized within a size range of 2 to 10 nm in diameter (Peng *et al.*, 2010). They have unique optical and electronic properties such as size-tuneable emission, excellent signal brightness and resistance to photobleaching; not exhibited by their organic fluorophore counterparts (Xing *et al.*, 2007, Zhang *et al.*, 2008). As a result Qdot can address some of the limitations of organic fluorophores in FISH and IHC. Furthermore, Qdot have several advantages in fluorescence imaging applications, such as a broad absorption spectrum and a narrow emission spectrum (Figure 1.4 A and B) (Tabatabaei-Panah *et al.*, 2013). The broad absorption spectra of Qdot allows for a wider range of excitation wavelengths (Jamieson *et al.*, 2007) compared to

organic fluorophores. The narrow emission spectra can allow for greater multiplexing possibilities (Pathak *et al.*, 2006). The implication for diagnostic techniques such as FISH and IHC is that the application of Qdot can facilitate the detection of multiple biomarkers simultaneously. This can significantly reduce the cost and time of diagnosis.

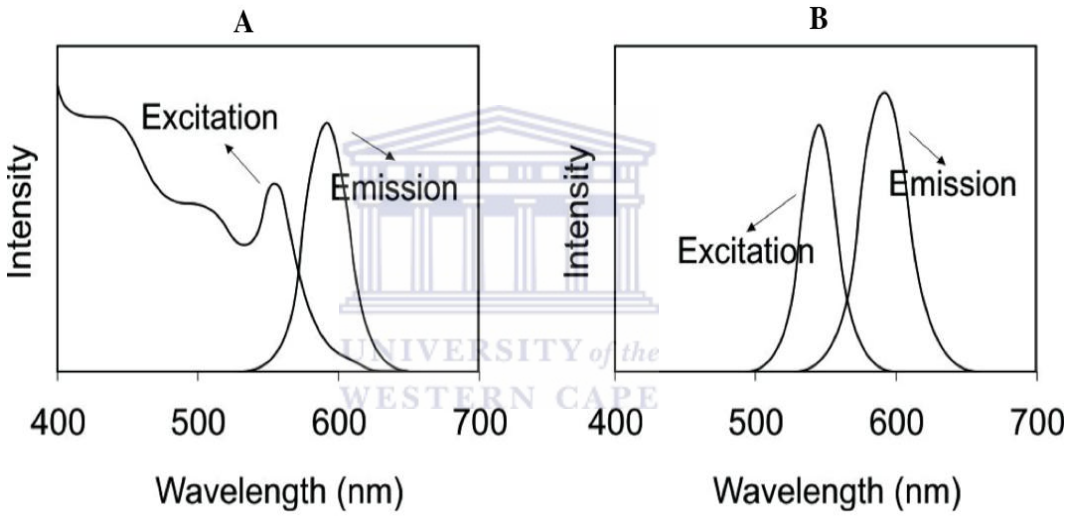


Figure 1. 4: Excitation and emission profiles between Qdot (A) and organic fluorophore (B). The absorption bands of Qdot extend from the gamma rays region and into the UV region; this width absorption band is related to the size of the Qdot, which gave flexibility to multiphoton microscopy. Whereas absorption bands of the organic fluorophore have a very narrow range as shown in figure 1.3 (B). (Fontes *et al*., 2012)

Qdots have huge potential in cancer diagnostics and imaging. The use of Qdots for fluorescent imaging has unique possibilities for cancer imaging at the molecular level (Fang *et al.*, 2012, Alam *et al.*, 2013). The application of Qdots as imaging agents has already been explored in cancer research by conjugating the Qdot to targeting biomolecules such as antibodies and peptides (Zhang *et al.,* 2008). Another study reported that multicolour Qdots have been used together with breast cancer biomarkers such as *HER-2* to detect the expression of this gene in tissues using Qdot/antibody profiling (Peng *et al.*, 2010, Tabatabaei-Panah *et al.*, 2013). The use of Qdots as imaging agents offers much lower detection limits with higher accuracy in comparison to traditional diagnostic techniques (Fang *et al.*, 2012).

Another application of Qdots is in the development of molecular beacons. Molecular beacons have been used in diagnostics assays to test for various infections in clinical samples (Tyagi *et al.*, 2012). Molecular beacons have also been applied in qRT-PCR assays for the detection of HIV, HBV and HCV (Tsourkas *et al*., 2003, Kim *et al*., 2008). Molecular beacons are molecular probes that target specific nucleotide sequence in the cells. Tyagi and Kramer developed molecular beacon technology in 1996 (Tyagi *et al.*, 1996). The use of Qdots in molecular beacons will be discussed in more detail in Chapter 4.

1.3 Biomarkers

Biomarkers play a very important role in the medical diagnostics field since these biomolecules are capable of discriminating between physiological and pathological conditions. These molecules are the result of altered biological processes due to the presence of a disease or due to the effect of a drug treatment. In the diagnosis and prognosis of a disease state, they are characteristic of a substance that can be objectively evaluated and measured as an indicator of a particular biological condition, disease response to medication, therapeutic outcome or disease progression and may be used to explore the disease mechanism (Strimbu *et al.*, 2010). Therefore, biomarkers can improve the quality of health through the early diagnosis and the monitoring of diseases following treatment. The sources of biomarkers can include tissues, cells and biological fluids (Nie *et al.*, 2007, Strimbu *et al.*, 2010).

The levels of cancer biomarkers vary with progression of the disease and this can give insight into various aspects of the disease, such as an individual's risk of developing cancer, the presence of cancer, the molecular mechanisms of cancer, the response of the disease to a particular treatment and the recurrence of the disease (Figure 1.5). In cancer, biomarkers can be classified as diagnostic, prognostic or predictive. Examples of such biomarkers are listed in Table 1.1. Diagnostic biomarkers play an important role in screening for individuals who have developed the disease or are susceptible to the disease at a later stage i.e. it can assess the risk of a patient to develop the disease. These biomarkers can be useful in prevention, early detection, intervention and eradication of the disease

and therefore reduce cancer mortality (Tainsky 2009). Prognostic biomarkers can distinguish different stages of the disease, predict clinical outcome and determine the course of therapy that must be applied to a particular patient after primary treatment (Riley *et al.*, 2009). Examples of prognostic biomarkers in breast cancer include BRCA1, *HER-2*/neu, estrogen receptor (ER) and progesterone receptor (PR), which are already used in the medical field (Gerhardus *et al.*, 2007, Mehta *et al.*, 2012, Sapino *et al.*, 2013).

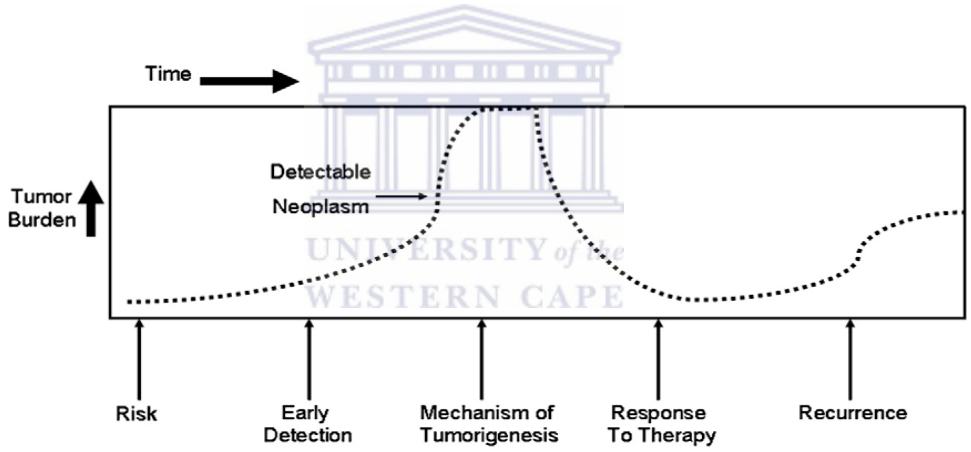


Figure 1. 5: The relevance of cancer biomarkers at different stages of cancer progression (Tainsky 2009)

Predictive biomarkers are used to determine relapse or the recurrence of the disease after the patient had undergone treatment (e.g. surgical removal of the tumours). A good example of a predictive biomarker for breast cancer is *HER-2*, which has been used to predict sensitivity to Herceptin treatment (Mehta *et al.*, 2010). The FDA approved this biomarker for use in the diagnosis of breast cancer. Several commercial diagnostic tests, which include HercepTest™, Pathway, Insite, PathVysion and SPOT-Light HER-2 CISH are available for this biomarker. These diagnostic tests most often use molecular diagnostic techniques such as IHC, FISH and qRT-PCR (van de Vijver *et al.*, 2007). These methods require a biopsy sample from the patient, which means that the diagnostic test is an invasive procedure, which can be very uncomfortable to the patient. Other disadvantages of biopsy samples include sampling error, which could lead to a false negative result, increasing the risk of complications such as hematoma and a more serious aspect is the possibility that the tumour cells may migrate into adjacent tissues following the procedure (Loughran *et al.*, 2010, Nassar 2011). Even if fine-needle aspiration biopsy is used there is still other limitations such as the high cost, the need for highly trained medical personnel for the preparation and analysis of the sample (Nassar 2011).

Diagnostic tests that are based on the detection of serum biomarkers do not require a biopsy sample, since the diagnosis can be done from a blood sample and less invasive diagnostic procedures such as ELISA can be used. Serum biomarkers are also suitable for application in POC diagnostic devices. Several molecular and blood-based assays for different types of serum cancer biomarkers

are available. Examples of such biomarkers are listed in Table 1.1 and includes prostate-specific antigen (PSA), which was approved by the FDA for use in the clinical diagnosis of prostate cancer (Uhl *et al.*, 1997, Chatterjee *et al.*, 2005). Bladder tumour antigen (BTA) and nuclear matrix protein-22 (NMP22) are amongst the other biomarkers approved by FDA for the screening of bladder cancer in urine samples (Mungan *et al.*, 2000, Chatterjee *et al.*, 2005). Human epididymis protein 4 (HE4) has been approved by the FDA for monitoring the recurrence or progression of ovarian cancer, while Cancer Antigen 125 (CA125) has been approved for the monitoring for ovarian cancer in serum (Molina *et al.*, 2011, Diamandis 2012). To date very few serum biomarkers that are specific for breast cancer are known and there is therefore a need to identify additional biomarkers for breast cancer.

Table 1. 1: A list of known breast cancer biomarkers.

Gene	Source	Sample	Biomarker type	Test	Genetic alteration	FDA-cleared test	Reference
ER	Breast	Serum	Prognostic and predictive	IHC	Elevated expression	ER/PR pharmDx	(Majewski *et al*., 2011)
PR	Breast	Serum	Prognostic and predictive	IHC	Elevated expression	ER/PR pharmDx	(Majewski *et al*., 2011)
CA15-3	Breast	Serum	Monitoring	-			(Gion *et al*., 1999, Bast *et al*., 2001)
CA27-29	Breast	Serum	Monitoring				(Gion *et al*., 1999, Bast *et al*., 2001)

1.3.1 Methods used to identify biomarkers

1.3.1.1 Transcriptomics

Changes in the expression patterns of genes can lead to the development of diseases. Various genetic diseases have been identified and can be passed on from one generation to the next. Hence, genetic profiling can be used in medical research for the identification of genetic markers that can be used to screen for the presence of certain diseases, including cancer (Davis *et al.*, 2006, Tainsky 2009). This was made possible through genomic technologies such as DNA sequencing, DNA microarrays, Serial Analysis of Gene Expression (SAGE), quantitative real-time PCR (qRT-PCR) and RNA-Sequence (Kawasaki 2006, Schroeder *et al.*, 2006).

Genomics is a study of an organism's entire gene complement including the gene structure, gene function and control of gene expression. Genomics also entails the study of gene networks with the aim of understanding how genes interact with each other and with the environment. The sequencing of the human genome has demonstrated that it is possible to identify genetic variations responsible for diseases and identify individuals that are at risk of developing genetic diseases. Understanding the human genome, gene functions and genetic interactions has opened avenues in molecular biology for the development of diagnostic and therapeutic strategies in the fight against human diseases, more specifically cancer (Venter *et al.*, 2001).

The development of cancer is characterised by genetic mutations. These genetic mutations can be used as genetic or molecular markers for cancer. These genes encode proteins that are involved in the regulation of cellular functions (cell growth, division and death). Irreversible changes in these gene sequences can cause mutations that promote the uncontrolled growth of cells leading to development of cancer (Greenman *et al.*, 2007). These mutations can be associated with the key identifiable features of cancer cells.

Independent breast cancer studies have identified some mutated genes responsible for its development such as breast cancer 1 (BRCA1) and breast cancer 2 (BRCA2) (Atchley *et al.*, 2008). These molecular markers are currently used for diagnostic and therapeutic intervention in breast cancer patients, and have proved their potential to help combat cancer (Mayeux 2004).

DNA microarray is one of the most commonly used high throughput techniques to study the genome wide gene expression profiles in cells or tissue. In this technique, small amounts of DNA or RNA are extracted from cells or tissues to determine the copy number of genes/mRNA or DNA region of interest. It has enabled researchers to query the mRNA expression levels of thousands of genes in an organism simultaneously (Karakach *et al*., 2010, Pulverer *et al.*, 2012) DNA microarrays are collections of microscopic spots created by robotic machines and arranged in a grid-like format on a solid support such as a glass slide. Each of these microscopic spots represents the cDNA derived from the mRNA of known genes. The cDNA's of several thousand genes can be spotted on a single slide.

The process of performing DNA microarray analysis involves a number of steps starting with the design of the experiment, the extraction of nucleic acids (usually mRNA) from the control (e.g. healthy cells or tissue) and experimental (e.g. beast cancer cells or tissue) samples, the transcription of the extracted mRNA into cDNA molecules that are differentially labelled (e.g. the controlled sample is labelled with Cy5 and experimental samples is labelled with Cy3) with fluorescent labels, hybridization of labelled cDNA molecules with the cDNA immobilized on the glass slides, scanning of the micro-array, image processing, normalization, ratio calculation, statistical analysis, and ending with the extraction of information and generation of knowledge from the results (Karakach *et al*., 2010). The analysis of the fluorescence signal (intensity and wavelength) of the spots on the slide can be used to assess the relative expression levels of thousands of genes simultaneously.

The transcriptomic signatures of cancers have been used as prognostic diagnostic tools to predict the clinical outcome for breast cancer patients (Wirapati *et al*. 2008). MammaPrint®, Oncotype DX™, MapQuant Dx™ and Mammostrat® are examples of such diagnostic tests that are currently licensed for commercial use (Arpino *et al*., 2013). Molecular technologies such as qRT-PCR and IHC form the bases of the diagnostic tests.

1.3.1.2 Proteomics

Proteomics refers to the study an entire proteome and protein expression profile in a given setting in the given cell type or tissue using high throughput technologies

such as HPLC and mass spectrometry (Baak *et al.*, 2005). The objective is to identify sets of proteins that are differentially expressed between a normal and pathological/disease state (Baak *et al.*, 2005), and thereby potentially creating an unique profile or fingerprint for a particular disease, for example breast cancer. Quantitative proteomics has been explored, especially in cancer research, to determine the changes and pathophysiology of this disease. For instance, in breast cancer, useful biomarkers for breast cancer diagnosis have been identified through proteomic studies (Liu *et al.,* 2013). Proteomic techniques are sensitive enough to detect changes that occur during the development of the disease and also the type of modifications induced by the disease state (Liu *et al.*, 2013). Proteomic technologies can also provide information on how the protein interacts during the development of a disease. Thus, the disease mechanisms and strategies on how to prevent, treat and diagnose may be established through proteomics (Baskın *et al.*, 2010, Wasinger *et al.*, 2013).

1.3.1.2.1 Proteomics technologies

Proteomics encompasses several technologies have been used for many years as for sample analysis tools in chemistry. This includes technologies such as Liquid Chromatography Mass Spectrometry (LC-MS), Matrix-assisted Laser Desorption Ionization coupled to a Time-of-flight Mass Spectrometer (MALDI-TOF/MS), Gas Chromatograph Mass Spectrometry (GC/MS) and Isotope Ration Mass Spectrometry (IR/MS) (Nair *et al.*, 2004). Gas Chromatograph Mass Spectrometry (GC/MS) and Isotope Ration Mass Spectrometry (IR/MS) (Nair *et*

al., 2004). These technologies are now used to characterise protein samples in an effort to identify protein biomarkers for diseases (Nair *et al.*, 2004).

New clinical proteomics technologies such as LC-MS/MS has proved to be useful in identifying potential biomarkers from variety of samples including cells, tissues and body fluids (Brase *et al.*, 2010). LC-MS/MS is an analytical technique used to determine the molecular mass with different sizes, and can be used to characterize complex mixtures ranging from proteins, carbohydrates, DNA, drugs, and biomolecules associated with disease development and progression and has a very high sensitivity of detective (Lee *et al.*, 2010, Grebe *et al.*, 2011). Unlike immunoassays, LC-MS/MS can be used to analyze low molecular weight samples and samples that is only available in low concentrations even in small amount of samples (Kang 2012).

LC-MS/MS has become a very valuable tool for analysis in the medical, clinical laboratory field and pharmaceutical analysis (Kang 2012, Haneef *et al.*, 2013), pertaining to drug metabolism and toxicology in the body, by following the pathways of metabolism of any drug through the identification of circulatory and excretory metabolites. During the last 10 to15 years, LC-MS/MS has become the instrument of choice in clinical laboratories for the analysis of target peptides and proteins.

One of the advantages of using LC-MS/MS in the clinical environment is the characterization of thousands of proteins in parallel in clinical samples, and it also

enables the study of complex mixtures of peptides, carbohydrate and proteins which are unique to relevant disease process (Grebe *et al.*, 2011, Haneef *et al.*, 2013).

1.3.1.2.2 Stable Isotope Labelling with Amino Acids in Cell Culture (SILAC)

SILAC involves the in vitro labelling of cellular proteins in cells to allow for the comparative analysis of proteomes. Healthy and diseased cells are cultured in the same conditions, except that the two populations are grown in media supplemented with either "light or "heavy" amino acids. During cell division the isotope labelled amino acids replaces the natural amino acids. The process takes 6-8 doubling times to ascertain 100 % labelling (Ong *et al.*, 2002, Munday *et al.*, 2012). Lysine and arginine are usually preferred because trypsin would cleave the polypeptides at the carboxyl terminal of arginine and lysine residues (Ong *et al.* 2003, Munday *et al.*, 2012). The labelled amino acids are similar to the natural ones in all respects and therefore the cell growth, cell morphology and biological activity of the cells are unaffected. Although chemically similar, the amino acids are isotopically different and induce a mass shift that can be distinguished through MS (Ong *et al.*, 2002, Ong *et al.*, 2003). The peptides generated by MS can be easily correlated to their respective proteins of origin through Bioinformatics tools, the ratio of their relative intensity can be used to quantify protein expression between normal and disease states (Liang *et al.*, 2006). Initially designed for cell culture labelling, SILAC has been used successful to study protein expression in brain tissue (Ishihama *et al.*, 2005), the secretome of primary tissues (Grønborg *et al.*, 2006) and changes in skin tissues (Zanivan *et al.*, 2013).

SILAC has been used successfully to study the proteome of several cell lines for the purpose of identifying disease biomarkers and understanding of systems biology. For example, the study by Everley *et al.* (2004) investigated the changes in the microsomal proteome of prostate cancer cell lines with varying metastatic potential with the aim of understanding the progression of prostate cancer. The authors compared the protein expression levels of more than 440 specific proteins from the microsomal fraction of prostate cell lines. Their data showed that 60 proteins were upregulated by a factor greater than 3-fold and 22 proteins were down-regulated in the metastatic prostate cell line. Their study suggested that SILAC proved to be a useful tool to study the progression of cancer in different cancer cell lines. Another study used SILAC to identify and quantify the differential expression of proteins between neoplastic and non-cancerous gastric cell lines (Marimuthu *et al.*, 2013). In this study 2205 proteins were identified from the secretomes of the cell lines. Two-hundred-and-sixty-three (263) proteins were overexpressed (greater than 4-fold) in the neoplastic cell line as compared the non-cancerous cell line. This study led to the identification of a number of potential novel biomarkers for gastric cancer. It is therefore clear that SILAC in combination with proteomics is a promising method for the identification of novel biomarkers in cancer and could also be used for the quantification of these newly identified biomarkers in the diagnosis and prognosis of a variety of cancers.

1.3.1.2.3 Mass spectrometry

Mass spectrometry (MS) is a standard platform employed in the identification of proteins isolated from biological samples (Hanash *et al.*, 2008). MS is an analytical technique that is used to sequence and identify proteins. The protein sample, upon introduction into the MS, is ionized; the resulting mass-to-charge ratio of the generated ions is then used to identify the protein molecule (Aebersold *et al.*, 2003). Various MS instruments are available and are classified as scanning, ion-bean or trapping MS (Yates *et al.*, 2009). The most commonly used MS methods include Fourier transform ion cyclotron (FT-MS), ion trap, guadrupole and Time of flight (TOF). The preferred methods, in protein biochemistry analysis, are matrix-assisted laser desorption/ionization-time of flight (MALDI–TOF), surface-enhanced laser desorption/ionization-time of flight (SELDI-TOF) and liquid chromatography-tandem mass spectrometry (LC-MS/MS) (Aebersold *et al.*, 2003, Baskın *et al.*, 2010).

The application of MS to analyze proteins led to the development of a variety of emerging proteomics methods for the analysis of clinical samples such as tissues, serum and plasma (Schaub *et al.*, 2009). LC-MS/MS is a technique used to identify and separate proteins and peptides from complex sample mixtures to allow for analysis of individual fractions (Hanash *et al.*, 2008). In terms of protein quantification, isotope labelling is currently a powerful strategy. It provides accurate measurements of the gene expression and protein-coding and is also widely used to analyse small-molecules and identify biomarkers for particular diseases (Ciccimaro *et al.*, 2010, Soufi *et al.*, 2010).

1.3.1.3 Limitations of proteomics

Sample preparation is the first and most crucial step when using mass spectrometry-based proteomics experiments. Sample preparation refers to all the stages that take place before sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis for gel-based proteomics analysis. Sample preparation for proteomic analysis is very delicate and crucial for the best outcome. Several factors must be taken into consideration prior to analysis; such as the sample type and source and their physicochemical properties since these factors will determine the complexity of the sample (Wang *et al.*, 2005). For this reason, proteomics is never a single step approach, which will help to determine the preparation steps required. The presence of high abundant proteins (HAPs) within the sample as well as the sample localization, solubility often complicates the sample analysis, therefore, additional steps must be included to successfully identify these proteins (Bellei *et al.*, 2011).

For complex samples such as blood, the removal of interfering HAPs from the sample is recommended. The sample preparation includes cell lysis, to release or extract the total protein; protein solubilization, removal of contaminants (interfering proteins) and fractionation to reduce the sample complexity and improve its resolution (Wang *et al.*, 2005). Fractionation is used for depletion of HAPs on the basis of size, charge, hydrophobicity or binding affinity, to enrich and improve the detection range of low abundance proteins (LAPs) by increasing the protein load. LAPs of clinical significance can be masked by the presence of HAPs. They can be hormones, cytokines, growth factors and high molecular

weight proteolytic protein fragments. LAPs are essential for normal biological processes and some of them may be potentially useful for diagnostic or therapeutic purposes (Wang *et al.*, 2005, Millioni *et al.*, 2011, Liu *et al.*, 2013). They can be concentrated to detectable levels necessary for analysis to discover their relevance, especially in pathological processes. Various sample preparation techniques are available and may also be combined to reduce protein convolution prior to its analysis (Wang *et al.*, 2005).

1.3.2 Examples of well characterized biomarkers for breast cancer

1.3.2.1 Breast cancer type 2 Susceptibility protein (BRCA2)

Breast cancer type 2 susceptibility locus was discovered by Richard Wooster in 1994 (Wooster *et al.*, 1994). BRCA2 gene codes for breast cancer type 2 susceptibility protein, which plays an important role in the maintenance of genome integrity during Double-Strand Break Repair (DSBR) replication. This protein interacts with proteins that are involved in DNA repair such as proteins RAD51 Recombinase (RAD51) and Partner And Localizer of BRCA2 (PALB2). Failure to repair DNA damage can lead to replication errors that cause mutated DNA, which is propagated, and result in cancer (Schlacher *et al.*, 2011).

BRCA2 is also classified as a tumour suppression gene as described in section 1.1 (Sharan *et al.,* 1997). This gene is involved in DNA repair pathways, such as homologous recombination (Stefansson *et al.*, 2009). The expression of BRCA2 was observed in different cancer tissues such as prostate cancer and ovarian cancer (Goodheart *et al.*, 2009, Castro *et al.*, 2012).

Kwong *et al.* (2009) reported that the clinical risk rate of BRCA2 mutation had a significantly high level among 226 Chinese women, and the prevalence of triple negative breast cancer in women who were carriers of the BRCA2 mutation was high when compared to the Caucasian cohort (Kwong *et al.*, 2009). A new deleterious mutation of BRCA2 in exon 15 was identified by Pisano *et al*. As a result of this mutation, the BRCA2 gene becomes completely inactive (Pisanò *et al.*, 2011). The BRCA2 functions as a gatekeeper of genomic integrity and it has critical step in homologous recombination by regulating the DNA repair protein RAD51 gene filament formation (Shuen *et al.*, 2011).

Malone et al studied the prevalence of BRCA2 among black and white American women (35 to 64 years old) and found within a total of 1628 women diagnosed with breast cancer 674 women without breast cancer; the BRCA2 mutation was slightly more prevalent in black women (2.6 %) than in white women (2.1 %). Also, the BRCA2 mutation in younger (aged 35 to 45) and older women (aged 45 to 64) was 4 % and 1.5 %, respectively (Malone *et al.*, 2006).

Recently, 360 breast cancer patients (aged 29 – 76 years) were used in a study in which the objective was to detect the BRCA2 mutation using a high resolution melting (HRM) assay. The BRCA2 mutation was observed in 27 of the 360 patients (Xu *et al.*, 2012). Kotsopoulos *et al* conducted a case-control study within seven countries, using women (1665 pairs) carrying the BRCA2 mutation. It was concluded that no significant association exists between breastfeeding and the risk of breast cancer for at least one year (Kotsopoulos *et al.*, 2012). De Bruin *et al.*

(2012) studied the prevalence of the BRCA mutation among Asian and white American women. Fifty Asian women and forty-nine white American women were included in the study. It was found that the BRCA2 mutation was more prevalent among Asian women 29/50 (58 %) than white American women 18/49 (37 %) (de Bruin *et al.*, 2012).

Lored-Pozos *et al*. (2009) reported on the gene expression of BRCA2 among premenopausal women, aged 22 - 45 years, diagnosed with breast cancer was assessed by RT-PCR. It was found that BRCA2 mRNA expression had no association with disease severity in young aged breast cancer women (Loredo-Pozos *et al.*, 2009).

Another study examined the relationship between BRCA2 expression and some clinical factors in breast cancer patients. It was found that the BRCA2 expression in patients younger than 50 years of age (53.3 %) was slightly lower than patients older than 50 years of age (59.5 %). Based on histological grade, positive BRCA2 expression marked 24/47 (51.1 %) and 14/20 (70 %) in grade I-II and grade III breast tumours, respectively. However, it's expression in axillary lymph node metastasis was 28 out of 50 (56 %) (Li *et al.*, 2011).

1.3.2.2 Human Epidermal Growth Factor Receptor 2 (*HER*-2/neu)

Human epidermal growth factor receptor 2 (*HER*-2), also known as C-erbB-2, is a transmembrane glycoprotein and member of the tyrosine kinase family. *HER-2* is a proto-oncogene that is located on chromosome 17q21 and encodes for a 185-

KDa protein (Yoshino *et al.*, 1994, Roşian *et al.*, 2005). *HER-2* was discovered in 1985 (Groudine *et al.*, 1985), and it was approved by the FDA in 1997 as diagnostic biomarker for breast cancer using the FISH technique (Ross *et al.*, 2009). It is over expressed in approximately 25-30 % of breast cancers and its expression is particularly high in invasive breast cancer specimens. Overexpression of *HER-2* has been associated with negative expression of ER (Groudine *et al.*, 1985, Lee *et al.*, 2007). HER-2 is also an important biomarker target for the treatment of this disease. Patients with high levels of *HER-2* expression are good candidates for treatment with trastuzumab, which is a monoclonal antibody-drug conjugate that acts by blocking HER-2 (Radojicic *et al.*, 2011).

At present, the only *HER-2* targeted therapy approved by the FDA for the treatment of metastatic breast cancer (MBC) is the drug, tastuzumab. The *HER-2* protein is an important therapeutic target in breast cancer for several reasons. Firstly, *HER-2* is known to be overexpressed in 20–25 % of invasive breast cancers (Slamon *et al.*, 1987, Slamon *et al.*, 1989, Nahta *et al.*, 2006) and is associated with poor prognosis in the disease. HER-2 overexpression is also associated with resistance to certain chemo- therapeutic agents and this decreases the effectiveness of trastuzumab. The overexpression of this receptor is generally due to gene amplification, which can result in up to a 25-fold increase in *HER-2* copy number (Press *et al.*, 2002, Nahta *et al.*, 2006). The normal levels of the HER membrane-bound receptor is reported to be 20,000 per cell. *HER-2* overexpression has been defined by immunohistochemistry and is classified as

being highest (reported as 3+) when the receptor levels are in the region of 2 million per cell, or medium (2+) when receptor levels are approximately 500,000 per cell (Ross *et al.*, 2004). Secondly, the pathogenesis and prognosis of breast cancer is also strongly correlated with elevated HER-2 levels) (Slamon *et al.,* 1987, Slamon *et al.*, 1989, Nahta *et al.*, 2006). Thirdly, in human cancer cells the level of gene amplification is much higher than in normal adult cells and thus, targeting the HER-2 protein with *HER-2*- targeted drugs could reduce the pathogenicity caused by HER-2 overexpression. Lastly, HER-2 protein levels are known to be relatively homogeneous among HER-2-overexpressing tumour cells (Eccles 2001).

HER-2 expression is an important predictive and prognostic factor in breast cancer. A study done in North Pakistan on a total of 1226 patients diagnosed with IDC found that 478 of these patients were positive for HER-2 expression (Faheem *et al.*, 2012). Kinsella *et al.* (2012) compared the IHC profile generated for *HER-2* in pre- and post- neoadjuvant chemotherapy among 37 females diagnosed with primary breast carcinoma and found that 32 % of the patients were positive for *HER-2* prior to the neoadjuvant chemotherapy. This dropped to 22 % after the treatment (Kinsella *et al.*, 2012).

In another study, the expression of *HER-2* was studied in 1134 female subjects (≥ 18 years old) diagnosed with invasive breast cancer. It was found that 18 % of the cases were *HER-2* positive (Onitilo *et al.*, 2009). Using IHC, Yanagawa *et al.* (2012) showed that 30 out of 363 primary invasive breast cancer patients (aged 30

to 87 years) were positive for *HER-2*; none of which had a family history of breast cancer (Yanagawa *et al.*, 2012). The presence of *HER-2* in CTCs in breast cancer patients was reported by Fehm *et al*. (2009); a study in which 431 blood samples were analysed. The detection rate, for *HER-2*, was 38 % (22/58 patients) in CTCs (Fehm *et al.*, 2009). The expression of HER-2 was investigated in human ductal breast carcinoma patients, a total of 69 samples were examined by IHC. A product of IHC as an indicator of *HER-2* expression was present on the cell surface of cancer cells. Staining for 21.74 % of the cell field was observed, while 28.98 %, 26.01 %, 23.19 % and 21.74 % exhibited strong, moderate and weak staining, respectively (Sopel *et al.*, 2011). Subik *et al.* (2010) analysed the level of *HER-2* in 17 common breast cancer cell lines using an IHC assay and HER-2 was shown to overexpress in 4 cell lines only (SKBR-3, MDA-MD-435, AU 565 and BT-474) (Subik *et al.*, 2010).

The expression of *HER-2* was evaluated using IHC. It was observed as positive (3+), weakly positive (2+), negative (1+) and negative (0), the expression was 3+ in 14 % (21 patients), 2+ in 10.7 % (16 patients) and 0/1 in 75.3 % (113 patients), but with FISH, the expression was detected in 18 specimens (all positive with 3+). (Schoppmann *et al.*, 2010). IHC and FISH techniques were used to detect *HER-2* expression in invasive breast cancer patients (n=1984). The expression of *HER-2* was found to be up-regulated in 75.4 of the cases (Caldarella *et al.*, 2011). Another study examined tissues from 518 invasive breast cancer patients and the results revealed that 155 patients scored positive (Moelans *et al.*, 2009). The amplification of *HER-2* in breast cancer was evaluated with using IHC and FISH,

30 out of 114 (26 %) breast cancer samples were amplified by FISH, whereas about (18 %) were IHC positive (Ellis *et al.* 2005). Al-Dujaily *et al*. (2008) conducted a study in which he collected 90 specimens of breast cancer tissues, including 20 normal tissue samples as a control and 25 benign breast lesions. *HER-2* expression level was detected in 67.8 % of the cases; and 73.2 % of IDC cases and 12.5 % of ILC cases, while the expression was negative in all normal and benign breast lesions (AL-Dujaily *et al.*, 2008). Stark *et al.* carried out a study, between 2001 and 2007, using women from African/white American origin. A total of 581 African American and 1008 white American women were tested for *HER-2* expression using the FISH assay. The assay indicated 82.9 % and 76 % of white American and African American women, respectively, were positive for *HER-2* expression; while (75.1 % 332/581) of the African American and 76.7 % of the white American women were negative for breast cancer (Stark *et al.*, 2010). Recently, Joensuu *et al.* (2013) reported that the *HER-2* gene level was evaluated in 72 females with primary breast cancer, of which 21 % of the cases were positive (Joensuu *et al.*, 2013). Georgescu *et al*. (2012) reported on the presence of the *HER-2* gene in 27.5 % of females with invasive breast carcinoma, while 72.5 % of the patients were *HER-2* negative (Enache *et al.*, 2012).

Another study evaluated *HER-2* status in 90 cases of mammary invasive carcinoma. The expression of the HER-2 oncoprotein was assessed as negative in 65.55 % cases, while positive cases were found to be 54.44 %, 11.11 %, 18,8 % and 15.56 % presented score 0, 1+, 2+ and 3+, respectively (Pătrană *et al.*, 2012).

1.3.2.3 Fork head

The fork head is a family of transcription factors, usually characterized by a 100-amino-acid, monomeric DNA- binding domain (Kaestner *et al.*, 2000). It was discovered on the basis of specific DNA binding activity present in liver nuclear extracts for a specific promoter region (Friedman *et al.*, 2006). Among this family is the Fork head box A1 (FOXA1), also known as hepatocyte nuclear factor 3 alpha (HNF3a) (Kaufmann *et al.*, 1996, Thorat *et al.*, 2008), and is located on chromosome 14q21.1 (Nucera *et al.*, 2009, Robinson *et al.*, 2013).

According to Wolf *et al*. (2002), the expression of FOXA1 was found in the liver, breast, bladder, pancreas, colon, lung and metastasized prostate cancer (Wolf *et al.*, 2007, Jain *et al.*, 2011). Wolf *et al.* (2002) reported that the expression level of FOXA1 in IDC is higher (94 %) than in pure DCIS (85 %) and those normal tissues adjacent to cancerous breast tissue expressed high levels of FOXA1. However normal breast tissues showed low expression and correlated with favorable prognostic factors such as tumour grade (Wolf *et al.*, 2007).

Furthermore, FOXA1's expression was correlated with the luminal breast cancer subtype (Thorat *et al.*, 2008). Hadashy *et al*. (2008) reported on the overexpression of FOXA1 in the nuclei of malignant and luminal ductal cells (Habashy *et al.*, 2008). Additionally, its expression was associated with tumour size (Habashy *et al.*, 2008, Albergaria *et al.*, 2009), however, no association between FOXA1 and the patients age in lymph node stage (Habashy *et al.*, 2008). Some studies show that the expression level of FOXA1 decreases with an increase

in tumour grade, tumour stage (Abe *et al.*, 2012), and nuclear grade (Hisamatsu *et al.*, 2012).

Yamaguchi *et al.* (2008) evaluated FOXA1 mRNA expression in a panel of 35 human breast cancer cell lines and positively correlated overexpressed FOXA1 with GATA3 in 63 % of the cell lines. In the same study they examined if any correlation occurs between FOXA1 and ERbB2 as a marker of poor prognosis in human breast cancer. The results revealed an up-regulation of FOXA1 in all of the ErbB2-positive breast cancer cell lines (Yamaguchi *et al.*, 2008). A number of studies had suggested that FOXA1 is a significant predictor in breast cancer (Albergaria *et al.*, 2009, Mehta *et al.*, 2012), and was associated with good prognosis (Thorat *et al.*, 2008). This was confirmed in a study carried out by Hisamatsu *et al*. (2012), in which the overexpression of FOXA1 in ER- and PR-positive breast cancer patients was recognized as a good prognostic indicator (Hisamatsu *et al.*, 2012).

