MOLECULAR AND CELLULAR EXPRESSION OF NOVEL GENES ASSOCIATED WITH BREAST CANCER

STEPHANIE ANNE-SOPHIE LAVERSIN

A thesis submitted in partial fulfilment of the requirements of Nottingham Trent University for the degree of Doctor of Philosophy

September 2009

Copyright Statement

J'aimerais dédier cette thèse tout particulièrement à ma maman, la meilleure maman du monde, mais aussi à toutes les personnes qui comptent ou ont compté pour moi. Merci d'être ou d'avoir été là pour moi, tout simplement.

"Soyons reconnaissants aux personnes qui nous donnent du bonheur, elles sont les charmants jardiniers par qui nos âmes sont fleuries." Marcel Proust

Acknowledgements

To start, I would like to thank Professor Robert Rees for allowing me to undertake this PhD within his research group and for his continuous guidance and support during the course of this work. I would also like to thank Dr Geng Li for his supervision at the beginning of the study. I am extremely grateful to my supervisors Dr Amanda Miles and Dr Graham Ball who have provided helpful suggestions, important advice and constant encouragement throughout the time of the research. I also wish to express my Appreciation to Dr Morgan Mathieu and Dr Stéphanie Mcardle who gave me constructive advice and valuable help in the lab.

Special thanks are due to Steve for providing great technical assistance during my research. Thank you to Dr Balwir Matharoo-Ball, Clare and Christophe for giving me great support for the mass spectrometry and bioinformatics side of the project.

An extended thank you to Professor Ian Ellis, Dr Des Powe and colleagues at the QMC in Nottingham as well as Dr Sonia Perez and Dr Angelos Gritzapis at Saint Savas Cancer Hospital (Greece) and Dr Aija Line at the Latvian University Biomedical Centre for the collaborative work and for providing samples which are crucial for this kind of expression analysis based project. Special Gratitude goes to the other group members of the John van Geest Cancer Research Centre, past and present, for offering me countless advice, valued conversations and laughter.

I would like to express my warmest thanks to my fellow PhD students, my dearest friend Florence and also Krish, Baharak, Gary, who shared with me the good times and the bad times that go with doing a PhD thesis and who gave me great moral support when needed and sincere friendship. My heartiest thanks also go to Richard for our comforting conversations and to my closest friends, in France and in England, Christelle, Natali, Marie and Olivier for their invaluable friendship and the great times we shared during all these years.

Finally, I would like to express my wholehearted special thanks to my family for their infinite encouragement and support: thank you to Alison, David and Graham; thank you to my mum Liliane, my dad Michel and my brother Christophe; thank you to my grandpa Robert, my brother Sébastien and my grandma Odette, all who sadly passed away and who I greatly miss; and last but not least, thank you to Alex, the love of my life, for his patience, understanding, care and love which are basically everything a girlfriend doing a PhD thesis could dream of.

Contents

List of Figures

List of Tables

Abbreviations

AICD	Activation-induced cell death
ANN	Artificial neural network
APC	Antigen-presenting cell
BAGE	B melanoma antigen
BCR-ABL	Breakpoint cluster region/Abelson
BLAST	Basic local alignment search tool
bp	Base pair
BRCA	Breast cancer susceptibility
BUC	Breast-associated UniGene clusters
CDK	Cyclin-dependent kinase
CDKN	Cyclin-dependent kinase inhibitor
cDNA	Complementary deoxyribonucleic acid
CEA	Carcinoembryonic antigen
CML	Chronic myeloid leukaemia
CpG	Cytosine and guanine separated by a phosphate
CT	Cancer testis
CTL	Cytotoxic T lymphocytes
DC	Dendritic cell
DCIS	Ductal carcinoma *in situ*
DISC	Disabled infectious single cycle
DNA	Deoxyribonucleic acid
EBV	Epstein-Barr virus
EGF	Epidermal growth factor
ELISA	Enzyme-linked immunosorbent assay
ER	Estrogen receptor
EST	Expressed sequence tag
FACS	Fluorescent-activated cell sorting
FCS	Foetal calf serum
FDA	Food and drug administration
FITC	Fluorescein isothicyanate
GAGE	G antigen
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase

GM-CSF	Granulocyte macrophage-colony stimulating factor
HAGE	Helicase antigen
HCV	Hepatitis c virus
HER2	Human epithelial growth factor receptor 2
HHV8	Human herpesvirus 8
HPRT1	Hypoxanthine guanine phosphoribosyl transferase 1
HRP	Horse radish peroxidase
HPV	Human papillomavirus
HSP	Heat shock protein
HSV	Herpes simplex virus
HTLV	Human T cell leukemia virus
Id	Idiotype
IFA	Incomplete Freund's adjuvant
IFN	Interferon
Ig	Immunoglobulin
IHC	Immunohistochemistry
IL	Interleukin
kDa	Kilodalton
L	Ligand
LCIS	Luminal carcinoma *in situ*
LDLR	Low density lipid receptor
LPS	Lipopolysaccharide
MAGE	Melanoma antigen
MALDI	Matrix-assisted laser desorption/ionisation
MHC	Major histocompatibility complex
MIC	MHC class I-related chain
miRNA	microribonucleic acid
MLP	Multiple layer perceptron
mRNA	messenger ribonucleic acid
NK	Natural killer cell
NKG2D	Natural killer cell activating receptor
NPI	Nottingham Prognostic Index
NY-ESO-1	New York esophageal squamous cell carcinoma antigen
PBMC	Peripheral blood mononuclear cell
PBS	Phosphate buffer saline
PDGF	Platelet-derived growth factor
PI	Propidium iodide
Poly I.C	Polyinosinic polycytidylic acid
PR	Progesterone receptor
PSA	Prostate-specific antigen

PSMA	Prostate-specific membrane antigen
Rb	Retinoblastoma
RNA	Ribonucleic acid
RP	Ribosomal protein
RT	Room temperature
RT-PCR	Reverse transcriptase-polymerase chain reaction
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SELDI	Surface-enhanced laser desorption/ionisation
SEREX	Serological analysis of recombinant tumour cDNA expression libraries
siRNA	Small interfering ribonucleic acid
TAA	Tumour-associated antigen
TBP	TATA box binding protein
TCR	T cell receptor
TGF	Transforming growth factor
Th	T helper lymphocytes
TLR	Toll-like receptor
TMA	Tissue microarray
TNF	Tumour necrosis factor
Treg	Regulatory T cell
UK	United Kingdom
ULPB	UL16-binding proteins
USA	United States of America
UV	Ultraviolet
VEGF	Vascular endothelial growth factor

Abstract

An improved and individualised breast cancer care can be pursued with the use of tumour markers that are associated with risk assessment, diagnosis, prognosis, prediction of treatment outcome, monitoring of the disease and development of novel therapies. As such, this study proposed to characterise and evaluate the potential for the management of breast cancer of four novel unpublished breast-associated genes, the breast-associated UniGene clusters (BUC6, BUC9, BUC10 and BUC11), which were identified in our laboratory by database mining. The first aspect of this research project was to undertake extensive *in silico* analysis and *in vitro* expression analysis for all genes using sequencing and RT-PCR based assays on a variety of tissues of normal and cancerous origin in order to further select the most promising candidates for potential use in cancer care. In light of all the data collected, the work was subsequently focused on the BUC11 gene. Since BUC11 mRNA was found to be expressed at very low levels in normal tissues except in some non-essential tissues where it is expressed at a higher level, BUC11 was not considered a cancer testis antigen. BUC11 mRNA was also found at high levels in all testicular cancer tissues tested as well as in a quarter of prostate cancer tissues tested and at lower levels in a small proportion of melanomas, and thus BUC11 was not considered to be a tumour-specific gene. Quantitative RT-PCR analysis of a large panel of normal/cancer breast tissues has demonstrated that up to 97% of all tissues expressed BUC11 mRNA and 60.6% of patients over-expressed BUC11 mRNA. Immunoassays performed with the custom-made anti-BUC11 antibody showed that BUC11 protein was preferentially found in the epithelial luminal type of malignant breast cells however no evidence was found to link BUC11 protein with breast tumour formation, state or progression of the disease. Therefore, the study also aimed to elucidate the molecular pathways that BUC11 may be involved in. Using a combination of *in vitro* gene silencing/induction based experiments and gene chip microarray technology, BUC11 was shown to potentially have a function in pathways controlling the proliferation of cells. The data obtained on the BUC11 gene during the course of this study provided a rationale to further investigate the potential immunogenicity of its protein in a transgenic mouse model however preliminary experiments did not lead to any conclusive data. Collectively, this study demonstrated the expression analysis of a novel gene set which can potentially have clinical utility for several cancers, especially the most promising candidate BUC11, and which require further investigation.

Chapter 1

Introduction

"Cancer is a word, not a sentence"

(John Diamond, British broadcaster and journalist, who died of throat cancer in 2001)

I have chosen this simple but meaningful quote, which has been used worldwide by physicians, charities and others, because it exposes that cancer is not necessarily a fatality as with appropriate care people can survive cancer.

1.1 Cancer: an overview

1.1.1 Cancer is a class of diseases which has a multifactorial, multistep causation

Cancer is a worldwide health concern that can affect anyone at any age. Worryingly one in three people will be diagnosed with cancer in their lifetime and one in four will die (Quinn *et al.*, 2001). According to the World Health Organisation, 7.9 million people died from cancer in the world in 2007, representing 13% of all worldwide deaths (http://www.who.int, 2009). It has been projected that cancer will cause approximately 12 million deaths in 2030. Almost any tissue of the body can be affected and the most common cancers worldwide for both men and women are lung, stomach, liver, colon and breast cancers.

The term "cancer" refers to a group of diseases, with over 100 types reported. Carcinomas represent approximately 85% of all cancers and originate from epithelial cells whereas sarcomas develop from mesodermal cells such as those found in bone and muscle and adenocarcinomas from glandular tissue such as the one found in breast. A cancer cell arises from an abnormal cell that has accumulated and sustained multiple defects in its DNA. Then, daughter cancer cells multiply in an unregulated manner, passing on these defects and aggregate in a mass called tumour. Abnormal cancer cells also have the ability to invade and destroy the surrounding healthy tissue. From the original site, cancer

cells can eventually spread to other parts of the body (metastasis) via the bloodstream or the lymphatic system causing further damage by compromising tissue functions. Surprisingly, the immune system which has a role in identifying and killing tumour cells does not always eliminate the abnormal cells. Sir Frank Macfarlane Burnet introduced the concept of immunological surveillance (or immunosurveillance) in the early seventies. In his theory, the normal immune system is able to detect and destroy newly mutated cells, however some of them can bypass this surveillance, thus causing cancer by growing and spreading (Burnet, 1970). This theory was supported by the observation of spontaneous tumour regressions (Hoption Cann, 2002). Cancer immunoediting also demonstrates immunosurveillance as this process was shown to have two antagonistic functions in immunocompetent hosts: to recognise and eliminate nascent transformed cells thus protecting the host against tumour progression and to assist a tumour that has reduced immunogenicity to grow and escape destruction by the immune system (Dunn *et al.*, 2002; Dunn *et al.*, 2005).

In 2000, Hanahan and Weinberg described six common features of all cancer types (Figure 1.1). The six hallmarks of cancer, vital for the process of carcinogenesis are: growth signal autonomy, evasion of growth inhibitory signals, evasion of apoptosis, unlimited replicative potential, the formation of new blood vessels (angiogenesis), invasion and metastasis (Hanahan and Weinberg, 2000). Researchers are studying and targeting each of these hallmarks to develop novel therapies for cancer patients.

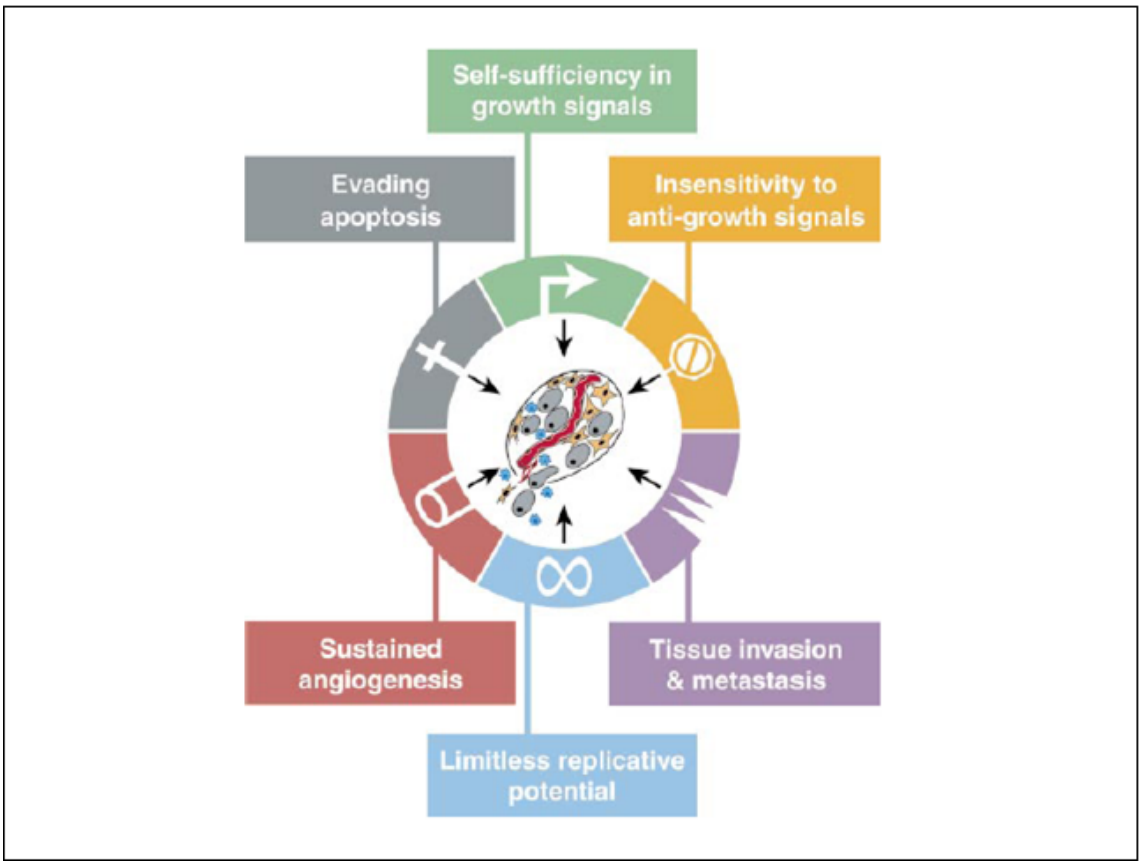


Figure 1.1: The six hallmarks of cancer.
This figure was taken from Hanahan and Weinberg (2000).

Cancers are genetic diseases at the cellular level. The majority of mutations responsible for carcinogenesis are somatic, as opposed to inherited. Inherited mutations (germline mutations) have occurred in the sperm or egg cells of the parents. Continuous or frequent exposure of DNA to carcinogens allows mutations to accumulate, increasing the chances that these mutations will eventually lead to cancer and explains the fact that the risk of developing cancer increases with age. Many factors contribute to carcinogenesis, whether they are endogenous (hereditary mutations, hormonal balance, strength of the immune system, by-products of the metabolism or errors during DNA replication) or exogenous (diet, reproductive life, exposure to carcinogens such as chemicals, Ultra Violet (UV) rays or viruses). Epidemiologic studies revealed that the four most important factors influencing transformation are the environment, smoking, diet and reproductive life. UV rays and chemicals can induce direct or indirect DNA mutations, for instance exposure to UVB radiation from the sun has been linked to melanoma (form of skin cancer) whereas hormones are not mutagenic but they can promote cell proliferation. Also, sexually transmitted viruses can promote cancer, for example, it is generally accepted that virtually all cervical cancers are positive for human papillomaviruses which cause the disease. Perhaps the most recognisable risk factor for a particular cancer is smoking, with approximately 81 carcinogens known to be present in cigarette smoke (Smith *et al.*,

2003). In addition, chronic inflammation, such as ulcers, has been shown to play a role in the development of up to 5% of all cancers. Some physiological factors that contribute to the development of cancer comprise of mutagenic oxygen radicals (by-products of aerobic metabolism) and mutations caused by error-prone polymerases during the normal cell processes of DNA replication or repair.

1.1.2 Oncogenes and tumour suppressor genes

Mutations acquired and sustained by a cell affect the regulation of its cell cycle and normal cellular processes during the development of the malignant cell phenotype, such as cell proliferation and survival, through the multi-step process of activation of proto-oncogenes into oncogenes and inactivation of tumour suppressor genes. In the model of tumour development by succession of clonal expansions that follows a Darwinian model of evolution, a cell that acquired and sustained activation of one oncogene or inactivation of one tumour suppressor gene gives rise to daughter cells that have growth and survival advantages thus proliferate greater than the neighbouring cells in the tissue. In that dominating clonal population, one cell will acquire an additional oncogene activated or tumour suppressor gene inactivated, thus gaining a further advantage for independent growth. It is believed that cancer progression can be explained by a sequence of four to six clonal expansions.

1.1.2.1 Oncogenes

Oncogenes (viral or cellular) are a class of growth-promoting genes involved in the pathogenesis of cancer, with more than 100 identified to date. Almost all oncogenes are the mutated forms of normal, pre-existing cellular genes called proto-oncogenes that have been activated either through retrovirus acquisition or somatic mutation. The protein product of an oncogene is either produced in higher quantities than normal or has a dramatically increased activity.

In the early 1970s, following animal work, the hypothesis that tumour viruses were responsible for initiating all human cancers led to extensive research on retrovirus-associated oncogenes and notably the discovery of the Sarcoma (Src) oncogene induced by the Rous retrovirus and found in colon cancer (Martin, 1970; Martin, 2004). Retroviruses need the host cellular mechanisms for energy and synthesis of new viral particals. The RNA of the virus is reverse-transcribed into DNA before being randomly integrated into the host cell genome. Over generations, viruses can acquire coding regions of the host cellular genes into their own RNA, resulting in the creation of oncogenes. For instance, the Rous Sarcoma Virus has integrated into its RNA a truncated form of the cellular proto-oncogene *c-src*. The viral oncogenes are transcribed in virus-infected cells under the control of the viral transcriptional promoter, independently to cellular signals that used to regulate the

proto-oncogenes and the resulting over-expression of the oncogenes promotes unregulated growth of the cells.

Using transfection experiments (DNA from chemically transformed cells introduced in normal cells), another type of oncogene was discovered that did not originate from viral infection but from somatic mutation in proto-oncogenes, either spontaneously or through exposure to endogenous or exogenous carcinogens. Somatic mutations affect the structure of the protein encoded by the oncogene or induce over-expression of the normal protein if the mutation occurred in the gene promoter region.

Taken together, the proteins encoded by oncogenes are involved at different levels in growth factor signal transduction pathways and include growth factors, growth factor receptors, intracellular signal transducers and transcription factors. During the course of tumour progression, activation of oncogenes, along with the inactivation of tumour suppressor genes, allows tumour cells to evade the growth inhibitory signals secreted by the surrounding normal cells thus acquiring growth signal autonomy and uncontrollable proliferation (Bertram, 2000). For example, the over-expression of the mutated form of the platelet-derived growth factor (PDGF) receptor and the constitutive activation of the epidermal growth factor receptor (EGFR) have been linked to uncontrolled proliferation of human glioma cells through an autocrine activation (Westermark *et al.*, 1995) and androgen-independent prostate cancer cells (Fong *et al.*, 1992), respectively. Also, the human epithelial growth factor receptor 2 (HER2), which has a role in signal transduction pathways leading to cell growth and differentiation, was found to be over-expressed in up to 25% of breast tumours and correlated with a worse prognosis for patients (Menard *et al.*, 2001; Slamon *et al.*, 1989). Constitutive activation of signalling pathways result in a high proliferative state of cells and perhaps the most cited example of intracellular signal inducers is the proto-oncogene RAS protein, a small GTPase, found in 20 to 25% of human tumours (Downward, 2003). In these tumours, cells acquire the neoplastic phenotype through the deregulation of RAS signalling pathways involved in cell growth and proliferation. Scholl and colleagues have demonstrated that transformation of mouse cells, by transfecting them *in vitro* with the Ras gene, is observed only following the immortalisation of the cells prior to the experiment (Scholl *et al.*, 2005); therefore Ras-induced malignant transformation can only occur with the involvement to the process of many other genes, acting altogether. Also, the constitutive and overexpression of the transcription factor c-Myc protein, due to chromosomal translocation, is responsible for the improper increase in transcription of Myc-regulated genes. This results in expression of proteins such as cyclins A and E which control the progression of cells through the cell cycle which then force tumour cells to divide uncontrollably (Ponzielli *et al.*, 2005).

1.1.2.2 Tumour suppressor genes

Tumour suppressor genes have an “anti-oncogenic” role and are involved in the control (limitation) of cell growth, cell survival and in DNA repair. Unrepaired genetic

changes in tumour suppressor genes lead to their inactivation, thus promoting survival of transforming cells and favouring progression of a malignant cell phenotype. The number of inactivated tumour suppressor genes in the genomes of human cancer cells is far greater than the number of activated oncogenes. According to Knudson's two-hit hypothesis, individuals are predisposed to cancer through the inheritance of a germline mutation in one allele of a tumour suppressor gene followed by the mutation in the second allele of that tumour suppressor gene during their life (Knudson, 1971). This hypothesis however only explains familial (inherited) forms of cancers as, in the case of sporadic forms of the disease, individuals can acquire subsequent mutations in both alleles of the tumour suppressor genes during their life or they can have an allelic loss followed by a mutation of the second allele. Other than mutation, the function of tumour suppressor genes can also be decreased or lost by epigenetic phenomenons (Esteller *et al.*, 2001) like promoter methylation which is frequently found in Wilm's tumours (Wagner *et al.*, 2002). Perhaps the most characterised tumour suppressor genes to date are the p53 and retinoblastoma (Rb) genes (Figure 1.2).

Retinoblastoma is a very treatable cancer developing in the cells of the retina and is the most common ocular malignancy in children. The disease results from either the genetic form or the non-genetic form (55% cases diagnosed with the latter). Defects in the Rb gene can either be inherited from a parent or is the result of somatic mutations during the development of the foetus. The nuclear protein encoded by the Rb gene (RB) controls the entry of cells into the S phase (DNA replication). Mutations in the Rb gene affect the struture of RB, compromising the interactions of the protein with the complex of transcription factors (E2F/DP). Under normal conditions, if RB is hypophosphorylated it will bind and inactivate E2F/DP. Defects in genes encoding proteins that normally control RB phosphorylation are also responsible for indirectly inactivating RB which then can not bind the E2F/DP transcription complex. Consequently, the activity of this complex is never repressed, the control over the cell cycle at the S phase entry is lost and cells can proliferate in an unrestricted manner. Overexpression of D-type cyclins, amplification of cyclin-dependent kinases 4 and 6 (CDK4, CDK6) and inactivation of cyclin-dependent kinase inhibitor p16 (CDKN2 gene) have all been associated with a decreased or lost function of the RB protein (Lukas *et al.*, 1995).

A tumour suppressor gene, which is a close "associate" of the Rb gene and has a very important role in regulating the stability of the genome (DNA repair mechanisms), replicative senescence, premature senescence and apoptosis, is the p53 gene (Bertram, 2000). The P53 protein, also referred as "guardian of the genome", is activated when DNA defects are detected and induces an arrest of the cell cycle and/or apoptosis so that the errors in the genome are not passed on to daughter cells. In response to stress stimuli, activated P53 induces an increase of the cyclin-dependent kinase inhibitor 1A, CDKN1A

(p21) expression level. The protein encoded by the p21 gene blocks the progression of the cell cycle at the G1 phase by binding and inhibiting the activity of various cyclin-CDK kinase activities (Bearss *et al.*, 2002). Over 70% of all human cancers have their p53 pathway inactivated and also present mutations in some genes upstream or downstream to this pathway (Levine, 1997). Inactivation of the tumour suppressor gene p53 is the result of mutations which cause changes in the protein structure, preventing the formation of hydrogen bonds between the DNA and P53 (Cho *et al.*, 1994). Inactivation of p53 pathway contributes to the progression of tumour development through evasion of apoptosis and uncontrolled cell growth.

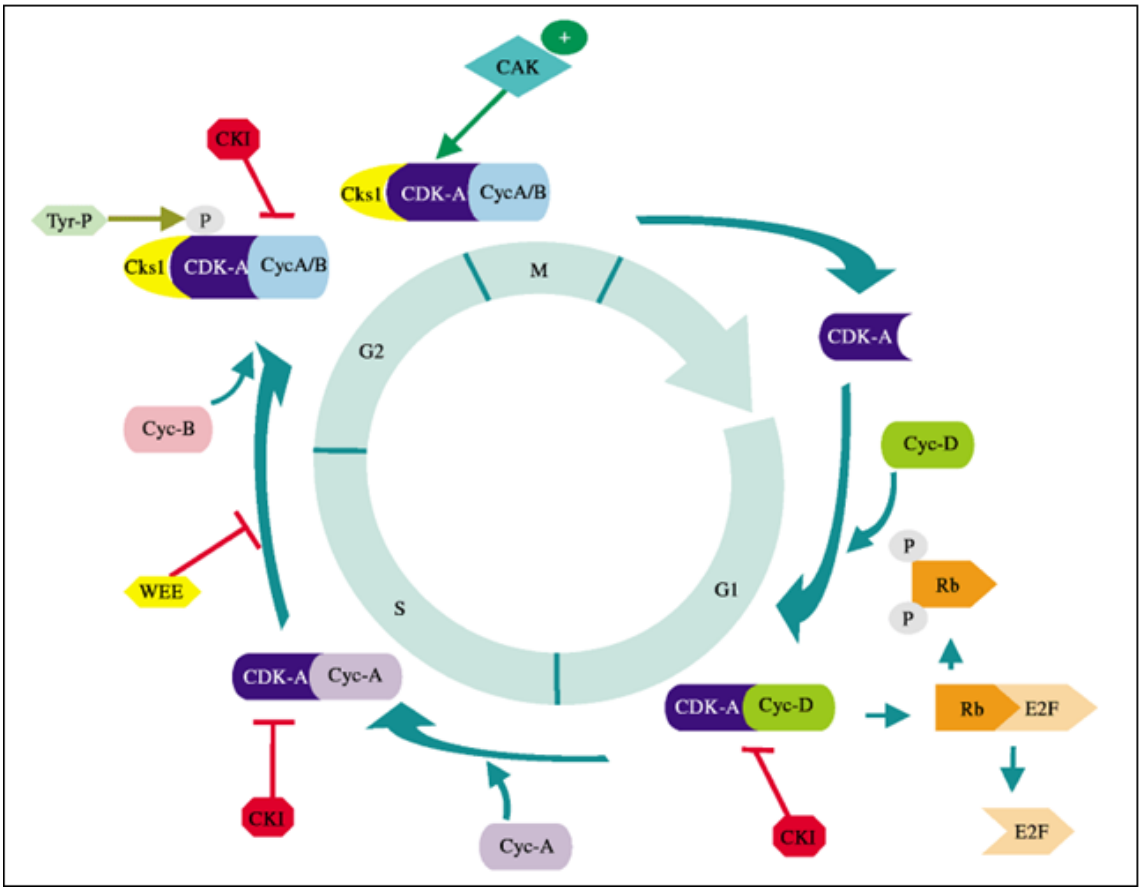


Figure 1.2: The mammalian cell cycle.
Normal growth and division of a mammalian cell are controlled by the processes of the cell cycle which is divided into four phases: G1, S (replication of DNA), G2 and M (mitosis). A fifth phase, G0 (not represented in the figure), is the state of nonproliferating cells that left the active cell cycle during G1. Progression through the cycle is rigorously guided by the activity of diverse molecules such as cyclins (Cyc), cyclin-dependent protein kinases (CDK) and the tumour suppressor gene retinoblastoma (Rb). Taken from Andrietta *et al.*, 2001.

1.1.3 Tumour-associated antigens

The accumulation of activated proto-oncogenes and inactivated tumour suppressor genes in a tumour cell has inevitably an effect on the expression of other genes involved in the same pathways or upstream and downstream to these pathways; the proteins of these affected genes are classified as Tumour Associated Antigens (TAA). The first human tumour antigen identified, called MAGE-1, was isolated from a melanoma patient by van der Bruggen and colleagues in 1991 (van der Bruggen *et al.*, 1991). They used a genetic approach in which constructed cDNA expression libraries from the melanoma tissue of a patient were transfected into target cells and then these target cells presented the antigen on the cell surface and any antigens that were able to induce an antigen-specific immune response were identified. Since the early nineties, many tumour antigens, recognised by T cells and capable of inducing a tumour-specific T cell response, have been identified in many different cancers by three approaches (Miles *et al.*, 2006): genetic, biochemical and "reverse immunology". These include techniques such as the serological analysis of recombinant tumour cDNA expression libraries (SEREX) (Li *et al.*, 2004; Miles *et al.*, 2007) and mass spectrometry (Matharoo-Ball *et al.*, 2007). The TAA can be classified into 5 main groups (Table 1.1): cancer testis antigens, antigens overexpressed in tumours, differentiation antigens, tumour antigens resulting from mutations such as fusion proteins and viral antigens. The characteristics that make a TAA an ideal target for the development of novel anti-cancer strategies include: to be solely or mainly expressed by tumour cells, to be vital to the tumour and to be immunogenic.

Categories	Tumour antigens	Subcellular location	Associated cancers	References
Cancer testis antigens	BAGE-1	Secreted (potential)	Tumours of various histological types	Boel *et al.*, 1995
	GAGE-1	Nucleus and/or cytoplasm	Tumours of various histological types	Van den Eynde *et al.*, 1995
	HAGE	Cytoplasm	Tumours of various histological types	Martelange *et al.*, 2000
	MAGE-1 and MAGE-3	Cytoplasm	Tumours of various histological types	van der Bruggen *et al.*, 1991
	NY-ESO-1	Cytoplasm	Tumours of various histological types	Scanlan *et al.*, 2004
Antigens overexpressed in tumours	HER-2/neu	Cell membrane	Ovary, breast, lung carcinomas	Fisk *et al.*, 1995
	MUC1	Cell membrane, secreted, cytoplasm or nucleus	Tumours of various histological types	Brossart *et al.*, 1999
	p53 (wild-type)	Nucleus or cytoplasm	Colon, breast and other carcinomas	Ropke *et al.*, 1996
	PSMA	Nucleus or cytoplasm	Prostate carcinoma	Horiguchi *et al.*, 2002
	Telomerase	Nucleus or cytoplasm (including mitochondria)	Tumours of various histological types	Vonderheide *et al.*, 1999
Differentiation antigens	gp100	Cell membrane or secreted (potential)	Melanoma	Bakker *et al.*, 1995
	NY-BR-1	Cell membrane and cytoplasm	Breast carcinoma	Wang *et al.*, 2006
	PSA	Secreted	Prostate carcinoma	Correale *et al.*, 1997
Fusion proteins	BCR-ABL	Cytoplasm or nuclear entrapment	Chronic myeloid leukemia	Bosch *et al.*, 1996
	LDLR-fucosyltransferase	Nucleus	Melanoma	Wang *et al.*, 1999
Viral antigens	EBNA-3A (EBV)	Nucleus	Burkitt's and Hodgkin's lymphoma	Murray *et al.*, 1992
	E2 (HCV)	Envelope	Hepatocellular carcinoma	Soldaini *et al.*, 2003
	Latent nuclear antigen-1 (HHV8)	Nucleus	Kaposi's sarcoma	Hong *et al.*, 2003
	E6, E7 (HPV16)	Nucleus	Cervical carcinoma	Feltkamp *et al.*, 1993
	HTLV-1 p40x (Tax)	Nucleus	T cell leukaemia	Koenig *et al.*, 1993

Table 1.1: Examples of tumour antigens classified into five categories.

1.1.4 The immune system: first line of defense against cancer

As stated previously, Burnet described in 1970 the concept of immune surveillance in which he hypothesised that the immune system can recognise cancer cells and act to stop their growth or prevent the formation of tumours (Burnet, 1970). The hypothesis has been long-discredited for many years until the mid-nineties when many studies on tumour immunology have fueled this assumption with the discovery of specialised cells of the immune system involved in this concept.

1.1.4.1 Innate immune system

The innate immune system, also referred as the "inbuilt immune protection", act rapidly upon infection of the body by external agents such as bacteria, fungi or viruses. The cells of the innate immune system circulate continuously and destroy almost immediately the pathogens upon encountering them. The innate immune system is constituted

of many types of cells that can bind to surface molecules "instinctively" recognised as foreign, displayed by the infectious agents or by the infected cells (Janeway, 1991). The most prominent players are natural killer cells (NK) and phagocytic cells (dendritic cells, macrophages), originating from lymphoid progenitor cells and myeloid progenitor cells respectively. Phagocytic cells engulf and lyse the foreign agents or infected cells whereas natural killer cells release cytotoxic granules which lyse the targets. Phagocytic cells can distinguish "self" (the body's own healthy cells) from "non-self" (the pathogens or infected cells) through the Toll-like receptors (TLR) that they express (Kopp and Medzhitov, 2003). These receptors bind to a variety of ligands whose patterns are common to most foreign agents, for example, double-stranded RNA (TLR3), unmethylated CpG DNA motifs (TLR9), peptidoglycans and lipoproteins (TLR2). Upon binding, the cells become activated and release signalling molecules (cytokines) to recruit other immune cells in order to increase the intensity of the response to the infection. If this first line of defense requires help during an infection, these cytokines can activate antigen-presenting cells (APC) belonging to the adaptive immune system which release co-stimulatory molecules that further engage other components of adaptive immunity.

This co-operation between the innate immune system and the adaptive immune system can also be observed in the case of cancer. The immune cells of the innate system are the first to "sense the danger" at the tumour site and alert the adaptive immune system. The mechanisms by which these cells can detect tumour cells are still unclear but are most likely linked to biological signals of dying cells (Shi *et al.*, 2003) or transformed cells (Biassoni, 2008). The combination of activation of oncogenes, inactivation of tumour suppressor genes and expression of tumour antigens in a cell result in biological characteristics that can be recognised and targeted as "abnormal" by the immune cells. Stressed cells release molecules like heat shock proteins (HSP) which are ligands to TLR (Balogh *et al.*, 2009). Furthermore, tumour cells developing under the stressful condition of hypoxia (environment deprived from oxygen supply) express specifically certain surface molecules such as the UL16-binding proteins (ULBPs), which are major histocompatibility complex (MHC) class I-related molecules (MICs), and which are ligands to the natural killer cell activating receptor NKG2D/DAP10 (Sutherland *et al.*, 2001). Activated NK cells then release the cytokine called interferon-gamma (IFN-γ) which has several roles in the immune system. It can target directly the cancer cells by either blocking important mechanisms required for the progression of the tumour or cause them to produce cytokines to recruit other immune cells, or indirectly by provoking the maturation of phagocytic cells which then secrete the cytokine Interleukin-12 (IL-12) that further activates specific cells of the adaptive immune system, namely cytotoxic T lymphocytes (Degli-Esposti and Smyth, 2005). Other receptors present on the cell surface of NK cells are the Killer Inhibitory Receptors (KIR) and C-type lectin-like receptors which specifically bind to MHC class I molecules, found on almost every nucleated cell of the

body, and induce self-tolerance; however, it also allows the escape from immune detection of tumour cells that have acquired down-regulation of the expression of MHC class I molecules (Moretta *et al.*, 1996; Restifo *et al.*, 1993). Approximately 90% of the animal species rely on the innate (nonspecific) immune system alone for protection against pathogens whereas the rest of the animal kingdom possess the adaptive immune system (antigen specific) as well.

1.1.4.2 Adaptive immune system

Dendritic cells (DC) are the most studied APC of the immune system as they efficiently mediate the collaboration between the innate and the adaptive immune pathways. Immature dendritic cells derived from hematopoietic bone marrow progenitor cells migrate to different tissues throughout the body via the blood. Dendritic cells engulf antigens, process them internally and present fragments of these on MHC class II molecules (found only on APC) as well as MHC class I molecules to lymphocytes. Dendritic cells become mature by up-regulating the expression of MHC molecules and co-stimulatory cell surface receptors (B7, CD40). Once activated, they migrate to lymph nodes where they encounter helper T cells and cytotoxic T cells (CTL). The process of "co-conditioning" of DCs and T cells takes place in the presence of IFN when the co-stimulatory proteins of DCs, CD40 and B7, bind respectively to the co-stimulatory molecules CD40 ligand (CD40L) and CD28 expressed on the cell surface of T cells (Ridge *et al.*, 1998). Mature dendritic cells produce diverse cytokines such as IL-12 to further activate T cells and they present antigens (in the form of peptides) on their MHC molecules to T cell receptors (TCR) in order to initiate specific immune responses. The adaptive immune system comprised two types of immune responses: humoral and cellular. An overview of the humoral and cellular arms of the immune system is given in Figure 1.3.

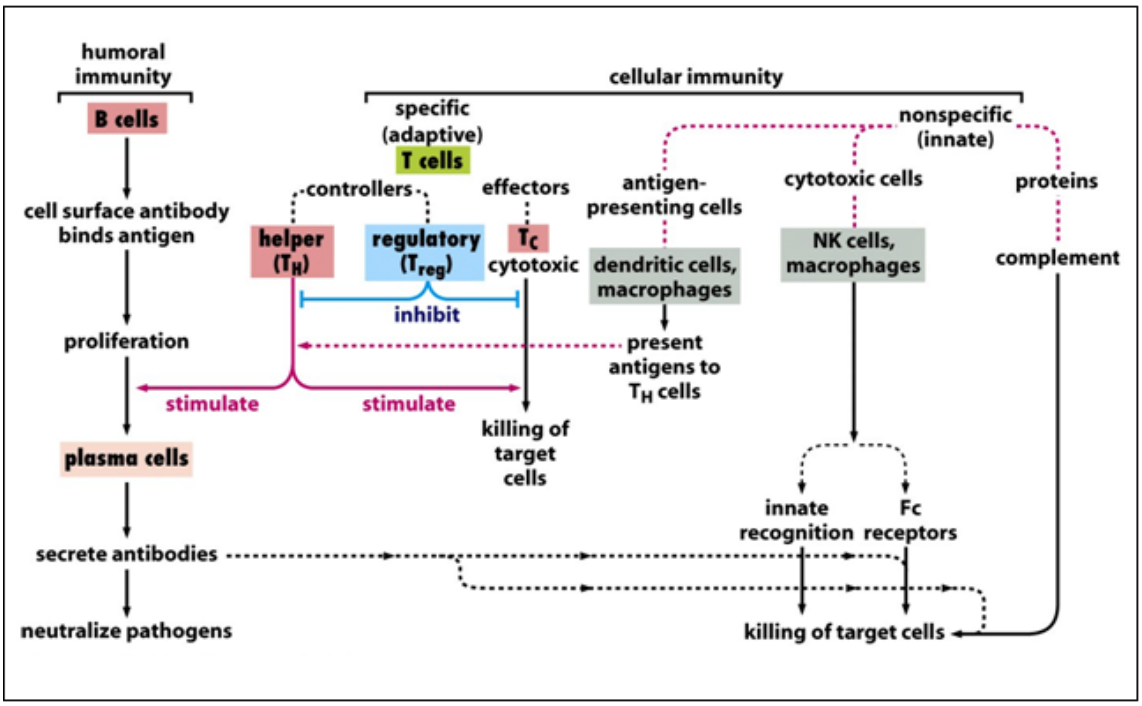


Figure 1.3: Overview of the humoral and cellular arms of the immune system. Taken from Weinberg (2007).

The main components of the humoral immune system are the B cells. They are lymphocytes that, upon stimulation by activated helper T cells, proliferate and differentiate into plasma cells. These plasma cells then secrete antibodies that recognise TAA present on tumour cells and help in the killing of the target cells. B cells act also as antigen-presenting cells thus activating T cells and making them produce cytokines which are important molecules for the initiation and maintenance of the humoral immune response by B cells. On the other hand, the main players of the cellular immune system are the T cells: helper T cells, CTLs and regulatory T cells. CTLs express the glycoprotein CD8 on their cell surface (CD4-/CD8+) whereas helper T cells can be divided into two groups: T helper 1 lymphocytes and T helper 2 lymphocytes. These cells are known respectively CD4+ Th1 cells and CD4+ Th2 cells and express the glycoprotein CD4 on their cell surface (CD4+/CD8-). CD4+ Th2 cells are the mediators of the humoral immune response (Th2) whereas CD4+ Th1 cells mediate the stimulation of CD8+ T cells.

T cells originate from lymphoid progenitors in the bone marrow. All T cells are initially double-negative (CD4-/CD8-) then develop into double-positive (CD4+/CD8+, low expression). The latter then become either CD4+/CD8- cells upon positive selection through the interaction TCR/MHC class II molecules with self-antigens or CD4-/CD8+ upon positive selection through the interaction TCR/MHC class I molecules with self-antigens (Fink and Bevan, 1978). Then, these selected T cells that have shown capability of binding to MHC molecules undergo a negative selection: the ones that react to self-antigens are removed in order to avoid the development of autoimmune diseases (Surh

and Sprent, 1994). In the lymph nodes, mature DCs interact with CD4+ Th1 cells through cytokines, co-stimulatory molecules and MHC molecules presenting tumour-derived peptides. Then, activated CD4+ Th1 cells stimulate CD8+ T cells to become CTLs. Activated CTLs then recognise MHC class I antigen and peptide and destroy the antigen-bearing tumour cells by releasing cytotoxic granules such as perforins which pierce the tumour cell membrane or triggering apoptosis through death receptors (Figure 1.4). According to several studies, the activity of CD4+ Th1 cells and CD8+ T cells appears to be crucial to the generation of a potent anti-tumour immune response.

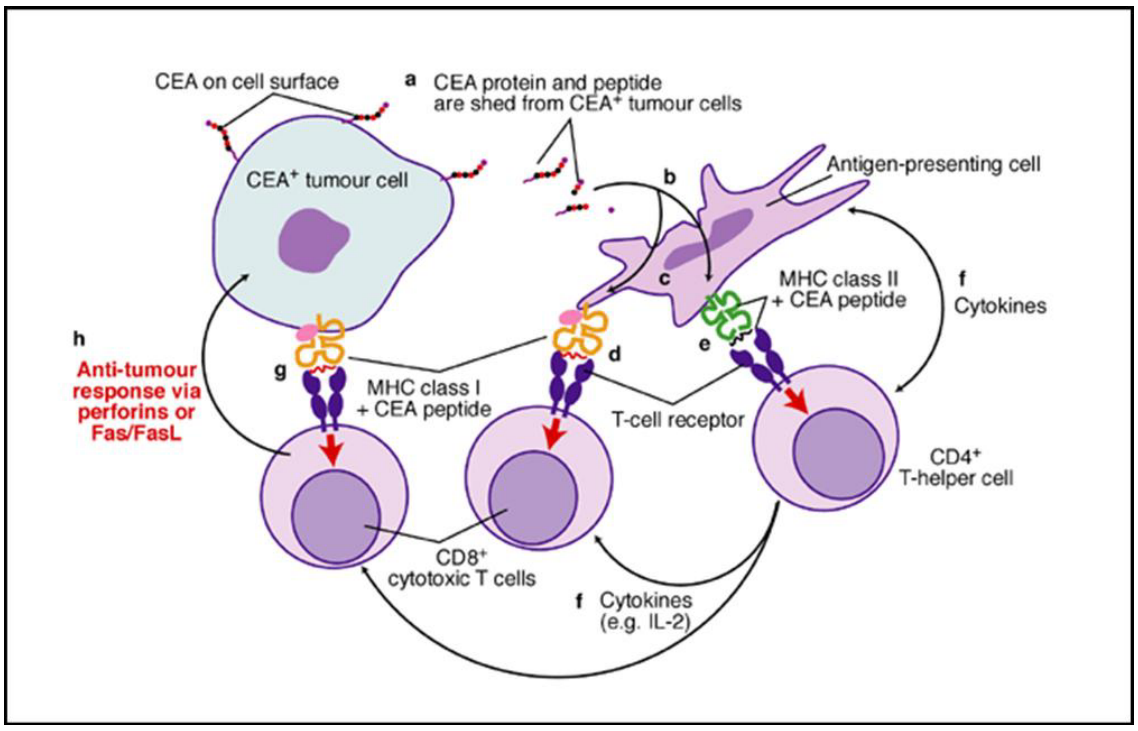


Figure 1.4: Anti-tumour responses mediated by CD4+ and CD8+ T cells.
Carcinoembryonic antigen (CEA) induces anti-tumour responses which are mediated by CD4+ and CD8+ T cells. (a) Tumour-derived CEA proteins and peptides are shed and (b) taken up by antigen-presenting cells (c) which process the antigen into small peptidic fragments (d) and presents these peptides to CD8+ T cells via their MHC class I molecules (e) and to CD4+ T cells via their MHC class II molecules. (f) This leads to the release of cytokines such as IL-2 which further activated CD8+ T cells. (g) Activated CD8+ T cells (CTL) target antigen-bearing tumour cells (h) by releasing perforins for tumour cell lysis or by inducing apoptosis of the tumour cells via Fas/FasL receptors. Taken from Hörig *et al.*, 2000.

1.1.5 Tumour evasion: immunological tolerance to cancer cells

The fact that patients are diagnosed everyday with growing malignant tumours demonstrates that the immune system of these cancer patients has somehow failed to detect tumour cells and generate an efficient immune response to kill them before the cancer is diagnosed. The tumour can bypass detection and elimination by the immune system as

a consequence of either deficiencies associated with the immune system or resourceful escape mechanisms from the tumour cells.

Down-regulation of cell surface molecules is an important mechanism for tumour cells to become less immunogenic. Tumour cells, especially metastatic cells and those found in cancers with an unfavourable prognosis, often show down-regulation or even loss of expression of MHC class I molecules (Gudmundsdottir *et al.*, 2000). Tumour cells having less MHC class I molecules on their cell surface increase their chances of avoiding detection by CTLs and to bias the immune system towards a humoral immune response, which in most studies has been shown to be incapable of destroying cancer cells. The decrease or loss of expression of MHC class I molecules, detected in approximately 16 to 50% of human solid tumours (Khong and Restifo, 2002), can be due to mutations (beta2-microglobulin, alpha chain, transcription factors), post-translational modifications (glycosylation), inhibitory mechanisms in their transport system or viruses (Benitez *et al.*, 1998; Garcia-Lora *et al.*, 2003; Paschen *et al.*, 2003). Several studies with murine models have shown the importance of a normal or high level (transfected) of expression of MHC class I molecules to induce tumour regression (Tanaka *et al.*, 1985; Ali *et al.*, 2002; Ahmad *et al.*, 2004). An event closely related to the down-regulation of MHC class I molecules is the processing of TAA which can be disregulated by mutations affecting genes involved in immunoproteasome processing (LMP2, LMP7) or genes involved in the transport of TAA; thus the tumour cell presents a reduced amount of antigenic peptides to cells of the immune system; hence favouring its escape from immunosurveillance (Lehmann *et al.*, 2000).

The decrease of expression of B7 co-stimulatory molecules can also benefit tumour development. *In vivo* and *in vitro* transfection studies have shown that a strong immune response by CTL, as well as tumour regression can be achieved when tumour cells express high levels of B7-1 molecules (Bueler and Mulligan, 1996; Heuer *et al.*, 1996). On the other hand, expression or up-regulation of other molecules on the tumour cell surface can be advantageous to the tumour. For example, the receptor FasL, a member of the Tumour Necrosis Factor (TNF)-receptor family, when binding to its receptor Fas (present on lymphocytes), activates a death signaling pathway leading to apoptosis of the immune cells (Hahne *et al.*, 1996; O'Connell *et al.*, 1999). Moreover, it has been shown that tumour cells can become resistant to cytotoxic cell-mediated lysis by decreasing their capability of binding perforins (Lehmann *et al.*, 2000), producing the serine protease inhibitor PI9 which acts against the pro-apoptotic molecule granzyme B (Medema *et al.*, 2001) or producing the anti-apoptotic transcription factor NF-kappaB in response to stress induced by chemotherapeutic drugs (Baldwin, 2001).

Other mechanisms of tumour escape are associated with the immune system itself. Immunocompromised individuals, such as HIV-infected patients or individuals having a weak immune system like elderly people, do not have fully effective defenses against cancer and thus are at a higher risk of developing a tumour. The CTL associated molecule 4 (CTLA-4) binds B7 molecules with a higher affinity than the receptor CD28, causing a decrease in proliferation of activated T lymphocytes (Carreno *et al.*, 2000). Also, the secretion of immunosuppressive cytokines at the tumour site by different types of T lymphocytes (CD4+ Th2 cells, Treg cells) can auto-inactivate these cells. For instance, in 1998, Asselin-Paturel and colleagues reported the production of IL-10 by CD4+ Th2 cells in non-small cell lung tumours (Asselin-Paturel *et al.*, 1998). Also, Treg cells can inhibit the activity of CTL by releasing the cytokines IL-10 and/or Tumour Growth Factor (TGF) (Von Boehmer, 2005). Several studies have demonstrated that the activity of Treg cells at the tumour site is detrimental to the antigen-specific anti-tumour immune response, thus beneficial to the development of the tumour (Curiel *et al.*, 2004; Kawaida *et al.*, 2005). Moreover, in 2002, Maher and colleagues discussed the process of Activation-Induced Cell Death (AICD) and its role in immune privilege. AICD controls immune system homeostasis by inducing apoptosis through Fas-FasL interactions, however tumour cells resistant to the effect of AICD can exploit this by expressing FasL on their cell surface and "counterattacking" T cells (Maher *et al.*, 2002). In conclusion, the list of all the described mechanisms of tumour escape here is non-exhaustive and cancer researchers are trying to gain a better understanding of the complexity of the processes involved in the confrontation between tumour cells and immune cells, in an attempt to develop more targeted anti-tumour therapies such as cancer vaccines.

1.1.6 Cancer vaccines: boosting the immune system to break immune tolerance to cancer

For many years, researchers have exploited the capability of the immune system to detect antigens and destroy antigen-bearing tumour cells in order to develop therapeutic cancer vaccines to boost immunosurveillance. Two preventive vaccines for cancers caused by viruses (human papillomavirus, hepatitis B virus) have been approved by the Food and Drug Administration (FDA) however, to date, no therapeutic vaccine for cancer has been authorised. These vaccines are only available to patients through clinical trials, which are essential to thoroughly test the safety and efficacy of the treatments. Where possible, vaccines are studied using animal models prior to clinical trials in humans and the lack of approved cancer vaccines for immunotherapies shows the challenging nature of the design of cancer vaccines for treatment compared to those used for prevention (Rosenberg *et al.*, 2004). Therapeutic cancer vaccines must demonstrate two required criteria for validation: they must induce specific immune responses targeting the tumour and these responses must "defeat" the mechanisms employed by tumour cells to escape

recognition. Many tumour antigens have been discovered, as described earlier, and their T-cell epitopes represent a source of potential new targets for inclusion in cancer vaccines. Cancer immunotherapy comprises three main categories which will be discussed: active immunotherapy, adoptive immunotherapy and passive immunotherapy.

1.1.6.1 Active immunotherapy

Active immunotherapy consists of several types of vaccines which will be considered: whole cell vaccines, heat shock protein vaccines, peptide vaccines, dendritic cell vaccines, DNA vaccines and viral vector based vaccines.

Whole cell vaccines were the first type of therapeutic cancer vaccine studied. They contain tumour cells taken from the cancer patient (autologous) or tumour cells from another cancer patient (allogeneic) which have been irradiated prior to injection into the patient together with an immunologic adjuvant. Before administration, these cells can be engineered to express a small molecule (hapten), designed to elicit an immune response when attached to a protein (Berd, 2001) or, to enhance T cell activity, the tumour cells can be cultured in cytokine-supplemented media such as IL-2, IL-6 or Granulocyte Monocyte-Colony Stimulating Factor (GM-CSF) (Zhang *et al.*, 2005) or transfected to express co-stimulatory molecules and cytokines (Ali *et al.*, 2000). This type of vaccine has proved to be partially effective, with generally 10-30% of cancer patients presenting a regressive tumour following immunotherapy (Weber *et al.*, 2000). For example, in clinical trials, 20 to 30% of the cancer (melanoma or kidney) patients injected with autologous tumour cells genetically modified to produce diverse cytokines (IL-2, IL-7, IFN-gamma or GM-CSF) have shown partial or complete response to the treatment (Parmiani *et al.*, 2000).

Heat shock protein vaccines consist of a mixture of adjuvants that are naturally produced by mammals, called heat shock proteins (HSP96, HSP70, HSP90, HSP110, grp170), which chaperone antigenic peptides to APC and thus improve the generation of anti-tumour immune responses (Srivastava *et al.*, 1998; Binder *et al.*, 2004). In clinical trials using autologous HSP96 vaccines, these adjuvants were proven to be non-toxic and to induce anti-tumour specific CTL responses in approximately 50% of the cancer patients (Janetzki *et al.*, 2000).

Peptides vaccines are developed according to identified immunogenic epitopes derived from TAA which are presented by APC to CTL in the context of MHC class I molecules. The first clinical trials using peptide vaccines were performed on melanoma patients using differentiation antigens (MelanA/MART-1, gp100, tyrosinase) or CT antigens (MAGE-1, MAGE-3). For example, a vaccine based on a MAGE-3 derived-peptide allowed the complete regression of melanomas in 28% of the patients treated, whereas

only 11% patients treated with a vaccine based on a MAGE-1 derived-peptide show remission (Marchand *et al.*, 1999) and a vaccine composed of MART-1 derived peptides, administrated with Incomplete Freund's Adjuvant (IFA), induced a peptide-specific CTL response, but a clinical response was only observed when IL-2 was co-administrated with this peptide vaccine (Rosenberg *et al.*, 1998). Also, a vaccine combining the CT antigen NY-ESO-1, considered the most immunogenic tumour antigen discovered to date, and GM-CSF promotes a peptide-specific CTL response and consequently regression of some metastases in 60% of the melanoma patients tested (Jager *et al.*, 2000). Not only melanoma but other cancers have been targeted with Breakpoint cluster region/Abelson (BCR/ABL)-derived, HER2-derived and Ras-derived peptides for chronic myeloid leukaemia (CML) (Pinilla-Ibartz *et al.*, 2000), breast cancer (Knutson *et al.*, 2001) and pancreatic adenocarcinoma (Gjertsen *et al.*, 2001), respectively. The choice of adjuvant and the dosage given to patients are important criteria to consider when designing peptide vaccines. Moreover, many immunological studies focus on the generation of immune responses through the enhanced activity of CD4+ T cells directed towards MHC class II peptides and not just CTL (Rojas *et al.*, 2005; Mathieu *et al.*, 2007) and some vaccines are composed of both MHC class I and MHC class II peptides (Zeng *et al.*, 2002). Other studies have focused on developing vaccine delivery systems using nano-particles (Kalkanidis *et al.*, 2006), CpG oligonucleotides as adjuvant (Haining *et al.*, 2008) or liposomes (Altin and Parish, 2006). During inflammatory responses and chronic autoimmune diseases, endogenous epitopes from released self antigens can re-direct existing immune responses whereby these epitopes become the focus targets of the immune attack, this detrimental phenomenon is called epitope (determinant) spreading (Vanderlugt *et al.*, 1998). Epitope spreading was shown to be an important process for a peptide vaccine strategy to be successful. Epitope spreading occurs when the T cell response, following vaccination of patients, is not specific to the peptide from the vaccine but targets also multiple epitopes derived from the TAA present in the vaccine or even a completely different TAA. This phenomenon was observed in breast cancer patients that were given a single HER2/neu peptide vaccine (Mittendorf *et al.*, 2006). Another important factor in the development of an efficient peptide vaccine is CTL avidity which is dependent on the concentration of the antigen. If the dose of peptide in the vaccine is too high, this may cause the inhibition of CTL proliferation (Alexander-Miller *et al.*, 1996). High avidity CTL recognise the antigen-derived peptide at low doses whereas low avidity CTL only recognises the peptide when present at high density (Snyder *et al.*, 2003). Peptide-based vaccines present many advantages compared to other vaccines, such as the low cost and simplicity of production as well as easy monitoring of immune responses. Also, using rapid molecular or immunohistochemical techniques, patients likely to be responsive to the treatment, that is to say who express the TAA of interest, can be selected.

Dendritic cell vaccines are based on the administration to cancer patients of autologous DC that have been cultured *in vitro* with cytokines such as GM-CSF or IL-4 (Siena *et al.*, 1995) and in the presence of antigenic peptides (Takahashi *et al.*, 2003), tumour cell lysates (Tamir *et al.*, 2007) in order to specifically activate T cell responses against the tumour cells. It has been reported that 50% of melanoma patients treated with autologous DC loaded with either MART-1 (Palucka *et al.*, 2006), tyrosinase (Schreurs *et al.*, 2000) or gp100-derived peptides (Linette *et al.*, 2005) have tested positive for antigen-specific CTL responses. Similarly, in another clinical trial, some melanoma patients in advanced stage IV of the disease showed specific CTL responses and tumour regressions following injection of autologous DC loaded with a MAGE-3 peptide (Godelaine *et al.*, 2003). Engineered dendritic cells include those transfected with mRNA isolated from tumour cells (Kyte and Gaudernack, 2006) or with vectors encoding TAA cDNA sequences (Chan *et al.*, 2004) and a cytokine cDNA sequence (Fu *et al.*, 2008) or those modified to produce DC maturating/activating agents such as CD40L (Lapteva *et al.*, 2007) or IL-12 (Vujanovic *et al.*, 2006).

DNA vaccines are not based on peptides derived from tumour antigens but on the whole sequence of the TAA. This has the advantage of presenting the full spectrum of antigenic MHC class I and class II molecules from one TAA. The DNA vaccines can also present CpG oligonucleotide sequences, cytokine sequences or co-stimulatory sequences to enhance the immunogenic potential of the targeted antigen-presenting cells (Leitner *et al.*, 1999). In animals, the DNA can be injected intramuscularly with a syringe and then transported in the blood to other parts of the body where it will mostly induce a Th1-orientated immune response. Alternatively the DNA can be introduced subcutaneously with a gene gun when coated on gold bullets. It is then taken up by dendritic cells of the epidermis, which migrate to the regional lymph nodes and induce primarily a Th2-orientated immune response (Weiss *et al.*, 2002). Unfortunately, studies to date have shown only poor immune responses generated following DNA immunisation. However, the prime-boost strategy (Woodland, 2004) has already shown promising results, for example with the use of electroporation (Buchan *et al.*, 2005). Finally, viral vector vaccines for cancer immunotherapy have considerable potential due to infectious mechanisms that produce the "danger signals" required to overcome tumour tolerance. Many different types of viral vectors have been used in animal models and in clinical trials such as Herpes simplex viruses, retroviruses, lentiviruses, poxviruses, adenoviruses and alphaviruses. Examples of recombinant viral vectors include those encoding TAA (PSA, CEA) or others encoding cytokines (GM-CSF) (Aarts *et al.*, 2002). Many studies aiming to develop viral vector-based immunotherapies have used disabled infectious single cycle-Herpes simplex virus (DISC-HSV) in murine tumour models (Rees *et al.*, 2002; Ahmad *et al.*, 2005). Ali and colleagues have shown regression of the tumour in up to 70% of mice which have been injected directly into the tumour with a DISC/HSV/GM-CSF vec-

tor (Ali *et al.*, 2002; Ali *et al.*, 2004). The limitations of viral vector vaccines are the patient's own antibodies directed against viruses that the immune system has previously encountered. The viral antigens, being more immunogenic than the TAA of interest, are recognised and targetted by existing antibodies. It is unlikely that any given viral vector could be injected more than once because of its ability to elicite neutralising antibody.

1.1.6.2 Adoptive immunotherapy

Adoptive immunotherapy consists of the isolation of autologous or allogeneic tumour infiltrating cells, supporting their proliferation *in vitro* and their re-introduction in the patient in large numbers in order to enhance the anti-tumour activity of immune cells. Allogeneic adoptive immunotherapy has proven very useful in treating leukaemias (June, 2007) whereas autologous adoptive immunotherapy is more applicable to solid tumours such as melanoma; for instance, up to 50% of patients treated with lymphoablative chemotherapy followed by vaccination with polyclonal T cell lines composed of tumour specific CD4+ and CD8+ T cells have shown anti-tumour responses (Dudley *et al.*, 2002; Dudley *et al.*, 2005). Finally, autologous T cells can be transfected *in vitro* with TAA-specific TCR genes such as MART-1 (Benlalam *et al.*, 2007), gp100 (Morgan *et al.*, 2003) or NY-ESO-1 (Kronig *et al.*, 2009) and form *in vivo* an important population of T cells specific to the TAA of interest leading to potentially more substantial clinical responses in cancer patients and regression of their tumours.

1.1.6.3 Passive immunotherapy

Passive immunotherapy is based on the use of tumour antigen-specific monoclonal antibodies (mAb). Monoclonal antibodies are monospecific, that is to say they will only recognise one specific antigen and are produced by clones derived from one single parent B cell. Importantly, many mAbs have been approved by the FDA as targeted cancer therapeutics. In 2007, the FDA approved, among others, the humanised mAb Cetuximab (produced by Imclone) targeting the EGFR in colorectal cancer, the humanised mAb Bevacizumab (produced by Genentech) targeting the Vascular Endothelial Growth Factor (VEGF) also for use in colorectal cancer and the humanised mAb Trastuzumab (produced by Genentech) targeting HER2 in breast cancer. Researchers are studying highly expressed tumour cell surface proteins in order to design novel mAbs to target these TAA in diverse cancers. The mechanisms of action of mAb either involves antibody opsonisation, which is the process of specific binding of the mAb to its surface tumour antigen that signal to phagocytic cells to ingest and destroy the target tumour cell, or the binding of mAb to growth factor receptors present on the surface of some tumour cells, which blocks the activity of these receptors (required for the survival of these tumours) and triggers the death of the tumour cell by apoptosis (Viani *et al.*, 2007). Studies are currently being undertaken to develop novel mAb conjugated to another molecule in order to enhance

the efficacy of conventional mAb. For instance, two antibody-radioisotype combinations, ^{90}Y-Ibritumumab tiuxetan and ^{131}I-Tositumomab, have been approved to date (Sharkey *et al.*, 2009). Other molecules used for labelling mAb include various toxins such as the anti-B4-blocked ricin targeting CD5 (Kreitman, 2001) or cytotoxic drugs such as gemtuzumab ozogamicin which has been approved for cancer therapy (Ricart and Tolcher, 2007). Another type of antibody-based vaccines is based on the use of anti-idiotypic (Id) mAb. These antibodies possess three-dimensional immunogenic regions (idiotopes) which mimic the structures and functions of tumour antigens and thus serve as substitute antigens in the induction of specific antitumour immunity in active immunotherapy. The murine anti-Id mAb 6D12 mimics a specific epitope of HER2 (Mohanty *et al.*, 2007) and the anti-Id mAb 1E10 has induced anti-tumour response in models of primary breast cancer and melanoma (Diaz *et al.*, 2009).

1.1.7 Cancer stem cells

It is now well recognised that tumour cells arise from normal cells that have increasingly acquired neoplastic phenotypes due to the accumulation of unrepaired changes involving the activation of proto-oncogenes into oncogenes and the inactivation of tumour suppressor genes. The course of tumour progression follows the process of successive clonal expansions in accordance to a Darwinian evolution of favouring the proliferation of daughter cells that have inherited an advantageous mutation. However, the increasing knowledge on stem cells and cancer stem cells challenges the whole model of multi-step tumour progression. Stem cells' main characteristics are their state of differentiation (less differentiated than non stem cell progeny), their self-renewing capability (some descendant cells remain stem cells), their ability to migrate to other tissues and their important unlimited ability to proliferate. According to the claims of many studies, a tumour appears to be composed of very few tumour stem cells and of a majority divided between progenitor cells (intermediates) and fully differentiated descendants which have limited proliferative potential *in vivo*. Therefore, it has been hypothesised that the tumour progression model actually takes place within the minority population of tumour stem cells. This theory is supported by several facts: the self-renewal capability of stem cells offers an extended window of time for mutations to take place; the turnover rate of fully differentiated cells driven by the Hayflick limit which explains that fully differentiated cells cannot divide indefinitely due to the telomeres reaching a critical length (Hayflick, 1965) and that it takes at least several months for an accumulation of mutations leading to a malignant phenotype; cancer cells that have been found tumorigenic present on their cell surface proteins that are characteristic of the normal stem cells of the tissue where the tumour originated and finally fast growing tumours contain more stem cell-like cells than slow growing tumours.

The first evidence of the existence of tumour stem cells came from a study on Acute Myelogenous Leukemia (AML) cells (Bonnet and Dick, 1997). Then, in 2003, Al-Hajj and colleagues prepared breast cancer cells from tumours and, using fluorescent-activated cell sorting (FACS), have enriched a subpopulation of cells, representing 12% of the whole cell population, that were cell surface antigens CD24 low and CD44 high. Afterwards, several hundred of these isolated cells successfully seeded new tumours following injection in immunocompromised mice, whereas approximately 20,000 cells from the major cell population failed to do so (Al-Hajj *et al.*, 2003). It became evident that only very few cells within a tumour were tumourigenic and the rest consisted of cells having limited capability for *in vivo* proliferation. Comparable results to those obtained in the study of Al-Hajj and colleagues have since been obtained in various tumours such as brain tumours (Singh *et al.*, 2003) and prostate tumours (Collins *et al.*, 2005). In 2005, Collins and colleagues identified and characterised a subpopulation of tumorigenic prostate cancer stem cells isolated from human prostate tumours which are $CD44^{+}$/alpha$_2$beta$_1$integrin^{High}/$CD133^{+}$, self-renewing and able to give rise to differentiated cells. These cells were isolated using the mentioned markers and similar techniques that are used in the isolation of normal prostate stem cells such as colony forming assays, immunofluorescence staining of target cells and invasion activity assay in Matrigel (Collins *et al.*, 2005). More recently and using similar methods, Engelmann and colleagues have isolated from the breast cancer cell line MCF-7 a breast cancer stem cell-like subpopulation that expressed the tumour antigen MUC1, which led them to reach the conclusion that these breast cancer stem cells would be good targets for MUC1-directed immunotherapy (Engelmann *et al.*, 2008). Indeed, cancer stem cells represent interesting targets for anti-cancer therapies. Researchers aim to develop novel drugs that specifically target and induce the death of cancer stem cells, which are essential for the proliferation of the solid tumour, thus regression of the tumour would be gradually observed and the risk of metastasis and recurrence of the disease would be reduced. Other methods for cancer therapy would include promoting differentiation of the cancer stem cells, inducing removal of cancer stem cells and targeting cancer stem cell-specific antigens. The discovery of the latter is likely to be important in order to develop novel polyvalent anti-tumour vaccine strategies (Morrison *et al.*, 2008).

1.2 Breast cancer

1.2.1 Breast cancer statistics

Breast cancer is a family of diseases with a high profile. The World Health Organisation stated that breast cancer accounted for 519,000 deaths worldwide in 2004, classifying it in fifth place in the list of types of cancer responsible for the overall cancer mortality (http://www.who.int, 2009). Although breast cancer is a disease associated with women

of all ages, carcinoma of the male breast can occur and accounts for less than 1% of all cases diagnosed. Consequently, little information is available in the literature on research studies involving male patients. According to the latest statistics given by Cancer Research UK, breast cancer is the most common cancer affecting women in the UK, with 45,660 new cases diagnosed in 2005. Breast cancer is the most commonly diagnosed cancer in British young women (under the age of 35 years old), with very few cases of patients being in their teens or early twenties. However, the majority of women diagnosed each year in the UK are post-menopausal. Current estimations suggest that the lifetime risk for women to have this disease is 1 in 9 and 11,990 women died in 2007 from breast cancer in the UK alone. The five-year survival rate is 80% nowadays and by comparison only half of the women diagnosed with breast cancer in the seventies survived the disease beyond five years (http://info.cancerresearchuk.org, 2009).

As shown in Figure 1.5, the incidence rate for age-standardised British females has increased by 57% between 1981 and 2005 while the mortality rates for the same group have decreased by 36% in the last 20 years. The number of breast cancer cases diagnosed is higher in the developed world than in Asia and Africa (http://info.cancerresearchuk.org, 2009). The number of patients diagnosed with breast cancer has been steadily increasing not only in the UK but in other economically developed countries as well, and if this pattern continues globally it has been estimated that by 2010 there will be approximately 1.5 million new cases (Parkin *et al.*, 2005). This increase can be mainly attributed to improved and increased screening and an increasingly aging population, along with research into earlier disease detection. Furthermore, improvements in detection and treatment of breast cancer have meant that the age-standardised mortality rate has been steadily declining since 1989 in the economically developed countries. However, despite this progress, large numbers of women are still either being diagnosed with late stage disease or go on to develop metastases and end up dying from their disease.

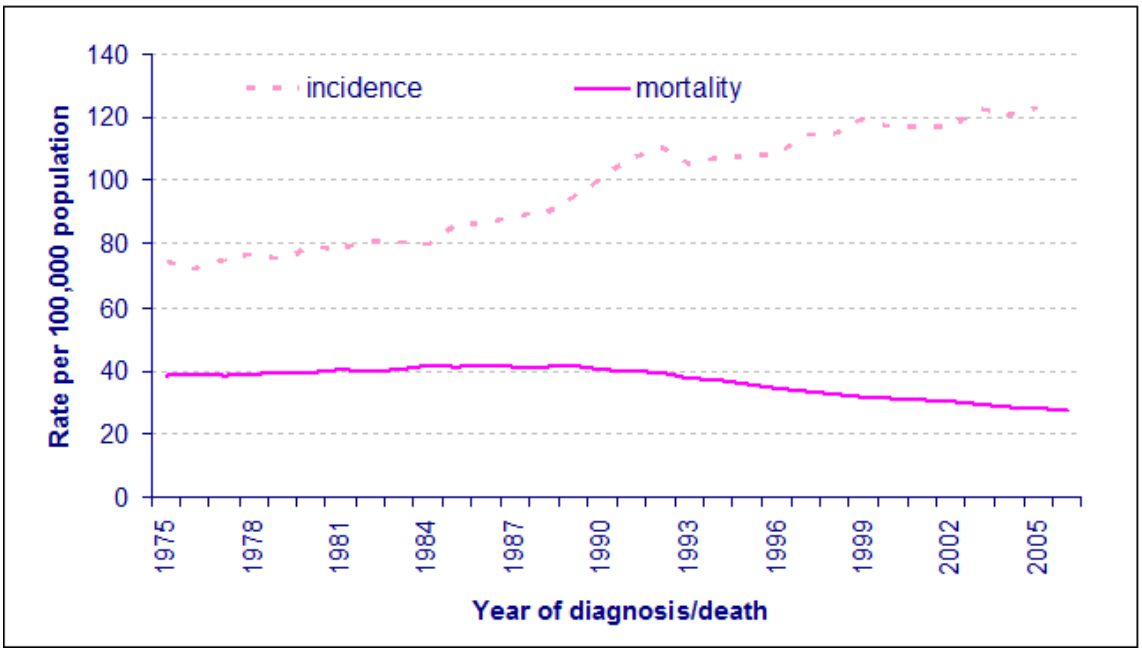


Figure 1.5: Age standardized (European) incidence and mortality rates, female breast cancer, Great Britain, 1975-2004.
Taken from Cancer Research UK (2009) at http://www.cancerresearchuk.org.

1.2.2 Epidemiology of breast cancer

As previously mentioned, the majority of breast cancer cases are diagnosed in women over 50 years old, this indicates that the risk of developing breast cancer increases with age. Routine screening of breast cancer within the UK is achieved by undertaking mammograms on women aged 50 to 70 years old as this age group has shown the highest number of cases of breast cancer diagnosed. Also, women who carry genetic predispositions or are exposed to other risk factors such as those outline in Table **??** have a higher risk of developing breast cancer. Risk factors associated with the family history of breast cancer in relatives, especially first degree of relationship and multiple cases on one side of the family, are strongly linked to an increased risk of developing breast cancer (Evans and Howell, 2007). In developed countries, these hereditary factors only make up to a quarter of all factors to breast cancer susceptibility (Key *et al.*, 2001), the remaining three-quarters are environmental, physiological and lifestyle-related. Prolonged exposure to the hormones oestrogen and progesterone linked to early menstruations, late menopause, oral contraceptive pills and hormone replacement therapies have been strongly associated with an increased risk of developing breast cancer (Travis and Key, 2003).

Other physiological factors that have been linked to breast cancer development include previous breast cancer (Chen *et al.*, 1999), certain types of benign (nonproliferative) breast lumps (Byrne *et al.*, 2000), high body mass index (BMI), high mammographic breast density (Vachon *et al.*, 2007) and taller individuals (van den Brandt *et al.*, 2000).

Higher BMI is associated with lower breast cancer incidence for pre-menopausal women, however higher BMI and weight gain before the menopause strongly increase the risk of developing breast cancer in post-menopausal women who don't use postmenopausal hormones (Huang *et al.*, 1997). Weight gain can be linked to hormonal changes or diet. Another dietary risk factor for breast cancer is excessive alcohol consumption (Holmes and Willett, 2004). Exercise is another lifestyle risk factor being investigated (Leitzmann *et al.*, 2008). An important environmental factor that increases the risk of breast cancer is the long-term exposure to radiation (Berrington de Gonzalez and Darby, 2004). On the other hand, reproductive risk factors (linked to hormonal mechanisms) that have a protective effect against the development of breast cancer include higher BMI in pre-menopausal women, early first pregnancy, number of children and breast-feeding (Evans and Howell, 2007). Furthermore, vitamin D, calcium, isothiocyanates, soy and marine n-3 fatty acids are among other dietary factors which appear to be beneficial in the prevention of breast cancer through an increase in the protective mechanism of lipid peroxidation-induced apoptosis (Gago-Dominguez *et al.*, 2007). Phyto-oestrogens might also be useful in the chemoprevention and treatment of breast cancer (Limer and Speirs, 2004).

1.2.3 Etiology of breast cancer

In 5 to 10% of all breast cancer cases, patients diagnosed have inherited from their parents one or several gene mutations that predispose to the disease. The genes BReast CAncer susceptibility 1 and 2 (BRCA1 and BRCA2), p53 and the Cowden disease gene PTEN/MMAC1 are tumour suppressor genes and their ineffective mutated forms have been linked to a high risk of developing breast cancer (Ellisen and Haber, 1998). BRCA1 and BRCA2 are "high risk" breast cancer susceptibility genes, identified in the 1990's (Miki *et al.*, 1994; Wooster *et al.*, 1995). These genes are involved in repairing DNA damage whereas their mutated versions are responsible for about 80% of the familial breast cancers and 5-6% of all breast cancers (Greene, 1997). Most cases of familial breast cancer actually carry inherited defects in at least one of the "moderate risk" or "low risk" genes. CHEK2, a cell cycle checkpoint regulator and putative tumour suppressor gene, has been identified as a moderate-risk gene and carrying an inherited alteration in this gene may double the risk of a woman developing breast cancer (CHEK2 Breast Cancer Case-Control Consortium, 2004; Vahteristo *et al.*, 2002).

Gene	Cancer syndrome	Associated tumours
BRCA1	Breast/ovarian predisposition	Breast, ovarian, bowel, prostate
BRCA2	Breast/ovarian predisposition	Breast (including male), ovarian, prostate, pancreatic
TP53	Li Fraumeni syndrome	Childhood sarcoma, brain, leukaemia, adrenocortical carcinoma, early-onset breast
PTEN	Cowden's syndrome	Breast, gastrointestinal, thyroid (benign and malignant)
STK11/LKB1	Peutz-Jeghers syndrome	Breast, gastrointestinal, pancreatic, ovarian
ATM	Ataxia telengiectasia	Non-Hodgkin lymphoma, ovarian, breast (in heterozygote carriers)
RAS		Sarcomas, carcinomas, leukemias, breast
Androgen receptor		Male breast
Neurofibromatosis 1		Nerve, brain, breast
hMSH2, hMLH1		Colon, skin, stomach, breast

Table 1.2: Rare familial cancer syndromes and genes associated with breast cancer susceptibility.
Taken from Laversin *et al.*, 2008.

The majority of breast cancers develops sporadically (no familial history), often due to somatic mutations or incorrectly regulated genes in the breast cells (Lerebours and Lidereau, 2002; Ross *et al.*, 2003; Widschwendter and Jones, 2002). Over-expression of oncogenic epidermal growth factor receptors (EGFR, HER2, ERBB3, ERBB4) and intracellular signaling molecules (c-Src, h-Ras) and inactivation of tumour suppressor genes such as cell cycle regulators (RB1, p53) are frequently found in sporadic cases of breast cancer. Also, over-expression of the EMSY gene, whose protein acts on BRCA2 protein, has been identified in many sporadic breast cancers and demonstrates the possibility of interlinked breast cancer pathways between the inherited and sporadic forms of the disease (Hughes-Davies *et al.*, 2003). As a conclusion, the predisposition to breast cancer for a woman is determined by the inheritance of different penetrance gene variants which may weaken the cell's defence mechanisms against carcinogens and the accumulation of unrepaired somatic mutations.

1.2.4 Pathological types and molecular classification of breast cancer

1.2.4.1 Pathological types

Breast tissue is a complex structure formed by an outer myoepithelial/basal layer, an inner luminal epithelial layer and non-epithelial components (fibroblasts, lymphocytes, endothelial cells, adipocytes, neurons and myocytes). In the breast, cancer develops from cells in the glands that produce milk or ducts that deliver milk to the nipples. Can-

cer Research UK provides an important source of information on the different types of breast cancer which will be looked at in this section (http://www.cancerhelp.org.uk, 2009). There are five main histological types of breast cancer: ductal carcinoma in situ (DCIS), lobular carcinoma in situ (LCIS, 10% cases diagnosed, more common in premenopausal women), invasive ductal breast cancer (70-80% cases diagnosed) (Robison *et al.*, 2004), invasive lobular breast cancer and inflammatory breast cancer (1-2% cases diagnosed). From the original site, breast cancer cells can eventually metastasise to other parts of the body, the most common being bone (Theriault and Hortobagyi, 1992). Other rare types of breast cancer include Paget's disease (1-2% cases diagnosed), breast cancer in men, medullary breast cancer (5% cases diagnosed), mucinous breast cancer (2% cases diagnosed) and tubular breast cancer (1% cases diagnosed). DCIS and LCIS are characterised by some cells inside the ducts and the lobes respectively that have become cancerous but have not invaded the surrounding healthy breast tissue. However, in some cases, if these carcinoma *in situ* are left untreated they might develop into invasive breast carcinoma in the future. The factors influencing the transition of breast carcinoma *in situ* to invasive breast carcinoma are still mainly unknown, however it appears that changes in the microenvironment play an important role in the process (Schnitt, 2009). In inflammatory breast cancer, the cancerous cells have invaded lymph ducts present in the breast which then can not drain properly excess tissue fluid, creating a localised inflammation. Figure 1.6 shows the current model of tumorigenesis in the breast tissue.

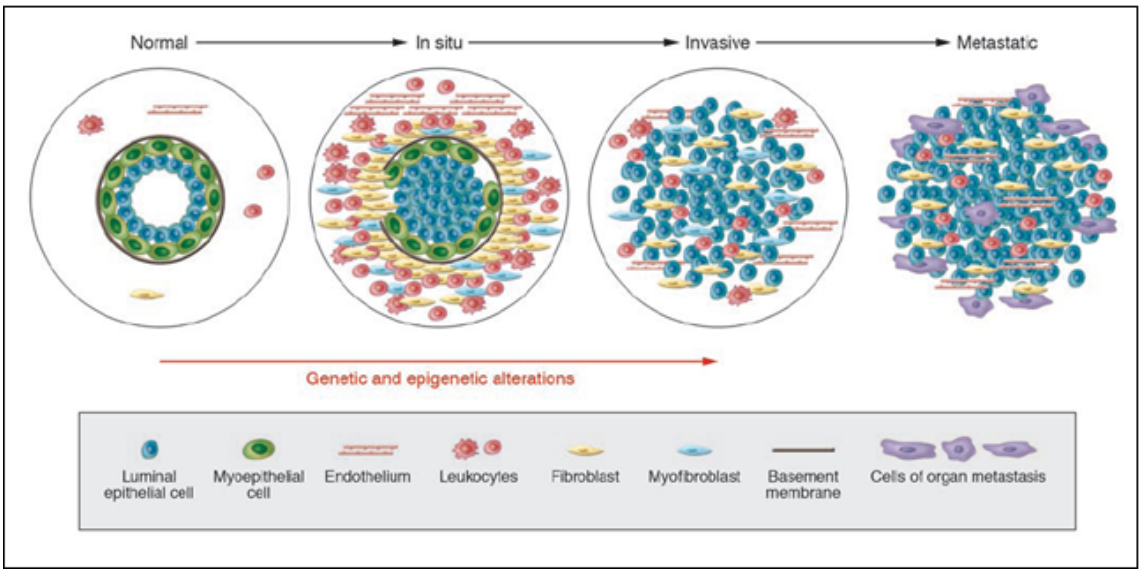


Figure 1.6: Hypothetical model of breast tumour progression.
A duct of the breast tissue is made of a basement membrane and layers of luminal epithelial cells and myoepithelial cells. The stroma surrounding the duct is composed of various endothelial cells, fibroblasts, myofibroblasts and leukocytes. Ductal carcinoma in situ (DCIS) comprises of genetically and phenotypically altered myoepithelial cells which have decreased in number and a degraded basement membrane. Furthermore, cell populations such as stromal fibroblasts, myofibroblasts, endothelial cells and lymphocytes have increased in number. In the case of invasive breast carcinoma, the basement membrane as well as the myoepithelial cell population have been lost and the malignant tumour cells have invaded the surrounding healthy tissue and potentially metastasised to other tissues and organs. Taken from Polyak (2007).

1.2.4.2 Molecular profiling of breast cancer

Breast tumours exhibit a characteristic cellular and molecular heterogeneity (Sorlie *et al.*, 2001) and it is globally accepted that the molecular differences among cancers are responsible for the different clinical courses in patients with histologically similar tumours (Pusztai *et al.*, 2007). An increasingly more accurate molecular classification of breast tumours into subtypes has been developed since the beginning of the century through the extensive use of DNA microarrays (Perou *et al.*, 2000; Sorlie *et al.*, 2001; Sorlie *et al.*, 2003). The definitive 5 clinically relevant molecular subgroups to date are listed in Table 1.3. In 2005, using tissue microarray (TMA) technology and hierarchical clustering methodology with 1,076 cases of invasive breast cancer, Abd El-Rehim and colleagues have uncovered 5 groups with distinct patterns of protein expression from a large panel of well-characterised commercially available biomarkers which illustrated the biological heterogeneity of breast cancer. Further analysis using multiple layer perceptron (MLP)-artificial neural network (ANN) analysis identified 2 large groups based on their expression of luminal epithelial cell phenotypic characteristics, hormone receptor positivity, absence of basal epithelial phenotype characteristics and lack of HER2 protein

overexpression. Two other groups were defined by high HER2 positivity and negative or weak hormone receptors expression. The fifth group was identified by strong basal epithelial characteristics, p53 positivity, absence of hormone receptors and weak to low luminal epithelial cytokeratin expression. According to the authors, this classification, which is not yet accepted in clinical practice, provides information for revision of the current traditional classification systems for breast cancer (Abd El-Rehim *et al.*, 2005). In 2008, Meijnen and colleagues have successfully categorised DCIS into two main groups and five subgroups using six published molecular markers defining breast cancer subtypes and a histological approach based on immunohistochemistry (Meijnen *et al.*, 2008).

Subgroup	ER	PR	HER2	Other characteristics
Luminal A	+	+	-	45% of breast cancer cases, good prognosis
Luminal B	+	+	-	Ki67≥10%, poor prognosis
HER2+	-	-	+	aggressive tumours
Basal	-	-	-	15% of breast cancer cases, aggressive tumours
Normal breast-like				

Table 1.3: The five clinically relevant molecular subgroups of breast cancer.
In the case of HER2, ”-” means ”not over-expressed” and ”+” means ”over-expressed”.

1.2.5 Detection of breast cancer

1.2.5.1 Breast self-examination

Breast awareness is the non-invasive, physical self-examination of the breasts in order to detect any symptoms of breast cancer such as a lump. This has proven effective in detecting tumour mass and is advised to all women, especially those not included in the UK breast cancer screening programme. Upon observation of breast abnormalities, women have an appointment with their General Practitioner (GP) who will firstly physically examine the breasts then recommend further clinical tests if judged necessary.

1.2.5.2 Imaging

In the UK, national breast cancer screening applies to all women aged 50 to 70 years old. These women, who are in an age group considered at higher risk of developing breast cancer, are invited every 3 years to undertake an X-ray of the breasts (mammogram). By 2012, the age group for the screening programme will have changed to 47-73 years old (http://www.cancerhelp.org.uk, 2009). Mammograms are also recommended to women of any age showing breast cancer symptoms identified by their GPs. Breast cancer can be detected at an early stage with mammograms. DCIS is nowadays more frequently detectable in the breast due to an increasing use of mammography that allows identification of typical microcalcification patterns (Kessar *et al.*, 2002) whereas *in situ* or invasive lobular carcinomas cannot be detected on a mammogram. Lobular carcinomas are often

defined by a thickened part of the breast tissue instead of a defined lump, making lobular carcinomas more difficult and longer to diagnose than ductal carcinomas. Most women who have no obvious lump in their breast have what is termed carcinoma *in situ* because tumour cells have not yet proliferated in the surrounding healthy tissue. Like many other medical tests, mammograms are not 100% accurate and false positives (calcification due to non-cancerous changes in the tissue) as well as false negatives occasionally occur.

A breast ultrasound scan is another imaging method for breast cancer screening. Younger women, who have denser breast tissues which is not recommended for mammography, can be screened for breast cancer using other imaging methods for example a breast ultrasound scan and an Magnetic Resonance Imaging (MRI) scan. Breast ultrasound scans are also used to assess if a detected lump is solid (tumour) or contains fluid (cyst) and MRI scans are also used to determine the surgical options for women diagnosed with lobular breast carcinoma. Knutson and Steiner have reviewed the usefulness of non-invasive methods traditionally used for breast cancer screening and reported that MRI scans have high false-positive rates and are costly; therefore they should not be used for general screening but as an alternative test if mammography failed (Knutson and Steiner, 2007).

Mammograms, ultrasound scans and MRI scans are non-invasive screening methods and according to the results of these tests, removal of cells or tissues for microscopic pathological examination of biopsy material might be carried out afterwards. The types of biopsy include, from the less to the more invasive methods, needle aspiration, needle (core) biopsy and surgical biopsy. Traditionally, nipple duct fluid and periductal needle aspiration are examined cytologically for the diagnosis of breast cancer (King *et al.*, 1983). LCIS is usually diagnosed with definite proof by biopsy and cytological analysis.

For some years, screening programmes have raised concerns about overdiagnosing breast carcinoma *in situ*, but not invasive breast carcinoma. Detecting breast carcinoma *in situ* that, if left untouched, was not going to develop into an invasive disease or detecting false positive exposes women to overtreatment, with the associated potential trauma and toxic side effects.

