Defining the regulation and function of SAFB1 and SAFB2 in human breast cancer cells

Elaine Hong

Institute of Cellular Medicine

Newcastle University

A thesis submitted for the degree of

Doctor of Philosophy

July 2013

Abstract

Scaffold attachment factor B1 (SAFB1) and SAFB2 are oestrogen receptor (ER) corepressors that bind and modulate ER activity through chromatin remodelling or interaction with the basal transcription machinery. However, little is known about the fundamental characteristics and function of SAFB1 and SAFB2 proteins in breast cancer.

In this study, an investigation of the characteristics and function(s) of SAFB1 and SAFB2 was undertaken; their expression profile was first assessed in ER-positive (MCF-7) and ER-negative (MDA-MB-231) breast cancer cell lines. Results show that SAFB1 and SAFB2 are themselves regulated by an active metabolite of oestrogen, 17β-oestradiol, in both ER positive and ER negative breast cancer cells.

Using a combined approach of RNA interference and gene expression profile studies, 12 novel targets closely linked to tumour progression were identified for SAFB1 and SAFB2. Expression levels of the following genes, *CDKN2A*, *CLU*, *ESR1*, *IGFBP2*, *IL2RA*, *ITGB4*, *KIT*, *KLK5*, *MT3*, *NGFR* and *SPRR1B* increased while *IL-6* expression and secretion decreased when cells were depleted of SAFB proteins. This observation supports their primary role as transcriptional repressors with SAFB2 playing a prominent role in transcriptional regulation in MDA-MB-231 cells. This study has also established a novel link between SAFB proteins and *ITGB4* and *IL-6* expression.

Both SAFB proteins have an internal RNA-recognition motif but little is known about the RNA-binding properties of SAFB1 or SAFB2. To investigate this, the concluding part of the project utilised crosslinking and immunoprecipitation (CLIP) coupled with high-throughput sequencing. This experimental approach enabled a transcriptome-wide mapping of SAFB1 protein-RNA interactions in breast cancer cells. SAFB1 crosslink sites are significantly enriched in ncRNAs, particularly within snRNAs. A putative RNA-binding motif for SAFB1 that contains adenine-rich sequences, highly similar to the RNA-binding motifs for SR proteins was also identified.

In summary, this study has defined the characteristics of SAFB1 and SAFB2 in both ER positive and ER negative breast cancer cell lines. It has also identified transcriptional targets and the RNA-binding ability of SAFB1 and SAFB2 in human breast cancer cells.

Acknowledgements

I would like to extend my thanks to my supervisor, Dr Alison Tyson-Capper for her constant support, patience, enthusiasm and motivation throughout my research. Also to my second supervisor, Prof David Elliott for his help and guidance especially in the CLIP work. I have benefited immensely from the strengths of this supervisory team and am truly grateful for the both of you.

I would also like to thank my PhD progress assessors, Prof Simi Ali and Prof Zofia Chrzanowska-Lightowlers for helping me progress from an average scientist to a critical thinker. The annual progress assessments have continuously challenged me to question my work, thus sharpening my critical thinking skills and giving me a clear objective to every experiment attempted.

My thanks also go to my fellow labmates: Dr David Cork, Dr Hannah Gautrey, Dr Rachel Eyre, Helen Lawrence, Andrew Best; and lab technician: Liz Shiells and Caroline Dalgiesh for all their support, advice and practical help in the lab. You have made my time in the lab a pleasant one and I'm thankful to have worked alongside each of you.

My sincere appreciation goes to all the close friends I've made throughout this journey in Newcastle: Stephanie Chin for being the best friend/housemate I could ever ask for. Thank you for those nights where you stayed up to accompany this mad scientist! You have been most supportive at my toughest times, you're such a gem! Cynthia Wong, Karen Chan and Lihyun Wong for being the friends I could trust and rely on when all else seem to fail. Thank you for bringing fun into my life and putting smiles on my face. All of you have been an integral part of my life here and our friendship is always precious to me.

My very special thanks go to my beloved family: Daddy and Mummy for always believing in me even though I've always doubted myself. Thank you for the continuous love, patience and grace that you've showered upon my entire life. Emily Hong for taking over my responsibility in the family while I'm away. Thank you for stepping up to the challenge and doing it so well. Adeline Hong for being here beside me and cheering me on at the final sprint of this journey. Thank you for your love and great support at my most crucial time.

Above all else, I would like to thank my Saviour and my Lord, Jesus Christ for being my constant and firm foundation in this temporal and fading world. Thank You for Your mercy, faithfulness and immeasurable love that has led me to discover the greater purpose of my existence. This life is no longer my own, I've been redeemed and transformed to pursue Your kingdom first. Let my life sing Your glory and praises. I am eternally grateful for everything You've done in and through me.

Table of contents

Abbreviations

×g	Times gravity
°C	Degree Celcius
µg	Microgram
µl	Microlitre
µM	Micromolar
AF	Transactivation domain
ATP	Adenosine triphosphate
CDKN2A	Cyclin-dependent kinase inhibitor 2A
cDNA	Complementary DNA
ChIP	Chromatin immunoprecipitation
CLIP	Crosslinking and immunoprecipitation
CLU	Clusterin
Cm	Centimeters
CO_2	Carbon dioxide
Co-IP	Co-immunoprecipitation
C_t	Cycle threshold
DAPI	4’,6-Diamidino-2-phenylindole
DBD	DNA-binding domain
DCIS	Ductal carcinoma in situ
DMEM	Dulbecco's Modified Eagles medium
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
ECL	Electrochemiluminescence
ELISA	Enzyme-linked immunosorbent assay
ER	Oestrogen receptor
ERE	Oestrogen response element
ERK-1	Extracellular signal-regulated kinase 1
ESR1	Oestrogen receptor 1
FBS	Foetal bovine serum
FDR	False discovery rate
FITC	Fluorescein isothiocyanate
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
HAT	Histone acetyltransferase
HDAC	Histone deacetylase
hnRNP	Heterogenous nuclear ribonucleoprotein
HRP	Horseradish peroxidase
HRT	Hormone replacement therapy
iCLIP	Individual nucleotide resolution CLIP
IDC	Invasive ductal carcinoma
IGFBP2	Insulin-like growth factor binding protein 2
IL2RA	Interleukin 2 receptor alpha
IL-6	Interleukin 6
ILC	Invasive lobular carcinoma
ITGB4	Integrin beta 4
kDa	KiloDalton

KIT	V-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog
KLK5	Kallikrein-related peptidase 5
LBD	Ligand-binding domain
LCIS	Lobular carcinoma in situ
Ml	Mililiter
mRNA	Messenger RNA
MT3	Metallothionein 3
ncRNA	Non-coding RNA
NGFR	Nerve growth factor receptor
NLS	Nuclear localisation sequence
nM	Nanomolar
ORF	Open reading frame
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PNK	Polynucleotide kinase
PR	Progestrone receptor
qRT-PCR	Quantitative real time PCR
RBD	RNA-binding domain
RBP	RNA-binding protein
RIPA	Radio immunoprecipitation assay
RNA	Ribonucleic acid
RNAi	RNA interference
RRM	RNA-recognition motif
S/MARs	Scaffold/matrix attachment regions
SAFB	Scaffold attachment factor B
SC35	Splicing component 35kDa
SDS	Sodium dodecyl sulphate
SDS-PAGE	SDS-Polyacrylamide gel electrophoresis
SERD	Selective oestrogen receptor downregulator
SERM	Selective oestrogen receptor modulator
SHF	Src homology 2 domain containing F
siRNA	Small interfering RNA
snRNA	Small nuclear RNA
SPRR1B	Small proline-rich protein 1B
SR	Serine/arginine rich
SRGN	Serglycin
SRSF	Serine/arginine rich splicing factor
SUMO	Small ubiquitin-like modifiers
TNM	Tumour, lymph nodes and metastasis
TRA2B	Transformer 2 protein homolog beta
UTR	Untranslated region
VEGF-A	Vascular endothelial growth factor A

List of Figures

List of Tables

Chapter 1 : Introduction

1.1 The mammary gland

1.1.1 Normal development of the mammary gland

The human female mammary gland develops through key stages in life particularly during fetal growth, infancy, puberty, pregnancy, lactation and post-menopausal regression (Russo *et al.* 1987). The formation of the mammary crest and primitive mammary buds begins during embryonic life through a series of highly ordered events regulated by growth factors and hormones [reviewed in (Hassiotou *et al.* 2012)]. Paired ectodermal thickenings, known as mammary ridges, develop from the axilla to the groin on the anterior surface of the embryo by 6 weeks of gestation. These mammary ridges initially appear in multiple pairs along the milk line but most ridges regress during development except for one pair in the pectoral region, where the two mammary glands subsequently develop. The mammary parenchyma then invades the stroma to form the mammary crest at week 7 and 8 of gestation. Between 10 and 12 weeks of gestation, the mammary ridge begins to differentiate and proliferate into mammary epithelial buds. The mammary buds continue to proliferate during weeks 13 to 20 of gestation to form branches and secondary buds. These secondary buds gradually elongate into 15 to 25 solid cords by 20 weeks gestation (Osborne 2004). At 32 weeks, placental hormones enter the fetal circulation to initiate canalisation of the solid cords and formation of the lactiferous ducts (Hovey *et al.* 2002). The ducts enlarge to form lactiferous sinuses and converge into the nipple area. The periductal stroma increases in density along with limited lobulo-alveolar development, while the areola acquires a slight pigmentation during the final 8 weeks of gestation (Naccarato *et al.* 2000). The infant's mammary gland produces small amount of colostrum shortly after birth as a result of maternal lactogenic hormones present in the fetal circulation. The decrease in infant prolactin levels coincide with the spontaneous regression of the infant's mammary gland within 4 weeks post-partum (Vorherr 1974).

The neonatal breast consists of rudimentary ducts that remain quiescent until early childhood. Growth of the mammary epithelium and stroma continues at puberty, between the age of 8 to 12 years (Russo *et al.* 2004). Puberty induces rapid breast growth as a result of increasing levels of secreted steroid hormones, primarily oestrogen in the form of 17β-oestradiol, driven by ovulation and the establishment of menstrual

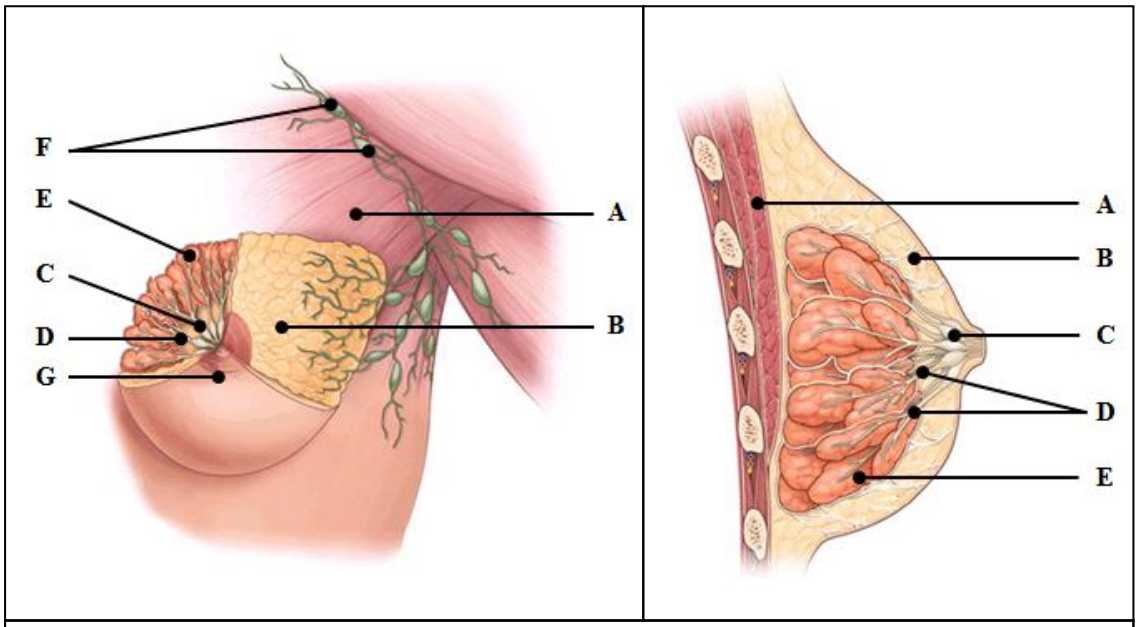
cycles (Vorherr 1974). The increased deposition of adipose tissue within the gland and vasculature enhancement causes an increase in breast size. Circulating oestrogen stimulates elongation of the existing ducts and branching into secondary ducts. The growing and dividing ducts form rounded terminal end buds, from which bi-layered epithelial buds appear and form lobules (Russo *et al.* 2004). The lobulo-alveolar development continues gradually during adolescence until approximately the age of 35 years, resulting in three distinct types of lobules in the mature adult breast (Russo *et al.* 2004).

Breast development culminates during pregnancy and lactation cycle when the mammary gland matures into a functional milk-secretory organ through complete remodelling of the breast. Increased levels of circulating lactogenic hormone complex (oestrogen, progesterone and prolactin) directly regulate this maturation stage by inducing ductal branching, alveolar morphogenesis and secretory differentiation (Pang *et al.* 2007). By the second trimester of pregnancy, mammary epithelial cells in the luminal layer of the alveolar differentiate into milk-producing lactocytes. Secretory activation and milk synthesis occurs after parturition, triggered by the decrease in circulating progesterone and further increase in prolactin. Cessation of milk removal causes the mammary gland to transition into a resting non-lactating state through post-lactational involution (Hurley 1989). Clearing of the mammary alveolar cells occurs during involution, allowing the breast to regress into a non-functional organ until the next pregnancy and lactation cycle. The breast undergoes a second phase of involution (post-menopausal involution) triggered by ovarian function decay during menopause, which is characterised by the decrease of glandular breast tissue and increase in the adipose surrounding tissue (Hutson *et al.* 1985).

1.1.2 Anatomy of the breast

The breast overlies the pectoralis major muscle positioned between the second and the sixth rib bone. The blood supply of the breast is derived from the internal mammary artery (60%) and the lateral thoracic artery (30%) (Vorherr 1974). Lymphatic drainage of the breast drains lymph from the breast to the axillary nodes and the internal mammary nodes (Hultborn *et al.* 1955, Turner-Warwick 1955).

The adult breast is composed of secretory and fatty tissue supported by loose fibrous connective tissue known as Cooper's ligaments (Cooper 1840). The secretory tissue is made of 15-25 lobes that converge at the nipple in a radial arrangement and a ductal system that drains the alveoli to transport milk to the nipple (Figure 1.1). Each lobe comprises of lobules that contain 10-100 alveoli that produce and store milk. Numerous small ductules that drain the alveoli merge into one main duct, which dilates to form the lactiferous sinus, then narrows and opens onto the nipple surface (Venta *et al.* 1994).



Legend:

A. Pectoralis major muscle
B. Adipose tissue
C. Lactiferous sinus
D. Ducts
E. Lobules
F. Lymph nodes
G. Areola

Figure 1.1 Anatomy of the human breast.

The left panel shows the front view and the right panel shows the side view of the human breast. Image adapted from (urmc.rochester.edu).

1.2 Breast cancer

1.2.1 Epidemiology

Breast cancer is the most common cancer in the UK since 1997 although its incidence is rare in men. In 2010, 49,564 women and 397 men in the UK were diagnosed with

invasive breast cancer, causing 11,633 deaths (Cancer Research UK 2010a). Breast cancer accounts for 31% of all female cancers and the lifetime risk for developing breast cancer is 1 in 8. Breast cancer is the second most common cause of cancer death among women in the UK after lung cancer; however, its survival rates have improved over these past 40 years. Five-year survival rate for breast cancer has increased from 52% patients diagnosed in 1971-1975 to 85.1% patients diagnosed in 2005-2009 (Cancer Research UK 2010b). Survival rate of this disease largely relies on the cancer stage at the point of diagnosis, where 90% of women diagnosed at stage I breast cancer survive beyond five years compared to only 10% when diagnosed at stage IV breast cancer (Cancer Research UK 2010a).

1.2.2 Risk factors

There is no single definitive factor that is responsible for the majority of breast cancer; however there are several established factors that strongly increase a woman's risk of developing the disease. The strongest risk factor for breast cancer is age, where incidence and mortality increases as women get older. Based on the estimates for 2008, the risk of developing breast cancer increases 10 times in women from age 29 to age 39, 4 times from age 39 to age 49, then doubles the next 10 years until menopause when the rate of increase slows dramatically (Cancer Research UK 2012, Sasieni *et al.* 2011). Environmental factors in the form of geographical variation presents as another important risk factor, as women in developed countries are five times more at risk of developing breast cancer compared to women from less developed countries. Interestingly, one study of migrants from eastern to western countries show that the risk of breast cancer in migrants assumes that of the host country within one or two generations, indicating that environmental factors outweigh genetic factors (McPherson *et al.* 2000). While genetic predisposition doubles the risk of breast cancer in a woman with an affected first degree relative, it only contributes up to 10% of breast cancer in western countries (Pharoah *et al.* 1997). Critical analysis of epidemiological studies for familial breast cancer has revealed that 8 out of 9 women who develop breast cancer do not have affected first degree relatives and although women with an affected close relative are at increased risk, most will never develop breast cancer (Collaborative Group on Hormonal Factors in Breast Cancer 2001). Hereditary factors only contribute to a quarter of an individual's susceptibility to breast cancer while

environmental and lifestyle factors contribute to the remaining three-quarters (Lichtenstein *et al.* 2000).

Genetic predisposition to breast cancer originally derived from evidence of cancer clustering in families and increased cancer risk in individuals with certain genetically linked diseases. Germline mutations in at least five genes have been identified to predispose an individual to breast cancer: breast cancer 1 (*BRCA1*), *BRCA2*, tumour protein 53 (*TP53*), phosphate and tensin homolog (*PTEN*) and ataxia telangiectasia mutated (*ATM*) (Malkin *et al.* 1990, Nelen *et al.* 1996, Peto *et al.* 1999, Swift *et al.* 1991). Mutations in *BRCA1* and *BRCA2* are high penetrance genes that increase the risk of breast and ovarian cancer, and account for 2% of all breast cancer cases. Germline mutations in *TP53* predispose to Li-Fraumeni cancer syndrome, a condition which encompasses childhood sarcoma, brain tumours and early onset breast cancer. Mutations in *PTEN* are responsible for Cowden syndrome, which is mainly characterised by breast cancer and heterozygous carriers of *ATM* have an increased risk of breast cancer.

Many reproductive factors that influence breast cancer risk are closely linked to prolonged exposure to hormones. Early menarche and late menopause extends the relative exposure of the breasts to elevated levels of endogenous oestradiol, thus increasing the risk of breast cancer. High oestradiol concentrations and obesity in post-menopausal women are also positively associated with breast cancer risk (Key *et al.* 1999). On the other hand, increased childbearing and prolonged breastfeeding have a protective effect that reduces the risk of breast cancer (Collaborative Group on Hormonal Factors in Breast Cancer 2002). The intake of exogenous hormones through oral contraceptive slightly increase the relative risk of developing breast cancer in current users, however the excess risk falls after cessation of use and becomes insignificant after more than 10 years cessation (Collaborative Group on Hormonal Factors in Breast Cancer 1996). Hormone replacement therapy (HRT) substantially increases oestrogen concentration in the serum and incidentally increases the relative risk of breast cancer. Current oestrogen-progestagen therapies increase the risk of breast cancer that is enhanced by the duration of use but diminished after cessation (Collaborative Group on Hormonal Factors in Breast Cancer 1997).

Clinical factors such as breast density and benign breast disease also contribute to the risk of developing breast cancer. Breast density is defined by its proportion of fatty

tissue, where low breast density correlates with high proportion of fatty tissue and vice versa. Women with high breast density are five times more likely to develop breast cancer than women with low breast density (McCormack *et al.* 2006). Benign breast diseases that stem from non-proliferative lesions are associated with little or no increase in breast cancer risk. Women with proliferative lesions without atypia have two-fold increased risk, while those with atypical hyperplasia have at least four-fold increased risk (Hartmann *et al.* 2005).

Other modifiable lifestyle factors that increase the risk of breast cancer include high body mass index, lack of physical activity, high alcohol consumption, high fat diet and smoking before the age of 20 (Cancer Research UK 2012).

1.2.3 Breast screening

The prognosis of breast cancer heavily depends on the stage of the disease during diagnosis; hence routine breast screening is important to detect early cancer changes with the aim of increasing survival rates. In the UK, the National Health Service Breast Screening Programme (NHSBSP) was set up in 1988 as a response to the recommendation of 'The Forrest Report' to implement a population based screening programme (Forrest 1986). The programme uses mammogram to screen for breast cancer every three years, initially to all women aged 50 to 64. Since 2004, this service has been extended to include women aged 65 to 70 and presently, all women aged 47 to 73 are eligible for breast screening. An assessment for the period 1990 to 1998 on the impact of NHSBSP revealed that the programme contributed to a substantial reduction on breast cancer mortality rate in women aged 55 to 69 (Blanks *et al.* 2000). Controversy about this programme has emerged in the recent years as questions were raised about the value of breast screening and possible harmful effects to some women. Mammographic screening is effective in detecting lesions in the breast but lacks the specificity to differentiate between low grade *in situ* carcinoma and high grade invasive cancer. This inevitably can lead to overdiagnosis and overtreatment of cancer in some women who would never have presented clinically. Mammography itself carries a small carcinogenic risk from radiation and has been estimated to cause 3 to 6 extra breast cancers for every 10,000 women enrolled in the screening programme (Cancer Research UK 2012). As a result, a group of independent experts were brought together to review the benefits and effectiveness of the programme. The review panel found that out

of 15,500 breast cancers diagnosed through screening, 4,000 cases were overdiagnosed while 1,300 breast cancer deaths are prevented each year (Cancer Research UK 2012). The panel concluded that the NHSBSP has significant benefit to the country as a whole and justifies the continuance of the screening programme.

1.2.4 Pathology

1.2.4.1 Types of breast cancer

Most breast cancers arise from an epithelial origin in cells lining the ducts or the lobules of the breast. Common changes in the breast begin with a rapid increase in breast epithelial cells (hyperplasia) that leads to the emergence of abnormal cells (atypical hyperplasia), followed by carcinoma *in situ* (non-invasive breast cancer) and infiltrating carcinoma (invasive breast cancer) [Figure 1.2].

Ductal carcinoma *in situ* (DCIS) is the most common type of non-invasive breast cancer, contributing to 20% of screen-detected cancers in the NHSBSP. DCIS is characterised by expanded breast ducts resulting from uncontrolled growth of ductal epithelial cells confined within the ducts. Several subtypes of DCIS have been described, ranging from low grade to high grade lesions which may give rise to invasive breast cancer. One third of patients with DCIS go on to develop invasive breast cancer in the same breast (Page *et al.* 1995).

Lobular carcinoma *in situ* (LCIS) is a rare type of non-invasive breast cancer, comprising only 1% of cancers detected by screening, presumably because it is not easily seen on a mammogram. LCIS presents as an expansion of the breast lobules by abnormal proliferation of epithelial cells within the lobules. Patients with LCIS confer an increased risk of invasive breast cancer, where 15% to 20% develops invasive carcinoma in the same breast and 10% to 15% in the contralateral breast (Page *et al.* 1995).

Invasive breast cancer is characterised by the infiltration of cancer cells outside the basement membrane of ducts or lobules into surrounding breast tissues and possibly spreading into lymph nodes or other body parts. Invasive breast cancers are classified into several distinct types depending on the characterisation of the tumour mass. Invasive ductal carcinoma (IDC) is the most common type of breast cancer, accounting for about 50% to 75% of all breast cancers, while invasive lobular carcinoma (ILC) is

the next most common and accounts for about 10% to 15% of cases (Dillon *et al.* 2010). Other less common types of invasive breast cancer are IDC subtypes characterised by their distinct pattern of growth and cell morphology: medullary carcinoma, mucinous carcinoma, tubular carcinoma, papillary carcinoma and cribriform carcinoma. Inflammatory breast cancer is another form of rare but aggressive breast cancer associated with breast inflammation due to the blockage of lymph ducts in the breast by cancer cells. Paget's disease is also associated with breast cancer, found only in 1% to 2% of patients. Paget's disease begins in the ducts of the nipple, spreads to the nipple surface and areola, and presents as an eczematoid change in the skin surrounding the nipple.

Legend:

A. Pectoralis major muscle
B. Adipose tissue
C. Lactiferous sinus
D. Ducts
E. Lobules
F. Ductal epithelial cells

(i) Normal cells
(ii) Hyperplasia
(iii) Atypical hyperplasia
(iv) Carcinoma *in situ*
(v) Microinvasion
(vi) Invasive breast cancer

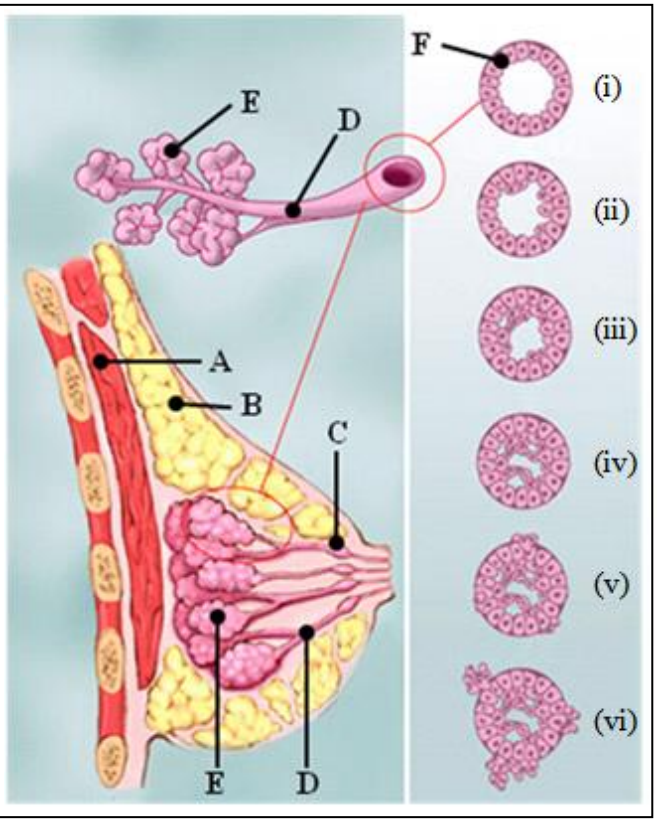


Figure 1.2 Development of ductal carcinoma *in situ* and invasive breast cancer.

The cross section of the breast is shown on the left panel. The right panel shows the organisation of epithelial cells in the mammary duct, progressing from hyperplasia, ductal carcinoma *in situ* to invasive breast carcinoma. Image adapted from (breastcancer.org).

1.2.4.2 Breast cancer grading and staging

The grading and staging of breast cancer is important to assess patient prognosis and determine treatment plan. Breast tumours are graded based on the morphology and proliferation rate of the cancer cells, with increased grade indicating poorer prognosis. Grade 1 (low grade) are slow growing, well differentiated cells; grade 2 (intermediate grade) are slighty faster growing, moderately differentiated cells; and grade 3 (high grade) are fast growing poorly differentiated cells. Staging of breast cancer describes the size and extent of spread beyond the site of a primary breast tumour. Breast cancer stages are frequently assessed using the Nottingham Prognostic Index (NPI) which predicts disease prognosis based on tumour size, number of lymph nodes involved and tumour grade (Haybittle *et al.* 1982). The commonly used breast cancer staging method is the tumour, lymph nodes and metastasis (TNM) staging system. This system considers the size of the primary tumour, the presence of lymph node metastasis and the presence of distant metastases to give an overall stage of the disease (Bateman 2006). Patients with stage I breast cancer have the best prognosis while patients with stage IV breast cancer, classified by distant metastases, have the poorest prognosis (Figure 1.3).

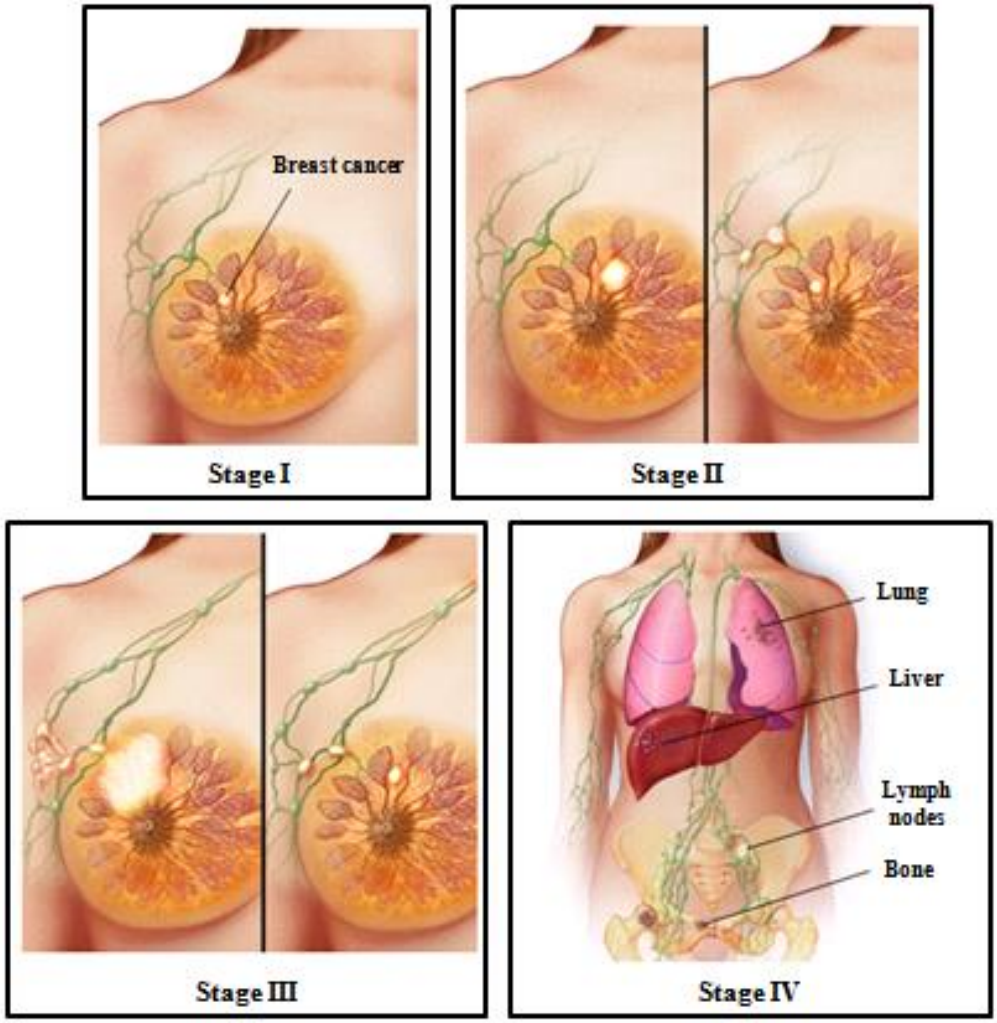


Figure 1.3 Illustration of breast cancer stages.

Stage I represents a tumour that is smaller than 2cm in diameter and has not spread outside the breast. Stage II represents a tumour that is larger than 2cm but not larger than 5cm and no cancer in the lymph nodes; or a tumour that is 2cm or smaller and found in 1 to 3 lymph nodes in the armpit or near the breast bone. Stage III represents a tumour that is larger than 5cm and small clusters of cancer cells in the lymph nodes; or a tumour that is smaller that 5cm but has spread to the lymph nodes above the collarbone. Stage IV represents a tumour of any size that has metastasised to other parts of the body such as lungs, liver, bones or brain. Image taken from (mayoclinic.com).

1.2.4.3 Tumour characteristics

The characteristics of breast cancer are also crucial prognostic factors that affect patient treatment and outcome. Specifically, hormone receptor and human epidermal growth factor receptor 2 (HER2) status are routinely assessed to assist in breast cancer

treatment strategies. Hormone receptor status refers to the significant presence or absence of oestrogen and progesterone receptors on the surface of tumour cells. Oestrogen receptor (ER) positive and progesterone receptor (PR) positive tumours generally have more favourable prognosis and lower chance of recurrence than ER negative and PR negative tumours because of their predicted response to hormone therapies. HER2 status indicates the presence or absence of cell surface receptor regulated by epidermal growth factor (EGF) to stimulate cell growth and proliferation. Although HER2-positive tumours are expected to respond to treatments that specifically target HER2, they are generally associated with an adverse prognosis (Chang *et al.* 2010).

1.2.5 Breast cancer treatment

The ultimate goal of breast cancer treatment is to eradicate the cancerous cells and minimise the possibility of recurrence. Complexity in treatment strategy arises from the fact that there is no single factor for breast cancer. Hence, decision on the course of treatment takes many factors into consideration, including the type, size and location of the cancer; the grade and stage of the cancer; the characteristics of the cancer cells; and the patient's menopausal status and overall health. The common approaches used in treating breast cancer are surgery, radiation therapy, chemotherapy, endocrine therapy and targeted therapy, in which patients could be given one or several treatments sequentially or in combination.

1.2.5.1 Surgery

Some patients begin their treatment with surgery to remove as much tumour as possible from the breast, while some others undergo chemotherapy to reduce the size or extent of the cancer prior to surgery. Surgery involves the removal of the tumour and surrounding normal tissue while leaving the breast intact (lumpectomy) or the removal of the entire breast (mastectomy). Lumpectomy is normally recommended to patients with early breast cancer in a single location, a tumour that is less than 5cm in diameter and patients who are willing to undergo follow up radiation therapy (Breast Cancer Review 2010). Mastectomy is usually recommended for women with more advanced breast cancer and involves removing the breast, nipple, areola and the skin of the breast but maintains the

function of the pectoralis major muscle. Patients with invasive breast cancer can also undergo a procedure known as axillary lymph node dissection, usually performed together with a lumpectomy or mastectomy, to determine the spread of the cancer from the breast and guide follow up treatment approach.

1.2.5.2 Radiation therapy

Radiation therapy is commonly administered as post-operative adjuvant treatment, utilising high energy x-rays to destroy any remaining cancer cells in the breast, chest wall or axilla. Clinical trials indicate that post-operative radiation therapy significantly reduces both 5-year recurrence and 15-year mortality rates (Darby *et al.* 2011). Radiation therapy is also used as pre-operative neo-adjuvant treatment in certain breast cancer patients to shrink large tumours to an operable size. The two main types of radiation method used today are external radiation therapy and internal radiation therapy. External radiation relies upon an external source to deliver high energy x-rays that targets a focused beam to the tumour affected area, while causing minimal damage on adjacent healthy tissues. Internal radiation, also known as brachytherapy, places a radioactive source inside the tumour site. Implanted radioisotopes deliver radiation directly into the site of surgical removal or into the tumour mass. Short term clinical analysis suggests that the efficacy of brachytherapy is comparable with whole breast irradiation and causes negligible late side effects (Strnad *et al.* 2011). However, brachytherapy is not yet considered a standard treatment method for breast cancer as its long term effects are still not known.

1.2.5.3 Chemotherapy

Chemotherapy uses pharmacologic or natural agents to inhibit the development of invasive breast cancer. Chemotherapy is used as adjuvant, neo-adjuvant or palliative treatment that reverses, suppresses or prevents carcinogenic progression of premalignant cells (Sporn 1976). It is also used as the primary treatment in the management of metastatic or recurrent breast cancer. Unlike radiation therapy and surgery which are localised treatments, chemotherapy utilises a systemic approach that affects all tissues and organs throughout the entire body. Types of chemotherapy drugs are divided into alkylating agents, anthracyclines, antimetabolites and taxanes; categorised according to

their effect on cell division or DNA synthesis and function. These drugs are usually given as a specific combination tailored to each patient individually and administered intravenously, intramuscularly, subcutaneously or orally.

1.2.5.4 Endocrine therapy

Steroid hormones such as oestrogen and progesterone have been implicated in the progression of breast cancer due to their effect in stimulating cell growth, differentiation and function in the breast. Understanding their mechanism of action has enabled the development of endocrine therapies for breast cancer that either reduce hormone production or block their action. Endocrine therapy is only effective in treating cancer cells that depend on these hormones to grow, commonly classified as ER and/or PR-positive tumours. Endocrine therapy can be used in addition to surgery, radiation therapy and chemotherapy; or as chemoprevention of breast cancer in high risk patients. The types of endocrine therapy used today are known as selective oestrogen receptor modulator (SERM), selective oestrogen receptor downregulator (SERD), aromatase inhibitors and ovarian ablation.

SERMs are drugs that act as receptor binding competitors of oestrogen and antagonise their downstream effects. Tamoxifen is a non-steroidal anti-oestrogen SERM most commonly used in both pre- and post-menopausal women as prevention and treatment of breast cancer (Cole *et al.* 1971). However, the partial agonist action of tamoxifen is associated with some serious side effects including endometrial cancer, thrombo-embolism and bone loss (Singh *et al.* 2005).

SERDs are anti-oestrogen drugs that antagonise ER by downregulating and degrading the receptor. Fulvestrant is a widely used steroidal anti-oestrogen SERD that has a higher affinity to ER with no agonist effect in the endometrium (Wakeling *et al.* 1992). It is administered to post-menopausal breast cancer patients who no longer respond to tamoxifen as fulvestrant is better tolerated and has less life-threatening side effects.

Aromatase inhibitors are drugs that block the synthesis of oestrogen from androgens by inhibiting the action of aromatase enzyme, which is the main source of oestrogen production in post-menopausal women. Aromatase inhibitors are used either as initial treatment or after tamoxifen and are associated with mild side effects.

Ovarian ablation or ovarian suppression is a form of therapy that stops the ovaries from producing oestrogen in pre-menopausal women. Oestrogen production could be temporarily suppressed by drugs that inhibit luteinising hormone-releasing hormone (LHRH) signalling or permanently suppressed by radiation and surgical removal of the ovaries (oophrectomy). This treatment shuts down the main source of the body's oestrogen, consequently depriving ER and/or PR-positive tumours of oestrogen that supplements its growth.

1.2.5.5 Targeted therapy

Greater understanding of the biology of breast cancer has led to the identification of molecular targets and the development of targeted therapies in treating the disease. Targeted therapy is directed to specific molecules or pathways in certain breast cancer cells while leaving the healthy cells unharmed. Most of the targeted therapies are not yet approved for clinical practice but some agents have been used as standard care in HER2-positive breast cancer patients. Two therapeutic agents targeting HER2-positive patients in standard breast cancer practice are trastuzumab and lapatinib. Trastuzumab is a monoclonal antibody that targets cancer cells overexpressing HER2 on their cell surface and is effectively used as first treatment for HER2-positive metastatic breast cancer in combination with chemotherapy. Lapatinib is a dual tyrosine kinase inhibitor that targets both epidermal growth factor receptor (EGFR) and HER2, currently approved for use in HER2-positive metastatic breast cancer in combination with trastuzumab or other chemotherapeutic drugs. However, the effectiveness of trastuzumab in metastatic breast cancer cases was only observed in one third of HER2-positive patients and resistance to this drug may be acquired by patients over a period of treatment [reviewed in (Barros *et al.* 2010)].

1.3 Oestrogen and breast cancer

Oestrogen is a potent steroid hormone with 17β-oestradiol being the most prominent metabolite in the human body, although lower levels of oestrone and oestriol are also present. Overwhelming epidemiological and experimental evidence implicates a crucial role of oestrogen in the progression of breast cancer. Levels of endogenous oestrogen are strongly associated with increased breast cancer risk in post-menopausal women

(Key *et al.* 2002), while anti-oestrogenic drugs reduce the risk of ER positive breast cancer (Cuzick *et al.* 2003).

1.3.1 Oestrogen receptors (ER)

ER belongs to the nuclear receptor superfamily of ligand-inducible transcription factors. Two major forms of ER have been identified to date, namely ER-alpha (ER-α) and ER-beta (ER-β), encoded by oestrogen receptor 1 (*ESR1*) and oestrogen receptor 2 (*ESR2*) respectively. Although they both share a high degree of overall homology, the two receptors have distinct physiological functions. ER-α knockout mice were reported to possess under-developed mammary glands and in contrast, ER-β knockout mice appear to undergo normal mammary development (Bocchinfuso *et al.* 1997, Krege *et al.* 1998). ER-α is essential for proliferation while ER-β has anti-proliferative role; and these two receptors regulate different target genes in response to oestrogen or anti-oestrogens (Monroe *et al.* 2003, Tee *et al.* 2004). Such evidence led ER-α to be the main form of receptor that determines ER status in breast tumours and the primary target for endocrine therapies in breast cancer.

The gene encoding ER-α consists of 8 exons spanning 140kb on human chromosome 6q25.1 and produces a 596 amino acid protein that consists of a six functional multidomain structure, as illustrated in Figure 1.4 (Green *et al.* 1986, Greene *et al.* 1986, Menasce *et al.* 1993). The variable N-terminal A/B domain contains the ligand-independent transactivation function (AF-1) which interacts with components of the core transcriptional machinery or other transactivators to activate target genes. The C domain contains a DNA-binding domain (DBD) composed of two zinc finger motifs, an essential component of ER that interacts specifically with DNA sequences in the oestrogen response element (ERE). Downstream of the DBD region is a hinge region (D domain) that acts as a linker peptide and contains the nuclear localisation signal. The E domain is a relatively large region that harbours the ligand-binding domain (LBD) and the ligand-dependent transactivation function (AF-2). The LBD is a globular domain that contains the ligand-binding site, a dimerisation interface and a coregulator interaction function. The LBD acts as a molecular switch that shifts the receptor to a transcriptionally active state upon hormone recognition and binding. The C-terminal end contains the F domain where its specific function has not been identified, as the deletion of F domain did not affect known ER function (Kumar *et al.* 1987).

Numerous mRNA splice variants of ER-α have been characterised, arising from alternative splicing or alternative promoter. The ER-α gene is transcribed from at least seven upstream promoters into multiple transcripts varying in their five prime untranslated region (5'-UTR) (Kos *et al.* 2001). Most ER-α splice variants in breast cancer cells are products of exon skipping (Poola *et al.* 2000). The classical full-length ER-α has a molecular weight of 66 kiloDalton (kDa), therefore alternatively referred to as ER-α66. Splice variant ER-α46 is a 46kDa protein, identified in MCF-7 breast cancer cell line, that lacks the first coding exon (exon 1A), therefore lacking the N-terminal AF-1 domain (Flouriot *et al.* 2000). ER-α46 appears to inhibit the AF-1 transactivating function of ER-α66 either through direct competition for ER-α DNA-binding site or in a dominant-negative manner involving heterodimerisation of the two isoforms (Fuqua *et al.* 1992). The expression ratio of ER-α46 and ER-α66 changes with cell growth status in MCF-7 breast cancer cells, suggesting a role of ER-α46 in cell proliferation (Flouriot *et al.* 2000).

More recently, Wang *et al.* have identified and cloned another variant of ER-α, in MCF-7 breast cancer cell line, that has a molecular weight of 36kDa and termed it ER-α36 (Wang *et al.* 2005). This ER-α36 transcript is a product of previously unidentified promoter within the first intron of *ESR1* gene that encodes a protein lacking both AF-1 and AF2 domains but retaining the DBD, LBD and partial dimerisation domains. The last 138 amino acids encoded by exons 7 and 8 of *ESR1* are replaced by a unique 27 amino acid domain (Wang *et al.* 2005). ER-α36 is primarily expressed in the cytoplasm and on the plasma membrane, where it transduces membrane-initiated effects of non-genomic oestrogen-mediated signalling. Further studies by Wang *et al.* have shown this in the oestrogen-dependent activation of the mitogen-activated protein kinase/extracellular signal-regulated kinase (MAPK/ERK) pathway (Wang *et al.* 2006). Tamoxifen treatment in ER-α36 overexpressing cells failed to block the ER-α36 mediated activation of MAPK/ERK pathway; instead it promoted cell growth (Wang *et al.* 2006). High levels of ER-α36 in ER positive breast cancer patients may contribute to tamoxifen resistance, while this promotes malignant growth in ER negative breast cancer cell lines (Shi *et al.* 2009, Zhang *et al.* 2011).

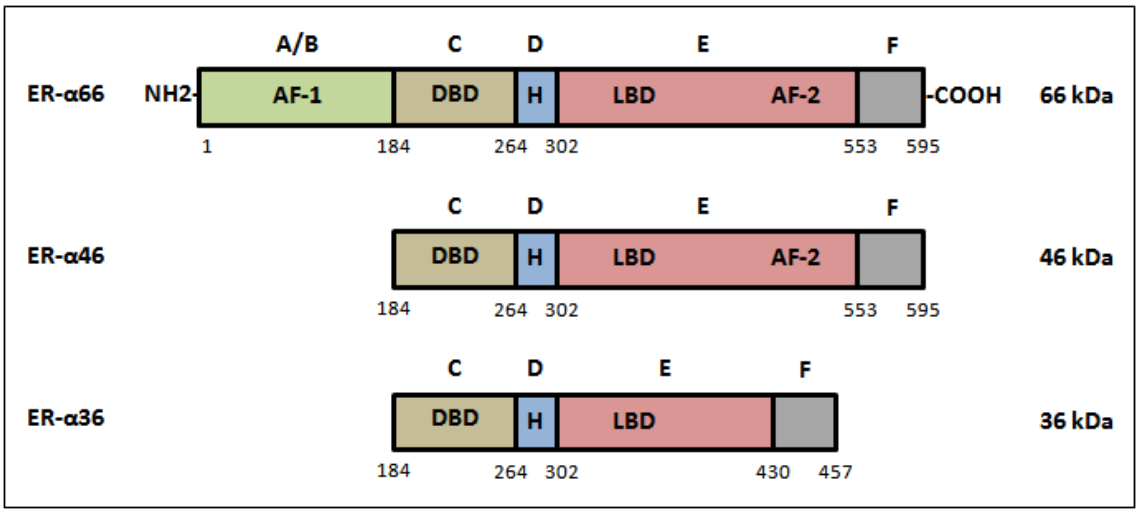


Figure 1.4 Structure of ER-α and its isoforms.

Similar to other members of the nuclear receptor family, ER-α display six conserved functional multidomain structure: A/B domain containing the AF-1, C domain containing the DBD, D domain (hinge region) containing the nuclear localisation signal, E domain containing the LBD and AF-2. ER-α46 is a truncated form of ER-α lacking the AF-1 function. ER-α36 isoform lacks both AF-1 and AF-2 functions. Image adapted from (Le Romancer *et al.* 2011).

1.3.2 Genomic action of oestrogen

The classical model of oestrogen action involves the activation of genomic signalling through high-affinity binding to ER and recognition of ERE. The non-active ER exists as a heterocomplex molecule associated with heat shock proteins and immunophilin chaperones, particularly heat shock protein 90 (Hsp90) and immunophilin-FK-binding protein 52 (FKBP52) (Gougelet *et al.* 2005, Pratt *et al.* 1997). These chaperones help to maintain ER in the appropriate conformation for rapid response to hormonal signals. Nuclear localisation and nuclear export signals enable the inactive ER complex to shuttle between the nucleus and cytoplasm.

Oestrogen diffuses through the plasma membrane and cytoplasm into the nucleus where it binds the ER LBD. Upon oestrogen binding, a ligand-specific conformational change of ER occurs, Hsp90 and FKBP52 dissociate and ER is transformed into its active form (Pratt *et al.* 1997). These modifications trigger homo- or hetero-dimerisation of ER and binding to specific ERE in target genes. The ERE-bound ER dimer complex facilitates the association of coactivators that stimulate gene transcription, including nucleosome remodelling, histone acetyltransferase (HAT) or methyltransferase (Figure 1.5,

pathway A). In contrast, the binding of ER to oestrogen antagonists induces a distinct conformational change that switches off gene transcription through the association of corepressors such as nuclear receptor corepressor (NCoR) and silencing mediator of retinoic acid and thyroid hormone receptor (SMRT) [reviewed in (Lonard *et al.* 2006, Perissi *et al.* 2010)].

Oestrogen-activated ER can also indirectly regulate gene promoters through protein-protein interactions with other DNA-bound transcription factors. In this non-classical genomic pathway, nuclear ER interacts with activator protein-1 (AP-1), specificity protein 1 (Sp1) or nuclear factor kappa B (NF-κB) through transcription factor crosstalk (Kushner *et al.* 2000, Quaedackers *et al.* 2007, Saville *et al.* 2000). This interaction stabilises the DNA-binding of the tethered transcription factor and alters the rate of gene transcription without direct ER DNA-binding (Figure 1.5, pathway B). Interestingly, increased AP-1 and NF-κB transcriptional activities in ER positive breast cancer are involved in endocrine resistance (Schiff *et al.* 2000, Zhou *et al.* 2007).

ER activities are also regulated by ligand-independent pathways which involve the modulation of ER gene transcription through the MAPK signalling pathway activated by growth factors such as EGF, insulin-like growth factor 1 (IGF-1) and transforming growth factor-β (TGF-β) (Figure 1.5, pathway C) (Edwards 2005, Kato *et al.* 1995). MAPK signalling cascade phosphorylates a serine residue within ER at position 118 and fully activates the AF-1 domain of ER without ligand-activation, triggering gene transcription (Kato *et al.* 1995). Recently, this ligand-independent ER activation pathway has been associated with acquired tamoxifen resistance in breast cancer (Rhodes *et al.* 2011).

1.3.3 Non-genomic action of oestrogen

Increasing evidence shows that oestrogen can elicit a rapid signalling response through a non-genomic action mediated by membrane-associated ER. Endogenous ER are localised in the nucleus and extra-nuclear compartments including the plasma membrane, mitochondria and endoplasmic reticulum. Approximately 5% to 10% of cellular ER is present at the plasma membrane and exhibits the same high-affinity for oestrogen as nuclear ER. Breast cancer cells express more ER-α than ER-β at the plasma membrane, consistent with the ER-α mediated rapid oestrogen signalling in

these cells (Marquez *et al.* 2006). Isolation of membrane-bound ER by affinity chromatography identified ER-α66 as the predominant protein at the plasma membrane (Pedram *et al.* 2006). Other truncated forms of ER-α, ER-α46 and ER-α36 have been reported to be present at the plasma membrane and these receptors also contribute to the rapid response of oestrogen (Marquez *et al.* 2006, Wang *et al.* 2006).

Caveolin-1 functions to facilitate transport of ER to the caveolae rafts in the plasma membrane and is then displaced from ER for productive signalling. In the absence of oestrogen, endogenous ER-α at the plasma membrane exists as a monomer, however it rapidly dimerises upon the presence of oestrogen (Razandi *et al.* 2004). Dimerised membrane ER-α interacts with various signalling proteins, including G proteins, tyrosine kinases (Src and Ras), growth factor receptors (EGFR, IGF-1), phosphatidylinositol 3-kinase (PI3K) and other adaptor proteins, to activate multiple signal transduction cascades (Bjornstrom *et al.* 2005, Migliaccio *et al.* 1996, Simoncini *et al.* 2000). The activation of signal transduction pathways triggers biological responses that are either dependent or independent of transcription (Figure 1.5, pathway D).

Signal transduction pathways connect this non-genomic action of oestrogen to genomic responses through the regulation of protein kinase-mediated phosphorylation of many transcription factors. This non-genomic-to-genomic signalling by oestrogen enables ER to regulate transcription at alternative response elements [reviewed in (Bjornstrom *et al.* 2005)]. In breast cancer cells, membrane ER can indirectly influence nuclear ER gene expression through MAPK signalling pathway activated by membrane ER interaction with growth factor receptors such as HER2, EGFR and IGF-1 (Kahlert *et al.* 2000, Keshamouni *et al.* 2002). Activation of MAPK signalling pathway also results in phosphorylation of AP-1 that enhances AP-1 DNA binding activity and transcriptional activation (Bjornstrom *et al.* 2004). The activation of Src kinase by membrane ER phosphorylates and recruits coactivators to the nuclear ER transcriptional complex, augmenting its transcriptional function (Zheng *et al.* 2005). PI3K activation by oestrogen targets NF-κB phosphorylation that leads to enhanced gene expression of NF-κB targets (Kawagoe *et al.* 2003).

Membrane ER mediated signal transduction pathways also trigger rapid oestrogen responses through post-translational modifications that affect cell distribution and function of the substrate protein. Oestrogen-mediated activation of PI3K leads to the

phosphorylation of B-cell lymphoma-2 (Bcl-2) associated death (BAD) protein; phosphorylation of BAD inactivates its pro-apoptotic function and favours cell survival, therefore abrogating apoptosis in breast cancer cells (Fernando *et al.* 2004). Oestrogen-activated PI3K/Akt signalling also restraints the ataxia telangiectasia-mutated and Rad-3 related (ATR) kinase cascade, inhibiting DNA damage repair and cell cycle checkpoints in breast cancer (Pedram *et al.* 2009). The interaction of histone deacetylase (HDAC) 6 with oestrogen-activated membrane ER causes rapid deacetylation of tubulin, which potentially contributes to cell migration and the aggressiveness of ER positive breast cancer cells (Azuma *et al.* 2009). Another role of non-genomic oestrogen signalling is observed in the rapid enhancement of aromatase enzymatic activity by oestrogen-stimulated Src that increases phosphorylation of aromatase protein in breast cancer cells (Catalano *et al.* 2009). This evidence revealed a short non-genomic autocrine loop between oestrogen and aromatase involved in breast tumourigenesis.

In addition to ER, a novel transmembrane protein known as G protein-coupled receptor 30 (GPR30) also mediates a rapid response to oestrogen (Filardo *et al.* 2002). GPR30 was initially known as an orphan receptor since no ligand was identified. Filardo *et al.* was the first to demonstrate a possible activation of the rapid non-genomic signalling through MAPK pathway by GPR30 in response to oestrogen and anti-oestrogens such as fulvestrant and tamoxifen (Filardo *et al.* 2000). This study reported that 17β-oestradiol rapidly activates ERK-1 and ERK-2 in ER negative SKBR breast cancer cells (Filardo *et al.* 2000). The association between rapid oestrogen response with the presence of GPR30 was confirmed when MDA-MB-231 breast cancer cells, which are GPR30 deficient and insensitive to oestrogen activation of ERK-1/2, display an oestrogen-responsive phenotype after GPR30 was transfected into the cells (Filardo *et al.* 2000). Subsequent reports support this notion and show that fulvestrant is able to bind GPR30 and mimic oestrogen effects, thus acting as an agonist to GPR30 (Filardo *et al.* 2002, Lucas *et al.* 2010, Thomas *et al.* 2005). Other studies suggest that GPR30 is strongly associated with cancer cell proliferation, migration, invasion, metastasis and drug resistance (Filardo *et al.* 2006, Filardo *et al.* 2002, He *et al.* 2009, Lapensee *et al.* 2009, Maggiolini *et al.* 2004). However, a recent study by Kang *et al.* showed that the reported activities of GPR30 were actually mediated through its ability to induce ER-α36 expression and non-genomic oestrogen signalling was mediated by ER-α36, not GPR30 (Kang *et al.* 2010)

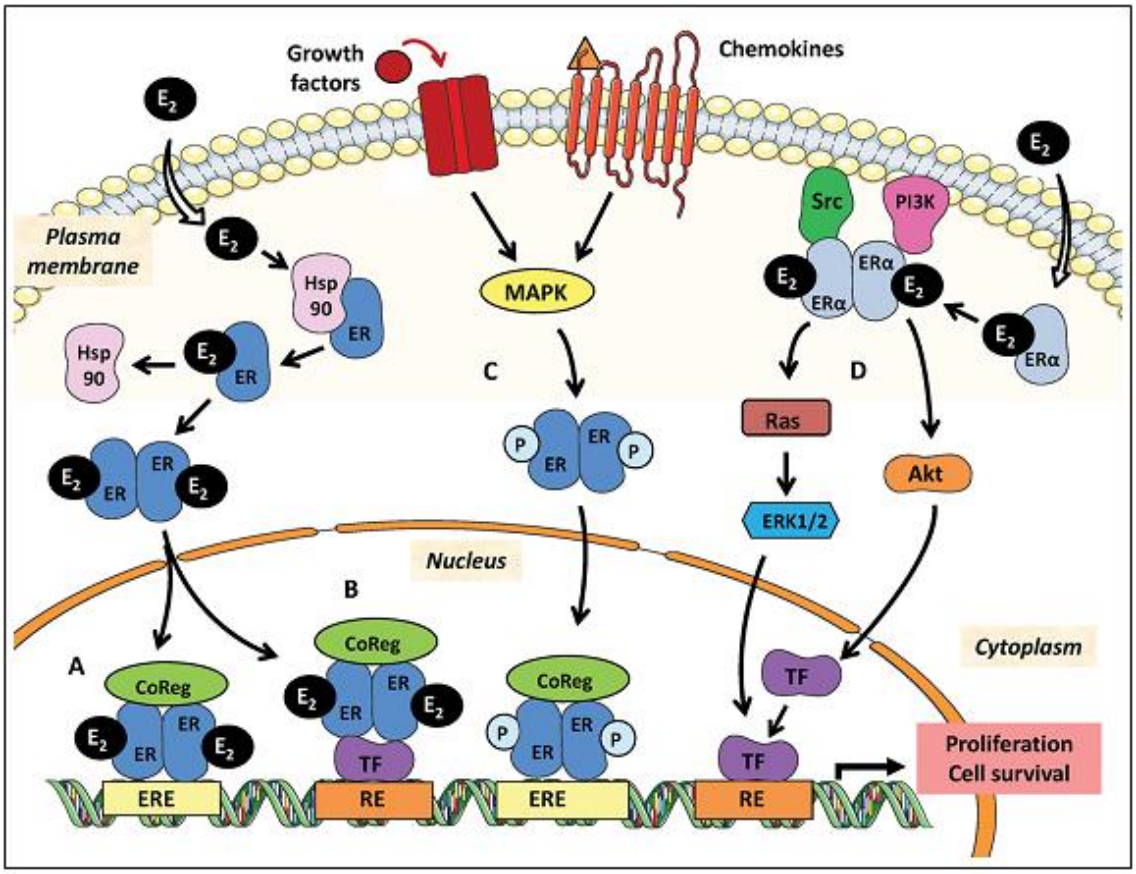


Figure 1.5 Genomic and non-genomic ER signalling pathways.

Illustration shows four distinct oestrogen signalling pathways through ER. Pathway A is ligand dependent, classical genomic pathway; where oestrogen-bound ER dimerises, enters the nucleus and binds directly to ERE to activate gene transcription. Pathway B is ligand dependent, non-classical genomic pathway; where ER interacts and tethers to other transcription factors to activate target genes. Pathway C is ligand independent, non-classical genomic pathway; where ER is activated through phosphorylation induced by growth factors. Pathway D is ligand dependent, non-genomic pathway; where extra-nuclear ER dimerises upon ligand binding and activates multiple signal transduction cascades through interactions with various proteins. Image taken from (Le Romancer *et al.* 2011).

1.3.4 Anti-oestrogen action in breast cancer

Approximately 70% of all breast cancer cases are ER-α positive and these breast tumours utilise oestrogen and functional ER-α for growth. These patients usually receive hormone reduction or anti-oestrogens such as SERMs or SERDs as part of their treatment plan (Cole *et al.* 1971, Wakeling *et al.* 1991). The anti-tumour effects of these anti-oestrogen drugs are mediated by the inhibition or alteration of ER activity that

subsequently affects gene transcription. SERMs such as tamoxifen are commonly thought to inactivate ER by competitive binding of the LBD, blocking the transactivation function of AF-2 while inducing an inactive conformational change which is unfavourable for coactivator interaction (MacGregor *et al.* 1998). Recent evidence suggests that tamoxifen also induces ligand-specific conformational change in ER that allows the exposure of unique surfaces for corepressor interaction to modulate the tamoxifen antagonist activity of ER (Huang *et al.* 2002, Oesterreich *et al.* 2000, Shang *et al.* 2002). Expression profiles of ER coregulators in specific cell types also determine the cellular response to tamoxifen-bound ER (Katzenellenbogen *et al.* 1996, Shang *et al.* 2002). These findings extend the common view of tamoxifen as an ER antagonist with an active anti-oestrogen signalling function.

The pure anti-oestrogen SERD, fulvestrant, is a pure ER antagonist that has a stronger affinity for ER than tamoxifen (Wakeling *et al.* 1991). Fulvestrant binds to newly synthesised ER in the cytoplasm and induces a ligand-specific conformational change that disrupts ER shuttling to the nucleus (Dauvois *et al.* 1993). The accumulated ER in the cytoplasm undergoes increased receptor turnover and the ligand-receptor complex is targeted for rapid destruction, thus reducing the cellular levels ER. In addition, fulvestrant blocks both AF-1 and AF-2 function of ER, thus completely abrogating the transcription of ER-regulated genes through the inhibition of receptor dimerisation and restriction of ER-binding to target DNA (Parker 1993, Wakeling *et al.* 1991). Consequently, any fulvestrant-bound ER that enters the nucleus is transcriptionally inactive and unable to exert its classical genomic action. This is a favourable characteristic of fulvestrant from the partial ER antagonist tamoxifen, where receptor activation could be achieved through the unrestricted AF-1 domain.

1.3.5 Anti-oestrogen resistance in breast cancer

Resistance to anti-oestrogen therapy in breast cancer patients can result from intrinsic or acquired mechanisms. Intrinsic resistance occurs when the patient lacks ER-α expression before any treatment is given and carries inactive alleles of cytochrome P450 2D6 (CYP2D6) that inhibits tamoxifen conversion to its active metabolite (Hoskins *et al.* 2009, Stewart *et al.* 1982). In contrast, a wide range of mechanisms have been postulated for acquired resistance of anti-oestrogen therapy. Deregulation of various components of the oestrogen signalling pathway, alterations in cell cycle signalling

molecules and activation of alternative survival pathways are among the mechanisms that confer anti-oestrogen resistance in breast cancer [reviewed in (Osborne *et al.* 2011)].

The deregulation of ER and its coregulators modulate the response to anti-oestrogen therapy. The loss of ER expression in ER positive breast cancer patients during tumour progression contributes to anti-oestrogen resistance (Encarnacion *et al.* 1993, Gutierrez *et al.* 2005). The presence of the ER splice variant, ER-α36, has been implicated in reduced responsiveness to anti-oestrogen drugs (Shi *et al.* 2009). To further investigate this claim, Wang *et al.* found that tamoxifen functions as an agonist of ER-α36 in ER positive breast cancer cells and activates the non-genomic ER-α36 mediated MAPK signalling pathway, hence conferring resistance to the drug (Rao *et al.* 2011). Post-translational modifications of ER and its coregulators also regulate ER function and may contribute to anti-oestrogen resistance. Specific phosphorylation of at least four serine residues within the A/B domain of ER induced by oestrogen, anti-oestrogens, growth factor receptors and protein kinases influences ER sensitivity to anti-oestrogen therapy (Ali *et al.* 1993, Arnold *et al.* 1994, Cho *et al.* 1993, Riggins *et al.* 2007). Changes in the expression pattern of ER coregulators may influence the balance of agonistic versus antagonistic effect of anti-oestrogens and contribute to a resistant phenotype. The ER coactivator, steroid receptor coactivator-3 (SRC-3) or commonly known as amplified in breast cancer 1 (AIB1), is overexpressed in 64% of breast tumours and is associated with worst disease outcome in tamoxifen treated patients (Anzick *et al.* 1997, Osborne *et al.* 2003). *In vitro* studies have shown that high levels of AIB1 enhances the agonist activity of tamoxifen and results in drug resistance (Smith *et al.* 1997, Su *et al.* 2008, Webb *et al.* 1998). The recruitment of corepressors to the ER usually results in a repression of anti-oestrogen mediated agonist activity. Lavinsky *et al.* have shown that the decreased levels of NCoR correlated with acquired tamoxifen resistance in breast cancer (Lavinsky *et al.* 1998). Increased levels of AP-1 and NF-κB transcription factors, that mediate ER tethering to specific genes, have also been associated with anti-oestrogen resistance (Zhou *et al.* 2007). Interestingly, Jameson *et al.* observed a potent agonist effect of fulvestrant in non-classical genomic pathways by mediating ER interaction with AP-1 transcription factor (Jakacka *et al.* 2001). The activation of AP-1 by this "pure" ER antagonist is independent of its AF-1 and AF-2 function and suggests a possible mechanism of anti-oestrogen resistance.

Growth factor receptor and other cellular kinase pathways have been implicated in anti-oestrogen resistance by providing alternative proliferation and survival stimuli to breast tumours. Multiple regulatory interactions between ER, growth factors and protein kinases may stimulate ER-independent growth and regulate resistance through their signalling crosstalk. Overexpression of EGFR and HER2 has been reported in cells stimulated by anti-oestrogen drugs. Continuous supplementation of tamoxifen or fulvestrant in cultured ER positive breast cancer cells revealed a parallel increase in EGFR and HER2 protein that contributes to tolerance toward these anti-oestrogens (Knowlden *et al.* 2003, McClelland *et al.* 2001). In addition, rapid activation of EGFR and HER2 has been observed in tamoxifen-treated breast cancer cells (Shou *et al.* 2004). Activated EGFR and HER2 receptors trigger the MAPK signalling pathway that result in ER-independent stimulation of proliferation and facilitates acquired resistance to anti-oestrogens (Gutierrez *et al.* 2005). Such crosstalk between ER and growth factor signalling pathways establishes a self-propagating autocrine growth regulatory loop that drives cell growth in anti-oestrogen resistant cells.

Understanding the ER and growth factor signalling crosstalk in developing anti-oestrogen resistance has prompted a combinatorial approach in breast cancer treatment, targeting both ER and growth factor signalling pathways. Pre-clinical study has shown that gefitinib, an EGFR tyrosine kinase inhibitor, in combination with tamoxifen or fulvestrant synergistically inhibits *in vitro* cancer cell growth and the development of anti-oestrogen resistance (Gee *et al.* 2003). Current clinical trials are ongoing to investigate the benefit of using fulvestrant in combination with trastuzumab to overcome, delay or prevent the onset of anti-oestrogen resistance in breast cancer patients.

1.3.6 ER coregulators and breast cancer

The ligand-bound ER is modulated by the recruitment and interaction with coregulatory proteins that either enhance (coactivators) or repress (corepressors) its transactivation functions. More than 350 coregulators have been identified to date; however most of these coregulators are not specific to only the ER but also other nuclear receptors (Lonard *et al.* 2007, McKenna *et al.* 2002). The association of ER with coactivators and corepressors is regulated by ligand-specific conformational changes; where agonistic ligands recruit coactivators while antagonistic ligands recruit corepressors.

Several classes of ER coactivators have been well described, especially the SRC family of coactivators. This family consists of SRC-1, SRC-2 and SRC-3 (also known as AIB1) (Leo *et al.* 2000). The binding of SRC family members to ligand-bound ER is mediated by the interaction of AF-2 region in the LBD of the ER and the nuclear receptor (NR) box within coactivators, which consists of a conserved leucine-rich motif, LXXLL (where L is leucine, X is any amino acid) (Ding *et al.* 1998). Coactivators from the SRC family possess histone acetyltransferase (HAT) activity that acetylates histone proteins in the chromatin and weakens the association of histones with DNA. This alters nucleosomal conformation and stability, enhancing the formation of a stable preinitiation complex, and facilitating RNA polymerase II transcriptional activation (Korzus *et al.* 1998, Sternglanz 1996). Steroid receptor RNA activator (SRA) is another class of unique coactivator that interacts with the SRC-1 complex and activates transcription through the AF-1 domain as an RNA molecule instead of a protein (Lanz *et al.* 1999).

Altered expression of coactivators has been correlated with breast cancer progression. Low levels of SRC-1 were observed in breast tumours compared to its relatively high levels in normal breast tissues (Berns *et al.* 1998). Although there was no correlation between SRC-1 levels with ER status, patients with low SRC-1 levels did not respond to tamoxifen treatment. In contrast, SRC-3 or AIB1 mRNA is overamplified in ER positive breast cancer cell lines and 64% primary breast tumours (Anzick *et al.* 1997). A larger case study on breast and ovarian cancer correlated SRC-3 expression with tumour size and ER-α positivity (Bautista *et al.* 1998). Overexpression of SRA has also been described in breast tumours compared to normal breast tissues (Murphy *et al.* 2000).