1.5 Aims of the study

The aims of the study were 1) to identify new potential biomarkers for breast cancer and 2) to develop molecular beacon technology (using quantum dots as the fluorescent signal) for applications in the diagnosis of breast cancer that can potentially be used in multiplex diagnostic assays.

1.5.1 The objectives of this study include:

- To use comparative proteomics analysis to identify membrane proteins that is differentially expressed between the human breast cancer cell line (MCF-7) and the non-cancerous human breast cell line (MCF-12A).
- To validate any new biomarkers in MCF-7 and MCF-12A cell lines as well as breast cancer patient samples using online bioinformatics tools such as GeneHub and molecular techniques such as qRT-PCR and IHC.
- To develop a molecular beacon probe for the detection of a known breast cancer biomarker

1.5.2 Hypothesis

Gene expression analysis such as DNA microarray has traditionally been used to identify disease biomarkers but this technology is prone to limitations. Proteomics on the other hand is a far more promising technology for the identification of disease biomarkers.

The identification of proteins that are over expressed in beast cancer cells can lead to the discovery of new biomarkers for this disease. Membrane proteins that are

involved in the pathology of breast cancer are over expressed in breast cancer tissues and cell lines derived from breast cancer tissue. Membrane proteins would be ideal biomarkers since these proteins are expressed on the surface of cells and can therefore facilitate the development of diagnostic methods that are less invasive. Proteomics can be used to identify these proteins.

The application of quantum dots as fluorescent signals in diagnostic techniques such as FISH, IHC and molecular beacons can lead to the production of diagnostic methods that are faster and cheaper, since quantum dots are more photostable and can potentially facilitate multiplex detection of several biomarkers in the same sample. Multiple genes that encode biomarkers, including membrane proteins can be detected using molecular beacon technology where quantum dots are used for the fluorescent signal.

CHAPTER 2

2.1 Materials and methods

Table 2. 1: General chemicals and materials suppliers

Item	Supplier
0.4 % trypan blue stain	Invitrogen
1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC)	Sigma-Aldrich
1 X Phosphate Buffer Saline PBS, pH 7.2	Invitrogen
6 X orange DNA loading dye	Fermentas
Acetone	Merck
Acrylamide-Bis (37.5:1), 40 %	Serva (GMbH)
Agarose	Lonza
Amicon Ultra-0.5 centrifugal filters	Merck
Ammonium persulphate (APS)	Merck
Bacteriological agar	Merck
Boric acid	Merck
Bovine serum albumin	Roche
Bromophenol blue	Sigma
Coomassie Brilliant Blue R250	Sigma

Dimethyl sulphoxide (DMSO)	Sigma
Dithiothreitol (DTT)	Thermo Scientific
DNA molecular weight marker ladder (100 bp)	Fermentas
Ethanol	Merck
Ethylene diamine tetra acetic acid (EDTA)	Merck
Formalin	Merck
Gel Red Nucleic acid stain	BioTIUM
Glacial acetic acid	Merck
Glycerol	Merck
Glycine	Merck
Hydrochloric acid	Merck
Isopropanol	Merck
Magnesium Chloride (50 mM)	Separation
Methanol	Merck
N,N,N',N'-Tetra methylethylene-diamine (TEMED)	Merck
Oligonucleotides	Integrated DNA technology (IDT)
PF Syringe Filter 0.8/0.2 μm	(PALL) life science

Polyvinylidene difluoride transfer membrane (PVDF)	BIO-RAD
Potassium chloride (KCl)	Merck
Protease inhibitor cocktail (complete, EDTA-free)	Roche
Qdot® 525 ITK™ Carboxyl Quantum Dots	Invitrogen Life technologies
Sodium dodecyl sulphate (SDS)	Merck
Sodium hydroxide	Merck
Tris (hydroxymethyl) aminoethane	Merck
Triton X-100	Sigma
Tween-20 (polyoxyethylene[20] sorbitan)	Merck

Table 2. 2: Cell culture material and suppliers

Item	Suppliers
DMEM: F12 Media	Lonza
Fetal Bovine serum (FBS)	Biochrom
Hydrocortisone	Sigma-Aldrich
Insulin	Roche
LONG® EGF human	SAFC Biosciences
Penicillin/Streptomycin (10.000 U/ml)	Lonza
Trypsin (2.5 %)	Gibco®

Table 2. 3: Biochemical assay kits and suppliers

Kit	Suppliers
BM Cyclin Kit	Roche
Clarity™ Western ECL Substrate	BIO-RAD
KAPA SYBER® FAST qPCR kit for LightCycler®480	KAPABiosystems
MycoFluor™ Mycoplasma detection kit (M-7006)	Life technology
NucleoSpin® RNA/protein extraction kit	Macherey-nagel GmbH
Pierce® Cell Surface Protein Isolation Kit	Thermo scientific
Pierce® SILAC Protein Quantitation Kits	Thermo scientific
SuperSignal™ West Pico Chemiluminescent Substrate	Bio-Rad
Transcriptor first strand cDNA synthesis Kit	Roche

Table 2. 4: Stock solutions and Buffers

Stock solution and Buffer	Composition
1 X Transfer Buffer	0.049 M Tris –HCl (pH 8.3), 0.38 M glycine and 20 % methanol. This solution was stored at 4 °C.
1 X Tris EDTA (TE)	10 mM Tris- HCl (pH8.0), 1 mM EDTA
10 X Tris, Boric acid and EDTA (TBE)	0.9 M Tris, 0.9 M Boric acid and 25 mM EDTA, pH 8.3
1 X Tris, Boric acid and EDTA (TBE)	10 X TBE was diluted 1:10 times
10 X SDS Electrophoresis Buffer	250 mM Tris-HCl pH 8.3 1.92 M Glycine and 10 % SDS.
12 % Separating gel	40 % Polyacrylamide, 1.5 M Tris–HCl (pH 8.8), 10 % Ammonium Persulphate, 10 % SDS, 0.002 % TEMED
12 X Separating gel Buffer	1.5 M Tis-HCl, pH 8.8.
1 X Phosphate Buffer Saline PBS	137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4 and 1.5 mM KH_2PO_4, pH 7.4.
1 X Tris-Buffered sodium chloride Tween-20-Bovine serum albumin	3 g BSA was diluted in 100 mL 1 X TBSTween-20
1 X Tris-Buffered sodium	5 g fat free milk powder was dissolved in 100

chloride Tween-20-Milk	mL TBSTween-20
2 X Sample Buffer	4 % SDS, 25 % glycerol, 10 mM DTT, 100 mM Tris-HCl pH 6.8, 0.01 bromophenol blue. This Buffer was stored at -20 °C and supplemented with 10 mM (final concentration) DTT when used.
5 % Stacking gel	40 % Polyacrylamide, 1 M Tris (pH 6.8), 10 % Ammonium Persulphate, 10 % SDS, 0.002 % TEMED
Ammonium persulphate	A 10 % Stock solution was prepared in dH_2O
Coomassie staining solution	0.25 % Coomassie Brilliant Blue (R250), 40 % Methanol, 10 % acetic acid and 50 % distilled H_2O.
Destaining solution	10 % methanol, 10 % acetic acid in distilled H_2O
Dithiothreitol (DTT)	A 1 M stock solution was prepared in dH_2O and sterilized by filtration (0.22 µm filter), the solution was aliquoted and stored at -20 °C .
Paraformaldehyde (PFA)	4 % (w/v) prepared in half the final volume distilled water, which was heated to 60 °C, the pH was adjusted to 7.4, the solution was made up to final volume, filter sterilized, and stored at 4 °C
Sodium dodecyl sulphate (SDS)	10 % in distilled water.

TBS/Tween-20	20 mM Tris-HCl pH 7.4, 150 mM NaCl and 0.1 % Tween-20. This solution was stored at 4 °C.

Table 2. 5: Equipment used

Instrument	Supplier
ABI StepOnePlus	Applied Biosystems
Axioplan 2 imaging Fluorescence microscope	Zeiss
BASIC 20 pH-Meter	Lasec
BioSepectrum® Imaging system	UVP
Countess™ Automated cell counter	Invitrogen
Eppendorf 5417 R microcentrifuge with rotor F45-30-11	Eppendorf
GeneAmp PCR system 2700	Applied Biosystem
Inverted light Microscope	Nikon TMS-F
Lieca EC3 Digital camera	Lieca Microsystems Ltd.
NanoDrop ND1000	Thermo Scientific
POLAR star Omega Microplate reader	BMG LABTECH

Power Pac Basic HC 300W	Bio-Rad
Qubit®2.0 Fluorometer	Invitrogen
Sorvall TC6 centrifuge H 400	American Instrument Exchange, Inc.
Thermo Fisher Incubator Shaker	Thermo Fisher Scientific Inc
Thermomixer comfort 1.5 mL	Eppendorf
Trans-Blot®Turbo^{TM} Transfer System	BIO-RAD
Vortex Mixer	Labnet international Inc.
Water Jacked CO_2 incubator	Forma Scientific

Table 2. 6: Cell lines and the media

Cell line	Organism	Tissue	Media
MCF-7	Human	Breast carcinoma, mammary gland	DMEM:F12 supplemented with 20 ng/ml EGF, 5 % FBS, 500 ng/ml hydrocortisone and 0.01 mg/ml bovine insulin
MCF12-A	Human	Breast epithelial mammary gland	DMEM:F12 supplemented with 20 ng/ml EGF, 5 % FBS, 500 ng/ml hydrocortisone and 0.01 mg/ml bovine insulin

Table 2. 7: List of the Antibodies

Antibody	Supplier
Goat anti-mouse antibody IgG-HRP: Sc-2055	Santa Cruz Biotechnology, Inc.
Pan-Cadherin antibody (4068)	Cell Signalling (Technology)
Anti GDNF Receptor alpha 1 antibody	Biocom-Biotech

2.2 Methodology

2.2.1 General cell culture procedure

The MCF-12A human mammary gland epithelial cell line has been derived from non-cancerous female breast tissue of a 60-year-old Caucasian female (Immortal human mammary epithelial cell sublines 1993). This cell line was obtained from American Type Culture Collection (ATCC) at passage 59. The MCF-7 human mammary gland breast adenocarcinoma was derived from the metastatic cancer site of a 69-year-old Caucasian female (A human cell line from a pleural effusion derived from a breast carcinoma 1971). This cell line was obtained from ATCC at passage 147. Cryovials were obtained from ATTC and stored at -150 °C.

The media for both cell lines was prepared as follow DMEM:F12 medium containing 2 mM L- glutamine was supplemented with 5 % Fetal Bovine serum (FBS), 100 µg/ml penicillin/streptomycin, 0.01 mg/ml insulin, 500 ng/ml hydrocortisone and 20 ng/ml epidermal growth factor (EGF).

2.2.2 Thawing of cells

Cryovials containing the cells were placed in a water bath at 37 °C and thawed for 1-2 min with constant agitation until the visible ice was melted. The cells were transferred into 15 mL sterile centrifuge tubes containing 9 mL of pre-warmed complete cell culture media. The tubes were centrifuged for 3 min at 1600 rpm in a bench top Sorvall TC6 centrifuge H400. The supernatant was discarded and the cellpellet was re-suspended in 5 mL of fresh pre-warmed complete media.

2.2.3 Culturing of cells

After thawing (2.2.2) the cells were seeded in 25 cm^2 cell culture flasks, and incubated in an atmosphere of 5 % CO_2 at 37 $^{\circ}C$ in a humidified incubator water Jacked CO_2 incubator. The cells were checked continuously for normal morphology, density and absence of bacterial and fungal contaminations using inverted light microscope with the 20 X objectives. The images were taken using a Lieca EC3 camera (Lieca). The media was changed three times a week until the density of the cells was about 80-90 %.

2.2.4 Trypsinization of cells

When the cell density reached 80-90 %, the media was discarded, the cells were washed twice with 5 mL sterile 1 X Phosphate Buffer Saline (PBS). The PBS was discarded; 3 mL of a 0.25 % trypsin solution was added to the flask. The flasks were returned to the incubator for 1- 5 min until cells were detached. Immediately there after 3 mL of pre-warmed medium was added to stop the trypsinization, the cells were transferred into a sterile 15 mL tube and centrifuged at 3000 rpm for 5 min in a Sorvall TC6 centrifuge H 400. The supernatant was discarded and the pellet was resuspended in pre-warmed media. The number and viability of the cells were evaluated as indicated in section 2.2.7 The cells were either sub-cultured or prepared for long-term storage at -150 $^{\circ}C$ in cryovials.

2.2.5 Mycoplasma screening of cell culture

The cells were routinely checked for mycoplasma contamination using MycoFluor^{TM} Mycoplasma detection kit A flask of cells was trypsinized as

described in 2.2.4, and seeded in 6 well plates. Each well contained a cover glass (22 X 22 mm) that was autoclaved prior to placing the cover glass into the well. The cells were incubated until about 75 % confluency was reached. The cells were incubated in 5 X MycoFluor reagent (prepared in growth media) for 10 min at room temperature. The coverslip was placed onto clean microscope. The edges of the coverslip were sealed using wax. The slide was examined by using the Axioplan 2 imaging-Zeiss, fluorescence microscope.

MB Cyclin (Pleurontilin and tetracycline derivatives) was used to eliminate the mycoplasma contamination. The stock solution was made by adding 10 mL of sterile dH_2O to BM Cyclin 1 (25 mg/ml) and BM Cyclin 2 (12.5 mg/ml), and the content was dissolved completely, aliquoted in to a sterile labelled Eppendorf tubes and kept at -20 $^{\circ}$C.

The cells were seeded as explained in sections 2.2.3. The cells were incubated for 3 days in BM Cyclin1 (10 µg/ml) followed by 4 days incubation in BM Cyclin2 (5 µg/ml). These steps were repeated twice and then the cells were checked for mycoplasma as explained above.

2.2.6 Cryovial preservation of cells

Once the cells were approximately 80-90 % confluent and free of fungal, bacterial or mycoplasma contamination. The cells were trypsinzed with 0.25 % trypsin as explained in section 2.2.4.

The cells pellets were resuspended in complete medium containing 10 % Dimethyl sulphoxide (DMSO).

The cells were counted and the viability was determined as described in section 2.2.7. If the viability of the cells were more than 90 %, the cells were considered healthy for freezing and storage. Cryovials were labelled appropriately with complete information such as date of freezing, cell type, the passage number of the cells, the name of the users (initials) and the media for the cell line. For each labelled vial, 1ml of the cells (at a concentration of 2-4 x 10^6 cells per ml) was aliquoted into the vials were sealed. The Cryovials were immediately placed at -150 $^{\circ}$C for long-term storage.

2.2.7 Cell counting with the countess Automated Cell Counter

The number and viability of the cells were determined as per the manufactures instructions using the Countess™ Automated cell counter. After trypsinization the cell suspension (40-50 µL) was mixed with an equal volume of 0.4 % trypan blue stain. The Countess chamber slides were loaded with 10 µL of the sample.

2.3. Stable Isotope Labelling of cells

Stable Isotope Labelling by Amino Acids in Cell Culture (SILAC) is a technique that depends on cellular proteins synthesis to incorporate isotope medium, which contain amino acids, such as lysine containing $^{13}C_6$ or arginine.

Two different populations of the cell (MCF-7 and MCF-12A) were seeded at 1 X 10^5 cells in 75 cm^2 cell culture flasks as described in section 2.2.3. In two separate medium formulations, SILAC was used for comparative, quantitative of proteome analysis between MCF-7 and MCF-12A cells based on incorporation of the labelled amino acid into newly synthesised proteins.

Pierce® SILAC protein quantitation Kit (Thermo scientific) was used for labelling of the cells with specific amino acid, following the manufacture's instructions. The MCF-7 were grown in DMEM:F12 media containing 50 mg/ml of $^{13}C_6$ L-Lysine-2HCl heavy label while MCF-12A cells were grown in the natural amino acid (L-Arginine-HCl) light. Cells were cultured in the SILAC medium for 8 doubling to allow for the incorporation of the amino acids.

2.3.1 Isolation of Plasma membrane proteins

In order to isolate the membrane protein from MCF7 and MCF-12A cells, Pierce® Cell surface protein isolation Kit (Thermo scientific) was used. The media was removed and the cells were washed twice with ice-cold 1 X PBS. The cell surface proteins were tagged with biotin by adding 10 mL biotin solution (EZ-link Sulfo-NHS-SS-Biotin Thermo Scientific) to each flask. The flasks were placed at 4 $^{\circ}$C with orbital shaking for 30 min. The biotin tagging reaction was quenched by adding 500 µL quenching solution (Thermo Scientific) to each flask. Cells were scraped off, transferred into 50 mL conical tubes and centrifuged at 500 xg in a bench top Sorvall TC6 centrifuge H400 for 3 min. The supernatant was discarded and the pellet was resuspended in 5 mL TBS. The suspension was

centrifuged at 500 xg for 3 min and the supernatant was discarded. Lysis Buffer (Thermo Scientific) consisting of 500 µL Lysis Buffer and 20 µL Protein inhibitor (Thermo Scientific) was added to the pellet.

The cells were lysed by sonication at low power for 5 Sec and incubated for 30 min on ice. The solubilization was improved by vortexing the tube every 5 min for 5 Sec using vortex mixer. The cell lysate was centrifuged at 10 000 xg for 2 min at 4 °C using a bench top centrifuge (Eppendorf 5417 R microcentrifuge with rotor F45-30-11). The supernatant (clarified cell lysate) was transferred to a 1.5 mL sterile tube. For isolation of the biotin tagged proteins, a gel column needed to be prepared. The column (Thermo Scientific) was assembled by placing the column into a 2 mL collection tube. The column was capped and 500 µL of NeutrAvodin Agarose was added and the column was centrifuged at 1000 xg for 1 min. The flow-through was discarded. The agarose column was washed by adding 500 µL Wash Buffer to the column and centrifuging it at 1000 xg for 1 min. The washing step was repeated twice.

The bottom of the column was tightly closed off and then the cell lysate was added to the column. The top of the column was closed off and the column was incubated for 60 min at 25 °C with end-over-end mixing using a rotator. Both caps were removed and the column was centrifuged at 1000 xg for 1 min. The flow-through was discarded and the bottom cap was replaced. To remove un-bound proteins (cytosolic and nuclear proteins) the column was washed three times with 500 µL Washing Buffer (Wash Buffer contained protease inhibitor). For protein

elution, 400 µL of sample Buffer (SDS-PAGE sample Buffer with 50 mM Dithiothreitol (DTT) was added to the column and the column was incubated for 60 min at 25 °C with end-over-end mixing using a rotator. The bottom cap of the column was removed and the column was placed into a new collection tube. The column was centrifuged at 1000 xg for 2 min. The membrane proteins were quantified described in section 2.3.1.1 and kept at -20 °C for future use.

2.3.1.1 Protein Quantification

The protein concentrations of the samples were determined using the Qubit® 2.0 Fluorometer (Invitrogen), according to the manufacture's instructions. All reagents were equilibrated to 25 °C before use. The following steps were taken in order to measure the concentration of protein using the Qubit™protein assay kit.

A 200 µL Qubit™ Working Solution for each standard and sample was prepared in a clear, thin-walled 0.5 mL PCR tube, by diluting the Qubit™ reagent in the Qubit Buffer (1:200). The assay tubes were then prepared as shown in table 2.8

Table 2. 8: Reagents utilized for protein quantification

	Standard 1	Standard 2	Standard 3	Sample tubes
Working solution	190 µL	190 µL	190 µL	190 µL
Standards	10 µL	10 µL	10 µL	-
Sample	-	-	-	10 µL
TOTAL volume	200 µL	200 µL	200 µL	200 µL

The standards and samples assay tubes were mixed by vortexing, followed by incubation for 2 min at 25 °C. The protein concentration for each standard was quantified and used to generate a calibration curve using the Qubit® 2.0 Fluorometer. The protein concentration for each sample was then extrapolated from the standard curve.

2.3.1.2 One-dimensional Sodium-dodecyl-sulfate-polyacrylamide gel electrophoresis (1D SDS-PAGE)

The concentration of the proteins were determined using Qubit® 2.0 Fluorometer (invitrogen) as described in section 2.3.1.1, protein samples were loaded onto gels at 100 µg/µL following heating at 95 °C for 10 min.

2.3.1.3 Preparation of PAGE

Proteins were separated by denaturing SDS-PAGE according to the method of (Laemmli 1970). Briefly the gels were made from 40 % stock of premix 37.5:1: bisacrylamide (Serva GMbH). The separating gel consisted of 12 % acrylamide: bisacrylamide (37.5:1), 0.375 M Tris-HCl, (pH 8.8), 0.5 % ammonium persulphate, 0.1 % SDS and 0.02 mL (0.198 %) TEMED. The stacking gel consisted of 5 % acrylamide: bisacrylamide (37.5:1), 0.125 M Tris-HCl (pH 6.8), 0.5 % ammonium persulphate, 0.1 % SDS and 0.01 mL (0.198 %) TEMED. The gels were prepared in 1 mm Hoeffer dual gel casters, and about 10 mL were enough for one gel. The separating gel was poured to about 1 cm below the level of the wells of the comb and about 2 mL of isopropanol was overlaid on top of the separating gel and the gel was allowed to set. Once set, the distilled water was poured off and the gel rinsed with deionized water. Stacking gel (~5 mL) was poured on top of the separating gel and a comb was inserted into the stacking gel. The gel was allowed to solidify. When the stacking gel was set, the comb was removed and the gel was transferred to the Mighty Small apparatus (Hoeffer).

The protein samples were prepared by mixing the samples1:1 (v:v) with 2 X SDS Sample Buffer. The samples were boiled for 10 min, centrifuged for 5 min at 10 000 xg, the samples were loaded into gels and electrophoresis in 1 X SDS Electrophoresis Buffer at 100 V (constant voltage) for 15 min. The voltage was increased to 120 V (constant voltage) until the *bromophenol* dye reaches the bottom of the gel. One the gels was transferred onto a polyvinylidene fluoride (PVDF) membrane while the other gel was stained with Coomasssie stain.

2.3.1.4 Staining and Destaining of the gels

After electrophoresis, the gel was removed from the casing. And incubated in Coomassie Stain Solution for 1 hour on an orbital shaker. The gel was destained over night in Destaining Solution on an orbital shaker at 25 °C. The gel was rinsed in dH_2O and the gel images were taken by using the BioImaging Systems.

2.3.1.5 Western Blotting analysis

Proteins samples were electrophoresed on SDS PAGE as described in section 2.3.1.3. The proteins were transferred onto PVDF membrane (Bio-Rad) using Bio-Rad MiniProtein Trans Blot system. Before transfer the membrane was equilibrated by pre-wetting the membrane in 100 % methanol for 30 Sec. the membrane was washed with distilled water for 5 min. the membrane, sponges, gels Whatman paper were equilibrated in pre-chilled transfer Buffer for 15 min and was assembled. Protein was electro-blotted onto the PVDF membrane at 100 V constant for 90 min in pre-cooled Transfer Buffer or at 4 °C, 35 V (constant voltage) overnight in pre-cooled Transfer Buffer using Power Pac Basic HC 300W (Bio-Rad).

2.3.1.6 Probing the blot with antibodies

After the protein was transferred into PVDF membrane. The PVDF membrane was incubated (with shaking) in the blocking Solution at room temperature for 60 min and then rinsed with TBStween-20 (3 times for 50 min). The PVDF membrane was then incubated (with shaking) over night at 4 °C in human Pan-Cadherin primary antibody (1:750 dilution prepared in TBSTween-20). The

membrane was rinsed for 5 min three times with 10 mL TBSTween-20. Then the membrane was incubated for 60 min in Goat anti mouse secondary antibody IgG-HRP (diluted at 1in 3000 times in TBSTween-20). The membrane was rinsed three times (7 min each) with TBSTween-20 for 7 min each while shacking. The membrane was placed into dark place in UVP BioImaging system making sure the side with the blotted protein is upright. Super Signal West Pico Chemiluminescent substrate was prepared by mixing 1:1 of solution A and Solution B as per manufacture's instructions (Thermo Scientific) and the substrate mixture was applied over the membrane making sure that the membrane was covered with substrate. The membrane was imaged using UVP Transilluminator and photographed with BioImaging Systems Chemi HR 410 Camera with different time 1-5 min.

2.3.1.7 Preparation of protein samples for by LC MS/MS analysis

The MCF-7 and MCF1-2A cell lines were cultured and labelled with $^{13}C_6$ L-Lysine-2HCl and L-Arginine-HCl amino acids, respectively as described in section 2.3. The plasma membrane proteins were extracted as described in section 2.3.1. 100 µg of the protein samples were resolved in 12 % SDS-PAGE following the method described in section 2.3.1.3, the gel (Figure 2.1) was stained using Coomassie staining solution for 60 min.

After destaining the gel in destaining solution, the lane containing the protein sample was cut into 11 sections using a sterile blade. The samples were transferred into sterile 1.5 mL Eppendorf tubes and sent (on dry ice) for LC

MS/MS analysis to the Central proteomics facility Sir William Dunn Pathology School, South Parks Rd, University of Oxford, UK.

2.3.1.7.1 Proteomics data analysis

Data analysis was preformed using Thermo Proteome Discoverer software (version1.4). Membrane, extracellular, cell surface proteins and unclassified proteins were generated using analysis software to identify global function of all available proteins. All the information such as IPI human identification numbers, score, coverage, cellular component and the quantitative of the proteins were received in excel spreadsheet file.

The data was analyzed using in silico bioinformatics tools (Figure 2.1).

The following tools were used: EMBL-EBI Genes Expression Atlas (http://www.ebi.ac.uk/gxa/home;jsessionid=257BDF22B95699B6F25E948C1F9C33C1), iHOP (http://www.ihop-net.org/UniPub/iHOP/), TiGER (http://bioinfo.wilmer.jhu.edu/tiger/), Uniprot (http://www.uniprot.org), The Maxquant Database (MaxQB) (http://maxqb.biochem.mpg.de/mxdb/) and GeneHub-GEPIS (http://research-public.gene.com/Research/genentech/genehub-gepis/genehub-gepis-search.html).

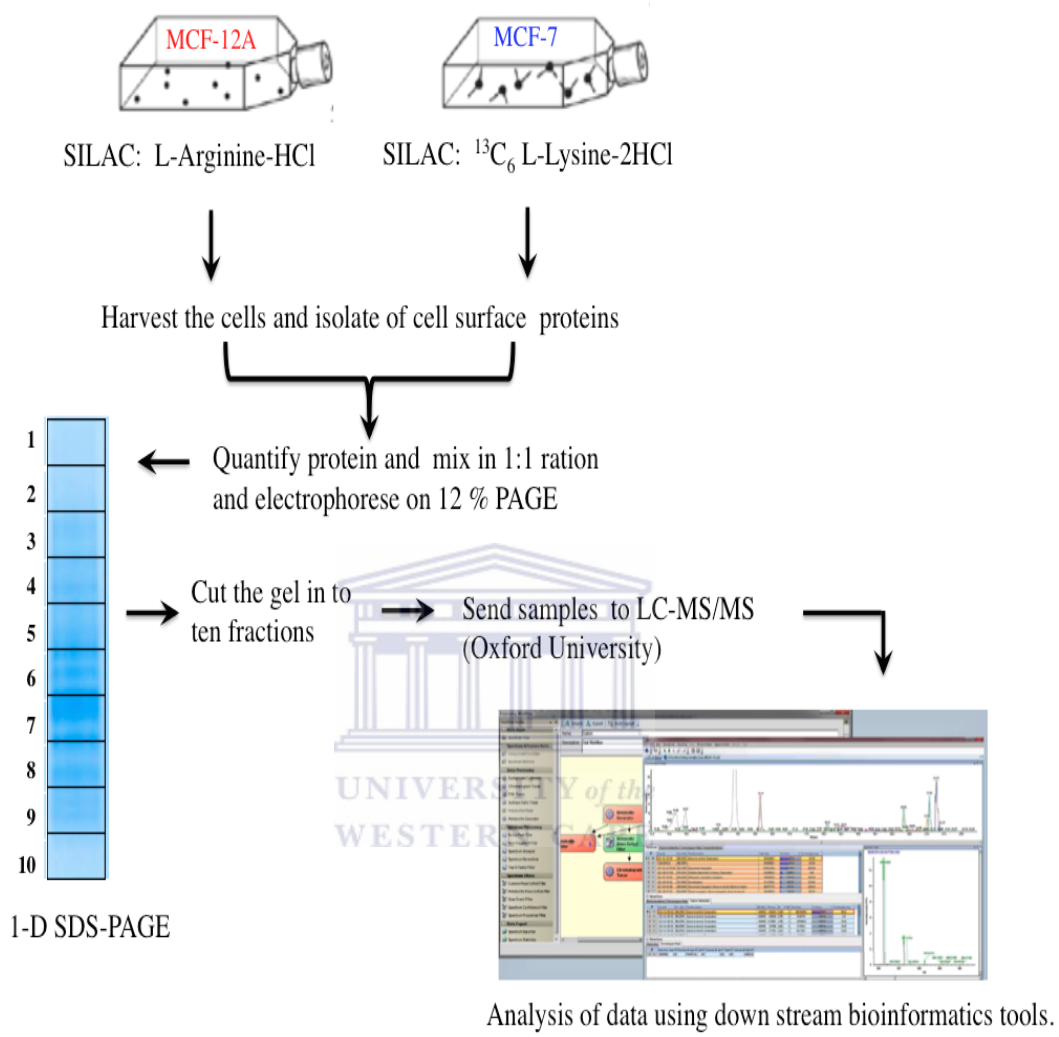


Figure 2. 1: Experimental strategy for the SILAC method for MCF-7 and MCF-12A. The cells were grown in media containing heavy and light isotope, respectively. The cell surface proteins were isolated and quantified, mixed in 1:1 a ratio and resolved by 1D SDS-PAGE. The protein fractions were excised and send to the Central proteomics facility Sir William Dunn Pathology School, South Parks Rd, University of Oxford, UK for LC-MS/MS analysis. The data received from the Central proteomics facility was analyzed by using bioinformatics tools.

2.4. Identification of known biomarkers for breast cancer

Computational tools were used to determine genes that were differentially expressed in breast cancer. The search engines used were PubMed, PubMed Central, Google Scholar, Science direct, springer Link and nature. The search terms included breast cancer, biomarkers and differential expression of breast cancer genes. The expression of identified breast cancer genes were further analyzed for using Database for Annotation, Visualization and Integrated Discovery (DAVID) version 6.7. Nineteen known biomarker genes were identified from literature searches and these were further analyzed using DAVID to determine the expression profile of these genes in different tissues and to validate their expression levels in breast cancer.

2.5 Primers design for qRT-PCR

In this study primers for the selected genes listed in Table 2.9 were designed to amplify a region of the coding sequence of each target gene using default parameters with the exception of PCR product size parameter, which was set to select primers to amplify product in the range 100 - 233 bp. The online Primer3 Plus software (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi) was used to design the primers. The mRNA sequences for all the genes were obtained from the National Centre for Biotechnology Information (NCBI). The primers were purchased from Inqaba Biotech (Pretoria, South Africa).

Table 2. 9: Oligonucleotide primers sets for genes encoding biomarkers. The expected product size and calculated annealing temperature for qRT-PCR reaction is also indicated.

Gene name	Ensemble Accession Number	Primer Name	Primer Sequence (5′-3′)	Product Size (bp)	T_A (°C)
FOXA1	ENSG00000129514	F P *FOXA1*	AAGGCATACGAACAGGCACT	161 bp	59
		R P *FOXA1*	GTGTTTAGGACGGGTCTGGA		
CK19	ENSG00000171345	F P *CK19*	TCGAGGGACAGGAAGATCAC	158 bp	59
		R P *CK19*	ATTGGCAGGTCAGGAGAAGA		
HER-2	ENSG00000146648	F P *HER-2*	AGTACCTGGGTCTGGACGTG	194 bp	59
		R P *HER-2*	CTGGGAACTCAAGCAGGAAG		
GFRA1	ENSG00000151892	F P *GFRA1*	AGACCACCACTGCCACTACC	219 bp	59
		R P *GFRA1*	TTGTGGTTATGTGGCTGGAA		
EPHB4	ENSG00000196411	F P *EPHB4*	TATTCGGACAAACACGGACA	233 bp	59
		R P *EPHB4*	TCTTGATTGCCACACAGCTC		
NENF	ENSG00000117691	F P *NENF*	GGTAGCCAAGATGTCCTTGGA	186 bp	59
		R P *NENF*	TCAGGCTTGAAGTCCAGGTTA		
PTGES2	ENSG00000148334	F P *PTGES2*	GCAGGGCTGAGATCAAGTTC	175 bp	59
		R P *PTGES2*	GCCTTCATGGCTGGGTAGTA		
GAPDH	ENSG00000111640	F P *GAPDH*	ACCCACTCCTCCACCTTTG	178 bp	59
		R P *GAPDH*	CTCTTGTGCTCTTGCTGGG		

Abbreviations: 5′= 5 prime; 3′= 3 prime; bp = base-pair F P = forward primer; R P = revers primer; TA = annealing temperature; A = adenosine; C = cytosine; G = guanosine; T = thymidine.

The primers were reconstituted in 1 X TE Buffer and stored at -20 °C until required.

2.6 RNA Isolation:

The isolation of total RNA was carried out with the use of NucleoSpin® RNA/protein Macherey-Nagel kit, according to the manufacture's data sheet.

The MCF-7 and MCF-12A cell lines were grown in DMEM:F12 media and upon reaching approximately 90 % confluency as described in section 2.2.3 The cell cells were trypsinzed with 3 mL 0.25 % trypsin solution as described in section 2.2.4. The pellet was washed with cooled 1 X PBS to remove trace trypsin prior to RNA isolation

The cells were lysed by adding 350 µL RP1 Buffer and 7 µL TCEP to the cells. The cells were vortexed vigorously for 1 min and the lysate was filtered through a NucleoSpin® Filter spin column (violet ring) into a collection tube by centrifugation at 11 000 xg for 1 min at 25 °C using an Eppendorf 5417 R microcentrifuge a F45-30-11 rotor. NucleoSpin® Filter was discarded and 350 µL 70 % ethanol was added to the filtrates (homogenized lysate). For each preparation, a NucleoSpin® RNA/protein column (light blue ring) was prepared by placing the column in a collection tube.

The lysate was added and centrifuged at 11 000 xg for 30 Sec at 25 °C. The NucleoSpin® RNA/protein column was placed into a new collection tube. The

RNA was purified by adding 350 µL of Membrane Desalting Buffer (MDB) to the column and centrifuged the column at 11 000 xg for 1 min. The DNA present on the column was digested by adding 95 µL of rDNase reaction mixture (10 µL reconstituted rDNase and 90 µL reaction Buffer rDNase) to column and incubating the column at 25 °C for 15 min. The first wash of the column was achieved by adding 200 µL RA2 Buffer to NucleoSpin® RNA/protein column and centrifuging at 11 000 xg for 30 Sec. The column was placed into a new 2 mL collection tube. The second wash was achieved by adding 600 µL RA3 Buffer onto NucleoSpin® RNA/protein column and centrifuging at 11 000 xg for 30 Sec and the flow-through was discarded and the column was returned to the collection tube. A volume of 250 µL RA3 Buffer was added to the NucleoSpin® RNA/protein column and centrifuged at 11 000 xg for 2 min.

The NucleoSpin® RNA/protein column was placed into a sterile 1.5 mL RNase-DNase free collection tube. RNase free water (40 µL) was added to the column and centrifuged at 11 000 xg for 1 min. The flow-through collected contained the total RNA. The RNA sample was quantified as described in section 2.6.4. The sample was stored at -80 °C for further use.

2.6.1 Preparing 1 % gel agarose for RNA electrophoresis

A 1 % (w/v) agarose gel was prepared by adding 1 g of agarose powder to 1 X TBE Buffer. The mixture was heated until the agarose was completely melted. The mixture was cooled down to about 50 - 60 °C. Upon cooling a GelRed nucleic acid stain (BioTIUM) was added to a final concentration 0.03 X. The casting tray

and the comb were washed with Sodium hypochlorite. The gel solution was poured and solidified in a horizontal gel-casting tray.

2.6.2 Agarose gel electrophoresis of RNA

The RNA was electrophoresed on 1 % agarose gels as mentioned in section 2.6.2. To assess the integrity of the RNA, the samples were prepared by adding 3 µg of RNA to and 1 µl of 6 X loading Buffer (Fermentas). The samples were heated at 65 °C for 4 min using Thermomixer comfort 1.5 mL before loading into the wells of the gel. GeneRuler 1Kb DNA ladder (Fermentas) was loaded to estimate the size of RNA fragments. The samples were loaded into 1 % agarose gel gel buffered with 1 X TBE Buffer (pH ~ 8.3) and electrophoresed at 90 V constant for 60 min using a Power Pac Basic HC 300W system (BIO-RAD). The gel was imaged using the UVP Transilluminator and photographed with the BioImaging Systems Chemi HR 410 Camera using SYBER Gold filter at 302 nm.