1.2.5.3 Genetic testing

Patients with a family history of breast cancer with known mutations in the genes BRCA1 and BRCA2 can be screened for these mutations to determine if they may benefit from tailored screening and prevention methods, such as prophylactic mastectomy (complete removal of the breast), which can reduce by 90% the incidence of breast cancer and by 81% the risk of death by the disease in women who have a high-risk family history (Hartmann *et al.*, 1999). Direct sequencing (DS) distinguishes carriers from noncarriers

however this method is expensive, time-consuming and subject to false negatives (Palma *et al.*, 2006). In high-risk families with no known inherited mutations, individuals with mutated BRCA genes can be also identified by this genetic test. However no other genetic breast cancer susceptibility tests are currently available to patients with a familial history of breast cancer. Research is focusing on the discovery of more frequently found moderate-risk and low-risk susceptibility genes in order to develop novel multiple-gene screening test that would be clinically available in the future for patients with or without a family history of the disease (Thompson *et al.*, 2008).

1.2.6 Diagnosis and prognosis of breast cancer

Diagnosis and prognosis are two very important interventions in breast cancer management and, as stated by Hu and colleagues (2005), early diagnosis of breast cancer is far more important than any therapy.

1.2.6.1 Diagnosis and staging

Once the cancer has been detected in the breast, it is staged using the globally recognised and standardised TNM cancer staging system which was primarily developed by Pierre Denoix between 1943 and 1952. Staging is very important in the decision-making process for best breast cancer treatment regime. This takes into account the histological characteristics of the malignant cells such as size of the tumour, spread to lymph nodes and spread to other organs. There are four main stages (http://www.cancerhelp.org.uk, 2009): stage I (tumour up to 2cm, no lymph nodes affected and no evidence of spread beyond the breast), stage II (tumour between 2cm and 5cm and/or lymph nodes in armpit affected and no evidence of spread beyond armpit), stage III (tumour more than 5cm, lymph nodes in armpit affected but no evidence of spread beyond armpit) and stage IV (tumour of any size, lymph nodes in armpit often affected and where the cancer has spread to other parts of the body). In addition, tumour cells are traditionally tested for hormonal receptors (ER, PR) and the growth factor receptor HER2 to determine if patients may benefit from tailored hormone therapies or molecular therapies, which will be discussed later on.

1.2.6.2 Prognosis

Galea and colleagues, in 1982, developed and further validated the Nottingham prognostic index (NPI) for prognosis for breast cancer patients who show a primary, operable cancer (Galea *et al.*, 1992). It is used in Europe and is based on histopathological features including tumour size, histological grade and lymph node status (Yu *et al.*, 2004). The aggressiveness of the invasive breast cancers is indicated by determining the histological grade of the tumour with the Elston and Ellis method which assesses tubule formation,

mitochondrial index and nuclear pleomorphism (variability in size and shape of cells and their nuclei) (Elston and Ellis, 2002). Grades are classified into three types: grade 1 (low grade, slow growing cells), grade 2 (intermediate grade) and grade 3 (high grade, fast growing cells). High grade tumour cells are more aggressively proliferating than low grade tumour cells and patients with grade 3 tumour cells are given a worse prognosis. DCIS cells can also be graded to determine the likeliness of the localised disease to become invasive in the future.

Furthermore, the age of the patient, the tumour grade and hormone receptor status are traditional characteristics that have been taken into account and incorporated into a program that is available online for the prognosis of breast cancer (www.adjuvantonline.com). Adjuvant online can predict 10 year recurrence of the cancer and overall survival of breast cancer patients by considering the predicted response to adjuvant chemotherapy (Ravdin *et al.*, 2001). Grading the tumour is useful to predict the breast cancer patient's outcome, thus determining the best treatment options according to the aggressiveness of the disease. However, traditional prognostic means remain an indication of the patient's outcome and are not 100% reliable.

1.2.7 Therapies for breast cancer

Following the process of detection, diagnosis and prognosis of the breast cancer, clinicians will determine the best options to offer a tailored treatment to the patient. Each of the five different histological types of breast cancer requires a specific mix of treatments.

Most women diagnosed with LCIS will not develop a malignant breast cancer so patients will usually not receive any treatments, not even surgery. They will be given regular monitoring in order to watch for an eventual progression of the disease into an invasive type that would then require an early intervention. Patients diagnosed with DCIS have a greater chance of developing invasive carcinoma than those diagnosed with LCIS (Fentiman, 2001) so the affected tissue is removed either by mastectomy (followed by breast reconstruction) or, in most cases, by conservative surgery (wide local excision) if the area affected is not too large. Surgery is usually followed by a course of radiotherapy in order to kill any remaining abnormal cells. If the DCIS or LCIS cells are tested positive for hormone receptors, a course of hormone therapy could also be given to the patient to try and prevent the formation of invasive breast cancer. Anastrozole is a hormone therapy currently used in a trial (IBIS 2) to assess its efficacy in preventing the development of breast cancer in women which have a higher than average risk, especially those diagnosed with LCIS or DCIS. Another study called GLACIER aims at the identification of genes in blood samples that increases the risk for women to develop LCIS and that could be used in a screening test for healthy women (http://www.cancerhelp.org.uk/trials, 2009).

Invasive ductal and lobular breast cancers are treated the same way, tailored according to the results of the diagnostic and prognostic tests of the individual: surgery usually followed by radiotherapy and chemotherapy or surgery usually followed by the less toxic hormone therapies if the tumour tested positive for ER/PR or molecular therapies if the tumour tested positive for HER2/EGFR. Alternatively, patients may be subjected to a combination of all these treatments. Inflammatory breast cancer is caused by lymph channels that are blocked in the breast, therefore neoadjuvant chemotherapy is given to reduce the swelling before surgery (local excision or mastectomy) is performed. In regards to late stage invasive breast cancers which have metastasised to other organs (not just lymph nodes), patients are considered almost incurable and therapies are carried out with the intention to extend the patient's life by decreasing tumour burden and palliating the symptoms. All breast cancer patients with any tumour type will be given follow up appointments after their course of treatment has ended to be monitored for their response to therapies and for early detection if recurrence of the disease was to occur. Five years post-treatment was shown to be the critical time period for relapse and is therefore used in assessing survival rates of breast cancer.

Oestrogen, essential for the development of the breast tissues, can promote breast cancer so a course of hormone therapy can be added for patients who have tumour cells testing positive for a high number of oestrogen receptors. Tamoxifen is a compound that competes with oestrogen for the oestrogen receptor thus inhibiting the effect of oestrogen as a growth factor for cancer cells (Ward, 1973). According to many trials, tamoxifen greatly diminishes the risk of a recurrent cancer and improves the overall survival for women by 10 years (Wishart *et al.*, 2002). The decline of the breast cancer age-standardised mortality rate at the end of the last century is partly due to the continuous extensive distribution of tamoxifen. Other hormonal therapies include progestogens (artificial progesterone) and aromatase inhibitors (given to postmenopausal women only), both block the production of oestrogen. For women with ER/PR negative tumours who cannot receive hormonal therapy, a course of chemotherapy would be advised for them. Chemotherapy can also be suggested to premenopausal women who have ER positive tumours. A combination of drugs is usually given during the course of the treatment and it includes the use of an anthracycline (DNA intercalating agent causing free radical damage of the ribose thus preventing cell division).

Approximately 25% of patients diagnosed with breast cancer present a high level of HER2 receptors on their tumours. HER2 positive tumours are ER and PR negative and present a more aggressive phenotype than the HER2 negative ones therefore patients have a poorer prognosis. HER2 positive patients can be treated with trastuzumab or lapatinib. Trastuzumab (Herceptin^{TM}), manufactured by Genentech, is a humanised monoclonal antibody directed against the growth factor receptor HER2. Lapatinib (Tykerb^{TM}), manu-

factued by GlaxoSmithKline, is a small-molecule inhibitor of the growth factor receptors HER2 and EGFR. Trastuzumab and lapatinib, both molecular therapies, were the only two targeted breast cancer therapeutics approved in 2007 and, to date, no immunotherapy for breast cancer has been approved. A UK phase I clinical trial is currently being conducted using a novel DNA-based vaccine called polyHER2neu for immunotherapy of metastasised breast cancer, with the aim of activating the immune system to recognise and kill the antigen-bearing breast cancer cells (http://www.cancerhelp.org.uk/trials, 2009).

Breast cancer patients diagnosed with triple negative tumours (ER negative, PR negative, HER2 not overexpressed) have a really poor prognosis since these will not respond to homornal therapies (tamoxifen, aromatase inhibitors) or targeted therapies (trastuzumab, lapatinib) and can only be treated with chemotherapies. Many studies are currently focusing on the molecular characteristics of the tumour cells of this group of patients in order to develop alternative targeted therapies (Irvin and Carey, 2008).

1.3 Towards personalised management of breast cancer

For the last 20 years, major breakthroughs in breast cancer research have resulted in improvements in the management of the disease which in turn has led to a steadily declining female age-standardised mortality rate (currently 34%, Cancer Research UK, Figure 1.5) and a steadily increasing female age-standardised incidence rate (13%, Cancer Research UK, Figure 1.5). However, despite this important progress, a late stage diagnosis or an unforseen development of metastases still ultimately claim the lives of large numbers of women. Early diagnosis is key to survival and is more significant than any treatment (Hu *et al.*, 2005). Up to 70% of breast cancer patients will have a recurrent disease following treatment (Chatterjee and Zetter, 2005) and Cancer Research UK communicated that out of all female breast cancer cases diagnosed in the UK in 2006 around 20% would not survive beyond five years (Cancer Research UK, 2009). These statistics highlight the limitations of accepted prognostic tools in predicting recurrence of the disease. Consequently, in a significant number of breast cancer cases, women are over-treated as they are given aggressive therapies in which associated side effects actually outweigh benefits (Chia *et al.*, 2001). Also, the major draw backs of a number of current therapies, such as chemotherapy and hormonal therapy, are undeniably the inefficacy and toxicity of treatments.

As previously mentioned, the majority of breast cancers arise sporadically from mutated or inadequately regulated genes in the breast cells. Breast cancer is described in the literature as a clinically heterogeneous and complex disease and, in the last ten years, researchers have used high-throughput technologies such as DNA microarrays to refine the classification of subtypes of breast tumours according to their molecular profiles (Perou

et al., 2000; Sorlie *et al.*, 2001; Sorlie *et al.*, 2003). The study of the differences in gene expression allows the identification of novel biomarkers which provide alternative, more reliable factors for the diagnosis, prognosis and therapy of breast cancer. The ultimate goal is the development of therapy-specific tests for the noninvasive screening of breast cancer patients at the clinic in order to provide an effective, personalised management of the disease.

A biomarker can be DNA-based, RNA-based or protein-based, used on its own or in a panel with others. The utility of the marker, the magnitude of effects (none, weak, moderate, strong), the analysis of the reliability of the marker, any technical issues (assay), any analytical issues (cutoff points, test/validation sets and multivariate analysis) and any trial design issues (appropriate patient population) are considered in the process of approval of any new breast cancer biomarkers for clinical use (Henry and Hayes, 2006). Very few biomarkers for breast cancer have been clinically validated so far, mainly due to low number of samples studied, inconsistent results and the long time duration between the discovery and validation processes (Ludwig and Weinstein, 2005).

For the last ten years or so, high throughput "omic" technologies such as genomics and proteomics have been extensively used in the identification of breast cancer specific biomarkers. Thanks to the sequencing of the human genome, completed at the beginning of the century (Venter *et al.*, 2001), large amounts of information have been obtained and are readily available in public databases. Microarray technology is an automated high-throughput technique that uses this information. The array can be a small chip (glass or silicon) or microscopic beads on which thousands of different DNA oligonucleotides or protein molecules (probes) have been fixed in an orderly approach at specific locations and whole human genome chips are commonly used nowadays. In DNA microarrays, the target sample (cDNA or cRNA from control or test samples) is hybridised to the probes, the interaction between the two is measured using fluorescence-based detection then the large amount of data generated is statistically analysed using bioinformatics. Researchers use DNA microarrays to quickly study the changes in expression levels between tumour and normal counterpart samples from breast cancer patients or between treated and control cells with the aim to discover novel biomarkers. Protein binding microarrays are useful in the study of protein interactions and for the discovery of novel drug targets and antibody microarrays are widely used to identify proteins in a given cell lysate.
Recent techniques to study the proteome also include protein mass spectrometry. It is used to compare the levels of protein expression between samples from healthy and cancer patients with the objective of discovering breast cancer associated biomarkers for diagnosis, prognosis and prediction of treatment outcome. The samples used can include serum, urine and tissue lysates. Ideally, lysates are produced from tissues that have been laser microdissected beforehand in the case of heterogeneous tumours like the breast

ones. Laser-capture microdissection is the process of isolating a pure, morphologically distinct cell population from a tissue section which is naturally made of heterogeneous types of cells (Emmert-Buck *et al.*, 1996); thus allowing a more targeted and accurate analysis of the samples by proteomic or genomic means. MALDI (matrix-assisted laser desorption/ionization mass spectrometry) and SELDI (surface-enhanced laser desorption/ionization mass spectrometry) technologies have been extensively used for biomarker discovery in the last few years (Laronga and Drake, 2007; Garrisi *et al.*, 2008; Matharoo-Ball *et al.*, 2007). The protein or peptide (tryptically digested proteins) profiles, combined with downstream bioinformatics analysis using relevant links and database interrogation, allows ions of high importance to be identified (Lancashire *et al*., 2009).

An expressed sequence tag (EST) is a short nucleotide sequence produced from randomly selected cDNA clones (Adams *et al.*, 1991) and a gene is usually represented by several ESTs. The identification of human-specific ESTs has increased dramatically following the completion of the human genome project. The number of ESTs of a given gene in a given library reflects the level of expression of the gene in the tissue or cell line which the library derived from. It is possible to mine databases, where the expression patterns of the studied genes in normal tissues and tumour tissues can be found, with the single objective of identifying *in silico* breast cancer related genes.

In 1995, Sahin and colleagues have developed SEREX, a novel serological approach to identify TAA that are able to induce humoral immune responses, which is based on the immunoscreening of cDNA expression libraries from tumour samples with sera from cancer patients (Sahin *et al.*, 1995). This technique has been successfully used in the identification of TAA from various tumour types. For example, NY-ESO-1 was discovered using an oesophageal cancer cDNA library (Jager *et al.*, 2000) and, in our laboratory, T21 and MTA1 were identified using prostate cancer cDNA libraries (Assudani *et al.*, 2006; Miles *et al.*, 2007; Li *et al.*, 2008).

1.3.1 Breast cancer biomarkers for screening and diagnosis

Hereditary breast cancer biomarkers have previously been discussed (section 1.2.3 Etiology of breast cancer). Around 80% of all the familial breast cancers and around 5 to 6% of all breast cancer cases diagnosed will show mutations in the BRCA1 and BRCA2 genes (Greene, 1997). Therefore, women with a family history of breast cancer will be advised by clinicians to be screened for these mutations by direct sequencing (DS). The test is unfortunately expensive, time-consuming (Palma *et al.*, 2006) and can give false negative results. To date, no other genetic breast cancer susceptibility tests using other high-risk genes for familial breast cancer are clinically available. However, several recent studies have reported that women diagnosed with breast cancer, who have a strong family history of breast cancer not linked to mutated BRCA1 or BRCA2, have mutations in ataxia telangiectasia mutated (ATM), BRCA1 interacting protein C-terminal helicase 1 (BRIP1) and partner and localiser of BRCA2 (PALB2) significantly more frequently

than the control women who were healthy, with or without a strong cancer family history (Renwick *et al.*, 2006; Seal *et al.*, 2006; Rahman *et al.*, 2006; Erkko *et al.*, 2007). ATM is a cell cycle checkpoint kinase which regulates the activity of many proteins including the tumour suppressor protein p53 and the checkpoint-like protein CHK2. Walsh and colleagues stated that around 5% of breast cancer patients with a strong family history of breast cancer, who have been tested positive for wild type BRCA1 and BRCA2, have mutations in the p53 and CHK2 genes (Walsh *et al.*, 2006). Therefore, it has been acknowledged that genetic testing for the mutations of these genes should be made available to women with a strong family history of breast cancer who have tested negative for mutations in BRCA1 and BRCA2.

To date, no biomarker has been validated for clinical use in breast cancer risk assessment and diagnosis, and some candidate biomarkers for screening failed the validation process as they showed a high false-positive rate. A breast cancer biomarker has to be very specific and be detectable through minimally invasive and cost effective techniques to be desirable for risk assessment and diagnosis (Ludwig and Weinstein, 2005). For the diagnosis of breast cancer, relatively non-invasive samples such as nipple duct fluid and periductal needle aspiration have been traditionally examined for abnormal cells (King *et al.*, 1983). Using patient's blood in molecular-based protocols for the detection of specific biomarkers is thought to be nowadays the best minimally invasive way to detect breast cancer (Levenson, 2007). The MUC1 glycoproteins (epitopes CA15-3, CA27-29) are components of cell-cell junctions in the breast and have been shown to be helpful in the detection of primary breast cancer (Gion *et al.*, 1999). Cytokeratins of intermediate filaments have been used successfully for testing for early breast cancer in several studies, however their lack of sensitivity meant they cannot be validated for clinical use (Eskelinen *et al.*, 1994). Mammaglobin is a protein specific to the breast tissue whose detection in serum or urine samples, measured by an ELISA assay, can indicate the presence of micrometastases in peripheral blood, lymph nodes and bone marrow of patients and could therefore be useful in the diagnosis of breast cancer (O'Brien *et al.*, 2002). Biomarkers that can also be used to distinguish between patients with early stage breast cancer and healthy individuals have been identified in serum and nipple duct fluid using SELDI (Li *et al.*, 2002; Sauter *et al.*, 2002).

1.3.2 Breast cancer biomarkers for prognosis and prediction of treatment outcome

Personalised management of breast cancer involves biomarker-guided therapy options in order to prevent unnecessary treatments and overtreatment of patients and to ensure high treatment success rate and reduced costs (Strand *et al.*, 2006).
Few biomarkers have been widely accepted and are routinely used for prediction of re-

sponse to therapy and in the prognosis of breast cancer. Some of the biomarkers which are indicative of a poor prognosis are: overexpression of HER2 (Bieche *et al.*, 1999), absence of ER (Bezwoda *et al.*, 1991), inactivated P53/PTEN (Borresen-Dale, 2003; Depowski *et al.*, 2001), RhoC (Kleer *et al.*, 2002), c-myc (Deming *et al.*, 2000), loss of expression of E-cadherin (Yoshida *et al.*, 2001), matrix metalloproteases (Egeblad and Werb, 2002), Ep-CAM (Gastl *et al.*, 2000), hTERT (Bieche *et al.*, 2000), a modified form of cyclin-E (Keyomarsi *et al.*, 2002), loss of expression of the tumour suppressor gene maspin (Maass *et al.*, 2002) and mutated BRCA1 (Robson *et al.*, 2004). Some of the biomarkers which play an important role in predicting the response to therapy are: HER2 whose expression levels can indicate the benefits of treating breast cancer patients with trastuzamab (Hunt and Keyomarsi, 2005), ER/PR whose expression makes tumours sensitive to treatments (Tamoxifen and aromatase inhibitors) (James *et al.*, 2007), PS2 which is linked to ER presence hence suggestive of a good response to hormonal therapy (Ardavanis *et al.*, 1997), Bcl-2 whose expression indicates increased survival with hormonal treatments (Yang *et al.*, 2003), MDR1 which is associated with resistance to chemotherapy (Trock *et al.*, 1997), BRCA1 whose absence means tumour cells are more sensitive to chemotherapy and the presence of which makes tumour cells more sensitive to antimicrotubule agents (James *et al.*, 2007), Glutathione S-transferase whose over-expression is associated with multi-drug resistance (Batist *et al.*, 1986) and mutated p53 which is associated with resistance to hormonal, adjuvant, neoadjuvant and combination chemotherapy (Bottini *et al.*, 2000; Daidone *et al.*, 1999; Geisler *et al.*, 2001 and Montero *et al.*, 2001). Also, high levels of the protein CA15-3 in the serum of breast cancer patients undergoing chemotherapy is usually indicative of a weak response to the treatment (Al-azawi *et al.*, 2006).

Widely available and recognised, yet expensive tests for prediction of treatment outcome and prognosis of breast cancer comprise the Mammaprint (a gene microarray-based test from Agendia BV), Oncotype DX (a quantitative-Polymerase Chain Reaction-based test from Genomic Health Inc) and Mammostrat (an immunohistochemistry-based test from Applied Genomics Inc) tests. It is thought that these microarray gene expression profilers have the same efficacy for the prediction of breast cancer outcome and response to therapies as the traditional clinical breast cancer biomarkers (Eden *et al.*, 2004). Mammaprint is used to categorise breast cancer patients (lymph-node negative, ER positive or negative tumours) who have only received surgery as treatment in a low or a high risk group for distant recurrence of the disease (van de Vijver *et al.*, 2002; van 't Veer *et al.*, 2002). Oncotype DX is used to test patients (lymph-node negative, ER positive tumours), who received surgery and tamoxifen, for prediction of cancer recurrence and to determine if they would benefit from certain types of chemotherapy (Cronin *et al.*, 2004; Paik *et al.*, 2004). Mammostrat is used to predict the risk (high, moderate or low) of breast cancer patients (lymph-node negative, ER positive tumours), who received only

surgery as treatment, to develop distant metastasis regardless of chemotherapy regimens (Ring *et al.*, 2006).

1.3.3 Breast cancer biomarkers for treatment

Hormonal therapies such as tamoxifen or aromatase inhibitors (Cristofanilli and Hortobagyi, 2002) and the monoclonal antibody trastuzumab (McKeage and Perry, 2002) are well accepted therapies as mentioned previously. The re-introduction of maspin expression (Maass *et al.*, 2002) or wild type-p53 expression (Liu *et al.*, 1995) are potential treatments for breast cancer; and EpCAM, cyclin E, the platelet-derived growth factor (PDGF) and Flk-1 (a growth factor receptor) are potential targets for immunotherapies (Ariad *et al.*, 1991; Gastl *et al.*, 2000; Hasan and Jayson, 2001; Hunt and Keyomarsi, 2005). Cancer cell growth can be inhibited in around 90% of breast cancer patients tested positive for the vitamin D receptor (nuclear steroid hormone receptor) when tumour cells are treated with the vitamin D analogue 1a,25-(OH)2D3 (Friedrich *et al.*, 1998). Wu and colleagues stated that receptor selective retinoids, which have a role in growth inhibition and induction of apoptosis, could be used in novel therapies for breast cancer patients whose tumour tested positive for these but negative for ER (Wu *et al.*, 2002). Also, ligands for the peroxisome proliferator-activated receptor (PPAR), such as the natural prostaglandin 15-deoxy-delta-12,14-prostaglandin J2 (PGJ2) and the synthetic anti-diabetic thiazolidenediones troglitazone (TGZ), could be of interest for use in novel therapies targeting breast cancer cell growth (Pignatelli *et al.*, 2001). Another receptor which could be a good candidate for therapy is the growth receptor flk-1. It has been shown that when flk-1 mRNA molecules are cleaved by ribozymes the growth rate of microvasculature endothelial cells decreases, making flk-1 a potential target in novel strategies against angiogenesis (Hasan and Jayson, 2001). In 2004, Castelli and colleagues stated that heat shock proteins which are potential biomarkers for breast cancer treatment due to their ability to chaperone tumour antigens, especially MHC peptides, could be an effective way of developing a vaccine strategy to break immune tolerance to these protein targets, thus inducing specific regression of the tumour (Castelli *et al.*, 2004).

1.3.4 Breast cancer biomarkers for monitoring the disease

The majority of deaths from breast cancer are due to the spread of metastases to other organs and the development of secondary tumours compromising the function of these organs. Therefore, it is of prime importance to detect metastasis and to monitor the disease following primary treatment. Serum or urine from breast cancer patients can be used in an ELISA assay to detect the presence of the protein mammaglobin which would indicate that breast tumour cells have spread through peripheral blood and lymph nodes to other parts of the body (O'Brien *et al.*, 2002). Ep-CAM and HER2 are also indicators of micrometastases in peripheral blood and bone marrow of breast cancer patients (Gastl *et*

al., 2000; Andersen *et al.*, 1995); and urokinase-type plasminogen activator (PAl-1 and PAl-2) and cathepsins B and L are biomarkers for breast cancer recurrence (Borstnar *et al.*, 2002; Foekens *et al.*, 1998).

1.4 Rationale and outline of the study

The development of novel cancer vaccines for immunotherapy, as well as novel diagnostic or prognostic tools, is based on the discovery and validation of TAA. For this purpose and as previously mentioned, EST database mining is one strategy thus has been used to identify potential tumour-specific antigens. Using this method, novel breast-associated genes (Breast-associated UniGene Clusters or BUC) have recently been discovered in our laboratory by limiting the search criteria to genes present in human normal breast and/or breast cancer tissues or cell lines but not in human normal tissues or organs except testis, ovary and placenta. Upon further characterisation using molecular techniques, four interesting and potential TAA, namely BUC6, BUC9, BUC10 and BUC11, were selected for further study (personal communication with Dr Li). This research project proposed to characterise the BUC genes in order to determine their potential applications in diagnosis, prognosis or therapy for cancer, particularly for breast carcinoma. Firstly, the full gene sequences of the four BUC genes had to be obtained to determine the similarities shared with published genes and a comprehensive mRNA expression analysis had been carried out in various invasive and non-invasive histological subtypes of breast cancer, benign breast tumours, breast cancer cell lines, other malignancies (melanoma, prostate cancer) and normal tissues. The gene which demonstrated promising expression patterns was then further characterised at the protein level using various immunoassays with diverse normal and tumour tissues. Furthermore, the study incorporated preliminary research into the biological functions of this gene and its potential involvement in carcinogenesis and/or tumour progression using gene expression induction experiments, gene expression knockdown experiments and microarrays. Finally, it was intended to evaluate the immunogenicity of the gene product in a HLA transgenic mouse model.

Chapter 2

Materials and Methods

2.1 General laboratory consumables and equipments

2.1.1 Reagents

Reagents were stored as per manufacturer instructions and used before the expiry date.

Culture media	**Supplier**
Dulbecco's Modified Eagle Medium (DMEM)	BioWhittaker Europe
Eagle's minimal essential medium (EMEM)	BioWhittaker Europe
RPMI 1640 medium (RPMI)	BioWhittaker Europe
McCoy's medium	Invitrogen
OPTI-MEM	Gibco Life Technologies
CTL media	Cellular Technology

Supplements to culture media	**Supplier**
Foetal calf serum (FCS)	BioWhittaker Europe
4-(2-hydroxyethyl)-1-piperazine- -ethanesulfonic acid (HEPES)	BioWhittaker Europe
Fungizone	Promega
Geneticin (G418)	Promega
L-glutamine	BioWhittaker Europe
Non essential amino-acids (NEAA)	BioWhittaker Europe
Penicillin/Streptomycin	BioWhittaker Europe
2-mercaptoethanol	Sigma

Other cell culture reagents	**Supplier**
Dimethyl sulfoxide (DMSO)	Sigma
Incomplete Freund's adjuvant (IFA)	Sigma

Lipopolysaccharide (LPS)	Sigma
Murine GM-CSF	PeproTech EC
Murine IL-2	PeproTech EC
Phosphate buffer saline (PBS)	BioWhittaker Europe
Polyinosinic polycytidylic acid (Poly I.C)	Sigma
Staphylococcus aureus enterotoxin B (SEB)	Sigma
Trypan blue	Sigma
Trypsin/Versene	BioWhittaker Europe

Chemical reagents	**Supplier**
Acetic acid	Fisher Scientific
Acetone	Acros Organics
Acetonitrile	Fisher Scientific
Acrylamide-bis	Geneflow
Agar	Oxoid Ltd
Agarose	Bioline
Alpha-chymotrypsin	Sigma
Ammonium bicarbonate	Sigma
Ammonium Persulphate	National Diagnostics
Ampicillin	Sigma
Anhydrous ethanol	Sigma
Aprotinin	Sigma
Aza-deoxycytidine (AZAC)	Sigma
BioTaq^{TM}DNA polymerase	Bioline
Bovine serum albumin (BSA)	ICN Biomedicals
Bromophenol blue sodium salt	Promega
Calcium Chloride ($CaCl_2$)	Sigma
Chloroform	BDH
Chromium-51	Amersham Biosciences
Citric acid	Sigma
Dextran sulphate	Sigma
Dithiothreitol (DTT)	Sigma
DNA ladder (1Kb plus)	Gibco Life Technologies
Deoxyribonucleotide triphosphate (dNTP)	Bioline
DPX mountant	Sigma
Embedding compound	Bright Instrument
Ethanol	BDH
Ethidium bromide	Sigma

Ethyldiamine tetraacetic acid (EDTA)	Sigma
4x ProtoGel Resolving Buffer	National Diagnostics
Fluorescent mounting media	Dako
Gills haematoxylin solution	Sigma
Glacial acetic acid	Fisher Scientific
Glucose	Sigma
Glutaraldehyde	Sigma
Glycerol	Sigma
Goat serum	Sigma
Gold microcarriers (1.0μm)	Bio-Rad
Harris haematoxylin solution	Sigma
Hydrochloric acid (HCl)	Fisher Scientific
Hydrogen peroxide (H_2O_2)	Sigma
INTERFERin^{TM}	Polyplus Transfection
Isopropanol	Sigma
Isoton	Beckman-Coulter
Liquid nitrogen	BOC
Lipofectamine 2000	Invitrogen
Marvel	Premier Brands
Magnesium chloride ($MgCl_2$)	Fischer Scientific or Promega
Methanol	Acros Organics
NANOpure Diamond distilled water	Barnstead
Nitrogen (gas)	Air Liquide
N,N,N',N'-tetramethyl-ethylenediamine (TEMED)	National Diagnostics
Octyl-beta-d-glucopyranoside (OGP)	Sigma
Orange G	Sigma
Paraformaldehyde	Sigma
Phenol/Chloroform/Isoamyl alcohol	Sigma
Phosphate buffer saline (PBS) (1X)	BioWhittaker Europe
Phosphate buffer saline (PBS) (Tablets)	Oxoid
Phusion^{TM}HF buffer	Finnzymes
Phusion^{TM}DNA polymerase	Finnzymes
Poly vinyl pyrrolidone (PVP)	Sigma
Ponceau S solution	Sigma
Potassium acetate (KOAc)	Sigma
Potassium ferricyanide ($K_3Fe(CN)_6$)	Sigma
Potassium ferrocyanide ($K_4Fe(CN)_6$)	Sigma

Propidium iodide (PI)	Sigma
Precision Plus Protein WesternC standard	Bio-Rad
ProtoGel Resolving Buffer	National Diagnostics
ProtoGel Stacking Buffer	National Diagnostics
Rabbit liver powder	Sigma
Restriction enzymes	Promega or New England Biolabs
Ribonuclease (RNase)	Qiagen
RNA-STAT 60	Biogenesis
Scintillation fluid	Packard
Silane	Sigma
Sodium azide (NaN_3)	Sigma
Sodium chloride (NaCl)	Fischer Scientific
Sodium citrate	Sigma
Sodium dodecyl sulphate (SDS)	Sigma
Sodium hydroxide (NaOH)	Fischer Scientific
Sodium phosphate dibasic (Na_2HPO_4)	Sigma
Sodium phosphate monobasic (NaH_2PO_4)	Sigma
Spermidine	Sigma
Substrate solution (for ELISA)	R&D Systems
Sulfuric acid (H_2SO_4)	Sigma
SYBR®green supermix	Bio-Rad
Trifluoroacetic acid	Acros Organics
Trichostatin A (TSA)	Sigma
Tris	Fischer Scientific
Tris-HCl (1.5M), pH 6.8 or pH 8.8	Geneflow
Tritiated (3H)-thymidine	Amersham Biosciences
Trizma base	Sigma
Trypsin Gold, Mass Spectrometry grade	Promega
Tryptone	Oxoid Ltd
Tween 20	Sigma
2-methylbutane (Isopentane)	Acros Organics
Ultra Pure 10x Tris/Glycine electrophoresis grade buffer	National Diagnostics
Ultra Pure 10x Tris/Glycine/SDS electro--phoresis grade buffer	National Diagnostics
Urea	Sigma
Vitamin E	Sigma

Xylene	Acros Organics
Yeast extracts	Oxoid Ltd
Zeocin™	Invitrogen

Immunochemical reagents	**Supplier**
Rabbit anti-BUC11 antibody	Pacific Immunology
BUC11 peptide to which Rabbit anti-BUC11 antibody has been raised	Pacific Immunology
Swine anti-rabbit Immunoglobulin G (IgG)-Fluorescein IsoThioCyanate (FITC)	DAKO
Swine anti-rabbit IgG-HorseRadish Pero--xidase (HRP)	DAKO
Goat anti-rabbit IgG biotinylated	DAKO
Polyclonal rabbit IgG isotype control	AbD serotec
Streptavidin conjugated HRP	AbD serotec
DRAQ5™	Biostatus

Plasmids	**Supplier**
pBudCE4.1 and pCRII-blunt-TOPO	Invitrogen
pMACS K^k.Tag(N)	Miltenyi Biotec

Kits	**Supplier**
ABC and DAB kits	Vector Laboratories
Cell line Nucleofector®Kit V	Amaxa
DNA isolation kit	Biological Industries
Dc Protein Assay kit	Bio-Rad
ECL Western Blotting reagents	Amersham Biosciences
M-MLV (Moloney Murine Leukemia Virus) Reverse Transcriptase kit	Promega
Mouse CD4: Dynabeads mouse CD4	Dynal
Mouse CD8: Dynabeads mouse CD8	Dynal
Mouse IFNγ Development Module	R&D Systems
ELISpot Blue color Module	R&D Systems
RNAeasy MiniKit	Qiagen
Kit for immunoprecipitation with μMACS^{TM}Protein G MicroBeads	Miltenyi Biotec
Kit for selecting cotransfected cells with pMACS vector	Miltenyi Biotec
Kit for isolating mouse leukocytes (Pan T Cell Isolation Kit)	Miltenyi Biotec
T4 ligase kit	Promega

2.1.2 Equipment

Glassware

Pyrex glassware was washed in teepol, rinsed twice in distilled water and autoclaved.

Disposable equipment and plastic-ware	**Supplier**
BD microlance 3 needles (0.5ml, 1ml)	Becton Dickenson
Bijou tubes (7ml)	Sterilin
Cell scrapers	TPP
Culture slides (8 chambers)	BD Falcon
Cryovials (1.2ml)	TPP
ELISA plates (96-well)	Costar
ELISPOT plates(96-well)	Millipore
Eppendorf tubes (0.5ml, 1.5ml)	Sarstedt
FACS tubes	Elkay
Filter tips (10μl, 200μl, 1ml)	Sarstedt
Flat bottom culture dishes(6-, 24-, 96-well)	Sarstedt
Gel blotting papers	Whatman
Hyperfilm ECL films	Amersham
Luma-Plate 96	Packard
96-well plate harvester filters	Perkin Elmer
Nitrocellulose membranes	GE Healthcare
Pasteur pipettes (plastic)	Sarstedt
PCR tubes	Axygen Scientific
Petri dishes (plastic)	Sarstedt
Real-Time PCR tubes (100μl)	Corbett Research
Round bottom culture dishes (96-well)	Sarstedt
Scalpels	Swann-Morton
Screw top tubes (15ml, 50ml)	Sarstedt
Serological pipettes (5ml, 10ml, 25ml)	Sarstedt
Syringes (10ml)	Becton Dickenson
Tefzel tubing	Bio-Rad
Tips (20μl, 200μl, 1ml)	Sarstedt
Tissue culture flasks (T25, T75, T175)	Sarstedt
Universal tubes (20ml)	Sterilin
0.2μm filters	Sartorius

Equipment	**Supplier**
Centrifuge, microcentrifuge	MSE
Class II safety cabinets	Walker
Cold room	Polysec
Confocal microscope	Leica
Cryostat (Leica CM 1900)	Leica
Distilled water system (ddH_2O)	Barnstead
Drying cabinet	Scientific Laboratory Supplies
Dynabeads separation unit	Dynal
Electrophoresis gel tanks	Bio-Rad
Flow cytometer (Coulter Epics XL.MCL)	Beckman-Coulter
Fridge	Denley
Glass cuvette for spectrophotometer	Hellma
Helios gene gun	Bio-Rad
Hypercassette	Amersham Biosciences
ImmunoSpot®Analyzer	Cellular Technology
Microscope	Nikon
Microwave	Matsui
-20°C freezer	Lec
-80°C freezer	Sanyo
Nanodrop ND1000 microspectrophotometer	Nanodrop Technologies
96-well plate harvester	Packard
96-well plate reader	Tecan
Nucleofector™	Amaxa
Orbital shaker	Stuart
Progene thermal cycler (for conventional PCR)	Techne
Power packs	Bio-Rad
ICycler iQ Multicolor Real-Time PCR Detection System	Bio-Rad
Rotor-Gene 6000	Corbett Research
Refrigerated microcentrifuge	Hettich Zentrifugen
Semi-dry transfer apparatus	Bio-Rad
Sonicator	VWR
Spectrophotometer (SP Bio)	Sanyo
37°C, 5% CO_2 incubator	Forma Scientific
Top count scintillation counter	Packard
Tubing prep station	Bio-Rad
UNO-Thermoblock	Biometra

UV spectrophotometer	Sanyo
UV transilluminator	Ultra Violet Products
Water baths	Grant instruments

2.1.3 Buffers

Buffers for bacterial cell culture:

LB broth	LB Agar
5g NaCl	5g NaCl
10g tryptone	10g tryptone
5g yeast extracts	5g yeast extracts
pH adjusted to 7.0 with NaOH	15g agar
Completed to 1,000ml with ddH_2O	pH adjusted to 7.0 with NaOH
Autoclaved	Completed to 1000ml with ddH_2O
	Autoclaved

Buffers for DNA extraction from bacterial cell culture:

GTE solution	1% SDS/0.2M NaOH solution
50mM glucose	100μl NaOH (10M)
25mM Tris-HCl (pH 8.0)	500μl 10% SDS
10mM EDTA	4.4ml ddH_2O

KOAc solution
60ml 5M KOAc
11.5ml glacial acetic acid
28.5ml ddH_2O

Buffers for tissue culture:

Trypan blue	White cell counting solution
Dilute the purchased 0.4% solution to 0.1% v/v in PBS	Dilute 100% acetic acid to a 0.6% v/v solution in PBS

Buffers for flow cytometry and confocal microscopy:

FACS buffer
0.1% w/v BSA
0.02% w/v NaN_3
1x PBS

Solution to permeabilise cells
Dilute 4% paraformaldehyde solution (kept at 4°C) to 1% v/v in PBS

Solution to fix cells
70% v/v ethanol in PBS

Buffers for immunohistochemistry:

Primary antibody solution (frozen sections)
2μg/ml rabbit anti-BUC11 antibody
5% (v/v) goat serum
Completed to final volume with PBS

Secondary antibody solution (frozen sections)
2μg/ml goat anti-rabbit IgG-biotin
1.5% (v/v) goat serum
Completed to final volume with PBS

2μg/ml Polyclonal rabbit IgG isotype control
5% (v/v) goat serum
Completed to final volume with PBS

Primary antibody solution (paraffin-embedded sections)
10μg/ml rabbit anti-BUC11 antibody
5% (v/v) goat serum
Completed to final volume with PBS

Secondary antibody solution (paraffin-embedded sections)
3.8μg/ml goat anti-rabbit IgG-biotin
1.5% (v/v) goat serum
Completed to final volume with PBS

ABC reagent (working solution)
2.5ml PBS
1 drop reagent A
1 drop reagent B

DAB reagent(working solution)
2.5ml ddH_2O
1 drop buffer
2 drops DAB
1 drop H_2O_2 30%

Citrate buffer
27ml 0.01M citric acid
123ml 0.1M sodium citrate
Completed to 1,500ml with ddH_2O
pH to 6.0

Buffers for Enzyme-Linked ImmunoSorbent Assay (ELISA):

Wash buffer
0.05% (v/v) Tween 20 in PBS

Blocking agent
3% (w/v) Marvel in PBS

Primary antibody solution
serial dilutions of rabbit anti-BUC11 antibody made in blocking agent

Secondary antibody solution
1:1000 dilution of Swine anti-rabbit IgG-HRP in blocking agent

Substrate solution
1 volume colour reagent A
1 volume colour reagent B

Stop solution
2.5M H_2SO_4

Buffers for Mass Spectrometry:

Lysis buffer for MALDI analysis
9.5M Urea
2% (w/v) DTT
1% (w/v) OGP
Aliquoted and stored at -80°C

Buffers for Enzyme-Linked ImmunoSorbent SPOT (ELISPOT):

Buffer for Pan T Cell Isolation
0.5% BSA
2mM EDTA
Completed to final volume with PBS pH 7.2 and degassed

Capture Antibody working solution
For each 96 well plate: 2 aliquots of Capture Antibody Concentrate
Completed to final volume of 10ml with PBS

Wash Buffer
0.05% Tween 20 in PBS

Blocking Buffer
1% BSA
5% sucrose
Completed to final volume in PBS

Reagent Diluent
1% BSA in PBS
pH 7.2-7.4, filtered

Detection Antibody working solution
For each 96 well plate: 2 aliquots of Detection Antibody Concentrate
Completed to final volume of 10ml with PBS

Streptavidin-AP working solution
For each 96 well plate:
167μl of Streptavidin-AP
Completed to final volume
of 10ml in Reagent Diluent

Buffers for Agarose gel electrophoresis or Western Blotting:

50x TAE
242g Tris
57.1ml glacial acetic acid
100ml 0.5M EDTA (pH 8.0)
Make up to 1000ml with ddH_2O
Dilute 1:50 in ddH_2O
for working solution

Orange G DNA loading buffer
12.5mg orange G powder
1.5ml glycerol
Completed to 5ml with ddH_2O

Ammonium Persulphate (APS)
100mg in 1ml of distilled water
Fresh frozen (-80°C) into aliquots

Reducing sample buffer
2.5 ml of 0.5M Tris-HCl pH 6.8
0.4g of 2% w/v SDS
2ml of 10% v/v Glycerol
200mg of 1% w/v DTT
0.04g of 0.2% w/v Bromophenol blue
Make up to 20ml with ddH_2O

Running buffer
100ml of Ultra Pure 10x Tris/Glycine/SDS

(0.25M Tris, 1.92M Glycine, 1%SDS)
Make up to 1000ml with ddH_2O
(0.025M Tris, 0.192M Glycine, 0.1% SDS)

Transfer buffer
100ml Ultra Pure 10x Tris/Glycine
(0.25M Tris, 1.92M Glycine)
200ml of methanol
Make up to 1000ml with ddH_2O
(0.025M Tris, 0.192M Glycine)

Tris Buffer Saline (TBS)
5x stock
12.11g of Trizma base
146.1g of NaCl
Make up to 1000ml with ddH_2O
Adjust pH to 7.5
Dilute 1:5 in ddH_2O for
working solution

TBS-Tween 20 (TBST)
500μl of Tween 20
Make up to 1000ml with TBS

TBS-Tween 20-Marvel
2.5g of Marvel
250μl of Tween 20
50ml of TBST

ECL chemiluminescence reagent
1 volume of solution A
1 volume of solution B

Primary antibody solutions
1:500 dilution of rabbit
anti-BUC11 antibody (1mg/ml)
in TBS-Tween 20-Marvel

Secondary antibody solutions
1:2,500 dilution of Swine
anti-rabbit IgG-HRP (0.76mg/ml)
in TBS-Tween 20-Marvel

1:1,000 dilution of rabbit
IgG isotype control (2mg/ml)
in TBS-Tween 20-Marvel

1:5,000 dilution of
Streptavidin conjugated HRP
(1mg/ml) in TBS-Tween 20-Marvel

2.1.4 Gels

Agarose gel:

1.5% agarose gel (small)	1.5% agarose gel (medium)
0.75g of agarose	2.25g of agarose
50ml of 1x TAE	150ml of 1x TAE
5μl of Ethidium bromide	15μl of Ethidium bromide

Protein gel:

15% resolving gel	4% stacking gel
To prepare 2 gels:	To prepare 2 gels:
7.5ml of Protogel	1.04ml of Protogel
3.74ml of resolving gel buffer	2ml of stacking gel buffer
3.5ml of ddH_2O	4.88ml of ddH_2O
150μl of APS	40μl of APS
15μl of TEMED	8μl of TEMED

2.1.5 Primary culture media

Culture media were prepared and used within one month. Complete BM-DC medium was prepared fresh, just before use.

Bone-Marrow Dendritic cell (BM-DC) medium:
500ml of RPMI
5% (v/v) of FCS
2mM of L-glutamine
10mM of HEPES
50μM of 2-mercaptoethanol
25U/ml of Penicillin/Streptomycin
0.25μg/ml of Fungizone

T cell medium:
500ml of RPMI
10% (v/v) of FCS
2mM of L-glutamine
20mM of HEPES
50μM of 2-mercaptoethanol
50U/ml of Penicillin/Streptomycin
0.25μg/ml of Fungizone

Complete BM-DC medium:
BM-DC medium
1ng/ml mGM-CSF

2.1.6 Cell lines and their media

The cell lines and their corresponding culture media used in this project are outlined in Table 2.1.

Name	Description	Media	Growth properties	Source
MDA231	Human breast adenocarcinoma	EMEM + 5% (v/v) FCS + 2mM L-glutamine + 1% (v/v) NEAA	Adherent	Dr. P. de Gremoux (Institut Curie, Paris, France)
T47D	Human breast invasive ductal carcinoma	EMEM + 5% (v/v) FCS + 2mM L-glutamine + 1% (v/v) NEAA	Adherent	ATCC (HTB-133)
MCF7	Human breast adenocarcinoma	DMEM + 10% (v/v) FCS + 2mM L-glutamine + 1% (v/v) NEAA	Adherent	ATCC (HTB-22)
BR293	Human breast carcinoma	EMEM + 10% (v/v) FCS + 2mM L-glutamine + 1% (v/v) NEAA	Adherent	Unknown
MDA468	Human breast adenocarcinoma	EMEM + 10% (v/v) FCS + 2mM L-glutamine + 1% (v/v) NEAA	Adherent	ATCC (HTB-132)
MDAP3	Human breast carcinoma	EMEM + 10% (v/v) FCS + 2mM L-glutamine + 1% (v/v) NEAA	Adherent	Unknown
SKBR3	Human breast adenocarcinoma	McCoy's medium + 10% (v/v) FCS + 2mM L-glutamine + 1% (v/v) NEAA	Adherent	ATCC (HTB-30)
MDA435	Human M14 melanoma cell line	DMEM + 10% (v/v) FCS + 2mM L-glutamine + 1% (v/v) NEAA	Adherent	Prof. I. Ellis (QMC, Nottingham, England)
FM3	Human malignant melanoma	RPMI + 10% (v/v) FCS + 2mM L-glutamine	Adherent	Dr. J. Zeuthen (University of Copenhagen, Denmark)
ALC	Murine lymphoma	RPMI + 10% (v/v) FCS + 2mM L-glutamine + 500mg/ml G418	Suspension	Dr. C. Baxevanis (Saint Savas Cancer Hospital, Athens, Greece)

Table 2.1: Cell lines and their description.
The cell line MDA435 was known as a breast cancer cell line until recently. That is why it was used as such during this study and appeared as such in Chapter 3. In their study published in 2007, Rae and colleagues concluded that the stocks of MDA435 cells are, without any doubt, all derived from the melanoma cell line M14 (Rae *et al.*, 2007).

2.2 Methods

2.2.1 Expression analysis

2.2.1.1 RNA extraction and cDNA synthesis

Dr. R. Ferris (USA) and Prof. D. Schadendorff (Germany) kindly provided the samples of head and neck carcinoma and the samples of melanoma, respectively. Dr. Aija Line (Latvia) kindly provided the samples of colon, gastric and breast carcinoma mRNA (with patient-matched normal tissue mRNA). Dr A. Gritzapis (Greece) generously provided a total of 218 frozen breast tissues which correspond to paired normal and tumour tissues from 109 breast cancer patients. The mRNA samples from normal tissues were purchased from Clontech (UK). The mRNAs from human cell lines were extracted from cells cultured at the laboratory (see Table 2.1).

Total RNAs were isolated from various cell lines and tissues using RNA STAT-60™ according to the manufacturer's protocol. For extracting RNA from frozen tissues, the tissue was grounded in liquid nitrogen in a Class II safety cabinet using a mortar and pestle to a powder which was then homogenised in 1ml of RNA STAT-60™, transferred to a 1.5ml Eppendorf tube and stored at room temperature (RT) for 5 minutes. For extracting RNA from cell lines, suspension cells were pelleted at 1,500rpm for 3 minutes and lysed in a Class II safety cabinet with 1ml RNA STAT-60™. For adherent cells, 1ml RNA STAT-60™was directly added to the culture vessel to lyse them. The homogenate was then transferred to a 1.5ml Eppendorf tube and stored at room temperature (RT) for 5 minutes. Then, under a fume hood, 0.2ml of chloroform was added to the homogenate and the tube was shaken vigorously by hand for 60 seconds and left at RT for 3 minutes. The samples were centrifuged at 14,000rpm for 15 minutes at 4°C. The colourless upper aqueous phase which contained the RNA was transferred to a fresh 1.5ml Eppendorf tube, mixed with 0.5ml of isopropanol and the tube was stored at RT for 8 minutes. The samples were centrifuged at 12,000rpm for 10 minutes at 4°C. The supernatant was discarded and the white pellet was washed with 1ml 75% ethanol. The RNA pellet was dried under the fume hood and re-suspended in double-distilled water (ddH_2O) by passing it few times through the pipette tip. The concentration and purity of the isolated RNA was measured with a UV spectrophotometer. Samples were diluted 1:100 in ddH_2O and absorbance was measured at 260nm and 280nm.

Total RNA was reverse transcribed into cDNA using M-MLV RT following manufacturer's instructions. The reverse transcription was either carried out with Oligo(dT)15 primers (Promega, UK) or random primers (Promega, UK). Random primers were the primers of choice when the cDNA was going to be used in quantitative RT-PCR assay screening several reference genes. The cDNA synthesis using Oligo(dT)15 primers was carried out as follows. In a 0.5ml Eppendorf tube, 2μg of total RNA (or all the RNA solution if less than 2μg available) was mixed with 1μl of Oligo(dT)15 primer solution and ddH_2O was added to make the final volume to 10μl. The tube was heated in a thermal block at 70°C for 5 minutes, immediately cooled on ice and briefly spun in a microcentifuge. The following mix of components was added to the tube at RT: 5μl of 5X M-MLV RT buffer (Promega, UK), 1μl of dNTP (12.5mM of each dATP, dTTP, dGTP and dCTP), 0.7μl of RNasin®ribonuclease inhibitor (Promega, UK), 1μl of M-MLV RT and 7.3μl of ddH_2O. The tube was gently mixed and incubated in a waterbath at 39.2°C for 80 minutes. The tube was then heated in a thermal block at 95°C for 5 minutes. The cDNA was cooled on ice, spun briefly and stored at -20°C. The cDNA synthesis using random primers was carried out as follows. In a 0.5ml Eppendorf tube, 2μg of total RNA (or all the RNA solution if less than 2μg available) was mixed with 0.5μl of random primer solution and ddH_2O was added to make the final volume to 10μl. The tube was heated in a thermal block at 70°C for 5 minutes, immediately cooled on ice and briefly spun in a microcen-

tifuge. The following mix of components was added to the tube at RT: 5μl of 5X M-MLV RT buffer (Promega, UK), 1μl of dNTP (12.5mM), 0.7μl of RNasin®ribonuclease inhibitor (Promega, UK), 1μl of M-MLV RT and 7.3μl of ddH_2O. The tube was gently mixed and incubated in a waterbath at 37°C for 60 minutes. The tube was heated in a thermal block at 95°C for 5 minutes and then immediately cooled on ice, spun briefly and stored at -20°C.

2.2.1.2 Primers for Polymerase Chain Reaction (PCR)

Primer sequences were either taken from literature when available or designed to cover two different exons in the sequence to eliminate amplification of trace amounts of genomic DNA in the cDNA samples where possible. Some of the primers were designed to generate PCR products of under 250 base pairs (bp) in size to optimise the quantitative RT-PCR. Primer design was carried out with the assistance of the Primer3 program (http://biotools.umassmed.edu/bioapps/primer3www.cgi).
The primers were supplied by Eurofins MWG Operon (UK). The name, the sequences, the temperature of annealing (Ta) and the extension time of all the primers used in this study are listed in Table 2.2 and Table 2.3.

Gene	Forward Primer (5' to 3')	Reverse Primer (5' to 3')	Ta/Time
GAPDH	ACCACCAACTGCTTAGCACC	CCATCCACAGTCTTCTGGGT	58°C/30s
TBP	TGCACAGGAGCCAAGAGTGAA	CACATCACAGCTCCCCACCA	56°C/30s
HPRT1	TGACACTGGCAAAACAATGCA	GGTCCTTTTCACCAGCAAGCT	55°C/30s
NY-BR-1	CGAAGTGCACAACAAGGCTA	GCTGAAGAAGCATGCCAACT	58°C/30s
BUC6	CCTTCGAGTTCCTTTTTCTGG	GGCAACACAAACTCAGAGCA	58°C/30s
BUC9	CCAGATTTTCACCGCTATGC	AGGCAAGCTCTCATCAGGAC	58°C/30s
BUC10	CACCGACGTTTAAAGGAGGA	GTCCTCTGCACCTTGGGATA	58°C/30s
BUC11	TCTTTCCCACAATCCCTGAC	CAGCTTGCCCCATGTATTTT	58°C/30s
BUC69	GAGGCCTTGCTAATTTCCTA	CTGTTGCAGTGAGCTCAAGT	58°C/30s
BUC910	GTCCTGATGAGAGCTTGCCT	CAAACTGGCCTTGATCTGGA	58°C/30s
BUC6Q(3)	CTCGAAGCCATCAATGACAA	TGAGATAATCCGCTCCTTGG	58°C/30s
BUC9Q	GCCACATGGGGTATGTTCTC	AGGCAAGCTCTCATCAGGAC	58°C/30s
BUC11 F3R3	TGCCATTCCACTGTTTTCTG	CAGCTTGCCCCATGTATTTT	58°C/30s
BUC11 F4R4	CTTGCCACCTCCCAGTAAAA	CTTGGTTTCCAGCTCTTTGC	58°C/30s
BUC11 F5R5	TCTAGTCGACCCACAATCCCTGACA	TCTATCTAGAGCCAGAAAACAGTGGA	58°C/30s

Table 2.2: Primers used for PCR of BUC genes, NY-BR-1 gene and reference genes.
The DNA polymerase used for non- or semi-quantitative RT-PCR with all these primers was the BioTaq DNA polymerase. The number of cycles used for non- or semi-quantitative RT-PCR was 30. "Ta" is the temperature of annealing in degree Celsius. "Time" is the time of the extension step during the PCR, in seconds (s). The primers "BUC11 FR3R3" were used for the semi-quantitative and quantitative RT-PCR of hCG25653 gene and designed just after the protein-coding region. The primers "BUC11 FR4R4" were also used for the quantitative RT-PCR of hCG25653 gene and designed within the protein-coding region. "BUC11 F5R5" primers were designed to flank the coding region of hCG25653 gene and used for the cloning in the mammalian expression vector pBudCE4.1. Each "BUC11 F5R5" primer has 4 common bases "TCA" at their 5' end then the forward primer has the recognition site for the restriction enzyme SalI (GˆTCGAC) and the reverse primer has the recognition site for the enzyme XbaI (TˆCTAGA).

Gene	Forward Primer (5' to 3')	Reverse Primer (5' to 3')	Ta/Time
MT2A	AAAGGGGCGTCGGACAAGTG	AACGGTCACGGTCAGGGTTG	60°C/30s
MTCH2	CTGCTGCTACCCTCATCACA	ATGCCCTCTTCCCGATAGAT	60°C/30s
RPS7	GTGGGAAGCATGTCGTCTTT	CCATCTAGTTTGACGCGGAT	60°C/30s
RPL30	GAAGACGAAAAAGTCGCTGG	CAGATTTCCTCAAAGCTGGG	60°C/30s
CRR9	TTTGTCTTTGACGGGTCCTC	AGCGGTTTATGACCATCAGG	60°C/30s
NDUFA1	CTGGCTACTGCGTACATCCA	TGCGCCTATCTCTTTCCATC	60°C/30s
COX7c	CACAACCTCTGTGGTCCGTA	GAAGGGTGTAGCAAATGCAGA	60°C/30s
MRPL11	ATTGGACAGCCCACTGTTTC	GTACATCCTGCAGGGCAAAT	60°C/30s
RPS11	GAAATGCCCCTTCACTGGTA	ACAGACATGTTCTTGTGGCG	60°C/30s
MT1F	TCCTGCAAGTGCAAAGAGTG	CACTTCTCTGACGCCCCTT	60°C/30s
LDHA	ACGTCAGCAAGAGGGAGAAA	CGCTTCCAATAACACGGTTT	60°C/30s
CYR61	CAGCTGACCAGGACTGTGAA	TGTAGAAGGGAAACGCTGCT	60°C/30s
SDC4	CCACCGAACCCAAGAAACTA	CAGTGCTGGACATTGACACC	60°C/30s
RHOA	CAGAAAAGTGGACCCCAGAA	ACTCCATGTACCCAAAAGCG	60°C/30s
PLAU	ATACAGACCATCTGCCTGCC	CACAGCATTTTGGTGGTGAC	60°C/30s
GPX1	CCAAGCTCATCACCTGGTCT	TCGATGTCAATGGTCTGGAA	60°C/30s
OAS1	CAAGCTCAAGAGCCTCATCC	TGGGCTGTGTTGAAATGTGT	59°C/30s
STAT1	AAATTCCTGGAGCAGGTTCA	TGGCCCCAGTCACTTAATC	59°C/30s

Table 2.3: Primers used for PCR of genes screened in the BUC11 gene knockdown study. "Ta" is the temperature of annealing in degree Celsius. "Time" is the time of the extension step during the PCR, in seconds (s).

2.2.1.3 Semi-quantitative Polymerase Chain Reaction

The non- or semi-quantitative RT-PCR reaction was performed in the Techne Progene PCR machine using either the BIOTAQ^{TM}DNA Polymerase or the Phusion^{TM}High-Fidelity DNA Polymerase. The RNA extracted from cell lines and tissues were pre-screened for the housekeeping gene glyceraldehyde-3-phosphate dehydrogenase (GAPDH). GAPDH was then used to normalise the mRNA levels between the samples.

For each reaction with the BIOTAQ^{TM}DNA Polymerase, thermocycling was done in a final volume of 25μl containing 1μl of template cDNA (or 1μl of ddH_2O for negative control), 2.5μl of the buffer for the polymerase, 0.75μl of MgCl2, 0.4μl of dNTPs (12.5mM of each dATP, dTTP, dGTP and dCTP), 1μl of gene-specific forward primer, 1μl of gene-specific reverse primer (20mM), 0.125μl of polymerase and 18.225μl of ddH_2O. The cycling conditions were as follows: melting step at 95°C for 3 minutes, (denaturation at 94°C for 1 minute, annealing at primer-specific Ta for 1 minute, extension at 72°C for primer-specific time) x appropriate number of cycles, final extension step at 72°C for 5 minutes and 4°C forever.

For each reaction with the Phusion^{TM}High-Fidelity DNA Polymerase, thermocycling was done in a final volume of 50μl containing 1.5μl of template cDNA (or 1.5μl of ddH_2O for negative control), 10μl of Phusion buffer HF, 1μl of dNTPs (12.5mM of each dATP, dTTP, dGTP and dCTP), 1μl of gene-specific forward primer, 1μl of gene-specific reverse primer (20mM), 0.5μl of polymerase and 35μl of ddH_2O. The cycling conditions were as follows: melting step at 98°C for 30 seconds, (denaturation at 98°C for 10 seconds, annealing at primer-specific Ta + 3°C for 20 seconds, extension at 72°C for primer-specific time) x appropriate number of cycles, final extension step at 72°C for 5 minutes and 4°C

forever.

2.2.1.4 Agarose gel electrophoresis

15μl minimum of the final PCR products were mixed with 5μl of 5X DNA loading buffer Orange G and subsequently size separated on 1.5% (m/v) agarose Tris-Acetate-EDTA (TAE) gel. The gel was prepared by repeated heating in a microwave and contained 0.0001% (v/v) of ethidium bromide for band visualisation. To approximate the size of the PCR products on electrophoresis, the samples were run along with 0.1μg of DNA ladder 1Kb Plus. The electrophoresis runs were conducted at 100 volts for 1 hour and 30 minutes to 2 hours. The PCR products were visualised under UV light.

2.2.1.5 Quantitative Polymerase Chain Reaction

The standard curve was prepared by serial dilution of cDNA from a sample (tissue or cell line) known to express well the gene of interest: undiluted sample, 1/10, 1/100, 1/1,000 and 1/10,000. The RT-Q-PCR reactions were performed in the i Cycler Q Multicolor Real Time PCR detection system (Bio-rad, USA) and later on in the study in the new Rotor-Gene 6000 (Corbett Research, UK), using iQ^{TM}SYBR®Green Supermix. Duplicates were carried out for each tissue or cell line to test. Thermocycling for each reaction was done in a final volume of 12.5μl containing 0.5μl of template or standard, 6.25μl of SYBR®Green Supermix, 0.5μl of gene-specific forward primer, 0.5μl of gene-specific reverse primer and 4.75μl of ddH_2O. In each experiment, 6 "no template" controls have been included to ensure no contamination has occurred and also to indicate the degree of amplification due to primer dimers. The cycling conditions were as follows: melting step at 95°C for 3 minutes, (denaturation at 95°C for 30 seconds, annealing at primer-specific Ta for 30 seconds, extension at 72°C for primer-specific time) x 45 cycles. Finally, a melting (dissociation) curve was acquired by slowly ramping the temperature from Ta to 95°C by 1°C increment.

The fluorescence of each sample was measured at the end of annealing step or extension step and at the end of each cycle in the case of the dissociation curve. The Ct value, which corresponds to the number of cycles at which the reaction crosses a threshold value (fluorescence exceeds the background level), was calculated by the software so as the standard curve. Following completion of the quantitative RT-PCR reaction, the Ct value for each sample and the standard curve were recorded. The standard curve method was used for the quantification of the results as follows. The standard quantity (SQ) mean for the gene of interest was divided by the SQ mean of the housekeeping gene which resulted in the normalised mRNA expression of the gene of interest in the tissues or cell lines.

Assuming that primer efficiencies are similar for the genes studied, the standard curve can be omitted from the assay, allowing more samples to be screened. The comparative Ct method is therefore used for the quantification of the results. Following completion of the quantitative RT-PCR that did not include a standard curve, the analysis of the results was carried out using the comparative Ct method, also known as the $2^{-[delta][delta]Ct}$ method, where [delta][delta]C_t=[delta]$C_{t,sample}$-[delta]$C_{t,reference}$ Here, [delta]$C_{t,sample}$ is the C_t value for any sample normalised to the housekeeping gene and [delta]$C_{t,reference}$ is the C_t value for the calibrator also normalised to the housekeeping gene. The result is given as a relative gene expression level of the gene of interest in the tissues or cell lines.

2.2.1.6 Cell lysate preparation

Cells, from cell lines cultured in the laboratory, were harvested and washed twice in ice cold PBS at 1,500rpm for 3 minutes at 4°C. Depending on the size of the pellets, cells were re-suspended in either 50μl or 100μl of ddH_2O for lysis and transferred to Eppendorf tubes. The lysates were kept on ice for 10 minutes, sonicated for 1 minute, placed on a rotating wheel in a cold room for 30 minutes, left on ice again for 30 minutes and finally centrifuged at 14,000rpm for 30 minutes at 4°C. Supernatants were transferred to fresh Eppendorf tubes and stored at -80°C until analysis (protein assay and SDS-PAGE) and further use (Western Blotting, Immunoassays, etc.).

2.2.1.7 Protein assay

The protein concentration in the lysates was measured by performing a Bio-Rad Dc protein assay in a 96-well plate according to the manufacturer's protocol. Briefly, the protein standard was made of BSA diluted in ddH_2O (buffer used for lysis of cells) in serial dilution, from 1.5mg/ml to 0.22mg/ml. 20μl of reagent S was added to each ml of reagent A. 25μl of reagent A (+S) was added to 5μl of samples (each sample run in duplicate) and 200μl of reagent B was then added to each well. The plate was wrapped in foil and the reaction was left to develop for 15 minutes at RT. Finally, the plate was read at 750nm on a Tecan 96-well plate reader.

2.2.1.8 Sodium Dodecyl Sulphate PolyAcrylamide Gel Electrophoresis (SDS-PAGE)

Protein samples were defrosted and diluted in reducing sample buffer at a ratio 3:1. The proteins were denatured by boiling the samples at 95°C for 5 minutes in a thermal block. Samples (about 30μg) were loaded to individual wells of the acrylamide gel placed in the electrophoresis chamber, along with biotinylated protein marker. The gel (100cm^2) was run at 100 volts (V) through the 4% stacking gel and at 150V through the 15% resolving gel.

2.2.1.9 Semi-Dry Western Blotting

Proteins, separated by size by SDS-PAGE, were then transferred at 13V for 30 minutes onto a nitrocellulose membrane using a semi-dry transfer unit according to the manufacturer's protocol. The membrane was then stained with Ponceau S and cut into the appropriate pieces for immunoprobing. The membrane was washed 3 times in TBS-Tween 20 to remove Ponceau S and blocked overnight in TBS-Tween 20-Marvel, on an orbital shaker in a cold room.
The peptide block was prepared by mixing the desired amount of primary antibody with 5 fold (by weight) excess of peptide, as follows. 20μg of antibody was gently mixed to 100μg of peptide, all diluted in TBS-Tween 20-Marvel to make up 500μl, and left at RT for 30 minutes to 2 hours prior adding the pre-incubated antibody/blocking peptide to 9,500μl of TBS-Tween 20-Marvel.
The membrane was washed 3 times for 15 minutes in TBS-Tween 20 at RT on an orbital shaker. The primary antibody solution (see Materials) was added to the membrane which was incubated for 3 hours on an orbital shaker at RT. Then, the membrane was washed 3 times for 15 minutes in TBS-Tween 20 at RT on an orbital shaker. The secondary antibody solution (see Materials) was added to the membrane which was incubated for 2 hours at RT on an orbital shaker. The membrane was then washed 4 times for 15 minutes in TBS-Tween 20 at RT on an orbital shaker, revealed with ECL Western Blotting kit reagents and the luminescence was detected on Hyperfilm ECL films in a dark room.

2.2.1.10 Fluorescent activated cell sorting (FACS)

The following protocol was used for staining the DNA for cell cycle analysis. Briefly, 72 hours after transfection experiment, adherent cells were trypsinised, transferred to FACS tubes and washed twice with FACS buffer for 3 minutes at 1,500rpm at 4°C. The supernatant was discarded and the pellet resuspended gently in the FACS buffer remaining in the tube. The tubes were placed on ice and 1ml ice cold 70% ethanol was mixed in order to fix the cells. Then, the cells were washed twice in FACS buffer as before and the pellet was resuspended in 1ml FACS buffer with 1μl of Rnase solution (100mg/ml) and 50μl of Propidium Iodide solution (1mg/ml). The tubes were incubated at 37°C for 20 minutes and placed on ice while performing the FACS analysis on the flow cytometer.

2.2.1.11 Enzyme-Linked ImmunoSorbent Assay (ELISA)

The following protocol corresponds to the heterologous, non-competitive, solid-phase ELISA. Briefly, BUC11 peptide (to which Rabbit anti-BUC11 antibody has been raised, 1mg/ml stock solution) was diluted in PBS in serial dilution. 200μl of peptide dilutions was added to the 96-well plate. The dilutions used were: 1:50 (20μg/ml), 1:100 (10μg/ml), 1:200 (5μg/ml), 1:400 (2.5μg/ml), 1:800 (1.25μg/ml), 1:1,600 (0.625μg/ml)

and 1:3,200 (0.3125μg/ml). A blank without any peptide was added. The plate was wrapped in foil and incubated at 4°C overnight. Then, the antigen solution was discarded by inversion in the sink and the wells washed 3 times with wash buffer. The plate was dried by inversion onto absorbent paper. 200μl of blocking agent was added to each well and the plate, wrapped in foil, was incubated for 3 hours at RT to block excess protein binding sites. The wells were washed as before 3 times with wash buffer and dried again. Rabbit anti-BUC11 antibody (3mg/ml, stock solution) was diluted in blocking agent in serial dilution. 100μl of antibody dilutions was added to the 96-well plate (duplicate wells). The dilutions used were: 1:2,500 (400ng/ml), 1:3,333 (300ng/ml), 1:5,000 (200ng/ml), 1:10,000 (100ng), 1:20,000 (50ng/ml) and 1:100,000 (10ng/ml). The plate was wrapped in foil and incubated for 2 hours at RT. The wells were washed 4 times with wash buffer as before. 100μl of swine anti-rabbit IgG-HRP solution was added to the wells and the plate, wrapped in foil, was incubated for 2 hours at RT. The wells were washed 4 times with wash buffer as before, 100μl of substrate solution was added to them and the plate, wrapped in foil, was incubated for 10 to 30 minutes, until an appropriate blue colour developed. The reaction was stopped with 50μl of stop solution. Finally, the plate was read at 450nm on a Tecan 96-well plate reader.

A second ELISA was subsequently performed, using the same protocol but with the following dilutions of antibody and peptide. The peptide dilutions used were: 1:500 (2μg/ml), 1:1,000 (1μg/ml), 1:2,000 (0.5μg/ml), 1:4,000 (0.25μg/ml), 1:8,000 (0.125μg/ml), 1:16,000 (0.0625μg/ml) and 1:32,000 (0.03125μg/ml). The antibody dilutions used were: 1:25,000 (40ng/ml), 1:100,000 (10ng/ml), 1:200,000 (5ng/ml), 1:400,000 (2.5ng/ml), 1:800,000 (1.25ng/ml) and 1:1,600,000 (0.63ng/ml).

2.2.1.12 Immunofluorescence

Briefly, 300μl of culture media containing 1.7x10^4 adherent cells were plated in each chamber of 8-chamber slides and incubated overnight at 37°C. The media was removed and the cells were washed twice with cold FACS buffer very gently, by pipetting on the side of the chamber. The cells were permeabilised with 100μl/chamber of 1% (w/v) paraformaldehyde for 10 minutes at RT, washed twice as before and then fixed with 100μl/chamber of ice cold 70% ethanol for 10 minutes at RT. Rabbit anti-BUC11 antibody was used at a concentration of 0.005μg of antibody per 1μl of FACS buffer. The peptide block was prepared by mixing 5μg of antibody to 100μg of peptide, all diluted in FACS buffer. 100μl of primary antibody solution and peptide block solution was added to designated chambers and the slides were incubated on ice for 45 minutes. Then, cells were washed twice as before and 100μl of swine anti-rabbit IgG-FITC (0.005μg of antibody per 1μl of FACS buffer) was added to the appropriate chambers. The slides were incubated on ice, in the dark, for 30 minutes. Cells were then washed 3 times with FACS buffer and the slides were dried at 37°C. Before the slides were dried, in some experi-

ments, DRAQ5^{TM}(far-red fluorescent DNA dye) was added to the corresponding chambers as follow: 125μl of PBS mixed with 0.25μl of neat DRAQ5^{TM}was added to each chamber, incubated at 37°C for 3 minutes then removed. The slides were mounted with fluorescent mounting media and left at 4°C for 3 hours to overnight. Finally, slides were studied under a confocal microscope. The laser (blue light) of the microscope excites the FITC molecules which consequently emit a visible green light.

2.2.1.13 Immunohistochemistry

Paraffin-embedded multiple normal and tumour tissue microarrays (BN241, BN242 and BR242) were purchased from US Biomax. Frozen melanoma tissue sections were kindly provided by Prof. D. Schadendorff (Germany). Frozen prostate tissues were kindly provided by clinicians at Nottingham City Hospital (Nottingham, UK) and were sectioned in our laboratory. Frozen sections were stored at -80°C and the paraffin-embedded tissue microarrays at RT, until immunohistochemistry assays were performed.

Dr A. Gritzapis (Greece) kindly provided a total of 218 frozen breast tissues which correspond to paired normal and tumour tissues from 109 breast cancer patients. All these tissues were mounted and sectioned as follows. Firstly, glass slides for sectioning were prepared with the following protocol. In a fume cupboard, slides were immersed in a bath of acetone for 5 minutes, in a bath of 2% (v/v) silane in acetone for 5 minutes and, after draining, in a different bath of acetone for 5 minutes. The slides were left to dry in the fume cupboard and stored at RT. The tissues were defrosted, quickly mounted in mounting liquid on cardboard pieces and re-frozen in liquid nitrogen. Mounted tissues were stored at -80°C until further use. In a cryostat, with the temperature of the blade set at -29°C and the temperature of the chamber set at -19°C, tissue sections of 7μm were applied on the glass slides. The slides were left at RT overnight. For mass spectrometry analysis using Matrix-assisted laser desorption/ionization (MALDI), about 10 sections (7μm) of each breast tissue were transferred into a 1.5ml Eppendorf tube and mixed with 100μl of lysis buffer, by pipetting up and down several times. The sample was sonicated for 30 seconds, vortexed for 30 seconds and stored at -80°C until further use. To fix the sections, the slides were immersed in a mix of 50% (v/v) methanol and 50% (v/v) acetone for 5 minutes at RT. The fixed sections were then frozen at -80°C until assayed by immunohistochemistry.

Antigen retrieval was performed on paraffin-embedded tissue microarrays as follows. The slides were baked for 2 hours at 60°C to melt the wax. The wax was then removed in a bath of xylene for 5 minutes followed by a different bath of xylene for 5 minutes. The sections were re-hydrated in graded ethanol (100%, 100% and 70% (v/v), each for 3 minutes) and rinsed in running tap water for 15 minutes. Citrate buffer was heated in the microwave for 10 minutes on full power or until boiling. The slides were added to the

near boiling citrate buffer which was further heated in the microwave on full power for 10 minutes. Slides were then rinsed in the running tap water for 15 minutes and finally rinsed with ddH_2O.
At this step, immunohistochemistry for both frozen tissues and paraffin-embedded tissues follows the same protocol. Frozen sections were taken out of the -80°C freezer, left to adjust to RT and placed in a bath of 1X PBS for 5 minutes. 0.3% (v/v) H_2O_2 diluted in 1X PBS was added to the tissue sections for 5 minutes. Tissues were then rinsed with 1X PBS and placed in a bath of 1X PBS for 5 minutes. 10% (v/v) goat serum (made from the species of the secondary antibody) diluted in 1X PBS was then added to block the sections for 20 minutes. The excess of serum was then shaken off and the sections were carefully blotted around. The primary antibody solution (see Materials) was added to the tissues and the slides were incubated overnight at 4°C. The staining of tissue sections with rabbit IgG isotype control and of sections from the same tissue (on a different slide) with anti-BUC11 antibody was always carried out at the same time. Tissues were then rinsed with 1X PBS and placed in a bath of 1X PBS for 5 minutes. The secondary antibody solution (see Materials) was added to the tissues and the slides were incubated at RT for 30 minutes. Meanwhile, ABC reagent was made up and left to stand 30 minutes. Tissues were then rinsed with 1X PBS and placed in a bath of 1X PBS for 5 minutes. The wash step was repeated once more. The ABC reagent was added to the tissues and left to react for 30 minutes at RT with the bound secondary antibodies. The slides were then washed twice as before and the DAB reagent was added onto the slides. The reagent was left to react with the ABC reagent until a suitable brown colour developed, then the slides were immediately rinsed in running tap water for 2.5 minutes and in ddH_2O for 2.5 minutes. Paraffin-embedded sections and frozen sections were finally counterstained respectively for 15 seconds in Gills haematoxylin solution and for 10 seconds in Harris haematoxylin solution. Tissues were then rinsed in running tap water and fixed consecutively, in a fume cupboard, in graded ethanol (70% 1 minute, 100% 1 minute, 100% 2 minutes, (v/v)) and in two different bath of xylene for 1 minute each. Slides were mounted with DPX mountant and a glass cover slip then left to air-dry for several hours. Fixed and mounted sections were subsequently stored at RT. Tissue staining was observed under the microscope. The scoring of the stained sections was not evaluated by a pathologist.

2.2.2 Cloning of BUC11

PCR products were either cloned into the mammalian expression vector pBudCE4.1 or into pCRII-blunt-TOPO vector following the manufacturer's protocol.

2.2.2.1 DNA band extraction

Following electrophoresis, in order to use the amplified gene products for sequencing or cloning into a vector, DNA was extracted and purified from the agarose using the DNA

Isolation Kit according to the manufacturer's protocol. Briefly, the agarose gel containing the DNA was dissolved in the NaI solution for 5 minutes at 55°C. Then, 30μl of glass powder was mixed to the solution and incubated at RT for 5 minutes, mixing occasionally during that time. The DNA bound to the glass powder was then washed three times with the wash buffer. After centrifugation at 14,000rpm, the pellet was resuspended in 30μl of ddH_2O and incubated at 55°C for 5 minutes. Following centrifugation at 14,000rpm for 45 seconds, the DNA in suspension within the supernatant was transferred to a fresh tube.

2.2.2.2 DNA ligation

Ligation of cDNA into the vectors was done using T4 ligase following the manufacturer's recommendations. Briefly, 7μl of the enzyme-digested insert was added to 1μl of the enzyme-double-digested plasmid with 1μl of T4 buffer and 1μl of T4 ligase for a final volume of 10μl. Then, to complete the ligation process, the mixture was incubated overnight at 4°C.

2.2.2.3 Transformation into competent XL1-Blue E.coli cells and bulking up of the plasmid

The ligated vector/gene product was used to transform XL1Blue cells (chemically competent E.coli cells). 2μl of the reaction was added onto cells on ice and gently mixed on ice for 10 minutes to 30 minutes. The cells were then heat shocked at 42°C for 3 minutes in a waterbath and immediately transferred on ice. 250μl of LB Broth was added at RT to the cells which were then shaken horizontally at 200rpm, 37°C for 1 hour. 150μl and 50μl of culture were subsequently spread onto two LB agar plates containing the required antibiotic (either 50μg/ml of ampicillin or 30μg/ml of zeocin, depending on the antibiotic resistance gene expressed by the plasmid of interest) and plates were incubated at 37°C overnight in the incubator. The following day, plates were checked for growth of vector positive colonies.

2.2.2.4 Enzyme digestion

Briefly, PCR products were digested either with one or two restriction enzymes to generate overhangs for the ligation process. The restriction enzymes used depended on the cDNA sequences to be cloned and on the recognition sites available in the multiple cloning sites of the plasmids. The separate digestion of pCRII-blunt-TOPO and gene-specific PCR products prior to ligation or the digestion, following cloning, of positive colonies vector pCRII-blunt-TOPO/gene of interest with the restriction enzyme EcoRI was carried out as follows. For each reaction, 0.1μg of plasmid DNA was mixed with 1μl of restriction enzyme buffer (H), 0.1μl of BSA (10μg/μl), the appropriate amount of ddH_2O to a final volume of 9.75μl and 0.25μl of EcoRI enzyme. The mix was incubated 2 hours in a waterbath at 37°C. The digested products were then loaded on an agarose

gel for electrophoresis as described previously. The separate digestion of pBudCE4.1 and BUC11-specific PCR products prior to ligation or the digestion, following cloning, of positive colonies vector pBudCE4.1/BUC11 using the two restriction enzymes SalI and XbaI was done as follow. For each reaction, 3μl of plasmid DNA was mixed with 0.5μl of restriction enzyme buffer (D), 0.5μl of BSA, 0.5μl of SalI enzyme and 0.5μl of XbaI enzyme. The mix was incubated 2 hours in a waterbath at 37°C. For one tube, the restriction enzymes have been replaced with ddH_2O to include a negative control. The digested products were then loaded on an agarose gel for electrophoresis as described previously.

2.2.2.5 DNA isolation and sequencing

Each picked colony which has grown on the LB agar plate is immersed into 3ml of LB broth containing antibiotic (concentrations as before) and incubated shaking overnight at 37°C. The following day, 1.5ml of culture was centrifuged at room temperature at 14,000rpm for 5 minutes. All the following centrifugations are performed under the same conditions. The supernatant was discarded, the pellet resuspended in 100μl cold GTE and incubated 5 minutes on ice. Then 200μl of 1% SDS/0.2M NaOH solution was added and the tube incubated 5 minutes on ice. Afterwards, 150μl of Potassium acetate (KOAc) solution was mixed and the tube was incubated 5 minutes on ice. The tube was centrifuged, the supernatant was collected in a fresh tube and the pellet discarded. 800μl of the bottom layer chloroform/isoamyl alcohol was added and the tube vortexed. Then the tube was centrifuged and the upper layer was collected in a fresh tube and 1ml of 100% molecular grade ethanol was added. The tube was then incubated at room temperature for 15 minutes and centrifuged. The supernatant was discarded and the pellet washed with 500μl of 70% molecular grade ethanol. The tube was centrifuged and the pellet was air dried for 30 minutes on the bench. Finally the pellet was resuspended in 30μl to 50μl ddH_2O with 1μl RNase. The concentration and purity of the isolated DNA was measured with a spectrophotometer. Samples were diluted 1:100 in ddH_2O and absorbance was measured at 260nm and 280nm. The purified PCR products were sent to MWG Biotech for sequencing (Value read).

The sequencing for the BUC genes was carried out using PCR products obtained with the following protocol. For each reaction with the Phusion^{TM}High-Fidelity DNA Polymerase, thermocycling was done in a final volume of 50μl containing 1.5μl of BR13 (breast cancer tissue) DNA (or 1.5μl of ddH_2O for negative control), 10μl of Phusion buffer HF, 1μl of dNTPs, 1μl of BUC6 forward primer, 1μl of BUC9 reverse primer (20mM), 0.5μl of polymerase and 35μl of ddH_2O. The cycling conditions for the reactions was as follow: 98°C 30 seconds, (98°C 10 seconds, 62°C 20 seconds, 72°C 1 minute) x 40 cycles, 72°C 5 minute, 4°C forever. This protocol was repeated in the same way for BUC9 forward primer with BUC10 reverse primer and BUC10 forward primer with BUC11 reverse primer. The PCR products were loaded on an agarose gel and, after

electrophoresis, PCR products were isolated as described previously. The purified PCR products were sent to MWG Biotech for sequencing (Value read). The sequencing results were aligned in SPIDEY alignment tool (http://www.ncbi.nlm.nih.gov/spidey/) against genomic DNA.

2.2.3 Gene induction or silencing

2.2.3.1 DNA transfection

Electroporation was the method of choice to transiently transfect cells as the experiments using Lipofectamine 2000 were proven unsuccessful. The transfection was carried out using the Amaxa®Cell line Nucleofector®Kit V for K562 cells as per the manufacturer's instructions. Briefly, 1x10^6 of non-adherent ALC cells used per reaction were centrifuged at 1,000rpm for 10 minutes. The pellet was resuspended in 100μl of Nucleofector®Solution V at RT and mixed with the required amount of plasmid DNA: pKKN2 with either pBudCE4.1/- or pBudCE4.1/BUC11, to a ratio of 1:2.5. The sample was then transferred into an Amaxa certified cuvette which was closed with a cap and inserted into the cuvette holder of the Nucleofector^{TM}. When the program for the electroporation has finished, 500μl of pre-warmed culture medium (no antibiotics) was added to the cells which were then immediately transferred using the provided plastic pipette to a well of a 6-well plate containing pre-warmed culture medium. When necessary, zeocin (50μg/ml) was added to the culture medium 48 hours following transfection. The following day, cells were harvested for expression analysis at different time points using MACSelect Microbeads and MS columns according to the manufaturer's protocol. Briefly, the transfected cells were centrifuged at 1,000rpm for 10 minutes and the pellet was resuspended in 320μl of the degassed, pre-cooled PBE solution. The cells were then magnetically labelled by incubating them with 80μl of MACSelect Microbeads on ice for 15 minutes (final volume adjusted to 2ml with PBE). The MS column was placed in the magnetic field of a suitable MAC separator and prepared by rinsing with 500μl of PBE. All of the cell suspension was then added to the column and unlabelled cells were collected in a fresh tube. The column was washed 4 times with 500μl of PBE with the unlabelled fraction being collected. The column was then removed from the separator and place in a suitable collection tube. Either 1ml of PBE or culture medium, depending on the subsequent use of the cells, was added onto the column and the transfected cells were immediately flushed out by firmly applying the supplied plunger.

2.2.3.2 Small interfering RNA (siRNA) transfection

The siRNA specific for BUC6 and BUC11 were designed using the online tool from Ambion siRNA Target Finder (http://www.ambion.com/techlib/misc/siRNA_finder). The siRNA were purchased from Eurogentec (Table 2.4). The transfection of breast cancer

cell lines SK-BR3-V(-) and MDA231 with gene-specific siRNA was carried out using INTERFERin^{TM}siRNA transfection reagent. Cells were plated out at a density of 5x10^4 cells per well in a 24 well plate in complete medium the day before transfection and placed in an incubator (37°C, 5% CO_2). The experiment was performed on duplicate wells. Each experiment comprised cells with gene-specific siRNA, cells with negative control siRNA, cells with INTERFERin^{TM}alone and cells alone. For each set of two wells, the procedure was as follow: 2.5μl (40μM) of gene-specific siRNA (or negative control siRNA) was added to 250μl of OptiMEM®I Reduced Serum medium and the tube was vortexed. 5μl of INTERFERin^{TM}was added and the tube was vortexed for 10 seconds. The mix was then incubated 10 minutes at RT. While incubating, the media from wells was removed and 100μl of fresh complete media was added to each well. After incubation, 100μl of INTERFERin^{TM}/siRNA mix was added to wells giving a final volume of 200μl, during this process the plate was gently rocked to ensure good coverage of the well. The cells were then placed in the incubator (37°C, 5% CO_2). After 6 hours incubation, 500μl of fresh complete medium was added to each well. After 24 hours, the cells were ready to be tested for gene knockdown. RNAs from cells were harvested at 24 hours (and at 48 hours and 72 hours when performing a time course) using the RNA STAT-60^{TM}method as described previously. RNA was reverse transcribed into cDNA and the gene expression was measured using quantitative RT-PCR as outlined previously. The Ct values for BUC6 or BUC11 were normalised to those of GAPDH. The Ct value obtained from cells with gene specific siRNA was compared to the Ct value obtained from cells with INTERFERin^{TM}alone which indicated the level of specific gene knockdown.

Gene	siRNA	Sense	Antisense
BUC6	1	CUU-UGU-GUG-UAC-CAA-UUA-A	UUA-AUU-GGU-ACA-CAC-AAA-G
BUC6	2	GCU-GAC-UCA-UGA-CAA-CUA-A	UUA-GUU-GUC-AUG-AGU-CAG-C
BUC11	2	GGC-CAA-GCA-CAU-CUG-CAA-AUU	UUU-GCA-GAU-GUG-CUU-GGC-CUU

Table 2.4: Sequences of siRNAs used for transfection in the gene knockdown study. The siRNA specific for BUC6 and BUC11 were designed using the online tool from Ambion siRNA Target Finder (http://www.ambion.com/techlib/misc/siRNA_finder). The siRNA were purchased from Eurogentec. The sense and antisense strands were annealed.