The number of corepressors identified over the years has been growing steadily. In contrast to ER coactivators, ER corepressors do not seem to share a common interaction domain or mechanism for ER repression. While corepressors appear to prominently act through chromatin modification, they may also function through more than one mechanism and regulate transcription at additional levels. The best characterised ER corepressors are NCoR and SMRT, originally identified as binding factors to thyroid and retinoic acid receptors (Chen *et al.* 1995, Horlein *et al.* 1995). Both NCoR and SMRT contain a conserved bipartite NR interaction domain (NRID) which consists of the L/IXXI/VI motif (where L is leucine, I is isoleucine, V is valine, X is any amino acid) termed the CoRNR box (Hu *et al.* 1999, Li *et al.* 1997). Although this motif is

similar to the NR box within coactivators, the CoRNR box is predicted to form an extended helical structure that differs from the coactivator NR box and facilitates its molecular mechanism of recruitment and ER repression in the presence of anti-oestrogens (Perissi *et al.* 1999). These corepressors interact with different components of the HDAC protein complexes to facilitate chromatin condensation and inhibition of gene transcription (McKenna *et al.* 1999). Unlike most corepressors that do not exclusively interact with ER, the repressor of ER-α activity (REA) function is ER specific and acts by competing for ER binding sites with coactivators (Montano *et al.* 1999). REA competes directly with SRC-1 for ER binding sites and has been shown to reverse enhanced ER activity mediated by coactivators (Delage-Mourroux *et al.* 2000). Scaffold attachment factor B1 (SAFB1) and SAFB2 are corepressors that also bind and modulate ER activity through chromatin remodelling or their interaction with the basal transcription machinery, RNA polymerase II (Nayler *et al.* 1998, Oesterreich *et al.* 2000, Townson *et al.* 2003).

The possible significance of these corepressors in breast cancer development and progression has been suggested. Overexpression of NCoR and SMRT have been reported in intraductal carcinomas compared to normal mammary glands, but both NCoR and SMRT levels were subsequently downregulated during progression from intraductal to invasive ductal carcinomas (Kurebayashi *et al.* 2000). NCoR and SMRT bind strongly to ER-α in the presence of tamoxifen and levels of NCoR are substantially decreased in breast cancer cells resistant to prolonged tamoxifen treatment (Lavinsky *et al.* 1998). In this same study, overexpression of NCoR and SMRT has been shown to reverse the agonistic activity of tamoxifen. Similar pattern of expression was observed for REA corepressor, where levels were lower in high-grade tumours compared with low-grade tumours (Simon *et al.* 2000). REA also binds strongly to ER-α in the presence of tamoxifen and its overexpression increased antagonist activity of anti-oestrogens (Montano *et al.* 1999). Analysis of SAFB1 expression in invasive breast cancers revealed a significant correlation between low SAFB1 levels with shorter overall patient survival (Oesterreich S 2002). Overexpression of SAFB1 resulted in increased antagonistic activity and reversed agonistic activity of tamoxifen (Oesterreich *et al.* 2000). A subsequent study revealed that SAFB1 overexpression was able to block cell cycle progression and inhibit anchorage-dependent and -independent growth of breast cancer cells (Townson *et al.* 2000).

1.4 Scaffold attachment factor B1 (SAFB1) and SAFB2

1.4.1 SAFB family, gene and protein

The SAFB protein family comprise of three protein members: SAFB1, SAFB2 and SAFB-like transcription modulator (SLTM). SAFB1, originally known as SAFB, was first identified by Renz and Fackelmayer as a protein that binds scaffold/matrix attachment regions (S/MARs) (Renz *et al.* 1996). At the same time, Oesterreich *et al.* reported the same protein binding to the promoter of heat shock protein 27 (Hsp27) and termed it Hsp27 ERE TATA-binding (HET) protein (Oesterreich *et al.* 1997). Meanwhile, Weighardt *et al.* discovered this identical protein in a yeast two-hybrid screen using heterogenous nuclear ribonucleoprotein (hnRNP) A1 as a bait and named it hnRNP A1 associated protein (HAP) (Weighardt *et al.* 1999). SAFB, HET and HAP are identical proteins encoded by the same gene but given different names based on their observed function from three independent studies. Since the identification of a second family member, SAFB2, SAFB1 has been used as the approved nomenclature to replace the original SAFB/HET/HAP protein. A third member of the SAFB family, SLTM, was recently characterised for its inhibitory effect on gene transcription associated with the induction of apoptosis (Chan *et al.* 2007).

Both SAFB1 and SAFB2 are encoded by two separate genes that map to the same locus on 19p13.3 of the human chromosome (DuPont *et al.* 1997). *SAFB1* gene lies adjacent to the *SAFB2* gene, separated by a 490bp bidirectional promoter and positioned in a head-to-head arrangement (Figure 1.6) (Townson *et al.* 2003). The *SLTM* gene is located on the human chromosome 15q22.1.

The SAFB family are large proteins (>100kDa) that consist of multi-functional domains highly conserved among its members (Figure 1.6) [reviewed in (Garee *et al.* 2010, Oesterreich 2003)]. The N-terminus contains a scaffold attachment factor-box (SAF-box) that interacts with S/MARs, often found in proteins involved in chromatin organisation and regulating gene expression (Aravind *et al.* 2000). The central domain consists of an RNA-recognition motif (RRM) and nuclear localisation signal (NLS). The C-terminus comprise of glutamic acid/arginine-rich and glycine-rich regions that act as a potent transcriptional repression domain (Townson *et al.* 2004). SAFB2 shares 74% similarity with SAFB1 at the amino acid level, while SLTM shares 34% overall identity with SAFB1 and 36% with SAFB2 (Chan *et al.* 2007, Townson *et al.* 2003).

The SAFB family members are ubiquitously expressed in most tissues, with very high expression in the brain (Townson *et al.* 2003). Interestingly, *in silico* studies have suggested unique functions of each family member in cancerous and normal tissues. For example, SAFB1 is only expressed in cancerous bone marrow tissue while SAFB2 is only expressed in normal tissue and SLTM is expressed in both (Garee *et al.* 2010). Both SAFB1 and SLTM have punctated nuclear distribution that excludes from the nucleoli, whereas SAFB2 is also found in the cytoplasm (Chan *et al.* 2007, Nayler *et al.* 1998, Townson *et al.* 2003, Weighardt *et al.* 1999).

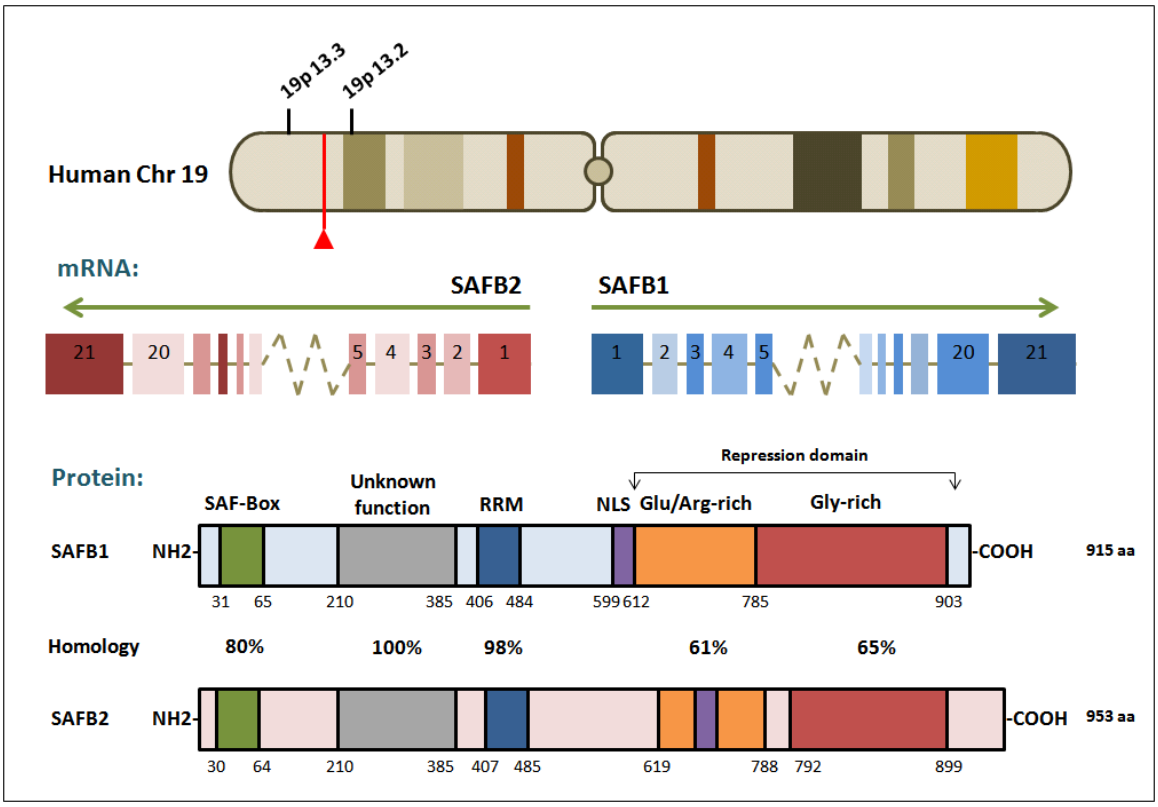


Figure 1.6 Schematic diagram of SAFB1 and SAFB2 protein structure.

SAFB1 and SAFB2 map to the same locus on 19p13.3, separated by a bidirectional promoter and oriented in a head-to-head arrangement. Each mRNA transcript contains 21 exons which translate to large multidomain proteins. SAFB1 and SAFB2 have homologous putative functional domains such as the scaffold attachment factor-box (SAF-box), RNA-recognition motif (RRM) and nuclear localisation signal (NLS), and Glu/Arg- and Gly-rich transcriptional repression domain. Image adapted from (Townson *et al.* 2003).

1.4.2 SAFB function

The presence of shared domains between SAFB family members suggests some common functions among SAFB1, SAFB2 and SLTM; however, limited information is available on SAFB2 and SLTM function. The major roles of the SAFB protein family will mainly be discussed based on the best characterised family member, SAFB1.

1.4.2.1 Chromatin organisation

The potential role of SAFB proteins in chromatin organisation has been speculated since its copurification with chromatin and nuclear matrix protein fractions, although ultimate proof has yet to be provided (Oesterreich *et al.* 1997, Renz *et al.* 1996). SAFB1 has been reported to bind AT-rich S/MARs that are important in the regulation of gene expression and disease progression including cancer (Gluch *et al.* 2008, Nayler *et al.* 1998). SAFB1 also binds to base unpairing regions (BURs), DNA regions similar to S/MARs that are critically important in higher order chromatin structure (Oesterreich, unpublished data). It was therefore proposed that SAFB1 may regulate gene expression through its effects on chromatin organisation via interaction with S/MARs and BURs (Figure 1.7, pathway A). Chromatin immunoprecipitation (ChIP) experiment showed significant SAFB1/SAFB2 enrichment on histone gene clusters of chromosome 1 and 6 (Hammerich-Hille *et al.* 2010). Further evidence has shown an interaction between SAFB1 and chromodomain/helicase/DNA-binding domain (CHD1) protein that binds AT-rich DNA motifs and involved in chromatin remodelling (Tai *et al.* 2003). Given that both CHD1 and SAFB1 associate with NCoR respectively, it is possible that a CHD1-SAFB1-NCoR repressive complex regulates the chromatin structure of SAFB1 target genes (Jiang *et al.* 2006).

1.4.2.2 Transcriptional regulation

The best described role for SAFB proteins is in transcriptional regulation of gene expression mediated by direct recruitment to the promoter region or indirect interaction with transcription factors [reviewed in (Hong *et al.* 2012)].

Oesterreich *et al.* was the first to observe SAFB1's ability to bind directly to regions of Hsp27 active promoter and significantly decrease its activity in several breast cancer

cell lines (Oesterreich *et al.* 1997). Affinity purification of the Enhancer-box (E-box) and Ku86 sites of the xanthine oxidoreductase (XOR) gene yielded SAFB1 as one of the binding protein (Lin *et al.* 2008). SAFB1 negatively regulates XOR gene expression through its binding to the E-box at the promoter and interaction with Ku86 sites. Omura *et al.* reported that SAFB1 binding to the promoter of sterol regulatory element binding protein-1c (SREBP-1c) positively regulate its expression when RNA-binding motif, X-linked (RBMX) is equally present in the cell (Omura *et al.* 2009). Although SAFB proteins may influence gene expression through direct binding to their target genes, Hammerich-Hille *et al.* observed a lack of significant similarities between SAFB1 and SAFB2 target genes identified through ChIP and genome-wide expression array (Hammerich-Hille *et al.* 2010). Therefore, they speculated that SAFB1 and SAFB2 recruitment to promoter regions alone may not completely mediate transcriptional regulation, but interaction with other transcription factors essentially contributes to their function (Figure 1.7, pathway B).

The protein-protein interaction between SAFB proteins with various transcription factors is facilitated by an intrinsic C-terminal repression domain (Townson *et al.* 2004). An investigation on the association of SAFB1 and ER-α action revealed a ligand-independent SAFB1 interaction with ER-α, notably strong at the DNA-binding/hinge regions of ER-α (Oesterreich *et al.* 2000). SAFB1 represses ER-α transcriptional activity through interaction with the ER-α DBD, although this doesn't appear to interfere with ER-α's ability to bind DNA. Similar observations were reported for SAFB2; hence SAFB1 and SAFB2 were categorised as ER-α corepressors (Townson *et al.* 2003). Interestingly, my previous findings show that SAFB1 and SAFB2 expression itself changes in response to 17β-oestradiol treatment in different breast cancer cell lines (Hong *et al.* 2010). A recent study reported that SAFB1 and SAFB2 interacts with ER-α in the presence of 17β-oestradiol and inhibit ER-α function by decreasing its intranuclear mobility (Hashimoto *et al.* 2012). Apart from ER-α, SAFB1 interacts promiscuously with several other nuclear receptors including peroxisome proliferator-activated receptor gamma (PPARγ), Farnesoid X receptor alpha (FXRα), RAR-related orphan receptor alpha 1 (RORα1), vitamin D receptor (VDR), liver receptor homolog-1 (LRH-1) and c-Jun (Debril *et al.* 2005). Interactions between the C-terminal region of SAFB1 and the multifunctional tumour suppressor p53 have also been reported (Peidis *et al.* 2011). This latter study showed that SAFB1 colocalised

with p53 under the treatment of 5-fluorouracil (5-FU) and repressed p53-dependent transcriptional activity.

1.4.2.3 RNA metabolism

Another role for SAFB proteins is found in RNA splicing and metabolism, speculated by their interaction with various RNA processing proteins (Figure 1.7, pathway C). Nayler *et al.* first discovered the interaction between SAFB1 with RNA polymerase II and a subset of serine/arginine-rich RNA processing factors (SR proteins); hence suggesting that SAFB1 serves as a molecular base for the assembly of a transcriptome complex that couples chromatin organising S/MARs elements with transcription and pre-mRNA processing (Nayler *et al.* 1998). SAFB1 has been classified as a novel member of the hnRNP protein family based on the presence of its RNA-binding domain (RBD) that is similar to the conserved residues in other hnRNP proteins (Weighardt *et al.* 1999). Protein-protein interactions between SAFB1 and a range of RNA-binding proteins including hnRNP A1, hnRNP D, hnRNP G, SR splicing regulatory protein 86 (SRrp86), SR protein kinase 1 (SRPK1) and Src-associated substrate in mitosis of 68kDa (Sam68) provide reasonable evidence to implicate a role in alternative splicing (Arao *et al.* 2000, Li *et al.* 2003, Nikolakaki *et al.* 2001, Sergeant *et al.* 2007, Weighardt *et al.* 1999). Indeed, SAFB1 was able to antagonise the exon inclusion of cell surface glycoprotein, CD44 when co-transfected with SRrp86 (Li *et al.* 2003). Sergeant *et al.* has also shown that SAFB1 and SAFB2 are able to act as negative regulator of transformer 2 protein homolog beta (*TRA2B*) alternative splicing (Sergeant *et al.* 2007). SAFB1 also inhibits SRPK1 enzymatic activity, thus affecting the downstream RNA splicing (Tsianou *et al.* 2009). It is still not known whether these SAFB proteins exert their effects on pre-mRNA splicing through direct RNA interaction or by tethering to other splicing factors.

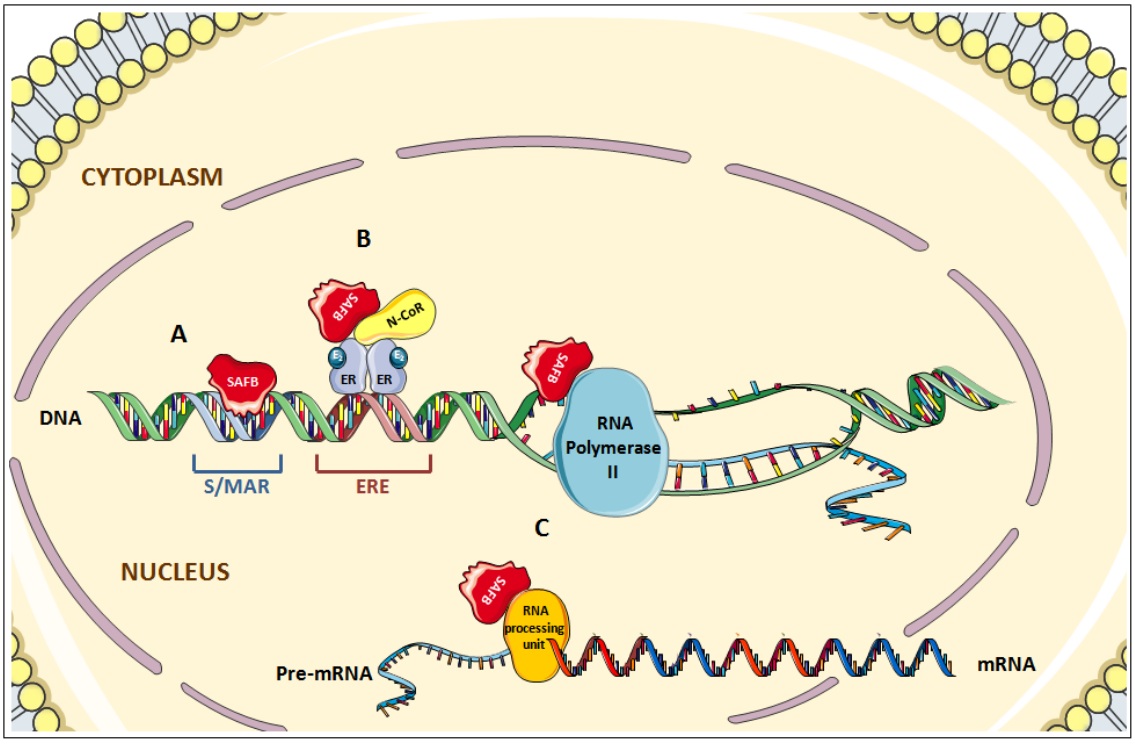


Figure 1.7 SAFB1 and SAFB2 cellular functions.

Illustration shows three known pathways by which SAFB1 and SAFB2 functions at the cellular level. Pathway A represents their function in chromatin organisation by binding to S/MARs that eventually effect target gene expression. Pathway B represents their function in gene regulation mediated by direct recruitment to the promoter region or indirect interaction with other transcription factors. Pathway C represents their interaction with various RNA processing proteins that consequently regulates RNA splicing and metabolism. The illustration was created using Servier Medical Art (servier.co.uk/medical-art-gallery).

1.4.3 SAFB mechanism of transcriptional regulation

Knowledge on the mechanism of transcriptional regulation by SAFB proteins has been established but remains limited (Hong *et al.* 2012). SAFB1 transcriptional repression executed by its intrinsic repression domain may be mediated by HDAC activity, as suggested by Oesterreich *et al.* (Oesterreich *et al.* 2000, Townson *et al.* 2004). Although SAFB1 has not been shown to directly interact with any known HDAC members, its transcriptional repression activity could partially be relieved by the HDAC inhibitor, trichostatin A (TSA) (Oesterreich *et al.* 2000). An indirect interaction between SAFB1 with HDAC complexes has been speculated and further investigation revealed NCoR as an interacting mediator between SAFB1 and HDAC3 (Jiang *et al.* 2006).

Jiang *et al.* reported that NCoR interacts directly with the C-terminal repression domain of SAFB1 and its absence diminishes SAFB1 repressive effect. This observation proposed that the repression ability of SAFB1 is partly a HDAC-dependent action mediated through NCoR (Jiang *et al.* 2006).

A more recent study by Garee *et al.* presented evidence for an alternative mechanism that could regulate SAFB action, i.e. through the post-translational modification by small ubiquitin-like modifiers (SUMO) known as SUMOylation (Garee *et al.* 2011). Similar to ubiquitination, SUMOylation is an enzymatic process that occurs at a unique four amino motif that includes a lysine residue for modification [reviewed in (Johnson 2004)]. SUMOylation of transcription factors and coregulators has been commonly shown to negatively regulate transcriptional activity. Potential lysine modification sites were identified at amino acid 231 and 294 within SAFB1 and these sites were 100% conserved in SAFB2 (Garee *et al.* 2011). SUMOylation of SAFB1 occurs predominantly at amino acid 294, a common effect that is not cell line specific. SUMOylation of SAFB1 did not indicate a change in subcellular localisation, protein half-life or alter interaction with NCoR; however SAFB1 SUMOylation did significantly decrease its interaction with HDAC3. Interestingly, SAFB1 transcriptional activity was abolished when SUMOylation sites were mutated, thus linking SAFB1 SUMOylation and its transcriptional repressive activity (Garee *et al.* 2011).

Song *et al.* has revealed another post-translational modification for SAFB2 in an attempt to identify BRCA1 ubiquitination substrates (Song *et al.* 2011). BRCA1 interacts with BRCA1-associated RING domain protein 1 (BARD1) to form a stable heterodimer complex that displays ubiquitin E3 ligase activity [reviewed in (Baer *et al.* 2002)]. Unlike most ubiquitin ligases, BRCA1/BARD1 catalyses an unconventional formation of lysine-6-linked chains which do not result in protein degradation (Wu-Baer *et al.* 2003). BRCA1 has been shown to induce ubiquitination of SAFB2 and result in increased SAFB2 protein expression. The overexpression of SAFB2 then significantly reduced the levels of BARD1 via its C-terminal repression domain, but did not affect BRCA1 expression. Taken together, these results showing the upregulation of SAFB2 through BRCA1/BARD1 mediated ubiquitination and consequent downregulation of BARD1 by SAFB2 overexpression suggest a possible feedback loop that regulates SAFB2 and BARD1 protein levels (Song *et al.* 2011). This finding also suggests a possible role for SAFB proteins as the missing link between BRCA1 and its function in repressing ER-α transcriptional activity.

1.4.4 SAFB and breast cancer

The strong correlation between SAFB proteins and breast cancer stems from their prominent role as ER-α transcriptional repressors and their interaction with key players in tumourigenesis. Functional and clinical studies have provided robust evidence that implicate the relevance of SAFB proteins in breast cancer.

SAFB1 expression were detected at varying levels in eight different breast cancer cell lines and SAFB1 levels in the cell inversely correlates with cell proliferation (Townson *et al.* 2000). SAFB1 protein expression assessed in 61 primary breast tumour tissues revealed widely varying levels, with 16% of the tumours lacking any detectable SAFB expression. High SAFB1 expression in these tumours was associated to low S-phase fraction and aneuploidy, a common phenotype of tumour cells (Townson *et al.* 2000). On the other hand, low levels of SAFB1 in invasive breast tumours were significantly associated to worse overall survival in patients (Hammerich-Hille *et al.* 2009, Oesterreich S 2002). Moreover, mouse embryonic fibroblasts with a genetic deletion of SAFB1 exhibit important characteristics of carcinogenesis, including cellular immortalisation, increased cell transformation, ability to proliferate in growth-restricting conditions and increased anchorage-independent growth (Dobrzycka *et al.* 2006). These observations suggest that a critical balance of SAFB proteins is important in breast tumourigenesis.

Low SAFB proteins expression could result from high rates of loss of heterozygosity (LOH) at the chromosomal locus 19p13, nearby both *SAFB1* and *SAFB2* genes (Oesterreich *et al.* 2001). The rate of LOH at this locus has been described as one of the highest LOH regions in the breast cancer genome, suggesting that SAFB proteins may also act as tumour suppressor proteins (Miller *et al.* 2003). However, SAFB1 germline mutation or inactivation was not observed in hereditary breast cancer (Bergman *et al.* 2008). These observations indicate that SAFB1 is not involved in hereditary breast cancer but may be important in sporadic breast cancer.

SAFB proteins have also been shown to be involved in apoptosis, a cellular response central to the development of cancer. Upon apoptosis induction by staurosporine, SAFB1 protein that normally displayed nuclear localisation excluding the nucleolus translocates into the nucleolus and thereafter, localised to peri-nucleolar ring structures (Lee *et al.* 2007). SAFB1 localisation, mediated by C-terminal protein-protein

interaction domain, may be associated with maturation of RNA at the initial stages and preparation of DNA cleavage at the later stages of apoptosis [reviewed in (Garee *et al.* 2010)]. Another study confirmed the participation of SAFB1 and SAFB2 in apoptosis when these proteins were shown to regulate transcription of apoptosis-related genes in breast cancer (Hammerich-Hille *et al.* 2010).

These clinical and *in vitro* observations provide convincing evidence that SAFB1 and SAFB2 are involved in breast cancer development, although to what extent is not understood. Further elucidation of their role in metastatic breast cancer and anti-oestrogen resistance is needed. Many other questions remain, such as: in what ways are SAFB1 and SAFB2 regulated? Do they function additively, synergistically or antagonistically? What are the consequences of inactivated SAFB1 and SAFB2 to breast tumourigenesis? What are the functions of their highly homologous RRM? In summary, further detailed study of SAFB1 and SAFB2 is needed to fully reveal their functions and involvement in breast cancer development.

1.5 Hypotheses and aims

1.5.1 Hypotheses

Despite increasing interest on the role of SAFB1 and SAFB2 in breast cancer, fundamental questions about the regulation and function of these proteins in breast cancer remain unanswered. The effect of SAFB proteins on ER function is well described but the effects of ER-ligands on SAFB proteins have yet to be explored. Recent work by Hammerich-Hille *et al.* has investigated the functional consequence of the loss of SAFB proteins in breast cancer cells containing ER, however their role in breast cancer cells lacking ER might contribute to the oestrogen-resistant phenotype of these cells. The role of SAFB proteins in RNA metabolism and processes has been speculated but the function of their highly homologous internal RRM has not yet been examined. Therefore, the overall hypotheses of this project are:

- SAFB1 and SAFB2 expression is regulated by ER-ligands in breast cancer cells.
- Loss of SAFB1 and SAFB2 alters expression of target genes that contribute to the invasive and metastatic characteristics of breast cancer cells.
- The homologous RRM within SAFB proteins is functional: SAFB proteins can bind RNA and regulate gene expression.

1.5.2 Aims

The specific aims of this project were to:

- Determine the effects of oestrogen, 17β-oestradiol, on the regulation and cellular localisation of SAFB1 and SAFB2 in ER positive (MCF-7) and ER negative (MDA-MB-231) breast cancer cells.
- Investigate whether loss of SAFB1 and SAFB2 by RNAi affects expression of breast cancer associated genes in MDA-MB-231 cells.
- Investigate the RNA-binding properties of SAFB1 using iCLIP technology to identify novel SAFB-target genes and a potential consensus RNA-binding motif.

Chapter 2 : Materials and methods

2.1 General laboratory practice

All experimental procedures were conducted in accordance to university standards for safe working with chemical substances in laboratories, which comply with the Control of Substances Hazardous to Health (COSHH) regulations. Tissue culture was carried out in compliance with regulations for containment of class II pathogens.

2.2 Source of tissue

Frozen breast tumour tissue was obtained from the Breast Cancer Campaign Tissue Bank (reference number: BCCTB13). Samples were taken from consenting patients and snap frozen in liquid nitrogen before storing at -80°C.

2.3 Cell culture

2.3.1 Cell lines

The breast cancer cell lines used in this study are MCF-7 (catalogue number: ATCC-HTB-22) and MDA-MB-231 cells (catalogue number: ATCC-HTB-26), both purchased from American Type Culture Collection (ATCC) and LGC Standards, Europe.

2.3.1.1 MCF-7

MCF-7 is a tumourigenic breast epithelial cell line originally derived from pleural effusion of a 69 year old female patient with breast adenocarcinoma. The phenotypic characteristics of this cell line is early-stage, non-invasive, ER and PR positive breast cancer (Soule *et al.* 1973).

2.3.1.2 MDA-MB-231

MDA-MB-231 is a tumourigenic breast epithelial cell line originally derived from pleural effusion of a 51 year old female patient with breast adenocarcinoma. The phenotypic characteristics of this cell line is invasive, ER and PR negative breast cancer (Cailleau *et al.* 1978).

2.3.2 Routine cell passage

Cell culture was performed under aseptic conditions in Class II laminar flow microbiological safety cabinet. MCF-7 and MDA-MB-231 were routinely cultured in 75cm^2 and 25cm^2 tissue culture flasks (Greiner) at 37°C in a humidified atmosphere containing 5% CO_2. Cell lines were cultured in Dulbecco's Modified Eagles medium (DMEM) without phenol-red (Sigma-Aldrich), supplemented with 10% foetal bovine serum (FBS) (Sigma-Aldrich), 2mM L-Glutamine (Sigma-Aldrich) and 1% penicillin-streptomycin (Sigma-Aldrich). All media was stored at 4°C and warmed in the water bath to 37°C prior to use.

2.3.3 Cell line maintenance

Cells were passaged every 3 to 5 days at 70-80% confluency. Cell passage was performed by removing growth media, rinsing the cell monolayer with sterile 1× phosphate buffered saline (PBS) (Sigma-Aldrich) and incubating with 2mM trypsin-EDTA (Sigma-Aldrich) for 5 minutes at 37°C. Complete growth media was added to neutralise the effect of trypsin-EDTA and detached cells were collected by centrifugation at 200×g for 5 minutes. The supernatant was discarded before the pelleted cells were resuspended in complete growth media and passaged at a ratio of 1:4.

2.3.4 Cryopreservation of cells

Cells were routinely cryopreserved at early passage numbers to generate a continuous stock of frozen cells. Cryopreservation was performed in 1ml aliquots of cryoprotective media and stored in cryovials (Nunc). Cryoprotective media consisted of FBS and

5% dimethyl sulphoxide (DMSO) (Sigma-Aldrich). Cells were stored in -80°C and liquid nitrogen for long term storage. When needed, frozen stocks were rapidly thawed in a 37°C water bath and DMSO removed by centrifugation at 200×g for 5 minutes. Cells were then resuspended in complete medium and plated in tissue culture flasks.

2.3.5 Cell counting

Cells were counted prior to experiments using an improved Neubauer chamber haemocytometer (Hawksley). Cell pellets were resuspended in complete growth media and the haemocytometer was filled with 10µl of single cell suspension. The number of cells overlying the ruled grid was counted using low power magnification (×10) on an inverted microscope (Leica). The number of cells per mililiter was calculated (cells in 25 squares of the grid multiplied by 10^4) and the cell suspension was diluted appropriately to seed the correct number of cells for each experiment.

2.3.6 Stimulation of cultured cells

2.3.6.1 Oestrogen treatment

MCF-7 and MDA-MB-231 cells were seeded at a density of 5×10^4 cells per well in 12-well culture plates (Greiner) and cultured in complete growth media until approximately 60% confluent. Cells were incubated in serum free media for 24 hours then replaced with phenol-free DMEM supplemented with 10% charcoal stripped FBS (Gibco) and 2mM L-Glutamine (Sigma-Aldrich). 17β-oestradiol (Sigma-Aldrich) was dissolved in absolute ethanol and added to the media at a range of final concentrations from 0.01µM to 10.0µM for a further 24 hours. Experimental control conditions such as untreated and ethanol only (vehicle) were included in each experiment. Oestrogen treatment experiments were performed at least three times with cells at different passage number.

2.3.6.2 Oestrogen and proteasome inhibitor treatment

MCF-7 and MDA-MB-231 cells were seeded at a density of 5×10^4 cells per well in 12-well culture plates and cultured in complete growth media until approximately

60% confluent. Cells were incubated in serum free media for 24 hours then replaced with phenol-free DMEM supplemented with 10% charcoal stripped FBS (Gibco) and 2mM L-Glutamine (Sigma-Aldrich). MG132 (Calbiochem) was dissolved in DMSO and added to the media to make a final concentration of 0.25μM. The previously used range of 17β-oestradiol dose was added together to the media for a further 24 hours. Experimental control conditions such as untreated and ethanol only (vehicle) were included in each experiment. Experiments were performed at least three times with cells at different passage number.

2.3.6.3 Anti-oestrogen treatment

MCF-7 and MDA-MB-231 cells were seeded at a density of 5×10^4 cells per well in 12-well culture plates and cultured in complete growth media until approximately 60% confluent. Cells were incubated in serum free media for 24 hours then replaced with phenol-free DMEM supplemented with 10% charcoal stripped FBS (Gibco) and 2mM L-Glutamine (Sigma-Aldrich). Fulvestrant (also known as ICI 182,780) (Sigma-Aldrich) was dissolved in absolute ethanol and added to the media at a range of final concentrations from 0.001μM to 1.0μM for a further 24 hours. Experimental control conditions such as untreated and ethanol only (vehicle) were included in each experiment. Anti-oestrogen treatment experiments were performed at least three times with cells at different passage number.

2.3.7 Transient transfection of siRNA in cultured cells

Gene knockdown by RNA interference (RNAi) was performed in MCF-7 and MDA-MB-231 cells using validated small interfering RNA (siRNA). Pre-designed siRNAs were purchased from Ambion (Life Technologies), targeting exon 12 of the SAFB1 transcript (RefSeq: NM_002967.2) and exon 14 of the SAFB2 transcript (RefSeq: NM_014649.2) (Figure 2.1). SAFB1 and SAFB2 siRNA were used to knockdown their respective expression to validate antibody specificity by Western immunoblotting. Cells were also transfected with GAPDH siRNA as a positive control or a negative control siRNA in each experiment. Untreated and transfection agent only controls were also included in parallel for comparison. Details of the siRNAs used in this thesis are listed in Table 2.1. Conditions for efficient siRNA transfection were

optimised by varying cell density and concentrations of siRNA or transfection reagent within the manufacturer's recommended guidelines. Reverse transfection method, where cells were transfected while still in suspension, was applied to improve transfection efficiency. For each siRNA, 8µl siPORT *NeoFX* Transfection Agent (Ambion) was diluted in 400µl DMEM lacking FBS and penicillin-streptomycin. In a separate 1.5ml microcentrifuge tube (Greiner), 24µl of 2µM siRNA was diluted in 400µl DMEM lacking FBS and penicillin-streptomycin to make up a final siRNA concentration of 12nM. For the double gene knockdown of SAFB1 and SAFB2, 24µl of each 2µM siRNA was diluted in 400µl DMEM lacking FBS and penicillin-streptomycin. The diluted transfection agent and siRNA were mixed together and incubated for 30 minutes to allow the formation of transfection complexes. Transfection complexes were dispensed into 25cm^2 tissue culture flasks along with MDA-MB-231 cells at the density of 50×10^4 cells per flask and supplemented with 4ml of DMEM lacking penicillin-streptomycin. Media containing transfection complexes was removed after 24 hours and replaced with fresh growth media to culture for a further 48 hours prior to harvesting. For the gene expression profile study (Section 2.8), transfected cells were cultured for 48 hours then followed by stimulation with 0.01µM 17β-oestradiol for 24 hours prior to harvesting.

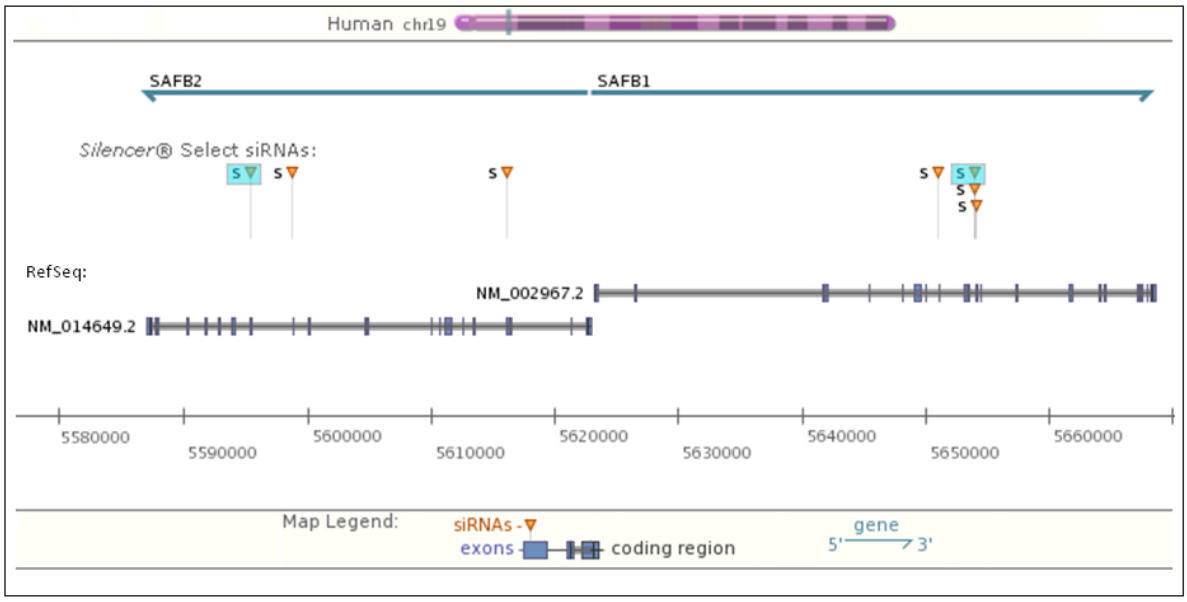


Figure 2.1 Mapping of the selected siRNAs to SAFB1 and SAFB2 transcripts.

Pre-designed Silencer® Select siRNAs were mapped to SAFB1 and SAFB2 transcripts to distinguish specific target sites. The selected siRNAs used in this study are highlighted in blue. Image taken from (lifetechnologies.com).

SiRNA target	Sense (5'-3')	Antisense (3'-5')
SAFB1	CCUUAAGAGGGAUGAUAAAtt	UUUAUCAUCCCUCUUAAGGtt
SAFB2	GAGUCAGGAUCGCAAGUCAtt	UGACUUGCGAUCCUGACUCtt
GAPDH	Sequence not provided	
Negative control	Sequence not provided	

Table 2.1 Targets and RNA sequences of siRNA used in this study.

2.4 Preparation of protein samples

2.4.1 Preparation of whole cell lysate

Whole cell lysates from MCF-7 and MDA-MB-231 cells were prepared by several different methods depending on their downstream application. Radio immuno-precipitation assay (RIPA) buffer was used to extract 17β-oestradiol and fulvestrant stimulated cells. Sodium dodecyl sulphate (SDS) lysis buffer was used to harvest siRNA transfected cells. The components in each buffer are listed in Table 2.2 and methods for each protocol are detailed below.

RIPA buffer	SDS lysis buffer
25mM Tris-HCl (pH7.6)	125mM Tris-HCl (pH6.8)
0.1% SDS	2% SDS
150mM NaCl	10% glycerol
1% NP-40	10% β-mercapthoethanol
0.1% sodium deoxycholate	0.1% bromophenol blue

Table 2.2 Components of RIPA and SDS lysis buffer used in this study.

2.4.1.1 RIPA buffer

MCF-7 and MDA-MB-231 cells were cultured and stimulated with 17β-oestradiol and fulvestrant as previously described. Prior to use, every 1ml of ice-chilled RIPA buffer was supplemented with 20µl of protease inhibitor cocktail (Sigma-Aldrich) and 10µl of phosphatase inhibitor cocktail (Thermo Scientific). Cells were firstly rinsed 3 times in

cold 1× PBS and left in 1ml PBS to enable cell collection into a 1.5ml microcentrifuge tube by a cell scraper. The detached cells were collected by centrifugation at 200×g for 5 minutes and the PBS supernatant was discarded. The pelleted cells were resuspended in 200µl of cold RIPA buffer with gentle agitation. Lysed cells were centrifuged for 15 minutes at 13000×g to pellet cell debris and supernatant were stored at -20°C until ready for use.

2.4.1.2 SDS lysis buffer

MCF-7 and MDA-MB-231 cells were transiently transfected with various siRNA as previously described. After transfection, cells were collected into a 1.5ml microcentrifuge tube as described above. The pelleted cells were resuspended in 100µl of SDS lysis buffer and stored at -20°C until ready for use.

2.4.2 Preparation of cellular fractions

MCF-7 and MDA-MB-231 cells stimulated with 17β-oestradiol were separated into subcellular fractions and phosphorylated/unphosphorylated fractions respectively. Cytoplasmic and nuclear fractions were prepared using the Universal Magnetic Co-IP kit (Active Motif). PhosphoProtein purification kit (Qiagen) was used to separate phosphorylated and unphosphorylated proteins within the whole cell lysate. The protocols used for each method are detailed below.

2.4.2.1 Universal Magnetic Co-IP kit

Cytoplasmic and nuclear fractions from MCF-7 and MDA-MB-231 cells stimulated with 17β-oestradiol were prepared according to the manufacturer's protocol. Briefly, cells were washed in 1× PBS and collected with a cell scraper into hypotonic lysis buffer containing protease inhibitors and phosphatase inhibitors. Cell lysate was subjected to centrifugation at 13,000×g and 4°C for 5 minutes to separate the cytoplasmic fraction. The resulting supernatant at this step contains the cytoplasmic fraction that was removed for storage at -20°C. The pellet was resuspended in digestion buffer containing protease inhibitors, phosphatase inhibitors and detergent. Lysate was

incubated for 20 minutes on ice before enzymatic shearing cocktail was added for an additional incubation at 37°C for 10 minutes. 0.5M ethylenediaminetetraacetic acid (EDTA) was added to inactivate the enzyme and lysate was pelleted by centrifugation at 13,000×g and 4°C for 5 minutes. The resulting contains the nuclear fraction that was removed for storage at -80°C.

2.4.2.2 PhosphoProtein purification kit

MDA-MB-231 cells were stimulated in culture with 1.0μM 17β-oestradiol or untreated for 24 hours prior to separation of phosphorylated proteins from whole cell lysate. Protein separation was performed according to the manufacturer's instruction. Briefly, cells were collected into a 1.5ml microcentrifuge tube as previously described. The pelleted cells were resuspended in 5ml of CHAPS or 3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonate lysis buffer containing protease inhibitors. Cell lysate was incubated for 30 minutes at 4°C, and then subjected to centrifugation at 13,000×g and 4°C for 30 minutes. Supernatant was allowed to pass through the separation column provided in the kit and the gel bed within the column binds to any phosphorylated protein present in the cell lysate. The flow-through fraction was collected for analysis of unphosphorylated proteins in the lysate. Phosphorylated and unphosphorylated lysate were stored at -20°C until ready for use.

2.4.3 Protein quantification

Protein concentrations of whole cell lysates in RIPA buffer and all eluate fractions were determined by a colourimetric assay following detergent solubilisation using DC Protein Assay (BioRad). Protein quantification was performed according to the manufacturer's protocol and absorbance were read at 750nm on SpectraMax 190 microplate reader (Molecular Devices) using parameters for the Lowry assay (Lowry *et al.* 1951) on the SoftMax Pro software (Molecular Devices).

2.5 Co-Immunoprecipitation

Co-Immunoprecipitation (Co-IP) was performed using the Universal Magnetic Co-IP kit (Active Motif) according to the manufacturer's protocol. Briefly, 100µg nuclear extract from MCF-7 cells was incubated with 5µg SAFB1 or SAFB2 antibody or an appropriate IgG control at 4°C for 4 hours on a rotating platform. Thereafter, Protein G magnetic beads (Active Motif) were added to each tube and left to incubate overnight at 4°C on a rotating platform. Antibody/protein complexes bound to the Protein G beads were precipitated using a magnetic tube stand and the beads were washed repeatedly for 4 times prior to resuspension in 20µl SDS lysis buffer. Protein G coupled antibody/protein complexes were denatured by heating to 95°C for 5 minutes. Protein resuspended in SDS lysis buffer were separated from the magnetic beads and stored at -20°C until ready for use.

2.6 Protein analysis

2.6.1 SDS-Polyacrylamide Gel Electrophoresis

SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE) was used to separate proteins for Western immunoblot analysis. Polyacrylamide gels were cast using a Mini-PROTEAN III system (BioRad) in the components as listed in Table 2.3 below. A 10% acrylamide resolving gel was allowed to set and overlaid by a 4% acrylamide stacking gel with an inserted comb to form 15 wells.

Resolving gel	Stacking gel
10% acrylamide	4% acrylamide
375mM Tris-HCl (pH7.6)	125mM Tris-HCl (pH6.8)
0.1% SDS	1% SDS
0.1% TEMED	0.1% TEMED
0.1% ammonium persulphate	0.5% ammonium persulphate

Table 2.3 Components of resolving and stacking gels used for SDS-PAGE.

Protein samples prepared in RIPA buffer and cell fractions were denatured by the addition of SDS lysis buffer and heating at 95°C for 5 minutes. Samples were loaded into the wells of the gel and BenchMark Pre-Stained Protein Ladder (Invitrogen) was used as a reference molecular weight marker. Electrophoresis was performed in electrode buffer at 175V for 1 hour using PowerPAC 3000 system (BioRad). Proteins were transferred onto nitrocellulose membrane (BioRad) in transfer buffer at 120V for 2 hours using PowerPAC 200 system (BioRad). The components of the electrode and transfer buffer are detailed in Table 2.4.

Electrode buffer	Transfer buffer
25mM Tris-HCl pH8.3	25mM Tris-HCl pH8.3
190mM glycine	150mM glycine
0.1% SDS	10% methanol

Table 2.4 Components of electrode and transfer buffer used for SDS-PAGE.

2.6.2 Western immunoblotting

All nitrocellulose membranes containing transferred proteins were routinely assessed for equal protein loading using Ponceau S protein staining (data not shown). Membranes were incubated in Ponceau S solution (Sigma-Aldrich) for 1 minute and rinsed in distilled water to visualise the protein bands. The membranes were destained by washing in 1× PBS prior to membrane blocking. Unbound membrane sites were blocked to prevent non-specific antibody binding by incubation in a blocking solution of 10% non-fat milk powder (Marvel) in 1× PBS at room temperature for 1 hour. Primary antibody incubations were performed at an appropriate dilution in 1% non-fat milk powder diluted in 1× PBS and 0.1% Tween-20 (PBST, Sigma-Aldrich). The nitrocellulose membranes were incubated in the primary antibody solution at 4°C overnight. Details of the primary antibodies used are detailed in Table 2.5. After incubation, membranes were washed twice for 5 minutes and a final 15 minutes in 1× PBS on a rocking platform. Horseradish peroxidase (HRP) conjugated secondary antibody (Dako) was diluted 1:3000 in 1× PBS and applied to the membranes for 1 hour at room temperature. Washes were repeated after secondary antibody incubation and proteins were detected by chemiluminescence. The ECL Western Blotting Substrate

(Thermo Scientific) or SuperSignal West Femto (Thermo Scientific) detection system was used to allow the development of luminescence on antibody bound proteins. Membranes were exposed to x-ray films (Kodak) and autoradiographs were developed using Compact X4 film processor (Xograph).

2.6.3 Immunofluorescent staining

MCF-7 and MDA-MB-231 cells were seeded at a density of 10×10^4 cells per well in 12-well culture plates containing glass coverslips (Scientific Laboratory Supplies). Cells were cultured in complete growth media until approximately 60% confluent and incubated in serum free media for 24 hours prior to stimulation with 0.01µM and 1.0µM 17β-oestradiol for a further 24 hours. Cells on coverslips were fixed by immersion in cold absolute methanol and incubated at -20°C for 20 minutes. Coverslips were then air dried at room temperature before being mounted onto microscope slides (Erie Scientific) using clear nail polish.

Non-specific binding was blocked by a blocking solution of 10% normal goat serum (Jackson ImmunoResearch Laboratories) in Tris-buffered saline (0.1M Tris, 0.05M NaCl, adjusted to pH7.6) and incubated at room temperature for 10 minutes. Primary antibody was added to the coverslips at a 1:200 dilution in 1% normal goat serum and incubated at 4°C overnight. Details of the primary antibodies used are detailed in Table 2.5. Following the removal of the primary antibody, coverslips were washed 3 times in TBS and fluorescein isothiocyanate (FITC) conjugated anti-mouse secondary antibody (Jackson ImmunoResearch Laboratories) was added to each coverslip at a 1:40 dilution in TBS for 30 minutes in a dark chamber at room temperature. The coverslips were stained in 4',6-diamidino-2-phenylindole (DAPI) nucleic acid stain for 5 minutes and mounted over using Fluorescent Mounting Media (Dako). Analysis of immunofluorescent staining was performed using Leica TCS SP2 UV confocal microscope at the Newcastle University Bioimaging Facility.

Primary antibodies	Company	Source	Dilution for Western blot	Dilution for immunofluorescence
SAFB1	Sigma-Aldrich	Mouse	1:1000	1:200
SAFB2	Sigma-Aldrich	Mouse	1:250	1:200
GAPDH	Santa Cruz	Rabbit	1:1000	-
β-actin	Santa Cruz	Rabbit	1:1000	-
ERK-1	BD Transduction Laboratories	Mouse	1:1000	-
c-Jun	Santa Cruz	Rabbit	1:200	-
SRSF1	Zymed	Mouse	1:1000	-
Phosphoserine	Millipore	Mouse	1:500	-
ER-α66	Vector Labs	Mouse	1:200	-
ER-α36	Gift from (Wang *et al.* 2006)	Rabbit	1:250	-
SC35	Abcam	Mouse	-	1:200
Sam68	Santa Cruz	Rabbit	1:1000	-

Table 2.5 Primary antibodies used in this study.

2.7 RNA isolation and analysis

2.7.1 RNA extraction and quantitation

Total RNA was extracted from cultured cells and frozen breast tumour tissue using the SV Total RNA Isolation System (Promega) following the manufacturer's protocol. Briefly, cells were collected in a 1.5ml microcentrifuge tube as described in Section 2.4.1.1, while 30mg of frozen breast tumour tissue was homogenised in liquid nitrogen using mortar and pestle. Cells or tissue were resuspended in 175µl RNA Lysis Buffer. RNA Dilution Buffer was added into the lysate, heated at 70°C for 3 minutes and centrifuged at 13,000×g for 10 minutes. The cleared lysate solution was transferred to a fresh 1.5ml microcentrifuge tube and added with 200µl of 95% ethanol before being applied to the spin column assembly. After several washes and centrifugation of the spin column assembly, the spin column membrane containing the RNA was treated with DNase I at room temperature for 15 minutes. Following several washes and centrifugation, RNA was eluted in nuclease free water provided in the kit.

RNA concentrations were measured using a NanoDrop ND-1000 spectrophotometer (Thermo Scientific) and RNA samples stored at -80°C until ready for use.

2.7.2 Reverse transcription

First strand cDNA was synthesised from total RNA by reverse transcription using the SuperScript III Reverse Transcriptase kit (Invitrogen). Components listed in Table 2.6(a) were added to a 1.5ml microcentrifuge tube and heated to 65°C for 5 minutes followed by at least 1 minute incubation on ice. The contents of the tube were collected by brief centrifugation and components listed in Table 2.6 (b) were added into the 1.5ml microcentrifuge tubes. Reverse transcription reaction was mixed by gentle pipetting and tubes were incubated in a thermal cycler (G-Storm) at 50°C for 45 minutes followed by 70°C for 15 minutes. Synthesised cDNA was stored at -20°C until ready for use.

RT reaction (a)	RT reaction (b)
1µl oligo $(dT)_{20}$ primers (0.5 µg)	4µl 5× First Strand buffer
1µl dNTP mix (10mM each)	1µl DTT (0.1M)
1µg RNA	1µl RNaseOUT (40units/µl)
Distilled water to 13µl volume	1µl SuperScript III reverse transcriptase

Table 2.6 Components in each reverse transcription reaction.

2.7.3 Quantitative real time polymerase chain reaction

Quantitative real time polymerase chain reaction (qRT-PCR) was performed using TaqMan gene expression assays on the StepOne Real Time PCR System (Applied Biosystems, Life Technologies) according to the manufacturer's instructions. Validated TaqMan assays were selected to target specific genes as listed in Table 2.7. The TaqMan assays were added to a reaction consisting of TaqMan Master Mix, cDNA template and RNase-free water. A total volume of 20µl of PCR reaction was transferred onto each well of a 96-well reaction plate and sealed with a clear adhesive film. The plate was loaded into the instrument and a standard cycling condition was performed as follows: 50°C for 2 minutes, 95°C for 10 minutes, 40 cycles of 95°C for 15 seconds and

60°C for 60 seconds. Data analysis was performed using the comparative C_t method normalised against β-actin expression. Experiments were performed in triplicate and statistical analysis was performed using student's t-test (Microsoft Excel). All effects at $p<0.05$ are reported as significant.

TaqMan assay	RefSeq	Product code	Exon boundary
SAFB1	NM_002967.2	Hs01561652_g1	3-4
SAFB2	NM_014649.2	Hs01006796_g1	7-8
β-actin	NM_001101.3	Hs99999903_m1	1-1
VEGF-A	NM_001025366.2	Hs03929046_s1	8-8
CLU	NM_001831.3	Hs00156548_m1	3-4
ITGB4	NM_000213.3	Hs00236216_m1	18-19
IL-6	NM_000600.3	Hs00985639_m1	2-3
SHF	NM_138356.2	Hs00403125_m1	3-4

Table 2.7 List of TaqMan assays used in this study.

2.8 Gene expression profile study

2.8.1 RT^2 Profiler PCR Array

SAFB1 and SAFB2 target genes were identified using a combined approach of siRNA gene knockdown and qRT-PCR in MDA-MB-231 cells. Cells were transiently transfected with negative control, SAFB1, SAFB2 and SAFB1 + SAFB2 siRNA as previously described (Section 2.3.7). Whole cell lysate and RNA were extracted (as described in Section 2.4.1.2 and 2.7.1) after transfection and analysed for the degree of gene knockdown by Western immunoblotting and qRT-PCR (as described in Section 2.6.2 and 2.7.3). Once verified for sufficient levels of gene knockdown, RNA was reverse transcribed and subjected to qRT-PCR using the RT^2 Profiler PCR Array system (SABiosciences) that utilises the SYBR green detection method. This qRT-PCR array system consists of 84 genes involved in breast cancer and ER signalling pathway, 5 housekeeping genes, 1 genomic DNA control and 6 PCR controls (Figure 2.2). The experiment was performed in three replicates and data generated was analysed using the web-based RT^2 Profiler PCR Array data analysis software (SABiosciences) that performs the comparative C_t calculations for fold change analysis. Following the

recommended fold change threshold, a fold change of ≥ 2.0 represents gene upregulation and a fold change of ≤ 0.5 represents gene downregulation. Relationship between target genes were analysed in an interaction network using the web-based GeneMANIA software (http://www.genemania.org). Several genes that were upregulated and downregulated beyond the fold change threshold were selected for further validation using qRT-PCR with TaqMan gene expression assays (Section 2.7.3).

Array Layout

AR A01	BAD A02	BAG1 A03	BCL2 A04	BCL2L2 A05	C3 A06	CCNA1 A07	CCNA2 A08	CCND1 A09	CCNE1 A10	CD44 A11	CDH1 A12
CDKN1A B01	CDKN1B B02	CDKN2A B03	CLDN7 B04	CLU B05	COL6A1 B06	CTNNB1 B07	CTSB B08	CTSD B09	CYP19A1 B10	DLC1 B11	EGFR B12
ERBB2 C01	ESR1 C02	ESR2 C03	FAS C04	FASLG C05	FGF1 C06	FLRT1 C07	FOSL1 C08	GABRP C09	GATA3 C10	GNAS C11	GSN C12
HMGB1 D01	HSPB1 D02	ID2 D03	IGFBP2 D04	IL2RA D05	IL6 D06	IL6R D07	IL6ST D08	ITGA6 D09	ITGB4 D10	JUN D11	KIT D12
KLF5 E01	KLK5 E02	KRT18 E03	KRT19 E04	MAP2K7 E05	MKI67 E06	MT3 E07	MUC1 E08	NFYB E09	NGF E10	NGFR E11	NME1 E12
PAPPA F01	PGR F02	PLAU F03	PTEN F04	PTGS2 F05	RAC2 F06	RPL27 F07	SCGB1D2 F08	SCGB2A1 F09	SERPINA3 F10	SERPINB5 F11	SERPINE1 F12
SLC7A5 G01	SPRR1B G02	STC2 G03	TFF1 G04	TGFA G05	THBS1 G06	THBS2 G07	TIE1 G08	TNFAIP2 G09	TOP2A G10	TP53 G11	VEGFA G12
B2M H01	HPRT1 H02	RPL13A H03	GAPDH H04	ACTB H05	HGDC H06	RTC H07	RTC H08	RTC H09	PPC H10	PPC H11	PPC H12

Figure 2.2 RT^2 Profiler PCR Array 96-well plate layout.

Wells A1 to G12 each contain a qRT-PCR assay for a gene related to breast cancer and ER signalling pathway. Wells H1 to H5 contain a housekeeping gene panel to normalise array data. Wells H6 contains a genomic DNA control. Wells H7 to H9 contain replicate reverse transcription controls and wells H10 to H12 contain replicate positive PCR controls.

2.8.2 Analysis of IL-6 secretion by enzyme-linked immunosorbent assay

Following the qRT-PCR results obtained for interleukin 6 (IL-6) from the RT^2 Profiler PCR Array and TaqMan gene expression assay, enzyme-linked immunosorbent assay (ELISA) was performed to analyse the protein secretion of IL-6. MDA-MB-231 cells were left untreated or transiently transfected with negative control, SAFB1, SAFB2 and SAFB1 + SAFB2 siRNA as previously described (Section 2.3.7). After transfection, the supernatant from the cultured cells were aspirated into a fresh 1.5ml microcentrifuge tube and cell debris was pelleted by refrigerated centrifuge at 13,000×g for 5 minutes. The cleared supernatant was subjected to ELISA using a Human IL-6 ELISA Development Kit (PeproTech) according to the manufacturer's protocol.

Briefly, Immulon ELISA microplates (ThermoFisher Scientific) were coated with 0.5µg/ml of capture antibody and incubated at 4°C overnight to allow antibody binding to the wells. Plates were washed with 1× PBST and blocked in 5% bovine serum albumin (Sigma-Aldrich) diluted in 1× PBST for 1 hour at room temperature on a shaking platform. After incubation, plates were washed again with 1× PBST and IL-6 standards were diluted (2,000pg/ml, 1,000pg/ml, 500pg/ml, 250pg/ml, 125pg/ml, 62pg/ml, 31pg/ml and 0pg/ml) before added into the wells in triplicate. Previously collected supernatant from the cultured cells were also added into the wells in triplicate and incubated at 4°C overnight on a shaking platform. The wash steps were repeated twice and 0.25µg/ml detection antibody was added to each well for 2 hours at room temperature on a shaking platform. Plates were washed twice and strepavidin HRP conjugate diluted 1:5500 was added to each well for 1 hour at room temperature on a shaking platform. After another two washes in 1× PBST, a colourimetric substrate solution was added and absorbance was read at 492nm on the MRX II microplate reader (Dynex Technologies). A standard curve was produced based on readings from the IL-6 standards and the amount of IL-6 secretion (pg/ml) from each experimental sample was calculated. IL-6 secretion was compared to the untreated sample to generate a relative fold change and statistical analysis was performed using student's t-test. All effects at $p<0.05$ are reported as significant.

2.9 Crosslinking and immunoprecipitation

2.9.1 Individual nucleotide resolution CLIP

2.9.1.1 UV crosslinking of MCF-7 cells

RNA targets for SAFB1 were identified by the UV crosslinking and immunoprecipitation (CLIP) technique. CLIP with individual nucleotide resolution (iCLIP) was performed using MCF-7 cells based on a published protocol (Konig *et al.* 2011). MCF-7 cells were cultured in 10cm tissue culture plates (Nunc) until approximately 80% confluent, then covered in ice-cold 1× PBS and irradiated with 150mJ/cm^2 of UV at 254nm in the UV Stratalinker (Stratagene). Cells were harvested with a cell scraper into 1.5ml microcentrifuge tubes and precipitated by centrifugation for 10 seconds at 13,000×g in a refrigerated centrifuge. Cell pellets were snap frozen on dry ice until ready for use.

2.9.1.2 Preparation of magnetic beads

Magnetic beads were prepared by washing 100µl protein A or protein G Dynabeads (Invitrogen) twice in lysis buffer [Table 2.9 (a)] and resuspending in 100µl lysis buffer containing 5µg SAFB1 antibody (Sigma-Aldrich). The magnetic beads were incubated at room temperature for 1 hour on a rotating platform and washed twice with lysis buffer before the addition of cell lysates.

2.9.1.3 Partial RNA digestion and immunoprecipitation

UV-irradiated cell pellets were resuspended in lysis buffer and treated with Turbo DNase I (Ambion) and high (1:10 dilution) or low (1:500 dilution) RNase I (Ambion). Cell lysate was incubated at 37°C for 3 minutes while shaking at 1100 rpm (Eppendorf Thermomixer) and immediately placed on ice. Cell debris was precipitated by centrifugation at 13,000×g and 4°C for 20 minutes, followed by careful collection of the supernatant. The cleared lysate was added to the magnetic beads for immuno-precipitation and incubated at 4°C for 2 hours on a rotating platform. After incubation, the supernatant was discarded and magnetic beads were washed twice in high-salt buffer [Table 2.9 (b)] followed by twice in wash buffer [Table 2.9 (c)].

2.9.1.4 Dephosphorylation and linker ligation of RNA 3' ends

The RNA 3' ends were dephosphorylated in 20µl of PNK mix [Table 2.9 (d)] and incubated at 37°C for 20 minutes. Samples were washed once in wash buffer and high-salt buffer, followed by two extended washes in wash buffer. RNA linker (Table 2.8) was ligated to the 3' ends by resuspending the magnetic beads in 20µl ligation mix [Table 2.9 (e)] and incubated overnight at 16°C.

2.9.1.5 Radioactive labeling of RNA 5' ends and protein separation

After a series of washes, the magnetic beads were resuspended in 8µl of PNK mix containing ^{32}P-γ-ATP and incubated at 37°C for 5 minutes to radioactively label the RNA 5' ends. The PNK mix was removed and the magnetic beads were resuspended

in 20µl NuPAGE loading buffer (Invitrogen). After incubation on a thermomixer at 70°C for 10 minutes, the empty magnetic beads were precipitated on a magnet and supernatant loaded on a 4-12% NuPAGE Bis-Tris gel (Invitrogen) with 1× MOPS running buffer (Invitrogen). BenchMark Pre-Stained Protein Ladder (Invitrogen) was used as a reference molecular weight marker and electrophoresis was performed at 180V for 1 hour. Protein and covalently bound RNAs were transferred onto a nitrocellulose membrane (BioRad) using a Novex wet transfer apparatus (Invitrogen) for 2 hours at 30V. After transfer, the membrane was rinsed in 1× PBS buffer and wrapped in cling film to be exposed to a BioMax XAR Film (Kodak) at -80°C for 1 hour or overnight.

2.9.1.6 RNA isolation

To isolate the protein-RNA complexes, the low RNase sample was cut out from the membrane above the molecular weight of SAFB1 protein (175kDa) and placed in a 1.5ml microcentrifuge tube. Membrane pieces were incubated in 2mg/ml Proteinase K (Roche) diluted in PK buffer [Table 2.9 (f)] at 37°C for 20 minutes while shaking at 1100 rpm. PK buffer containing 7M urea was then added into the tube and incubated at 37°C for another 20 minutes. Samples were collected and added with 400µl RNA phenol/chloroform (Ambion) to a Phase Lock Gel Heavy tube (VWR). The solution was incubated at 30°C for 5 minutes while shaking at 1100 rpm prior to phase separation by centrifugation at 13,000×g for 5 minutes. The aqueous layer was transferred into a fresh tube and mixed with 0.5µl GlycoBlue (Ambion), 40µl 3M sodium acetate pH 5.5 and 1ml 100% ethanol to precipitate the RNA overnight at -20°C.

2.9.1.7 Reverse transcription

The precipitated RNA was reverse transcribed in 7.25µl RNA/primer mix [Table 2.9 (g)] containing different Rclip primers with individual barcode sequences for each replicate (Table 2.8). Samples were incubated at 70°C for 5 minutes and cooled to 25°C before 2.75µl RT mix containing SuperScript III reverse transcriptase (Invitrogen) was added. Reverse transcription reaction was performed at the following conditions: 25°C for 5 minutes, 42°C for 20 minutes, 50°C for 40 minutes and 80°C for 5 minutes before cooling to 4°C. cDNAs were precipitated by the addition of 90µl TE buffer,

0.5µl GlycoBlue, 10µl 3M sodium acetate pH 5.5 and 250µl 100% ethanol incubated overnight at -20°C.

2.9.1.8 cDNA purification

The precipitated cDNA was resuspended in 6µl of water and 2× TBE-urea loading buffer (Invitrogen) and incubated at 80°C for 3 minutes. Samples were loaded onto a 6% TBE-urea gel (Invitrogen) beside a low molecular weight marker and electrophoresis performed at 180V for 40 minutes. Three gel fragments corresponding to cDNA size were cut at 120-200 nucleotides (high), 85-120 nucleotides (medium) and 70-85 nucleotides (low). Gel fragments were mixed with 400µl TE buffer and crushed using a 1ml syringe plunger. The crushed gel mixture was incubated at 37°C for 2 hours while shaking at 1100 rpm. The liquid portion of the sample was transferred into a Costar SpinX column (Corning Incorporated) containing two glass pre-filters (Whatman). The tubes were centrifuged at 13,000×g for 1 minute into a fresh 1.5ml microcentrifuge tube. The samples were added with 0.5µl GlycoBlue, 40µl 3M sodium acetate pH 5.5 and 1ml 100% ethanol to precipitate again overnight at -20°C.

2.9.1.9 Ligation of primer to cDNA 5’ ends

Precipitated cDNAs were resuspended in 8µl ligation mix [Table 2.9 (h)] and incubated at 60°C for 1 hour to circularise the cDNAs. In order to subsequently linearise the cDNAs, a primer complementary to BamHI restriction site (Table 2.8) was annealed to the 5’ end of the cDNAs by adding 30µl oligo annealing mix [Table 2.9 (i)] and incubated at the following conditions: 95°C for 1 minute, then temperature decreased by 1°C every 20 seconds until 25°C was reached. BamHI cleavage was performed by adding 2µl BamHI (Fermentas) and incubated at 37°C for 30 minutes. Samples were mixed with 50µl TE buffer, 0.5µl GlycoBlue, 10µl 3M sodium acetate pH 5.5 and 250µl 100% ethanol to precipitate overnight at -20°C.

2.9.1.10 Incorporation of sequencing primers by PCR amplification

cDNAs were amplified by adding a PCR mix [Table 2.9 (j)] containing P5/P3 Solexa primer mix (Table 2.8) and PCR was performed at the following conditions: 94°C for 2 minutes, 30 cycles of [94°C for 15 seconds, 65°C for 30 seconds, 68°C for 30 seconds], 68°C for 3 minutes and hold at 4°C. PCR products were loaded onto the QIAxcel analyser system (Qiagen) for an automated gel electrophoresis analysis of the cDNA fragments.

2.9.2 High-throughput sequencing and mapping

Prior to sequencing of the iCLIP libraries, the success of the experiment was monitored at two crucial steps: the autoradiograph of protein-RNA complex after membrane transfer and the gel image of the amplified PCR products. Once verified, samples were submitted for high-throughput sequencing at the Institute of Genetic Medicine (Newcastle University, UK), in collaboration with Professor Bernard Keavney's group. The samples were prepared for sequencing using the TruSeq Sample Preparation kit (Illumina) and the three replicates were sequenced on one lane of the Genome Analyser II system (GAIIx, Illumina). Bioinformatics analyses were performed on the sequencing results with the help of Tomaz Curk (University of Ljubljana, Slovenia) on the web-based iCount software (http://icount.fri.uni-lj.si/). Mapping of SAFB1 crosslink sites to regions of respective genes were visualised in UCSC Genome Browser (http://genome.ucsc.edu/) and a graphical representation of the novel SAFB1 consensus binding motif was designed using the web-based WebLogo software (http://weblogo.berkeley.edu/). Two genes that contain SAFB1 binding sites were selected for further validation using qRT-PCR with TaqMan gene expression assays (Section 2.7.3) and conventional PCR.