2.6.3 Synthesis of cDNA

cDNA was synthesized from total RNA of MCF7 and MCF12A using transcriptor First Strand cDNA synthesis Kit (Roche) according to manufacturer's data sheet instructions. Briefly, the following components were mixed added as follows; Achored-oligo (dT) (2.5 µM) was added to 1 µg of RNA and made up to a total volume of 13 µL was made by adding PCR grade water. The template-primer mixture was denatured by heating the tube at a 65 °C for 10 min in a thermal block cycler with a heated lid (GeneAmp PCR system 2700, Applied Biosystems) and then placed on ice. Transcriptor reverse transcriptase reaction Buffer (1 X),

Protector RNase Inhibitor (20 U), Deoxynuleotide Mix (1 mM) and Transcriptor Reverse Transcriptase (10 U) were added, respectively, to make a final volume of 20 µL (appendix Table 1). The sample mixture was collected at the bottom of the tube by brief centrifugation, after which the tube was heated at 55 °C for 30 min. Transcriptor Reverse Transcriptase was inactivated by heating at 85 °C for 5 min, after which the tube was put on ice to stop the reaction. The cDNA was quantified and kept at -20 °C until further use.

2.6.4 Quantification of RNA and cDNA

2.6.4.1 Determination of RNA quality and quantity using Agilent BioAnalzyer

For RNA Quantification and determination of RNA integrity number (RIN), RNA samples were analyzed on an Agilent 2100 BioAnalyzer (Agilent Technologies) at the Central Analytical Facility (CAF) in the Department of Genetics, Stellenbosch University (appendix Figure: 1 and 2). Briefly, 1 ul of extracted total RNA was loaded on a Bioanalyzer chip, and the quality and quantity was measured using the default parameters for RNA on the BioAnalyzer. This was done for both MCF-7 and MCF-12A.

2.6.4.2 Determination of cDNA and RNA concentration by Nanodrop

The cDNA and RNA concentrations were determined on a Nanodrop-ND1000 Spectrophotometer (Thermo Scientific). Briefly, the Nanodrop was setup for nucleic acid, and then the calibration was done using distilled water. One microletter of this sample was loaded for measurement onto lower measurement pedestal. The sample was measured using absorbance at 260/280 to assess the

purity of DNA and RNA nm and the ratio of RNA was ~ 2.0 and for cDNA was ~ 1.8. The concentration of the cDNA and RNA was also determined using this method. The measurement of each sample was repeated three times.

2.6.5 Quantitative Real Time Plymerase Chain Reaction (qRT-PCR)

The cDNA from MCF-7 and MCF-12A cell lines was synthesized through reverse transcription method as described in Section 2.3.6, and used as a template for amplification of genes using the primers listed in table 2.9.

qRT-PCR reactions were performed on a ABI Prism StepOnePlus™ Instrument (Applied Biosystems) using a KAPA SYBR® FAST for ABI Prism® PCR Mix (2 X) (Kapa Biosystems), according to manufacturer's datasheet and was optimized for all primer combinations. Serial dilutions of known cDNA concentration was made from a stock solution and ranged from 200 ng to 0.02 ng of input RNA. This was done for each gene to construct a standard curve, which could be used to calculate PCR efficiency (indicated in Table 2.10) and to determine template concentrations. 200 nM Forward primer and 200 nM Reverse primer were used for each reaction in a final reaction volume of 20 μL as indicated below.

Table 2. 10: Quantitative Real-Time PCR reaction reagents

Reagent	Final concentration
KAPA q-PCR Mix (2 X)	1 X
F. Primers	200 nM
R. Primers	200 nM
Template (cDNA)	20 ng/ µL
PCR-Grade deionized H_2O	Variable~
Final Volume	20 µL

The qRT-PCR was set up according to the standard three-step protocol for qRT-PCR (denaturation, annealing and extension) and subjected to 45 cycles consistently. These steps had some variations in annealing temperature depending on nucleotide sequences of the gene of interest. The reactions were performed for all selected genes with two housekeeping genes for both cell lines MCF-7 and MCF-12A. The reactions were performed in a 96-well PCR plate on an ABI Step One Plus; at the Central Analytical Facility at the University of Stellenbosch, the system was cycled according to the standard parameters (appendix Table 2). The reactions were prepared according to the following steps table 2.10. The products

were stored at – 20 °C for possible future downstream analysis or electrophoresed on 1 % agarose gels.

2.7 Immunohistochemistry (IHC)

2.7.1 Cell cultures

The cells were seeded on positive charged glass slides (76 mm x 26 mm; 1.0 mm to 1.2 mm thick). The slides were autoclaved prior to seeding the cells. The cells were cultured as described in section 2.2.3 until about 80 % confluency were reached. The cells were washed three times with 1 X PBS, 3 min each. The cells were fixed in 4 % paraformaldehyde (PFA) at 25 °C for 20 min. The cells were washed with 1 X PBS, for 5 min. The cells were stained with Anti-GFRA1 antibody (1:500 dilution). All staining, washing and antibody incubations were performed in a Leica Bond Autostainer (Leica) at the Histology Laboratory at Tygerberg Hospital.

2.7.2 Breast cancer tissues

The breast cancer tissues were obtained from National Cancer Institute Misurata (Libya). Two experienced pathologists (Dr Y. Topov and Dr Fathi Abdalla) performed the diagnosis of the tissue samples. Three tissue samples diagnosed as stage I, II and III was used in this study.

2.7.2.1 Ethical Approval

The Ethics Committee of the Stellenbosch University granted ethics approval for the use of the human samples in this study (Reference Number: N13/08/118).

2.7.3 Fixation and paraffin wax embedding of the tissues

Briefly; after the cancer tissues was removed surgically the samples were kept over night at 25 °C in 10 % formalin solution. After the fixation period, the tissues were moved to labelled histology cassettes.

2.7.4 Tissues processing

The tissues were processed through series of solutions using as described in Table 2.11.

Table 2. 11: Processing protocol for FFPE tissues

Step	Solution	Temperature	Time
1	70 % Ethanol	25	60 min
2	90 % Ethanol	25	120 min
3	100 % Ethanol	25	60 min
4	100 % Ethanol	25	60 min
5	100 % Ethanol	25	6 hrs
6	100 % Xylene 1	25	60 min
7	100 % Xylene 2	25	2 hrs
8	Paraffin wax	58	60 min
9	Paraffin wax	58	60 min

2.7.5 Embedding

Processed tissues were embedded in paraffin wax at 60 °C to obtain tissues blocks. The tissues blocks were kept at 25 °C until sectioning was performed.

2.7.6 Sectioning

The tissues blocks were quickly frozen and then sectioned using a Leica RM 2125 RT microtome (Leica) to obtain uniform 5µm section. The sections were placed on positive charged glass slides. Three slides were prepared for each tissue sample.

2.7.7 Haematoxylin and eosin (H&E) staining

Prior to staining, slides were placed in an incubator for the wax to melt of the tissues. The tissues or cell cultures were stained using a Leica Bond Autostainer (Leica) at the Histology Laboratory at Tygerberg Hospital as shown in Table 2.12

Table 2. 12: The processes and the solutions for H&E staining.

Step	Solution	Time	Repetitions
]	Xylene	10 min	X 2
2	Ethanol (99 %)	5 min	X 2
3	Ethanol (95 %)	2 min	X 1
4	Ethanol (70 %)	2 min	X 1
5	Distilled water	5 sec	X 1
6	Haematoxylin	8 min	X 1
7	Running water	5 min	X 1
8	Ethanol (1 % acid alcohol)	30 sec	X 1
9	Running water	1 min	X 1
10	Ammonia (0.2 %)	45 sec	X 1
11	Running water	5 min	X 2
12	Ethanol (95 %)	10 dips	X 1
13	Eosin	45 sec	X 1
14	Ethanol (95 %)	5 min	X 2
15	Xylene	5 min	X 2

2.7.8 Immunohistochemical (IHC) staining

The slides were stained with Anti-GFRA1 antibody (1:500 dilution). All staining, washing and antibody incubations were performed in a Leica Bond Autostainer (Leica) at the Histology Laboratory at Tygerberg Hospital as shown in Table 2.13.

Table 2. 13: The IHC staining protocol

Step	Type	Incubation Time	Temperature	Dispense Type
1	Peroxide Block	5 min	Ambient	Selected vol.
2	Bond Wash Solution	0 min	Ambient	Selected vol.
3	Bond Wash Solution	0 min	Ambient	Open
4	Bond Wash Solution	0 min	Ambient	Selected vol.
5	Primary Antibody	15 min	Ambient	Selected vol.
6	Bond Wash Solution	0 min	Ambient	Selected vol.
7	Bond Wash Solution	0 min	Ambient	Selected vol.
8	Bond Wash Solution	0 min	Ambient	Selected vol.
9	Post Primary	8 min	Ambient	Selected vol.
10	Bond Wash Solution	2 min	Ambient	Selected vol.
11	Bond Wash Solution	2 min	Ambient	Selected vol.
12	Bond Wash Solution	2 min	Ambient	Selected vol.
13	Polymer	8 min	Ambient	Selected vol.

14	Bond Wash Solution	2 min	Ambient	Selected vol.
15	Bond Wash Solution	2 min	Ambient	Selected vol.
16	Deionized Water	0 min	Ambient	Selected vol.
17	Deionized Water	0 min	Ambient	Selected vol.
18	Mixed DAB Refine	10 min	Ambient	Selected vol.
19	Deionized Water	0 min	Ambient	Selected vol.
20	Deionized Water	0 min	Ambient	Selected vol.
21	Deionized Water	0 min	Ambient	Selected vol.
22	Hematoxylin	5 min	Ambient	Selected vol.
23	Deionized Water	0 min	Ambient	Selected vol.
24	Deionized Water	0 min	Ambient	Selected vol.
25	Deionized Water	0 min	Ambient	Selected vol.

2.7.9 Rehydration

After the staining the tissues or cells, the samples were rehydrated and cleared manually as shown in Table 2.14.

Table 2. 14: The rehydration protocol

Step	Solution	Duration
1	70 % alcohol	5 dips
2	96 % alcohol	5 dips
3	96 % alcohol	5 dips
4	99 % alcohol	5 dips
5	99 % alcohol	5 dips
6	Xylene	Dip for 1 min
7	Xylene	Dip for 1 min

2.7.10 Mounting

PDX mounting medium was used to cover the glass slide. The slide was left to dry at 25 °C.

2.7.11 The evaluation of the protein expression

The protein expression in cell lines or cancer tissues was evaluated by microscopy using a Zeiss Microscope with objectives at 20 X and 40 X magnifications for FFPE slides and 100 X for both cell cultures.

2.8 Design and synthesis of the Oligonucleotides with/ Biotion TEG of Molecular beacon probe.

The use of MBs for RNA detection and localization requires designing the MB for specific region of RNA inside the cell. Designing MB involves four steps, the loop, stem, fluorophore and the quencher. Loop region is the primary concern for designing MBs for intracellular RNA.

The loop of MBs of specific gene was designed by selected an appropriate target region on the mRNA sequence obtained from the NCBI database. The probe sequence, 20 bp in length, was designed to be complementary to a target sequence. The obtained sequence was queried against the NCBI sequence database using BLAST 2.2.28 (www.ncbi.nlm.nih.gov/BLAST/) to confirm its specificity for the target sequence.

The stem portion sequence was designed to contain a 5-7 bp region that was complementary to each other to allow hybridization to one another. This region was confirmed to be non-complementary to the loop sequence to prevent hybridization to the loop.

The sequences were sent to IDT (Integrated DNA technology) for synthesizing and modification with Biotin-tetraethylene glycol (TEG-) at 3′ end. The MBs were synthesized to nmole scale and purified by HPLC. The oligos were prepared in 1 X TE buffer (pH 8.0) as per the manufactures instructions. Oligonucleotides which represented the complementary sequence of the MB were designed and synthesized at Inqaba Biotech (Pretoria, South Africa).

The oligonucleotides were designed using the Primer3Plus software package (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi). It was designed to hybridize to 20 base sequence based on selected breast cancer gene sequences. The nucleotide codon sequence for each target gene was translated to protein sequences by using Expasy bioinformatics resource Portal (http://web.expasy.org/translate/) to ensure that the oligo sequence portion is located in the same protein sequence. This analysis was also performed to ensure that the sequences were completely matched between two sequences (the query sequences to a target sequences). The primers sequences were then selected and searched using BLASTN 2.2.28 (www.ncbi.nlm.nih.gov/BLAST/). The sequences were sent to IDT (Integrated DNA technology) for synthesizing and modification with Biotin -tetraethylene glycol (TEG-) at 3′ end. The MBs were synthesized to

nmole scale and purified by HPLC. The oligos were prepared in 1 X TE buffer (pH 8.0) as per the manufactures instructions. Oligonucleotide, which represented the complementary sequence of the MB, were designed and synthesized at Inqaba Biotech (Pretoria, South Africa).

2.8.1 Activation of molecular beacon oligo's

The oligo's (MB-CK19/Short and MB-CK19/Long) were activated by re-suspending the oligo's in 1 X TE buffer (pH 7.5), containing 10 mM DTT. Using the Amicon Ultra-0.5 centrifugal filters, CK19 Oligo (500 µL) was centrifuged at high speed at 4 °C for 5 min. The flow-through was discarded. The remaining amount of Oligo was added and the process was repeated yielding about 30 µL of the concentrated activated Oligo in the filter. For the washing step, 500 µL of 1 X TE was added to the activated Oligo and centrifuged at high speed for 10 min at 4 °C. To recover the concentrated Oligo, the centrifugal filter was placed upside down into clean microcentrifuge tube and centrifuged at 2000 RPM for 2 min at 4 °C. The volume of the recovered concentrated Oligo was measured and finally made up to total volume of 640 µL to achieve concentration of 100 µM of the CK19-Oligo. The Oligo was kept at -20 °C until used.

2.8.2 Preparation of MB-CK19/Short and MB-CK19/Long molecular beacon

Modified Iowa Black-Oligo (2 µL, 100 µM) with final concentration 200 nM was added to Qdot® 525 ITK™ Carboxyl Quantum Dot (100 µL, 1 000 pM) with final concentration 100 pM, and dissolved in PBS (892 µL, 1 X). The solution was incubated for 5 min at 37 °C with shaking. ***N-Hydroxysuccinimide (NHS)*** **(**3

μL, 100 μM) was added and incubated for 5 min with shaking at 37 °C. Then 1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC) (3 μL, 100 μM) incubated at 37 °C for 60 min with shaking. The Qdot-MB-CK19 was kept at 4 °C until used.

2.8.3 Fluorescence measurement for Qdot

To measure the fluorescence intensity of Qdot-MB-CK19 used POLAR star Omega Microplate reader. The fluorescence intensity of molecular beacon was measured with excitation 355 nm and emission 525 nm.

CHAPTER 3

3.1 Identification and validation of biomarkers for breast cancer

3.1.1 Introduction

One of the limitations of current molecular biomarkers for breast cancer is that most of the biomarkers can only be used in invasive diagnostic systems. There is therefore a need to discover new molecular biomarkers for breast cancer that can facilitate the development of non-invasive diagnostic techniques. Currently used molecular diagnostic methods for breast cancer such as FISH, IHC and qRT-PCR requires biopsy samples, which make the diagnosis highly invasive. Proteins that are present on the surface of cancer cells or are shed by cancer cells would make ideal biomarkers for the development of non-invasive diagnostic procedures, since these biomarkers can be detected in bodily fluids such as blood or urine.

Although numerous studies have been performed to identify biomarkers for diseases such as cancer, the most common methods used involved functional genomic studies using techniques such as DNA microarrays (Zhang *et al.,* 2013). DNA microarrays have made it possible for researchers to estimate the level of expression of thousands of genes in a sample of cells or tissue and identify genes that can be linked to diseased states. Researchers in the biomedical research fields and those working in the clinical diagnostic fields have been quick to adopt this powerful research tool. Due to the vast amount of data obtained from these studies

accurate analysis and interpretation of the data obtained from DNA microarrays studies have provided unique challenges to investigators.

There is some evidence that microarray technology can produce inconsistent results mainly due to the associated computational and statistical difficulties. For example, in diffuse large B-cell lymphoma it was shown that the survival rate for this disease can be predicted based on the expression levels of six genes (Lossos *et al.*, 2004, Tarca *et al.*, 2009). However, microarray analysis could not identify the same six-gene panel in diseased samples even though PCR results demonstrated that these genes were up-regulated in the diseased samples (Rosenwald *et al.*, 2002, Shipp *et al.,* 2002).

DNA microarrays assess gene expression at the level of transcription, while proteomics assess gene expression at a translational level. Since it is proteins and not mRNA molecules that carry out biochemical processes within cells and are directly involved in the pathology of diseases, proteomics is a more reliable technology to use for the identification of biomarkers.

The aim of this chapter was therefore to identify cell surface proteins that are differentially expressed between a human breast cancer cell line (MCF-7) and a non-cancerous human breast cell line (MCF-12A), since these proteins can be potential biomarkers for breast cancer. The first objective towards this aim was to isotopically label cellular proteins of MCF-7 and MCF-12A cells in cell culture using SILAC. The second objective was to extract the cell surface proteins from

the labelled cells and to perform a comparative proteomics analysis of the expression levels of the proteins isolated from MCF-7 and MCF-12A cells. The third objective was to use online bioinformatics tools such as GeneHub-GEPIS to investigate the expression levels of the putative biomarkers in cancer tissues. The fourth objective was to confirm the up-regulation of some of these putative biomarkers in the MCF-7 cells by expression analysis using qRT-PCR and IHC.

3.2 Results and Discussion

3.2.1 The isotopic labelling of MCF-7 and MCF-12A cells

MCF-12A is a non-cancerous cell line that was derived from human mammary gland epithelial cells of a 60-year old Caucasian female, while the MCF-7 cell line is a human mammary gland breast adenocarcinoma that was derived from the metastatic cancer site of a 69-years old Caucasian female (information obtained from ATCC).

MCF-12A cells exhibit typical luminal epithelial morphology, three-dimensional growth in collagen, and form domes in confluent cultures. The cells are positive for epithelial cytokeratins 8, 14 and 18, and negative for cytokeratin 19. MCF-7 cells are useful for in vitro breast cancer studies because the cell line has retained several ideal characteristics particular to the mammary epithelium. These include the ability to process estrogen, in the form of estradiol, via estrogen receptors in the cell cytoplasm. This makes the MCF-7 cell line an estrogen receptor (ER) positive control cell line (information obtained from ATCC).

The successful SILAC labelling of cells in cell culture medium requires that the cells be cultured for a period equivalent to 6 to 8 doubling events (Ong *et al*., 2002, Munday *et al*., 2012). To optimise the culture period in SILAC medium so that optimal labelling can be achieved, MCF-7 cells were cultured in "heavy" labelling media (media containing $^{3}C_{6}$ L-Lysine-2HCl) for four weeks.

The cells were harvested as described in section 2.3.1 and the total cellular proteins were isolated from the cells as described in section 2.3.1. The protein sample was quantified using the Qubit® 2.0 Fluorometer (Invitrogen) and 20µg of the protein sample was analysed by 1D SDS PAGE (Figure 3.1) as described in section 2.3.1.3. To confirm that the cells were sufficiently labelled, protein bands that stained intensely with Coomassie Brilliant blue (Figure 3.1, lane A) was excised from the gel and the gel slice was sent to the Central Proteomics Facility at the Sir William Dunn Pathology School (University of Oxford, UK) for LC-MS/MS analysis. This analysis demonstrated MCF-7 cells were sufficiently labelled in SILAC medium after a four-week culture period. Consequently, MCF-7 and MCF-12A cells were metabolically labelled in cell culture with "heavy" $^{3}C_{6}$ L-Lysine-2HCl and "light" L-Arginine-HCl stable isotopes, respectively (as described in section 2.3).

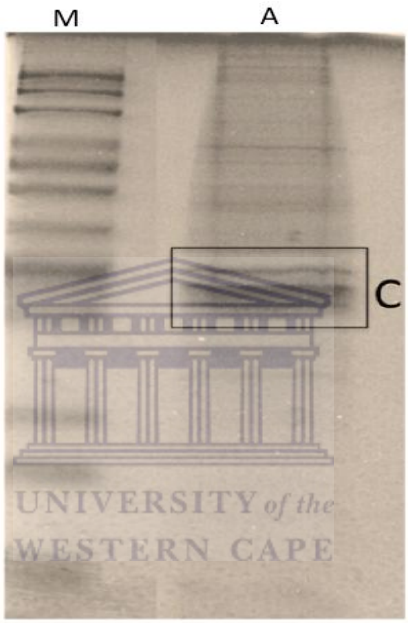


Figure 3. 1: SDS-PAGE of SILAC labelled proteins isolated from MCF-7 cells. MCF-7 were labelled with $^{3}C_{6}$ L-Lysine-2HCl. Lane M represents the protein ladder. Lane A shows the proteins isolated from MCF-7. Section C represents the area in the gel that was excised for LC-MS/MS analysis.

3.2.2 The extraction of cell membrane proteins

The Pierce® Cell surface protein isolation Kit from Thermo Scientific was used to isolate cell surface proteins from MCF-7 and MCF-12A cells following SILAC labelling as described in section 2.3.1. This Kit allows for the isolation of cell surface proteins based on biotin-avidin affinity. The cell surface proteins are first biotinylated by incubating the cells with EZ-Link Sulfo-NHS-SS-Biotin. The cells are lysed with a mild detergent and the labelled proteins are pulled down with NeutrAvidin Agarose. The bound proteins represent the cell membrane or cell surface proteins.

The cell or plasma membrane is the outermost layer of a cell and consist of a phospholipid bilayer. The membrane proteins are anchored in membranes and each protein has specific function and that often required them to extend beyond the membrane surface (Shen *et al*., 1997).

Some membrane proteins (peripheral proteins) are on the outside of the membrane. These proteins are usually involved in cell-to-cell communications (Mader 2011, Shukla *et al.*, 2012). Other membrane proteins are positioned on the inside of the cell membrane and can be involved intracellular signaling processes or anchoring cytoskeletal proteins to the cell membrane. Some membrane proteins (integral proteins) are located within the phospholipid bilayer (Mader 2011). Membrane proteins serve a variety of key biological functions, such as cell-to-cell communication, receptor-mediated signal transduction, and selective transport (Shukla *et al.*, 2012). During disease development plasma membrane proteins can

undergo posttranslational modifications (PTMs) which is crucial for pathological development of the diseased cell or tissue (Shukla *et al.*, 2012). The expression levels of membrane proteins can also be affected to facilitate the new pathophysiology of the diseased tissue. Since these proteins are exposed on the surface of the cell, they are ideal biomarkers for the diagnosis of the disease. Membrane proteins with functional domains exposed to the extracellular environment such as integral membrane proteins has been exploited in prognostic and predictive diagnosis of patients (Kohnke *et al.*, 2009). These proteins can potentially be detected on the surface of circulating tumour cells or be shed from the cells ending up in the serum and can thus facilitate the development of non-invasive diagnostic methods. These proteins can also serve as targets for the delivery of therapeutic agents.

To confirm the extraction of cell membrane proteins Western blotting (as described in section 2.3.1.6) was performed using the pan-cadherin antibody. Pan-cadherin is a transmembrane glycoprotein, belonging the cadherin family of proteins. The cadherin proteins are calcium-dependent and acts in cell-cell adhesion. Pan-cadherin was chosen as cell membrane reference protein as it is expressed as an integral protein in many cell types (Patel *et al*., 2003, Wheelock *et al.*, 2003, Paredes *et al.*, 2004, Hulpiau *et al.*, 2009, Paredes *et al.*, 2012).

The Pierce® Cell Surface Protein Isolation Kit separate the total cellular protein lysate into a membrane and cytosolic fraction. Twenty microgram (20 μg) of the cell surface and cytosolic proteins were resolved by 1D SDS-PAGE and stained

using Coomassie Brilliant Blue (Figure 3.2). The Coomassie Brilliant Blue stained gel shows equal protein loadings for the membrane and cytosolic fractions for both MCF-7 and MCF-12A cell lines. The Western blot (Figure 3.3) using the pan-cadherin antibody shows the detection of a 132 kDa protein in the membrane fractions of MCF-7 and MCF-12A cells (Figure 3.3; lanes A and C, respectively). This protein was not detected in the cytosolic fractions of both MCF-7 and MCF-12A cells (Figure 3.3; lanes B and D, respectively). The molecular weight of the protein detected in the membrane fractions of MCF-7 and MCF-12A cells correspond to the expected molecular weight of pan-cadherin. This suggested that the cytosolic fraction was free of cell membrane proteins and that the extraction of cell membrane proteins was successful.

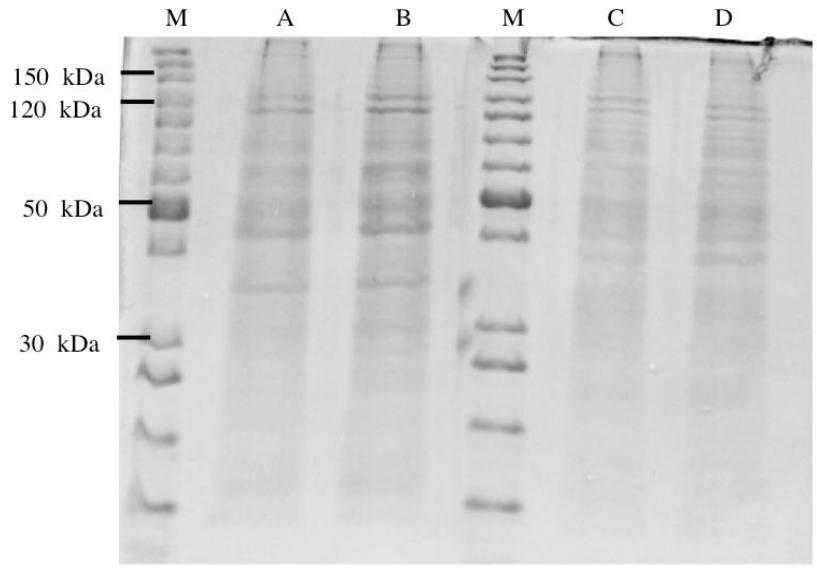


Figure 3. 2: 1D SDS-PAGE of membrane and cytosolic proteins isolated from MCF-7 and MCF-12 cells. Lane M represents the protein molecular weight marker, lane A shows the membrane protein fraction of MCF-7 cells, lane B shows the cytosolic protein fraction of MCF-7 cells. Lane C shows the membrane protein fraction of MCF-12A cells, lane D shows the cytosolic protein fraction of MCF-12A cells.

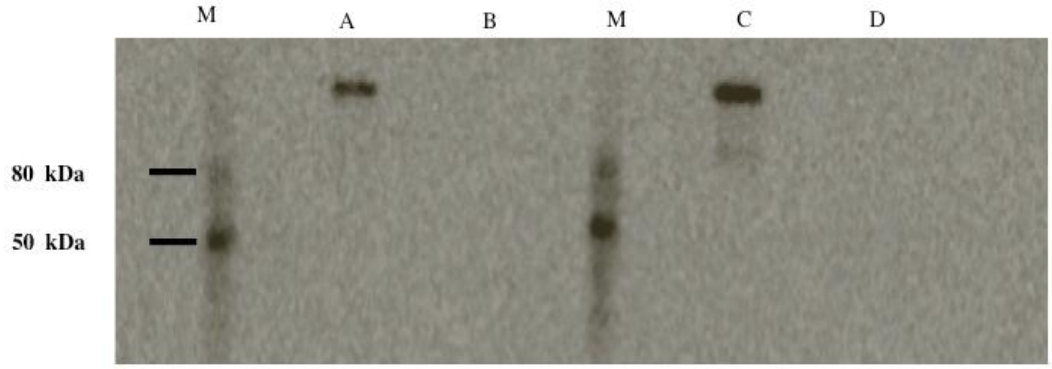


Figure 3. 3: Western blot analysis of membrane and cytosolic fractions of MCF-7 and MCF-12A cells using the Pan-cadherin antibody. Lane M represents the protein molecular weight marker, lane A shows the membrane protein fraction of MCF-7 cells, lane B shows the cytosolic protein fraction of MCF-7 cells. Lane C shows the membrane protein fraction of MCF-12A cells, lane D shows the cytosolic protein fraction of MCF-12A cells.

3.3 Characterization of membrane protein samples by LC-MS/MS

An equal amount (i.e. 20 μg of each) of the membrane protein fraction isolated from MCF-7 (labelled with $^{3}C_{6}$ L-Lysine-2HCl) and MCF-12A (labelled with L-Arginine-HCl) cells were mixed in a 1:1 ratio. The protein mixtures were analysed on a 12 % SDS-PAGE gel (Figure 3.4). The lane representing the mixed protein sample was divided into 10 sections and excised as shown in Figure 3.4; lane B. The gel slices were sent to the Central Proteomics Facility at the Sir William Dunn Pathology School (University of Oxford, UK) for LC-MS/MS analysis.

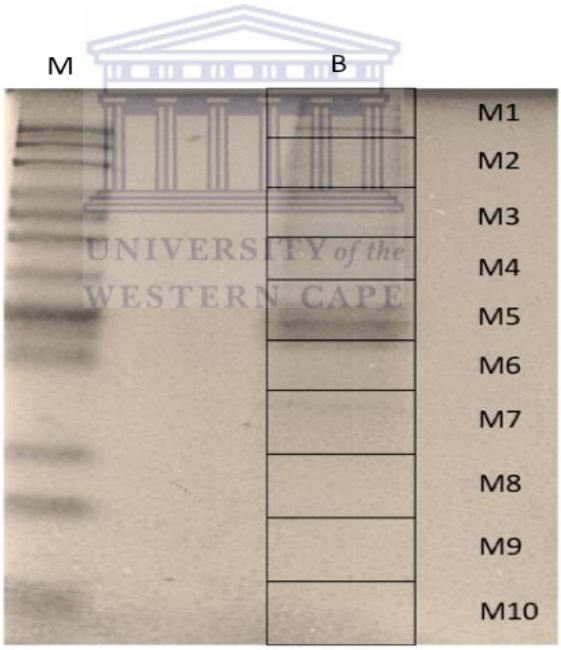


Figure 3. 4: SDS-PAGE analysis of mixed membrane protein fractions isolated from MCF-7 and MCF-12A cells. MCF-7 and MCF-12A cells were differentially labelled in SILAC medium, the membrane proteins of the cells were isolated, mixed in a 1:1 ratio and analysed by SDS-PAGE. Lane M represents the protein marker and lane B shows a mixture of the labelled of MCF-7 and MCF-12A proteins. M1 to M10 represent the individual sections of gel that was excides and sent for LC MS/MS analysis.

3.3.1 Identification and quantitation of proteins in the membrane fraction

A total of 2831 proteins were identified by the LC-MS/MS from the 10 gel sections. This analysis was done at the Central Proteomics Facility at the Sir William Dunn Pathology School (University of Oxford, UK) using MaxQuant Software. MaxQuant analysis is a quantitative proteomics software package designed for analyzing large mass-spectrometric data sets and provides a large amount of information on the protein sample, but of importance in this study is the information on the identity of the proteins in the sample, the cellular location of each protein and the relative expression levels of each protein in the two samples (MCF-7 and MCF-12A). The proteins were categorized based on their locations within the cells as shown in Figure 3.5. The locations of 13 % of the proteins were unknown (Unclassified Proteins), while 46 % (or 1289 proteins) were classified as Cell surface or Extracellular or Membrane (CEM) Proteins and 41 % (or 1170 proteins) were Cytoplasmic Proteins. Of the 46 % CEM proteins, 10 % (or 271 proteins) are exclusively located on the cell surface, extracellular or membrane and 36 % (or 1018 proteins) can also be located in the cytoplasm. Although the Western blot in Figure 3.3 showed that the cytosolic fraction did not contain membrane proteins, this result suggests that the membrane fractions contained Cytoplasmic proteins as well. Based on the sequence analysis of several eukaryotic genomes it is estimated that 30 % of the total protein content of a cell are membrane proteins (Qoronfleh *et al.*, 2003). Since 46 % of the proteins in membrane fractions were CEM proteins, it can be concluded that the attempt to isolate the membrane proteins from MCF-7 and MCF-12A cells resulted in a membrane fraction that was significantly enriched for membrane proteins, but still

contained some cytosolic proteins. It can also be concluded that no membrane proteins were lost in the separation of the membrane and cytosolic proteins since the marker for membrane proteins (pan-cadherin) indicated that no membrane proteins were present in the cytosolic fraction.

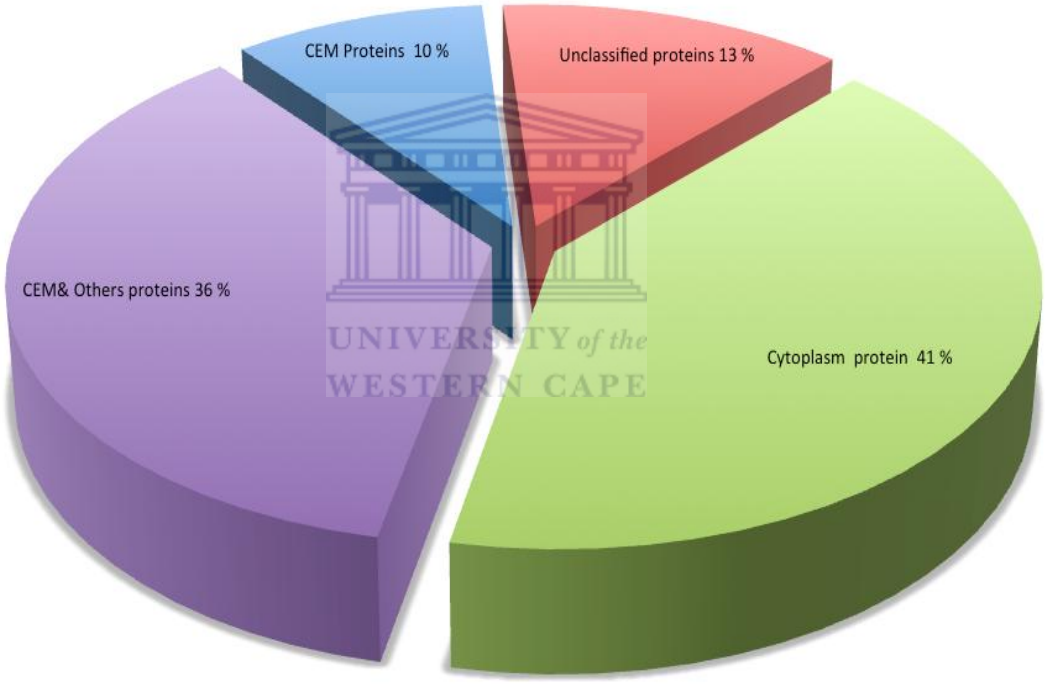


Figure 3. 5: Intracellular locations of proteins identified by LC-MS/MS. The proteins from 10 fractions were analysed and classified according to their location in the cells as Cell surface or Extracellular or Membrane (CEM) and Cytoplasmic Proteins. Unclassified proteins are proteins with unknown cellular localization.

The focus of this study was to identify membrane proteins that are differentially expressed between MCF-7 and MCF-12A cells. It is for this reason that proteins that are exclusively expressed in the cytoplasm were excluded from further study. Figure 3.6 is a Venn diagram showing the distribution of the CEM proteins into the three categories (Cell surface, Extracellular and Membrane). The majority (212) of these proteins are exclusively classified as cell membrane proteins and 16 proteins are Extracellular proteins. The other proteins are classified into more than one of these grouping.