2.2.3.3 Proliferation assay

Transfected MDA231 cells or ALC cells were used in a tritiated thymidine-based proliferation assay. Tritiated thymidine was added to the cells (final concentration of 0.037MBq/ml) and incubated at 37°C in a 5% CO_2 incubator for 18 hours. Then, cells were harvested onto a 96-well filter plate (Luma-Plate 96) with the harvester. After the plate was dried in an oven for about 2 hours, each well was covered with 40μl of scintillation fluid. The radioactivity, in counts per minute for each well, was detected in a Top-count scintillation counter.

2.2.3.4 Affymetrix GeneChip® Microarray

The expression analysis of MDA231 siRNA-transfected cells was carried out using the GeneChip® Human Genome U133 Plus 2.0 Array and the NASC Affymetrix service provided by the University of Nottingham, Sutton Bonington Campus, Loughrough (information on the methodology they used is available at http://affymetrix.arabidopsis.info/). The experiment was conducted in duplicates for both BUC11 siRNA-transfected cells and negative control siRNA-transfected cells. The quantity and purity of the RNA extracted from the cells were assessed using a Nanodrop ND1000 microspectrophotometer. For each array, at least 5μg of highly pure RNA was required. The NASC Affymetrix service checked the quality of the RNA provided, using the Agilent Bioanalyzer, when the samples were received to ensure there was no degradation during transport and then they check again when cRNA has been made prior to fragmentation and hybridisation.

2.2.4 Animals and immunisation

2.2.4.1 Animals

Double transgenic C57BL/6 HLA-A2.1/-DR1 (HHDII-DR1) mice were received as a generous gift from Dr. F. Lemonnier (Institut Pasteur, Paris, France). HHDII-DR1 mating positive mice were maintained inbred by making sure they have a common F0 ancestor. Animals were bred at Nottingham Trent University animal house according to the Home Office Codes of Practice for the housing and care of animals.

2.2.4.2 DNA bullets and DNA immunisation

The expression vectors pBudCE4.1/- (control) and pBudCE4.1/BUC11 were coated respectively onto 1μm gold microcarriers according to the manufacturer's instructions. Briefly, 18μg of plasmid DNA was mixed with 200μl of 0.05M spermidine containing 0.0083g of gold by sonicating for a few seconds. During sonication, 200μl of 1M $CaCl_2$ was added drop wise then the tube was incubated at RT for 10 minutes. The mixture was washed three times in anhydrous ethanol and resuspended in 1ml of 0.025mg/ml PVP. Whilst sonicating, the mixture was loaded into a dried Tefzel tubing and left to stand in a Tubing Prep Station for 30 minutes. The dry ethanol was removed using a syringe and the tubing was spinned for 5 minutes with the nitrogen. The dried tubing was then removed from the station and cut into small pieces using a guillotine. These DNA bullets were stored at 4°C. Each HHDDII-DR1 mouse was either immunised with one gold bullet using a Helios gene gun or intra-muscularly in both legs with 25μl of RPMI medium (no G418, no FCS) containing 25μg of naked plasmid DNA either pBudCE4.1/- or pBudCE4.1/BUC11 in each leg. Three rounds of immunisation for the immunisation with the gene gun and 2 rounds of immunisation for the immunisation intra-muscularly were carried out on the same animals at 7-day intervals.

2.2.5 Cytotoxicity assay, proliferation assay and Enzyme-Linked ImmunoSorbent SPOT (ELISPOT) in mice

2.2.5.1 Generation of murine BM-DC

The protocol for the preparation of murine dendritic cells was adapted from the one provided by Inaba and colleagues (Inaba *et al.*, 1992). Briefly, the hind limbs of the mouse were harvested and the muscles and knuckles were removed using a scalpel. The bones were then flushed with BM-DC media, the cells were washed, resuspended in 2ml of BM-DC media and counted in 0.6% (v/v) acetic acid and 0.1% (v/v) Trypan blue. Cells were then plated in a 24-well plate at $0.5x10^6$ cells per well per ml of BM-DC media which contains 1ng/ml of murine GM-CSF. The BM-DC were pulsed with either 25μg of cell lysate per well (ALC/pBudCE4.1/BUC11 as the relevant gene or ALC/pBudCE4.1/- as a negative control or ALC/pBudCE4.1/HAGE as an irrelevant gene) or 12.5μg of peptide per well (BUC11 specific-peptide or HAGE specific-peptide as an irrelevant peptide) which were pre-incubated with the BUC11 antibody for 10 minutes at 37°C in a 5% CO_2 incubator. After incubation of the plate for 6 hours in the incubator, 1μl of LPS per well was added to induce complete maturation and the plate was incubated overnight in the incubator. The following day, the BM-DC were scraped with a plastic tip and centrifuged at at 1,500rpm for 3 minutes at 4°C. The pellet was resuspended in 1ml of BM-DC medium, 12.5μl of Poly I.C was added and BM-DC were incubated 4 hours at 37°C in a 5% CO_2 incubator prior their use in Enzyme-Linked ImmunoSorbent SPOT (ELISPOT) or in proliferation assay. Poly I.C is a synthetic Polyinosinic polycytidylic acid dsRNA which induces the production of IFNγ.

2.2.5.2 Isolation of murine leukocytes

Spleens were harvested from immunised mice and a naive mouse seven days after the last immunisation and the cells were flushed out with T cell medium in sterile condition. Some T cells were cultured without isolation for use in proliferation assays. These were cultured in T cell medium containing vitamin E (1μl/ml of culture medium) and 10μl (1mg/ml) per well of either BUC11 relevant peptide or T21 irrelevant peptide (*in vitro* re-stimulation). Three days later, 25U of IL-2 was added per well to the culture medium. The isolation of mouse leucocytes was done using the Pan T Cell Isolation Kit in accordance to the manufacturer's instructions. Briefly, following cell count, cells freshly isolated from spleens were spun down at 1,500rpm for 10 minutes at 4°C. The pelleted cells were resuspended in the buffer solution and the Biotin-AbCocktail was added up to 10μl per $1x10^7$ cells. The mix was incubated for 10 minutes at 4°C. Then, the buffer solution (30μl per $1x10^7$ cells) and the anti-biotin microbeads (19μl per $1x10^7$ cells) were added and the mix was incubated for 15 minutes at 4°C. Cells were washed with 6ml of

buffer and centrifuged at 1,500rpm for 10 minutes at 4°C. The pellet was resuspended in 500μl of buffer and the cell suspension was applied to a MS column that was placed in the magnetic field of a MACS separator and prepared by rinsing with 500μl of buffer. The unlabelled cells were collected and represented the enriched T cell fraction. The column was then washed 3 times with 500μl of buffer and the effluent was collected in the same tube of the previous step. The total effluent was centrifuged at 1,500rpm for 10 minutes at 4°C and the cells were counted in 0.1% (v/v) Trypan blue. Cells were then plated out in a 24 well-plate at 3x10^6 cells per ml per well in T cell medium and incubated overnight at 37°C in a 5% CO_2 incubator.

2.2.5.3 Proliferation assay for murine T cells

BM-DC were prepared as described before and counted in 0.1% (v/v) Trypan blue. The T cells used in this assay were the T cells restimulated *in vitro* with relevant (BUC11) and irrelevant (T21) peptides, prepared as described before and counted in 0.1% (v/v) Trypan blue. The ratio 10 T cells for every BM-DC was used so 100μl of T cells suspension (5x10^5 cells/ml of culture medium) and 100μl of BM-DC suspension (5x10^4 cells/ml of culture medium) were plated per well. The co-cultured cells were incubated 48 hours at 37°C in a 5% CO_2 incubator. Tritiated thymidine was then added to each well (final concentration of 0.037MBq/ml) and the plate was incubated for 18 hours at 37°C in a 5% CO_2 incubator. Cells were harvested using a 96-well harvester onto a 96-well filter plate (Luma-Plate 96) and the plate was dried in an oven for 2 hours. Then, 40μl of scintillation fluid was added to each well and the radioactivity was measured in counts per minute for each well in a Top-count scintillation counter.

2.2.5.4 ELISPOT

The IFNγ ELISPOT was carried out using the Mouse IFNγ ELISPOT Development Module according to the manufacturer's instructions. A 96-well plate was coated with 100μl of the diluted capture antibody, sealed with the lid and incubated at 4°C overnight. The plate was washed 3 times with 200μl of wash buffer per well then inverted on a clean paper towel. The membranes were blocked with 200μl of blocking buffer per well. The plate was covered with foil paper and incubated for 2 hours at RT. Meanwhile, the T cells prepared as outlined before were harvested, counted and resuspended in CTL culture medium containing glutamine. The plate was rinsed with CTL culture medium and the wells were immediately filled with 100μl of the T cell suspensions. As a control, some T cells were contionned with Staphylococcus aureus enterotoxin B (SEB promotes TH2 cell polarisation). The plate was incubated for about 2 hours at 37°C in a 5% CO_2 incubator (until the incubation of BM-DC described before has finished). The BM-DC are counted and plated at 100μl per well of the 96-well plate. The ratio used was 10 T cells for every BM-DC. The plate was incubated for 24 hours at 37°C in a 5% CO_2 incubator.

The plate was washed 4 times with wash buffer as described before then the 100μl of diluted detection antibody was added per well. The plate was covered with the lid and foil paper and incubated overnight at 4°C. The plate was washed 3 times with wash buffer as described before and 100μl of diluted Streptavidin-AP was added per well. The plate was then cover in foil paper and incubated for 2 hours at RT. The plate was washed 3 times with wash buffer as described before, rinsed with ddH_2O and inverted on a clean paper towel. Then, 100μl of BCIP/NBT solution was added per well, the plate was covered with the lid and foil paper and incubated in the dark for 30 minutes at RT. The plate was then rinsed once with ddH_2O, inverted on a clean paper towel and left to dry at 37°C. Once the plate was dry, it was stored at RT prior to analysis in the ImmunoSpot®Analyzer.

2.2.5.5 Chromium release cytotoxicity assay

For this assay, the T cells used were prepared as described before but have not been re-stimulated *in vitro* and the targets used were ALC/pBudCE4.1/BUC11 cells as relevant targets, ALC/pBudCE4.1/- cells as negative control and ALC/pBudCE4.1/HAGE cells as irrelevant targets. The ratio used was 100 T cells for every ALC cell. ALC cells were counted in 0.1% (v/v) Trypan blue and centrifuged at 1,500rpm for 3 minutes. The pellet was resuspended in residual medium, 56μl of Chromium-51 was added (final concentration of 1.85MBq/ml) and the tubes were incubated in the waterbath at 37°C for 1 hour. T cells were counted in 0.1% (v/v) Trypan blue and resuspended in T cell medium to a concentration of 2.5x10^6 cells/ml. After incubation, ALC cells were washed in 10ml of T cell medium and incubated for another hour in the waterbath at 37°C. ALC cells were then washed as before, counted and resuspended to a final concentration of 2.5x10^4 cells/ml in T cell medium. T cells and ALC cells were co-plated at 100μl per well (200μl final volume per well). The "maximum release" wells were prepared by mixing 70μl of T cell medium with 100μl of ALC cell suspension and 30μl of 1% SDS solution. The "spontaneous release" wells were prepared by mixing 100μl of T cell medium with 100μl of ALC cell suspension. The plate was incubated for 4 hours at 37°C in a 5% CO_2 incubator. Then, 50μl of each well was replated onto a 96-well filter plate (Luma-Plate 96) in the same order and this plate was incubated in an oven at 37°C to dry overnight. The radioactivity was then measured in counts per minute (CPM) per well in a Topcount scintillation counter. The percentage of specific cytotoxicity was determined using the following equation: percentage of cytotoxicity = (CPM experimental release - CPM spontaneous release)/(CPM maximum release - CPM spontaneous release).

Chapter 3

BUC Family Identification and mRNA Expression Analysis

3.1 Introduction

3.1.1 The importance of breast cancer-associated markers

The most common cancer for women in the UK is breast cancer, with around 46,000 new cases diagnosed each year (Cancer Research UK, 2008). The majority of breast cancers develop sporadically, often due to mutations or incorrectly regulated genes in the breast cells (Lerebours and Lidereau, 2002; Ross *et al.*, 2003; Widschwendter and Jones, 2002). Over the last ten years, from a better understanding of the complexicity of breast cancer and major developments in genomics and proteomics using high-throughput technologies, alternative, more trustworthy factors, called biomarkers, for the clinical management of breast cancer patients have emerged. General opinion is that biomarkers can bring more accurate and sensitive tools to the clinical settings than conventional approved histopathological parameters. A biomarker can be DNA-based, RNA-based or protein-based and can be used individually or in combination with other biomarkers. As stated by Ludwig and Weinstein (2005), the process of clinical validation for most of the candidate breast cancer biomarkers is only starting and almost none of them have been integrated into the formal TNM staging system. The low number of validated biomarkers for breast cancer is the consequence of often inconsistent results, too few patients' samples taken for studies and the enormous time it takes to validate any new medical innovation (Ludwig and Weinstein, 2005).

To move towards individualised and improved breast cancer care, there is a need for tumour markers that are associated with risk assessment, diagnosis, prognosis, prediction of treatment outcome, monitoring of the disease and development of novel therapies. The complete human genome has been sequenced and mapped in the Human Genome Project, launched in 1990. Following its completion, a large volume of information has been obtained and deposited into publicly available databases. This resource of information opened the door for researchers to understand the function of our 20,000-25,000 protein-coding genes and to cancer genomics/proteomics with the comparison of the genomes and proteomes of a normal cell and of a cancer cell of a given tissue.

The main methods for biomarker discovery are expression profiling by gene microarrays, proteomics and Expressed Sequence Tag (EST) database search. Microarray technology can be defined as an automated high-throughput technique for the analysis of thousands of different DNA sequences or proteins fixed to a really small chip of glass or silicon. Proteomics is the use of protein mass spectrometry to analyse and draw comparisons between protein sets of interest e.g. all proteins found in a breast cancer sample versus all proteins found in a normal breast sample. It can be combined with bioinformatics to determine relevant links and database interrogation in order to identify those proteins (Lancashire *et al.*, 2009). One of the proteomics techniques which has been used to identify breast cancer biomarkers is MALDI (Laronga and Drake, 2007). ESTs are generally short sequences produced from randomly selected cDNA clones and very often reflect the transcriptional and functional activity of the tissue/cell line which they were derived from. The presence or absence of ESTs in different libraries provides information about the organ, cell type, or tumour specificity of expressed genes. A gene is usually represented by several ESTs. The more ESTs of the same gene is found in a given library the more the gene is expressed in the tissue which the library derived from. Several studies have focused on the search for cancer related genes by whole genome *in silico* subtraction using EST entries.

3.1.2 The discovery of the breast-associated UniGene clusters family

The National Center for Biology Information (NCBI) website provides the UniGene system (http://www.ncbi.nlm.nih.gov/sites/entrez?db=unigene) (Figure 3.1) to automatically divide the collection of all DNA sequences available (GenBank database) into a non-redundant set of gene-oriented “clusters”. The information provided for each UniGene cluster include the ESTs of one gene, the tissues of origin and the chromosome map location.

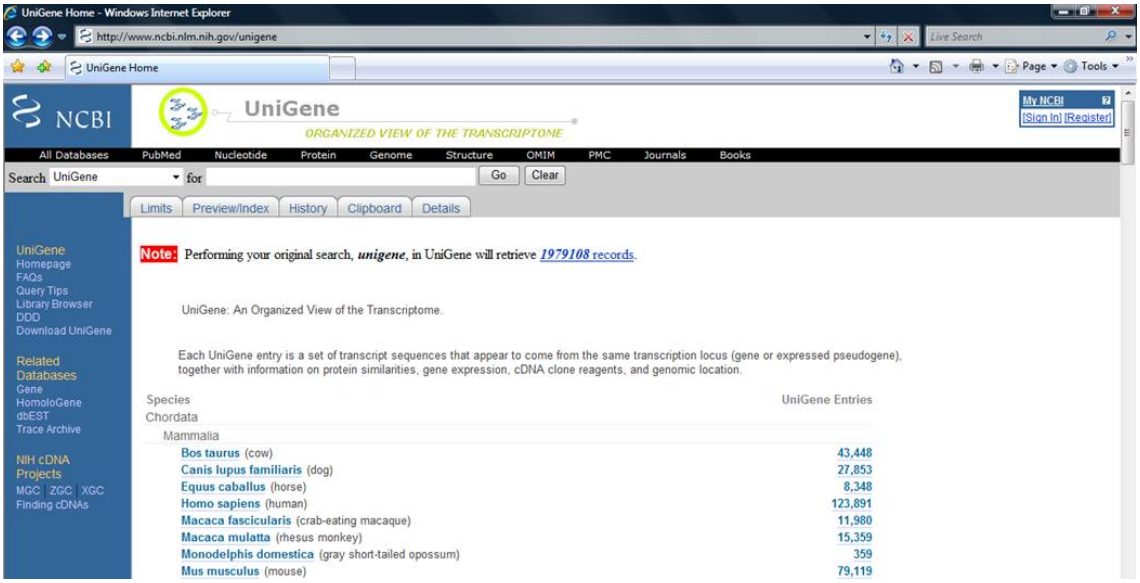


Figure 3.1: UniGene Homepage.
The UniGene Homepage shows the list of Species and their respective UniGene Entries. Each entry is a group of transcript sequences which originate from the same gene or expressed pseudogene and gives information such as gene expression, gene location and protein similarities (http://www.ncbi.nlm.nih.gov/sites/entrez?db=unigene, January 2009)

Previously in our laboratory, Dr Geng Li mined the UniGene database with the objective to identify new breast-associated genes, limiting his search to ESTs which originate from human normal breast and/or breast cancer tissues or cell lines but not from normal tissues or organs except testis, ovary and placenta (personal communication with Dr Li). He subsequently selected 228 breast-associated UniGene clusters (BUC1-228) and used RT-PCR to validate the expression patterns. According to preliminary work, four BUC genes (BUC6, BUC9, BUC10 and BUC11) were found highly expressed in normal breast, breast cancer and testes tissues but not in other normal tissues. In order to evaluate the potential of BUC6, BUC9, BUC10 and BUC11 for the management of breast cancer patients, this project was undertaken firstly to obtain the full sequence of the genes and determine their similarity with published genes and subsequently to conduct an exhaustive mRNA expression analysis. Gene expression was investigated using conventional RT-PCR. Afterwards, the expression patterns obtained with conventional RT-PCR were validated and further investigated using quantitative RT-PCR. The samples used in the study were from various normal breast tissues, breast cancer tissues, breast cancer cell lines, malignancies (e.g. melanoma and prostate cancer) and normal tissues. Some cancer tissues used were provided with patient-matched normal tissues and all material collected had full ethical approval granted.

3.2 Results

3.2.1 Characterisation of the full length mRNA sequences of the BUC family

The Basic Local Alignment Search Tool (BLAST) is a program available on the NCBI website (http://blast.ncbi.nlm.nih.gov/Blast.cgi) that identifies regions of local similarity between sequences, comparing nucleotide or protein sequences of interest to sequences deposited within databases. The program gives the statistical significance of matches and is very helpful in the identification of novel sequences or members of gene families and in the studies of the relationship between sequences such as function or evolution. The mRNA sequences of the ESTs BUC6, BUC9, BUC10 and BUC11, formatted in FASTA, were uploaded in the option "BLAST Human Sequences" of the "BLAST Assembled Genomes" found on the BLAST Homepage. The BLAST results indicated that all four ESTs appear to be located on the human chromosome 10.

In order to determine the precise location on the human chromosome 10 of the ESTs BUC6, BUC9, BUC10 and BUC11 and unlock information on their identity (e.g. gene homologies), a powerful alignment program called Spidey (Figure 3.2) was used.

Figure 3.2: Spidey Homepage.
Spidey is a mRNA-to-genomic alignment program invented by Sarah Wheelan and located on the NCBI website. The program allows the alignment of mRNA sequences of interest against the genomic sequence, regardless of intron size, using two alignment tools: BLAST and Dot View. It can also perform interspecies alignments (http://www.ncbi.nlm.nih.gov/spidey/index.html, January 2009).

Spidey is a program that can be found on the NCBI website (http://www.ncbi.nlm.nih.gov/spidey/index.html). The purpose of the program is to align mRNA sequences of interest against the genomic sequence, regardless of intron size, using two alignment tools: BLAST and Dot View. The input for Spidey are a unique genomic sequence and a set of mRNA accession numbers or sequences in FASTA format. The data is analysed one mRNA sequence at a time. Briefly, a high-stringency BLAST against the genomic sequence is firstly performed for each mRNA sequence and the alignments are sorted by score. Then, another BLAST search is performed with the entire mRNA and a high-scoring, non-overlapping subset of the alignments is created. The percent identity per exon is finally determined, as well as the number of gaps per exon, the overall percent identity, the percent coverage of the mRNA, the number of splice donor sites, the presence of an aligning or non-aligning poly(A) tail, the acceptor sites for each exon and the occurence of a mRNA that has a 5' or 3' end (or both) that does not align to the genomic sequence. The program can also perform interspecies alignments (http://www.ncbi.nlm.nih.gov/spidey/spideydoc.html).

The mRNA sequences, formatted in FASTA, of the ESTs BUC6, BUC9, BUC10 and BUC11 were aligned against the genomic Human DNA sequence from clone RP11-20F24 on chromosome 10p11.21-12.1. This region contains the 3' end of the NY-BR-1 gene for breast cancer antigen NY-BR-1, two novel genes, a pseudogene similar to part of ATP8A2 (ATPase, aminophospholipid transporter-like, Class I, type 8A, member 2), a novel pseudogene and a vomeronasal receptor pseudogene, complete sequence (GenBank accession number AL157387.9, information from NCBI website, updated on the 6th May 2008). When an "All Databases" search is performed on the NCBI website using the accession number for this genomic sequence, two gene entries can be found: ANKRD30A which is the Homo sapiens ankyrin repeat domain 30A (mRNA of 4405 base-pairs, validated protein coding), also known as the breast cancer antigen NY-BR-1 (Jager *et al.*, 2001) and the unpublished gene hCG25653 (LOC646360, mRNA of 653 base-pairs, predicted protein coding) located next on the chromosome. As shown in Figure 3.3, the BUC sequences don't overlap and they are aligned on this part of the chromosome, which strongly suggested that the BUC genes may be related.

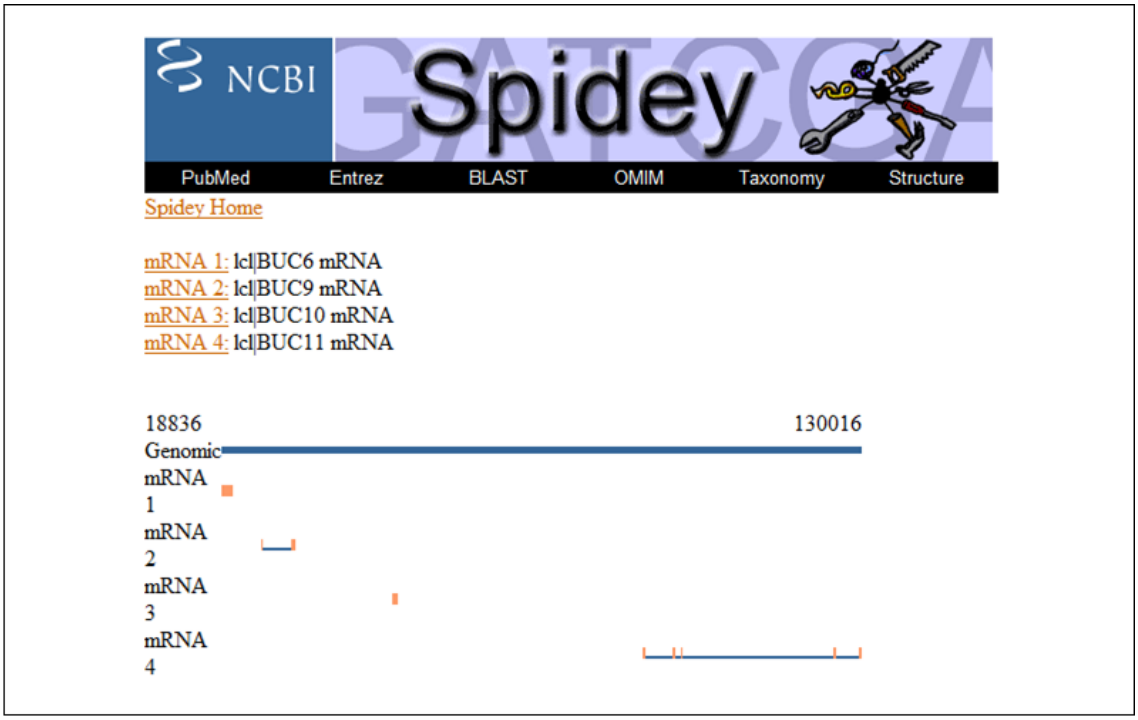
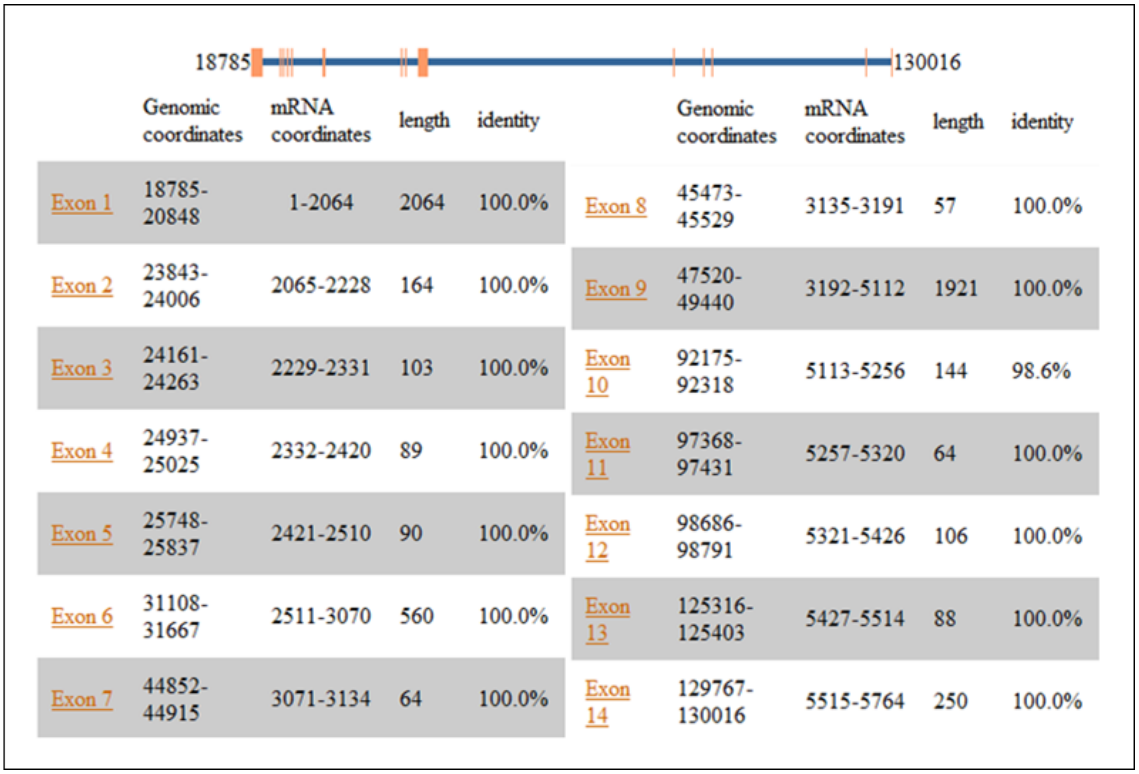


Figure 3.3: Spidey alignment of the BUC original mRNA sequences against genomic Human DNA sequence from clone RP11-20F24 on chromosome 10p11.21-12.1.
The mRNA sequences of BUC6, BUC9, BUC10 and BUC11, in FASTA format, were aligned in Spidey alignment tool against the genomic Human DNA sequence from clone RP11-20F24 on chromosome 10p11.21-12.1. Following the alignment, it became apparent that the sequences are not overlapping and that they are aligned one after another on this part of the chromosome.

The ESTs, as stated previously, are incomplete sequences of expressed genes, therefore it was decided to amplify and sequence the full length of the mRNA molecules between BUC6 and BUC11. The sequencing was carried out using PCR products obtained with the following protocol. The following primers were designed especially for this sequencing: BUC69 forward and reverse primers as well as BUC910 forward and reverse primers. BUC69 forward primer anneals in the 3' end of BUC6 sequence and BUC69 reverse primer anneals in the 5' end of BUC9 sequence. BUC910 forward primer anneals in the 3' end of BUC9 sequence and BUC910 reverse primer anneals in the 5' end of BUC10 sequence. Briefly, for each reaction with the Phusion^{TM} High-Fidelity DNA Polymerase, thermocycling was done in a final volume of 50μl containing 1.5μl of BR13 (breast cancer tissue) cDNA (or 1.5μl of double distilled water for negative control), 10μl of Phusion buffer HF, 1μl of dNTPs, 1μl of BUC69 forward primer, 1μl of BUC69 reverse primer, 0.5μl of polymerase and 35μl of double distilled water. The cycling conditions for the reactions was as follow: 98°C 30sec, (98°C 10sec, 58°C 30sec, 72°C 1min) x 40 cycles, 72°C 7min, 4°C hold. This protocol was repeated in the same way for BUC910 forward primer with BUC910 reverse primer, for BUC9 forward primer with BUC11 reverse primer, for BUC69 forward primer with BUC910 reverse primer and

for BUC69 forward primer with BUC11 reverse primer. The PCR products were loaded on an agarose gel and after electrophoresis PCR products were isolated as described in Chapter 2. The purified PCR products were used for a second round of PCR amplification and the purified products were sent to MWG Biotech (UK) for sequencing (Value read). The sequencing results were aligned in Spidey against the genomic DNA corresponding to Human DNA sequence from clone RP11-20F24 on chromosome 10p11.21-12.1 (GenBank accession number AL157387.9). The overlapping sequences were merged together to create one single sequence of the "BUC gene" which was re-aligned against the same genomic DNA sequence (Figure 3.4). The figure shows the "BUC gene" mRNA sequence to be 5,764 base-pair-long and located on chromosome 10.



	Genomic coordinates	mRNA coordinates	length	identity		Genomic coordinates	mRNA coordinates	length	identity
Exon 1	18785-20848	1-2064	2064	100.0%	Exon 8	45473-45529	3135-3191	57	100.0%
Exon 2	23843-24006	2065-2228	164	100.0%	Exon 9	47520-49440	3192-5112	1921	100.0%
Exon 3	24161-24263	2229-2331	103	100.0%	Exon 10	92175-92318	5113-5256	144	98.6%
Exon 4	24937-25025	2332-2420	89	100.0%	Exon 11	97368-97431	5257-5320	64	100.0%
Exon 5	25748-25837	2421-2510	90	100.0%	Exon 12	98686-98791	5321-5426	106	100.0%
Exon 6	31108-31667	2511-3070	560	100.0%	Exon 13	125316-125403	5427-5514	88	100.0%
Exon 7	44852-44915	3071-3134	64	100.0%	Exon 14	129767-130016	5515-5764	250	100.0%

Figure 3.4: Spidey alignment of the BUC sequencing results against genomic Human DNA sequence from clone RP11-20F24 on chromosome 10p11.21-12.1.
After sequencing, the mRNA sequences of BUC6, BUC9, BUC10 and BUC11, in FASTA format, were aligned in Spidey alignment tool against the genomic Human DNA sequence from clone RP11-20F24 on chromosome 10p11.21-12.1. The overlapping sequences were merged together and the single sequence "BUC gene" was re-aligned against the same genomic DNA sequence which resulted in this figure. The "BUC genes" are expressed as a mRNA sequence of 5,764 base-pair-long which does not include any obvious protein coding frame.

Following sequencing and analysis, the nucleotide sequence of each BUC gene was uploaded on the Translator tool of the JustBio website in order to get the predicted protein sequence (http://www.justbio.com/translator/index.php). The nucleotide sequences for BUC6, BUC9, BUC10 and BUC11 can be found, respectively, in Figure 3.5, Figure 3.7,

Figure 3.9 and Figure 3.11 and the predicted protein sequences for BUC6, BUC9, BUC10 and BUC11 can be found, respectively, in Figure 3.6, Figure 3.8, Figure 3.10 and Figure 3.12.

BUC6 nucleotide sequence

AATCTCGCTCTGTCACTCAGGCTGGAGTGCAGTGGCGCAATCTCGGCTCACTGCAAACTCTGCCTCCCA
GGTTCACGCCATTCTCCTGCCACAGCCTCCCGAGTAGCTGGGACTACAGGCGCCCGCCACCAAGCACAG
CTAAGTTTTTTATTTATAGTAGAGACGGGGTTTCACTGTGTTAGCCAGGATGGTCTCGATCTCCTGACC
TCGTGATCTGCCTGCCTCGGCCTCCCAAAGTGCCGGGATTACAGGCGTGAGCCACCGCGCCGGGCCTGA
TTTCAGTTTCTTCCAGCCCTTCCTATTGTTAACATGGGGGTTGTGTTGAAGAATATAAAGTTACAAAGT
CAAGGAAGTAGGAAACATTTTTACAAGTATTATGTAGCCATCTTGGTGGGGCTGTGGTGAGGTAGGCTG
CAAATGATTCTCCTATTTCTTTCCCTGAGTTCAGAACATAGGAATTAGATTGATAGACATCAACATACC
CGCTTTATTGCTGACTCATGACAACTAATGGGAAGACATGGCTCAGATGTGCAGCCACAGTGAGCTTCT
GAACATTTCTTCTCAGACTAAGCTCTTACACACAGTTGCAGTTGAAAGAAAGAATTGCTTGACATGGCC
ACAGGAGCAGGCAGCTTCCTGCAGACATGACAGTCAACGCAAACTCATGTCACTGTGGGCAGACACATG
TTTGCAAAGAGACTCAGAGCCAAACAAGCACACTCAATGTGCTTTGCCCAAATTTACCCATTAGGTAAA
TCTTCCCTCCTCCCAAGAAGAAAGTGGAGAGAGCATGAGTCCTCACATGGAAACTTGAAGTCAGGGAAA
TGAAGGCTCACCAATTATTTGTGCATGGGTTTAAGTTTTCCTTGAAATTAAGTTCAGGTTTGTCTTTGT
GTGTACCAATTAATGACAAGAGGTTAGATAGAAGTATGCTAGATGGCAAAGAGAAATATGTTTTGTGTC
TTCAATTTTGCTAAAAATAACCCAGAACATGGATAATTCATTTATTAATTGATTTTGGTAAGCCAAGTC
CTATTTGGAGAAAATTAATAGTTTTTCTAAAAAAGAATTTTCTCAATATCACCTGGCTTGATAACATTT
TTCTCCTTCGAGTTCCTTTTTCTGGAGTTTAACAAACTTGTTCTTTACAAATAGATTATATTGACTACC
ACTCACTGATGTTATGATATTAGTTTCTATTGCTTACTTTGTATTTCTAATTTTAGGATTCACAATTTA
GCTGGAGAACTATTTTTTAACCTGTTGCACCTAAACATGATTGAGCTAGAAGACAGTTTTACCATATGC
ATGCATTTTCTCTGAGTTATATTTTAAAATCTATACATTTCTCCTAAATATGGAGGAAATCACTGGCAT
CAAATGCCAGTCTCAGACGGAAGACCTAAAGCCCATTTCTGGCCTGGAGCTACTTGGCTTTGTGACCTA
TGGTGAGGCATAAGTGCTCTGAGTTTGTGTTGCCTCTTTTGTAAAATGAGGGTTTGACTTAATCAGTGA
TTTTCATAGCTTAAAATTTTTTTGAAGAACAGAACTTTTTTTAAAAACAGTTAGATGCAACCATATTAT
ATAAAACAGAACAGATACAAGTAGAGCTAACTTGCTAAAGAAAGGATGGAGGCTCTGAAGCTGTGACTT
CATTATCCCTTAATACTGCTATGTCCTCTGTAGTACCTTAGATTTCTATGGGACATCGTTTAAAAACTA
TTGTTTATGCGAGAGCCTTGCTAATTTCCTAAAAATTGTGGATACATTTTTTCTCCCATGTATAATTTT
CTCACCTTCTATTTAAAAAGAAAAAAAAAGTCAGTGTAGTATTTACATATTTTACCCTATAAGGAGCTA
ACATAACTTTTGATTTAGTGTTATTCATAAAATTAGGTTAGCAGTTTATTAACCTTTTGTATTTGCTCT
GGCAATGTTTAATATCTCATAAGCTATACACACCTCGAAGCCATCAATGACAACCTTTTCTTGCTGAAT
AGAACAGTGATTGATGTCATGAAGACAATTTTATCTCCTTTTGCCTTCCATAATTTGTACCAGCAGGTT
ATATAATAGTATAACACTGCCAAGGAGCGGATTATCTCATCTTCATCCTGTAATTCCAGTGTTTGTCAC
GTGGTTGTTGAATAAATGAATAAAGAATGAGAAAACCAGAAGCTCTGATACATAATCATAATGATAATT
ATTTCAATGCACAACTACGGGTGGTGCTGAACTAGAATCTATATTTTCTGAAACTGGCTCCTCTAGGAT
CTACTAATGATTTAAATCTAAAAGATGAAGTTAGTAAAGCATCAGAAAAAAAAAGGTGGGTATTCCTACA
AGTCAGGACATTCTACGTGACTACAATATAATCTCACAGAAATTTAACATTAATACATTCTAAGATTTA
ATTCT

Figure 3.5: BUC6 nucleotide sequence.
This is the full mRNA sequence for BUC6, obtained following sequencing and alignments in Spidey alignment tool. The mRNA sequence is 2,420 base-pair-long.

BUC6 translated protein sequence (3 phase translation)

Phase 1:

NLALSLRLECSGAISAHCKLCLPGSRHSPATASRVAGTTGARHQAQLSFLFIVETGFHCVSQDGLDLLT
S*SACLGLPKCRDYRREPPRRA*FQFLPALPIVNMGVVLKNIKLQSQGSRKHFYKYYVAILVGLW*GRL
QMILLFLSLSSEHRN*IDRHQHTRFIADS*QLMGRHGSDVQPQ*ASEHFFSD*ALTHSCS*KKELLDMA
TGAGSFLQT*QSTQTHVTVGRHMFAKRLRAKQAHSMCFAQIYPLGKSSLLPRRKWREHESSHGNLKSGK
*RLTNYLCMGLSFP*N*VQVCLCVYQLMTRG*IEVC*MAKRNMFCVFNFAKNNPEHG*FIY*LILVSQV
LFGEN**FF*KRIFSISPGLITFFSFEFLFLEFNKLVLYK*IILTTTH*CYDISFYCLLCISNFRIHNL
AGELFFNLLHLNMIELEDSFTICMHFL*VIF*NLYISPKYGGNHWHQMPVSDGRPKAHFWPGATWLCDL
W*GISALSLCCLFCKMRV*LNQ*FS*LKIFLKNRTFFKNS*MQPYYIKQNRYK*S*LAKERMEALKL*L
HYPLILLCPL*YLRFLWDIV*KLLFMREPC*FPKNCGYIFSPMYNFLTFYLKRKKKSV*YLHILPYKEL
T*LLI*CYS*N*VSSLLTFCICSGNV*YLISYTHLEAINDNLFLLNRTVIDVMKTILSPFAFHNLYQQV
I**YNTAKERIISSSSCNSSVCHVVVE*MNKE*ENQKL*YIIIMIIISMHNYGWC*TRIYIF*NWLL*D
LLMI*I*KMKLVKHQKKKVGIPTSQDILRDYNIISQKFNINTF*DLIX

Phase 2:

ISLCHSGWSAVAQSRLTANSASQVHAILLPQPPE*LGLQAPATKHS*VFYL**RRGFTVLARMVSIS*P
RDLPASASQSAGITGVSHRAGPDFSFFQPFLLLTWGLC*RI*SYKVKEVGNIFTSIM*PSWWGCGEVGC
K*FSYFFP*VQNIGIRLIDINIPALLLTHDN*WEDMAQMCSHSELLNISSQTKLLHTVAVERKNCLTWP
QEQAASCRHDSQRKLMSLWADTCLQRDSEPNKHTQCALPKFTH*VNLPSSQEESGESMSPHMET*SQGN
EGSPIICAWV*VFLEIKFRFVFVCTN**QEVR*KYARWQREICFVSSILLKITQNMDNSFIN*FW*AKS
YLEKINSFSKKEFSQYHLA**HFSPSSSFFWSLTNLFFTNRLY*LPLTDVMILVSIAYFVFLILGFTI*
LENYFLTCCT*T*LS*KTVLPYACIFSELYFKIYTFLLNMEEITGIKCQSQTEDLKPISGLELLGFVTY
GEA*VL*VCVASFVK*GFDLISDFHSLKFF*RTELFLKTVRCNHII*NRTDTSRANLLKKGWRL*SCDF
IIP*YCYVLCSTLDFYGTSFKNYCLCESLANFLKIVDTFFLPCIIFSPSI*KEKKSQCSIYIFYPIRS*
HNF*FSVIHKIRLAVY*PFVFALAMFNIS*AIHTSKPSMTTFSC*IEQ*LMS*RQFYLLLPSIICTSRL
YNSITLPRSGLSHLHPVIPVFVTWLLNK*IKNEKTRSSDT*S***LFQCTTTGGAELESIFSETGSSRI
Y**FKSKR*S**SIRKKRWVFLQVRTFYVTTI*SHRNLTLIHSKI*FX

Phase 3:

SRSVTQAGVQWRNLGSLQTLPPRFTPFSCHSLPSSWDYRRPPPSTAKFFIYSRDGVSLC*PGWSRSPDL
VICLPRPPKVPGLQA*ATAPGLISVSSSPSYC*HGGCVEEYKVTKSRK*ETFLQVLCSHLGGAVVR*AA
NDSPISFPEFRT*ELD**TSTYPLYC*LMTTNGKTWLRCAATVSF*TFLLRLSSYTQLQLKERIA*HGH
RSRQLPADMTVNANSCHCGQTHVCKETQSQTSTLNVLCPNLPIR*IFPPPKKKVERA*VLTWKLEVREM
KAHQLFVHGFKFSLKLSSGLSLCVPINDKRLDRSMLDGKEKYVLCLQFC*K*PRTWIIHLLIDFGKPSP
IWRKLIVFLKKNFLNITWLDNIFLLRVPFSGV*QTCSLQIDYIDYHSLML*Y*FLLLTLYF*F*DSQFS
WRTIF*PVAPKHD*ARRQFYHMHAFSLSYILKSIHFS*IWRKSLASNASLRRKT*SPFLAWSYLAL*PM
VRHKCSEFVLPLL*NEGLT*SVIFIA*NFFEEQNFF*KQLDATILYKTEQIQVELTC*RKDGGSEAVTS
LSLNTAMSSVVP*ISMGHRLKTIVYARALLIS*KLWIHFFSHV*FSHLLFKKKKKVSVVFTYFTL*GAN
ITFDLVLFIKLG*QFINLLYLLWQCLISHKLYTPRSHQ*QPFLAE*NSD*CHEDNFISFCLP*FVPAGY
IIV*HCQGADYLIFIL*FQCLSRGC*INE*RMRKPEALIHNHNDNYFNAQLRVVLN*NLYFLKLAPLGS
TNDLNLKDEVSKASEKKGGYSYKSGHST*LQYNLTEI*H*YILRFNSX

Figure 3.6: BUC6 predicted protein sequence.
The nucleotide sequence was translated into a protein sequence using the Translator tool of the JustBio website (http://www.justbio.com/translator/index.php). The translation was carried out in 3 phases from 5' end of nucleotide sequence, using the standard genetic code. Stars represent stop codons.

BUC9 nucleotide sequence

GTAAACAAATTGCTCCTGTGGAGATGATTGGCATCACATGGTGTTTTGAGCTGATACACCCAACACTTG
AGCTCACTGCAACAGTACCAGATTTTCACCGCTATGCCTCCTTTCACTCTGGGAGTCTTCCAGAGGTCT
TGCACTCGGGAGAGCATGCTCAGGTTTCCCCAGCTCTACAAAATCACCCAGAATGCCAAAGACTTCAAC
ACAAGGGTAAATAAGGTTGATCTCAGAATTGTCACCTCAAAAAGGCCCTGCCTTCCACTGTTCAGTTCT
GGTCATCTGCCTATGAGATATCTGAAGCTTGAAAGAGAACACTTGAAAATCACTGAGACCGTGACTCCC
ATCCCAGCACACACAGCAAGCCAAGTAGGTTACAGAGATTTCTTCTTGGGTGATGAGTTCACGCCACAT
GGGGTATGTTCTCCAGTTCCAGTGTGTTGACTCCTTCCTGCTTCCCCCCATCAGCCCCTGAGGTCAATG
TGGGCAGCAGCCCCCATGGTCCAAGTTCAAGATCCGCTGTGGAAGACTTTTTTAGGCAATCACACAGCC
CTGCATGGAGTCCTGATGAGAGCTTGCCTAATTGTTGCTAGGTTTGTCATTTTAAATACAGTGTTTCTT
TAGCTAGTGAGTAAAATTGGCTATATAGGAGATACTGTGTTGACCAGTGGTCATGCCACTGCCTGTTGA
TTTGTTGAAAATATTGTTTACACGTATGTTCTTGTTACTGATTGTCAGAAAGCTGGTTTTGAGACTGCA
GCTTGGACTAAA

Figure 3.7: BUC9 nucleotide sequence.
This is the full mRNA sequence for BUC9, obtained following sequencing and alignments in Spidey alignment tool. The mRNA sequence is 771 base-pair-long.

BUC9 translated protein sequence (3 phase translation)

Phase 1:

VNKLLLWR*LASHGVLS*YTQHLSSLQQYQIFTAMPPFTLGVFQRSCTRESMLRFPQLYKITQNAKDFN
TRVNKVDLRIVTSKRPCLPLFSSGHLPMRYLKLEREHLKITETVTPIPAHTASQVGYRDFFLGDEFTPH
GVCSPVPVC*LLPASPHQPLRSMWAAAPMVQVQDPLWKTFLGNHTALHGVLMRACLIVARFVILNTVFL
*LVSKIGYIGDTVLTSGHATAC*FVENIVYTYVLVTDCQKAGFETAAWTKX

Phase 2:

*TNCSCGDDWHHMVF*ADTPNT*AHCNSTRFSPLCLLSLWESSRGLALGRACSGFPSSTKSPRMPKTST
QG*IRLISELSPQKGPAFHCSVLVICL*DI*SLKENT*KSLRP*LPSQHTQQAK*VTEISSWVMSSRHM
GYVLQFQCVDSFLLPPISP*GQCGQQPPWSKFKIRCGRLF*AITQPCMES**ELA*LLLGLSF*IQCFF
S**VKLAI*EILC*PVVMPLPVDLLKILFTRMFLLLIVRKLVLRLQLGLX

Phase 3:

KQIAPVEMIGITWCFELIHPTLELTATVPDFHRYASFHSGSLPEVLHSGEHAQVSPALQNHPECQRLQH
KGK*G*SQNCHLKKALPSTVQFWSSAYEISEA*KRTLENH*DRDSHPSTHSKPSRLQRFLLG**VHATW
GMFSSSSVLTPSCFPPSAPEVNVGSSPHGPSSRSAVEDFFRQSHSPAWSPDESLPNCC*VCHFKYSVSL
ASE*NWLYRRYCVDQWSCHCLLIC*KYCLHVCSCY*LSESWF*DCSLD*X

Figure 3.8: BUC9 predicted protein sequence.
The nucleotide sequence was translated into a protein sequence using the Translator tool of the JustBio website (http://www.justbio.com/translator/index.php). The translation was carried out in 3 phases from 5' end of nucleotide sequence, using the standard genetic code. Stars represent stop codons.

BUC10 nucleotide sequence

TTCAGTCATCTGGCTGTCTGGGGAAGCATGCTGACCAGTCTGGTGTTCTTTGGCATCTACTCAGCCATC
TGGTCCACCATTCTCATTGCCCCAAATATGAGAGGACAGGTAAGTATGCCAGATTGGAAGTTTGCCTTC
TGTGTCTTGTGATAACAATGTTTCTCCATGTGTAGGCAAGCAGGCTCTCTGTGTTTTACAACTGTATTG
AGGTTTGATAGACGCTGGTGCTACCTCTTTCTGTCAGGATCTATCAGGAGGTAAACCAAATTAAAGCAC
CTGGAAGATGTCATCCAATCCACAGAAATGGCATTTCAAATTTAAAGTATTATAAAAAGTCAACCTAAA
TCCCCTTACTTACCTGGCATGTCCACATTTTCATAATAAATTATGTGTTAATTGCACATAGTACTGATA
TGTGCAGTTTATGGCCCTTTGGAAACTGGAGCTATTTTGAGTTGCTCATAATCCAGTGATTATACTGGT
TACAGTTTTGATGTTGTACTCTGATTAGTAGTTATCCTCTGCTGAGTCATCTGTGGAAGTTAAATCTAA
CAGCTCTATGAACCAGCATCTCTCATTGTTCCGTCTCTTTGAACTAGAAAGACTGAGATTAAAGGTCTT
TATACTCACTTGAATATCTTTTTATTTTTTTCTTTTATATTACCTTTTTCTAGTGCATGAGATTCAGCC
ATCAACCGTGATGAACTGGATCATGGAATCTAACAATATCAAGTTTGGATAGAAAGGAGCTTGGCAATC
TTGTGGTCCATGCCTCCTCAATTTACCAAGGAAGCTGTGCAAGCTCAGGTCTCCAGGTTCATCCCAGAG
ACCCCACCCTGGGGTTGCCCGCAGCATTCCTGTCCTTAGACCCCAAAAACAGGCTGCTGGGCCCTGTAT
TTAGGGACCTCTGCACTGGCTCTTCTTCTCTGGGTGTATTTCTCTGCACAGAGTGAAAAGCTGTGATGC
ACACAGCCAGTCTTCCCACTTCTAATCTGTTCTCTGGGCTCCCAGGAACCCAAGCCTCTTTCTCAAACA
GCCAAATGCCCACTGCCCAGGCCTGCCTGGGCCTCCTTCCAGATCAAGGCCAGTTTGCACAACCTTAGG
CCCAAGGGACAGTTGTTTGCAAGGGGTGTGCACTGAGTAGATTGATGTTTGCACTGGTGTTTCCACCGA
CGTTTAAAGGAGGATGTTCATGCTGGGGCACAGAACCCAGGTGGAAAGGGAAGAACTGTGGGTCAGACT
TTAATTGGTGTTTCCCACCCAGGAATACATATTTTAGGGGCAGGTCTGGGCATGGCACAGCTGTCTGCT
ATCTCTGCCTCACTTACTCTGCGGAAGTAAACAAACCCCCCTGGGTTAACCCAAGTCTAGACTGTGTCC
AGCAGCTGTTTCTTGTGAGGTAGCCCTGAGTAGTGTTGCTTGTCTTTTGACCTGGTGGCTCAATTGGCT
AATATCCCAAGGTGCAGAGGACCCCCACTGAGGATGCTGAGTGGGATGGACCTGTCCTGCAGGAGCGGG
CGTGCTTGCCCTCTAATGCATTTTTGCTTTTTGTAAGAAAAGAGGACAAATGTCTAATGTGCTTTTCAG
TTCTTTTCCAAATTCAGACAATAGAGGAAATTAGAGGGGATTCAATAAACCCAGAATATTGTTGTGTAG
CAAAATGAGCATTTCCAACAAAGAAACTTTGTGTGTTTAAAGAAAGCAATTCTATTTTAAAGCAGCATG
ATTGAGCCCCACAAAACAAAGTGTCAGCTACAAGACAACAATCTTTTGGTAAGGTATTGGGGCTTACGA
TATCTGTTTTGTGCCCTCTTCCTTCTGGTTCAGAGTTGATCTGCTATGATAAATATTTTATTTGGGGAG
TTTTTTGGTACTGGACAATAAACAGGGCCACAGCTGATTGCCAGAGACAATTTTCTAA

Figure 3.9: BUC10 nucleotide sequence.
This is the full mRNA sequence for BUC10, obtained following sequencing and alignments in Spidey alignment tool. The mRNA sequence is 1,921 base-pair-long.

BUC10 translated protein sequence (3 phase translation)

Phase 1:

FSHLAVWGSMLTSLVFFGIYSAIWSTILIAPNMRGQVSMPDWKFAFCVL**QCFSMCRQAGSLCFTTVL
RFDRRWCYLFLSGSIRR*TKLKHLEDVIQSTEMAFQI*SIIKSQPKSPYLPGMSTFS**IMC*LHIVLI
CAVYGPLETGAILSCS*SSDYTGYSFDVVL*LVVILC*VICGS*I*QLYEPASLIVPSL*TRKTEIKGL
YTHLNIFLFFSFILPFSSA*DSAINRDELDHGI*QYQVWIERSLAILWSMPPQFTKEAVQAQVSRFIPE
TPPWGCPQHSCP*TPKTGCWALYLGTSALALLLWVYFSAQSEKL*CTQPVFPLLICSLGSQEPKPLSQT
AKCPLPRPAWASFQIKASLHNLRPKGQLFARGVH*VD*CLHWCFHRRLKEDVHAGAQNPGGKGRTVGQT
LIGVSHPGIHILGAGLGMAQLSAISASLTLRK*TNPPGLTQV*TVSSSCFL*GSPE*CCLSFDLVAQLA
NIPRCRGPPLRMLSGMDLSCRSGRACPLMHFCFL*EKRTNV*CAFQFFSKFRQ*RKLEGIQ*TQNIVV*
QNEHFQQRNFVCLKKAILF*SSMIEPHKTKCQLQDNNLLVRYWGLRYLFCALFLLVQS*SAMINILFGE
FFGTGQ*TGPQLIARDNFLX

Phase 2:

SVIWLSGEAC*PVWCSLASTQPSGPPFSLPQI*EDR*VCQIGSLPSVSCDNNVSPCVGKQALCVLQLY*
GLIDAGATSFCQDLSGGKPN*STWKMSSNPQKWHFKFKVL*KVNLNPLTYLACPHFHNKLCVNCT*Y*Y
VQFMALWKLELF*VAHNPVIILVTVLMLYSD**LSSAESSVEVKSNSSMNQHLSLFRLFELERLRLKVF
ILT*ISFYFFLLYYLFLVHEIQPSTVMNWIMESNNIKFG*KGAWQSCGPCLLNLPRKLCKLRSPGSSQR
PHPGVARSIPVLRPQKQAAGPCI*GPLHWLFFSGCISLHRVKSCDAHSQSSHF*SVLWAPRNPSLFLKQ
PNAHCPGLPGPPSRSRPVCTTLGPRDSCLQGVCTE*IDVCTGVSTDV*RRMFMLGHRTQVEREELWVRL
*LVFPTQEYIF*GQVWAWHSCLLSLPHLLCGSKQTPLG*PKSRLCPAAVSCEVALSSVACLLTWWLNWL
ISQGAEDPH*GC*VGWTCPAGAGVLAL*CIFAFCKKRGQMSNVLFSSFPNSDNRGN*RGFNKPRILLCS
KMSISNKETLCV*RKQFYFKAA*LSPTKQSVSYKTTIFW*GIGAYDICFVPSSFWFRVDLL**IFYLGS
FLVLDNKQGHS*LPETIF*X

Phase 3:

QSSGCLGKHADQSGVLWHLLSHLVHHSHCPKYERTGKYARLEVCLLCLVITMFLHV*ASRLSVFYNCIE
V**TLVLPLSVRIYQEVNQIKAPGRCHPIHRNGISNLKYYKKST*IPLLTWHVHIFIINYVLIAHSTDM
CSLWPFGNWSYFELLIIQ*LYWLQF*CCTLISSYPLLSHLWKLNLTAL*TSISHCSVSLN*KD*D*RSL
YSLEYLFIFFFYITFF*CMRFSHQP**TGSWNLTISSLDRKELGNLVVHASSIYQGSCASSGLQVHPRD
PTLGLPAAFLSLDPKNRLLGPVFRDLCTGSSSLGVFLCTE*KAVMHTASLPTSNLFSGLPGTQASFSNS
QMPTAQACLGLLPDQGQFAQP*AQGTVVCKGCALSRLMFALVFPPTFKGGCSCWGTEPRWKGKNCGSDF
NWCFPPRNTYFRGRSGHGTAVCYLCLTYSAEVNKPPWVNPSLDCVQQLFLVR*P*VVLLVF*PGGSIG*
YPKVQRTPTEDAEWDGPVLQERACLPSNAFLLFVRKEDKCLMCFSVLFQIQTIEEIRGDSINPEYCCVA
K*AFPTKKLCVFKESNSILKQHD*APQNKVSATRQQSFGKVLGLTISVLCPLPSGSELICYDKYFIWGV
FWYWTINRATADCQRQFSX

Figure 3.10: BUC10 predicted protein sequence.
The nucleotide sequence was translated into a protein sequence using the Translator tool of the JustBio website (http://www.justbio.com/translator/index.php). The translation was carried out in 3 phases from 5' end of nucleotide sequence, using the standard genetic code. Stars represent stop codons.

BUC11 nucleotide sequence

GTTGAAGATATGAGTGCGGGGCTCATCTATCCCTGGAATTGTCTTTCCCACAATCCCTGACACAGAATA
TGAGCCATACAGGAATTCTGAAGAAATGGGTCTCTTGCCACCTCCCAGTAAAAGATTATTTTTTAAAAA
AAAAAGGCTCTGCTTTGACCTGAAGTATTTTATCTATCCTCAGTCTCAGGACACTGTTGATGGAATTAA
GGCCAAGCACATCTGCAAAAAAGACATTGCTGGAGGAGGTGCAAAGAGCTGGAAACCAAGTCTCCAGTC
CTGGGAAAAGCAGTGGTATGGAAAAGCAATGGAAAGAGCATTTTGAAAATGCCATTCCACTGTTTTCTG
GCCTTTATGATTTCTGCTGAGAAATCCACTGTTAGTCTGATGGGGTCTCCTTCATAGCACCAATGACCT
GAAGAGCCTTGTTGAAGGAAGACTCCATCTGATGACTCAGAGCAAGTATTTTTTAGTGTGTTATTGTTA
TTAGCAGAAAGAGGGCCATAAAATACATGGGGCAAGCTGAATATATCTTAGGCAAAAGAAGAAAATATT
CAAATTCTTATGTTATTTTATCTAATTATTTTATCTCTTTTTGTGTGTGACTTATAATGTGTGTATTGT
ATTAATAAAAGTATATAAACATGTAGTTTAC

Figure 3.11: BUC11 nucleotide sequence.
This is the full mRNA sequence for BUC11, obtained following sequencing and alignments in Spidey alignment tool. The mRNA sequence is 652 base-pair-long.

BUC11 translated protein sequence (3 phase translation)

Phase 1:

VEDMSAGLIYPWNCLSHNP*HRI*AIQEF*RNGSLATSQ*KIIF*KKKALL*PEVFYLSSVSGHC*WN*
GQAHLQKRHCWRRCKELETKSPVLGKAVVWKSNGKSILKMPFHCFLAFMISAEKSTVSLMGSPS*HQ*P
EEPC*RKTPSDDSEQVFFSVLLLLAERGP*NTWGKLNIS*AKEENIQILMLFYLIILSLFVCDL*CVYC
INKSI*TCSLX

Phase 2:

LKI*VRGSSIPGIVFPTIPDTEYEPYRNSEEMGLLPPPSKRLFFKKKRLCFDLKYFIYPQSQDTVDGIK
AKHICKKDIAGGGAKSWKPSLQSWEKQWYGKAMERAF*KCHSTVFWPL*FLLRNPLLV*WGLLHSTNDL
KSLVEGRLHLMTQSKYFLVCYCY*QKEGHKIHGAS*IYLRQKKKIFKFLCYFI*LFYLFLCVTYNVCIV
LIKVYKHVVYX

Phase 3:

*RYECGAHLSLELSFPQSLTQNMSHTGILKKWVSCHLPVKDYFLKKKGSALT*SILSILSLRTLLMELR
PSTSAKKTLLEEVQRAGNQVSSPGKSSGMEKQWKEHFENAIPLFSGLYDFC*EIHC*SDGVSFIAPMT*
RALLKEDSI**LRASIF*CVIVISRKRAIKYMGQAEYILGKRRKYSNSYVILSNYFISFCV*LIMCVLY
**KYINM*FX

Figure 3.12: BUC11 predicted protein sequence.
The nucleotide sequence was translated into a protein sequence using the Translator tool of the JustBio website (http://www.justbio.com/translator/index.php). The translation was carried out in 3 phases from 5' end of nucleotide sequence, using the standard genetic code. Stars represent stop codons.

The nucleotide sequences as well as the predicted protein sequences were compared to database entries in order to ascertain whether or not the identified BUC genes had already been published or were novel. The sequences were uploaded on the BLAST search page (http://blast.ncbi.nlm.nih.gov/Blast.cgi). For the nucleotide sequences, the blastn (somewhat similar sequences) program was used with the options: "all GenBank+EMBL+DDBJ +PDB sequences (but no EST, STS, GSS,environmental samples or phase 0, 1 or 2 HTGS sequences)" and the search set "Nucleotide collection (nr/nt)". For the predicted protein sequences, the blastp (protein-protein BLAST) program was used with the options: "All

non-redundant GenBank CDS translations+PDB+SwissProt+PIR+PRF excluding environmental samples from WGS projects" and the search set "Non-redundant protein sequences (nr)". The top 3 most significant search results for the similarity of the BUC nucleotide sequences to other entries in the database and the top 2 most significant search results for the similarity of the BUC predicted protein sequences to other entries in the database have been compiled in Table 3.1, Table 3.2, Table 3.3 and Table 3.4 for BUC6, BUC9, BUC10 and BUC11 respectively.

Nucleotide similarity (GenBank accession number)		
Description	Query coverage	Maximum identity
Human DNA sequence from clone RP11-20F24 on chromosome 10p11.21-12.1 Contains the 3' end of the NY-BR-1 gene for breast cancer antigen NY-BR-1, two novel genes, a pseudogene similar to part of ATP8A2 (ATPase, aminophospholipid transporter-like, Class I, type 8A, member 2), a novel pseudogene and a vomeronasal receptor pseudogene, complete sequence (AL157387.9)	100%	100%
MACACA MULATTA BAC clone CH250-103G5 from chromosome unknown, complete sequence (AC209749.3)	92%	96%
MACACA MULATTA BAC clone CH250-14B20 from chromosome unknown, complete sequence (AC205191.3)	68%	96%
Phase 1 Protein similarity (GenBank accession number)		
Description	Query coverage	Maximum identity
unnamed protein product [Homo sapiens] (BAC85788.1)	70/807	77%
PREDICTED: hypothetical protein isoform 2 [Pan troglodytes] (XP_515750.2)	70/807	72%
Phase 2 Protein similarity (GenBank accession number)		
Description	Query coverage	Maximum identity
LAMA5 protein [Homo sapiens] (AAH85017.1)	43/807	76%
hCG2042294 [Homo sapiens] (EAW95054.1)	46/807	69%
Phase 3 Protein similarity (GenBank accession number)		
Description	Query coverage	Maximum identity
LAG1 longevity assurance homolog 5 (S. cerevisiae), isoform CRA_a [Homo sapiens] (EAW58131.1)	84/807	70%
unnamed protein product [Macaca fascicularis] (BAE89454.1)	83/807	75%

Table 3.1: Nucleotide and predicted protein sequence similarity of BUC6 to nucleotide and protein sequences previously deposited in the GenBank database.
The nucleotide sequence was uploaded in blastn and the predicted protein sequence was uploaded in blastp (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The top 3 most significant search results for the nucleotide sequence and the top 2 most significant search results for each of the 3 phases of translated protein sequence are shown. The Query Coverage in the Protein Similarity section is shown as the number of amino acids of the query matching the amino acids of the database entries/total number of amino acids of the query.

Nucleotide similarity (GenBank accession number)		
Description	Query coverage	Maximum identity
Human DNA sequence from clone RP11-20F24 on chromosome 10p11.21-12.1 Contains the 3' end of the NY-BR-1 gene for breast cancer antigen NY-BR-1, two novel genes, a pseudogene similar to part of ATP8A2 (ATPase, aminophospholipid transporter-like, Class I, type 8A, member 2), a novel pseudogene and a vomeronasal receptor pseudogene, complete sequence (AL157387.9)	100%	100%
Human DNA sequence from clone RP11-467D10 on chromosome 13 Contains part of the ATP8A2 gene for ATPase, aminophospholipid transporter-like, Class I, type 8A, member 2 (ML-1 protein (IB, ATP, ML-1, ATPIB, DKFZP434B1913)) and a novel pseudogene, complete sequence (AL157366.12)	88%	87%
Homo sapiens cDNA clone IMAGE:9052749 (BC144229.1)	32%	91%
Phase 1 Protein similarity (GenBank accession number)		
Description	Query coverage	Maximum identity
hCG1789874 [Homo sapiens] (EAW85905.1)	113/258	100%
rCG52191 [Rattus norvegicus] (EDL85250.1)	43/258	81%
Phase 3 Protein similarity (GenBank accession number)		
Description	Query coverage	Maximum identity
PREDICTED: similar to ATPase, aminophospholipid transporter (APLT) [Rattus norvegicus] (XP_223390.4)	42/257	42%
hypothetical protein IL0605 [Idiomarina loihiensis L2TR] (YP_154995.1)	51/257	39%

Table 3.2: Nucleotide and predicted protein sequence similarity of BUC9 to nucleotide and protein sequences previously deposited in the GenBank database.
The nucleotide sequence was uploaded in blastn and the predicted protein sequence was uploaded in blastp (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The top 3 most significant search results for the nucleotide sequence and the top 2 most significant search results for phase 1 and phase 3 of translated protein sequence are shown. Search for phase 2 did not result in any significant similarity to any known protein sequences. The Query Coverage in the Protein Similarity section is shown as the number of amino acids of the query matching the amino acids of the database entries/total number of amino acids of the query.

Nucleotide similarity (GenBank accession number)		
Description	Query coverage	Maximum identity
Human DNA sequence from clone RP11-20F24 on chromosome 10p11.21-12.1 Contains the 3' end of the NY-BR-1 gene for breast cancer antigen NY-BR-1, two novel genes, a pseudogene similar to part of ATP8A2 (ATPase, aminophospholipid transporter-like, Class I, type 8A, member 2), a novel pseudogene and a vomeronasal receptor pseudogene, complete sequence (AL157387.9)	100%	100%
MACACA MULATTA BAC clone CH250-14B20 from chromosome unknown, complete sequence (AC205191.3)	100%	92%
MACACA MULATTA BAC clone CH250-103G5 from chromosome unknown, complete sequence (AC209749.3)	100%	92%
Phase 1 Protein similarity (GenBank accession number)		
Description	Query coverage	Maximum identity
unnamed protein product [Homo sapiens] (BAH13866.1)	39/641	82%
PREDICTED: ATPase, aminophospholipid transporter-like, Class I, type 8A, member 2 isoform 1 [Macaca mulatta] (XP_001092554.1)	39/641	82%

Table 3.3: Nucleotide and predicted protein sequence similarity of BUC10 to nucleotide and protein sequences previously deposited in the GenBank database.
The nucleotide sequence was uploaded in blastn and the predicted protein sequence was uploaded in blastp (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The top 3 most significant search results for the nucleotide sequence and the top 2 most significant search results for phase 1 of translated protein sequence are shown. Searches for phases 2 and 3 did not result in any significant similarity to any known protein sequences. The Query Coverage in the Protein Similarity section is shown as the number of amino acids of the query matching the amino acids of the database entries/total number of amino acids of the query.

Nucleotide similarity (GenBank accession number)		
Description	Query coverage	Maximum identity
PREDICTED: Homo sapiens hCG25653 (LOC646360), mRNA (XM_001724231.1)	100%	99%
PREDICTED: Homo sapiens hCG25653 (LOC646360), mRNA (XM_944750.2)	100%	99%
PREDICTED: Homo sapiens hCG25653 (LOC646360), mRNA (XM_933553.2)	100%	99%
Phase 1 Protein similarity (GenBank accession number)		
Description	Query coverage	Maximum identity
PREDICTED: ATPase, aminophospholipid transporter-like, Class I, type 8A, member 2 isoform 2 [Macaca mulatta] (XP_001092668.1)	34/218	55%
PREDICTED: ATPase, aminophospholipid transporter-like, Class I, type 8A, member 2 isoform 1 [Macaca mulatta] (XP_001092554.1)	34/218	55%
Phase 2 Protein similarity (GenBank accession number)		
Description	Query coverage	Maximum identity
PREDICTED: similar to hCG25653 [Homo sapiens] (XP_938646.1)	75/218	100%
PREDICTED: T32E20.19 [Arabidopsis thaliana] (AC020646_27)	101/218	27%
Phase 3 Protein similarity (GenBank accession number)		
Description	Query coverage	Maximum identity
hCG25653, isoform CRA_a [Homo sapiens] (EAW85902.1)	30/217	100%

Table 3.4: Nucleotide and predicted protein sequence similarity of BUC11 to nucleotide and protein sequences previously deposited in the GenBank database.
The nucleotide sequence was uploaded in blastn and the predicted protein sequence was uploaded in blastp (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The top 3 most significant search results for the nucleotide sequence and the top 2 most significant search results for each of the 3 phases of translated protein sequence are shown. There was only one search result for phase 3. The Query Coverage in the Protein Similarity section is shown as the number of amino acids of the query matching the amino acids of the database entries/total number of amino acids of the query.

To summarise, the common first hit for BUC6, BUC9 and BUC10 nucleotide sequences is, as expected, the genomic Human DNA sequence from clone RP11-20F24 on chromosome 10p11.21-12.1 (GenBank accession number AL157387.9). As shown in Table 3.1, BUC6 nucleotide sequence is highly similar to clone CH250-103G5 and clone CH250-14B20, both DNA isolated from the blood cells of a healthy rhesus monkey (Macaca mulatta). Also, BUC6 predicted protein sequence has weak similarities with the human laminin alpha 5 (LAMA5), an extracellular glycoprotein. Most impor-

tantly, BUC6 sequence matches the 3' end of the NY-BR-1 gene but not in the coding region (intron). Regarding BUC9, its nucleotide sequence is strongly similar to a human DNA sequence on chromosome 13 and weakly similar to an uncharacterised human cDNA clone. BUC9 predicted protein sequence is quite similar to the predicted protein hCG1789874. As for BUC6 nucleotide sequence, BUC10 nucleotide sequence is highly similar to CH250-103G5 and clone CH250-14B20 of the rhesus monkey. BUC10 protein sequence is weakly similar to a human unnamed protein product. More interestingly, 99% of BUC11 nucleotide sequence matches the complete mRNA sequence of the unpublished human-specific gene hCG25653 (mentioned earlier), as shown in Figure 3.13. The predicted BUC11 protein (phase 2 of translation) is 100% homologous to the predicted protein of the gene hCG25653 (75 amino-acid-long). It is important to mention that the sequencing results for BUC11 (data not shown) originally matched only up to 47% of hCG25653 gene sequence, with 100% sequence coverage. However, just after sequencing of the BUC genes and some mRNA expression analysis, specific primers for the coding region of hCG25653 were designed and used to amplidy cDNA for cloning into an expression vector (Chapter 5). Only one PCR product was obtained so it was concluded that the sequencing results of BUC11 included artifact sequences such as genomic. Therefore, the correct nucleotide sequence of BUC11 (652bp) shown here without the artifacts matched up to 99% hCG25653 nucleotide sequence thus the predicted protein for BUC11 and the predicted protein of hCG25653 are identical. The two nucleotide difference is due to artifacts included in the original sequencing.

To conclude, the "BUC gene" mRNA sequence is composed, at its 5' end of the full BUC6 nucleotide sequence in the 3' end of the NY-BR-1 gene, but in the non-coding region, and, at its 3' end, of the BUC11 nucleotide sequence which 99% of it corresponds to the mRNA sequence of the gene hCG25653. A final Spidey alignment result of the mRNA sequences for the BUC genes, the NY-BR-1 gene and the hCG25653 gene against the genomic Human DNA sequence from clone RP11-20F24 on chromosome 10p11.21-12.1 is presented in Figure 3.14.

hCG25653 mRNA sequence

GTTGAAGAGATGAGTGCGGGGCTCATCTATCCCTGGAATTGTCTTTCCCACAATCCCTGACACAGAATA
TGAGCCATACAGGAATTCTGAAGAAATGGGTCTCTTGCCACCTCCCAGTAAAAGATTATTTTTTAAAAA
AAAAAGGCTCTGCTTTGACCTGAAGTATTTTATCTATCCTCAGTCTCAGGACACTGTTGATGGAATTAA
GGCCAAGCACATCTGCAAAAAAGACATTGCTGGAGGAGGTGCAAAGAGCTGGAAACCAAGTCTCCAGTC
CTGGGAAAAGCAGTGGTATGGAAAAGCAATGGAAAGAGCATTTTGAAAATGCCATTCCACTGTTTTCTG
GCCTTTATGATTTCTGCTGAGAAATCCACTGTTAGTCTGATGGGGTCTCCTTCATAGCACCAATGACCT
GAAGAGCCTTGTTGAAGGAAGACTCCATCTGATGACTCAGAGCAAGTATTTTTTAGTGTGTTATTGTTA
TTAGCAGAAAGAGGGCCATAAAATACATGGGGCAAGCTGAATATATCTTAGGCAAAAGAAGAAAATATT
CAAATTCTTATGTTATTTTATCTAATTATTTTATCTCTTTTTGTGTGTGACTTATAATGTGTGTATTGT
ATTAATAAAAGTATATAAACATGTAGTTTACA

hCG25653 predicted protein sequence

MGLLPPPSKRLFFKKKRLCFDLKYFIYPQSQDTVDGIKAKHICKKDIAGGGAKSWKPSLQSWEKQWYGK
AMERAF

Figure 3.13: hCG25653 mRNA and predicted protein sequences.
The GenBank accession number is XM_001724231.1. The complete mRNA sequence of the gene hCG25653 is 653 base-pair-long. The coding region is predicted to start at nucleotide 95 and end at nucleotide 322. The predicted protein sequence is 75 amino-acid-long. The nucleotides highlighted in red shows the differences between hCG25653 mRNA sequence and BUC11 mRNA sequence. In BUC11 mRNA sequence, the first nucleotide highlighted was found a Thymine and the last nucleotide is not present.

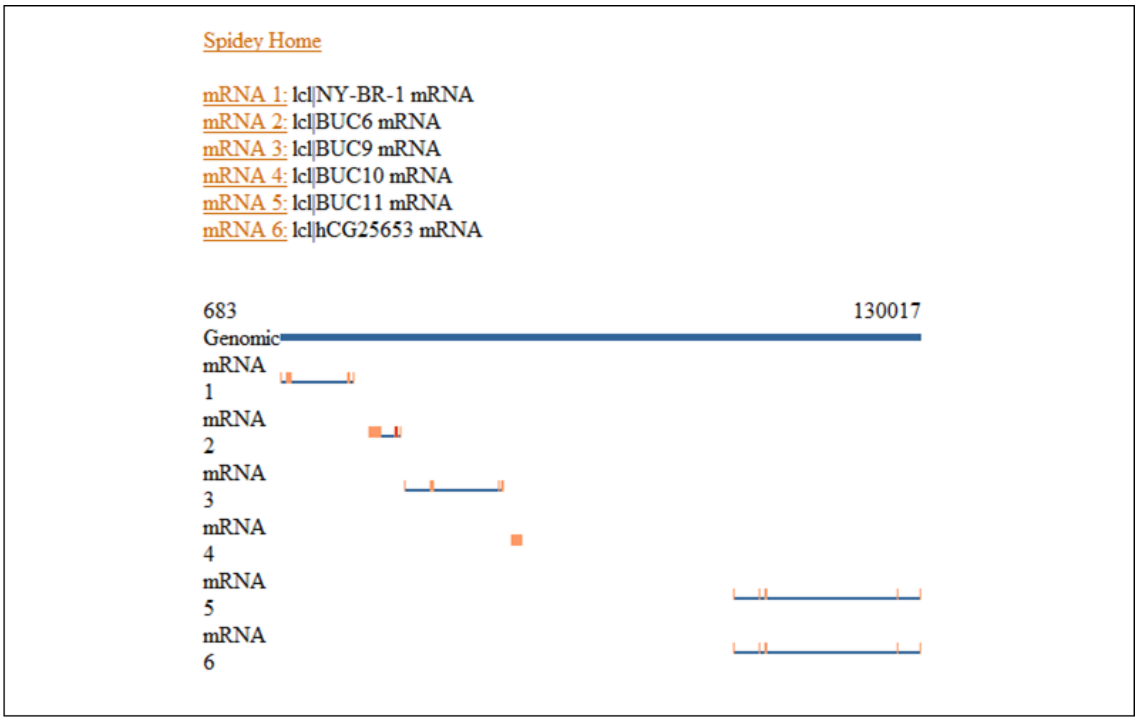


Figure 3.14: Spidey alignment of the BUC, NY-BR-1 and hCG25653 mRNA sequences.
After sequencing and BLAST analysis, the mRNA sequences of BUC6, BUC9, BUC10 and BUC11, along with the mRNA sequences of NY-BR-1 and hCG25653 genes, in FASTA format, were aligned in Spidey alignment tool against the genomic Human DNA sequence from clone RP11-20F24 on chromosome 10p11.21-12.1.

3.2.2 mRNA expression analysis of the BUC family

As stated previously, Dr Geng Li has mined the UniGene database for ESTs from normal breast, and/or breast cancer tissues or cell lines but not from normal essential tissues or organs except testis, ovary and placenta. Prior to this project, Dr Geng Li used both semi-quantitative and quantitative Real-Time RT-PCR to analyse the expression patterns of the 228 BUC genes selected in a panel of RNAs from normal, malignant tissues and breast tumour cell lines. Out of the 228 BUC genes, BUC6, BUC9, BUC10 and BUC11 were validated and further selected because of their high expression in breast cancer, normal breast and testes but not in other normal tissues. The semi-quantitative and quantitative Real-Time RT-PCR results from his own study (personal communication with Dr Li) are summarised in Figure 3.15, Table 3.5 and Figure 3.16. The reference gene GAPDH was used to normalise the mRNA levels between the samples. Briefly, the semi-quantitative RT-PCR results indicated that BUC6 was not expressed in 21 different normal foetal and adult tissues including brain, lung, liver and peripheral blood mononuclear cells (PBMC) but highly expressed in normal breast tissues, testis cancer tissues, breast cancer tissues and one breast cancer cell line (T47D). Moreover, the analysis of the results of the quantitative Real-Time RT-PCR performed with BUC6 primers on a RNA panel derived from various normal and tumour specimens revealed very low levels of BUC6 mRNA in normal tissues except testis tissues. BUC6 mRNA expression level in all breast cancer tissues tested was high with, as an example, BUC6 mRNA level in Br4 that was found to be 131,072 times the BUC6 mRNA level detected in liver. Using semi-quantitative RT-PCR, a similar pattern of mRNA expression for BUC9, BUC10 and BUC11 was uncovered. BUC9 and BUC10 mRNA were highly expressed in the normal testis tissue and all the breast cancer tissues tested. BUC9 and BUC10 mRNA were not expressed in brain, heart, liver and lung however they were found at a low level in PBMC. BUC11 mRNA was not expressed in the normal tissues tested, including the one testis tissue used. However, it was expressed in 3 out of 4 breast cancer tissues tested.

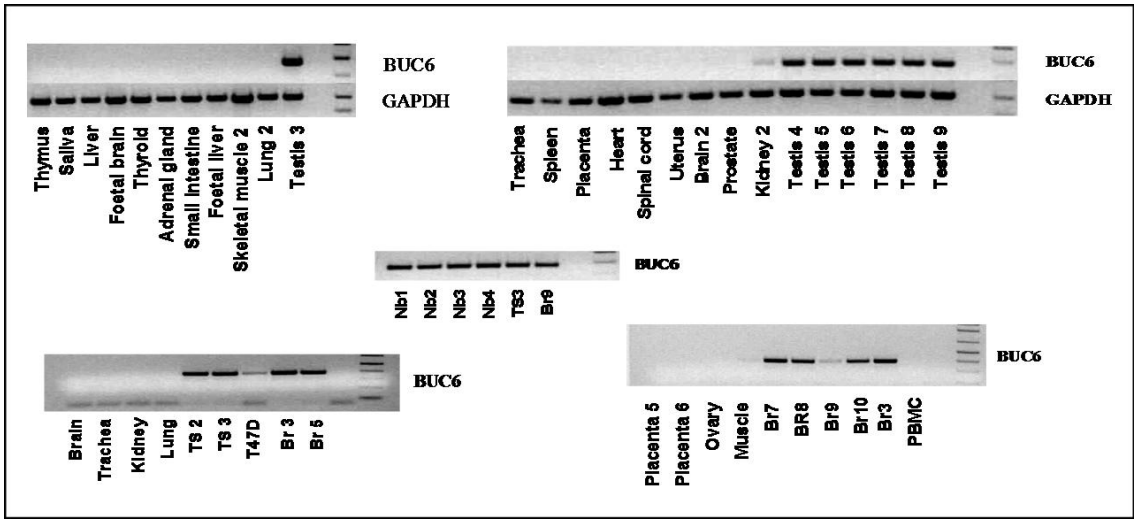


Figure 3.15: BUC6 mRNA expression analysis by semi-quantitative RT-PCR.
BUC6 mRNA expression analysis in a panel of normal tissues, breast cancer tissues and one breast cancer cell line. The housekeeping gene GAPDH was used to normalise the mRNA levels between the samples. The top left and top right pictures show GAPDH and BUC6 mRNA expression in a panel of normal tissues. The picture in the middle shows BUC6 mRNA expression in four normal breast tissues (Nb1-4), in one normal testis tissue (TS3) and one breast cancer tissue (Br9). The bottom left picture shows BUC6 mRNA expression in four normal tissues (brain, trachea, kidney and lung), two cancer testis tissues (TS2 and TS3), one breast cancer cell line (T47D) and two breast cancer tissues (Br3 and Br5). The bottom right picture shows BUC6 mRNA expression in five normal tissues (placenta5, placenta6, ovary, muscle and peripheral blood mononuclear cells (PBMC)) and five breast cancer tissues (Br7, BR8, Br9, Br10 and Br3).

Tissues	GAPDH	ΔCt (GAPDH)	BUC6	ΔCt (BUC6)	Δ ΔCt	BUC6 expression‡
Brain	17.6*	-2.3	33.8	-3.0	-0.7	×2
Lung	18.8	-1.1	34.9	-1.9	-0.8	×2
Trachea	17.7	-2.2	32.2	-4.6	-2.3	×4
Heart	17.8	-2.1	36.8	0.0	2.1	×0.25
Liver†	19.9	0.0	36.8	0.0	0.0	×1
Kidney	16.7	-3.2	32.9	-3.9	-0.7	×2
TS 2	18.7	-1.2	22.4	-14.4	-13.2	×8192
TS 3	17.7	-2.2	22.5	-14.3	-12.1	×4096
TS 4	17.7	-2.2	22.7	-14.1	-11.9	×4096
TS 5	18.9	-1.0	28.3	-8.5	- 7.5	×256
TS 6	18.1	-1.9	23.5	-13.3	-11.4	×2048
TS 7	17.8	-2.1	22.9	-13.9	-11.8	×4096
Br 1	20.0	0.1	29.6	-7.2	-7.3	×128
Br 2	18.3	-1.6	29.4	-7.4	-5.8	×64
Br 3	15.8	-4.1	20.3	-16.5	-12.4	×4096
Br 4	19.7	-0.2	19.6	-17.2	-17.0	×131072
T47D	16.0	-3.9	27.4	-9.4	-5.5	×64

* Ct value
†Reference tissue
‡Compared with liver

Table 3.5: BUC6 mRNA expression analysis by quantitative RT-PCR.
BUC6 mRNA expression analysis in a panel of normal tissues, breast cancer tissues and one breast cancer cell line. The samples used include 6 essential normal tissues (brain, lung, trachea, heart, liver and kidney), 3 normal testis tissues (TS2-4), 3 testis cancer tissues (TS5-7), 4 breast cancer tissues (Br1-4) and one breast cancer cell line (T47D). The analysis of the results was carried out using the comparative Ct method. The reference gene GAPDH was used to normalise changes in specific gene expression. The last column shows the normalised level of BUC6 mRNA expression in the tissues and cell line, relative to its expression level in normal liver (reference tissue).

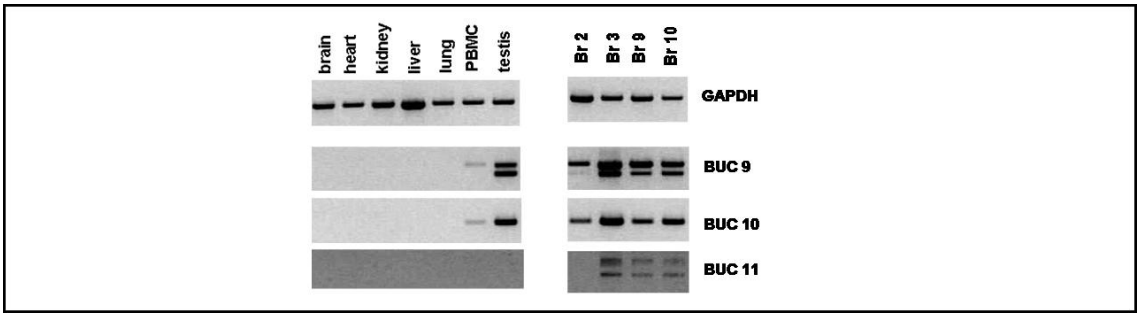


Figure 3.16: BUC9, BUC10 and BUC11 mRNA expression analysis by semi-quantitative RT-PCR.
BUC9, BUC10 and BUC11 mRNA expression analysis in a panel of normal tissues and breast cancer tissues. The reference gene GAPDH was used to normalise the mRNA levels between the samples. The samples used include seven normal tissues (brain, heart, kidney, liver, lung, peripheral blood mononuclear cells (PBMC) and testes) and four breast cancer tissues (Br2, Br3, Br9 and Br10).