2.9.3 Polymerase chain reaction

Conventional polymerase chain reaction (PCR) was used to investigate any alternative splicing of exon 2 in the serglycin (*SRGN*) mRNA transcript. Primers spanning the whole of SRGN exon 2 region were designed using Primer-BLAST software (NCBI) based on the SRGN mRNA sequence (NCBI accession: NM_002727.2) and synthesised

by Invitrogen (Figure 2.3). The exact primer sequences used for this experiment are 5'-CAAATGCAGTCGGCTTGTCC-3' forward primer and 5'-CGTTAGGAAGCCA CTCCCAG-3' reverse primer. PCR reactions consisted of 3µl cDNA template, 0.25µl of each primer (forward and reverse, 0.5µg/µl), 12.5 µl PCR Master Mix (Promega) and 10µl DNase-free water (Promega). PCR reaction conditions were as follows: initial denaturation step of 95°C for 5 minutes, 35 cycles of denaturation at 95°C for 30 seconds, annealing at 57°C for 30 seconds and extension at 72°C for 60 seconds. This was followed by a final extension step of 72°C for 5 minutes and cooling to 4°C for storage. A 'no template control' was included and contained 3µl DNase-free water instead of cDNA template. Blue/Orange Loading Dye (Promega) was added to the PCR products and 10µl 1kb DNA ladder (Promega) before loading on to a 2% agarose gel (Sigma-Aldrich) made in Tris-Borate-EDTA (TBE) buffer (Sigma-Aldrich) containing 2µl ethidium bromide (10mg/ml; Sigma-Aldrich). PCR products were separated by electrophoresis at 90V for 45 minutes using PowerPAC 200 system (BioRad). The gel was visualised under ultraviolet (UV) light using BioSpectrum Multi Spectral Imager System (UVP).

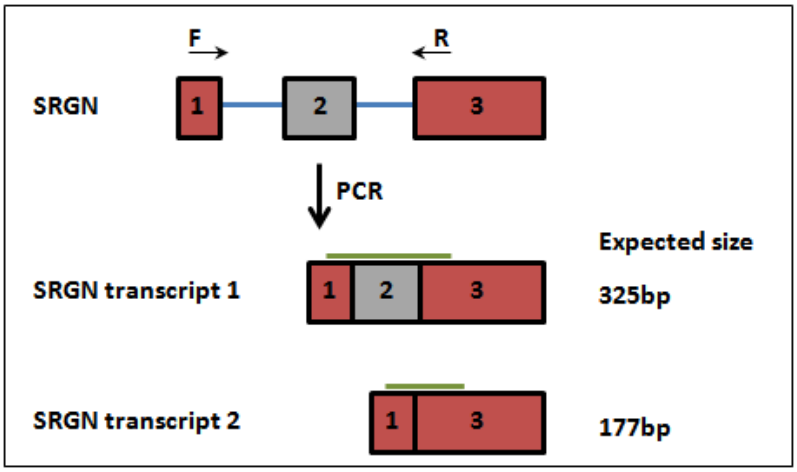


Figure 2.3 Schematic illustration of SRGN primers and PCR products.

Conventional PCR primers were designed within exon 1 and exon 3 to span across exon 2 (grey box). This primer design enables the identification of exon skipping event based on the size of PCR products obtained. F represents the forward primer and R represents the reverse primer.

Primer	Sequence
L3 linker	/5rApp/AGATCGGAAGAGCGGTTCAG/3ddC/
Rclip 4	X33NNAGGTNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC
Rclip 6	X33NNCCGGNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC
Rclip 8	X33NNCATTNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC
X33	5’ Phosphate
Cut oligo	GTTCAGGATCCACGACGCTCTTCaaaa
P5 Solexa	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT
P3 Solexa	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT

Table 2.8 List of primers used in iCLIP experiment.

Highlighted nucleotides represent unique barcode sequences for each replicate, red fonts represent BamH1 endonuclease recognition site.

(a) Lysis buffer	(b) High-salt buffer	(c) Wash buffer	(d) PNK mix	(e) Ligation mix
50mM Tris-HCl, pH 7.4 100mM NaCl 1% NP-40 0.1% SDS 0.5% sodium deoxycholate Protease inhibitors	50mM Tris-HCl, pH 7.4 1M NaCl 1mM EDTA 1% NP-40 0.1% SDS 0.5% sodium deoxycholate	20mM Tris-HCl, pH 7.4 10mM $MgCl_2$ 0.2% Tween-20	15µl water 4µl 5× PNK pH 6.5 buffer [350mM Tris-HCl, pH 6.5; 50mM $MgCl_2$; 25mM dithiothreitol] 0.5µl PNK enzyme 0.5µl RNasin	9µl water 4µl 4× ligation buffer [200mM Tris-HCl; 40mM $MgCl_2$; 40mM dithiothreitol] 1µl RNA ligase 0.5µl RNasin 1.5µl pre-adenylated linker L3 [20µM] 4µl PEG400
(f) PK buffer	**(g) RNA/primer mix**	**(h) Ligation mix**	**(i) Oligo annealing mix**	**(j) PCR mix**
100mM Tris-HCl pH 7.4 50mM NaCl 10mM EDTA	6.25µl water 0.5µl Rclip primer [0.5 pmol/µl] 0.5µl dNTP mix [10mM]	6.5µl water 0.8µl 10× CircLigase Buffer II 0.4µl 50mM $MnCl_2$ 0.3µl Circligase II	26µl water 3µl FastDigest Buffer 1µl cut oligo [10µM]	19µl cDNA 1µl primer mix P5/P3 Solexa [10µM each] 20µl Accuprime Supermix 1 enzyme

Table 2.9 Components of each buffer used in iCLIP experiment.

Chapter 3 : Characterisation of SAFB1 and SAFB2 expression in breast cancer cell lines

3.1 Introduction

SAFB proteins, particularly SAFB1 and SAFB2, have been widely studied and implicated in breast tumourigenesis (Section 1.4.4). The growing interest with SAFB1 and SAFB2 in relation to cancer is generated from their well described ability to bind to and modulate ER-α, a central player in breast cancer development.

Studies have indicated that the association of SAFB proteins with ER-α could be affected by the presence of ER-α ligands. Although SAFB proteins have been known to interact with ER-α regardless of oestradiol, the presence of the anti-oestrogen drug, tamoxifen, enhances the corepression effect of SAFB proteins (Oesterreich *et al.* 2000). Hashimoto *et al.* has shown that the presence of 17β-oestradiol initiates protein colocalisation of SAFB1 and SAFB2 with ER-α which decreases the intranuclear mobility of ER-α, thus inhibiting its function (Hashimoto *et al.* 2012). Apart from studies that investigate the functional effects of SAFB proteins on ER-α activity, there is a scarcity of knowledge on the regulation of SAFB1 and SAFB2 themselves in breast cancer cells.

My previous investigation showed that 17β-oestradiol is capable of differentially regulating SAFB1 and SAFB2 expression in ER positive and negative breast cancer cell lines (Hong *et al.* 2010). Interestingly, 17β-oestradiol influenced SAFB1 and SAFB2 expression in the breast cancer cell line MDA-MB-231, which is commonly known to be ER-α negative. This speculates a possible non-classical regulation potentially mediated by ER-α isoform present in this cell line (Wang *et al.* 2005). This preliminary observation suggests that SAFB1 and SAFB2 themselves may be regulated by the presence of ER ligand, in addition to their ability to regulate ER-α activity. However, this implication on SAFB1 and SAFB2 cellular function in breast cancer remains to be characterised. Therefore, the central focus of this chapter is to further understand the regulation of SAFB1 and SAFB2 in ER positive and ER negative breast cancer cell lines.

3.2 Aims

The aims of this chapter were to:

1. Validate the specificity of commercially available SAFB1 and SAFB2 antibodies used in further experimentation
2. Determine the effect of 17β-oestradiol stimulation on the expression of SAFB1 and SAFB2 in breast cancer cell lines
3. Examine the effect of 17β-oestradiol stimulation on cellular localisation of SAFB1 and SAFB2 in breast cancer cell lines
4. Investigate the possible post-translational modifications involved in the regulation of SAFB1 and SAFB2 by 17β-oestradiol
5. Study the effect of the drug fulvestrant on the expression of SAFB1 and SAFB2 in breast cancer cell lines

3.3 Results

3.3.1 Validation of SAFB1 and SAFB2 antibodies and siRNA oligonucleotides

The purification of SAFB1 and SAFB2 specific antibodies by previous groups has proven to be challenging due to the high sequence homology that they share. Most of the protein work relating SAFB expression to breast cancer has been performed using a pan-antibody, which recognises both SAFB1 and SAFB2 proteins but is unable to reliably distinguish them separately (personal communication). Commercial SAFB1- and SAFB2-specific antibodies have recently been made available; however the similarities between these two proteins necessitate the need to validate their specificity.

3.3.1.1 Knockdown of SAFB1 and SAFB2 using siRNA oligonucleotides

Due to the homology between SAFB1 and SAFB2, the specificity of their siRNA oligonucleotides also had to be validated to ensure knockdown using SAFB1 siRNA doesn't affect SAFB2 and vice versa. Knockdown of SAFB1 and

SAFB2 expression was performed using transient transfection of gene specific siRNA in MDA-MB-231 cells (Section 2.3.7). Whole cell lysate was extracted and analysed by immunoblotting using SAFB1 and SAFB2 antibodies respectively.

Using a SAFB1 specific antibody, a significant knockdown of SAFB1 protein was observed only in the sample transfected with SAFB1 siRNA and no knockdown was seen in the SAFB2 siRNA transfected sample (Figure 3.1). A similar outcome was observed using the SAFB2 antibody, whereby a significant knockdown of SAFB2 protein was observed only in the sample transfected with SAFB2 siRNA and not in the SAFB1 siRNA transfected sample. MDA-MB-231 cells were also transfected with GAPDH siRNA as a positive transfection control and a significantly reduced GAPDH level was observed in these cells. This validation included other experimental controls such as untreated cells, transfection reagent only (vehicle) treated cells and negative control siRNA treated cells to enable reliable comparison of protein levels within the same experiment.

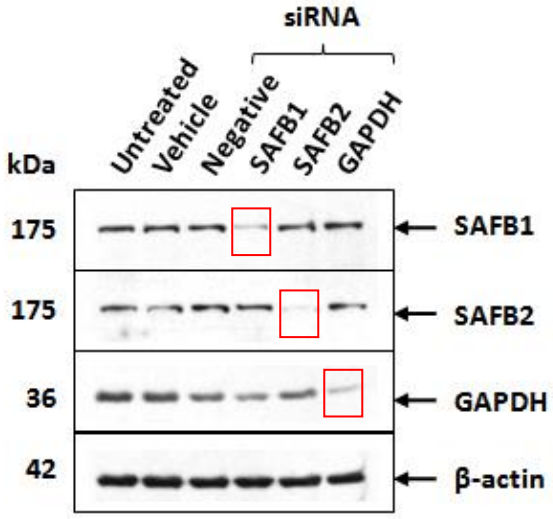


Figure 3.1 Validation of SAFB1 and SAFB2 antibody and siRNA specificity.

MDA-MB-231 cells in culture were transiently transfected with 12nM of siRNA targeting SAFB1 and SAFB2 for 72 hours prior to whole cell extraction for SDS-PAGE and immunoblotting. Gene knockdown was performed in parallel with experimental control conditions including untreated cells, vehicle-only treated cells, negative siRNA as negative control and GAPDH siRNA as positive control. Successful gene knockdowns are highlighted in the red boxes. β-actin expression was analysed to ensure equal loading between the samples. Data shown is a representative immunoblot result from at least three biological replicates.

3.3.1.2 Immunoprecipitation using SAFB1 antibody

The specificity of SAFB1 and SAFB2 antibodies was further tested by immunoprecipitation to rule out any antibody cross reactivity. MCF-7 cells under normal culture conditions were fractionated to obtain nuclear extracts for this part of the study (Section 2.4.2.1). SAFB1 and SAFB2 antibodies were used as 'capture' antibodies linked to Protein G magnetic beads to precipitate their respective epitope in 100µg of nuclear extracts (Section 2.5). Precipitated products were separated by SDS-PAGE and analysed by immunoblotting with SAFB1 and SAFB2 antibodies respectively, as 'detection' antibodies.

Immunoprecipitated protein was detected at approximately 175kDa by the SAFB1 antibody in the corresponding nuclear extract containing SAFB1 antibody as a capture [Figure 3.2 (a), lane 2] but not found in lane 3 containing SAFB2 immunoprecipitant. Mouse anti-IgG was included as an experimental control to distinguish background signal. SAFB2 antibody was not able to detect any protein through this experimental design [Figure 3.2 (b)]. This could possibly be linked to its intracellular localisation, as nuclear extracts were used for this experiment, however SAFB2 is not exclusively present in the nucleus but also in the cytoplasm (Townson *et al.* 2003, Weighardt *et al.* 1999).

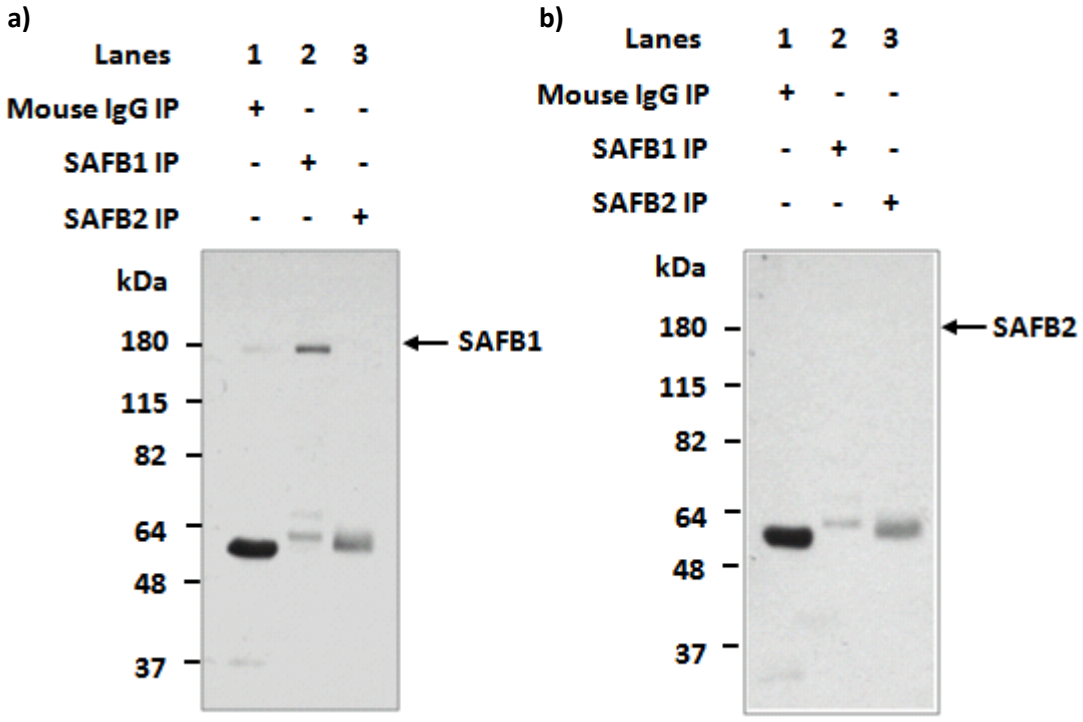


Figure 3.2 Immunoprecipitation of SAFB1 and SAFB2 proteins in MCF-7 cells.

Nuclear fractions of MCF-7 cells were used in immunoprecipitation with 3µg of SAFB1 (lane 2) and SAFB2 (lane 3) as the capture antibody. Mouse IgG (lane 1) was included as an experimental control. Precipitated products were separated on SDS-PAGE and analysed by immunoblotting with SAFB1 or SAFB2 antibodies. **(a)** Immunoblotting for SAFB1 specifically detects immuno-precipitated protein at approximately 175kDa in lane 2, a low background level in lane 1 but none in lane 3. **(b)** Immunoblotting using SAFB2 antibody fails to detect any immunoprecipitated protein at approximately 175kDa in any of the samples. Data shown is a representative immunoblot result from at least three biological replicates.

3.3.2 Oestrogen responsive SAFB1 and SAFB2 expression in breast cancer cell lines

Oestrogen is well recognised for its role in breast cancer development, exerting its effect on cell proliferation and differentiation via ER action (see Section 1.3). *SAFB1* and *SAFB2* have yet to be identified as oestrogen-response genes, however preliminary results from my previous study suggest this possibility (Hong *et al.* 2010). An interesting observation showed that the active metabolite, 17β-oestradiol was able to regulate SAFB1 and SAFB2 expression in MDA-MB-231 cells, reportedly an

ER negative breast cancer cell line. Therefore, it was of interest to further examine the effects of 17β-oestradiol stimulation on the regulation of SAFB1 and SAFB2 in breast cancer cell lines.

3.3.2.1 Effect of 17β-oestradiol on SAFB1 and SAFB2 expression

Two breast cancer cell lines with different ER status were chosen to establish and confirm my preliminary observations; MCF-7 cells are known to be an ER positive breast cancer cell line, whilst MDA-MB-231 cells are reported to be an ER negative breast cancer cell line. These cells were stimulated in culture with 17β-oestradiol using a range of concentration (0.01μM-10.0μM) for 24 hours prior to total RNA and whole cell lysate extraction (Section 2.3.6.1). qRT-PCR using validated TaqMan gene expression assays was performed to analyse *SAFB1* and *SAFB2* mRNA expression and whole cell lysates were used in immunoblotting to examine their protein expression.

In MCF-7 cells, increasing doses of 17β-oestradiol progressively increased the expression of SAFB1 and SAFB2 proteins [Figure 3.3 (a)]. A similar trend was observed for SAFB1 and SAFB2 mRNA transcript expression when the relative fold change was compared to untreated cells [Figure 3.3 (b)]. Interestingly, 17β-oestradiol has an opposite effect on SAFB1 and SAFB2 protein expression in the oestrogen non-responsive MDA-MB-231 cells. Increasing doses of 17β-oestradiol caused a progressive decrease in the expression of both SAFB1 and SAFB2 proteins [Figure 3.4 (a)]. However, this pattern of expression was not reflected at the mRNA level [Figure 3.4 (b)], suggesting that 17β-oestradiol may have an effect at the post-transcriptional or translational level of SAFB1 and SAFB2 in this particular ER negative cell line.

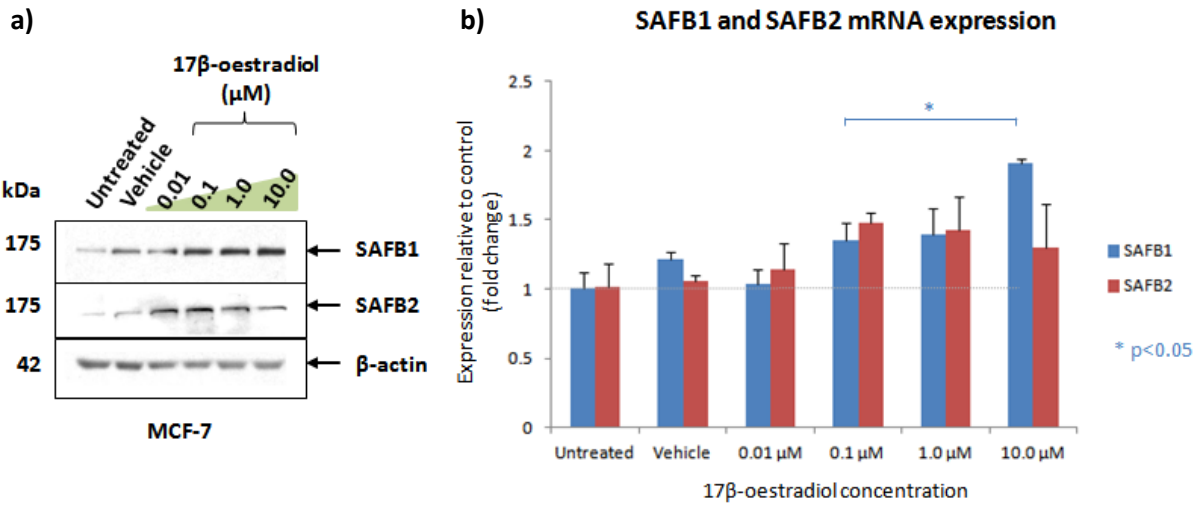


Figure 3.3 The effect of 17β-oestradiol on SAFB1 and SAFB2 expression in MCF-7 cells.

MCF-7 cells were cultured in the presence of 17β-oestradiol at the concentration of 0.01μM to 10.0μM for 24 hours prior to total RNA and whole cell lysate extraction. Experimental control conditions included were untreated cells and ethanol only (vehicle) treated cells. **(a)** A total of 50μg whole cell lysate were separated in SDS-PAGE and analysed by immunoblotting for SAFB1 and SAFB2 expression. β-actin expression was analysed to ensure equal loading between the samples. Data shown is a representative immunoblot result from at least three biological replicates. **(b)** mRNA extracted from MCF-7 cells was reverse transcribed into cDNA, then subjected to qRT-PCR using TaqMan gene expression assays targeting *SAFB1* and *SAFB2* gene. mRNA expression of each sample was normalised against a housekeeping gene, *β-actin*. *SAFB1* and *SAFB2* expression were compared against their corresponding expression in untreated cells using the comparative C_t method to generate a relative fold change. Data represents the average of three biological replicates ± S.D. Statistical significance for *SAFB1* expression was calculated using a student's t-test. * = $p<0.05$.

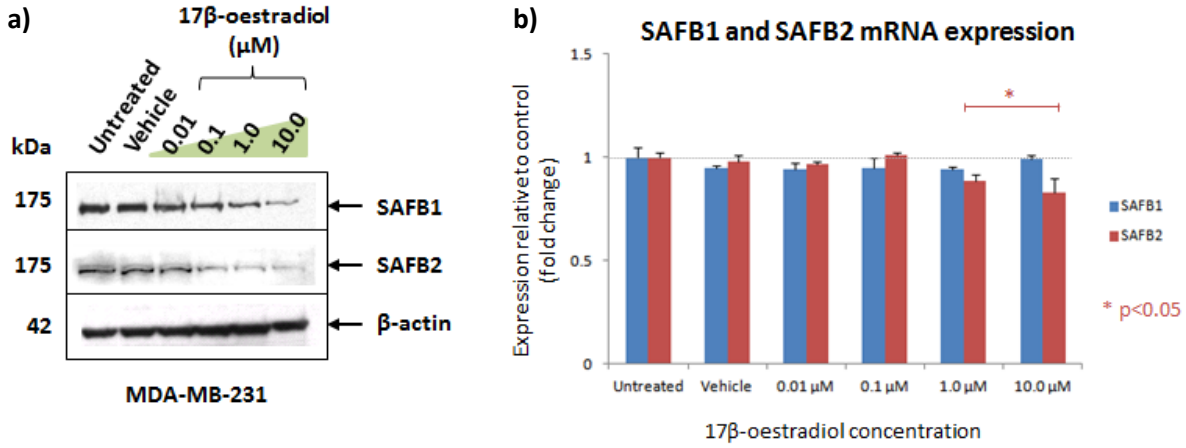


Figure 3.4 The effect of 17β-oestradiol on SAFB1 and SAFB2 expression in MDA-MB-231 cells.

MDA-MB-231 cells were cultured in the presence of 17β-oestradiol at the concentration of 0.01μM to 10.0μM for 24 hours prior to total RNA and whole cell lysate extraction. Experimental control conditions included were untreated cells and ethanol only (vehicle) treated cells. **(a)** A total of 50μg whole cell lysate were separated in SDS-PAGE and analysed by immunoblotting for SAFB1 and SAFB2 expression. β-actin expression was analysed to ensure equal loading between the samples. Data shown is a representative immunoblot result from at least three biological replicates. **(b)** mRNA extracted from MDA-MB-231 cells was reverse transcribed into cDNA, then subjected to qRT-PCR using TaqMan gene expression assays targeting *SAFB1* and *SAFB2* gene. mRNA expression of each sample was normalised against a housekeeping gene, *β-actin*. *SAFB1* and *SAFB2* expression were compared against their corresponding expression in untreated cells using the comparative C_t method to generate a relative fold change. Data represents the average of three biological replicates ± S.D. Statistical significance for *SAFB2* expression was calculated using a student's t-test. * = $p<0.05$.

3.3.2.2 Effect of 17β-oestradiol on SAFB1 and SAFB2 cellular localisation

The classical ligand bound ER-α66 is known to predominantly reside in the nucleus, however ER-α isoforms are shown to have extra-nuclear localisation (King *et al.* 1984, Monje *et al.* 2001, Wang *et al.* 2006). The cellular localisation of SAFB1 and SAFB2 proteins may affect their repressive function on ER; as previously shown by Hashimoto *et al.* (Hashimoto *et al.* 2012). This association between SAFB proteins localisation and ER-α has only been established in osteosarcoma (Saos-2) and primate kidney (Cos-1) cells; therefore the next step of this study was to examine the effect of

oestrogen on the cellular distribution of SAFB1 and SAFB2 in breast cancer cell lines using immunoblotting and immunofluorescent staining. Cultured MCF-7 and MDA-MB-231 cells were stimulated with two selected concentrations of 17β-oestradiol, 0.01µM and 1.0µM, or untreated for 24 hours (Section 2.3.6.1).

Cells were fractionated to separate the cytoplasmic and nuclear proteins, and then subjected to immunoblotting using SAFB1 and SAFB2 antibodies (Section 2.4.2.1). SAFB1 protein was detected specifically in the nuclear fractions of both MCF-7 [Figure 3.5 (a), top panel] and MDA-MB-231 [Figure 3.5 (b), top panel] in the absence and presence of 17β-oestradiol. Hormone stimulation did not to have any effect on SAFB1 cellular localisation in both these cell lines. In the absence and presence of 17β-oestradiol, SAFB2 protein was detected in the cytoplasmic and nuclear fractions of MCF-7 cells [Figure 3.5 (a), second panel] but only in the nuclear fractions of MDA-MB-231 cells [Figure 3.5 (b), second panel]. Higher dose of 17β-oestradiol (1.0µM) appears to induce a slight increase of SAFB2 nuclear localisation in MCF-7 cells but did not show such an effect in MDA-MB-231 cells. The purity of the fractions was confirmed in both cell lines by the detection of ERK-1 protein in the cytoplasmic fractions and c-Jun specifically in the nuclear fractions.

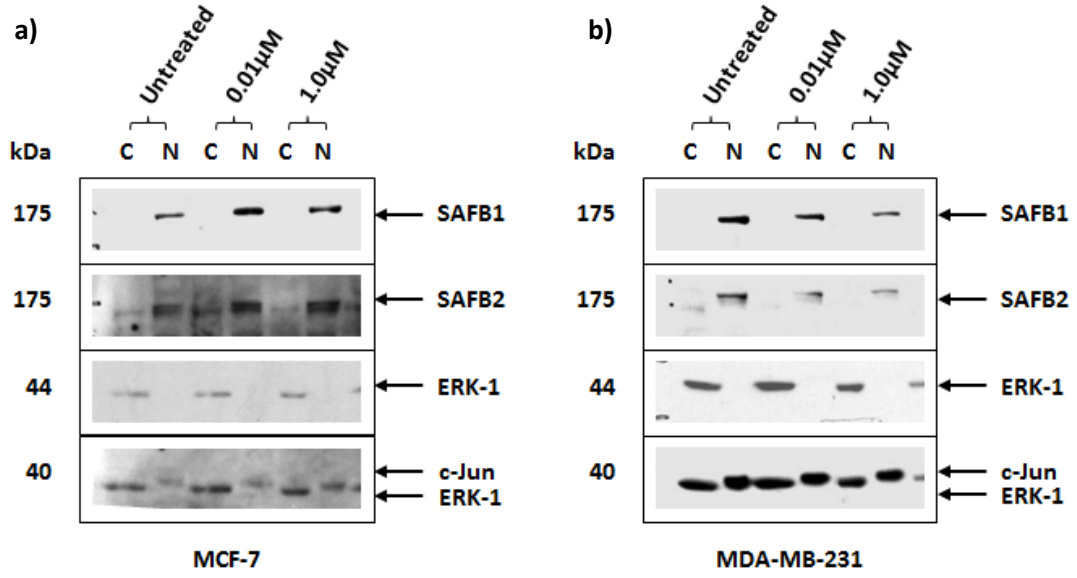


Figure 3.5 Detection of SAFB1 and SAFB2 in cytoplasmic and nuclear fractions of MCF-7 and MDA-MB-231 cells.

(a) MCF-7 and **(b)** MDA-MB-231 cells were stimulated with 0.01µM and 1.0µM 17β-oestradiol or untreated for 24 hours prior to cell fractionation to separate cytoplasmic and nuclear proteins. An equal amount of 50µg protein was used for

SDS-PAGE and immunoblotting to detect SAFB1 and SAFB2 expression. Each membrane was probed sequentially with ERK-1 and c-Jun to verify the purity of the fractions. Data shown is a representative immunoblot result from at least three biological replicates. C=cytoplasmic fraction, N=nuclear fraction.

To further validate the cellular distribution of SAFB1 and SAFB2 in these two breast cancer cell lines, immunofluorescent staining was performed using SAFB1 and SAFB2 antibodies to examine their endogenous expression within the cells (Section 2.6.3). Confocal laser microscopy revealed punctate patterns of SAFB1 distribution in the nucleus of MCF-7 in the absence of 17β-oestradiol [Figure 3.6, top panel]. Addition of 0.01μM 17β-oestradiol enhanced the distribution of SAFB1 protein in distinct nuclear speckles [Figure 3.6, middle panel], while 1.0μM 17β-oestradiol changed the punctate pattern to a diffuse nucleoplasmic pattern of SAFB1 [Figure 3.6, bottom panel].

In MDA-MB-231 cells, a punctate pattern of SAFB1 distribution was also observed in the absence of 17β-oestradiol [Figure 3.7, top panel]. Stimulation with 0.01μM and 1.0μM 17β-oestradiol induced redistribution of SAFB1 from punctate nuclear speckles to a diffuse nucleoplasmic pattern [Figure 3.7, middle and bottom panel].

In untreated conditions, endogenous SAFB2 showed a diffusely homogenous pattern in the nucleoplasmic region and slight cytoplasmic expression in MCF-7 cells [Figure 3.8, top panel] and MDA-MB-231 cells [Figure 3.9, top panel]. Concentrated nuclear staining was observed following the addition of 17β-oestradiol (0.01μM or 1.0μM) in both cell lines [Figure 3.8 and Figure 3.9].

The observations on SAFB1 distribution in the nuclear speckles prompted further investigation for a correlation with splicing factor, splicing component 35kDa (SC35) which is known to reside in nuclear speckles (Fu *et al.* 1990). Limitations of the SAFB1 antibody made co-localisation study challenging, however SC35 protein distribution showed similar punctate pattern indicative of nuclear speckles [Figure 3.10]; as observed for SAFB1 localisation in both MCF-7 and MDA-MB-231 cells [Figure 3.6 and Figure 3.7].

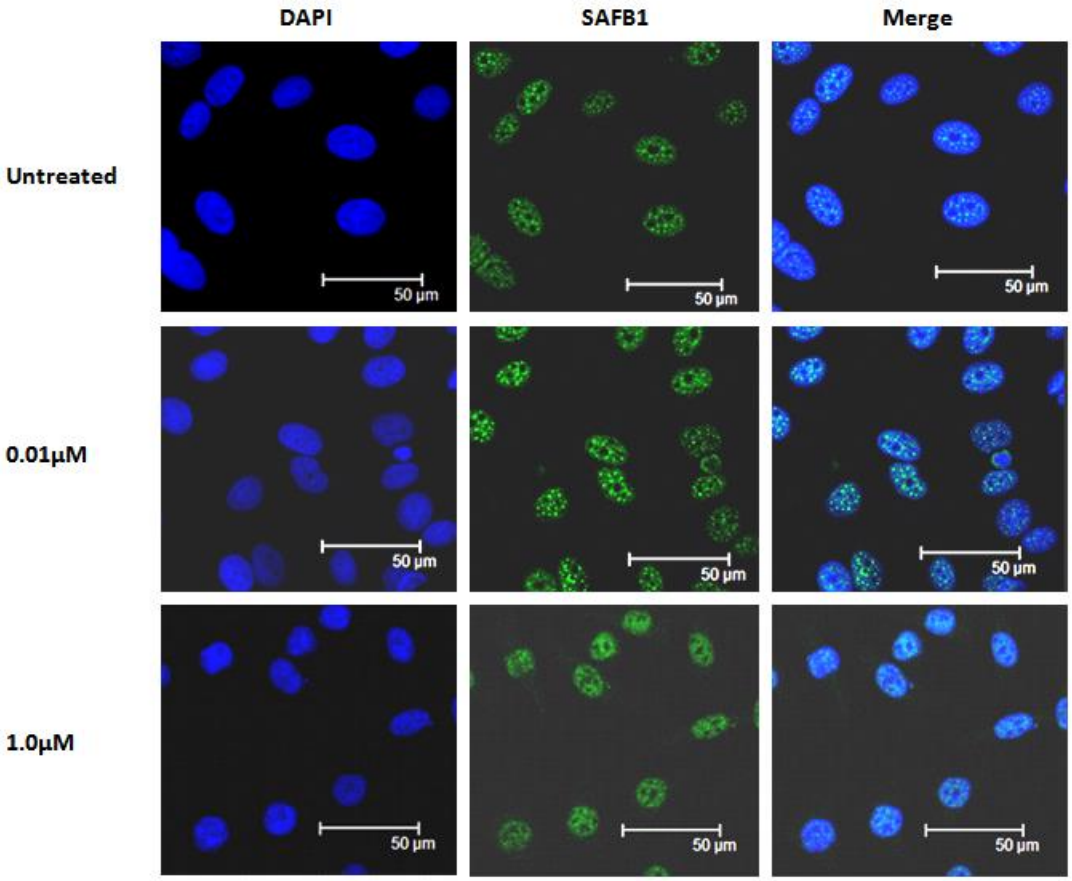


Figure 3.6 Distribution of SAFB1 in 17β-oestradiol stimulated MCF-7 cells.

MCF-7 cells were cultured on glass coverslips in the presence of 0.01µM or 1.0µM 17β-oestradiol for 24 hours prior to fixing and immunofluorescent staining. Nuclear staining was performed using DAPI (blue) and SAFB1 (green) staining using validated SAFB1 antibody. Untreated cells were also stained for comparison. Confocal laser microscopy at 63× magnification revealed punctate patterns of SAFB1 in nucleus speckles were enhanced by 0.01µM 17β-oestradiol and redistributed into a diffuse nucleoplasmic pattern by 1.0µM 17β-oestradiol.

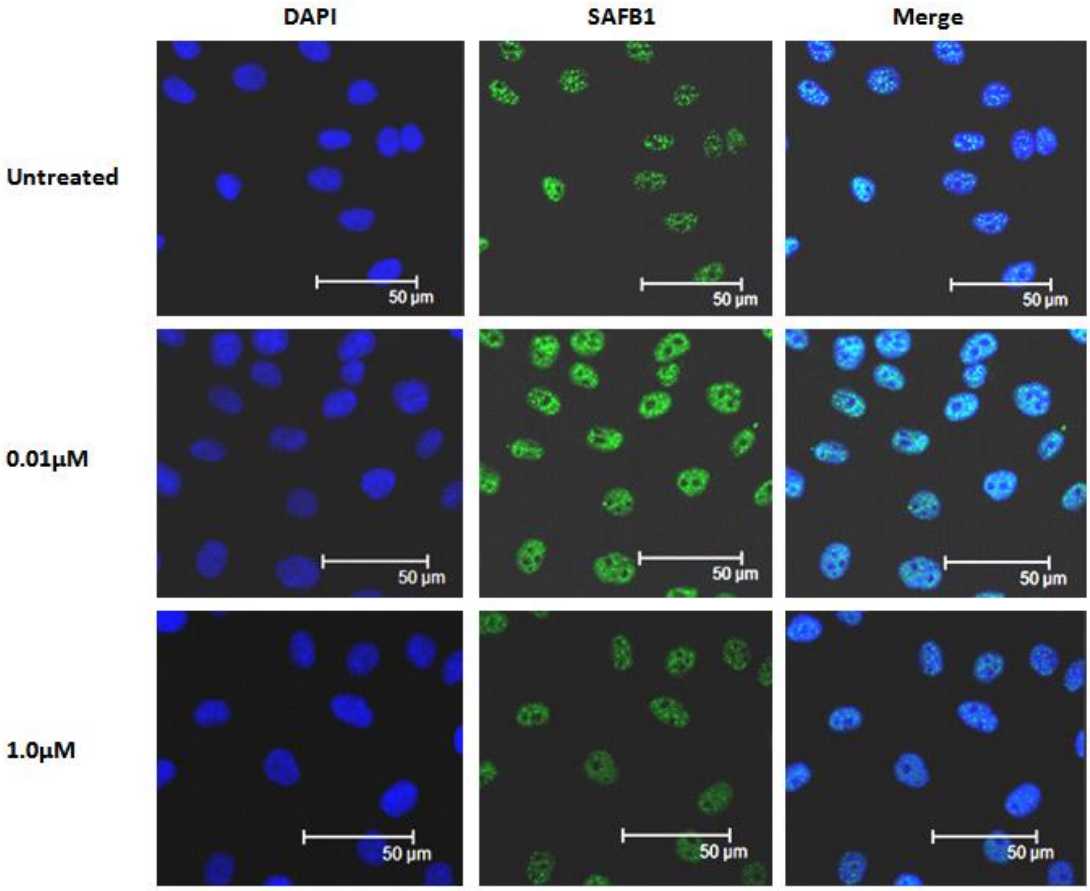


Figure 3.7 Distribution of SAFB1 in 17β-oestradiol stimulated MDA-MB-231 cells.

MDA-MB-231 cells were cultured on glass coverslips in the presence of 0.01µM or 1.0µM 17β-oestradiol for 24 hours prior to fixing and immunofluorescent staining. Nuclear staining was performed using DAPI (blue) and SAFB1 (green) staining using validated SAFB1 antibody. Untreated cells were also stained for comparison. Confocal laser microscopy at 63× magnification revealed punctate patterns of SAFB1 in nucleus speckles were redistributed to a diffuse nucleoplasmic pattern in response to 0.01µM and 1.0µM 17β-oestradiol stimulation.

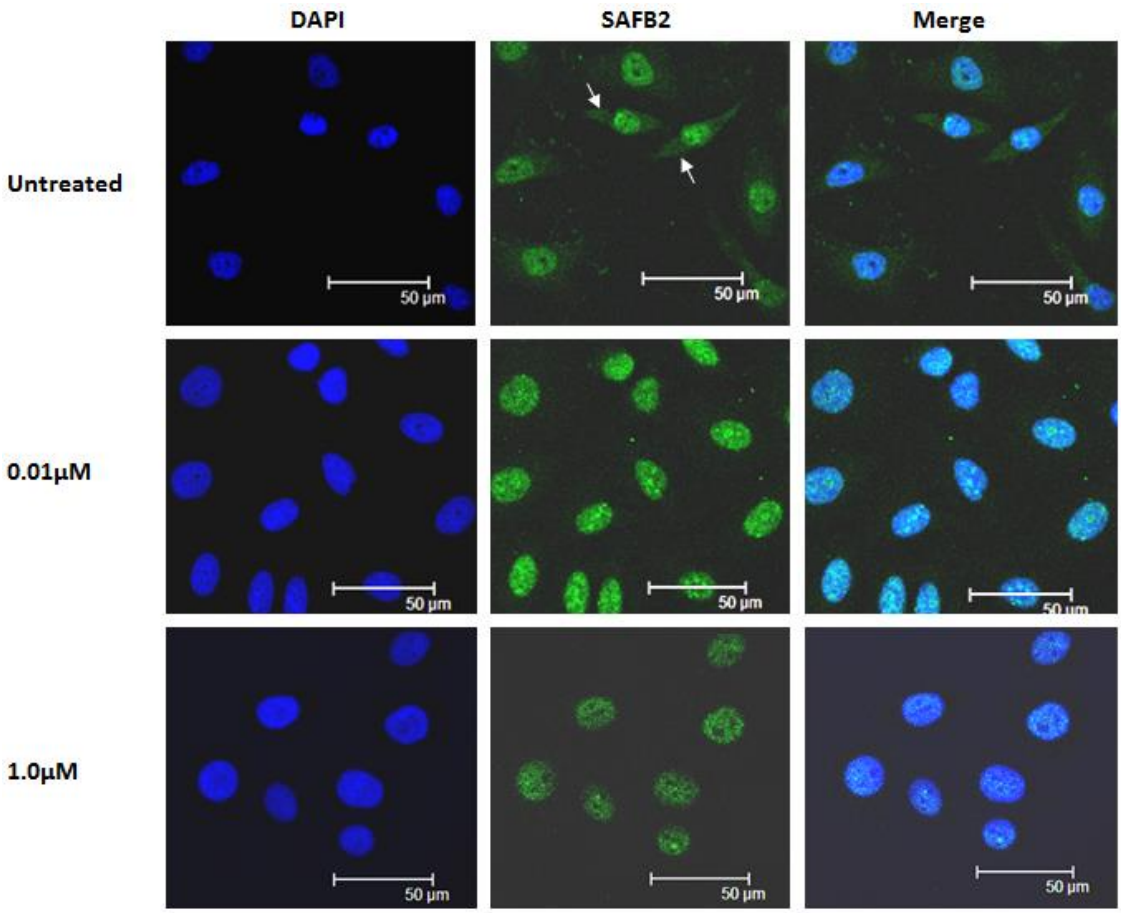


Figure 3.8 Distribution of SAFB2 in 17β-oestradiol stimulated MCF-7 cells.

MCF-7 cells were cultured on glass coverslips in the presence of 0.01µM or 1.0µM 17β-oestradiol for 24 hours prior to fixing and immunofluorescent staining. Nuclear staining was performed using DAPI (blue) and SAFB2 (green) staining using validated SAFB2 antibody. Untreated cells were also stained for comparison. Confocal laser microscopy at 63× magnification revealed a diffused pattern of SAFB2 in the nucleoplasmic and slight expression in the cytoplasmic region (arrow).

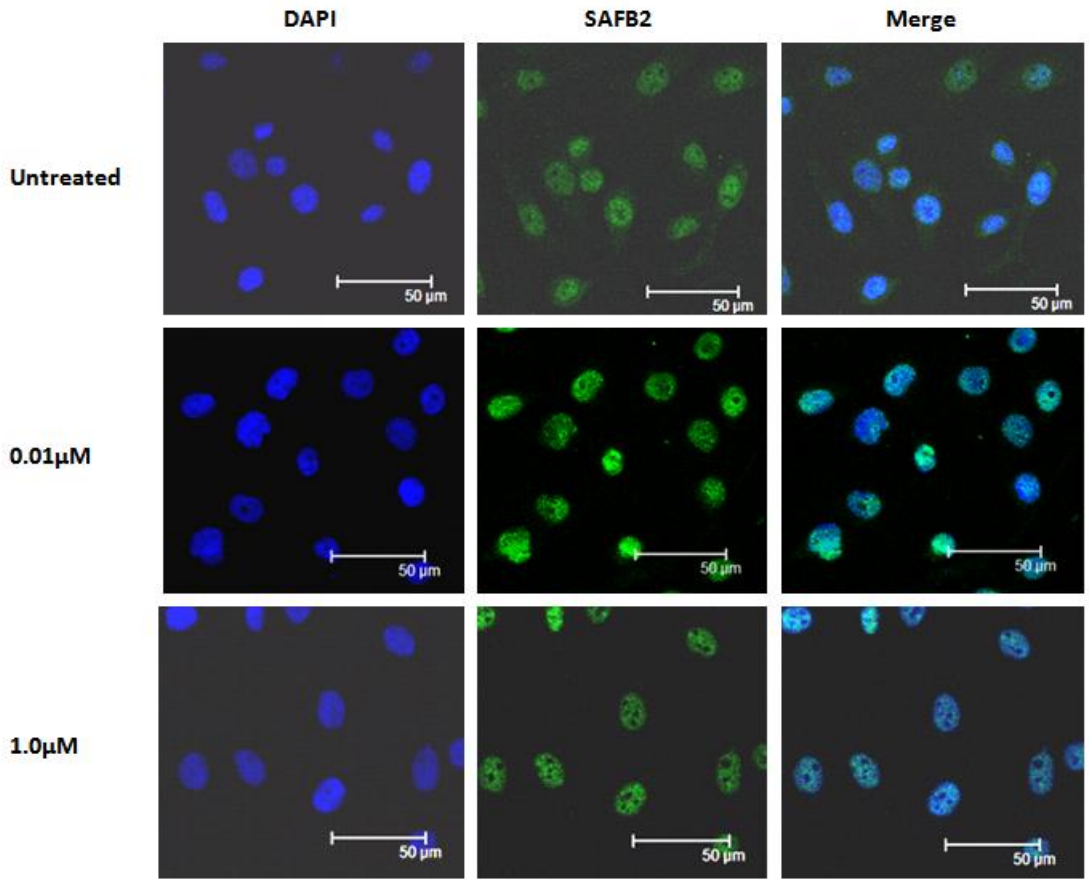


Figure 3.9 Distribution of SAFB2 in 17β-oestradiol stimulated MDA-MB-231 cells.

MDA-MB-231 cells were cultured on glass coverslips in the presence of 0.01µM or 1.0µM 17β-oestradiol for 24 hours prior to fixing and immunofluorescent staining. Nuclear staining was performed using DAPI (blue) and SAFB2 (green) staining using validated SAFB2 antibody. Untreated cells were also stained for comparison. Confocal laser microscopy at 63× magnification revealed a diffused pattern of SAFB2 in the nucleoplasmic and slight expression in the cytoplasmic region.

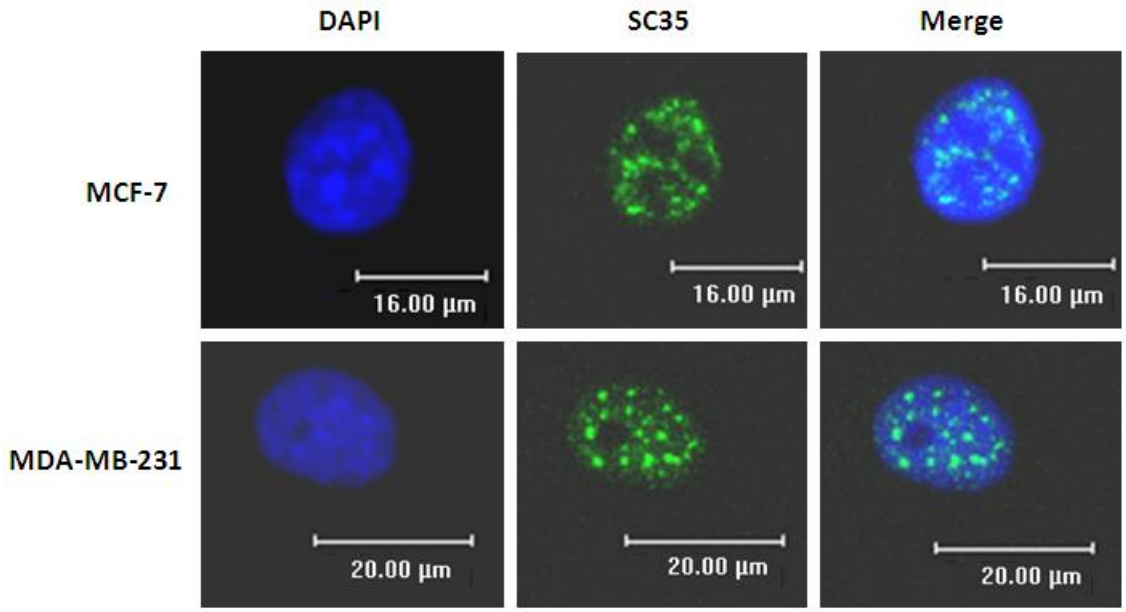


Figure 3.10 Intranuclear distribution of SC35 in MCF-7 and MDA-MB-231 cells.

MCF-7 and MDA-MB-231 cells were cultured on glass coverslips for 24 hours prior to fixing and immunofluorescent staining. Nuclear staining was performed using DAPI and SC35 staining using SC35 antibody. Confocal laser microscopy revealed a punctate pattern of SC35 in nuclear speckles similar to the observed SAFB1 distribution in these cells.

3.3.3 Oestrogen affects SAFB1 and SAFB2 protein stability in breast cancer cells

Oestrogen is known to regulate post-translational modifications and protein turnover that consequently affect the stability of a protein (Dery *et al.* 2003, Horner-Glister *et al.* 2005, Rebas *et al.* 2005). The ability of 17β-oestradiol to alter SAFB1 and SAFB2 proteins but not mRNA expression in MDA-MB-231 cells prompted investigations on the effect of oestrogen stimulation on SAFB1 and SAFB2 protein stability especially in this cell type (Section 3.3.2.1).

3.3.3.1 Effect of 17β-oestradiol on SAFB1 and SAFB2 protein phosphorylation

Post-translational modifications can regulate protein activity, cellular localisation and stability (Nishi *et al.* 2011, Ptacek *et al.* 2006). Numerous putative post-translational modifications such as phosphorylation, acetylation, methylation, ubiquitination and

sumoylation sites have been identified in SAFB1 and SAFB2 proteins (Garee *et al.* 2011, Lin *et al.* 2008, Song *et al.* 2011). Renz and Fackelmayer identified phosphorylated SAFB1 as its naturally-occuring state and suggest that post-translational modifications could possibly explain the high discrepancy between its apparent molecular weight (150-175kDa) and actual calculated molecular weight (102kDa) (Renz *et al.* 1996). Various reports have shown that 17β-oestradiol can modify the degree of protein phosphorylation while others have shown that phosphorylation can affect protein stability and function (Auger *et al.* 2001, Auricchio *et al.* 1987, Dery *et al.* 2003, Nishi *et al.* 2011, Ptacek *et al.* 2006, Rebas *et al.* 2005). Altered phosphorylation status of SAFB1 and SAFB2 proteins may affect their protein stability and potentially impede their mechanism of action; therefore the effect of 17β-oestradiol stimulation on SAFB1 and SAFB2 phosphorylation status was assessed.

MDA-MB-231 cells were stimulated in culture with 1.0μM 17β-oestradiol or untreated for 24 hours prior to separation of phosphorylated proteins using the PhosphoProtein purification kit (Section 2.4.2.2). Purified phosphorylated and unphosphorylated lysates were used in immunoblotting to examine SAFB1 and SAFB2 protein expression. SAFB1 and SAFB2 proteins were readily present in their phosphorylated form in MDA-MB-231 cells in the absence of hormone and the stimulation with 1.0μM 17β-oestradiol did not alter their state of protein phosphorylation [Figure 3.11(a)]. This demonstrates that 17β-oestradiol stimulation appears not to affect SAFB1 and SAFB2 phosphorylation modification. Equal loading for each fraction was validated using β-actin and serine/arginine-rich splicing factor 1 (SRSF1) antibodies. β-actin is an unphosphorylated protein while SRSF1 protein is extensively phosphorylated *in vivo* on serine residues within its serine/arginine-rich region (Colwill *et al.* 1996). The purity of the fractions was confirmed using a phosphoserine-specific antibody [Figure 3.11(b)].

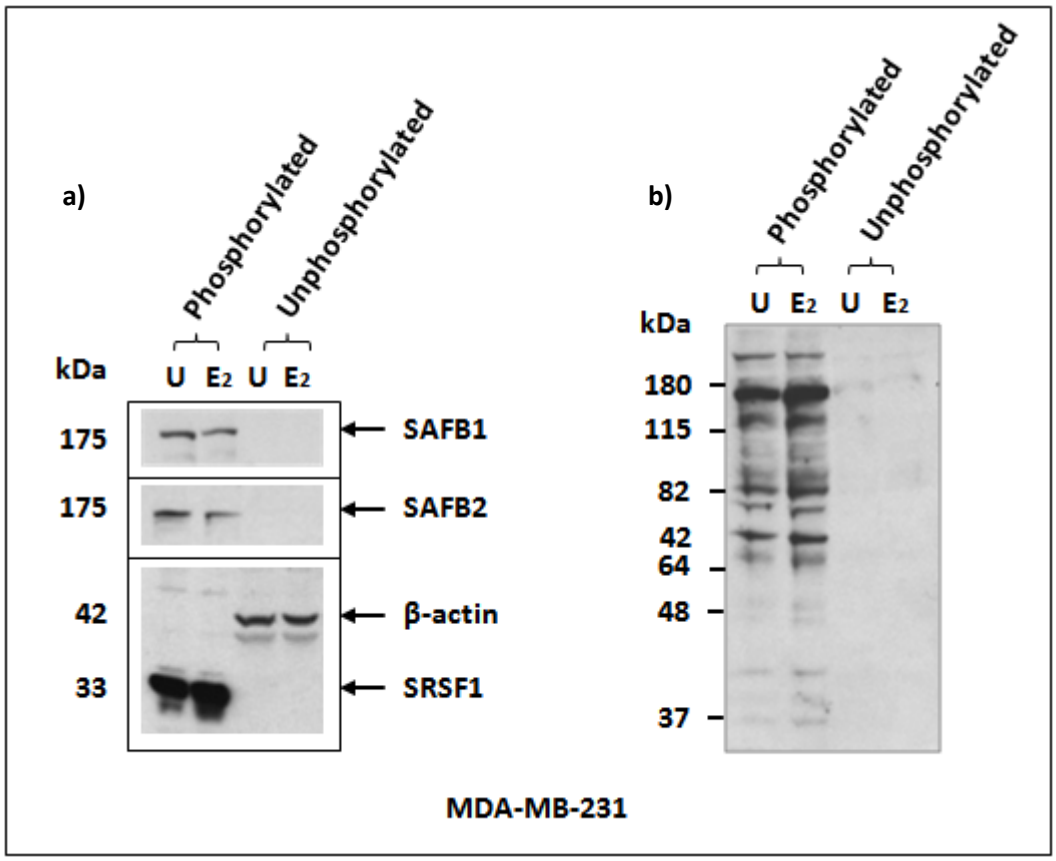


Figure 3.11 The effect of 17β-oestradiol on SAFB1 and SAFB2 protein phosphorylation.

MDA-MB-231 cells were cultured in the presence of 1.0μM 17β-oestradiol (E_2) or untreated (U) for 24 hours prior to the purification of phosphorylated and unphosphorylated protein fractions. A total of 15μg of each purified fractions were separated in SDS-PAGE and analysed by immunoblotting for SAFB1 and SAFB2 expression. **(a)** SAFB1 and SAFB2 were detected only in the phosphorylated fractions of both untreated and hormone stimulated cells; not in the unphosphorylated fractions. β-actin and SRSF1 expression were analysed in these samples to ensure equal loading. **(b)** The purity of the fractions was verified on the same membrane using a phosphoserine-specific antibody.

3.3.3.2 Effect of 17β-oestradiol on SAFB1 and SAFB2 protein turnover

The steady-state level of any protein is the outcome of the change in its rate of synthesis compared with its rate of degradation. The balance between these opposing processes, known as protein turnover, determine the concentration of a protein (Benaroudj 2005). Damaged or unneeded proteins are marked for destruction by the attachment of ubiquitin and subsequently degraded by proteasomes (Berg *et al.* 2002). Recent work by Song *et al.* has shown that BRCA1 induces SAFB2 ubiquitination *in vivo*

(Song *et al.* 2011). On the other hand, BRCA1 expression is positively regulated by oestrogen in breast cancer cells (Gudas *et al.* 1995, Romagnolo *et al.* 1998). Considering these reports together with the oestrogen induced downregulation of SAFB proteins observed in Section 3.3.2.1, the effect of 17β-oestradiol stimulation on SAFB1 and SAFB2 protein degradation was examined.

MDA-MB-231 cells were stimulated in culture with 17β-oestradiol at previously used concentrations (0.01μM-10.0μM) in the presence of 0.25μM MG132 proteasome inhibitor for 24 hours (Section 2.3.6.2). Whole cell lysate was extracted and used in immunoblotting to examine SAFB1 and SAFB2 protein expression. Oestrogen-induced decrease of SAFB1 and SAFB2 expression was blocked in the presence of MG132 [Figure 3.12 (right panel)].

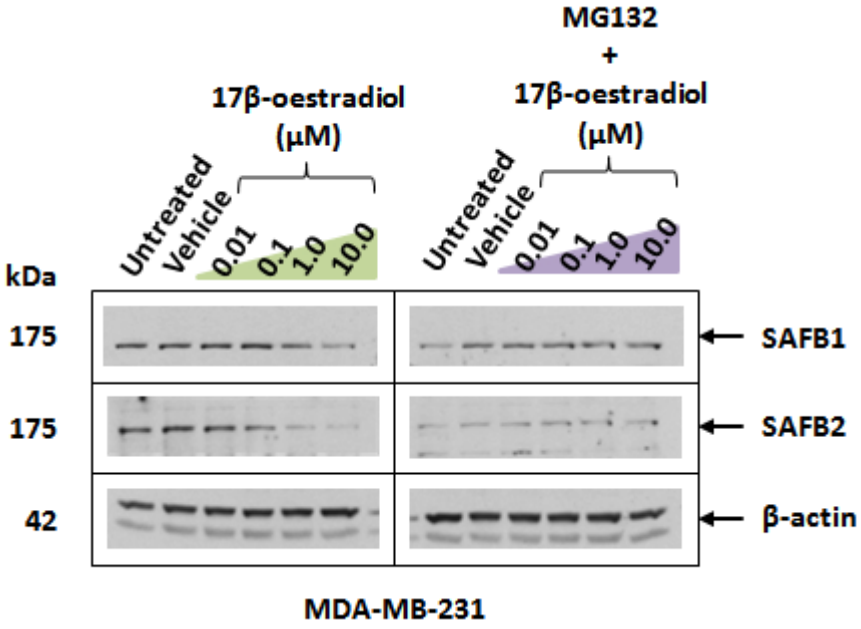


Figure 3.12 The effect of 17β-oestradiol on SAFB1 and SAFB2 pretreated with MG132.

MDA-MB-231 cells were stimulated with 17β-oestradiol at the concentration of 0.01μM to 10.0μM in the presence of 0.25μM MG132 proteasome inhibitor for 24 hours prior to whole cell lysate extraction (right panel). Cells were also stimulated with 17β-oestradiol only for parallel comparison (left panel). Experimental control conditions included were untreated cells and ethanol only (vehicle) treated cells. A total of 50μg whole cell lysate were analysed for SAFB1 and SAFB2 expression. As previously observed, SAFB1 and SAFB2 antibodies detected a gradual decrease in protein expression in response to 17β-oestradiol stimulation (left panel). This 17β-oestradiol induced effect was abolished in the presence of the MG132 proteasome inhibitor (right panel). β-actin expression was

analysed to ensure equal loading between the samples and a representative immunoblot from at least three biological replicates.

3.3.4 Anti-oestrogen alters SAFB1 and SAFB2 expression in breast cancer cell lines

Oestrogen responsiveness of breast tumours play a key role in treatment selection as it determines patients' responses to anti-oestrogen drugs. Conventionally, anti-oestrogen therapies are only used to treat patients with ER positive breast cancer (Section 1.2.5.4). However, the recent detection of truncated but functional ER-α (Section 1.3.1) and observed oestrogen-stimulated response in ER negative breast cancer (Section 3.3.2.1) beckons further investigation on the effect of anti-oestrogen drug in ER positive and ER negative breast cancer cells.

The 17β-oestradiol stimulated upregulation of SAFB1 and SAFB2 expression in MCF-7 cells and the opposite effect of protein downregulation in MDA-MB-231 cells mimic the characteristic of oestrogen-response genes (Section 3.3.2.1). To examine if their response to oestrogen was mediated by ER, the anti-oestrogen drug fulvestrant was selected for further experimentation. Fulvestrant is a routinely used pure ER antagonist that functions by downregulating and degrading ER-α (Dauvois *et al.* 1993).

Both ER positive and ER negative cell lines were treated with fulvestrant to study its effect on SAFB1 and SAFB2 expression. MCF-7 and MDA-MB-231 cells were stimulated in culture with fulvestrant at a range of concentration (0.001μM-1.0μM) for 24 hours prior to total RNA and whole cell lysate extraction (Section 2.3.6.3). qRT-PCR using validated TaqMan gene expression assays was performed to analyse *SAFB1* and *SAFB2* mRNA expression and whole cell lysates were used in immunoblotting to examine their protein expression.

In MCF-7 cells, increasing doses of fulvestrant decreased the expression of SAFB1 and SAFB2 proteins [Figure 3.13 (a)]. SAFB1 and SAFB2 mRNA transcript expression mirrors this progressively decreasing trend when the relative fold change was compared to untreated cells [Figure 3.13 (b)]. Conversely, fulvestrant treatment did not appear to substantially alter SAFB1 and SAFB2 protein or mRNA expression in MDA-MB-231 cells (Figure 3.14).

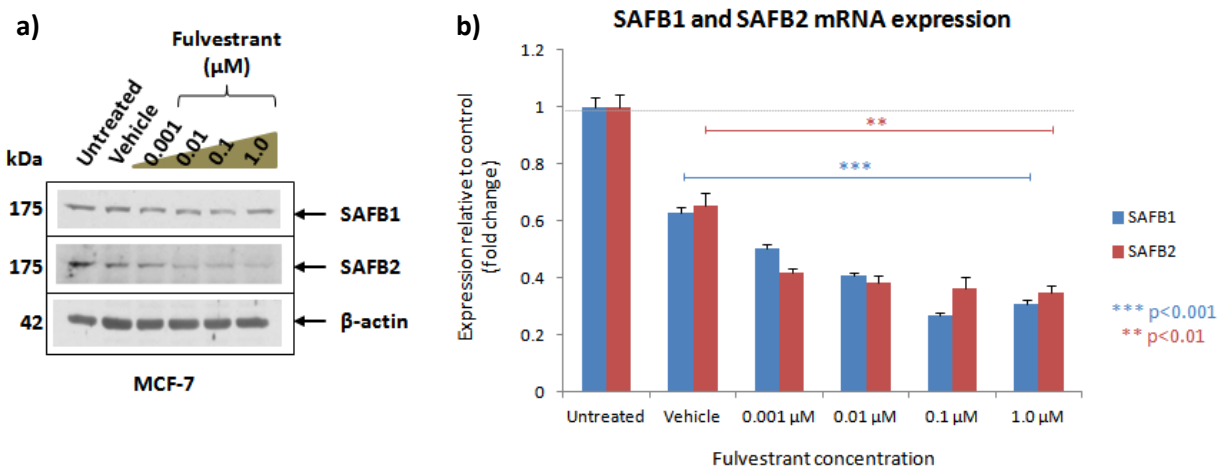


Figure 3.13 The effect of fulvestrant on SAFB1 and SAFB2 expression in MCF-7 cells.

MCF-7 cells were cultured in the presence of fulvestrant at the concentration of 0.001μM to 1.0μM for 24 hours prior to total RNA and whole cell lysate extraction. Experimental control conditions included were untreated cells and ethanol only (vehicle) treated cells. **(a)** A total of 50μg whole cell lysate were separated in SDS-PAGE and analysed by immunoblotting for SAFB1 and SAFB2 expression. SAFB1 and SAFB2 antibodies detected a gradually decreasing protein expression in response to fulvestrant stimulation. β-actin expression was analysed to ensure equal loading between the samples. Data shown is a representative immunoblot result from at least three biological replicates. **(b)** mRNA extracted from cells was reverse transcribed into cDNA, then subjected to qRT-PCR using TaqMan gene expression assays targeting *SAFB1* and *SAFB2* gene. mRNA expression of each sample was normalised against a housekeeping gene, *β-actin*. *SAFB1* and *SAFB2* expression were compared against their corresponding expression in untreated cells to generate a relative fold change. Data represents the average of three biological replicates ± S.D. Statistical significance of *SAFB1* and *SAFB2* expression was calculated using a student's t-test. **=p<0.01, *** = p<0.001.

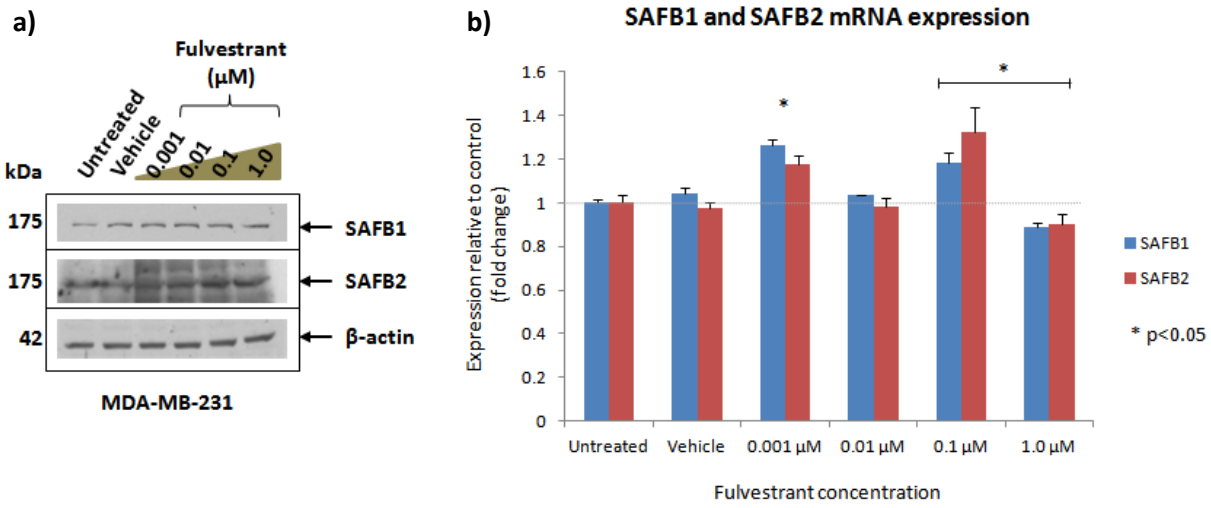


Figure 3.14 The effect of fulvestrant on SAFB1 and SAFB2 expression in MDA-MB-231 cells.

MDA-MB-231 cells were cultured in the presence of fulvestrant at the concentration of 0.001μM to 1.0μM for 24 hours prior to total RNA and whole cell lysate extraction. Experimental control conditions included were untreated cells and ethanol only (vehicle) treated cells. **(a)** A total of 50μg whole cell lysate were separated in SDS-PAGE and analysed by immunoblotting for SAFB1 and SAFB2 expression. SAFB1 and SAFB2 antibodies detected unaltered protein expression in response to fulvestrant stimulation. β-actin expression was analysed to ensure equal loading between the samples. Data shown is a representative immunoblot result from at least three biological replicates. **(b)** mRNA extracted from cells was reverse transcribed into cDNA, then subjected to qRT-PCR using TaqMan gene expression assays targeting *SAFB1* and *SAFB2* gene. mRNA expression of each sample was normalised against a housekeeping gene, *β-actin*. *SAFB1* and *SAFB2* expression were compared against their corresponding expression in untreated cells to generate a relative fold change. Data represents the average of three biological replicates ± S.D. Statistical significance of *SAFB1* and *SAFB2* expression was calculated using a student's t-test. $*$ = $p<0.05$.

3.3.5 Oestrogen and anti-oestrogen mediate a response in both ER positive and ER negative breast cancer cell lines

3.3.5.1 Oestrogen and anti-oestrogen affects ER protein expression

Interesting evidence from my findings in MDA-MB-231 cells led to extended investigations to examine its ability to mediate a hormone response. Although commonly described as an ER negative breast cancer cell line, the truncated ER-α36 variant is present in MDA-MB-231 cells and able to mediate non-genomic oestrogen

signalling (Wang *et al.* 2006, Zhang *et al.* 2012). As previously observed, differential expression of SAFB1 and SAFB2 was regulated by 17β-oestradiol and fulvestrant in MCF-7 and MDA-MB-231 cells (Section 3.3.2.1 and Section 3.3.4). To further elucidate the involvement of ER-α in the responsiveness to oestrogen and its antagonist, the presence and regulation of ER-α and its isoforms in both these cell lines were characterised.