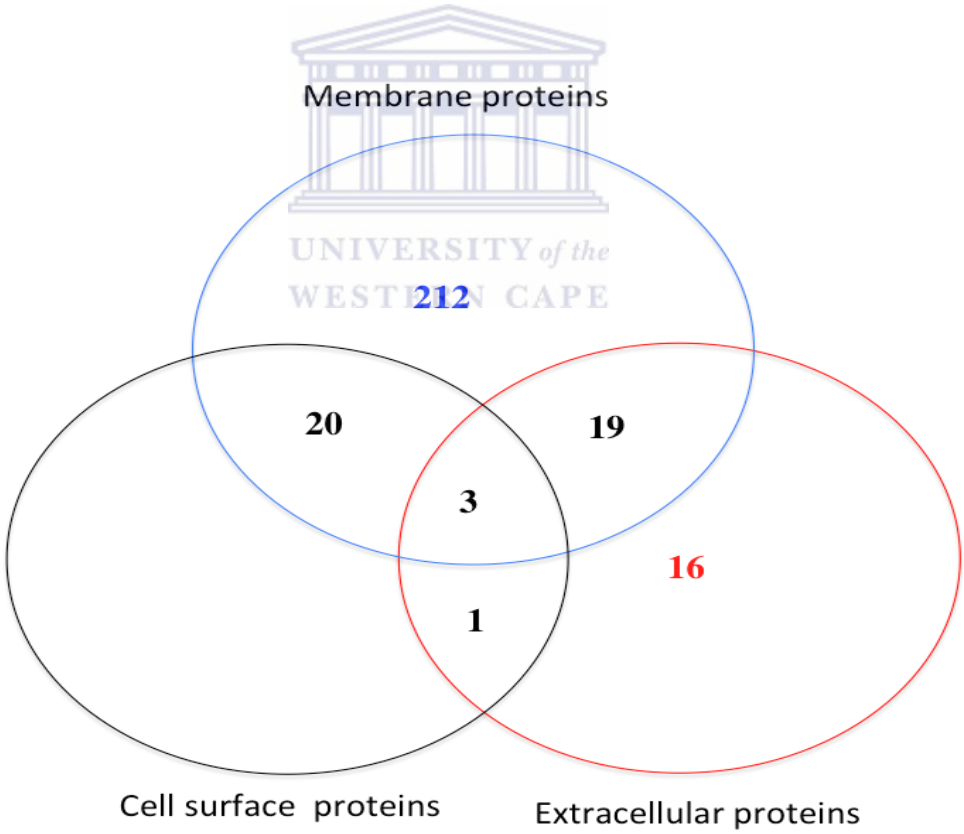


Figure 3. 6: Venn diagram showing the distribution of the 271 CEM proteins.

In order to identify CEM proteins that are highly overexpressed in MCF-7 cells as compared to MCF-12A cells, a further analysis of the LC-MS/MS data was performed. This was done using two parameters (the ratio of "Heavy/Light" counts and the "Score") generated by MaxQuant analysis.

The ratio of "Heavy/Light" counts gives an indication of the comparative expression levels of the proteins between the two samples (MCF-7 and MCF-12A). A high "Heavy/Light" ratio, suggest that the expression of the "Heavy" labelled proteins (isolated from MCF-7 cells) was significantly higher compared to the "Light" labelled proteins (isolated from MCF-12A cells). If the ratio of "Heavy/Light" count is equal to 1, it means that the relative expression levels of the "Heavy" labelled proteins (isolated from MCF-7 cells) were equal to the "Light" labelled proteins (isolated from MCF-12A cells) and that the expression of the these proteins are not different between the two cell lines. For the purpose of this study, if the "Heavy/Light" ratio was between 1 and 2, the protein expression levels were considered to be equivalent in the two cell lines. If the "Heavy/Light" ratios were lower than 1, then the expression of the protein was considered to be lower in MCF-7 cells as compared to MCF-12A cells. If the "Heavy/Light" ratios were 2 or higher, the protein expression levels were considered to be higher in MCF-7 cells as compared to MCF-12A cells. Of the 2831 proteins identified in this study, 19 % had a "Heavy/Light" ratio of 2 or higher. This 19 % represents the proteins that are overexpressed in MCF-7 cells as compared to MCF-12A cells. Thirty-five percentile (35 %) had a "Heavy/Light" ratio below 1 and 16 % had "Heavy/Light" ratio below between 1 and 2. This

implies that expression levels 35 % of the proteins were down regulated and 16 % of the proteins were unchanged in MCF-7 cells relative to MCF-7 cells. Although this study identified 2831 proteins from the cells, LC-MS/MS analysis produced "Heavy/Light" counts for only 1969 (or 70 %) of the proteins. This implies that 30 % of the proteins were not labelled with the SILAC media.

The second parameter that was considered was the "Score", which is a confidence indicator for the accurate identification of a protein by MS. The higher the "Score", the higher the probability that the identification was accurate and vice versa. The "Score" also gives an indication of the relative abundance of the protein in the sample. Proteins that are more abundant in the sample will have a higher "Score". It was essential to consider the relative abundance of the potential biomarker proteins, since low abundant proteins will not be suitable biomarkers proteins.

To identify potential biomarker proteins that are over expressed in MCF-7 cells, the 271 CEM proteins that are exclusively located on the cell surface were ranked based on "Heavy/Light" ratios. Table 3.1 is showing a list of proteins with the highest (top 10) "Heavy/Light" ratio. Isoform 2 of GDNF family receptor alpha-1 (GFRA1) had the highest "Heavy/Light" ratio of 131.06, while Ectonucleotide pyrophosphatase/ phosphodiesterase family member 1 (ENPP1) had the lowest "Heavy/Light" ratio of 2.4. The abundance score for GFRA1 was 456.42, however other proteins e.g. Ephrin type-B receptor 4 (EPHB4) had higher

(1079.9) abundance scores, but lower "Heavy/Light" ratios as compared to GFRA1.

Table 3. 1: Top 10 genes with the highest ratio of "Heavy/Light" counts.

NO	Accession	Description	Score (abundance)	Heavy/ Light
1	P56159-2	Isoform 2 of GDNF family receptor alpha-1 GN=GFRA1 (GDNF).	465.2	131.066
2	P54760	Ephrin type-B receptor 4 GN=EPHB4.	1079.99	6.338
3	Q8N2K0	Monoacylglycerol lipase ABHD12 GN=ABHD12.	117.38	5.541
4	F5GYD3	Neuronal cell adhesion molecule GN=NRCAM (KIAA0343; MGC138845; MGC138846).	1027.16	5.399
5	H7C5L1	Prostaglandin E synthase 2 (Fragment) GN=PTGES2.	575.18	3.877
6	F8VR84	UPF0160 protein MYG1, mitochondrial GN=C12orf10.	44.92	3.859
7	Q9UMX5	Neudesin GN=NENF, CIR2; NEUDESIN; SCIRP10; SPUF.	253.8	3.738
8	Q9Y6I9	Testis-expressed sequence 264 protein GN=TEX264 (FLJ13935, ZSIG11).	367.93	3.551
9	Q00765	Receptor expression-enhancing protein 5 GN=REEP5 (DP1, TB2, D5S346, C5orf18).	384.27	2.977
10	P22413	Ectonucleotide pyrophosphatase/phosphodiesterase family member 1 GN=ENPP1 (M6S1; NPP1; NPPS; PC-1; PCA1; PDNP1).	265.17	2.402

3.4 Biomarker validation

3.4.1 In silico analysis of the expression levels of potential biomarkers

The candidate biomarkers obtained using proteomics analysis were subjected to further study using bioinformatics tools, in particular GeneHub-GEPIS, in an effort to investigate the potential of these CEM proteins as biomarkers for breast cancer. GenHub-GEPIS is an online bioinformatics tool that can be used to study gene expression patterns in healthy and cancer tissues based on human and mouse expression sequence tag (EST) abundance (Zhang *et al.*, 2007). GeneHub-GEPIS calculates (represented as digital expression unit or DEU) the normalized gene expression levels across a large panel of tumour and healthy tissues. This tool defines the genomic structures based on mRNA transcript sequence and protein-coding genes using transcripts from several reliable sources, such as sequences from RefSeq, Ensemble genes and proteome (http://research-public.gene.com/Research/genentech/genehub-gepis/index.html).

The expression patterns of all 10 potential biomarkers (listed in Table 3.1) were investigated in 40 different healthy (normal) and tumour tissues using GeneHub-GEPIS analysis. Although all these biomarkers were up-regulated in breast cancer tissue, 6 of these potential biomarkers (*ABHD12*, *NRCAM*, *MYG1*, *TEX264*, *REEP5* and *ENPP1*) also demonstrated expression in the normal breast tissue (appendix Figure 3). Consequently, these genes were not studied further. Figure 3.7 show the expression patterns of the genes encoding the 4 potential biomarkers (*GFRA1*, *NENF*, *EPHB4* and *PTGES2*) and genes encoding 3 known biomarkers (*HER-2/ERBB2, FOXA1* and *CK19*). Figure 3.7 show that all these genes are over

expressed in breast cancer tissue as compared to normal breast tissue. *GFRA1* (Figure 3.7. A) is highly expressed (DEU = 102.57) in breast cancer tissue. Although, this gene is also up-regulated in 4 other cancers (bone marrow, brain, kidney and testis), the expression levels in breast cancer is significantly higher in beast cancer compared to the other cancers. *GFRA1* is up-regulated 6 normal tissues (brain, eye, muscle, nervous system, skin and spleen), but the expression level in breast cancer is still significantly higher in beast cancer. This expression pattern of *GFRA1* compares favourably to that of the known biomarkers. Although *HER-2/ERBB2* is highly expressed (DEU = 315.03) in breast cancer tissue, it is also expressed in 14 normal tissues and over expressed in 21 other cancer tissues. *FOXA1* is expressed in only 3 normal tissues, but is also over expressed in 9 other cancer tissues. However, the expression level for *FOXA1* is higher in prostate cancer and some normal tissues (e.g. peripheral nervous system) as compared to breast cancer tissue. *CK19* shows expression in normal beast tissue as well. In addition it is also expressed in 14 other normal tissues and 21 cancer tissues.

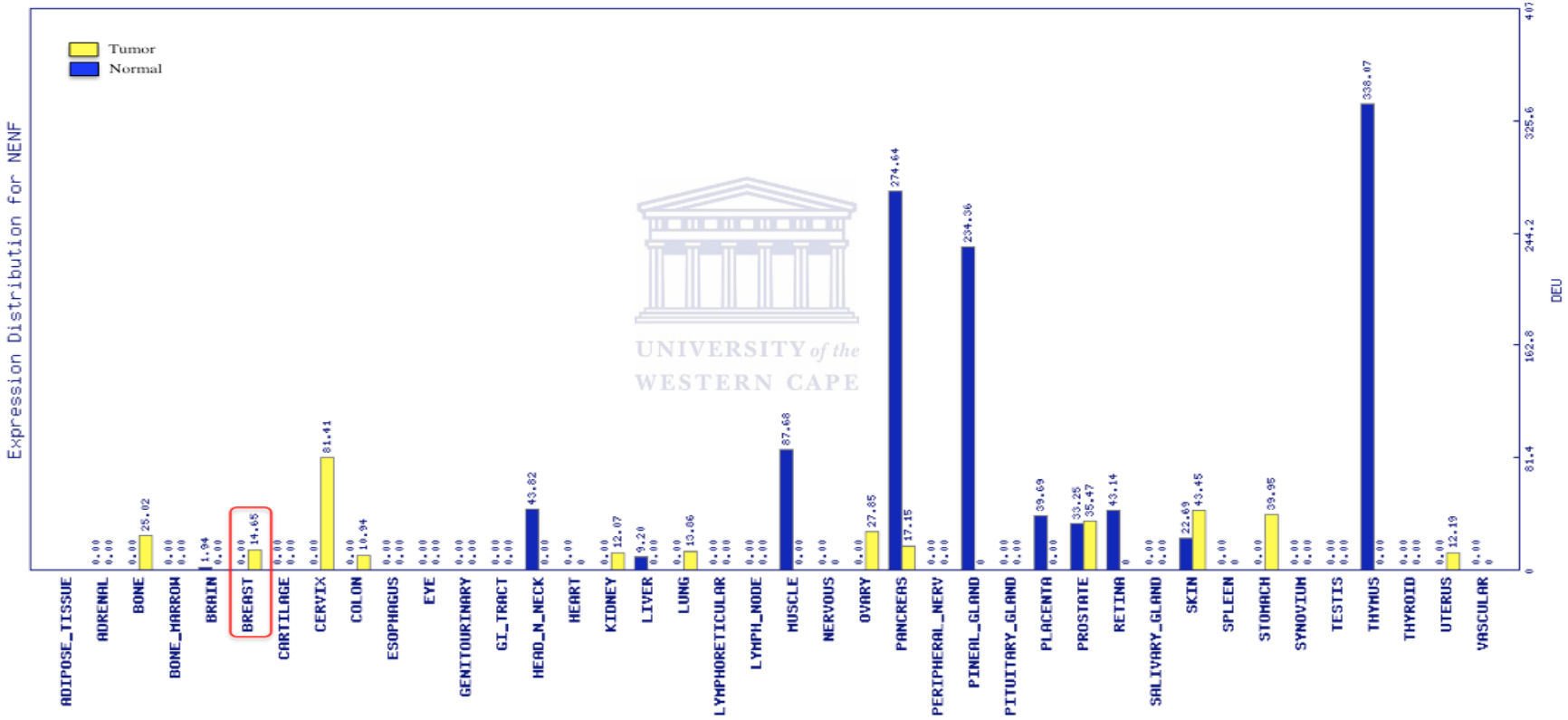
The DEU values for *NENF*, *EPHB4* and *PTGES2* were 14.65, 40.96 and 51.28, respectively. Although *NENF*, *EPHB4* and *PTGES2* was over expressed in breast cancer tissue as compared to normal breast tissue, these genes were also up-regulated in other cancerous and normal tissues. *NENF* was up-regulated in 11 normal tissues and 10 other cancer tissues. *EPHB4* was up-regulated in 12 normal tissues and 20 other cancer tissues. *PTGES2* was up-regulated in 20 normal tissues and 22 other cancer tissues.

Based on GeneHub-GEPIS analysis, *GFRA1* would be a more selective biomarker for breast cancer compared to *NENF*, *EPHB4* and *PTGES2*. This data also suggest that *GFRA1* would be a more promising biomarker than the known biomarker genes *HER-2/ERBB2, FOXA1* and *CK19*. *BRCA2* is a very prominent breast cancer biomarker; unfortunately, no expression analysis data was available for this gene and therefore the comparison to *GFRA1* could not be done.

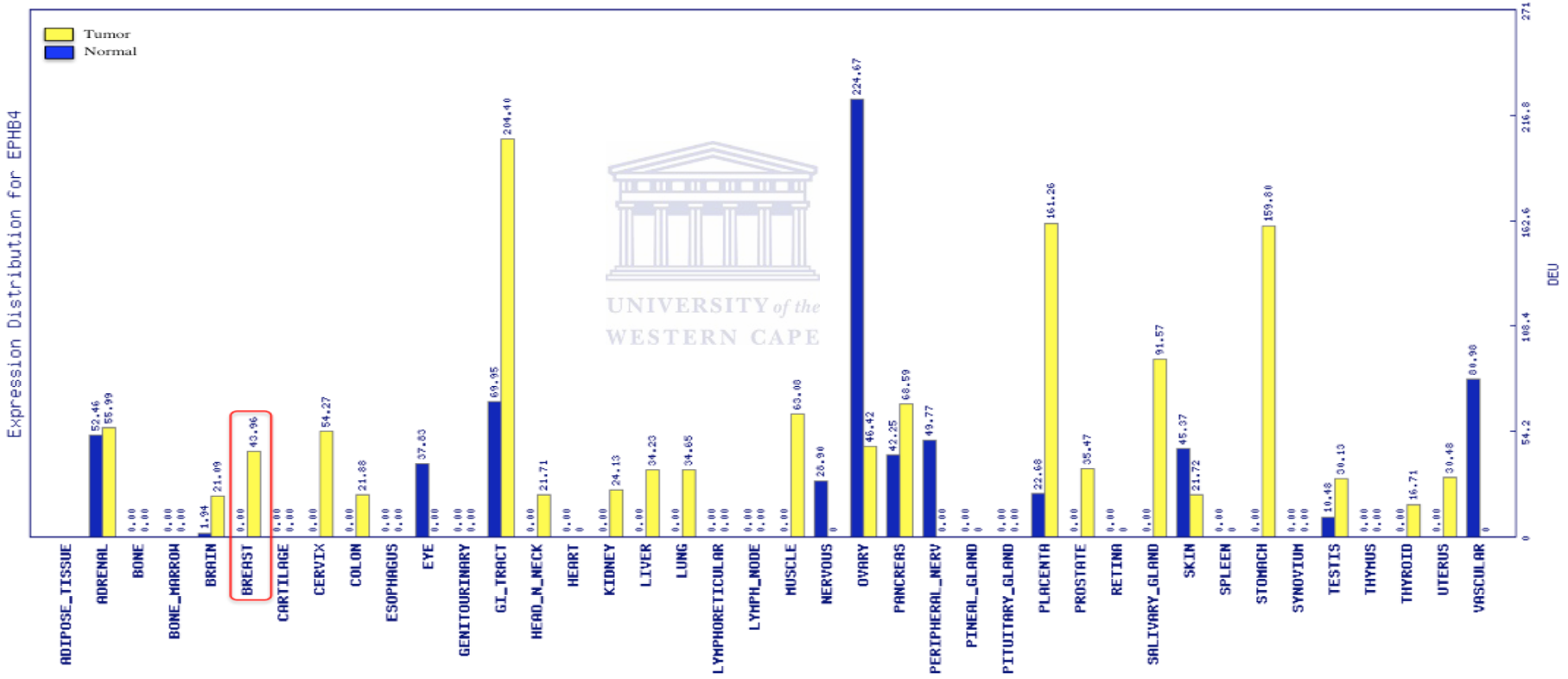
A

Expression Distribution for GFRA1
Tumor
Normal
ADIPOSE_TISSUE
ADRENAL
BONE
BONE_MARROW
BRAIN
BREAST
CARTILAGE
CERVIX
COLON
ESOPHAGUS
EYE
GENITOURINARY
GI_TRACT
HEAD_N_NECK
HEART
KIDNEY
LIVER
LUNG
LYMPHORETICULAR
LYMPH_NODE
MUSCLE
NERVOUS
OVARY
PANCREAS
PERIPHERAL_NERV
PINEAL_GLAND
PITUITARY_GLAND
PLACENTA
PROSTATE
RETINA
SALIVARY_GLAND
SKIN
SPLEEN
STOMACH
SYNOVIUM
TESTIS
THYMUS
THYROID
UTERUS
VASCULAR
0.00
1.94
10.55
102.57
12.07
35.97
37.63
29.23
28.90
22.69
61.21
45.20
0
24.8
49.6
74.4
99.2
124
DEU
UNIVERSITY of the
WESTERN CAPE

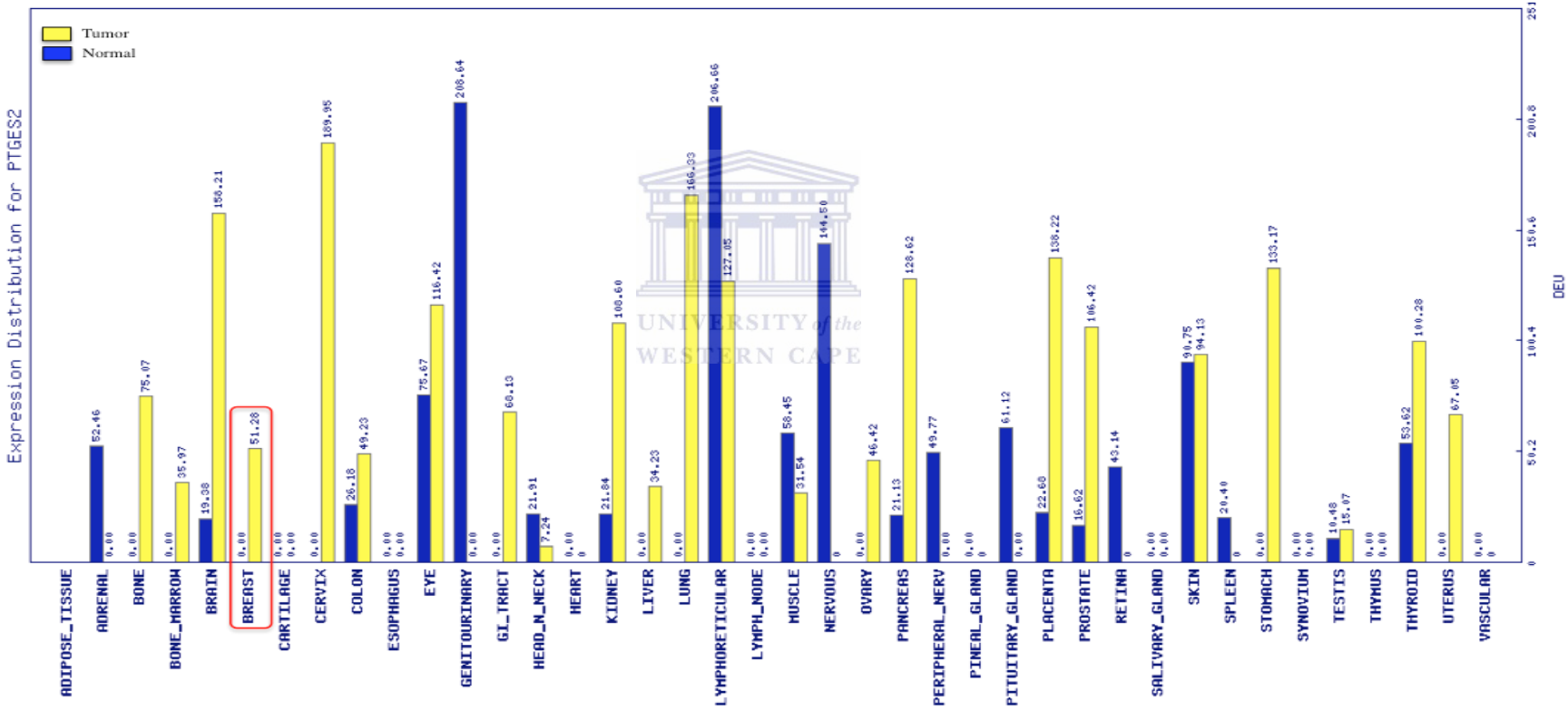
B


Expression Distribution for NENF
Tumor
Normal
DEU
0
81.4
162.8
244.2
325.6
407
ADIPOSE_TISSUE 0.00 0.00
ADRENAL 0.00 0.00
BONE 0.00 25.02
BONE_MARROW 0.00 0.00
BRAIN 1.94 0.00
BREAST 0.00 14.65
CARTILAGE 0.00 0.00
CERVIX 0.00 81.41
COLON 0.00 10.94
ESOPHAGUS 0.00 0.00
EYE 0.00 0.00
GENITOURINARY 0.00 0.00
GI_TRACT 0.00 0.00
HEAD_N_NECK 0.00 43.82
HEART 0.00 0
KIDNEY 0.00 12.07
LIVER 9.20 0.00
LUNG 0.00 13.66
LYMPHORETICULAR 0.00 0.00
LYMPH_NODE 0.00 0.00
MUSCLE 0.00 87.68
NERVOUS 0.00 0
OVARY 0.00 27.85
PANCREAS 274.64 17.15
PERIPHERAL_NERV 0.00 0.00
PINEAL_GLAND 234.36 0
PITUITARY_GLAND 0.00 0.00
PLACENTA 0.00 39.69
PROSTATE 33.25 35.47
RETINA 43.14 0
SALIVARY_GLAND 0.00 0.00
SKIN 22.69 43.45
SPLEEN 0.00 0
STOMACH 0.00 39.95
SYNOVIUM 0.00 0.00
TESTIS 0.00 0.00
THYMUS 338.07 0.00
THYROID 0.00 0.00
UTERUS 0.00 12.19
VASCULAR 0.00 0
UNIVERSITY of the WESTERN CAPE

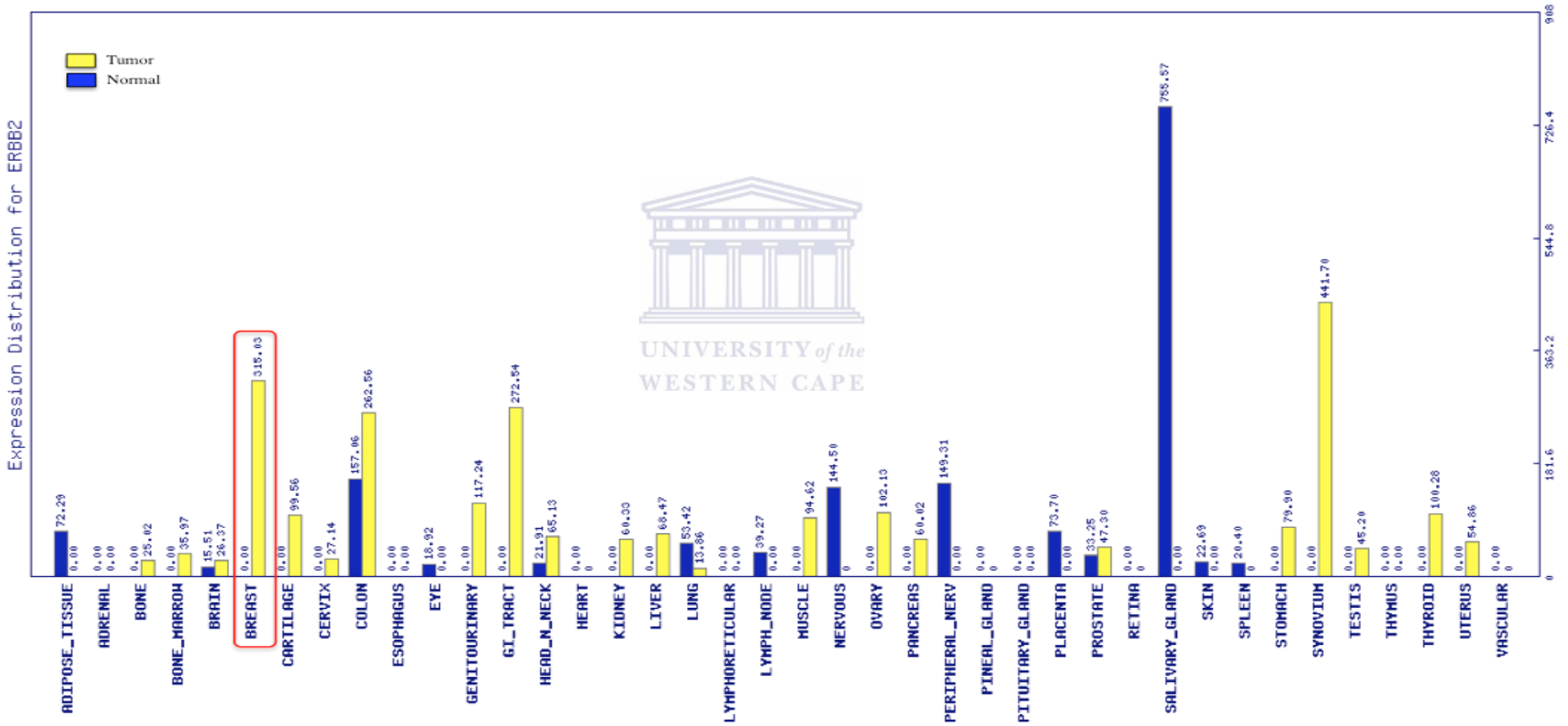
C



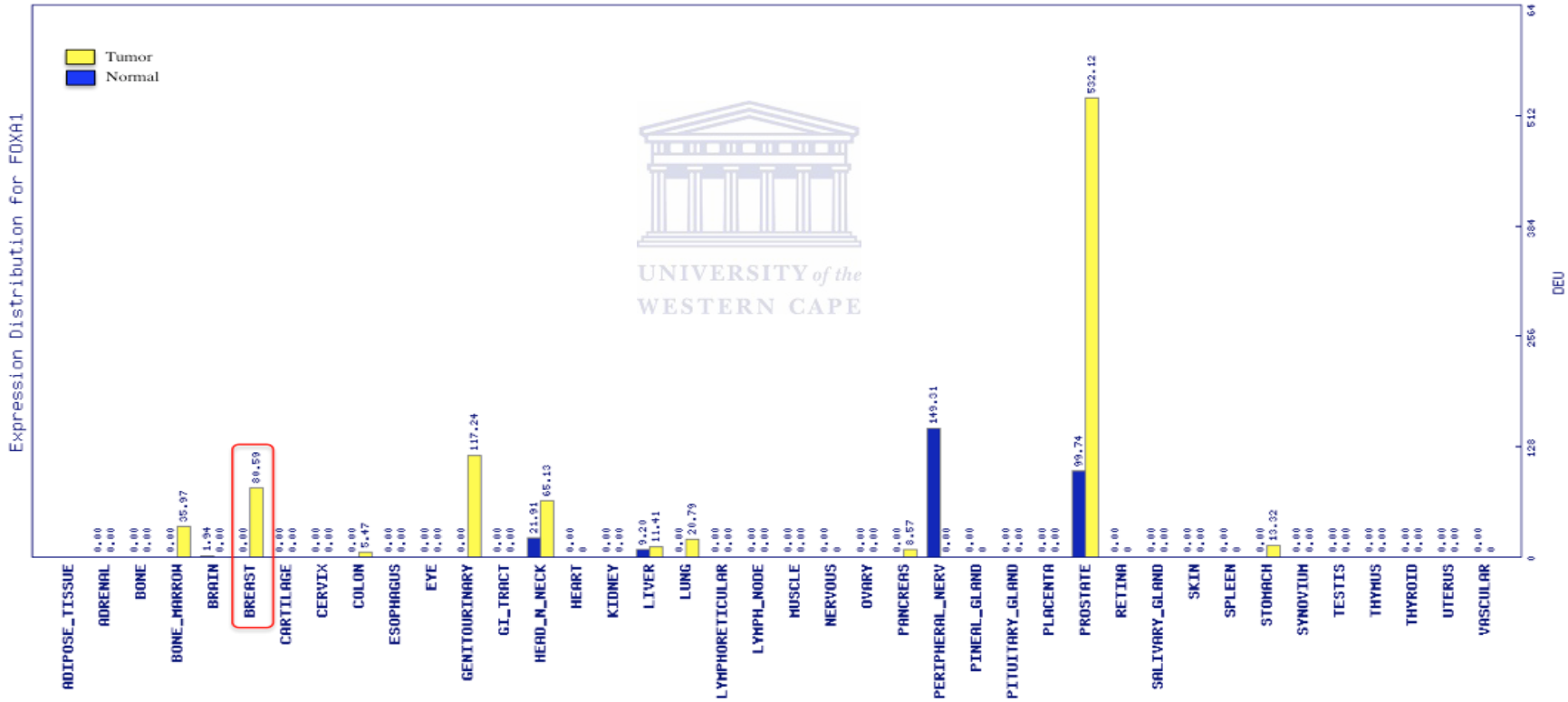
D


Expression Distribution for PTGES2
Tumor
Normal
DEU
0
50.2
100.4
150.6
200.8
251
ADIPOSE_TISSUE
ADRENAL 0.00 52.46
BONE 0.00 75.07
BONE_MARROW 0.00 35.97
BRAIN 19.38 156.21
BREAST 0.00 51.28
CARTILAGE 0.00 0.00
CERVIX 0.00 189.95
COLON 26.18 49.23
ESOPHAGUS 0.00 0.00
EYE 75.67 116.42
GENITOURINARY 0.00 208.64
GI_TRACT 0.00 60.13
HEAD_N_NECK 21.91 7.24
HEART 0.00 0
KIDNEY 21.84 100.60
LIVER 0.00 34.23
LUNG 0.00 160.33
LYMPHORETICULAR 206.66 127.05
LYMPH_NODE 0.00 0.00
MUSCLE 58.45 31.54
NERVOUS 164.50 0
OVARY 0.00 46.42
PANCREAS 21.13 128.62
PERIPHERAL_NERV 0.00 49.77
PINEAL_GLAND 0.00 0
PITUITARY_GLAND 61.12 0.00
PLACENTA 22.60 138.22
PROSTATE 16.62 106.42
RETINA 43.14 0
SALIVARY_GLAND 0.00 0.00
SKIN 90.75 94.13
SPLEEN 20.40 0
STOMACH 0.00 133.17
SYNOVIUM 0.00 0.00
TESTIS 10.48 15.07
THYMUS 0.00 0.00
THYROID 53.62 100.28
UTERUS 0.00 67.05
VASCULAR 0.00 0
UNIVERSITY of the WESTERN CAPE

E


Expression Distribution for ERBB2
Tumor
Normal
ADIPOSE_TISSUE 72.29 0.00
ADRENAL 0.00 0.00
BONE 0.00 25.02
BONE_MARROW 0.00 35.97
BRAIN 15.51 26.37
BREAST 0.00 315.03
CARTILAGE 0.00 99.56
CERVIX 0.00 27.14
COLON 157.06 262.56
ESOPHAGUS 0.00 0.00
EYE 18.92 0.00
GENITOURINARY 0.00 117.24
GI_TRACT 0.00 272.54
HEAD_N_NECK 21.91 65.13
HEART 0.00 0
KIDNEY 0.00 60.33
LIVER 0.00 68.47
LUNG 53.42 13.86
LYMPHORETICULAR 0.00 0.00
LYMPH_NODE 39.27 0.00
MUSCLE 0.00 94.62
NERVOUS 144.50 0
OVARY 0.00 102.13
PANCREAS 0.00 60.02
PERIPHERAL_NERV 149.31 0.00
PINEAL_GLAND 0.00 0
PITUITARY_GLAND 0.00 0.00
PLACENTA 73.70 0.00
PROSTATE 33.25 47.30
RETINA 0.00 0
SALIVARY_GLAND 755.57 0.00
SKIN 22.69 0.00
SPLEEN 20.40 0
STOMACH 0.00 79.90
SYNOVIUM 0.00 441.70
TESTIS 0.00 45.20
THYMUS 0.00 0.00
THYROID 0.00 100.28
UTERUS 0.00 54.86
VASCULAR 0.00 0
0
181.6
363.2
544.8
726.4
908
UNIVERSITY of the WESTERN CAPE

F


Expression Distribution for FOXA1
Tumor
Normal
ADIPOSE_TISSUE
ADRENAL
BONE
BONE_MARROW
BRAIN
BREAST
CARTILAGE
CERVIX
COLON
ESOPHAGUS
EYE
GENITOURINARY
GI_TRACT
HEAD_N_NECK
HEART
KIDNEY
LIVER
LUNG
LYMPHORETICULAR
LYMPH_NODE
MUSCLE
NERVOUS
OVARY
PANCREAS
PERIPHERAL_NERV
PINEAL_GLAND
PITUITARY_GLAND
PLACENTA
PROSTATE
RETINA
SALIVARY_GLAND
SKIN
SPLEEN
STOMACH
SYNOVIUM
TESTIS
THYMUS
THYROID
UTERUS
VASCULAR
DEU
0
128
256
384
512
64
UNIVERSITY of the
WESTERN CAPE

G

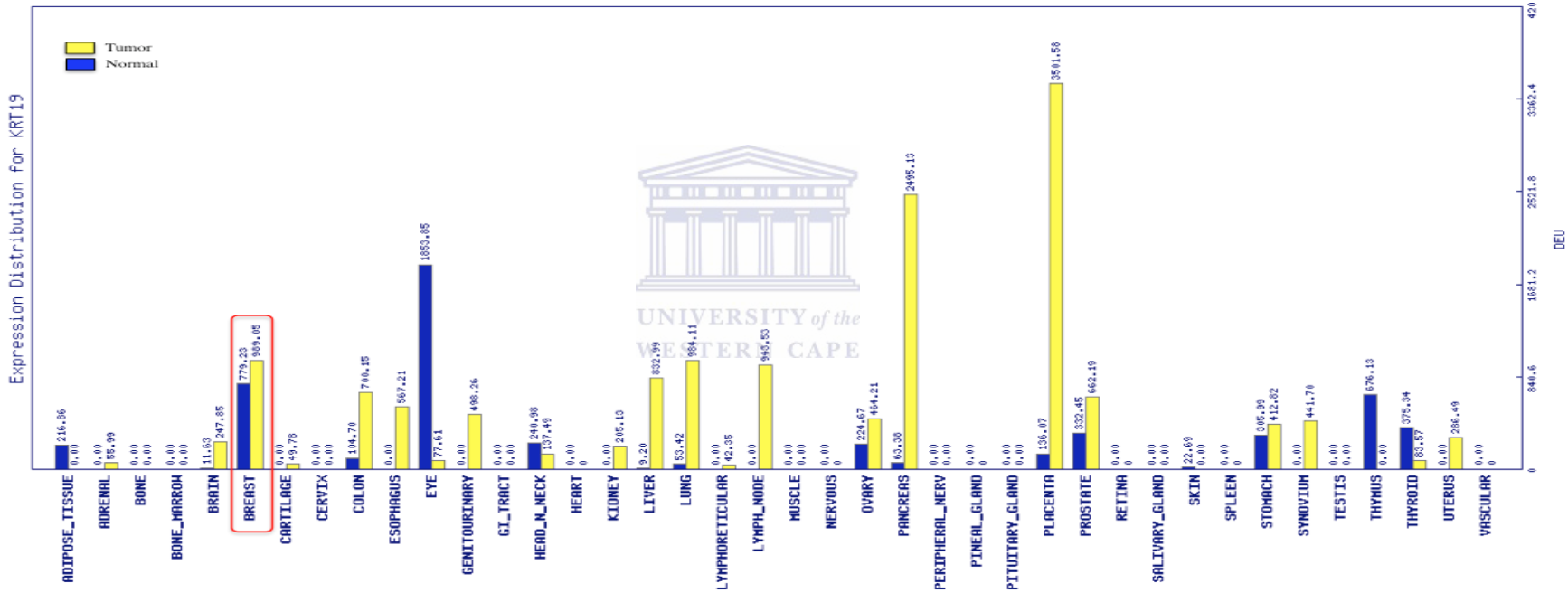


Figure 3. 7: GeneHUB-GEPIS charts showing the expression profile of 4 potential biomarker genes and 4 control biomarkers in 40 different tissues. A, B, C and D shows the expression patterns of *GFRA1*, *NENF*, *EPHB4* and *PTGES2*, respectively. E, F and G charts showing the expression of 3 control biomarkers (*HER-2*/*ERBB2, FOXA* and *CK19)*. The blue bar displays the expression in normal tissue and the yellow bar shows the expression in tumour tissues. The expression level is represented as a digital expression unit (DEU) values. The expression levels in normal and cancerous breast tissue are indicates in a red box.