As illustrated, Dr Geng Li has studied the four BUC genes mRNA expression patterns in a variety of normal and tumour specimens prior to this project however a more thorough analysis was required before considering that one or more BUC genes could be of interest for breast cancer management. The BUC genes potential for diagnosis, therapy (gene therapy or immunotherapy) and detection of metastatic breast cancer cells (in peripheral blood or lymph nodes of patients) depends partly on the restricted expression patterns of the genes in certain tissues. As stated before, ideally, they must not be expressed in essential normal tissues and must be highly expressed and up-regulated in tumour tissues compared to normal counterpart tissues. Consequently, the mRNA expression analysis of BUC6, BUC9, BUC10 and BUC11 was conducted using a large number of normal tissue RNAs (purchased) and tumour tissue RNAs of varied origins (generous donations of collaborators). The results of the parallel study of the full length BUC sequences and their similarities with known genes or proteins (results shown in section 3.2.1) justified that the mRNA expression study became gradually focused on BUC11 and thus explains the larger number of analysis carried out on BUC11 compared to BUC6, BUC9 and BUC10. Non-, semi-quantitative and quantitative RT-PCR were performed with the primers and conditions detailed in Chapter 2.
Briefly, 2μg of RNA (or entire sample if the RNA concentration was low) and Oligo(dT)$_{15}$ or random primers were used to generate cDNA by reverse-transcription. Following completion of the semi-quantitative RT-PCR reaction, products were loaded and run on a 1.5% agarose gel. The housekeeping gene GAPDH was used to normalise the mRNA levels between the samples. Following completion of the quantitative RT-PCR reaction, the Ct value for each sample and the standard curve were generated. The standard quantity (SQ) mean for the BUC gene was divided by the SQ mean of the housekeeping gene GAPDH which resulted in the normalised mRNA expression of the BUC gene in the tissues. The summary of the mRNA expression analysis, conducted during the project and prior to

the project by Dr Geng Li, for all the BUC genes is presented in Table 3.6. Examples of the results obtained in the non- or semi-quantitative RT-PCR analysis are presented in Figure 3.17, Figure 3.18, Figure 3.19, Figure 3.20, Figure 3.21 and Figure 3.22. These experiments are semi-quantitative as cDNA concentrations have been adjusted to ensure a similar copy number for GAPDH, except the ones presented in Figure 3.17 and (B) of Figure 3.22.

Normal tissues	**BUC6**	**BUC9**	**BUC10**	**BUC11**
Panel of essential tissues	0/22	0/6	6/14	0/6
Placenta	0/3	1/1	1/1	0/1
Breast	4/4	2/2	nd	1/1
Testes	6/6	1/1	nd	3/3
Stomach	nd	2/4	nd	0/4
Colon	nd	nd	nd	nd
Oesophagus	nd	4/4	nd	0/3
Kidney	nd	nd	nd	0/4
PBMC	0/1	1/1	1/1	0/6
Cancer tissues	**BUC6**	**BUC9**	**BUC10**	**BUC11**
Melanoma	7/7	7/7	7/7	0/7
Mesothelioma	nd	nd	nd	0/8
Breast	9/9	6/9	10/11	10/11
Testes	nd	nd	nd	3/3
Stomach	3/3	2/4	5/5	0/4
Colon	2/2	nd	nd	nd
Oesophagus	nd	4/4	4/5	0/3
Kidney	nd	nd	nd	0/4
Head and Neck	nd	nd	6/8	nd
Breast cell lines	4/5	4/7	4/5	1/5

Table 3.6: Summary table of the exhaustive mRNA analysis of the BUC genes.
This table sums up the results from the intensive mRNA expression analysis of all BUC genes, using the non- or semi-quantitative and quantitative RT-PCR techniques. It illustrates the number of individual tissues showing positive expression, for each of the genes of interest. “nd” means “not determined”. RNA samples from gastric tissues (cancer and normal), colon tissues (cancer and normal) and breast tissues (cancer and normal, different stages) have been provided by Dr. Aija Line from the Latvian University Biomedical Centre (Latvia). The tissues are paired cancer and normal counterpart, each pair from the same patient. Oesophageal and kidney tissues were also provided as pairs. PBMC: peripheral blood mononuclear cells. Breast cancer cell lines: SKBR3, MDAP3, MDA435, T47D, BR293, MDA231 and MDA468.

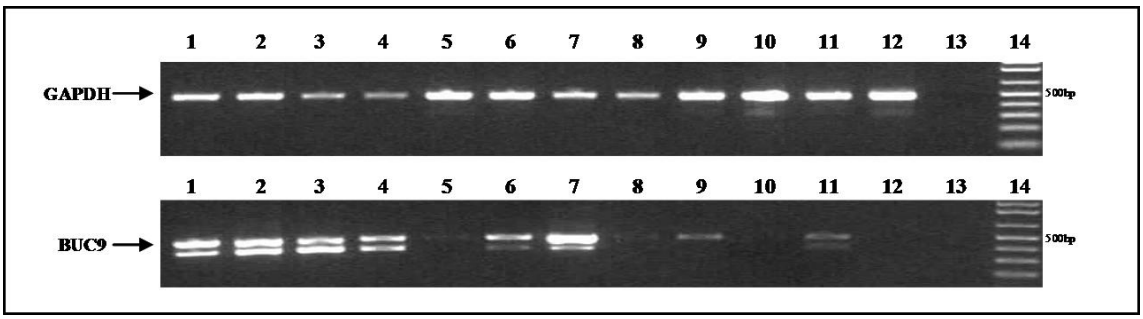


Figure 3.17: BUC9 mRNA expression analysis in breast cancer tissues and cell lines.
BUC9 mRNA expression analysis in breast cancer tissues and cell lines by semi-quantitative RT-PCR. The housekeeping gene GAPDH was used to normalise the mRNA levels between the samples. Lanes: (1) to (8) Breast cancer tissues: (1) BR11, (2) BR12, (3) BR13, (4) BR15, (5) BR16, (6) BR19, (7) BR25, (8) BR28. (9) to (12): breast cancer cell lines: (9) MDAP3, (10) MDA435, (11) T47D, (12) BR293. (13) Negative control, (14) DNA ladder. Bands were present at approximately 500 base pairs which correspond to the expected size (495 base pairs) of the target for the BUC9 primers used. This work was undertaken by Julieta De Silva (BSc student).

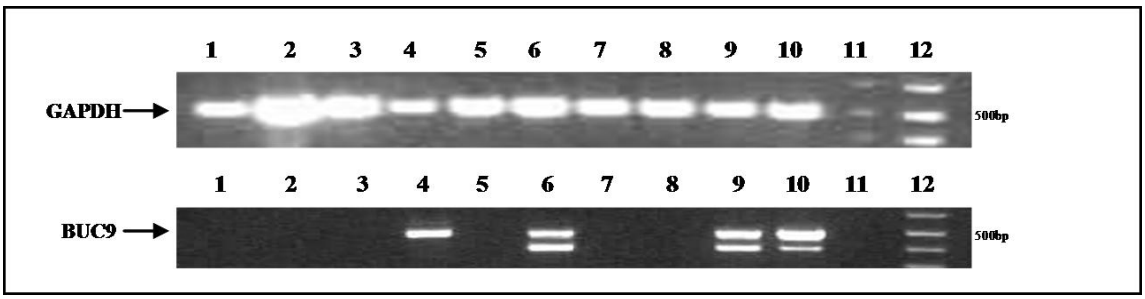


Figure 3.18: BUC9 mRNA expression analysis in normal tissues.
BUC9 mRNA expression analysis in normal tissues by semi-quantitative RT-PCR. The housekeeping gene GAPDH was used to normalise the mRNA levels between the samples. Lanes: (1) Brain, (2) Skeletal muscle, (3) Kidney, (4) Placenta, (5) Heart, (6) Testis, (7) Lung, (8) Liver, (9) Breast 1, (10) Breast 2, (11) Negative control, (12) DNA ladder. Bands were present at approximately 500 base pairs which correspond to the expected size (495 base pairs) of the target for the BUC9 primers used. This work was undertaken by Julieta De Silva (BSc student).

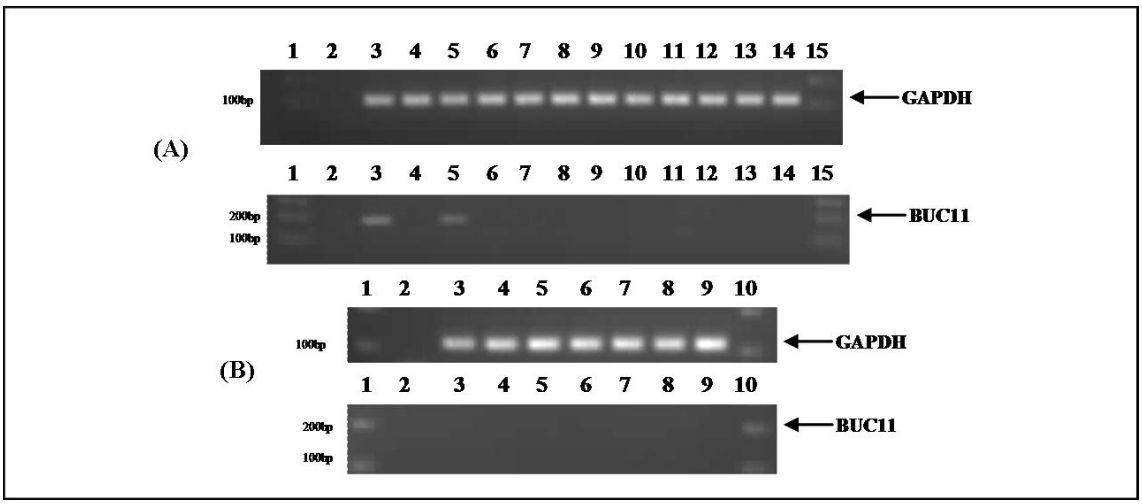


Figure 3.19: BUC11 mRNA expression analysis in normal tissues.
BUC11 mRNA expression analysis in normal tissues by semi-quantitative RT-PCR. The housekeeping gene GAPDH was used to normalise the mRNA levels between the samples. Figure (A) lanes: (1) DNA ladder, (2) Negative control, (3) Breast, (4) Heart, (5) Testis, (6) Liver, (7) Prostate, (8) Brain, (9) Uterus, (10) Spleen, (11) Skeletal muscle, (12) Lung, (13) Kidney, (14) Placenta, (15) DNA ladder. Figure (B) lanes: (1) DNA ladder, (2) Negative control, (3) Fetal brain, (4) Fetal liver, (5) Trachea, (6) Thyroid, (7) Spinal cord, (8) Salivary gland, (9) Thymus, (10) DNA ladder. Bands were present at approximately 200 base pairs which correspond to the expected size (197 base pairs) of the target for the BUC11 primers used.

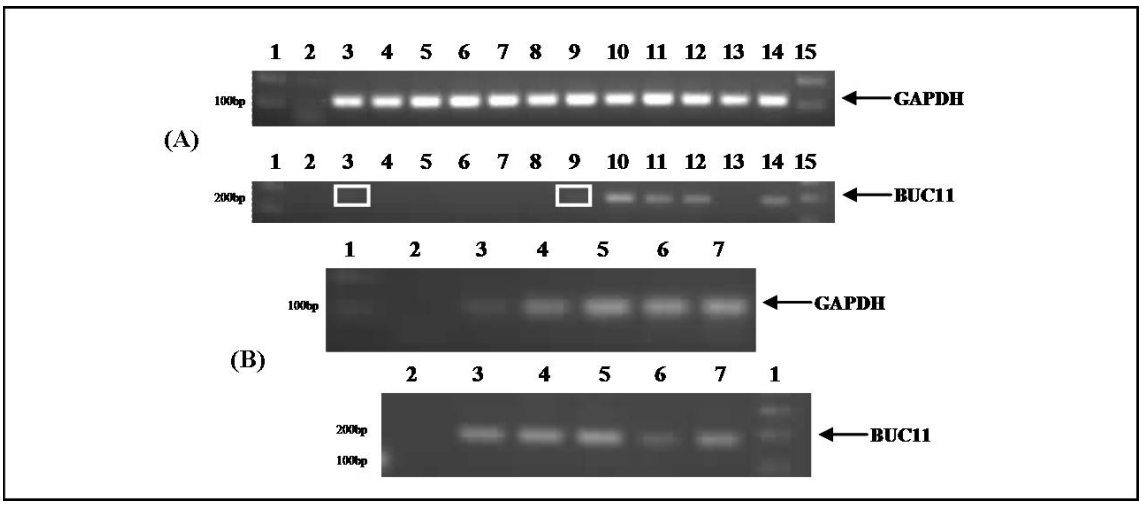


Figure 3.20: BUC11 mRNA expression analysis in breast cancer tissues and cell lines. BUC11 mRNA expression analysis in breast cancer tissues and cell lines by semi-quantitative RT-PCR. The housekeeping gene GAPDH was used to normalise the mRNA levels between the samples. Figure (A) lanes: (1) DNA ladder, (2) Negative control, (3) to (8) cell lines: (3) T47D, (4) MDA231, (5) MDA468, (6) BR293, (7) MCF7, (8) MDAP3, (9) to (14) tissues: (9) BR2, (10) BR15, (11) BR19, (12) BR11, (13) BR20, (14) BR13, (15) DNA ladder. Figure (B) lanes: (1) DNA ladder, (2) Negative control, (3) to (7) tissues: (3) BR25, (4) BR26, (5) BR27, (6) BR9, (7) BR12. Bands were present at approximately 200 base pairs which correspond to the expected size (197 base pairs) of the target for the BUC11 primers used.

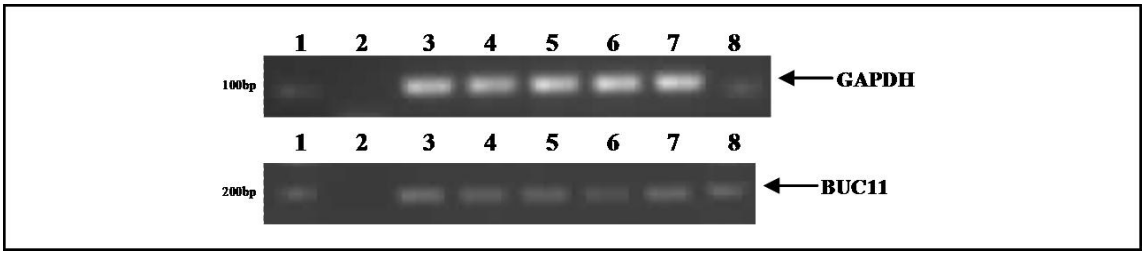


Figure 3.21: BUC11 mRNA expression analysis in testis (normal and cancer) tissues. BUC11 mRNA expression analysis in in testis (normal and cancer) tissues by semi-quantitative RT-PCR. The housekeeping gene GAPDH was used to normalise the mRNA levels between the samples. Lanes: (1) DNA ladder, (2) Negative control, (3) to (5) testicular cancer tissues: (3) TS2, (4) TS3, (5) TS4, (6) to (7) normal testis tissues: (6) TS6, (7) TS7, (8) DNA ladder. Bands were present at approximately 200 base pairs which correspond to the expected size (197 base pairs) of the target for the BUC11 primers used.

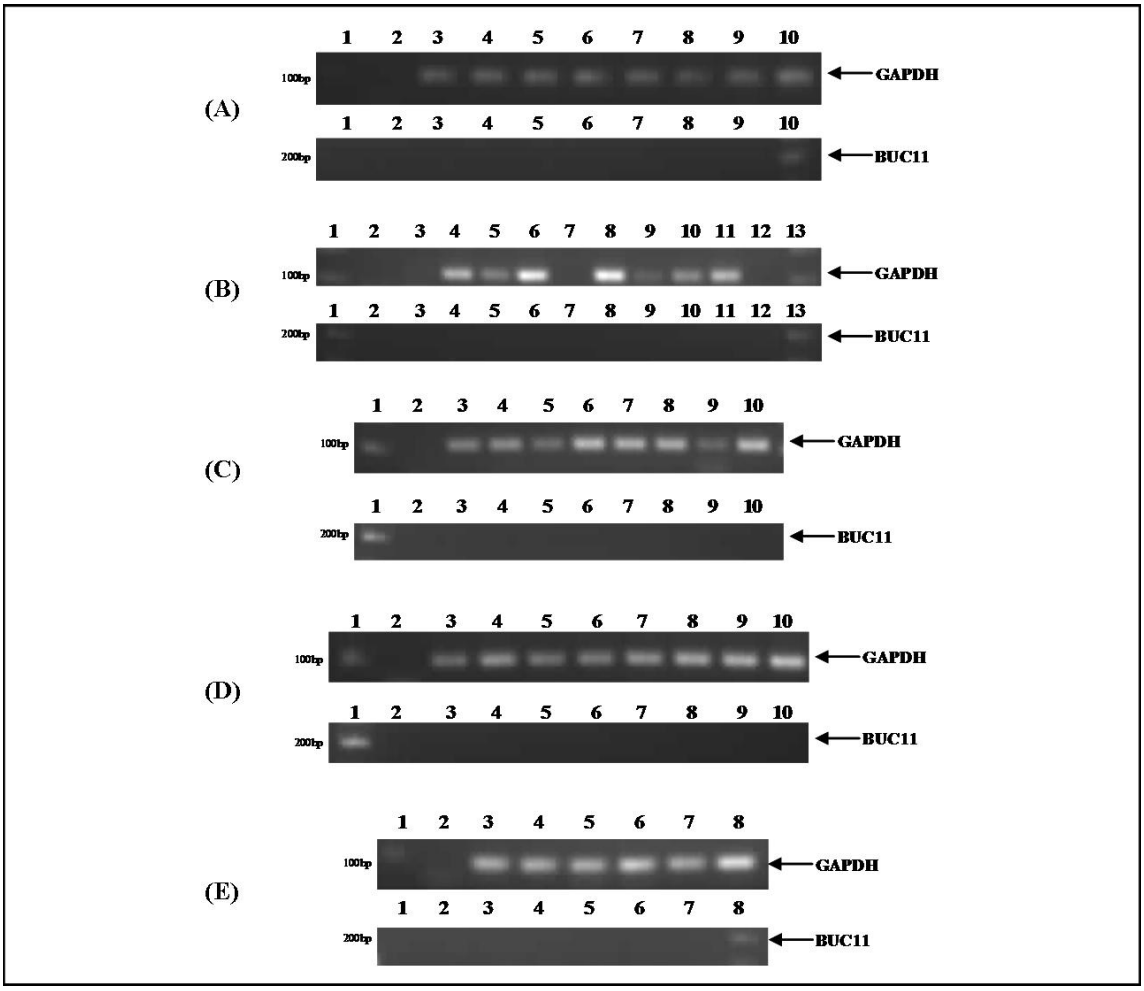


Figure 3.22: BUC11 mRNA expression analysis in diverse cancer tissues.
BUC11 mRNA expression analysis in diverse cancer tissues by semi-quantitative RT-PCR. The housekeeping gene GAPDH was used to normalise the mRNA levels between the samples. Figure (A) lanes: (1) DNA ladder, (2) Negative control, (3) to (10) Mesothelioma tissues: (3) Meso1, (4) Meso2, (5) Meso3, (6) Meso4, (7) Meso5, (8) Meso6, (9) Meso7, (10) Meso8. Figure (B) lanes: (1) DNA ladder, (2) Negative control, (3) to (12) Melanoma samples 1 to 10, (13) DNA ladder. Figure (C) lanes: (1) DNA ladder, (2) Negative control, (3) to (10) Gastric paired normal (N) and cancer (T) tissues: (3) BN2, (4) BT2, (5) BN4, (6) BT4, (7) BN5, (8) BT5, (9) BN6, (10) BT6. Figure (D) lanes: (1) DNA ladder, (2) Negative control, (3) to (10) Kidney paired normal (N) and cancer (T) tissues: (3) KN1, (4) KT1, (5) KN2, (6) KT2, (7) KN10, (8) KT10, (9) KN12, (10) KT12. Figure (E) lanes: (1) DNA ladder, (2) Negative control, (3) to (8) Oesophageal paired normal (N) and cancer (T) tissues: (3) EN1, (4) ET1, (5) EN3, (6) ET3, (7) EN4, (8) ET4.

A more detailed account of the results (Table 3.6) for each individual gene of interest, completing the analysis undertaken by Dr Geng Li prior to this project, is given in the following paragraph. Following semi-quantitative RT-PCR, it was found that BUC6 is expressed in all the melanoma samples tested. BUC6 mRNA is expressed in all of the breast cancer cell lines tested (SKBR3, T47D, MDA468 and MDA231) except BR293, with levels of expression varying from one cell line to another. The quantitative RT-PCR analysis revealed that BUC6 mRNA expression is significantly higher in breast cancer,

when compared to expression in gastric and colon (normal and cancer) samples tested. As seen in Figure 3.17, the two gene products observed during semi-quantitative RT-PCR analysis suggests that BUC9 mRNA may exist as two splice variants in breast cancer tissues and cell lines tested positive except for the breast cancer cell lines MDAP3 and SKBR3. BUC9 mRNA is expressed in 6 out of 9 breast cancer tissues and in the breast cancer cell lines MDAP3, T47D, SKBR3 and MDA231 but not MDA435, BR293 and MDA468. As seen in Figure 3.18, BUC9 mRNA expression was positive only in the following normal tissues: placenta, testis and breast. The results suggest that BUC9 may exist as two splice variants in testis and breast tissues but as a single spliced mRNA in placenta. However, another semi-quantitative RT-PCR experiment (figure not shown) brought to light that BUC9 mRNA is also expressed in all the tested paired oesophageal normal and cancer tissues, with a similar level. BUC9 mRNA is expressed in 2 paired gastric normal and cancer tissues out of 4, with BUC9 overexpressed in normal compared to cancer tissue. Both gastric and oesophageal tissues only expressed the longest splice variant.

Comparable to BUC6 analysis, the melanoma samples all tested positive for the expression of BUC9 mRNA as well as BUC10 mRNA. BUC10 mRNA molecules were found in all the breast cancer cell lines tested (SKBR3, T47D, MDA468 and MDA231) with the exception of BR293. BUC10 mRNA expression was proved to be detectable in 6 out of 14 essential normal tissues. BUC10 mRNA was also tested positive in all gastric cancer tissues, breast cancer tissues, oesophageal cancer tissues and head and neck (normal and cancer) tissues tested, using quantitative RT-PCR. BUC10 mRNA is expressed at different levels in all the samples tested and to a much weaker extend in breast cancer than seen in the other cancer and normal tissues (figure not shown). Following semi-quantitative RT-PCR analysis, it appears that BUC11 mRNA is only expressed in the breast cancer cell line T47D (Figure 3.20) and is found in about 90% of the breast cancer tissues tested. Figure 3.21 shows that all testis (normal and cancer) tissues tested positive for BUC11 mRNA. Only breast and testis tissues tested positive for BUC11 mRNA molecules (Figure 3.19). This observation was validated using quantitative RT-PCR (figure not shown) showing that BUC11 mRNA is not expressed or is at very low levels in the normal tissues tested (heart, liver, prostate, brain, uterus, spleen, skeletal muscle, lung, kidney, placenta, trachea, thyroid, spinal cord, salivary gland, thymus and peripheral blood mononuclear cell). BUC11 mRNA is expressed at similar levels in the normal testis and testicular cancer tissues tested. Quantitative RT-PCR using specific BUC11 primers also revealed that BUC11 mRNA is expressed at widely different levels in the 10 breast cancer samples tested with apparent higher levels in early stages of the disease when compared to later stages. In the normal counterpart tissue, BUC11 mRNA expression follows the same trend. Furthermore, the more sensitive technique of quantitative RT-PCR proved that BUC11 mRNA is not only expressed in the breast cancer

cell line T47D but also at low level in the breast cancer cell line MDA231 which was not detected using semi-quantitative RT-PCR. BUC11 mRNA was undetectable in the 6 different samples of peripheral blood mononuclear cells tested (figure not shown) nor in any of the diverse cancer tissues tested (mesothelioma, melanoma, gastric cancer, kidney cancer and oesophageal cancer) as seen in Figure 3.22.

As discussed earlier, BUC9 mRNA expression studies showed either one or two PCR products depending on the tissue type. From this observation, it was concluded that BUC9 might exist as two mRNA splice variants in some tissue types but as a single spliced mRNA in others. This was investigated further by cloning the two PCR products for sequencing. Briefly, following RT-PCR with the breast cancer sample BR15 cDNA, the gel extracted and purified upper band (about 500 base pairs) and lower band (about 400 base pairs) were individually cloned into pCRII-blunt-TOPO (Invitrogen) vectors and XL1B cells were transformed and cultured. After DNA Miniprep and EcoRI restriction digest, positive clones for BUC9 genes were sent for sequencing to MWG Biotech (UK). The sequences were then aligned to the original BUC9 sequence using the JustBio Aligner tool (http://www.justbio.com). The data, presented in Figure 3.23 and Figure 3.24, clearly shows that the smaller PCR product lacks 103 base pairs which are present in the larger PCR product. This finding confirms that BUC9 mRNA molecules can at least be spliced into two different molecules, depending on the tissue of origin. As shown in Figure 3.25, the predicted protein sequences for both splice variants have been generated using the Translator tool of the JustBio website (http://www.justbio.com/translator/index.php) and potential polypeptides uploaded in a protein-protein BLAST (http://blast.ncbi.nlm.nih.gov /Blast.cgi) using the same conditions as previously mentioned. No significant similarity was found with human nor mouse proteins when database searches were performed.

Figure 3.23: Alignment of BUC9 nucleotide sequence with the sequencing result for the larger PCR product.
"BUC9" is the reference nucleotide sequence for BUC9 and "8H1M13Forward" is the sequencing result of the approximate 500 base pair PCR product, following semi-quantitative RT-PCR using cDNA from a breast cancer tissue that tested positive for the two potential BUC9 splice variants. The two sequences perfectly align and prove the specificity of the PCR product sequenced.

Figure 3.24: Alignment of BUC9 nucleotide sequence with the sequencing result for the smaller PCR product.
"BUC9" is the reference nucleotide sequence for BUC9 and "4LM13Forward" is the sequencing result of the approximate 400 base pair PCR product, following semi-quantitative RT-PCR using cDNA from a breast cancer tissue tested positive for the two potential BUC9 splice variants. The two sequences align except that the sequencing result is missing 103 base pairs which are present in the reference, suggesting that the shorter PCR product is a splice variant of the original gene and is not a result of contamination or artifact.

BUC9 longest splice variant mRNA sequence (771bp)

ctaaacaaattgctcctgtggagatgattggcatcacatggtgttttgagctgatacacccaacacttg
agctcactgcaacagtaccagattttcaccgctatgcctcctttcactctgggagtcttccagaggtct
tgcactcgggagagcatgctcaggtttccccagctctacaaaatcacccagaatgccaaagacttcaac
acaagggtaaataaggttgatctcagaattgtcacctcaaaaaggccctgccttccactgttcagttct
ggtcatctgcctatgagatatctgaagcttgaaagagaacacttgaaaatcactgagaccgtgactccc
atcccagcacacacagcaagccaagtaggttacagagatttcttcttgggtgatgagttcacgccacat
ggggtatgttctccagttccagtgtgttgactccttcctgcttccccccatcagcccctgaggtcaatg
tgggcagcagcccccatggtccaagttcaagatccgctgtggaagacttttttaggcaatcacacagcc
ctgcatggagtcctgatgagagcttgcctaattgttgctaggtttgtcattttaaatacagtgtttctt
tagctagtgagtaaaattggctatataggagatactgtgttgaccagtggtcatgccactgcctgttga
tttgttgaaaatattgtttacacgtatgttcttgttactgattgtcagaaagctggttttgagactgca
gcttggactaaa

BUC9 shortest splice variant mRNA sequence (668bp)

ctaaacaaattgctcctgtggagatgattggcatcacatggtgttttgagctgatacacccaacacttg
agctcactgcaacagtaccagattttcaccgctatgcctcctttcactctgggagtcttccagaggtct
tgcactcgggagagcatgctcaggtttccccagctctacaaaatcacccagaatgccaaagacttcaac
acaagggtaaataaggttgatctcagaattgtcacctcaaaaaggccctgccttccactgttcagttct
ggtcatctgcctatgagatatctgaagcttgaaagagaacacttgaaaatcactgagaccgtgactccc
atcccagcacacacagcaagccaatgaggtcaatgtgggcagcagcccccatggtccaagttcaagatc
cgctgtggaagacttttttaggcaatcacacagccctgcatggagtcctgatgagagcttgcctaattg
ttgctaggtttgtcattttaaatacagtgtttctttagctagtgagtaaaattggctatataggagata
ctgtgttgaccagtggtcatgccactgcctgttgatttgttgaaaatattgtttacacgtatgttcttg
ttactgattgtcagaaagctggttttgagactgcagcttggactaaa

BUC9 longest splice variant predicted protein sequence

MIGITWCFELIHPTLELTATVPDFHRYASFHSGSLPEVLHSGEHAQVSPALQNHPECQRLQHKGK*G*S
QNCHLKKALPSTVQFWSSAYEISEA*KRTLENH*DRDSHPSTHSKPSRLQRFLLG**VHATWGMFSSSS
VLTPSCFPPSAPEVNVGSSPHGPSSRSAVEDFFRQSHSPAWSPDESLPNCC*VCHFKYSVSLASE*NWL
YRRYCVDQWSCHCLLIC*KYCLHVCSCY*LSESWF*DCSLD*

BUC9 shortest splice variant predicted protein sequence

MIGITWCFELIHPTLELTATVPDFHRYASFHSGSLPEVLHSGEHAQVSPALQNHPECQRLQHKGK*G*S
QNCHLKKALPSTVQFWSSAYEISEA*KRTLENH*DRDSHPSTHSKPMRSMWAAAPMVQVQDPLWKTFLG
NHTALHGVLMRACLIVARFVILNTVFL*LVSKIGYIGDTVLTSGHATAC*FVENIVYTYVLVTDCQKAG
FETAAWTKX

Figure 3.25: BUC9 two potential splice variants mRNA and predicted protein sequences. The nucleotide sequence was translated into a protein sequence using the Translator tool of the JustBio website (http://www.justbio.com/translator/index.php). The translation was carried out in one phase from the first ATG of the nucleotide sequence, using the standard genetic code. Stars represent stop codons. The polypeptide highlighted in blue (65 amino acids) is common to both splice variants predicted protein sequences. The sequence highlighted in green is another potential polypeptide for each splice variant which has a unique amino acid sequence. For the longest splice variant, the polypeptide in green is 57 amino-acid-long whereas the polypeptide in green for the shortest splice variant is 50 amino-acid-long.

3.2.3 Focus on BUC11 mRNA expression analysis

Based on the data collected for BUC11 compared to the data obtained for the other BUC genes so far (section 3.2.2), BUC11 was identified as being the most promising gene of the BUC family. Further work presented in this thesis was consequently focused on this gene and primarily on a more thorough mRNA expression analysis. In our study, we designed specific primers to hCG25653 gene for use in conventional RT-PCR and other specific primers for use in quantitative RT-PCR. Some experiments were conducted on normal tissue RNAs (purchased from Clontech, UK), melanoma tissue RNAs (generously donated by Prof. D. Schadendorff, Germany), prostate tissue RNAs (generously donated by Nottingham City Hospital, UK) and RNA from breast cancer cell lines cultured in the laboratory. Quantitative Real-Time RT-PCR was the method of choice because it is less time consuming, more specific and more sensitive (especially when low levels of mRNA are obtained from tissues gained through different collaborations) compared to semi-quantitative RT-PCR. Briefly, 2μg of RNA (or entire sample if the RNA concentration was low) and Oligo(dT)$_{15}$ or random primers were used to generate cDNA by reverse-transcription. The quantitative RT-PCR was performed as described in Chapter 2. The housekeeping gene GAPDH was used to normalise changes in specific gene expression. Following completion of the quantitative RT-PCR reaction, the analysis of the results was carried out using the comparative Ct method. The graphs showing the relative gene expression levels of BUC11 in the breast cancer cell lines, the normal tissues, the melanoma tissues and the prostate tissues are presented, respectively, in Figure 3.26, Figure 3.27, Figure 3.28 and Figure 3.29 (the Ct values can be found in Appendix I).

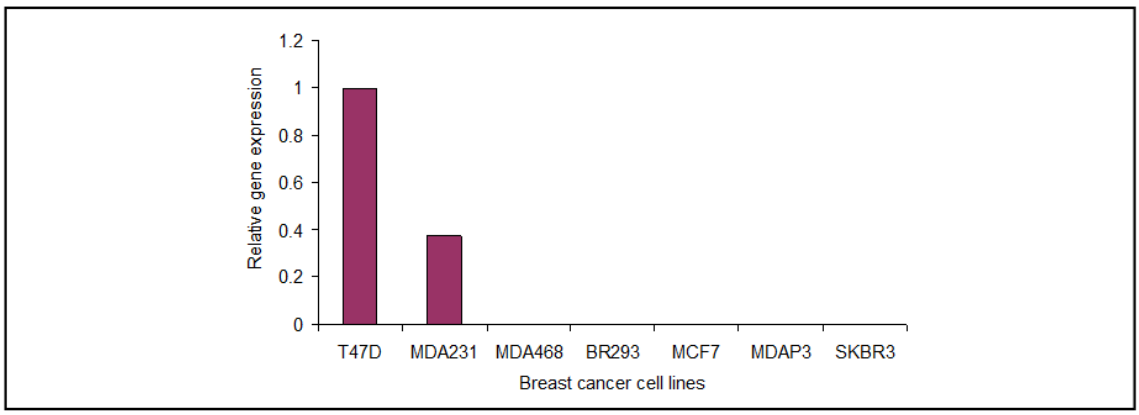


Figure 3.26: BUC11 mRNA expression in human breast cancer cell lines by quantitative RT-PCR.
Quantitative RT-PCR analysis was carried out on 7 breast cancer cell lines. The graph shows the normalised level of BUC11 mRNA expression in the cell lines, relative to its expression level in T47D, arbitrarily set as 1.

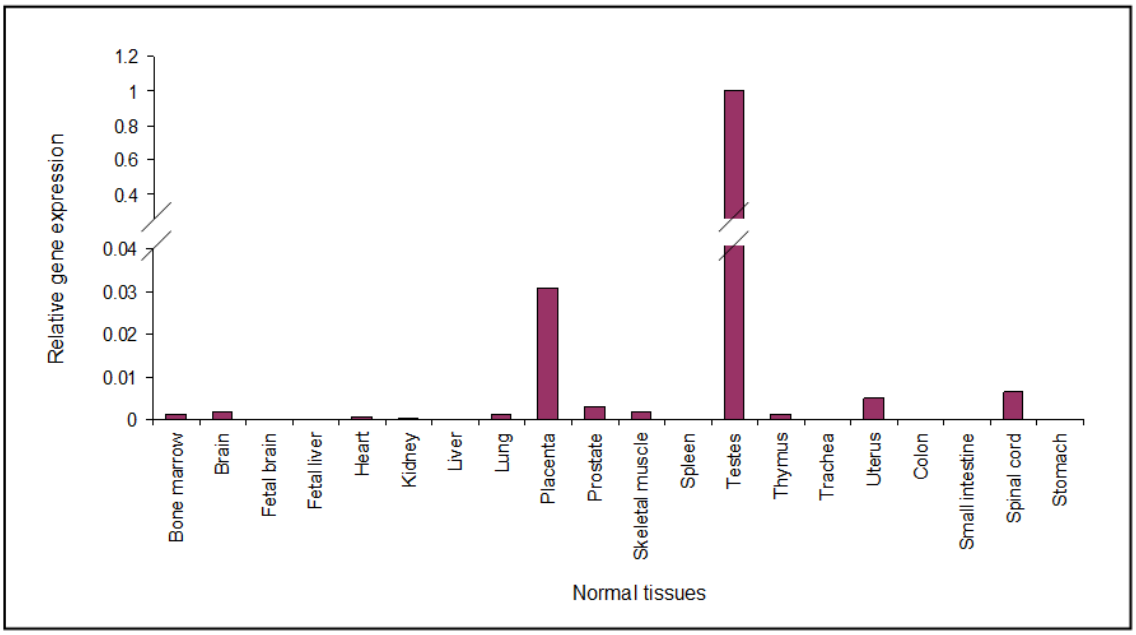


Figure 3.27: BUC11 mRNA expression in human normal tissues by quantitative RT-PCR. Quantitative RT-PCR analysis was carried out on a panel of 20 normal tissues. The graph shows the normalised level of BUC11 mRNA expression in the tissues, relative to its expression level in testis tissue, arbitrarily set as 1.

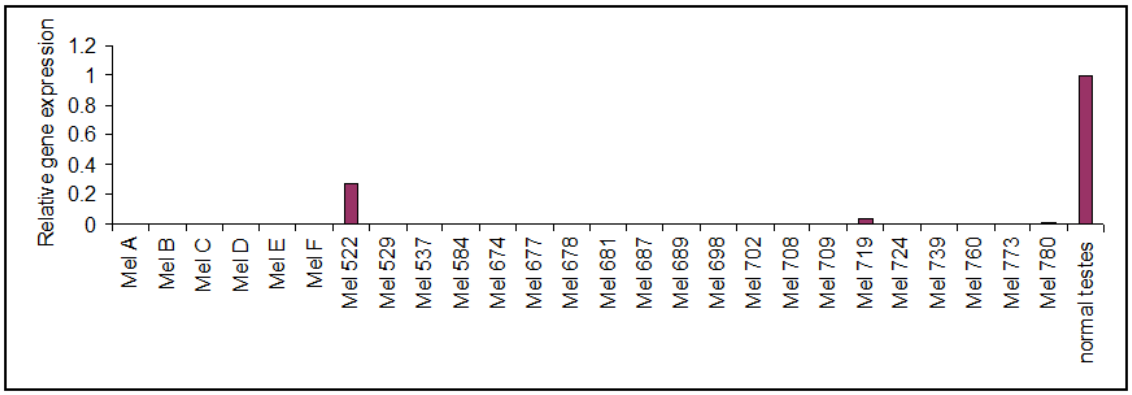


Figure 3.28: BUC11 mRNA expression in human melanoma tissues by quantitative RT-PCR.
Quantitative RT-PCR analysis was carried out on 26 melanoma tissues. The clinical information for these samples is unknown. The graph shows the normalised level of BUC11 mRNA expression in the tissues, relative to its expression level in normal testis tissue, arbitrarily set as 1.

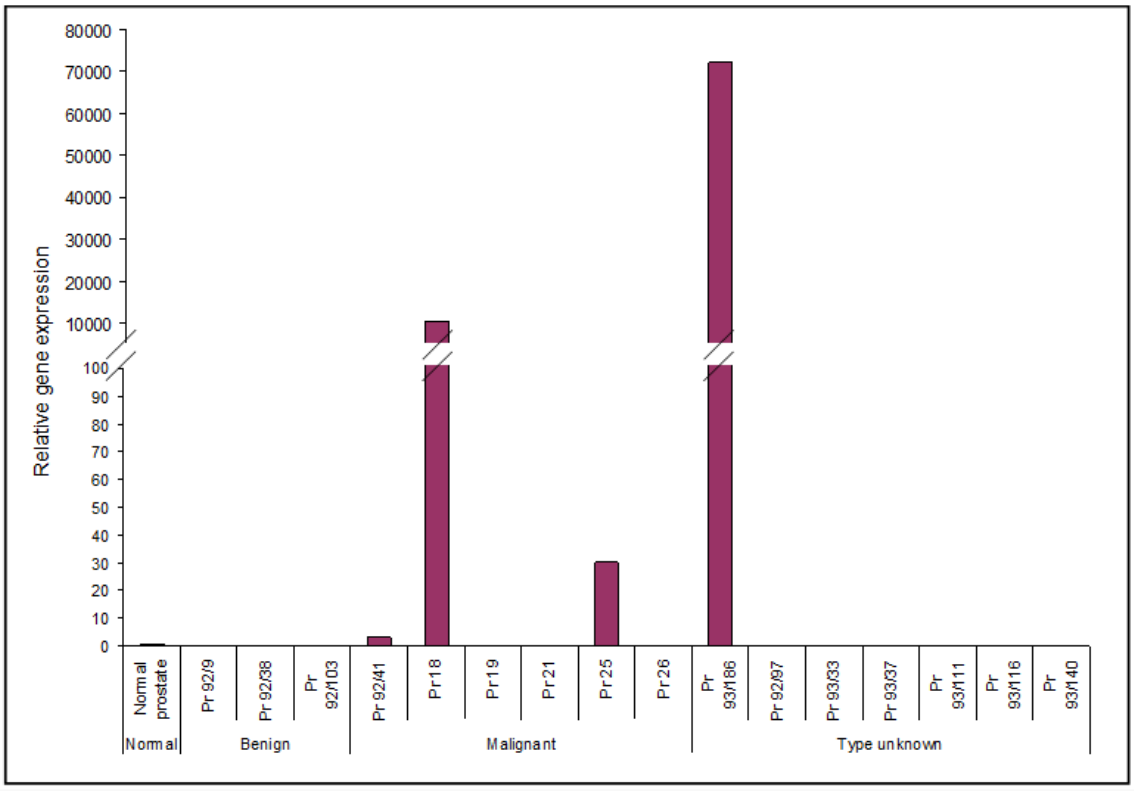


Figure 3.29: BUC11 mRNA expression in human prostate tissues by quantitative RT-PCR.
Quantitative RT-PCR analysis was carried out on 3 known benign prostate tissues, 6 known malignant prostate tissues and 7 prostate tissues of unknown tumour type. The graph shows the normalised level of BUC11 mRNA expression in the tissues, relative to its expression level in normal prostate tissue, arbitrarily set as 1.

As shown in Figure 3.26, BUC11 mRNA was only detected in 2 out of 7 breast cancer cell lines tested. BUC11 mRNA is found at a higher level in T47D compared to MDA231. It can not be observed on the graph but extremely low levels of BUC11 mRNA were detected in MCF7, SKBR3 and MDAP3 cells, with relative gene expression respectively 0.0009, 0.0009 and 0.0004. With semi-quantitative RT-PCR, BUC11 mRNA was found only in breast and testis tissues. Quantitative RT-PCR also gave a high level of BUC11 mRNA in these two tissues but low levels were also found in other normal tissues (Figure 3.27), however the expression in these compared to the expression in testis and placenta is negligible. Only 3 out of 26 melanoma patients tested positive for BUC11 mRNA however the level in these patients compared to the level in normal testis tissue is again very low (Figure 3.28). Finally, as shown in Figure 3.29, BUC11 mRNA was found expressed in 5 out of 16 prostate tissues tested (1 normal prostate tissue, 3 malignant tumour tissues and 1 unknown tumour type tissues), with high levels in tumour when compared to the level in normal prostate. Also, none of the known benign tissues expressed BUC11 mRNA.

In order to assess the potential of BUC11 for the care of breast cancer patients, a large scale study was conducted on cDNA from patients where matched breast cancer tissues and normal breast tissues had been obtained. Dr Angelos Gritzapis (The Cancer Immunology and Immunotherapy Center, Saint Savas Cancer Hospital, Athens, Greece) made available a total of 218 frozen breast tissues which correspond to paired normal and tumour tissues from 109 patients (a list of the tissues can be found in Appendix II). The clinical information provided include, for each patient, the age at diagnosis of the disease, the stage of the tumour, the grade of the tumour and the HER2 receptor expression status of the tumour. Briefly, RNA was extracted from frozen tissues using liquid nitrogen, a mortar, a pestle and RNA-STAT 60 under a class II safety cabinet. Reverse-transcription using random primers was carried out on 2μg of RNA and cDNA was used in quantitative RT-PCR. It is noteworthy that three different control genes were used: GAPDH, hypoxanthine ribosyltransferase (HPRT1) and TATA-box-binding protein (TBP). In quantitative RT-PCR analysis, a internal control gene has to be used to correct expression data for differences between samples in the quality of the RNA used, the quantity of the RNA used and the efficiency of the RT. As suggested by Jacques de Kok and colleagues (2005), a lot of studies wrongly use a single reference gene for normalisation even if no gold standard control gene has been characterised to date for use with clinical tumour samples (de Kok *et al.*, 2005). They compared 13 frequently used reference genes in cancer research (large ribosomal protein, b-actin, cyclophilin A, GAPDH, phosphoglycerokinase 1, b-2-microglobin, b-glucuronidase, HPRT1, TBP, transferrin receptor, porphobilinogen deaminase, ATP synthase 6 and 18S ribosomal RNA) and came to the conclusion that, for normalisation, the use of the mean expression of several reference genes is the best substitute to the use of a single control gene. However they found out that HPRT1 is the best out of all of the tested genes and can alternatively be used on its own.

Following quantitative RT-PCR, the geNorm VBA applet for Microsoft Excel (free download at http://medgen.ugent.be/genorm/) was used to determine the most stable control genes from GAPDH, HPRT1 and TBP in the panel of breast tissues tested. As stated in the user manual, the tool allows the calculation of "the gene expression stability measure M for a reference gene as the average pairwise variation V for that gene with all other tested reference genes. Stepwise exclusion of the gene with the highest M value allows ranking of the tested genes according to their expression stability". Vandesompele and colleagues (2002) give an account of the calculations and principles in their paper published in Genome Biology under the title "Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal reference genes". The result of the calculation of the average expression stability values of control genes is given in Figure 3.30. HPRT1 and TBP were determined as the most stable genes therefore the mean of their Ct values were used for the determination of relative gene expressions in the analysis carried out using the comparative Ct method. The data were plotted on

four graphs using Prism version 3.03. Each graph shows the normalised level of BUC11 mRNA expression in the normal tissue, relative to its expression level in the tumour tissue on the Y axis and one of the clinical information of the X axis, either the age of patients at diagnosis, the grade of the tumour, the stage of the tumour or the HER2 expression status of the tumour tissues.

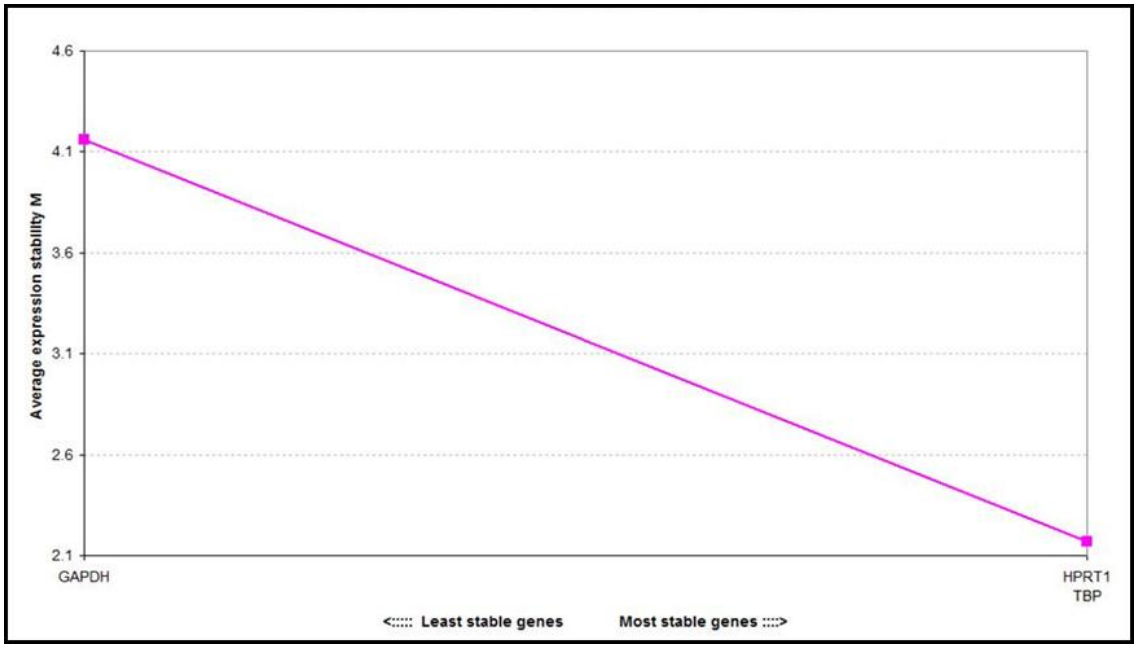


Figure 3.30: Average expression stability values of control genes.
The most stable reference genes from a set of tested genes in the panel of cDNA from the paired normal and tumour breast tissues were determined using the software geNorm (http://medgen.ugent.be/genorm/).

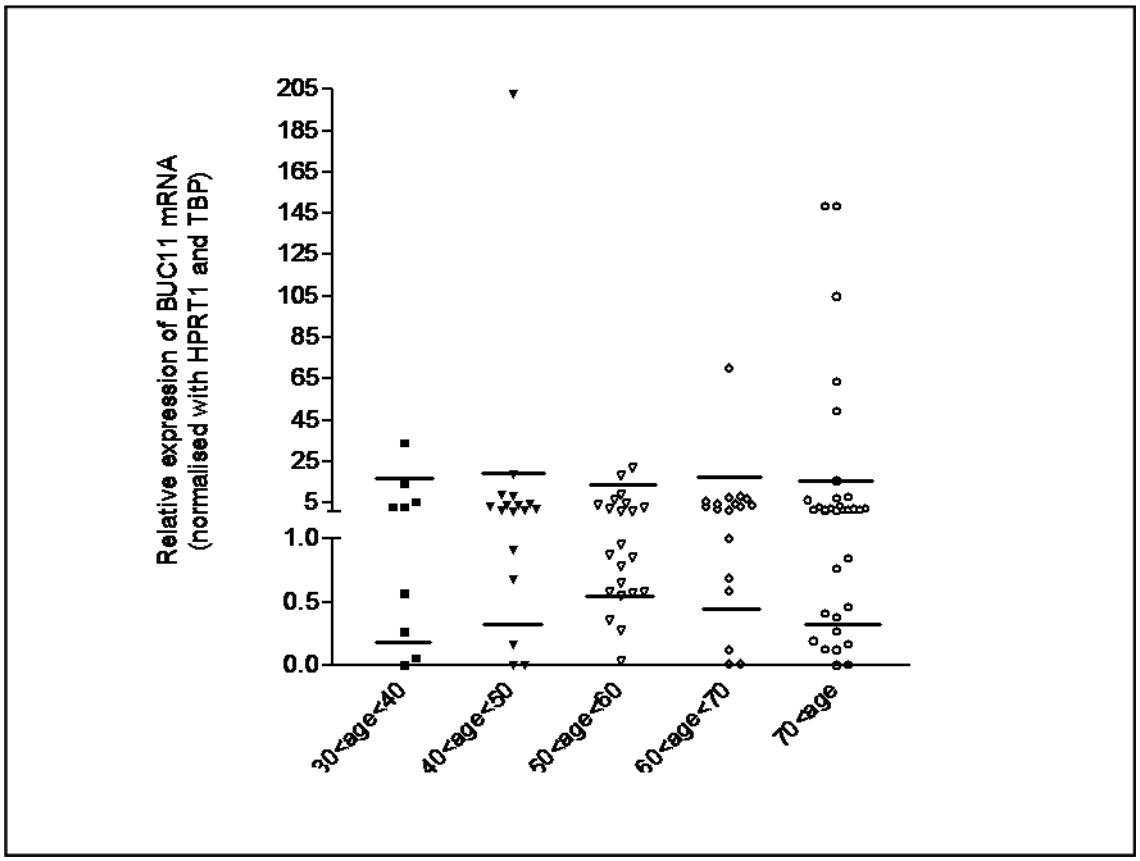


Figure 3.31: BUC11 mRNA expression in paired normal and tumour breast tissues by quantitative RT-PCR, sorted according to the age of patients at diagnosis.
Quantitative RT-PCR analysis was carried out on the cDNA panel of paired normal and tumour breast tissues. The analysis of the results was carried out using the comparative Ct method. The control genes HPRT1 and TBP were used for normalisation. The graph shows the normalised level of BUC11 mRNA expression in the normal tissue, relative to its expression level in the tumour tissue. The values above 1 represent the patients whose BUC11 mRNA expression is upregulated in the tumour tissue compared to the normal counterpart tissue. The values below 1 represent the patients where BUC11 mRNA expression is downregulated in the tumour tissue compared to the normal counterpart tissue. The average for each group is given by the horizontal line.

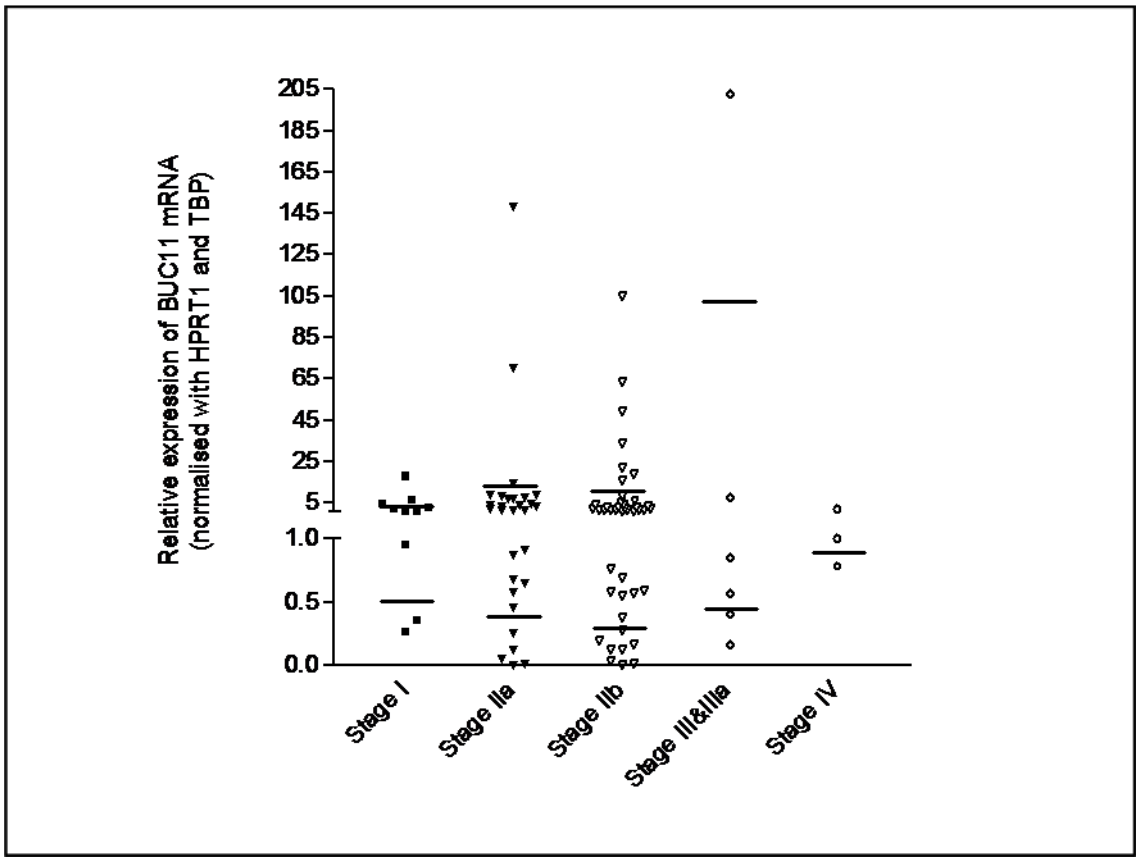


Figure 3.32: BUC11 mRNA expression in paired normal and tumour breast tissues by quantitative RT-PCR, sorted according to the stage of the tumour tissues.
Quantitative RT-PCR analysis was carried out on the cDNA panel of paired normal and tumour breast tissues. The analysis of the results was carried out using the comparative Ct method. The control genes HPRT1 and TBP were used for normalisation. The graph shows the normalised level of BUC11 mRNA expression in the normal tissue, relative to its expression level in the tumour tissue. The values above 1 represent the patients whose BUC11 mRNA expression is upregulated in the tumour tissue compared to the normal counterpart tissue. The values below 1 represent the patients where BUC11 mRNA expression is downregulated in the tumour tissue compared to the normal counterpart tissue. The average for each group is given by the horizontal line. Stage I: the tumour is less than 2cm, the lymph nodes in the armpit are not affected and the cancer has not spread. Stage IIa: the tumour is less than 2cm, the lymph nodes in the armpit are affected and the cancer has not spread. Stage IIb: the tumour is less than 5cm, the lymph nodes in the armpit are affected and the cancer has not spread. Stage III&IIIa:the tumour is more than 5cm, the lymph nodes in the armpit are affected and the cancer has not spread. Stage IV: the tumour can be any size, the lymph nodes may or may not be affected and the cancer has metastasised to other parts of the body.

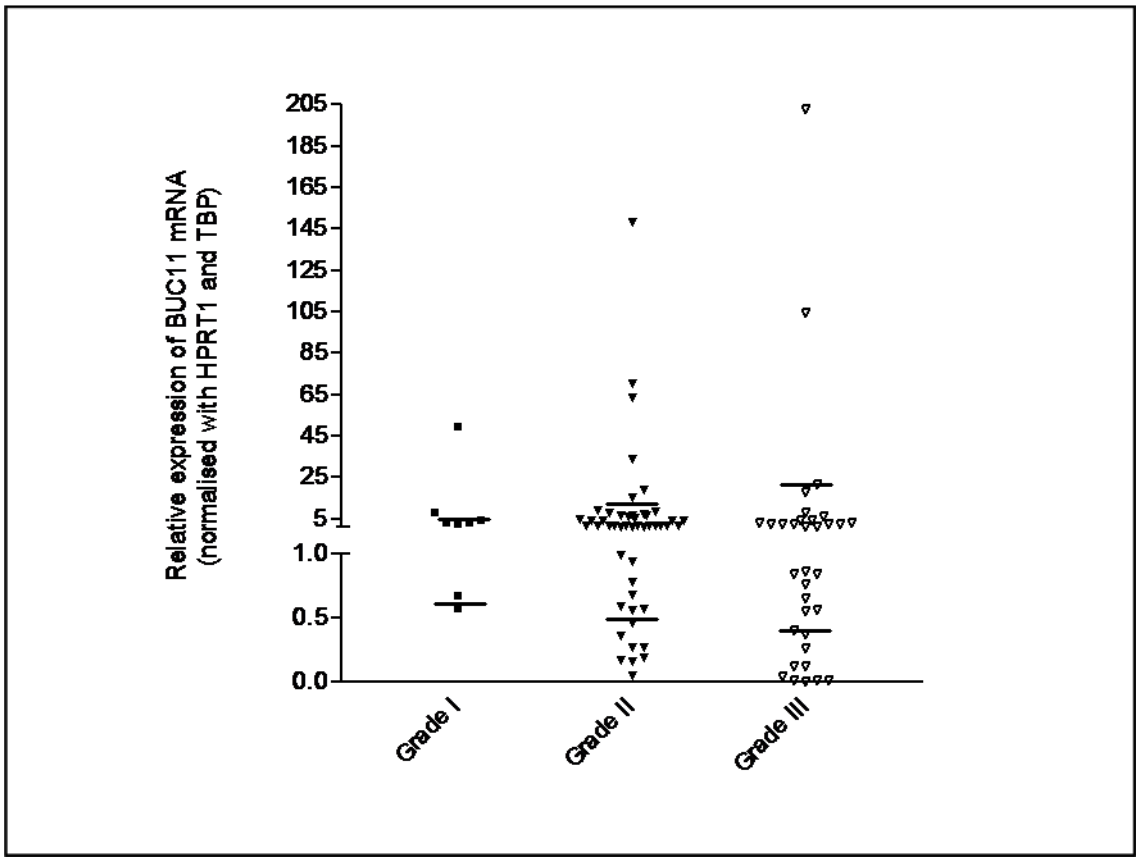


Figure 3.33: BUC11 mRNA expression in paired normal and tumour breast tissues by quantitative RT-PCR, sorted according to the grade of the tumour tissues.
Quantitative RT-PCR analysis was carried out on the cDNA panel of paired normal and tumour breast tissues. The analysis of the results was carried out using the comparative Ct method. The control genes HPRT1 and TBP were used for normalisation. The graph shows the normalised level of BUC11 mRNA expression in the normal tissue, relative to its expression level in the tumour tissue. The values above 1 represent the patients whose BUC11 mRNA expression is upregulated in the tumour tissue compared to the normal counterpart tissue. The values below 1 represent the patients where BUC11 mRNA expression is downregulated in the tumour tissue compared to the normal counterpart tissue. The average for each group is given by the horizontal line. Grade I: slow growing tumour cells, Grade II: intermediate growing tumour cells , Grade III: fast growing tumour cells.

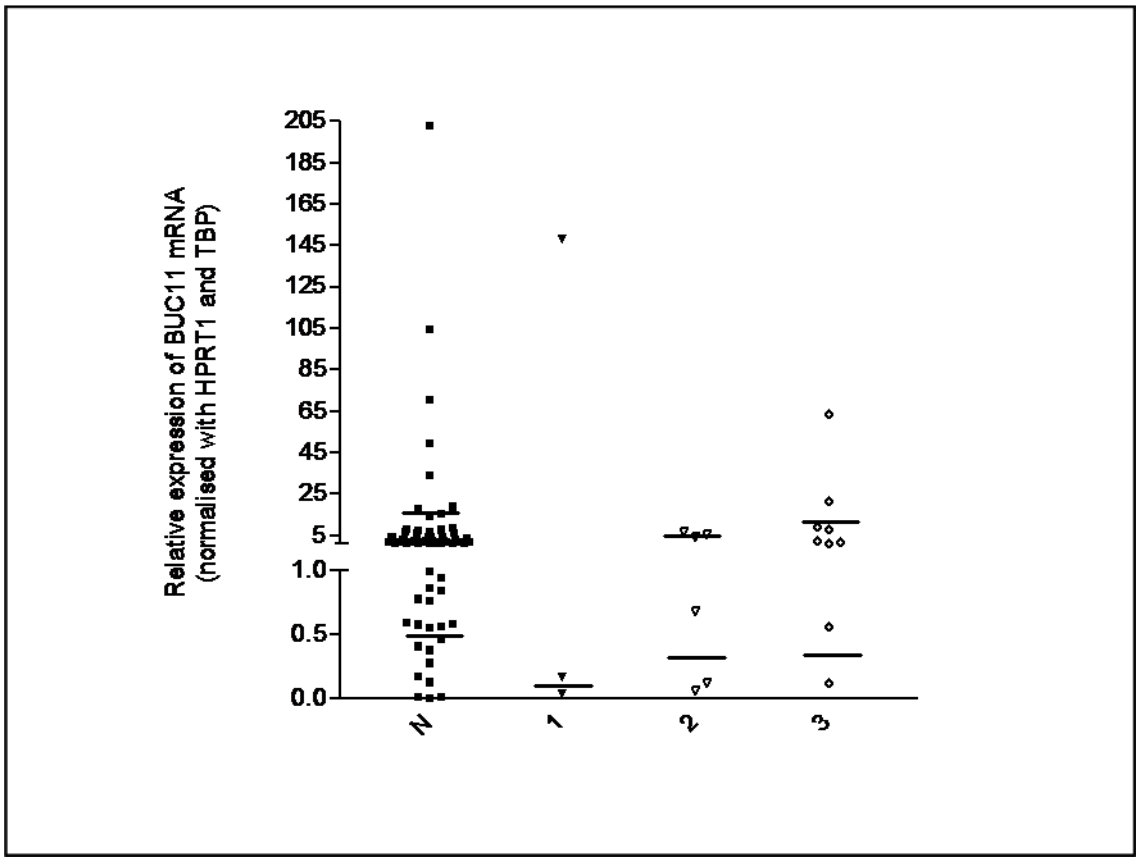


Figure 3.34: BUC11 mRNA expression in paired normal and tumour breast tissues by quantitative RT-PCR, sorted according to the HER2 expression status of the tumour tissues.
Quantitative RT-PCR analysis was carried out on the cDNA panel of paired normal and tumour breast tissues. The analysis of the results was carried out using the comparative Ct method. The control genes HPRT1 and TBP were used for normalisation. The graph shows the normalised level of BUC11 mRNA expression in the normal tissue, relative to its expression level in the tumour tissue. The values above 1 represent the patients whose BUC11 mRNA expression is upregulated in the tumour tissue compared to the normal counterpart tissue. The values below 1 represent the patients where BUC11 mRNA expression is downregulated in the tumour tissue compared to the normal counterpart tissue. The average for each group is given by the horizontal line. N: negative, 1: dim positive, 2: intermediate positive, 3: highly positive.

First of all, it is worth noting that almost all breast cancer patients expressed BUC11 mRNA in the normal and tumour tissues except three patients who did express BUC11 mRNA in the normal breast tissue but no expression was found in the tumour tissue. The first patient was 71 years old with the tumour at stage III, of grade II and did not express HER2. The second patient was 40 years old with the tumour at stage II, of grade II and highly expressed HER2. The third patient was 45 years old with the tumour at stage III, of grade II and did not express HER2. The regulation (down- or up-) of BUC11 mRNA expression is not linked to the age of the patient at diagnosis (Figure 3.31). Similarly, as shown in Figure 3.32, the data analysed according to the stage of the tumour indicated that BUC11 mRNA levels do not specifically vary with the advancement of the disease

and does not follow any trend. Furthermore, the BUC11 mRNA levels in tissues is not connected to the expression status of the HER2 receptor (Figure 3.34). Finally, BUC11 mRNA expression levels appeared to be unrelated to the grade of the breast tumours (Figure 3.33).

3.3 Discussion

The BUC genes are novel breast-associated genes recently identified in our lab by mining EST databases, on the basis of their specific expression spectrum including testis, normal breast and breast cancer tissue. In this chapter, the aims were to obtain the full length sequence of the BUC genes and determine the clinical relevance of the BUC genes by investigating their expression spectrum at the mRNA level, before making further consideration on the potential applications of these genes. Genes which expression is limited to non vital organs such as breast and testis are targets for immunotherapeutic strategies. These aims have been achieved in that interesting expression analysis data have been obtained for all the BUC genes.

In this study, following exhaustive sequencing of the BUC genes, it was shown that all the sequences are mapping to chromosome 10p1121.12.1. The sequencing results revealed that, in fact, BUC6 mRNA sequence, which form exon 1 of the "BUC gene", is part of the 3' end of NY-BR-1 gene, away from the coding region. Furthermore, exons 10 to 14 of the "BUC gene" forms the entire sequence of BUC11 and 99% of this sequence matches perfectly the mRNA sequence of the unpublished (hypothetical coding protein) gene hCG25653, which is next to NY-BR-1 on the chromosome. NY-BR-1 is a tumour antigen identified by SEREX (serological analysis of recombinant cDNA expression libraries) (Jager *et al.*, 2001). Using RT-PCR, Jager and colleagues found that NY-BR-1 mRNA is present solely in the ductal epithelium of the breast, prostate, prostate cancer (32%) and testis (Jager *et al.*, 2001 and 2005). NY-BR-1 was expressed in 70% of all tested breast cancer tissues in the study published by Jager and Knuth in 2005 (Jager and Knuth, 2005). During the research on NY-BR-1 expression at the protein level conducted by Jager, Seil and colleagues in 2007, it was found that NY-BR-1 is present in the epithelia of ducts and acini of the mammary gland and that all other normal tissues, including testis, did not express the protein. The expression was more frequent in ductal carcinomas in situ and/or the intraductal component of invasive carcinomas than in invasive carcinomas. The majority of the breast cancer cell lines tested expressed only low NY-BR-1 levels and the NY-BR-1 protein was found localised to the cytoplasm and the cell membrane. Researchers came to the conclusion that NY-BR-1 is a differentiation antigen of the mammary gland with potential use for diagnosis and immunotherapy of breast cancer (Jager *et al.*, 2007 and Seil *et al.*, 2007). Jager and others identified two

HLA-A2 restricted NY-BR-1 peptides which are CD8+ T cell specific and naturally processed (Jager *et al.*, 2005 and Wang *et al.*, 2006). These findings support the idea of using NY-BR-1 for antigen-based therapy of HLA-A2 positive breast cancer patients whose tumours express NY-BR-1. Furthermore, in 2007, Theurillat and colleagues studied the co-expression of NY-BR-1 with HLA class I antigens in a tissue microarray. About 15% of primary breast tumours and 6% of distant metastases highly co-expressed NY-BR-1 and HLA class I, suggesting that about 10% of breast cancer patients could receive T-cell-based immunotherapy (Theurillat *et al.*, 2007). In the same study, they interestingly uncovered that NY-BR-1 expression was more frequent in grade I than in grade II or grade III breast cancers and it was correlated with estrogen receptor expression and inversely correlated with HER2-status and EGFR expression (Theurillat *et al.*, 2007). Also, NY-BR-1 expression is more frequent in lymph node-negative primary carcinomas and is lost with tumor progression (Varga *et al.*, 2006). In 2008, Theurillat and colleagues published a study stating that NY-BR-1 protein can not be detected in testis and prostate even if mRNA can be detected in prostate, prostate cancer and testis as previous published work demonstrated. They also found NY-BR-1 mRNA expression in primary spermatocytes, in 2% of prostate carcinomas and 70% of primary breast carcinomas and normal breast epithelium, with strong correlation to its protein expression. Interestingly, they hypothesised that, because of NY-BR-1 restricted expression to sex-related organs, the 4 estrogen response elements (ERE)-like sequences nearby the promoter region, the strong co-expression of ER alpha and NY-BR-1 (mRNA and protein levels) and the important decrease of NY-BR-1 expression in recurrences after treatment with tamoxifen, ER alpha may control NY-BR-1 transcription (Theurillat *et al.*, 2008).

The nucleotide and protein BLAST results for the four BUC genes indicated that no significant homology to any known genes or proteins was found, except in the case of BUC11 with its similarity to hCG25653. Nevertheless, it is known that the production of a particular mRNA molecule is not always associated with the assembly of the corresponding amino acid sequence. Many gene sequences which are not transcribed into proteins are non-coding RNA transcripts (ncRNA) and the complex tertiary structure of these non-coding RNA transcripts give them structural, catalytic or regulatory functions (Szymanski *et al.*, 2003). However, as previously mentioned, BUC11 generates a predicted protein with homology to hCG25653. The predicted protein could be of immunotherapeutic interest if it is proven that the gene product is actually produced and has a vital role in breast cancer cells.

Preliminary RT-PCR results (personal communication with Dr. Li) appeared promising therefore mRNA expression analyses were undertaken in our study for all the BUC genes. RT-PCR has been widely recognised as the most sensitive tool available for mRNA detection and quantitation. It is generally assumed that the absence of detectable RT-PCR

products in one given tissue implies a level of expression lower than 0.1% when compared to the counterpart tissue (cancerous versus normal) (Boel *et al.*, 1995). Most of the mRNA expression analysis has been carried out using conventional, non- or semi-quantitative RT-PCR depending on the experiment. However, the issue with conventional RT-PCR is that the intensity of the signal is not necessarily directly proportional to the amount of DNA being amplified, however sensitivity can be increased with the correct optimisation of the PCR conditions. The strategy was also to include quantitative RT-PCR, which was developed over the last few years and has the advantage of being highly quantitative and reliable and is now the preferred method to measure changes in gene expression. This study has confirmed that quantitative RT-PCR is more reliable and sensitive for mRNA expression analysis than conventional RT-PCR. Quantitative RT-PCR allows a quantitative analysis and can detect very low levels of expression (with a low amount of sample i.e. 0.5μl) that can not be seen on an agarose gel electrophoresis loaded with the product from conventional RT-PCR thus leading to an inaccurate interpretation of the results.

Genes which are tissue-specific manifest an expression profile restricted to certain tissue types (e.g. reproductive tissues) and are supervised by rigorous and specific transcriptional control mechanisms. It is through preferential activation or silencing of tissue-specific genes that cell differentiation mechanisms take place in multi-cellular eukaryotes. Differentiated cells then show their individual phenotypic characters and their roles in the organism. A strictly restricted expression spectrum, especially to non vital tissues such as breast, is of particular interest for immunotherapeutic strategies because the normal vital tissues would not be affected by the treatments. After the reproductive years, the breast is not essential hence unintended destruction of normal breast tissues of breast cancer patients through novel therapies should have negligible consequences. To determine whether the BUC genes display an appropriate expression profile, various invasive and non-invasive histological subtypes of breast cancer, breast cancer cell lines, malignancies (e.g. melanoma and mesothelioma) and normal tissues were used for the mRNA expression studies. It is worth noting that, as explained in Chapter 2, MDA435 cell line was used in the study of the expression of the BUC genes in breast cancer cell lines before it was known that MDA435 is actually a melanoma cell line as stated in a recent publication in the journal Breast Cancer Research and Treatment in 2007, entitled "MDA-MB-435 cells are derived from M14 Melanoma cells–a loss for breast cancer, but a boon for melanoma research" (Rae *et al.*, 2007).

In the preliminary studies using RT-PCR (personal communication with Dr. Li), the BUC6 gene demonstrated a specific expression pattern restricted mainly to normal testis and breast tissues. However, traces of expression in PBMC and placenta were also found. BUC6 mRNA was found expressed in normal breast (100%), breast cancer tissues (89%)

and testis tissues (100%). BUC6 mRNA was also found expressed in the breast cancer cell line T47D. Interestingly, low expression of BUC6 was found in mesothelioma tissues (50%) and mesothelioma cell lines (75%). In this study, BUC6 mRNA expression was observed in all the melanoma samples tested. However, the level of expression for each individual sample was not ascertained as the RT-PCR was non-quantitative. This data suggests that BUC6 is not only expressed in breast cancer and mesothelioma but also is highly expressed in melanoma. Therefore, due to its high occurrence in breast cancer, melanoma and mesothelioma, BUC6 could be of interest as a biomarker for these types of cancer hence further investigations should be carried out. Mesothelioma is a rare cancer almost always caused by exposure to asbestos and as a result the most common site of development is the lungs. Current therapies have proven unsuccessful and as a result the median survival time after diagnosis of the malignant form is 6 to 12 months. Therefore, novel mesothelioma associated biomarkers could be of high interest for the development of more effective therapies. BUC6 was expressed in breast cancer cell lines with high levels in T47D (which confirmed the result in preliminary studies) and SKBR3, slightly lower level in MDA231 and at a low level in MDA468. BR293 does not express BUC6 mRNA and this might be due to transcriptional silencing in this particular cell line.

Gene silencing is a common process that can occur under normal conditions at some point during the development by, for instance, the downregulation caused by transacting factors, the deactivation of tissue-specific promoter or by silencer regulatory molecules. Mutations of the cis and trans-regulatory elements are responsible for changes in the gene expression profile in cancerous cells compared to normal cells i.e. for the reduction of level of transcription thus the loss of cell differentiation. Moreover, other mechanisms can be responsible for the change of expression. For instance, promoter hypermethylation through methylation of the CpG islands is associated with transcriptional silencing (Ohtani-Fujita *et al.*, 1993). Also, under *in vitro* conditions, cell lines divide more rapidly than the cells from their tissues of origin and it was demonstrated that the observed gene expression differences comprise an up-regulation of genes that have a role in proliferation (Perou *et al.*, 1999). Sandberg and Ernberg (2005) showed that in cell lines the tissue-specific up-regulation of genes is generally lost and observed that within cell lines which originated from the same tumour type, large variations in the expression of tumour and tissue-specific genes can be found (Sandberg and Ernberg, 2005). Therefore, studies stating the percentage of cell lines expressing the gene of interest do not necessarily reflect the percentage of patients who may express the gene. Finally, BUC6 mRNA expression was analysed using quantitative RT-PCR in some gastric tissues (normal and cancer), colon tissues (normal and cancer) and breast cancer tissues. BUC6 mRNA expression was significantly higher (at least 80 times) in the breast cancer samples than in the gastric and colon samples, however the low number of samples used should moderate this affirmation. Therefore, it is suggested that at least 10 more samples for each type of cancers

should be used to repeat this experiment. The level of expression of BUC6 mRNA in vital tissues remains to be confirmed.

In the preliminary studies using RT-PCR (personal communication with Dr Li), BUC9 mRNA was not present in vital tissues, was expressed at very low level in PBMC and was highly expressed in testis and breast cancer tissues. Two gene products were observed in breast cancer and testis tissues but only one gene product was shown in PBMC. This data was confirmed in this study as BUC9 mRNA expression was only found in placenta, testis and breast when tested in a panel of 9 normal tissues. The two PCR products only observed for testis and breast suggest that BUC9 may exist as a splice variant in these tissues. The two PCR products were true findings and not a result of artifact or contamination. Alternative splicing occurs in at least 30% of human genes and is defined as the re-arrangement of exons of the pre-mRNA (transcribed from the same gene) in alternative manners. These mRNAs can then code for proteins which functions can differ a lot or not at all. The presence and/or level of a specific variant in a specific tissue may be an indication of disease state. The two forms of the gene were found highly expressed in all the breast cancer tissues tested except for BR16 which exhibits a low level of expression for both forms.

BUC9 splice variants were also observed in the breast cancer cell lines T47D (high level), MDA468 (low level) and MDA231 (low level). On the other hand, the breast cancer cell lines MDAP3 and SKBR3 only exhibit one variant of the gene (largest variant). The splicing occurs in the middle of the longest BUC9 mRNA sequence, with 103 base pairs missing. All oesophageal paired normal and cancer tissues tested expressed BUC9 mRNA, with a high level in the majority of the samples. Two out of four paired normal and gastric cancer tissues expressed the longest variant of BUC9, with BU9 mRNA level being significantly higher in normal compared to tumour tissues. As a conclusion, BUC9 is expressed in normal oesophageal and gastric normal tissues therefore this gene is less desirable as a target for immunotherapy because of the possible toxic side effects. All the melanoma samples tested present the longest variant of BUC9. The data collected so far suggest that the shortest BUC9 variant is only expressed in the majority but not all of the normal breast tissues, breast cancer tissues, breast cancer cell lines and normal testis tissues. Therefore, this particular variant could be a target in immunotherapies which will only affect these non-essential tissues and not the vital organs. However, the expression in testicular cancer tissues has to be ascertained and expression of the variant in other cancers tested to confirm the restriction of expression of this variant. The BLAST analysis of the predicted proteins of both variants did not show any significant homology to known proteins.

In the preliminary studies using RT-PCR (personal communication with Dr. Li),

BUC10 mRNA expression pattern is comparable to BUC9 mRNA expression that is to say PBMC expressed very low levels of BUC10 mRNA whereas testis and breast cancer tissues expressed high levels of BUC10 mRNA. However, in the case of BUC10, only one gene product was shown. The results of the present study showed that BUC10 mRNA is highly expressed in melanoma. Furthermore, regarding expression in breast cancer cell lines, BUC10 mRNA is expressed at very low level in most of the breast cancer cell lines tested (T47D, MDA468, and MDA231) with the exception of the high level of expression found in SKBR3 cells. BUC10 mRNA is expressed at lower level in breast cancer cell lines and tissues than in normal tissues, gastric cancer, oesophageal cancer and head and neck paired normal and cancer tissues. Higher expression in normal tissues (including adrenal, thyroid and PBMC) than in breast cancer tissues is a drawback to consider using BUC10 in immunotherapies. However, further experiments can be undertaken to determine if BUC10 is suitable as a target for immunotherapies for any of the gastric, oesophageal and head and neck cancers.

Preliminary studies on BUC11 mRNA expression (personal communication with Dr. Li) suggested that BUC11 mRNA is not expressed in the following normal tissues: brain, heart, kidney, liver, PBMC and testis. Moreover, BUC11 mRNA was found expressed in breast cancer tissues, with two gene products visible on agarose gel electrophoresis. The conventional RT-PCR results show that BUC11 mRNA is expressed in normal testis and breast tissues but not in any of the other 17 normal tissues tested which includes vital organs and not in any of the PBMC samples tested. BUC11 mRNA is highly expressed in breast cancer tissues (90% positive so far) but the majority of the breast cancer cell lines tested did not express BUC11 mRNA and only a low BUC11 level was found in T47D. All the testis cancer and normal tissues tested expressed similar levels of BUC11 mRNA. BUC11 mRNA was not detectable using conventional RT-PCR in any of the mesothelioma samples, melanoma samples, gastric (paired normal and cancer) tissues, kidney (paired normal and cancer) tissues and oesophageal (paired normal and cancer) tissues.

Using quantitative RT-PCR, the expression pattern obtained by conventional RT-PCR was confirmed and further investigated. Quantitative RT-PCR results corroborated the finding that normal and testis cancer tissues present similar levels of BUC11 expression however more samples would be required to have a better knowledge of the expression levels in these tissues.

High levels of BUC11 mRNA in normal breast and testis tissues were found, with higher expression in breast compared to testes, however expression was also detected in other normal tissues due to the higher sensitivity of the quantitative RT-PCR compared to the conventional RT-PCR. The level of expression in these other normal tissues such as brain and heart compared to the level of expression in testis and placenta is extremely low. Therefore, one can conclude that BUC11 mRNA expression is limited to non vital

organs, making BUC11 a potential attractive vaccine target. The testis are an immunopriviledged organ that is to say the antigen coded by any of the BUC genes would not be expressed on the surface of testicular cells in conjunction with MHC class I (spermatogonia lack MHC) therefore would not elicit an immune response. However, the normal breast tissues could be affected by the treatment but immune responses directed against normal breast cells would have similar effects to complete surgical removal of the breast. Furthermore, the level BUC11 mRNA expression confirmed in placenta suggests that pregnant women would not be able to receive immunotherapies that use BUC11.

Regarding the breast cancer cell lines, quantitative RT-PCR results confirmed the conventional RT-PCR data but also showed that a detectable but lower level of BUC11 mRNA in MDA231 than in T47D exists. The lack or the extremely low level of expression in most of the cell lines could be explained by transcriptional silencing as mentioned before. Also, the administration of 5-azacytidine (Taylor and Jones, 1982) a DNA-demethylating agent, to a breast cancer cell line that does not expressed BUC11 could potentially restore the gene expression but only in the case of BUC11 promoter hypermethylation. The result interestingly correlates with a published study in which the majority of the breast cancer cell lines tested expressed only low NY-BR-1 levels (Seil *et al.*, 2007).

Using quantitative RT-PCR, BUC11 mRNA was found to be expressed in 25% of all prostate tissues tested. The analysis established for the first time the positive expression of BUC11 mRNA in prostate tumour tissues. The levels were high in tumour tissues (malignant and unknown tumour type) when compared to the level found in normal prostate during the expression analysis of normal tissues. BUC11 mRNA was not detected in the known benign tissues. The low number of tissues tested and also the lack of information on most of the prostate tumour samples (7 out of 16) suggest that the study would need to be repeated on a larger scale in order to confirm the pattern of BUC11 expression seen in prostate tissues. This is necessary prior to determining if BUC11 could be of interest in the management of prostate cancer. In published studies, NY-BR-1 mRNA was found expressed in up to 32% of prostate cancers, sporadically expressed in normal prostate tissues (Jager *et al.*, 2005), however NY-BR-1 protein could not be detected in prostate samples (Theurillat *et al.*, 2008). The mammary gland (along with the ovary and uterus) in females and the testis in males (along with the prostate gland) are human reproductive tissues and the most important target tissues of sex steroid hormones which include estrogen, progesterone and androgen (Gao *et al.*, 2002). These hormones exert their roles in the target tissues via their specific intracellular receptors which are part of the nuclear receptor family and function as transcription factors to regulate the target gene expression (McKenna *et al.*, 1999). Tumours of the human reproductive tissues (male and female) have been linked to the abnormal expression or function of the sex steroid hormone receptors and the abnormal functioning (or expression) and mutations of these receptors has

been linked to resistance to hormonal therapies for breast and prostate cancers (Gao *et al.*, 2002). One can conclude that BUC11 could be a differentiation antigen (regulated by sex steroid hormones) of the breast and testis if antigenicity of its protein is established and if its protein expression correlates strongly with the expression of hormones receptors such as ER-alpha (Theurillat *et al.*, 2008).

Regarding mRNA expression of BUC11 in melanoma, only 11.5% of patients (unknown clinical information) tested were positive for a very low level of expression when compared to the level in normal testes. The levels in melanoma samples had to be compared to the level in testes because no "normal skin" samples were available for this study. It would be interesting to determine whether the 3 positively tested melanoma samples have a common clinical feature such as stage of the disease in order to speculate on the role of BUC11 mRNA in these tumour cells at a certain point of the development of the disease. This result proves that BUC11 mRNA expression is actually not restricted to tumours of the breast, the testes or the prostate however the low mRNA levels in melanoma samples and the low number of patients positive for mRNA molecules indicate that BUC11 might not be of interest for skin cancer.

Since it appears irrelevant to therapeutically target or exploit a gene expressed in low abundance or in a low number of patients, an accurate quantification of BUC11 in breast cancer tissues was thought indispensable to provide information on the significance of BUC11 for the management of patients. Furthermore, the hypothesis that BUC11 might be upregulated in the early stages of breast cancer and downregulated at later stages had to be challenged with expression analysis in a large panel of tissues from all stages and grades. A number of normal breast tissues were also tested for the expression of BUC11 to determine if the silencing of gene expression observed in 10% of breast cancer tissues is linked to cancer mechanisms or to normal molecular or cellular mechanisms. For this purpose, cDNA from paired normal and tumour tissues from 109 breast cancer patients were utilised for quantitative RT-PCR analysis, along with primers for BUC11 and the control genes GAPDH, HPRT1 and TBP. Since the choice of the best control genes for a given panel of cDNA samples using quantitative RT-PCR is essential, the geNorm VBA applet for Microsoft Excel (free download at http://medgen.ugent.be/genorm/) was applied to select HPRT1 and TBP as the most stable control genes for this study. The results were sorted according to the clinical information provided by Dr Angelos Gritzapis (The Cancer Immunology and Immunotherapy Center, Saint Savas Cancer Hospital, Athens, Greece): age at diagnosis, stage of the tumour, grade of the tumour and HER2-status, but the clinical information made available for this study was restricted. In order to further classify the samples and have a more refined BUC11 expression pattern, it would have been of interest to know the status of other biomarkers (e.g. ER, PR, p53 mutation), the method used to determine the stage and grade of the tumour, the cut off point in

determining the HER2 status as well as the distance in the breast between the extracted tumour tissues and their normal counterparts. Unfortunately, as the tissues were harvested recently (less than 5 years ago), the follow-up of patients (survival, tumour recurrence, distance metastasis formation) was not available and did therefore not allow the prognostic potential of BUC11 to be determined.

The tissues used in this study have not been laser-capture microdissected and the distance between the harvested tumour and normal tissue is unknown and it is very possible that the tumour tissue contained some normal cells and vice versa, therefore, the gene expression of BUC11 might have been diluted. It is well-known that breast tissue is heterogeneous and that the use of whole heterogeneous tissues in molecular analytical techniques such as quantitative RT-PCR poses the problem of interpretation of the data because the true genomic state of a given cell population (disease or normal) can only be ascertained using a very pure cell population (Espina *et al.*, 2006). Laser-capture microdissection can overcome this issue and is the process of isolating a pure, morphologically distinct cell population from a tissue section, cytological preparation or live cell culture which is made, by nature, of heterogeneous types of cells (Emmert-Buck *et al.*, 1996). Recently, Walton and colleagues showed for the first time upregulation of the expression of the gene ERbeta in prostate tumours which have been laser microdissected (Walton *et al.*, 2009). Consequently, one can doubt that the results obtained in this study show the true biological profile of the disease and disease-free cells, however the data should still provide some clues towards the understanding of BUC11 mRNA expression pattern in breast tissues. In these breast samples, BUC11 mRNA is expressed at widely varying levels. It was observed that 60.6% of patients had BUC11 mRNA expression upregulated in their tumour when compared to BUC11 mRNA expression in the matched normal adjacent breast tissue. It appears that, in almost half of the breast cancer patients tested in this study, molecular control mechanisms downregulated the transcription of BUC11 gene in the tumour cells during the process of tumorigenesis. Transcriptional silencing (BUC11 mRNA molecules not detectable) was found in the tumour tissue but not in the normal matched tissue of 3 patients, representing only about 3% of the total number of patients tested. The age group, stage of the tumour and HER2-status of these 3 patients did not seem to be significant and only one clinical factor, the grade (II), appeared to show a trend with the expression of BUC11 (increased in grade II and decreased in grade III). In these patients, it appears that BUC11 gene expression was repressed during tumorigenesis. The transcriptional silencing observed can not be due to mutagenic inactivation because the patients express BUC11 mRNA expression in their normal breast tissue, however it may be explained by hypermethylation mechanisms, which has been reported in various human tumours. In this study, BUC11 mRNA regulation could not be correlated with the age of the patient at diagnosis, thus it is likely that age has no effect on the expression of BUC11 in breast tissues. In women, a naturally-occurring fall of oestrogen

levels leads to the menopause. As menopausal status is usually linked to the age of an individual (although premature menopause can happen), one can then hypothesize that oestrogen levels do not have a direct effect on the regulation of BUC11 expression in breast tissues.