Firstly, the effect of 17β-oestradiol stimulation on ER-α66 and ER-α36 in MCF-7 and MDA-MB-231 cells was investigated. Cells were stimulated in culture with 17β-oestradiol using a range of concentrations (0.01µM-10.0µM) for 24 hours prior to whole cell lysate extraction (Section 2.3.6.1). Immunoblotting was performed to analyse ER-α66 and ER-α36 protein expression. In MCF-7 cells, ER-α66 was readily present in the untreated sample but ER-α36 protein level was low and negligible [Figure 3.15 (a)]. 17β-oestradiol stimulation did not appear to significantly alter the expression of these receptors. As expected in MDA-MB-231 cells, ER-α66 was not present but ER-α36 protein was present in the untreated sample. The expression of these receptors was also not significantly altered by 17β-oestradiol stimulation [Figure 3.15 (b)].

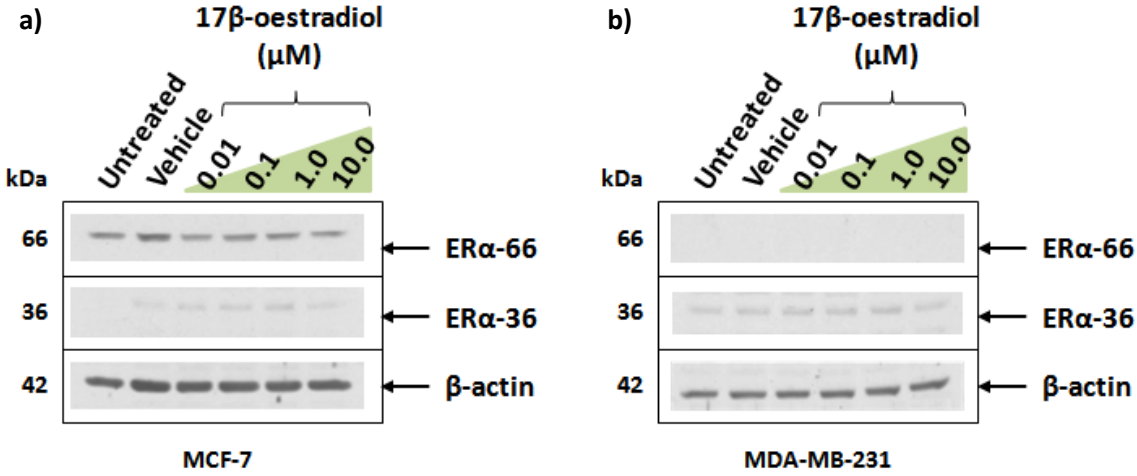


Figure 3.15 The effect of 17β-oestradiol on ER-α expression in MCF-7 and MDA-MB-231 cells.

(a) MCF-7 and **(b)** MDA-MB-231 cells were cultured in the presence of 17β-oestradiol at the concentration of 0.01µM to 10.0µM for 24 hours prior to whole cell lysate extraction. Experimental control conditions included were untreated cells and ethanol only (vehicle) treated cells. A total of 50µg whole cell lysate were separated in SDS-PAGE and analysed by immunoblotting for ER-α66 and ER-α36 protein expression. β-actin expression was analysed to ensure equal

loading between the samples. Data shown is a representative immunoblot result from at least three biological replicates.

The effect of fulvestrant treatment on ER-α66 and ER-α36 was also examined in the ER positive and ER negative cells. MCF-7 and MDA-MB-231 cells were stimulated in culture with fulvestrant at a range of concentration (0.001µM-1.0µM) for 24 hours prior to whole cell lysate extraction (Section 2.3.6.3). Immunoblotting was performed to analyse ER-α66 and ER-α36 protein expression. Increasing doses of fulvestrant stimulation in MCF-7 cells progressively downregulates ER-α66 expression, consistent with its function as an ER downregulator [Figure 3.16 (a)] (Peekhaus *et al.* 2004). ER-α36 protein expression was slightly increased in the presence of increasing doses of fulvestrant. In MDA-MB-231 cells, stimulation with increasing doses of fulvestrant gradually enhanced ER-α36 expression [Figure 3.16 (a)]. ER-α66 was again not detected in this cell line.

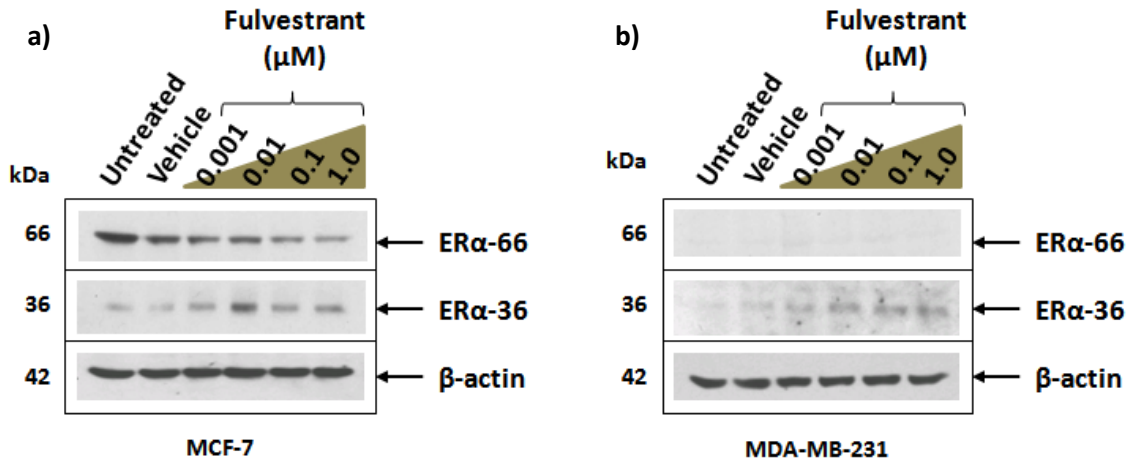


Figure 3.16 The effect of fulvestrant on ER-α expression in MCF-7 and MDA-MB-231 cells.

(a) MCF-7 and **(b)** MDA-MB-231 cells were cultured in the presence of fulvestrant at the concentration of 0.001µM to 1.0µM for 24 hours prior to whole cell lysate extraction. Experimental control conditions included were untreated cells and ethanol only (vehicle) treated cells. A total of 50µg whole cell lysate were separated in SDS-PAGE and analysed by immunoblotting for ER-α66 and ER-α36 protein expression. β-actin expression was analysed to ensure equal loading between the samples. Data shown is a representative immunoblot result from at least three biological replicates.

3.3.5.2 Oestrogen and anti-oestrogen affects oestrogen-response gene in MDA-MB-231 cells

The presence of ER-α36 and observed upregulation by anti-oestrogen in MDA-MB-231 cells prompted further assessment on its function and downstream effect on oestrogen-response genes. A bi-functional role of ER-α36 has been described by Wang *et al.* even at the onset of its discovery (Wang *et al.* 2006). They have shown that ER-α36 could effectively block and suppress ERE-dependent gene transactivation; while on the other hand, it could also elicit gene transcriptional activation via the MAPK/ERK pathway. To determine the significance of increased ER-α36 protein in MDA-MB-231 cells, an oestrogen-response gene was selected to establish a potential link to the anti-oestrogen induced ER-α36 expression (Section 3.3.5.1). Vascular endothelial growth factor A (VEGF-A) transcriptional activity is regulated by oestrogen stimulation in an ERE-dependent manner and therefore, was chosen as the oestrogen-response target gene for this part of the study (Applanat *et al.* 2008, Hyder *et al.* 1999, Mueller *et al.* 2000).

VEGF-A expression was examined in MDA-MB-231 cells stimulated with 17β-oestradiol. Cells were incubated in 17β-oestradiol at a range of concentration (0.01μM-10.0μM) for 24 hours prior to total RNA extraction (Section 2.3.6.1). qRT-PCR using validated TaqMan gene expression assays was performed to analyse *VEGF-A* mRNA expression. 17β-oestradiol stimulation did not appear to alter expression of *VEGF-A* in MDA-MB-231 cells (Figure 3.17).

VEGF-A expression was also examined in breast cancer cells stimulated with fulvestrant. MCF-7 and MDA-MB-231 cells were stimulated with fulvestrant at a range of concentration (0.001μM-1.0μM) for 24 hours prior to total RNA extraction (Section 2.3.6.3). qRT-PCR using TaqMan gene expression assay was performed to analyse *VEGF-A* mRNA expression. As expected in MCF-7 cells, the mRNA expression of *VEGF-A* inversely correlates with the dose concentration of fulvestrant [Figure 3.18 (a)]. Interestingly, a similar pattern of expression was observed in MDA-MB-231 cells, showing that fulvestrant could significantly downregulate *VEGF-A* mRNA expression in ER-negative cells [Figure 3.18 (b)]. *VEGF-A* protein expression were not examined due to challenges with antibody in immunoblotting.

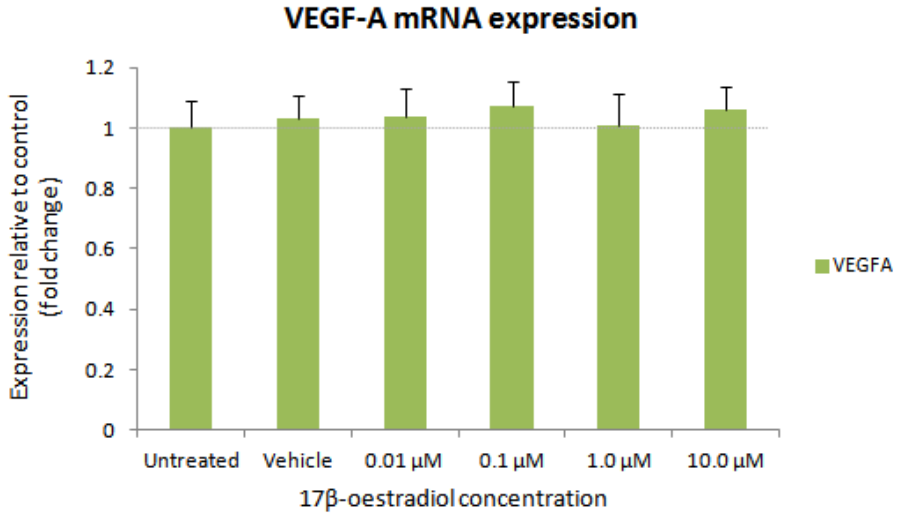


Figure 3.17 The effect of 17β-oestradiol on VEGF-A expression in MDA-MB-231 cells.

MDA-MB-231 cells were cultured in the presence of 17β-oestradiol at the concentration of 0.01µM to 10.0µM for 24 hours prior to total RNA extraction. Experimental control conditions included were untreated cells and ethanol only (vehicle) treated cells. mRNA extracted was reverse transcribed and subjected to qRT-PCR using TaqMan gene expression assays targeting *VEGF-A*. mRNA expression of each sample was normalised against a housekeeping gene, *β-actin*. *VEGF-A* expression were compared against its corresponding expression in untreated cells using the comparative C_t method to generate a relative fold change. Data represents the average of three biological replicates ± S.D. Statistical significance for *VEGF-A* expression was calculated using a student's t-test.

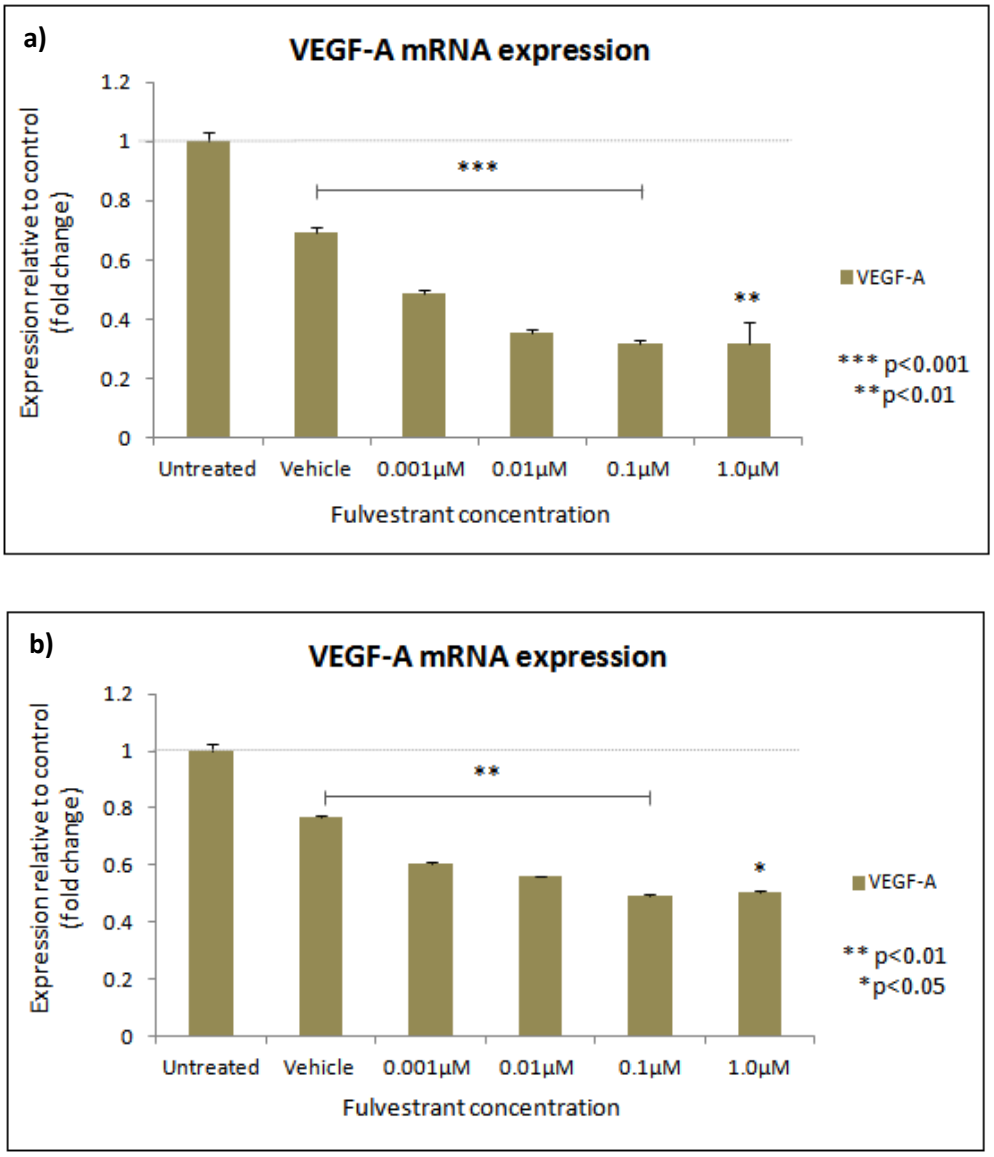


Figure 3.18 The effect of fulvestrant on VEGF-A expression in MCF-7 and MDA-MB-231 cells.

(a) MCF-7 and **(b)** MDA-MB-231 cells were cultured in the presence of fulvestrant at the concentration of 0.001µM to 1.0µM for 24 hours prior to total RNA extraction. Experimental control conditions included were untreated cells and ethanol only (vehicle) treated cells. mRNA extracted was reverse transcribed and subjected to qRT-PCR using TaqMan gene expression assays targeting *VEGF-A*. mRNA expression of each sample was normalised against a housekeeping gene, *β-actin*. *VEGF-A* expression were compared against its corresponding expression in untreated cells using the comparative C_t method to generate a relative fold change. Data represents the average of three biological replicates ± S.D. Statistical significance for *VEGF-A* expression was calculated using a student's t-test. $*$ = $p<0.05$, $**$ = $p<0.01$, $***$ = $p<0.001$.

3.4 Discussion

Many of the function and regulation of SAFB proteins are still unknown. In regards to breast cancer development, their role as a potential cause or consequence of the disease remains to be elucidated. Nevertheless, unanswered questions about their fundamental characteristics and regulation in breast cancer provide a vast and fascinating area of study. Preliminary work has shown interesting observations that suggest oestrogen-induced SAFB1 and SAFB2 regulation in breast cancer cells (Hong *et al.* 2010). The central focus of this chapter is to advance our current understanding on the regulation of SAFB proteins by oestrogen in breast cancer cells.

MCF-7 and MDA-MB-231 cells, representing ER positive and ER negative breast cancer cells respectively, were used as a model system to determine if the regulation of SAFB1 and SAFB2 is oestrogen-responsive. This was examined predominantly by stimulating cultured cells with the ER ligand (17β-oestradiol) or ER antagonist (fulvestrant) to observe the change in SAFB1 and SAFB2 expression. A novel discovery from this study revealed an ER-α66 dependent mechanism for SAFB1 and SAFB2 regulation in ER positive MCF-7 cells. Oestrogen-induced an increase in both protein and mRNA expression that mimics the classic effect of oestrogen-responsive genes (Figure 3.3); the anti-oestrogen induced a decrease in both protein and mRNA confirms the participation of ER-α66 in the regulation of these genes (Figure 3.13). The evaluation of *VEGF-A* mRNA expression and ER-α66 stability in response to fulvestrant further validates this notion (Figure 3.16 and Figure 3.18). An initial bioinformatics analysis using the Dragon ERE Finder web-based software predicted the presence of at least nine ERE on *SAFB1* and *SAFB2* genome sequence (Bajic *et al.* 2003). Considering these evidences, *SAFB1* and *SAFB2* could potentially be oestrogen-responsive genes; although further experimentation is needed to validate this speculation.

Another novel key finding from this study revealed that oestrogen may induce degradation of SAFB1 and SAFB2 proteins potentially via the ubiquitin-mediated protein degradation pathway in the ER negative MDA-MB-231 cells. 17β-oestradiol induced a progressive decrease of SAFB expression at the protein level which was not mirrored at the mRNA level (Figure 3.4). Similar to this inability to stimulate a change at the transcriptional level, the evaluation of *VEGF-A* mRNA expression confirms that 17β-oestradiol does not have an effect on the transcriptional activity of oestrogen-

responsive genes (Figure 3.17). Further investigations at the post-translational level revealed that 17β-oestradiol regulates SAFB proteins by disrupting their protein stability, but not by altering their protein phosphorylation (Figure 3.11). Post-translational modification such as phosphorylation is achieved by the adding or removing of a phosphate group on a protein that might alter its protein activity, subcellular localisation and stability (Nishi *et al.* 2011, Ptacek *et al.* 2006). Many tumour suppressors are regulated by phosphorylation and although 17β-oestradiol is capable of modifying the degree of protein phosphorylation, it did not exert the same effect on SAFB1 and SAFB2 in this system (Heilmann *et al.* 2012, Rizzolio *et al.* 2012, Trzepacz *et al.* 1997). On the other hand, 17β-oestradiol induced degradation of SAFB1 and SAFB2 was sensitive and protected by proteasome inhibition, suggesting that proteasome inhibition increases SAFB protein stability in MDA-MB-231 cells (Figure 3.12). Repression of ER corepressor regulated by oestrogen induced proteolysis can lead indirectly to an enhancement of transcriptional activation due to the elimination of a repressor. Several nuclear receptor corepressors are known to be modulated by proteasome-dependent proteolysis and oestrogen stimulated protein degradation through the ubiquitin-proteasome pathway have also been reported in other proteins (Alarid *et al.* 1999, Dace *et al.* 2000, Hoyt 1997, Zhang *et al.* 1998, Zhao *et al.* 2011). Specifically, proteasome inhibitor has been shown to block oestrogen-induced ER degradation by inhibiting proteasome activity, suggesting that oestrogen is capable of mediating protein turnover through the ubiquitin-proteasome pathway (Nawaz *et al.* 1999). As such, this evidence supports the possible occurrence for oestrogen-stimulated SAFB degradation through proteasome activity, as observed in the MDA-MB-231 cells. Furthermore, this result indicates that SAFB protein degradation may be linked to BRCA1 induced ubiquitination of SAFB2 potentially facilitated by oestrogen mediated BRCA1 upregulation (Romagnolo *et al.* 1998, Song *et al.* 2011). However, this remains a speculation and calls for further investigation.

Taking into account the data showing unaltered SAFB expression in response to the anti-oestrogen fulvestrant in MDA-MB-231 cells (Figure 3:14), it was initially unclear if this effect resulted from the inability of this ER negative cells to respond to anti-oestrogen or an antagonist-mediated opposing counter reaction that stabilises SAFB protein expression. The anti-oestrogen response was examined and data shows that MDA-MB-231 cells are capable of exerting an anti-oestrogen response as fulvestrant mediated a decrease in *VEGF-A* mRNA similar to that observed in MCF-7 cells

(Figure 3.18). The ability of this cell line to respond to anti-oestrogen stimulation suggests that fulvestrant appears to induce a 'protective' effect on SAFB protein stability opposed to the oestrogen-induced repression of these proteins.

As seen in this study, MDA-MB-231 cells that lack ER-α66 expression were able to mediate a response to oestrogen and anti-oestrogen stimulation. This ability to respond to oestrogen and anti-oestrogen stimulation has been linked to the presence of truncated ER-α36 in MDA-MB-231 cells (Wang *et al.* 2006, Zhang *et al.* 2012, Zhang *et al.* 2012). This study has also investigated the effects of oestrogen and anti-oestrogen stimulation on ER and its related isoforms. In previous reports, 17β-oestradiol has been shown to decrease the steady-state level of ER-α66 in MCF-7 cells (Alarid *et al.* 1999, Saceda *et al.* 1988). A similar effect was not observed in this study possibly because ER-α66 was examined after 24 hours of 17β-oestradiol stimulation [Figure 3.15 (a)], while significant ER-α66 downregulation was normally seen after 17β-oestradiol stimulation for more than 48 hours (Pink *et al.* 1996, Ree *et al.* 1989).

Fulvestrant is a potent ER-antagonist that has a binding affinity to ER-α66 that is 89% of 17β-oestradiol (Wakeling *et al.* 1987). Fulvestrant binding to ER impairs receptor dimerisation and blocks nuclear localisation of the receptor, resulting in an unstable complex that accelerates its degradation (Dauvois *et al.* 1993, Fawell *et al.* 1990). An expected effect of ER-antagonist was observed in MCF-7 cells, where increasing doses of fulvestrant successfully decreased ER-α66 expression [Figure 3.16 (a)]. Consistently, the expression of the oestrogen-response gene, *VEGF-A* mRNA decreased accordingly with fulvestrant stimulation [Figure 3.18 (a)]. Interestingly, fulvestrant mediated an increased ER-α36 protein expression in MDA-MB-231 cells, a previously unknown effect of this ER-antagonist [Figure 3.16 (b)]. This fulvestrant-induced ER-α36 expression is inversely correlated with *VEGF-A* mRNA expression, however their association is yet to be explored [Figure 3.18 (b)].

It has been well reported that SAFB1 and SAFB2 differ in their intracellular localisation. While SAFB1 is exclusively present in the nucleus, SAFB2 has shown cytoplasmic and nuclear staining as well as an interaction with vinexin (Chiodi *et al.* 2000, Townson *et al.* 2003, Weighardt *et al.* 1999). Consistent with previous reports, SAFB2 has also been found in the cytoplasmic fractions of MCF-7 and MDA-MB-231 cells (Figure 3.5). Considering the promiscuity of these

SAFB proteins, multiple levels of control of their actions would be expected (Debril *et al.* 2005). Their potential to influence a broad spectrum of cellular processes may be reliant on their spatio-temporal regulation, even within the nucleus (Hermanson *et al.* 2002). Although corepressors can translocate between the nucleus and cytoplasm, very little is understood regarding the regulation of shuttling corepressors. In this study, SAFB1 translocation did not occur in the presence of 17β-oestradiol; however high dose of 17β-oestradiol could induce SAFB2 translocation into the nucleus (Figure 3.5, Figure 3.8 and Figure 3.9).

Within the nucleus, SAFB1 is localised exclusively in a punctuate pattern in the nucleoplasm excluding the nucleoli (Figure 3.6 and Figure 3.7). The addition of a low dose of 17β-oestradiol enhances SAFB1 punctate staining in MCF-7 cells to a similar pattern observed in SC35 splicing factor (Figure 3.6 and Figure 3.10). SC35 is known to reside in nuclear speckles which are enriched for pre-mRNA metabolic factors (Fu *et al.* 1990, Hall *et al.* 2006). As SAFB1 has been reported to be involved in alternative splicing (Section 1.4.2.3), this observation suggests a possibility of enhanced SAFB1 localisation in nuclear speckles induced by 17β-oestradiol. An initial study has shown that SAFB1 co-localises with SC35, although subsequent studies showed that the SAFB1 speckles were distinct from SC35 nuclear speckles in the cervical cancer, HeLa cells (Chiodi *et al.* 2000, Nayler *et al.* 1998, Weighardt *et al.* 1999). However, co-localisation studies have yet to be performed for SAFB1 and SC35 in the MCF-7 cell line. For this study, both SAFB1 and SC35 antibodies were produced in a mouse host, therefore limiting the ability to perform specific co-localisation studies.

In summary, this study has revealed novel mechanisms of SAFB1 and SAFB2 regulation in ER positive and ER negative breast cancer cells; and provided evidence that both SAFB proteins are regulated by oestrogen. Collectively, data from this study forms an essential basis for the following work in this thesis.

Chapter 4 : Loss of SAFB1 and SAFB2 mediates expression of ER target genes in ER negative breast cancer cell line

4.1 Introduction

The role of SAFB proteins in transcriptional regulation has been well characterised (Section 1.4.2.2). SAFB1 and SAFB2 have the ability to control gene transcription through direct interaction at the promoter region or indirect interaction via transcription factors [reviewed in (Hong *et al.* 2012)]. Comparison of SAFB1 and SAFB2 target genes identified through ChIP and genome-wide expression array revealed a lack of significant similarities, suggesting a vital contribution of indirect protein-protein interaction with other transcription factors in mediating their role in transcriptional regulation (Hammerich-Hille *et al.* 2010).

The promiscuous nature of SAFB proteins has been observed through their association with various transcription factors. However, their interaction with the hormone receptor ER-α remains the most extensively studied and understood, especially in relation to breast cancer. SAFB proteins have been described as ER corepressors that bind and negatively modulate ER-α transcriptional activity (Oesterreich *et al.* 2000, Townson *et al.* 2003). The direct role of SAFB proteins on gene regulation in ER positive breast cancer cells has been investigated by Hammerich-Hille *et al*. In the ER positive MCF-7 breast cancer cells, oestrogen-mediated repression of ER-α target genes appears to be SAFB1 and ER-α dependent (Hammerich-Hille *et al.* 2010). In this study, most of the identified target genes were upregulated in the absence of SAFB1 or SAFB2, consistently reflecting their role as ER-α corepressors. SAFB1 appears to have more unique target genes than SAFB2 and many of these oestrogen-repressed genes are critical immune regulators and apoptotic genes (Hammerich-Hille *et al.* 2010).

Triple negative breast cancer patients, defined by ER, PR and HER2 negativity, have poor prognosis due to the aggressive tumour biology and lack of targeted therapy. Endocrine therapies are typically not administered to triple negative breast cancer patients based on the premise that it lacks steroid hormone receptors to mediate a treatment response. The ER status of breast cancer patients is determined by the screening for its full-length receptor, ER-α66, thus neglecting the presence of any other truncated forms of ER-α or PR (Cork *et al.* 2012, Pelekanou *et al.* 2012). The ability of the truncated ER-α36 to mediate anti-oestrogen signalling in triple negative breast cancer cells and endocrine therapy resistance in ER-α66 positive patients may challenge

the competency of current screening strategies and potentially lead to the consideration of ER-α36 status in patient selection for endocrine therapy (Lee *et al.* 2008, Shi *et al.* 2009, Zhang *et al.* 2012).

MDA-MB-231 cells are known as a triple negative metastatic breast cancer cell line and it has been demonstrated in this thesis that oestrogen and anti-oestrogen stimulation on these cells can trigger a response in the regulation of SAFB1 and SAFB2 expression (Section 3.3.2.1 and 3.3.4). It was also demonstrated in Section 3.3.5.1 and 3.3.5.2 that the anti-oestrogen drug fulvestrant could negatively regulate the ER-α target gene *VEGF-A* and also increase ER-α36 expression in MDA-MB-231 cells. These observations, in addition to other reported evidence of hormone responsiveness in MDA-MB-231 cells, present it as an ideal candidate to identify potential SAFB target genes from a panel of ER-regulated genes.

4.2 Aims

The aims of this chapter were to:

1. Identify potential oestrogen-mediated SAFB1 and SAFB2 target genes related to ER signalling pathway and breast cancer regulation in MDA-MB-231 cells
2. Validate the expression profile of selected SAFB1 and SAFB2 target genes using specific TaqMan assays
3. Investigate the effect of loss of SAFB1 and SAFB2 on the protein expression of target genes in MDA-MB-231 breast cancer cells

4.3 Results

4.3.1 Identification of SAFB1 and SAFB2 target genes using gene expression array

The role of SAFB proteins in the triple negative metastatic breast cancer cell line, MDA-MB-231 cells, was investigated using the Human Breast Cancer and Estrogen Receptor Signalling RT^2 Profiler PCR Array System (SABiosciences). This experiment was designed to identify changes in the levels of gene transcripts within a focused panel of ER-related genes which are altered upon the decrease of SAFB1 or SAFB2 by RNAi. This experimental strategy was selected as it mimics the *in vivo* effects of decreased SAFB expression in invasive breast tumours (Hammerich-Hille *et al.* 2009, Oesterreich S 2002).

4.3.1.1 Experimental design and optimisation of SAFB1 and SAFB2 RNAi in MDA-MB-231 cells

A combined approach of siRNA gene knockdown and qRT-PCR was employed to identify SAFB1 and SAFB2 target genes in MDA-MB-231 cells (Section 2.8). Cells were transiently transfected with siRNA targeting *SAFB1* or *SAFB2* to decrease their expression prior to gene expression profile study. Cells were also cotransfected with siRNA targeting both *SAFB1* and *SAFB2* to address the combined effects of these two paralogs. Transient transfection was performed for 48 hours, followed by stimulation with 0.01μM 17β-oestradiol for another 24 hours. The concentration of 17β-oestradiol and stimulation time was selected based on optimal effects previously observed (Section 3.3.2.1). RNA and whole cell lysate were extracted after transfection and analysed for the degree of gene knockdown in all experiments before they were subjected to qRT-PCR. The gene profiler array consists of 84 genes related to breast cancer regulation and ER signalling (Section 2.8.1). Once the samples were verified for sufficient levels of gene knockdown (50% decrease or more), mRNA was reverse transcribed into cDNA and subjected to qRT-PCR using the SYBR green detection method. This experiment was performed in three biological replicates to generate sufficient data to allow statistical analysis using comparative C_t method for relative fold change against negative control siRNA. The complete workflow of this experimental design is illustrated in Figure 4.1.

Level of SAFB1 and SAFB2 gene knockdown was examined at transcript and protein level prior to qRT-PCR. Significant knockdown of SAFB1 expression was observed in SAFB1 siRNA transfected and cotransfected samples; similarly for SAFB2 expression (Figure 4.2). qRT-PCR analysis confirmed that SAFB1 and SAFB2 mRNA levels were decreased by approximately 60% and 85% [Figure 4.2(a)]. Immunoblotting data showed a significant decrease in SAFB1 and SAFB2 protein levels as very low intensity or no bands were visible in these experimental conditions [Figure 4.2(b)]. The mRNA from biological triplicates was subsequently used for cDNA synthesis and qRT-PCR in the ER signalling pathway focused gene expression array.

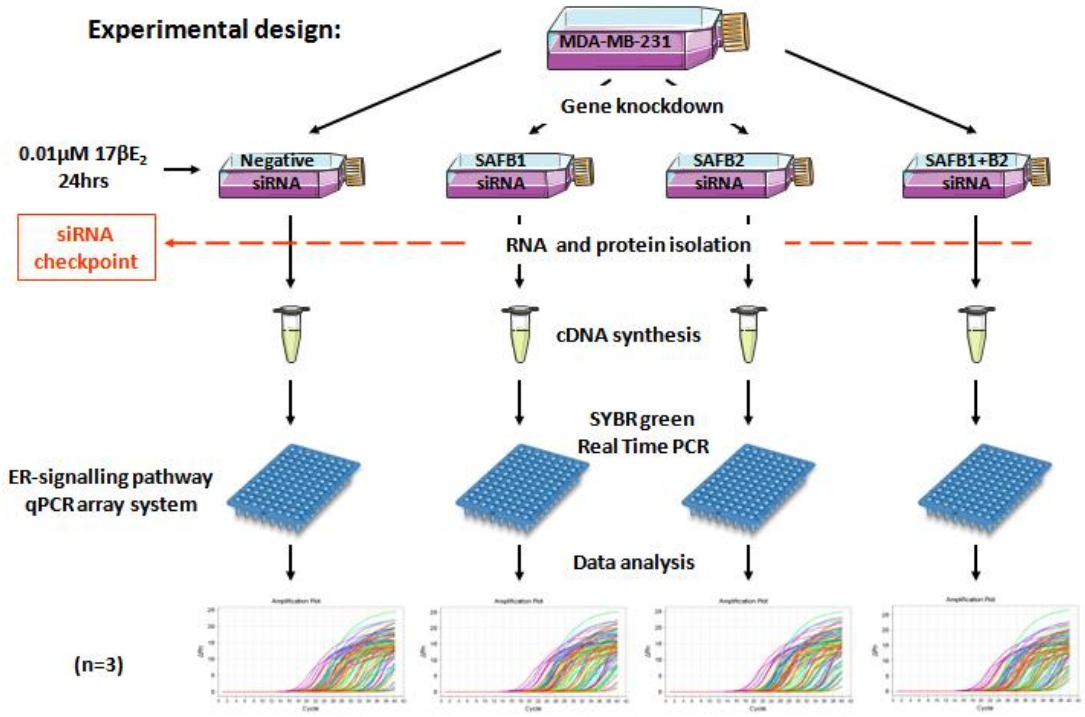


Figure 4.1 Gene expression profile study design.

MDA-MB-231 cells in culture were transiently transfected with 12nM of siRNA targeting *SAFB1* and *SAFB2* for 48 hours followed by stimulation with 0.01μM 17β-oestradiol for another 24 hours. Cells were also cotransfected with siRNA targeting both SAFB1 and SAFB2, and negative siRNA as control. RNA and whole cell lysate were analysed for the quality of gene knockdown before they were used in the gene expression array. Once verified, mRNA was reverse transcribed into cDNA and subjected to qRT-PCR in ER signalling pathway focused gene expression array using the SYBR green detection method. This procedure was performed in three biological replicates. Data generated was analysed using comparative C_t method for relative fold change against negative siRNA.

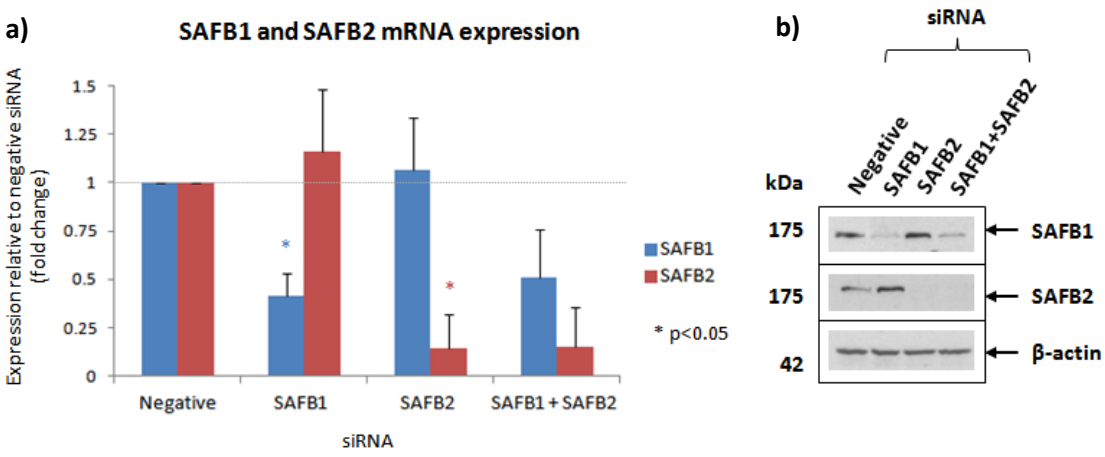


Figure 4.2 SAFB1 and SAFB2 downregulation by siRNA in MDA-MB-231 cells.

MDA-MB-231 cells were transiently transfected with negative, SAFB1, SAFB2 or SAFB1 and SAFB2 siRNA. Significant and specific downregulation by siRNA was observed at the mRNA and protein level. **(a)** mRNA levels were measured by qRT-PCR. Data represents the average of three biological replicates ± S.D. Statistical significance of mRNA expression was calculated using a student's t-test. * = $p<0.05$. **(b)** Protein levels were analysed by immunoblotting using SAFB1 and SAFB2 antibodies. β-actin expression was analysed to ensure equal loading between the samples. Data shown is a representative immunoblot result from at least three biological replicates.

4.3.1.2 SAFB1 and SAFB2 contributes greatly to gene regulation in MDA-MB-231 cells

The gene expression array data was analysed using RT^2 Profiler PCR Array Data Analysis software (SABiosciences) for gene expression. Analysis of gene expression was performed in comparison to negative siRNA and presented as relative fold change using the comparative C_t method. Using the recommended fold change threshold of more than 2.0 for gene upregulation and less than 0.5 for gene downregulation, the expression of genes across three experimental groups (SAFB1 siRNA, SAFB2 siRNA or SAFB1 + SAFB2 siRNA) were analysed to identify the SAFB1 or SAFB2 regulated gene set.

From this study, expression of 12 candidate genes altered in the loss of SAFB1 or/and SAFB2 [Figure 4.3(a)]. Specifically, expression of 11 genes increased while expression of only one gene decreased; reflecting the known role of SAFB1 and SAFB2 in transcriptional repression (Hammerich-Hille *et al.* 2010, Oesterreich *et al.* 2000,

Townson *et al.* 2003). The genes that were upregulated in the absence of SAFB1 or/and SAFB2 are cyclin-dependent kinase inhibitor 2A (*CDKN2A*), clusterin (*CLU*), oestrogen receptor 1 (*ESR1*), insulin-like growth factor binding protein 2 (*IGFBP2*), interleukin 2 receptor alpha (*IL2RA*), integrin beta 4 (*ITGB4*), v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog (*KIT*), kallikrein-related peptidase 5 (*KLK5*), metallothionein 3 (*MT3*), nerve growth factor receptor (*NGFR*) and small proline-rich protein 1B (*SPRR1B*). The only gene that was downregulated in the absence of SAFB1 or/and SAFB2 is interleukin 6 (*IL-6*). These genes are listed in Table 4.1 with their corresponding fold change in each experimental condition.

As shown in Table 4.1, the loss of SAFB1 altered the expression of 3 genes (*CLU*, *KIT* and *NGFR*) and the loss of SAFB2 altered the expression of 8 genes (*CLU*, *KIT*, *ESR1*, *IL2RA*, *ITGB4*, *KLK5*, *SPRR1B* and *IL-6*). Combined knockdown of SAFB1 and SAFB2 regulated the expression of 10 genes (*CDKN2A*, *CLU*, *KIT*, *ESR1*, *IGFBP2*, *IL2RA*, *KLK5*, *MT3*, *SPRR1B* and *IL-6*), suggesting a possible relationship between SAFB1 and SAFB2 in gene regulation.

The overlapping genes between these three experimental groups (genes regulated by SAFB1 siRNA, SAFB2 siRNA or SAFB1 + SAFB2 siRNA) were investigated to identify individual contributors and combined effects of SAFB1 and SAFB2 in gene regulation [Figure 4.3(b)]. *CLU* and *KIT* appear to be upregulated in all three experimental groups, suggesting that the loss of either SAFB1 or SAFB2 is sufficient to alter the expression of both these genes. On the other hand, altered *NGFR* expression appears to be specific to the loss of SAFB1 while the upregulation of *ITGB4* is specific in the absence of SAFB2. Interestingly, the expression of *CDKN2A*, *IGFBP2* and *MT3* is only upregulated in the absence of both SAFB1 and SAFB2; suggesting a possible compensatory effect between SAFB1 and SAFB2 in the regulation of these genes.

	RefSeq	Official gene symbol	Gene name	Fold change		
				SAFB1 siRNA	SAFB2 siRNA	SAFB1 +SAFB2 siRNA
1	NM_000077	CDKN2A	Cyclin-dependent kinase inhibitor 2A	1.0953	1.179	3.0851
2	NM_001831	CLU	Clusterin	3.0384	4.2004	3.266
3	NM_000125	ESR1	Oestrogen receptor 1	1.3817	2.3556	3.4701
4	NM_000597	IGFBP2	Insulin-like growth factor binding protein 2	1.5266	1.6529	2.4272
5	NM_000417	IL2RA	Interleukin 2 receptor alpha	1.9672	2.9273	4.5411
6	NM_000213	ITGB4	Integrin beta 4	1.5481	2.4671	1.7421
7	NM_000222	KIT	V-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog	2.1898	2.4395	2.2595
8	NM_012427	KLK5	Kallikrein-related peptidase 5	1.3505	2.0235	3.6878
9	NM_005954	MT3	Metallothionein 3	1.6914	0.9267	3.2198
10	NM_002507	NGFR	Nerve growth factor receptor	2.5525	1.8195	1.9933
11	NM_003125	SPRR1B	Small proline-rich protein 1B	1.5543	2.2703	4.4874
12	NM_000600	IL-6	Interleukin 6	0.5521	0.3912	0.3515

Table 4.1 SAFB1 and SAFB2 target genes identified by gene expression array in MDA-MB-231 cells.

The table contains 12 candidate genes identified through the three experimental groups (SAFB1 siRNA, SAFB2 siRNA or SAFB1 + SAFB2 siRNA). The fold change of ≥ 2.0 is highlighted in red, while ≤ 0.5 is highlighted in green. The expression of 11 genes was upregulated while the expression of only 1 gene (*IL-6*) was downregulated in the absence of SAFB1 or/and SAFB2.

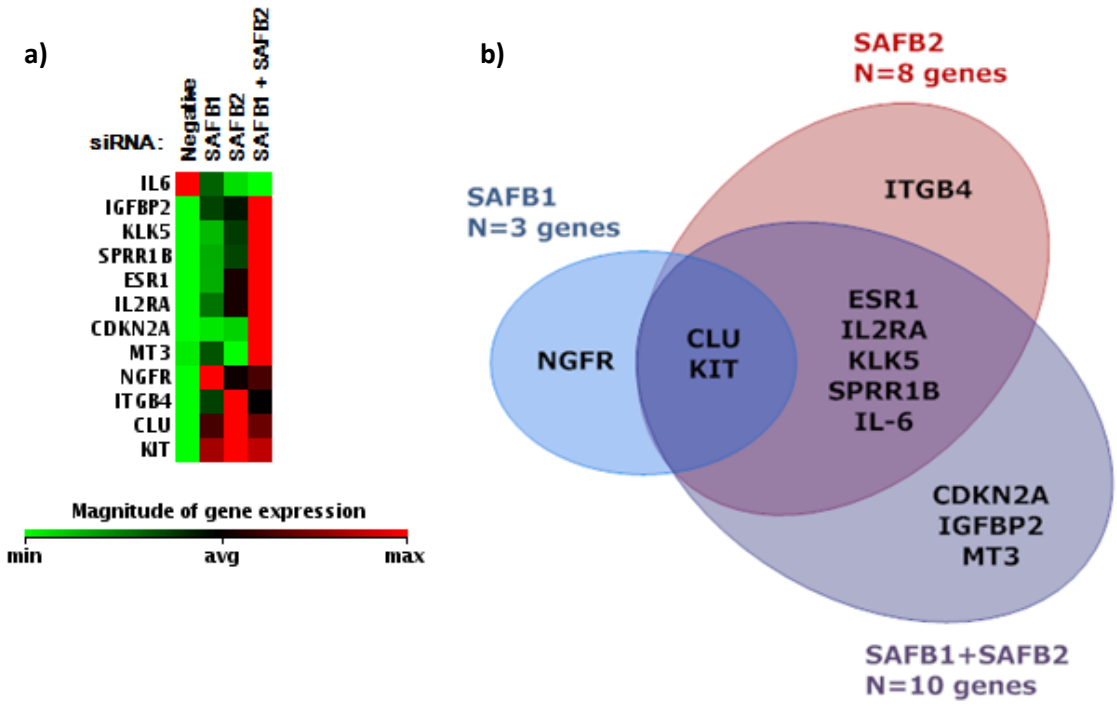


Figure 4.3 Identification of SAFB1, SAFB2, and SAFB1 and SAFB2 target genes in MDA-MB-231 cells.

Gene expression array data were analysed against negative siRNA and presented as fold change. The fold change threshold of ≥ 2.0 signifies gene upregulation and ≤ 0.5 signifies gene downregulation. Gene expression across three experimental groups (SAFB1 siRNA, SAFB2 siRNA or SAFB1 + SAFB2 siRNA) was compared. **(a)** The heat map represents 12 candidate genes regulated by SAFB1 siRNA, SAFB2 siRNA or the combination of SAFB1 and SAFB2 siRNA. Expression levels are shown in red and green, representing levels above and below the median respectively. **(b)** The Venn diagram represents genes that are regulated by SAFB1 siRNA, SAFB2 siRNA or SAFB1 + SAFB2 siRNA. While the relationship between SAFB1 and SAFB2 is important in the regulation of most ER-related genes, SAFB2 regulates more target genes than SAFB1 in MDA-MB-231 cells.

4.3.1.3 SAFB1 and SAFB2 target genes in MDA-MB-231 cells are key players in breast tumourigenesis

The identity of the 12 candidate genes were further investigated to gain insight into their functions. These genes were evaluated for their associated Gene Ontology terms (UniProt-GOA) and analysis revealed that SAFB1 and SAFB2 target genes are involved in important biological processes including apoptosis (*CDKN2A, CLU, NGFR*), cell proliferation (*IGFBP2, MT3, KIT*), immune response *(IL-6, IL2RA*), cellular adhesion

(*ITGB4*, *KLK5*, *SPRR1B*) and transcriptional regulation (*ESR1*) (Figure 4.4). Considering the pivotal role of these genes in tumourigenesis, the loss of SAFB1 or/and SAFB2 that leads to the de-repression of these oestrogen-regulated genes may contribute to the invasive properties of this metastatic breast cancer cell line.

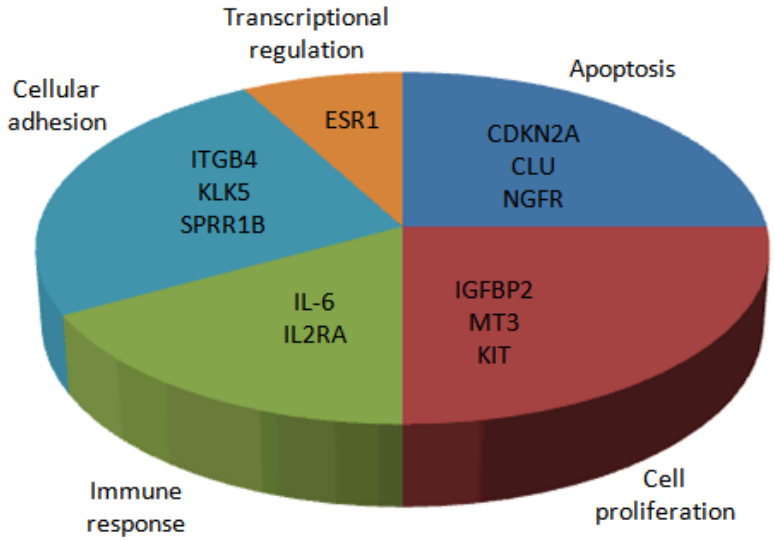


Figure 4.4 SAFB1 and SAFB2 regulate key players of tumourigenesis in MDA-MB-231 cells.

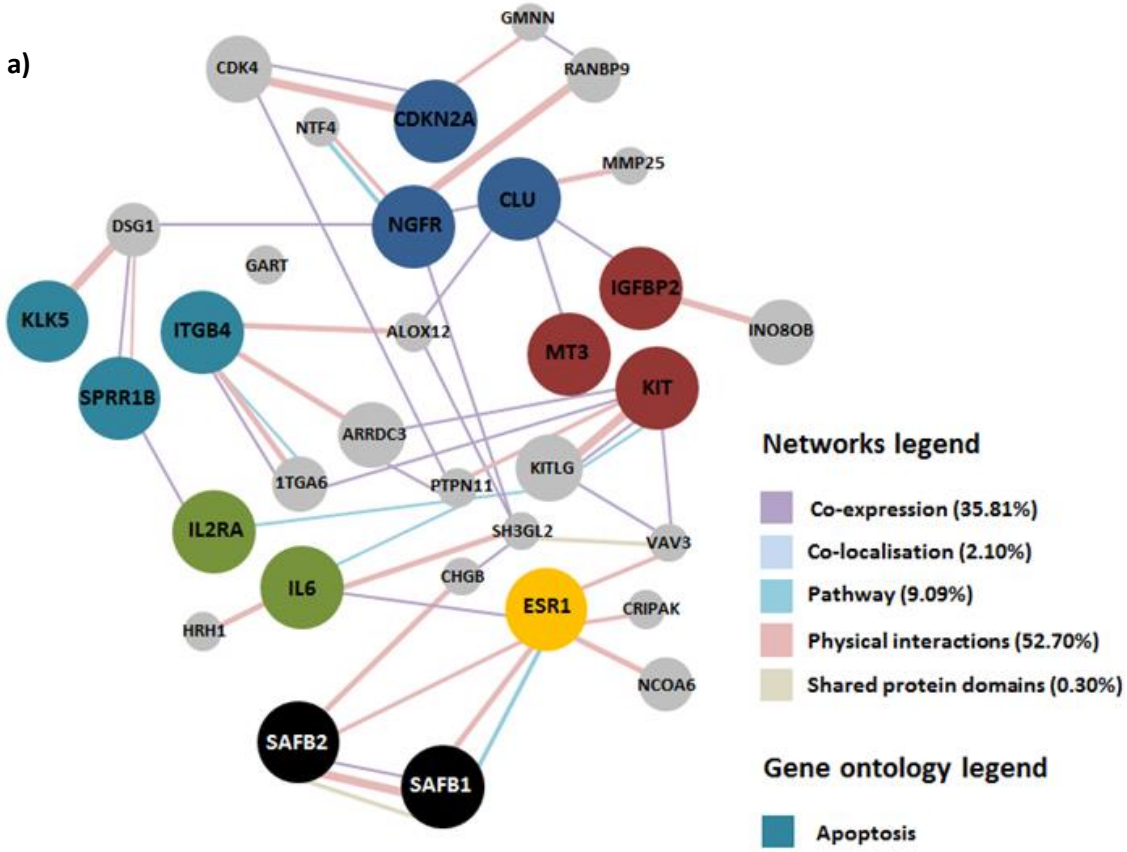
Gene ontology analysis for SAFB1 and SAFB2 regulated genes revealed important key players in apoptosis, cell proliferation, immune response, cellular adhesion and transcriptional regulation. The gene ontology annotations were classified using UniProt-GOA database (ebi.ac.uk/GOA).

The relationship between these genes was examined in a bioinformatics analysis using GeneMANIA prediction server (Warde-Farley *et al.* 2010). GeneMANIA utilises a large collection of genomics and proteomics interaction network to make predictions about gene function, analyse gene lists and reveal additional genes that share the same function. GeneMANIA makes predictions based on the scoring of each gene on its interaction network and assigns a weight to each network based on how well connected these genes are. It also connects the gene list to a predicted top 20 most similar genes.

When examining the interaction network between the 12 candidate genes together with *SAFB1* and *SAFB2*, it is clear that these genes share many physical interactions either directly or indirectly with each other, encompassing 52.70% of the network [Figure 4.5 (a), pink network]. It is interesting to note that the network weight assigned

for the physical interaction between *SAFB1* and *SAFB2* exceeds the maximum weight of 1.0, indicating a strong connection between these two paralogs [Figure 4.5 (b)]. This could be a possible explanation for the highest number of altered genes observed in the absence of both SAFB1 and SAFB2 [Figure 4.3 (b)]. Network interactions between *ESR1* with *SAFB1* and *SAFB2* were also observed in this analysis, consistent with previous studies reporting their role as ER-α corepressors (Oesterreich *et al.* 2000, Townson *et al.* 2003, Townson *et al.* 2004). Another 35.81% of the network defines the co-expression link between the 12 candidate genes and their top 20 most similar genes [Figure 4.5 (a), purple network]. The assigned network weights and connections between specific genes are summarised in the table below [Figure 4.5 (b)].

a)



b)

Gene 1	Gene 2	Weight	Network group
SAFB1	SAFB2	1.080	Physical interactions
ESR1	SAFB1	0.579	Physical interactions
ESR1	SAFB2	0.183	Physical interactions
SAFB1	SAFB2	0.166	Co-expression
CLU	IGFBP2	0.027	Co-expression
CLU	NGFR	0.021	Co-expression
IL2RA	SPRR1B	0.018	Co-expression
KIT	SPRR1B	0.013	Co-expression
IL6	ESR1	0.007	Co-expression
NGFR	MT3	0.007	Co-expression
ESR1	SAFB1	0.176	Pathway
SAFB1	SAFB2	0.020	Shared protein domains
CLU	MT3	0.020	Co-localization

Figure 4.5 Analysis of the interaction network between the 12 candidate genes of SAFB1 and SAFB2.

(a) Graphical view of the gene interaction network generated by GeneMANIA when the names of 14 genes (*CDKN2A*, *CLU*, *ESR1*, *IGFBP2*, *IL2RA*, *ITGB4*, *KIT*, *KLK5*, *MT3*, *NGFR*, *SPRRB1*, *IL-6* plus *SAFB1* and *SAFB2*) were input as the query list (http://www.genemania.org). The black nodes correspond to genes in the query list and the grey nodes indicate the extended gene list consisting of the top 20 neighbour gene predictions. The background colours for the 12 candidate genes correspond to previously described gene ontology annotations. The combined network is constructed from co-expression, co-localisation, pathway, physical interactions and shared protein domains. **(b)** A table summary of the assigned network weights for each connections between specific genes from the query list.

4.3.2 Validation of transcriptional target genes from gene expression profile study

4.3.2.1 *Clusterin (CLU)*

Considering the critical role of *CLU* in tumourigenesis and its rate of fold change in the three experimental groups (SAFB1 siRNA, SAFB2 siRNA or SAFB1 + SAFB2 siRNA), *CLU* was selected as the first candidate for further validation. A specific TaqMan assay targeting *CLU* was used for qRT-PCR. Surprisingly, this different experimental approach on the same sample set was unable to confirm the induced *CLU* expression previously observed; instead data revealed that *CLU* was suppressed in the absence of SAFB1 or/and SAFB2 (Figure 4.6). The possible explanation for this observation will be discussed later in Section 4.4.

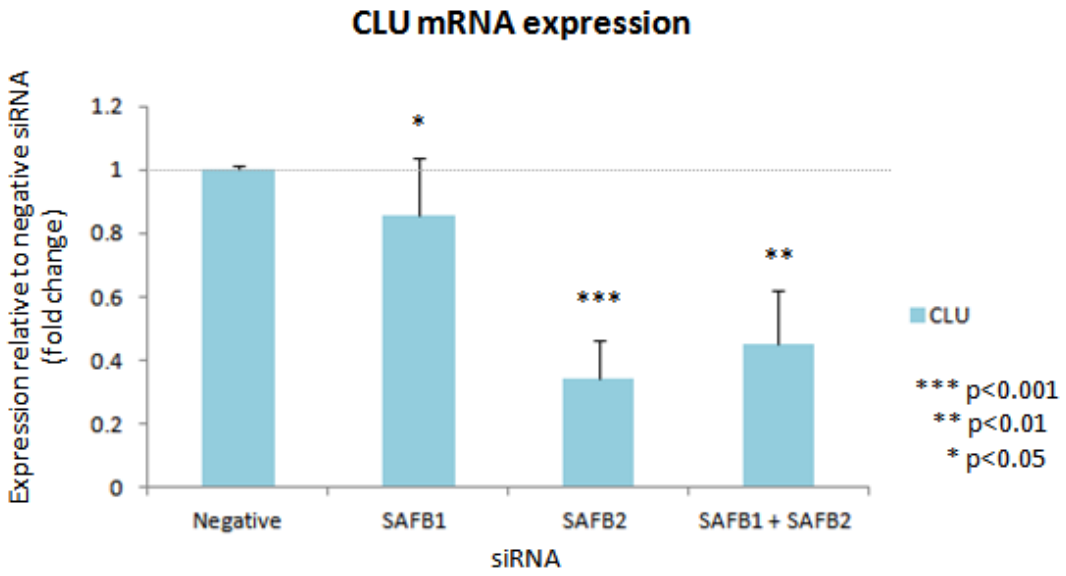


Figure 4.6 Expression of *CLU* in the absence of SAFB1 or/and SAFB2.

qRT-PCR was performed on mRNA from MDA-MB-231 cells transfected with negative, SAFB1, SAFB2 or SAFB1 and SAFB2 siRNA using validated TaqMan probes specifically targeting *CLU*. In contrast to previously observed data, *CLU* expression significantly decreased in the absence of SAFB1 or/and SAFB2. Data represents the average of three biological replicates ± S.D. Statistical significance of mRNA expression was calculated using a student's t-test. $^{***}p<0.001$, $^{**}p<0.01$, $^{*}p<0.05$.

4.3.2.2 *Integrin beta 4 (ITGB4)*

The prospect of SAFB proteins affecting the transcription of *ITGB4* attracted an interest due to its association with aggressive behaviour of breast tumours (Lu *et al.* 2008). Therefore, *ITGB4* was selected as the next candidate for validation study. A specific TaqMan assay targeting *ITGB4* was used for qRT-PCR. Consistent with gene expression array data, the loss of SAFB1 did not have a significant effect on *ITGB4* mRNA expression but the loss of SAFB2 and both SAFB proteins significantly increased *ITGB4* expression (Figure 4.7).

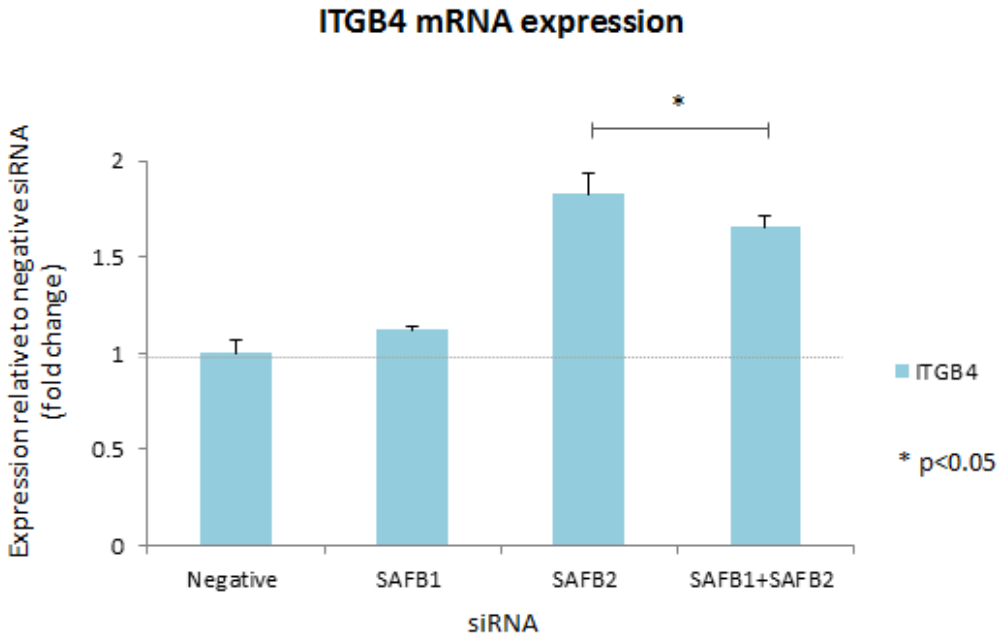


Figure 4.7 Expression of *ITGB4* in the absence of SAFB1 or/and SAFB2.

qRT-PCR was performed on mRNA from MDA-MB-231 cells transfected with negative, SAFB1, SAFB2 or SAFB1 and SAFB2 siRNA using validated TaqMan probes specifically targeting *ITGB4*. Data represents the average of three biological replicates ± S.D. Statistical significance of mRNA expression was calculated using a student's t-test. * = $p<0.05$.

After successful validation of *ITGB4* mRNA expression in MDA-MB-231 cells, an initial qRT-PCR experiment was performed using breast tissue samples. Primary breast tumours from two patients with ER negative, infiltrating ductal carcinoma was obtained according to the procedure described in Section 2.2. Total RNA was extracted from the non-involved 'normal' tissue and 'tumour' tissue from each patient (Section 2.7.1) and qRT-PCR performed to examine *ITGB4* expression (Section 2.7.2 and Section 2.7.3).

Data shows that *ITGB4* mRNA expression significantly increased in tumour tissue compared to their normal counterpart [Figure 4.8 (a)]. This prompted a second qRT-PCR experiment to investigate *SAFB1* and *SAFB2* mRNA levels in relation to the observed *ITGB4* mRNA expression. The expression of *SAFB1* and *SAFB2* in tumour 1 did not differ significantly compared to its normal counterpart but significantly increased in tumour 2 [Figure 4.8 (b)]. Comparison between *SAFB1* and *SAFB2* with *ITGB4* mRNA levels shows that lower levels of *SAFB1* and *SAFB2* correlates with higher *ITGB4* expression, whilst increased levels of *SAFB1* and *SAFB2* correlates with lower levels of *ITGB4* mRNA.

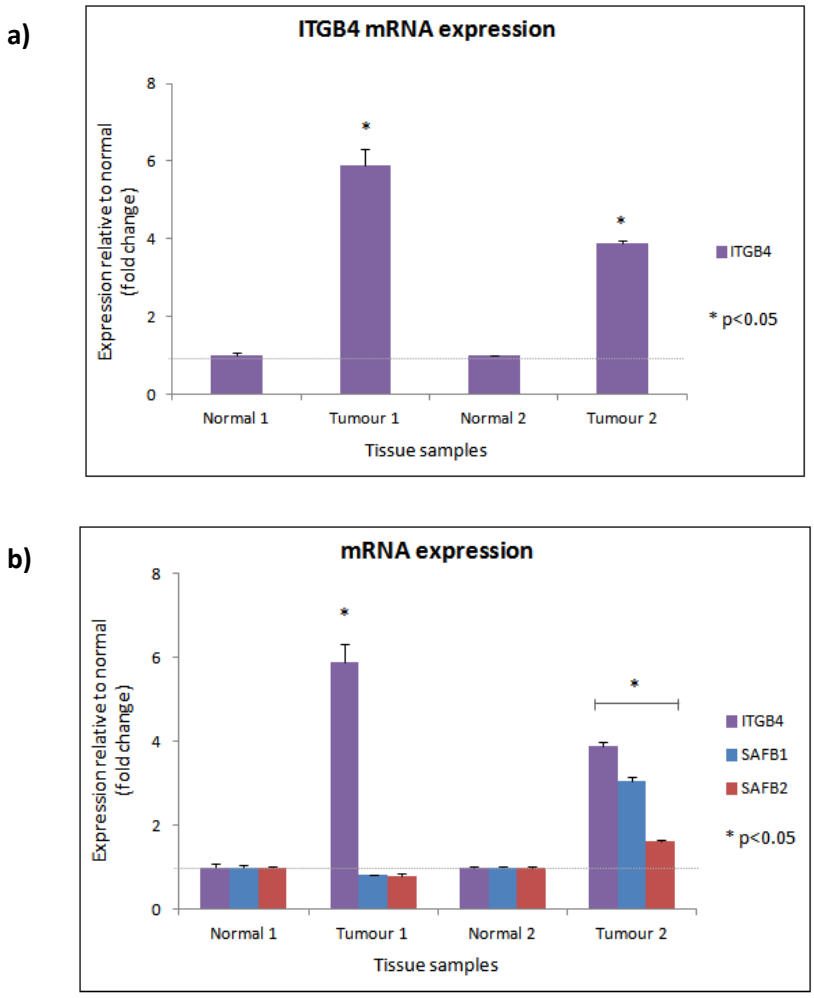


Figure 4.8 Expression of *ITGB4*, *SAFB1* and *SAFB2* in breast tissue samples.

qRT-PCR was performed on mRNA from non-involved 'normal' and 'tumour' tissue samples validated TaqMan assays specifically targeting **(a)** *ITGB4*, **(b)** *SAFB1* and *SAFB2*. Data represents the average of three experimental replicate ± S.D. Statistical significance of mRNA expression was calculated using a student's t-test. * = $p<0.05$.

4.3.2.3 *Interleukin 6 (IL-6)*

The effect of SAFB1 and SAFB2 on immune regulatory genes in MDA-MB-231 cells was also of interest due to the ability of immune evasion as an emerging hallmark of cancer and the association between SAFB1 and immune regulatory genes in MCF-7 cells (Hammerich-Hille *et al.* 2010, Hanahan *et al.* 2011). *IL-6* was the only unique candidate whose gene expression decreased in the absence of SAFB1 or/and SAFB2 and therefore was selected as another candidate gene for further validation. Specific validated TaqMan assay targeting *IL-6* was used in qRT-PCR to validate the previous observation generated from the gene expression array (Section 2.7.3). Data confirms the significant repression of *IL-6* in the absence of SAFB1 or/and SAFB2 (Figure 4.9). Loss of SAFB1 or/and SAFB2 that led to the repression of *IL-6* suggest a potentially novel role for SAFB1 and SAFB2 in transcriptional activation that has yet to be reported.

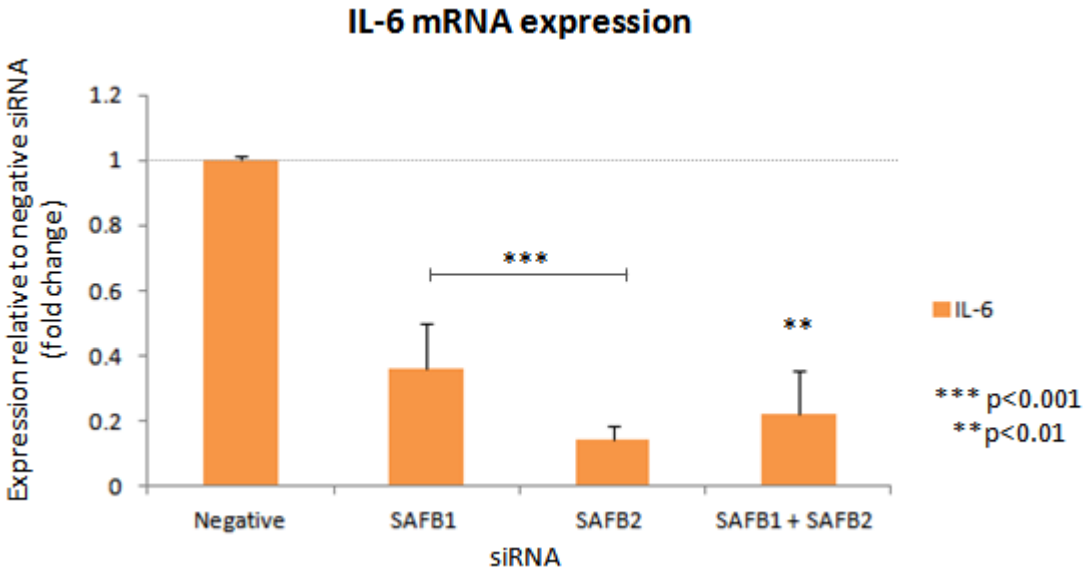


Figure 4.9 Expression of *IL-6* in the absence of SAFB1 or/and SAFB2.

qRT-PCR was performed on mRNA from MDA-MB-231 cells transfected with negative, SAFB1, SAFB2 or SAFB1 and SAFB2 siRNA using validated TaqMan probes specifically targeting *IL-6*. Consistent with previously observed data, *IL-6* expression significantly decreased in the absence of SAFB1 or/and SAFB2. Data represents the average of three biological replicates ± S.D. Statistical significance of mRNA expression was calculated using a student's t-test. $_{***}$=p<0.001, $_{**}$ = p<0.01.