3.4.2 Expression analysis of potential biomarkers using qRT-PCR

The expression levels of 4 potential biomarker genes (*GFRA1, NENF*, *EPHB4* and *PTGES2*) were investigated in MCF-7 and MCF-12A cells using qRT-PCR. The expression levels of five positive control genes *GAPDH, BRCA2, HER-2, CK19* and *FOXA1* were also evaluated. *GAPDH* is a commonly used reference gene in qRT-PCR analysis (Jain *et al.*, 2006), while *BRCA2, HER-2, CK19* and *FOXA1* are known molecular biomarkers for breast cancer (Grunewald *et al.*, 2000, Radojicic *et al.*, 2011, Faheem *et al.*, 2012). Primer sequences for the biomarker genes (*FOXA1, CK19, HER-2, BRCA2, PTGES2, GFRA1, NENF* and *EPHB4)* were designed based on the gene sequences as shown in section 2.5, while primer sequences for the reference gene, *GAPDH*, were obtained from literature (Lecke *et al.*, 2013).

MCF-7 and MCF-12A cells were cultured as discussed in section 2.2.3 and mRNA was extracted from these cell lines and converted to cDNA as described in section 2.6.3 The RNA integrity, quality and quantity was determined using a BioAnalyzer.

qRT-PCR reactions were done using serial dilutions (with concentrations ranging from 0.02 ng to 200 ng) of the synthesized cDNA to establish a standard curve for each gene and to determine the efficiency of the PCR reactions. Figure 3.8 shows the amplification curves for *GAPDH* using serial dilutions of cDNA produced from MCF-7 and MCF-12A mRNA. The 1/10, 1/00 and 1/1000 cDNA dilutions produced superimposable amplification curves for both MCF-7 and MCF-12A

cells. This suggests that the expression levels of *GAPDH* in these two cell lines were the same. Melting curve analysis was also performed to verify the specific amplification of the target genes. Figure 3.9 shows the melting curves of PCR products produced for *GAPDH*. The presence of only one melting peak in the range of 75 °C to 90 °C indicates specific amplification. Additional peaks will indicate the amplification of non-specific PCR fragments, which could be as a result of contamination, mis-priming (referring to the annealing of primers to cDNA sequences other than the target cDNA) or primer-dimers (primers annealing to themselves). The melting curves for *FOXA1, CK19, HER-2, BRCA2, PTGES2, GFRA1, NENF* and *EPHB4* also gave a single peak (data not shown), suggesting the specific amplification of these genes.

In order to normalize the relative quantification of the biomarkers to reference gene, the sets of crossing points for the reference genes were imported into Relative Expression Software Tool (REST). The chosen crossing points were considered during the calculation process of the software, which allowed for a normalization of the target genes in relation to the reference gene, *GAPDH*. The randomization test were performed to determine whether normalization via the reference gene were useful. The results of the randomization test displayed the factor of regulation and level of significance, which indicated that the reference gene were suitable in this experiment. *GAPDH* is a suitable reference gene as it is present in all nucleated cell types and mRNA synthesis of this gene has been found to be stable in the two cell lines used in this study.

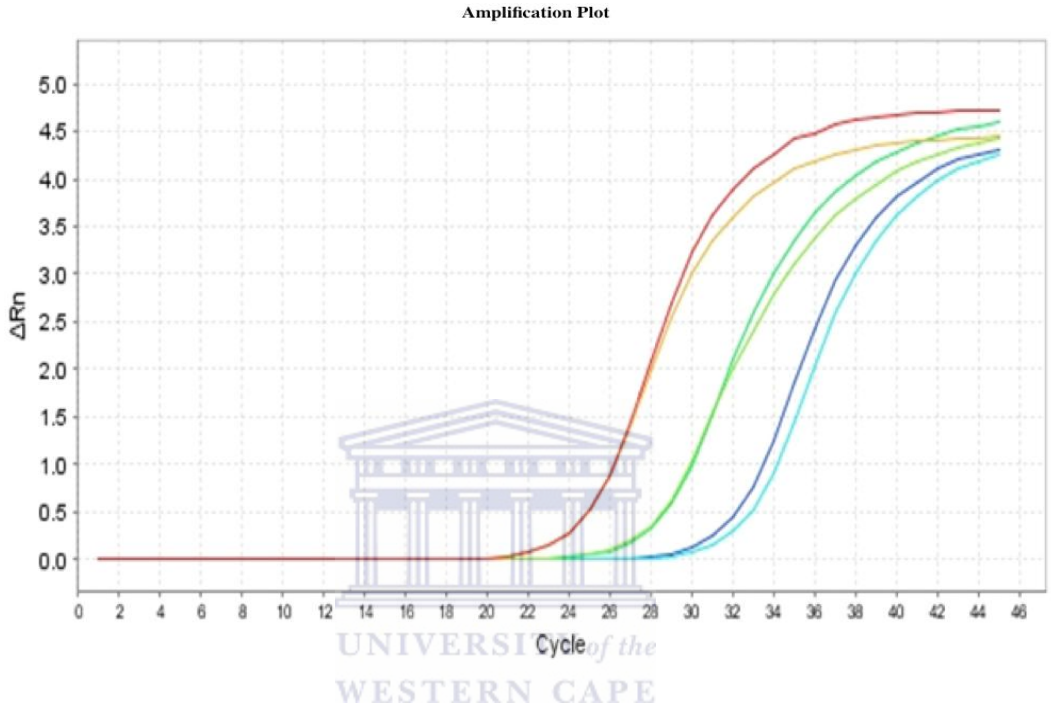


Figure 3. 8: Amplification plot for the reference gene, *GAPDH*. The red, dark green, and dark blue slopes indicate the amplification of *GAPDH* from MFC-7 cDNA at the 1:10, 1:100, and 1:1000 dilutions, respectively. The yellow, light green, and light blue slopes indicate the amplification of *GAPDH* from MFC-12A cDNA at the 1:10, 1:100, and 1:1000 dilutions, respectively.

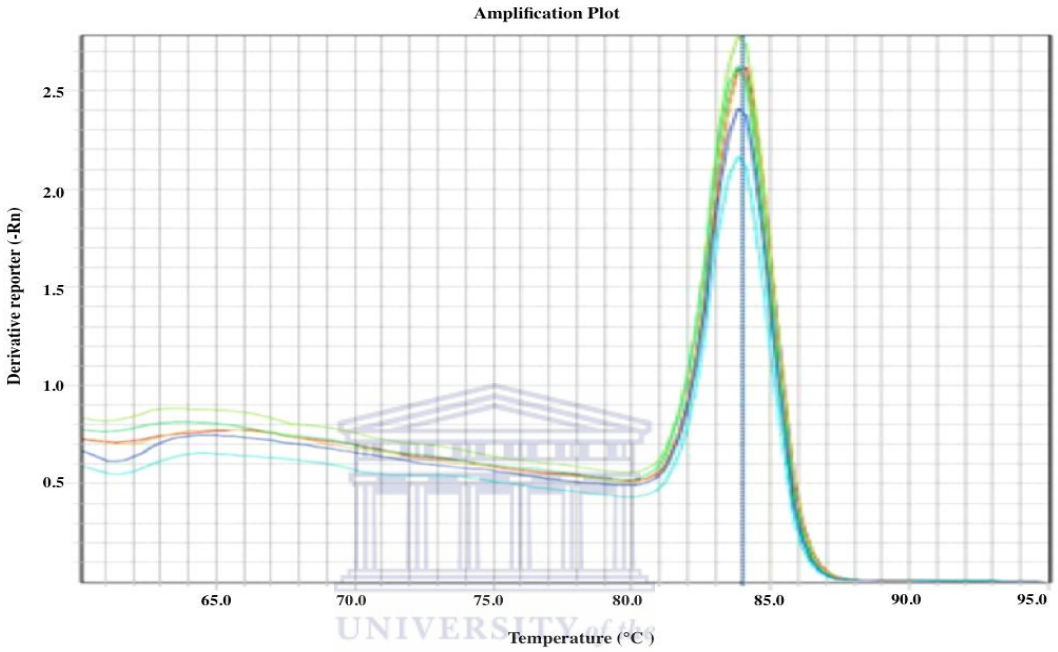


Figure 3. 9: Melting curve analysis for *GAPDH*. Melting curve for the housekeeping gene, *GAPDH*, ranging from 65 °C to 95 °C. The red, dark green, and dark blue peaks indicate the melting curves of *GAPDH* from MFC-7 cDNA at the 1:10, 1:100, and 1:1000 dilutions, respectively.. The yellow, light green, and light blue peaks indicate the melting curves of *GAPDH* from MFC-12A cDNA at the 1:10, 1:100, and 1:1000 dilutions, respectively.

The amplification efficiencies of the qRT-PCR reactions, performed with serial dilutions of the templates, are based on the slopes of the standard curves obtained from the experiment performed for each gene. Table 3.2 shows the amplification efficiencies for *GAPDH*, *FOXA1, GFRA1, NENF*, *EPHB4* and *PTGES2*. These values ranged from 91 to 98 % and are highly similar, indicating that a similar rate of amplification occurred for all these genes.

Table 3. 2: PCR amplification efficiencies (EFF) for the potential biomarker genes and the reference gene *GAPDH*.

GENE	EFF in MCF-12A	EFF in MCF-7
GAPDH	96 %	95 %
FOXA1	95 %	95 %
CK19	93 %	92 %
HER-2	91 %	94 %
BRCA2	96 %	93 %
GFRA1	94 %	98 %
EPHB4	96 %	94 %
NENF	94 %	96 %
PTGES2	98 %	97 %

qRT-PCR was performed for *FOXA1, CK19, HER-2, BRCA2, PTGES2, GFRA1, NENF* and *EPHB4* as described in section 2.6.5 on cDNA produced from mRNA isolated form MCF-7 and MCF-12A cells. Table 3.3 shows a summary of the qRT-PCR results for all 8 genes. Indicated in this table are the crossing points (C_p) for the amplifications for all the genes and the fold change in the expression levels of these genes in MCF-7 cells relative to the expression levels in MCF-12A cells. The results show that *FOXA1* is significantly up-regulated in MCF-7 cells, with a 40.77 fold increase in the expression level of this gene. The other three known biomarker genes *(CK19, HER-2* and *BRCA2*) were also up-regulated with fold increases of 26.89, 23.82 and 8.13, respectively as shown in Figure 3.10. The results are in agreement with the GeneHub-GEPIS analysis, which also showed that *FOXA1, CK19* and *HER-2* were up-regulated in breast cancer tissues (Figure 3.7 E, F and G). Amongst the 4 putative biomarker genes identified in this study, 3 genes (*GFRA1, NENF* and *EPHB4*) were up-regulated and *PTGES2* was down-regulated. The fold increase in the expression levels of *GFRA1, EPHB4* and *NENF* was 4.59, 3.91 and 1.18, respectively. *PTGES2* was down-regulated in MCF-7 cells relative to MCF-12A cells, with a fold change of -1.52 (Figure 3.10). qRT-PCR results for *GFRA1, NENF* and *EPHB4* supports the proteomics and GeneHub-GEPIS data, however, the qRT-PCR results for *PTGES2* is not in agreement with the proteomics data, since the proteomics results suggest that the expression levels of the protein encoded by this gene is significantly up-regulated in MCF-7 cells relative to MCF-12A cells. Based on the qRT-PCR results, *GFRA1, EPHB4* and *NENF* can potentially be used as biomarkers for breast cancer. However, the fold increase in the expression levels for these potential

biomarkers were significantly lower compared to the known biomarkers (*FOXA1, CK19, HER-2* and *BRCA2)*.

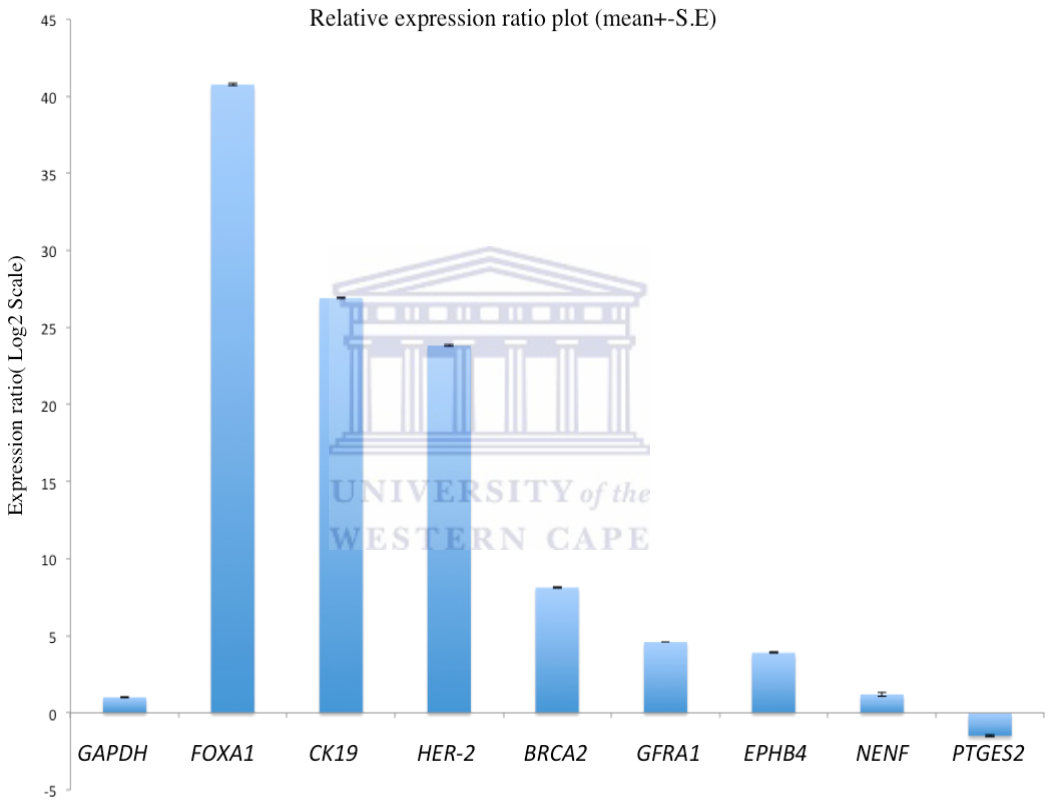


Figure 3. 10: Relative expression ration of 4 known and 4 putative biomarkers. Four biomarkers that selected from proteomic databases for cell surface protein expression in MCF-7 breast cancer line compare to MCF-12A non-cancer cell line with *GAPDH* as internal control.

Table 3. 3: Descriptive statistics of 8 cancer genes showing differential expression between MCF-7 and MCF-12A cell lines.

Gene	Sample	Mean C_T	Std. error	Fold change	p-value
FOXA1	MCF-12A	21.05	0.05	40.765	0.003
	MCF-7	15.75	0.10		
CK19	MCF-12A	24.23	0.05	26.895	0.001
	MCF-7	19.53	0.05		
HER-2	MCF-12A	21.23	0.07	23.823	0.005
	MCF-7	16.70	0.06		
BRCA2	MCF-12A	18.88	0.08	8.136	0.021
	MCF-7	15.90	0.06		
GFRA1	MCF-12A	33.4	0.06	4.593	0.001
	MCF-7	31.23	0.03		
EPHB4	MCF-12A	28.27	0.07	3.907	0.001
	MCF-7	26.33	0.03		
NENF	MCF-12A	24.27	0.09	1.175	0.258
	MCF-7	24.07	0.13		
PTGES2	MCF-12A	33.57	0.09	-1.516	0.145
	MCF-7	34.20	0.05		

3.4.3 Immunohistochemistry (IHC)

To assess the expression levels of GFRA1 protein in MCF-7 and MCF-12A cells immunohistochemistry (IHC) was performed on both cell lines using an antibody to GFRA1 as described in section 2.7.8. As discussed in section 1.2.4.2.2, ICH is often used in the diagnosis of cancer. Figure 3.11 show that the expression of GFRA1 was up regulated in the MCF-7 cells as compared to MCF-12A cells. This is indicated by the more intense brown stain observed in MCF-7 cells.

This study also evaluated the expression of GFRA1 in FFPE breast tissues of three breast cancer patients using ICH. A pathologist (Dr Y Topov, the National Cancer Institute Misurata Libya) confirmed the three patient samples used were cancerous. The three patients were diagnosed at stage I, stage II and stage III breast cancer, respectively. Figure 3.12 A/B, C/D and E/F show immunostaining for the stage I, stage II and stage III breast, respectively. The brown stain represents the presence of the GFRA1 protein, while blue stain represents the cell nuclei. Figure 3.12 show that the GFRA1 protein is expressed in epithelial layer of the milk ducts. It is also evident that the expression levels of the GFRA1 protein increase from stage I to stage III.

This result supports the proteomics data as well as the qRT-PCR data and the in silico gene expression analysis. The result obtained further strengthens the evidence that GFRA1 could be a potential biomarker for breast cancer. However, the expression level of this protein needs to be evaluated in other cancer cell lines and cancer patient tissue to determine its specificity for breast cancer.

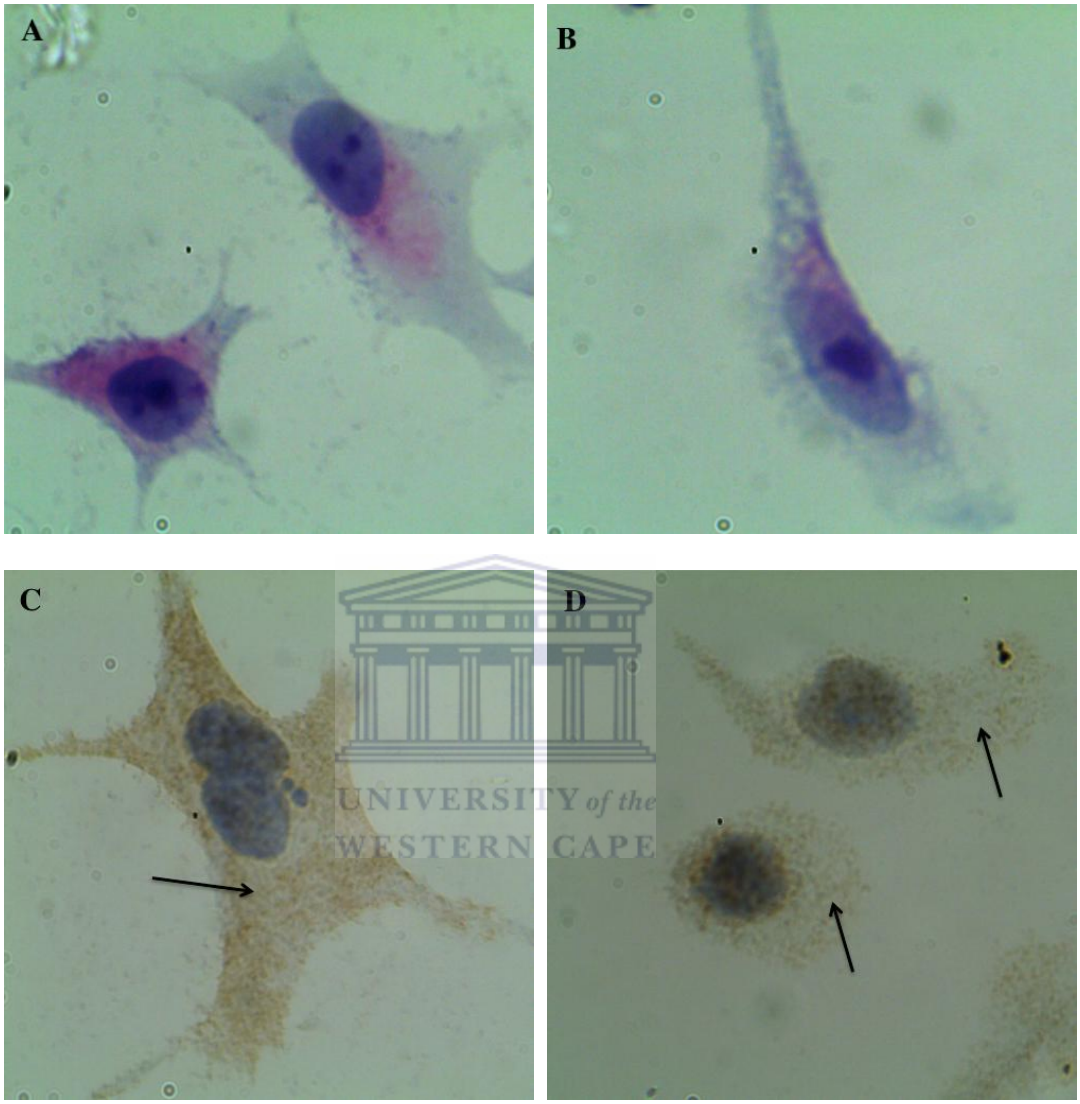


Figure 3. 11: IHC analysis of GFRA1 expression in MCF-7 and MCF-12A cells. A and B show MCF-7 and MCF-12A cells, respectively; stained with hematoxylin and eosin (H&E). C and D show MCF-7 and MCF-12A cells, respectively; stained with H&E and the GFRA1 antibody. The dark purple stain represents the nuclei, while the brown stain represents GFRA1. The black arrows point to the localisation of GFRA1. Images were taken at 1000 X magnification

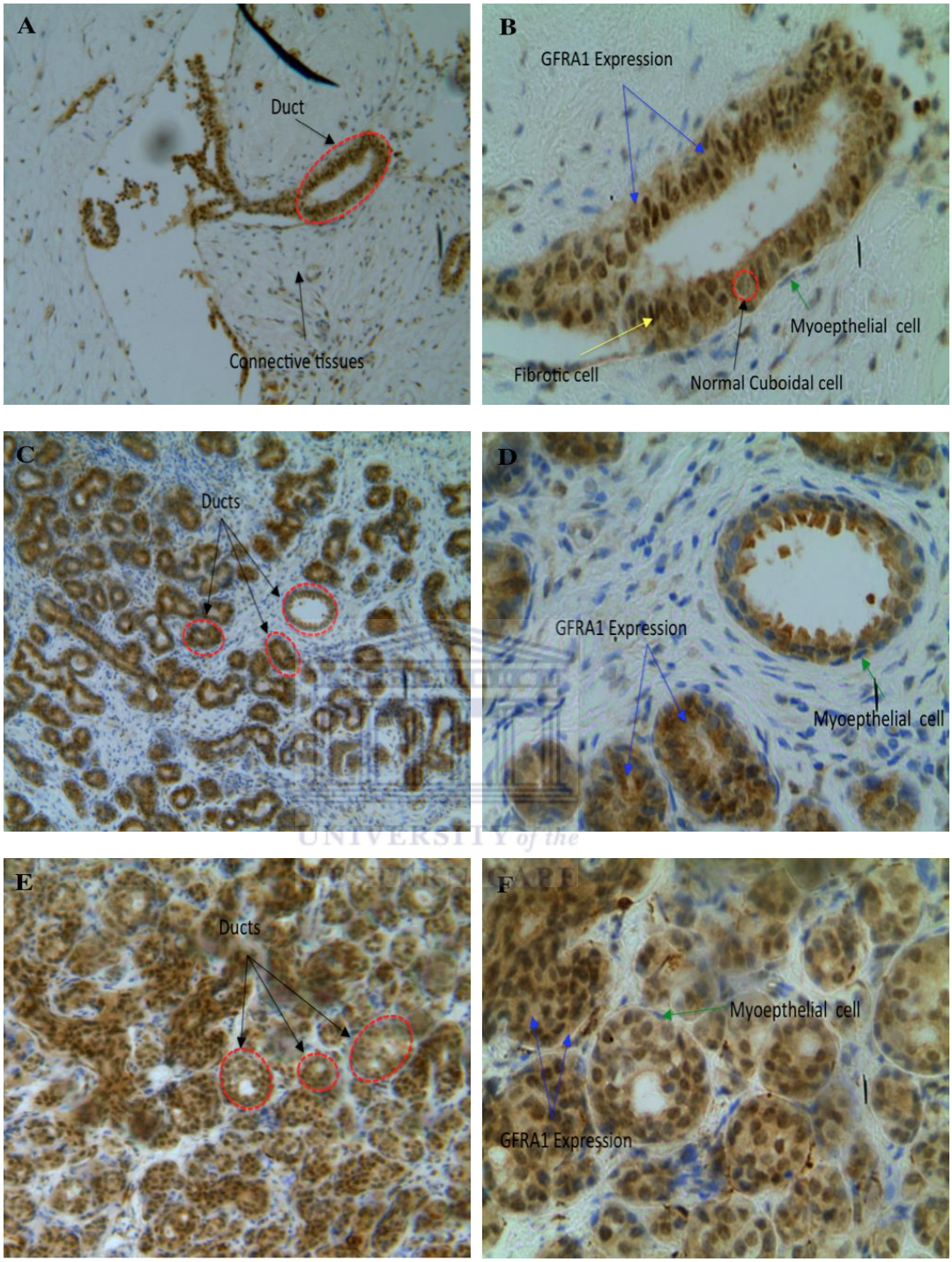


Figure 3. 12: IHC analysis of GFRA1 expression in different FFPE tissues from three breast cancer patients. A/B, C/D and E/F show immunostaining for the stage I, stage II and stage III breast, respectively. The ducts are labelled with red circles in figures (A, C and E). The normal cuboidal cells was labelled with red circles in figure B. Myoepthelial cell were indicated with green arrows in figures (B, D and F) and the fibrotic cell was appointed with yellow arrow in Figure B, whereas the GFRA1 expression in the cells in Figures (B, D and F) was shown with blue arrows. The images in Figures A, C and E were taken at 200 X magnifications, whereas the images in figures B, D and F were taken at 400 X magnifications.

3.5 Summary and Conclusion:

In this study the cellular proteins of MCF-7 and MCF-12A cells were successfully labelled in SILAC medium and the membrane proteins of the labelled cells were successfully isolated. The study identified 2831 proteins from these cells. LC-MS/MS analysis showed that only 70 % of the proteins were labelled and that some cytosolic proteins were isolated together with membrane proteins.

A comparative analysis of the relative expression levels of proteins in MCF-7 and MCF-12A cells found that 19 % (or 538) of the proteins were over expressed in MCF-7 cells. Of the 2831 proteins identified, 46 % were classified as CEM proteins localised to the membrane of the cells. Ten percentile (10 %) or 271 of the CEM proteins are only expressed in the membrane of cells, while 36 % or 1018 is expressed in both the membrane and cytoplasm of cells.

This study focussed only on proteins expressed in the membranes of cells and therefore proteins that are expressed in both the membrane and cytoplasm were excluded. The proteins with the highest (the top 10) differential expression between MCF-7 and MCF-12A cells were identified. LC-MS/MS analysis showed that GFRA1 was expressed at significantly higher levels in the membranes MCF-7 cells as compared to MCF-12A cells.

In silico gene expression analysis using GeneHub-GEPIS online tools was performed for the genes encoding the 10 proteins with the highest differential expression. GeneHub-GEPIS analysis found that all 10 genes were over expressed

in breast cancer tissue. However, 6 of these genes (*ABHD12*, *NRCAM*, *MYG1*, *TEX264*, *REEP5* and *ENPP1*) were also expressed in normal breast tissue. In contrast the other 4 genes (*GFRA1*, *NENF*, *EPHB4* and *PTGES2*) were only expressed in breast cancer tissue and not in normal breast tissue. *ABHD12*, *NRCAM*, *MYG1*, *TEX264*, *REEP5* and *ENPP1* were excluded as potential biomarkers since these genes are also expressed in normal breast tissue, leaving only 4 potential biomarker genes (*GFRA1*, *NENF*, *EPHB4* and *PTGES2*). Amongst the 4 genes, *GFRA1* stood out as a good candidate biomarker for breast cancer, since GeneHub-GEPIS analysis showed that the relative expression level of this gene was higher than the other 3 potential biomarkers. The expression profile of *GFRA1* generated by GeneHub-GEPIS analysis also compared favourably with other known breast cancer biomarkers (*HER-2/ERBB2, FOXA1* and *CK19*).

An expression analysis of GFRA1 using both qRT-PCR and IHC demonstrated that both the gene and the protein encoded by the gene was over expressed in MCF-7 cells as compared to MCF-12A cells. These results were in agreement with the proteomics data. The expression levels of GFRA1 were also investigated by IHC analysis in breast tissue samples of three breast cancer patients. The patients represented stage I, II and II breast cancer. IHC analysis showed that GFRA1 was expressed in the myoepithelium layer of the milk ducts. There is also evidence to suggest that GFRA1 expression levels increase from stage I to III.

It can be concluded from this study that GFRA1 is potentially a good candidate biomarker for breast cancer. This study show that the gene and protein encoded by this gene is over expressed in human breast cancer cell lines and tissue samples of breast cancer patients. This is supported by a previous study conducted by Esseghir and co-workers which showed that mRNA transcripts of GFRA1 or GFRα1 was up-regulated in invasive breast carcinomas (Esseghir *et al*., 2007). GFRA1 is a glycosylphosphatidylinositol(GPI)-linked transmembrane cell surface receptor for Glial cell line derived neurotrophic factor (GDNF) and mediates the activation of the Rearranged during Transfection (RET) tyrosine kinase receptor resulting in pro-survival signals. GFRA1 belongs to a GDNF receptor family comprising of GFRα1 – 4 (Carey *et al*., 2007, Lu *et al*., 2011). Another study by Esseghir and co-workers aimed at identifying transmembrane and secreted proteins that are overexpressed in breast cancer tissue and cells provided evidence that GDNF may play a role in promoting breast cancer growth by signalling via the RET and GFRA1 (Esseghir *et al*., 2007). A more recent study by Wu and co-workers investigated the prognostic significance of GFRα1, GFRα3, Syndecan-3 and artemin (ARTN) in the survival outcome of breast cancer patients (Wu *et al.*, 2013). ARTN is up-regulated in cancers and was shown to bind both GFRα1 and GFRα3. The study by Wu and co-workers demonstrated that the co-expression of GFRα1, GFRα3 and ARTN, but not Syndecan-3 was associated with a poor patient outcome, such as lymph node metastases and tumour stage (Wu *et al.*, 2013). There is therefore strong evidence that GFRA1 can be used as a biomarker for breast cancer.

3.5.1 Future considerations:

This study identified 4 membrane proteins as possible biomarkers for breast cancer. However, only one of these proteins (GFRA1) was investigated by IHC expression analysis. Although GeneHub-GEPIS analysis suggested that the other 3 membrane proteins (NENF, EPHB4 and PTGES2) are not good candidates for breast cancer biomarkers, expression analysis by IHC should be done for these membrane proteins as well.

Furthermore, even though IHC was performed for GFRA1, it was only done in two cell lines (MCF-7 and MCF-12A) and tissue from three breast cancer patients. Additional cell lines and a larger number of breast cancer patients should be screened. Moreover, tissue from healthy individuals should also be tested.

This study identified 271 CEM proteins that are up-regulated in MCF-7 cells. A large number of these proteins were excluded from the study based on the fact that these proteins are also expressed in normal breast tissue as demonstrated by the GeneHub-GEPIS analysis. However, it is possible that some of these proteins can still be used as biomarkers for breast cancer if the expression of these proteins is limited to a few tissues and if these proteins are significantly up-regulated breast cancer tissue. Further GeneHub-GEPIS analysis and expression analysis of these proteins in breast cancer samples should therefore be performed.

CHAPTER 4

4.1 Development of a molecular beacon probe for the detection of a breast cancer biomarker

4.1.1 Introduction

Molecular beacons are single stranded hairpin DNA probes that can be used to detect the presence of specific DNA sequences (Yang *et al.*, 2005). As such they can be used in the diagnosis of diseases, including breast cancer. The principle of detection of specific DNA sequences using molecular beacons is based on a fluorescent signal, where the presence of the target DNA sequence result in an increase in the fluorescent signal and the absence of the specific DNA sequence result in a low fluorescent signal (as demonstrated in Figure 4.1) (Yang *et al.*, 2005). The molecular beacon consists of a Loop structure that is complementary to the specific target DNA sequence, which can be a gene (or biomarker) implicated in a disease or genetic material of an infectious organism (a bacterium or virus). The Loop structure is flanked on each side by a region of reverse complementary DNA sequences (the Stem structure), which has the ability to hybridise to each other and form a region of double stranded DNA, the Stem. One end of the molecular beacon is conjugated to a fluorescent molecule such as texas red or cyanine3 (Cy3), while the other end is conjugated to a quencher molecule such as Iowa black or dimethylaminoazobenzen aminoexal-3-acryinido (Dabcyl) (Fang *et al*., 1999, Marras and Salvatore. 2006). In the absence of the target DNA sequence, the fluorescent molecule is in close proximity to the quencher molecule

because of hybridisation of the Stem structures and as a result the fluorescence signal is suppressed. In the presence of the target DNA sequence, the fluorescent molecule is separated from the quencher molecule because the target DNA sequence hybridise to the Loop structure of the molecular beacon and as a result the fluorescence signal is increased. The fluorescent signal typically comes from an organic fluorophore such as texas red, cyanine3 (Cy3), fluorescein, etc. As pointed out previously, there are a number of limitations when organic fluorophores are used in fluorescent diagnostic assays. Some molecular beacon studies have been reported where the organic fluorophores have been replaced with quantum dots (Bryers *et al.*, 2011). Molecular beacons using quantum dots would be more stable and can enable the development of diagnostic systems that can detect multiple biomarkers in the same sample. Although, this kind of multiplex detection can be done by qRT-PCR, it is very costly, requiring high-end instruments and highly trained technicians. Diagnostic assays using molecular beacons can be analysed using a plate reader and the cost of the assay would be significantly lower compared to qRT-PCR.

It was demonstrated in Chapter 3 by qRT-PCR that the known biomarker for breast cancer, CK19 was significantly up-regulated in the breast cancer cell line, MCF-7 as compared to the non-cancerous cell line, MCF-12A. The aim in this part of the study was to develop a molecular beacon probe for CK19 and to investigate the application of this molecular beacon probe to assess the expression levels of CK19 in MCF-7 and MCF-12A cells.

4.2 Results and Discussion

4.2.1 Design and development of molecular beacon probe

Molecular beacon probes consisting of a 20 base pair (bp) region (the Loop structure) flanked by either a 5 or 7 bp region (the Stem structure) was designed (Figure 4.2) to detect mRNA of the human CK19 gene (ENSG00000171345). The Loop structure was a single-stranded DNA oligonucleotide designed based on the mRNA sequence for the CK19 gene as described in section 2.8. This single-stranded DNA oligonucleotide was 20 base pairs long and was complementary to region spanning bp 1292 to 1312 of the human CK19 mRNA. BLAST analysis against the human genome database was used to confirm the specificity of this sequence for the detection of human CK19 mRNA. The undesired hybridization and denaturation of the Stem structure due to temperature fluctuations can negatively affect the usefulness of the molecular beacon. Therefore two molecular beacon probes, one with a Stem of 5 bp (MB-CK19/Short) and one with a Stem of 7 bp (MB-CK19/Long) were designed. The DNA sequence for the Stem region was adapted from the study by (Li *et al.*, 2008), in their study new signal-amplifying mechanism and nicking enzymes signal amplification (NESA) was created to improve the sensitivity of molecular beacons.

The oligonucleotides probes were synthesised at Integrated DNA Technologies. The quencher molecule, Iowa Black was coupled to the 5'end of the oligonucleotides probes and the 3'end was modified with and amino group as described in section 2.8. Qdot525 ITK carboxyl quantum dots were obtained from

Life Technologies. These quantum dots (100 pM) were conjugated to the oligonucleotides probes (200 nM) as described in section 2.8.2.