HER2 is an epidermal growth factor receptor which is over-expressed in up to 30% of breast cancers and is an important prognostic factor as it is associated with aggressive phenotype of the tumour, non-responsiveness to hormonal treatment and reduced sensitivity to conventional chemotherapies (Slamon *et al.*, 1989). In this study, BUC11 mRNA expression and HER2-status did not correlate, thus it appears that HER2 and BUC11 are not co-expressed to the same levels in breast tissues. The stage of the breast cancer is a diagnostic factor and the grade of the tumour is a factor of aggressiveness, thus prognosis. It is worth noting that histological grade has a very subjective prognostic value as suggested by Volpi and colleagues in 2004 (Volpi *et al.*, 2004). Some tumours are categorised by clinicians as grade II only because the results of its three components (tubule formation, nuclear pleomorphism and mitotic count) do not clearly fall in the category of grade I or grade III. In these samples, BUC11 mRNA levels did not significantly follow any pattern regarding the progression or the aggressiveness of the disease. It is worth noting that the distribution of patients in the different categories of clinical information about the tumour is unbalanced. For instance, there is a low number of these patients tested for a grade I tumour compared to the number of patients whose tumours are of grade II and III. Similarly, few patients have been diagnosed with a stage I, III or IV tumour compared to the patients who had a stage II disease. Before ruling out BUC11 as a good candidate for diagnosis and prognosis of breast cancer, a comparison of BUC11 expression in breast tumour tissues with the expression of other genes recognised as biomarkers for breast cancer prognosis and diagnosis would be necessary.

In light of all the mRNA expression data collected for BUC11, this gene was identified as being the most promising of the BUC gene family. Cancer/testis (CT) antigens are tumour antigens with an expression in normal tissues confined to male germ cells in the testis but not in adult somatic tissues (with the exception of ovary and placenta in some cases) hence BUC6, BUC9, BUC10 and BUC11 genes do not fall into the category of CT antigens. BUC11 is highly expressed in breast and testis tissues and found at high levels in up to 97% of the breast cancer tissues tested. The tissue specificity of BUC11 expression profile provides rationale to consider BUC11 as a tissue-specific gene involved in the differentiation of breast and testis tissues. BUC11 is a potential target for vaccine strategies however most of the therapeutic relevance of the BUC11 breast and testis-associated gene relies on the study of the protein encoded by this gene. It is of prime importance to confirm that the hypothetical protein is produced using an antibody specific to this antigen in immunoassays.

Chapter 4

BUC11 Protein Expression Analysis

4.1 Introduction

The BUC family comprises novel unpublished breast-associated genes identified by mining databases in our laboratory. Their respective mRNA expression levels in a variety of tissues of normal and cancerous origin were thoroughly characterised in this study using RT-PCR based assays. Out of all of the results obtained, BUC11 gene appeared to be a promising candidate for breast cancer management. BUC11 mRNA was found to be highly expressed in breast and testis tissues (normal and tumour) and high mRNA levels were obtained in up to 97% of breast cancer cases. However, all the information gathered previoulsy on BUC11 is restricted to the messenger RNA level and the therapeutic applicability of BUC11 depends on proof that the BUC11 protein is naturally produced by the cells and, if present, that BUC11 protein expression patterns show relevance to the management of the disease. BUC11 could potentially be a target for novel molecular therapies or for immunotherapies and it could also be a marker for diagnosis or prognosis of breast cancer. For instance, for an antigen to be effective as a target in T-cell-based immunotherapies, it has to be co-expressed with HLA class I antigens (Theurillat *et al.*, 2007). In addition, there have been many reports of non-protein-coding RNA that might have important regulatory roles in cells which can be assessed by specific gene knockdown experiments (Huttenhofer *et al.*, 2005).

Quantification of a gene at both the mRNA and protein levels is therefore of prime importance in order to gain a comprehensive view of the mechanisms controlling gene expression. Whether mRNA and protein levels correlates or not can provide valuable information on how the cell regulates the expression of the gene of interest (Greenbaum *et al.*, 2003). Therefore, it was necessary in the present study to corroborate mRNA quantitation (Chapter 3) with protein quantitation. The later was performed using immunoassays with the custom-made anti-BUC11 monospecific polyclonal antibody. Firstly, information regarding the sequence of BUC11 predicted protein was obtained by means of database mining to highlight any similarities with known proteins, potential roles, subcel-

lular localisations and post-translational modifications. Secondly, anti-BUC11 antibody was used in a series of immunoassays in order to determine the expression of BUC11 protein, to characterise its expression patterns in diverse normal and tumour tissues and to compare protein levels with mRNA levels previously obtained. These assays included immunoblotting, immunofluorescence and immunohistochemistry.

4.2 Results

4.2.1 Data mining with BUC11 predicted protein sequence

As shown in Chapter 3, the BUC11 mRNA sequence shares 99% similarity with hCG25653 mRNA sequence and the 2 nucleotide difference is due to errors in sequencing. For this research programme, hCG25653 mRNA and hCG25653 predicted protein will be referred as BUC11 mRNA and BUC11 predicted protein respectively.
The predicted protein sequence of BUC11 gene was compared to database entries by a blastp (protein-protein BLAST) search (http://blast.ncbi.nlm.nih.gov/Blast.cgi). When the blastp search was undertaken using "all species", only one entry was found and this correspond to hCG25653 predicted protein sequence (Table 4.1). This result suggested that BUC11 predicted protein sequence does not significantly match any other deposited protein sequences in the database. However, when the blastp search was applied using the option "Homo sapiens only", surprisingly additional entries were shown (Table 4.1). All entries, apart from the expected hCG25653 predicted protein, belong to the Baculoviral IAP repeat-containing protein 4 (XIAP). The matched amino acids between the two proteins are found in the region 49-67 of BUC11 predicted protein. XIAP is a chromosome X-linked protein; it binds to tumour necrosis factor receptor-associated factors TRAF1 and TRAF2, inhibits the cell-death associated proteases caspase-3 and caspase-7 and inhibits apoptosis induced by menadione (causes oxidative stress) and interleukin-1 beta converting enzyme (ICE, also known as caspase-1) (Entrez Gene, NCBI). Studies have highlighted the interest in targeting XIAP protein for the treatment of breast cancer (Lee and Cho, 2008; Choi *et al.*, 2008; Aird *et al.*, 2008). For instance, Hong and colleagues showed that suppression of a protein (p34SE-1) that normally protects XIAP from being degraded and which is highly expressed in breast cancer tissues induces apoptosis, making this protein a target in the development of novel breast cancer treatments (Hong *et al.*, 2009). Also, some clinical trials for the treatment of cancer are ongoing using an antisense oligonucleotide against XIAP mRNA (Tamm, 2008). The blastp search was repeated with another species, Mus musculus, and the sequence of the mouse equivalent for human XIAP, Xiap (90.9% protein sequence similarity between XIAP human and Xiap mouse), was the only significant match to BUC11 predicted protein sequence, within the same region 49-67 (Table 4.1). Following this analysis, another narrowing blastp search was generated by comparing the region 49-67 of BUC11 predicted protein against all de-

posited sequences for Homo sapiens proteins. The entries are shown in Tables 4.2, 4.3, 4.4 and 4.5. This not only resulted in hCG25653 and XIAP as significant hits but also other proteins, with varying percentage and regions of similarity. Tables 4.6 and 4.7 show the role, if known, of these proteins. The proteins XIAP and BIRC8 inhibits apoptosis (Schimmer *et al.*, 2006; Lagace *et al.*, 2001) and the proteins CAPRIN2 and PPP2R1A inhibits cell growth (Aerbajinai *et al.* 2004; Lin *et al.*, 2006). The proteins FGD3, PKN2, DNAH2 and PLEKHH3 are associated with the structure of the cell (Entrez Gene, 2009; Schmidt *et al.*, 2007; Chapelin *et al.*, 1997). Interestingly, JMJD3 has been linked with metastatic prostate cancer, CHRM3 with invasive migration of melanomas, the lack of MAP2K7 with resistance to treatment of colon cancer cells and FUT4 as well as MLL2 with cancer cell growth (Xiang *et al.*, 2007; Oppitz *et al.*, 2008; Vasilevskaya *et al.*, 2008; Zhang *et al.*, 2008; Mo *et al.*, 2006).

BUC11 predicted protein Blastp (all species)					
Description (GenBank accession number)	Gene ID	Function	Query coverage	Amino acids	Maximum identity
PREDICTED: similar to hCG25653 [Homo sapiens] (XP_938646)	hCG25653	Unknown	75/75	1-75	100%
BUC11 predicted protein Blastp (Homo sapiens)					
Description (GenBank accession number)	Gene ID	Function	Query coverage	Amino acids	Maximum identity
PREDICTED: similar to hCG25653 [Homo sapiens] (XP_938646)	hCG25653	Unknown	75/75	1-75	100%
unnamed protein product [Homo sapiens] (BAG36609)	X-linked inhibitor of apoptosis (XIAP)	Inhibits apoptosis	11/21	49-67	61%
IAP-like protein ILP (AAC50518)	X-linked inhibitor of apoptosis (XIAP)	Inhibits apoptosis	11/21	49-67	61%
baculoviral IAP repeat-containing protein 4 [Homo sapiens] (NP_001158)	X-linked inhibitor of apoptosis (XIAP)	Inhibits apoptosis	11/21	49-67	61%
BUC11 predicted protein Blastp (Mus musculus)					
Description (GenBank accession number)	Gene ID	Function	Query coverage	Amino acids	Maximum identity
baculoviral IAP repeat-containing 4 [Mus musculus] (NP_033818)	X-linked inhibitor of apoptosis (Xiap)	Inhibits apoptosis	11/21	49-67	61%

Table 4.1: BUC11 predicted protein sequence similarity to protein sequences previously deposited in the GenBank database.
The predicted protein sequence of BUC11 was uploaded in the blastp tool (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Three searches were conducted: all species proteins, Homo sapiens only proteins and Mus musculus only proteins. The significant search results are shown. “Query coverage” gives the number of amino acids of the query matching the amino acids of the database entries/total number of amino acids of the query. “Amino acids” gives the region of BUC11 predicted protein.

Gene ID [GenBank accession number]	Query coverage	Amino acids	Maximum identity
hCG25653 [XP_938646]	19/19	49-67	100%
X-linked inhibitor of apoptosis (XIAP) [NP_001158]	11/21	49-67	61%
coiled-coil domain containing 87 (CCDC87) [NP_060689]	9/13	9-18	69%
FYVE, RhoGEF and PH domain containing 3 (FGD3) [NP_001077005]	9/14	7-19	64%
protein kinase N2 (PKN2) [AAI25200]	8/11	6-15	81%
caprin family member 2 (CAPRIN2) [AAI17673]	7/9	5-13	88%
chaperonin containing TCP1, subunit 7 (eta) pseudogene 2 (CCT7P2) [XP_001727067]	8/13	6-18	61%
dynein, axonemal, heavy chain 2 (DNAH2) [BAA96027]	8/11	7-16	81%
	6/11	9-17	72%
	3/4	4-7	100%
	3/3	5-7	100%
protein phosphatase 2 (formerly 2A), regulatory subunit A, alpha isoform (PPP2R1A) [BAC03652]	8/12	6-17	75%
	3/3	5-7	100%
zinc finger protein 397 (ZNF397) [AAN65175]	8/16	5-15	56%
	2/2	17-18	100%

Table 4.2: BUC11 predicted protein (amino acids 49-67) sequence similarity to Homo sapiens protein sequences previously deposited in the GenBank database (1).
The sequence of amino acids 49-67 of BUC11 predicted protein was uploaded in the blastp tool (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The search was conducted on Homo sapiens only proteins. The significant search results are shown. “Query coverage” gives the number of amino acids of the query matching the amino acids of the database entries/total number of amino acids of the query. “Amino acids” gives the region of BUC11 predicted protein.

Gene ID [GenBank accession number]	Query coverage	Amino acids	Maximum identity
EMI domain containing 1 (EMID1) [CAQ10944]	9/12	3-13	75%
	5/7	1-7	85%
	3/6	2-7	83%
jumonji domain containing 3, histone lysine demethylase (JMJD3) [BAA21572]	7/9	4-12	77%
	4/4	10-13	100%
mitogen-activated protein kinase kinase 7 (MAP2K7) [AAC16273]	5/6	13-18	100%
baculoviral IAP repeat-containing 8 (BIRC8) [AAH39318]	10/21	1-19	57%
ubiquitin specific peptidase 11 (USP11) [AAH63668]	6/11	12-19	63%
cholinergic receptor, muscarinic 3 (CHRM3) [AAK68114]	6/6	5-10	100%
fucosyltransferase 4 (FUT4) [NP_002024]	7/11	8-18	72%
piggyBac transposable element derived 1 (PGBD1) [BAF83917]	8/15	5-14	53%
pleckstrin homology domain containing, family H (with MyTH4 domain) member 3 (PLEKHH3) [EAW60863]	6/8	11-18	87%
	7/14	1-14	71%
	4/4	1-4	100%
myeloid/lymphoid or mixed-lineage leukemia 2 (MLL2) [NP_003473]	6/7	10-16	85%
	4/4	8-11	100%
	3/3	14-16	100%

Table 4.3: BUC11 predicted protein (amino acids 49-67) sequence similarity to Homo sapiens protein sequences previously deposited in the GenBank database (2).
The sequence of amino acids 49-67 of BUC11 predicted protein was uploaded in the blastp tool (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The search was conducted on Homo sapiens only proteins. The significant search results are shown. “Query coverage” gives the number of amino acids of the query matching the amino acids of the database entries/total number of amino acids of the query. “Amino acids” gives the region of BUC11 predicted protein.

Gene ID	Amino acid sequence																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
hCG25653	G	G	G	A	K	S	W	K	P	S	L	Q	S	W	E	K	Q	W	Y
XIAP	G	G	G				W	K	P	S				W	E			W	Y
CCDC87									P	S	L	Q	S	W	E	K		W	
FGD3								W	K		L	Q		W		K	Q	W	Y
PKN2						S	W	K	P	S		Q	S	W					
CAPRIN2					K	S	W		P	S		Q	S						
CCT7P2								K	P	S	L	Q							
DNAH2							W	K	P			Q	S	W	E	K			
									P	S	L	Q			E	K			
				A	K		W												
					K	S	W												
PPP2R1A					K	S	W												
ZNF397					K	S	W	K	P		L			W	E				
																	Q	W	

Table 4.4: BUC11 predicted protein (amino acids 49-67) sequence similarity to Homo sapiens protein sequences previously deposited in the GenBank database (3).
The sequence of amino acids 49-67 of BUC11 predicted protein was uploaded in the blastp tool (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The search was conducted on Homo sapiens only proteins. The significant search results are shown.

Gene ID	Amino acid sequence																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
hCG25653	G	G	G	A	K	S	W	K	P	S	L	Q	S	W	E	K	Q	W	Y
EMID1			G		K	S	W		P	S	L	Q	S						
	G	G	G	A			W												
		G	G				W												
JMJD3				A	K		W	K	P		L	Q							
MAP2K7													S	W	E		Q	W	
BIRC8	G	G	G				W	K	P					W	E			W	Y
USP11													S	W	E	K		W	Y
CHRM3					K	S	W	K	P	S									
FUT4								K	P	S				W	E	K		W	
PGBD1					K	S	W												
PLEKHH3											L	Q	S	W	E			W	
	G	G	G	A			W		P					W					
	G	G	G	A															
MLL2										S	L	Q		W	E	K			
								K	P	S	L								

Table 4.5: BUC11 predicted protein (amino acids 49-67) sequence similarity to Homo sapiens protein sequences previously deposited in the GenBank database (4).
The sequence of amino acids 49-67 of BUC11 predicted protein was uploaded in the blastp tool (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The search was conducted on Homo sapiens only proteins. The significant search results are shown.

Gene ID	Protein function	Reference
hCG25653	Unknown	Entrez Gene (NCBI), 2009
XIAP	Inhibits apoptosis, binds to caspases 3, 7, and/or 9, but not caspase 8	Schimmer *et al*., 2006
CCDC87	Protein-protein interaction domain	Entrez Gene (NCBI), 2009
FGD3	Role in actin cytoskeleton organisation and regulation of Rho protein signal transduction	Entrez Gene (NCBI), 2009
PKN2	Regulated by Rho GTPases to control entry into mitosis and exit from cytokinesis	Schmidt *et al*., 2007
CAPRIN2	Role in negative regulation of cell growth	Aerbajinai *et al*., 2004
CCT7P2	Role in protein folding	Entrez Gene (NCBI), 2009
DNAH2	Dyneins are microtubule-associated motor protein complexes	Chapelin *et al*., 1997
PPP2R1A	Role in the negative control of cell growth and division, mediates dephosphorylation of BCL-2	Lin *et al*., 2006
ZNF397	Interphase to early prophase-specific mammalian centromere protein	Bailey *et al*., 2008

Table 4.6: Known function of Homo sapiens proteins which sequences deposited in the GenBank database are similar to peptide 49-67 of BUC11 predicted protein (1).
A blastp search (http://blast.ncbi.nlm.nih.gov/Blast.cgi) was carried out using the sequence of amino acids 49-67 of BUC11 predicted protein, on Homo sapiens only proteins. The Gene ID and the published function(s) of the proteins which sequences were found similar to amino acids 49-67 of BUC11 predicted protein are given.

Gene ID	Protein function	Reference
EMID1	Protein binding	Entrez Gene (NCBI), 2009
JMJD3	Histone demethylase, upregulated in prostate cancer, expression higher in metastatic prostate cancer	Xiang *et al*., 2007
MAP2K7	Role in cell responses to proinflammatory cytokines and environmental stresses, MAP2K7-deficient hypoxic colon cancer cells resistant to treatment with oxaliplatin	Vasilevskaya *et al*., 2008
BIRC8	Role in the control of apoptosis in the testis by direct inhibition of caspase 9	Lagace *et al*., 2001
USP11	Role in DNA damage repair functions within the BRCA2 pathway independently of BRCA2 deubiquitination	Schoenfeld *et al*., 2004
CHRM3	Controls smooth muscle contraction, associated with invasive migration of melanomas	Oppitz *et al*., 2008
FUT4	Key enzyme for the synthesis of LeY (highly expressed in a variety of human carcinomas of epithelial cell origin), suppressed expression linked to inhibition of cancer cell growth	Zhang *et al*., 2008
PGBD1	Unknown, expressed only in brain	Entrez Gene (NCBI), 2009
PLEKHH3	Component of the cytoskeleton	Entrez Gene (NCBI), 2009
MLL2	MLL2 complex (MML2, ASH2, RBQ3, and WDR5) is a coactivator for estrogen receptor alpha, role in ERalpha-positive cancer cell growth	Mo *et al*., 2006

Table 4.7: Known function of Homo sapiens proteins which sequences deposited in the GenBank database are similar to peptide 49-67 of BUC11 predicted protein (2).
A blastp search (http://blast.ncbi.nlm.nih.gov/Blast.cgi) was carried out using the sequence of amino acids 49-67 of BUC11 predicted protein, on Homo sapiens only proteins. The Gene ID and the published function(s) of the proteins which sequences were found similar to amino acids 49-67 of BUC11 predicted protein are given.

The *in silico* analysis of post-translational modifications, predicted motifs and structure of BUC11 predicted protein was carried out using internet-based tools. Firstly, the BUC11 predicted amino acid sequence was uploaded in the program PSORT II which predicts the possible localisations of the protein in cells by comparing the sequence of interest with the different sequence characteristics of known protein sorting signals (http://psort.ims.u-tokyo.ac.jp/form2.html). To enter the nucleus, proteins either have nuclear localisation signal (NLS) or they are co-transported with a protein that has its own NLS. In 1995, Hicks and Raikhel reviewed the three categories of NLSs (Hicks and Raikhel, 1995). One category comprised pat4 (4 basic amino acids K or R; or 3 basic amino acids, K or R followed by H or P) and pat7 (P followed within 3 residues by 3 K/R residues out of 4) of SV40 large T antigen. Another category is the bipartite NLS which is formed by 2 basic residues, 10 residue spacer and a region of at least 3 basic residues

out of 5. It was found, as shown in Table 4.8, that BUC11 predicted protein has one pat4 NLS, corresponding to the amino acids 14-17. Furthermore, the protein content of basic residues is slightly above 20% and the consensus NLS score equals the threshold. These results suggests the possibility that the protein may localise to the nucleus. Additionaly, the results obtained for Reinhardt's method for both cytoplasmic/nuclear discrimination and the *k*-NN prediction imply, with up to 56.5% confidence, that the protein is cytoplasmic. Therefore, it is possible that BUC11 predicted protein, if produced, is likely to preferentially reside in the cytoplasm.

Subprogram	Result
Signal sequence recognition	No N-terminal signal peptide
Nuclear localization signals (pat4)	KKKR (14-17)
Nuclear localization signals (pat7)	None
Nuclear localization signals (bipartite)	None
Nuclear localization signals (content of basic residues)	21.3%
Nuclear localization signals (NLS score)	0.79
ER retention motif in the C-terminus	None
ER membrane retention signals	None
Peroxisomal targeting signal in the C-terminus	None
Second peroxisomal targeting signal	None
RNA-binding motif	None
Actinin-type actin-binding motif	None
N-myristoylation pattern	None
DNA binding motifs	None
Ribosomal protein motifs	None
Reinhardt's method for cytoplasmic/nuclear discrimination	Cytoplasmic (55.5%)
Coiled-coil regions	None
k-NN prediction	Cytoplasmic (56.5%) [nuclear (26.1%), mitochondrial (13%) and cytoskeletal (4.3%)]

Table 4.8: PSORT analysis of BUC11 predicted protein sequence.
PSORT is a computer program (http://psort.ims.u-tokyo.ac.jp/form2.html) that predicts the localisation of proteins in cells by comparing the protein sequences with the different sequence characteristics of known protein sorting signals. Proteins where the NLS score is equal to or above the 0.8 threshold are predicted to spend time in the nucleus. Proteins which comprise of basic residues higher than 20% are predicted to be more nuclear than cytoplasmic (nuclear proteins are rich in basic residues in general).

The sequence of BUC11 predicted protein was also analysed using the program PROSITE (http://www.expasy.ch/prosite/) which compares protein sequences with patterns of documented protein domains, protein families, functional sites and structural sites. The search only revealed 4 matches, as seen in Table 4.9. The protein has 2 protein kinase C phosphorylation sites (8-10 and 54-56), one tyrosine kinase phosphorylation site (17-24) and one N-myristoylation site (50-55). The match for the N-myristoylation site is actually in conflict with the data from the analysis with PSORT. It is known that N-myristoylation

promotes protein-membrane and protein-protein interactions and that phosphorylation by kinases is essential for the regulation of enzyme activity.

Motif	Amino acid position
Protein kinase C phosphorylation site	8-10, 54-56
Tyrosine kinase phosphorylation site	17-24
N-myristoylation site	50-55

Table 4.9: PROSITE analysis of BUC11 predicted protein sequence.
PROSITE (http://www.expasy.ch/prosite/) compares the protein sequence with patterns of documented protein domains, protein families, functional sites and structural sites.

In addition, the sequence of BUC11 predicted protein was uploaded into the program Pfam (http://pfam.sanger.ac.uk/) which is a protein family database comprising sequence alignments and hidden Markov models (HMMs). Pfam compares the sequence of the protein of interest to protein domains in order to provide clues on the function of the protein. Pfam analysis of BUC11 predicted protein did not produced any match with sequence alignments of the database or HMMs which exceeded the defined cut off value. Further information on the predicted protein was obtained through the JustBio tool ProtCalc (http://www.justbio.com/protcalc/index.php) such as the amino acid composition, the molecular weight (8724.36 Daltons) and the theoretical isoelectric point (pI, 9.91). The predicted protein is quite small, with a molecular weight of about 8.7kDa. The theoretical pI of 9.91 implies that the protein is basic and rich in lysine (composing 17.33%).

The *in silico* software packages used in the BUC11 protein sequence analysis are only predictors, therefore the interpretation and results generated should be used as guidance only.

4.2.2 ELISA using anti-BUC11 antibody

To date, hCG25653 (BUC11) has not been the subject of any publication. hCG25653 has only been described in the databases as a protein-coding gene which RefSeq status (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene&cmd=search&term=hCG25653, updated 05/12/2008) is “model”, that is to say the sequence has not been curated and has been derived from a whole-genome computational annotation pipeline. The predicted coding region has 228bp which starts from base number 95 and finishes base number 322. The protein (75 amino acids) is “predicted”(http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?val=XP_938646.1, updated 29/02/08). The protein sequence is given in Figure 3.13 (Chapter 3).
So far in this study, BUC11 has been detected at the messenger RNA level in various tissues and cell lines. In order to investigate the possibility of BUC11 expression at the protein level, a polyclonal monospecific antibody was custom-made by Pacific Immunology

(USA). The antibody was raised in rabbit against the following 13-amino acid peptide ($BUC11_{7-19}$): PSKRLFFKKKRLC. Peptide chemists at Pacific Immunology analysed the protein sequence and recommended the most immunogenic sequences, taking into account the characteristics of the predicted protein such as its hydrophobicity, hydrophilicity and folding. However, it was possible that the selected peptide sequence (linear epitope) would not correspond to an exposed region of the native protein. This particular peptide sequence ($BUC11_{7-19}$) was selected after protein BLAST analysis (blastp, Homo sapiens only, all non-redundant sequences, http://blast.ncbi.nlm.nih.gov/Blast.cgi) to assure a minimal risk of cross-reaction with similar epitopes of non-targeted proteins. As expected, the peptide sequence was 100% homologous to a region of the BUC11 predicted protein. The other matched proteins, matching 3 or more amino acids (minimum length of binding site for antibody) of the 13-amino acid peptide and their molecular weight are listed in Table 4.10. Western Blotting analysis could potentially show anti-BUC11 antibody binding to one or more of these hCG proteins of similar molecular weight to BUC11 protein instead of the targetted protein, however differences in the molecular weight and hence the band observe on the Western Blot should provide evidence as to the identity of the binding protein. In addition, anti-BUC11 antibody is polyclonal therefore it should recognise a panel of different epitopes (11 epitopes of 3-amino acid long) from the 13-amino acid peptide, increasing the chances of specifically binding to BUC11 protein and not to the same epitopes on non-targetted proteins. The serum was affinity purified against the peptide sequence thus the antibodies specific to the targetted protein (monospecific) were isolated. An ELISA assay was performed to quantify peptide-specific antibody present in the serum, flow-through and purified antibody samples.

Protein ID [GenBank accession number]	Molecular weight (Da)	Matched sequences
11176 BAZ2A [BAG51228.1]	94,570.43	SKR, LFF, PSKR
23500 DAAM2 [EAX03995.1]	124,510.61	LFFKKK, SKRL
89853 FAM125B [AAH28675.1]	24,516.03	KKRLC
9184 BUB3 [NP_004716.1]	37,154.90	KKRLC
11047 ADRM1 [NP_008933.2]	42,153.34	SKRLFF
hCG2018963 [EAX10079.1]	11,206.15	KKRL
hCG2039006 [EAW63870.1]	11,277.37	KKKRLC
22989 MYH15 [NP_055796.1]	224,620.33	RLFFK
137835 TMEM71 [NP_001138625.1]	26,291.31	KKKR
hCG1780483 [EAX06649.1]	35,655.51	KRL, FFKK
153643 FAM81B [NP_689761.2]	52,022.84	RLFFK
51202 DDX47 [NP_957518.1]	45,169.40	PSK, KRL, KKKR
199920 C1orf168 [NP_001004303.3]	82,070.38	KRL, FFKK
4803 NGF [NP_002497.2]	26,958.79	RLF
7398 USP1 [BAG35333.1]	88,208.63	PSK
hCG2039120 [EAX11523.1]	9,036.56	SKRL, FKK
266811 NPSA [AAO16090.1]	29,233.44	KRLF
7273 TTN [NP_597676.2]	3,006,837.27	FFKKK, RLFF, KRLC
728858 C12orf71 [A8MTZ7.2|CL071_HUMAN]	33,989.98	KRLF
64111 NPVF [NP_071433.3]	22,326.72	FKK
53944 CSNK1G1 [BAH13567.1]	43,508.70	FFK
57644 MYH7B [A7E2Y1.2|MYH7B_HUMAN]	221,389.44	LFFK
9881 TRANK1 [NP_055646.2]	336,222.88	KRLFF
201625 DNAH12 [NP_848599.3]	356,944.17	KRLFF, KKRL
27068 PPA2 [BAG61178.1]	30,008.51	RLFFK
9692 KIAA0391 [BAG64973.1]	65,459.58	RLFFK
64427 TTC31 [BAG64213.1]	22,127.15	KRL, KKKR, KKRL
54816 ZNF280D [BAG58173.1]	116,368.54	PSKR, FKK
353116 RILPL1 [EAW98415.1]	47,242.58	KRLF, KKRL

Table 4.10: BUC11 peptide sequence (7-19) similarity to protein sequences previously deposited in the GenBank database.
The sequence of amino acids 7-19 of BUC11 predicted protein was uploaded into the blastp tool (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The search was conducted on Homo sapiens only proteins. The "matched sequences" are the matched amino acid sequences of the proteins to the amino acid sequence of BUC11 peptide and only sequences of 3 or more consecutive amino acids are shown.

Upon receiving and aliquoting rabbit anti-BUC11 polyclonal antibody and BUC11 peptide, an ELISA was carried out to test the binding strength of the antibodies to the antigen (see Chapter 2). Briefly, the microtiter plate was coated with the peptide and the unbound sites were blocked to avoid false positive results. Then, BUC11 antibody was added to the wells followed by anti-rabbit IgG-HRP. The colored positive reaction was measured at 450nm. As seen in Figure 4.1, the ideal concentration of BUC11 antibody for the widest range of peptide concentration is 10ng/ml. The result of this ELISA was confirmed by a second ELISA (Figure 4.2) although 40ng/ml of BUC11 antibody is more appropriate for the concentration of peptide equal or below 0.25μg/ml.

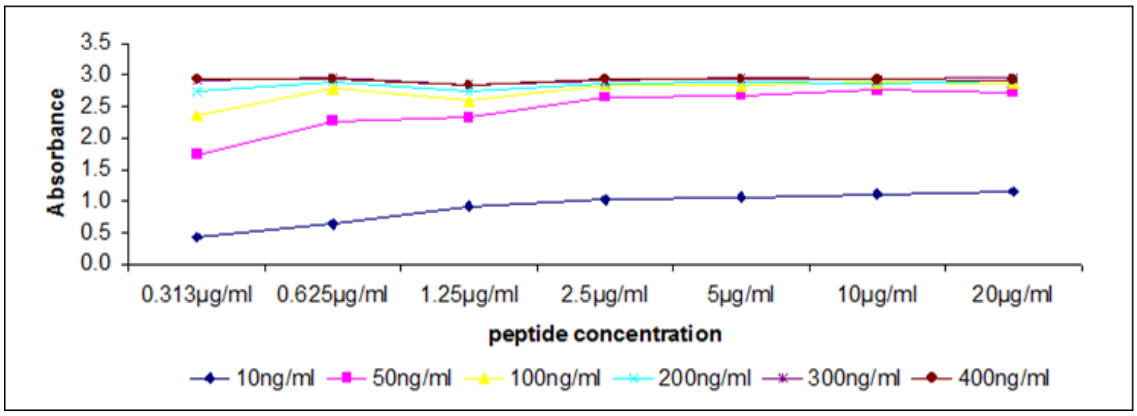


Figure 4.1: Binding strength of rabbit anti-BUC11 polyclonal antibody to BUC11 peptide by ELISA (1).
ELISA was carried out using the dilutions of antigen given on the graph in μg/ml and the dilutions of antibody given in the legend of the graph in ng/ml. Absorbances were measured at 450nm. The results are shown as Absorbance(sample)-Absorbance(blank).

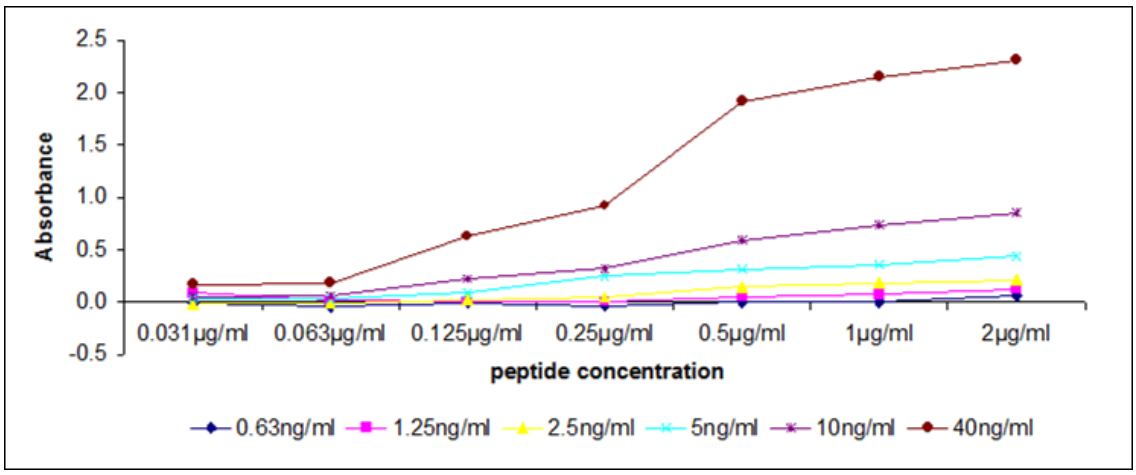


Figure 4.2: Binding strength of rabbit anti-BUC11 polyclonal antibody to BUC11 peptide by ELISA (2).
ELISA was carried out using the dilutions of antigen given on the graph in μg/ml and the dilutions of antibody given in the legend of the graph in ng/ml. Absorbances were measured at 450nm. The results are shown as Absorbance(sample)-Absorbance(blank).

4.2.3 Western Blotting using anti-BUC11 antibody

The anti-BUC11 polyclonal antibody was validated using Semi-Dry Western Blotting as described in Chapter 2. The antibody was used to probe a cell lysate produced from a previously identified BUC11 mRNA-positive breast cancer cell line MDA231. The controls for the experiment were "secondary antibody only", "rabbit IgG antibody (isotype control) with secondary antibody" and "peptide block". As shown in Figure 4.3, a strong band was observed on the developed membrane with a molecular weight in the region of 15kDa and a weaker band in the region of 30kDa. The controls gave expected re-

sults: the secondary antibody as well as the rabbit IgG isotype control did not show any non-specific background staining and the BUC11 peptide (7-19) disrupted interactions between the anti-BUC11 antibody and the BUC11 protein.

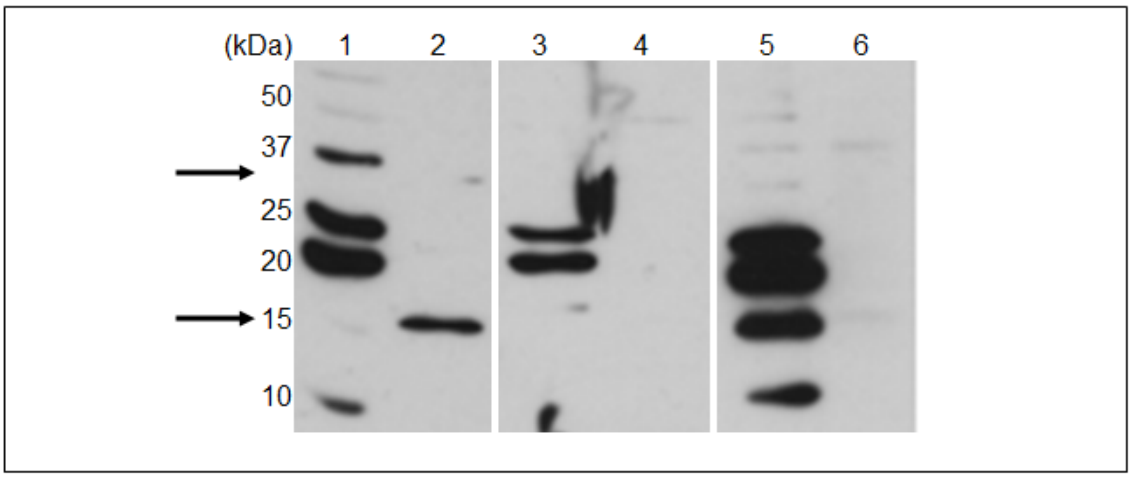


Figure 4.3: Western Blot of human breast cancer cell line MDA231.
Western Blotting was carried out by probing with anti-BUC11 antibody. The cell lysate used was produced from the human breast cancer cell line MDA231 (positive for BUC11 mRNA expression). Results are from the same experiment. (1), (3) and (5): protein marker. (2) MDA231 with primary and secondary antibodies, (4) MDA231 with rabbit IgG isotype control and secondary antibody, and (6) MDA231 with peptide block.

Anti-BUC11 antibody was also used to probe cell lysates produced from human breast cancer cell lines MDA468 (no BUC11 mRNA molecules detected, Chapter 3) and MDAP3 (very low levels of BUC11 mRNA, Chapter 3) as well as on cell lysates produced from mouse lymphoma cell line ALC which was tested negative for BUC11 mRNA expression and used for transfection with the recombinant plasmid pBudCE4.1/BUC11 (quantitative RT-PCR, cloning and transfection data are given in Chapter 5). Figure 4.4 shows that the two bands observed with the MDA231 cell lysate also appeared with ALC/pBudCE4.1/BUC11 cell lysate, MDAP3 cell lysate and MDA468 cell lysate but not with ALC/- cell lysate. Thus, ALC/pBudCE4.1/BUC11 cells transfected with the construct plasmid pBudCE4.1/BUC11 seemed to successfully express BUC11 predicted protein and the protein can be detected on the blot with anti-BUC11 antibody. However, detection of BUC11 protein expression in MDA468 cells, similar to the expression found in MDA231 cells, is unexpected as no mRNA molecules could be detected by quantitative RT-PCR. This could be due to a very low number of target mRNA molecules present in the sample and it is well documented that less than 10 (or 100) copies of target cDNA per sample can only be quantified after optimisation of the assay (choice of master mix, number of replicates and PCR machine conditions).

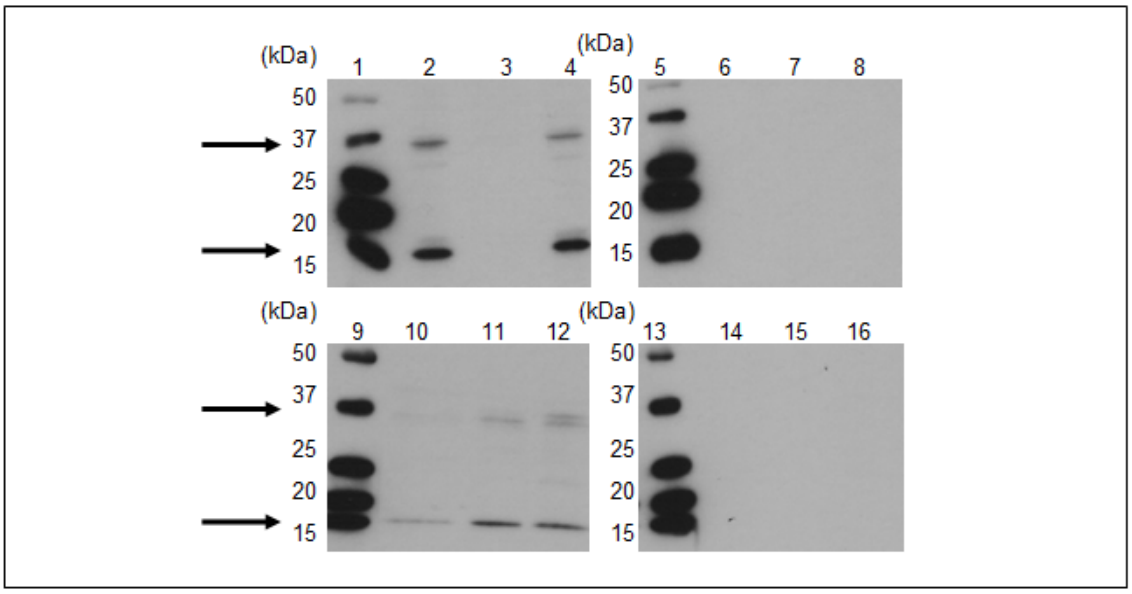


Figure 4.4: Western Blot of breast cancer cell lines and ALC cell lines.
Western Blotting was carried out by probing with anti-BUC11 antibody. The cell lysates used were produced from the human breast cancer cell line MDA231 (positive for BUC11 mRNA expression), the human breast cancer cell lines MDAP3 and MDA468 (negative for BUC11 mRNA expression), the untransfected mouse lymphoma cell line ALC (negative for BUC11 mRNA expression) and the transfected ALC cell line with the recombinant plasmid pBudCE4.1/BUC11 (positive for BUC11 mRNA expression). Lanes with protein marker: (1), (5), (9) and (13). Lanes with primary and secondary antibodies: (2) MDA231, (3) ALC/-, (4) MDAP3, (10) ALC/pBudCE4.1/BUC11, (11) MDA231 and (12) MDA468. Lanes with secondary antibody only: (6) MDA231, (7) ALC/-, (8) MDAP3, (14) ALC/pBudCE4.1/BUC11, (15) MDA231 and (16) MDA468.

The expected molecular weight of BUC11 predicted gene product is in the region of 8.7kDa. Two hypotheses were then proposed for the observed difference in mass in the case of the stronger band: either anti-BUC11 antibody is preferentially binding another protein other than the BUC11 gene product which could be in high abundance in the sample and could be one of those mentioned in Table 4.10, although the peptide block and the blots with ALC/pBudCE4.1/BUC11 lysate contradict this affirmation, or anti-BUC11 antibody binds specifically to BUC11 protein which has post-translational modifications. One has then to question if the same post-translational modifications can occur in cells with or without detectable BUC11 mRNA, including cells engineered to produce the protein (ALC/pBudCE4.1/BUC11). The weaker band, in the region of 30kDa, which is approximately twice the size of the protein(s) detected at 15kDa, could be due to protein dimers even though the conditions for SDS-PAGE were denaturing (boiling in loading buffer containing SDS and DTT). Under native conditions it is possible that BUC11 protein has the ability to build dimers and dimeric aggregates could actually be consolidated following the denaturation process. In order to confirm this and reduce the formation of dimers, the samples could be incubated with the powerful, irreversible reducing agent tris-

(2-carboxyethyl)-phosphine (TCEP) to break disulfide bonds prior to SDS-PAGE (Bond-Breaker TCEP solution, http://www.piercenet.com/products).

To investigate the true identity of the proteins bound to anti-BUC11 antibody on the blots, cell lysates (ALC/- and MDA231) and anti-BUC11 antibody were used to immunoprecipitate the proteins in columns purchased from Miltenyi Biotec (data not shown). The eluted immunoprecipitates, along with non-immunoprecipitated (original) cell lysates, were used in SDS-PAGE. One gel was kept for Western Blotting which showed, as expected, a band at 15 kDa in the lane loaded with the original MDA231 lysate and no band in the lane loaded with the original ALC/- lysate. However, only non-specific, high molecular weight bands, at approximately 50kDa, were visible in the lanes loaded with the immunopurified MDA231 and ALC/- samples. Meanwhile, another gel was stained with Comassie blue and after selection of several bands (at approximately 10kDa and 15kDa), proteins were extracted, digested with trypsin and spotted on a plate for MALDI analysis. Further analysis of the peaks with LTQ mass spectrometer did not identify BUC11 protein which is as expected due to the lack of visible band on the Western Blot that is specific for the size of BUC11 protein. Future work is now required as it is crucial to confirm that the proteins bound by the anti-BUC11 antibody in the Western Blotting is the targeted BUC11 protein and is not another non-specific protein. This could be achieved by 2D gel analysis, or other attempts at optimising the immunoprecipitation protocol, followed by tryptic digestion and mass spectrometry.

4.2.4 Immunofluorescence using anti-BUC11 antibody

Several attempts were made using anti-BUC11 antibody for flow cytometry analysis, however, these did not demonstrate BUC11 staining because the two rabbit IgG isotype controls gave high levels of non-specific background staining. The shift in fluorescence for the control samples was actually greater than the shift found for the test samples (data not shown). This finding suggested that the rabbit IgG isotype controls were not suitable for FACS analysis therefore no further FACS analysis was carried out to study BUC11 protein expression.

In order to determine the localisation pattern of BUC11 protein within cells, immunofluorescence experiments were performed. Briefly, human breast cancer cell lines (positive for BUC11 mRNA) and a melanoma cell line (negative for BUC11 mRNA) were grown in multi-chamber slides, permeabilised, fixed and stained with antibodies (mentioned in Chapter 2) and the DNA dye DRAQ5^{TM}and then observed using confocal microscopy. High levels of BUC11 protein could be detected in MDA231 cells (pictures B, C and F of Figure 4.5) but no protein was detectable in FM3 cells (picture H of Figure 4.5). None or extremely low non-specific background fluorescence was obtained in cells stained with the secondary antibody only (pictures A, D and G of Figure 4.5) and with

the rabbit IgG isotype control (picture E of Figure 4.5). In Figure 4.5 B and C, BUC11 proteins appeared to be localised in the cytoplasm. An accumulation of them appeared to be located in the Golgi apparatus but this would need to be confirmed with future experiments at a higher magnification and by using co-localisation with an antibody directed against a Golgi-specific protein, such as the cellular trans-Golgi network membrane protein, golgin-87, to which the anti-human golgin-97 mouse monoclonal antibody (Sigma) binds. However, when the staining was duplicated, there appeared to be same nuclear staining for BUC11 (Figure 4.5, F). In correlation with the conflicting *in silico* analysis, BUC11 could represent a protein translocated from one cellular compartment to another upon activation. Additional work to confirm this would include subcellular fractionation of cellular components followed by Western Blotting. Because of time constraints, this aspect of the research could not be taken forward.

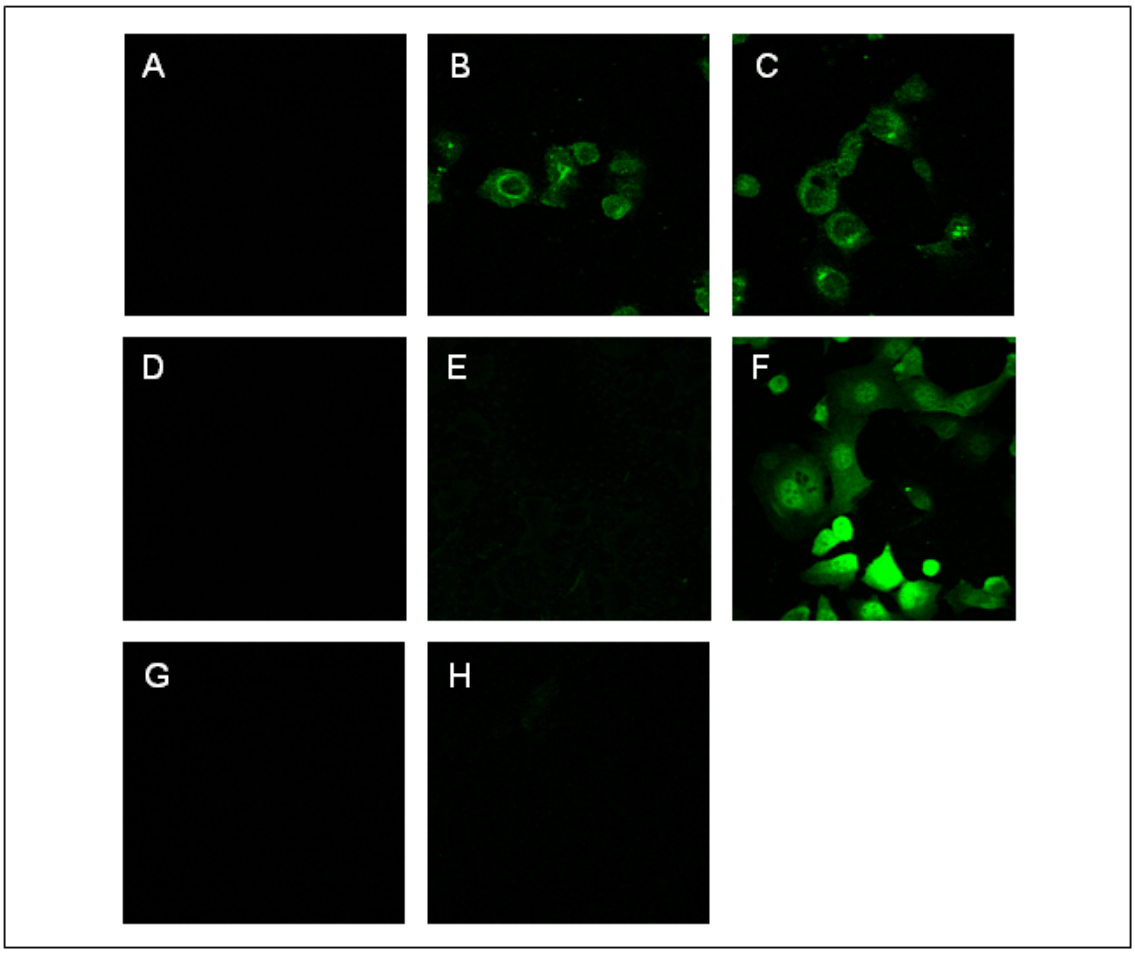


Figure 4.5: Immunofluorescence assay for BUC11 protein expression and localisation analyses in human cell lines.
Immunofluorescence assays were carried out using anti-BUC11 antibody on the human breast cancer cell line MDA231 (positive for BUC11 mRNA) and the human melanoma cell line FM3 (negative for BUC11 mRNA). Immunofluorescence was observed using confocal microscopy in MDA231 cells (B), (C) and (F) but not in FM3 cells (H). Non-specific secondary antibody staining was not observed for MDA231 cells (A and D) and FM3 cells (G). Non-specific rabbit IgG isotype control staining was not observed for MDA231 cells (E). Objective magnification: x40.

Staining of MDA231 cells and T47D cells is shown in Figure 4.6, where BUC11 protein is clearly localised in the cytoplasm and not in the nucleus; as the overlay picture does not show any bright yellow colour in the nucleus, to be expected if BUC11 was also localised in the nucleus (mixing of green colour of BUC11 staining with red colour of DRAQ5^{TM}staining). The same result was found for T47D stained cells, which express BUC11 mRNA (picture A, B and C of Figure 4.6). Without the overlay of staining, Figure 4.6, A could be misleading and suggest that BUC11 is located in the nucleus. Unexpectedly, BUC11 protein expression was found in the cytoplasm of MDA468 cells (negative for BUC11 mRNA molecule under the limits of detection by the quantitative RT-PCR), as shown in Figure 4.6. Also, MDAP3 cells seemed to express high levels of BUC11 protein, whereas BUC11 mRNA levels were low following quantitative RT-PCR.

This would suggest that the translation of mRNA to protein in MDAP3 cells is efficient or that the protein is not easily degraded and remains intact within the cell for an extended period of time. These results however correlate with the results obtained during Western Blotting analysis.

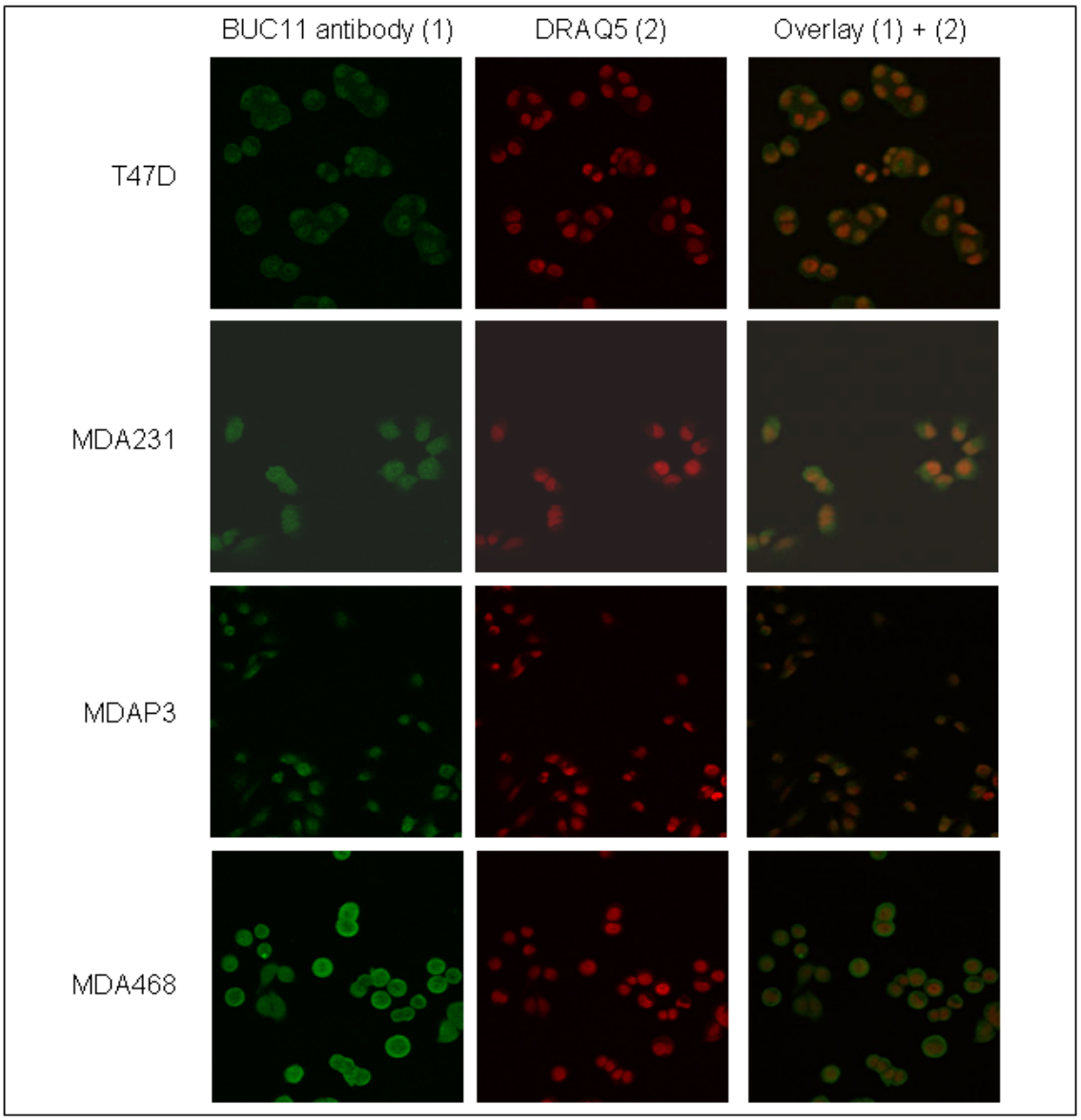


Figure 4.6: Immunofluorescence assay for BUC11 protein expression and localisation analyses in human breast cancer cell lines.
Immunofluorescence assay was carried out using anti-BUC11 antibody and DRAQ5^{TM}DNA dye on the human breast cancer cell lines MDA231, T47D (both positive for BUC11 mRNA), MDAP3 and MDA468 (both negative for BUC11 mRNA). Non-specific rabbit IgG isotype control staining was not observed for any of the cell lines tested (pictures not shown). Objective magnification: x40. 2D-image processing was carried out using the software Volocity version 4.3.2.

4.2.5 Immunohistochemistry using anti-BUC11 antibody

As well as *in vitro* analysis of BUC11 protein expression in cell lines, the anti-BUC11 antibody was used for *in vivo* expression analysis of BUC11 protein in immunohistochemical studies using a diverse array of normal and tumour tissue sections as described in Chapter 2. Paraffin-embedded multiple normal tissue microarrays were purchased from US Biomax (USA). Prior to antibody probing (anti-BUC11 antibody or isotype control) the microarrays had to be subjected to an antigen retrieval process as described in Chapter 2. Each slide was then stained in duplicate; one with anti-BUC11 antibody and the other with isotype control antibody to determine any background staining (Figures 4.7 to 4.10). The BUC11 specific staining was then normalised using Mirax viewer software package to take into account any background staining observed with the isotype control antibody. At objective magnification x10, high levels of BUC11 protein expression were observed in the cytoplasm of cells in testes, prostate, breast, ovary and uterus (Figure 4.7). Therefore, the protein expression in testis, breast, uterus and prostate tissues correlate with mRNA levels detected previously by quantitative RT-PCR (Chapter 3). Some positive BUC11 protein staining was obtained in the brain, heart, skin, spleen and kidney tissues however the isotype control also stained these tissues, to a weaker extent, which suggests a certain degree of non-specific staining in these tissues (Figure 4.8). Previous quantitative RT-PCR results showed low amounts of BUC11 mRNA molecules in brain, heart and kidney but not in spleen. Similarly, positive BUC11 protein staining was observed in the cytoplasm of cells in liver, lung, colon, intestine and stomach tissues (Figure 4.9). A weaker isotype control staining was obtained in liver, lung and colon and a similar isotype control staining was found in intestine and stomach. Some BUC11 mRNA expression was found previously in lung but not in liver, colon, intestine and stomach tissues (quantitative RT-PCR). Therefore, BUC11 protein expression analysis matches BUC11 mRNA expression analysis for lung, intestine and stomach tissues. Figure 4.10 demonstrates that the rabbit IgG isotype control and anti-BUC11 antibody produced the same staining pattern on larynx, tonsil, oesophagus, pancreas and bladder tissue sections, thus one can conclude that BUC11 protein cannot be visually detected in these tissues and that this rabbit isotype control antibody is not suitable for *in vivo* immunohistochemical staining on these particular tissues. The comparison of the protein and the mRNA levels in these tissues has to be taken with caution because both analyses were not carried out on tissues of the same origin.

Immunohistochemical staining on purchased paraffin embedded multiple breast (normal and tumour) tissue microarrays (US Biomax) was undertaken following the same procedure. The stained sections were subjected to scanning, however the slide stained with the isotype control antibody could not be analysed and the pictures were not suitable for further image processing. This was due to a fault in mounting. Figure 4.11 demonstrates high levels of BUC11 protein in the tumour cells of patients with invasive ductal

carcinoma of grade II (A, B, C and D) and patients with invasive ductal carcinoma of grade III (E and F). Positive staining was also observed in normal duct cells (G) although to a lesser extent. These results correlate with the high levels of BUC11 mRNA expression previously observed in breast tissues by quantitative RT-PCR.

Cryostat sections taken from fresh frozen breast tissues (paired normal and tumour, generously donated by Dr A. Gritzapis), melanoma tissues (generously donated by Prof. D. Schadendorff) and prostate tissues (collaboration with Nottingham City Hospital) were prepared on silanised glass slides for immunohistochemical assays as described in Chapter 2. These tissues had previously had RNA extracted from them and had been used to assess BUC11 mRNA levels in quantitative RT-PCR analysis. Unfortunately, the melanoma tissue sections taken were of poor quality and no useful information regarding BUC11 protein expression could be obtained following staining therefore the data is not presented. The tissue sections for breast and prostate are shown. Figure 4.12 shows stronger staining of cells for all 3 patients, thus higher level of BUC11 protein expression was observed in the breast tumour tissues compared to the normal breast tissues. To corroborate these findings, all 3 patients were found to have BUC11 mRNA expression upregulated in the tumour compared to the normal adjacent tissue. Figure 4.13 demonstrates that the level of BUC11 protein expression in the malignant prostate tumour tissue Pr18 is higher than in the malignant prostate tumour tissue 92/41. Positive staining correlated with the detection of BUC11 mRNA in these two tissues and quantitative RT-PCR analysis also showed that BUC11 mRNA level in Pr18 is approximately 3,000 fold higher than the level in 92/41.

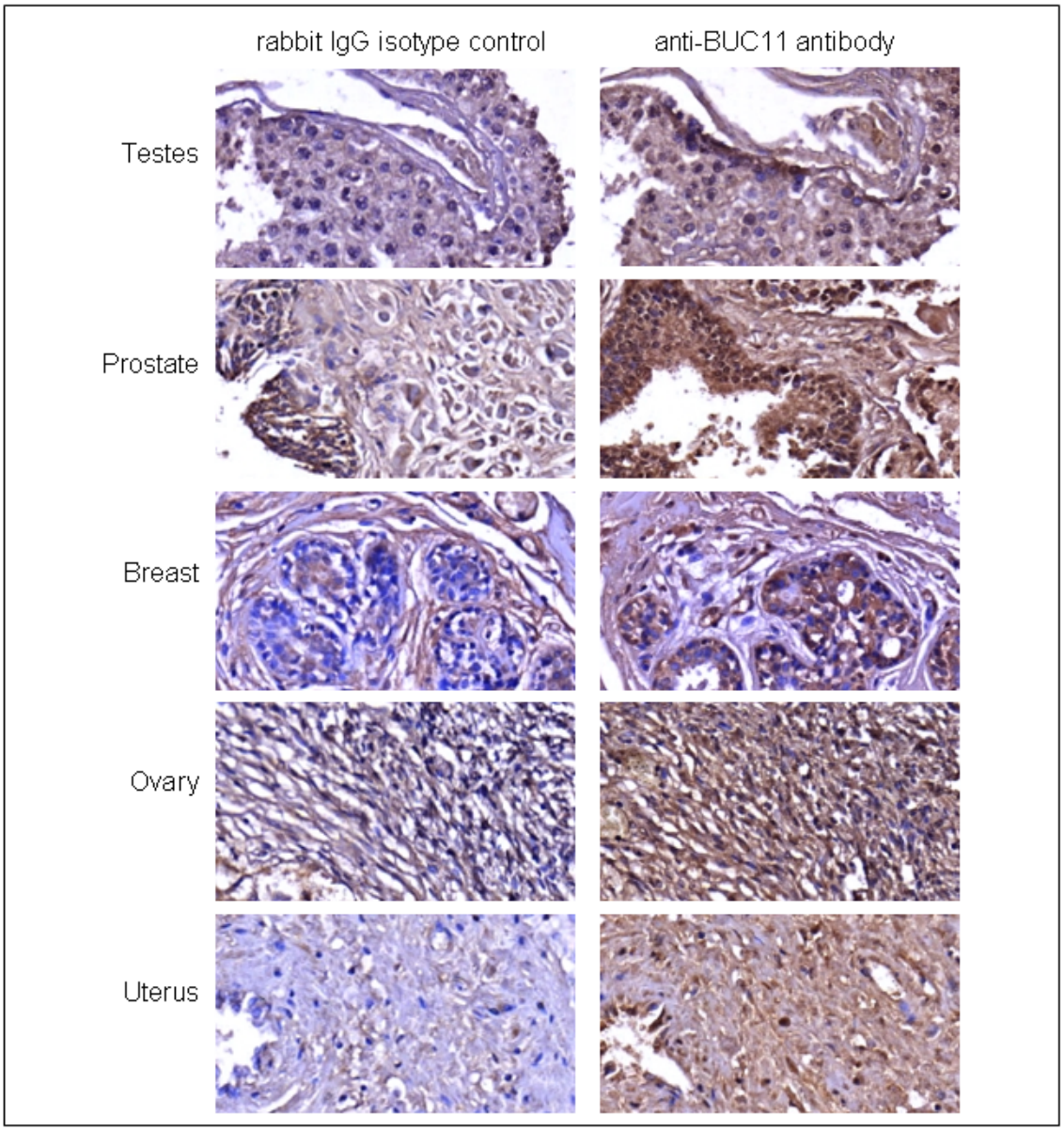


Figure 4.7: Immunohistochemical staining for BUC11 protein expression in multiple normal tissue microarrays (1).
Positive staining obtained with anti-BUC11 antibody was observed in testes, prostate, breast, ovary and uterus, when compared to the staining of the same tissues with rabbit IgG isotype control. Tissue sections were scanned and image processing was carried out using the software Mirax viewer. Objective magnification: x10.

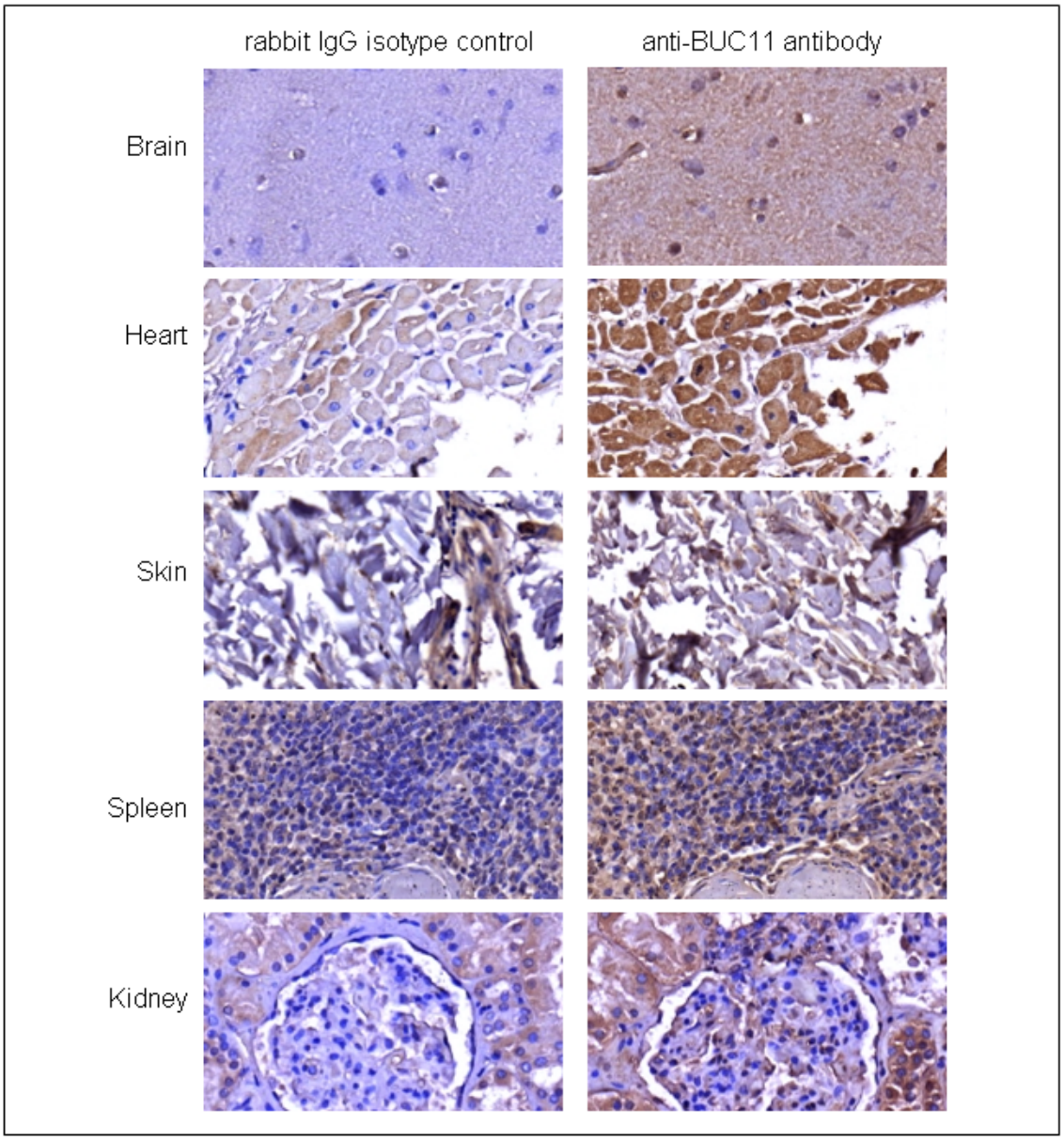


Figure 4.8: Immunohistochemical staining for BUC11 protein expression in multiple normal tissue microarrays (2).
Staining obtained with anti-BUC11 antibody was observed in brain, heart, spleen, skin and kidney. Weaker staining obtained with the isotype control antibody was also observed in these tissues. Tissue sections were scanned and image processing was carried out using the software Mirax viewer. Objective magnification: x10.

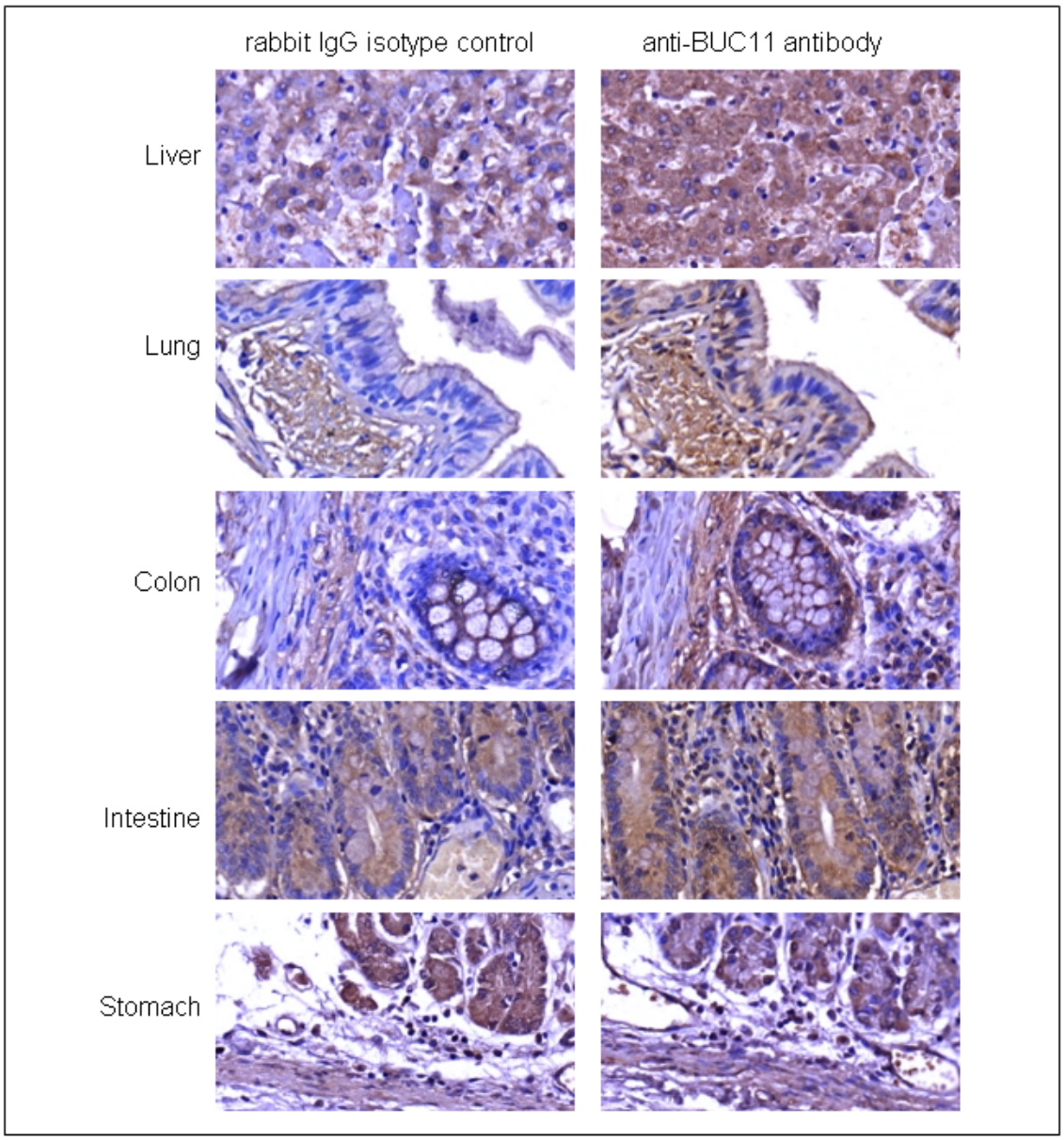


Figure 4.9: Immunohistochemical staining for BUC11 protein expression in multiple normal tissue microarrays (3).
Staining obtained with anti-BUC11 antibody was observed in liver, lung, colon, intestine and stomach. Weaker to similar staining as shown with anti-BUC11 antibody was obtained with the isotype control antibody in these tissues. Tissue sections were scanned and image processing was carried out using the software Mirax viewer. Objective magnification: x10.

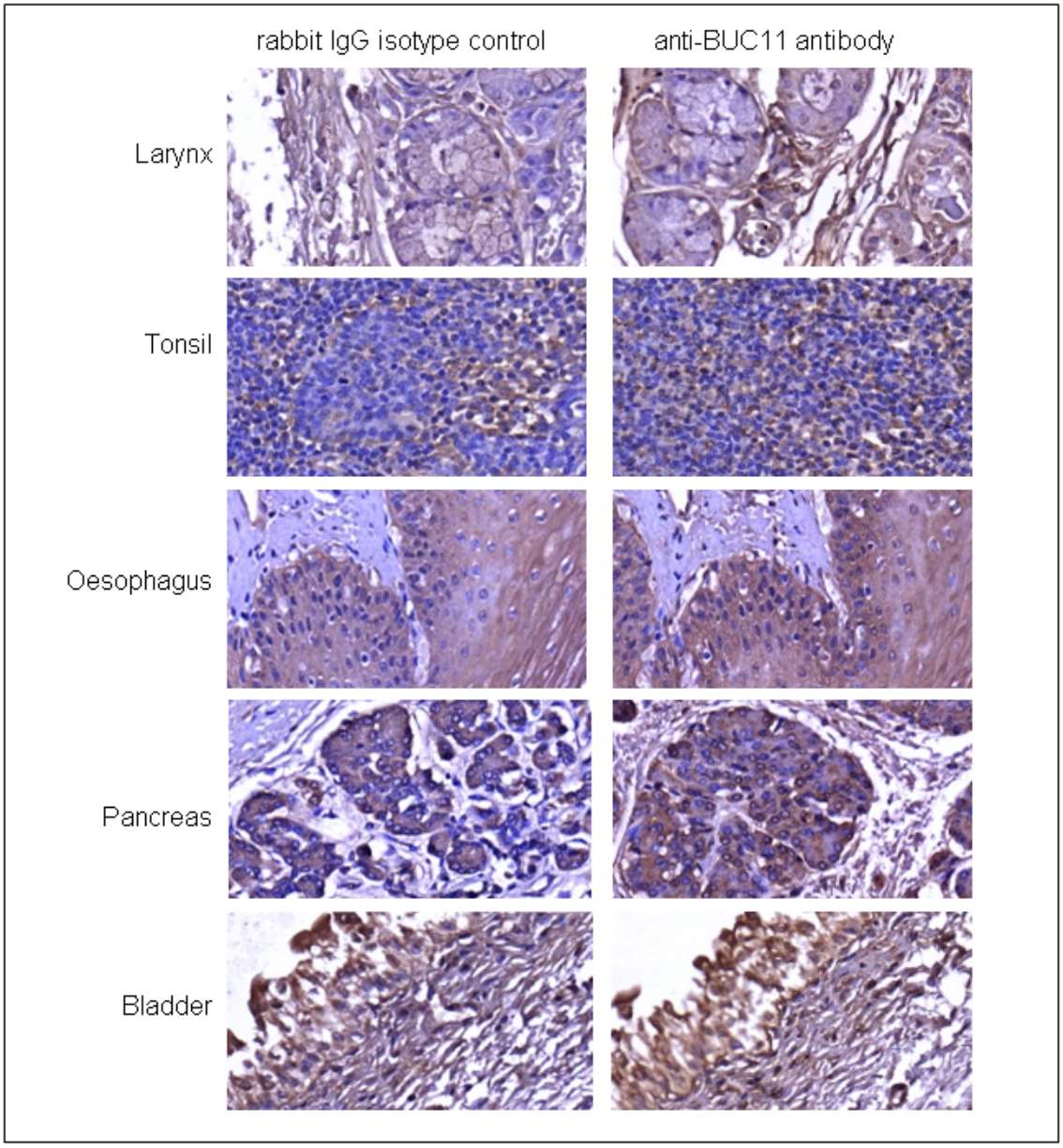


Figure 4.10: Immunohistochemical staining for BUC11 protein expression in multiple normal tissue microarrays (4).
Staining obtained with anti-BUC11 antibody was observed in liver, lung, colon, intestine and stomach. Similar staining obtained with the isotype control antibody was also observed in these tissues. Tissue sections were scanned and image processing was carried out using the software Mirax viewer. Objective magnification: x10.

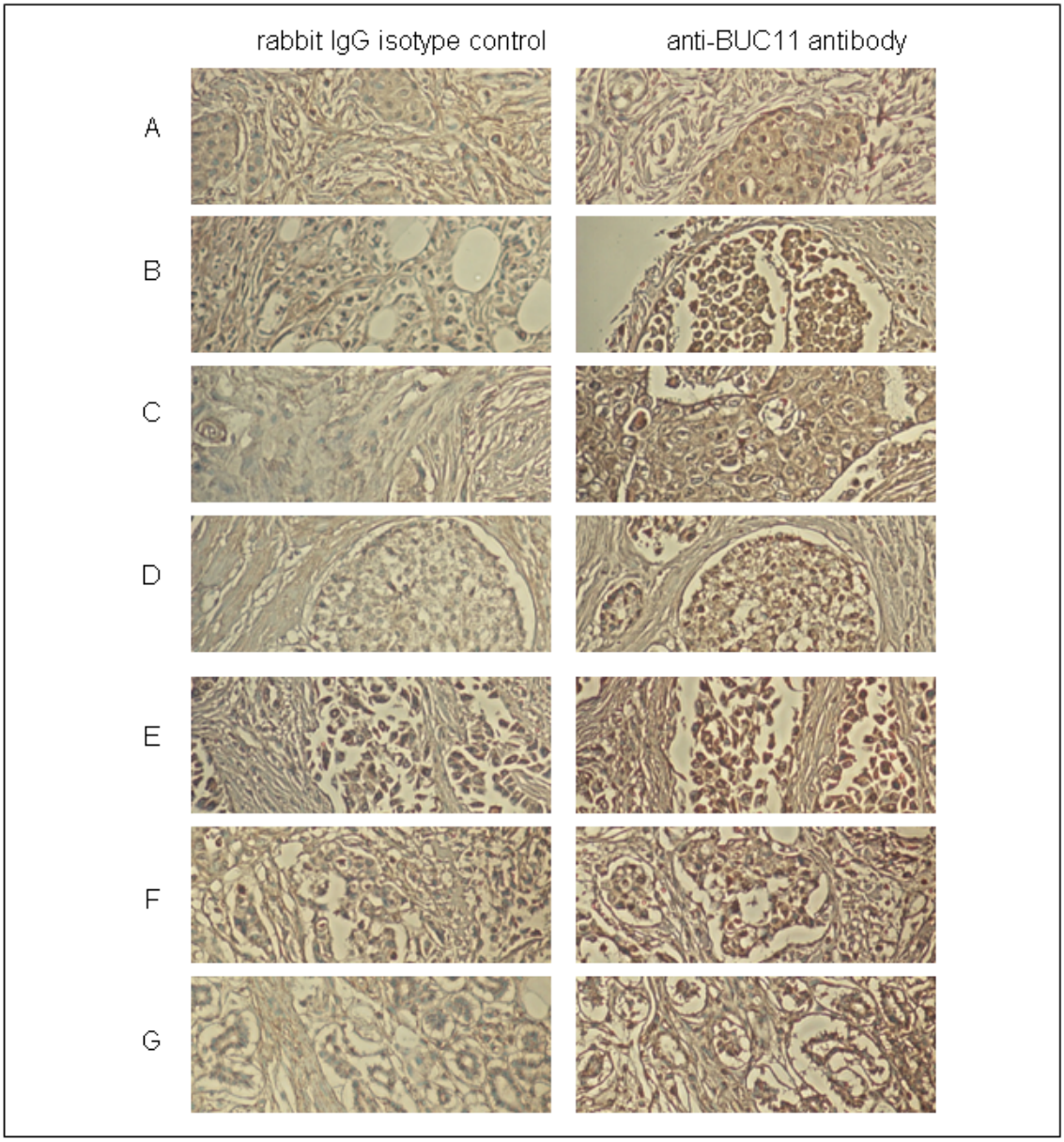


Figure 4.11: Immunohistochemical staining for BUC11 protein expression in multiple breast tissue microarrays.
(A), (B), (C) and (D): breast cancer tissues from four different patients with invasive ductal carcinoma of grade II. (E) and (F): breast cancer tissues from two different patients with invasive ductal carcinoma of grade III. (G): cancer adjacent normal tissue of a different patient. Positive staining obtained with anti-BUC11 antibody was observed in all these tissues when compared to the staining of the same tissues with rabbit IgG isotype control. Pictures of stained tissue sections were taken with a camera attached to the microscope, under the same conditions. Objective magnification: x10.

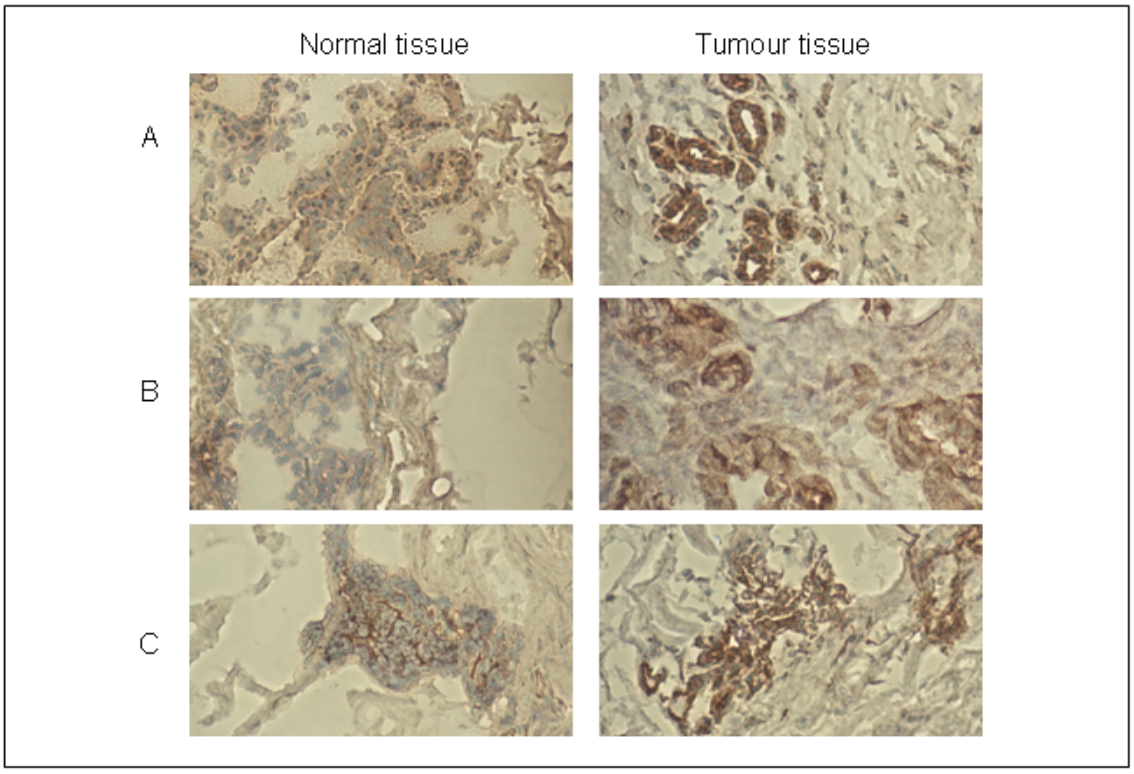


Figure 4.12: Immunohistochemical staining for BUC11 protein expression in breast frozen sections.
Immunohistochemical assay was carried out on breast tumour tissue sections and normal counterpart tissue sections. The mRNA expression analysis of these tissues using quantitative RT-PCR had been carried out. (A) 50-year-old breast cancer patient with tumour diagnosed at stage I, grade II and negative HER2 expression. (B) 41-year-old breast cancer patient with tumour diagnosed at stage IIIa, grade III and negative HER2 expression. (C) 52-year-old breast cancer patient with tumour diagnosed at stage I, grade III and negative HER2 expression. Positive staining obtained with anti-BUC11 antibody was observed in all these tissues, with greater intensity of staining in the tumour tissues compared to the normal tissues. Pictures of stained tissue sections were taken with a camera attached to the microscope, under the same conditions. Objective magnification: x10.

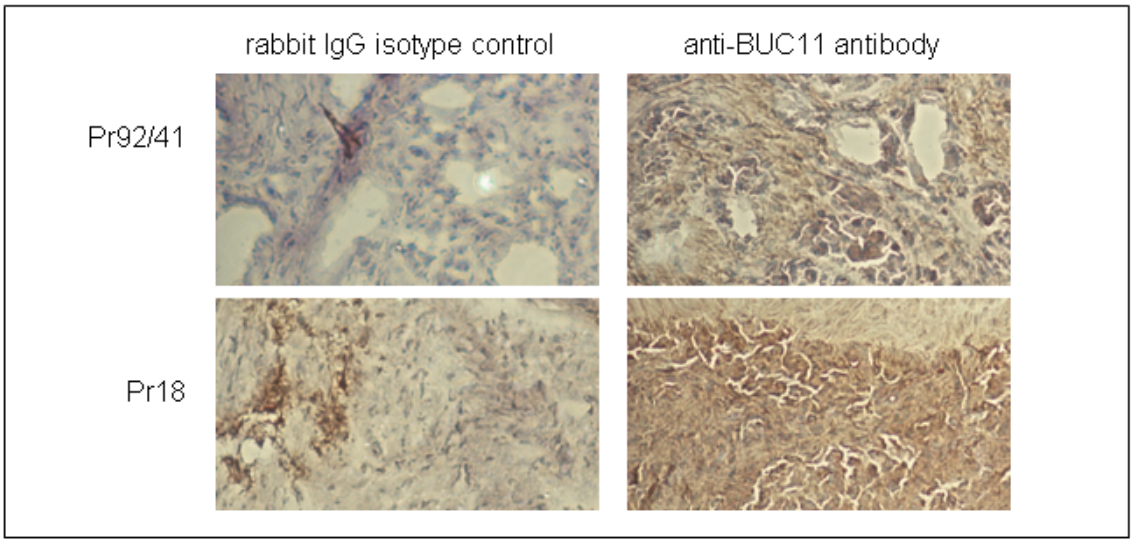


Figure 4.13: Immunohistochemical analysis for BUC11 protein expression in prostate cancer frozen sections.
Immunohistochemical assay was carried out on malignant prostate carcinoma tissues which tested positive for BUC11 mRNA expression using quantitative RT-PCR. Positive staining obtained with anti-BUC11 antibody was observed in the tissues, when compared to the staining of the same tissues with rabbit IgG isotype control. There was also greater intensity of staining in Pr18 compared to Pr92/41. Pictures of stained tissue sections were taken with a camera attached to the microscope, under the same conditions. Objective magnification: x10.