The next step in this study was to examine whether loss of SAFB1 and SAFB2 affected expression of IL-6 at the protein level. The production and secretion of IL-6 was assessed by ELISA using culture medium from MDA-MB-231 cells that were transiently transfected with SAFB1 or/and SAFB2 siRNA (Section 2.8.2). In agreement with the mRNA data, untreated and negative siRNA transfected cells express high levels of IL-6 while SAFB1 or SAFB2 siRNA transfected cells express significantly lower amount of IL-6 (Figure 4.10). Cells that express decreased levels of both SAFB1 and SAFB2 produced the lowest concentration of IL-6, suggesting that a correlation between these two paralogs may be important in the regulation of *IL-6* gene expression which leads to the reduction in IL-6 secretion.

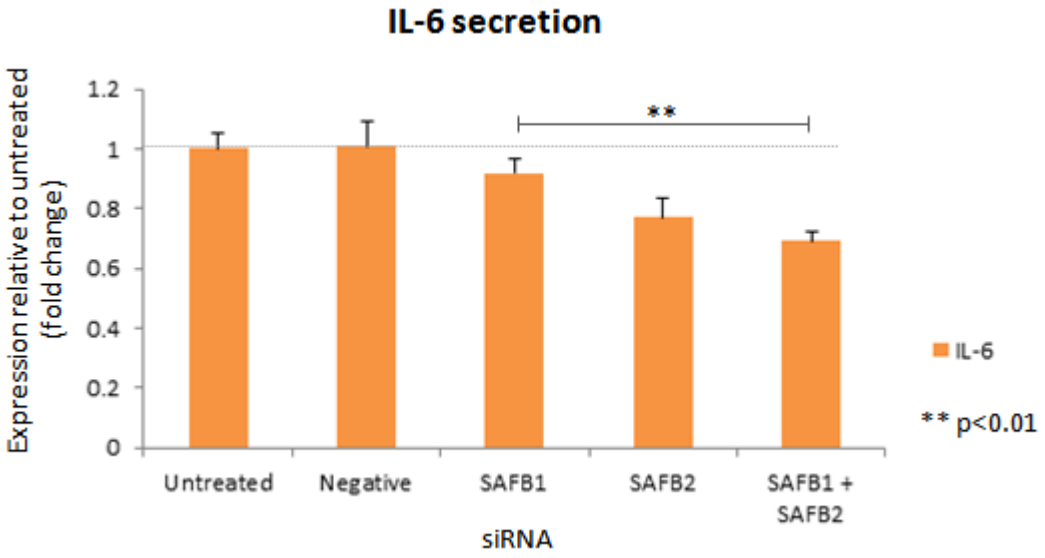


Figure 4.10 Secretion of IL-6 in the absence of SAFB1 or/and SAFB2.

MDA-MB-231 cells were transfected with negative, SAFB1, SAFB2 or SAFB1 and SAFB2 siRNA to assess their effect on IL-6 production. Untreated cells were included as control. Aliquots of supernatant from the culture medium were removed 72 hours after transfection and IL-6 concentration was determined by ELISA. The values obtained were compared to that of the untreated cells and represented here as a relative fold change. IL-6 production was significantly repressed in the absence of SAFB1 or/and SAFB2. Data represents the average of three biological replicates ± S.D. Statistical significance was calculated using a student's t-test. ** = $p<0.01$.

4.4 Discussion

The association between SAFB proteins and ER-α nuclear receptor in transcriptional regulation has gathered much attention in the recent years, particularly due to their relevance to breast cancer. Since being identified as ER-α corepressors (Oesterreich *et al.* 2000, Townson *et al.* 2003), substantial evidence has confirmed the direct correlation of SAFB1 and SAFB2 in ER-α mediated transcriptional repression, specifically in an ER positive breast cancer cell line, MCF-7 (Hammerich-Hille *et al.* 2010). In consideration of previously reported results, this part of the study was performed to determine the role of SAFB1 and SAFB2 in a breast cancer cell line that displays characteristics of the triple negative subtype.

The MDA-MB-231 cell line was utilised to investigate the effects of SAFB1 and SAFB2 in the regulation of ER-related genes. Although generally referred to as an ER and PR negative cell line that should not mediate a hormonal response, increasing evidence including the work reported in this thesis has shown the contrary to be true [discussed in Chapter 3 and (Cork *et al.* 2012)]. Twelve novel target genes for SAFB1 and SAFB2 were identified using RNAi and a gene expression array analysis approach on a focused panel of ER signalling candidates. Consistent with the work of Hammerich-Hille *et al.* in MCF-7 cells, the primary role of SAFB1 and SAFB2 as transcriptional repressors was also confirmed in MDA-MB-231 cells (Hammerich-Hille *et al.* 2010). Contrary to the observation by Hammerich-Hille *et al.* in MCF-7 cells, SAFB2 regulates more target genes than SAFB1 in MDA-MB-231 cells [Figure 4.3(b)]. The findings from this study confirm previous speculations regarding the distinct molecular role between SAFB1 and SAFB2 (Hammerich-Hille *et al.* 2010, Ivanova *et al.* 2005, Sergeant *et al.* 2007).

Another important observation from this study revealed an importance of the physical interaction between SAFB1 and SAFB2 in oestrogen-mediated repression of ER-target genes. Comparing the list of regulated genes and overlapping them according to their experimental groups (SAFB1 siRNA, SAFB2 siRNA or SAFB1 + SAFB2 siRNA) showed that the absence of both SAFB1 and SAFB2 affects the expression of the most number of genes [Figure 4.3(b)]. It has been established since the discovery of SAFB2 that both these SAFB proteins physically interact with each other, as observed in co-immunoprecipitation experiments (Townson *et al.* 2003). Bioinformatics data from Section 4.3.1.3 also supports a strong physical interaction between SAFB1 and SAFB2

(Figure 4.5). Another study has also revealed that SAFB1 and SAFB2 interact and colocalise with ER-α in the presence of 17β-oestradiol and work together to cause a synergistic reduction of ER-α mobility and co-operatively inhibit ER-α mediated transcription (Hashimoto *et al.* 2012). Taken together with these reports, the results from this study suggest a co-operative SAFB1 and SAFB2 effect on the repression of ER-target genes in MDA-MB-231 breast cancer cells.

The data from this study also revealed SAFB1 or/and SAFB2 mediated regulation of apoptotic, cell proliferation, immune response, cellular adhesion and transcriptional regulation genes in the triple negative breast cancer cells (Figure 4.4). A number of these genes are closely linked to the development of breast cancer.

CDKN2A

CDKN2A or more commonly known as the tumour suppressor p16, is a cyclin-dependent kinase inhibitor that negatively regulates cell cycle and induces apoptosis in tumour cells (Liggett *et al.* 1998, Shapiro *et al.* 1996). The inactivation of *CDKN2A* through aberrant DNA methylation appears to be a common event in many human cancers, including breast cancer (Cairns *et al.* 1995, Herman *et al.* 1995, Merlo *et al.* 1995, Nobori *et al.* 1994). Knockout mouse model has shown that *CDKN2A* deficiency is implicated with spontaneous tumour development, rapid cell proliferation and high colony-formation efficiency; hence directly confirming its role as a tumour suppressor (Serrano *et al.* 1996). Interestingly, despite its correlation to anti-tumour and anti-proliferative effects, a comprehensive study of *CDKN2A* mRNA expression in a cohort of primary breast tumour samples revealed an inverse relationship between *CDKN2A* expression and ER status; speculating that overexpression of *CDKN2A* may be a marker of poor prognosis in breast cancer patients (Hui *et al.* 2000). Coincidentally, the loss of SAFB proteins in invasive breast tumours was also associated with worse patient survival (Hammerich-Hille *et al.* 2009, Oesterreich S 2002). The evidence from this study revealed that the loss of both SAFB1 and SAFB2 led to a 3.0 fold increase in *CDKN2A* (Table 4.1). Taken together with evidence from other reports, SAFB1 and SAFB2 may be important players in the regulation of *CDKN2A* expression that could contribute to worse outcome in ER negative breast cancer patients.

IGFBP-2

Another candidate that has been extensively studied in relation to breast cancer is IGFBP-2, an insulin-like growth factor binding protein that modulates IGF action. IGFBP-2 is highly expressed in breast cancer compared to non-malignant tissue and levels of IGFBP-2 positively correlate with disease progression (Busund *et al.* 2005, So *et al.* 2008). In the ER positive breast cancer cell line, MCF-7, IGFBP-2 is positively regulated by 17β-oestradiol while the loss of IGFBP-2 in this cell line inhibits cell proliferation and enhances chemosensitivity (Clemmons *et al.* 1990, Foulstone *et al.* 2013, Juncker-Jensen *et al.* 2006, Yee *et al.* 1991). Several investigations have reported undetectable levels of IGFBP-2 expression in MDA-MB-231 cells, however; exogenous overexpression of IGFBP-2 in this ER negative cell line conferred growth advantage and chemoresistance (Clemmons *et al.* 1990, Dubois *et al.* 1995, Kim *et al.* 1991, So *et al.* 2008, Yee *et al.* 1991). Other evidence has implicated IGFBP-2 as a novel therapeutic target and useful marker to predict lymph node metastasis in invasive breast carcinoma patients (So *et al.* 2008, Wang *et al.* 2008). To date, there still remains a scarcity of knowledge regarding the molecular mechanisms that underlie the regulation of IGFBP-2 in breast cancer. Recent evidence has shown that IGFBP-2 expression is regulated by the PI3K/Akt/mTOR pathway through Sp1-induced transcription (Mireuta *et al.* 2010). Data from my study show that *IGFBP-2* transcription is induced in the absence of both SAFB1 and SAFB2 in MDA-MB-231 cells, previously shown to be negative for *IGFBP-2* expression, therefore suggesting that SAFB1 and SAFB2 may be involved in the repression of this gene (Table 4.1).

ESR1

Unexpectedly, *ESR1* was detected from the gene expression array as a potential transcriptional target for SAFB1 and SAFB2 (Table 4.1). The link between *ESR1* and breast cancer is unprecedented and has been extensively described in Section 1.3.1. However, the MDA-MB-231 cell line used in this study is known to be ER negative and therefore the detection of this gene provoked further analysis. The location of the SYBR green primers used in the gene expression array was interrogated using Primer-BLAST software (NCBI). This revealed its primer location on exon 1 that generates a 160bp PCR product starting at nucleotide position 254 of *ESR1* mRNA transcript (NCBI accession: NM_000125.3). Although ER-α variants were detected in this particular cell line, they usually lack the first coding exon and result in a truncated

ER-α protein [Section 3.3.5.1 and previous reports by (Flouriot *et al.* 2000, Wang *et al.* 2005)]. The ambiguity between these observations prompted a preliminary experiment to examine the specificity of the *ESR1* primers used in the gene expression array. Primers were designed to target the exact location as the ones used in the array and generated for use in conventional PCR. These primers were tested on RNA extracted from untreated MCF-7 and MDA-MB-231 cells. As shown in Figure 4.11, the *ESR1* primers were able to generate one specific PCR product in MCF-7 cells of the expected size (160bp) but multiple PCR products in MDA-MB-231 cells of around 100bp, 160bp and 500bp. This data suggests that the *ESR1* primers used in the SYBR green gene expression array may also generate non-specific PCR products in MDA-MB-231 cells which could possibly explain the observed fold change in *ESR1* (Table 4.1).

Computational analysis of these primers using the Primer-BLAST software (NCBI) revealed a long list of PCR products from potential target mRNA templates. When this list of target templates was filtered for mRNA templates that would generate PCR products between 400bp to 600bp, at least 14 target templates were identified (Appendix I), indicating the lack of specificity of the *ESR1* primers used in the gene expression array.

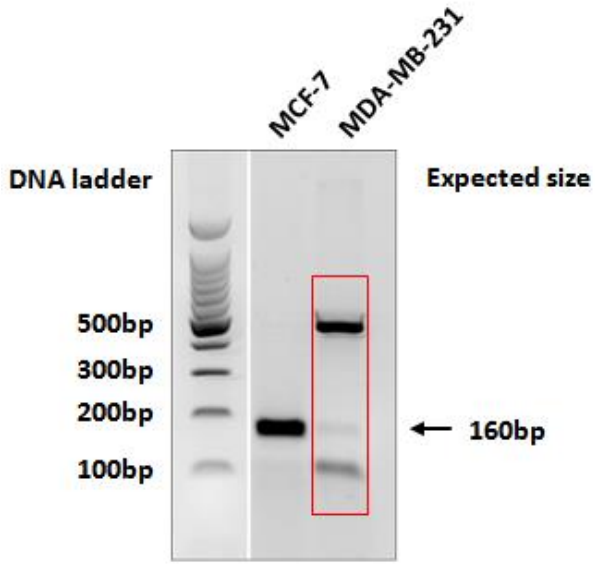


Figure 4.11 Analysis of ESR1 primers used in the gene expression array.

Conventional PCR was performed on mRNA from MCF-7 and MDA-MB-231 using the exact primers used in the SYBR green gene expression array (SABiosciences). PCR products were separated by electrophoresis on an agarose gel. Data represents preliminary experiment that was only performed once.

CLU

Gene expression array data from this study identified the increase of *CLU* in the absence of SAFB1 or/and SAFB2 (Table 4.1). This observation suggests that the repression of *CLU* is highly dependent on SAFB1 and SAFB2, and their function is independent of any synergistic action with each other. *CLU* is a single gene that expresses numerous mRNA transcripts as a result of alternative pre-mRNA splicing and several protein isoforms with different sub-cellular localisation and diverse biological functions. CLU protein isoforms have been associated with contradictory functions in apoptosis; the secreted form of CLU protein (sCLU) is anti-apoptotic and the nuclear form of CLU protein (nCLU) is pro-apoptotic [reviewed in (Shannan *et al.* 2006)]. sCLU is translated from the full-length *CLU* mRNA to produce a precursor protein that is directed to the endoplasmic reticulum, undergoes cleavage and extensive glycosylation before being transported to the Golgi for secretion (de Silva *et al.* 1990). In contrast, nCLU is synthesised from an alternatively spliced *CLU* mRNA transcript to form a precursor protein that does not undergo cleavage or glycosylation but localised to the cytoplasm in unstressed conditions (Leskov *et al.* 2003). In response to cytotoxic stress, such as ionising radiation or TGF-β stimulation, the precursor cytoplasmic CLU protein is post-translationally modified and translocated to the nucleus to induce apoptosis (Reddy *et al.* 1996, Yang *et al.* 2000).

CLU upregulation has been reported in various human cancers, especially in hormone-dependent malignancies such as prostate and breast cancer (Bettuzzi *et al.* 2002, Leskov *et al.* 2003, Miyake *et al.* 2004, Redondo *et al.* 2000). The precise relationship between *CLU* gene expression and programmed cell death has not been clearly elucidated, due to the recognition of different protein forms and their apparent opposing functions. Several studies have demonstrated that overexpression of nCLU acts as pro-death signal and inhibit cell survival, while the expression of sCLU exert cytoprotective properties (Leskov *et al.* 2003, Li *et al.* 2012, Niu *et al.* 2012, Wang *et al.* 2012, Zhang *et al.* 2012). The activation of tumour suppressor p53 in MCF-7 cells has been implicated in the suppression of sCLU secretion by repressing *CLU* promoter activity and transcription (Criswell *et al.* 2003). On the other hand, loss of functional p53 result in lost of nCLU function. Additionally, stable knockdown of CLU inhibits tumour cell invasion and metastasis in MDA-MB-231 breast cancer cells (Li *et al.* 2012). These evidences propose that tumour cell survival is associated with

overexpression of sCLU and loss of nCLU; and the function of CLU in tumour growth may be related to a pattern shift in its isoform production (Pucci *et al.* 2004).

In this study, *CLU* expression was induced in the loss of SAFB1, SAFB2 or SAFB1 and SAFB2 in the gene expression profile study using SYBR green detection method (Table 4.1). When using a highly specific TaqMan gene expression assay to validate this observation, results revealed that the expression of *CLU* was significantly decreased in the absence of SAFB1, SAFB2 or SAFB1 and SAFB2 (Figure 4.6). To explain these contradicting observations, the location of both SYBR green primers and TaqMan gene expression assay was analysed and mapped to the *CLU* gene to identify the regions of *CLU* mRNA transcripts detected by qRT-PCR (Figure 4.12). Analysis of *CLU* mRNA and its alternative isoforms revealed that the SYBR green primers were located within exon 10 of *CLU* gene while the TaqMan probe was located within exon 6, a region that is prone to multiple alternative splicing events (Figure 4.12). Although the only transcript currently known to encode a functional CLU protein is transcript variant 1 (NCBI accession: NM_001831.3), several other transcript variants have been detected in human tissues. The position of the SYBR green primers indicates that most, if not all, of the *CLU* mRNA transcript variants would be detected and amplified; while the TaqMan probe, though specific, is more restrictive and unable to detect mRNA transcripts that have a skipped or truncated exon 6 (Figure 4.12). The fact that the two qRT-PCR assays are targeting different regions within the *CLU* transcript provides a possible explanation for the conflicting results obtained using two different qRT-PCR approaches. Irrespective of this, loss of SAFB1 and SAFB2 did have a significant effect on the transcriptional regulation of *CLU*.

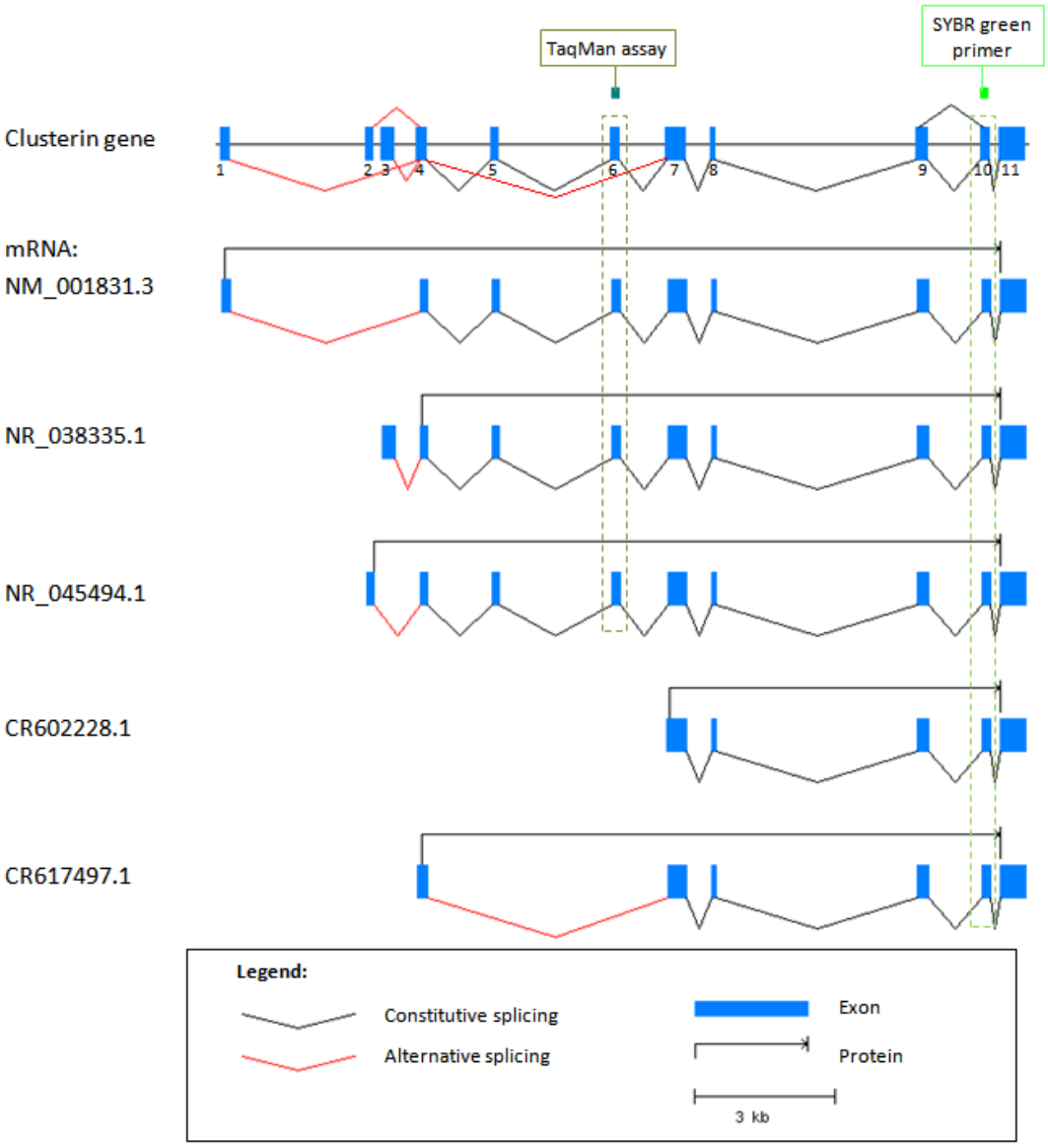


Figure 4.12 Analysis of CLU transcript variants and location of SYBR green primers and TaqMan gene expression assay.

The *CLU* gene was analysed in the Alternative Splicing Annotation Project (ASAP) database for alternative splicing information of this gene (Lee *et al.* 2003). Several transcript variants listed here have been reported in different human tissues. Mapping of the SYBR green primer and TaqMan gene expression assay used in this study revealed their position on exon 10 and exon 6 respectively. Exon skipping (CR617497.1) and alternate transcription start site (CR602228.1) result in transcript variants that lack exon 6, consequently undetected by TaqMan probes. However, exon 10 remains intact in all transcript variants and therefore could be detected by SYBR green primer.

ITGB4

The gene expression array data identified the increase of *ITGB4* mRNA expression particularly in the absence of SAFB2 (Table 4.1). This suggests that the regulation of *ITGB4* transcription may be dependent on the corepressor role of SAFB2. The *ITGB4* gene encodes the beta4 (ITGβ4) subunit that associates with the alpha6 (α6) subunit to form the α6/β4 integrin laminin receptor. ITGβ4 plays a pivotal role in pathways associated with cancer progression by facilitating the migration, invasion and survival of carcinoma cells (Lipscomb *et al.* 2005, Wilhelmsen *et al.* 2006). A comprehensive study using mouse models revealed that loss of ITGβ4 signalling suppresses tumour progression and metastasis, while ITGβ4 signalling promotes cell proliferation, invasive growth and inhibits apoptosis (Guo *et al.* 2006). ITGβ4 contributes to anchorage-independent growth and its gene expression significantly correlates with breast cancer size and grade, as well as basal-like breast cancer (Diaz *et al.* 2005, Lipscomb *et al.* 2005, Lu *et al.* 2008). Arrestin domain-containing 3 (*ARRDC3*), a gene inversely correlates with breast tumour progression, has been shown to function as a novel regulator of ITGβ4 internalisation, ubiquitination and ultimate degradation (Draheim *et al.* 2010). In ovarian cancer cells, *ITGB4* expression was upregulated in response to 17β-oestradiol (Parker *et al.* 2009). *ITGB4* transcription is also induced by the depletion of homeodomain-interacting protein kinase 2 (HIPK2), suggesting that HIPK2 corepresses its transcription in *ITGB4*-null cancer cells (Bon *et al.* 2009). Despite our current understanding on *ITGB4* and its significant contribution in tumour progression, the mechanism underlying *ITGB4* transcriptional regulation remains unclear.

In this study, successful validation by TaqMan gene expression assay has shown that the loss of SAFB2 and both SAFB proteins significantly increased *ITGB4* mRNA expression (Figure 4.7). Despite the very small sample size (n=2), comparison between *SAFB1* and *SAFB2* with *ITGB4* mRNA levels in normal and breast tumour samples reflects the data observed *in vitro* (Figure 4.8). Although this analysis was performed in only two tumour samples, further bioinformatics analysis using the newly available multidimensional cBioPortal cancer genomics database (www.cBioPortal.org) revealed that amplified *ITGB4* expression inversely correlate with *SAFB* expression in at least 69 breast cancer case studies (data from cBioPortal). These observations collectively show that the loss of SAFB proteins, particularly SAFB2, result in the loss of *ITGB4* repression.

IL-6

In this study, *IL-6* was the only candidate that was significantly suppressed in the absence of SAFB1 and SAFB2 (Table 4.1). This interesting finding highlights a novel role of SAFB1 and SAFB2 in transcriptional activation. IL-6 is a pleiotropic inflammatory cytokine that acts in a paracrine or autocrine fashion to alter the function of its target cells. *In vitro* studies on breast cancer cells revealed controversial effects of IL-6 that indicate its tumour promoting and tumour inhibitory role in breast cancer [reviewed in (Knupfer *et al.* 2007)]. It has been established that MCF-7 cells are incapable of expressing IL-6 due to the presence of an active gene repression mechanism on the IL-6 gene in this cell line (Faggioli *et al.* 1996). The mechanism of IL-6 gene silencing in MCF-7 cells involves local chromatin remodelling causing the surrounding chromatin to be in a repressed state (Armenante *et al.* 1999). In this non-expressing cell line, production of IL-6 could be induced in an autocrine fashion by exogenous IL-6 and its expression increases resistance to doxorubicin chemotherapy drug, indicating a protective effect of IL-6 against drug-induced cell death (Conze *et al.* 2001). Interestingly, in the triple negative MDA-MB-231 cells a positive feedback mechanism in IL-6 secretion has been proposed, where increased cytokine concentrations cause more cell aggregation and proliferation that further stimulated cytokine secretion (Geng *et al.* 2013).

Several studies have also shown that IL-6 can promote breast cancer cell motility, contributing to tumour metastasis. IL-6 induces significant cell migration and spreading in ER negative breast cancer cells, MDA-MB-231 and SK-BR-3 (Arihiro *et al.* 2000, Verhasselt *et al.* 1992). The effects of IL-6 were also investigated in four breast cancer cell lines and results showed that IL-6 decreased cell adhesion in three ER positive breast cancer cell lines, however did not further affect the already poorly adherent cell line MDA-MB-231 (Asgeirsson *et al.* 1998). Although IL-6 on its own did not show significant effect on the proliferation of MCF-7 cells, the simultaneous stimulation of IL-6 and oestrone sulphate significantly increased cell proliferation (Honma *et al.* 2002). On the other hand, Chiu *et al.* reported that ER negative breast cancer cell lines secrete high levels of biologically active IL-6 that could inhibit proliferation of IL-6 sensitive breast cancer cells through paracrine signalling induced apoptosis, but had no effect on autocrine signalling (Chiu *et al.* 1996). Growth inhibitory effects of IL-6 were also observed in other cancers, especially in prostate cancer (Wang *et al.* 2004). IL-6 has also been demonstrated to contribute in cell cycle arrest by inhibiting IGF-induced

DNA synthesis (Shen *et al.* 2002). Taken together, the *in vitro* data regarding IL-6 in breast cancer cells are not uniformly consistent and highlights the pleiotropic nature of this cytokine.

Regulation of IL-6 expression in breast cancer cell lines has been examined and reported. An early study presented evidence that 17β-oestradiol inhibits the expression of IL-6 via an indirect interaction with its receptor, ER-α (Pottratz *et al.* 1994). Another later study substantiates this observation when wild type ER-α was shown to repress IL-6 expression through its trans-acting process (Bhat-Nakshatri *et al.* 2004). Post-transcriptional regulation also plays a vital role in *IL-6* expression by modulating its mRNA stability [reviewed in (Palanisamy *et al.* 2012)]. *IL-6* transcripts have abundant adenine- and uracil-rich (AU-rich) elements (AREs) in the 3'-untranslated region (UTR) that promote its mRNA degradation (Hao *et al.* 2009). RNA-binding proteins such as tristetraprolin (TTP) and AT-rich interactive domain containing protein 5a (Arid5a) have been shown to regulate *IL-6* mRNA stability through their interaction with AREs (Masuda *et al.* 2013, Van Tubergen *et al.* 2011).

Data presented in this thesis study reveals a novel correlation between loss of SAFB1 and SAFB2 and regulation of IL-6 expression and secretion in MDA-MB-231 cells (Figure 4.9 and Figure 4.10). Transient transfection experiments using human *IL-6* promoter reporter constructs showed that the loss of SAFB1 and SAFB2 had no direct effect on *IL-6* promoter activity (data not shown). Considering the presence of AREs in *IL-6* and the speculative role of SAFB as RNA-binding proteins, it is plausible that SAFB2 may affect *IL-6* mRNA stability rather than its transcription. This mechanism of action could potentially be investigated in future work. Irrespective of the mechanism involved, this study is the first associate SAFB1 and SAFB2 proteins in the regulation of IL-6 expression.

In summary, this study has provided further evidence for the role of SAFB1 and SAFB2 as transcriptional repressors. While SAFB1 appeared to be the main transcriptional regulator in MCF-7 cells (Hammerich-Hille *et al.* 2010), this work shows that SAFB2 has a more prominent role in transcriptional repression in MDA-MB-231 cells. This also highlights the difference in gene expression and transcriptional control that exists between ER positive and ER negative breast cancer cells. Evidence also suggests that SAFB1 and SAFB2 may function synergistically in mediating the repression of their

target genes. For the first time, this study has reported a link between SAFB proteins in the regulation of *ITGB4* and IL-6 expression.

Chapter 5 : Transcriptome-wide identification for RNA-binding sites of SAFB1 in breast cancer cells

5.1 Introduction

The role of SAFB1 in RNA processing, metabolism and splicing has been well described (Section 1.4.2.3). SAFB1 and SAFB2 proteins share a highly conserved RNA-recognition motif (RRM) with 98% similarity in the central region although their direct RNA-binding potential is still currently unknown. Sequence analysis revealed the alignment of SAFB RBD residues with the consensus sequence of RBD as defined by Birney *et al.* (Birney *et al.* 1993). SAFB1 was classified as a RNA-binding protein and a novel hnRNP protein due to its similarity to the highly conserved RBD found in the hnRNP protein family (Weighardt *et al.* 1999).

Subsequent studies have implicated both SAFB proteins in alternative splicing, as overexpression of SAFB1 and SAFB2 was shown to inhibit the splicing of a *TRA2B* variable exon (Sergeant *et al.* 2007, Stoilov *et al.* 2004). Further investigation using mutants lacking the RRM domain revealed that SAFB1's ability to inhibit *TRA2B* exon skipping was independent of its RNA-binding ability (Stoilov *et al.* 2004). This evidence suggests that SAFB1 may not bind directly to *TRA2B* pre-mRNA to regulate exon skipping but could possibly mediate an indirect effect through its interaction with various splicing factors (Arao *et al.* 2000, Li *et al.* 2003, Nikolakaki *et al.* 2001, Sergeant *et al.* 2007, Weighardt *et al.* 1999). In an unrelated study, *in vitro* evidence has shown that the RRM domain of SAFB1 was able to bind RNA isolated from MCF-7 cells, although the identity of the RNA targets was not described (Townson *et al.* 2004). Although implicated in alternative splicing, it is still largely unknown whether SAFB1 exerts its RNA processing functions through direct RNA interaction or by tethering to other protein factors. This is an unexplored avenue of research that beckons a collaborative effort to expand our knowledge of the function of SAFB proteins. Therefore, much interest has been generated for an in depth investigation to identify possible direct RNA targets of SAFB1, especially in the context of breast cancer.

CLIP combined with high-throughput sequencing is an emerging powerful tool to study protein-RNA interactions in cells or tissues. This method has been utilised to identify transcriptome-wide binding maps of several RNA-binding proteins (Licatalosi *et al.* 2008, Sanford *et al.* 2008, Ule *et al.* 2005, Ule *et al.* 2003, Yeo *et al.* 2009). Despite the

high specificity of the CLIP data, primer extension assays frequently truncate prematurely before the crosslink nucleotide, causing truncated cDNA to be lost during CLIP library preparation (Urlaub *et al.* 2002). Konig *et al.* have recently improved this method and developed iCLIP that captures the truncated cDNA and provides insights into the position of the crosslink site at single nucleotide resolution (Konig *et al.* 2011). This improved method has been applied to study RNA-binding proteins such as hnRNP C, SRSF3, SRSF4 and T-cell intracellular antigen 1 (TIA1) (Anko *et al.* 2012, Konig *et al.* 2010, Wang *et al.* 2010).

Irradiation of cells with UV light creates a covalent bond between proteins and the RNAs to which they are bound *in vivo*. This physical bond is used to isolate the protein-bound RNAs using protein immunoprecipitation followed by denaturing gel electrophoresis. After the isolation of RNAs that are crosslinked to the protein of interest, the proteins are digested by Proteinase K while a residue polypeptide remains attached at the RNA crosslink site, causing cDNA products generated to truncate at this site (Konig *et al.* 2011). cDNA is generated by reverse transcription using primers with random barcodes to enable discrimination between unique cDNA products and PCR duplicates. A second adapter is introduced via self-circularisation of the cDNA to prepare a quantitative cDNA library that allows for high-throughput sequencing to identify RNA targets from transcriptome-wide binding maps and precise mapping of protein-RNA interaction sites at single nucleotide resolution (Konig *et al.* 2010, Ule *et al.* 2003).

A brief report has shown that SAFB1 was able to bind RNA in MCF-7 cells (Townson *et al.* 2004), therefore this cell line was selected as a candidate for this part of the study. iCLIP was performed for SAFB1 (Section 2.9.1) followed by high-throughput sequencing and mapping (Section 2.9.2) to generate a transcriptome-wide binding map for SAFB1.

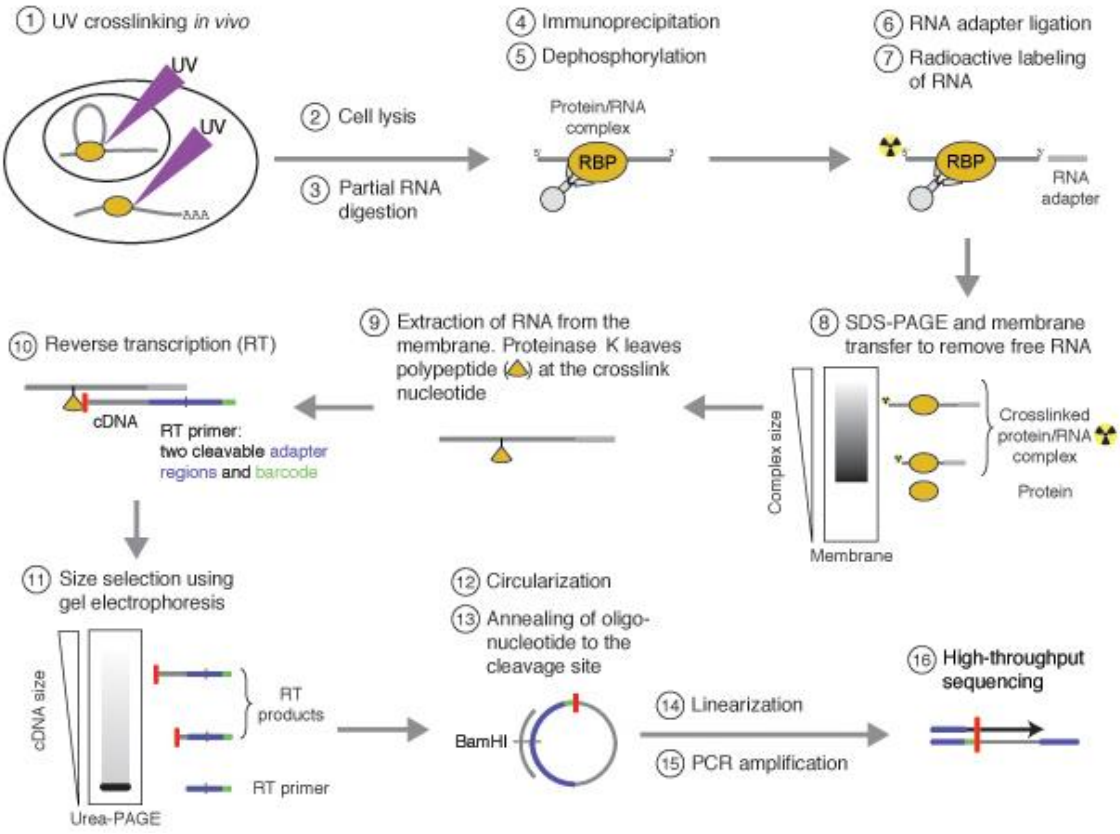


Figure 5.1 Schematic representation of the iCLIP protocol.

MCF-7 cells are UV-irradiated on ice to enable the formation of protein-RNA covalent bonds *in vivo*. Protein is purified by immunoprecipitation together with the bound RNA. An RNA adapter is ligated to the 3' end to allow sequence specific priming of reverse transcription, while the 5' end is radioactively labelled. The crosslinked protein-RNA complexes are purified from unbound RNA using SDS-PAGE and membrane transfer. The recovered RNA is subjected to protein digestion by Proteinase K, leaving a short polypeptide at the crosslink nucleotide. Reverse transcription incorporates two cleavable adapter regions and barcode sequences into the newly synthesised cDNA that truncates at the crosslink nucleotide. The cDNA undergoes size separation to remove primers prior to self-circularisation. Linearisation of cDNA at known restriction enzyme site produces suitable templates for PCR amplification. Finally, high-throughput sequencing generates reads that contain the barcode sequences followed by the last nucleotide of the cDNA, allowing the binding site to be deduced with high resolution [image taken from (Konig *et al.* 2011)].

5.2 Aims

The aims of this chapter were to:

1. Identify RNA-binding sites for SAFB1 protein in MCF-7 cells
2. Investigate the distribution of SAFB1 binding sites
3. Determine a potential consensus binding motif for SAFB1
4. Identify novel RNA targets from transcriptome-wide binding maps generated by iCLIP
5. Validate selected RNA targets in SAFB1 knockdown cells using qRT-PCR and conventional PCR

5.3 Results

5.3.1 Identification of SAFB1 binding sites in MCF-7 cells

5.3.1.1 Optimisation and sequencing of iCLIP library for SAFB1 in MCF-7 cells

An initial experiment was performed to examine the compatibility of the SAFB1 antibody with the magnetic beads used for protein-RNA immunoprecipitation. MCF-7 cells were irradiated with 150mJ/cm^2 of UV at 254nm and lysed (Section 2.9.1.1), prior to immunoprecipitation with protein A or protein G magnetic beads coupled with SAFB1 antibody (Section 2.9.1.3). Once the crosslinked RNAs have been radioactively-labelled, the samples were subjected to denaturing gel electrophoresis and membrane transfer (Section 2.9.1.5). The autoradiograph revealed that both SAFB1-coupled protein A and protein G magnetic beads were capable of precipitating the protein-RNA complex with different binding affinity (Figure 5.2). Protein G magnetic beads appear to have higher affinity binding to protein-RNA complex; however RNA digestion with high RNase concentration was unable to have any effect on the radioactive signal on the autoradiograph [Figure 5.2 (a), lane 3]. This observation suggests that the protein G samples may potentially contain high levels of background noise. Comparison of SAFB1-coupled protein A magnetic beads revealed a difference in radioactive signals between samples with high and low RNase concentration [Figure 5.2 (a), lanes 1 and 2]. This preliminary experiment confirms the

suitability of protein A magnetic beads in the application of the iCLIP experiments, based on the low background noise seen in these samples.

To further validate the efficiency and specificity of the immunoprecipitation step, the membrane was analysed for SAFB1 protein by a conventional immunoblotting method (Section 2.6.2). The SAFB1 antibody successfully detected SAFB1 protein in all immunoprecipitants using either protein A or protein G magnetic beads [Figure 5.2 (b)]. This observation validates the accuracy for SAFB1 immunoprecipitation in this crucial step of the iCLIP experiment.

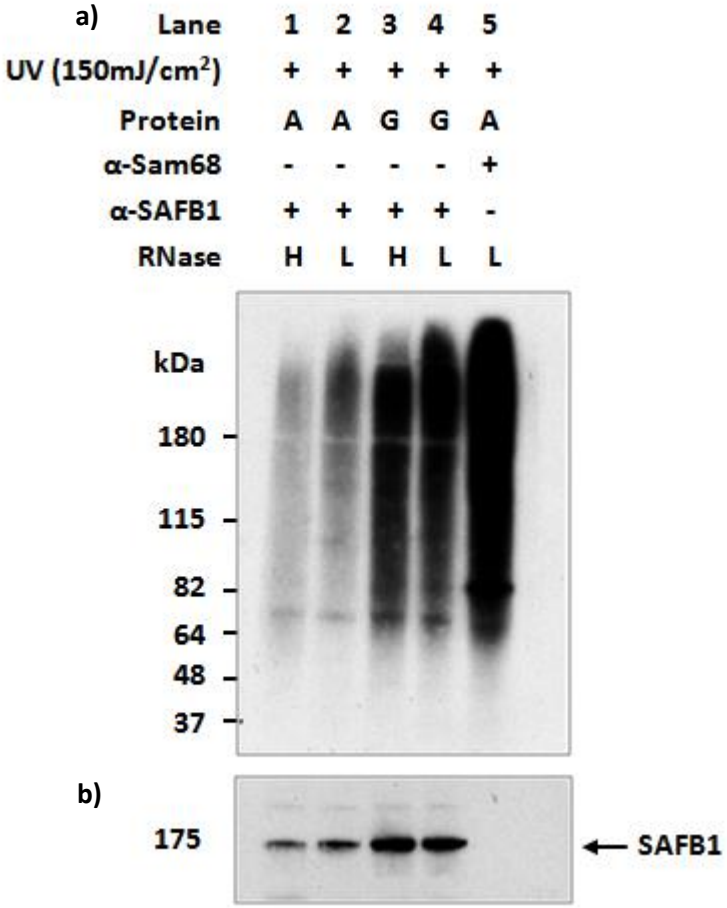


Figure 5.2 Optimisation of magnetic beads for the use of iCLIP in MCF-7 cells.

UV-irradiated MCF-7 cells were lysed and treated with high (H) or low (L) concentration of RNase to allow for partial RNA digestion. High RNase digestion is necessary as a control for antibody specificity, while low RNase digestion is needed for sequencing library preparation. Cell lysates were subjected to immunoprecipitation using SAFB1-coupled protein A or protein G magnetic beads, followed by radioactive labelling of the 5' end of the RNAs. Samples were separated by denaturing gel electrophoresis and transferred onto a membrane to

allow analysis using autoradiography and immunoblotting. **(a)** Autoradiograph revealed that high RNase dilution could reduce the radioactive signal in protein A immunoprecipitants (lane 1) but not in protein G immunoprecipitants (lane 3), when compared to their corresponding samples treated with low RNase dilution (lanes 2 and 4 respectively). Sam68-coupled protein A beads were included as a control for antibody specificity (lane 5). **(b)** The lower panel shows immunoblot analysis of protein extracts used as input for the immunoprecipitation. Antibody against SAFB1 reveals its protein was present in the corresponding immunoprecipitant samples (lanes 1-4), but not in the Sam68 samples (lane 5).

To identify potential RNA targets of SAFB1 *in vivo*, three biological replicates of iCLIP experiments were performed using the SAFB1 antibody on MCF-7 cell lysates. Prior to sequencing of the iCLIP library, the success of the experiments were monitored during the purification of protein-RNA complex (Section 2.9.1.5) and amplification of the cDNA (Section 2.9.1.10). Conditions were altered in several samples to provide a parallel comparison for experimental controls. In reactions when UV crosslinking or the use of SAFB1 antibody was omitted during sample preparation, purified protein-RNA complex was absent in the autoradiograph (Figure 5.3, lanes 3-6). This validates the presence of purified protein-RNA complex observed in the actual iCLIP replicates (Figure 5.3, lanes 1-2).

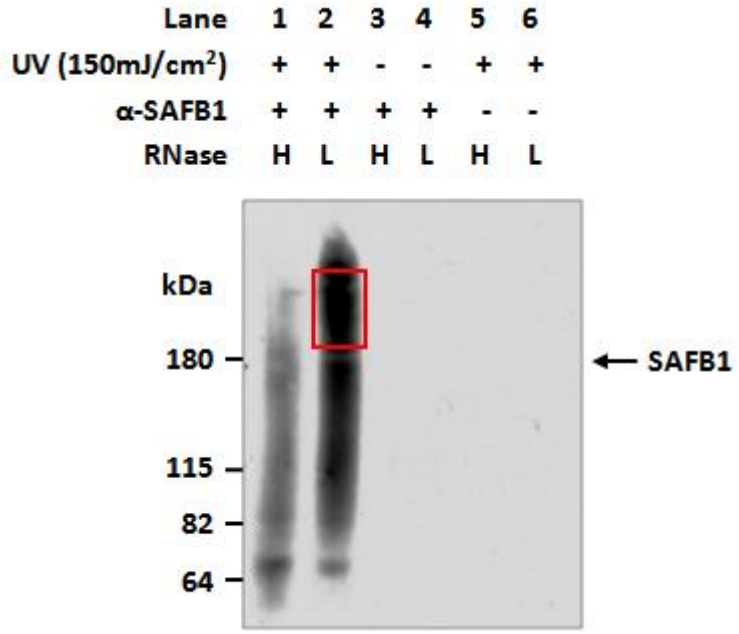


Figure 5.3 Analysis of crosslinked SAFB1-RNA complexes.

Cell extracts were prepared from UV crosslinked and control MCF-7 cells, then RNA was partially digested using high (H) or low (L) concentration of RNase. SAFB1-RNA complexes were immuno-purified from cell extracts using SAFB1 antibody (α-SAFB1) and the RNAs were ligated to RNA adapters at the 3' ends before radioactively labelled with ^{32}P-γ-ATP at the 5' ends. SAFB1-RNA complexes were separated using denaturing gel electrophoresis and transferred to a membrane before being exposed on an autoradiograph. The shift of SAFB1-RNA complexes upwards from the size of the protein was observed (lane 2), while this shift was less pronounced when high concentrations of RNase were used (lane 1). The radioactive signal disappears when cells were not UV crosslinked (lanes 3 and 4) or no antibody was used in the immunoprecipitation (lanes 5 and 6). The red box marks a region of the membrane that was cut out for subsequent purification steps. This image is a representative of the three iCLIP biological replicates.

RNA recovered from the membrane was purified and reverse transcribed. The cDNA generated was size-purified using denaturing gel electrophoresis into three size fractions [high (H): 120-200 nucleotides, medium (M): 85-120 nucleotides and low (L): 70-85 nucleotides]. The recovered cDNA was amplified by PCR using P5/P3 Solexa sequencing primers that introduces an additional 76 nucleotides to the cDNA. PCR products of different size distribution corresponding to the sizes of the input fractions were observed for all replicates (Figure 5.4, lanes 4-12). An experiment that lacked the SAFB1 antibody during immunoprecipitation was included as a control for

PCR specificity and no corresponding PCR products were detected in this sample (Figure 5.4, lanes 1-3). Following the analysis of PCR products by denaturing gel electrophoresis, the three iCLIP replicates were submitted for high-throughput sequencing using the Illumina Genome Analyser II system (Illumina Inc; San Diego, USA) (Section 2.9.2).

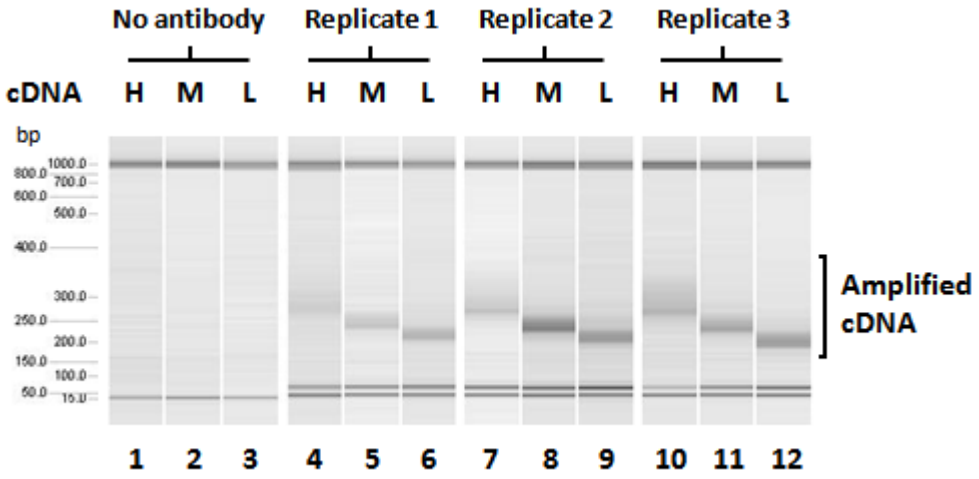


Figure 5.4 Analysis of PCR amplified iCLIP cDNA libraries.

RNA recovered from the membrane was reverse transcribed and size-purified into three size fractions [high (H): 120-200 nucleotides, medium (M): 85-120 nucleotides and low (L): 70-85 nucleotides]. Purified cDNA was recovered, self-circularised, relinearised and PCR amplified using sequencing primers that introduces an additional 76 nucleotides to the cDNA. Amplified cDNA of different size distribution were observed in all three replicates (lanes 4-12) while PCR products were absent when no antibody was used for immunoprecipitation (lanes 1-3). PCR products were submitted for high-throughput sequencing.

Bioinformatics analyses were performed on the sequencing results by Tomaz Curk (University of Ljubljana, Slovenia) and uploaded on the web-based iCount software (icount.fri.uni-lj.si/). High-throughput sequencing of all three biological replicates for SAFB1 iCLIP generated a total of 1,145,271 unique cDNA reads with single-hits mapping to the human genome. In order to reduce false positive hits and increase the resolution of the dataset, clusters of SAFB1 crosslink sites were identified by the iCount software according to a previous CLIP study (Yeo *et al.* 2009). This filtering approach removed almost 49% of all crosslink nucleotides and identified 587,119 significant

SAFB1 crosslink clusters (FDR<0.05). The number of sequences at each crosslink nucleotide were summarised into a 'cDNA count' to represent a quantitative measure of the amount of SAFB1 crosslinking at each position. The random barcode introduced into the iCLIP cDNAs allowed for the distribution of SAFB1 on human RNAs to be analysed in a quantitative and reproducible manner. An example of the view for SAFB1 crosslink sites on the UCSC Genome Browser (genome.ucsc.edu/) is shown below (Figure 5.5).

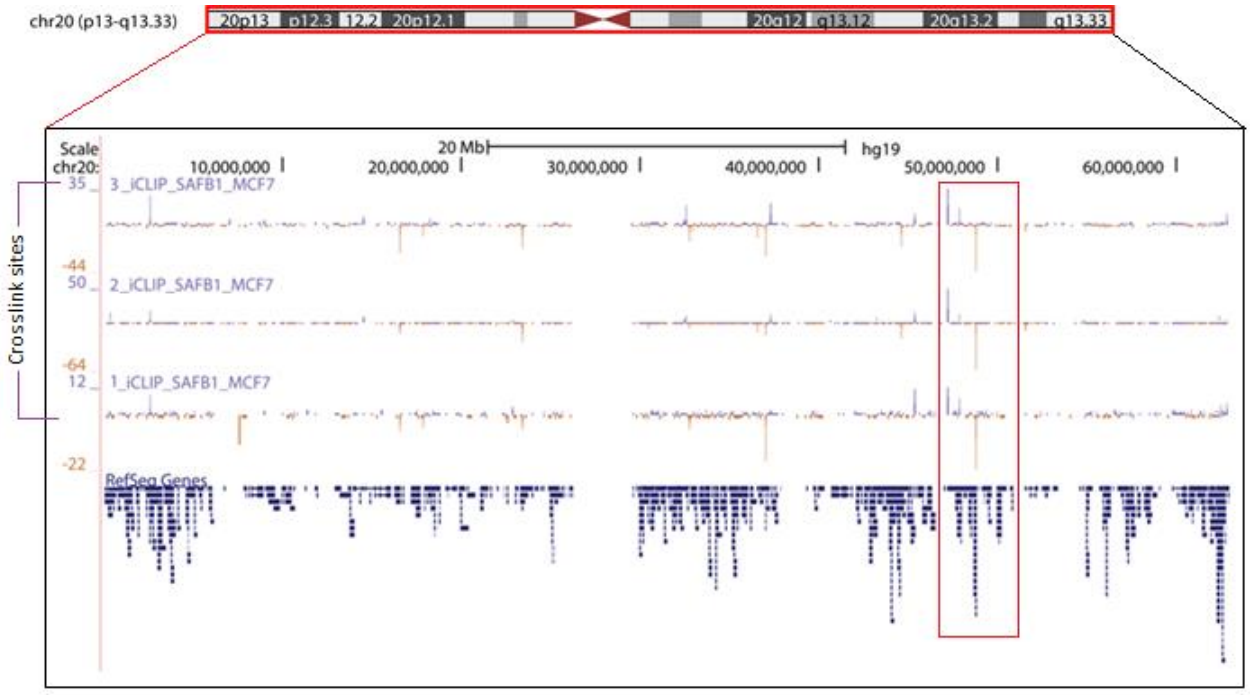


Figure 5.5 Global view of SAFB1 crosslink nucleotides on chromosome 20.

The three individual biological replicates of SAFB1 iCLIP (labelled as 'crosslink sites') are shown in BedGraph format in the UCSC hg19 Genome Browser. The peaks within each replicate represent cDNA counts at crosslink sites. Purple peaks denote crosslink sites on the sense strand while orange peaks represent crosslink sites on the antisense strand. The bottom track refers to known gene annotations based on the RefSeq database. Similar pattern of crosslink sites was observed between each replicate and an example is highlighted in the red box.

5.3.1.2 Mapping of SAFB1 crosslink sites to the transcriptome

iCLIP identified binding of SAFB1 across the whole transcriptome, where 100% of significant cDNA reads mapped to the sense orientation to annotated genes. This confirms the high strand specificity of iCLIP also observed in other studies

(Konig *et al.* 2010, Wang *et al.* 2010). Analysis of crosslinking frequency mapped to transcript regions revealed that SAFB1 binds to coding and non-coding RNAs (ncRNAs). The highest proportion of SAFB1 crosslink sites mapped to ncRNAs, followed by intergenic regions, open reading frames (ORF), introns and 3' or 5' untranslated regions (UTR) [Figure 5.6 (a)]. When the cDNA density for each transcript regions was analysed relative to the cDNA density in the whole genome, the highest density enrichment was detected in ncRNAs [Figure 5.6 (b)]. The distribution of SAFB1 crosslink sites within ncRNA subclasses was also analysed. SAFB1 crosslink sites were most abundant in small nuclear RNA (snRNA), mitochondrial RNA (Mt RNA) and small nucleolar RNA (snoRNA) (Figure 5.7).

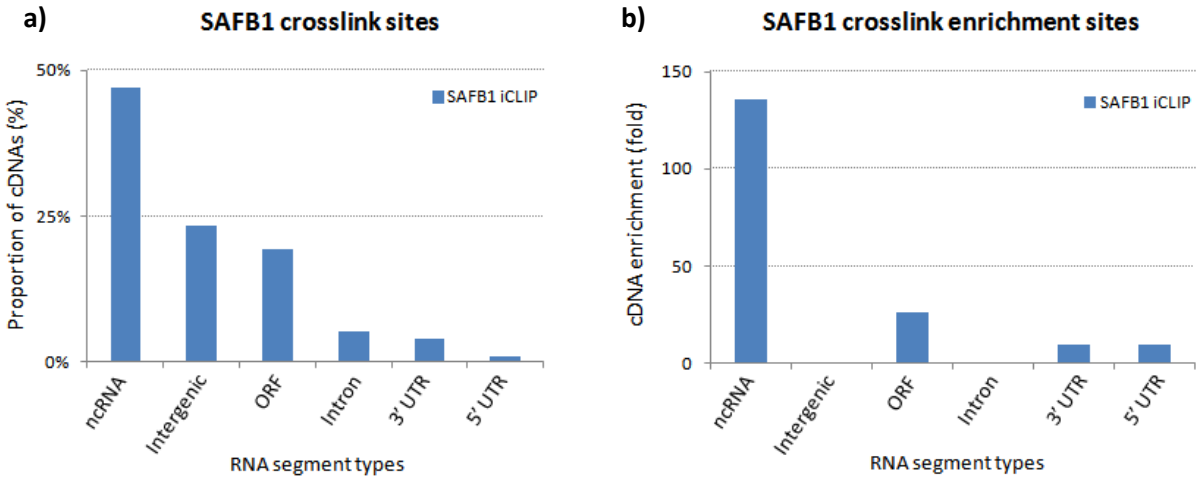


Figure 5.6 Distribution of significant SAFB1 crosslink sites within RNA segment types.

(a) The proportion of cDNAs mapped to different transcript regions relative to the total number of cDNA reads revealed that the highest percentage of cDNAs was mapped to ncRNA (47.08%), followed by intergenic regions (23.24%), ORFs (19.38%), introns (5.23%), 3' UTRs (3.83%) and 5' UTRs (0.86%). **(b)** The fold enrichment of cDNA density in different types of RNAs relative to cDNA density in the whole genome highest density enrichment in ncRNAs.

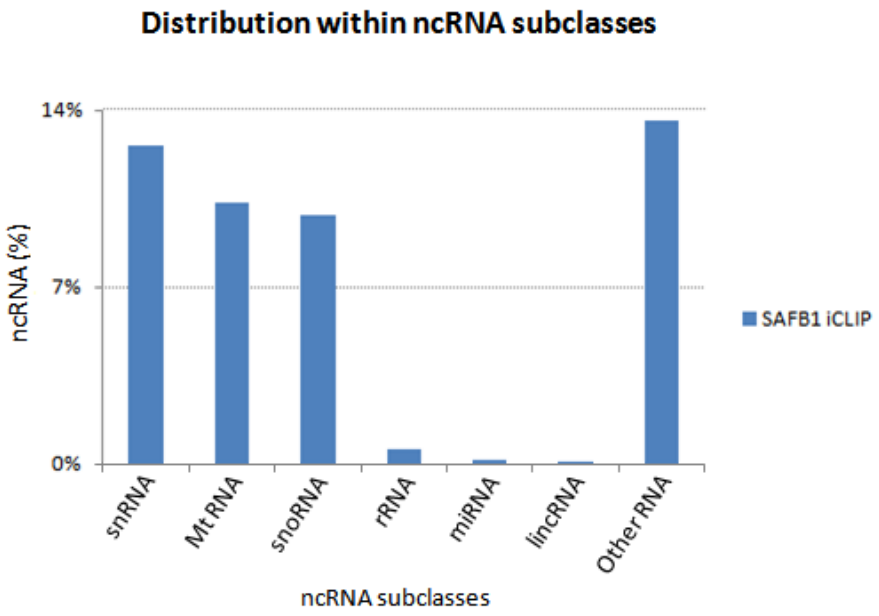


Figure 5.7 Distribution of significant SAFB1 crosslink sites within ncRNA subclasses.

The distribution of SAFB1 crosslink sites within different ncRNA subclasses revealed significant abundance in snRNA, Mt RNA and snoRNA. 'Other RNA' consists of pseudogenes and processed transcripts with no known ORF or function.

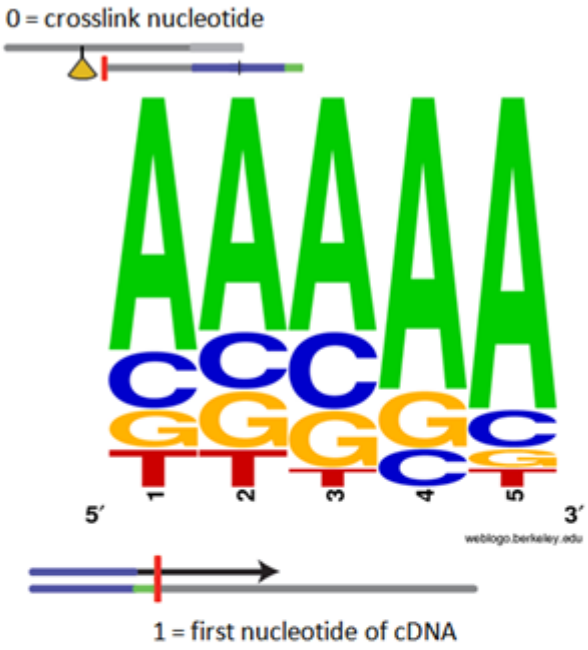
5.3.1.3 Identification of consensus binding motif for SAFB1

The *in vivo* binding specificity of SAFB1 is still currently unknown. The advantage of single nucleotide resolution provided by iCLIP method enabled the assessment of sequence specificity for SAFB1 binding. To derive whether a consensus binding motif exists for SAFB1, enriched pentamer sequences surrounding the crosslink sites were identified. The frequencies of each pentamer were analysed to determine the top 20 pentamers for SAFB1. Strikingly, adenine appeared as the most frequent nucleotide in the top 20 pentamers and represents 68% of the enriched pentamers [Figure 5.8 (a)]. The predicted SAFB1 consensus binding motif contains adenine-rich sequences derived from the pentamers [Figure 5.8 (b)]. When the frequency of each nucleotide in the cDNA libraries was analysed relative to its base position, a strong inclusion of adenine at base position 5 was observed (80%) while thymine (uracil in RNA) was excluded at base position 4 of the putative RNA-binding motif [Figure 5.8 (c)]. The consensus binding motif for SAFB1 has not been described before, thus this novel finding is likely to be of great importance to further our current understanding of SAFB1 RNA-binding specificity.

a)

Pentamer	Frequency
GAAAA	34.644
AGAAA	31.201
AAACA	30.533
AACAA	30.072
AAAAC	29.955
ATGAA	29.857
AAGAA	28.988
GAAGA	26.496
AAAGA	26.144
TGAAA	25.850
AACCA	25.815
ACAAC	25.653
AATGA	25.522
ATCAA	25.514
ACCAA	25.450
ACAAA	25.227
CAGAA	24.147
CAAAT	23.917
CAAAA	23.660
TGAAG	23.323

b)



c)

Base	Base position 1	2	3	4	5
A	65%	60%	60%	75%	80%
C	15%	15%	20%	10%	10%
G	10%	15%	15%	15%	5%
T	10%	10%	5%	0%	5%

Figure 5.8 ***In vivo*** **consensus binding motif of SAFB1.**

(a) The frequency (million k-mers) of pentamers surrounding SAFB1 crosslink sites was determined. Adenine represents 68% of the 20 pentamers that has the highest frequencies. **(b)** Weblogo showing base frequencies of each base at respective positions of the pentamer. SAFB1 binds to adenine-rich motifs. **(c)** The frequency of each base relative to its position within the pentamer was summarised in this table. Highest frequency of adenine was observed at base position 5, thymine was excluded at base position 4 of the consensus binding motif. This consensus binding motif was predicted from iCLIP cDNA libraries, therefore the uracil base is referred as thymine in these sequences.

5.3.2 Identification of novel RNA targets from data generated by iCLIP

Data analysis of bound RNAs revealed the number of SAFB1 crosslink sites within each RNA target (Appendix II). When the top 10 RNA targets with the largest number of crosslink sites were listed according to each RNA segment, the position of SAFB1 binding within each gene was visualised using the USCS Genome Browser.

This enabled the identification of several interesting RNA targets that were selected for validation. Further experimentations were performed using qRT-PCR or conventional PCR on RNAi transfected MCF-7 and MDA-MB-231 cells to verify the effect of loss of SAFB1 on the expression of these selected RNA targets (Section 2.3.7). Since SAFB2 shares 98% sequence homology to the RRM of SAFB1, these cells were also depleted of SAFB2 in this part of the study. Double knockdown of SAFB1 and SAFB2 were also included as previous study in this thesis has shown the importance of their interaction in regulating gene expression (Section 4.3.1.2).

5.3.2.1 *Src homology 2 domain containing F (SHF)*

Analysis of the RNA map revealed a large number of SAFB1 binding sites on the *SHF* mRNA, particularly accumulated around the alternative promoter (Figure 5.9). The use of an alternative promoter plays a significant role in gene expression control [reviewed in (Ayoubi *et al.* 1996, Davuluri *et al.* 2008, Koch *et al.* 2008)]. More importantly, the aberrant use of a alternative promoter has been linked to a number of diseases, including cancer (Singer *et al.* 2008). Therefore, the identification of *SHF* as a potential RNA target for SAFB1 warrants further investigation.

MCF-7 and MDA-MB-231 cells were transfected with SAFB1 siRNA, SAFB2 siRNA or SAFB1 and SAFB2 siRNA and analysed for *SHF* mRNA expression using specific validated TaqMan probes targeting *SHF* in qRT-PCR (Section 2.7.3). Interestingly, the loss of SAFB1 had opposite effects in both cell lines. In MCF-7 cells, loss of SAFB1 did not appear to significantly alter *SHF* mRNA expression; whereas in MDA-MB-231 cells there was a significant increase in *SHF* mRNA expression when SAFB1 was absent (Figure 5.10). Loss of SAFB2 and both SAFB proteins increased *SHF* expression, again supporting their role as transcriptional repressors.

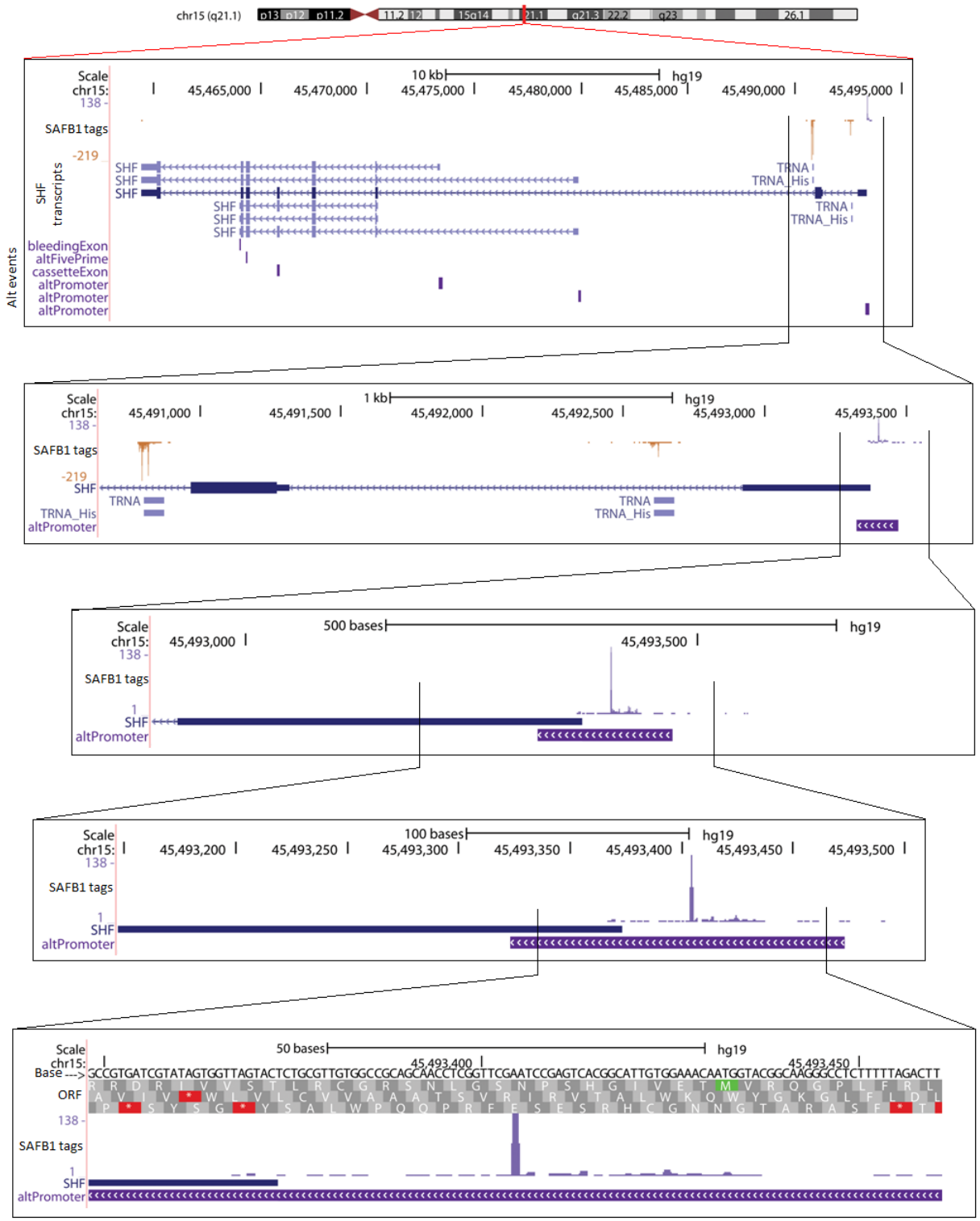


Figure 5.9 Distribution of SAFB1 crosslink sites on *SHF* mRNA.