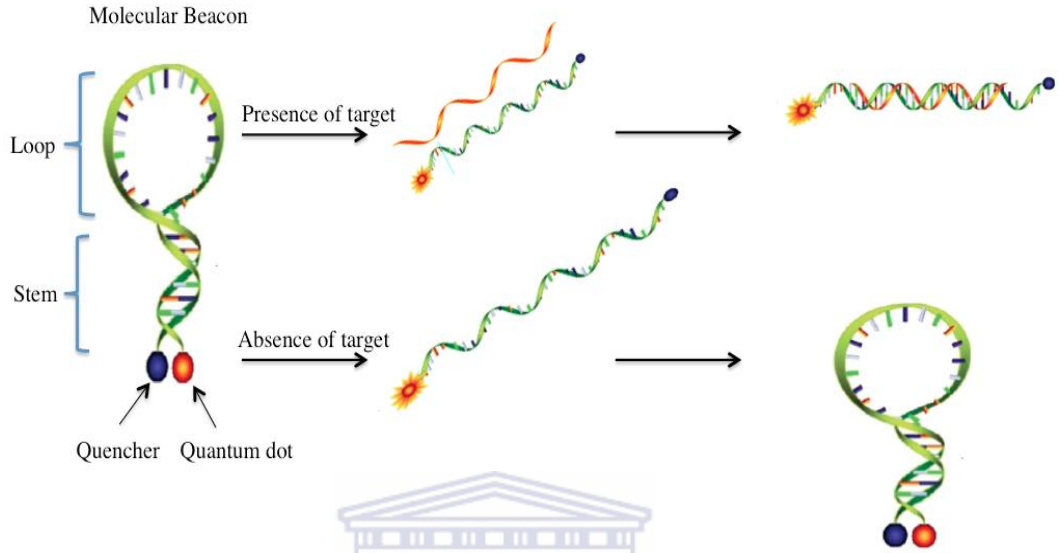


Figure 4. 1: Illustration of molecular beacon in the presence or absence of target.

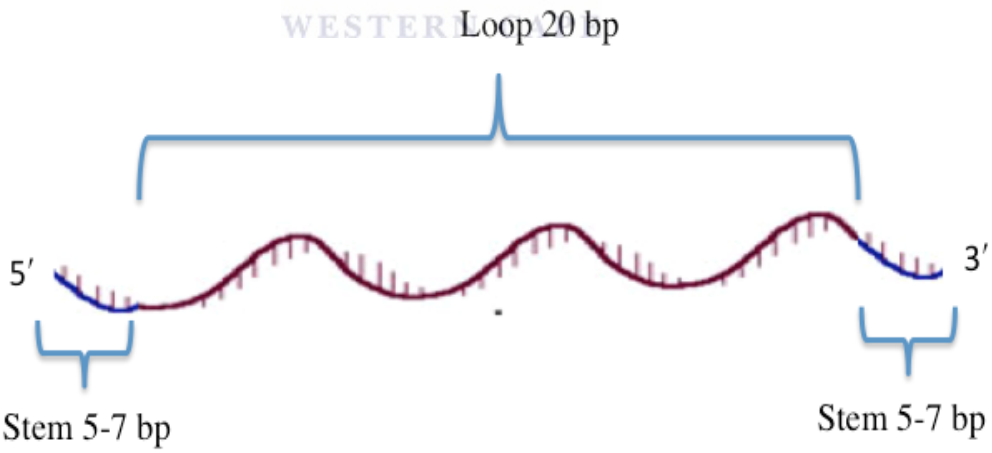


Figure 4. 2: Design of the molecular beacon probe.

4.2.2 Testing the ability of the CK19 molecular beacon probes to detect a synthetic complementary DNA sequence

An oligonucleotide (CK19-Com-Seq) with a sequence that was reverse complementary to the Loop structure of the CK19 molecular beacon probe was acquired from Integrated DNA Technologies and used as a target sequence for the testing of the CK19 molecular beacon probes. This oligonucleotide was used to test the ability of the CK19 molecular beacon probes, to detect complementary DNA sequences. In addition, this complementary oligonucleotide was also used to do a comparative analysis between the effectiveness of a molecular beacon probe with a short Stem (MB-CK19/Short) and one with a long Stem (MB-CK19/Long), as well as the effect of temperature on the result.

The fluorescence intensity of 10 μM MB-CK19/Short and 10 μM MB-CK19/Long were measured in the presence and absence of 25 μM CK19-Com-Seq. MB-CK19/Short and MB-CK19/Long were incubated in the presence and absence of CK19-Com-Seq at 60°C (the calculated annealing temperature for the CK19-Com-Seq to the Loop structure) for 5 min, cooled down to either 20 °C or 25°C and the fluorescence intensity was then measured using a POLARstar Omega microplate reader as described in section 2.8.3. The calculated annealing temperatures for the Stem structure of MB-CK19/Short and MB-CK19/Long was 20 °C and 25 °C, respectively and therefore the fluorescence intensities were evaluated at both temperatures. Figure 4.3 shows that the fluorescence intensity for both MB-CK19/Short and MB-CK19/Long increased significantly in the presence of CK19-Com-Seq. In comparison, the fluorescence intensity for MB-CK19/Short and MB-CK19/Long in the absence of CK19-Com-Seq was very low.

This suggested that CK19-Com-Seq binds to the Loop structure of the molecular beacons. The p-values for MB-CK19/Short were 0.01668 and 0.000352 at 20 °C and 25 °C, respectively. The p-values for MB-CK19/Long were 0.080285 and 0.00081 at 20 °C and 25 °C, respectively. This suggested that 25 °C is more suitable than 20 °C for both MB-CK19/Short and MB-CK19/Long.

No significant difference in the fluorescence intensity between MB-CK19/Short and MB-CK19/Long was observed at these temperatures. However, a greater degree of variation in the fluorescence intensity was observed for MB-CK19/Short when compared to MB-CK19/Long. This suggested that MB-CK19/Long was more stable than MB-CK19/Short.

To determine the lower detection limit of MB-CK19/Short and MB-CK19/Long, the fluorescence intensity of these molecular beacons were measured in the presence of CK19-Com-Seq at concentrations lower than 25 μM. Figure 4.4 show that both MB-CK19/Short and MB-CK19/Long can still detect CK19-Com-Seq at concentrations as low as 5 μM. However no significant differences could be detected for MB-CK19/Short between the different concentrations. A significant difference (p-value = 0.0091) was detected for MB-CK19/Long between 10 to 25 μM concentrations of CK19-Com-Seq. This suggested that the detection limit for MB-CK19/Long was 10 μM.

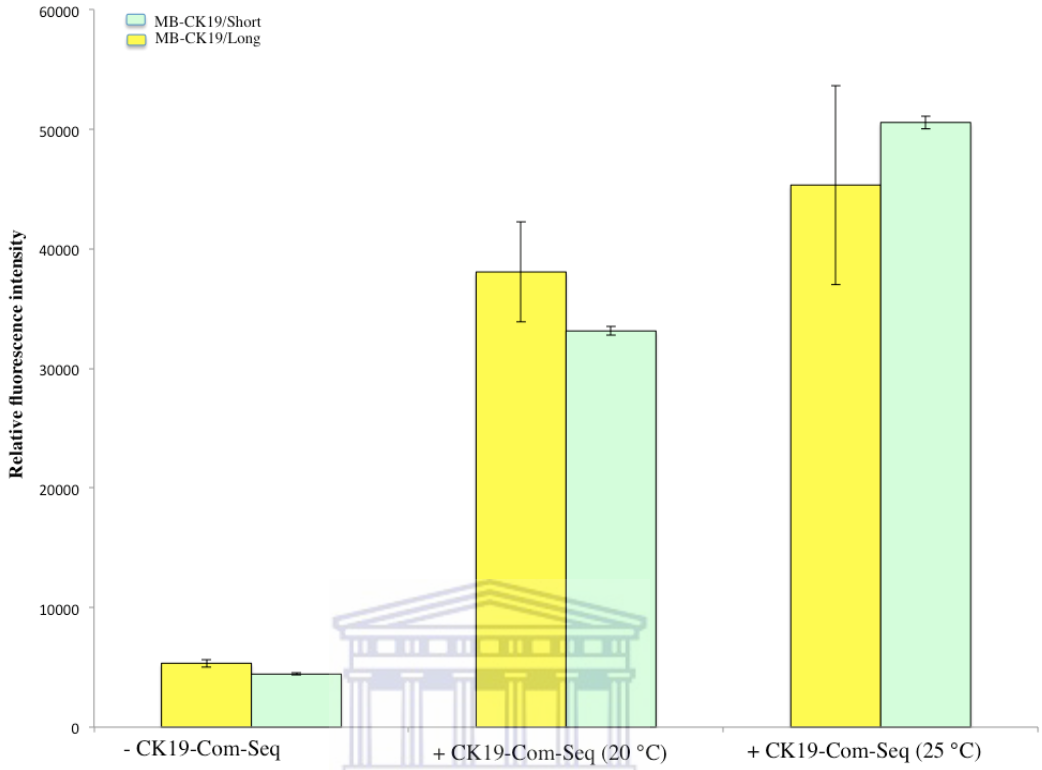


Figure 4. 3: Comparison between MB-CK19/Short and MB-CK19/Long at 20 °C and 25 °C. The fluorescence intensity of MB-CK19/Short and MB-CK19/Long (10 µM) was measured in the presence and absence of 25 µM CK19-Com-Seq. Fluorescence intensity was measured using the POLAR star Omega Microplate reader. (-) indicate the absence of CK19 complementary, (+) indicates the presence of CK19 complementary. n=3.

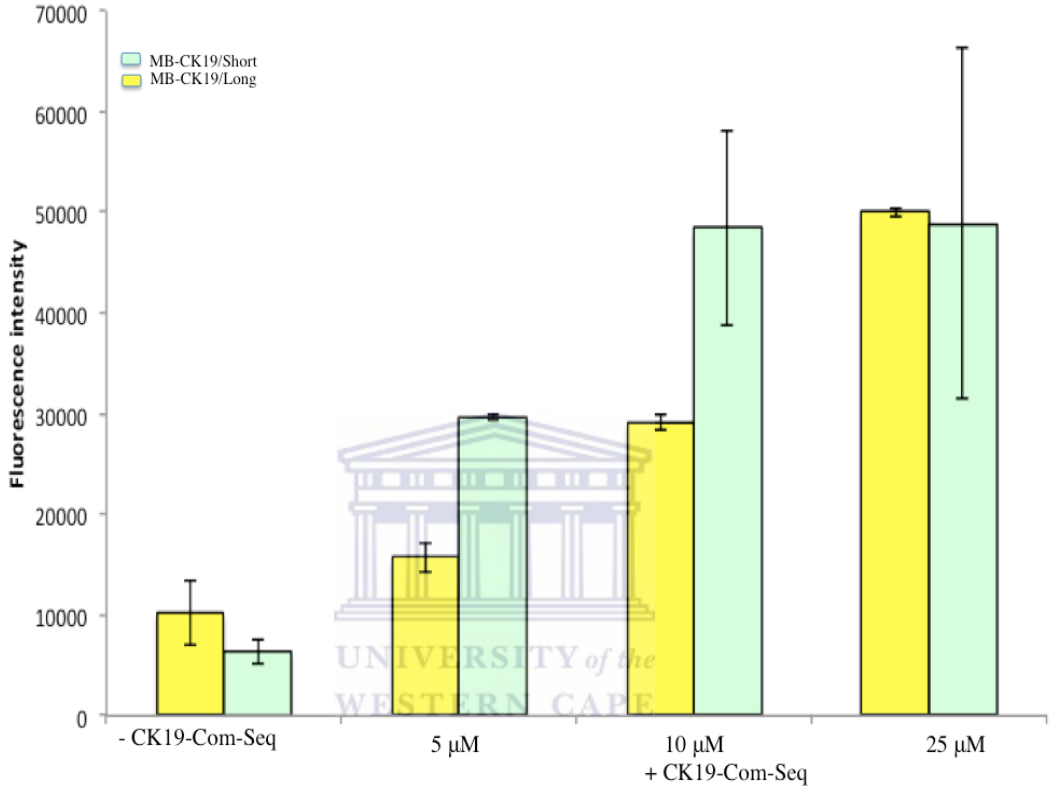


Figure 4. 4: Evaluating the lower detection limit of MB-CK19/Short and MB-CK19/Long. The fluorescence intensity of MB-CK19/Short and MB-CK19/Long (10 µM) was measured in the presence and absence of 5 µM, 10 µM and 25 µM CK19-Com-Seq. Fluorescence intensity was measured using the POLAR star Omega Microplate reader. (-) indicate the absence of CK9 complementary, (+) indicates the presence of CK19 complementary. n=3.

4.2.3 Evaluating the application of MB-CK19/Long to detect CK19 in MCF7 and MCF-12A samples

The MB-CK19/Long molecular beacon probe proved to be more stable than the MB-CK19/Short probe, as demonstrated in section 4.2.2. To evaluate the application of the MB-CK19/Long molecular beacon probe, cDNA samples produced from the mRNA isolated from MCF-7 and MCF-12A cells were used as the testing material to assess relative expression levels of CK19 in these two cell lines.

It was already demonstrated in Chapter 3 that the expression levels of CK19 was significantly higher in MCF-7 cells. qRT-PCR analysis demonstrated a 26.895 fold increase in the expression levels of CK19 in MCF-7 cells in comparison to MCF-12A cells. The same cDNA samples that were used in the qRT-PCR analysis (section 2.6.3) were used as testing material for the evaluation of the MB-CK19/Long molecular beacon probe. The fluorescence intensity of 10 μM MB-CK19/Long was measured in the presence and absence of 10 μM cDNA produced from MCF-7 and MCF-12A cells. Figure 4.5 shows a significant (p-value = 0.0483) increase in the fluorescence intensity of MB-CK19/Long in the presence of MCF-7 cDNA. This increase in fluorescence intensity was not observed in the presence of MCF-12A cDNA (p-value = 0.418). A significant difference (p-value = 0.0462) in the fluorescence intensity for MCF-7 and MCF-12A cDNA was also observed. This demonstrated that the MB-CK19/Long molecular beacon probe was able to detect CK19 cDNA in a complex samples such as the cDNA library prepared from a cell culture. This probe was also able to demonstrate that the

expression level of CK19 in MCF-7 cells was higher as compared to MCF-12A cells.

This essentially means that the MB-CK19/Long molecular beacon probe can give the same result as qRT-PCR analysis, but at a fraction of the cost of the qRT-PCR assay. However, the sensitivity of qRT-PCR assay is still significantly higher than the molecular beacon probe assay, since much higher concentrations of the cDNA were required. In addition, although the assay with the molecular beacon probe was able to differentiate between samples with low and high concentrations of the target DNA, more research is needed to use this molecular beacon probe quantitatively. Since this molecular beacon probe makes use of quantum dots for the fluorescence signal, there is a possibility that such molecular beacon probes can be used in multiplex diagnostic systems to evaluate the presence of several biomarkers in the same sample. As an example, several breast cancer biomarkers (e.g. *CK19*, *FOXA1*, *BRCA2* and *HER-2*) can be detected in a patient sample simultaneously. This can significantly reduce the cost and time of the diagnostic assay.

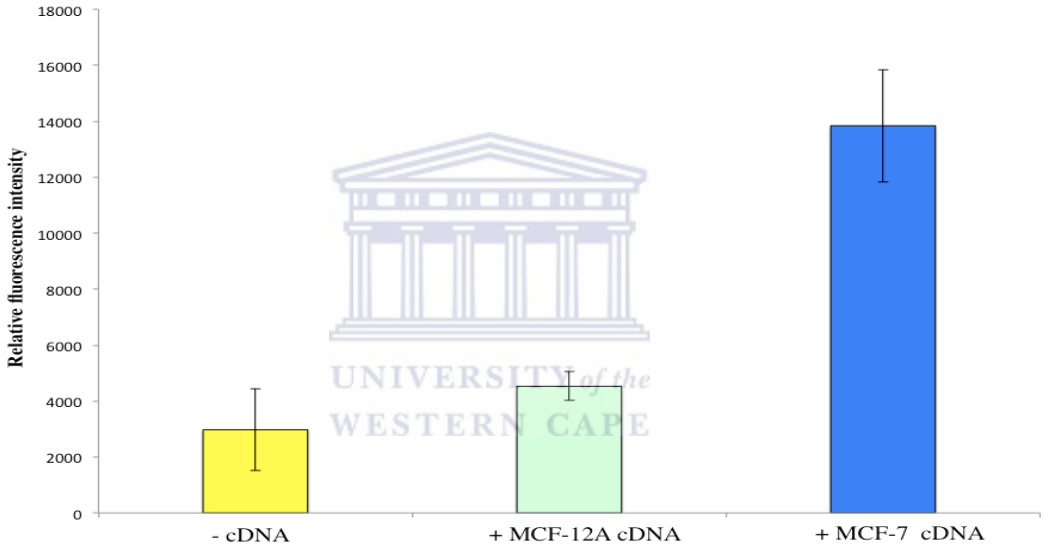


Figure 4. 5: Evaluating the expression levels of CK19 in MCF-7 and MCF-12A cDNA libraries using MB-CK19/Long. Fluorescence intensity was measured using the POLAR star Omega Microplate reader. n=3. (-) indicates the absence of cDNA; (+) indicates the presence cDNA.

4.3 Summary and Conclusion:

This study aimed to develop a molecular beacon probe for the detection of cDNA sequences for the breast cancer biomarker, CK19. The molecular probes developed here used Iowa Black as the quencher molecule and Qdot525 ITK carboxyl quantum dot as the fluorescent molecule.

Two probes (MB-CK19/Short and MB-CK19/Long) were developed to test the effects of the Stem length on molecular beacon. A series of control experiments using a synthetic oligonucleotide (CK19-Com-Seq) that was complementary to the Loop structure of the molecular beacon were performed to test the effects of the length of the Stem structure and to determine the lower detection limit of the molecular beacon. This study demonstrated that both MB-CK19/Short and MB-CK19/Long was able to detect the presence of CK19-Com-Seq. The measurement of the fluorescence intensity at a temperature of 25 $^{\circ}$C produced more reliable results for both MB-CK19/Short and MB-CK19/Long. The variance in fluorescence intensity for MB-CK19/Short was significantly higher compared to MB-CK19/Long. This can only be ascribed to the difference in length of the Stem structure. This proved that MB-CK19/Long was more suitable than MB-CK19/Short to detect CK19-Com-Seq. This study also demonstrated that the lower detection limit for MB-CK19/Long was 10 μM.

MB-CK19/Long was also used to evaluate presence of CK19 cDNA in a cDNA library prepared from MCF-7 and MCF-12A cells. This study showed that MB-CK19/Long was able to detect the presence of CK19 cDNA in a cDNA library

prepared from MCF-7 cells, but not MCF-12A cells. This result is in agreement with the qRT-PCR result for CK19, which showed that this gene is over expressed in MCF-7 cells. The fact that MB-CK19/Long was able to detect the presence of CK19 cDNA in a cDNA library was unexpected since experiments with CK19-Com-Seq showed that the detection limit for this probe is 10 μM and it is expected that the concentration of CK19 cDNA would be much lower than 10 μM. However, this study does provide evidence that MB-CK19/Long can be used to detect CK19 cDNA in a cDNA library.

4.3.1 Future considerations:

Although this study demonstrates the potential of molecular beacons probes such as MB-CK19/Long to detect specific DNA targets, the detection limits of these probes are very high and would not be able to detect DNA targets that are expressed at low levels. Therefore, the sensitivity of these probes to detect their targets should be improved. One possible solution is to use more sensitive instrumentation to detect the fluorescence.

Molecular beacon probes such as MB-CK19/Long can potentially lead to the development of multiplex detection systems, capable of detecting several breast cancer biomarkers in a single sample. This is possible because these probes use quantum dots as fluorescence signals. This can be investigated by MB-CK19/Long in combination with other molecular beacon probes that are designed to detect other breast cancer biomarkers such as *HER-2*, BRCA, FOXA1, etc.

CHAPTER 5

5.1 General conclusion

This study identified 4 putative novel cell surface protein biomarkers (*GFRA1*, *NENF*, *EPHB4* and *PTGES2*) for the diagnosis of breast cancer. The study demonstrated that *GFRA1* was significantly overexpressed in a human breast cancer cell line (MCF-7) and breast cancer patient tissue. Although the other 3 genes are also up-regulated in breast cancer cells, the expression levels of these genes are not as high as the expression levels of *GFRA1*. In silico expression analysis also shows that although *GFRA1* is also over expressed in some other types of cancer, the expression levels of this gene is significantly higher breast cancer. These biomarkers can potentially be used in the development of POC diagnostic systems for the detection of breast cancer.

This study also investigated the development of molecular beacons, which were designed using quantum dots as the fluorescent tags, for applications in the diagnosis of breast cancer using the known breast cancer biomarker, CK19. The data shows that this molecular beacon can distinguish between cDNA samples generated from a cancerous (MCF-7) (which overexpress this biomarker) and a non-cancerous cell line (MCF-12A).

This molecular beacon technology can potentially also be used to develop diagnostic tools for the novel biomarkers (*GFRA1*, *NENF*, *EPHB4* and *PTGES2*) identified in this study.

Appendix

Table 1: Reagents utilized in cDNA synthesis reaction in order

Reagent	**Final Concentration**
RNA	1 µg
Anchored-Oligo–p(dT)$_{15}$ Primer	2.5 µM
PCR grade water	Variable
Reaction Buffer	1 X
RNase Inhibitor	20 U
Deoxynucleotide Mix	1 mM
Trancriptor Reverse Transcriptor	10 U
Total volume	20 µL

Table 2: Quantitative real-time PCR cycling parameters.

Programs		
Program Name	**Cycles**	**Analysis Mode**
pre incubation	1	None
Amplification	40	Quantification
melting	1	Melting Curves
cooling	1	None

Temperature Targets

Target (°C)	Acquisition Mode	Hold (hh:mm:ss)	Ramp Rate (°C/s)	Acquisitions (per °C)	Sec Target (°C)	Step Size (°C)	Step Delay (cycles)
95	None	00:03:00	4.4		0	0	0

Amplification Temperature Targets

Target (°C)	Acquisition Mode	Hold (hh:mm:ss)	Ramp Rate (°C/s)	Acquisitions (per °C)	Sec Target (°C)	Step Size (°C)	Step Delay (cycles)
95	None	00:00:10	4.4		0	0	0
* 61	None	00:00:20	2.2		0	0	0
72	Single	00:00:01	4.4		0	0	0

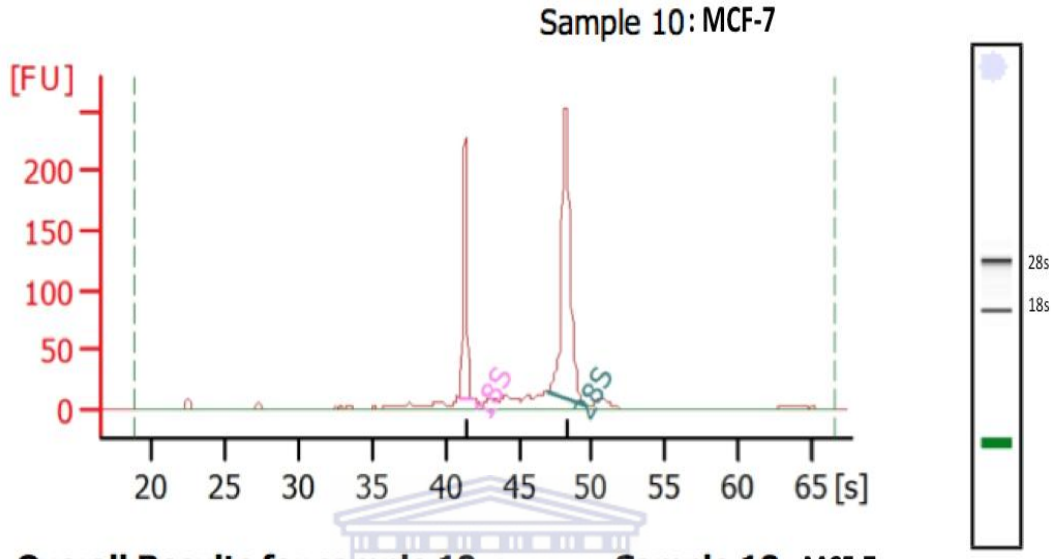
melting Temperature Targets

Target (°C)	Acquisition Mode	Hold (hh:mm:ss)	Ramp Rate (°C/s)	Acquisitions (per °C)	Sec Target (°C)	Step Size (°C)	Step Delay (cycles)
95	None	00:00:05	4.4				
65	None	00:01:00	2.2				
97	Continuous		0.11	5			

cooling Temperature Targets

Target (°C)	Acquisition Mode	Hold (hh:mm:ss)	Ramp Rate (°C/s)	Acquisitions (per °C)	Sec Target (°C)	Step Size (°C)	Step Delay (cycles)
40	None	00:00:40	2.2		0	0	0

*Annealing temperature is primer dependent, the range between 57 to 63C



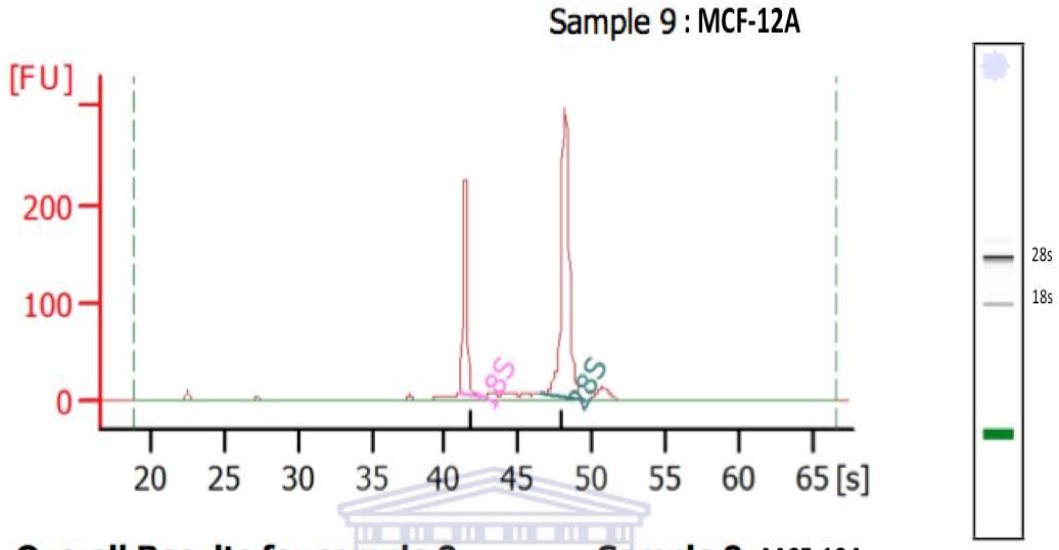
Overall Results for sample 10 : Sample 10 : MCF-7

RNA Area:	781.0
RNA Concentration:	292 ng/µl
rRNA Ratio [28s / 18s]:	2.2
RNA Integrity Number (RIN):	9.7 (B.02.08)
Result Flagging Color:	
Result Flagging Label:	RIN: 9.70

Fragment table for sample 10 : Sample 10 : MCF-7

Name	Start Time [s]	End Time [s]	Area	% of total Area
18S	40.99	41.86	157.9	20.2
28S	46.83	49.59	349.9	44.8

Figure 1: Total RNA sample of MCF-7 was analyzed by using bio-analyzer**.** The electropherogram used to obtain RNA integrity Number (RIN).. The RIN value of 9.7 was measured for MCF-7.



Overall Results for sample 9 : Sample 9 : MCF-12A

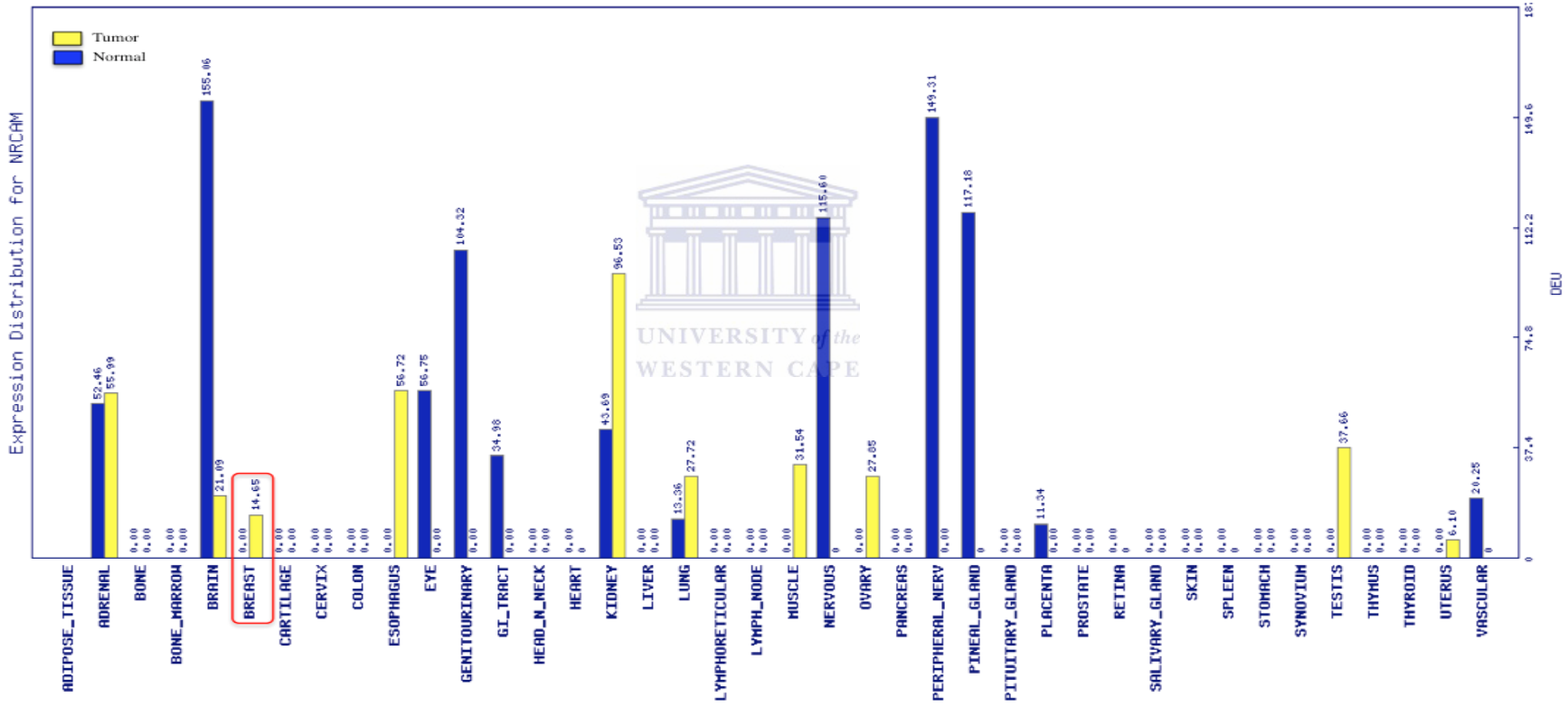
RNA Area:	771.3
RNA Concentration:	288 ng/µl
rRNA Ratio [28s / 18s]:	2.3
RNA Integrity Number (RIN):	10 (B.02.08)
Result Flagging Color:	
Result Flagging Label:	RIN:10

Fragment table for sample 9 : Sample 9 : MCF-12A

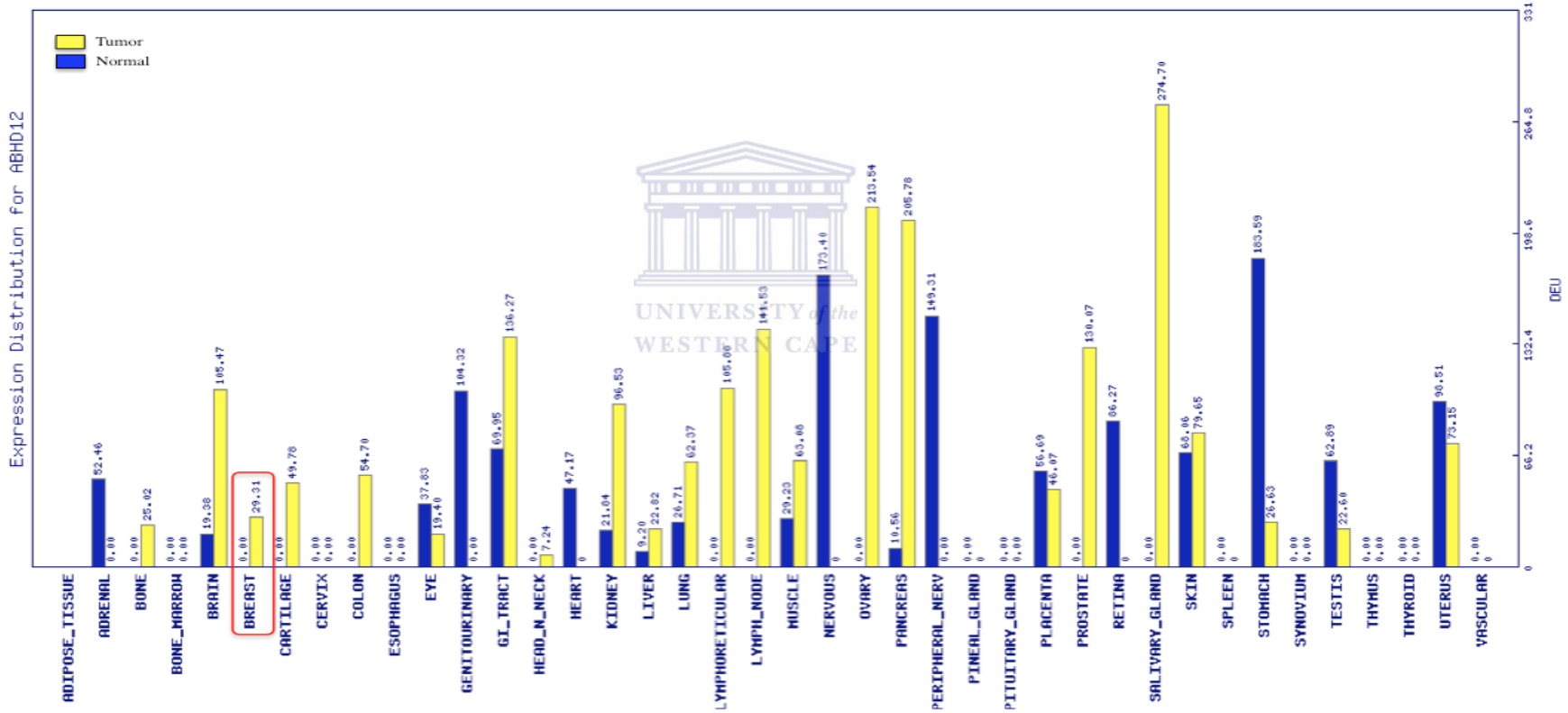
Name	Start Time [s]	End Time [s]	Area	% of total Area
18S	41.08	42.48	167.7	21.7
28S	46.50	49.60	387.0	50.2

Figure 2: Total RNA sample of MCF-12A was analyzed by using bio-analyzer**.** The electropherogram used to obtain RNA integrity Number (RIN).. The RIN value of 9.7 was measured for MCF-12A.