Following encouraging preliminary results of tissue staining with anti-BUC11 antibody, a collaboration with Dr Des Powe (Department of Histopathology, Queen's Medical Centre, Nottingham University Hospitals NHS Trust, Nottingham, UK) was undertaken. Using anti-BUC11 antibody (stock provided at 10μg/ml by NTU), immunohistochemistry (IHC) was performed by Dr Powe on formalin fixed paraffin embedded (FFPE) tissue microarray (TMA) blocks of breast carcinomas. The patient cohort comprised seven TMA FFPE blocks constructed from consecutive primary operable invasive breast carcinomas taken from patients involved in the Nottingham Tenovus Primary Breast Carcinoma Series between 1986 and 1993. This patient cohort has been extensively characterised and contains at least 10 years follow up history with additional clinicopathology and cancer-relevant biomarker data. Briefly, the BUC11 antibody concentration was optimised on full face breast tumour excisional and TMA FFPE sections. Microwave heat antigen retrieval was carried out using 0.01M citrate (pH6) and a labelled streptavidin biotin IHC technique used on a DakoCytomation (Denmark) TechMate autostainer. The optimal working dilution of anti-BUC11 antibody for staining was determined to be 1:900 from the stock provided. For scoring of the TMAs, the levels of cytoplasmic and nuclear staining intensity in malignant epithelium was determined using the H-score where intensity was graded 0: no staining, 1: weak, 2: moderate, 3: strong staining. The association between BUC11 and clinical outcome (survival, disease free interval and tumour recur-

rence) was modelled using Kaplan-Meier plots and Cox regression hazards risk model (SPSS software from SPSS Inc., USA). In addition, association between BUC11 and other key markers of tumour relevance were investigated using a chi square test. A statistical significance level of $p \leq 0.05$ was applied to the data.

In the TMA, BUC11 staining was detected at varying levels in the cytoplasm and nuclei of breast cancer tissues. These findings corroborate previous observations and provide a clearer insight into the confocal results obtained previously; confocal localisation of BUC11 appeared to generate contradicting results in duplicate experiments, localisation appeared to be cytoplasmic in one experiment and nuclear in subsequent experiments. The IHC findings of dual localisation of BUC11 to the cytoplasm and nucleus in some tissue cores suggests that the confocal findings were actually valid. One explanation could be that BUC11 is a nuclear protein that is translocated to the cytoplasm following activation. The TMA staining was scored on 670 breast cancer cases and the distribution of cytoplasmic staining was as follows: 81 patients scored 0 (12.1%), 199 patients scored 1 (29.7%), 250 patients scored 2 (37.3%) and 139 patients scored 3 (20.7%). The distribution of nuclear staining was as follows: 205 patients positive (30.5%) and 467 patients negative (69.5%). The H-score cutoff for low nuclear staining was 0-24 and the cutoff for high nuclear protein expression was 25-300.

The results of the statistical analysis performed are presented in the following five tables (Table 4.11, Table 4.12, Table 4.13, Table 4.14 and Table 4.15). To summarise, no significant association was seen between BUC11 protein expression and, among others (shown in the tables), ER and HER2 status, size of the tumour, grade of the tumour and clinical outcome in this cohort of breast cancer patients. However, BUC11 protein expression significantly correlated with the expression of the androgen receptor (AR), the progesterone receptor (PgR), cytokeratins 18 and 19 (CK18 and CK19), the transferrin receptor CD71 and the transcription factors AP2 and YY1. Interestingly, these results suggest that cytoplasmic BUC11 protein is found in endocrine hormonally positive luminal-type breast cancers.

Significant associations		
Tumour characteristic tested	**p value**	**Correlation**
AR	p=0.007	(+)
PgR	p=0.002	(+)
CK18	p=0.027	(+)
CK19	p=0.002	(+)
CD71	p=0.013	(-)
AP2	p=0.008	(-)
YY1	p=0.004	(-)
Non-significant associations (p>0.05)		
Tumour size	CK5/6	TFF
Tumour grade	CK14	FOXA1
Menopausal status	P53 mutation	Cyclin D1
Mitotic count	FHIT	Basal phenotype
ER	P21, P27	*BCSS
E-cadherin	HER2, 3, 4	*DFI
P-cadherin	Mib1	*DM
BRCA-1 mutation	Bcl2	*Tumour recurrence

Table 4.11: Statistical analysis of cases showing negative BUC11 staining versus any positive BUC11 cytoplasmic protein expression.
*The clinical outcome is defined with breast cancer specific survival (BCSS), disease free interval (DFI), distant metastasis formation (DM) and tumour recurrence.

Significant associations		
Tumour characteristic tested	**p value**	**Correlation**
PgR	p=0.015	(+)
CK19	p=0.002	(+)
AP2	p=0.025	(+)
TFF	p=0.029	(-)
Non-significant associations (p>0.05)		
Tumour size	CK5/6	FOXA1
Tumour grade	CK14	Cyclin D1
Menopausal status	CK18	Basal phenotype
Mitotic count	P53 mutation	*BCSS
AR	FHIT	*DFI
ER	P21, P27	*DM
E-cadherin	HER2, 3, 4	*Tumour recurrence
P-cadherin	Mib1	
BRCA-1 mutation	Bcl2	

Table 4.12: Statistical analysis of cases showing negative BUC11 staining versus weak BUC11 cytoplasmic protein expression.
Weak protein expression corresponds to H-score intensity=1. *The clinical outcome is defined with breast cancer specific survival (BCSS), disease free interval (DFI), distant metastasis formation (DM) and tumour recurrence.

Significant associations		
Tumour characteristic tested	**p value**	**Correlation**
AR	p=0.007	(+)
PgR	p=0.002	(+)
CK18	p=0.027	(+)
CK19	p=0.002	(+)
AP2	p=0.008	(+)
YY1	p=0.004	(+)
CD71	p=0.013	(+)
Non-significant associations (p>0.05)		
Tumour size	CK5/6	TFF
Tumour grade	CK14	FOXA1
Menopausal status	P53 mutation	Cyclin D1
Mitotic count	FHIT	Basal phenotype
ER	P21, P27	*BCSS
E-cadherin	HER2, 3, 4	*DFI
P-cadherin	Mib1	*DM
BRCA-1 mutation	Bcl2	*Tumour recurrence

Table 4.13: Statistical analysis of cases showing negative BUC11 staining versus moderate-strong BUC11 cytoplasmic protein expression.
Moderate-strong protein expression corresponds to H-score intensity=2&3. *The clinical outcome is defined with breast cancer specific survival (BCSS), disease free interval (DFI), distant metastasis formation (DM) and tumour recurrence.

Significant associations		
Tumour characteristic tested	**p value**	**Correlation**
YY1	p=0.04	(+)
Non-significant associations (p>0.05)		
Tumour size	CK5/6	TFF
Tumour grade	CK14	FOXA1
Menopausal status	P53 mutation	Cyclin D1
Mitotic count	FHIT	Basal phenotype
ER	P21, P27	*BCSS
E-cadherin	HER2, 3, 4	*DFI
P-cadherin	Mib1	*DM
BRCA-1 mutation	Bcl2	*Tumour recurrence

Table 4.14: Statistical analysis of cases showing negative BUC11 nuclear staining versus any positive BUC11 nuclear protein expression.
*The clinical outcome is defined with breast cancer specific survival (BCSS), disease free interval (DFI), distant metastasis formation (DM) and tumour recurrence.

Significant associations		
Tumour characteristic tested	**p value**	**Correlation**
PgR	p=0.022	(+)
CK18	p=0.002	(+)
CD71	p=0.027	(+)
AP2	p=0.03	(+)
YY1	p=0.003	(+)
Non-significant associations (p>0.05)		
Tumour size	CK5/6	TFF
Tumour grade	CK14	FOXA1
Menopausal status	P53 mutation	Cyclin D1
Mitotic count	FHIT	Basal phenotype
ER	P21, P27	*BCSS
E-cadherin	HER2, 3, 4	*DFI
P-cadherin	Mib1	*DM
BRCA-1 mutation	Bcl2	*Tumour recurrence

Table 4.15: Statistical analysis of cases showing low BUC11 nuclear staining versus BUC11 nuclear protein expression.
The H-score cutoff for low nuclear staining was 0-24 and the cutoff for high BUC11 nuclear protein expression was 25-300. *The clinical outcome is defined with breast cancer specific survival (BCSS), disease free interval (DFI), distant metastasis formation (DM) and tumour recurrence.

4.3 Discussion

The BUC11 gene, a member of the unpublished BUC family of novel breast-associated genes found by database mining in our laboratory, was previously identified as a potential candidate for breast cancer management based on the findings of an exhaustive analysis of its mRNA expression pattern in diverse normal tissues, tumour tissues and cell lines. All data obtained earlier were limited to the detection of BUC11 at the messenger RNA level, however it has been shown in the literature that divergence between protein and RNA levels is commonly found and can be due to factors, for examples the stability of RNA molecules, the translation and post-translation associated regulatory processes as well as protein degradation (Rogel *et al.*, 1985; Chen *et al.*, 2002). Consequently, investigation as to whether the BUC11 mRNA is translated into protein was essential to establish its possible clinical uses and the results of this research are presented in this chapter. The main focuses of this chapter were, firstly, to mine databases for any information on the BUC11 predicted protein sequence for examples sequence similarities, protein function and structural features; secondly, to produce a custom-made anti-BUC11 antibody for use in *in vitro* and *in vivo* immunoassays to determine whether or not BUC11 protein is actually naturally produced and to elucidate its subcellular localisation and its protein expression pattern in diverse cell lines and tissues. Finally, where possible, to compare the protein expression pattern to the mRNA expression pattern obtained with RT-PCR analyses.

To determine the novelty of BUC11 protein, a thorough study of the predicted protein sequence by database mining was conducted. Blast analyses demonstrated that BUC11 protein sequence has similarities with the sequences of the human and mouse chromosome X-linked protein respectively XIAP and Xiap. XIAP has been proven to have an important role in inhibiting cell death and generated a great interest in the research community in studying it as a potential target for breast cancer therapies. This could be important when deciphering a potential function of BUC11 protein. Other proteins which have sequence similarities with BUC11 included an inhibitor of apoptosis (BIRC8), inhibitors of cell growth (CAPRIN2 and PPP2R1A) and structural proteins (FGD3, PKN2, DNAH2 and PLEKHH3); also, several have been linked to cancer (JMJD3, CHRM3, MAP2K7, FUT4 and MLL2). All the references used for the proteins cited above can be found in the results section of this chapter. Based on the results of the BLAST alignments it can be concluded that BUC11 is not 100% homologous to any previously published genes. This maintains BUC11's attractiveness as a breast cancer marker/target candidate. Additionaly, proteins demonstrating sequence similarities to BUC11 are all involved in some way to cell growth or survival, thus providing additional support for the potential function of the BUC11 protein. *In silico* analyses of the BUC11 predicted protein sequence suggested that the BUC11 protein could localised to the cytoplasm or nucleus. Also, the protein was predicted to have a mass of 8.724kDa, to be basic and quite lysine-rich. The search for structural or functional sites in BUC11 predicted protein sequence that would provide clues on the potential role(s) of the protein did not give any similarity with known functional sites or motifs.

In conclusion, the *in silico* analyses of BUC11 predicted protein did not prove informative. The lack of predicted functional motifs could be due to the small length of the BUC11 protein (8.7kDa); or, more likely, that the BUC11 protein does not encode any common functional motifs. It must be kept in mind that these *in silico* analysis tools should only be used as a guide to the potential functions of a particular protein. More comprehensive assays should always be performed to validate the predictions. For absolute confirmation of BUC11 protein production, a monospecific BUC11 antibody was custom produced. Upon receiving anti-BUC11 antibody, the affinity of the antibody for the peptide to which it has been raised against was confirmed by ELISA. Following Western Blotting of cell lysates with the BUC11 antibody, the blots showed that the antibody binds to proteins with a mass in the range of 15kDa and to a weaker extend to proteins with a mass in a range of 30kDa. These proteins were detected not only in lysates of cell lines which have been previously shown to express BUC11 mRNA (MDA231, MDAP3 and ALC/pBudCE4.1/BUC11) but also in lysates of cells which have been previously shown to not express any detectable BUC11 mRNA levels (MDA468). In order to determine whether or not the 15kDa/30kDa bands observed on the blots were specific to BUC11 protein, a control blot was performed; prior to addition of BUC11 antibody to

the blot, BUC11 antibody was preincubated with BUC11 peptide that the antibody was raised against. In theory if the band observed on the Western Blot is BUC11 specific then it should be abolished by preincubation of the peptide and antibody. This binding was abolished by preincubation informing that the products were BUC11 specific. Absolute proof could be obtained by mass spectrometry analysis.

The difference in size observed between the predicted protein and the proteins detected on the blots could well reflect post-translational modifications of BUC11 protein even if these were not predicted *in silico* with the predicted protein sequence. Post-translational modifications result in mass changes that can be detected during analysis in which individual products of the same gene migrate to different locations on SDS-PAGE or 2D-PAGE gels (Anderson and Anderson, 1998). Post-translational modifications include the addition of functional groups (e.g. biotinylation, glycosylation, phosphorylation), the addition of other peptides or proteins (e.g. ubiquitination) and structural changes of the protein (proteolytic cleavage and disulfide bridges). Thus, the size of the proteins detected (15kDa/30kDa) could be explained by the post-translational formation of dimers due to addition of one BUC11 protein to another BUC11 protein, making the proteins weakly detected in the region of 30kDa (possible quadrimers of BUC11 protein). To test this hypothesis, important denaturation of the linked proteins would be necessary to observe a change of location of the detected proteins on SDS-PAGE gels. It is also possible that the difference in size is due to the addition of one (or several) functional group or the addition of another protein (e.g. ubiquitin is 8.5kDda) and these modifications could be detected with the use of readily available kits such as the ones produced by the company Sigma-Aldrich.

Confocal microscopy, which allows the detection of fluorescently labelled proteins, was carried out to determine the subcellular localisation of BUC11 protein. Upon visualisation, BUC11 protein was detected homogeneously in the cytoplasm of the positive cells and it appears that, in some cases, proteins may also aggregate closely around the nucleus. This finding correlated well with the *in silico* prediction as to cellular localisation of BUC11 protein. BUC11 protein was predicted to translocate to a different cellular compartment upon activation.
BUC11 protein expression levels in T47D and MDA231 cells, as determined by confocal microscopy, correlated with the BUC11 mRNA expression levels observed in the same cell lines by quantitative RT-PCR. However, the high level of expression of BUC11 protein in MDAP3 cells does not correlate with the extremely low level of BUC11 mRNA expression detected by quantitative RT-PCR. As mentioned earlier, there is no direct correlation between mRNA expression and protein expression, emphasising the importance of studying a gene of interest at the mRNA and protein expression levels. Lower levels of mRNA molecules compared to proteins could be explained by an increase in gene

transcription, low protein degradation rate and high protein stability. More surprisingly, but correlating with the Western Blotting results, MDA468 cells showed good levels of BUC11 protein expression whereas no mRNA molecules could be detected previously by quantitative RT-PCR. One possible explanation on the divergence of mRNA and protein expression is that RNA extraction for quantitative RT-PCR assays and immunofluorescence assays or cell lysates production for Western Blotting were not conducted at the same time during the study therefore the number of passages and even batches of the cells used were different and this could potentially have affected the gene copy number. Also, another possible explanation is the limitations of mRNA molecules detected by quantitative RT-PCR. The level of gene expression (mRNA and protein) always vary according to the state (stressed, growing, etc.) and needs related to this particular state during the life of a cell. The study should be repeated with RNA extraction (avoiding possible RNA degradation thus loss of detectable mRNA molecules) and isolation of cells (for lysates and fluorescence) from the same culture in order to provide a true and precise "snapshot" of mRNA and protein expression at a given time in a given state of the cells. Also, the immunofluorescence assay should be repeated by including control chambers with peptide block. Most of the research undertaken on breast cancer is carried out using breast cancer cell lines as *in vitro* models. The majority of these cell lines were established from tumour metastases and cell lines undergo genotypic and phenotypic changes with intensive culture. The accumulation of such changes ultimately means that these cell lines will differ from the original characteristics of the tumours they were established from (Burdall *et al.*, 2003). Good alternatives to overcome these limitations include the use of recently established cell lines from primary tumours or the more challenging primary cultures (Burdall *et al.*, 2003).

Hall and colleagues stated that studies focusing on gene expression in breast cancer cell lines only partially reflect the gene expression in breast tissues (Hall *et al.*, 1986). Immunohistochemistry is an *in vivo* assay that allows both the absolute presence of protein to be determined as well as showing the subcellular localisation of the protein of interest. It also has the advantages of showing which particular cell types in whole heterogeneous tissues express the protein, which is something not possible in tissue lysates unless tissues have been laser micro-dissected beforehand. A normal FFPE tissue microarray was purchased from US Biomax and probed with BUC11 antibody. Varying levels of BUC11 protein were detected in these normal tissues, however one has to be cautious with these findings due to the high level of background staining of the isotype control antibody after optimisation. High levels of BUC11 protein were found in the hormone-dependent tissues: breast, prostate, testes, ovary and uterus. This result correlated with the quantitative RT-PCR data. Some additional vital tissues showed positive staining of BUC11 protein too, but again high background staining with the isotype control antibody was also observed. High levels of background staining were only problematic with the FFPE normal

TMAs. This would suggest that more careful optimisation of the staining protocol on these TMAs is required in the future. Also, the possible future use of monoclonal BUC11 antibody would reduce background staining. It is possible that DAB staining was left longer on the sections stained with the isotype control compared to the sections stained with anti-BUC11 antibody even if the experiment was undertaken with care. Also, positive staining of some tissues with the rabbit IgG isotype control suggests that either the DAB staining should be left for a shorter period of time (reduce background staining to a minimum) or more likely that this particular isotype is not suitable for the FFPE TMAs. If high levels of BUC11 protein expression is confirmed in vital tissues then BUC11 might not be considered for molecular therapies and immunotherapies where there could be serious toxic effects as the normal tissues would also be targeted for destruction. High levels of BUC11 protein expression were found in diverse breast tumour and matched normal tissues, confirming the BUC11 mRNA pattern of expression observed previously by RT-PCR. The positive staining obtained on the small cohort of fresh frozen tissues (breast and prostate) correlated with the quantitative RT-PCR data generated with RNA isolated from the same tissues. Furthermore, staining of the fresh frozen breast cancer tissues showed very specific localisation of BUC11 protein in the glandular or ductal cells and not the surrounding stroma or fatty tissue. Also, the staining observed in the fresh frozen normal breast tissues is not as intense as in the cancer tissue and does not appear to be restricted to glands but seems to be more heterogeneously diffused across the section.

In collaboration with Dr Powe, immunostaining was performed on FFPE TMA blocks comprising 670 invasive breast carcinoma patients using anti-BUC11 antibody. This provided extensive information in the BUC11 protein expression pattern in a large cohort of fully quality assured breast cancer samples. BUC11 protein expression was observed in 87.9% of the cores on the TMA. This result correlates with both the semi-quantitative and quantitative RT-PCR data previsouly obtained on BUC11 mRNA expression in breast tumour tissues in which silencing of gene expression was observed in about 10% of breast cancer cases. BUC11 protein was detected in varying levels in the cytoplasm, with homogenous distribution, which is similar to the staining obtained previously with anti-BUC11 antibody both using our IHC tissues and also using confocal microscopy. Moreover, it was demonstrated that BUC11 protein expression was also detectable in the nuclei of 30.5% of the tissues tested. *In silico* analyses of BUC11 predicted protein sequence showed the possibility of BUC11 protein to be directed to the nucleus which was not confirmed by confocal microscopy nor preliminary immunohistochemical staining in our laboratory. Ubiquitous cellular expression of BUC11 protein might reflect the changing needs of the breast cell at the time of resection of the tumour. BUC11 protein does not appear to be present on the membrane of the cells, making it an unlikely target for antibody-based therapies in breast cancer patients. BUC11 protein expression was not found significantly associated with ER status, HER2 status, tumour size, tumour

grade and clinical outcome. However BUC11 protein was positively correlated to expression of progesterone receptor (PgR), cytokeratin 18 (CK18), the transferrin receptor CD71 and the transcription factors AP2 and YY1. Luminal-type breast cancers are usually classified by showing positive expression of ER and PgR (Loi, 2008); and CK18 has been reported to be a marker for this type of breast cancers (Tang *et al.*, 2006). Furthermore, CD71 might be associated with the growth of endocrine resistant phenotypes within luminal-type breast cancers (Habashy *et al.*, 2009). Finally, AP2 nuclear expression has solely been found in epithelial and myoepithelial of normal and cancerous breast cells (Pellikainen and Kosma, 2007) and YY1 has been associated with AP2 transcriptional activity (Allouche *et al.*, 2008). An interesting finding about BUC11 protein expression and its potential role in breast cells was the fact that cytoplasmic and/or nuclear BUC11 protein was widely found in endocrine hormonally positive epithelial luminal-type breast cancers. Therefore, BUC11 protein could potentially be considered as a novel luminal epithelial marker and further confirmation studies on larger cohorts of patients need to be performed.

Breast tissue is a complex structure formed by an outer myoepithelial/basal layer, an inner luminal epithelial layer and non-epithelial components (fibroblasts, lymphocytes, endothelial cells, adipocytes, neurons and myocytes). For the last ten years or so, distinctive patterns of gene expression, obtained with DNA microarrays, have been widely used in order to determine the most accurate molecular classification of breast tumours into subtypes (Perou *et al.*, 2000; Sorlie *et al.*, 2001; Sorlie *et al.*, 2003). The existence of at least 5 clinically relevant molecular subgroups is known: luminal subtype A (ER+, PR+, HER2-, 45% of breast cancer cases, better prognosis), luminal subtype B (ER+, PR+, HER2-, Ki67$\geq$10%, worse prognosis), HER2+ (ER-, PR-, HER2+, aggresive tumours), basal subtype (ER-, PR-, HER2-, 15% of breast cancer cases, aggressive tumours) and normal breast-like. It is of high importance to discover epithelial cell-type specific proteins in order to increase knowledge on the multistep cancer formation and advancement from normal cells to invasive breast carcinoma as well as to inform clinicians with regard to novel diagnostic and prognostic markers of breast cancer (Jones *et al.*, 2004; Loi, 2008). Furthermore, it is thought that breast cancer may arise from mammary epithelial stem cells (Smith, 2002) and that quiescent cancer stem cells could be responsible for resistance to therapy and recurrence of the disease (Reya *et al.*, 2001). This theory is based on the fact that most anti-cancer therapies such as chemotherapies target rapidly dividing cells thus having no killing effect on resting cells like stem cells. Cells with stem cell-like properties (quiescent, long life, high clonogenicity, self-replicating potential, plasticity and drug resistance) have been found in various human tissues and cell lines; for examples normal breast (Stingl *et al.*, 2005), breast cancer (Al-Hajj *et al.*, 2003), human brain tumours (Singh *et al.*, 2003), prostate cancer (Collins *et al.*, 2005), melanoma (Fang *et al.*, 2005), lung cancer (Ho *et al.*, 2007), pancreatic cancer (Olempska *et al.*, 2007), nasopha-

ryngeal cancer (Wang *et al.*, 2007) and thyroid cancer (Mitsutake *et al.*, 2007). Farnie and Clarke have provided evidence on the existence of breast cancer stem cells; in their studies, they showed the possibility of enriching stem cell like and breast cancer initiating populations through selection by the cell surface markers $CD44^{+}/CD24^{-}$, CD24 being a marker associated with normal luminal epithelial cells (Farnie and Clarke, 2007). It was also shown recently that MCF7 breast cancer cell line has a side population comprising cells that present biological characteristics of quiescent tumour stem cells and which express the transmembrane glycoprotein MUC1, a potential immunotherapeutic target; therefore, these tumour stem cells should also be targetted and destroyed during immunotherapy (Engelmann *et al.*, 2008). As a conclusion, novel breast cancer stem cell-associated markers could represent potential targets for novel therapies to treat and prevent recurrence of breast cancer. It would be interesting and of value to the current study to determine whether or not BUC11 is expressed in breast cancer stem cells.

Chapter 5

BUC11 Functional Analysis

5.1 Introduction

Gene function analysis can be defined as the study of cellular changes due to varying levels of expression of a gene of interest in order to uncover the biological function(s) that a particular gene has in the targeted cells and the pathway(s) the gene might be involved in. The analysis is based on the observation of specific effects on cell morphology, cell proliferation or cell behaviour, generated after the induction of expression in cells that do not normally express the gene of interest or after the repression of expression in cells that normally express the gene of interest. Transfection of expression vectors into cells and RNA interference-mediated gene silencing are extensively used as tools for research in gene function studies and, more importantly, can be used in gene therapies for cancer patients in order to replace defective tumour suppressor genes or inhibit the activity of tumour oncogenes (Toloza *et al.*, 2006; Takeshita and Ochiya, 2006; Huang *et al.*, 2008). Gene delivery, with vectors encoding a tumour antigen that has previously shown to be immunogenic in *in vitro* and *in vivo* studies, can also be used to vaccinate patients, with the objective of triggering a specific and strong immune response against tumour cells expressing the tumour antigen.

Microarray technology is a way of undertaking functional analysis on a genome wide scale and its popularity has increased dramatically over the last 10 years. Expression microarrays have been used to classify breast tumours (Perou *et al.*, 2000; Sorlie *et al.*, 2001; Sorlie *et al.*, 2003) and uncover genes involved in tumour development, progression and recurrence of the disease. The genes identified can then be profiled and used to develop diagnostic and prognostic tests. For example, MammaPrint is a 70-gene signature issued from data analysis of microarrays containing 25,000 60-mer oligonucleotides (Glas *et al.*, 2006) and used by clinicians for predicting outcome of neoadjuvant chemotherapy (Straver *et al.*, 2009). Also, several gene expression profiles discovered by Kok and colleagues can predict the progression of breast cancer in patients who received tamoxifen (Kok *et al.*, 2009). Microarrays are also important for the discovery of targets for

novel breast cancer therapies and the development of targeting drugs (Foekens *et al.*, 2008; Caldwell, 2007). For studies on a particular gene of interest, microarrays allow researchers a snapshot of the entire genome of targeted cells according to the state of cells and whether the gene of interest is up or downregulated with respect to a particular question.

As BUC11 is an unpublished gene, no function for its predicted protein has been characterised. Previously in this study, immunohistochemical analysis showed that BUC11 protein appears to be preferentially found in the epithelial luminal type of breast cells (predominantly in the cytoplasm) however no evidence was found to link BUC11 protein with breast tumour formation, state or progression. In this chapter, the aim was to provide evidence to support the hypothesis that BUC11 might be involved in carcinogenesis and/or tumour progression and to elucidate some of the potential molecular pathways that BUC11 may be involved in. In order to investigate the role(s) that BUC11 plays in the cell, several experiments were conducted following "switching on" BUC11 expression in mammalian cells, that do not normally express this gene and "switching off" BUC11 expression in mammalian cells by short interfering RNA (siRNA)-mediated gene silencing. These experiments included proliferation assays, cell cycle analysis by flow cytometry and microarray technology. Furthermore, conclusions on preliminary experiments (data not shown) on the potential immunogenicity of BUC11 protein in a mouse model are presented in this chapter.

5.2 Results

5.2.1 Cloning of BUC11 DNA into pBudCE4.1

The cloning of the full open reading frame of BUC11 into a mammalian expression vector was necessary to carry out transfection of BUC11 DNA into mammalian cells that do not normally express BUC11 mRNA. The murine lymphoma cell line (ALC), which is stably transfected to express a HLA-A2 construct (containing neomycin resistance gene), was chosen for this purpose and also for the eventual future use in a mouse tumour model (not performed in this study). The mammalian expression vector pBudCE4.1 (Invitrogen) includes the zeocin resistance gene and allows a high level of expression of genes in stable and transient studies in most mammalian cell types.

Briefly, the cDNA of the full open reading frame of BUC11 protein was amplified using specifically designed flanking primers that also included the cutting sites for the restriction enzymes *Sal* I and *Xba* I found in the cytomegalovirus (CMV) multiple cloning site of the pBudCE4.1 plasmid (Figure 5.1). Amplified BUC11 cDNA with flanking restriction enzyme sites and the pBudCE4.1 plasmid were double digested with the specific

restriction enzymes, gel purified and ligated at 4°C overnight. Ligation products were used to transform chemically competent *E.coli* cells which were then plated on LB agar plates containing the antibiotic zeocin. As shown in Figure 5.2, following selection of clones subjected to restriction enzyme double digestion, positive clones that contained BUC11 cDNA (300bp) were obtained. After purification, pBudCE4.1/BUC11 plasmids were sequenced (MWG Biotech, UK) in order to confirm that the BUC11 sequence was correct (Figure 5.3). The pBudCE4.1/BUC11 plasmid was subsequently used in the transfection of ALC cells to determine any effects on cell proliferation and for immunogenicity studies in MHC transgenic mice.

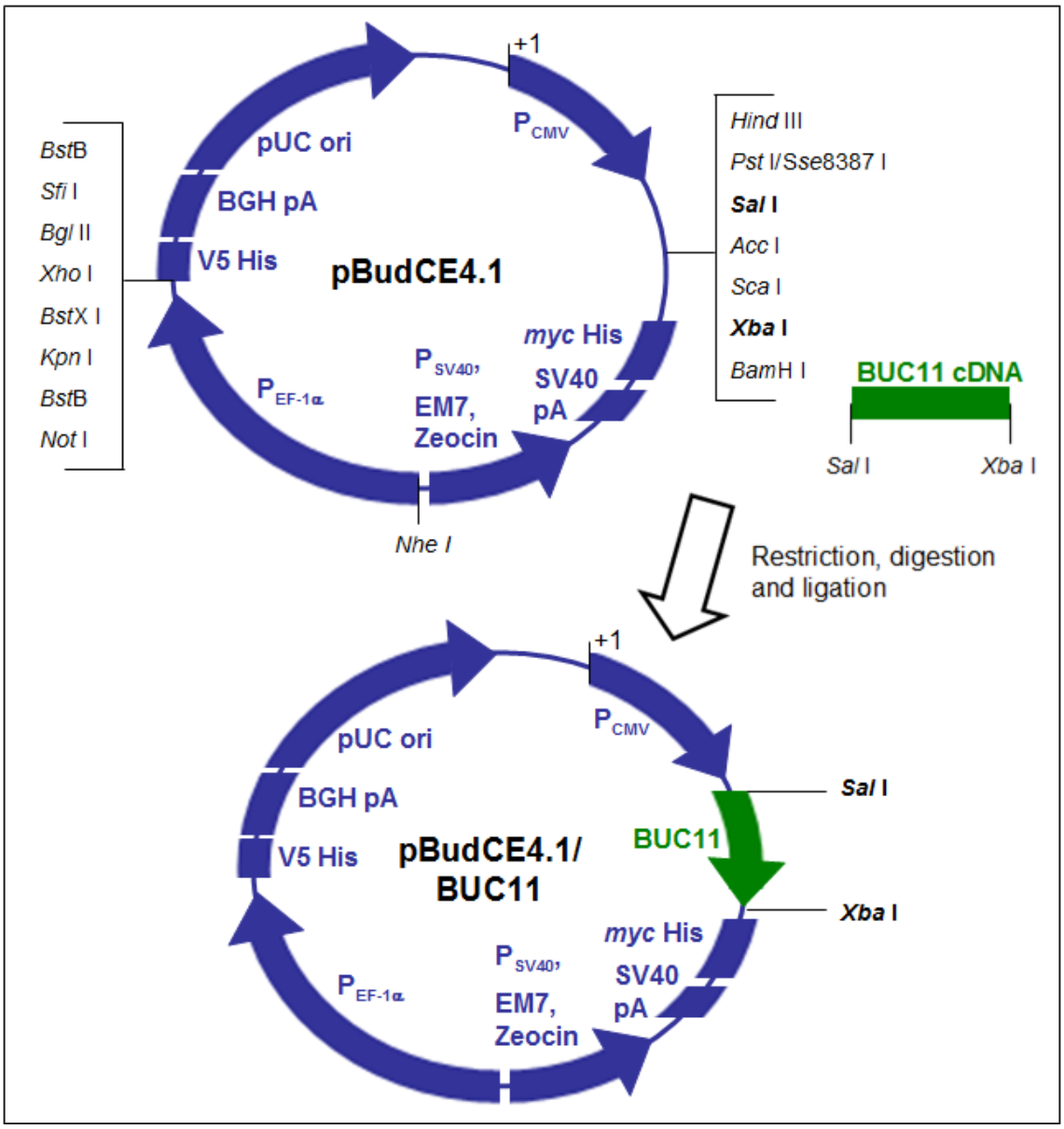


Figure 5.1: Diagrammatic representation of the construction of the plasmid pBudCE4.1/BUC11.
Specific primers were designed to amplify the full open reading frame of the BUC11 gene with the addition of the restriction enzyme cutting sites for *Sal* I and *Xba* I. Both pBudCE4.1 vector and BUC11 cDNA were digested with the restriction enzymes *Sal* I and *Xba* I. Following purification of desired digested products, BUC11 cDNA and pBudCE4.1 were ligated.

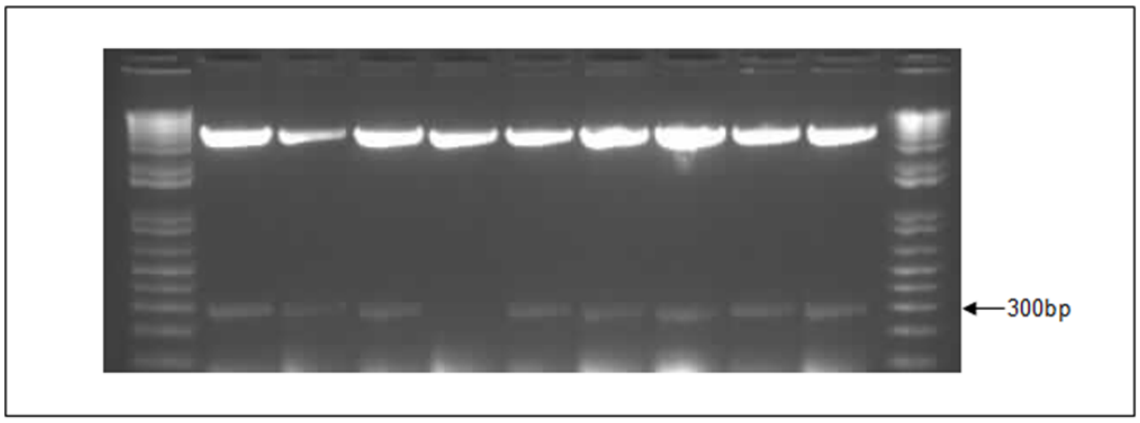


Figure 5.2: Agarose gel electrophoresis of pBudCE4.1/BUC11 clones after double digestion.
Following double digestion of the selected clones with the restriction enzymes *Sal* I and *Xba* I, DNA measuring 300bp could be detected in 8 out of 9 selected clones corresponding to the expected size of cloned BUC11 cDNA.

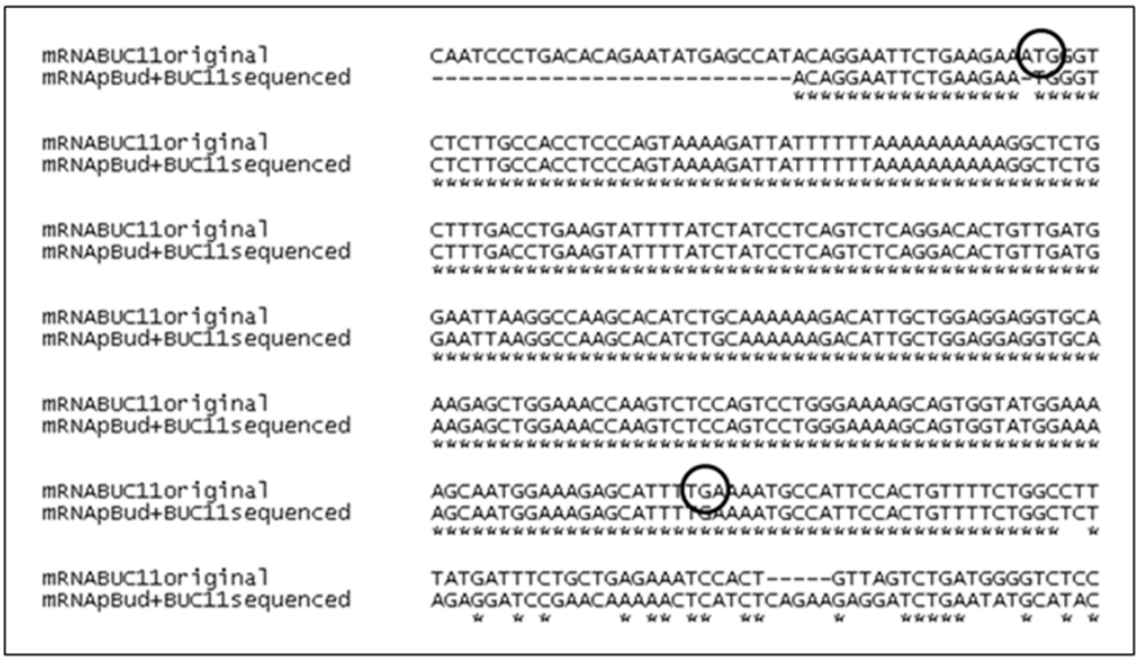


Figure 5.3: Alignment of BUC11 nucleotide sequence with the sequencing result for one of the positive pBudCE4.1/BUC11 clones.
”mRNABUC11original” is the reference nucleotide sequence for BUC11 and ”mRNApBud+BUC11sequenced” is the sequencing result of one of the positive pBudCE4.1/BUC11 clones. The clone sequence presents the correct coding region for BUC11 predicted protein. The start (ATG) and stop (TGA) codons are circled. There is a deletion of an adenine in the clone 6 sequence at the start codon site but the reading frame is not affected.

5.2.2 Transfection of pBudCE4.1/BUC11 in mouse lymphoma cells

Following the successful cloning of BUC11, the pBudCE4.1/BUC11 plasmid was used to transiently transfect, by electroporation, the murine lymphoma cell line (ALC) which does not naturally express BUC11 mRNA. Attempts to create stable transfected ALC cells using the transfection reagent Lipofectamine 2000 were unsuccessful. The MACSelect - Transfected Cell Selection system (Miltenyi Biotec, Germany) enables a quick magnetic enrichment of transfected cells, without the use of antibiotics. In this study, a co-transfection methodology was chosen and carried out using the plasmid of interest (pBudCE4.1/BUC11 or pBudCE4.1/- empty vector as control) and the purchased plasmid pMACS K^k.II. The optimal co-transfection ratio of pMACS and the expression vector of interest and the optimal selection time were determined following recommendation from the manufacturer's protocol. A time course (15 hours, 21 hours and 24 hours) was carried out on cells transfected with 3 different ratios (1:1.5, 1:2.5 and 1:5) of pMACS:pBudCE4.1/- and pMACS:pBudCE4.1/BUC11.

According to quantitative RT-PCR analysis, BUC11 mRNA expression was detectable in ALC/pBudCE4.1/BUC11 cells and not in the control cells 15 hours following transfection. The highest BUC11 mRNA expressions were found in cells transfected with the plasmid ratio 1:2.5. In these cells, the highest BUC11 mRNA expression was obtained 21 hours following transfection and a decreased expression was observed after 24 hours due to the nature of a transient transfection without antibiotic selection (Figure 5.4). Therefore, in subsequent experiments, the selection of co-transfected cells using the MACS Columns was carried out on cells transfected with the plasmid ratio 1:2.5 21 hours following electroporation. In addition to the detection of BUC11 mRNA expression by quantitative RT-PCR analysis, ALC/pBudCE4.1/- and ALC/pBudCE4.1/BUC11 lysates were used for Western blotting analysis to demonstrate BUC11 at the protein level (data shown in Chapter 4).

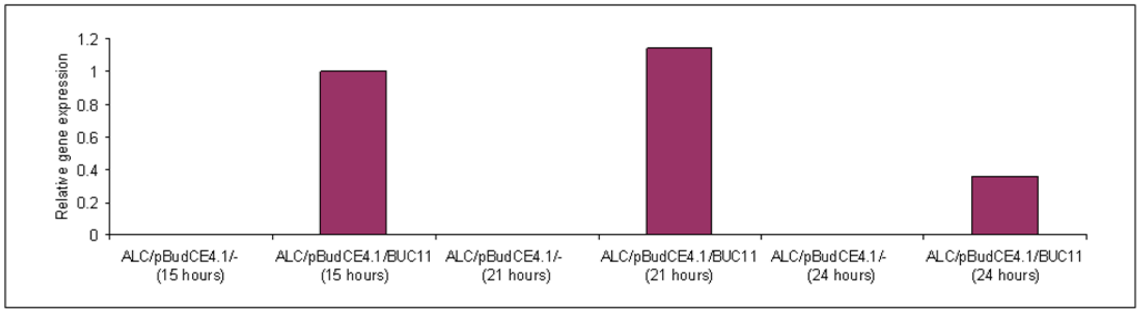


Figure 5.4: Quantitative RT-PCR analysis of BUC11 transfection in ALC cells.
Following co-transfection with 3 different ratios of the plasmids pKKN2 and pBudCE4.1/- (control) or pBudCE4.1/BUC11, RNAs from ALC cells were harvested at three time points (no antibiotic added). The highest expression of BUC11 mRNA was observed with the ratio 1:2.5. The relative BUC11 gene expressions following the transient transfection with the plasmid ratio 1:2.5 of the BUC11-negative mouse lymphoma cell line are shown.

5.2.3 siRNA-mediated BUC11 gene silencing

BUC11 specific siRNA (targeting the coding region of BUC11 gene) was designed for use in BUC11 gene silencing experiments on the BUC11 mRNA-expressing breast cancer cell line MDA231. Attempts were made to transfect BUC11 siRNA in the breast cancer cell line T47D first, which showed a higher BUC11 mRNA expression than MDA231 cells (data shown in chapter 3), but specific gene knockdown never reached more than 70% 24 hours following transfection (data not shown). It was later discovered, when MDA231 cells were used instead, that specific BUC11 gene knockdown above 70% were only obtained 48 hours following transfection. Briefly, MDA231 cells were plated in 24-well plates the day before the transfection so that 60% confluence was reached. Cells were then transfected with either BUC11 siRNA or negative control siRNA (Eurogentec, Belgium) using the siRNA transfection reagent INTERFERin^{TM}. A time course was carried out to confirm the duration of specific BUC11 gene silencing: RNA was extracted 24, 48 and 72 hours following transfection for quantitative RT-PCR analysis. As seen in Figure 5.5, BUC11 siRNA-specific decreased BUC11 mRNA expression was observed 24 hours following transfection, with an optimal silencing (above 70%) obtained after 48 hours. A decrease in gene silencing was found thereafter which reflects the transient nature of siRNA-mediated gene knockdown. It was therefore decided to use the transfected cells for subsequent experiments at the optimal time of 48 hours following transfection. Neither negative control siRNA nor INTERFERin had a significant effect on BUC11 mRNA expression. Moreover, microscopic examination of transfected cells at different time points did not show any obvious changes in cell morphology or cell adherence, however it was evident that cells transfected with BUC11 siRNA proliferated at a slower rate than the control cells (images not shown).

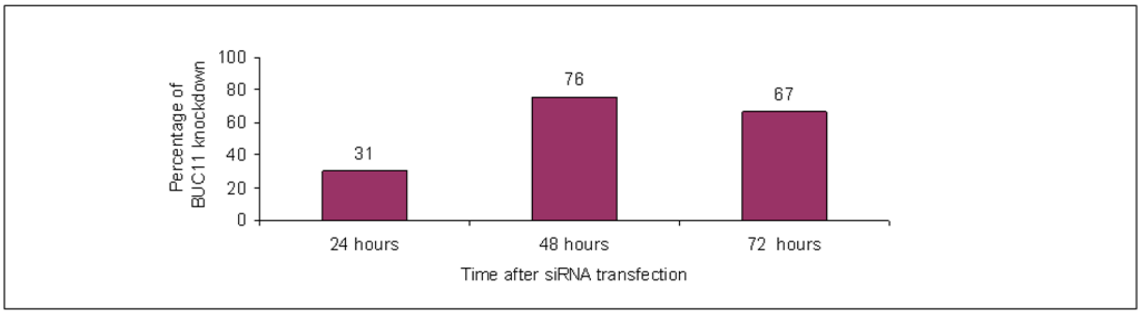


Figure 5.5: Quantitative RT-PCR analysis of BUC11 silencing in MDA231 cells.
A time course was carried out to determine the nature of BUC11 silencing overtime. The percentages of BUC11 expression following the transient transfection with BUC11 specific siRNA of the BUC11-positive MDA231 breast cancer cell line are shown. The highest percentage of BUC11 gene knockdown was observed 48 hours following transfection. Gene silencing was considered successful when more than 70% gene knockdown is observed.

In many cells, the production of interferon proteins is triggered in response to the presence of double-stranded RNA (dsRNA) suggesting viral infection. It was previously reported that *in vitro* transfection of cells with dsRNA such as the siRNA used in this study can be responsible for an interferon-based response, altering not only the expression of interferon-responsive genes, but also the expression of the targeted gene of interest. Quantitative RT-PCR analysis showing a reduction in the expression of the gene of interest would therefore be misleading in the interpretation of non specific effects. For instance, expression of both interferon-responsive genes 2'-5'-oligoadenylate synthase 1 (OAS1) and signal transducer and activator of transcription 1 (STAT1) were found to be affected in mammalian cells transfected with siRNA (Fish and Kruithof, 2004; Ma *et al.*, 2005). In this study, OAS1 and STAT1 primers were used in quantitative RT-PCR analysis. No significant variations in the mRNA expression of these genes were observed, suggesting that decreased expression of BUC11 mRNA obtained in this study was indeed the result of specific BUC11 silencing by siRNA and not as a result of so called "off gene" silencing (data not shown).

5.2.4 Proliferation assays

Proliferation assays were performed to determine the potential functional effects of BUC11 in cell lines. Firstly, the BUC11-positive cell line MDA231, which was successfully used previously in siRNA-based gene knockdown studies, was again used in siRNA-based gene knockdown studies and cell proliferation was measured. Secondly, a naturally BUC11-negative cell line, ALC mouse lymphoma cell line, which has previously been successfully transfected with the expression vector pBudCE4.1/BUC11, was also used in the proliferation study. The proliferation assay was based on the incubation of cells with ^{3}H-thymidine. During each cell division, the cells incorporate radioactive thymidine

into their DNA, thus the more proliferative the cells are the higher is the radioactivity of the DNA, which can then be collected on a filter membrane and counted in a scintillation counter (counts per minute). In some instances, DNA or siRNA transfection effects on mRNA expression can be measured as early as 6 hours following transfection, however, effects on protein production and function are not generally observed earlier than 48 hours following transfection. For this reason, proliferation assays were performed 48 hours or longer following transfection, usually at 2, 4 and 8 days after the transfection of ALC/pBudCE4.1/BUC11 cells and ALC/pBudCE4.1/- cells (control). Figure 5.6 shows a statistically significant increase in proliferation of the cells transfected with BUC11 gene when compared with the cells transfected with the empty vector, both at 2 and 4 days. On day 8, there was still a statistically significant difference, however this observation requires caution because at this time point the cells had become confluent and the rate of overall proliferation was reduced for both transfections (this can be observed visually and by the drop in the proliferation rate as compared to days 2 and 4).

The effect of BUC11 gene silencing on the proliferation of MDA231 cells, test and control cells were labelled with ^{3}H-thymidine, 3 days, 7 days and 10 days following siRNA transfection. Figure 5.7 shows a statistically significant decrease of proliferation of the cells transfected with BUC11 siRNA when compared with the cells transfected with negative control siRNA, both 3 days and 7 days after transfection. However, the proliferation rate of the cells transfected with BUC11 siRNA had returned to a rate that was very similar to the control cells, due mainly to the transient effect of siRNA transfection.

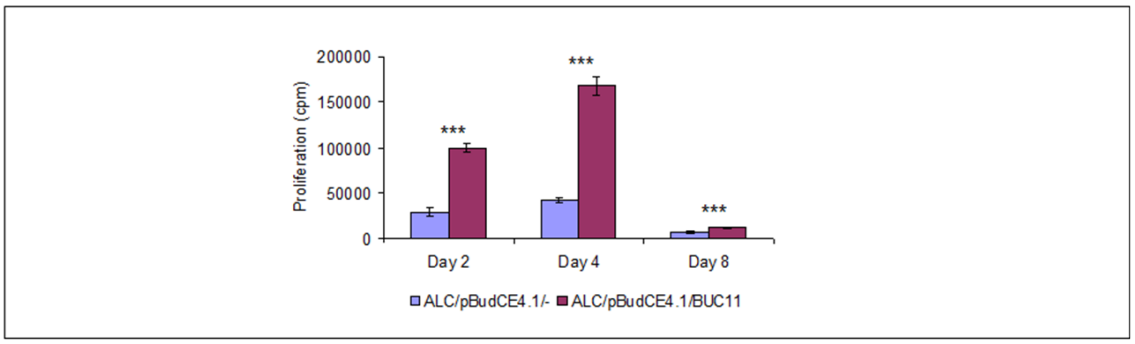


Figure 5.6: Proliferation assay following BUC11 transfection.
A time course was carried out to determine the rate of cell proliferation following transient transfection of BUC11 mRNA-negative mouse lymphoma (ALC) cells with pBudCE4.1/- and pBudCE4.1/BUC11. Zeocin was added to the media 48 hours following transfection. Four replicates were used for each condition and each time point in this experiment. A significant increase in proliferation of the BUC11-negative ALC cells was obtained 2 days after transfection with the vector expressing BUC11. *** $p<0.001$ represents the statistical difference between ALC cells transfected with pBudCE4.1 empty vector and those transfected with pBudCE4.1/BUC11 vector, determined by Student's t-test.

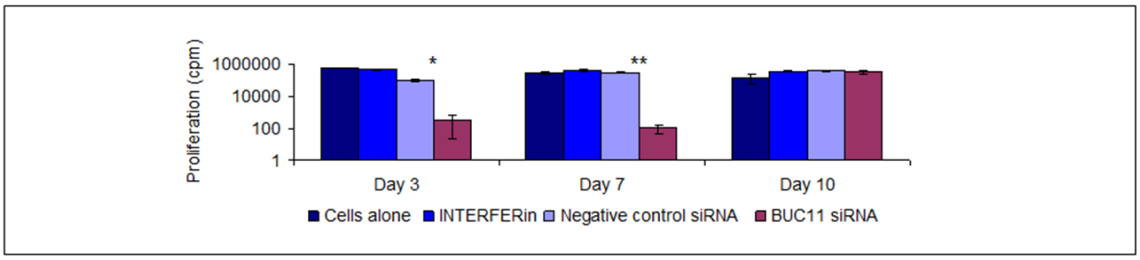


Figure 5.7: Proliferation assay following BUC11 silencing.
A time course was carried out to determine the rate of cell proliferation following transient transfection of BUC11 mRNA-positive MDA231 breast cancer cell line with BUC11 siRNA or negative control siRNA. Other controls included cells treated only with INTERFERin and non-treated cells. Duplicates were used for each condition and each time point in this experiment. A significant decrease in proliferation of the BUC11-positive MDA231 cells was obtained 3 days after transfection with BUC11 siRNA. * $p<0.05$ and ** $p<0.01$ represent the statistical difference between MDA231 cells transfected with negative control siRNA and those transfected with BUC11 siRNA, determined by Student's t-test.

5.2.5 Cell cycle analysis

The proliferation analysis suggested that BUC11 might have a role in a pathway involved in the control of cell division. Uncontrolled cellular proliferation can result from disorganisation of the cell cycle, therefore cell cycle analysis by propidium iodide (PI) staining with flow cytometry (FACS) was carried out on the MDA231 breast cancer cell line. Briefly, 72 hours following transfection with either control siRNA or BUC11 siRNA, adherent cells were trypsinised, washed several times and fixed with cold ethanol. The cells were washed, incubated with a mix of RNase and PI solutions and DNA fluorescence was measured by FACS. Table 5.1 shows the percentage of cells in each of the 3 phases of the cell cycle (G1, G2 and S) for both test and control cells. From these results it appears that BUC11 gene silencing was responsible for an increase of cells in the G1 phase of the cell cycle with a comparable drop of the number of cells in the G2 and S phases. Therefore, knockdown of BUC11 gene causes growth arrest in these cells, most likely due to cell cycle inhibition at the G0-G1 checkpoint or G1/S checkpoint, potentially due to BUC11 having an influence on the activity of G1 regulatory proteins. This result correlates with the decrease of BUC11 gene silenced-MDA231 cell proliferation observed and provides further evidence for a potential role of BUC11 in the regulation of cell growth.

Cells	Percentage of cells in:		
	G1	G2	S
Control	23.3	30.9	45.8
BUC11 knockdown	50.7	13.5	35.8

Table 5.1: Cell cycle analysis by flow cytometry of BUC11 siRNA-transfected MDA231 cells.
"Control" cells are MDA231 cells transfected with negative control siRNA and "BUC11 knockdown" cells are MDA231 cells transfected with BUC11 siRNA. Cells were stained with propidium iodide 72 hours following transfection and measurement of DNA fluorescence was carried out by FACS. This experiment was carried out twice (n=2) with only one replicate in each, the percentages are given as mean of both experiments. Analysis of fluorescence using the software Cylchred gave the percentage of cells in each of the 3 phases of the cell cycle (G1, G2 and S) for both test and control cells.

5.2.6 Affymetrix GeneChip®Microarray

As BUC11 is a novel unpublished gene, no information is available on the pathways or cellular processes it is involved in. Following the observations made in the siRNA-mediated BUC11 silencing experiments, analysis of the effects of BUC11 gene knockdown on a wider scale, that is to say on the whole genome of MDA231 cells, was carried out using Affymetrix GeneChip® Microarray technology. Briefly, in the same experiment, RNA was extracted from duplicates of control (control siRNA) and test (BUC11 siRNA) transfected MDA231 cells. Significant specific BUC11 gene silencing in the test samples was confirmed by quantitative RT-PCR and the purity of the four RNA samples was confirmed using a Nanodrop microspectrophotometer. The duplicate control and test RNA samples were analysed at the Nottingham Arabidopsis Stock Centre (School of Biosciences, University of Nottingham, Sutton Bonington Campus, Loughborough) on four different GeneChip®Human Genome U133 Plus 2.0 Arrays, which include more than 47,000 RNA transcripts (Affymetrix, USA).

Upon receiving the data, the gene expression profile data was plotted control versus silenced. Regression analysis was conducted and the standard residual value determined for each transcript spot. This was repeated for all replicates. The average and standard deviation of standard residual values for each gene was calculated. Mean standard residuals were ranked most negative to most positive. These values varied between -31.07 to 24.32 and the standard deviation values were within acceptable limits. Any genes that had an average value above 2 were considered significantly upregulated in the test samples compared to the control samples whereas any genes that had an average value below -2 were considered significantly downregulated in the test samples compared to the control samples. On the array, the U133 Plus 2.0 Probe set 237339_at was created from the genes associated to the Genbank accession AI668620 which includes the gene

hCG_25653 (BUC11)(http://www.genecards.org/cgi-bin/carddisp.pl?gene=hCG_25653). The average value for the transcript representing BUC11 was -0.12 therefore the BUC11 gene is not significantly downregulated in this analysis.

The 10 most significantly upregulated genes and the 10 most significantly downregulated genes found in the microarray analysis are shown in Table 5.2 and Table 5.3 respectively. The top 100 most upregulated genes (their average values range from 24 to 6.9) and the top 100 most downregulated genes (their average values range from -31 to -7) were studied by searching relevant literature. It was possible to establish links between genes of the study based on information reported in the literature that they have a direct effect on each other or that they have an indirect effect through a cascade of events in the cell. Genes could also be grouped according to their biological function(s). The list of the most upregulated genes is mainly composed of ribosomal proteins, heat shock proteins, respiratory chain proteins, proteins involved in apoptotic pathways and proteins involved in cell growth and protection of cell against hydrogen peroxide. The list of the most downregulated genes is mainly composed of ribosomal proteins, metallothioneins, angiogenic inducers and proteins involved in the cytoskeletal organisation, cell adhesion, cell migration and cell proliferation.

Probe set	Average	Stdev	Descriptions
217466_x_at	24.32	6.41	RPS2 - Ribosomal protein S2
200680_x_at	17.48	0.59	HMG1 - High-mobility group protein 1
225190_x_at	16.51	2.38	DBNDD2 - Dysbindin domain containing 2
223034_s_at	15.21	1.39	C1orf43 - Chromosome 1 open reading frame 43
201195_s_at	14.56	2.46	LAT1 - L-type amino acid transporter 1
201463_s_at	14.37	0.91	TALDO1 - Transaldolase 1
226091_s_at	14.00	1.25	MRPL11 - Mitochondrial ribosomal protein L11
221691_x_at	13.96	4.93	NPM1 - Nucleophosmin
200679_x_at	13.46	0.67	HMG1 - High-mobility group protein 1
223020_at	13.42	0.68	CRR9 - CLPTM1-like

Table 5.2: Top 10 most upregulated genes in the microarray analysis.
The 10 most upregulated genes in the BUC11 silenced-MDA231 breast cancer cells when compared to the control cells found in the regression analysis of the microarray data are shown. CLPTM1 stands for cleft lip and palate associated transmembrane protein 1. “Average” is the average of standard residual values.

Probe set	Average	Stdev	Descriptions
212185_x_at	-31.07	1.33	MT2A - Metallothionein 2A
200062_s_at	-30.83	3.38	RPL30 - Ribosomal protein L30
200650_s_at	-29.99	2.03	LDHA - Lactate Dehydrogenase A
209101_at	-28.24	2.34	CCN6 - Connective tissue growth factor
216241_s_at	-26.94	2.56	TFIIS - mRNA for transcription elongation factor
201761_at	-26.49	2.55	MTHFD2 - Methylenetetrahydrofolate
202686_s_at	-25.54	2.50	AXL - Receptor tyrosine kinase
208636_at	-24.92	1.34	ACTN1 - Actinin alpha 1
208581_x_at	-24.89	0.55	MT1X - Metallothionein 1X
213350_at	-23.03	1.77	RPS11 - Ribosomal protein S11

Table 5.3: Top 10 most downregulated genes in the microarray analysis.
The 10 most downregulated genes in the BUC11 silenced-MDA231 breast cancer cells when compared to the control cells found in the regression analysis of the microarray data are shown. “Average” is the average of standard residual values.

The literature search of the top 100 significantly upregulated genes in MDA231 breast cancer cells when BUC11 gene expression is knocked-down revealed a number of interesting genes. As stated earlier, several ribosomal proteins were found. These include ribosomal protein S2 (RPS2), ribosomal protein L15 (RPL15), ribosomal protein S19 (RPS19), ribosomal protein L13a (RPL13a), ribosomal protein L7a (RPL7a), ribosomal protein S7 (RPS7) and mitochondrial ribosomal protein L11 (MRPL11). Several studies have shown that the expression of some ribosomal proteins can be linked with the development of prostate and colon cancer (Ohkia *et al.*, 2004; Vaarala *et al.*, 1998; Bee *et al.*, 2006; Pogue-Geile *et al.*, 1991). Wang and colleagues showed that over-expression of RPS2 might be involved in the development of prostate cancer and that RPL15 promotes the proliferation of gastric cancer cells, making these proteins a good target for novel cancer therapies (Wang *et al.*, 2009; Wang *et al.*, 2006). Loss of RPS19 expression was associated with a decreased cell proliferation and an increase in apoptosis in a cell line model of Diamond-Blackfan anemia (Miyake *et al.*, 2008). On the other hand, RPL13a and RPL7 expressions were associated with the induction of apoptosis by arresting cell growth respectively in the G2/M phase and G1 phase of the cell cycle (Chen and Ioannou, 1999). Also, increased expression of RPS7 and of MRPL11 were shown to induce apoptosis and inhibit proliferation of cells (Chen *et al.*, 2007; Dai *et al.*, 2007). Two heat shock (stress) proteins were also found in the most upregulated genes: Heat Shock 90kd protein 1 and Hsp89 alpha Delta-N and apoptosis can induce the accumulation of stress proteins in cells. HSP90 is a good target for cancer therapies as it was shown to maintain, under stress conditions, the functions of proteins associated with survival and proliferation of cancer cells (Messaoudi *et al.*, 2008). Several respiratory chain proteins were found in the top 100 upregulated genes such as voltage-dependent anion channel 3 (VDAC3), mitochondrial ubiquinone binding protein, uniquinol-cytochrome c reductase binding protein (UQCRB), NADH dehydrogenase (ubiquinone) 1 alpha subcomplex 1 (NDUFA1), NDUFS6, NDUFS5, NDUFB4, cytochrome c oxidases (COX7c, COX4, COX7B), ATP

synthases (ATP5E and ATP5B). Mitochondria are essential cellular organelles for the production of energy but also have functions in apoptosis. For example, VDAC3 was shown to play an important role in mitochondria-mediated apoptosis (Shoshan-Barmatz *et al.*, 2006). Other interesting proteins found in the most upregulated genes that did not belong to major protein groups include the mitochondrial carrier homolog 2 protein (MTCH2), the glutathione peroxidase 1 protein (GPX1) and the cleft lip and palate associated transmembrane protein 1-like protein (CRR9). MTCH2 was found expressed in apoptotic cells (Grinberg *et al.*, 2005), CRR9 was associated with apoptosis in an ovarian cancer cell line (Yamamoto *et al.*, 2001) and overexpression of GPX1 was associated with delayed cell growth and detoxification of hydrogen peroxide from cells (Faucher *et al.*, 2003).

The literature search of the top 100 significantly downregulated genes in MDA231 breast cancer cells when BUC11 gene expression is knockdown also uncovered many interesting genes. The ribosomal proteins L30 (RPL30) and S11 (RPS11) were found to be amongst the 10 most downregulated genes. RPS11 was found downregulated in apoptotic breast cancer cell lines such as MCF-7 (Nadano *et al.*, 2001) and RPL30 was associated with tumourigenesis in colon cancer (Jiang *et al.*, 2008). Moreover, the following metallothioneins were found: MT2A, MT1F, MT-1E, MT-1X, MT-1H, MT1L and MT1G. The family of metallothioneins is composed of at least 10 functional genes that encode low molecular weight metal binding proteins that have been widely linked with proliferation of breast cancer cells and have been associated with poorer prognosis in breast cancer (Bay *et al.*, 2006). Their functions include the maintenance of transition metal ion homeostasis, redox balance, cell proliferation and apoptosis (Jin *et al.*, 2004). In particular, MT2A has been associated with higher histological grade of breast cancer (Jin *et al.*, 2002). In this study, MT2A was found to be the most significantly downregulated gene. Yang and colleagues stated that metallothionein genes are good targets for antisense therapies in breast cancer (Yang *et al.*, 2002). As previously mentioned, the top 100 most downregulated genes included proteins involved in the cytoskeletal organisation, cell adhesion, cell migration and cell proliferation such as actinin (alpha) 1 (ACTN1), syndecan 4 (SDC4), syndecan binding protein, clathrin (LCA, actin assembly role), actin gamma 1, calponin 3 (CNN3, cytoskeleton-associated), tubulin beta 2, tropomyosin (actin-binding protein), actin related protein 2/3 complex (PNAS-139), plasminogen activator urokinase (PLAU) and ras homolog gene family member A (RHOA). Alpha-actinin was shown to interact with syndecan 4 in the process of stabilisation of focal adhesions for cell survival (Choi *et al.*, 2008) and Urokinase-type plasminogen activator (PLAU) was associated with breast cancer cell invasion (Hildenbrand and Schaaf, 2009). Syndecan 4 binds and activates protein kinase C (PKC) alpha and was associated with RHOA in cell spreading mechanisms (Thodeti *et al.*, 2003). It was previously shown that PKC alpha promotes cell survival and cell proliferation (Reyland, 2009). Thus, decreased expression of syndecan 4 would lead to decreased activity of PKC alpha, thus increased apoptosis. Also, lactate

dehydrogenase A (LDHA) and cysteine-rich angiogenic inducer 61 (CYR61) were found significantly downregulated in the BUC11 silenced-MDA231 cells. Lactate dehydrogenase has a role in angiogenesis (Koukourakis *et al.*, 2003) and CYR61, also an angiogenic inducer, is associated with integrins that play a role in RHOA-mediated cell proliferation (Walsh *et al.*, 2008). CYR61 expression was also associated with resistance to apoptosis in the breast cancer cell line MCF-7 (Lin *et al.*, 2004). Finally, the protein ubiquitin, shown to have a role in labelling proteins for proteosomal degradation, in stabilisation of other proteins and in apoptosis (Yang *et al.*, 2009), was also found significantly downregulated in the microarray study.

In order to validate the changes in gene expression observed in the microarray analysis, primers were designed for 7 significantly upregulated genes (GPX1, RPS7, CRR9, MRPL11, NDUFA1, COX7c and MTCH2) and 9 significantly downregulated genes (LDH A, MT2A, RSP11, RPL30, RHOA, MT1F, CYR61). These genes were chosen according to their rank order in the list of the most up- or down-regulated genes as well as their interesting cell function(s). Quantitative RT-PCR analysis was carried out using the same RNA samples that were sent for microarray analysis. As seen in Figure 5.8, 8 out of the 9 tested genes for downregulation in BUC11 silenced-MDA231 cells compared to control cells and 4 out of the 7 tested genes for upregulation in BUC11 silenced-MDA231 cells compared to control cells were validated (the Ct values can be found in Appendix III). Overall, 75% of the genes tested showed the same change in expression observed in the analysis of the microarray data. The change in gene expression for RPS11, RPS7, COX7c and NDUFA1 did not correlate to the microarray analysis. For COX7c and NDUFA1, one of the replicates did not correlate to the microarray data so the genes were not validated whereas both replicates for the ribosomal proteins RPS7 and RPS11 did not correlate to the microarray data.

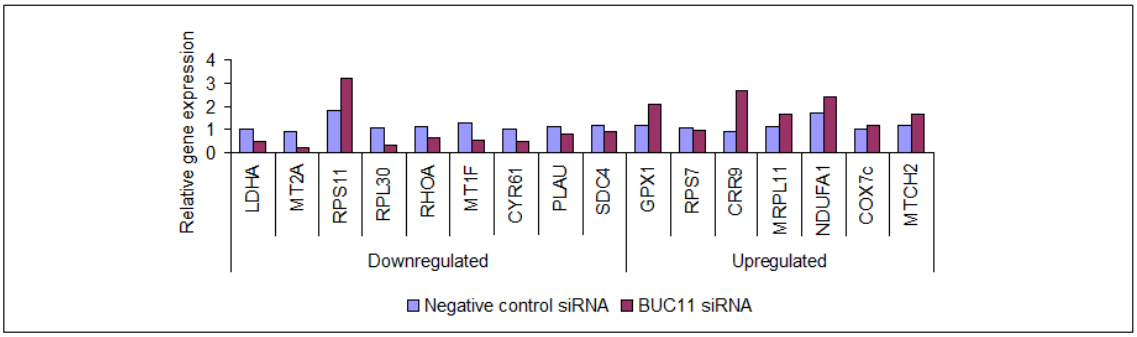


Figure 5.8: Validation of the microarray data by quantitative RT-PCR.
The 16 genes selected for validation that were found to be significantly upregulated or downregulated in the analysis of the microarray data are shown. Primers were designed and used in quantitative RT-PCR analysis using the same RNA samples sent to microarray in order to confirm the change in gene expression between control and test samples. The valued are given as mean of the duplicates. The change in gene expression for RPS11, RPS7, COX7c and NDUFA1 did not correlate with the microarray analysis.

As stated previously, the metallothionein 2A was found to be the most downregulated gene in the BUC11 silenced-MDA231 cells. Therefore, levels of MT2A expression and levels of BUC11 expression might be controlled together by the same regulatory mechanism according to the state of the breast cells. MT2A mRNA expression was measured using quantitative RT-PCR in a panel of the breast tissues (paired normal and tumour tissues) generously provided by Dr A. Gritzapis (Greece) and previously used in the quantitative RT-PCR analysis of BUC11 (Chapter 3). MT2A mRNA expression analysis was carried out on cDNA from 23 patients which were selected so that the panel had a large range of age, stage of the cancer, grade of the cancer and HER2 status. As seen in Figure 5.9, no significant correlation between BUC11 mRNA expression and MT2A mRNA expression was observed in these tissues. Only 10 patients out of 23 showed an increase in mRNA expression for both BUC11 and MT2A in the tumour tissue compared to the normal tissue (patient numbers 78, 80, 64, 77, 63, 90, 60, 62, 84 and 67, grouped as group A). No patients tested showed a decrease in mRNA expression for both genes. It was found that 11 patients out of 23 showed an increase in MT2A mRNA expression and a decrease in BUC11 mRNA expression in the tumour tissue compared to the normal tissue (patient numbers 65, 76, 74, 83, 88, 89, 61, 86, 81, 75 and 87, grouped as group B). In particular, patient 81 showed a silenced BUC11 mRNA expression and an above average increase in MT2A mRNA expression in the tumour tissue compared to the normal tissue. Patient 59 showed decreased MT2A and increased BUC11 mRNA expressions and patient 79 showed an unchanged BUC11 and increased MT2A mRNA expressions, in the tumour tissue compared to the normal tissue. In group A, 8 out of 10 patients were over 50 years of age at the time of diagnosis whereas only for 4 out of 11 patients of group B were over 50 years of age. In group A, 9 out of 10 patients tested negative for overexpression of the HER2 receptor in their tumour whereas 6 out of 11 patients tested negative in group B. In

group A, 6 out of 9 patients had a grade II tumour and 3 out of 9 patients had a grade III tumour. In group B, 6 out of 10 patients had a grade II tumour and 4 out of 10 patients had a grade III tumour. In group A, 2 out of 10 patients had a Stage I tumour and 8 out of 10 patients had a stage II tumour. In group B, 2 out of 10 patients had a Stage I tumour, 7 out of 10 patients had a stage II tumour and 1 out of 10 patients had a stage III tumour.

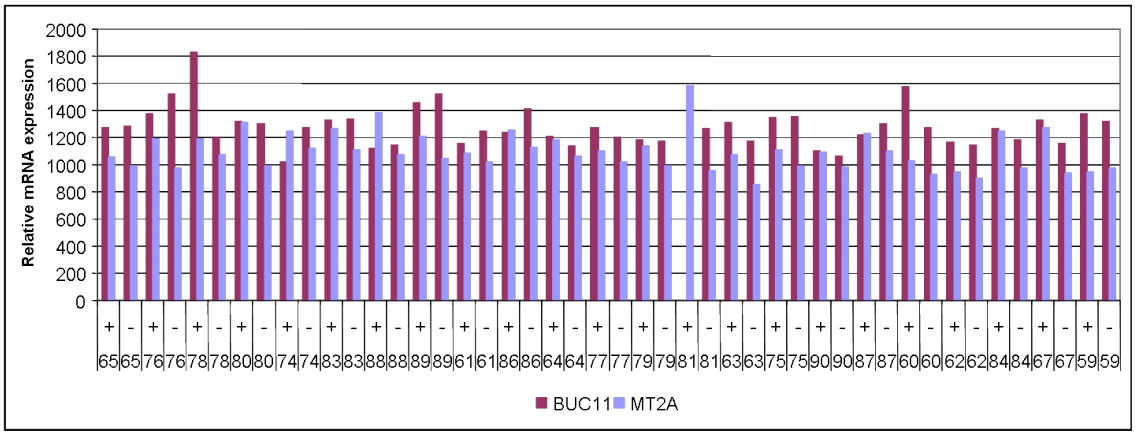


Figure 5.9: Comparison of BUC11 mRNA expression with MT2A mRNA expression in paired normal and tumour breast tissues by quantitative RT-PCR.
Quantitative RT-PCR analysis was carried out on 23 patients (numbers) with paired normal (-) and tumour (+) tissues. No significant trend was found between BUC11 mRNA expression and MT2A mRNA expression in these tissues.

Ingenuity Pathways Analysis is a software application made by Ingenuity®systems which allows researchers to quickly analyse experimental biological and chemical data by building dynamic pathway models linked to the scientific literature. Pathways based on the analysis of the microarray data obtained in this study have been built using this software. For this purpose, the 926 significantly upregulated genes (averages range from 24.32 to 2) and the 592 significantly downregulated genes (averages range from -31.07 to -2) were selected however, as the number of genes that can be associated in one single pathway analysis is limited to about 500, only the first 500 of each list were used.

Regarding the network build with the most upregulated genes of the microarray analysis (Figure 5.10, Figure 5.11 and Figure 5.12), several respiratory chain proteins were found linked to each other. For instance, the respiratory chain transporter protein ATP5B binds to YWHAG which is a protein associated with genomic instability that might promote tumourigenesis (Qi *et al.*, 2007). Another protein mentioned earlier, the ion channel protein VDAC3, binds to beta-actin (ACTB) which is important in cell shape, motility and growth (Dormoy-Raclet *et al.*, 2007) and could play a role in apoptosis (Tang *et al.*, 2006). Moreover, annexin ANXA2 binds to ACTB and was shown to be involved in apoptosis of lung cancer cells (Huang *et al.*, 2008), to have a role in the malignant phenotype of prostate cancer (Inokuchi *et al.*, 2009) and to be associated with breast tumour

invasion (Sharma *et al.*, 2006). ACTB acts on the SWI/SNF related, matrix associated, actin dependent regulator of chromatin (SMARCA4) which is an actin-dependent regulator of chromatin and which loss of expression is associated with lung cancer development (Reisman *et al.*, 2003). The ribosomal proteins RPS9 and RPL31 bind to both ACTB and the Ly1 antibody reactive homolog (mouse) (LYAR). LYAR is not well characterised to date and its limited literature suggests metal ion and protein binding. Several heat shock proteins (HSP90B1, HSP90AA1 and HSP90AB1) involved in stabilising proteins under cellular stress and in cell survival to chemotherapy of acute myeloid leukemia cells when overexpressed (Flandrin*et al.*, 2008) and their interactions with each other are shown. As seen in Figure 5.11 and Figure 5.12, several proteins indirectly act on Cyclin D1 (CCND1). CCND1 activity is essential for the cell cycle transition at G1/S and its overexpression has been associated with cancer (Fu *et al.*, 2008; Zhao *et al.*, 2008; Shi *et al.*, 2008). Finally, the small nuclear ribonucleoprotein polypeptide E (SNRPE) is a contributor of oncogene-mediated reorganisation of the actin cytoskeleton and focal adhesions (Klein *et al.*, 2008).

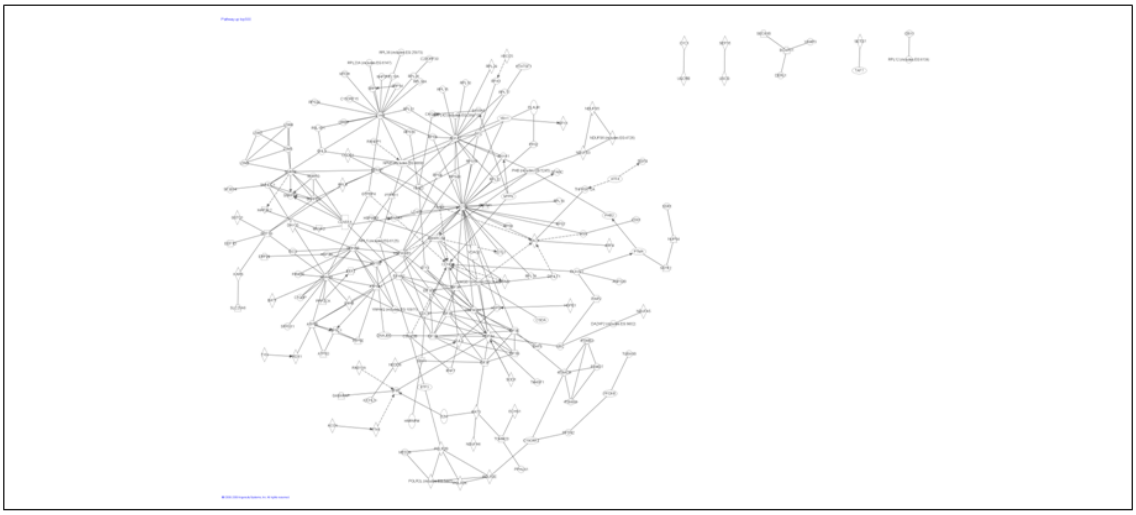


Figure 5.10: Microarray upregulated gene network build in the software Ingenuity Pathways Analysis.
The analysis was carried out with the top 500 most upregulated genes, when BUC11 gene expression is downregulated in MDA231 breast cancer cells, found during the analysis of the microarray data. Some genes of the microarray analysis did not show any interaction with other genes of the microarray analysis so they are not shown on this figure. The software searches in the literature to build the network and to inform on the functions of the proteins (network shapes) and the relationship between the proteins (network arrows).

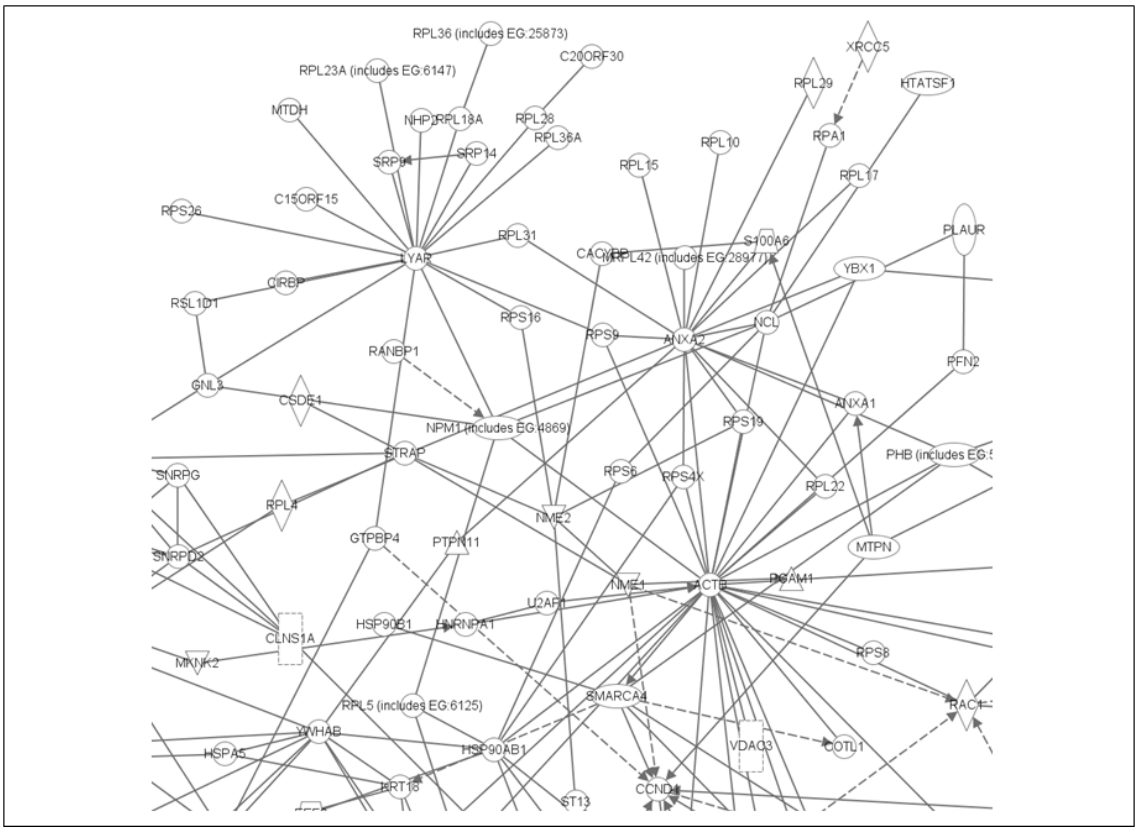


Figure 5.11: Focus on the microarray upregulated gene network build in the software Ingenuity Pathways Analysis (1).
The software searches the literature to build the network and to inform on the functions of the proteins (network shapes) and the relationship between the proteins (network arrows). The proteins highly involved in the network such as LYAR, ANXA2, ACTB, CCND1, YWHAG and HSP90 (B1, A5, AB1) are shown.

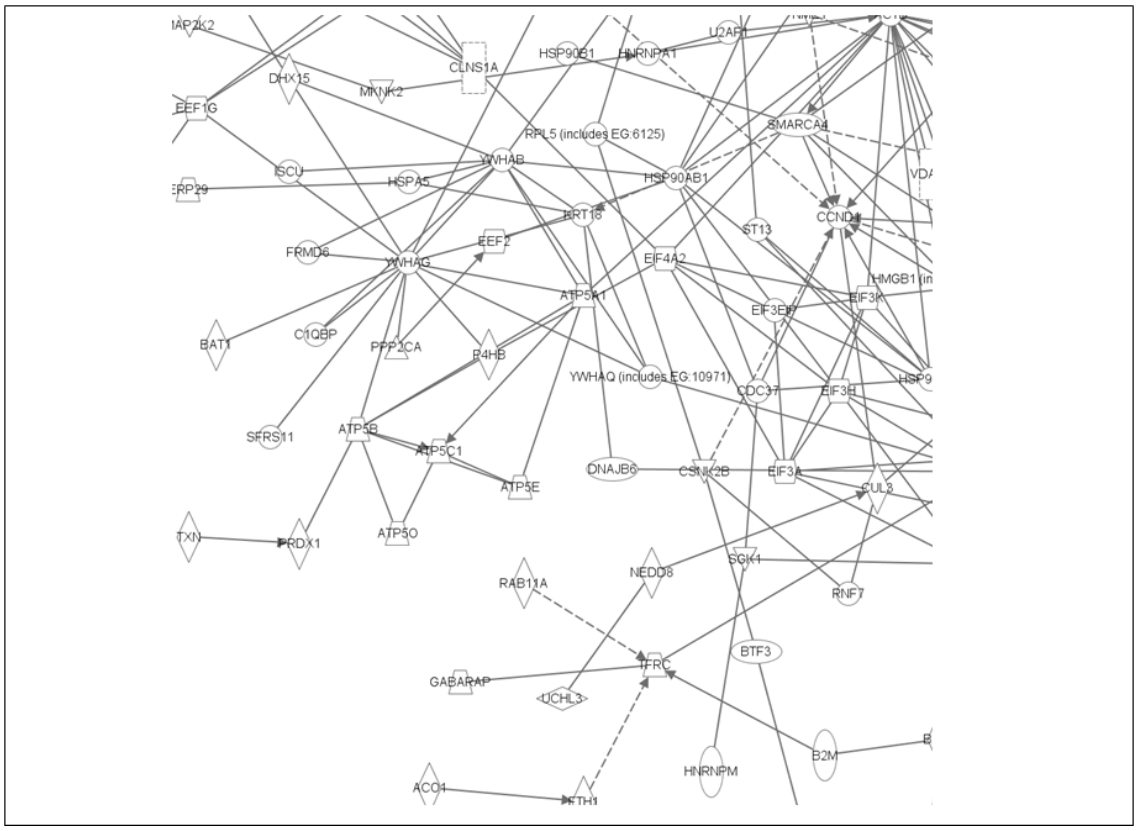


Figure 5.12: Focus on the microarray upregulated gene network build in the software Ingenuity Pathways Analysis (2).
The software searches the literature to build the network and to inform on the functions of the proteins (network shapes) and the relationship between the proteins (network arrows). The proteins highly involved in the network such as CCND1, YWHAG, ATP5B and HSP90 (B1, A5, AB1) are shown.

Regarding the network built with the most downregulated genes of the microarray analysis (Figure 5.13, Figure 5.14 and Figure 5.15), the major histocompatibility complex, class I antigen (HLA-B) was found to be associated with a number of proteins in the network. Another important protein in the network is the catenin (cadherin-associated protein), beta 1 (CTNNB1) which mediates cell adhesion, anchoring the actin cytoskeleton and was associated with tumour progression and metastasis (Goulioumis *et al.*, 2008; Gaujoux *et al.* 2008; Chen *et al.*, 2005). The tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta polypeptide (YWHAZ) is another major player in the network and acts on CTNNB1. It was found to be a marker of recurrence in breast cancer (Frasor *et al.*, 2006). YWHAZ binds to the previously mentioned angiogenic inducer in breast cancer LDHA. RHOA, PLAU, SDC4 and CYR61 are other proteins previously mentioned. Here in the network, it was shown that CYR61 indirectly acts on PLAU whereas CTNNB1 directly acts on PLAU. SDC4 is indirectly linked to CYR61 and CTNNB1. RHOA was found indirectly linked to CTNNB1. CTNNB1 is indirectly linked to the previously mentioned ACTN1 and directly acts on cyclin D2 (CCND2). Methylation of cyclin D2 has been associated with 41% of breast cancers and has been linked to

an increased cell proliferation (Euhus *et al.*, 2008). Finally, the hypoxia inducible factor 1 alpha subunit (HIF1A) acts directly on LDHA and CTNNB1 and its overexpression has been associated with tumour progression and poor outcome for breast cancer patients (Kim *et al.*, 2008).

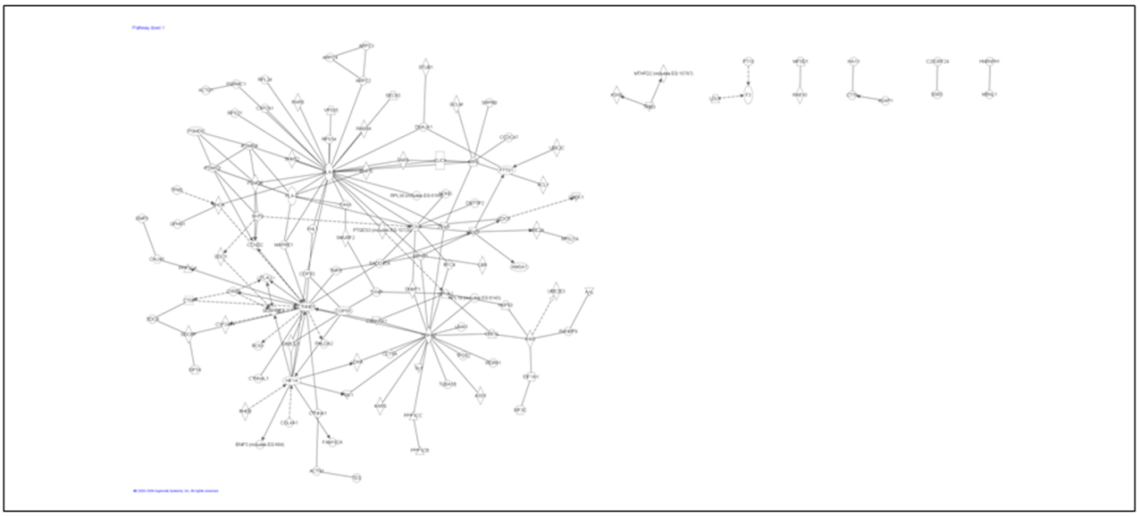


Figure 5.13: Microarray downregulated gene network build in the software Ingenuity Pathways Analysis.
The analysis was carried out with the top 500 most downregulated genes, when BUC11 gene expression is downregulated in MDA231 breast cancer cells, found during the analysis of the microarray data. Some genes of the microarray analysis did not show any interaction with other genes of the microarray analysis so they are not shown on this figure. The software searches in the literature to build the network and to inform on the functions of the proteins (network shapes) and the relationship between the proteins (network arrows).