Overview of the crosslink nucleotides on chromosome 15 shows the location of *SHF* mRNA. SAFB1 crosslink sites are enriched within the alternative promoter of *SHF* pre-mRNA (SAFB1 tags). The bottom tracks represent known *SHF* transcripts and alternative splicing events that occur within this gene. Figure represents a modified image of the UCSC genome browser (human genome, version hg19, chromosome 15, nucleotides 45,459,412 to 45,493,373).

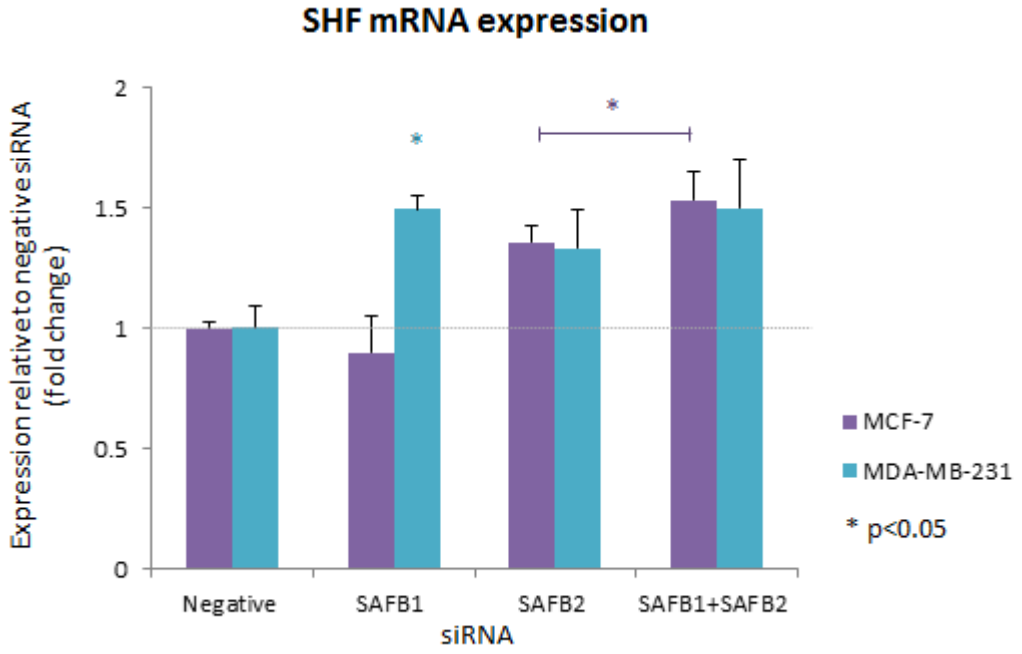


Figure 5.10 Expression of *SHF* in the absence of SAFB1 or/and SAFB2.

qRT-PCR was performed on mRNA from MCF-7 and MDA-MB-231 cells transfected with negative, SAFB1, SAFB2 or SAFB1 and SAFB2 siRNA using validated TaqMan probes specifically targeting *SHF*. Data represents the average of three biological replicates ± S.D. Statistical significance of mRNA expression was calculated using a student's t-test. * = $p<0.05$.

5.3.2.2 *Serglycin (SRGN)*

The identification of *SRGN* in the list of RNA targets also sparked an interest for further validation due to concentrated SAFB1 crosslink sites along all the exons, especially around the cassette exon (Figure 5.11). Previous evidence has shown the ability of SAFB1 and SAFB2 to promote the skipping of a variable exon, although this splicing event was independent of their RNA-binding ability (Nayler *et al.* 1998, Sergeant *et al.* 2007, Stoilov *et al.* 2004). The interaction between SAFB1 and *SRGN* RNA may shed some light on the direct or indirect involvement of SAFB proteins in alternative splicing, therefore was chosen as a RNA target for validation.

MCF-7 and MDA-MB-231 cells were transfected with SAFB1 siRNA, SAFB2 siRNA or SAFB1 and SAFB2 siRNA as described, and analysed for *SRGN* exon skipping event using specially designed primers spanning the cassette exon in conventional PCR (Section 2.9.3). The primers were designed to allow a distinct size separation of PCR products dependent on the inclusion or exclusion of the cassette exon (Figure 2.2).

The full length *SRGN* transcript produces a PCR product of 325bp while the transcript that lacks the cassette exon produces a 177bp PCR product. Conventional PCR only detected the 325bp PCR product in all MCF-7 and MDA-MB-231 samples tested, while the 177bp PCR product was not detected (Figure 5.12).

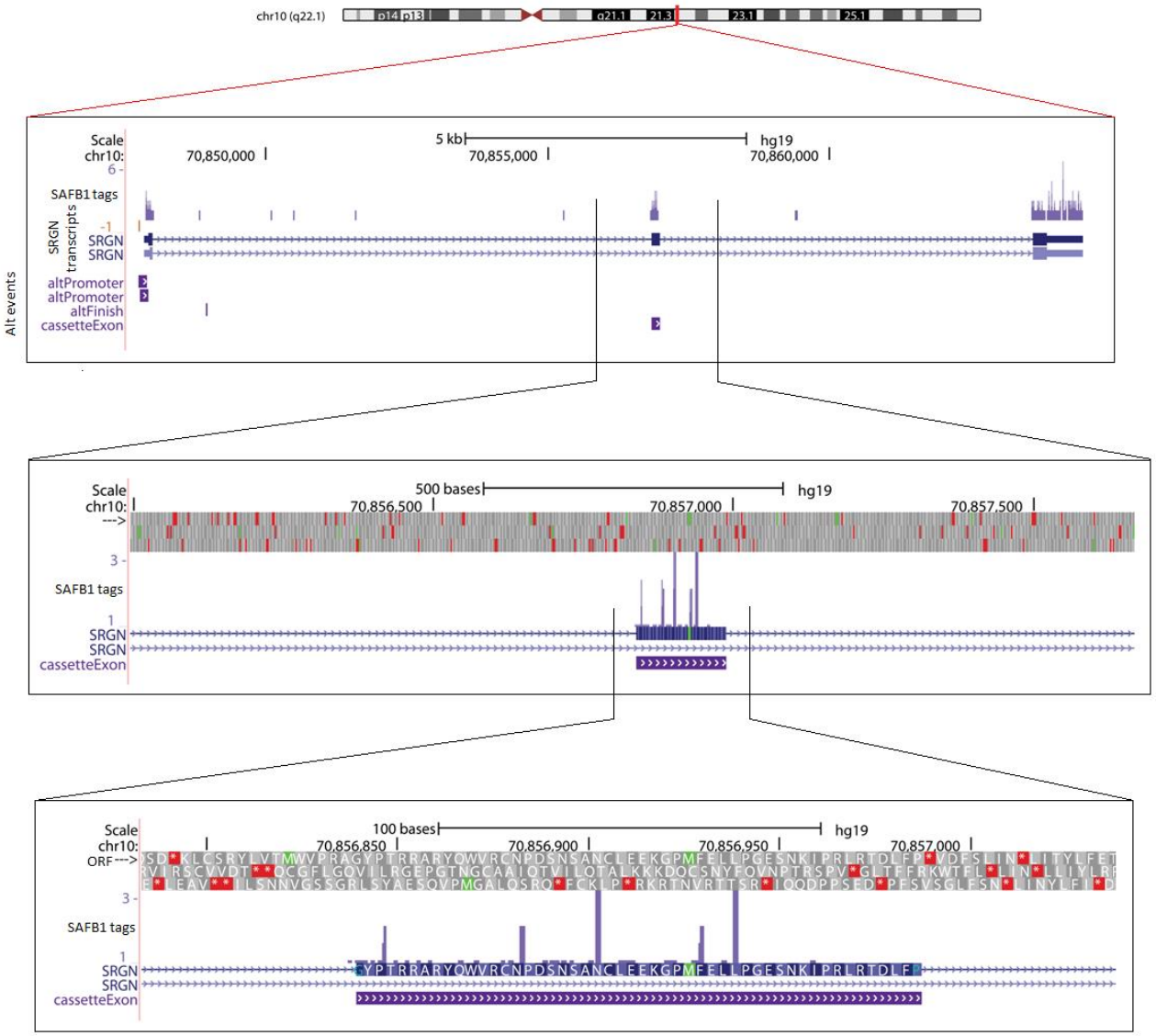


Figure 5.11 Distribution of SAFB1 crosslink sites on *SRGN* mRNA.

Overview of the crosslink nucleotides on chromosome 10 shows the location of *SRGN* mRNA. SAFB1 crosslink sites are enriched on exons, particularly around the cassette exon (SAFB1 tags). The bottom tracks represent known *SRGN* transcripts and alternative splicing events that occur within this gene. Figure represents a modified image of the UCSC genome browser (human genome, version hg19, chromosome 10, nucleotides 70,847,828 to 70,864,567).

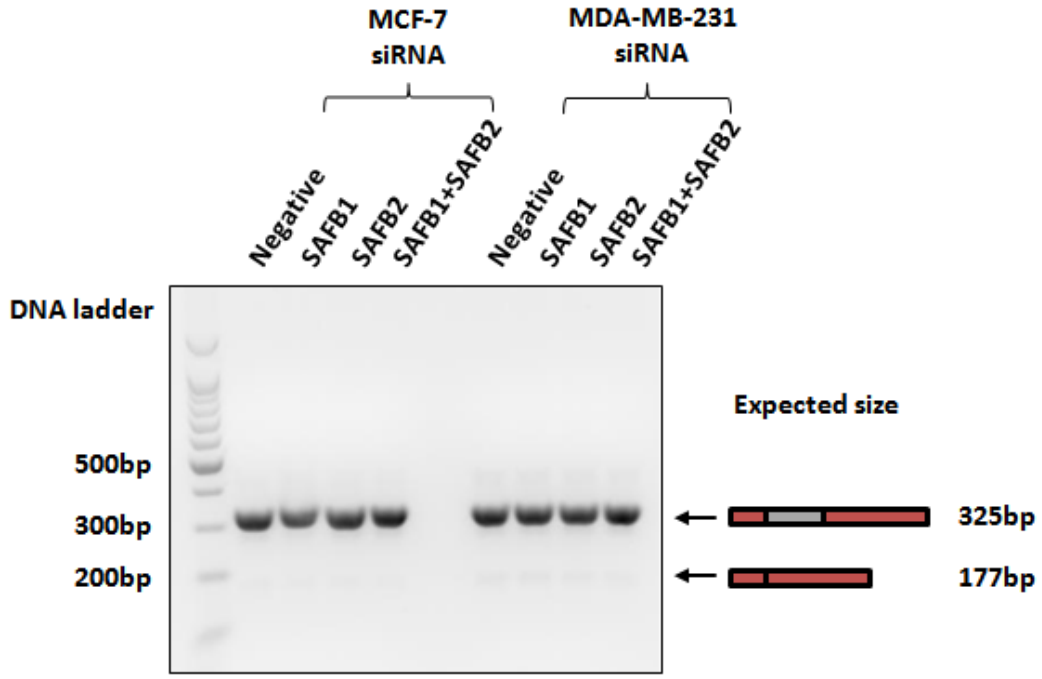


Figure 5.12 Analysis of *SRGN* exon skipping event in the absence of SAFB1 or/and SAFB2.

Conventional PCR was performed on mRNA from MCF-7 and MDA-MB-231 cells transfected with negative, SAFB1, SAFB2 or SAFB1 and SAFB2 siRNA using primers spanning across the cassette exon. PCR products were separated by electrophoresis on an agarose gel. Data represents a gel image of at least three biological replicates.

5.3.2.3 *Integrin beta 4 (ITGB4)*

The list of 12 SAFB-regulated genes obtained from the gene expression profile study in Chapter 4 were examined in the RNA map using the UCSC genome browser to identify direct SAFB1 binding sites in these genes. Interestingly, overlapping of the dataset revealed that only *ITGB4* mRNA contains SAFB1 crosslink sites. Due to time constraints, further experimentation was not performed to examine this observation. However, preliminary computational analysis of the iCLIP data shows that SAFB1 crosslink nucleotides were present along the entire length of the *ITGB4* mRNA, suggesting that it may be involved in the expression of *ITGB4* through direct RNA interaction (Figure 5.13).

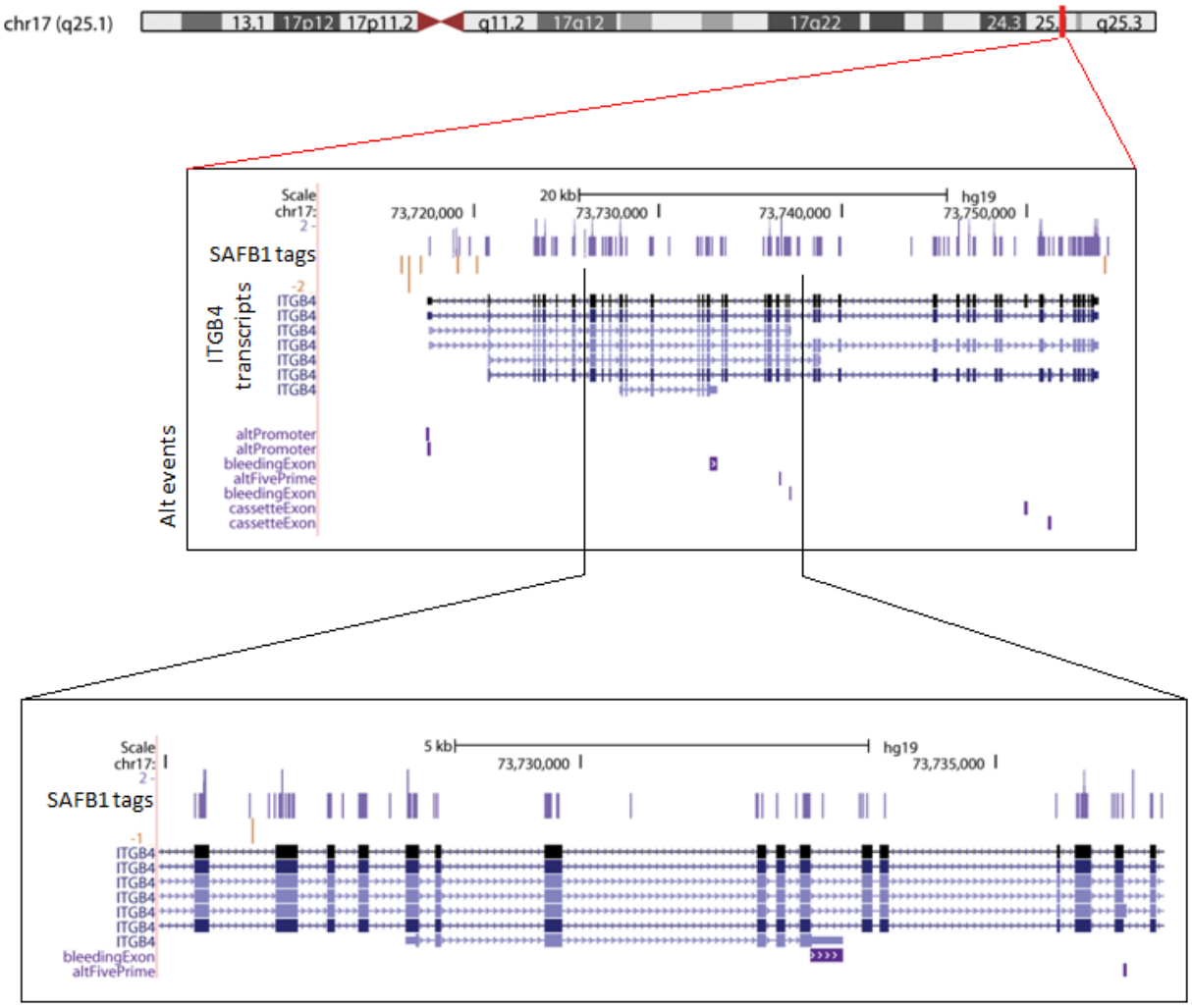


Figure 5.13 Distribution of SAFB1 crosslink sites on *ITGB4* mRNA.

Overview of the crosslink nucleotides on chromosome 17 shows the location of *ITGB4* mRNA. SAFB1 crosslink sites are enriched along the entire length of the exons (SAFB1 tags). The bottom tracks represent known *ITGB4* transcripts and alternative splicing events that occur within this gene. Figure represents a modified image of the UCSC genome browser (human genome, version hg19, chromosome 17, nucleotides 73,717,516 to 73,753,899).

5.3.2.4 *Transformer 2 protein homolog beta (TRA2B)*

Both SAFB1 and SAFB2 have been implicated in the alternative splicing of *TRA2B* (Sergeant *et al.* 2007, Stoilov *et al.* 2004). Interestingly, Stoilov *et al.* also revealed that SAFB1 is able to promote *TRA2B* exon skipping independent of its RNA-binding ability (Stoilov *et al.* 2004). Using the iCLIP data obtained from this study, *TRA2B* was analysed for SAFB1 binding sites to further validate these reported observations. Computer analysis of *TRA2B* mRNA revealed that SAFB1 binding sites were not

present in the *TRA2B* variable exons, supporting the observation by Stoilov *et al.* (Figure 5.14).

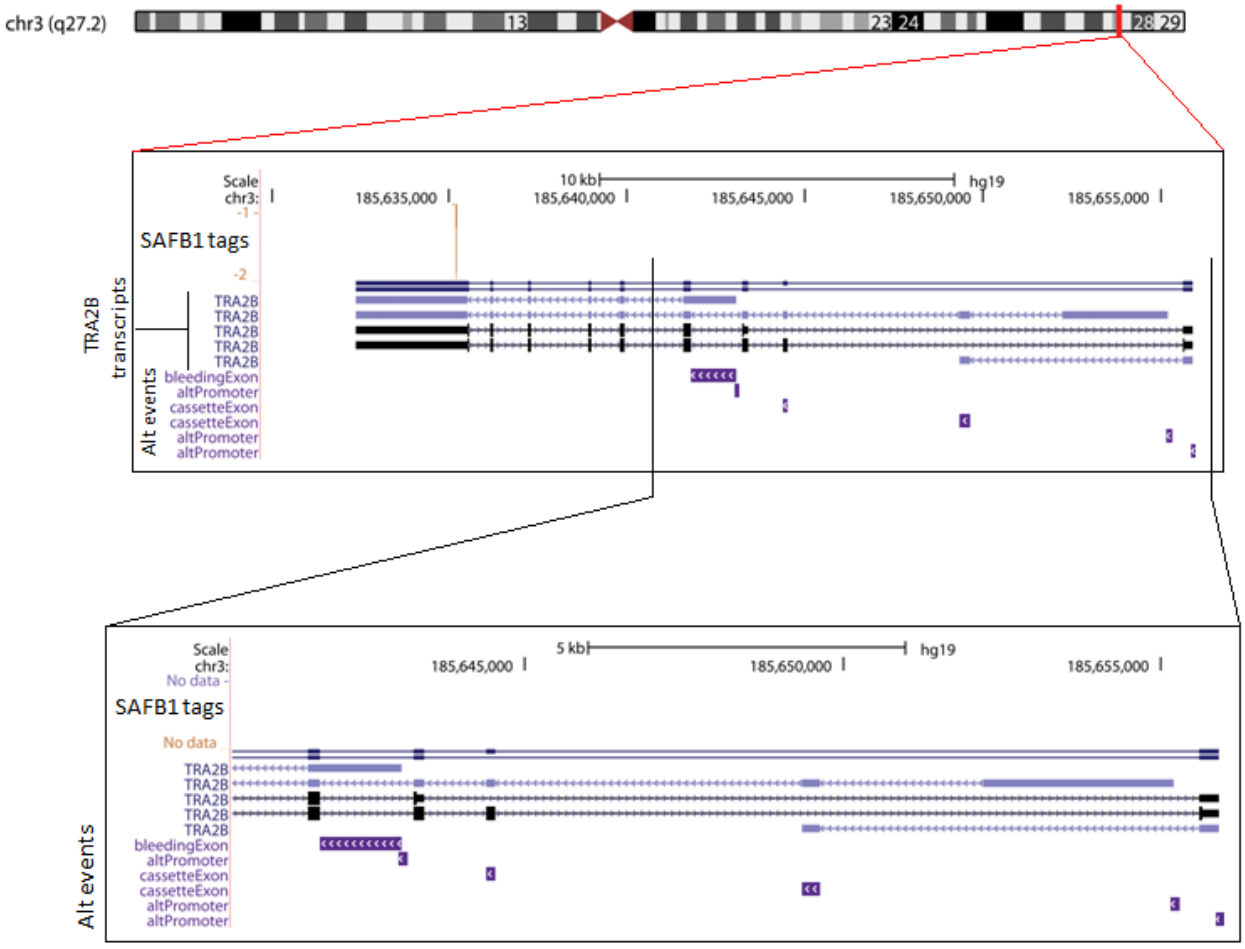


Figure 5.14 Distribution of SAFB1 crosslink sites on *TRA2B* mRNA.

Overview of the crosslink nucleotides (SAFB1 tags) on chromosome 3 shows the location of *TRA2B* mRNA. SAFB1 crosslink sites are not present in the variable exons (SAFB1 tags). The bottom tracks represent known *TRA2B* transcripts and alternative splicing events that occur within this gene. Figure represents a modified image of the UCSC genome browser (human genome, version hg19, chromosome 3, nucleotides 185,632,358 to 185,655,924).

5.4 Discussion

The presence of the highly conserved RRM within SAFB1 and SAFB2 proteins has been a subject of interest since their discovery, especially in relation to their RNA-binding potential. Despite the fascination surrounding this question, very little has been undertaken and reported to describe their RNA-binding capabilities. Initial *in vitro* evidence has shown that the RRM of SAFB1 was able to bind RNA when glutathione S-transferase (GST)-tagged SAFB1 protein combined with total RNA from MCF-7 cells could generate a PCR product when reverse transcribed and PCR amplified (Townson *et al.* 2004). In that report, the identity of the RNA targets was not described and important questions with respect to the role of SAFB1 and SAFB2 in RNA processing remained unanswered. This area of research has not yet been investigated, therefore this work will serve as a pioneering study to advance our current knowledge on the RNA-binding potential of SAFB1 in breast cancer.

iCLIP has been proven as a powerful method to determine protein-RNA interactions *in vivo* on a global scale and identify the positions of crosslink sites at nucleotide resolution (Konig *et al.* 2010). The random barcode incorporated to individual cDNA molecules addresses the problem of PCR artifacts faced by all high-throughput sequencing methods. iCLIP has generated a huge dataset and this is an initial analysis of the RNA-binding data for SAFB1. In this study, a global view comparison of the complete dataset from each individual biological replicate showed that all datasets generated consistent and reproducible results, underlining the high quality iCLIP data achieved by high stringency purification and library preparation (Figure 5.5). The use of a highly specific SAFB1 antibody that has been tested multiple times prior to the onset of the iCLIP experiments also greatly contributed to the quality of the iCLIP data (Section 3.3.1 and Section 5.3.1.1).

The identification of *in vivo* targets by iCLIP enabled the mapping of transcript regions and RNA classes bound by SAFB1. An overview of the iCLIP results showed that the important class of RNAs bound by SAFB1 was ncRNAs. Although SAFB1 binds to both coding and ncRNAs, highest proportion of SAFB1 crosslinked to ncRNAs and the highest density enrichment was also detected in ncRNAs (Figure 5.6). Interestingly, this binding distribution of SAFB1 is similar to the RNA-binding distribution of SRSF3 and SRSF4 splicing factors rather than hnRNP C protein, even though SAFB1 was initially classified as a novel member of the hnRNP protein family (Weighardt *et al.* 1999).

Recent work by Anko *et al.* utilised the iCLIP method to reveal that concentrated SRSF3 and SRSF4 binding sites were also in ncRNAs (Anko *et al.* 2012), while Konig *et al.* showed that hnRNP C binding sites were most abundant within introns (Konig *et al.* 2010). This observation raises the possibility that SAFB1 protein may have similar characteristics to SR proteins rather than hnRNP protein members, although at this stage this is only speculative.

The term ncRNA is commonly used for RNA that does not encode a protein but appears to comprise internal signals that control various levels of gene expression, including chromatin organisation, transcription, RNA splicing, editing, translation and turnover [reviewed in (Mattick *et al.* 2006)]. Consistent with already known functions of SAFB proteins, concentrated SAFB1 binding in ncRNAs observed from the iCLIP data could possibly contribute to its various role in chromatin organisation, transcription and RNA metabolism (Section 1.4.2). Analysis of SAFB1 distribution within ncRNA subclasses revealed most abundant SAFB1 binding in snRNAs (Figure 5.7). snRNAs are a class of small RNA molecules found to be uridylate-rich and localised within the nucleus (Busch *et al.* 1982). The most common members of snRNAs are the U1, U2, U4, U5 and U6 snRNAs that form the spliceosome along with many other protein factors and primarily function in pre-mRNA splicing [reviewed in (Wahl *et al.* 2009)]. The high distribution of SAFB1 binding sites in snRNAs observed in this study supports previously identified interactions between SAFB1 with various RNA processing factors and splicing machinery (Li *et al.* 2003, Nayler *et al.* 1998, Rappsilber *et al.* 2002, Weighardt *et al.* 1999).

Another interesting observation from the iCLIP data revealed significant SAFB1 binding sites to metastasis associated lung adenocarcinoma transcript 1 (*MALAT-1*) (data not shown). *MALAT-1* is a highly conserved long ncRNA enriched in nuclear speckles that regulates alternative splicing by modulating splicing factor phosphorylation (Tripathi *et al.* 2010). Several studies have shown that *MALAT-1* is overexpressed in many different cancers including breast cancer and suggest that *MALAT-1* is an oncogenic long ncRNA (Li *et al.* 2009, Perez *et al.* 2008). It has also been shown that *MALAT-1* co-localises with SC35 in nuclear speckles (Hutchinson *et al.* 2007). In addition, abundant SAFB1 crosslink sites were also found on *7SK* RNA (data not shown), another nuclear speckle localised ncRNA (Prasanth *et al.* 2010). These findings support the previously observed punctuate pattern of SAFB1 distribution indicative of its ‘nuclear speckles’ localisation (Section 3.3.2.2). Furthermore, other

splicing factors that localise in nuclear speckles such as SRSF1, SRSF3 and SRSF4 also bind to *MALAT-1* and *7SK* (Anko *et al.* 2012, Sanford *et al.* 2009). It is therefore conceivable that SAFB1 may possess other typical characteristics of a splicing factor that supports its observed function in pre-mRNA splicing (Nayler *et al.* 1998, Sergeant *et al.* 2007, Stoilov *et al.* 2004).

The genome-wide, single nucleotide resolution of iCLIP data enabled the prediction of *in vivo* consensus binding sequences for SAFB1 based on the enriched pentamer sequences surrounding the crosslink sites. This study is the first to report that SAFB1 binds a consensus adenine-rich sequence *in vivo* [Figure 5.8 (b)]. Closer examination of the putative consensus binding sequence revealed the exclusion of thymine at base position 4 and a strong inclusion of adenine at base position 5 [Figure 5.8 (c)]. Interestingly, the predicted SAFB1consensus binding motif shows high similarity to the purine-rich sequences found in RNA-binding motifs for other SR proteins [reviewed in (Long *et al.* 2009)].

When analysing SAFB1 crosslink sites within protein-coding transcripts, SAFB1 binding density was also enriched in regions encompassing the ORF, 3' and 5' UTR [Figure 5.6 (b)]. The list of RNA targets was filtered according to the region and density of SAFB1 binding to identify targets that are relevant to tumourigenesis. Several interesting genes were highlighted in this study.

SHF

SHF is a member of a family of adaptor protein characterised by their ability to mediate protein-protein interactions through their Src homology 2 domain (Lindholm *et al.* 2000, Welsh *et al.* 1994). Although the function of SHF is not fully understood, evidence has shown that overexpression of SHF significantly decreases the rate of growth factor-induced apoptosis in neuroblastoma cells (Lindholm *et al.* 2000). Subsequently, Ohira *et al.* showed that *SHF* mRNA was highly expressed in non-metastatic neuroblastoma compared to metastatic tumour samples (Ohira *et al.* 2003). Another recent study provided evidence that loss of SHF increased cellular mobility and the invasive capability of neuroblastoma cells (Takagi *et al.* 2013).

Initial iCLIP data from this study revealed enriched SAFB1 binding sites at the alternative promoter of *SHF* (Figure 5.9). As the aberrant expression of alternative promoters is linked to cancer, SAFB1 binding surrounding this region gathered an

interest for further examination. Interestingly, the knockdown of SAFB1 in MCF-7 cells did not significantly alter *SHF* expression while the knockdown of SAFB2 or both SAFB proteins significantly increased *SHF* expression (Figure 5.10). This suggests that direct SAFB1 binding to the alternative promoter did not affect the expression of this gene. MDA-MB-231 cells were included in this part of the study for comparison and in this cell type, increased *SHF* expression was observed in the loss of SAFB1 or SAFB2 and both SAFB proteins. Although the RNA-binding pattern of SAFB1 in MDA-MB-231 cells is not yet known, this data shows that SAFB1 differentially regulate *SHF* expression in non-invasive (MCF-7) and invasive (MDA-MB-231) breast cancer cells.

SRGN

SRGN encodes the protein serglycin, a proteoglycan composed of a core peptide with serine-glycine dipeptide repeats and various glycosaminoglycan side chains [reviewed in (Kolset *et al.* 2008)]. Overexpression of serglycin has been observed in nasopharyngeal carcinoma (NPC) and shown to promote motility, invasion and metastasis of NPC cells (Li *et al.* 2011). The core protein of serglycin is encoded by three exons that form three functional domains. Exon 1 encodes a signal peptide domain that is commonly present in secreted proteins, exon 2 encodes the N-terminal domain and exon 3 encodes a glycosylation domain that enables glycosaminoglycan attachment (Nicodemus *et al.* 1990). Exon 2 of *SRGN* has been identified as a cassette exon although the significance of exon 2 deletion is still uncertain. However, Castronuevo *et al*. has also speculated that a deletion would generate a different N-terminal sequence that could influence interactions of serglycin with other proteins (Castronuevo *et al.* 2003).

iCLIP data revealed abundant binding of SAFB1 within the entire exonic region of *SRGN* that dropped towards the exon-intron boundaries (Figure 5.11). This pattern of binding resembles that of SR proteins that are known to bind to exonic splicing enhancers, where they influence adjacent splice sites [reviewed in (Blencowe 2000)]. Considering the similarities that have been observed in this study between SAFB1 and SR proteins, the effect of both SAFB paralogs on the alternative splicing of *SRGN* exon 2 was examined in breast cancer cells. Data shows that the knockdown of SAFB1 or SAFB2 and both SAFB proteins in MCF-7 and MDA-MB-231 cells did not alter the inclusion or exclusion of *SRGN* exon 2 (Figure 5.12). This observation

suggests that SAFB1 may not be involved in the alternative splicing of the *SRGN* cassette exon.

ITGB4

As previously discussed in Chapter 4 (Section 4.4), ITGβ4 plays an important role in cancer progression by influencing the migration, invasion and survival of cancer cells (Lipscomb *et al.* 2005, Wilhelmsen *et al.* 2006). Interestingly, the comparison between the gene expression array dataset (Section 4.3.1.2) and iCLIP data in this study highlights only one common target, *ITGB4*. Analysis of the *ITGB4* mRNA revealed abundant SAFB1 binding sites along the entire length of the RNA, suggesting that SAFB1 may be involved in post-transcriptional modification of *ITGB4* (Figure 5.13). Multiple alternatively spliced transcript variants encoding distinct isoforms have been found for *ITGB4*, although the full function of most variants remains to be defined (Clarke *et al.* 1994, Tamura *et al.* 1990, van Leusden *et al.* 1997). Alternative splicing mechanism has been indicated to subtly regulate the ligand binding and signalling activity of many integrin subunits [reviewed in (de Melker *et al.* 1999)]. Although the mechanism and significance of alternative splicing in *ITGB4* has not been elucidated, the discovery of SAFB1 binding sites in its exonic regions may provide a new perspective to further understand the mRNA processing of *ITGB4*.

TRA2B

TRA2B encodes the Tra2β splicing regulator protein which functions predominantly in mRNA splicing [reviewed in (Elliott *et al.* 2012)]. Tra2β activates the splicing inclusion of bound exons by direct binding to target sites within an exon (Clery *et al.* 2011, Grellscheid *et al.* 2011, Tacke *et al.* 1998). The activation of splicing events by Tra2β is concentration dependent, as increased Tra2β concentration correlates with increased levels of target exon inclusion (Elliott *et al.* 2012, Grellscheid *et al.* 2011). Overexpression of Tra2β has been observed in several cancers, including breast cancer and associated with tumour metastasis (Best *et al.* 2013, Fischer *et al.* 2004, Gabriel *et al.* 2009, Ouyang *et al.* 2011, Watermann *et al.* 2006). Tra2β protein concentration is known to be influenced by the alternative splicing of a *TRA2B* variable exon. Inclusion of the *TRA2B* variable exon introduces a premature stop codon and prevents translation of the full length protein, thus reducing the concentration of Tra2β protein. Tra2β has been shown to autoregulate its protein concentration by influencing the inclusion of its own variable exon (Stoilov *et al.* 2004). SAFB1 and SAFB2 proteins

have also been implicated in the alternative splicing of the *TRA2B* variable exon (Sergeant *et al.* 2007, Stoilov *et al.* 2004). Specifically in both studies, SAFB proteins inhibit the inclusion of *TRA2B* variable exon in embryonic kidney cells (HEK293) transfected with *TRA2B* minigene thereby increasing the concentration of Tra2β protein. However, this splicing inhibition by SAFB1 was mediated independent of its RNA-binding ability, suggesting that SAFB1 may not bind directly to *TRA2B* pre-mRNA (Stoilov *et al.* 2004). Using the iCLIP data for SAFB1 obtained from this study, *TRA2B* mRNA was analysed for the presence of SAFB1 binding sites. Computational analysis shows that SAFB1 binding sites were not present at the *TRA2B* variable exon (Figure 5.14). Taken together, these evidences provide additional rationale to speculate that the effect of SAFB1 on *TRA2B* variable exon splicing is a result of an indirect RNA interaction possibly mediated through its interaction with other splicing factors.

In summary, this study is the first to investigate the RNA-binding ability of SAFB1 in breast cancer cells. The data revealed that SAFB1 binds directly to RNA and its binding is particularly enriched at adenine-rich sequences. Utilising the powerful single nucleotide resolution iCLIP method, a putative SAFB1 consensus binding motif was derived for the first time. SAFB1 shares multiple similarities in RNA-binding pattern and characteristics with SR proteins although it was initially identified as a novel member of the hnRNP protein family. Analysis of SAFB1 crosslink regions and RNA targets confirm previous reports regarding its interaction with other RNA processing machinery and function.

Chapter 6 : Summary and future work

6.1 Summary

At the onset of this study, very little was known about the fundamental characteristics and biological effect of the two related SAFB proteins, SAFB1 and SAFB2. The work presented in this thesis, utilising RNAi and both medium and high throughput experimental strategies, provides evidence to explain how SAFB1 and SAFB2 are regulated and also what impact they have on the regulation of SAFB target genes. Novel SAFB1 and SAFB2 target genes, with defined roles in breast cancer biology, were identified in this study.

In Chapter 3, ER positive and ER negative breast cancer epithelialcells were used to characterise the regulation of SAFB1 and SAFB2 in response to steroid hormone; results presented here show that these ER coregulatory proteins are themselves regulated by oestrogen in ER positive MCF-7 cells and also ER negative MDA-MB-231 cells. In MCF-7 cells, the regulation of SAFB1 and SAFB2 by 17β-oestradiol appears to be via a classical ER-signalling pathway suggesting that they are typical oestrogen-responsive genes. The participation of ER-α66 in this mechanism of regulation was confirmed by the stimulation with pure ER antagonist, fulvestrant. Oestrogen response elements (ERE) were also identified in *SAFB1* and *SAFB2* genomic sequences, another characteristic of ER target genes.

In MDA-MB-231 cells, oestrogen regulates SAFB1 and SAFB2 protein expression at the post-translational level. 17β-oestradiol appears to disrupt SAFB1 and SAFB2 protein turnover rate by reducing their protein stability. The participation of ubiquitin-mediated protein degradation pathway was indicated when this effect by 17β-oestradiol was protected by proteasome inhibition. *In vitro* evidence obtained from this chapter suggests that SAFB1 and SAFB2 are regulated by different mechanisms in cell lines that display distinct breast cancer phenotypes. These findings led to the proposition that SAFB1 and SAFB2 are oestrogen-responsive genes. Another outcome highlighted in this part of the study is the ability of fulvestrant to increase expression of the ER-α variant, ER-α36 in the MDA-MB-231 breast cancer cells.

In Chapter 4, a medium throughput gene expression profile study was undertaken to investigate the role of SAFB1 and SAFB2 in the regulation of known ER target genes.

From a panel of 84 pre-determined ER-related candidate genes, twelve potential targets closely linked to tumour progression were identified for SAFB1 and SAFB2 in the ER negative MDA-MB-231 cells. Eleven of the targets increased in their mRNA expression when cells were depleted of SAFB1 or/and SAFB2 by RNAi, thus confirming their primary role as transcriptional repressors. This study has also established a new link between SAFB1 and SAFB2 in the regulation of *ITGB4* and *IL-6*. Another novel finding that is contrary to the observation in non-invasive MCF-7 breast cancer cells is that SAFB2, not SAFB1, plays a more prominent role in transcriptional regulation in the invasive MDA-MB-231 breast cancer cells. In addition, data presented from this study raises the possibility that SAFB1 and SAFB2 function synergistically to repress the transcription of their target genes.

In Chapter 5, the RNA-binding ability of SAFB1 was examined in MCF-7 breast cancer cells using the iCLIP high throughput method that identifies direct protein-RNA interactions *in vivo*. Mapping of transcript regions bound by SAFB1 revealed high density enrichment in ncRNAs, particularly within snRNAs. These snRNA are localised within the nucleus and form the spliceosome, thus confirming reports that identified interactions between SAFB1 and various RNA processing factors. Significant SAFB1 binding sites were identified in *MALAT-1*, a highly conserved and potentially oncogenic ncRNA that co-localises with SC35 in nuclear speckles. This provides further explanation to the nuclear speckle-like localisation of SAFB1 observed in Chapter 3 (Section 3.3.2.2). Data also revealed the first description of a putative consensus RNA-binding sequence for SAFB1. The proposed consensus RNA-binding motif for SAFB1 is an adenine-rich sequence, highly similar to the RNA-binding motifs for other SR proteins. The iCLIP data also revealed the presence of SAFB1 on *ITGB4* mRNA, supporting the role of SAFB proteins in regulating *ITGB4* expression observed in Chapter 4 (Section 4.3.2.2). This initial analysis of SAFB1 RNA-binding patterns, binding targets and binding sequences shows that SAFB proteins share multiple characteristics with the SR protein family.

Taken together, the evidence presented in this thesis show that oestrogen exerts its effect on SAFB expression via the genomic and non-genomic ER-signalling pathway in MCF-7 and MDA-MB-231 cells respectively. ER-α66 upregulates the expression of SAFB mRNA through the classical genomic pathway in ER positive breast cancer cells, while ER-α36 downregulates SAFB protein expression through the ubiquitin-mediated protein degradation pathway in ER negative breast cancer cells (Figure 6.1).

The downregulation of SAFB proteins in ER negative cells led to significant transcriptional alterations of 12 novel target genes, notably in *ITGB4* and *IL-6*. More importantly, the transcriptional changes observed in *ITGB4* is associated with direct SAFB1 binding to its mRNA. SAFB1 also binds directly to other RNAs and its binding sites are particularly enriched in ncRNAs. The consensus RNA-binding sequence for SAFB1 is defined as adenine-rich pentamers within the transcriptome. The model below proposes the mechanism of SAFB regulation and function that summarises the key findings highlighted in this thesis.

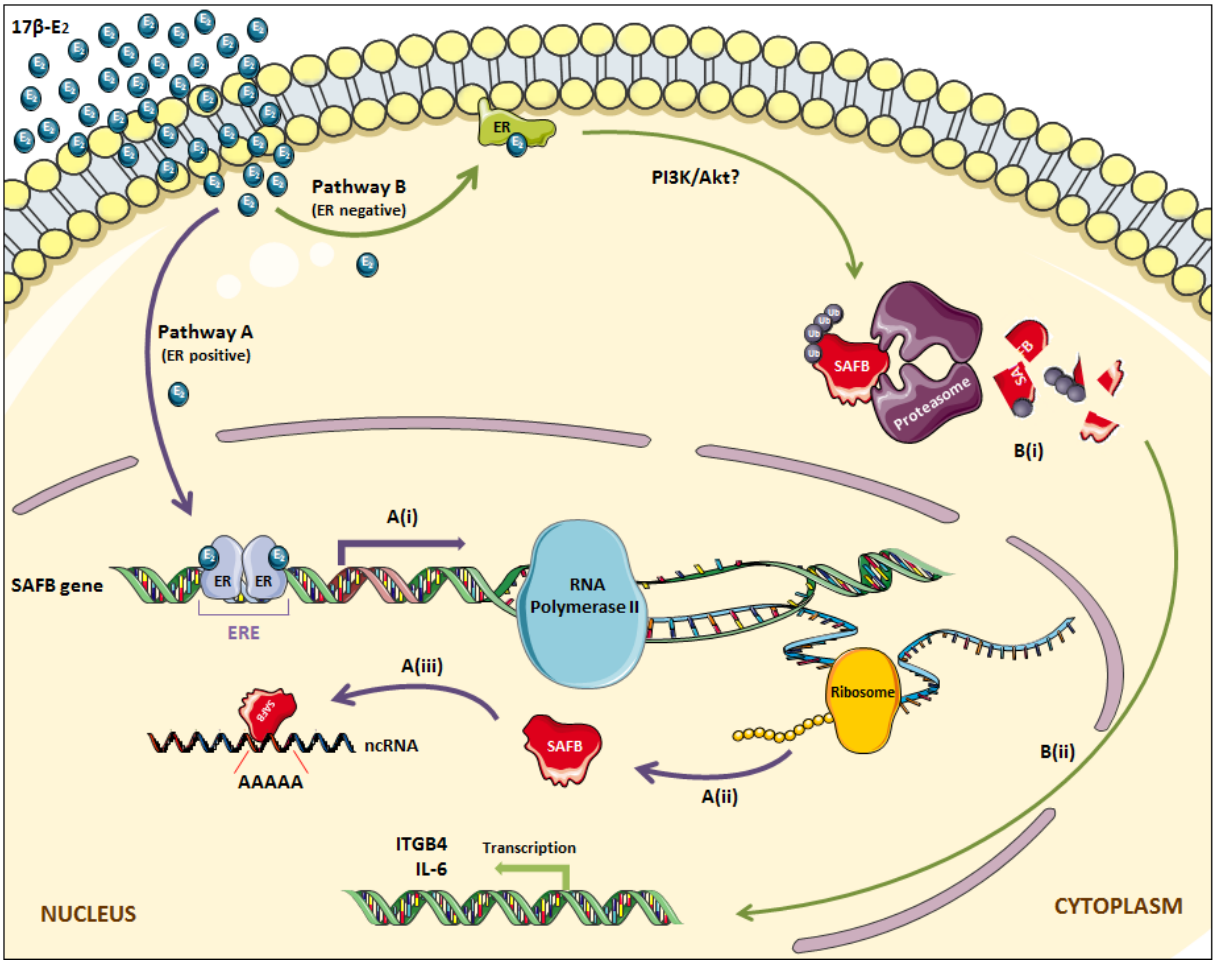


Figure 6.1 Proposed mechanism of SAFB regulation and function in breast cancer cells.

Diagramatic illustration shows the proposed mechanism of action for SAFB regulation and function observed in this thesis. Prolonged exposure to the active 17β-oestradiol metabolite exerts its action via two different mechanisms. Pathway A refers to its ability to activate genomic signalling through ER-α66 present in ER positive cells. **A(i)** Activated ER-α66 is able to initiate the transcription of SAFB mRNA possibly through several ERE that are present within this

gene. **A(ii)** Transcribed mRNA is subsequently translated into SAFB protein. **A(iii)** SAFB protein then exerts its function in the cell through direct binding to RNA at adenine-rich sites. Pathway B refers to oestrogen's ability to activate non-genomic signalling through ER-α36 present in ER negative cells. **B(i)** Activated membrane bound ER-α36 could lead to the degradation of SAFB protein possibly through the PI3K/Akt signalling pathway. **B(ii)** The progressive loss of SAFB protein, especially SAFB2, in the cell leads to the transcriptional change of several target genes, particularly *ITGB4* and *IL-6*. The illustration was created using Servier Medical Art (servier.co.uk/medical-art-gallery).

Several limitations in this work have been acknowledged. In Chapter 3, the SAFB2 antibody was unable to successfully immunoprecipitate its own peptide (Section 3.3.1.2). Although this observation is linked to its intracellular localisation in MCF-7 cells, a pan-SAFB antibody could have been included as a positive control in this experiment. When investigating the effects of 17β-oestradiol in SAFB protein turnover, the rate of protein degradation was examined using the proteasome inhibitor MG132 (Section 3.3.3.2). However, the steady-state level of any protein is the outcome of the change in its rate of synthesis compared with its rate of degradation. The effect of 17β-oestradiol on the rate of SAFB protein synthesis should also be taken into account to fully understand SAFB protein turnover. This could be performed by inhibiting protein synthesis using the translation inhibitor cycloheximide.

In Chapter 4, the gene expression array used was based on the SYBR green detection method. This method utilises a double-stranded DNA binding dye (SYBR green) to detect and quantify PCR products as it accumulates during qRT-PCR. The limitations of using this method is the possibility of false positive signals generated through the binding of the SYBR green dye to any double-stranded DNA sequences and its inability to distinguish between specific and non-specific PCR products. The SYBR green detection method is therefore highly dependent on the specificity of the primers to not amplify non-target sequences. Experience from using this SYBR green gene expression array revealed a weakness in the primers targeting *ESR1*. Analysis of the melting curve revealed multiple peaks using this primer, suggesting amplification of more than one PCR product. *In silico* analysis revealed a plethora of non-specific targets that could be amplified by these primers (Appendix I). Preliminary experiment was performed to examine the specificity of the *ESR1* primers using conventional PCR (Figure 4.11). The PCR product generated could be sequenced to confirm the specificity of the *ESR1*

primers. Nevertheless, accumulated data from these analyses show the weakness in the design of *ESR1* primers used in this array.

This example highlights the limitation in direct interpretation of the SYBR green qRT-PCR results and emphasises the need for caution when analysing data. It also indicates the necessity of validation work with a more specific method such as TaqMan gene expression assays that contain both primers and gene-specific probes. It is noteworthy to mention that the gene expression array plate used in this study has since been improved and the primers for both *ESR1* and *CLU* have been exchanged.

Another limitation noted in this work is the lack of 'normal' breast cells and tissues to assist the further analysis of the data. Primary breast cells derived from healthy individuals could be included for *in vitro* studies, while tissue from breast reduction surgery could be used as a control for *in vivo* studies. In the *ITGB4* gene expression studies, a brief experiment was performed with only two breast tumour samples (Section 4.3.2.2). Although 'normal' samples were included in this work, they were derived from non-involved tissues within the same tumour sample. This indicates that even though these tissues were not involved in tumourigenesis, they came from diseased patients and may not truly reflect a normal phenotype. Furthermore, incorporating truly normal breast phenotype as experimental controls (e.g. from cosmetic breast reduction surgery) could facilitate the analysis for biological significance of the observed data.

In conclusion, despite the limitations aforementioned, the work presented in this thesis provides new insights into understanding the regulation, function and RNA-binding ability of SAFB1 and SAFB2. It has also paved the way for new opportunities and interesting scope of study for future research into the role of SAFB1 and SAFB2 in breast cancer.

6.2 Future work

Future work that will follow on from this study includes investigating the *in vivo* implications and physiological relevance of SAFB1 and SAFB2 regulated genes in different breast cancer subtypes. Experimental strategies will also be included to evaluate the mechanism(s) involved in SAFB1- and SAFB2- mediated regulation of *ITGB4* and *IL-6*. Another direction will be to verify the predicted consensus binding motif for SAFB1 using computational and experimental strategies. The iCLIP work has generated a vast iCount dataset, therefore future work will continue analysing the dataset and also perform RNA sequencing using breast cancer cells expressing or lacking SAFB1 or/and SAFB2.

References

Alarid, E. T., N. Bakopoulos and N. Solodin (1999). "Proteasome-mediated proteolysis of estrogen receptor: a novel component in autologous down-regulation." Mol Endocrinol **13**(9): 1522-1534.

Ali, S., D. Metzger, J. M. Bornert and P. Chambon (1993). "Modulation of transcriptional activation by ligand-dependent phosphorylation of the human oestrogen receptor A/B region." EMBO J **12**(3): 1153-1160.

Anko, M. L., M. Muller-McNicoll, H. Brandl, T. Curk, C. Gorup, I. Henry, J. Ule and K. M. Neugebauer (2012). "The RNA-binding landscapes of two SR proteins reveal unique functions and binding to diverse RNA classes." Genome Biol **13**(3): R17.

Anzick, S. L., J. Kononen, R. L. Walker, D. O. Azorsa, M. M. Tanner, X. Y. Guan, G. Sauter, O. P. Kallioniemi, J. M. Trent and P. S. Meltzer (1997). "AIB1, a steroid receptor coactivator amplified in breast and ovarian cancer." Science **277**(5328): 965-968.

Applanat, M. P., H. Buteau-Lozano, M. A. Herve and A. Corpet (2008). "Vascular endothelial growth factor is a target gene for estrogen receptor and contributes to breast cancer progression." Adv Exp Med Biol **617**: 437-444.

Arao, Y., R. Kuriyama, F. Kayama and S. Kato (2000). "A nuclear matrix-associated factor, SAF-B, interacts with specific isoforms of AUF1/hnRNP D." Arch Biochem Biophys **380**(2): 228-236.

Aravind, L. and E. V. Koonin (2000). "SAP - a putative DNA-binding motif involved in chromosomal organization." Trends Biochem Sci **25**(3): 112-114.

Arihiro, K., H. Oda, M. Kaneko and K. Inai (2000). "Cytokines facilitate chemotactic motility of breast carcinoma cells." Breast Cancer **7**(3): 221-230.

Armenante, F., M. Merola, A. Furia, M. Tovey and M. Palmieri (1999). "Interleukin-6 repression is associated with a distinctive chromatin structure of the gene." Nucleic Acids Res **27**(22): 4483-4490.

Arnold, S. F., J. D. Obourn, H. Jaffe and A. C. Notides (1994). "Serine 167 is the major estradiol-induced phosphorylation site on the human estrogen receptor." Mol Endocrinol **8**(9): 1208-1214.

Asgeirsson, K. S., K. Olafsdottir, J. G. Jonasson and H. M. Ogmundsdottir (1998). "The effects of IL-6 on cell adhesion and e-cadherin expression in breast cancer." Cytokine **10**(9): 720-728.

Auger, A. P., J. M. Meredith, G. L. Snyder and J. D. Blaustein (2001). "Oestradiol increases phosphorylation of a dopamine- and cyclic AMP-regulated phosphoprotein (DARPP-32) in female rat brain." J Neuroendocrinol **13**(9): 761-768.

Auricchio, F., A. Migliaccio, M. Di Domenico and E. Nola (1987). "Oestradiol stimulates tyrosine phosphorylation and hormone binding activity of its own receptor in a cell-free system." EMBO J **6**(10): 2923-2929.

Ayoubi, T. A. and W. J. Van De Ven (1996). "Regulation of gene expression by alternative promoters." FASEB J **10**(4): 453-460.

Azuma, K., T. Urano, K. Horie-Inoue, S. Hayashi, R. Sakai, Y. Ouchi and S. Inoue (2009). "Association of estrogen receptor alpha and histone deacetylase 6 causes rapid deacetylation of tubulin in breast cancer cells." Cancer Res **69**(7): 2935-2940.

Baer, R. and T. Ludwig (2002). "The BRCA1/BARD1 heterodimer, a tumor suppressor complex with ubiquitin E3 ligase activity." Curr Opin Genet Dev **12**(1): 86-91.

Bajic, V. B., S. L. Tan, A. Chong, S. Tang, A. Strom, J. A. Gustafsson, C. Y. Lin and E. T. Liu (2003). "Dragon ERE Finder version 2: A tool for accurate detection and analysis of estrogen response elements in vertebrate genomes." Nucleic Acids Res **31**(13): 3605-3607.

Barros, F. F., D. G. Powe, I. O. Ellis and A. R. Green (2010). "Understanding the HER family in breast cancer: interaction with ligands, dimerization and treatments." Histopathology **56**(5): 560-572.

Bateman, A. C. (2006). "Pathology of breast cancer." Women's Health Medicine **3**(1): 18-21.

Bautista, S., H. Valles, R. L. Walker, S. Anzick, R. Zeillinger, P. Meltzer and C. Theillet (1998). "In breast cancer, amplification of the steroid receptor coactivator gene AIB1 is correlated with estrogen and progesterone receptor positivity." Clin Cancer Res **4**(12): 2925-2929.

Benaroudj, N. (2005). "[Proteasome, a cellular machinery for protein degradation]." Med Sci (Paris) **21**(2): 115-116.

Berg, J. M., J. L. Tymoczko and L. Stryer (2002). Protein Turnover and Amino Acid Catabolism. Biochemistry. New York, W H Freeman.

Bergman, A., F. Abel, A. Behboudi, M. Yhr, J. Mattsson, J. H. Svensson, P. Karlsson and M. Nordling (2008). "No germline mutations in supposed tumour suppressor genes SAFB1 and SAFB2 in familial breast cancer with linkage to 19p." BMC Med Genet **9**: 108.

Berns, E. M., I. L. van Staveren, J. G. Klijn and J. A. Foekens (1998). "Predictive value of SRC-1 for tamoxifen response of recurrent breast cancer." Breast Cancer Res Treat **48**(1): 87-92.

Best, A., C. Dagliesh, I. Ehrmann, M. Kheirollahi-Kouhestani, A. Tyson-Capper and D. J. Elliott (2013). "Expression of Tra2β in Cancer Cells as a Potential Contributory Factor to Neoplasia and Metastasis." International Journal of Cell Biology **Volume 2013**.

Bettuzzi, S., F. Scorcioni, S. Astancolle, P. Davalli, M. Scaltriti and A. Corti (2002). "Clusterin (SGP-2) transient overexpression decreases proliferation rate of SV40-immortalized human prostate epithelial cells by slowing down cell cycle progression." Oncogene **21**(27): 4328-4334.

Bhat-Nakshatri, P., R. A. Campbell, N. M. Patel, T. R. Newton, A. J. King, M. S. Marshall, S. Ali and H. Nakshatri (2004). "Tumour necrosis factor and PI3-kinase control oestrogen receptor alpha protein level and its transrepression function." Br J Cancer **90**(4): 853-859.

Birney, E., S. Kumar and A. R. Krainer (1993). "Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors." Nucleic Acids Res **21**(25): 5803-5816.

Bjornstrom, L. and M. Sjoberg (2005). "Mechanisms of estrogen receptor signaling: convergence of genomic and nongenomic actions on target genes." Mol Endocrinol **19**(4): 833-842.

Bjornstrom, L. and M. Sjoberg (2004). "Estrogen receptor-dependent activation of AP-1 via non-genomic signalling." Nucl Recept **2**(1): 3.

Blanks, R. G., S. M. Moss, C. E. McGahan, M. J. Quinn and P. J. Babb (2000). "Effect of NHS breast screening programme on mortality from breast cancer in England and Wales, 1990-8: comparison of observed with predicted mortality." BMJ **321**(7262): 665-669.

Blencowe, B. J. (2000). "Exonic splicing enhancers: mechanism of action, diversity and role in human genetic diseases." Trends Biochem Sci **25**(3): 106-110.

Bocchinfuso, W. P. and K. S. Korach (1997). "Mammary gland development and tumorigenesis in estrogen receptor knockout mice." J Mammary Gland Biol Neoplasia **2**(4): 323-334.

Bon, G., S. E. Di Carlo, V. Folgiero, P. Avetrani, C. Lazzari, G. D'Orazi, M. F. Brizzi, A. Sacchi, S. Soddu, G. Blandino, M. Mottolese and R. Falcioni (2009). "Negative regulation of beta4 integrin transcription by homeodomain-interacting protein kinase 2 and p53 impairs tumor progression." Cancer Res **69**(14): 5978-5986.

Breast Cancer Review. (2010). "Treatment options."

breastcancer.org. "Diagnosis of DCIS." Types of breast cancer, from http://www.breastcancer.org/symptoms/types/dcis/diagnosis.

Busch, H., R. Reddy, L. Rothblum and Y. C. Choi (1982). "SnRNAs, SnRNPs, and RNA processing." Annu Rev Biochem **51**: 617-654.

Busund, L. T., E. Richardsen, R. Busund, T. Ukkonen, T. Bjornsen, C. Busch and H. Stalsberg (2005). "Significant expression of IGFBP2 in breast cancer compared with benign lesions." J Clin Pathol **58**(4): 361-366.

Cailleau, R., M. Olive and Q. V. Cruciger (1978). "Long-term human breast carcinoma cell lines of metastatic origin: preliminary characterization." In Vitro **14**(11): 911-915.

Cairns, P., T. J. Polascik, Y. Eby, K. Tokino, J. Califano, A. Merlo, L. Mao, J. Herath, R. Jenkins, W. Westra and et al. (1995). "Frequency of homozygous deletion at p16/CDKN2 in primary human tumours." Nat Genet **11**(2): 210-212.

Cancer Research UK (2012). Breast Cancer - risk factors. http://www.cancerresearchuk.org/cancer-info/cancerstats/types/breast/risk factors/breast-cancer-risk-factors#source3.

Cancer Research UK (2012). What the breast screening review means, http://www.cancerresearchuk.org/cancer-help/about-cancer/cancer-questions/what-the-breast-screening-review-means.

Cancer Research UK (2010b). Breast cancer survival statistics. http://www.cancerresearchuk.org/cancer-info/cancerstats/types/breast/survival/.

Cancer Research UK (2010a). Breast cancer incidence statistics. http://www.cancerresearchuk.org/cancer-info/cancerstats/types/breast/.

Castronuevo, P., M. A. Thornton, L. E. McCarthy, J. Klimas and B. P. Schick (2003). "DNase I hypersensitivity patterns of the serglycin proteoglycan gene in resting and phorbol 12-myristate 13-acetate-stimulated human erythroleukemia (HEL), CHRF 288-11, and HL-60 cells compared with neutrophils and human umbilical vein endothelial cells." J Biol Chem **278**(49): 48704-48712.

Catalano, S., I. Barone, C. Giordano, P. Rizza, H. Qi, G. Gu, R. Malivindi, D. Bonofiglio and S. Ando (2009). "Rapid estradiol/ERalpha signaling enhances aromatase enzymatic activity in breast cancer cells." Mol Endocrinol **23**(10): 1634-1645.

Chan, C. W., Y. B. Lee, J. Uney, A. Flynn, J. H. Tobias and M. Norman (2007). "A novel member of the SAF (scaffold attachment factor)-box protein family inhibits gene expression and induces apoptosis." Biochem J **407**(3): 355-362.

Chang, J. C. and S. G. Hilsenbeck (2010). Chapter 31: Prognostic and Predictive Markers. Diseases of the Breast. J. R. Harris, M. E. Lippman, M. Morrow and C. K. Osborne, Lippincott Williams & Wilkins.

Chen, J. D. and R. M. Evans (1995). "A transcriptional co-repressor that interacts with nuclear hormone receptors." Nature **377**(6548): 454-457.

Chiodi, I., M. Biggiogera, M. Denegri, M. Corioni, F. Weighardt, F. Cobianchi, S. Riva and G. Biamonti (2000). "Structure and dynamics of hnRNP-labelled nuclear bodies induced by stress treatments." J Cell Sci **113 (Pt 22)**: 4043-4053.

Chiu, J. J., M. K. Sgagias and K. H. Cowan (1996). "Interleukin 6 acts as a paracrine growth factor in human mammary carcinoma cell lines." Clin Cancer Res **2**(1): 215-221.

Cho, H. and B. S. Katzenellenbogen (1993). "Synergistic activation of estrogen receptor-mediated transcription by estradiol and protein kinase activators." Mol Endocrinol **7**(3): 441-452.

Clarke, A. S., M. M. Lotz and A. M. Mercurio (1994). "A novel structural variant of the human beta 4 integrin cDNA." Cell Adhes Commun **2**(1): 1-6.

Clemmons, D. R., C. Camacho-Hubner, E. Coronado and C. K. Osborne (1990). "Insulin-like growth factor binding protein secretion by breast carcinoma cell lines: correlation with estrogen receptor status." Endocrinology **127**(6): 2679-2686.

Clery, A., S. Jayne, N. Benderska, C. Dominguez, S. Stamm and F. H. Allain (2011). "Molecular basis of purine-rich RNA recognition by the human SR-like protein Tra2-beta1." Nat Struct Mol Biol **18**(4): 443-450.

Cole, M. P., C. T. Jones and I. D. Todd (1971). "A new anti-oestrogenic agent in late breast cancer. An early clinical appraisal of ICI46474." Br J Cancer **25**(2): 270-275.

Collaborative Group on Hormonal Factors in Breast Cancer (2002). "Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50302 women with breast cancer and 96973 women without the disease." Lancet **360**(9328): 187-195.

Collaborative Group on Hormonal Factors in Breast Cancer (2001). "Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease." Lancet **358**(9291): 1389-1399.

Collaborative Group on Hormonal Factors in Breast Cancer (1997). "Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer. Collaborative Group on Hormonal Factors in Breast Cancer." Lancet **350**(9084): 1047-1059.

Collaborative Group on Hormonal Factors in Breast Cancer (1996). "Breast cancer and hormonal contraceptives: collaborative reanalysis of individual data on 53 297 women with breast cancer and 100 239 women without breast cancer from 54 epidemiological studies. Collaborative Group on Hormonal Factors in Breast Cancer." Lancet **347**(9017): 1713-1727.

Colwill, K., T. Pawson, B. Andrews, J. Prasad, J. L. Manley, J. C. Bell and P. I. Duncan (1996). "The Clk/Sty protein kinase phosphorylates SR splicing factors and regulates their intranuclear distribution." EMBO J **15**(2): 265-275.

Conze, D., L. Weiss, P. S. Regen, A. Bhushan, D. Weaver, P. Johnson and M. Rincon (2001). "Autocrine production of interleukin 6 causes multidrug resistance in breast cancer cells." Cancer Res **61**(24): 8851-8858.

Cooper, A. P. (1840). The Anatomy of the Breast. London, Longman.

Cork, D. M., T. W. Lennard and A. J. Tyson-Capper (2012). "Progesterone receptor (PR) variants exist in breast cancer cells characterised as PR negative." Tumour Biol **33**(6): 2329-2340.

Criswell, T., D. Klokov, M. Beman, J. P. Lavik and D. A. Boothman (2003). "Repression of IR-inducible clusterin expression by the p53 tumor suppressor protein." Cancer Biol Ther **2**(4): 372-380.

Cuzick, J., T. Powles, U. Veronesi, J. Forbes, R. Edwards, S. Ashley and P. Boyle (2003). "Overview of the main outcomes in breast-cancer prevention trials." Lancet **361**(9354): 296-300.

Dace, A., L. Zhao, K. S. Park, T. Furuno, N. Takamura, M. Nakanishi, B. L. West, J. A. Hanover and S. Cheng (2000). "Hormone binding induces rapid proteasome-mediated degradation of thyroid hormone receptors." Proc Natl Acad Sci U S A **97**(16): 8985-8990.

Darby, S., P. McGale, C. Correa, C. Taylor, R. Arriagada, M. Clarke, D. Cutter, C. Davies, M. Ewertz, J. Godwin, R. Gray, L. Pierce, T. Whelan, Y. Wang and R. Peto (2011). "Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: meta-analysis of individual patient data for 10,801 women in 17 randomised trials." Lancet **378**(9804): 1707-1716.

Dauvois, S., R. White and M. G. Parker (1993). "The antiestrogen ICI 182780 disrupts estrogen receptor nucleocytoplasmic shuttling." J Cell Sci **106 (Pt 4)**: 1377-1388.

Davuluri, R. V., Y. Suzuki, S. Sugano, C. Plass and T. H. Huang (2008). "The functional consequences of alternative promoter use in mammalian genomes." Trends Genet **24**(4): 167-177.

de Melker, A. A. and A. Sonnenberg (1999). "Integrins: alternative splicing as a mechanism to regulate ligand binding and integrin signaling events." Bioessays **21**(6): 499-509.

de Silva, H. V., W. D. Stuart, Y. B. Park, S. J. Mao, C. M. Gil, J. R. Wetterau, S. J. Busch and J. A. Harmony (1990). "Purification and characterization of apolipoprotein J." J Biol Chem **265**(24): 14292-14297.

Debril, M. B., L. Dubuquoy, J. N. Feige, W. Wahli, B. Desvergne, J. Auwerx and L. Gelman (2005). "Scaffold attachment factor B1 directly interacts with nuclear receptors in living cells and represses transcriptional activity." J Mol Endocrinol **35**(3): 503-517.

Delage-Mourroux, R., P. G. Martini, I. Choi, D. M. Kraichely, J. Hoeksema and B. S. Katzenellenbogen (2000). "Analysis of estrogen receptor interaction with a repressor of estrogen receptor activity (REA) and the regulation of estrogen receptor transcriptional activity by REA." J Biol Chem **275**(46): 35848-35856.

Dery, M. C., V. Leblanc, C. Shooner and E. Asselin (2003). "Regulation of Akt expression and phosphorylation by 17beta-estradiol in the rat uterus during estrous cycle." Reprod Biol Endocrinol **1**: 47.

Diaz, L. K., M. Cristofanilli, X. Zhou, K. L. Welch, T. L. Smith, Y. Yang, N. Sneige, A. A. Sahin and M. Z. Gilcrease (2005). "Beta4 integrin subunit gene expression correlates with tumor size and nuclear grade in early breast cancer." Mod Pathol **18**(9): 1165-1175.

Dillon, D. A., A. J. Guidi and S. J. Schnitt (2010). Chapter 28: Pathology of Invasive Breast Cancer. Diseases of the Breast. J. R. Harris, M. E. Lippman, M. Morrow and C. K. Osborne, Lippincott Williams & Wilkins.

Ding, X. F., C. M. Anderson, H. Ma, H. Hong, R. M. Uht, P. J. Kushner and M. R. Stallcup (1998). "Nuclear receptor-binding sites of coactivators glucocorticoid receptor interacting protein 1 (GRIP1) and steroid receptor coactivator 1 (SRC-1): multiple motifs with different binding specificities." Mol Endocrinol **12**(2): 302-313.

Dobrzycka, K. M., K. Kang, S. Jiang, R. Meyer, P. H. Rao, A. V. Lee and S. Oesterreich (2006). "Disruption of scaffold attachment factor B1 leads to TBX2 up-regulation, lack of p19ARF induction, lack of senescence, and cell immortalization." Cancer Res **66**(16): 7859-7863.

Draheim, K. M., H. B. Chen, Q. Tao, N. Moore, M. Roche and S. Lyle (2010). "ARRDC3 suppresses breast cancer progression by negatively regulating integrin beta4." Oncogene **29**(36): 5032-5047.

Dubois, V., D. Couissi, E. Schonne, C. Remacle and A. Trouet (1995). "Intracellular levels and secretion of insulin-like-growth-factor-binding proteins in MCF-7/6, MCF-7/AZ and MDA-MB-231 breast cancer cells. Differential modulation by estrogens in serum-free medium." Eur J Biochem **232**(1): 47-53.