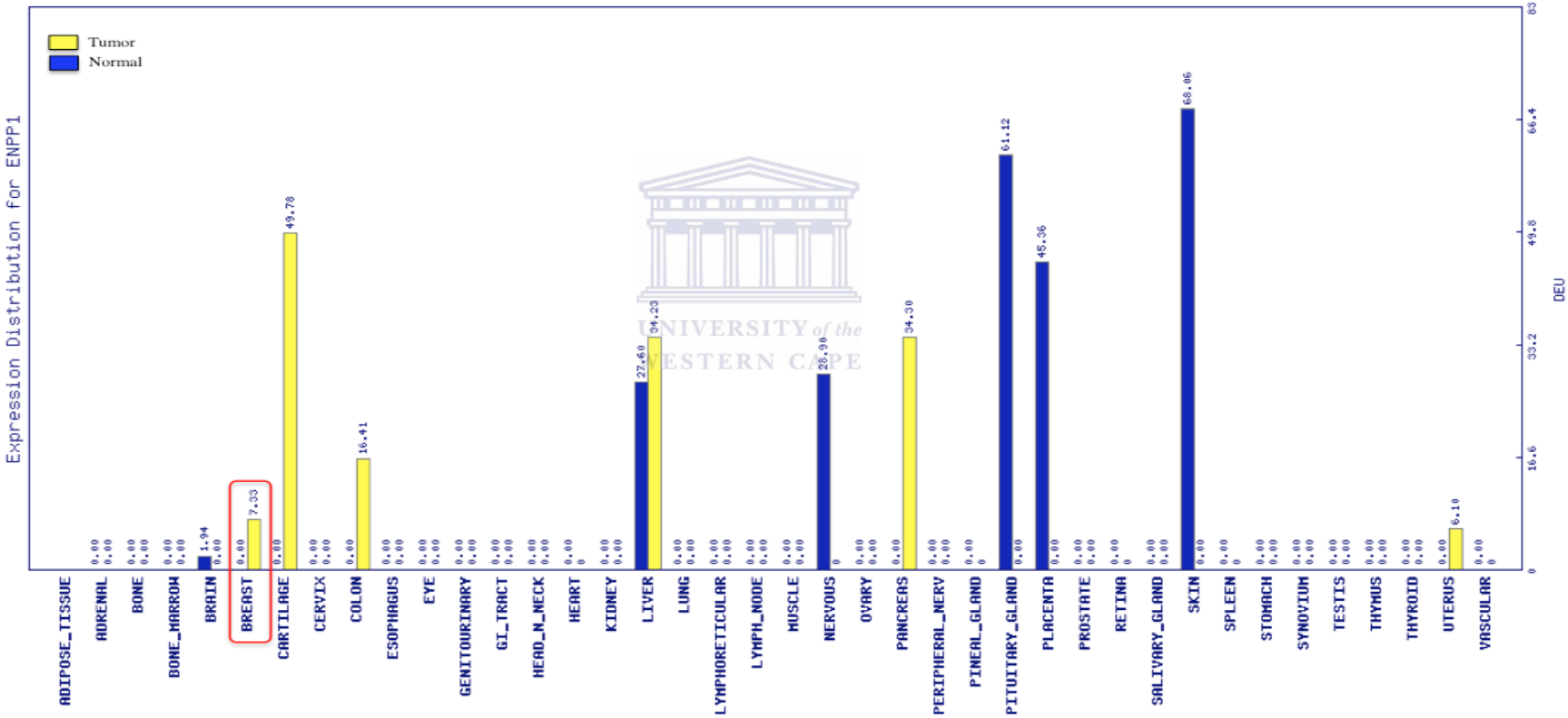
A


Expression Distribution for NRCAM
Tumor
Normal
ADIPOSE_TISSUE
ADRENAL
BONE
BONE_MARROW
BRAIN
BREAST
CARTILAGE
CERVIX
COLON
ESOPHAGUS
EYE
GENITOURINARY
GI_TRACT
HEAD_N_NECK
HEART
KIDNEY
LIVER
LUNG
LYMPHORETICULAR
LYMPH_NODE
MUSCLE
NERVOUS
OVARY
PANCREAS
PERIPHERAL_NERV
PINEAL_GLAND
PITUITARY_GLAND
PLACENTA
PROSTATE
RETINA
SALIVARY_GLAND
SKIN
SPLEEN
STOMACH
SYNOVIUM
TESTIS
THYMUS
THYROID
UTERUS
VASCULAR
52.46
55.99
155.06
21.09
14.65
56.72
56.75
104.32
34.98
43.69
96.53
13.36
27.72
31.54
135.60
27.85
149.31
117.18
11.34
37.66
6.10
20.25
0
37.4
74.8
112.2
149.6
DEU
UNIVERSITY of the
WESTERN CAPE

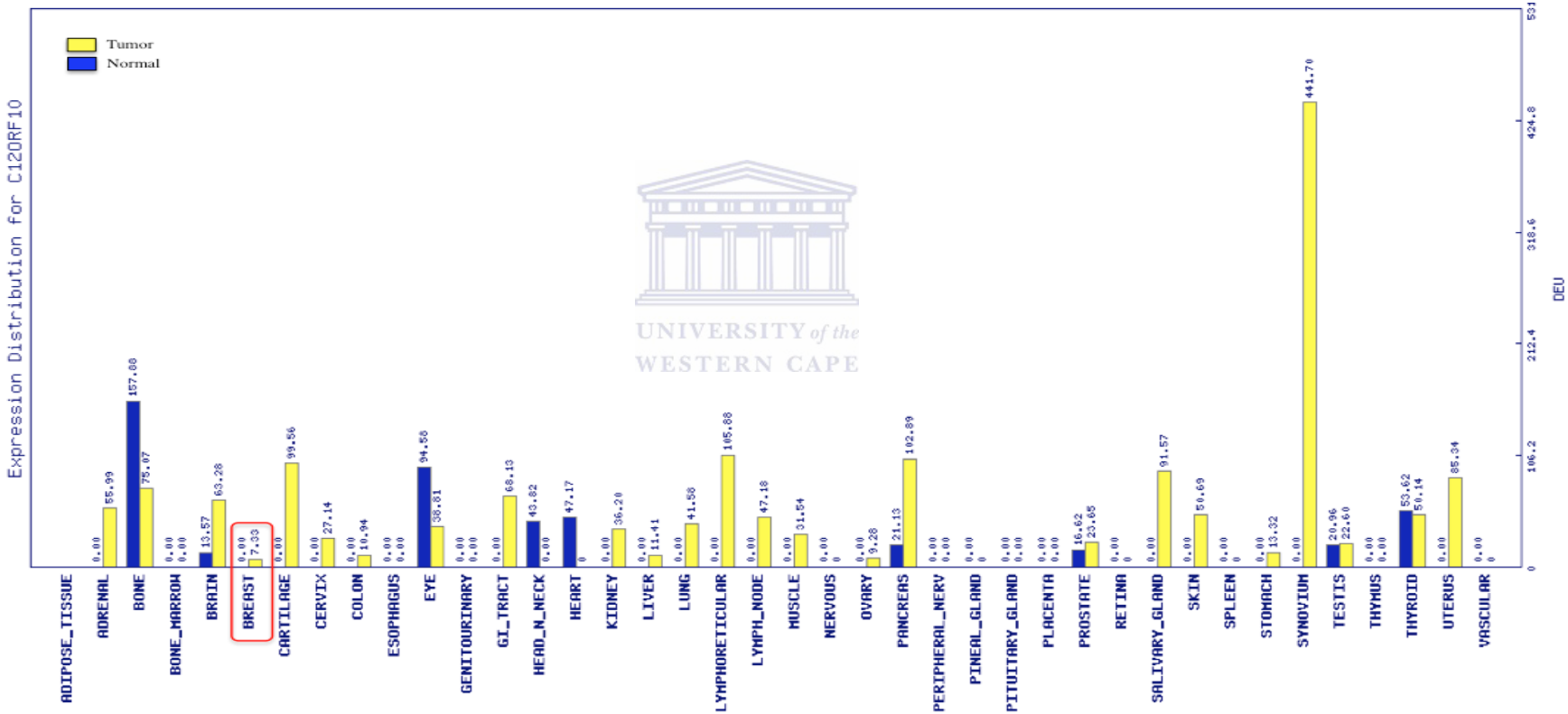
B


Expression Distribution for ABHD12
Tumor
Normal
ADIPOSE_TISSUE
ADRENAL
BONE
BONE_MARROW
BRAIN
BREAST
CARTILAGE
CERVIX
COLON
ESOPHAGUS
EYE
GENITOURINARY
GI_TRACT
HEAD_N_NECK
HEART
KIDNEY
LIVER
LUNG
LYMPHORETICULAR
LYMPH_NODE
MUSCLE
NERVOUS
OVARY
PANCREAS
PERIPHERAL_NERV
PINEAL_GLAND
PITUITARY_GLAND
PLACENTA
PROSTATE
RETINA
SALIVARY_GLAND
SKIN
SPLEEN
STOMACH
SYNOVIUM
TESTIS
THYMUS
THYROID
UTERUS
VASCULAR
0
66.2
132.4
198.6
264.8
331
DEU
UNIVERSITY of the
WESTERN CAPE

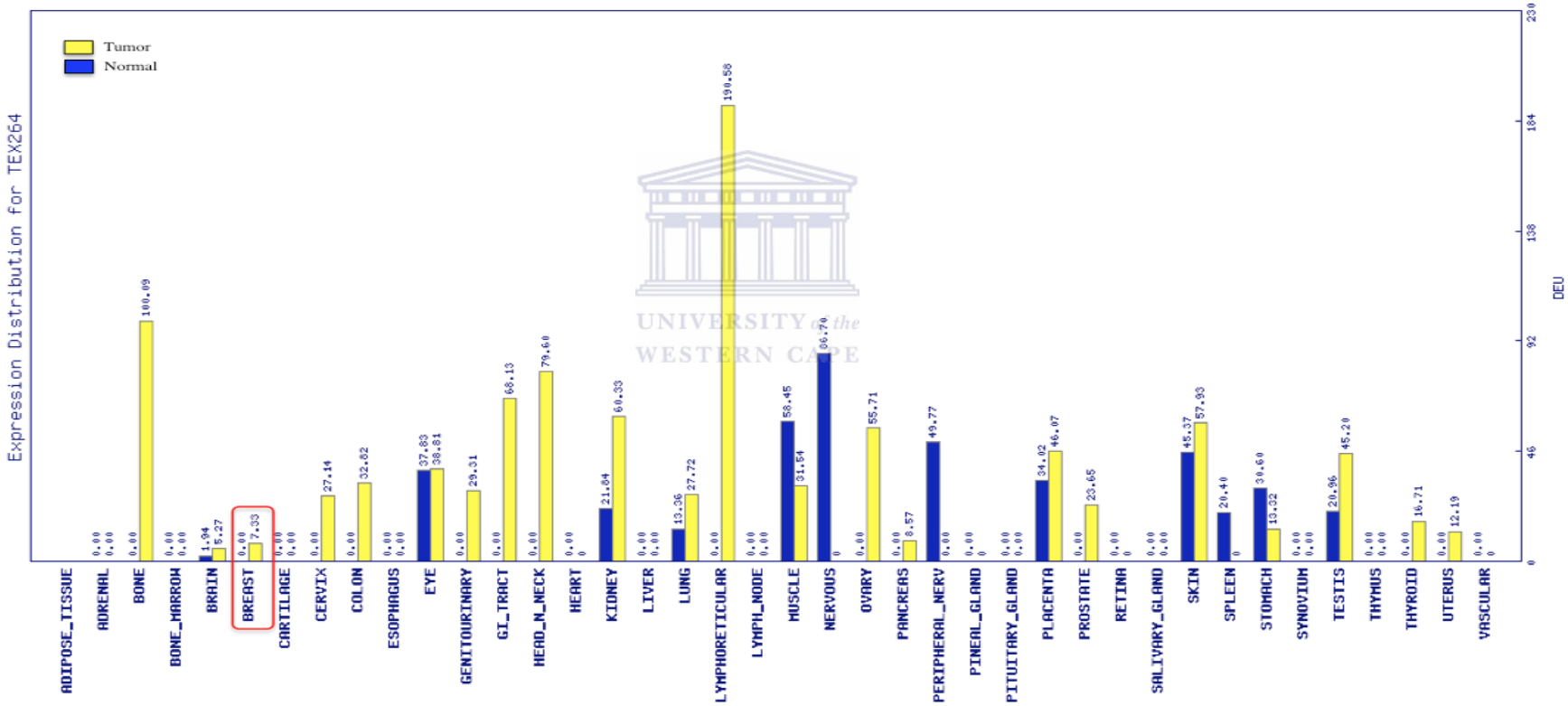
C



D


Expression Distribution for C120RF10
Tumor
Normal
ADIPOSE_TISSUE
ADRENAL
BONE
BONE_MARROW
BRAIN
BREAST
CARTILAGE
CERVIX
COLON
ESOPHAGUS
EYE
GENITOURINARY
GI_TRACT
HEAD_N_NECK
HEART
KIDNEY
LIVER
LUNG
LYMPHORETICULAR
LYMPH_NODE
MUSCLE
NERVOUS
OVARY
PANCREAS
PERIPHERAL_NERV
PINEAL_GLAND
PITUITARY_GLAND
PLACENTA
PROSTATE
RETINA
SALIVARY_GLAND
SKIN
SPLEEN
STOMACH
SYNOVIUM
TESTIS
THYMUS
THYROID
UTERUS
VASCULAR
0.00 55.99 157.66 0.00 13.57 63.28 0.00 7.33 99.56 0.00 27.14 0.00 10.94 0.00 94.58 38.81 0.00 68.13 43.82 0.00 47.17 0 0.00 36.20 0.00 11.41 0.00 41.58 0.00 105.88 0.00 47.18 0.00 31.54 0.00 0 0.00 9.28 21.13 102.89 0.00 0.00 0.00 0 0.00 0.00 0.00 0.00 0.00 0.00 16.62 23.65 0.00 0 0.00 91.57 0.00 50.69 0.00 0 0.00 13.32 0.00 441.70 20.96 22.60 0.00 0.00 53.62 50.14 0.00 85.34 0.00 0
0 106.2 212.4 318.6 424.8 531
DEU
UNIVERSITY of the
WESTERN CAPE

E



F

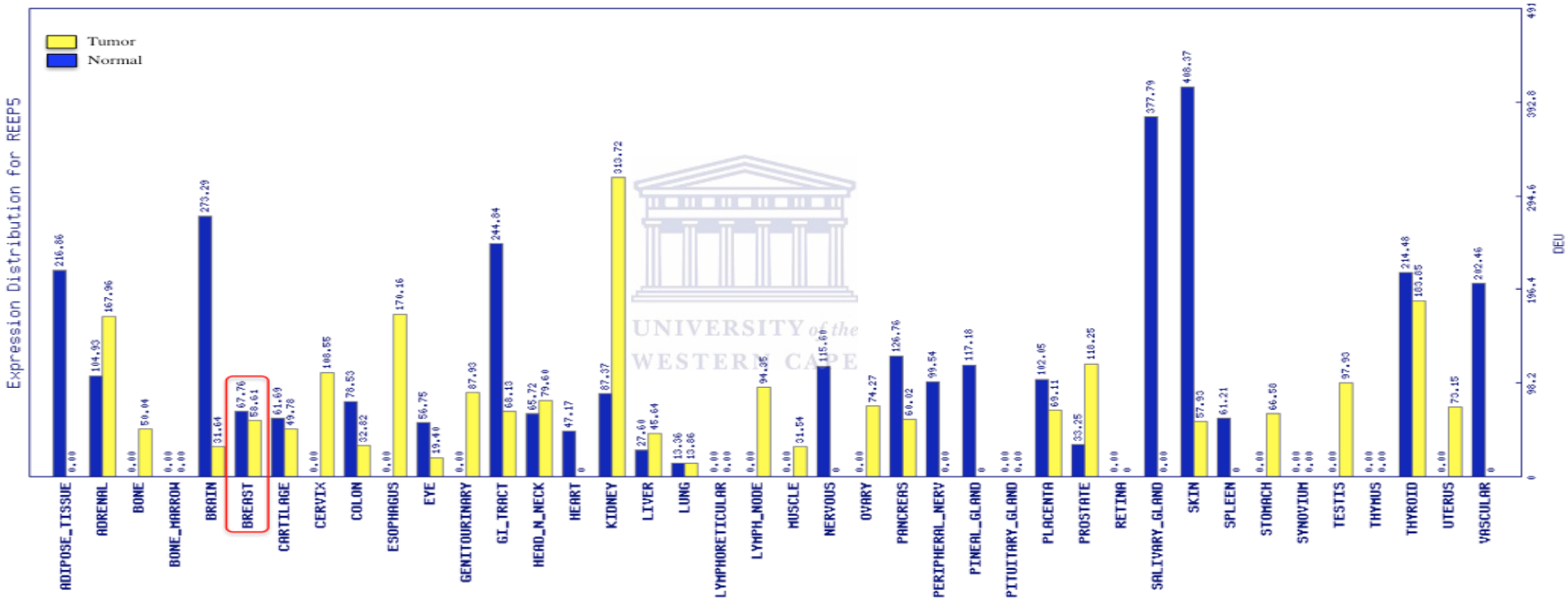


Figure 3. GeneHUb-GEPIS charts showing the expression profile of 6 of 10 slected biomarker genes in 40 different tissues. A, B, C, D, E and F shows the expression patterns of *NRCAM, ABHD12, ENPP1, C12orf10, TEX264* and *REEP5* respectively. The blue bar displays the expression in normal tissue and the yellow bar shows the expression in tumour tissues. The expression level is represented as a digital expression unit (DEU) values. The expression levels in normal and cancerous breast tissue are indicates in a red box.

References

Abe, Ijichi, Ikeda, Kayano, Horie- Inoue, Takeda and Inoue (2012). "Forkhead box transcription factor, forkhead box A1, shows negative association with lymph node status in endometrial cancer, and represses cell proliferation and migration of endometrial cancer cells." Cancer science **103**(4): 806-812.

Aebersold and Mann (2003). "Mass spectrometry-based proteomics." Nature **422**(6928): 198-207.

Agarwal, Ramakant, Forgach, Rendón, Chaparro, Basurto and Margaritoni (2009). "Breast cancer care in developing countries." World journal of surgery **33**(10): 2069-2076.

Ahmed and Douek (2013). "The role of magnetic nanoparticles in the localization and treatment of breast cancer." BioMed research international **2013**.

AL-Dujaily, Al-Janabi, Pierscionek and Yasseen (2008). "High prevalence of *HER-2*/neu overexpression in female breast cancer among an Iraqi population exposed to depleted uranium." Journal of carcinogenesis **7**(1): 8.

Alam and Yadav (2013). "Potential Applications of Quantum Dots in Mapping Sentinel Lymph Node and Detection of Micrometastases in Breast Carcinoma." Journal of breast cancer **16**(1): 1-11.

Albergaria, Paredes, Sousa, Milanezi, Carneiro, Bastos, Costa, Vieira, Lopes and Lam (2009). "Expression of FOXA1 and GATA-3 in breast cancer: the prognostic significance in hormone receptor-negative tumours." Breast Cancer Res **11**(3): R40.

American Cancer Society (2014). "Cancer facts& Figure 2014" access date 7/1/2015
http://www.cancer.org/acs/groups/content/@research/documents/webcontent/acspc-042151.pdf.

Antoniou, Pharoah, Narod, Risch, Eyfjord, Hopper, Loman, Olsson, Johannsson and Borg (2003). "Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case series unselected for family history: a combined analysis of 22 studies." The American Journal of Human Genetics **72**(5): 1117-1130.

Arora and Scholar (2005). "Role of tyrosine kinase inhibitors in cancer therapy." Journal of Pharmacology and Experimental Therapeutics **315**(3): 971-979.

Arpino, Bardou, Clark and Elledge (2004). "Infiltrating lobular carcinoma of the breast: tumor characteristics and clinical outcome." Breast Cancer Res **6**(3): R149-R156.

Arps, Healy, Zhao, Kleer and Pang (2013). "Invasive ductal carcinoma with lobular features: a comparison study to invasive ductal and invasive lobular carcinomas of the breast." Breast cancer research and treatment **138**(3): 719-726.

Arpino, Generali, Sapino, Del Matro, Frassoldati, de Laurentis, Pronzato, Mustacchi, Cazzaniga, De Placido, Pierfranco Conte, Cappelletti, Vanessa Zanoni, Antonelli, Martinotti, Puglisi, Berruti Bottini, Dogliotti (2013). "Gene expression profiling in breast cancer: A clinical perspective." The Breast **22**(2): 109-120.

Atchley, Albarracin, Lopez, Valero, Amos, Gonzalez-Angulo, Hortobagyi and Arun (2008). "Clinical and pathologic characteristics of patients with BRCA-positive and BRCA-negative breast cancer." Journal of clinical oncology **26**(26): 4282-4288.

Baak, Janssen, Soreide and Heikkilae (2005). "Genomics and proteomics–the way forward." Annals of Oncology **16**(S2): 30-44.

Baptista, Pereira, Eaton, Doria, Miranda, Gomes, Quaresma and Franco (2008). "Gold nanoparticles for the development of clinical diagnosis methods." Analytical and bioanalytical chemistry **391**(3): 943-950.

Baskın and Yiğitbaşı (2010). "Clinical proteomics of breast cancer." Current genomics **11**(7): 528.

Bast, Ravdin, Hayes, Bates, Fritsche, Jessup, Kemeny, Locker, Mennel and Somerfield (2001). "2000 Update of Recommendations for the Use of Tumor Markers in Breast and Colorectal Cancer: Clinical Practice Guidelines of the American Society of Clinical Oncology*." Journal of clinical oncology 19(6): 1865-1878.

Bellei, Bergamini, Monari, Fantoni, Cuoghi, Ozben and Tomasi (2011). "High-abundance proteins depletion for serum proteomic analysis: concomitant removal of non-targeted proteins." Amino acids **40**(1): 145-156.

Berg (2009). "Tailored supplemental screening for breast cancer: what now and what next?" American Journal of Roentgenology **192**(2): 390-399.

Bishop (2010). "Applications of fluorescence in situ hybridization (FISH) in detecting genetic aberrations of medical significance." Bioscience Horizons: hzq009.

Brase, Wuttig, Kuner and Sultmann (2010). "Serum microRNAs as non-invasive biomarkers for cancer." Mol Cancer **9**(9): 306.

Bratu, Catrina and Marras (2011). Tiny molecular beacons for in vivo mRNA detection. RNA Detection and Visualization, Springer**:** 141-157.

Bravaccini, Granato, Medri, Foca, Falcini, Zoli, Ricci, Lanzanova, Masalu and Serra (2013). "Biofunctional characteristics of in situ and invasive breast carcinoma." Cellular Oncology **36**(4): 303-310.

Brodersen, Jorgensen and Gotzsche (2010). "The benefits and harms of screening for cancer with a focus on breast screening." Pol Arch Med Wewn **120**(3): 89-94.

Brunelli, Manfrin, Martignoni, Bersani, Remo, Reghellin, Chilosi and Bonetti (2008). "*HER-2*/neu Assessment in Breast Cancer Using the Original FDA and New ASCO/CAP Guideline Recommendations Impact on Selecting Patients for Herceptin Therapy." American journal of clinical pathology **129**(6): 907-911.

Bryce, Morrisont, Hoare, Muir and Keith (2000). "Mapping of the gene for the human telomerase reverse transcriptase, hTERT, to chromosome 5p15. 33 by fluorescence in situ hybridization." Neoplasia **2**(3): 197-201.

Bryers and Hitchman (2011). "Quantum Dots Brighten Biological Imaging." Progress in histochemistry and cytochemistry **45**(4): 201-237.

Buchholz (2009). "Radiation therapy for early-stage breast cancer after breast-conserving surgery." New England Journal of Medicine **360**(1): 63-70.

Byers, Di Vizio, O'Connell, Tholouli, Levenson, Gossard, Twomey, Yang, Benedettini and Rose (2007). "Semiautomated multiplexed quantum dot-based in situ hybridization and spectral deconvolution." The Journal of Molecular Diagnostics **9**(1): 20-29.

Caldarella, Crocetti, Bianchi, Vezzosi, Urso, Biancalani and Zappa (2011). "Female breast cancer status according to ER, PR and HER-2 expression: a population based analysis." Pathology & Oncology Research **17**(3): 753-758.

Callesen, Vach, Jørgensen, Cold, Mogensen, Kruse, Jensen and Madsen (2008). "Reproducibility of mass spectrometry based protein profiles for diagnosis of breast cancer across clinical studies: a systematic review." Journal of proteome research **7**(4): 1395-1402.

Cao, Huang, Wu, Lu, Liu, Shen, Shao and Di (2012). "Tumor characteristics and the clinical outcome of invasive lobular carcinoma compared to infiltrating ductal carcinoma in a Chinese population." World J SurgOncol **10**: 152.

Cancer Association of South Africa (2014). "Fact sheet on the top ten cancers per population group." Access date 1/4/2015. http://www.cansa.org.za/files/2014/11/Fact-Sheet-Top-Ten-Cancers-per-Population-Group-NCR-2008-Stats.pdf.

Cancer Association of South Africa (2015). "Position Statement On Breast Cancer." Access date 1/4/2015. http://www.cansa.org.za/files/2015/02/Position-Statement-Breast-Cancer-Feb-2015.pdf.

Carey and Viskochil (2007). "Status of the Human Malformation Map: 2007." American journal of medical genetics part A **143**(24): 2868-2885.

Castro and Eeles (2012). "The role of BRCA1 and BRCA2 in prostate cancer." Asian journal of andrology **14**(3): 409.

Chatterjee and Zetter (2005). "Cancer biomarkers: knowing the present and predicting the future."

Chen, Silver, Walpita, Cantor, Gazdar, Tomlinson, Couch, Weber, Ashley and Livingston (1998). "Stable interaction between the products of the BRCA1 and BRCA2 tumor suppressor genes in mitotic and meiotic cells." Molecular cell **2**(3): 317-328.

Ciccimaro and Blair (2010). "Stable-isotope dilution LC-MS for quantitative biomarker analysis." Bioanalysis **2**(2): 311-341.

Collins, Liu, Schootman, Aft, Yan, Dean, Eilers and Jeffe (2011). "Effects of breast cancer surgery and surgical side effects on body image over time." Breast cancer research and treatment **126**(1): 167-176.

Connor and Attai (2013). "Adjuvant endocrine therapy for the surgeon: options, side effects, and their management." Annals of surgical oncology **20**(10): 3188-3193.

Coopey, Tang, Lei, Freer, Kansal, Colwell, Gadd, Specht, Austen Jr and Smith (2013). "Increasing eligibility for nipple-sparing mastectomy." Annals of surgical oncology **20**(10): 3218-3222.

Croce (2008). "Oncogenes and cancer." New England Journal of Medicine **358**(5): 502-511.

Davar, Beumer, Hamieh and Tawbi (2012). "Role of PARP inhibitors in cancer biology and therapy." Current medicinal chemistry **19**(23): 3907.

Davis and Hanash (2006). "High-throughput genomic technology in research and clinical management of breast cancer. Plasma-based proteomics in early detection and therapy." Breast Cancer Research **8**(6): 217.

de Bruin, Kwong, Goldstein, Lipson, Ikeda, McPherson, Sharma, Kardashian, Schackmann and Kingham (2012). "Breast cancer risk factors differ between Asian and white women with BRCA1/2 mutations." Familial cancer **11**(3): 429-439.

Del Bianco, Zavagno, Burelli, Scalco, Barutta, Carraro, Pietrarota, Meneghini, Morbin and Tacchetti (2008). "Morbidity comparison of sentinel lymph node biopsy versus conventional axillary lymph node dissection for breast cancer patients: results of the Sentinella–GIVOM Italian randomised clinical trial." European Journal of Surgical Oncology (EJSO) **34**(5): 508-513.

Dellapasqua, Colleoni, Gelber and Goldhirsch (2005). "Adjuvant endocrine therapy for premenopausal women with early breast cancer." Journal of clinical oncology **23**(8): 1736-1750.

Deng and Scott (2000). "Role of the tumor suppressor gene Brca1 in genetic stability and mammary gland tumor formation." Oncogene **19**(8): 1059-1064.

Derveaux, Vandesompele and Hellemans (2010). "How to do successful gene expression analysis using real-time PCR." Methods **50**(4): 227-230.

Dhanasekaran, Doherty, Kenneth and Group (2010). "Comparison of different standards for real-time PCR-based absolute quantification." Journal of immunological methods **354**(1): 34-39.

Diamandis (2012). "The failure of protein cancer biomarkers to reach the clinic: why, and what can be done to address the problem?" BMC medicine **10**(1): 87.

Eccles (2001). "The role of c-erbB-2/HER-2/neu in breast cancer progression and metastasis." Journal of mammary gland biology and neoplasia **6**(4): 393-406.

Ellis, Dyson, Stephenson and Maltby (2005). "HER-2 amplification status in breast cancer: a comparison between immunohistochemical staining and fluorescence in situ hybridisation using manual and automated quantitative image analysis scoring techniques." Journal of clinical pathology **58**(7): 710-714.

Enache, Georgescu and Pătrană (2012). "Negative estrogen-receptor invasive breast carcinoma: mammographic aspects, correlations with HER-2/neu oncoprotein status." Rom J Morphol Embryol **53**(3 Suppl): 755-762.

Espina and Liotta (2011). "What is the malignant nature of human ductal carcinoma in situ?" Nature Reviews Cancer **11**(1): 68-75.

Esseghir Selma, Todd S. Katrina, Hunt Toby, Poulsom Richard, Plaza-Menacho Ivan, S. Reis-Filho Jorge and M. Isacke Clare (2007). "A Role for Glial Cell–Derived Neurotrophic Factor–Induced Expression by Inflammatory Cytokines and RET/GFRA1 Receptor Up-regulation in Breast Cancer." Cancer research **67**(24): 11732-11741.

Evans (2012). "Breast cancer screening: successes and challenges." CA: a cancer journal for clinicians **62**(1): 5-9.

Everley, Krijgsveld, Zetter and Gygi (2004). "Quantitative cancer proteomics: stable isotope labeling with amino acids in cell culture (SILAC) as a tool for prostate cancer research." Molecular & Cellular Proteomics **3**(7): 729-735.

Faheem, Mahmood, Khurram, Qasim and Irfan (2012). "Estrogen receptor, progesterone receptor, and Her 2 Neu positivity and its association with tumour characteristics and menopausal status in a breast cancer cohort from northern Pakistan." Ecancermedicalscience **6**: 283.

Fang, Chen and Weigel (2009). "GATA-3 as a marker of hormone response in breast cancer." Journal of Surgical Research **157**(2): 290-295.

Fang, Peng, Pang and Li (2012). "Quantum dots for cancer research: current status, remaining issues, and future perspectives." Cancer Biology & Medicine **9**(3): 151.

Fang Xiaohong, Liu Xiaojing, Schuster Sheldon and Tan Weihong (1999). "Designing a Novel Molecular Beacon for Surface-Immobilized DNA Hybridization Studies." Journal of the American Chemical Society **121**(12): 2921-2922.

Fehm, Hoffmann, Aktas, Becker, Solomayer, Wallwiener, Kimmig and Kasimir-Bauer (2009). "Detection and characterization of circulating tumor cells in blood of primary breast cancer patients by RT-PCR and comparison to status of bone marrow disseminated cells." Breast Cancer Res **11**(4): R59.

Ferlay, Autier, Boniol, Heanue, Colombet and Boyle (2007). "Estimates of the cancer incidence and mortality in Europe in 2006." Annals of oncology **18**(3): 581-592.

Fernandez, Robertson, Pei, Aburto-Chumpitaz, Mu, Chu, Alpaugh, Huang, Cao and Ye (2013). "Inflammatory breast cancer (IBC): clues for targeted therapies." Breast cancer research and treatment **140**(1): 23-33.

Fleige and Pfaffl (2006). "RNA integrity and the effect on the real-time qRT-PCR performance." Molecular aspects of medicine **27**(2): 126-139.

Fontes, Bezerra de Lira, Barreto Lopes Seabra, Gomes da Silva, Antônio Gomes and Santos (2012). "Quantum Dots in Biomedical Research." Biomedical Engineering-Technical Applications in Medicine. Access date 5/5/2015 http://www.intechopen.com/books/biomedical-engineering-technical-applications-in-medicine/quantum-dots-in-biomedical-research.

Friedman and Kaestner (2006). "The Foxa family of transcription factors in development and metabolism." Cellular and Molecular Life Sciences CMLS **63**(19-20): 2317-2328.

Funahashi, Hitoshi; Takeyama, Hiromitsu; Sawai, Hirozumi; Furuta, Akiyoshi;

Gampenrieder, Rinnerthaler and Greil (2013). "Neoadjuvant chemotherapy and targeted therapy in breast cancer: past, present, and future." Journal of oncology **2013**.

Garrison, Lalla, Brammer, Babigumira, Wang and Perez (2013). "Assessing the potential cost- effectiveness of retesting IHC0, IHC1+, or FISH- negative early stage breast cancer patients for HER-2 status." Cancer **119**(17): 3113-3122.

Gast, Schellens and Beijnen (2009). "Clinical proteomics in breast cancer: a review." Breast cancer research and treatment **116**(1): 17-29.

Gauger and Schneider (2014). "Tumour supressor secreted frizzled related protein 1 regulates p53- mediated apoptosis." Cell biology international **38**(1): 124-130.

Gerhardus, Schleberger, Schlegelberger and Gadzicki (2007). "Diagnostic accuracy of methods for the detection of BRCA1 and BRCA2 mutations: a systematic review." European Journal of Human Genetics **15**(6): 619-627.

Gion, Mione, Leon and Dittadi (1999). "Comparison of the diagnostic accuracy of CA27. 29 and CA15. 3 in primary breast cancer." Clinical chemistry **45**(5): 630-637.

Gonzaga (2010). "How accurate is ultrasound in evaluating palpable breast masses?" Pan African Medical Journal **7**(1).

Goodheart, Rose, Hattermann- Zogg, Smith, De Young and Buller (2009). "BRCA2 alteration is important in clear cell carcinoma of the ovary." Clinical genetics **76**(2): 161-167.

Grebe and Singh (2011). "LC-MS/MS in the Clinical Laboratory–Where to From Here?" The Clinical Biochemist Reviews **32**(1): 5.

Greenman, Stephens, Smith, Dalgliesh, Hunter, Bignell, Davies, Teague, Butler and Stevens (2007). "Patterns of somatic mutation in human cancer genomes." Nature **446**(7132): 153-158.

Grønborg, Kristiansen, Iwahori, Chang, Reddy, Sato, Molina, Jensen, Hruban and Goggins (2006). "Biomarker discovery from pancreatic cancer secretome using a differential proteomic approach." Molecular & Cellular Proteomics **5**(1): 157-171.

Groudine and Casimir (1985). "Tyrosine kinase receptor with extensive homology to EGF receptor shares chromosomal location with neu oncogene."

Grünewald, Haun, Urbanek, Fiegl, Müller-Holzner, Gunsilius, Dünser, Marth and Gastl (2000). "Mammaglobin gene expression: a superior marker of breast cancer cells in peripheral blood in comparison to epidermal-growth-factor receptor and cytokeratin-19." Laboratory investigation **80**(7): 1071-1077.

Grünewald Kurt, Haun Margot, Urbanek Martina, Fiegl Michael, Müller-Holzner Elisabeth, Gunsilius Eberhard, Dünser Martina, Marth Christain, Gastl Günther (2000). " Mammaglobin gene expression: a superior marker of breast cancer cells in peripheral blood in comparison to epidermal-growth-factor receptor and cytokeratin-19."Laboratory investigation **80**(7): 1071-1077.

Guénin, Mauriat, Pelloux, Van Wuytswinkel, Bellini and Gutierrez (2009). "Normalization of qRT-PCR data: the necessity of adopting a systematic, experimental conditions-specific, validation of references." Journal of experimental botany **60**(2): 487-493.

Habashy, Powe, Rakha, Ball, Paish, Gee, Nicholson and Ellis (2008). "Forkhead-box A1 (FOXA1) expression in breast cancer and its prognostic significance." European journal of cancer **44**(11): 1541-1551.

Hall, Lee, Newman, Morrow, Anderson, Huey and King (1990). "Linkage of early-onset familial breast cancer to chromosome 17q21." Science **250**(4988): 1684-1689.

Hamada, Nomura, Kamikawa, Yamada, Batra, Yonezawa and Sugihara (2012). "DF3 epitope expression on MUC1 mucin is associated with tumor aggressiveness, subsequent lymph node metastasis, and poor prognosis in patients with oral squamous cell carcinoma." Cancer **118**(21): 5251-5264.

Hammer, Fanning and Crowe (2008). "Overview of breast cancer staging and surgical treatment options." Cleve land and Clinic Journal of medicine **75 Suppl 1**: S10-16.

Hanahan and Weinberg (2000). "The hallmarks of cancer." cell **100**(1): 57-70.

Hanash, Pitteri and Faca (2008). "Mining the plasma proteome for cancer biomarkers." Nature **452**(7187): 571-579.

Haneef, Shaharyar, Husain, Rashid, Mishra, Parveen, Ahmed, Pal and Kumar (2013). "Application of LC–MS/MS for quantitative analysis of glucocorticoids and stimulants in biological fluids." Journal of Pharmaceutical Analysis **3**(5): 341-348.

Harris and Kinsinger (2002). "Routinely teaching breast self-examination is dead. What does this mean?" Journal of the National Cancer Institute **94**(19): 1420-1421.

Hickey, Peate, Saunders and Friedlander (2009). "Breast cancer in young women and its impact on reproductive function." Human Reproduction Update **15**(3): 323-339.

Hisamatsu, Tokunaga, Yamashita, Akiyoshi, Okada, Nakashima, Aishima, Morita, Kakeji and Maehara (2012). "Impact of FOXA1 expression on the

prognosis of patients with hormone receptor-positive breast cancer." Annals of surgical oncology **19**(4): 1145-1152.

Hulpiau and Van Roy (2009). "Molecular evolution of the cadherin superfamily." The international journal of biochemistry & cell biology **41**(2): 349-369.

Ishihama, Sato, Tabata, Miyamoto, Sagane, Nagasu and Oda (2005). "Quantitative mouse brain proteomics using culture-derived isotope tags as internal standards." Nature biotechnology **23**(5): 617-621.

Jain, Mehta, Nakshatri, Idrees and Badve (2011). "High- level expression of forkhead- box protein A1 in metastatic prostate cancer." Histopathology **58**(5): 766-772.

Jain, Nijhawan, Tyagi and Khurana (2006). "Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR." Biochemical and biophysical research communication **345** (2): 646-651.

Jamieson, Bakhshi, Petrova, Pocock, Imani and Seifalian (2007). "Biological applications of quantum dots." Biomaterials **28**(31): 4717-4732.

Jemal, Siegel, Ward, Murray, Xu and Thun (2007). "Cancer statistics, 2007." CA: a cancer journal for clinicians **57**(1): 43-66.