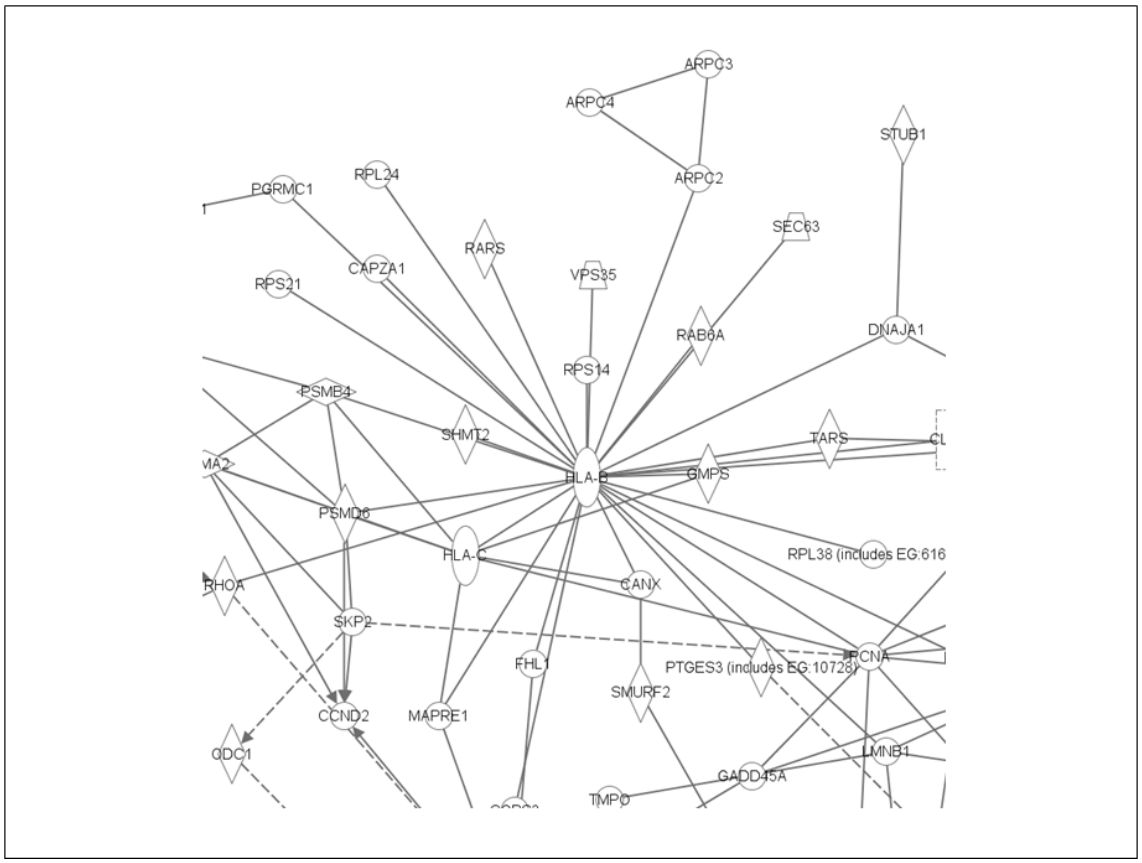


Figure 5.14: Focus on the microarray downregulated gene network build in the software Ingenuity Pathways Analysis (1).
The software searches the literature to build the network and to inform on the functions of the proteins (network shapes) and the relationship between the proteins (network arrows). The proteins highly involved in the network such as HLA-B, CCND2 and PCNA.

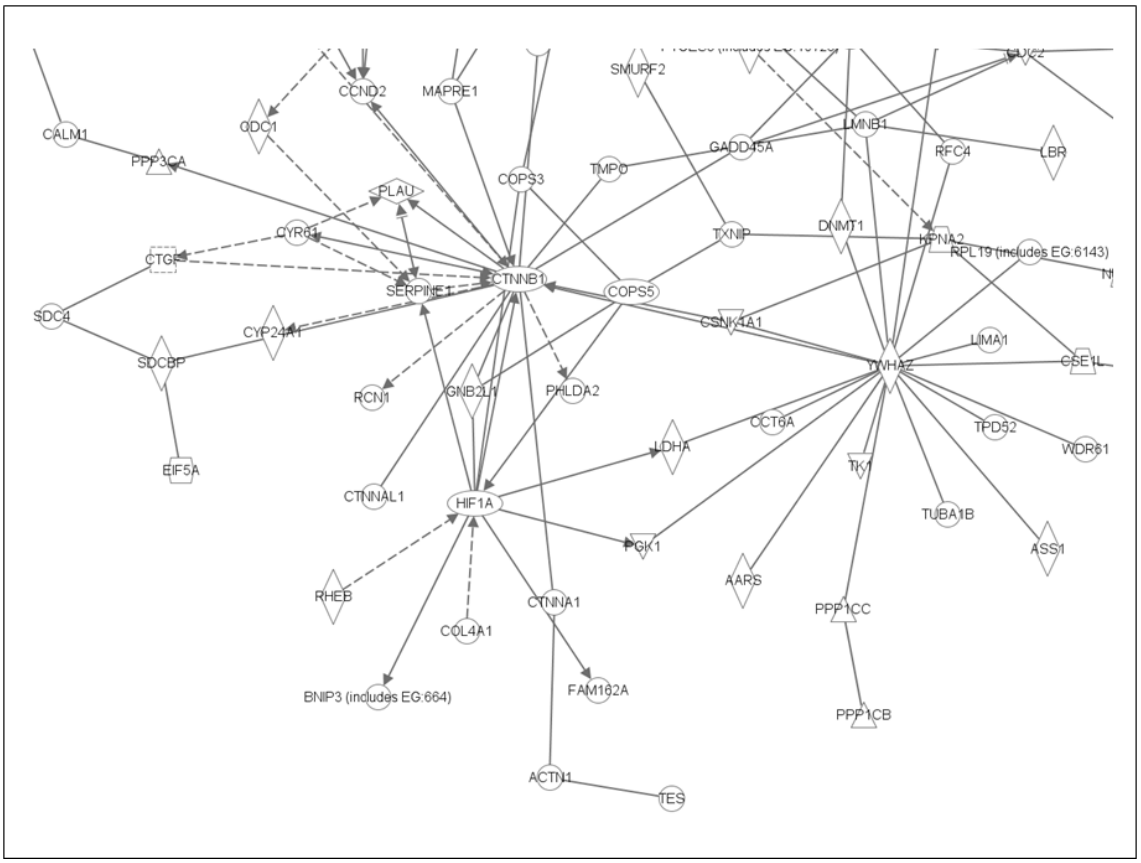


Figure 5.15: Focus on the microarray downregulated gene network build in the software Ingenuity Pathways Analysis (2).
The software searches the literature to build the network and to inform on the functions of the proteins (network shapes) and the relationship between the proteins (network arrows). The proteins highly involved in the network such as YWHAZ, PLAU, SDC4, CYR61, CTNNB1, ACTN1 and HIF1A.

5.2.7 Preliminary studies on the immunogenicity of BUC11 protein

Conventional therapies (surgery, radiotherapy and chemotherapy) can prove successful in the treatment of small and accessible tumours. However, these can be associated with side effects, resistance to drugs and inefficacy to eliminate distant metastatic tumours. Immunotherapy is an alternative therapy that can overcome these issues and is based on targeting antigens of interest which ideal characteristics include: to be solely or mainly expressed by tumour cells, to be vital to the tumour and to be immunogenic. In the previous chapters, BUC11 was shown to be highly expressed in tumour cells and not in vital normal tissues and seemed to have an essential role in tumour cell growth. It was then of interest to find out whether BUC11 was immunogenic in view that it could be used as a target. Antigen presenting cells such as dendritic cells ingest abnormal or foreign cells and present the digested contents which are small peptides (18-22 amino acids long) to other cellular components of the immune system via the Major histocompatibility complex (MHC) Class I molecules with the CD8+ T cell receptor or via the MHC class II molecules with the CD4+ T cell receptor. Upon MHC class I presentation to CD8+ T cells and with the help of IFN gamma, CD8+ T cells will be activated into CTL. Upon MHC class II presentation to CD4+ T cells (helper T cells), CD4+ will also be activated. Activated CTL will recognise the antigen-displaying tumour cell and kill it. In order to have all this cascade of events, the participation of these three cell populations is necessary in the majority of cases. The identification of class I and class II peptides that can be used in vaccination-based therapies of HLA-A2 and HLA-DR1/DR4 cancer patients would demonstrate the immunogenic potential of BUC11 protein. For this purpose, preliminary studies on the natural processing and the immunogenicity of peptides derived from BUC11 protein were conducted using a DNA immunisation strategy of HHDII/DR1 double transgenic mice which rely on the human MHC molecules to present peptide; the mice are knock-out for mouse MHC class I and II antigens.

At the time of this study, we did not have a mouse tumour model to study tumour regression nor BUC11 overlapping peptides to test individually for immunogenicity. The mammalian expression vector pBudCE4.1/BUC11 plasmid was used in this study to ascertain that the whole BUC11 protein could induce a specific immune response by showing that, in the immunised mice, BUC11-derived peptides are naturally endogenously processed and presented by HLA molecules on the cell surface. Two methods of immunisation were tested in this study. The first method was the injection of gold-coated pBudCE4.1/BUC11 plasmid DNA with a gene gun so that BUC11 protein is expressed in the HHDII/DR1 mice (three rounds of gene gun immunisation at seven-day intervals). The second method was the injection intramuscularly (IM) of pBudCE4.1/BUC11 plasmid DNA (two rounds of IM immunisation at seven-day intervals). In both cases, other HHDII/DR1 mice were injected with the empty vector pBudCE4.1/- as controls. Briefly, after the rounds of immunisation, splenocytes were harvested from the immu-

nised mice. Meanwhile, BM-DC from naive mice were harvested and stimulated *in vitro* with BUC11 peptide or an irrelevant peptide. In this study, the irrelevant peptide chosen was HAGE, which is a CT antigen expressed in several human cancer types. Then, the two cell populations were plated together. The CTL activity was assessed by monitoring their ability to kill target cells and the release of IFN gamma. Also, proliferation of T helper cells was measured by thymidine incorporation. BM-DC were pulsed with either ALC/pBudCE4.1/BUC11 cell lysate, ALC/pBudCE4.1/HAGE cell lysate, ALC/pBudCE4.1/- cell lysate, HAGE peptide or BUC11 peptide. Each of these were individually mixed with BUC11 antibody before plating them with BM-DC cultures.

In the study, the CTL activity (the ability of T cells generated in immunised mice against BUC11-derived peptides to specifically recognised and kill ALC/pBudCE4.1/BUC11 cells (HLA-A2 positive cells)) was measured by ^{15}Cr release assay. ALC/pBudCE4.1 /BUC11 cells, ALC/pBudCE4.1/HAGE cells and ALC/pBudCE4.1/- cells were plated with T cells generated and isolated from naive, control and test mice to determine the specificity of the immune response. Preliminary results obtained after one experiment were not found to be statistically significant, with large standard deviations. There was generally more killing of ALC/pBudCE4.1/BUC11 cells compared to ALC/pBudCE4.1/ HAGE cells but this difference was not significant. Therefore, no peptide-specific cytotoxic response was achieved in this experiment (data not shown). This observation suggests that none of the BUC11-derived peptides were naturally processed and presented on HLA molecules to T cells in the immunised mice.
Also, the peptide-specific cytokine production (IFN gamma) was measured by ELISPOT. Preliminary results obtained from three experiments showed that controls (T cells alone and BM-DC alone) showed no significant difference in production of IFN gamma from test mice (data not shown).
Finally, CD4+ cell proliferation was measured by ^{3}H-thymidine incorporation assay (data not shown). After one experiment, a significant difference in increased proliferation of CD4+ cells in contact with BUC11 peptide compared to CD4+ cells in contact with the irrelevant HAGE peptide was observed. This result was obtained for both test mice immunised by gene gun with pBudCE4.1/BUC11 plasmid DNA when compared with the two control mice immunised by gene gun with pBudCE4.1/- plasmid DNA ($p<0.05$ for mouse test 1 and $p<0.01$ for mouse test 2). Unfortunately, with the IM immunised mice, CD4+ cells isolated from both test and control mice (no replicate due to death of one control mouse and one test mouse during the experiment) showed highly significant increase of proliferation (respectively $p<0.01$ and $p<0.001$) when incubated with BUC11 peptide-pulsed DC compared to incubation with HAGE peptide-pulsed DC. Furthermore, CD4+ cells isolated from the naive mouse used as negative control also showed a significant increased in cell proliferation ($p<0.001$). These observations lead to the hypothesis that gene gun immunisation with pBudCE4.1/BUC11 might have produced an immune re-

sponse whereby BUC11-derived peptide(s) specifically primed CD4+ cells to proliferate when cells were re-stimulated *in vitro* and that IM immunisation with pBudCE4.1/BUC11 was insufficient when used alone but stimulation *in vitro* was apparently sufficient to induce a significant CD4+ cell proliferation in IM-immunised and naive mice.

5.3 Discussion

Previously, the structure and conserved domains of the unpublished BUC11 predicted protein was studied by database mining and its localisation was shown by immunohistochemistry. However, no conclusions could be drawn on the potential function of BUC11 in the cells from these studies. In this chapter, an attempt to provide information on the role of BUC11 gene in the proliferation, morphology and behaviour of cells by conducting experiments based on the transfection or silencing of BUC11 was described.
The coding region for the 75 amino acid BUC11 protein (hCG25653 nucleotides 95 to 322) was successfully cloned into the mammalian expression vector pBudCE4.1. At this point, which was early in the study (after sequencing of BUC genes and prior BUC11 protein analysis), it was clearly understood that the sequencing result of BUC11 included artifact sequences which can be introduced during sequencing experiments. The correct nucleotide sequence of BUC11 therefore matched the hCG25653 nucleotide sequence thus the predicted protein is the same. The plasmid pBudCE4.1/BUC11 was then used to transfect, by electroporation, a BUC11-negative mouse lymphoma (ALC) cell line which was previously used in our laboratory for similar experiments. Quantitative RT-PCR analysis on the RNA extracted from transiently transfected cells demonstrated significant expression of BUC11 mRNA by these cells, as early as 15 hours following transfection. Also, lysates of pBudCE4.1/BUC11-transfected ALC cells were used in a Western blotting experiment and appeared to contain a protein specifically recognised by BUC11 antibody (Chapter 4). These results demonstrated that BUC11 DNA has been successfully transfected into cells that do not naturally express BUC11 and this cell line can be used in further studies. It has previously been noted that mRNA and protein levels do not necessarily correlate, therefore the transfection experiment should be repeated with a time course and the levels of expression of BUC11 protein confirmed by confocal microscopy using BUC11 specific antibody. Also, attempts to produce stably transfected cells were unsuccessful in this study and further work should be undertaken to establish a cell line with constitutive BUC11 expression.

Specific siRNA to the BUC11 protein coding region allowed a significant knockdown of BUC11 mRNA expression in a BUC11-positive MDA231 human breast cancer cell line, measured 48 hours following transient transfection by quantitative RT-PCR. Morphological comparisons of BUC11 siRNA-transfected cells with control siRNA transfected cells did not reveal any noticeable effects of BUC11 silencing on cell morphology,

however cells appeared to proliferate slower when BUC11 expression was inhibited. The observed decrease in the rate of proliferation was not linked to an interferon-based or transfection reagent-based response but to a specific BUC11 siRNA-mediated response to the knockdown of BUC11 mRNA expression. Similar to BUC11 DNA transfection in ALC cells, the siRNA transfection experiment should be repeated with a longer time course to establish the levels of mRNA expression by quantitative RT-PCR in parallel to the confirmation of BUC11 knockdown at the protein level by immunofluorescence. In this study, preliminary gene knockdown experiments were also carried out with BUC6 specific siRNAs but specific gene silencing following transfection into SKBR3 cells could not be observed (data not shown).

Evasion of apoptosis and unregulated growth are two of the six hallmarks of cancer (Hanahan and Weinberg, 2000). Upon confirmation of specific BUC11 gene induction and repression obtained in targeted cells, proliferation assays using incorporation of ^{3}H-thymidine were carried out in order to determine whether BUC11 could have a role in the proliferative pathways of the cell. A significant decrease in cell proliferation was shown in BUC11-silenced cells whereas the reverse phenomenon was observed in BUC11-induced cells. These correlating effects strongly suggested that BUC11 plays a role in pathways controlling the proliferation of cells. However, the proliferation assays would need to be successfully repeated several times in the future to absolutely confirm the findings. Also, transfection of ALC cells should be repeated with the same time course (8 days) in order to correlate the proliferation assay results with the levels of expression of BUC11 mRNA and protein by quantitative RT-PCR and immunofluorescence and confirm the transient nature of transfection without selection with antibiotics. Furthermore, as part of future work, stably transfected ALC cells with pBudCE4.1/BUC11 plasmid should be used in siRNA-mediated BUC11 silencing experiments to demonstrate decrease in cell proliferation by proliferation assays and in BUC11 protein expression by immunofluorescence. This would confirm that the effects of BUC11 gene induction can be reversed by specific gene knockdown. Also, for a more accurate comparison between the targeted cell lines, BUC11 DNA transfection should be repeated into a BUC11-negative human breast cancer cell line. Finally, BUC11 induction and silencing experiments could be repeated in the future in normal human breast cell lines, if available, to establish if the same effects can be obtained in non-cancerous cells.

According to the findings of the cell cycle analysis carried out on BUC11-silenced cells, BUC11 seems to play a role in a pathway involved in the regulation of cells at the DNA damage G1/S checkpoint. Cells enter the G1 phase of the cell cycle once they have received signals for cell division. According to intracellular and extracellular signals at the G1/S checkpoint of the cell cycle, cells are directed either to die by apoptosis or to undergo cell division. It is at this point that damaged cells undergo cell cycle arrest

while repair mechanisms help to restore a healthy state. Defects in cell cycle arrest at the G1/S checkpoint has been linked to genome instability and tumourigenesis (Nojima, 2004). Therefore, BUC11 could be involved in a pathway regulating cell proliferation which allows damaged or cancer cells to bypass surveillance mechanisms (checkpoints), avoid cellular repairs and apoptosis and ultimately divide uncontrollably.
The effects of downregulating BUC11 mRNA expression were studied at the whole genome level using Affymetrix GeneChip®Microarray. Microarray technology has been widely used for studies on the classification of breast cancer and on breast tumour development and progression, following the completion of the Human Genome Project almost 10 years ago (Cooper, 2001). As expected, the genechips produced an overwhelming amount of data so to make the data more manageable the top 100 most upregulated and top 100 most downregulated genes when BUC11 mRNA expression is downregulated were selected initially for analysis by searching the literature. Overall, when BUC11 gene is downregulated in MDA231 breast cancer cells, a significantly increased expression of genes associated with growth arrest and apoptosis and a significantly decreased expression of genes associated with the cytoskeleton, cell migration, angiogenesis, cell proliferation (especially metallothioneins) and cell survival were observed. These observations were confirmed by studying networks of the up- and down-regulated genes respectively with the software application Ingenuity Pathways Analysis. This software is a useful tool and provides a snapshot of the significant microarray findings in just two networks, therefore adding weight in the hypothesis of pathways that BUC11 might be involved in. It is worth noting that RNA microarrays do not reflect the proteome, as mRNA and protein levels do not always correlate. A validation study on key genes found in the microarray analysis was carried out and showed validity of the data although the experiment should be repeated on a set of RNAs extracted from additional gene knockdown experiments. To summarise, the microarray analysis of the whole genome of a breast cancer cell line, in which BUC11 mRNA expression has been downregulated, correlated with the cell proliferation analysis and highlighted genes that might be interacting, directly and indirectly, with BUC11 in cells. However, this microarray analysis did not associate BUC11 with any particular pathway, and only provided insight into genes and pathways to target in the future. Also, the experiment was only carried out with duplicates and would need to be repeated with more replicates in order to obtain more statistically significant data. Furthermore, the experiment should be extended to other breast cancer cell lines expressing BUC11 and with cell lines transfected to express BUC11, in order to compare different genomes.

Reverse immunology and transgenic mice expressing HLA molecules are used to identify peptides from tumour antigens that are of interest in developing vaccination strategies for tumour immunotherapies. For instance, Gritzapis and colleagues have shown protection and therapeutic potentials of cytotoxic T lymphocyte (CTL) epitopes

derived from the antigen HER2, when HER2 transgenic mice were immunised with a HER2 vaccine (Gritzapis *et al.*, 2006). Also, vaccination of breast cancer patients with a mix of the same HER2 peptides and GM-CSF led to specific activation of naive and memory CD4+ and CD8+ T cells (Hueman *et al.*, 2007). CD4+ T cells are now known to be essential players in the induction and continuity of an immune response directed against a tumour as well as being involved in tumour regression through apoptotic pathway initiation or IFN-mediated mechanisms (Assudani *et al.*, 2007). Previously, our laboratory used gene gun technology in the study of class I p53-derived peptides, which consists in immunising mice with gold particles coated with a plasmid encoding the gene of interest (Tuting, 1999). Immunisation with plasmid DNA has many advantages, including the production of the protein of interest *in vivo* which can last for months, contrary to direct injection of proteins which effects only last a few days. Unfortunately, preliminary studies on the immunogenicity of BUC11 protein provided inconclusive data. No specific killing of ALC/pBudCE4.1/BUC11 cells by CTL and no specific production of IFN gamma was observed. Specific CD4+ T cell proliferation was obtained following gene gun immunisation however the non-specific proliferation of CD4+ T cells in the IM study and naive mouse disproved the validity of the phenomenon due to immunisation and not subsequent re-stimulation *in vitro*. The experiments would need to be repeated in the future to assess the immunogenic potential of BUC11 protein and of individual overlapping peptides which could form targets for antigen-specific vaccine or cell-based therapy in breast cancer. Overlapping peptides of a antigen of interest are commonly used in our laboratory to identify peptide targets that have low-to-moderate affinity with HLA molecules, but nevertheless show immunogenic potential. DNA vaccine efficiency can be improved by cloning the gene of interest together with cytokine genes or genes coding for co-stimulatory molecules in the same plasmid (Guinn *et al.*, 1999; Liu *et al.*, 2004). The expression vector pBudCE4.1 has two cloning sites and two independent promoters, thus future cloning for immunogenicity studies using pBudCE4.1/BUC11 is possible. Another potential immunisation strategy for future work could be the use of viral or bacterial vectors which are widely used for gene delivering in both *in vivo* and *in vitro* immunological studies and have been shown to increase the intensity of the immune response to antigens (Aarts *et al.*, 2002).

Genes that play a role in tumourigenesis and tumour progression (such as cell proliferation) and which peptides derived from the proteins are able to induce a strong immune response against the tumour cells and not the normal cells make appealing targets for novel cancer therapies respectively gene therapy and tumour immunotherapy. In conclusion, BUC11 appears to be involved in a proliferative pathway however further investigation is necessary to refine the role(s) of BUC11 in the breast cancer cells as well as in the normal breast cells and to assess the immunogenic potential of BUC11 protein-derived peptides.

Chapter 6

Discussion

6.1 The BUC genes - novel breast-associated genes

Since the rapid development and use of new technologies in the last 20 years or so, considerable advances in breast cancer research have been made possible. In accordance with the latest statistical figures, these have in turn supported greater improvements in the tailored management of patients, such as the exciting use of gene signatures for prediction of treatment outcome and prognosis of breast cancer (van de Vijver *et al.*, 2002; van 't Veer *et al.*, 2002; Cronin *et al.*, 2004; Paik *et al.*, 2004; Ring *et al.*, 2006). However, despite the substantial amount of work being undertaken and the abundance of promising results obtained in pre-clinical studies, very few biomarkers for the management of breast cancer patients have shown significant utility in clinical trials and hence been approved for use in humans. Thus, improvements still need to be made in all aspects of breast cancer, from prevention to treatment of late-stage diseases (Thompson *et al.*, 2008). Conventionally, the treatment options available for breast cancer patients include surgery, chemotherapy, radiotherapy, hormonal therapies and molecular targeted therapies, however, to date, no vaccine for immunotherapy has been approved. Also, not all patients can receive hormonal or molecular therapies as demonstrated in the case of patients bearing triple negative tumours (ER negative, PR negative, HER2 not overexpressed) who consequently have an extremely poor prognosis and can only be referred for chemotherapy and/or radiotherapy with their respective toxic side effects. Therefore, the identification and validation of novel, highly specific antigens for breast cancer, which could be used to develop novel targeted treatment strategies, is of prime importance for many research groups. Ideally, following screening of the antigens presented by the tumour cells, a patient would receive a personalised combination of surgery to remove the tumour mass and targeted immunotherapy to boost the immune system to overcome tumour escape and seek and kill invasive, metastatic and cancer stem cells, the latter thought to be involved in the recurrent nature of the disease (Morrison *et al.*, 2008). This tailored management would therefore spare breast cancer patients from the burden, unspecificity and associated toxic side effects of chemotherapy and radiotherapy as well as improve the success rate

and reduce the cost of treatments.

It is well recognised that a long and painstaking process is undertaken between the discovery and the clinical approval of tumour antigen-based therapy or new drugs. An ideal tumour antigen for molecular therapy or immunotherapy has to conform to several criteria, briefly: it must not be expressed in normal tissues, it has to be overexpressed in malignant tissues, it should be involved in tumorigenesis making it essential for tumour cell survival, it must be immunogenic and be able to induce tumour regression. Importantly for the use in diagnosis or prognosis, the antigen must be easily detectable in the patient's fluids (serum or urine) or biopsy, isolated using minimally invasive procedures and be highly expressed in metastatic cells.

Prior to this study, novel unpublished breast-associated genes (BUC6, BUC9, BUC10 and BUC11) were identified in our laboratory by UniGene database mining and preliminary mRNA expression analysis in various normal and malignant tissues had been performed. The scope of this research project was to validate these genes as breast cancer-associated antigens prior to pre-clinical *in vivo* studies. The first aspect of the study was to undertake extensive *in silico* analysis and *in vitro* expression analysis for all the BUC genes in order to further select the most promising ones for potential use in the management of breast cancer patients. For instance, potential targets for cancer vaccines should not be expressed in normal tissues in order to avoid the toxicity of auto-immune responses but overexpressed in malignant tissues and be essential for the survival of the tumour cells.

The research was concerned as to whether any of the BUC genes meet the first requirements for an "ideal tumour antigen", i.e. no expression in normal tissues and overexpression in malignant tissues. To address this question, the data collected for each gene will be summarised individually. The techniques used in this study for the expression analysis of the BUC genes included sequencing, database mining (gene similarities and gene function), semi-quantitative RT-PCR and quantitative RT-PCR. Firstly, it was ascertained that the BUC genes were novel genes as they did not correspond or were not highly similar to already published and well-characterised genes. The BUC genes are located on chromosome 10p11.21 and they appear to be human-specific genes. The mRNA expression analysis of BUC6 revealed that it was highly expressed in the melanoma, breast cancer and mesothelioma samples tested; and expressed to a comparatively lower level in gastric and colon carcinoma samples. Thus, BUC6 is not tumour-specific to breast cancer but further studies with larger cohort of cancer patients is justified to reveal its potential as a diagnostic, prognostic or therapeutic marker for these malignancies. Similarly, BUC9 was not found to be specific to any particular cancer, but interestingly, two splice variants of its mRNA have been characterised and the semi-quantitative RT-PCR data collected during this study suggests that the shortest BUC9 variant is solely expressed in the major-

ity but not all of the normal breast tissues, breast cancer tissues, breast cancer cell lines and normal testis tissues. Therefore, this particular variant could potentially be a breast cancer target antigen present in non-vital tissues. The expression pattern of the two splice variants could also be an indication of the disease state (diagnosis) thus further work on larger cohorts of patients is required. BUC10 also had a similar mRNA expression pattern to BUC9 but did not present splice variants. However, its high mRNA expression level in some normal tissues is a major drawback when considering BUC10 as a target for novel breast cancer therapies.

Undoubtedly, BUC11 was found to be the most promising gene of the BUC family, therefore much of the research presented in this thesis focused, at a relatively early stage, on the characterisation of this gene. BUC11 is not considered to be a tumour-specific gene either, as its mRNA was found at high levels in prostate cancer tissues (25%), breast cancer tissues (up to 97%) and testicular cancer tissues (100%) and at lower levels in melanoma (11.5%). BUC11 mRNA was not detected in any other cancer tissues tested and is expressed at very low levels in normal tissues except in the non-essential tissues of the prostate, breast, testis and placenta where it is expressed at a higher level and therefore not truly cancer-testis specific. However, it has been recently shown that cancer testis antigens, such as NY-ESO-1, which expression was thought to be restricted to gametogenic tissues (normal and tumour), can actually be expressed at low levels in some normal tissues, predominantly pancreatic tissues and this does not affect their value in immunotherapy as these low levels are unlikely to induce an autoimmune response (Scanlan *et al.*, 2004). In comparison, the extensively studied breast tissue differentiation antigen NY-BR-1, which is located next to BUC11 on the chromosome, was found to be expressed in more than 80% of breast tumours (Wang *et al.*, 2006), in up to 32% of prostate cancers and sporadically expressed in normal prostate tissues (Jager *et al.*, 2005) which is similar to BUC11 mRNA expression pattern. However, quantitative RT-PCR analysis on a small cohort of breast samples suggested that NY-BR-1 gene expression is not related to BUC11 gene expression (data not shown). The data collected from the analysis of BUC11 mRNA expression in paired tumours (of various clinical features) and normal surrounding tissues from 109 breast cancer patients demonstrated that up to 97% of all tissues expressed BUC11 (3% silencing in tumour tissues only), however only 60.6% of patients over-expressed BUC11 mRNA in their tumours compared to the normal counterpart tissues, at widely varying levels, with an overall average of 14 times more BUC11 mRNA molecules in the tumour tissue; whereas the remaining 39.4% had downregulated BUC11 mRNA expression with an overall average of 3 times less than in the normal tissue. These percentages could represent a drawback when considering BUC11 as a valid target for breast cancer immunotherapies, however, the importance of this result has to be moderated because the tissues have not been subjected to laser-capture microdissection to remove the stromal cell component (Walton *et al.*, 2009). Also, BUC11 mRNA levels

could not be significantly associated with the progression or aggressiveness of the disease, and from the overall analysis of the data collected in this study, it may be concluded that BUC11 may not represent the ideal candidate biomarker for diagnosis, prognosis or treatment of breast cancer, nevertheless its mRNA expression patterns warrant further validation. Furthermore, most of the therapeutic applicability of the BUC11 gene depends on the translation of the gene product into protein, its role in the tumour cell behaviour and the protein expression patterns in normal versus malignant tissues.

Important recent research has shown that short, single-stranded, non-coding RNA molecules (MicroRNAs or miRNAs) have crucial roles in the post-transcriptional regulation of gene expression through the degradation of target mRNAs or inhibition of the translation of target proteins. They have been associated with many processes in tumour development and tumour differentiation in various human cancers, making them very useful to classify tumours for diagnostic and prognostic purposes (Lu *et al.*, 2005). Several breast cancer-associated miRNAs which expression correlates with ER status, PR status, HER2 status and stage of the tumour have been identified (Iorio *et al.*, 2005; Mattie *et al.*, 2006; Lowery *et al.*, 2007; Adams *et al.*, 2007). MiRNAs can deregulate apoptotic or differentiation pathways in the cell and act as oncogenes by inhibiting the function of tumour suppressor genes (Lowery *et al.*, 2008). For examples, miR-17-5p is thought to be a tumour suppressor miRNA which down-regulates the AIB1 protein, previously found over-expressed in breast cancer (Hossain *et al.*, 2006), whereas miR-21 has been characterised as an oncogenic miRNA whereby its inhibition *in vitro* in MCF-7 cells caused a significant decrease in cell proliferation and increase in apoptosis (Si *et al.*, 2007). The regulatory nature of miRNAs generated great interest in targeting specific pathways of cancer stem cells such as self-renewal pathways, especially since it has been published that stem cells differentially expressed particular miRNAs (Suh *et al.*, 2004). Specific miRNAs have been shown to suppress the activity of multiple inhibitors of the G1/S restriction point of the cell cycle in embryonic stem cells, enabling rapid growth of the embryo in normal development (Wang and Blelloch, 2009). Some of these miRNAs, shown to be overexpressed in cancers, could potentially be targeted to reduce tumour growth in a therapeutic setting (Shi and Guo, 2009). Even if some of the BUC genes (6, 9 and 10) are not translated into proteins, one can not exclude that they could act in a similar manner to miRNAs, having an important role in the regulation of crucial pathways for the cancer cells and this justifies further studies on the BUC family to assess their potential as markers for diagnosis or prognosis and as targets for treatment of cancer patients. This forms a potential focus for further research.

6.2 BUC11 - a novel marker of epithelial luminal-type breast cancers?

It is widely accepted that mRNA expression levels and protein expression levels do not necessarily correlate (Greenbaum *et al.*, 2003). Also, as mentioned previously, the most valuable markers are present in the tumour cell as expressed protein, ideally as cell surface protein antigens that are over-expressed. Consequently, BUC11 protein expression analysis and protein localisation analysis were carried out during this project using immunoassays with the "custom-designed" anti-BUC11 monospecific antibody. Although the identity of the protein to which the antibody was binding to could not absolutely be confirmed by mass spectrometry during this study, the immunoassays showed the presence of BUC11 protein and validated the results of quantitative RT-PCR. *In silico* analysis of BUC11 protein sequence (predicted mass of 8.724kDa) revealed that it is dissimilar to any other deposited protein sequence. Interestingly, the region 49-67 of BUC11 protein sequence shares similarity to the X-linked XIAP protein, an inhibitor of apoptosis (Lee and Cho, 2008; Choi *et al.*, 2008; Aird *et al.*, 2008), previously studied as a target for the treatment of breast cancer using inhibitors of proteins protecting XIAP from degradation and antisense oligonucleotides (Hong *et al.*, 2009; Tamm, 2008). Moreover, proteins demonstrating sequence similarities to BUC11 have all been linked to cell growth or survival, thus providing additional support for the potential function of the BUC11 protein.

Immunohistochemical staining of tissue microarrays composed of archival tumour samples is a reliable, high-throughput technology commonly used in the expression analysis of multiple markers of the subgroups of breast cancers in order to determine their respective relevance in diagnosis and prognosis of the disease. These immunohistochemical techniques applied to TMA preparations are frequently used by Professor Ian Ellis's breast cancer research group and proven successful in the expression study of basal and luminal cell cytokeratins (Abd El-Rehim *et al.*, 2004), in the validation study of markers identified through cDNA microarrays (Abd El-Rehim *et al.*, 2005), in the evaluation of the prognostic relevance of mucins (Rakha *et al.*, 2005) and in the characterisation of breast cancers with myoepithelial and basal phenotypes (Rakha *et al.*, 2006). An extensive immunohistochemical study on tumour tissues from a large cohort of 670 invasive breast carcinoma patients confirmed the homogenous distribution of BUC11 protein in different tumour cell compartments, either the nucleus (30.5% tested positive for BUC11 protein presence in at least the nucleus) or the cytoplasm or both but predominantly in the cytoplasm. The levels varied depending on the tissue type, thus suggesting BUC11 could be involved in different processes depending on the cell requirements, for example, at particular time points in the cell cycle or in response to external stimuli. Following post-translational modification processes, proteins synthesised in the cytoplasm which have a role to fulfill in the nucleus present in their sequence a nuclear localisation sig-

nal (NLS) which allows their transport to their target compartment through nuclear pore complexes (Gerace and Burke, 1988; Garcia-Bustos *et al.*, 1991). Thus, BUC11 protein does not appear to be present on the outer membrane of the tumour cells, making it an unlikely target for antibody-based therapies in breast cancer patients. Also, its expression was not significantly associated to the clinical outcome of breast cancer patients (from the TMA analysis performed) however it was significantly correlated with the expression of the progesterone receptor and other markers of endocrine hormonally positive epithelial luminal-type breast cancers, without any distinction between luminal A or non-luminal A (luminal B and potentially other subtypes associated with poor prognosis) subtypes. Consequently, BUC11 is a potential marker of the luminal type of breast cancers and a potential antigen regulated by sex steroid hormones (Theurillat *et al.*, 2008). The basal and HER2 breast cancer subtypes are already well defined and can be readily identified through the molecular profiling of their respective validated markers, however clinically relevant markers defining the luminal subgroups have yet to be ascertained. Luminal-specific markers are of high interest for the stratification of breast cancer patients likely to respond to therapies (patients bearing luminal A subtype tumours have good prognosis when treated with tamoxifen) as well as targets for novel therapies themselves (Loi, 2008). Also, in the past few years, researchers have provided important information on the luminal B subtype such as the low expression of the ER and ER-regulated genes, high expression of proliferation genes and deregulated apoptotic, HER2 and PI3K/ AKT pathways; however further pre-clinical studies, using *in vitro* and *in vivo* models, are necessary to refine the knowledge on cellular pathways affected in luminal B as well as luminal A subtypes of breast cancer and to validate clinically relevant luminal subtype-associated markers (Loi, 2008).

6.3 BUC11 - a gene involved in cell proliferation pathways?

As previously discussed, evasion of apoptosis and unregulated growth are two of the six hallmarks of cancer (Hanahan and Weinberg, 2000) thus over-expressed genes that are involved in pathways affecting tumorigenesis, cell growth and proliferation represent excellent potential molecular targets for new diagnostic, prognostic or therapeutic strategies for the management of breast cancer patients. No conclusions could be drawn on the potential function of BUC11 in the cells from the *in silico* analysis, however experiments based on the transfection and silencing of BUC11 gene were conducted and the results provided clues on the biological and cellular role of the BUC11 gene product. Using the incorporation of ^{3}H-thymidine, a significant decrease in cell proliferation was observed in BUC11-silenced cells whereas the reverse phenomenon was observed in BUC11-induced cells. This interdependence suggested that BUC11 might have a function in pathways

controlling the proliferation of cells and a possible anti-apoptotic effect. This hypothesis was supported by cell cycle analysis performed on BUC11-silenced cells which informed that BUC11 could be involved in a pathway regulating cell proliferation which allows damaged or cancer cells to bypass surveillance mechanisms at checkpoints, especially G1/S, to avoid apoptosis and grow in an unrestrained manner (Nojima, 2004). Furthermore, microarray analysis of the expressed genes of a breast cancer cell line in which BUC11 mRNA expression has been significantly downregulated correlated with the cell cycle and proliferation analyses. The interpretation of the microarray data showed that, under normal circumstances, BUC11 is co-expressed with various genes associated with cell proliferation (especially metallothioneins), cell survival, cell structure, cell migration and even angiogenesis. As mentioned before, metallothioneins are associated with proliferation of breast cancer cells (Bay *et al.*, 2006), making them good targets for antisense therapies (Yang *et al.*, 2002). Although the BUC11 gene could not be assigned to any specific functional pathways, the genes identified by modulating BUC11 expression and supposedly co-expressed with BUC11 represent targets for future exploration to uncover signalling events associated with BUC11, in both normal and malignant breast cells. Antigens involved in cell proliferation, such as HER2 which is one of the most studied (Menard *et al.*, 2001; Slamon *et al.*, 1989), are often associated to a poor clinical outcome. In this study, BUC11 expression could not be linked to an adverse prognosis for breast cancer patients (quantitative RT-PCR and immunohistochemical experiments) even though it appears to be critical to the proliferation of the breast cancer cells tested.

6.4 BUC11 - a potential candidate for immunotherapy?

Proteins associated with the malignant process and having an important role in tumorigenesis and tumour progression can provide MHC-restricted peptides effective in promoting immune responses specifically targeted against the tumour cells and promoting tumour regression. Such antigens make desirable targets for novel molecular therapies, based on knowledge of cancer-associated signalling pathways and immunotherapies (Ostrand-Rosenberg, 2004). The discovery and validation of novel targets for cancer vaccine therapy is particularly crucial for the reason that, as previously stated, no therapeutic vaccine for the treatment of human cancers has been FDA approved to date. Trastuzumab and lapatinib, both molecular therapies, were the only two targeted breast cancer therapeutics approved in 2007 and a novel DNA-based vaccine for immunotherapy of metastasised breast cancer, called polyHER2neu, is currently being tested in phase 1 of a UK clinical trial (http://www.cancerhelp.org.uk/trials, 2009). The results obtained during the course of this project provide a rationale to further investigate BUC11 as a target for immunotherapy: the protein is expressed; low mRNA expression in normal essential tissues; high mRNA expression in non-vital normal tissues (placenta, breast, prostate and testis); high mRNA expression in prostate, testis and breast cancer tissues; appears to have a

vital role in cell proliferation. Over-expression of the antigen in tumours is an essential characteristic to be considered as a potential target for therapy. Although BUC11 is only over-expressed in the tumour of approximately 61% of the breast cancer patients tested in this study, patients could be pre-screened for BUC11 over-expression following biopsy and those testing positive for BUC11 over-expression could be referred for BUC11 targeted therapies. Theoretically, more than half of breast cancer patients could benefit from such treatments.

Due to time restriction, only preliminary immunological experiments to determine the suitability of BUC11 as a potential target for immunotherapy were performed. Briefly, the approach was to immunise mice transgenic for HLA molecules with a BUC11 DNA construct encoding for the protein to induce specific immune responses that could be monitored *in vitro*. The data generated demonstrated no specific killing of target cells by CTL, no specific production of IFN-gamma and no specific CD4+ T cell proliferation. These results show that the BUC11 construct might not be suitable or that BUC11 protein might not be immunogenic, therefore the investigation would have to be repeated, especially with BUC11-specific peptides.

6.5 Conclusions and future work

The findings from this study suggest that BUC11 is a novel marker of epithelial luminal-type breast cancers. It was shown to have a role in cell proliferation and meet the criteria of an antigen that could potentially be used in novel targeted therapies of breast and possibly other cancers such as testicular or prostate cancers.

It is essential that any novel candidate antigen is rigorously characterised during the early stages of the research prior to undertaking time consuming and costly work with pre-clinical animal models followed by clinical trials. Importantly, the candidate antigen has to present additional benefits to already published and validated antigens. Here, initial studies should focus on further validation of the BUC11 gene. The vital role for tumour progression, expression patterns as well as the immunogenicity and endogenous processing of BUC11 protein needs to be unquestionably ascertained in future experiments. However, one should consider also the potential of other BUC genes, which should be studied in larger cohorts of cancer patients using quantitative RT-PCR, ideally with RNA isolated from laser micro-dissected tissues to obtain representative expression patterns in cells from normal and malignant glands as well as the tumour stroma. In particular, the mRNA expression patterns of BUC9 splice variants should be further investigated in order to confirm its relevance in patient management. The quantity, quality and accessibility of materials remain limiting factors in breast cancer research but these issues need to be addressed so that appropriate clinical samples, collected before and

during treatments to primary tumours as well as metastatic samples, are used in expression studies (Thompson *et al.*, 2008). If BUC6, BUC9 or BUC10 do not translate into proteins, screening of their respective expression at the RNA level could be used in the context of diagnosis or prognosis if they present valid expression patterns associated with clinical disease parameters. It would also be essential to assess the species specificity of the BUC genes using *in silico* analysis and PCR-based assays with RNA from diverse species, to validate the full length of mRNA sequences of the BUC genes by sequencing the 5' and 3' ends using a RACE-PCR approach and to confirm the organisation of the BUC genes on chromosome 10 (one large gene having several splice variants or several individual genes?). Furthermore, recombinant BUC proteins could be produced through the cloning of the coding sequences into an expression vector which would give, for example, high-level expression of Histidine-tagged proteins. These proteins could be easily purified using colums containing a resin with immobilised cobalt or nickel ions which Histidine has an affinity for.

Prior to any further immunoassays, mass spectrometry could be used to validate the identity of the protein BUC11 antibody is binding to on the Western Blot. Once the tryptic MS fragments are characterised and the key ions identified, the data could be used in the mass spectrometry analysis to establish the presence of BUC11 in the serum of cancer patients. It was presumed that the size of proteins detected on the Western Blots (15kDa/30kDa) could be explained by post-translational formation of dimers or other modifications. Important denaturation of the protein mixtures to detect post-translational modifications in BUC11 should be used to further clarify this point. The study of BUC11 mRNA and protein expression patterns should be extended to larger cohorts of melanoma, breast and prostate cancer patients in order to validate the expression pattern and obtain statistical significance. This would also support the hypothesis that BUC11 is a novel marker of epithelial luminal-type breast cancers, along with the analysis of the co-expression of BUC11 with validated epithelial luminal markers by immunohistochemistry and quantitative RT-PCR. Also, transcriptional silencing affecting BUC11 gene and protein expression were observed in up to 10% of the breast tumours tested but not in the normal counterpart tissues. It would therefore be of interest to identify the processes underlying this observation as well as identify the BUC11 promoter.

Regarding the biological role of BUC11 in cells, transfection experiments with either the expression vector or siRNA silencing should be repeated in both malignant and normal human breast cell lines (BUC11-negative and BUC11-positive cell lines) to confirm the role of BUC11 in cell proliferation. One could use as a negative control cells that would have been tranfected with a plasmid encoding a mutated BUC11 protein. Furthermore, the cause of BUC11 protein translocation to the nucleus of the breast cancer cells should be ascertained with the use of stress-related assays and immunofluorescence. The

phenomenon should also be studied in normal breast cells using immunohistochemistry to determine whether expression could be cancer-specific. It has recently been demonstrated that breast cancer stem cells are more resistant to radiation (Phillips *et al.*, 2006) and to cytotoxic agents (Dean *et al.*, 2005) than differentiated breast cancer cells, resulting in failure of current therapies due to the relapse and metastasis of the disease. Therefore, following conventional therapies that debulk most of the tumour mass, surviving breast cancer stem cells should be targeted using novel therapies aimed at either inducing differentiation of those cells or elimination through immunotherapy. In the latter case, targets for immunotherapy must be antigens that are expressed by the breast cancer stem cells, ideally on their cell surface, but not by the normal stem cells but in association with MHC class I antigen (Morrison *et al.*, 2008). It would thus be interesting to determine whether or not BUC11 is expressed in breast cancer stem cells as a novel cancer stem cell-associated marker that might represent a potential target for novel therapies to treat and/or prevent recurrence of the disease. Finally, the immunogenicity of BUC11 protein can be assessed using MHC class I and class II overlapping peptides derived from the protein using a MHC transgenic mouse model and a tumour cell line expressing MHC class I and the appropriate class I-restricted epitope(s).

This study has demonstrated for the first time the expression of a novel gene set in breast/breast cancer cells that potentially have clinical utility and merit further detailed study. This is considerable scope for research, since the function and biological significance of the BUC genes/proteins have not been elucidated and based on the results of the present study, research to address these issues can be proposed.

The Ct values of the BUC11 mRNA expression analysis in cell lines and tissues

: not specific (either no amplification or melt curve not specific)

Breast cancer cell line	GAPDH Ct	BUC11 Ct
T47D	13.74	24.02
MDA231	12.31	24.01
MDA468	11.86	
BR293	11.51	
MCF7	11.78	32.17
MDAP3	10.88	32.35
SKBR3	12.41	32.76

Normal tissue	BUC11 Ct	GAPDH Ct	HPRT1 Ct	TBP Ct
Bone marrow	31.29	13.50	20.28	22.35
Brain	28.69	11.54	17.38	20.73
Fetal brain		11.80	18.41	19.47
Fetal liver		13.30	21.19	23.34
Heart	31.82	11.39	19.66	22.52
Kidney	31.74	11.82	19.27	21.18
Liver		14.14	22.05	23.76
Lung	31.67	14.25	20.51	21.69
Placenta	26.91	13.60	21.75	21.45
Prostate	30.38	14.38	20.62	21.93
Skeletal muscle	30.44	10.02	22.03	22.64
Spleen		15.34	21.47	22.61
Testes	19.49	13.43	18.02	18.17
Thymus	30.36	13.20	19.41	20.75
Trachea		20.79	22.50	23.40
Uterus	30.32	15.15	22.01	21.94
Colon		14.36	21.51	22.71
Small intestine		14.31	20.97	23.12
Spinal cord	28.21	12.48	20.22	21.13
Stomach		14.89	21.86	22.87

Melanoma tissue	BUC11 Ct	GAPDH Ct	HPRT1 Ct	TBP Ct
Mel A		18.14	25.8	27.57
Mel B		17.33	26.15	27.65
Mel C		17.31	25.42	26.95
Mel D	32.98	21.85		31.18
Mel E	17.10	18.97	26.96	
Mel F		18.56		26.73
Mel 522	30.61	19.57	28.32	29.43
Mel 529		18.59		28.5
Mel 537		20.57	29.22	30.09
Mel 584				
Mel 674	34.89	12.89	18.96	22.26
Mel 677		22.42		
Mel 678		18.48		27.44
Mel 681		19.77		28.3
Mel 687	21.29	20.42	28.64	
Mel 689		18.85	27.88	28.34
Mel 698	15.86	14.86	23.32	25.48
Mel 702		12.38	20.49	23.91
Mel 708		16.69	21.5	26.31
Mel 709		14.49	24.88	29.09
Mel 719	30.62	16.68	25.18	26.77
Mel 724		19.83		28.74
Mel 739		16.97	24.52	26.62
Mel 760	31.93	19.01		28.69
Mel 773		15.11	22.64	27.66
Mel 780	30.18	12.26	23.69	25.14

Prostate tissue		BUC11 Ct	GAPDH Ct	HPRT1 Ct	TBP Ct
Normal	Normal prostate	30.38	14.38	20.62	21.93
Benign	Pr 92/9		18.48	27.43	25.79
	Pr 92/38		17.29	27.04	27.56
	Pr 92/103		19.15	28.41	28.83
Malignant	Pr 92/41	31.75	14.65	27.4	24.28
	Pr 18	23.49	15.95	30.1	30.21
	Pr 19		14.75	26.19	27.86
	Pr 21		19.35	31.95	29.78
	Pr 25	29.37	16.4	24.78	27.46
	Pr 26		18.73	28.28	28.21
Type unknown	Pr 93/186	18.93	14.21	29.18	27.61
	Pr 92/97		21.09	29.16	30.71
	Pr 93/33		16.22	25.26	26.95
	Pr 93/37		18.57		28.14
	Pr 93/111		17.09	25.94	25.85
	Pr 93/116		16.17		26.56
	Pr 93/140		18.47	31.99	30.36

The list of the breast tissues collected at the Saint Savas Cancer Hospital

Tissue number	Patient number	Tissue	Age	HER	Grade	Stage	Comments
BRCA1	1	+	45	N	II	IIb	
BRCA2	1	-					
BRCA3	2	+	50		III	IIb	
BRCA4	2	-					
BRCA5	3	+	70		II	IIb	
BRCA6	3	-					
BRCA7	4	+	74	N	III	IIb	
BRCA8	4	-					
BRCA9	5	+	56		III	IIa	
BRCA10	5	-					
BRCA11	6	+	50		II	II	
BRCA12	6	-					
BRCA13	7	+	66	N	I	IIb	
BRCA14	7	-					
BRCA15	8	+	63	N	II	IIIa	
BRCA16	8	-					
BRCA17	9	+	55	N	II	Metastasis	Tumour from relapse
BRCA18	9	-					
BRCA19	10	+	67	3	III	IIb	
BRCA20	10	-					
BRCA21	11	+	54	N	II	IIb	
BRCA22	11	-					
BRCA23	12	+	58	3	III	IIb	
BRCA24	12	-					
BRCA25	13	+	34	N	II	IIb	
BRCA26	13	-					
BRCA27	14	+	41	N	III	IIIa	
BRCA28	14	-					
BRCA29	15	+	34	3	III	IIb	
BRCA30	15	-					
BRCA31	16	+	74	1	II	IIa	
BRCA32	16	-					
BRCA33	17	+	60	3	III	IIb	
BRCA34	17	-					
BRCA35	18	+	42	N	II	IIa	
BRCA36	18	-					
BRCA37	19	+	71	3	II	IIb	
BRCA38	19	-					
BRCA39	20	+	65	N	II	IIb	
BRCA40	20	-					
BRCA41	21	+	71	N	III	IIIa	
BRCA42	21	-					
BRCA43	22	+	65	N	III	IIb	
BRCA44	22	-					
BRCA45	23	+	53	N	I	IIb	
BRCA46	23	-					
BRCA47	24	+	75	N	II	IIa	
BRCA48	24	-					
BRCA49	25	+	52	N	III	I	
BRCA50	25	-					
BRCA51	26	+	65	N	II	IIa	
BRCA52	26	-					
BRCA53	27	+	52	N	II	IIb	
BRCA54	27	-					
BRCA55	28	+	68	N	II	IIb	
BRCA56	28	-					
BRCA57	29	+	73	N	II	IIa	
BRCA58	29	-					
BRCA59	30	+	71	N	II	III	
BRCA60	30	-					
BRCA61	31	+	50	N	III	IIa	
BRCA62	31	-					
BRCA63	32	+	40	3	I	I	
BRCA64	32	-					
BRCA65	33	+	36		III	Iia	
BRCA66	33	-					
BRCA67	34	+	72	N	III	IIb	
BRCA68	34	-					
BRCA69	35	+	49			IIb	
BRCA70	35	-					

BRCA71	36	+	42	N	I	IIa	
BRCA72	36	-					
BRCA73	37	+	68	N	I	IIa	
BRCA74	37	-					
BRCA75	38	+	46	N	I	IIa	
BRCA76	38	-					
BRCA77	39	+	80	N	II	IIb	
BRCA78	39	-					
BRCA79	40	+	81	N	I	IIb	
BRCA80	40	-					
BRCA81	41	+	80	N	III	IIa	
BRCA82	41	-					
BRCA83	42	+	79	N	III	IIb	
BRCA84	42	-					
BRCA85	43	+	74	N	II	IIb	
BRCA86	43	-					
BRCA87	44	+	70	N	II	IIa	
BRCA88	44	-					
BRCA89	45	+	66	N	II	Metastasis	
BRCA90	45	-					
BRCA91	46	+	57	1	III	IIb	
BRCA92	46	-					
BRCA93	47	+	53	N	III	IIb	
BRCA94	47	-					
BRCA109	55	+	76	N	III		Tumour from relapse
BRCA110	55	-					
BRCA111	56	+	50	N	II	I	
BRCA112	56	-					
BRCA113	57	+	75	N	III	IIb	
BRCA114	57	-					
BRCA115	58	+	73	N	II	IIb	
BRCA116	58	-					
BRCA117	59	+	80	N	II	IIb	
BRCA118	59	-					
BRCA119	60	+	72	2	III	IIb	
BRCA120	60	-					
BRCA121	61	+	44	N	II	I	
BRCA122	61	-					
BRCA123	62	+	73	N	II	IIb	
BRCA124	62	-					
BRCA125	63	+	58	N	II	IIa	
BRCA126	63	-					
BRCA127	64	+	52	N		I	
BRCA128	64	-					
BRCA129	65	+	35	3	II	III	
BRCA130	65	-					
BRCA131	66	+	48			Iia	
BRCA132	66	-					
BRCA133	67	+	78	N	II	IIa	
BRCA134	67	-					
BRCA135	68	+	56	N	II	I	
BRCA136	68	-					
BRCA137	69	+	68	2	II	IIb	
BRCA138	69	-					
BRCA139	70	+	66		III	IIa	
BRCA140	70	-					
BRCA141	71	+	37	2	II	IIb	
BRCA142	71	-					
BRCA143	72	+	66	2	II	IIa	
BRCA144	72	-					
BRCA145	73	+	50	N	III	III	
BRCA146	73	-					
BRCA147	74	+	40	3	II	IIa	
BRCA148	74	-					
BRCA149	75	+	63	1--2	III	IIa	
BRCA150	75	-					
BRCA151	76	+	36	N		IIa	
BRCA152	76	-					
BRCA153	77	+	56	N	II	I	
BRCA154	77	-					

BRCA155	78	+	36	N	III	IIb	
BRCA156	78	-					
BRCA157	79	+	57	N	III	IIb	
BRCA158	79	-					
BRCA159	80	+	36	N	III	IIb	
BRCA160	80	-					
BRCA161	81	+	57	N	III	IIb	
BRCA162	81	-					
BRCA163	82	+	71	N	III	IIb	
BRCA164	82	-					
BRCA165	83	+	40	2	III		Tumour from relapse
BRCA166	83	-					
BRCA167	84	+	77	N	II	IIa	
BRCA168	84	-					
BRCA169	85	+	68	N	I	I	
BRCA170	85	-					
BRCA171	86	+	50	N	II	IIa	
BRCA172	86	-					
BRCA173	87	+	68		II	I	
BRCA174	87	-					
BRCA175	88	+	43	N	II	IIa	
BRCA176	88	-					
BRCA177	89	+	43	N	III	IIa	
BRCA178	89	-					
BRCA179	90	+	65	N	II	IIb	
BRCA180	90	-					
BRCA181	91	+	75	N	III	II	
BRCA182	91	-					
BRCA183	92	+	38	2	II	IIa	
BRCA184	92	-					
BRCA189	95	+	40	3	II	II	
BRCA190	95	-					
BRCA191	96	+	72	3		IIa	
BRCA192	96	-					
BRCA193	97	+	46	2	II	IIa	
BRCA194	97	-					
BRCA195	98	+	75	N	II	III	
BRCA196	98	-					
BRCA197	99	+	49		I	IIa	
BRCA198	99	-					
BRCA199	100	+	56	N	II	IIb	
BRCA200	100	-					
BRCA201	101	+	45	N	II	III	
BRCA202	101	-					
BRCA203	102	+	54		II	I	
BRCA204	102	-					
BRCA205	103	+	82	N	III	IIa	
BRCA206	103	-					
BRCA207	104	+	47	1	II	IIb	
BRCA208	104	-					
BRCA209	105	+	32	N	III	I	
BRCA210	105	-					
BRCA213	107	+	55	N	III	IIa	
BRCA214	107	-					
BRCA215	108	+	61	N	III	IIb	
BRCA216	108	-					
BRCA217	109	+	54		II	IIb	
BRCA218	109	-					
BRCA219	110	+	80		III	Ia	
BRCA220	110	-					
BRCA221	111	+	80		II	I	
BRCA222	111	-					
BRCA223	112	+	71	N	II	I	
BRCA224	112	-					
BRCA225	113	+	75	N	II	IIb	
BRCA226	113	-					
BRCA227	114	+	41	3	III	IIb	
BRCA228	114	-					
BRCA229	115	+	70		I+	IIb	
BRCA230	115	-					

BRCA231	116	+	74	3	II	IIa	
BRCA232	116	-					
BRCA233	117	+	NO CA				Eventually free of cancer
BRCA234	117	-	NO CA				
BRCA235	118	+	62	N	II	IIb	
BRCA236	118	-					
BRCA239	120	+	42	N	II	Metastasis	
BRCA240	120	-					

The Ct values of the validation of the microarray data by qPCR

Delta Delta CT Method analysis of GPX1/GAPDH

Replicate Name	GOI CT	Norm. CT	Relative Conc.	Calibrator
BUC11 (1)	14.59	11.17	2.32	
BUC11 (3)	14.17	10.43	1.85	
Ctrl (1)	15.33	10.69	1	Yes
Ctrl (3)	15	10.75	1.3	
Negative control	43.43	33.39	0.02	

Delta Delta CT Method analysis of LDHA/GAPDH

Replicate Name	GOI CT	Norm. CT	Relative Conc.	Calibrator
BUC11 (1)	13.07	11.17	0.56	
BUC11 (3)	12.86	10.43	0.38	
Ctrl (1)	11.75	10.69	1	Yes
Ctrl (3)	11.82	10.75	0.99	
Negative control	36.87	33.39	0.19	

Delta Delta CT Method analysis of RPS7/GAPDH

Replicate Name	GOI CT	Norm. CT	Relative Conc.	Calibrator
BUC11 (1)	13.74	11.17	1.09	
BUC11 (3)	13.42	10.43	0.81	
Ctrl (1)	13.38	10.69	1	Yes
Ctrl (3)	13.29	10.75	1.11	
Negative control	40.19	33.39	0.06	

Delta Delta CT Method analysis of MT2A/GAPDH

Replicate Name	GOI CT	Norm. CT	Relative Conc.	Calibrator
BUC11 (1)	12.91	11.17	0.20	
BUC11 (3)	11.92	10.43	0.24	
Ctrl (1)	10.11	10.69	1	Yes
Ctrl (3)	10.55	10.75	0.77	
Negative control	38.13	33.39	0.02	

Delta Delta CT Method analysis of CRR9/GAPDH

Replicate Name	GOI CT	Norm. CT	Relative Conc.	Calibrator
BUC11 (1)	14.36	11.17	3.13	
BUC11 (3)	14.1	10.43	2.24	
Ctrl (1)	15.53	10.69	1	Yes
Ctrl (3)	15.8	10.75	0.86	
Negative control	35.73	33.39	5.64	

Delta Delta CT Method analysis of RPS11/GAPDH

Replicate Name	GOI CT	Norm. CT	Relative Conc.	Calibrator
BUC11 (1)	12.23	11.17	4.15	
BUC11 (3)	12.37	10.43	2.26	
Ctrl (1)	13.81	10.69	1	Yes
Ctrl (3)	12.46	10.75	2.63	
Negative control	39.54	33.39	0.12	

Delta Delta CT Method analysis of MRPL11/GAPDH

Replicate Name	GOI CT	Norm. CT	Relative Conc.	Calibrator
BUC11 (1)	17.37	11.17	1.85	
BUC11 (3)	16.97	10.43	1.46	
Ctrl (1)	17.77	10.69	1	Yes
Ctrl (3)	17.52	10.75	1.24	
Negative control		33.39		

Delta Delta CT Method analysis of RPL30/GAPDH

Replicate Name	GOI CT	Norm. CT	Relative Conc.	Calibrator
BUC11 (1)	12.87	11.17	0.33	
BUC11 (3)	12.32	10.43	0.29	
Ctrl (1)	10.81	10.69	1	Yes
Ctrl (3)	10.73	10.75	1.1	
Negative control	41.01	33.39	0.01	

Delta Delta CT Method analysis of NDUFA1/GAPDH

Replicate Name	GOI CT	Norm. CT	Relative Conc.	Calibrator
BUC11 (1)	18.8	11.17	2.82	
BUC11 (3)	18.58	10.43	1.97	
Ctrl (1)	19.82	10.69	1	Yes
Ctrl (3)	18.59	10.75	2.43	
Negative control		33.39		

Delta Delta CT Method analysis of RHOA/GAPDH

Replicate Name	GOI CT	Norm. CT	Relative Conc.	Calibrator
BUC11 (1)	13.17	11.17	0.84	
BUC11 (3)	13.29	10.43	0.46	
Ctrl (1)	12.43	10.69	1	Yes
Ctrl (3)	12.15	10.75	1.26	
Negative control	40.04	33.39	0.03	

Delta Delta CT Method analysis of COX7c/GAPDH

Replicate Name	GOI CT	Norm. CT	Relative Conc.	Calibrator
BUC11 (1)	15.71	11.17	1.46	
BUC11 (3)	15.73	10.43	0.87	
Ctrl (1)	15.78	10.69	1	Yes
Ctrl (3)	15.84	10.75	1	
Negative control	38.05	33.39	1.35	

Delta Delta CT Method analysis of MT1F/GAPDH

Replicate Name	GOI CT	Norm. CT	Relative Conc.	Calibrator
BUC11 (1)	18.16	11.17	0.58	
BUC11 (3)	17.55	10.43	0.52	
Ctrl (1)	16.88	10.69	1	Yes
Ctrl (3)	16.29	10.75	1.56	
Negative control	31.69	33.39	236.7	

Delta Delta CT Method analysis of MTCH2/GAPDH

Replicate Name	GOI CT	Norm. CT	Relative Conc.	Calibrator
BUC11 (1)	15.64	11.17	1.72	
BUC11 (3)	14.98	10.43	1.62	
Ctrl (1)	15.94	10.69	1	Yes
Ctrl (3)	15.57	10.75	1.34	
Negative control	38.38	33.39	1.19	

Delta Delta CT Method analysis of CYR61/GAPDH

Replicate Name	GOI CT	Norm. CT	Relative Conc.	Calibrator
BUC11 (1)	14.91	11.17	0.54	
BUC11 (3)	14.74	10.43	0.37	
Ctrl (1)	13.56	10.69	1	Yes
Ctrl (3)	13.54	10.75	1.05	
Negative control	48.32	33.39	0	

Delta Delta CT Method analysis of PLAU/GAPDH

Replicate Name	GOI CT	Norm. CT	Relative Conc.	Calibrator
BUC11 (1)	15.11	11.17	0.87	
BUC11 (3)	14.66	10.43	0.71	
Ctrl (1)	14.43	10.69	1	Yes
Ctrl (3)	14.21	10.75	1.21	
Negative control	36.53	33.39	1.51	

Delta Delta CT Method analysis of SDC4/GAPDH

Replicate Name	GOI CT	Norm. CT	Relative Conc.	Calibrator
BUC11 (1)	16.02	11.17	0.94	
BUC11 (3)	15.44	10.43	0.84	
Ctrl (1)	15.45	10.69	1	Yes
Ctrl (3)	15.02	10.75	1.4	
Negative control	42.81	33.39	0.04	

References

Aarts, W. M., Schlom, J. & Hodge, J. W. (2002). Vector-based vaccine/cytokine combination therapy to enhance induction of immune responses to a self-antigen and antitumor activity. *Cancer Res*, **62**, 5770-7.

Abd El-Rehim, D. M., Ball, G., Pinder, S. E., Rakha, E., Paish, C., Robertson, J. F., Macmillan, D., Blamey, R. W. & Ellis, I. O. (2005). High-throughput protein expression analysis using tissue microarray technology of a large well-characterised series identifies biologically distinct classes of breast cancer confirming recent cDNA expression analyses. *Int J Cancer*, **116**, 340-50.

Abd El-Rehim, D. M., Pinder, S. E., Paish, C. E., Bell, J., Blamey, R. W., Robertson, J. F., Nicholson, R. I. & Ellis, I. O. (2004). Expression of luminal and basal cytokeratins in human breast carcinoma. *J Pathol*, **203**, 661-71.

Adams, B. D., Furneaux, H. & White, B. A. (2007). The micro-ribonucleic acid (miRNA) miR-206 targets the human estrogen receptor-alpha (ERalpha) and represses ERalpha messenger RNA and protein expression in breast cancer cell lines. *Mol Endocrinol*, **21**, 1132-47.

Adams, M. D., Kelley, J. M., Gocayne, J. D., Dubnick, M., Polymeropoulos, M. H., Xiao, H., Merril, C. R., Wu, A., Olde, B., Moreno, R. F. & et al. (1991). Complementary DNA sequencing: expressed sequence tags and human genome project. *Science*, **252**, 1651-6.

Aerbajinai, W., Lee, Y. T., Wojda, U., Barr, V. A. & Miller, J. L. (2004). Cloning and characterization of a gene expressed during terminal differentiation that encodes a novel inhibitor of growth. *J Biol Chem*, **279**, 1916-21.

Ahmad, M., Rees, R. C. & Ali, S. A. (2004). Escape from immunotherapy: possible mechanisms that influence tumor regression/progression. *Cancer Immunol Immunother*, **53**, 844-54.

Ahmad, M., Rees, R. C., McArdle, S. E., Li, G., Mian, S., Entwisle, C., Loudon, P. & Ali, S. A. (2005). Regulation of CTL responses to MHC-restricted class I peptide of the gp70 tumour antigen by splenic parenchymal CD4+ T cells in mice failing immunotherapy with DISC-mGM-CSF. *Int J Cancer*, **115**, 951-9.

Aird, K. M., Ding, X., Baras, A., Wei, J., Morse, M. A., Clay, T., Lyerly, H. K. & Devi, G. R. (2008). Trastuzumab signaling in ErbB2-overexpressing inflammatory breast cancer correlates with X-linked inhibitor of apoptosis protein expression. *Mol Cancer Ther*, **7**, 38-47.

Al-azawi, D., Kelly, G., Myers, E., McDermott, E. W., Hill, A. D., Duffy, M. J. & Higgins, N. O. (2006). CA 15-3 is predictive of response and disease recurrence following treatment in locally advanced breast cancer. *BMC Cancer*, **6**, 220.

Alexander-Miller, M. A., Leggatt, G. R., Sarin, A. & Berzofsky, J. A. (1996). Role of antigen, CD8, and cytotoxic T lymphocyte (CTL) avidity in high dose antigen induction of apoptosis of effector CTL. *J Exp Med*, **184**, 485-92.

Al-Hajj, M., Wicha, M. S., Benito-Hernandez, A., Morrison, S. J. & Clarke, M. F. (2003). Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci U S A*, **100**, 3983-8.

Ali, S. A., Ahmad, M., Lynam, J., McLean, C. S., Entwisle, C., Loudon, P., Choolun, E., McArdle, S. E., Li, G., Mian, S. & Rees, R. C. (2004). Anti-tumour therapeutic efficacy of OX40L in murine tumour model. *Vaccine*, **22**, 3585-94.

Ali, S. A., McLean, C. S., Boursnell, M. E., Martin, G., Holmes, C. L., Reeder, S., Entwisle, C., Blakeley, D. M., Shields, J. G., Todryk, S., Vile, R., Robins, R. A. & Rees, R. C. (2000). Preclinical evaluation of "whole" cell vaccines for prophylaxis and therapy using a disabled infectious single cycle-herpes simplex virus vector to transduce cytokine genes. *Cancer Res*, **60**, 1663-70.

Ali, S. A., Lynam, J., McLean, C. S., Entwisle, C., Loudon, P., Rojas, J. M., McArdle, S. E., Li, G., Mian, S. & Rees, R. C. (2002). Tumor regression induced by intratumor therapy with a disabled infectious single cycle (DISC) herpes simplex virus (HSV) vector, DISC/HSV/murine granulocyte-macrophage colony-stimulating factor, correlates with antigen-specific adaptive immunity. *J Immunol*, **168**, 3512-9.

Allouche, A., Nolens, G., Tancredi, A., Delacroix, L., Mardaga, J., Fridman, V., Winkler, R., Boniver, J., Delvenne, P. & Begon, D. Y. (2008). The combined immunodetection of AP-2alpha and YY1 transcription factors is associated with ERBB2 gene overexpression in primary breast tumors. *Breast Cancer Res*, **10**, R9.

Altin, J. G. & Parish, C. R. (2006). Liposomal vaccines–targeting the delivery of antigen. *Methods*, **40**, 39-52.

Andersen, T. I., Paus, E., Nesland, J. M., McKenzie, S. J. & Borresen, A. L. (1995). Detection of c-erbB-2 related protein in sera from breast cancer patients. Relationship to ERBB2 gene amplification and c-erbB-2 protein overexpression in tumour. *Acta Oncol*, **34**, 499-504.

Anderson, N. L. & Anderson, N. G. (1998). Proteome and proteomics: new technologies, new concepts, and new words. *Electrophoresis*, **19**, 1853-61.

Andrietta, M. H., Eloy, N. B., Hemerly, A. S. & Ferreira, P. C. G. (2001). Identification of sugarcane cDNAs encoding components of the cell cycle machinery. *Genetics and Molecular Biology*, **24**, 61-68.

Ardavanis, A., Gerakini, F., Amanatidou, A., Scorilas, A., Pateras, C., Garoufali, A., Pissakas, G., Stravolemos, K., Apostolikas, N. & Yiotis, I. (1997). Relationships between cathepsin-D, pS2 protein and hormonal receptors in breast cancer cytosols: inconsistency with their established prognostic significance. *Anticancer Res*, **17**, 3665-9.

Ariad, S., Seymour, L. & Bezwoda, W. R. (1991). Platelet-derived growth factor (PDGF) in plasma of breast cancer patients: correlation with stage and rate of progression. *Breast Cancer Res Treat*, **20**, 11-7.

Asselin-Paturel, C., Echchakir, H., Carayol, G., Gay, F., Opolon, P., Grunenwald, D., Chouaib, S. & Mami-Chouaib, F. (1998). Quantitative analysis of Th1, Th2 and TGF-beta1 cytokine expression in tumor, TIL and PBL of non-small cell lung cancer patients. *Int J Cancer*, **77**, 7-12.

Assudani, D. P., Ahmad, M., Li, G., Rees, R. C. & Ali, S. A. (2006). Immunotherapeutic potential of DISC-HSV and OX40L in cancer. *Cancer Immunol Immunother*, **55**, 104-11.

Assudani, D. P., Horton, R. B., Mathieu, M. G., McArdle, S. E. & Rees, R. C. (2007). The role of CD4+ T cell help in cancer immunity and the formulation of novel cancer vaccines. *Cancer Immunol Immunother*, **56**, 70-80.

Bakker, A. B., Schreurs, M. W., Tafazzul, G., de Boer, A. J., Kawakami, Y., Adema, G. J. & Figdor, C. G. (1995). Identification of a novel peptide derived from the melanocyte-specific gp100 antigen as the dominant epitope recognized by an HLA-A2.1-restricted anti-melanoma CTL line. *Int J Cancer*, **62**, 97-102.

Baldwin, A. S. (2001). Control of oncogenesis and cancer therapy resistance by the transcription factor NF-kappaB. *J Clin Invest*, **107**, 241-6.

Balogh, S., Kiss, I. & Csaszar, A. (2009). Toll-like receptors: link between "danger" ligands and plaque instability. *Curr Drug Targets*, **10**, 513-8.

Batist, G., Tulpule, A., Sinha, B. K., Katki, A. G., Myers, C. E. & Cowan, K. H. (1986). Overexpression of a novel anionic glutathione transferase in multidrug-resistant human breast cancer cells. *J Biol Chem*, **261**, 15544-9.

Bay, B. H., Jin, R., Huang, J. & Tan, P. H. (2006). Metallothionein as a prognostic biomarker in breast cancer. *Exp Biol Med (Maywood)*, **231**, 1516-21.

Bearss, D. J., Lee, R. J., Troyer, D. A., Pestell, R. G. & Windle, J. J. (2002). Differential effects of p21(WAF1/CIP1) deficiency on MMTV-ras and MMTV-myc mammary tumor properties. *Cancer Res*, **62**, 2077-84.

Bee, A., Ke, Y., Forootan, S., Lin, K., Beesley, C., Forrest, S. E. & Foster,

C. S. (2006). Ribosomal protein l19 is a prognostic marker for human prostate cancer. *Clin Cancer Res*, **12**, 2061-5.

Benitez, R., Godelaine, D., Lopez-Nevot, M. A., Brasseur, F., Jimenez, P., Marchand, M., Oliva, M. R., van Baren, N., Cabrera, T., Andry, G., Landry, C., Ruiz-Cabello, F., Boon, T. & Garrido, F. (1998). Mutations of the beta2-microglobulin gene result in a lack of HLA class I molecules on melanoma cells of two patients immunized with MAGE peptides. *Tissue Antigens*, **52**, 520-9.

Benlalam, H., Vignard, V., Khammari, A., Bonnin, A., Godet, Y., Pandolfino, M. C., Jotereau, F., Dreno, B. & Labarriere, N. (2007). Infusion of Melan-A/Mart-1 specific tumor-infiltrating lymphocytes enhanced relapse-free survival of melanoma patients. *Cancer Immunol Immunother*, **56**, 515-26.

Berd, D. (2001). Autologous, hapten-modified vaccine as a treatment for human cancers. *Vaccine*, **19**, 2565-70.

Berrington de Gonzalez, A. & Darby, S. (2004). Risk of cancer from diagnostic X-rays: estimates for the UK and 14 other countries. *Lancet*, **363**, 345-51.

Bertram, J. S. (2000). The molecular biology of cancer. *Mol Aspects Med*, **21**, 167-223.

Bezwoda, W. R., Esser, J. D., Dansey, R., Kessel, I. & Lange, M. (1991). The value of estrogen and progesterone receptor determinations in advanced breast cancer. Estrogen receptor level but not progesterone receptor level correlates with response to tamoxifen. *Cancer*, **68**, 867-72.

Biassoni, R. (2008). Natural killer cell receptors. *Adv Exp Med Biol*, **640**, 35-52.

Bieche, I., Onody, P., Laurendeau, I., Olivi, M., Vidaud, D., Lidereau, R. & Vidaud, M. (1999). Real-time reverse transcription-PCR assay for future management of ERBB2-based clinical applications. *Clin Chem*, **45**, 1148-56.

Bieche, I., Nogues, C., Paradis, V., Olivi, M., Bedossa, P., Lidereau, R. & Vidaud, M. (2000). Quantitation of hTERT gene expression in sporadic breast tumors with a real-time reverse transcription-polymerase chain reaction assay. *Clin Cancer Res*, **6**, 452-9.

Binder, R. J., Vatner, R. & Srivastava, P. (2004). The heat-shock protein receptors: some answers and more questions. *Tissue Antigens*, **64**, 442-51.

Boel, P., Wildmann, C., Sensi, M. L., Brasseur, R., Renauld, J. C., Coulie, P., Boon, T. & van der Bruggen, P. (1995). BAGE: a new gene encoding an antigen recognized on human melanomas by cytolytic T lymphocytes. *Immunity*, **2**, 167-75.

Bonnet, D. & Dick, J. E. (1997). Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nature*

Medicine, **3**, 730-737.

Borresen-Dale, A. L. (2003). TP53 and breast cancer. *Hum Mutat*, **21**, 292-300.

Borstnar, S., Vrhovec, I., Svetic, B. & Cufer, T. (2002). Prognostic value of the urokinase-type plasminogen activator, and its inhibitors and receptor in breast cancer patients. *Clin Breast Cancer*, **3**, 138-46.

Bosch, G. J., Joosten, A. M., Kessler, J. H., Melief, C. J. & Leeksma, O. C. (1996). Recognition of BCR-ABL positive leukemic blasts by human CD4+ T cells elicited by primary in vitro immunization with a BCR-ABL breakpoint peptide. *Blood*, **88**, 3522-7.

Bottini, A., Berruti, A., Bersiga, A., Brizzi, M. P., Brunelli, A., Gorzegno, G., DiMarco, B., Aguggini, S., Bolsi, G., Cirillo, F., Filippini, L., Betri, E., Bertoli, G., Alquati, P. & Dogliotti, L. (2000). p53 but not bcl-2 immunostaining is predictive of poor clinical complete response to primary chemotherapy in breast cancer patients. *Clin Cancer Res*, **6**, 2751-8.

Brossart, P., Heinrich, K. S., Stuhler, G., Behnke, L., Reichardt, V. L., Stevanovic, S., Muhm, A., Rammensee, H. G., Kanz, L. & Brugger, W. (1999). Identification of HLA-A2-restricted T-cell epitopes derived from the MUC1 tumor antigen for broadly applicable vaccine therapies. *Blood*, **93**, 4309-17.

Buchan, S., Gronevik, E., Mathiesen, I., King, C. A., Stevenson, F. K. & Rice, J. (2005). Electroporation as a "prime/boost" strategy for naked DNA vaccination against a tumor antigen. *J Immunol*, **174**, 6292-8.

Bueler, H. & Mulligan, R. C. (1996). Induction of antigen-specific tumor immunity by genetic and cellular vaccines against MAGE: enhanced tumor protection by coexpression of granulocyte-macrophage colony-stimulating factor and B7-1. *Mol Med*, **2**, 545-55.

Burdall, S. E., Hanby, A. M., Lansdown, M. R. & Speirs, V. (2003). Breast cancer cell lines: friend or foe?. *Breast Cancer Res*, **5**, 89-95.

Burnet, F.M. (1970). The concept of immunological surveillance. *Prog Exp Tumor Res* , **13**, 1-27.

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

Byrnes, G. B., Southey, M. C. & Hopper, J. L. (2008). Are the so-called low penetrance breast cancer genes, ATM, BRIP1, PALB2 and CHEK2, high risk for women with strong family histories?. *Breast Cancer Res*, **10**, 208.

Caldwell, J. S. (2007). Cancer cell-based genomic and small molecule screens. *Adv Cancer Res*, **96**, 145-73.

Carreno, B. M., Bennett, F., Chau, T. A., Ling, V., Luxenberg, D., Jussif, J., Baroja, M. L. & Madrenas, J. (2000). CTLA-4 (CD152) can inhibit T cell activation by two different mechanisms depending on its level of cell surface expression. *J Immunol*, **165**, 1352-6.

Castelli, C., Rivoltini, L., Rini, F., Belli, F., Testori, A., Maio, M., Mazzaferro, V., Coppa, J., Srivastava, P. K. & Parmiani, G. (2004). Heat shock proteins: biological functions and clinical application as personalized vaccines for human cancer. *Cancer Immunol Immunother*, **53**, 227-33.

Chan, R. C., Pang, X. W., Wang, Y. D., Chen, W. F. & Xie, Y. (2004). Transduction of dendritic cells with recombinant adenovirus encoding HCA661 activates autologous cytotoxic T lymphocytes to target hepatoma cells. *Br J Cancer*, **90**, 1636-43.

Chapelin, C., Duriez, B., Magnino, F., Goossens, M., Escudier, E. & Amselem, S. (1997). Isolation of several human axonemal dynein heavy chain genes: genomic structure of the catalytic site, phylogenetic analysis and chromosomal assignment. *FEBS Lett*, **412**, 325-30.

Chatterjee, S. K. & Zetter, B. R. (2005). Cancer biomarkers: knowing the present and predicting the future. *Future Oncol*, **1**, 37-50.

CHEK2 Breast Cancer Case-Control Consortium (2004). CHEK2*1100delC and susceptibility to breast cancer: a collaborative analysis involving 10,860 breast cancer cases and 9,065 controls from 10 studies. *Am J Hum Genet*, **74**, 1175-82.

Chen, G., Gharib, T. G., Huang, C. C., Taylor, J. M., Misek, D. E., Kardia, S. L., Giordano, T. J., Iannettoni, M. D., Orringer, M. B., Hanash, S. M. & Beer, D. G. (2002). Discordant protein and mRNA expression in lung adenocarcinomas. *Mol Cell Proteomics*, **1**, 304-13.

Chen, F. W. & Ioannou, Y. A. (1999). Ribosomal proteins in cell proliferation and apoptosis. *Int Rev Immunol*, **18**, 429-48.

Chen, Y., Thompson, W., Semenciw, R. & Mao, Y. (1999). Epidemiology of contralateral breast cancer. *Cancer Epidemiol Biomarkers Prev*, **8**, 855-61.

Chen, X. Y., Wang, Z. C., Li, H., Cheng, X. X., Sun, Y., Wang, X. W., Wu, M. L. & Liu, J. (2005). Nuclear translocations of beta-catenin and TCF4 in gastric cancers correlate with lymph node metastasis but probably not with CD44 expression. *Hum Pathol*, **36**, 1294-301.

Chen, D., Zhang, Z., Li, M., Wang, W., Li, Y., Rayburn, E. R., Hill, D. L., Wang, H. & Zhang, R. (2007). Ribosomal protein S7 as a novel modulator of p53-MDM2 interaction: binding to MDM2, stabilization of p53 protein, and activation of p53 function. *Oncogene*, **26**, 5029-37.

Chia, S. K., Wykoff, C. C., Watson, P. H., Han, C., Leek, R. D., Pastorek,

J., Gatter, K. C., Ratcliffe, P. & Harris, A. L. (2001). Prognostic significance of a novel hypoxia-regulated marker, carbonic anhydrase IX, in invasive breast carcinoma. *J Clin Oncol*, **19**, 3660-8.

Cho, Y., Gorina, S., Jeffrey, P. D. & Pavletich, N. P. (1994). Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. *Science*, **265**, 346-55.

Choi, W. Y., Kim, G. Y., Lee, W. H. & Choi, Y. H. (2008). Sanguinarine, a benzophenanthridine alkaloid, induces apoptosis in MDA-MB-231 human breast carcinoma cells through a reactive oxygen species-mediated mitochondrial pathway. *Chemotherapy*, **54**, 279-87.

Choi, Y., Kim, S., Lee, J., Ko, S. G., Lee, W., Han, I. O., Woods, A. & Oh, E. S. (2008). The oligomeric status of syndecan-4 regulates syndecan-4 interaction with alpha-actinin. *Eur J Cell Biol*, **87**, 807-15.

Collins, A. T., Berry, P. A., Hyde, C., Stower, M. J. & Maitland, N. J. (2005). Prospective identification of tumorigenic prostate cancer stem cells. *Cancer Res*, **65**, 10946-51.

Cooper, C. S. (2001). Applications of microarray technology in breast cancer research. *Breast Cancer Res*, **3**, 158-75.

Correale, P., Walmsley, K., Nieroda, C., Zaremba, S., Zhu, M., Schlom, J. & Tsang, K. Y. (1997). In vitro generation of human cytotoxic T lymphocytes specific for peptides derived from prostate-specific antigen. *J Natl Cancer Inst*, **89**, 293-300.

Cristofanilli, M. & Hortobagyi, G. N. (2002). Molecular targets in breast cancer: current status and future directions. *Endocr Relat Cancer*, **9**, 249-66.

Cronin, M., Pho, M., Dutta, D., Stephans, J. C., Shak, S., Kiefer, M. C., Esteban, J. M. & Baker, J. B. (2004). Measurement of gene expression in archival paraffin-embedded tissues: development and performance of a 92-gene reverse transcriptase-polymerase chain reaction assay. *Am J Pathol*, **164**, 35-42.

Curiel, T. J., Coukos, G., Zou, L., Alvarez, X., Cheng, P., Mottram, P., Evdemon-Hogan, M., Conejo-Garcia, J. R., Zhang, L., Burow, M., Zhu, Y., Wei, S., Kryczek, I., Daniel, B., Gordon, A., Myers, L., Lackner, A., Disis, M. L., Knutson, K. L., Chen, L. & Zou, W. (2004). Specific recruitment of regulatory T cells in ovarian carcinoma fosters immune privilege and predicts reduced survival. *Nat Med*, **10**, 942-9.

Dai, M. S., Sears, R. & Lu, H. (2007). Feedback regulation of c-Myc by ribosomal protein L11. *Cell Cycle*, **6**, 2735-41.

Dean, M., Fojo, T. & Bates, S. (2005). Tumour stem cells and drug resistance. *Nat Rev Cancer*, **5**, 275-84.

Degli-Esposti, M. A. & Smyth, M. J. (2005). Close encounters of different kinds: dendritic cells and NK cells take centre stage. *Nat Rev Immunol*, **5**, 112-24.

de Kok, J. B., Roelofs, R. W., Giesendorf, B. A., Pennings, J. L., Waas, E. T., Feuth, T., Swinkels, D. W. & Span, P. N. (2005). Normalization of gene expression measurements in tumor tissues: comparison of 13 endogenous control genes. *Lab Invest*, **85**, 154-9.