DuPont, B. R., D. K. Garcia, T. M. Sullivan, S. L. Naylor and S. Oesterreich (1997). "Assignment of SAFB encoding Hsp27 ERE-TATA binding protein (HET)/scaffold attachment factor B (SAF-B) to human chromosome 19 band p13." Cytogenet Cell Genet **79**(3-4): 284-285.

Edwards, D. P. (2005). "Regulation of signal transduction pathways by estrogen and progesterone." Annu Rev Physiol **67**: 335-376.

Elliott, D. J., A. Best, C. Dalgliesh, I. Ehrmann and S. Grellscheid (2012). "How does Tra2beta protein regulate tissue-specific RNA splicing?" Biochem Soc Trans **40**(4): 784-788.

Encarnacion, C. A., D. R. Ciocca, W. L. McGuire, G. M. Clark, S. A. Fuqua and C. K. Osborne (1993). "Measurement of steroid hormone receptors in breast cancer patients on tamoxifen." Breast Cancer Res Treat **26**(3): 237-246.

Faggioli, L., C. Costanzo, M. Merola, E. Bianchini, A. Furia, A. Carsana and M. Palmieri (1996). "Nuclear factor kappa B (NF-kappa B), nuclear factor interleukin-6 (NFIL-6 or C/EBP beta) and nuclear factor interleukin-6 beta (NFIL6-beta or C/EBP delta) are not sufficient to activate the endogenous interleukin-6 gene in the human breast carcinoma cell line MCF-7. Comparative analysis with MDA-MB-231 cells, an interleukin-6-expressing human breast carcinoma cell line." Eur J Biochem **239**(3): 624-631.

Fawell, S. E., R. White, S. Hoare, M. Sydenham, M. Page and M. G. Parker (1990). "Inhibition of estrogen receptor-DNA binding by the "pure" antiestrogen ICI 164,384 appears to be mediated by impaired receptor dimerization." Proc Natl Acad Sci U S A **87**(17): 6883-6887.

Fernando, R. I. and J. Wimalasena (2004). "Estradiol abrogates apoptosis in breast cancer cells through inactivation of BAD: Ras-dependent nongenomic pathways requiring signaling through ERK and Akt." Mol Biol Cell **15**(7): 3266-3284.

Filardo, E. J., C. T. Graeber, J. A. Quinn, M. B. Resnick, D. Giri, R. A. DeLellis, M. M. Steinhoff and E. Sabo (2006). "Distribution of GPR30, a seven membrane-spanning estrogen receptor, in primary breast cancer and its association with clinicopathologic determinants of tumor progression." Clin Cancer Res **12**(21): 6359-6366.

Filardo, E. J., J. A. Quinn, K. I. Bland and A. R. Frackelton, Jr. (2000). "Estrogen-induced activation of Erk-1 and Erk-2 requires the G protein-coupled receptor homolog, GPR30, and occurs via trans-activation of the epidermal growth factor receptor through release of HB-EGF." Mol Endocrinol **14**(10): 1649-1660.

Filardo, E. J., J. A. Quinn, A. R. Frackelton, Jr. and K. I. Bland (2002). "Estrogen action via the G protein-coupled receptor, GPR30: stimulation of adenylyl cyclase and cAMP-mediated attenuation of the epidermal growth factor receptor-to-MAPK signaling axis." Mol Endocrinol **16**(1): 70-84.

Fischer, D. C., K. Noack, I. B. Runnebaum, D. O. Watermann, D. G. Kieback, S. Stamm and E. Stickeler (2004). "Expression of splicing factors in human ovarian cancer." Oncol Rep **11**(5): 1085-1090.

Flouriot, G., H. Brand, S. Denger, R. Metivier, M. Kos, G. Reid, V. Sonntag-Buck and F. Gannon (2000). "Identification of a new isoform of the human estrogen receptor-alpha (hER-alpha) that is encoded by distinct transcripts and that is able to repress hER-alpha activation function 1." EMBO J **19**(17): 4688-4700.

Forrest, P. (1986). Breast Cancer Screening: Report to the Health Minister of England, Wales, Scotland and Northern Ireland. London, Her Majesty's Stationery Office.

Foulstone, E. J., L. Zeng, C. M. Perks and J. M. Holly (2013). "Insulin-Like Growth Factor Binding Protein 2 (IGFBP-2) Promotes Growth and Survival of Breast Epithelial Cells: Novel Regulation of the Estrogen Receptor." Endocrinology **154**(5): 1780-1793.

Fu, X. D. and T. Maniatis (1990). "Factor required for mammalian spliceosome assembly is localized to discrete regions in the nucleus." Nature **343**(6257): 437-441.

Fuqua, S. A., S. D. Fitzgerald, D. C. Allred, R. M. Elledge, Z. Nawaz, D. P. McDonnell, B. W. O'Malley, G. L. Greene and W. L. McGuire (1992). "Inhibition of estrogen receptor action by a naturally occurring variant in human breast tumors." Cancer Res **52**(2): 483-486.

Gabriel, B., A. Zur Hausen, J. Bouda, L. Boudova, M. Koprivova, M. Hirschfeld, M. Jager and E. Stickeler (2009). "Significance of nuclear hTra2-beta1 expression in cervical cancer." Acta Obstet Gynecol Scand **88**(2): 216-221.

Garee, J. P., R. Meyer and S. Oesterreich (2011). "Co-repressor activity of scaffold attachment factor B1 requires sumoylation." Biochem Biophys Res Commun.

Garee, J. P. and S. Oesterreich (2010). "SAFB1's multiple functions in biological control-lots still to be done!" J Cell Biochem **109**(2): 312-319.

Gee, J. M., M. E. Harper, I. R. Hutcheson, T. A. Madden, D. Barrow, J. M. Knowlden, R. A. McClelland, N. Jordan, A. E. Wakeling and R. I. Nicholson (2003). "The antiepidermal growth factor receptor agent gefitinib (ZD1839/Iressa) improves antihormone response and prevents development of resistance in breast cancer in vitro." Endocrinology **144**(11): 5105-5117.

Geng, Y., S. Chandrasekaran, J. W. Hsu, M. Gidwani, A. D. Hughes and M. R. King (2013). "Phenotypic switch in blood: effects of pro-inflammatory cytokines on breast cancer cell aggregation and adhesion." PLoS One **8**(1): e54959.

Gluch, A., M. Vidakovic and J. Bode (2008). "Scaffold/matrix attachment regions (S/MARs): relevance for disease and therapy." Handb Exp Pharmacol(186): 67-103.

Gougelet, A., C. Bouclier, V. Marsaud, S. Maillard, S. O. Mueller, K. S. Korach and J. M. Renoir (2005). "Estrogen receptor alpha and beta subtype expression and transactivation capacity are differentially affected by receptor-, hsp90- and immunophilin-ligands in human breast cancer cells." J Steroid Biochem Mol Biol **94**(1-3): 71-81.

Green, S., P. Walter, V. Kumar, A. Krust, J. M. Bornert, P. Argos and P. Chambon (1986). "Human oestrogen receptor cDNA: sequence, expression and homology to v-erb-A." Nature **320**(6058): 134-139.

Greene, G. L., P. Gilna, M. Waterfield, A. Baker, Y. Hort and J. Shine (1986). "Sequence and expression of human estrogen receptor complementary DNA." Science **231**(4742): 1150-1154.

Grellscheid, S., C. Dalgliesh, M. Storbeck, A. Best, Y. Liu, M. Jakubik, Y. Mende, I. Ehrmann, T. Curk, K. Rossbach, C. F. Bourgeois, J. Stevenin, D. Grellscheid, M. S. Jackson, B. Wirth and D. J. Elliott (2011). "Identification of evolutionarily conserved exons as regulated targets for the splicing activator tra2beta in development." PLoS Genet **7**(12): e1002390.

Gudas, J. M., H. Nguyen, T. Li and K. H. Cowan (1995). "Hormone-dependent regulation of BRCA1 in human breast cancer cells." Cancer Res **55**(20): 4561-4565.

Guo, W., Y. Pylayeva, A. Pepe, T. Yoshioka, W. J. Muller, G. Inghirami and F. G. Giancotti (2006). "Beta 4 integrin amplifies ErbB2 signaling to promote mammary tumorigenesis." Cell **126**(3): 489-502.

Gutierrez, M. C., S. Detre, S. Johnston, S. K. Mohsin, J. Shou, D. C. Allred, R. Schiff, C. K. Osborne and M. Dowsett (2005). "Molecular changes in tamoxifen-resistant breast cancer: relationship between estrogen receptor, HER-2, and p38 mitogen-activated protein kinase." J Clin Oncol **23**(11): 2469-2476.

Hall, L. L., K. P. Smith, M. Byron and J. B. Lawrence (2006). "Molecular anatomy of a speckle." Anat Rec A Discov Mol Cell Evol Biol **288**(7): 664-675.

Hammerich-Hille, S., V. J. Bardout, S. G. Hilsenbeck, C. K. Osborne and S. Oesterreich (2009). "Low SAFB levels are associated with worse outcome in breast cancer patients." Breast Cancer Res Treat **121**(2): 503-509.

Hammerich-Hille, S., B. A. Kaipparettu, A. Tsimelzon, C. J. Creighton, S. Jiang, J. M. Polo, A. Melnick, R. Meyer and S. Oesterreich (2010). "SAFB1 mediates repression of immune regulators and apoptotic genes in breast cancer cells." J Biol Chem **285**(6): 3608-3616.

Hanahan, D. and R. A. Weinberg (2011). "Hallmarks of cancer: the next generation." Cell **144**(5): 646-674.

Hao, S. and D. Baltimore (2009). "The stability of mRNA influences the temporal order of the induction of genes encoding inflammatory molecules." Nat Immunol **10**(3): 281-288.

Hartmann, L. C., T. A. Sellers, M. H. Frost, W. L. Lingle, A. C. Degnim, K. Ghosh, R. A. Vierkant, S. D. Maloney, V. S. Pankratz, D. W. Hillman, V. J. Suman, J. Johnson, C. Blake, T. Tlsty, C. M. Vachon, L. J. Melton, 3rd and D. W. Visscher (2005). "Benign breast disease and the risk of breast cancer." N Engl J Med **353**(3): 229-237.

Hashimoto, T., K. Matsuda and M. Kawata (2012). "Scaffold attachment factor B (SAFB)1 and SAFB2 cooperatively inhibit the intranuclear mobility and function of ERalpha." J Cell Biochem **113**(9): 3039-3050.

Hassiotou, F. and D. Geddes (2012). "Anatomy of the human mammary gland: Current status of knowledge." Clin Anat.

Haybittle, J. L., R. W. Blamey, C. W. Elston, J. Johnson, P. J. Doyle, F. C. Campbell, R. I. Nicholson and K. Griffiths (1982). "A prognostic index in primary breast cancer." Br J Cancer **45**(3): 361-366.

He, Y. Y., B. Cai, Y. X. Yang, X. L. Liu and X. P. Wan (2009). "Estrogenic G protein-coupled receptor 30 signaling is involved in regulation of endometrial carcinoma by promoting proliferation, invasion potential, and interleukin-6 secretion via the MEK/ERK mitogen-activated protein kinase pathway." Cancer Sci **100**(6): 1051-1061.

Heilmann, A. M. and N. J. Dyson (2012). "Phosphorylation puts the pRb tumor suppressor into shape." Genes Dev **26**(11): 1128-1130.

Herman, J. G., A. Merlo, L. Mao, R. G. Lapidus, J. P. Issa, N. E. Davidson, D. Sidransky and S. B. Baylin (1995). "Inactivation of the CDKN2/p16/MTS1 gene is frequently associated with aberrant DNA methylation in all common human cancers." Cancer Res **55**(20): 4525-4530.

Hermanson, O., C. K. Glass and M. G. Rosenfeld (2002). "Nuclear receptor coregulators: multiple modes of modification." Trends Endocrinol Metab **13**(2): 55-60.

Hong, E. A., H. L. Gautrey, D. J. Elliott and A. J. Tyson-Capper (2012). "SAFB1- and SAFB2-mediated transcriptional repression: relevance to cancer." Biochem Soc Trans **40**(4): 826-830.

Hong, E. A. L., D. J. Elliott and A. J. Tyson-Capper (2010). "Effects of Oestrogen on RNA-binding Splicing Regulators in Breast Cancer Cells." Reproductive Sciences **17**(3 (Supplement)): 109A-110A.

Honma, S., K. Shimodaira, Y. Shimizu, N. Tsuchiya, H. Saito, T. Yanaihara and T. Okai (2002). "The influence of inflammatory cytokines on estrogen production and cell proliferation in human breast cancer cells." Endocr J **49**(3): 371-377.

Horlein, A. J., A. M. Naar, T. Heinzel, J. Torchia, B. Gloss, R. Kurokawa, A. Ryan, Y. Kamei, M. Soderstrom, C. K. Glass and et al. (1995). "Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor." Nature **377**(6548): 397-404.

Horner-Glister, E., M. Maleki-Dizaji, C. J. Guerin, S. M. Johnson, J. Styles and I. N. White (2005). "Influence of oestradiol and tamoxifen on oestrogen receptors-alpha and -beta protein degradation and non-genomic signalling pathways in uterine and breast carcinoma cells." J Mol Endocrinol **35**(3): 421-432.

Hoskins, J. M., L. A. Carey and H. L. McLeod (2009). "CYP2D6 and tamoxifen: DNA matters in breast cancer." Nat Rev Cancer **9**(8): 576-586.

Hovey, R. C., J. F. Trott and B. K. Vonderhaar (2002). "Establishing a framework for the functional mammary gland: from endocrinology to morphology." J Mammary Gland Biol Neoplasia **7**(1): 17-38.

Hoyt, M. A. (1997). "Eliminating all obstacles: regulated proteolysis in the eukaryotic cell cycle." Cell **91**(2): 149-151.

Hu, X. and M. A. Lazar (1999). "The CoRNR motif controls the recruitment of corepressors by nuclear hormone receptors." Nature **402**(6757): 93-96.

Huang, H. J., J. D. Norris and D. P. McDonnell (2002). "Identification of a negative regulatory surface within estrogen receptor alpha provides evidence in support of a role for corepressors in regulating cellular responses to agonists and antagonists." Mol Endocrinol **16**(8): 1778-1792.

Hui, R., R. D. Macmillan, F. S. Kenny, E. A. Musgrove, R. W. Blamey, R. I. Nicholson, J. F. Robertson and R. L. Sutherland (2000). "INK4a gene expression and methylation in primary breast cancer: overexpression of p16INK4a messenger RNA is a marker of poor prognosis." Clin Cancer Res **6**(7): 2777-2787.

Hultborn, K. A., L. G. Larsson and I. Ragnhult (1955). "The lymph drainage from the breast to the axillary and parasternal lymph nodes, studied with the aid of colloidal Au198." Acta radiol **43**(1): 52-64.

Hurley, W. L. (1989). "Mammary gland function during involution." J Dairy Sci **72**(6): 1637-1646.

Hutchinson, J. N., A. W. Ensminger, C. M. Clemson, C. R. Lynch, J. B. Lawrence and A. Chess (2007). "A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains." BMC Genomics **8**: 39.

Hutson, S. W., P. N. Cowen and C. C. Bird (1985). "Morphometric studies of age related changes in normal human breast and their significance for evolution of mammary cancer." J Clin Pathol **38**(3): 281-287.

Hyder, S. M. and G. M. Stancel (1999). "Regulation of angiogenic growth factors in the female reproductive tract by estrogens and progestins." Mol Endocrinol **13**(6): 806-811.

Ivanova, M., K. M. Dobrzycka, S. Jiang, K. Michaelis, R. Meyer, K. Kang, B. Adkins, O. A. Barski, S. Zubairy, J. Divisova, A. V. Lee and S. Oesterreich (2005). "Scaffold attachment factor B1 functions in development, growth, and reproduction." Mol Cell Biol **25**(8): 2995-3006.

Jakacka, M., M. Ito, J. Weiss, P. Y. Chien, B. D. Gehm and J. L. Jameson (2001). "Estrogen receptor binding to DNA is not required for its activity through the nonclassical AP1 pathway." J Biol Chem **276**(17): 13615-13621.

Jiang, S., R. Meyer, K. Kang, C. K. Osborne, J. Wong and S. Oesterreich (2006). "Scaffold attachment factor SAFB1 suppresses estrogen receptor alpha-mediated transcription in part via interaction with nuclear receptor corepressor." Mol Endocrinol **20**(2): 311-320.

Johnson, E. S. (2004). "Protein modification by SUMO." Annu Rev Biochem **73**: 355-382.

Juncker-Jensen, A., A. E. Lykkesfeldt, J. Worm, U. Ralfkiaer, U. Espelund and J. S. Jepsen (2006). "Insulin-like growth factor binding protein 2 is a marker for antiestrogen resistant human breast cancer cell lines but is not a major growth regulator." Growth Horm IGF Res **16**(4): 224-239.

Kahlert, S., S. Nuedling, M. van Eickels, H. Vetter, R. Meyer and C. Grohe (2000). "Estrogen receptor alpha rapidly activates the IGF-1 receptor pathway." J Biol Chem **275**(24): 18447-18453.

Kang, L., X. Zhang, Y. Xie, Y. Tu, D. Wang, Z. Liu and Z. Y. Wang (2010). "Involvement of estrogen receptor variant ER-alpha36, not GPR30, in nongenomic estrogen signaling." Mol Endocrinol **24**(4): 709-721.

Kato, S., H. Endoh, Y. Masuhiro, T. Kitamoto, S. Uchiyama, H. Sasaki, S. Masushige, Y. Gotoh, E. Nishida, H. Kawashima, D. Metzger and P. Chambon (1995). "Activation of the estrogen receptor through phosphorylation by mitogen-activated protein kinase." Science **270**(5241): 1491-1494.

Katzenellenbogen, J. A., B. W. O'Malley and B. S. Katzenellenbogen (1996). "Tripartite steroid hormone receptor pharmacology: interaction with multiple effector sites as a basis for the cell- and promoter-specific action of these hormones." Mol Endocrinol **10**(2): 119-131.

Kawagoe, J., M. Ohmichi, T. Takahashi, C. Ohshima, S. Mabuchi, K. Takahashi, H. Igarashi, A. Mori-Abe, M. Saitoh, B. Du, T. Ohta, A. Kimura, S. Kyo, M. Inoue and H. Kurachi (2003). "Raloxifene inhibits estrogen-induced up-regulation of telomerase activity in a human breast cancer cell line." J Biol Chem **278**(44): 43363-43372.

Keshamouni, V. G., R. R. Mattingly and K. B. Reddy (2002). "Mechanism of 17-beta-estradiol-induced Erk1/2 activation in breast cancer cells. A role for HER2 AND PKC-delta." J Biol Chem **277**(25): 22558-22565.

Key, T., P. Appleby, I. Barnes and G. Reeves (2002). "Endogenous sex hormones and breast cancer in postmenopausal women: reanalysis of nine prospective studies." J Natl Cancer Inst **94**(8): 606-616.

Key, T. J. and P. K. Verkasalo (1999). "Endogenous hormones and the aetiology of breast cancer." Breast Cancer Res **1**(1): 18-21.

Kim, I., A. Manni, J. Lynch and J. M. Hammond (1991). "Identification and regulation of insulin-like growth factor binding proteins produced by hormone-dependent and -independent human breast cancer cell lines." Mol Cell Endocrinol **78**(1-2): 71-78.

King, W. J. and G. L. Greene (1984). "Monoclonal antibodies localize oestrogen receptor in the nuclei of target cells." Nature **307**(5953): 745-747.

Knowlden, J. M., I. R. Hutcheson, H. E. Jones, T. Madden, J. M. Gee, M. E. Harper, D. Barrow, A. E. Wakeling and R. I. Nicholson (2003). "Elevated levels of epidermal growth factor receptor/c-erbB2 heterodimers mediate an autocrine growth regulatory pathway in tamoxifen-resistant MCF-7 cells." Endocrinology **144**(3): 1032-1044.

Knupfer, H. and R. Preiss (2007). "Significance of interleukin-6 (IL-6) in breast cancer (review)." Breast Cancer Res Treat **102**(2): 129-135.

Koch, F., F. Jourquin, P. Ferrier and J. C. Andrau (2008). "Genome-wide RNA polymerase II: not genes only!" Trends Biochem Sci **33**(6): 265-273.

Kolset, S. O. and H. Tveit (2008). "Serglycin--structure and biology." Cell Mol Life Sci **65**(7-8): 1073-1085.

Konig, J., K. Zarnack, G. Rot, T. Curk, M. Kayikci, B. Zupan, D. J. Turner, N. M. Luscombe and J. Ule (2011). "iCLIP--transcriptome-wide mapping of protein-RNA interactions with individual nucleotide resolution." J Vis Exp(50).

Konig, J., K. Zarnack, G. Rot, T. Curk, M. Kayikci, B. Zupan, D. J. Turner, N. M. Luscombe and J. Ule (2010). "iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution." Nat Struct Mol Biol **17**(7): 909-915.

Korzus, E., J. Torchia, D. W. Rose, L. Xu, R. Kurokawa, E. M. McInerney, T. M. Mullen, C. K. Glass and M. G. Rosenfeld (1998). "Transcription factor-specific requirements for coactivators and their acetyltransferase functions." Science **279**(5351): 703-707.

Kos, M., G. Reid, S. Denger and F. Gannon (2001). "Minireview: genomic organization of the human ERalpha gene promoter region." Mol Endocrinol **15**(12): 2057-2063.

Krege, J. H., J. B. Hodgin, J. F. Couse, E. Enmark, M. Warner, J. F. Mahler, M. Sar, K. S. Korach, J. A. Gustafsson and O. Smithies (1998). "Generation and reproductive phenotypes of mice lacking estrogen receptor beta." Proc Natl Acad Sci U S A **95**(26): 15677-15682.

Kumar, V., S. Green, G. Stack, M. Berry, J. R. Jin and P. Chambon (1987). "Functional domains of the human estrogen receptor." Cell **51**(6): 941-951.

Kurebayashi, J., T. Otsuki, H. Kunisue, K. Tanaka, S. Yamamoto and H. Sonoo (2000). "Expression levels of estrogen receptor-alpha, estrogen receptor-beta, coactivators, and corepressors in breast cancer." Clin Cancer Res **6**(2): 512-518.

Kushner, P. J., D. A. Agard, G. L. Greene, T. S. Scanlan, A. K. Shiau, R. M. Uht and P. Webb (2000). "Estrogen receptor pathways to AP-1." J Steroid Biochem Mol Biol **74**(5): 311-317.

Lanz, R. B., N. J. McKenna, S. A. Onate, U. Albrecht, J. Wong, S. Y. Tsai, M. J. Tsai and B. W. O'Malley (1999). "A steroid receptor coactivator, SRA, functions as an RNA and is present in an SRC-1 complex." Cell **97**(1): 17-27.

Lapensee, E. W., T. R. Tuttle, S. R. Fox and N. Ben-Jonathan (2009). "Bisphenol A at low nanomolar doses confers chemoresistance in estrogen receptor-alpha-positive and -negative breast cancer cells." Environ Health Perspect **117**(2): 175-180.

Lavinsky, R. M., K. Jepsen, T. Heinzel, J. Torchia, T. M. Mullen, R. Schiff, A. L. Del-Rio, M. Ricote, S. Ngo, J. Gemsch, S. G. Hilsenbeck, C. K. Osborne, C. K. Glass, M. G. Rosenfeld and D. W. Rose (1998). "Diverse signaling pathways modulate nuclear receptor recruitment of N-CoR and SMRT complexes." Proc Natl Acad Sci U S A **95**(6): 2920-2925.

Le Romancer, M., C. Poulard, P. Cohen, S. Sentis, J. M. Renoir and L. Corbo (2011). "Cracking the estrogen receptor's posttranslational code in breast tumors." Endocr Rev **32**(5): 597-622.

Lee, C., L. Atanelov, B. Modrek and Y. Xing (2003). "ASAP: the Alternative Splicing Annotation Project." Nucleic Acids Res **31**(1): 101-105.

Lee, L. M., J. Cao, H. Deng, P. Chen, Z. Gatalica and Z. Y. Wang (2008). "ER-alpha36, a novel variant of ER-alpha, is expressed in ER-positive and -negative human breast carcinomas." Anticancer Res **28**(1B): 479-483.

Lee, Y. B., S. Colley, M. Norman, G. Biamonti and J. B. Uney (2007). "SAFB re-distribution marks steps of the apoptotic process." Exp Cell Res **313**(18): 3914-3923.

Leo, C. and J. D. Chen (2000). "The SRC family of nuclear receptor coactivators." Gene **245**(1): 1-11.

Leskov, K. S., D. Y. Klokov, J. Li, T. J. Kinsella and D. A. Boothman (2003). "Synthesis and functional analyses of nuclear clusterin, a cell death protein." J Biol Chem **278**(13): 11590-11600.

Li, H., C. Leo, D. J. Schroen and J. D. Chen (1997). "Characterization of receptor interaction and transcriptional repression by the corepressor SMRT." Mol Endocrinol **11**(13): 2025-2037.

Li, J., I. C. Hawkins, C. D. Harvey, J. L. Jennings, A. J. Link and J. G. Patton (2003). "Regulation of alternative splicing by SRrp86 and its interacting proteins." Mol Cell Biol **23**(21): 7437-7447.

Li, J., L. Jia, P. Zhao, Y. Jiang, S. Zhong and D. Chen (2012). "Stable knockdown of clusterin by vectorbased RNA interference in a human breast cancer cell line inhibits tumour cell invasion and metastasis." J Int Med Res **40**(2): 545-555.

Li, L., T. Feng, Y. Lian, G. Zhang, A. Garen and X. Song (2009). "Role of human noncoding RNAs in the control of tumorigenesis." Proc Natl Acad Sci U S A **106**(31): 12956-12961.

Li, X. J., C. K. Ong, Y. Cao, Y. Q. Xiang, J. Y. Shao, A. Ooi, L. X. Peng, W. H. Lu, Z. Zhang, D. Petillo, L. Qin, Y. N. Bao, F. J. Zheng, C. S. Chia, N. G. Iyer, T. B. Kang, Y. X. Zeng, K. C. Soo, J. M. Trent, B. T. Teh and C. N. Qian (2011). "Serglycin is a theranostic target in nasopharyngeal carcinoma that promotes metastasis." Cancer Res **71**(8): 3162-3172.

Licatalosi, D. D., A. Mele, J. J. Fak, J. Ule, M. Kayikci, S. W. Chi, T. A. Clark, A. C. Schweitzer, J. E. Blume, X. Wang, J. C. Darnell and R. B. Darnell (2008). "HITS-CLIP yields genome-wide insights into brain alternative RNA processing." Nature **456**(7221): 464-469.

Lichtenstein, P., N. V. Holm, P. K. Verkasalo, A. Iliadou, J. Kaprio, M. Koskenvuo, E. Pukkala, A. Skytthe and K. Hemminki (2000). "Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland." N Engl J Med **343**(2): 78-85.

Liggett, W. H., Jr. and D. Sidransky (1998). "Role of the p16 tumor suppressor gene in cancer." J Clin Oncol **16**(3): 1197-1206.

Lin, J., P. Xu, P. LaVallee and J. R. Hoidal (2008). "Identification of proteins binding to E-Box/Ku86 sites and function of the tumor suppressor SAFB1 in transcriptional regulation of the human xanthine oxidoreductase gene." J Biol Chem **283**(44): 29681-29689.

Lindholm, C. K., J. D. Frantz, S. E. Shoelson and M. Welsh (2000). "Shf, a Shb-like adapter protein, is involved in PDGF-alpha-receptor regulation of apoptosis." Biochem Biophys Res Commun **278**(3): 537-543.

Lipscomb, E. A. and A. M. Mercurio (2005). "Mobilization and activation of a signaling competent alpha6beta4integrin underlies its contribution to carcinoma progression." Cancer Metastasis Rev **24**(3): 413-423.

Lipscomb, E. A., K. J. Simpson, S. R. Lyle, J. E. Ring, A. S. Dugan and A. M. Mercurio (2005). "The alpha6beta4 integrin maintains the survival of human breast carcinoma cells in vivo." Cancer Res **65**(23): 10970-10976.

Lonard, D. M. and W. O'Malley B (2007). "Nuclear receptor coregulators: judges, juries, and executioners of cellular regulation." Mol Cell **27**(5): 691-700.

Lonard, D. M. and B. W. O'Malley (2006). "The expanding cosmos of nuclear receptor coactivators." Cell **125**(3): 411-414.

Long, J. C. and J. F. Caceres (2009). "The SR protein family of splicing factors: master regulators of gene expression." Biochem J **417**(1): 15-27.

Lowry, O. H., N. J. Rosebrough, A. L. Farr and R. J. Randall (1951). "Protein measurement with the Folin phenol reagent." J Biol Chem **193**(1): 265-275.

Lu, S., K. Simin, A. Khan and A. M. Mercurio (2008). "Analysis of integrin beta4 expression in human breast cancer: association with basal-like tumors and prognostic significance." Clin Cancer Res **14**(4): 1050-1058.

Lucas, T. F., C. Royer, E. R. Siu, M. F. Lazari and C. S. Porto (2010). "Expression and signaling of G protein-coupled estrogen receptor 1 (GPER) in rat sertoli cells." Biol Reprod **83**(2): 307-317.

MacGregor, J. I. and V. C. Jordan (1998). "Basic guide to the mechanisms of antiestrogen action." Pharmacol Rev **50**(2): 151-196.

Maggiolini, M., A. Vivacqua, G. Fasanella, A. G. Recchia, D. Sisci, V. Pezzi, D. Montanaro, A. M. Musti, D. Picard and S. Ando (2004). "The G protein-coupled receptor GPR30 mediates c-fos up-regulation by 17beta-estradiol and phytoestrogens in breast cancer cells." J Biol Chem **279**(26): 27008-27016.

Malkin, D., F. P. Li, L. C. Strong, J. F. Fraumeni, Jr., C. E. Nelson, D. H. Kim, J. Kassel, M. A. Gryka, F. Z. Bischoff, M. A. Tainsky and et al. (1990). "Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms." Science **250**(4985): 1233-1238.

Marquez, D. C., H. W. Chen, E. M. Curran, W. V. Welshons and R. J. Pietras (2006). "Estrogen receptors in membrane lipid rafts and signal transduction in breast cancer." Mol Cell Endocrinol **246**(1-2): 91-100.

Masuda, K., B. Ripley, R. Nishimura, T. Mino, O. Takeuchi, G. Shioi, H. Kiyonari and T. Kishimoto (2013). "Arid5a controls IL-6 mRNA stability, which contributes to elevation of IL-6 level in vivo." Proc Natl Acad Sci U S A **110**(23): 9409-9414.

Mattick, J. S. and I. V. Makunin (2006). "Non-coding RNA." Hum Mol Genet **15 Spec No 1**: R17-29.

mayoclinic.com. "Slide show: Stages of breast cancer." Mayo Clinic, from http://www.mayoclinic.com/health/stage-of-breast-cancer/BR00011&slide=2.

McClelland, R. A., D. Barrow, T. A. Madden, C. M. Dutkowski, J. Pamment, J. M. Knowlden, J. M. Gee and R. I. Nicholson (2001). "Enhanced epidermal growth factor receptor signaling in MCF7 breast cancer cells after long-term culture in the presence of the pure antiestrogen ICI 182,780 (Faslodex)." Endocrinology **142**(7): 2776-2788.

McCormack, V. A. and I. dos Santos Silva (2006). "Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis." Cancer Epidemiol Biomarkers Prev **15**(6): 1159-1169.

McKenna, N. J., R. B. Lanz and B. W. O'Malley (1999). "Nuclear receptor coregulators: cellular and molecular biology." Endocr Rev **20**(3): 321-344.

McKenna, N. J. and B. W. O'Malley (2002). "Combinatorial control of gene expression by nuclear receptors and coregulators." Cell **108**(4): 465-474.

McPherson, K., C. M. Steel and J. M. Dixon (2000). "ABC of breast diseases. Breast cancer-epidemiology, risk factors, and genetics." BMJ **321**(7261): 624-628.

Menasce, L. P., G. R. White, C. J. Harrison and J. M. Boyle (1993). "Localization of the estrogen receptor locus (ESR) to chromosome 6q25.1 by FISH and a simple post-FISH banding technique." Genomics **17**(1): 263-265.

Merlo, A., J. G. Herman, L. Mao, D. J. Lee, E. Gabrielson, P. C. Burger, S. B. Baylin and D. Sidransky (1995). "5' CpG island methylation is associated with transcriptional silencing of the tumour suppressor p16/CDKN2/MTS1 in human cancers." Nat Med **1**(7): 686-692.

Migliaccio, A., M. Di Domenico, G. Castoria, A. de Falco, P. Bontempo, E. Nola and F. Auricchio (1996). "Tyrosine kinase/p21ras/MAP-kinase pathway activation by estradiol-receptor complex in MCF-7 cells." EMBO J **15**(6): 1292-1300.

Miller, B. J., D. Wang, R. Krahe and F. A. Wright (2003). "Pooled analysis of loss of heterozygosity in breast cancer: a genome scan provides comparative evidence for multiple tumor suppressors and identifies novel candidate regions." Am J Hum Genet **73**(4): 748-767.

Mireuta, M., A. Darnel and M. Pollak (2010). "IGFBP-2 expression in MCF-7 cells is regulated by the PI3K/AKT/mTOR pathway through Sp1-induced increase in transcription." Growth Factors **28**(4): 243-255.

Miyake, H., I. Hara, M. E. Gleave and H. Eto (2004). "Protection of androgen-dependent human prostate cancer cells from oxidative stress-induced DNA damage by overexpression of clusterin and its modulation by androgen." Prostate **61**(4): 318-323.

Monje, P., S. Zanello, M. Holick and R. Boland (2001). "Differential cellular localization of estrogen receptor alpha in uterine and mammary cells." Mol Cell Endocrinol **181**(1-2): 117-129.

Monroe, D. G., B. J. Getz, S. A. Johnsen, B. L. Riggs, S. Khosla and T. C. Spelsberg (2003). "Estrogen receptor isoform-specific regulation of endogenous gene expression in human osteoblastic cell lines expressing either ERalpha or ERbeta." J Cell Biochem **90**(2): 315-326.

Montano, M. M., K. Ekena, R. Delage-Mourroux, W. Chang, P. Martini and B. S. Katzenellenbogen (1999). "An estrogen receptor-selective coregulator that potentiates the effectiveness of antiestrogens and represses the activity of estrogens." Proc Natl Acad Sci U S A **96**(12): 6947-6952.

Mueller, M. D., J. L. Vigne, A. Minchenko, D. I. Lebovic, D. C. Leitman and R. N. Taylor (2000). "Regulation of vascular endothelial growth factor (VEGF) gene transcription by estrogen receptors alpha and beta." Proc Natl Acad Sci U S A **97**(20): 10972-10977.

Murphy, L. C., S. L. Simon, A. Parkes, E. Leygue, H. Dotzlaw, L. Snell, S. Troup, A. Adeyinka and P. H. Watson (2000). "Altered expression of estrogen receptor coregulators during human breast tumorigenesis." Cancer Res **60**(22): 6266-6271.

Naccarato, A. G., P. Viacava, S. Vignati, G. Fanelli, A. G. Bonadio, G. Montruccoli and G. Bevilacqua (2000). "Bio-morphological events in the development of the human female mammary gland from fetal age to puberty." Virchows Arch **436**(5): 431-438.

Nawaz, Z., D. M. Lonard, A. P. Dennis, C. L. Smith and B. W. O'Malley (1999). "Proteasome-dependent degradation of the human estrogen receptor." Proc Natl Acad Sci U S A **96**(5): 1858-1862.

Nayler, O., W. Stratling, J. P. Bourquin, I. Stagljar, L. Lindemann, H. Jasper, A. M. Hartmann, F. O. Fackelmayer, A. Ullrich and S. Stamm (1998). "SAF-B protein couples transcription and pre-mRNA splicing to SAR/MAR elements." Nucleic Acids Res **26**(15): 3542-3549.

Nelen, M. R., G. W. Padberg, E. A. Peeters, A. Y. Lin, B. van den Helm, R. R. Frants, V. Coulon, A. M. Goldstein, M. M. van Reen, D. F. Easton, R. A. Eeles, S. Hodgsen, J. J. Mulvihill, V. A. Murday, M. A. Tucker, E. C. Mariman, T. M. Starink, B. A. Ponder, H. H. Ropers, H. Kremer, M. Longy and C. Eng (1996). "Localization of the gene for Cowden disease to chromosome 10q22-23." Nat Genet **13**(1): 114-116.

Nicodemus, C. F., S. Avraham, K. F. Austen, S. Purdy, J. Jablonski and R. L. Stevens (1990). "Characterization of the human gene that encodes the peptide core of secretory granule proteoglycans in promyelocytic leukemia HL-60 cells and analysis of the translated product." J Biol Chem **265**(10): 5889-5896.

Nikolakaki, E., R. Kohen, A. M. Hartmann, S. Stamm, E. Georgatsou and T. Giannakouros (2001). "Cloning and characterization of an alternatively spliced form of SR protein kinase 1 that interacts specifically with scaffold attachment factor-B." J Biol Chem **276**(43): 40175-40182.

Nishi, H., K. Hashimoto and A. R. Panchenko (2011). "Phosphorylation in protein-protein binding: effect on stability and function." Structure **19**(12): 1807-1815.

Niu, Z., X. Li, B. Hu, R. Li, L. Wang, L. Wu and X. Wang (2012). "Small interfering RNA targeted to secretory clusterin blocks tumor growth, motility, and invasion in breast cancer." Acta Biochim Biophys Sin (Shanghai) **44**(12): 991-998.

Nobori, T., K. Miura, D. J. Wu, A. Lois, K. Takabayashi and D. A. Carson (1994). "Deletions of the cyclin-dependent kinase-4 inhibitor gene in multiple human cancers." Nature **368**(6473): 753-756.

Oesterreich, S. (2003). "Scaffold attachment factors SAFB1 and SAFB2: Innocent bystanders or critical players in breast tumorigenesis?" J Cell Biochem **90**(4): 653-661.

Oesterreich, S., D. C. Allredl, S. K. Mohsin, Q. Zhang, H. Wong, A. V. Lee, C. K. Osborne and P. O'Connell (2001). "High rates of loss of heterozygosity on chromosome 19p13 in human breast cancer." Br J Cancer **84**(4): 493-498.

Oesterreich S, K. K., Townson SM, Clark GM, Hilsenbeck SG, Osborne CK, Bardou V (2002). "Critical balance of scaffold attachment factor SAFB levels play important role in breast tumor suppression." 25th Annual San Antonio Breast Cancer Symposium.

Oesterreich, S., A. V. Lee, T. M. Sullivan, S. K. Samuel, J. R. Davie and S. A. Fuqua (1997). "Novel nuclear matrix protein HET binds to and influences activity of the HSP27 promoter in human breast cancer cells." J Cell Biochem **67**(2): 275-286.

Oesterreich, S., Q. Zhang, T. Hopp, S. A. Fuqua, M. Michaelis, H. H. Zhao, J. R. Davie, C. K. Osborne and A. V. Lee (2000). "Tamoxifen-bound estrogen receptor (ER) strongly interacts with the nuclear matrix protein HET/SAF-B, a novel inhibitor of ER-mediated transactivation." Mol Endocrinol **14**(3): 369-381.

Oesterreich, S., Q. P. Zhang and A. V. Lee (2000). "Inhibition of oestrogen receptor activity by the co-repressor HET/SAF-B is relieved by blockade of histone deacetylase activity." Eur J Cancer **36 Suppl 4**: S43-44.

Ohira, M., A. Morohashi, H. Inuzuka, T. Shishikura, T. Kawamoto, H. Kageyama, Y. Nakamura, E. Isogai, H. Takayasu, S. Sakiyama, Y. Suzuki, S. Sugano, T. Goto, S. Sato and A. Nakagawara (2003). "Expression profiling and characterization of 4200 genes cloned from primary neuroblastomas: identification of 305 genes differentially expressed between favorable and unfavorable subsets." Oncogene **22**(35): 5525-5536.

Omura, Y., Y. Nishio, T. Takemoto, C. Ikeuchi, O. Sekine, K. Morino, Y. Maeno, T. Obata, S. Ugi, H. Maegawa, H. Kimura and A. Kashiwagi (2009). "SAFB1, an RBMX-binding protein, is a newly identified regulator of hepatic SREBP-1c gene." BMB Rep **42**(4): 232-237.

Osborne, C. K., V. Bardou, T. A. Hopp, G. C. Chamness, S. G. Hilsenbeck, S. A. Fuqua, J. Wong, D. C. Allred, G. M. Clark and R. Schiff (2003). "Role of the estrogen receptor coactivator AIB1 (SRC-3) and HER-2/neu in tamoxifen resistance in breast cancer." J Natl Cancer Inst **95**(5): 353-361.

Osborne, C. K. and R. Schiff (2011). "Mechanisms of endocrine resistance in breast cancer." Annu Rev Med **62**: 233-247.

Osborne, M. P. (2004). Breast Anatomy and Development. Diseases of the Breast. J. R. Harris, M. E. Lippman, M. Morrow and C. K. Osborne. Philadelphia, Lippincott Williams & Wilkins. **3rd edition**.

Ouyang, Y. Q., A. zur Hausen, M. Orlowska-Volk, M. Jager, H. Bettendorf, M. Hirschfeld, X. W. Tong and E. Stickeler (2011). "Expression levels of hnRNP G and hTra2-beta1 correlate with opposite outcomes in endometrial cancer biology." Int J Cancer **128**(9): 2010-2019.

Page, D. L., C. M. Steel and J. M. Dixon (1995). "ABC of breast diseases. Carcinoma in situ and patients at high risk of breast cancer." BMJ **310**(6971): 39-42.

Palanisamy, V., A. Jakymiw, E. A. Van Tubergen, N. J. D'Silva and K. L. Kirkwood (2012). "Control of cytokine mRNA expression by RNA-binding proteins and microRNAs." J Dent Res **91**(7): 651-658.

Pang, W. W. and P. E. Hartmann (2007). "Initiation of human lactation: secretory differentiation and secretory activation." J Mammary Gland Biol Neoplasia **12**(4): 211-221.

Parker, L. P., D. D. Taylor, S. Kesterson and C. Gercel-Taylor (2009). "Gene expression profiling in response to estradiol and genistein in ovarian cancer cells." Cancer Genomics Proteomics **6**(3): 189-194.

Parker, M. G. (1993). "Action of "pure" antiestrogens in inhibiting estrogen receptor action." Breast Cancer Res Treat **26**(2): 131-137.

Pedram, A., M. Razandi, A. J. Evinger, E. Lee and E. R. Levin (2009). "Estrogen inhibits ATR signaling to cell cycle checkpoints and DNA repair." Mol Biol Cell **20**(14): 3374-3389.

Pedram, A., M. Razandi and E. R. Levin (2006). "Nature of functional estrogen receptors at the plasma membrane." Mol Endocrinol **20**(9): 1996-2009.

Peekhaus, N. T., T. Chang, E. C. Hayes, H. A. Wilkinson, S. W. Mitra, J. M. Schaeffer and S. P. Rohrer (2004). "Distinct effects of the antiestrogen Faslodex on the stability of estrogen receptors-alpha and -beta in the breast cancer cell line MCF-7." J Mol Endocrinol **32**(3): 987-995.

Peidis, P., N. Voukkalis, E. Aggelidou, E. Georgatsou, M. Hadzopoulou-Cladaras, R. E. Scott, E. Nikolakaki and T. Giannakouros (2011). "SAFB1 interacts with and suppresses the transcriptional activity of p53." FEBS Lett **585**(1): 78-84.

Pelekanou, V., G. Notas, M. Kampa, E. Tsentelierou, J. Radojicic, G. Leclercq, E. Castanas and E. N. Stathopoulos (2012). "ERalpha36, a new variant of the ERalpha is expressed in triple negative breast carcinomas and has a specific transcriptomic signature in breast cancer cell lines." Steroids **77**(10): 928-934.

Perez, D. S., T. R. Hoage, J. R. Pritchett, A. L. Ducharme-Smith, M. L. Halling, S. C. Ganapathiraju, P. S. Streng and D. I. Smith (2008). "Long, abundantly expressed non-coding transcripts are altered in cancer." Hum Mol Genet **17**(5): 642-655.

Perissi, V., K. Jepsen, C. K. Glass and M. G. Rosenfeld (2010). "Deconstructing repression: evolving models of co-repressor action." Nat Rev Genet **11**(2): 109-123.

Perissi, V., L. M. Staszewski, E. M. McInerney, R. Kurokawa, A. Krones, D. W. Rose, M. H. Lambert, M. V. Milburn, C. K. Glass and M. G. Rosenfeld (1999). "Molecular determinants of nuclear receptor-corepressor interaction." Genes Dev **13**(24): 3198-3208.

Peto, J., N. Collins, R. Barfoot, S. Seal, W. Warren, N. Rahman, D. F. Easton, C. Evans, J. Deacon and M. R. Stratton (1999). "Prevalence of BRCA1 and BRCA2 gene mutations in patients with early-onset breast cancer." J Natl Cancer Inst **91**(11): 943-949.

Pharoah, P. D., N. E. Day, S. Duffy, D. F. Easton and B. A. Ponder (1997). "Family history and the risk of breast cancer: a systematic review and meta-analysis." Int J Cancer **71**(5): 800-809.

Pink, J. J. and V. C. Jordan (1996). "Models of estrogen receptor regulation by estrogens and antiestrogens in breast cancer cell lines." Cancer Res **56**(10): 2321-2330.

Poola, I., S. Koduri, S. Chatra and R. Clarke (2000). "Identification of twenty alternatively spliced estrogen receptor alpha mRNAs in breast cancer cell lines and tumors using splice targeted primer approach." J Steroid Biochem Mol Biol **72**(5): 249-258.

Pottratz, S. T., T. Bellido, H. Mocharla, D. Crabb and S. C. Manolagas (1994). "17 beta-Estradiol inhibits expression of human interleukin-6 promoter-reporter constructs by a receptor-dependent mechanism." J Clin Invest **93**(3): 944-950.

Prasanth, K. V., M. Camiolo, G. Chan, V. Tripathi, L. Denis, T. Nakamura, M. R. Hubner and D. L. Spector (2010). "Nuclear organization and dynamics of 7SK RNA in regulating gene expression." Mol Biol Cell **21**(23): 4184-4196.

Pratt, W. B. and D. O. Toft (1997). "Steroid receptor interactions with heat shock protein and immunophilin chaperones." Endocr Rev **18**(3): 306-360.

Ptacek, J. and M. Snyder (2006). "Charging it up: global analysis of protein phosphorylation." Trends Genet **22**(10): 545-554.

Pucci, S., E. Bonanno, F. Pichiorri, C. Angeloni and L. G. Spagnoli (2004). "Modulation of different clusterin isoforms in human colon tumorigenesis." Oncogene **23**(13): 2298-2304.

Quaedackers, M. E., C. E. van den Brink, P. T. van der Saag and L. G. Tertoolen (2007). "Direct interaction between estrogen receptor alpha and NF-kappaB in the nucleus of living cells." Mol Cell Endocrinol **273**(1-2): 42-50.

Rao, J., X. Jiang, Y. Wang and B. Chen (2011). "Advances in the understanding of the structure and function of ER-alpha36,a novel variant of human estrogen receptor-alpha." J Steroid Biochem Mol Biol **127**(3-5): 231-237.

Rappsilber, J., U. Ryder, A. I. Lamond and M. Mann (2002). "Large-scale proteomic analysis of the human spliceosome." Genome Res **12**(8): 1231-1245.

Razandi, M., A. Pedram, I. Merchenthaler, G. L. Greene and E. R. Levin (2004). "Plasma membrane estrogen receptors exist and functions as dimers." Mol Endocrinol **18**(12): 2854-2865.

Rebas, E., L. Lachowicz and A. Lachowicz (2005). "Estradiol modulates the synapsins phosphorylation by various protein kinases in the rat brain under in vitro and in vivo conditions." J Physiol Pharmacol **56**(1): 39-48.

Reddy, K. B., G. Jin, M. C. Karode, J. A. Harmony and P. H. Howe (1996). "Transforming growth factor beta (TGF beta)-induced nuclear localization of apolipoprotein J/clusterin in epithelial cells." Biochemistry **35**(19): 6157-6163.

Redondo, M., E. Villar, J. Torres-Munoz, T. Tellez, M. Morell and C. K. Petito (2000). "Overexpression of clusterin in human breast carcinoma." Am J Pathol **157**(2): 393-399.

Ree, A. H., B. F. Landmark, W. Eskild, F. O. Levy, H. Lahooti, T. Jahnsen, A. Aakvaag and V. Hansson (1989). "Autologous down-regulation of messenger ribonucleic acid and protein levels for estrogen receptors in MCF-7 cells: an inverse correlation to progesterone receptor levels." Endocrinology **124**(5): 2577-2583.

Renz, A. and F. O. Fackelmayer (1996). "Purification and molecular cloning of the scaffold attachment factor B (SAF-B), a novel human nuclear protein that specifically binds to S/MAR-DNA." Nucleic Acids Res **24**(5): 843-849.

Rhodes, L. V., S. P. Short, N. F. Neel, V. A. Salvo, Y. Zhu, S. Elliott, Y. Wei, D. Yu, M. Sun, S. E. Muir, J. P. Fonseca, M. R. Bratton, C. Segar, S. L. Tilghman, T. Sobolik-Delmaire, L. W. Horton, S. Zaja-Milatovic, B. M. Collins-Burow, S. Wadsworth, B. S. Beckman, C. E. Wood, S. A. Fuqua, K. P. Nephew, P. Dent, R. A. Worthylake, T. J. Curiel, M. C. Hung, A. Richmond and M. E. Burow (2011). "Cytokine receptor CXCR4 mediates estrogen-independent tumorigenesis, metastasis, and resistance to endocrine therapy in human breast cancer." Cancer Res **71**(2): 603-613.

Riggins, R. B., R. S. Schrecengost, M. S. Guerrero and A. H. Bouton (2007). "Pathways to tamoxifen resistance." Cancer Lett **256**(1): 1-24.

Rizzolio, F., C. Lucchetti, I. Caligiuri, I. Marchesi, M. Caputo, A. J. Klein-Szanto, L. Bagella, M. Castronovo and A. Giordano (2012). "Retinoblastoma tumor-suppressor protein phosphorylation and inactivation depend on direct interaction with Pin1." Cell Death Differ **19**(7): 1152-1161.

Romagnolo, D., L. A. Annab, T. E. Thompson, J. I. Risinger, L. A. Terry, J. C. Barrett and C. A. Afshari (1998). "Estrogen upregulation of BRCA1 expression with no effect on localization." Mol Carcinog **22**(2): 102-109.

Russo, J. and I. H. Russo (2004). "Development of the human breast." Maturitas **49**(1): 2-15.

Russo, J. and I. H. Russo (1987). Development of the human mammary gland. The mammary gland. M. C. Neville and C. W. Daniel. New York, Plenum**:** 67-93.

Saceda, M., M. E. Lippman, P. Chambon, R. L. Lindsey, M. Ponglikitmongkol, M. Puente and M. B. Martin (1988). "Regulation of the estrogen receptor in MCF-7 cells by estradiol." Mol Endocrinol **2**(12): 1157-1162.

Sanford, J. R., P. Coutinho, J. A. Hackett, X. Wang, W. Ranahan and J. F. Caceres (2008). "Identification of nuclear and cytoplasmic mRNA targets for the shuttling protein SF2/ASF." PLoS One **3**(10): e3369.

Sanford, J. R., X. Wang, M. Mort, N. Vanduyn, D. N. Cooper, S. D. Mooney, H. J. Edenberg and Y. Liu (2009). "Splicing factor SFRS1 recognizes a functionally diverse landscape of RNA transcripts." Genome Res **19**(3): 381-394.

Sasieni, P. D., J. Shelton, N. Ormiston-Smith, C. S. Thomson and P. B. Silcocks (2011). "What is the lifetime risk of developing cancer?: the effect of adjusting for multiple primaries." Br J Cancer **105**(3): 460-465.

Saville, B., M. Wormke, F. Wang, T. Nguyen, E. Enmark, G. Kuiper, J. A. Gustafsson and S. Safe (2000). "Ligand-, cell-, and estrogen receptor subtype (alpha/beta)-dependent activation at GC-rich (Sp1) promoter elements." J Biol Chem **275**(8): 5379-5387.

Schiff, R., P. Reddy, M. Ahotupa, E. Coronado-Heinsohn, M. Grim, S. G. Hilsenbeck, R. Lawrence, S. Deneke, R. Herrera, G. C. Chamness, S. A. Fuqua, P. H. Brown and C. K. Osborne (2000). "Oxidative stress and AP-1 activity in tamoxifen-resistant breast tumors in vivo." J Natl Cancer Inst **92**(23): 1926-1934.

Sergeant, K. A., C. F. Bourgeois, C. Dalgliesh, J. P. Venables, J. Stevenin and D. J. Elliott (2007). "Alternative RNA splicing complexes containing the scaffold attachment factor SAFB2." J Cell Sci **120**(Pt 2): 309-319.

Serrano, M., H. Lee, L. Chin, C. Cordon-Cardo, D. Beach and R. A. DePinho (1996). "Role of the INK4a locus in tumor suppression and cell mortality." Cell **85**(1): 27-37.

Shang, Y. and M. Brown (2002). "Molecular determinants for the tissue specificity of SERMs." Science **295**(5564): 2465-2468.

Shannan, B., M. Seifert, K. Leskov, J. Willis, D. Boothman, W. Tilgen and J. Reichrath (2006). "Challenge and promise: roles for clusterin in pathogenesis, progression and therapy of cancer." Cell Death Differ **13**(1): 12-19.

Shapiro, G. I. and B. J. Rollins (1996). "p16INK4A as a human tumor suppressor." Biochim Biophys Acta **1242**(3): 165-169.

Shen, W. H., J. H. Zhou, S. R. Broussard, G. G. Freund, R. Dantzer and K. W. Kelley (2002). "Proinflammatory cytokines block growth of breast cancer cells by impairing signals from a growth factor receptor." Cancer Res **62**(16): 4746-4756.

Shi, L., B. Dong, Z. Li, Y. Lu, T. Ouyang, J. Li, T. Wang, Z. Fan, T. Fan, B. Lin, Z. Wang and Y. Xie (2009). "Expression of ER-{alpha}36, a novel variant of estrogen receptor {alpha}, and resistance to tamoxifen treatment in breast cancer." J Clin Oncol **27**(21): 3423-3429.

Shou, J., S. Massarweh, C. K. Osborne, A. E. Wakeling, S. Ali, H. Weiss and R. Schiff (2004). "Mechanisms of tamoxifen resistance: increased estrogen receptor-HER2/neu cross-talk in ER/HER2-positive breast cancer." J Natl Cancer Inst **96**(12): 926-935.

Simon, S. L., A. Parkes, E. Leygue, H. Dotzlaw, L. Snell, S. Troup, A. Adeyinka, P. H. Watson and L. C. Murphy (2000). "Expression of a repressor of estrogen receptor activity in human breast tumors: relationship to some known prognostic markers." Cancer Res **60**(11): 2796-2799.

Simoncini, T., A. Hafezi-Moghadam, D. P. Brazil, K. Ley, W. W. Chin and J. K. Liao (2000). "Interaction of oestrogen receptor with the regulatory subunit of phosphatidylinositol-3-OH kinase." Nature **407**(6803): 538-541.

Singer, G. A., J. Wu, P. Yan, C. Plass, T. H. Huang and R. V. Davuluri (2008). "Genome-wide analysis of alternative promoters of human genes using a custom promoter tiling array." BMC Genomics **9**: 349.

Singh, R. R. and R. Kumar (2005). "Steroid hormone receptor signaling in tumorigenesis." J Cell Biochem **96**(3): 490-505.

Smith, C. L., Z. Nawaz and B. W. O'Malley (1997). "Coactivator and corepressor regulation of the agonist/antagonist activity of the mixed antiestrogen, 4-hydroxytamoxifen." Mol Endocrinol **11**(6): 657-666.

So, A. I., R. J. Levitt, B. Eigl, L. Fazli, M. Muramaki, S. Leung, M. C. Cheang, T. O. Nielsen, M. Gleave and M. Pollak (2008). "Insulin-like growth factor binding protein-2 is a novel therapeutic target associated with breast cancer." Clin Cancer Res **14**(21): 6944-6954.

Song, M., K. Hakala, S. T. Weintraub and Y. Shiio (2011). "Quantitative proteomic identification of the BRCA1 ubiquitination substrates." J Proteome Res **10**(11): 5191-5198.

Soule, H. D., J. Vazguez, A. Long, S. Albert and M. Brennan (1973). "A human cell line from a pleural effusion derived from a breast carcinoma." J Natl Cancer Inst **51**(5): 1409-1416.

Sporn, M. B. (1976). "Approaches to prevention of epithelial cancer during the preneoplastic period." Cancer Res **36**(7 PT 2): 2699-2702.

Sternglanz, R. (1996). "Histone acetylation: a gateway to transcriptional activation." Trends Biochem Sci **21**(10): 357-358.

Stewart, J., R. King, J. Hayward and R. Rubens (1982). "Estrogen and progesterone receptors: correlation of response rates, site and timing of receptor analysis." Breast Cancer Res Treat **2**(3): 243-250.

Stoilov, P., R. Daoud, O. Nayler and S. Stamm (2004). "Human tra2-beta1 autoregulates its protein concentration by influencing alternative splicing of its pre-mRNA." Hum Mol Genet **13**(5): 509-524.

Strnad, V., G. Hildebrandt, R. Potter, J. Hammer, M. Hindemith, A. Resch, K. Spiegl, M. Lotter, W. Uter, M. Bani, R. D. Kortmann, M. W. Beckmann, R. Fietkau and O. J. Ott (2011). "Accelerated partial breast irradiation: 5-year results of the German-Austrian multicenter phase II trial using interstitial multicatheter brachytherapy alone after breast-conserving surgery." Int J Radiat Oncol Biol Phys **80**(1): 17-24.

Su, Q., S. Hu, H. Gao, R. Ma, Q. Yang, Z. Pan, T. Wang and F. Li (2008). "Role of AIB1 for tamoxifen resistance in estrogen receptor-positive breast cancer cells." Oncology **75**(3-4): 159-168.

Swift, M., D. Morrell, R. B. Massey and C. L. Chase (1991). "Incidence of cancer in 161 families affected by ataxia-telangiectasia." N Engl J Med **325**(26): 1831-1836.

Tacke, R., M. Tohyama, S. Ogawa and J. L. Manley (1998). "Human Tra2 proteins are sequence-specific activators of pre-mRNA splicing." Cell **93**(1): 139-148.

Tai, H. H., M. Geisterfer, J. C. Bell, M. Moniwa, J. R. Davie, L. Boucher and M. W. McBurney (2003). "CHD1 associates with NCoR and histone deacetylase as well as with RNA splicing proteins." Biochem Biophys Res Commun **308**(1): 170-176.

Takagi, D., Y. Tatsumi, T. Yokochi, A. Takatori, M. Ohira, T. Kamijo, S. Kondo, Y. Fujii and A. Nakagawara (2013). "Novel adaptor protein Shf interacts with ALK receptor and negatively regulates its downstream signals in neuroblastoma." Cancer Sci **104**(5): 563-572.

Tamura, R. N., C. Rozzo, L. Starr, J. Chambers, L. F. Reichardt, H. M. Cooper and V. Quaranta (1990). "Epithelial integrin alpha 6 beta 4: complete primary structure of alpha 6 and variant forms of beta 4." J Cell Biol **111**(4): 1593-1604.

Tee, M. K., I. Rogatsky, C. Tzagarakis-Foster, A. Cvoro, J. An, R. J. Christy, K. R. Yamamoto and D. C. Leitman (2004). "Estradiol and selective estrogen receptor modulators differentially regulate target genes with estrogen receptors alpha and beta." Mol Biol Cell **15**(3): 1262-1272.

Thomas, P., Y. Pang, E. J. Filardo and J. Dong (2005). "Identity of an estrogen membrane receptor coupled to a G protein in human breast cancer cells." Endocrinology **146**(2): 624-632.

Townson, S. M., K. M. Dobrzycka, A. V. Lee, M. Air, W. Deng, K. Kang, S. Jiang, N. Kioka, K. Michaelis and S. Oesterreich (2003). "SAFB2, a new scaffold attachment factor homolog and estrogen receptor corepressor." J Biol Chem **278**(22): 20059-20068.

Townson, S. M., K. Kang, A. V. Lee and S. Oesterreich (2004). "Structure-function analysis of the estrogen receptor alpha corepressor scaffold attachment factor-B1: identification of a potent transcriptional repression domain." J Biol Chem **279**(25): 26074-26081.

Townson, S. M., T. Sullivan, Q. Zhang, G. M. Clark, C. K. Osborne, A. V. Lee and S. Oesterreich (2000). "HET/SAF-B overexpression causes growth arrest and multinuclearity and is associated with aneuploidy in human breast cancer." Clin Cancer Res **6**(9): 3788-3796.

Tripathi, V., J. D. Ellis, Z. Shen, D. Y. Song, Q. Pan, A. T. Watt, S. M. Freier, C. F. Bennett, A. Sharma, P. A. Bubulya, B. J. Blencowe, S. G. Prasanth and K. V. Prasanth (2010). "The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation." Mol Cell **39**(6): 925-938.

Trzepacz, C., A. M. Lowy, J. J. Kordich and J. Groden (1997). "Phosphorylation of the tumor suppressor adenomatous polyposis coli (APC) by the cyclin-dependent kinase p34." J Biol Chem **272**(35): 21681-21684.

Tsianou, D., E. Nikolakaki, A. Tzitzira, S. Bonanou, T. Giannakouros and E. Georgatsou (2009). "The enzymatic activity of SR protein kinases 1 and 1a is negatively affected by interaction with scaffold attachment factors B1 and 2." FEBS J **276**(18): 5212-5227.

Turner-Warwick, R. T. (1955). "The demonstration of lymphatic vessels." Lancet **269**(6905): 1371.

Ule, J., K. Jensen, A. Mele and R. B. Darnell (2005). "CLIP: a method for identifying protein-RNA interaction sites in living cells." Methods **37**(4): 376-386.

Ule, J., K. B. Jensen, M. Ruggiu, A. Mele, A. Ule and R. B. Darnell (2003). "CLIP identifies Nova-regulated RNA networks in the brain." Science **302**(5648): 1212-1215.

Urlaub, H., K. Hartmuth and R. Luhrmann (2002). "A two-tracked approach to analyze RNA-protein crosslinking sites in native, nonlabeled small nuclear ribonucleoprotein particles." Methods **26**(2): 170-181.

urmc.rochester.edu. "Anatomy of the breast." Health Encyclopedia, from https://www.urmc.rochester.edu/encyclopedia/content.aspx?ContentTypeID=92&ContentID=P07789.

van Leusden, M. R., I. Kuikman and A. Sonnenberg (1997). "The unique cytoplasmic domain of the human integrin variant beta4E is produced by partial retention of intronic sequences." Biochem Biophys Res Commun **235**(3): 826-830.

Van Tubergen, E., R. Vander Broek, J. Lee, G. Wolf, T. Carey, C. Bradford, M. Prince, K. L. Kirkwood and N. J. D'Silva (2011). "Tristetraprolin regulates interleukin-6, which is correlated with tumor progression in patients with head and neck squamous cell carcinoma." Cancer **117**(12): 2677-2689.

Venta, L. A., C. M. Dudiak, C. G. Salomon and M. E. Flisak (1994). "Sonographic evaluation of the breast." Radiographics **14**(1): 29-50.

Verhasselt, B., J. Van Damme, N. van Larebeke, W. Put, M. Bracke, C. De Potter and M. Mareel (1992). "Interleukin-1 is a motility factor for human breast carcinoma cells in vitro: additive effect with interleukin-6." Eur J Cell Biol **59**(2): 449-457.

Vorherr, H. (1974). The breast: morphology, physiology and lactation. New York, Academic Press.

Wahl, M. C., C. L. Will and R. Luhrmann (2009). "The spliceosome: design principles of a dynamic RNP machine." Cell **136**(4): 701-718.

Wakeling, A. E. and J. Bowler (1992). "ICI 182,780, a new antioestrogen with clinical potential." J Steroid Biochem Mol Biol **43**(1-3): 173-177.

Wakeling, A. E. and J. Bowler (1987). "Steroidal pure antioestrogens." J Endocrinol **112**(3): R7-10.

Wakeling, A. E., M. Dukes and J. Bowler (1991). "A potent specific pure antiestrogen with clinical potential." Cancer Res **51**(15): 3867-3873.

Wang, H., B. K. Arun, G. N. Fuller, W. Zhang, L. P. Middleton and A. A. Sahin (2008). "IGFBP2 and IGFBP5 overexpression correlates with the lymph node metastasis in T1 breast carcinomas." Breast J **14**(3): 261-267.

Wang, Q., D. Horiatis and J. Pinski (2004). "Interleukin-6 inhibits the growth of prostate cancer xenografts in mice by the process of neuroendocrine differentiation." Int J Cancer **111**(4): 508-513.

Wang, Y., X. Wang, H. Zhao, B. Liang and Q. Du (2012). "Clusterin confers resistance to TNF-alpha-induced apoptosis in breast cancer cells through NF-kappaB activation and Bcl-2 overexpression." J Chemother **24**(6): 348-357.

Wang, Z., M. Kayikci, M. Briese, K. Zarnack, N. M. Luscombe, G. Rot, B. Zupan, T. Curk and J. Ule (2010). "iCLIP predicts the dual splicing effects of TIA-RNA interactions." PLoS Biol **8**(10): e1000530.

Wang, Z., X. Zhang, P. Shen, B. W. Loggie, Y. Chang and T. F. Deuel (2006). "A variant of estrogen receptor-{alpha}, hER-{alpha}36: transduction of estrogen- and antiestrogen-dependent membrane-initiated mitogenic signaling." Proc Natl Acad Sci U S A **103**(24): 9063-9068.

Wang, Z., X. Zhang, P. Shen, B. W. Loggie, Y. Chang and T. F. Deuel (2005). "Identification, cloning, and expression of human estrogen receptor-alpha36, a novel variant of human estrogen receptor-alpha66." Biochem Biophys Res Commun **336**(4): 1023-1027.

Warde-Farley, D., S. L. Donaldson, O. Comes, K. Zuberi, R. Badrawi, P. Chao, M. Franz, C. Grouios, F. Kazi, C. T. Lopes, A. Maitland, S. Mostafavi, J. Montojo, Q. Shao, G. Wright, G. D. Bader and Q. Morris (2010). "The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function." Nucleic Acids Res **38**(Web Server issue): W214-220.

Watermann, D. O., Y. Tang, A. Zur Hausen, M. Jager, S. Stamm and E. Stickeler (2006). "Splicing factor Tra2-beta1 is specifically induced in breast cancer and regulates alternative splicing of the CD44 gene." Cancer Res **66**(9): 4774-4780.

Webb, P., P. Nguyen, J. Shinsako, C. Anderson, W. Feng, M. P. Nguyen, D. Chen, S. M. Huang, S. Subramanian, E. McKinerney, B. S. Katzenellenbogen, M. R. Stallcup and P. J. Kushner (1998). "Estrogen receptor activation function 1 works by binding p160 coactivator proteins." Mol Endocrinol **12**(10): 1605-1618.

Weighardt, F., F. Cobianchi, L. Cartegni, I. Chiodi, A. Villa, S. Riva and G. Biamonti (1999). "A novel hnRNP protein (HAP/SAF-B) enters a subset of hnRNP complexes and relocates in nuclear granules in response to heat shock." J Cell Sci **112 (Pt 10)**: 1465-1476.

Welsh, M., J. Mares, T. Karlsson, C. Lavergne, B. Breant and L. Claesson-Welsh (1994). "Shb is a ubiquitously expressed Src homology 2 protein." Oncogene **9**(1): 19-27.

Wilhelmsen, K., S. H. Litjens and A. Sonnenberg (2006). "Multiple functions of the integrin alpha6beta4 in epidermal homeostasis and tumorigenesis." Mol Cell Biol **26**(8): 2877-2886.

Wu-Baer, F., K. Lagrazon, W. Yuan and R. Baer (2003). "The BRCA1/BARD1 heterodimer assembles polyubiquitin chains through an unconventional linkage involving lysine residue K6 of ubiquitin." J Biol Chem **278**(37): 34743-34746.