Jimenez- Andrade, Mantyh, Bloom, Ferng, Geffre and Mantyh (2010). "Bone cancer pain." Annals of the New York Academy of Sciences **1198**(1): 173-181.

Joensuu, Leidenius, Kero, Andersson, Horwitz and Heikkilä (2013). "eR, pR, HER-2, Ki-67 and cK5 in early and Late Relapsing Breast cancer—Reduced cK5 expression in Metastases." Breast cancer: basic and clinical research **7**: 23.

Johnson, Gorringe, Thompson, Opeskin, Boyle, Wang, Hill, Mann and Campbell (2012). "Identification of copy number alterations associated with the progression of DCIS to invasive ductal carcinoma." Breast cancer research and treatment **133**(3): 889-898.

Jonjić, Mustać, Bekafigo, Molek, Lovasić, Lukanović, Ivanković, Girotto, Lovasić and Petković (2012). "[Analysis of sentinel lymph node in early breast cancer patient-Clinical Hospital Center Rijeka]." Lijecnicki vjesnik **135**(7-8): 201-205.

Kaestner, Knöchel and Martínez (2000). "Unified nomenclature for the winged helix/forkhead transcription factors." Genes & development **14**(2): 142-146.

Kang (2012). "Principles and Applications of LC-MS/MS for the Quantitative Bioanalysis of Analytes in Various Biological Samples." Tandem Mass Spectrometry–Applications and Principles: 441-492.

Karakach, Flight, Douglas and Wentzell (2010). "An introduction to dna microarrays for gene expression analysis." chemometrics and Intelligent laboratory system **104** (1): 28-52.

Kaufmann and Knöchel (1996). "Five years on the wings of fork head." Mechanisms of development **57**(1): 3-20.

Kawasaki (2006). "The end of the microarray Tower of Babel: will universal standards lead the way?" Journal of biomolecular techniques: JBT **17**(3): 200.

Kelly, Dean, Comulada and Lee (2010). "Breast cancer detection using automated whole breast ultrasound and mammography in radiographically dense breasts." European radiology **20**(3): 734-742.

Keune, Jeffe, Schootman, Hoffman, Gillanders and Aft (2010). "Accuracy of ultrasonography and mammography in predicting pathologic response after neoadjuvant chemotherapy for breast cancer." The American Journal of Surgery **199**(4): 477-484.

Kim, Sohn and Tan (2008). "Molecular Beacons in Biomedical Detection and Clinical Diagnosis." International journal of clinical and experimental pathology**1**(2): 105-116.

Kinsella, Nassar, Siddiqui and Cohen (2012). "Estrogen receptor (ER), progesterone receptor (PR), and HER-2 expression pre-and post-neoadjuvant chemotherapy in primary breast carcinoma: a single institutional experience." International journal of clinical and experimental pathology **5**(6): 530.

Kohnke, Mulligan and Christopherson (2009). "Membrane proteomics for leukemia classification and drug target identification." Current opinion in molecular therapeutics **11**(6): 603-610.

Königsberg, Obermayr, Bises, Pfeiler, Gneist, Wrba, de Santis, Zeillinger, Hudec and Dittrich (2011). "Detection of EpCAM positive and negative circulating tumor cells in metastatic breast cancer patients." Acta oncologica **50**(5): 700-710.

Kotsopoulos, Lubinski, Salmena, Lynch, Kim-Sing, Foulkes, Ghadirian, Neuhausen, Demsky and Tung (2012). "Breastfeeding and the risk of breast cancer in BRCA1 and BRCA2 mutation carriers." Breast Cancer Res **14**(2): R42.

Krag, Anderson, Julian, Brown, Harlow, Costantino, Ashikaga, Weaver, Mamounas, Jalovec, Frazier, Noyes, Robidoux, Scarth and Wolmark (2010). "Sentinel-lymph-node resection compared with conventional axillary-lymph-node dissection in clinically node-negative patients with breast cancer: overall survival findings from the NSABP B-32 randomised phase 3 trial." Lancet Oncology **11**(10): 927-933.

Kuhl, Kuhn, Braun and Schild (2007). "Pre-operative staging of breast cancer with breast MRI: one step forward, two steps back?" The breast **16**: 34-44.

Kwong, Wong, Wong, Law, Ng, Tang, Chan, Suen, Choi and Ho (2009). "Clinical and pathological characteristics of Chinese patients with BRCA related breast cancer." The HUGO journal **3**(1-4): 63-76.

Laemmli (1970). "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." Nature **227**(5259): 680-685.

Lecke, Morsch and Spritzer (2013). "Circulating levels and subcutaneous adipose tissue gene expression of pigment epithelium-derived factor in polycystic ovary syndrome and normal women: a case control study." Reproductive Biology and Endocrinology **11**(1): 77.

Lee and Kohn (2010). "Proteomics as a guiding tool for more effective personalized therapy." Annals of oncology **21**(suppl 7): vii205-vii210.

Lee, Park, Yim, Lee, Park and Lee (2007). "Expression of c-erbB2, cyclin D1 and estrogen receptor and their clinical implications in the invasive ductal carcinoma of the breast." Japanese journal of clinical oncology **37**(9): 708-714.

Li Jianwei Jeffery, Chu Yizhuo, Lee Benjamin Yi-Hung and Xie Xiaoliang Sunney (2008). "Enzymatic signal amplification of molecular beacons for sensitive DNA detection." Nucleic acid research **36**(6): e36.

Li, Yan, Zhou, Dang, Wang and Zhang (2011). "Effects of music therapy on pain among female breast cancer patients after radical mastectomy: results from a randomized controlled trial." Breast cancer research and treatment **128**(2): 411-419.

Li, Zhu, Chen, Dai and Zou (2011). "TGF-β1 and BRCA2 expression are associated with clinical factors in breast cancer." Cell biochemistry and biophysics **60**(3): 245-248.

Liang, Zhao, Hajivandi, Wu, Tao, Amshey and Pope (2006). "Quantification of membrane and membrane-bound proteins in normal and malignant breast cancer cells isolated from the same patient with primary breast carcinoma." Journal of proteome research **5**(10): 2632-2641.

Liberman (2004). "Breast cancer screening with MRI-what are the data for patients at high risk?" New England Journal of Medicine **351**: 497-497.

Liu, Dekker, Stingl, Güzel, De Marchi, Martens, Foekens, Luider and Umar (2013). "Quantitative proteomic analysis of microdissected breast cancer tissues: comparison of label-free and SILAC-based quantification with shotgun, directed, and targeted MS approaches." Journal of proteome research **12**(10): 4627-4641.

Liu, Peng, Chen, Yang, Hu, Xia, Liu, Pang and Li (2011). "Quantum dots-based double-color imaging of HER-2 positive breast cancer invasion." Biochemical and biophysical research communications **409**(3): 577-582.

Loredo-Pozos, Chiquete, Oceguera-Villanueva, Panduro, Siller-López and Ramos-Márquez (2009). "Expression profile of BRCA1 and BRCA2 genes in premenopausal Mexican women with breast cancer: clinical and immunohistochemical correlates." Medical Oncology **26**(3): 269-275.

Lossos, Czerwinski, Alizadeh, Wechser, Tibshirani, Botstein and Levy (2004). "Prediction of survival in diffuse large-B-cell lymphoma based on the expression of six genes." New England Journal of Medicine **350**(18): 1828-1837.

Loughran and Keeling (2010). "Seeding of tumour cells following breast biopsy: a literature review." Breast Cancer Research **12**(Suppl 3): P43.

Loukas, Tubbs, Mirzayan, Shirak, Steinberg and Shoja (2011). "The history of mastectomy." The American Surgeon **77**(5): 566-571.

Lu, Mishra, Zhu, Meltzer and Cheng (2011). "Global expression profiling reveals gain-of-function onco- genic activity of a mutated thyroid hormone receptor in thyroid carcinogenesis." **1**(2): 168.

Luparia, Mariscotti, Durando, Ciatto, Bosco, Campanino, Castellano, Sapino and Gandini (2013). "Accuracy of tumour size assessment in the preoperative staging of breast cancer: comparison of digital mammography, tomosynthesis, ultrasound and MRI." La radiologia medica **118**(7): 1119-1136.

Ma, Han, Zhang, Cui and Wang (2013). "[Differences and clinical significance of receptor expression between primary and locally recurrent breast tumor tissues]." Zhonghua zhong liu za zhi [Chinese journal of oncology] **35**(5): 372-376.

Mader (2011). “Inquriy into life.” 13th edition. McGraw-Hill Higher Education, pp125-141.

Madjar (2010). "Role of breast ultrasound for the detection and differentiation of breast lesions." Breast Care **5**(2): 109.

Mahmoud, Paish, Powe, Macmillan, Grainge, Lee, Ellis and Green (2011). "Tumor-infiltrating CD8+ lymphocytes predict clinical outcome in breast cancer." Journal of clinical oncology **29**(15): 1949-1955.

Majewski and Bernards (2011). "Taming the dragon: genomic biomarkers to individualize the treatment of cancer." Nature medicine: 304-312.

Malhotra, Zhao, Band and Band (2010). "Histological, molecular and functional subtypes of breast cancers." Cancer biology & therapy **10**(10): 955-960.

Malone, Daling, Doody, Hsu, Bernstein, Coates, Marchbanks, Simon, McDonald and Norman (2006). "Prevalence and predictors of BRCA1 and BRCA2 mutations in a population-based study of breast cancer in white and black American women ages 35 to 64 years." Cancer Research **66**(16): 8297-8308.

Mansel, Fallowfield, Kissin, Goyal, Newcombe, Dixon, Yiangou, Horgan, Bundred and Monypenny (2006). "Randomized multicenter trial of sentinel node biopsy versus standard axillary treatment in operable breast cancer: the ALMANAC Trial." Journal of the National Cancer Institute **98**(9): 599-609.

Marimuthu, Subbannayya, Sahasrabuddhe, Balakrishnan, Syed, Sekhar, Katte, Pinto, Srikanth and Kumar (2013). "SILAC- based quantitative proteomic analysis of gastric cancer secretome." PROTEOMICS-Clinical Applications **7**(5-6): 355-366.

Markovsky, Baabur-Cohen, Eldar-Boock, Omer, Tiram, Ferber, Ofek, Polyak, Scomparin and Satchi-Fainaro (2012). "Administration, distribution, metabolism and elimination of polymer therapeutics." Journal of Controlled Release **161**(2): 446-460.

Marras, Salvatore AE (2006). "Selection of Fluorophore and Quencher Pairs for Fluorescent Nucleic Acid Hybridization Probes." Fluoresecent Energy Transfer Nucleic Acid Probes: 3-16

Martín, Hernández, Avendaño, Rodríguez and Martínez (2011). "Manual lymphatic drainage therapy in patients with breast cancer related lymphoedema." BMC cancer **11**(1): 94.

Mayeux (2004). "Biomarkers: potential uses and limitations." NeuroRx **1**(2): 182-188.

McCahill, Single, Bowles, Feigelson, James, Barney, Engel and Onitilo (2012). "Variability in reexcision following breast conservation surgery." Jama **307**(5): 467-475.

Medley (2007). Method and Material Development for the Detection and Analysis of Cancer Cells, University of Florida.

Mehta, Jain, Leung, Choo, Nielsen, Huntsman, Nakshatri and Badve (2012). "FOXA1 is an independent prognostic marker for ER-positive breast cancer." Breast cancer research and treatment **131**(3): 881-890.

Mehta, Shelling, Muthukaruppan, Lasham, Blenkiron and Laking (2010). "Predictive and prognostic molecular markers for cancer medicine." Therapeutic advances in medical oncology **2**(2): 125-148.

Millioni, Tolin, Puricelli, Sbrignadello, Fadini, Tessari and Arrigoni (2011). "High abundance proteins depletion vs low abundance proteins enrichment:

comparison of methods to reduce the plasma proteome complexity." PloS one **6**(5): e19603.

Moelans, de Weger, van Blokland, Ezendam, Elshof, Tilanus and van Diest (2009). "HER-2/neu amplification testing in breast cancer by multiplex ligation-dependent probe amplification in comparison with immunohistochemistry and in situ hybridization." Analytical Cellular Pathology **31**(1): 1-10.

Mohamed, Krajewski, Cakar and Ma (2013). "Targeted therapy for breast cancer." The American journal of pathology **183**(4): 1096-1112.

Molina, Escudero, Augé, Filella, Foj, Torné, Lejarcegui and Pahisa (2011). "HE4 a novel tumour marker for ovarian cancer: comparison with CA 125 and ROMA algorithm in patients with gynaecological diseases." Tumor Biology **32**(6): 1087-1095.

Munday, Surtees, Emmott, Dove, Digard, Barr, Whitehouse, Matthews and Hiscox (2012). "Using SILAC and quantitative proteomics to investigate the interactions between viral and host proteomes." Proteomics **12**(4- 5): 666-672.

Mungan, Vriesema, Thomas, Kiemeney and Witjes (2000). "Urinary bladder cancer test: a new urinary tumor marker in the follow-up of superficial bladder cancer." Urology **56**(5): 787-791.

Muthusamy, Bosenberg and Wajapeyee (2010). "Redefining regulation of DNA methylation by RNA interference." Genomics **96**(4): 191-198.

Nahta, Yu, Hung, Hortobagyi and Esteva (2006). "Mechanisms of disease: understanding resistance to HER-2-targeted therapy in human breast cancer." Nature clinical practice Oncology **3**(5): 269-280.

Nair, Jaleel, Asmann, Short and Raghavakaimal (2004). "Proteomic research: potential opportunities for clinical and physiological investigators." American Journal of Physiology-Endocrinology and Metabolism **286**(6): E863-E874.

Nass and Dickson (1997). "Defining a role for c-Myc in breast tumorigenesis." Breast cancer research and treatment **44**(1): 1-22.

Nassar (2011). "Core needle biopsy versus fine needle aspiration biopsy in breast—a historical perspective and opportunities in the modern era." Diagnostic cytopathology **39**(5): 380-388.

Nelson, Tyne, Naik, Bougatsos, Chan and Humphrey (2009). "Screening for breast cancer: an update for the US Preventive Services Task Force." Annals of internal medicine **151**(10): 727-737.

Network (2012). "Comprehensive molecular portraits of human breast tumours." Nature **490**(7418): 61-70.

Nie, Xing, Kim and Simons (2007). "Nanotechnology applications in cancer." Annu. Rev. Biomed. Eng. **9**: 257-288.

Niikura, Kimura, Iwamoto, Hayashi, Shintoku, Saito, Suzuki and Tokuda (2013). "Women prefer adjuvant endocrine therapy to chemotherapy for breast cancer treatment." Breast Cancer **20**(1): 67-74.

Nijhawan, Votava and Mariwalla (2012). "Clinical application and limitations of the fluorescence in situ hybridization (FISH) assay in the diagnosis and management of melanocytic lesions: a report of 3 cases." Cutis **90**(4): 189-195.

Nothacker, Duda, Hahn, Warm, Degenhardt, Madjar, Weinbrenner and Albert (2009). "Early detection of breast cancer: benefits and risks of supplemental breast ultrasound in asymptomatic women with mammographically dense breast tissue. A systematic review." BMC cancer **9**(1): 335.

Nucera, Eeckhoute, Finn, Carroll, Ligon, Priolo, Fadda, Toner, Sheils and Attard (2009). "FOXA1 is a potential oncogene in anaplastic thyroid carcinoma." Clinical Cancer Research **15**(11): 3680-3689.

Ong, Blagoev, Kratchmarova, Kristensen, Steen, Pandey and Mann (2002). "Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics." Molecular & Cellular Proteomics **1**(5): 376-386.

Ong, Foster and Mann (2003). "Mass spectrometric-based approaches in quantitative proteomics." Methods **29**(2): 124-130.

Onitilo, Engel, Greenlee and Mukesh (2009). "Breast cancer subtypes based on ER/PR and HER-2 expression: comparison of clinicopathologic features and survival." Clinical medicine & research **7**(1-2): 4-13.

Osman, Saleh, Jackson, Corrigan and Cil (2013). "Increased postoperative complications in bilateral mastectomy patients compared to unilateral mastectomy: an analysis of the NSQIP database." Annals of surgical oncology **20**(10): 3212-3217.

Paredes, Figueiredo, Albergaria, Oliveira, Carvalho, Ribeiro, Caldeira, Costa, Simões-Correia and Oliveira (2012). "Epithelial E-and P-cadherins: role and clinical significance in cancer." Biochimica et Biophysica Acta (BBA)-Reviews on Cancer **1826**(2): 297-311.

Paredes, Stove, Stove, Milanezi, Van Marck, Derycke, Mareel, Bracke and Schmitt (2004). "P-cadherin is up-regulated by the antiestrogen ICI 182,780 and promotes invasion of human breast cancer cells." Cancer Research **64**(22): 8309-8317.

Parveen, Misra and Sahoo (2012). "Nanoparticles: a boon to drug delivery, therapeutics, diagnostics and imaging." Nanomedicine: Nanotechnology, Biology and Medicine **8**(2): 147-166.

Pataky, Armstrong, Chia, Coldman, Kim-Sing, McGillivray, Scott, Wilson and Peacock (2013). "Cost-effectiveness of MRI for breast cancer screening in BRCA1/2 mutation carriers." BMC cancer **13**(1): 339.

Patel, Madan, Getsios, Bertrand and MacCalman (2003). "Cadherin switching in ovarian cancer progression." International journal of cancer **106**(2): 172-177.

Pathak, Cao, Davidson, Jin and Silva (2006). "Quantum dot applications to neuroscience: new tools for probing neurons and glia." The Journal of neuroscience **26**(7): 1893-1895.

Pătrană, Georgescu, Fota, Enache and Pirici (2012). "Invasive mammary carcinoma: assessment of HER-2/neu status by immunohistochemistry." Romanian journal of morphology and embryology= Revue roumaine de morphologie et embryologie **53**(3 Suppl): 781.

Peng and Li (2010). "Application of quantum dots-based biotechnology in cancer diagnosis: current status and future perspectives." Journal of Nanomaterials **2010**: 5.

Perche and Torchilin (2013). "Recent trends in multifunctional liposomal nanocarriers for enhanced tumor targeting." Journal of drug delivery **2013**.

Petro-Nustas, Tsangari, Phellas and Constantinou (2013). "Health beliefs and practice of breast self-examination among young Cypriot women." Journal of Transcultural Nursing **24**(2): 180-188.

Pfaffl (2001). "A new mathematical model for relative quantification in real-time RT–PCR." Nucleic Acids Research **29**(9): e45-e45.

Pfaffl (2004). "Quantification strategies in real-time PCR." AZ of quantitative PCR **1**: 89-113.

Pietras and Östman (2010). "Hallmarks of cancer: interactions with the tumor stroma." Experimental cell research **316**(8): 1324-1331.

Pisanò, Mezzolla, Galante, Alemanno, Manca, Lorusso, Malvasi and Tinelli (2011). "A new mutation of BRCA2 gene in an Italian healthy woman with familial breast cancer history." Familial cancer **10**(1): 65-71.

Polyak (2007). "Breast cancer: origins and evolution." The Journal of clinical investigation **117**(117 (11)): 3155-3163.

Prat and Perou (2011). "Deconstructing the molecular portraits of breast cancer." Molecular oncology **5**(1): 5-23.

Press, Slamon, Flom, Park, Zhou and Bernstein (2002). "Evaluation of HER-2/neu gene amplification and overexpression: comparison of frequently used assay methods in a molecularly characterized cohort of breast cancer specimens." Journal of clinical oncology **20**(14): 3095-3105.

Pulverer, Weinhaeusel, Noehammer and Vierlinger (2012). "Principles and Application of Microarray Technology in Thyroid Cancer Research." Intech open access publisher: 119-146.

Qoronfleh, Benton, Ignacio and Kaboord (2003). "Selective enrichment of membrane proteins by partition phase separation for proteomic studies." BioMed research international **2003**(4): 249-255.

Radojicic, Zaravinos, Vrekoussis, Kafousi, Spandidos and Stathopoulos (2011). "MicroRNA expression analysis in triple-negative (ER, PR and HER-2/neu) breast cancer." Cell Cycle **10**(3): 507-517.

Riley, Sauerbrei and Altman (2009). "Prognostic markers in cancer: the evolution of evidence from single studies to meta-analysis, and beyond." British journal of cancer **100**(8): 1219-1229.

Robinson, Holmes and Carroll (2013). "FOXA1 mutations in hormone-dependent cancers." Frontiers in oncology **3**.

Romond, Perez, Bryant, Suman, Geyer Jr, Davidson, Tan-Chiu, Martino, Paik and Kaufman (2005). "Trastuzumab plus adjuvant chemotherapy for operable HER-2-positive breast cancer." New England Journal of Medicine **353**(16): 1673-1684.

Rosenthal, Chang, Kovtun, McBride and Tomlinson (2011). "Biocompatible quantum dots for biological applications." Chemistry & biology **18**(1): 10-24.

Rosenwald, Wright, Chan, Connors, Campo, Fisher, Gascoyne, Muller-Hermelink, Smeland and Giltnane (2002). "The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma." New England Journal of Medicine **346**(25): 1937-1947.

Roşian, Lazăr and Dema (2005). "C-erbB2 oncoprotein: prognostic marker in breast cancer." Rom J Morphol Embryol **46**: 99-104.

Ross, Fletcher, Bloom, Linette, Stec, Symmans, Pusztai and Hortobagyi (2004). "Targeted therapy in breast cancer the HER-2/neu gene and protein." Molecular & Cellular Proteomics **3**(4): 379-398.

Ross, Slodkowska, Symmans, Pusztai, Ravdin and Hortobagyi (2009). "The HER-2 receptor and breast cancer: ten years of targeted anti–HER-2 therapy and personalized medicine." The oncologist **14**(4): 320-368.

Saadeh, Leung, Vyas, Chaturvedi, Perumal and Vyas (2014). "Applications of nanomedicine in breast cancer detection, imaging, and therapy." Journal of nanoscience and nanotechnology **14**(1): 913-923.

Sabel, Rogers, Griffith, Jagsi, Kleer, Diehl, Breslin, Cimmino, Chang and Newman (2009). "Residual disease after re- excision lumpectomy for close margins." Journal of surgical oncology **99**(2): 99-103.

Sapino, Goia, Recupero and Marchiò (2013). "Current challenges for HER-2 testing in diagnostic pathology: state of the art and controversial issues." Frontiers in oncology **3**.

Saslow, Boetes, Burke, Harms, Leach, Lehman, Morris, Pisano, Schnall and Sener (2007). "American Cancer Society guidelines for breast screening with MRI as an adjunct to mammography." CA: a cancer journal for clinicians **57**(2): 75-89.

Schaub, Jones, Nyalwidhe, Cazares, Karbassi, Semmes, Feliberti, Perry and Drake (2009). "Serum proteomic biomarker discovery reflective of stage and obesity in breast cancer patients." Journal of the American College of Surgeons **208**(5): 970-978.

Schlacher, Christ, Siaud, Egashira, Wu and Jasin (2011). "Double-strand break repair-independent role for BRCA2 in blocking stalled replication fork degradation by MRE11." cell **145**(4): 529-542.

Schmittgen and Livak (2008). "Analyzing real-time PCR data by the comparative CT method." Nature Protocols **3**(6): 1101-1108.

Schoppmann, Tamandl, Roberts, Jomrich, Schoppmann, Zwrtek, Dubsky, Gnant, Jakesz and Birner (2010). "HER-2/neu expression correlates with vascular endothelial growth factor-C and lymphangiogenesis in lymph node-positive breast cancer." Annals of oncology **21**(5): 955-960.

Schroeder, Mueller, Stocker, Salowsky, Leiber, Gassmann, Lightfoot, Menzel, Granzow and Ragg (2006). "The RIN: an RNA integrity number for assigning integrity values to RNA measurements." BMC molecular biology **7**(1): 3.

Sharan and Bradley (1997). "MurineBrca2: Sequence, Map Position, and Expression Pattern." Genomics **40**(2): 234-241.

Shen, Bassolino and Stouch (1997). "Membrane Structure and Function of the plasma membrane." Bristol-myers squibb research institute, biophysical journal **73**: 120-18.

Sherr (2004). "Principles of tumor suppression." cell **116**(2): 235-246.

Shipp, Ross, Tamayo, Weng, Kutok, Aguiar, Gaasenbeek, Angelo, Reich and Pinkus (2002). "Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning." Nature medicine **8**(1): 68-74.

Shuen and Foulkes (2011). "Inherited mutations in breast cancer genes—risk and response." Journal of mammary gland biology and neoplasia **16**(1): 3-15.

Shukla, Vaitiekunas and Cotter (2012). "Advances in membrane proteomics and cancer biomarker discovery: current status and future perspective." Proteomics **12**(19-20): 3085-3104.

Siegel, Ma, Zou and Jemal (2014). "Cancer statistics, 2014." CA: a cancer journal for clinicians **64**(1): 9-29.

Siegel, Naishadham and Jemal (2012). "Cancer statistics, 2012." CA: a cancer journal for clinicians **62**(1): 10-29.

Siegel, Ward, Brawley and Jemal (2011). "The impact of eliminating socioeconomic and racial disparities on premature cancer deaths." Ca-a Cancer Journal for Clinicians **61**(4): 212-236.

Slamon, Clark, Wong, Levin, Ullrich and McGuire (1987). "Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene." Science **235**(4785): 177-182.

Slamon, Godolphin, Jones, Holt, Wong, Keith, Levin, Stuart, Udove and Ullrich (1989). "Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer." Science **244**(4905): 707-712.

Smeets, Yoshihara, Laenen, Reynders, Soens, Wildiers, Paridaens, Van Ongeval, Floris and Neven (2013). "Is the sentinel lymph node biopsy more sensitive for the identification of positive lymph nodes in breast cancer than the axillary lymph node dissection?" SpringerPlus **2**(1): 1-5.

Smith, Heys, Hutcheon, Miller, Payne, Gilbert, Ah-See, Eremin, Walker and Sarkar (2002). "Neoadjuvant chemotherapy in breast cancer: significantly enhanced response with docetaxel." Journal of clinical oncology **20**(6): 1456-1466.

Soerjomataram, Lortet-Tieulent, Parkin, Ferlay, Mathers, Forman and Bray (2012). "Global burden of cancer in 2008: a systematic analysis of disability-adjusted life-years in 12 world regions." The Lancet **380**(9856): 1840-1850.

Sopel, Kasprzyk and Berdowska (2011). "Maspin and c-erbB-2 expression in correlation with microvessel density in invasive ductal breast cancer." Folia Histochemica et Cytobiologica **43**(2): 109-108.

Soufi, Kumar, Gnad, Mann, Mijakovic and Macek (2010). "Stable isotope labeling by amino acids in cell culture (SILAC) applied to quantitative proteomics of Bacillus subtilis." Journal of proteome research **9**(7): 3638-3646.

Stark, Kleer, Martin, Awuah, Nsiah- Asare, Takyi, Braman, E Quayson, Zarbo and Wicha (2010). "African ancestry and higher prevalence of triple- negative breast cancer." Cancer **116**(21): 4926-4932.

Stefansson, Jonasson, Johannsson, Olafsdottir, Steinarsdottir, Valgeirsdottir and Eyfjord (2009). "Genomic profiling of breast tumours in relation to BRCA abnormalities and phenotypes." Breast Cancer Res **11**(4): R47.

Stratton, Campbell and Futreal (2009). "The cancer genome." Nature **458**(7239): 719-724.

Strimbu and Tavel (2010). "What are biomarkers?" Current opinion in HIV and AIDS **5**(6): 463.

Subik, Lee, Baxter, Strzepek, Costello, Crowley, Xing, Hung, Bonfiglio and Hicks (2010). "The expression patterns of ER, PR, HER-2, CK5/6, EGFR, Ki-67 and AR by immunohistochemical analysis in breast cancer cell lines." Breast cancer: basic and clinical research **4**: 35.

Sue, Lannin, Killelea and Chagpar (2013). "Predictors of microinvasion and its prognostic role in ductal carcinoma in situ." The American Journal of Surgery **206**(4): 478-481.

Switzer, Lim, Du, Al-Sairafi, Tonkin and Schiller (2014). "Case series of 21 patients with extrahepatic metastatic lobular breast carcinoma to the gastrointestinal tract." Cancer Treatment Communications.

Tabatabaei-Panah, Jeddi-Tehrani, Ghods, Akhondi, Mojtabavi, Mahmoudi, Mirzadegan, Shojaeian and Zarnani (2013). "Accurate sensitivity of quantum dots for detection of HER-2 expression in breast cancer cells and tissues." Journal of fluorescence **23**(2): 293-302.

Tainsky (2009). "Genomic and proteomic biomarkers for cancer: a multitude of opportunities." Biochimica et Biophysica Acta (BBA)-Reviews on Cancer **1796**(2): 176-193.

Tanaka, Decuzzi, Cristofanilli, Sakamoto, Tasciotti, Robertson and Ferrari (2009). "Nanotechnology for breast cancer therapy." Biomedical microdevices **11**(1): 49-63.

Tanke, Dirks and Raap (2005). "FISH and immunocytochemistry: towards visualising single target molecules in living cells." Current opinion in biotechnology **16**(1): 49-54.

Tarca, Draghici and Romero (2009). Developing classifiers for the detection of cancer using multi-analytes. Tumor Biomarker Discovery, Springer**:** 259-272.

Temel, Jackson, Billings, Dahlin, Block, Buss, Ostler, Fidias, Muzikansky and Greer (2007). "Phase II study: Integrated palliative care in newly diagnosed advanced non–small-cell lung cancer patients." Journal of clinical oncology **25**(17): 2377-2382.

Tharkar, Madani, Lasham, Shelling and Al-Kassas (2014). "Nanoparticulate carriers: an emerging tool for breast cancer therapy." Journal of drug targeting **23**(2): 97-108.

Thorat, Marchio, Morimiya, Savage, Nakshatri, Reis-Filho and Badve (2008). "Forkhead box A1 expression in breast cancer is associated with luminal subtype and good prognosis." Journal of clinical pathology **61**(3): 327-332.

Tsourkas, Behlke, Rose and Bao (2003). "Hybridization kinetics and thermodynamics of molecular beacons."Nucleic acids research **31**(4):1319-1330

Tyagi Sanjay and Kramer Fred Russell(1996). "Molecular beacons: Probes that fluoresce upon hybridization." Nature biotechnol **14**(3): 303-308.

Tyagi and Kramer (2012). "Molecular beacons in diagnostics." F1000 medicine reports **4**.

Uhl, Chan, Jones, Kelley, Assmann, Von Eckardstein, Sägers, Yvert, Schneider and Torralba (1997). "Elecsys CEA, PSA and AFP. Clinical results of a multicentre evaluation." Wiener klinische Wochenschrift **110**: 51-61.

van de Vijver, Bilous, Hanna, Hofmann, Kristel, Penault-Llorca and Rüschoff (2007). "Chromogenic in situ hybridisation for the assessment of HER-2 status in breast cancer: an international validation ring study." Breast Cancer Res **9**(5): R68.

Venter, Adams, Myers, Li, Mural, Sutton, Smith, Yandell, Evans and Holt (2001). "The sequence of the human genome." Science **291**(5507): 1304-1351.

Vet, and Marras (2005). "Design and optimization of molecular beacon real-time polymerase chain reaction assays." Methods in molecular biology. **288**: 273-290.

Vorobiof, Sitas and Vorobiof (2001). "Breast cancer incidence in South Africa." Journal of clinical oncology **19**(suppl 1): 125s-127s.

Wang and Hanash (2005). "Intact- protein based sample preparation strategies for proteome analysis in combination with mass spectrometry." Mass spectrometry reviews **24**(3): 413-426.

Wang, Yang, Chen and Shin (2008). "Application of nanotechnology in cancer therapy and imaging." CA: a cancer journal for clinicians **58**(2): 97-110.

Warner (2011). "Breast-cancer screening." New England Journal of Medicine **365**(11): 1025-1032.

Wasinger, Zeng and Yau (2013). "Current status and advances in quantitative proteomic mass spectrometry." International journal of proteomics **2013**.

Weedon-Fekjær, Lindqvist, Vatten, Aalen and Tretli (2008). "Breast cancer tumor growth estimated through mammography screening data." Breast Cancer Res **10**(3): R41.

Weigelt, Peterse and Van't Veer (2005). "Breast cancer metastasis: markers and models." Nature Reviews Cancer **5**(8): 591-602.

Weil and Chen (2011). "PARP inhibitor treatment in ovarian and breast cancer." Current problems in cancer **35**(1): 7-50.

Wheelock and Johnson (2003). "Cadherins as modulators of cellular phenotype." Annual review of cell and developmental biology **19**(1): 207-235.

Whelan, Pignol, Levine, Julian, MacKenzie, Parpia, Shelley, Grimard, Bowen and Lukka (2010). "Long-term results of hypofractionated radiation therapy for breast cancer." New England Journal of Medicine **362**(6): 513-520.

Winslow (2011). "Anatomy of the female breast." access date 5/4/2015 http://www.pic2fly.com/viewimage/Female%20Breast%20Structure/i_966a18cc92ea104809952334a85144cc

Wirapati, Sotiriou, Kunkel, Farmer, Pradervand, Haibe-Kains, Desmedi, Ignatiadis, Sengstag and Schutz (2008). "Meta-analysis of gene expression profiles in breast cancer: toward a unifiedunderstanding of breast cancer subtyping and prognosis signatures." Breast cancer research **10**(4): R65.

Wolf, Bose, Williamson, Miller, Karlan and Koeffler (2007). "FOXA1: Growth inhibitor and a favorable prognostic factor in human breast cancer." International journal of cancer **120**(5): 1013-1022.

Wong and Medrano (2005). "Real-time PCR for mRNA quantitation." Biotechniques **39**(1): 75.

Wood, Shapiro and Recht (2001). "Side effects of adjuvant treatment of breast cancer." New England Journal of Medicine **344**(26): 1997-2008.

Wooster, Neuhausen, Mangion, Quirk, Ford, Collins, Nguyen, Seal, Tran and Averill (1994). "Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13." Science **265**(5181): 2088-2090.

Wu, Pandey, WU, Ye, Zhu and Lobie (2013). "Prognostic significance of the expression of GFRα1, GFRα3 and syndecan-3, proteins binding ARTEMIN, in mammary carcinoma." Bio med central **13**(1): 34

Xing, Chaudry, Shen, Kong, Zhau, Chung, Petros, O'Regan, Yezhelyev and Simons (2007). "Bioconjugated quantum dots for multiplexed and quantitative immunohistochemistry." Nature Protocols **2**(5): 1152-1165.

Xu, Wang, Zhang, Li, Wang and Zhang (2012). "Clinical implications for BRCA gene mutation in breast cancer." Molecular biology reports **39**(3): 3097-3102.

Xue, Lu and Lai (2008). "The cancer secretome: a reservoir of biomarkers." J Transl Med **6**(52): 20.

Yamaguchi, Ito, Azuma, Honma, Yanagisawa, Nishikawa, Kawamura, Imai, Tatsuta and Inoue (2008). "FoxA1 as a lineage-specific oncogene in luminal type breast cancer." Biochemical and biophysical research communications **365**(4): 711-717.

Yanagawa, Ikemot, Kawauchi, Furuya, Yamamoto, Oka, Oga, Nagashima and Sasaki (2012). "Luminal A and luminal B (HER-2 negative) subtypes of breast cancer consist of a mixture of tumors with different genotype." BMC research notes **5**(1): 376.

Yang and Wang (2013). "Perspectives of nanotechnology in minimally invasive therapy of breast cancer." Journal of healthcare engineering **4**(1): 67-86.

Yates, Ruse and Nakorchevsky (2009). "Proteomics by mass spectrometry: approaches, advances, and applications." Annual review of biomedical engineering **11**: 49-79.

Yoshino, Peoples, Goedegebuure, Maziarz and Eberlein (1994). "Association of HER-2/neu expression with sensitivity to tumor-specific CTL in human ovarian cancer." The Journal of Immunology **152**(5): 2393-2400.

Zanivan, Meves, Behrendt, Schoof, Neilson, Cox, Tang, Kalna, van Ree and van Deursen (2013). "In vivo SILAC-based proteomics reveals phosphoproteome changes during mouse skin carcinogenesis." Cell reports **3**(2): 552-566.

Zhang, Chen, Wang and Drabier (2013). A neural network approach to multi-biomarker panel discovery by high-throughput plasma proteomics profiling of breast cancer. BMC Proc, BioMed Central Ltd.

Zhang, Luoh, Hon, Baertsch, Wood and Zhang (2007). "GeneHub-GEPIS: digital expression profiling for normal and cancer tissues based on an integrated gene database." Nucleic Acids Research **35**(suppl 2): W152-W158.

Zhang, Yee and Wang (2008). "Quantum dots for cancer diagnosis and therapy: biological and clinical perspectives."

Ziogas and Roukos (2009). "Genetics and personal genomics for personalized breast cancer surgery: progress and challenges in research and clinical practice." Annals of surgical oncology **16**(7): 1771-1782.