Deming, S. L., Nass, S. J., Dickson, R. B. & Trock, B. J. (2000). C-myc amplification in breast cancer: a meta-analysis of its occurrence and prognostic relevance. *Br J Cancer*, **83**, 1688-95.

Depowski, P. L., Rosenthal, S. I. & Ross, J. S. (2001). Loss of expression of the PTEN gene protein product is associated with poor outcome in breast cancer. *Mod Pathol*, **14**, 672-6.

Diaz, Y., Gonzalez, A., Lopez, A., Perez, R., Vazquez, A. M. & Montero, E. (2009). Anti-ganglioside anti-idiotypic monoclonal antibody-based cancer vaccine induces apoptosis and antiangiogenic effect in a metastatic lung carcinoma. *Cancer Immunol Immunother*, **58**, 1117-28.

Dormoy-Raclet, V., Menard, I., Clair, E., Kurban, G., Mazroui, R., Di Marco, S., von Roretz, C., Pause, A. & Gallouzi, I. E. (2007). The RNA-binding protein HuR promotes cell migration and cell invasion by stabilizing the beta-actin mRNA in a U-rich-element-dependent manner. *Mol Cell Biol*, **27**, 5365-80.

Downward, J. (2003). Targeting RAS signalling pathways in cancer therapy. *Nat Rev Cancer*, **3**, 11-22.

Dudley, M. E., Wunderlich, J. R., Yang, J. C., Hwu, P., Schwartzentruber, D. J., Topalian, S. L., Sherry, R. M., Marincola, F. M., Leitman, S. F., Seipp, C. A., Rogers-Freezer, L., Morton, K. E., Nahvi, A., Mavroukakis, S. A., White, D. E. & Rosenberg, S. A. (2002). A phase I study of nonmyeloablative chemotherapy and adoptive transfer of autologous tumor antigen-specific T lymphocytes in patients with metastatic melanoma. *J Immunother*, **25**, 243-51.

Dudley, M. E., Wunderlich, J. R., Yang, J. C., Sherry, R. M., Topalian, S. L., Restifo, N. P., Royal, R. E., Kammula, U., White, D. E., Mavroukakis, S. A., Rogers, L. J., Gracia, G. J., Jones, S. A., Mangiameli, D. P., Pelletier, M. M., Gea-Banacloche, J., Robinson, M. R., Berman, D. M., Filie, A. C., Abati, A. & Rosenberg, S. A. (2005). Adoptive cell transfer therapy following non-myeloablative but lymphodepleting chemotherapy for the treatment of patients with refractory metastatic melanoma. *J Clin Oncol*, **23**, 2346-57.

Dunn, G. P., Bruce, A. T., Ikeda, H., Old, L. J. & Schreiber, R. D. (2002). Cancer immunoediting: from immunosurveillance to tumor escape. *Nat Immunol*, **3**, 991-8.

Dunn, G. P., Bruce, A. T., Sheehan, K. C., Shankaran, V., Uppaluri, R.,

Bui, J. D., Diamond, M. S., Koebel, C. M., Arthur, C., White, J. M. & Schreiber, R. D. (2005). A critical function for type I interferons in cancer immunoediting. *Nat Immunol*, **6**, 722-9.

Eden, P., Ritz, C., Rose, C., Ferno, M. & Peterson, C. (2004). "Good Old" clinical markers have similar power in breast cancer prognosis as microarray gene expression profilers. *Eur J Cancer*, **40**, 1837-41.

Egeblad, M. & Werb, Z. (2002). New functions for the matrix metalloproteinases in cancer progression. *Nat Rev Cancer*, **2**, 161-74.

Ellisen, L. W. & Haber, D. A. (1998). Hereditary breast cancer. *Annu Rev Med*, **49**, 425-36.

Elston, C. W. & Ellis, I. O. (2002). Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology*, **41**, 154-61.

Emmert-Buck, M. R., Bonner, R. F., Smith, P. D., Chuaqui, R. F., Zhuang, Z., Goldstein, S. R., Weiss, R. A. & Liotta, L. A. (1996). Laser capture microdissection. *Science*, **274**, 998-1001.

Engelmann, K., Shen, H. & Finn, O. J. (2008). MCF7 side population cells with characteristics of cancer stem/progenitor cells express the tumor antigen MUC1. *Cancer Res*, **68**, 2419-26.

Erkko, H., Xia, B., Nikkila, J., Schleutker, J., Syrjakoski, K., Mannermaa, A., Kallioniemi, A., Pylkas, K., Karppinen, S. M., Rapakko, K., Miron, A., Sheng, Q., Li, G., Mattila, H., Bell, D. W., Haber, D. A., Grip, M., Reiman, M., Jukkola-Vuorinen, A., Mustonen, A., Kere, J., Aaltonen, L. A., Kosma, V. M., Kataja, V., Soini, Y., Drapkin, R. I., Livingston, D. M. & Winqvist, R. (2007). A recurrent mutation in PALB2 in Finnish cancer families. *Nature*, **446**, 316-9.

Eskelinen, M., Hippelainen, M., Kettunen, J., Salmela, E., Penttila, I. & Alhava, E. (1994). Clinical value of serum tumour markers TPA, TPS, TAG 12, CA 15-3 and MCA in breast cancer diagnosis; results from a prospective study. *Anticancer Res*, **14**, 699-703.

Espina, V., Wulfkuhle, J. D., Calvert, V. S., VanMeter, A., Zhou, W., Coukos, G., Geho, D. H., Petricoin, E. F., 3rd & Liotta, L. A. (2006). Laser-capture microdissection. *Nat Protoc*, **1**, 586-603.

Esteller, M., Corn, P. G., Baylin, S. B. & Herman, J. G. (2001). A gene hypermethylation profile of human cancer. *Cancer Res*, **61**, 3225-9.

Euhus, D. M., Bu, D., Milchgrub, S., Xie, X. J., Bian, A., Leitch, A. M. & Lewis, C. M. (2008). DNA methylation in benign breast epithelium in relation to age and breast cancer risk. *Cancer Epidemiol Biomarkers Prev*, **17**, 1051-9.

Evans, D. G. & Howell, A. (2007). Breast cancer risk-assessment models.

Breast Cancer Res, **9**, 213.

Fang, D., Nguyen, T. K., Leishear, K., Finko, R., Kulp, A. N., Hotz, S., Van Belle, P. A., Xu, X., Elder, D. E. & Herlyn, M. (2005). A tumorigenic subpopulation with stem cell properties in melanomas. *Cancer Res*, **65**, 9328-37.

Farnie, G. & Clarke, R. B. (2007). Mammary stem cells and breast cancer–role of Notch signalling. *Stem Cell Rev*, **3**, 169-75.

Faucher, K., Rabinovitch-Chable, H., Barriere, G., Cook-Moreau, J. & Rigaud, M. (2003). Overexpression of cytosolic glutathione peroxidase (GPX1) delays endothelial cell growth and increases resistance to toxic challenges. *Biochimie*, **85**, 611-7.

Feltkamp, M. C., Smits, H. L., Vierboom, M. P., Minnaar, R. P., de Jongh, B. M., Drijfhout, J. W., ter Schegget, J., Melief, C. J. & Kast, W. M. (1993). Vaccination with cytotoxic T lymphocyte epitope-containing peptide protects against a tumor induced by human papillomavirus type 16-transformed cells. *Eur J Immunol*, **23**, 2242-9.

Fentiman, I. S. (2001). 8. The dilemma of in situ carcinoma of the breast. *Int J Clin Pract*, **55**, 680-3.

Fink, P. J. & Bevan, M. J. (1978). H-2 antigens of the thymus determine lymphocyte specificity. *J Exp Med*, **148**, 766-75.

Fish, R. J. & Kruithof, E. K. (2004). Short-term cytotoxic effects and long-term instability of RNAi delivered using lentiviral vectors. *BMC Mol Biol*, **5**, 9.

Fisk, B., Blevins, T. L., Wharton, J. T. & Ioannides, C. G. (1995). Identification of an immunodominant peptide of HER-2/neu protooncogene recognized by ovarian tumor-specific cytotoxic T lymphocyte lines. *J Exp Med*, **181**, 2109-17.

Flandrin, P., Guyotat, D., Duval, A., Cornillon, J., Tavernier, E., Nadal, N. & Campos, L. (2008). Significance of heat-shock protein (HSP) 90 expression in acute myeloid leukemia cells. *Cell Stress Chaperones*, **13**, 357-64.

Foekens, J. A., Kos, J., Peters, H. A., Krasovec, M., Look, M. P., Cimerman, N., Meijer-van Gelder, M. E., Henzen-Logmans, S. C., van Putten, W. L. & Klijn, J. G. (1998). Prognostic significance of cathepsins B and L in primary human breast cancer. *J Clin Oncol*, **16**, 1013-21.

Foekens, J. A., Wang, Y., Martens, J. W., Berns, E. M. & Klijn, J. G. (2008). The use of genomic tools for the molecular understanding of breast cancer and to guide personalized medicine. *Drug Discov Today*, **13**, 481-7.

Fong, C. J., Sherwood, E. R., Mendelsohn, J., Lee, C. & Kozlowski, J. M. (1992). Epidermal growth factor receptor monoclonal antibody inhibits constitutive receptor phosphorylation, reduces autonomous growth, and sensitizes

androgen-independent prostatic carcinoma cells to tumor necrosis factor alpha. *Cancer Res*, **52**, 5887-92.

Frasor, J., Chang, E. C., Komm, B., Lin, C. Y., Vega, V. B., Liu, E. T., Miller, L. D., Smeds, J., Bergh, J. & Katzenellenbogen, B. S. (2006). Gene expression preferentially regulated by tamoxifen in breast cancer cells and correlations with clinical outcome. *Cancer Res*, **66**, 7334-40.

Friedrich, M., Rafi, L., Tilgen, W., Schmidt, W. & Reichrath, J. (1998). Expression of 1,25-dihydroxy vitamin D3 receptor in breast carcinoma. *J Histochem Cytochem*, **46**, 1335-7.

Fu, C. L., Chuang, Y. H., Huang, H. Y. & Chiang, B. L. (2008). Induction of IL-10 producing CD4+ T cells with regulatory activities by stimulation with IL-10 gene-modified bone marrow derived dendritic cells. *Clin Exp Immunol*, **153**, 258-68.

Fu, Z. J., Ma, Z. Y., Wang, Q. R., Lei, D. P., Wang, R., Liu, C. X. & Pan, X. L. (2008). Overexpression of CyclinD1 and underexpression of p16 correlate with lymph node metastases in laryngeal squamous cell carcinoma in Chinese patients. *Clin Exp Metastasis*, **25**, 887-92.

Gago-Dominguez, M., Jiang, X. & Castelao, J. E. (2007). Lipid peroxidation, oxidative stress genes and dietary factors in breast cancer protection: a hypothesis. *Breast Cancer Res*, **9**, 201.

Galea, M. H., Blamey, R. W., Elston, C. E. & Ellis, I. O. (1992). The Nottingham Prognostic Index in primary breast cancer. *Breast Cancer Res Treat*, **22**, 207-19.

Gao, X., Loggie, B. W. & Nawaz, Z. (2002). The roles of sex steroid receptor coregulators in cancer. *Mol Cancer*, **1**, 7.

Garcia-Bustos, J., Heitman, J. & Hall, M. N. (1991). Nuclear protein localization. *Biochim Biophys Acta*, **1071**, 83-101.

Garcia-Lora, A., Algarra, I. & Garrido, F. (2003). MHC class I antigens, immune surveillance, and tumor immune escape. *J Cell Physiol*, **195**, 346-55.

Garrisi, V. M., Abbate, I., Quaranta, M., Mangia, A., Tommasi, S. & Paradiso, A. (2008). SELDI-TOF serum proteomics and breast cancer: which perspective?. *Expert Rev Proteomics*, **5**, 779-85.

Gastl, G., Spizzo, G., Obrist, P., Dunser, M. & Mikuz, G. (2000). Ep-CAM overexpression in breast cancer as a predictor of survival. *Lancet*, **356**, 1981-2.

Gaujoux, S., Tissier, F., Groussin, L., Libe, R., Ragazzon, B., Launay, P., Audebourg, A., Dousset, B., Bertagna, X. & Bertherat, J. (2008). Wnt/beta-catenin and 3',5'-cyclic adenosine 5'-monophosphate/protein kinase A signaling

pathways alterations and somatic beta-catenin gene mutations in the progression of adrenocortical tumors. *J Clin Endocrinol Metab*, **93**, 4135-40.

Gjertsen, M. K., Buanes, T., Rosseland, A. R., Bakka, A., Gladhaug, I., Soreide, O., Eriksen, J. A., Moller, M., Baksaas, I., Lothe, R. A., Saeterdal, I. & Gaudernack, G. (2001). Intradermal ras peptide vaccination with granulocyte-macrophage colony-stimulating factor as adjuvant: Clinical and immunological responses in patients with pancreatic adenocarcinoma. *Int J Cancer*, **92**, 441-50.

Gerace, L. & Burke, B. (1988). Functional organization of the nuclear envelope. *Annu Rev Cell Biol*, **4**, 335-74.

Gion, M., Mione, R., Leon, A. E. & Dittadi, R. (1999). Comparison of the diagnostic accuracy of CA27.29 and CA15.3 in primary breast cancer. *Clin Chem*, **45**, 630-7.

Glas, A. M., Floore, A., Delahaye, L. J., Witteveen, A. T., Pover, R. C., Bakx, N., Lahti-Domenici, J. S., Bruinsma, T. J., Warmoes, M. O., Bernards, R., Wessels, L. F. & Van't Veer, L. J. (2006). Converting a breast cancer microarray signature into a high-throughput diagnostic test. *BMC Genomics*, **7**, 278.

Godelaine, D., Carrasco, J., Lucas, S., Karanikas, V., Schuler-Thurner, B., Coulie, P. G., Schuler, G., Boon, T. & Van Pel, A. (2003). Polyclonal CTL responses observed in melanoma patients vaccinated with dendritic cells pulsed with a MAGE-3.A1 peptide. *J Immunol*, **171**, 4893-7.

Goulioumis, A. K., Bravou, V., Varakis, J., Goumas, P. & Papadaki, H. (2008). Integrin-linked kinase cytoplasmic and nuclear expression in laryngeal carcinomas. *Virchows Arch*, **453**, 511-9.

Greenbaum, D., Colangelo, C., Williams, K. & Gerstein, M. (2003). Comparing protein abundance and mRNA expression levels on a genomic scale. *Genome Biol*, **4**, 117.

Greene, M. H. (1997). Genetics of breast cancer. *Mayo Clin Proc*, **72**, 54-65.

Grinberg, M., Schwarz, M., Zaltsman, Y., Eini, T., Niv, H., Pietrokovski, S. & Gross, A. (2005). Mitochondrial carrier homolog 2 is a target of tBID in cells signaled to die by tumor necrosis factor alpha. *Mol Cell Biol*, **25**, 4579-90.

Gritzapis, A. D., Mahaira, L. G., Perez, S. A., Cacoullos, N. T., Papamichail, M. & Baxevanis, C. N. (2006). Vaccination with human HER-2/neu (435-443) CTL peptide induces effective antitumor immunity against HER-2/neu-expressing tumor cells in vivo. *Cancer Res*, **66**, 5452-60.

Gudmundsdottir, I., Gunnlaugur Jonasson, J., Sigurdsson, H., Olafsdottir, K., Tryggvadottir, L. & Ogmundsdottir, H. M. (2000). Altered expression of HLA class I antigens in breast cancer: association with prognosis. *Int J Cancer*, **89**, 500-5.

Guinn, B. A., DeBenedette, M. A., Watts, T. H. & Berinstein, N. L. (1999). 4-1BBL cooperates with B7-1 and B7-2 in converting a B cell lymphoma cell line into a long-lasting antitumor vaccine. *J Immunol*, **162**, 5003-10.

Habashy, H. O., Powe, D. G., Staka, C. M., Rakha, E. A., Ball, G., Green, A. R., Aleskandarany, M., Paish, E. C., Douglas Macmillan, R., Nicholson, R. I., Ellis, I. O. & Gee, J. M. (2009). Transferrin receptor (CD71) is a marker of poor prognosis in breast cancer and can predict response to tamoxifen. *Breast Cancer Res Treat*,

Hahne, M., Rimoldi, D., Schroter, M., Romero, P., Schreier, M., French, L. E., Schneider, P., Bornand, T., Fontana, A., Lienard, D., Cerottini, J. & Tschopp, J. (1996). Melanoma cell expression of Fas(Apo-1/CD95) ligand: implications for tumor immune escape. *Science*, **274**, 1363-6.

Haining, W. N., Davies, J., Kanzler, H., Drury, L., Brenn, T., Evans, J., Angelosanto, J., Rivoli, S., Russell, K., George, S., Sims, P., Neuberg, D., Li, X., Kutok, J., Morgan, J., Wen, P., Demetri, G., Coffman, R. L. & Nadler, L. M. (2008). CpG oligodeoxynucleotides alter lymphocyte and dendritic cell trafficking in humans. *Clin Cancer Res*, **14**, 5626-34.

Hall, L., Henney, A., Ralphs, D. N., Herries, D. G. & Craig, R. K. (1986). Changes in gene expression in established human mammary tumor cell lines when compared with normal breast and breast tumor tissue. *Cancer Res*, **46**, 5786-94.

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

Hankinson, S. E., Colditz, G. A. & Willett, W. C. (2004). Towards an integrated model for breast cancer etiology: the lifelong interplay of genes, lifestyle, and hormones. *Breast Cancer Res*, **6**, 213-8.

Hartmann, L. C., Schaid, D. J., Woods, J. E., Crotty, T. P., Myers, J. L., Arnold, P. G., Petty, P. M., Sellers, T. A., Johnson, J. L., McDonnell, S. K., Frost, M. H. & Jenkins, R. B. (1999). Efficacy of bilateral prophylactic mastectomy in women with a family history of breast cancer. *N Engl J Med*, **340**, 77-84.

Hasan, J. & Jayson, G. C. (2001). VEGF antagonists. *Expert Opin Biol Ther*, **1**, 703-18.

Hayflick, L. (1965). The Limited in Vitro Lifetime of Human Diploid Cell Strains. *Exp Cell Res*, **37**, 614-36.

Henry, N. L. & Hayes, D. F. (2006). Uses and abuses of tumor markers in the diagnosis, monitoring, and treatment of primary and metastatic breast cancer. *Oncologist*, **11**, 541-52.

Heuer, J. G., Tucker-McClung, C., Gonin, R. & Hock, R. A. (1996). Retrovirus-

mediated gene transfer of B7-1 and MHC class II converts a poorly immunogenic neuroblastoma into a highly immunogenic one. *Hum Gene Ther*, **7**, 2059-68.

Hicks, G. R. & Raikhel, N. V. (1995). Protein import into the nucleus: an integrated view. *Annu Rev Cell Dev Biol*, **11**, 155-88.

Hildenbrand, R. & Schaaf, A. (2009). The urokinase-system in tumor tissue stroma of the breast and breast cancer cell invasion. *Int J Oncol*, **34**, 15-23.

Ho, M. M., Ng, A. V., Lam, S. & Hung, J. Y. (2007). Side population in human lung cancer cell lines and tumors is enriched with stem-like cancer cells. *Cancer Res*, **67**, 4827-33.

Holmes, M. D. & Willett, W. C. (2004). Does diet affect breast cancer risk?. *Breast Cancer Res*, **6**, 170-8.

Hong, A., Davies, S. & Lee, C. S. (2003). Immunohistochemical detection of the human herpes virus 8 (HHV8) latent nuclear antigen-1 in Kaposi's sarcoma. *Pathology*, **35**, 448-50.

Hong, S. W., Kim, C. J., Park, W. S., Shin, J. S., Lee, S. D., Ko, S. G., Jung, S. I., Park, I. C., An, S. K., Lee, W. K., Lee, W. J., Jin, D. H. & Lee, M. S. (2009). p34SEI-1 inhibits apoptosis through the stabilization of the X-linked inhibitor of apoptosis protein: p34SEI-1 as a novel target for anti-breast cancer strategies. *Cancer Res*, **69**, 741-6.

Hoption Cann, S.A., van Netten, J.P., van Netten, C. And Glover, D.W. (2002). Spontaneous regression: a hidden treasure buried in time. *Med Hypotheses*, **58**, 115-119.

Horig, H., Medina, F. A., Conkright, W. A. & Kaufman, H. L. (2000). Strategies for cancer therapy using carcinoembryonic antigen vaccines. *Expert Rev Mol Med*, **2**, 1-24.

Horiguchi, Y., Nukaya, I., Okazawa, K., Kawashima, I., Fikes, J., Sette, A., Tachibana, M., Takesako, K. & Murai, M. (2002). Screening of HLA-A24-restricted epitope peptides from prostate-specific membrane antigen that induce specific antitumor cytotoxic T lymphocytes. *Clin Cancer Res*, **8**, 3885-92.

Hossain, A., Kuo, M. T. & Saunders, G. F. (2006). Mir-17-5p regulates breast cancer cell proliferation by inhibiting translation of AIB1 mRNA. *Mol Cell Biol*, **26**, 8191-201.

Hu, Y., Zhang, S., Yu, J., Liu, J. & Zheng, S. (2005). SELDI-TOF-MS: the proteomics and bioinformatics approaches in the diagnosis of breast cancer. *Breast*, **14**, 250-5.

Huang, Z., Hankinson, S. E., Colditz, G. A., Stampfer, M. J., Hunter, D. J., Manson, J. E., Hennekens, C. H., Rosner, B., Speizer, F. E. & Willett, W. C. (1997). Dual effects of weight and weight gain on breast cancer risk. *Jama*, **278**,

1407-11.

Huang, Y., Jin, Y., Yan, C. H., Yu, Y., Bai, J., Chen, F., Zhao, Y. Z. & Fu, S. B. (2008). Involvement of Annexin A2 in p53 induced apoptosis in lung cancer. *Mol Cell Biochem*, **309**, 117-23.

Huang, C., Li, M., Chen, C. & Yao, Q. (2008). Small interfering RNA therapy in cancer: mechanism, potential targets, and clinical applications. *Expert Opin Ther Targets*, **12**, 637-45.

Hueman, M. T., Stojadinovic, A., Storrer, C. E., Dehqanzada, Z. A., Gurney, J. M., Shriver, C. D., Ponniah, S. & Peoples, G. E. (2007). Analysis of naive and memory CD4 and CD8 T cell populations in breast cancer patients receiving a HER2/neu peptide (E75) and GM-CSF vaccine. *Cancer Immunol Immunother*, **56**, 135-46.

Hughes-Davies, L., Huntsman, D., Ruas, M., Fuks, F., Bye, J., Chin, S. F., Milner, J., Brown, L. A., Hsu, F., Gilks, B., Nielsen, T., Schulzer, M., Chia, S., Ragaz, J., Cahn, A., Linger, L., Ozdag, H., Cattaneo, E., Jordanova, E. S., Schuuring, E., Yu, D. S., Venkitaraman, A., Ponder, B., Doherty, A., Aparicio, S., Bentley, D., Theillet, C., Ponting, C. P., Caldas, C. & Kouzarides, T. (2003). EMSY links the BRCA2 pathway to sporadic breast and ovarian cancer. *Cell*, **115**, 523-35.

Hunt, K. K. & Keyomarsi, K. (2005). Cyclin E as a prognostic and predictive marker in breast cancer. *Semin Cancer Biol*, **15**, 319-26.

Huttenhofer, A., Schattner, P. & Polacek, N. (2005). Non-coding RNAs: hope or hype?. *Trends Genet*, **21**, 289-97.

Inaba, K., Inaba, M., Romani, N., Aya, H., Deguchi, M., Ikehara, S., Muramatsu, S. & Steinman, R. M. (1992). Generation of large numbers of dendritic cells from mouse bone marrow cultures supplemented with granulocyte/macrophage colony-stimulating factor. *J Exp Med*, **176**, 1693-702.

Inokuchi, J., Narula, N., Yee, D. S., Skarecky, D. W., Lau, A., Ornstein, D. K. & Tyson, D. R. (2009). Annexin A2 positively contributes to the malignant phenotype and secretion of IL-6 in DU145 prostate cancer cells. *Int J Cancer*, **124**, 68-74.

Iorio, M. V., Ferracin, M., Liu, C. G., Veronese, A., Spizzo, R., Sabbioni, S., Magri, E., Pedriali, M., Fabbri, M., Campiglio, M., Menard, S., Palazzo, J. P., Rosenberg, A., Musiani, P., Volinia, S., Nenci, I., Calin, G. A., Querzoli, P., Negrini, M. & Croce, C. M. (2005). MicroRNA gene expression deregulation in human breast cancer. *Cancer Res*, **65**, 7065-70.

Irvin, W. J., Jr. & Carey, L. A. (2008). What is triple-negative breast cancer?. *Eur J Cancer*, **44**, 2799-805.

Jager, D., Filonenko, V., Gout, I., Frosina, D., Eastlake-Wade, S., Castelli,

S., Varga, Z., Moch, H., Chen, Y. T., Busam, K. J., Seil, I., Old, L. J., Nissan, A., Frei, C., Gure, A. O., Knuth, A. & Jungbluth, A. A. (2007). NY-BR-1 is a differentiation antigen of the mammary gland. *Appl Immunohistochem Mol Morphol*, **15**, 77-83.

Jager, D., Karbach, J., Pauligk, C., Seil, I., Frei, C., Chen, Y. T., Old, L. J., Knuth, A. & Jager, E. (2005). Humoral and cellular immune responses against the breast cancer antigen NY-BR-1: definition of two HLA-A2 restricted peptide epitopes. *Cancer Immun*, **5**, 11.

Jager, D. & Knuth, A. (2005). Antibodies and vaccines–hope or illusion?. *Breast*, **14**, 631-5.

Jager, E., Nagata, Y., Gnjatic, S., Wada, H., Stockert, E., Karbach, J., Dunbar, P. R., Lee, S. Y., Jungbluth, A., Jager, D., Arand, M., Ritter, G., Cerundolo, V., Dupont, B., Chen, Y. T., Old, L. J. & Knuth, A. (2000). Monitoring CD8 T cell responses to NY-ESO-1: correlation of humoral and cellular immune responses. *Proc Natl Acad Sci U S A*, **97**, 4760-5.

Jager, D., Stockert, E., Gure, A. O., Scanlan, M. J., Karbach, J., Jager, E., Knuth, A., Old, L. J. & Chen, Y. T. (2001). Identification of a tissue-specific putative transcription factor in breast tissue by serological screening of a breast cancer library. *Cancer Res*, **61**, 2055-61.

James, C. R., Quinn, J. E., Mullan, P. B., Johnston, P. G. & Harkin, D. P. (2007). BRCA1, a potential predictive biomarker in the treatment of breast cancer. *Oncologist*, **12**, 142-50.

Janetzki, S., Palla, D., Rosenhauer, V., Lochs, H., Lewis, J. J. & Srivastava, P. K. (2000). Immunization of cancer patients with autologous cancer-derived heat shock protein gp96 preparations: a pilot study. *Int J Cancer*, **88**, 232-8.

Janeway, C. (1991). Immune recognition. Mls: makes a little sense. *Nature*, **349**, 459-61.

Jiang, W., Li, X., Rao, S., Wang, L., Du, L., Li, C., Wu, C., Wang, H., Wang, Y. & Yang, B. (2008). Constructing disease-specific gene networks using pair-wise relevance metric: application to colon cancer identifies interleukin 8, desmin and enolase 1 as the central elements. *BMC Syst Biol*, **2**, 72.

Jin, R., Chow, V. T., Tan, P. H., Dheen, S. T., Duan, W. & Bay, B. H. (2002). Metallothionein 2A expression is associated with cell proliferation in breast cancer. *Carcinogenesis*, **23**, 81-6.

Jin, R., Huang, J., Tan, P. H. & Bay, B. H. (2004). Clinicopathological significance of metallothioneins in breast cancer. *Pathol Oncol Res*, **10**, 74-9.

Jones, C., Mackay, A., Grigoriadis, A., Cossu, A., Reis-Filho, J. S., Fulford, L., Dexter, T., Davies, S., Bulmer, K., Ford, E., Parry, S., Budroni, M., Palmieri, G., Neville, A. M., O'Hare, M. J. & Lakhani, S. R. (2004). Expression

profiling of purified normal human luminal and myoepithelial breast cells: identification of novel prognostic markers for breast cancer. *Cancer Res*, **64**, 3037-45.

June, C. H. (2007). Adoptive T cell therapy for cancer in the clinic. *J Clin Invest*, **117**, 1466-76.

Kalkanidis, M., Pietersz, G. A., Xiang, S. D., Mottram, P. L., Crimeen-Irwin, B., Ardipradja, K. & Plebanski, M. (2006). Methods for nano-particle based vaccine formulation and evaluation of their immunogenicity. *Methods*, **40**, 20-9.

Kawaida, H., Kono, K., Takahashi, A., Sugai, H., Mimura, K., Miyagawa, N., Omata, H., Ooi, A. & Fujii, H. (2005). Distribution of CD4+CD25high regulatory T-cells in tumor-draining lymph nodes in patients with gastric cancer. *J Surg Res*, **124**, 151-7.

Kessar, P., Perry, N., Vinnicombe, S. J., Hussain, H. K., Carpenter, R. & Wells, C. A. (2002). How significant is detection of ductal carcinoma in situ in a breast screening programme?. *Clin Radiol*, **57**, 807-14.

Key, T. J., Verkasalo, P. K. & Banks, E. (2001). Epidemiology of breast cancer. *Lancet Oncol*, **2**, 133-40.

Keyomarsi, K., Tucker, S. L., Buchholz, T. A., Callister, M., Ding, Y., Hortobagyi, G. N., Bedrosian, I., Knickerbocker, C., Toyofuku, W., Lowe, M., Herliczek, T. W. & Bacus, S. S. (2002). Cyclin E and survival in patients with breast cancer. *N Engl J Med*, **347**, 1566-75.

Khong, H. T. & Restifo, N. P. (2002). Natural selection of tumor variants in the generation of "tumor escape" phenotypes. *Nat Immunol*, **3**, 999-1005.

Kim, H. O., Jo, Y. H., Lee, J., Lee, S. S. & Yoon, K. S. (2008). The C1772T genetic polymorphism in human HIF-1alpha gene associates with expression of HIF-1alpha protein in breast cancer. *Oncol Rep*, **20**, 1181-7.

King, E. B., Chew, K. L., Petrakis, N. L. & Ernster, V. L. (1983). Nipple aspirate cytology for the study of breast cancer precursors. *J Natl Cancer Inst*, **71**, 1115-21.

Kleer, C. G., van Golen, K. L., Zhang, Y., Wu, Z. F., Rubin, M. A. & Merajver, S. D. (2002). Characterization of RhoC expression in benign and malignant breast disease: a potential new marker for small breast carcinomas with metastatic ability. *Am J Pathol*, **160**, 579-84.

Klein, R. M., Spofford, L. S., Abel, E. V., Ortiz, A. & Aplin, A. E. (2008). B-RAF regulation of Rnd3 participates in actin cytoskeletal and focal adhesion organization. *Mol Biol Cell*, **19**, 498-508.

Knudson, A. G., Jr. (1971). Mutation and cancer: statistical study of retinoblas-

toma. *Proc Natl Acad Sci U S A*, **68**, 820-3.

Knutson, K. L., Schiffman, K. & Disis, M. L. (2001). Immunization with a HER-2/neu helper peptide vaccine generates HER-2/neu CD8 T-cell immunity in cancer patients. *J Clin Invest*, **107**, 477-84.

Knutson, D. & Steiner, E. (2007). Screening for breast cancer: current recommendations and future directions. *Am Fam Physician*, **75**, 1660-6.

Koenig, S., Woods, R. M., Brewah, Y. A., Newell, A. J., Jones, G. M., Boone, E., Adelsberger, J. W., Baseler, M. W., Robinson, S. M. & Jacobson, S. (1993). Characterization of MHC class I restricted cytotoxic T cell responses to tax in HTLV-1 infected patients with neurologic disease. *J Immunol*, **151**, 3874-83.

Kok, M., Linn, S. C., Van Laar, R. K., Jansen, M. P., van den Berg, T. M., Delahaye, L. J., Glas, A. M., Peterse, J. L., Hauptmann, M., Foekens, J. A., Klijn, J. G., Wessels, L. F., Van't Veer, L. J. & Berns, E. M. (2009). Comparison of gene expression profiles predicting progression in breast cancer patients treated with tamoxifen. *Breast Cancer Res Treat*, **113**, 275-83.

Kopp, E. & Medzhitov, R. (2003). Recognition of microbial infection by Toll-like receptors. *Curr Opin Immunol*, **15**, 396-401.

Koukourakis, M. I., Giatromanolaki, A. & Sivridis, E. (2003). Lactate dehydrogenase isoenzymes 1 and 5: differential expression by neoplastic and stromal cells in non-small cell lung cancer and other epithelial malignant tumors. *Tumour Biol*, **24**, 199-202.

Kreitman, R. J. (2001). Toxin-labeled monoclonal antibodies. *Curr Pharm Biotechnol*, **2**, 313-25.

Kronig, H., Hofer, K., Conrad, H., Guilaume, P., Muller, J., Schiemann, M., Lennerz, V., Cosma, A., Peschel, C., Busch, D. H., Romero, P. & Bernhard, H. (2009). Allorestricted T lymphocytes with a high avidity T-cell receptor towards NY-ESO-1 have potent anti-tumor activity. *Int J Cancer*, **125**, 649-55.

Kyte, J. A. & Gaudernack, G. (2006). Immuno-gene therapy of cancer with tumour-mRNA transfected dendritic cells. *Cancer Immunol Immunother*, **55**, 1432-42.

Lagace, M., Xuan, J. Y., Young, S. S., McRoberts, C., Maier, J., Rajcan-Separovic, E. & Korneluk, R. G. (2001). Genomic organization of the X-linked inhibitor of apoptosis and identification of a novel testis-specific transcript. *Genomics*, **77**, 181-8.

Lancashire, L. J., Lemetre, C. & Ball, G. R. (2009). An introduction to artificial neural networks in bioinformatics–application to complex microarray and mass spectrometry datasets in cancer studies. *Brief Bioinform*, **10**, 315-29.

Lapteva, N., Seethammagari, M. R., Hanks, B. A., Jiang, J., Levitt, J. M., Slawin, K. M. & Spencer, D. M. (2007). Enhanced activation of human dendritic cells by inducible CD40 and Toll-like receptor-4 ligation. *Cancer Res*, **67**, 10528-37.

Laronga, C. & Drake, R. R. (2007). Proteomic approach to breast cancer. *Cancer Control*, **14**, 360-8.

Laversin, S.A., Miles, A.K., Ball, G.R.& Rees, R.C. (2008). Emerging Breast Cancer Biomarkers. *Current Cancer Therapy Reviews*, **4**, 79-85.

Lee, J. W. & Cho, M. K. (2008). Phenethyl isothiocyanate induced apoptosis via down regulation of Bcl-2/XIAP and triggering of the mitochondrial pathway in MCF-7 cells. *Arch Pharm Res*, **31**, 1604-12.

Lehmann, C., Zeis, M., Schmitz, N. & Uharek, L. (2000). Impaired binding of perforin on the surface of tumor cells is a cause of target cell resistance against cytotoxic effector cells. *Blood*, **96**, 594-600.

Leitner, W. W., Ying, H. & Restifo, N. P. (1999). DNA and RNA-based vaccines: principles, progress and prospects. *Vaccine*, **18**, 765-77.

Leitzmann, M. F., Moore, S. C., Peters, T. M., Lacey, J. V., Jr., Schatzkin, A., Schairer, C., Brinton, L. A. & Albanes, D. (2008). Prospective study of physical activity and risk of postmenopausal breast cancer. *Breast Cancer Res*, **10**, R92.

Lerebours, F. & Lidereau, R. (2002). Molecular alterations in sporadic breast cancer. *Crit Rev Oncol Hematol*, **44**, 121-41.

Levenson, V. V. (2007). Biomarkers for early detection of breast cancer: what, when, and where?. *Biochim Biophys Acta*, **1770**, 847-56.

Levine, A. J. (1997). p53, the cellular gatekeeper for growth and division. *Cell*, **88**, 323-31.

Li, G., Miles, A., Line, A. & Rees, R. C. (2004). Identification of tumour antigens by serological analysis of cDNA expression cloning. *Cancer Immunol Immunother*, **53**, 139-43.

Li, G., Assudani, D. P., Line, A., Cao, F., Miles, A., Rees, R. C. & McArdle, S. E. B. (2008). Identification of Metastasis Associated Antigen 1 (MTA1) by Serological Screening of Prostate Cancer cDNA Libraries. *Open Biochem J*, **2**, 100-7.

Li, J., Zhang, Z., Rosenzweig, J., Wang, Y. Y. & Chan, D. W. (2002). Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer. *Clin Chem*, **48**, 1296-304.

Limer, J. L. & Speirs, V. (2004). Phyto-oestrogens and breast cancer chemoprevention. *Breast Cancer Res*, **6**, 119-27.

Lin, S. S., Bassik, M. C., Suh, H., Nishino, M., Arroyo, J. D., Hahn, W. C., Korsmeyer, S. J. & Roberts, T. M. (2006). PP2A regulates BCL-2 phosphorylation and proteasome-mediated degradation at the endoplasmic reticulum. *J Biol Chem*, **281**, 23003-12.

Lin, M. T., Chang, C. C., Chen, S. T., Chang, H. L., Su, J. L., Chau, Y. P. & Kuo, M. L. (2004). Cyr61 expression confers resistance to apoptosis in breast cancer MCF-7 cells by a mechanism of NF-kappaB-dependent XIAP up-regulation. *J Biol Chem*, **279**, 24015-23.

Linette, G. P., Zhang, D., Hodi, F. S., Jonasch, E. P., Longerich, S., Stowell, C. P., Webb, I. J., Daley, H., Soiffer, R. J., Cheung, A. M., Eapen, S. G., Fee, S. V., Rubin, K. M., Sober, A. J. & Haluska, F. G. (2005). Immunization using autologous dendritic cells pulsed with the melanoma-associated antigen gp100-derived G280-9V peptide elicits CD8+ immunity. *Clin Cancer Res*, **11**, 7692-9.

Liu, M., Acres, B., Balloul, J. M., Bizouarne, N., Paul, S., Slos, P. & Squiban, P. (2004). Gene-based vaccines and immunotherapeutics. *Proc Natl Acad Sci U S A*, **101 Suppl 2**, 14567-71.

Liu, T. J., el-Naggar, A. K., McDonnell, T. J., Steck, K. D., Wang, M., Taylor, D. L. & Clayman, G. L. (1995). Apoptosis induction mediated by wild-type p53 adenoviral gene transfer in squamous cell carcinoma of the head and neck. *Cancer Res*, **55**, 3117-22.

Loi, S. (2008). Molecular analysis of hormone receptor positive (luminal) breast cancers: what have we learnt?. *Eur J Cancer*, **44**, 2813-8.

Lowery, A. J., Miller, N., McNeill, R. E. & Kerin, M. J. (2008). MicroRNAs as prognostic indicators and therapeutic targets: potential effect on breast cancer management. *Clin Cancer Res*, **14**, 360-5.

Lu, J., Getz, G., Miska, E. A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B. L., Mak, R. H., Ferrando, A. A., Downing, J. R., Jacks, T., Horvitz, H. R. & Golub, T. R. (2005). MicroRNA expression profiles classify human cancers. *Nature*, **435**, 834-8.

Ludwig, J. A. & Weinstein, J. N. (2005). Biomarkers in cancer staging, prognosis and treatment selection. *Nat Rev Cancer*, **5**, 845-56.

Lukas, J., Parry, D., Aagaard, L., Mann, D. J., Bartkova, J., Strauss, M., Peters, G. & Bartek, J. (1995). Retinoblastoma-protein-dependent cell-cycle inhibition by the tumour suppressor p16. *Nature*, **375**, 503-6.

Ma, Z., Li, J., He, F., Wilson, A., Pitt, B. & Li, S. (2005). Cationic lipids enhance siRNA-mediated interferon response in mice. *Biochem Biophys Res Commun*, **330**, 755-9.

Maass, N., Nagasaki, K., Ziebart, M., Mundhenke, C. & Jonat, W. (2002). Expression and regulation of tumor suppressor gene maspin in breast cancer. *Clin Breast Cancer*, **3**, 281-7.

Maher, S., Toomey, D., Condron, C. & Bouchier-Hayes, D. (2002). Activation-induced cell death: the controversial role of Fas and Fas ligand in immune privilege and tumour counterattack. *Immunol Cell Biol*, **80**, 131-7.

Marchand, M., van Baren, N., Weynants, P., Brichard, V., Dreno, B., Tessier, M. H., Rankin, E., Parmiani, G., Arienti, F., Humblet, Y., Bourlond, A., Vanwijck, R., Lienard, D., Beauduin, M., Dietrich, P. Y., Russo, V., Kerger, J., Masucci, G., Jager, E., De Greve, J., Atzpodien, J., Brasseur, F., Coulie, P. G., van der Bruggen, P. & Boon, T. (1999). Tumor regressions observed in patients with metastatic melanoma treated with an antigenic peptide encoded by gene MAGE-3 and presented by HLA-A1. *Int J Cancer*, **80**, 219-30.

Martelange, V., De Smet, C., De Plaen, E., Lurquin, C. & Boon, T. (2000). Identification on a human sarcoma of two new genes with tumor-specific expression. *Cancer Res*, **60**, 3848-55.

Martin, G. S. (1970). Rous sarcoma virus: a function required for the maintenance of the transformed state. *Nature*, **227**, 1021-3.

Martin, G. S. (2004). The road to Src. *Oncogene*, **23**, 7910-7.

Matharoo-Ball, B., Ratcliff, L., Lancashire, L., Ugurel, S., Miles, A.K., Weston, D.J., Rees, R., Schadendorf, D., Ball, G. and Creaser, C.S. (2007). Diagnostic biomarkers differentiating metastatic melanoma patients from healthy controls identified by an integrated MALDI-TOF mass spectrometry/bioinformatic approach. *Proteomics Clin. Appl.*, **1**, 605-620.

Mathieu, M. G., Knights, A. J., Pawelec, G., Riley, C. L., Wernet, D., Lemonnier, F. A., Straten, P. T., Mueller, L., Rees, R. C. & McArdle, S. E. (2007). HAGE, a cancer/testis antigen with potential for melanoma immunotherapy: identification of several MHC class I/II HAGE-derived immunogenic peptides. *Cancer Immunol Immunother*, **56**, 1885-95.

Mattie, M. D., Benz, C. C., Bowers, J., Sensinger, K., Wong, L., Scott, G. K., Fedele, V., Ginzinger, D., Getts, R. & Haqq, C. (2006). Optimized high-throughput microRNA expression profiling provides novel biomarker assessment of clinical prostate and breast cancer biopsies. *Mol Cancer*, **5**, 24.

McKeage, K. & Perry, C. M. (2002). Trastuzumab: a review of its use in the treatment of metastatic breast cancer overexpressing HER2. *Drugs*, **62**, 209-43.

McKenna, N. J., Lanz, R. B. & O'Malley, B. W. (1999). Nuclear receptor coregulators: cellular and molecular biology. *Endocr Rev*, **20**, 321-44.

Medema, J. P., de Jong, J., Peltenburg, L. T., Verdegaal, E. M., Gorter, A.,

Bres, S. A., Franken, K. L., Hahne, M., Albar, J. P., Melief, C. J. & Offringa, R. (2001). Blockade of the granzyme B/perforin pathway through overexpression of the serine protease inhibitor PI-9/SPI-6 constitutes a mechanism for immune escape by tumors. *Proc Natl Acad Sci U S A*, **98**, 11515-20.

Meijnen, P., Peterse, J. L., Antonini, N., Rutgers, E. J. & van de Vijver, M. J. (2008). Immunohistochemical categorisation of ductal carcinoma in situ of the breast. *Br J Cancer*, **98**, 137-42.

Menard, S., Casalini, P., Campiglio, M., Pupa, S., Agresti, R. & Tagliabue, E. (2001). HER2 overexpression in various tumor types, focussing on its relationship to the development of invasive breast cancer. *Ann Oncol*, **12 Suppl 1**, S15-9.

Messaoudi, S., Peyrat, J. F., Brion, J. D. & Alami, M. (2008). Recent advances in Hsp90 inhibitors as antitumor agents. *Anticancer Agents Med Chem*, **8**, 761-82.

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

Miles, A. K., Matharoo-Ball, B., Li, G., Ahmad, M. & Rees, R. C. (2006). The identification of human tumour antigens: Current status and future developments. *Cancer Immunol Immunother*, **55**, 996-1003.

Miles, A. K., Rogers, A., Li, G., Seth, R., Powe, D., McArdle, S. E., McCulloch, T. A., Bishop, M. C. & Rees, R. C. (2007). Identification of a novel prostate cancer-associated tumor antigen. *Prostate*, **67**, 274-87.

Mittendorf, E. A., Gurney, J. M., Storrer, C. E., Shriver, C. D., Ponniah, S. & Peoples, G. E. (2006). Vaccination with a HER2/neu peptide induces intra- and inter-antigenic epitope spreading in patients with early stage breast cancer. *Surgery*, **139**, 407-18.

Mitsutake, N., Iwao, A., Nagai, K., Namba, H., Ohtsuru, A., Saenko, V. & Yamashita, S. (2007). Characterization of side population in thyroid cancer cell lines: cancer stem-like cells are enriched partly but not exclusively. *Endocrinology*, **148**, 1797-803.

Miyake, K., Utsugisawa, T., Flygare, J., Kiefer, T., Hamaguchi, I., Richter, J. & Karlsson, S. (2008). Ribosomal protein S19 deficiency leads to reduced proliferation and increased apoptosis but does not affect terminal erythroid differentiation in a cell line model of Diamond-Blackfan anemia. *Stem Cells*, **26**, 323-9.

Mo, R., Rao, S. M. & Zhu, Y. J. (2006). Identification of the MLL2 complex as a coactivator for estrogen receptor alpha. *J Biol Chem*, **281**, 15714-20.

Mohanty, K., Saha, A., Pal, S., Mallick, P., Chatterjee, S. K., Foon, K. A. & Bhattacharya-Chatterjee, M. (2007). Anti-tumor immunity induced by an

anti-idiotype antibody mimicking human Her-2/neu. *Breast Cancer Res Treat*, **104**, 1-11.

Moretta, L., Mingari, M. C., Pende, D., Bottino, C., Biassoni, R. & Moretta, A. (1996). The molecular basis of natural killer (NK) cell recognition and function. *J Clin Immunol*, **16**, 243-53.

Morgan, R. A., Dudley, M. E., Yu, Y. Y., Zheng, Z., Robbins, P. F., Theoret, M. R., Wunderlich, J. R., Hughes, M. S., Restifo, N. P. & Rosenberg, S. A. (2003). High efficiency TCR gene transfer into primary human lymphocytes affords avid recognition of melanoma tumor antigen glycoprotein 100 and does not alter the recognition of autologous melanoma antigens. *J Immunol*, **171**, 3287-95.

Morrison, B. J., Schmidt, C. W., Lakhani, S. R., Reynolds, B. A. & Lopez, J. A. (2008). Breast cancer stem cells: implications for therapy of breast cancer. *Breast Cancer Res*, **10**, 210.

Murray, R. J., Kurilla, M. G., Brooks, J. M., Thomas, W. A., Rowe, M., Kieff, E. & Rickinson, A. B. (1992). Identification of target antigens for the human cytotoxic T cell response to Epstein-Barr virus (EBV): implications for the immune control of EBV-positive malignancies. *J Exp Med*, **176**, 157-68.

Nadano, D., Aoki, C., Yoshinaka, T., Irie, S. & Sato, T. A. (2001). Electrophoretic characterization of ribosomal subunits and proteins in apoptosis: specific downregulation of S11 in staurosporine-treated human breast carcinoma cells. *Biochemistry*, **40**, 15184-93.

Nojima, H. (2004). G1 and S-phase checkpoints, chromosome instability, and cancer. *Methods Mol Biol*, **280**, 3-49.

O'Brien, N., Maguire, T. M., O'Donovan, N., Lynch, N., Hill, A. D., McDermott, E., O'Higgins, N. & Duffy, M. J. (2002). Mammaglobin a: a promising marker for breast cancer. *Clin Chem*, **48**, 1362-4.

O'Connell, J., Bennett, M. W., O'Sullivan, G. C., O'Callaghan, J., Collins, J. K. & Shanahan, F. (1999). Expression of Fas (CD95/APO-1) ligand by human breast cancers: significance for tumor immune privilege. *Clin Diagn Lab Immunol*, **6**, 457-63.

Ohkia, A., Hu, Y., Wang, M., Garcia, F. U. & Stearns, M. E. (2004). Evidence for prostate cancer-associated diagnostic marker-1: immunohistochemistry and in situ hybridization studies. *Clin Cancer Res*, **10**, 2452-8.

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

Olempska, M., Eisenach, P. A., Ammerpohl, O., Ungefroren, H., Fandrich, F. & Kalthoff, H. (2007). Detection of tumor stem cell markers in pancreatic

carcinoma cell lines. *Hepatobiliary Pancreat Dis Int*, **6**, 92-7.

Oppitz, M., Busch, C., Garbe, C. & Drews, U. (2008). Distribution of muscarinic receptor subtype M3 in melanomas and their metastases. *J Cutan Pathol*, **35**, 809-15.

Ostrand-Rosenberg, S. (2004). Animal models of tumor immunity, immunotherapy and cancer vaccines. *Curr Opin Immunol*, **16**, 143-50.

Paik, S., Shak, S., Tang, G., Kim, C., Baker, J., Cronin, M., Baehner, F. L., Walker, M. G., Watson, D., Park, T., Hiller, W., Fisher, E. R., Wickerham, D. L., Bryant, J. & Wolmark, N. (2004). A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med*, **351**, 2817-26.

Palma, M., Ristori, E., Ricevuto, E., Giannini, G. & Gulino, A. (2006). BRCA1 and BRCA2: the genetic testing and the current management options for mutation carriers. *Crit Rev Oncol Hematol*, **57**, 1-23.

Palucka, A. K., Ueno, H., Connolly, J., Kerneis-Norvell, F., Blanck, J. P., Johnston, D. A., Fay, J. & Banchereau, J. (2006). Dendritic cells loaded with killed allogeneic melanoma cells can induce objective clinical responses and MART-1 specific CD8+ T-cell immunity. *J Immunother*, **29**, 545-57.

Parkin, D. M., Bray, F., Ferlay, J. & Pisani, P. (2005). Global cancer statistics, 2002. *CA Cancer J Clin*, **55**, 74-108.

Parmiani, G., Rodolfo, M. & Melani, C. (2000). Immunological gene therapy with ex vivo gene-modified tumor cells: a critique and a reappraisal. *Hum Gene Ther*, **11**, 1269-75.

Paschen, A., Mendez, R. M., Jimenez, P., Sucker, A., Ruiz-Cabello, F., Song, M., Garrido, F. & Schadendorf, D. (2003). Complete loss of HLA class I antigen expression on melanoma cells: a result of successive mutational events. *Int J Cancer*, **103**, 759-67.

Pellikainen, J. M. & Kosma, V. M. (2007). Activator protein-2 in carcinogenesis with a special reference to breast cancer–a mini review. *Int J Cancer*, **120**, 2061-7.

Perou, C. M., Jeffrey, S. S., van de Rijn, M., Rees, C. A., Eisen, M. B., Ross, D. T., Pergamenschikov, A., Williams, C. F., Zhu, S. X., Lee, J. C., Lashkari, D., Shalon, D., Brown, P. O. & Botstein, D. (1999). Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. *Proc Natl Acad Sci U S A*, **96**, 9212-7.

Perou, C. M., Sorlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L. A., Fluge, O., Pergamenschikov, A., Williams, C., Zhu, S. X., Lonning, P. E., Borresen-Dale, A. L., Brown, P. O. & Botstein, D. (2000). Molecular portraits of human breast tumours. *Nature*, **406**, 747-52.

Phillips, T. M., McBride, W. H. & Pajonk, F. (2006). The response of CD24(-/low)/CD44+ breast cancer-initiating cells to radiation. *J Natl Cancer Inst*, **98**, 1777-85.

Pignatelli, M., Cortes-Canteli, M., Lai, C., Santos, A. & Perez-Castillo, A. (2001). The peroxisome proliferator-activated receptor gamma is an inhibitor of ErbBs activity in human breast cancer cells. *J Cell Sci*, **114**, 4117-26.

Pinilla-Ibarz, J., Cathcart, K. & Scheinberg, D. A. (2000). CML vaccines as a paradigm of the specific immunotherapy of cancer. *Blood Rev*, **14**, 111-20.

Polyak, K. (2007). Breast cancer: origins and evolution. *J Clin Invest*, **117**, 3155-63.

Pogue-Geile, K., Geiser, J. R., Shu, M., Miller, C., Wool, I. G., Meisler, A. I. & Pipas, J. M. (1991). Ribosomal protein genes are overexpressed in colorectal cancer: isolation of a cDNA clone encoding the human S3 ribosomal protein. *Mol Cell Biol*, **11**, 3842-9.

Ponzielli, R., Katz, S., Barsyte-Lovejoy, D. & Penn, L. Z. (2005). Cancer therapeutics: targeting the dark side of Myc. *Eur J Cancer*, **41**, 2485-501.

Pusztai, L., Cristofanilli, M. & Paik, S. (2007). New generation of molecular prognostic and predictive tests for breast cancer. *Semin Oncol*, **34**, S10-6.

Qi, W., Liu, X., Chen, W., Li, Q. & Martinez, J. D. (2007). Overexpression of 14-3-3gamma causes polyploidization in H322 lung cancer cells. *Mol Carcinog*, **46**, 847-56.

Quinn, M., Babb, P., Brock, A., Kirkby, L. & Jones, J. (2001). Cancer Trends in England and Wales 1950-1999. Office for National Statistics. http://www.statistics.gov.uk/downloads/theme_health/cancertrends_5099.pdf.

Rae, J. M., Creighton, C. J., Meck, J. M., Haddad, B. R. & Johnson, M. D. (2007). MDA-MB-435 cells are derived from M14 melanoma cells–a loss for breast cancer, but a boon for melanoma research. *Breast Cancer Res Treat*, **104**, 13-9.

Rahman, N., Seal, S., Thompson, D., Kelly, P., Renwick, A., Elliott, A., Reid, S., Spanova, K., Barfoot, R., Chagtai, T., Jayatilake, H., McGuffog, L., Hanks, S., Evans, D. G., Eccles, D., Easton, D. F. & Stratton, M. R. (2007). PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nat Genet*, **39**, 165-7.

Rakha, E. A., Boyce, R. W., Abd El-Rehim, D., Kurien, T., Green, A. R., Paish, E. C., Robertson, J. F. & Ellis, I. O. (2005). Expression of mucins (MUC1, MUC2, MUC3, MUC4, MUC5AC and MUC6) and their prognostic significance in human breast cancer. *Mod Pathol*, **18**, 1295-304.

Rakha, E. A., Putti, T. C., Abd El-Rehim, D. M., Paish, C., Green, A. R.,

Powe, D. G., Lee, A. H., Robertson, J. F. & Ellis, I. O. (2006). Morphological and immunophenotypic analysis of breast carcinomas with basal and myoepithelial differentiation. *J Pathol*, **208**, 495-506.

Ravdin, P. M., Siminoff, L. A., Davis, G. J., Mercer, M. B., Hewlett, J., Gerson, N. & Parker, H. L. (2001). Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. *J Clin Oncol*, **19**, 980-91.

Rees, R. C., McArdle, S., Mian, S., Li, G., Ahmad, M., Parkinson, R. & Ali, S. A. (2002). Disabled infectious single cycle-herpes simplex virus (DISC-HSV) as a vector for immunogene therapy of cancer. *Curr Opin Mol Ther*, **4**, 49-53.

Reisman, D. N., Sciarrotta, J., Wang, W., Funkhouser, W. K. & Weissman, B. E. (2003). Loss of BRG1/BRM in human lung cancer cell lines and primary lung cancers: correlation with poor prognosis. *Cancer Res*, **63**, 560-6.

Renwick, A., Thompson, D., Seal, S., Kelly, P., Chagtai, T., Ahmed, M., North, B., Jayatilake, H., Barfoot, R., Spanova, K., McGuffog, L., Evans, D. G., Eccles, D., Easton, D. F., Stratton, M. R. & Rahman, N. (2006). ATM mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles. *Nat Genet*, **38**, 873-5.

Restifo, N. P., Kawakami, Y., Marincola, F., Shamamian, P., Taggarse, A., Esquivel, F. & Rosenberg, S. A. (1993). Molecular mechanisms used by tumors to escape immune recognition: immunogenetherapy and the cell biology of major histocompatibility complex class I. *J Immunother Emphasis Tumor Immunol*, **14**, 182-90.

Reya, T., Morrison, S. J., Clarke, M. F. & Weissman, I. L. (2001). Stem cells, cancer, and cancer stem cells. *Nature*, **414**, 105-11.

Reyland, M. E. (2009). Protein kinase C isoforms: Multi-functional regulators of cell life and death. *Front Biosci*, **14**, 2386-99.

Ricart, A. D. & Tolcher, A. W. (2007). Technology Insight: cytotoxic drug immunoconjugates for cancer therapy. *Nature Reviews Clinical Oncology*, **4**, 245-255.

Ridge, J. P., Di Rosa, F. & Matzinger, P. (1998). A conditioned dendritic cell can be a temporal bridge between a CD4+ T-helper and a T-killer cell. *Nature*, **393**, 474-8.

Ring, B. Z., Seitz, R. S., Beck, R., Shasteen, W. J., Tarr, S. M., Cheang, M. C., Yoder, B. J., Budd, G. T., Nielsen, T. O., Hicks, D. G., Estopinal, N. C. & Ross, D. T. (2006). Novel prognostic immunohistochemical biomarker panel for estrogen receptor-positive breast cancer. *J Clin Oncol*, **24**, 3039-47.

Robison, J. E., Perreard, L. & Bernard, P. S. (2004). State of the science:

molecular classifications of breast cancer for clinical diagnostics. *Clin Biochem*, **37**, 572-8.

Robson, M. E., Chappuis, P. O., Satagopan, J., Wong, N., Boyd, J., Goffin, J. R., Hudis, C., Roberge, D., Norton, L., Begin, L. R., Offit, K. & Foulkes, W. D. (2004). A combined analysis of outcome following breast cancer: differences in survival based on BRCA1/BRCA2 mutation status and administration of adjuvant treatment. *Breast Cancer Res*, **6**, R8-R17.

Rogel, A., Popliker, M., Webb, C. G. & Oren, M. (1985). p53 cellular tumor antigen: analysis of mRNA levels in normal adult tissues, embryos, and tumors. *Mol Cell Biol*, **5**, 2851-5.

Rojas, J. M., McArdle, S. E., Horton, R. B., Bell, M., Mian, S., Li, G., Ali, S. A. & Rees, R. C. (2005). Peptide immunisation of HLA-DR-transgenic mice permits the identification of a novel HLA-DRbeta1*0101- and HLA-DRbeta1*0401-restricted epitope from p53. *Cancer Immunol Immunother*, **54**, 243-53.

Ropke, M., Hald, J., Guldberg, P., Zeuthen, J., Norgaard, L., Fugger, L., Svejgaard, A., Van der Burg, S., Nijman, H. W., Melief, C. J. & Claesson, M. H. (1996). Spontaneous human squamous cell carcinomas are killed by a human cytotoxic T lymphocyte clone recognizing a wild-type p53-derived peptide. *Proc Natl Acad Sci U S A*, **93**, 14704-7.

Rosenberg, S. A., Yang, J. C. & Restifo, N. P. (2004). Cancer immunotherapy: moving beyond current vaccines. *Nat Med*, **10**, 909-15.

Rosenberg, S. A., Yang, J. C., Schwartzentruber, D. J., Hwu, P., Marincola, F. M., Topalian, S. L., Restifo, N. P., Dudley, M. E., Schwarz, S. L., Spiess, P. J., Wunderlich, J. R., Parkhurst, M. R., Kawakami, Y., Seipp, C. A., Einhorn, J. H. & White, D. E. (1998). Immunologic and therapeutic evaluation of a synthetic peptide vaccine for the treatment of patients with metastatic melanoma. *Nat Med*, **4**, 321-7.

Ross, J. S., Linette, G. P., Stec, J., Clark, E., Ayers, M., Leschly, N., Symmans, W. F., Hortobagyi, G. N. & Pusztai, L. (2003). Breast cancer biomarkers and molecular medicine. *Expert Rev Mol Diagn*, **3**, 573-85.

Sahin, U., Tureci, O., Schmitt, H., Cochlovius, B., Johannes, T., Schmits, R., Stenner, F., Luo, G., Schobert, I. & Pfreundschuh, M. (1995). Human neoplasms elicit multiple specific immune responses in the autologous host. *Proc Natl Acad Sci U S A*, **92**, 11810-3.

Sandberg, R. & Ernberg, I. (2005). Assessment of tumor characteristic gene expression in cell lines using a tissue similarity index (TSI). *Proc Natl Acad Sci U S A*, **102**, 2052-7.

Sauter, E. R., Zhu, W., Fan, X. J., Wassell, R. P., Chervoneva, I. & Du Bois, G. C. (2002). Proteomic analysis of nipple aspirate fluid to detect biologic

markers of breast cancer. *Br J Cancer*, **86**, 1440-3.

Scanlan, M. J., Simpson, A. J. & Old, L. J. (2004). The cancer/testis genes: review, standardization, and commentary. *Cancer Immun*, **4**, 1.

Schimmer, A. D., Dalili, S., Batey, R. A. & Riedl, S. J. (2006). Targeting XIAP for the treatment of malignancy. *Cell Death Differ*, **13**, 179-88.

Schmidt, A., Durgan, J., Magalhaes, A. & Hall, A. (2007). Rho GTPases regulate PRK2/PKN2 to control entry into mitosis and exit from cytokinesis. *Embo J*, **26**, 1624-36.

Schnitt, S. J. (2009). The transition from ductal carcinoma in situ to invasive breast cancer: the other side of the coin. *Breast Cancer Res*, **11**, 101.

Scholl, F. A., Dumesic, P. A. & Khavari, P. A. (2005). Effects of active MEK1 expression in vivo. *Cancer Lett*, **230**, 1-5.

Schreurs, M. W., Eggert, A. A., de Boer, A. J., Vissers, J. L., van Hall, T., Offringa, R., Figdor, C. G. & Adema, G. J. (2000). Dendritic cells break tolerance and induce protective immunity against a melanocyte differentiation antigen in an autologous melanoma model. *Cancer Res*, **60**, 6995-7001.

Seal, S., Thompson, D., Renwick, A., Elliott, A., Kelly, P., Barfoot, R., Chagtai, T., Jayatilake, H., Ahmed, M., Spanova, K., North, B., McGuffog, L., Evans, D. G., Eccles, D., Easton, D. F., Stratton, M. R. & Rahman, N. (2006). Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles. *Nat Genet*, **38**, 1239-41.

Seil, I., Frei, C., Sultmann, H., Knauer, S. K., Engels, K., Jager, E., Zatloukal, K., Pfreundschuh, M., Knuth, A., Tseng-Chen, Y., Jungbluth, A. A., Stauber, R. H. & Jager, D. (2007). The differentiation antigen NY-BR-1 is a potential target for antibody-based therapies in breast cancer. *Int J Cancer*, **120**, 2635-42.

Sharkey, R. M., Press, O. W. & Goldenberg, D. M. (2009). A re-examination of radioimmunotherapy in the treatment of non-Hodgkin lymphoma: prospects for dual-targeted antibody/radioantibody therapy. *Blood*, **113**, 3891-5.

Sharma, M. R., Koltowski, L., Ownbey, R. T., Tuszynski, G. P. & Sharma, M. C. (2006). Angiogenesis-associated protein annexin II in breast cancer: selective expression in invasive breast cancer and contribution to tumor invasion and progression. *Exp Mol Pathol*, **81**, 146-56.

Shi, X. Y., Bhagwandeen, B. & Leong, A. S. (2008). p16, cyclin D1, Ki-67, and AMACR as markers for dysplasia in Barrett esophagus. *Appl Immunohistochem Mol Morphol*, **16**, 447-52.

Shi, Y., Evans, J. E. & Rock, K. L. (2003). Molecular identification of a danger signal that alerts the immune system to dying cells. *Nature*, **425**, 516-21.

Shi, M. & Guo, N. (2009). MicroRNA expression and its implications for the diagnosis and therapeutic strategies of breast cancer. *Cancer Treat Rev*, **35**, 328-34.

Shoshan-Barmatz, V., Israelson, A., Brdiczka, D. & Sheu, S. S. (2006). The voltage-dependent anion channel (VDAC): function in intracellular signalling, cell life and cell death. *Curr Pharm Des*, **12**, 2249-70.

Si, M. L., Zhu, S., Wu, H., Lu, Z., Wu, F. & Mo, Y. Y. (2007). miR-21-mediated tumor growth. *Oncogene*, **26**, 2799-803.

Siena, S., Di Nicola, M., Bregni, M., Mortarini, R., Anichini, A., Lombardi, L., Ravagnani, F., Parmiani, G. & Gianni, A. M. (1995). Massive ex vivo generation of functional dendritic cells from mobilized CD34+ blood progenitors for anticancer therapy. *Exp Hematol*, **23**, 1463-71.

Singh, S. K., Clarke, I. D., Terasaki, M., Bonn, V. E., Hawkins, C., Squire, J. & Dirks, P. B. (2003). Identification of a cancer stem cell in human brain tumors. *Cancer Res*, **63**, 5821-8.

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

Smith, G. H. (2002). Mammary cancer and epithelial stem cells: a problem or a solution?. *Breast Cancer Res*, **4**, 47-50.

Smith, C. J., Perfetti, T. A., Garg, R. & Hansch, C. (2003). IARC carcinogens reported in cigarette mainstream smoke and their calculated log P values. *Food Chem Toxicol*, **41**, 807-17.

References

Snyder, J. T., Alexander-Miller, M. A., Berzofskyl, J. A. & Belyakov, I. M. (2003). Molecular mechanisms and biological significance of CTL avidity. *Curr HIV Res*, **1**, 287-94.

Soldaini, E., Wack, A., D'Oro, U., Nuti, S., Ulivieri, C., Baldari, C. T. & Abrignani, S. (2003). T cell costimulation by the hepatitis C virus envelope protein E2 binding to CD81 is mediated by Lck. *Eur J Immunol*, **33**, 455-64.

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

Sorlie, T., Tibshirani, R., Parker, J., Hastie, T., Marron, J. S., Nobel, A., Deng, S., Johnsen, H., Pesich, R., Geisler, S., Demeter, J., Perou, C. M., Lonning, P. E., Brown, P. O., Borresen-Dale, A. L. & Botstein, D. (2003). Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A*, **100**, 8418-23.

Srivastava, P. K., Menoret, A., Basu, S., Binder, R. J. & McQuade, K. L. (1998). Heat shock proteins come of age: primitive functions acquire new roles in an adaptive world. *Immunity*, **8**, 657-65.

Stingl, J., Raouf, A., Emerman, J. T. & Eaves, C. J. (2005). Epithelial progenitors in the normal human mammary gland. *J Mammary Gland Biol Neoplasia*, **10**, 49-59.

Strand, K. J., Khalak, H., Strovel, J. W., Ebner, R. & Augustus, M. (2006). Expression biomarkers for clinical efficacy and outcome prediction in cancer. *Pharmacogenomics*, **7**, 105-15.

Straver, M. E., Glas, A. M., Hannemann, J., Wesseling, J., van de Vijver, M. J., Rutgers, E. J., Vrancken Peeters, M. J., van Tinteren, H., Van't Veer, L. J. & Rodenhuis, S. (2009). The 70-gene signature as a response predictor for neoadjuvant chemotherapy in breast cancer. *Breast Cancer Res Treat*,

Suh, M. R., Lee, Y., Kim, J. Y., Kim, S. K., Moon, S. H., Lee, J. Y., Cha, K. Y., Chung, H. M., Yoon, H. S., Moon, S. Y., Kim, V. N. & Kim, K. S. (2004). Human embryonic stem cells express a unique set of microRNAs. *Dev Biol*, **270**, 488-98.

Surh, C. D. & Sprent, J. (1994). T-cell apoptosis detected in situ during positive and negative selection in the thymus. *Nature*, **372**, 100-3.

Sutherland, C. L., Chalupny, N. J. & Cosman, D. (2001). The UL16-binding proteins, a novel family of MHC class I-related ligands for NKG2D, activate natural killer cell functions. *Immunol Rev*, **181**, 185-92.

Szymanski, M., Barciszewska, M. Z., Zywicki, M. & Barciszewski, J. (2003). Noncoding RNA transcripts. *J Appl Genet*, **44**, 1-19.

Takahashi, T., Tanaka, Y., Nieda, M., Azuma, T., Chiba, S., Juji, T., Shibata, Y. & Hirai, H. (2003). Dendritic cell vaccination for patients with chronic myelogenous leukemia. *Leuk Res*, **27**, 795-802.

Takeshita, F. & Ochiya, T. (2006). Therapeutic potential of RNA interference against cancer. *Cancer Sci*, **97**, 689-96.

Tamir, A., Basagila, E., Kagahzian, A., Jiao, L., Jensen, S., Nicholls, J., Tate, P., Stamp, G., Farzaneh, F., Harrison, P., Stauss, H., George, A. J., Habib, N., Lechler, R. I. & Lombardi, G. (2007). Induction of tumor-specific T-cell responses by vaccination with tumor lysate-loaded dendritic cells in colorectal cancer patients with carcinoembryonic-antigen positive tumors. *Cancer Immunol Immunother*, **56**, 2003-16.

Tamm, I. (2008). AEG-35156, an antisense oligonucleotide against X-linked inhibitor of apoptosis for the potential treatment of cancer. *Curr Opin Investig Drugs*, **9**, 638-46.

Tanaka, K., Isselbacher, K. J., Khoury, G. & Jay, G. (1985). Reversal of oncogenesis by the expression of a major histocompatibility complex class I gene. *Science*, **228**, 26-30.

Tang, H. L., Le, A. H. & Lung, H. L. (2006). The increase in mitochondrial association with actin precedes Bax translocation in apoptosis. *Biochem J*, **396**, 1-5.

Tang, P., Wang, X., Schiffhauer, L., Wang, J., Bourne, P., Yang, Q., Quinn, A. & Hajdu, S. I. (2006). Relationship between nuclear grade of ductal carcinoma in situ and cell origin markers. *Ann Clin Lab Sci*, **36**, 16-22.

Taylor, S. M. & Jones, P. A. (1982). Changes in phenotypic expression in embryonic and adult cells treated with 5-azacytidine. *J Cell Physiol*, **111**, 187-94.

Theriault, R. L. & Hortobagyi, G. N. (1992). Bone metastasis in breast cancer. *Anticancer Drugs*, **3**, 455-62.

Theurillat, J. P., Zurrer-Hardi, U., Varga, Z., Barghorn, A., Saller, E., Frei, C., Storz, M., Behnke, S., Seifert, B., Fehr, M., Fink, D., Rageth, C., Linsenmeier, C., Pestalozzi, B., Chen, Y. T., Knuth, A., Jager, D. & Moch, H. (2008). Distinct

expression patterns of the immunogenic differentiation antigen NY-BR-1 in normal breast, testis and their malignant counterparts. *Int J Cancer*, **122**, 1585-91.

Theurillat, J. P., Zurrer-Hardi, U., Varga, Z., Storz, M., Probst-Hensch, N. M., Seifert, B., Fehr, M. K., Fink, D., Ferrone, S., Pestalozzi, B., Jungbluth, A. A., Chen, Y. T., Jager, D., Knuth, A. & Moch, H. (2007). NY-BR-1 protein expression in breast carcinoma: a mammary gland differentiation antigen as target for cancer immunotherapy. *Cancer Immunol Immunother*, **56**, 1723-31.

Thodeti, C. K., Albrechtsen, R., Grauslund, M., Asmar, M., Larsson, C., Takada, Y., Mercurio, A. M., Couchman, J. R. & Wewer, U. M. (2003). ADAM12/syndecan-4 signaling promotes beta 1 integrin-dependent cell spreading through protein kinase Calpha and RhoA. *J Biol Chem*, **278**, 9576-84.

Thompson, A., Brennan, K., Cox, A., Gee, J., Harcourt, D., Harris, A., Harvie, M., Holen, I., Howell, A., Nicholson, R., Steel, M. & Streuli, C. (2008). Evaluation of the current knowledge limitations in breast cancer research: a gap analysis. *Breast Cancer Res*, **10**, R26.

Toloza, E. M., Morse, M. A. & Lyerly, H. K. (2006). Gene therapy for lung cancer. *J Cell Biochem*, **99**, 1-22.

Travis, R. C. & Key, T. J. (2003). Oestrogen exposure and breast cancer risk. *Breast Cancer Res*, **5**, 239-47.

Trock, B. J., Leonessa, F. & Clarke, R. (1997). Multidrug resistance in breast cancer: a meta-analysis of MDR1/gp170 expression and its possible functional significance. *J Natl Cancer Inst*, **89**, 917-31.

Tuting, T., Gambotto, A., Robbins, P. D., Storkus, W. J. & DeLeo, A. B. (1999). Co-delivery of T helper 1-biasing cytokine genes enhances the efficacy of gene gun immunization of mice: studies with the model tumor antigen beta-galactosidase and the BALB/c Meth A p53 tumor-specific antigen. *Gene Ther*, **6**, 629-36.

Vaarala, M. H., Porvari, K. S., Kyllonen, A. P., Mustonen, M. V., Lukkarinen, O. & Vihko, P. T. (1998). Several genes encoding ribosomal proteins are over-expressed in prostate-cancer cell lines: confirmation of L7a and L37 over-expression in prostate-cancer tissue samples. *Int J Cancer*, **78**, 27-32.

Vachon, C. M., van Gils, C. H., Sellers, T. A., Ghosh, K., Pruthi, S., Brandt, K. R. & Pankratz, V. S. (2007). Mammographic density, breast cancer risk and risk prediction. *Breast Cancer Res*, **9**, 217.

Vahteristo, P., Bartkova, J., Eerola, H., Syrjakoski, K., Ojala, S., Kilpivaara, O., Tamminen, A., Kononen, J., Aittomaki, K., Heikkila, P., Holli, K., Blomqvist, C., Bartek, J., Kallioniemi, O. P. & Nevanlinna, H. (2002). A CHEK2 genetic variant contributing to a substantial fraction of familial breast cancer. *Am J Hum Genet*, **71**, 432-8.

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

Van den Eynde, B., Peeters, O., De Backer, O., Gaugler, B., Lucas, S. & Boon, T. (1995). A new family of genes coding for an antigen recognized by autologous cytolytic T lymphocytes on a human melanoma. *J Exp Med*, **182**, 689-98.

van der Bruggen, P., Traversari, C., Chomez, P., Lurquin, C., De Plaen, E., Van den Eynde, B., Knuth, A. & Boon, T. (1991). A gene encoding an antigen recognized by cytolytic T lymphocytes on a human melanoma. *Science*, **254**, 1643-7.

Vanderlugt, C. L., Begolka, W. S., Neville, K. L., Katz-Levy, Y., Howard, L. M., Eagar, T. N., Bluestone, J. A. & Miller, S. D. (1998). The functional significance of epitope spreading and its regulation by co-stimulatory molecules. *Immunol Rev*, **164**, 63-72.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A. & Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol*, **3**, RESEARCH0034.

van de Vijver, M. J., He, Y. D., van't Veer, L. J., Dai, H., Hart, A. A., Voskuil, D. W., Schreiber, G. J., Peterse, J. L., Roberts, C., Marton, M. J., Parrish, M., Atsma, D., Witteveen, A., Glas, A., Delahaye, L., van der Velde, T., Bartelink, H., Rodenhuis, S., Rutgers, E. T., Friend, S. H. & Bernards, R. (2002). A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med*, **347**, 1999-2009.

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

Varga, Z., Theurillat, J. P., Filonenko, V., Sasse, B., Odermatt, B., Jungbluth, A. A., Chen, Y. T., Old, L. J., Knuth, A., Jager, D. & Moch, H. (2006). Preferential nuclear and cytoplasmic NY-BR-1 protein expression in primary breast cancer and lymph node metastases. *Clin Cancer Res*, **12**, 2745-51.

Vasilevskaya, I. A., Selvakumaran, M. & O'Dwyer, P. J. (2008). Disruption of signaling through SEK1 and MKK7 yields differential responses in hypoxic colon cancer cells treated with oxaliplatin. *Mol Pharmacol*, **74**, 246-54.

Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., Gocayne, J. D., Amanatides, P., Ballew, R. M., Huson, D. H., Wortman, J. R., Zhang, Q., Kodira,

C. D., Zheng, X. H., Chen, L., Skupski, M. *et al.* (2001). The sequence of the human genome. *Science*, **291**, 1304-51.

Viani, G. A., Afonso, S. L., Stefano, E. J., De Fendi, L. I. & Soares, F. V. (2007). Adjuvant trastuzumab in the treatment of her-2-positive early breast cancer: a meta-analysis of published randomized trials. *BMC Cancer*, **7**, 153.

Volpi, A., Bacci, F., Paradiso, A., Saragoni, L., Scarpi, E., Ricci, M., Aldi, M., Bianchi, S., Muretto, P., Nuzzo, F., Simone, G., Mangia, A., Schittulli, F. & Amadori, D. (2004). Prognostic relevance of histological grade and its components in node-negative breast cancer patients. *Mod Pathol*, **17**, 1038-44.

von Boehmer, H. (2005). Mechanisms of suppression by suppressor T cells. *Nat Immunol*, **6**, 338-44.

Vonderheide, R. H., Hahn, W. C., Schultze, J. L. & Nadler, L. M. (1999). The telomerase catalytic subunit is a widely expressed tumor-associated antigen recognized by cytotoxic T lymphocytes. *Immunity*, **10**, 673-9.

Vujanovic, L., Ranieri, E., Gambotto, A., Olson, W. C., Kirkwood, J. M. & Storkus, W. J. (2006). IL-12p70 and IL-18 gene-modified dendritic cells loaded with tumor antigen-derived peptides or recombinant protein effectively stimulate specific Type-1 CD4+ T-cell responses from normal donors and melanoma patients in vitro. *Cancer Gene Ther*, **13**, 798-805.

Wagner, K. J., Cooper, W. N., Grundy, R. G., Caldwell, G., Jones, C., Wadey, R. B., Morton, D., Schofield, P. N., Reik, W., Latif, F. & Maher, E. R. (2002). Frequent RASSF1A tumour suppressor gene promoter methylation in Wilm's tumour and colorectal cancer. *Oncogene*, **21**, 7277-82.

Walsh, T., Casadei, S., Coats, K. H., Swisher, E., Stray, S. M., Higgins, J., Roach, K. C., Mandell, J., Lee, M. K., Ciernikova, S., Foretova, L., Soucek, P. & King, M. C. (2006). Spectrum of mutations in BRCA1, BRCA2, CHEK2, and TP53 in families at high risk of breast cancer. *Jama*, **295**, 1379-88.

Walsh, C. T., Radeff-Huang, J., Matteo, R., Hsiao, A., Subramaniam, S., Stupack, D. & Brown, J. H. (2008). Thrombin receptor and RhoA mediate cell proliferation through integrins and cysteine-rich protein 61. *Faseb J*, **22**, 4011-21.

Walton, T. J., Li, G., McCulloch, T. A., Seth, R., Powe, D. G., Bishop, M. C. & Rees, R. C. (2009). Quantitative RT-PCR analysis of estrogen receptor gene expression in laser microdissected prostate cancer tissue. *Prostate*, **69**, 810-9.

Wang, Y. & Blelloch, R. (2009). Cell cycle regulation by MicroRNAs in embryonic stem cells. *Cancer Res*, **69**, 4093-6.

Wang, W., Epler, J., Salazar, L. G. & Riddell, S. R. (2006). Recognition of breast cancer cells by CD8+ cytotoxic T-cell clones specific for NY-BR-1. *Cancer Res*, **66**, 6826-33.

Wang, J., Guo, L. P., Chen, L. Z., Zeng, Y. X. & Lu, S. H. (2007). Identification of cancer stem cell-like side population cells in human nasopharyngeal carcinoma cell line. *Cancer Res*, **67**, 3716-24.

Wang, M., Hu, Y. & Stearns, M. E. (2009). RPS2: a novel therapeutic target in prostate cancer. *J Exp Clin Cancer Res*, **28**, 6.

Wang, W., Epler, J., Salazar, L. G. & Riddell, S. R. (2006). Recognition of breast cancer cells by CD8+ cytotoxic T-cell clones specific for NY-BR-1. *Cancer Res*, **66**, 6826-33.

Wang, R. F., Wang, X. & Rosenberg, S. A. (1999). Identification of a novel major histocompatibility complex class II-restricted tumor antigen resulting from a chromosomal rearrangement recognized by CD4(+) T cells. *J Exp Med*, **189**, 1659-68.

Ward, H. W. (1973). Anti-oestrogen therapy for breast cancer: a trial of tamoxifen at two dose levels. *Br Med J*, **1**, 13-4.

Weber, J. (2000). Melanoma peptide vaccines: from preclinical background to clinical trials. *Curr Oncol Rep*, **2**, 38-47.

Weinberg, R. A., The biology of cancer, Garland Science, 2007

Weiss, R., Scheiblhofer, S., Freund, J., Ferreira, F., Livey, I. & Thalhamer, J. (2002). Gene gun bombardment with gold particles displays a particular Th2-promoting signal that over-rules the Th1-inducing effect of immunostimulatory CpG motifs in DNA vaccines. *Vaccine*, **20**, 3148-54.

Westermark, B., Heldin, C. H. & Nister, M. (1995). Platelet-derived growth factor in human glioma. *Glia*, **15**, 257-63.

Widschwendter, M. & Jones, P. A. (2002). DNA methylation and breast carcinogenesis. *Oncogene*, **21**, 5462-82.

Wishart, G. C., Gaston, M., Poultsidis, A. A. & Purushotham, A. D. (2002). Hormone receptor status in primary breast cancer–time for a consensus?. *Eur J Cancer*, **38**, 1201-3.

Woodland, D. L. (2004). Jump-starting the immune system: prime-boosting comes of age. *Trends Immunol*, **25**, 98-104.

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

Wu, K., Zhang, Y., Xu, X. C., Hill, J., Celestino, J., Kim, H. T., Mohsin, S. K., Hilsenbeck, S. G., Lamph, W. W., Bissonette, R. & Brown, P. H. (2002). The retinoid X receptor-selective retinoid, LGD1069, prevents the development of estrogen receptor-negative mammary tumors in transgenic mice. *Cancer Res*, **62**,

6376-80.

Xiang, Y., Zhu, Z., Han, G., Lin, H., Xu, L. & Chen, C. D. (2007). JMJD3 is a histone H3K27 demethylase. *Cell Res*, **17**, 850-7.

Yamamoto, K., Okamoto, A., Isonishi, S., Ochiai, K. & Ohtake, Y. (2001). A novel gene, CRR9, which was up-regulated in CDDP-resistant ovarian tumor cell line, was associated with apoptosis. *Biochem Biophys Res Commun*, **280**, 1148-54.

Yang, D. C., Elliott, R. L. & Head, J. F. (2002). Gene targets of antisense therapies in breast cancer. *Expert Opin Ther Targets*, **6**, 375-85.

Yang, Y., Kitagaki, J., Wang, H., Hou, D. X. & Perantoni, A. O. (2009). Targeting the ubiquitin-proteasome system for cancer therapy. *Cancer Sci*, **100**, 24-8.

Yang, Q., Sakurai, T., Yoshimura, G., Suzuma, T., Umemura, T., Nakamura, M., Nakamura, Y., Mori, I. & Kakudo, K. (2003). Prognostic value of Bcl-2 in invasive breast cancer receiving chemotherapy and endocrine therapy. *Oncol Rep*, **10**, 121-5.

Yoshida, R., Kimura, N., Harada, Y. & Ohuchi, N. (2001). The loss of E-cadherin, alpha- and beta-catenin expression is associated with metastasis and poor prognosis in invasive breast cancer. *Int J Oncol*, **18**, 513-20.

Yu, K., Lee, C. H., Tan, P. H., Hong, G. S., Wee, S. B., Wong, C. Y. & Tan, P. (2004). A molecular signature of the Nottingham prognostic index in breast cancer. *Cancer Res*, **64**, 2962-8.

Zeng, G., Li, Y., El-Gamil, M., Sidney, J., Sette, A., Wang, R. F., Rosenberg, S. A. & Robbins, P. F. (2002). Generation of NY-ESO-1-specific CD4+ and CD8+ T cells by a single peptide with dual MHC class I and class II specificities: a new strategy for vaccine design. *Cancer Res*, **62**, 3630-5.

Zhang, W. G., Liu, S. H., Cao, X. M., Cheng, Y. X., Ma, X. R., Yang, Y. & Wang, Y. L. (2005). A phase-I clinical trial of active immunotherapy for acute leukemia using inactivated autologous leukemia cells mixed with IL-2, GM-CSF, and IL-6. *Leuk Res*, **29**, 3-9.

Zhang, Z., Sun, P., Liu, J., Fu, L., Yan, J., Liu, Y., Yu, L., Wang, X. & Yan, Q. (2008). Suppression of FUT1/FUT4 expression by siRNA inhibits tumor growth. *Biochim Biophys Acta*, **1783**, 287-96.

Zhao, P., Lu, Y., Zhong, M., Liu, L. & Li, B. (2008). Inverse correlation of aberrant expression of fragile histidine triad (FHIT) protein with cyclin D1 protein and prognosis in Chinese patients with cholangiocarcinoma. *Acta Oncol*, **47**, 1557-63.

Communications

Oral presentations

Oral presentations on the progress of the study at research group meetings and collaborator research group meetings.

Oral presentation "Expression analysis of genes for both diagnosis of breast cancer stage and prediction of outcome" at the School of Biomedical and Natural Sciences Postgraduate Research Conference, 5 June 2007 (second year of the PhD talk).

School seminar presentation "Expression analysis of novel biomarkers for breast cancer", 12 December 2008 (final year of the PhD talk).

Posters and published poster abstracts

"Expression analysis of genes for both diagnosis of breast cancer stage and prediction of outcome", presented and published in the Conference proceedings of the School of Biomedical and Natural Sciences Postgraduate Research Conference (5 June 2007, Nottingham Trent University).

"Expression Analysis of Novel Biomarkers for Breast Cancer" presented and published in the Conference Proceedings of the AACR Special Conference in Cancer Research "Advances in Breast Cancer Research: Genetics, Biology, and Clinical Applications" (17-20 October 2007, San Diego).

"Expression Analysis of Novel Biomarkers for Breast Cancer" presented and published in Breast Cancer Research Volume Supplement 2, 2008 "Meeting Abstracts Breast Cancer Research 2008" (13 May 2008, The Royal Society, London).

Publications

Review:
Laversin, S.A., Miles, A.K., Ball, G.R.& Rees, R.C. (2008). Emerging Breast Cancer Biomarkers. *Current Cancer Therapy Reviews*, **4**, 79-85.

Research article: in preparation.