Yang, C. R., K. Leskov, K. Hosley-Eberlein, T. Criswell, J. J. Pink, T. J. Kinsella and D. A. Boothman (2000). "Nuclear clusterin/XIP8, an x-ray-induced Ku70-binding protein that signals cell death." Proc Natl Acad Sci U S A **97**(11): 5907-5912.

Yee, D., R. E. Favoni, M. E. Lippman and D. R. Powell (1991). "Identification of insulin-like growth factor binding proteins in breast cancer cells." Breast Cancer Res Treat **18**(1): 3-10.

Yeo, G. W., N. G. Coufal, T. Y. Liang, G. E. Peng, X. D. Fu and F. H. Gage (2009). "An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells." Nat Struct Mol Biol **16**(2): 130-137.

Zhang, D., B. Sun, X. Zhao, Y. Cui, S. Xu, X. Dong, J. Zhao, J. Meng, X. Jia and J. Chi (2012). "Secreted CLU is associated with the initiation of triple-negative breast cancer." Cancer Biol Ther **13**(5): 321-329.

Zhang, J., M. G. Guenther, R. W. Carthew and M. A. Lazar (1998). "Proteasomal regulation of nuclear receptor corepressor-mediated repression." Genes Dev **12**(12): 1775-1780.

Zhang, X., L. Ding, L. Kang and Z. Y. Wang (2012). "Estrogen receptor-alpha 36 mediates mitogenic antiestrogen signaling in ER-negative breast cancer cells." PLoS One **7**(1): e30174.

Zhang, X. T., L. Ding, L. G. Kang and Z. Y. Wang (2012). "Involvement of ER-alpha36, Src, EGFR and STAT5 in the biphasic estrogen signaling of ER-negative breast cancer cells." Oncol Rep **27**(6): 2057-2065.

Zhang, X. T., L. G. Kang, L. Ding, S. Vranic, Z. Gatalica and Z. Y. Wang (2011). "A positive feedback loop of ER-alpha36/EGFR promotes malignant growth of ER-negative breast cancer cells." Oncogene **30**(7): 770-780.

Zhao, K. W., D. Sikriwal, X. Dong, P. Guo, X. Sun and J. T. Dong (2011). "Oestrogen causes degradation of KLF5 by inducing the E3 ubiquitin ligase EFP in ER-positive breast cancer cells." Biochem J **437**(2): 323-333.

Zheng, F. F., R. C. Wu, C. L. Smith and B. W. O'Malley (2005). "Rapid estrogen-induced phosphorylation of the SRC-3 coactivator occurs in an extranuclear complex containing estrogen receptor." Mol Cell Biol **25**(18): 8273-8284.

Zhou, Y., C. Yau, J. W. Gray, K. Chew, S. H. Dairkee, D. H. Moore, U. Eppenberger, S. Eppenberger-Castori and C. C. Benz (2007). "Enhanced NF kappa B and AP-1 transcriptional activity associated with antiestrogen resistant breast cancer." BMC Cancer **7**: 59.

Appendix I

ESR1 Primer pair

	Sequence (5'->3')	Length	Tm	GC%	Self complement	Self 3' complement
Forward primer	CCAAAGCATCTGGGATGG	18	55.36	55.56	5.00	3.00
Reverse primer	TAGGCGGCGCCCTCGG	16	64.31	81.25	7.00	2.00

Products on target templates

>NM_002333.3 Homo sapiens low density lipoprotein receptor-related protein 3 (LRP3), mRNA

```
product length = 416
Forward primer  1    CCAAAGCATCTGGGATGG  18
Template        478  TG..G...GG........  461

Reverse primer  1   TAGGCGGCGCCCTCGG  16
Template        63  G...........G...  78
```

>NM_001171816.1 Homo sapiens ring finger protein 166 (RNF166), transcript variant 3, mRNA

```
product length = 595
Reverse primer  1     TAGGCGGCGCCCTCGG  16
Template        1350  A.......CT.....T  1335

Reverse primer  1    TAGGCGGCGCCCTCGG  16
Template        756  ........C...C...  771
```

>NM_033452.2 Homo sapiens tripartite motif containing 47 (TRIM47), mRNA

```
product length = 440
Reverse primer  1    TAGGCGGCGCCCTCGG  16
Template        440  AG...C..........  425

Reverse primer  1  TAGGCGGCGCCCTCGG  16
Template        1  --...........G..  14
```

>NM_004472.2 Homo sapiens forkhead box D1 (FOXD1), mRNA

```
product length = 449
Reverse primer  1     TAGGCGGCGCCCTCGG  16
Template        1466  GC.......G..G...  1451

Reverse primer  1     TAGGCGGCGCCCTCGG  16
Template        1018  ..C...C.........  1033
```

>NM_001040653.2 Homo sapiens ZXD family zinc finger C (ZXDC), transcript variant 2, mRNA

```
product length = 430
Reverse primer  1    TAGGCGGCGCCCTCGG  16
Template        533  GC..........G...  518
```

```
Reverse primer  1    TAGGCGGCGCCCTCGG  16
Template        104  AT.......G..C...  119
```

>NM_002332.2 Homo sapiens low density lipoprotein receptor-related protein 1 (LRP1), mRNA

```
product length = 453
Reverse primer  1     TAGGCGGCGCCCTCGG  16
Template        7996  G.TT....C.......  7981

Reverse primer  1     TAGGCGGCGCCCTCGG  16
Template        7544  CG........T.....  7559
```

>NM_025112.4 Homo sapiens ZXD family zinc finger C (ZXDC), transcript variant 1, mRNA

```
product length = 430
Reverse primer  1    TAGGCGGCGCCCTCGG  16
Template        533  GC.........G...  518

Reverse primer  1    TAGGCGGCGCCCTCGG  16
Template        104  AT.......G..C...  119
```

>NM_006602.2 Homo sapiens transcription factor-like 5 (basic helix-loop-helix) (TCFL5), mRNA

```
product length = 517
Reverse primer  1    TAGGCGGCGCCCTCGG  16
Template        569  .T.........G....  554

Reverse primer  1   TAGGCGGCGCCCTCGG  16
Template        53  CCC...T........C  68
```

>NM_001098834.1 Homo sapiens gastrulation brain homeobox 1 (GBX1), mRNA

```
product length = 522
Reverse primer  1     TAGGCGGCGCCCTCGG  16
Template        1030  .GT........G....  1015

Reverse primer  1    TAGGCGGCGCCCTCGG  16
Template        509  C....T.T.......A  524
```

>NM_005853.5 Homo sapiens iroquois homeobox 5 (IRX5), transcript variant 1, mRNA

```
product length = 494
Reverse primer  1    TAGGCGGCGCCCTCGG  16
Template        653  GCTT...G........  638

Reverse primer  1    TAGGCGGCGCCCTCGG  16
Template        160  .TCA..........C  175
```

>NM_001252197.1 Homo sapiens iroquois homeobox 5 (IRX5), transcript variant 2, mRNA

```
product length = 494
Reverse primer  1    TAGGCGGCGCCCTCGG  16
Template        653  GCTT...G........  638

Reverse primer  1    TAGGCGGCGCCCTCGG  16
Template        160  .TCA..........C  175
```

>NM_006237.3 Homo sapiens POU class 4 homeobox 1 (POU4F1), mRNA

```
product length = 547
Reverse primer  1     TAGGCGGCGCCCTCGG  16
Template        1256  CGCTG...........  1241

Reverse primer  1    TAGGCGGCGCCCTCGG  16
Template        710  CG.......G..C...  725
```

>NM_138444.3 Homo sapiens potassium channel tetramerization domain containing 12 (KCTD12), mRNA

```
product length = 624
Reverse primer  1    TAGGCGGCGCCCTCGG  16
Template        681  AG.......G..C...  666

Reverse primer  1   TAGGCGGCGCCCTCGG  16
Template        58  ACA.......G....C  73
```

>NM_024963.4 Homo sapiens F-box and leucine-rich repeat protein 18 (FBXL18), mRNA

```
product length = 589
Forward primer  1    CCAAAGCATCTGGGATGG  18
Template        144  .TGC....G.C.......  161

Reverse primer  1    TAGGCGGCGCCCTCGG  16
Template        732  AG.AT..........C  717
```

Appendix II

Distribution of x-link sites and counts in RNAmaps.
[x-link sites] total: 579977
[x-link sites] used in any map: 257504 (44.40%)
[unknown_counts] total: 579977
[unknown_counts] used in any map: 505177 (50.73%)

map	span (upstream; downstream)	[x-link sites] used in map	[x-link sites] % of total	[x-link sites] % of used in any map
exon-intron	-300; 300	58460	10.08%	22.70%
intron-exon	-300; 300	105228	18.14%	40.86%
5-ncRNA	-1000; 300	18027	3.11%	7.00%
ncRNA-3	-300; 1000	18863	3.25%	7.33%
ORF-3UTR	-300; 300	23302	4.02%	9.05%
5UTR-ORF	-300; 300	6313	1.09%	2.45%
inter-5UTR	-1000; 300	7807	1.35%	3.03%
3UTR-inter	-300; 1000	17043	2.94%	6.62%
inter-5	-1000; 300	21103	3.64%	8.20%
3-inter	-300; 1000	27219	4.69%	10.57%
BP best in intron	-300; 100	17956	3.10%	6.97%

RNAmap exon-intron

x-links and junctions distributions

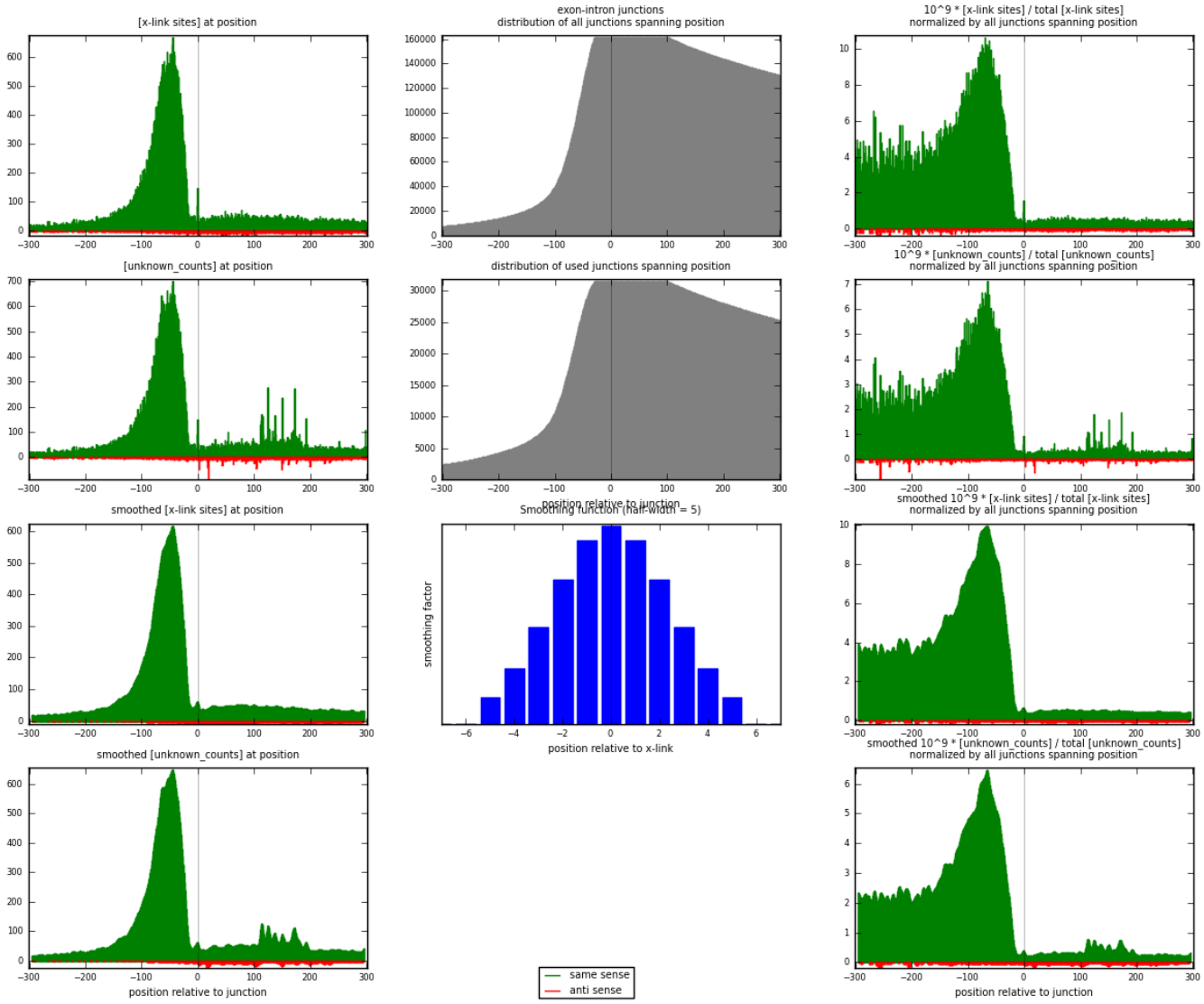
strand	[x-link sites] used in map	[x-link sites] % of total	[x-link sites] % of used in any map	[x-link sites] % of used in map	unique [junctions] in map	[x-link sites] per unique [junction]
same	56213	9.69%	21.83%	96.16%	30192	1.78812
anti	2247	0.39%	0.87%	3.84%	1616	0.0714763
both	58460	10.08%	22.70%	100.00%	31437	1.85959

Top 10 junctions with largest x-link sites sum (both strands)

position of [junction]	span (upstream; downstream)	associated gene	[x-link sites] (same; anti)
chr2:85133785@-	-173; 300	C2orf89	0; 101
chr5:76129052@+	-269; 287	F2RL1	87; 0
chr15:45490968@-	-165; 300	SHF	68; 0
chr7:139025365@+	-85; 256	C7orf55	66; 1
chr14:102453100@+	-300; 300	DYNC1H1	63; 0
chr11:62296095@-	-300; 300	AHNAK	62; 0
chr1:45241812@+	-253; 174	RPS8	51; 0
chr14:56079289@+	-277; 300	KTN1	46; 0
chr11:62288933@-	-61; 300	AHNAK	43; 0
chr17:38600336@+	-175; 300	IGFBP4	40; 0

RNAmap exon-intron

junction is at position zero.
upstream is **regions of type: ORF, 3UTR, 5UTR**
downstream is **regions of type: intron**



RNAmap intron-exon

x-links and junctions distributions

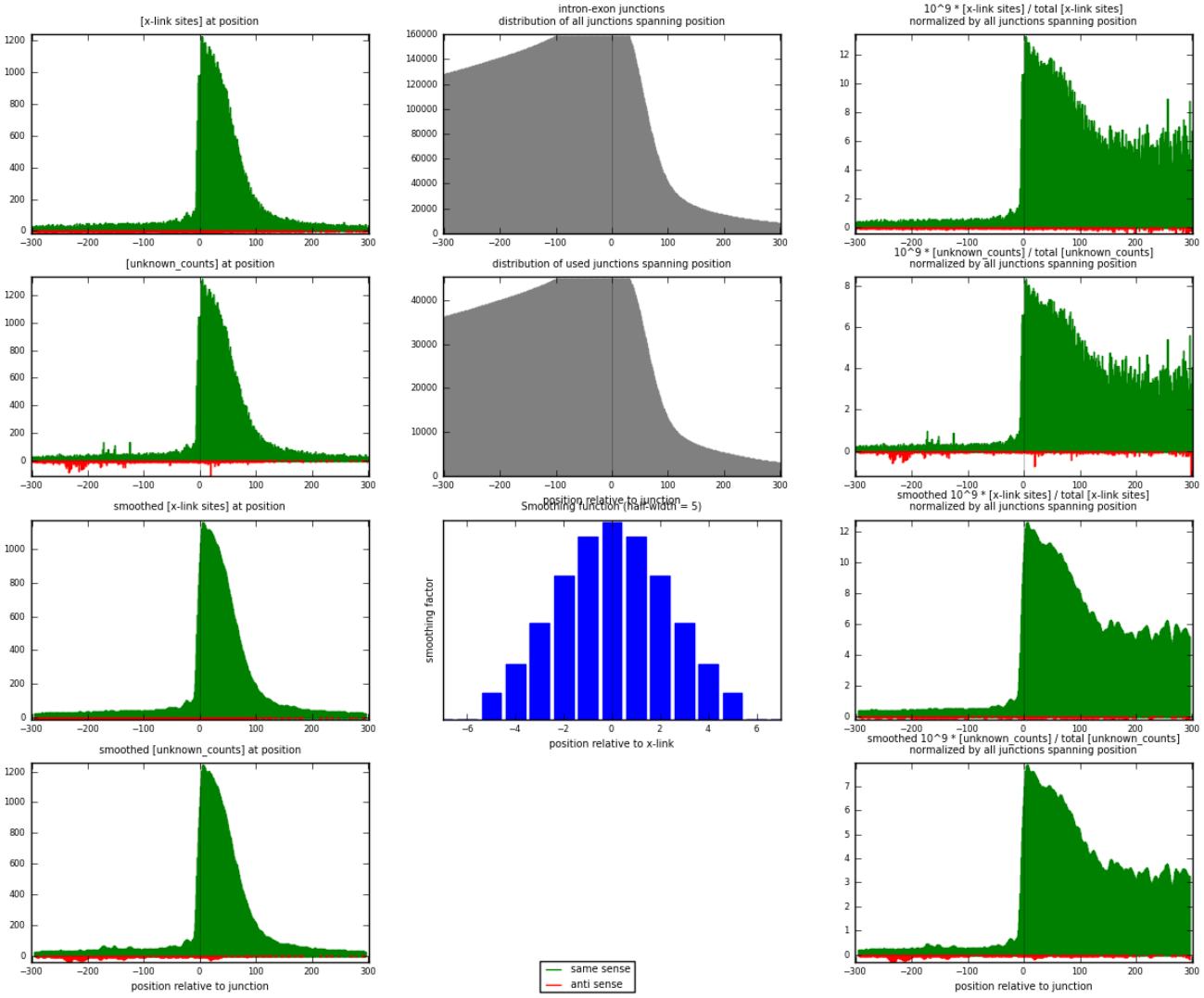
strand	[x-link sites] used in map	[x-link sites] % of total	[x-link sites] % of used in any map	[x-link sites] % of used in map	unique [junctions] in map	[x-link sites] per unique [junction]
same	103045	17.77%	40.02%	97.93%	43574	2.30361
anti	2183	0.38%	0.85%	2.07%	1547	0.0488018
both	105228	18.14%	40.86%	100.00%	44732	2.35241

Top 10 junctions with largest x-link sites sum (both strands)

position of [junction]	span (upstream; downstream)	associated gene	[x-link sites] (same; anti)
chr10:17279228@+	-300; 182	VIM	121; 0
chr2:122288320@-	-300; 249	CLASP1	0; 112
chr7:5568349@-	-150; 219	ACTB	107; 0
chr10:17271274@+	-124; 300	VIM	98; 0
chr1:153507850@-	-206; 141	S100A6	98; 0
chr5:76129626@+	-287; 300	F2RL1	94; 0
chr1:153507305@-	-131; 116	S100A6	93; 0
chr15:45009804@+	-300; 277	B2M	91; 0
chr14:102452023@+	-300; 300	DYNC1H1	78; 0
chr17:79478651@-	-138; 219	ACTG1	77; 0

RNAmap intron-exon

junction is at position zero.
upstream is **regions of type: intron**
downstream is **regions of type: ORF, 3UTR, 5UTR**



RNAmap 5-ncRNA

x-links and junctions distributions

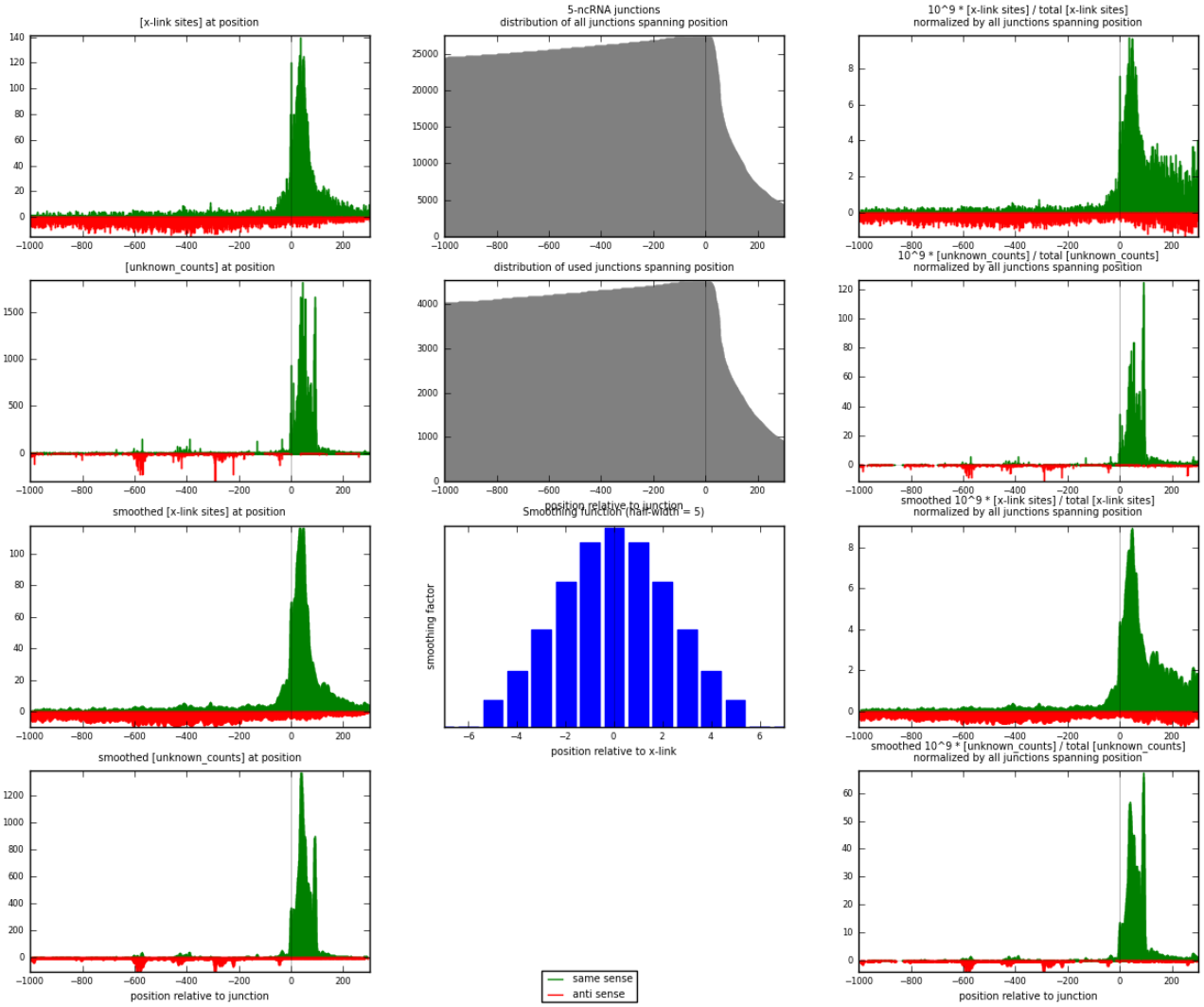
strand	[x-link sites] used in map	[x-link sites] % of total	[x-link sites] % of used in any map	[x-link sites] % of used in map	unique [junctions] in map	[x-link sites] per unique [junction]
same	11518	1.99%	4.47%	63.89%	2809	2.56183
anti	6509	1.12%	2.53%	36.11%	1966	1.44773
both	18027	3.11%	7.00%	100.00%	4496	4.00956

Top 10 junctions with largest x-link sites sum (both strands)

position of [junction]	span (upstream; downstream)	associated gene	[x-link sites] (same; anti)
chr11:65269806@+	-1000; 29	MALAT1	220; 0
chr14:20811565@-	-1000; 166	RNaseP_nuc-ENSG00000252678	146; 1
chr9:35658013@-	-429; 132	RNase_MRP-ENSG00000199916	113; 2
chr10:17276831@-	-1000; 66	RP11-124N14.3	0; 111
chr3:58156362@-	-1000; 70	RP11-456N14.2	0; 103
chr16:3202679@-	-1000; 36	AC108134.5	0; 102
chr2:27273131@-	-1000; 42	AC013472.6	0; 101
chr11:62609280@-	-1000; 95	U2-ENSG00000222328	98; 1
chr6:28864306@+	-1000; 100	HCG14	1; 93
chr8:67025246@-	-1000; 36	AC084082.1	0; 84

RNAmap 5-ncRNA

junction is at position zero.
upstream is **regions of type: intron, inter, telo**
downstream is **regions of type: ncRNA**



RNAmap ncRNA-3

x-links and junctions distributions

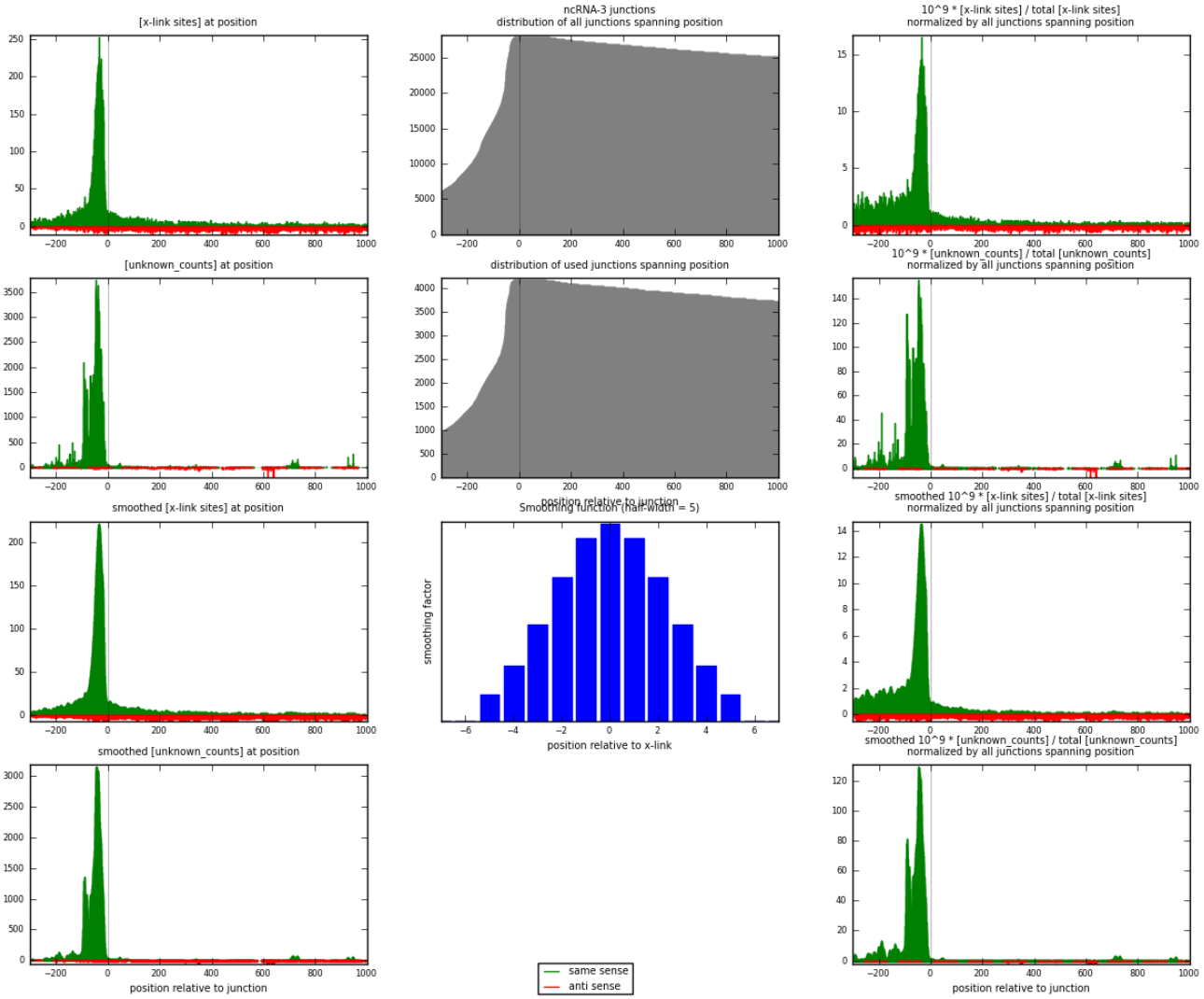
strand	[x-link sites] used in map	[x-link sites] % of total	[x-link sites] % of used in any map	[x-link sites] % of used in map	unique [junctions] in map	[x-link sites] per unique [junction]
same	14326	2.47%	5.56%	75.95%	2834	3.43138
anti	4537	0.78%	1.76%	24.05%	1511	1.08671
both	18863	3.25%	7.33%	100.00%	4175	4.51808

Top 10 junctions with largest x-link sites sum (both strands)

position of [junction]	span (upstream; downstream)	associated gene	[x-link sites] (same; anti)
chrUn_gl000220:118417@+	-300; 1000	28S	463; 2
chr11:65271080@+	-300; 1000	MALAT1	338; 0
chr1:28835512@-	-182; 1000	AL513497.3	0; 174
chr1:173833079@+	-300; 1000	RP5-1198E17.1	0; 154
chr14:20811232@-	-167; 1000	RNaseP_nuc-ENSG00000252678	143; 0
chr9:35657749@-	-132; 1000	RNase_MRP-ENSG00000199916	131; 0
chr11:62609089@-	-96; 978	U2-ENSG00000222328	120; 1
chr6:52860748@+	-166; 1000	7SK-ENSG00000202198	112; 0
chr1:153507591@+	-106; 1000	BX470102.3	0; 112
chr14:50053596@+	-150; 1000	RN7SL1	97; 0

RNAmap ncRNA-3

junction is at position zero.
upstream is **regions of type: ncRNA**
downstream is **regions of type: intron, inter, telo**



RNAmap ORF-3UTR

x-links and junctions distributions

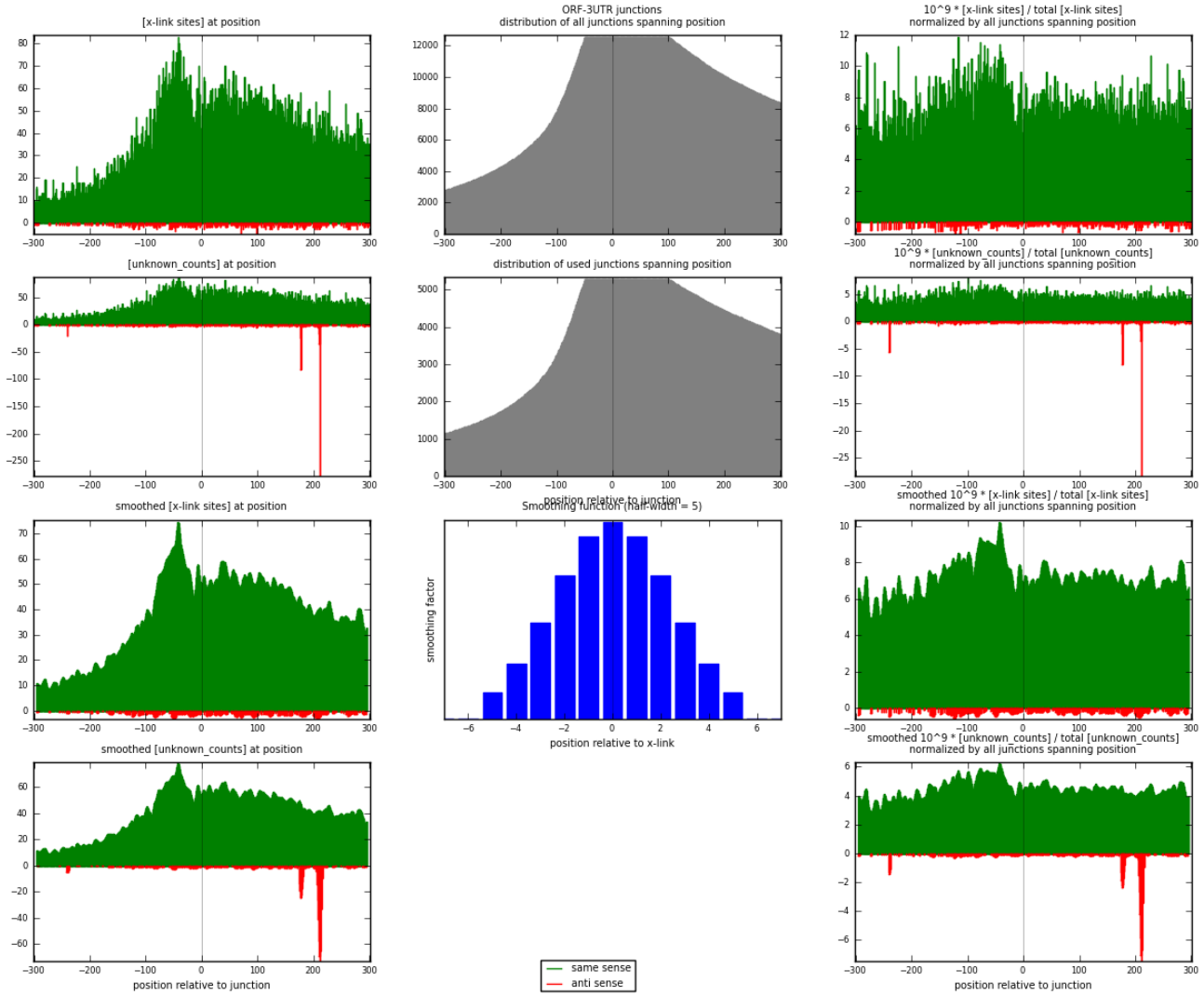
strand	[x-link sites] used in map	[x-link sites] % of total	[x-link sites] % of used in any map	[x-link sites] % of used in map	unique [junctions] in map	[x-link sites] per unique [junction]
same	22875	3.94%	8.88%	98.17%	5114	4.34143
anti	427	0.07%	0.17%	1.83%	266	0.08104
both	23302	4.02%	9.05%	100.00%	5269	4.42247

Top 10 junctions with largest x-link sites sum (both strands)

position of [junction]	span (upstream; downstream)	associated gene	[x-link sites] (same; anti)
chr11:18428923@+	-130; 300	LDHA	105; 0
chr7:116199341@+	-171; 300	CAV1	90; 0
chr17:79477712@-	-73; 300	ACTG1	90; 0
chr17:38612835@+	-68; 300	IGFBP4	84; 0
chr10:70864297@+	-300; 135	SRGN	83; 0
chr9:127998867@-	-283; 300	HSPA5	78; 0
chr2:55199719@-	-300; 198	RTN4	77; 0
chr19:13054727@+	-101; 288	CALR	74; 0
chr6:26046049@+	-206; 300	HIST1H3C	73; 0
chr11:1774729@-	-85; 223	CTSD	73; 0

RNAmap ORF-3UTR

junction is at position zero.
upstream is **regions of type: ORF**
downstream is **regions of type: 3UTR**



RNAmap 5UTR-ORF

x-links and junctions distributions

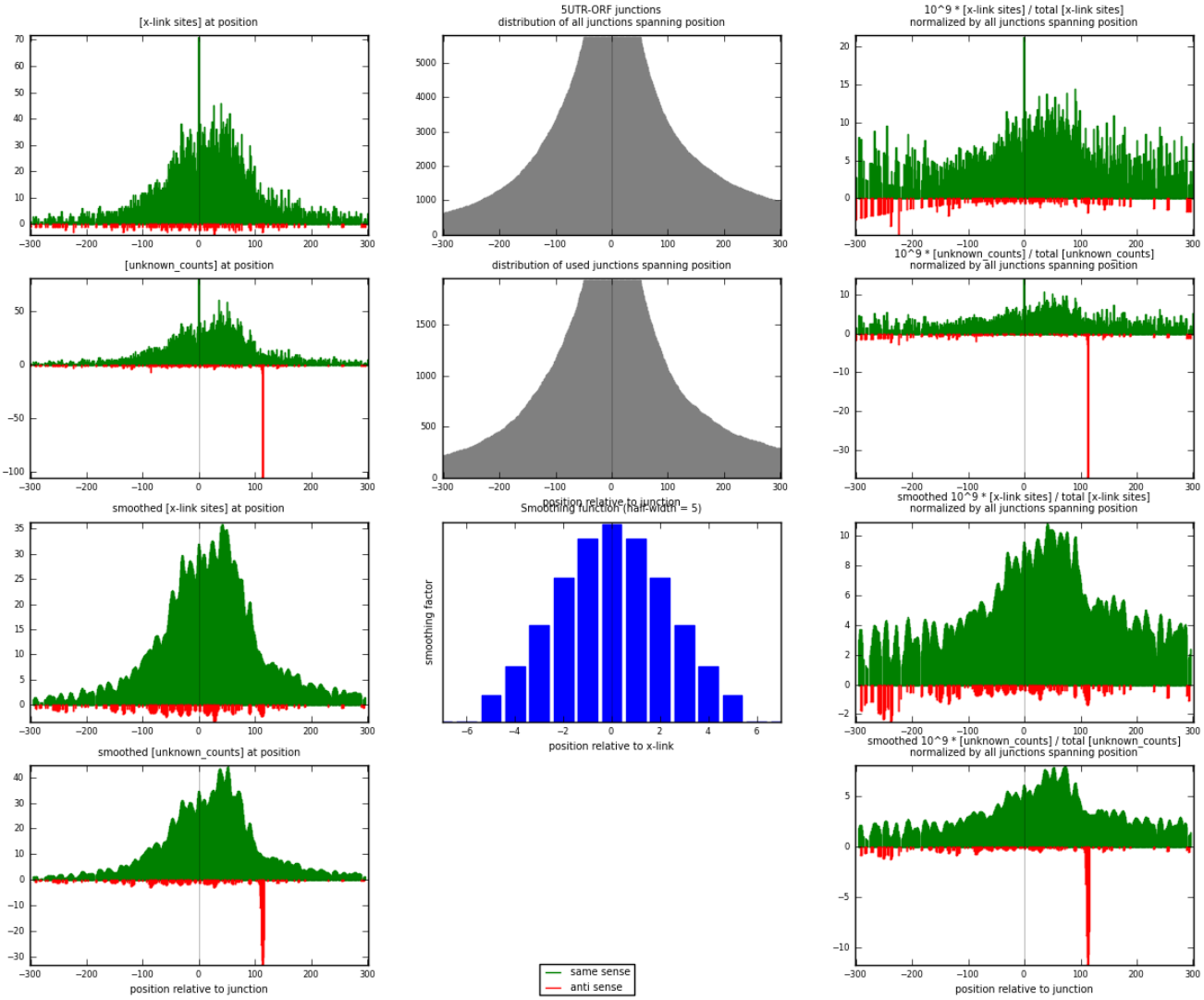
strand	[x-link sites] used in map	[x-link sites] % of total	[x-link sites] % of used in any map	[x-link sites] % of used in map	unique [junctions] in map	[x-link sites] per unique [junction]
same	6106	1.05%	2.37%	96.72%	1847	3.16866
anti	207	0.04%	0.08%	3.28%	140	0.107421
both	6313	1.09%	2.45%	100.00%	1927	3.27608

Top 10 junctions with largest x-link sites sum (both strands)

position of [junction]	span (upstream; downstream)	associated gene	[x-link sites] (same; anti)
chr17:38599987@+	-138; 174	IGFBP4	72; 0
chr17:4851688@-	-300; 66	PFN1	55; 0
chr2:113404405@+	-133; 167	SLC20A1	53; 0
chr6:31508971@-	-66; 265	BAT1	51; 0
chr3:57994291@+	-83; 146	FLNB	46; 0
chr2:55277435@-	-149; 278	RTN4	43; 0
chr8:119123284@-	-300; 300	EXT1	34; 0
chr14:102431028@+	-82; 128	DYNC1H1	31; 0
chr7:116166578@+	-124; 82	CAV1	30; 0
chr22:38201551@+	-219; 292	H1F0	30; 0

RNAmap 5UTR-ORF

junction is at position zero.
upstream is **regions of type: 5UTR**
downstream is **regions of type: ORF**



RNAmap inter-5UTR

x-links and junctions distributions

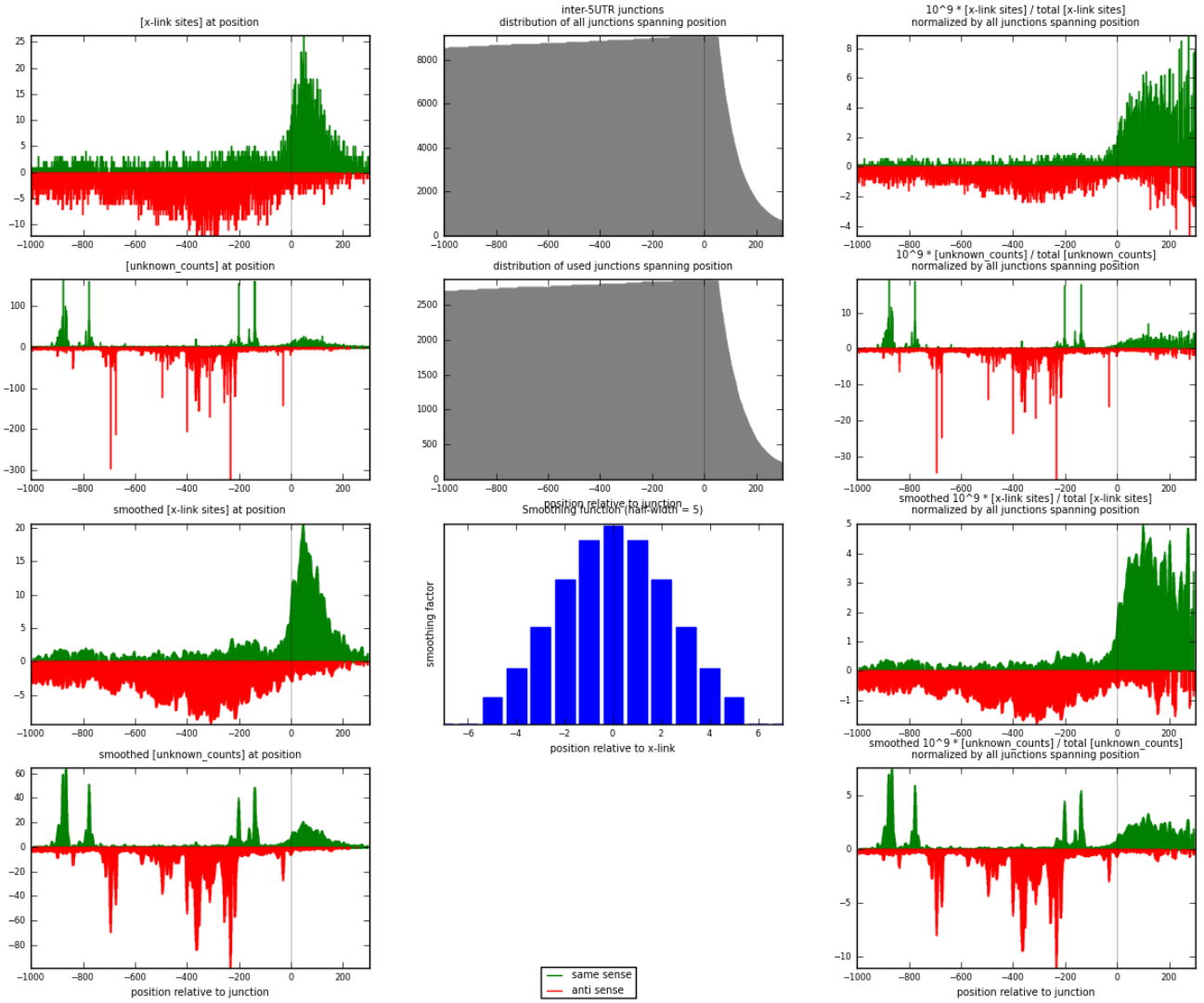
strand	[x-link sites] used in map	[x-link sites] % of total	[x-link sites] % of used in any map	[x-link sites] % of used in map	unique [junctions] in map	[x-link sites] per unique [junction]
same	3275	0.56%	1.27%	41.95%	1579	1.15561
anti	4532	0.78%	1.76%	58.05%	1769	1.59915
both	7807	1.35%	3.03%	100.00%	2834	2.75476

Top 10 junctions with largest x-link sites sum (both strands)

position of [junction]	span (upstream; downstream)	associated gene	[x-link sites] (same; anti)
chr14:21152371@+	-1000; 158	RNASE4	149; 0
chr11:62623517@+	-1000; 112	SLC3A2	0; 129
chr8:33370702@-	-1000; 286	C8orf41	0; 119
chr1:153643523@-	-1000; 63	ILF2	0; 106
chr20:47894962@-	-1000; 248	ZNFX1	0; 83
chr5:180649498@+	-1000; 141	TRIM41	0; 74
chr16:22308729@+	-1000; 64	POLR3E	68; 1
chr11:66115160@-	-1000; 72	AP001107.2	0; 67
chr17:37026255@+	-1000; 93	LASP1	0; 63
chr14:21078042@-	-1000; 170	RNASE11	0; 63

RNAmap inter-5UTR

junction is at position zero.
upstream is **regions of type: inter, telo**
downstream is **regions of type: 5UTR**



RNAmap 3UTR-inter

x-links and junctions distributions

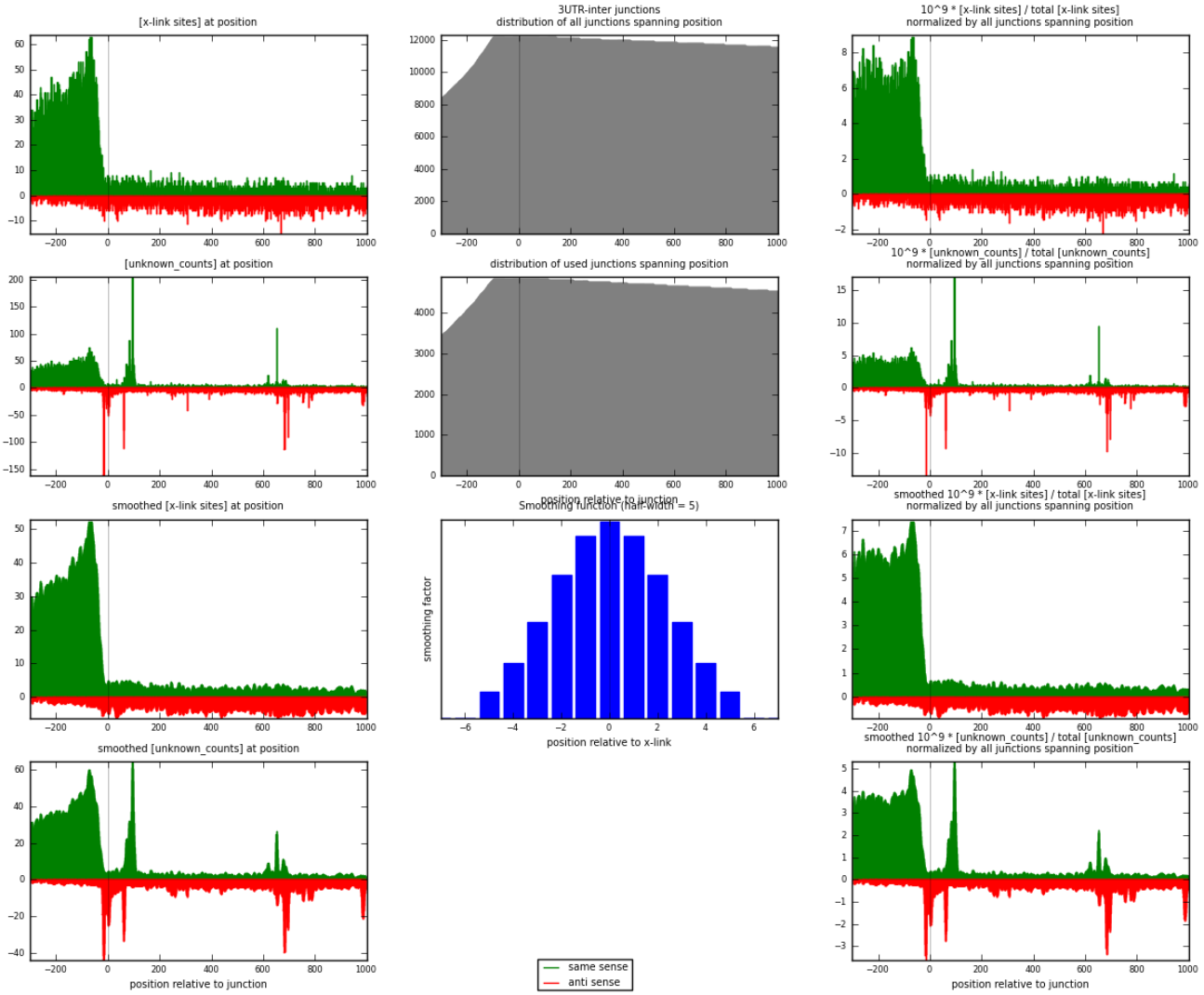
strand	[x-link sites] used in map	[x-link sites] % of total	[x-link sites] % of used in any map	[x-link sites] % of used in map	unique [junctions] in map	[x-link sites] per unique [junction]
same	12690	2.19%	4.93%	74.46%	4142	2.63223
anti	4353	0.75%	1.69%	25.54%	1090	0.902925
both	17043	2.94%	6.62%	100.00%	4821	3.53516

Top 10 junctions with largest x-link sites sum (both strands)

position of [junction]	span (upstream; downstream)	associated gene	[x-link sites] (same; anti)
chr17:8024297@-	-294; 1000	HES7	105; 64
chr17:8091650@-	-300; 1000	C17orf59	0; 91
chr4:119201191@-	-300; 1000	PRSS12	0; 75
chr19:39220830@-	-300; 1000	CAPN12	0; 70
chr17:38613983@+	-300; 1000	IGFBP4	70; 0
chr15:72491368@-	-300; 616	PKM2	64; 0
chr17:79476997@-	-300; 1000	ACTG1	61; 0
chr6:28912314@+	-125; 1000	C6orf100	60; 0
chr10:75880014@-	-300; 1000	AP3M1	0; 57
chr15:45010359@+	-278; 1000	B2M	56; 0

RNAmap 3UTR-inter

junction is at position zero.
upstream is **regions of type: 3UTR**
downstream is **regions of type: inter, telo**



RNAmap inter-5

x-links and junctions distributions

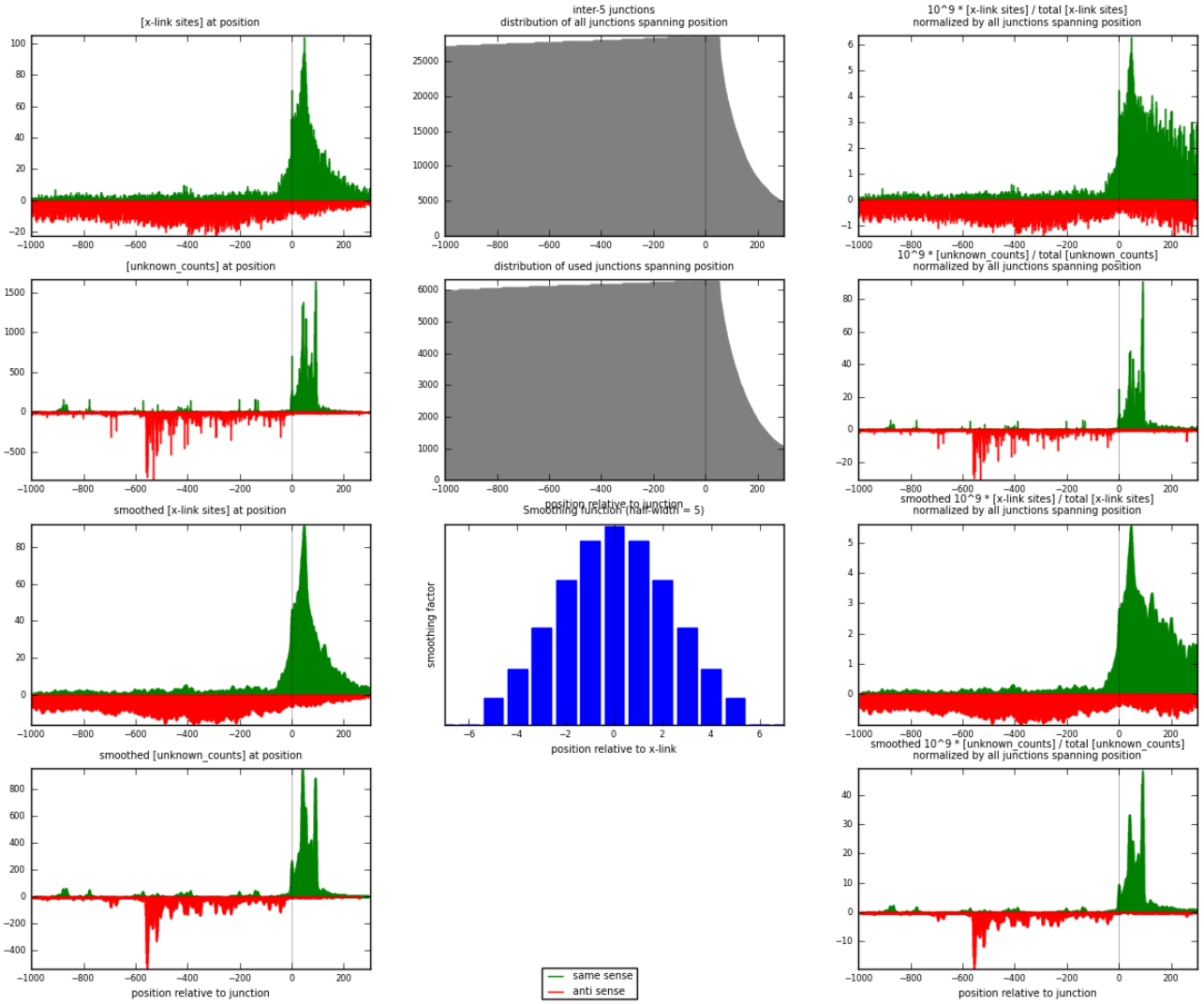
strand	[x-link sites] used in map	[x-link sites] % of total	[x-link sites] % of used in any map	[x-link sites] % of used in map	unique [junctions] in map	[x-link sites] per unique [junction]
same	10302	1.78%	4.00%	48.82%	3501	1.64385
anti	10801	1.86%	4.19%	51.18%	3569	1.72347
both	21103	3.64%	8.20%	100.00%	6267	3.36732

Top 10 junctions with largest x-link sites sum (both strands)

position of [junction]	span (upstream; downstream)	associated gene	[x-link sites] (same; anti)
chrUn_gl000220:117814@-	-1000; 67	AL592188.5	1; 755
chr14:21152371@+	-1000; 158	RNASE4	149; 0
chr14:20811565@-	-1000; 166	RNaseP_nuc-ENSG00000252678	146; 1
chr11:62623517@+	-1000; 112	SLC3A2	0; 129
chr9:35658013@-	-429; 132	RNase_MRP-ENSG00000199916	113; 2
chr10:17276831@-	-1000; 66	RP11-124N14.3	0; 111
chr1:45241481@-	-1000; 127	RP11-269F19.2	0; 107
chr1:153643523@-	-1000; 63	ILF2	0; 106
chr3:58156362@-	-1000; 70	RP11-456N14.2	0; 103
chr11:62609280@-	-1000; 95	U2-ENSG00000222328	98; 1

RNAmap inter-5

junction is at position zero.
upstream is **regions of type: inter, telo**
downstream is **regions of type: ORF, intron, 3UTR, 5UTR, ncRNA**



RNAmap 3-inter

x-links and junctions distributions

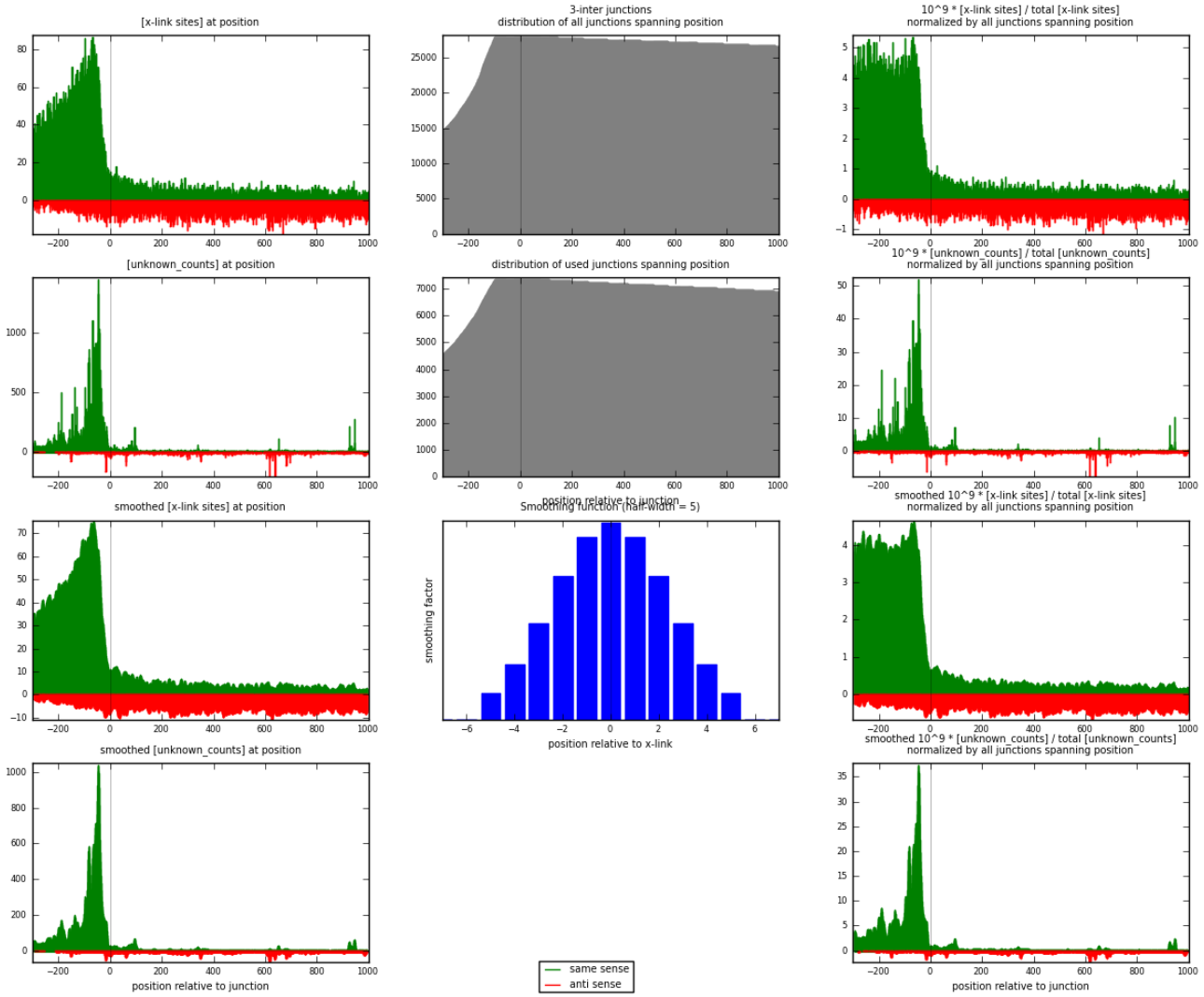
strand	[x-link sites] used in map	[x-link sites] % of total	[x-link sites] % of used in any map	[x-link sites] % of used in map	unique [junctions] in map	[x-link sites] per unique [junction]
same	18799	3.24%	7.30%	69.07%	5725	2.56082
anti	8420	1.45%	3.27%	30.93%	2165	1.14698
both	27219	4.69%	10.57%	100.00%	7341	3.70781

Top 10 junctions with largest x-link sites sum (both strands)

position of [junction]	span (upstream; downstream)	associated gene	[x-link sites] (same; anti)
chrUn_gl000220:118417@+	-300; 1000	28S	463; 2
chr11:65271080@+	-300; 1000	MALAT1	338; 0
chr1:28835512@-	-182; 1000	AL513497.3	0; 174
chr17:8024297@-	-294; 1000	HES7	105; 64
chr2:55199157@+	-142; 1000	EML6	0; 156
chr1:173833079@+	-300; 1000	RP5-1198E17.1	0; 154
chr14:20811232@-	-167; 1000	RNaseP_nuc-ENSG00000252678	143; 0
chr9:35657749@-	-132; 1000	RNase_MRP-ENSG00000199916	131; 0
chr17:16344889@-	-300; 1000	C17orf76	0; 121
chr6:52860748@+	-166; 1000	7SK-ENSG00000202198	112; 0

RNAmap 3-inter

junction is at position zero.
upstream is **regions of type: ORF, intron, 3UTR, 5UTR, ncRNA**
downstream is **regions of type: inter, telo**



RNAmap "BP best in intron"

x-links and BPs distributions

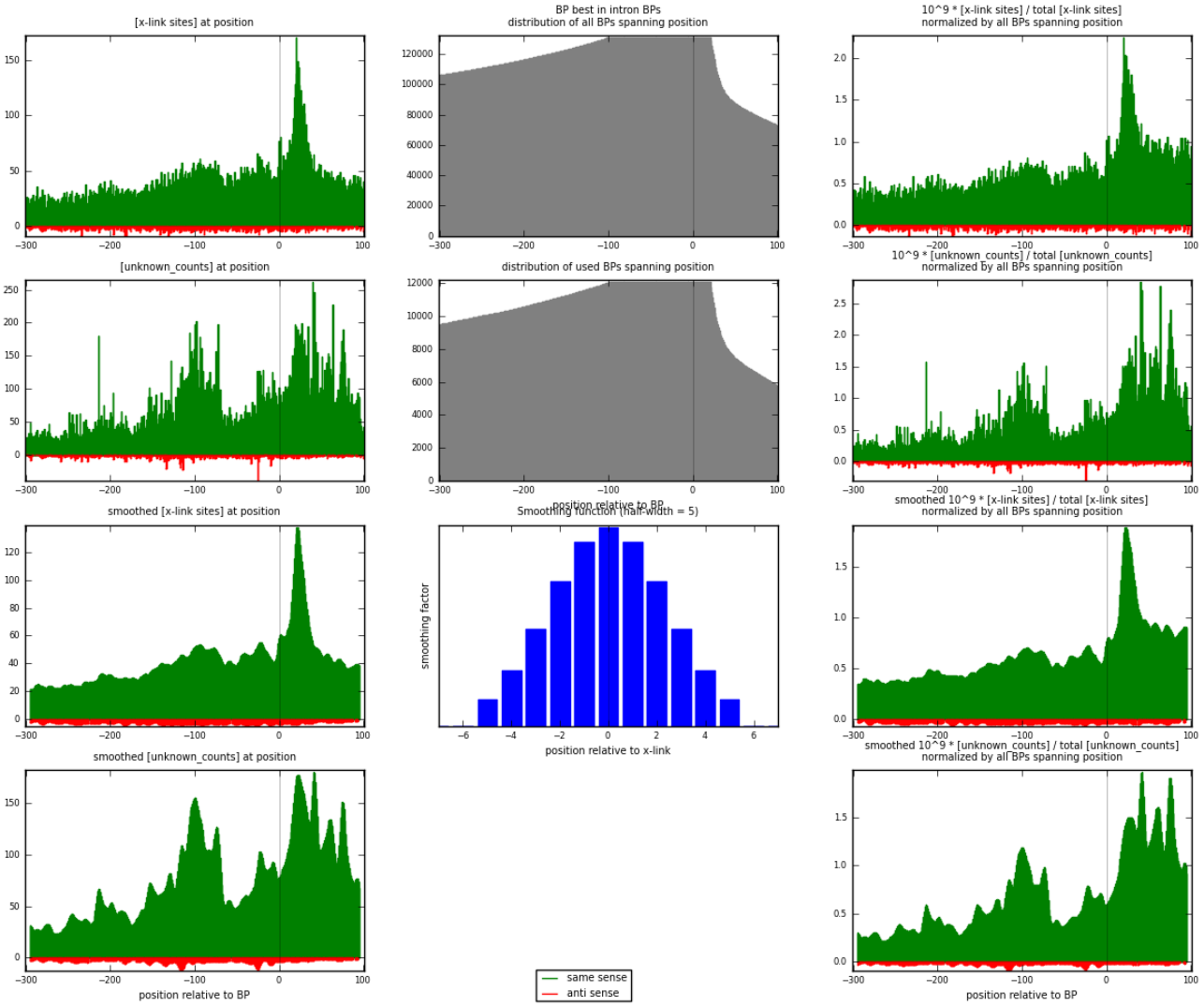
strand	[x-link sites] used in map	[x-link sites] % of total	[x-link sites] % of used in any map	[x-link sites] % of used in map	unique [BPs] in map	[x-link sites] per unique [BP]
same	16881	2.91%	6.56%	94.01%	11281	1.40231
anti	1075	0.19%	0.42%	5.99%	865	0.0893005
both	17956	3.10%	6.97%	100.00%	12038	1.49161

Top 10 BPs with largest x-link sites sum (both strands)

position of [BP]	span (upstream; downstream)	associated gene	[x-link sites] (same; anti)
chr1:28835317@+	-300; 22	SNHG3	148; 0
chr3:39449972@+	-288; 100	SNORA62-ENSG00000206760	98; 0
chr10:101996964@-	-287; 100	SNORA12-ENSG00000212464	97; 0
chr5:138614713@+	-297; 24	MATR3	93; 0
chr10:120819634@-	-109; 100	SNORA19-ENSG00000207468	92; 0
chr12:6619768@+	-224; 24	NCAPD2	90; 0
chr14:103804475@+	-300; 100	EIF5	85; 0
chr9:130210892@-	-238; 100	SNORA65	83; 0
chr11:8707006@+	-177; 100	SNORA3-ENSG00000212607	82; 0
chr11:75111460@+	-283; 100	SNORD15-ENSG00000206941	81; 0

RNAmap "BP best in intron"

BP is at position zero.
upstream is **usptream of branch point**
downstream is **downstream of branch point**



Appendix III

Publications

Hong, E.A., Elliott, D.J. and Tyson-Capper, A.J. (In preparation) *Loss of oestrogen-responsive SAFB1 and SAFB2 alters expression of ER-target genes that contribute to the invasive and metastatic charecteristics of breast cancer cells*

Hong, E.A., Gautrey, H.L., Elliott, D.J. and Tyson-Capper, A.J. (2012) *SAFB1 and SAFB2 mediated transcriptional repression: relevance in cancer.* Biochem Soc Trans 40(4):826-30

Hong, E.A., Elliott, D.J. and Tyson-Capper, A.J. (2011) *Identification of SAFB1 and SAFB2 target genes within the oestrogen receptor signalling pathway in breast cancer cells.* Reproductive Sciences (Supplement) A359.

Hong, E.A., Elliott, D.J. and Tyson-Capper, A.J. (2010) *The effects of oestrogen on alternative splicing regulators in breast cancer cells.* (http://www.isge2010.com)

Hong, E.A., Elliott, D.J. and Tyson-Capper, A.J. (2009) *Effect of oestrogen on RNA-binding splicing regulators in breast cancer cells.* Reproductive Sciences, 17, (3) (Supplement) 109A-110A.

Presentations

21/01/2012 An investigation into the regulation and function of scaffold attachment factors, SAFB1 and SAFB2, in breast cancer cells. (Oral presentation, RNA UK 2012 conference, Lake District, UK)

04/10/2011 An investigation into the regulation and function of scaffold attachment factors, SAFB1 and SAFB2, in breast cancer cells. (Oral presentation, ICM Seminar, Newcastle University)

05/03/2010 The effects of oestrogen on alternative splicing regulators in breast cancer cells. (Poster presentation, 14th World Congress of Gynecological Endocrinology, Florence, Italy)

Prizes and Awards

04/03/2010 Young Scientist Travel and Conference Registration Award, International Society for Gynecological Endocrinology Abstract Competition for the 14th World Congress of Gynecological Endocrinology.

21/09/2008 Overseas Research Students Award Scheme (ORSAS), Newcastle University