THE EXPRESSION OF ALTERNATIVE SPLICE VARIANTS OF HUMAN EPIDERMAL GROWTH FACTOR RECEPTOR 2 IN INVASIVE BREAST CANCER AND CELL LINE MODELS

EMMA ASSAM NKWAM

A thesis submitted in partial fulfilment of the requirements of the University of the West of England, Bristol for the degree of Doctor of Philosophy

Faculty of Health and Life Sciences

May 2015

ACKNOWLEDGEMENTS

My very sincere appreciation goes to my Director of Studies, Professor Anthony Rhodes, and to my second supervisor Dr Michael Ladomery, for all their invaluable help and support, advice and training. I thank them both also for the opportunity to do this research.

I thank the people who have been very helpful during my time as a PhD student; Dave Corry for teaching me tissue culture; Rachel Hagen for her invaluable help with extraction of RNA from FFPE samples; Patricia Adamo for all her help with my gene analysis; Dann Turner for his help with bioinformatics; and Jonathon Hull for his help with western blotting.

I thank all past and present members of the CRIB lab, as well as the team of technicians, and especially my friends Sarah Dean, Keith Page and Hanan Alabouh for all their help and support.

I thank my family and friends, especially my siblings, for their continued support throughout this journey. I would never have made it this far without the people who love me the most.

My parents believed in me enough to fund this research. For this and so much more I am eternally grateful.

I thank my husband Michael, the wind beneath my wings. Thank you for always telling me this was possible. Words are not enough.

I dedicate this thesis to our little angel and my queen, Olivia-Grace Nkwam.

ABBREVIATIONS

ABC	Avidin-Biotin Complex
AKT	Protein Kinase B
bp	Base Pair
cDNA	Complementary DNA
CISH	Chromogenic *In Situ* Hybridization
DAB	Diaminobenzidine
DEPC	Diethylprocarbonate
DMEM	Dolbecco's Modified Eagle Medium
DNA	Deoxyribonucleic Acid
dNTPs	Deoxynucleotide Triphosphates
ECD	Extracellular Domain
ETDA	Ethylenediaminetetraacetic Acid
EGFR	Epidermal Growth Factor Receptor
ELISA	Enzyme-Linked Immunosorbent Assay
ER	Oestrogen Receptor
ERK	Extracellular Signal-Regulated Kinases
ESS	Exonic Splice Silencers
ESE	Exonic Splice Enhancers
FDA	Food And Drug Agency
FFPE	Formalin-Fixed And Paraffin-Embedded
FISH	Fluorescence In Situ Hybridisation
GRB2	Growth Factor Receptor Bound Protein 2
GTP	Guanosine Triphosphate
HER	Human Epidermal Growth Factor Receptor
HER2	Human Epidermal Growth Factor Receptor 2
hnRNPs	Heterogeneous Nuclear Ribonucleoproteins
HRP	Horseradish Peroxidase
IGFR	Insulin-Like Growth Factor Receptor
IHC	Immunohistochemistry

ISE	Intronic Splice Enhancers
ISS	Intronic Splice Silencers
kDa	Kilodalton
LB	*Luria Bertani* Medium
ng	Nanogram
NICE	National Institute For Health And Clinical Care
ml	Millilitre
MAPK	Mitogen-Activated Protein Kinase
µl	Microlitre
µg	Migrogram
mRNA	Messenger Ribonucleic Acid
mTOR	Mammalian Target Of Rapamycin
PARP-1	Poly [ADP-Ribose] Polymerase 1
PBS	Phosphate Buffered Saline
PI3K	Phosphatoinositide 3-Kinae
PIP3	Phosphatidylinositol (3,4,5)-Triphosphate
PR	Progesterone Receptor
PTEN	Phosphatease And Tensin Homologue
RNA	Ribonucleic Acid
rpm	Revolutions Per Minute
PCR	Polymerase Chain Reaction
RT	Reverse Transcription
RT-PCR	Reverse Transcription Polymerase Chain Reaction
siRNA	Small Interfering Ribonucleic Acid
SOS	Sons Of Sevenless
SR protein	Serine And Arginine-Rich Protein
snRNP	Small Nuclear Ribonucleic Particles
TAE	Tris-Acetate-EDTA
TBS	Tris Buffered Saline
TNBC	Triple Negative Breast Cancer

UTR	Untranslated Region
VEGF	Vascular Endothelial Growth Factor
WT1	Wilm's Tumour 1

CONTENT

TABLES

FIGURES

ABSTRACT

The Human Epidermal Growth Factor Receptor 2 (*HER2*) is an oncogene expressed in 25-30% of invasive breast cancers. The *HER2* gene encodes an 185kDa transmembrane protein with tyrosine kinase activity. Gene amplification or protein expression of *HER2* is a predictor of poor prognosis in women with breast cancer, and also indicates a favourable response to *Trastuzumab* (Herceptin) therapy, or a combinational therapy comprising Herceptin plus chemotherapy. However, resistance to *Trastuzumab* remains the case in approximately 50% of *HER2* amplified/overexpressing tumours. Understanding the molecular mechanisms of *Trastuzumab* resistance is critical in the treatment of patients whose breast cancers express this aggressive disease phenotype. In this study, it is postulated that the abnormal generation of mRNA splice variants may be responsible for the continued tumour growth and progression.

The aim of this study is to investigate the expression of alternative splice variants in invasive breast cancer, and to increase our understanding of the regulation of *HER2* and *HER2* splice variants in invasive breast cancer.

The coding region of *HER2* cDNA was PCR amplified in *HER2* positive cell lines (SKOV-3, SKBR-3, and MDA-MB-453). The regulation of *HER2* expression was investigated by siRNA silencing of the splice factor *SRSF1* and its phosphorylating gene *SRPK1*. The role of hypoxia and the inhibition of *SRPK1* via *SRPIN340* were also investigated for its effects of *HER2* expression in cell lines. Human cancer tissues known to be positive for *HER2* were tested for the expression of alternative splice variants of *HER2*.

RT-PCR results reveal new alternative splice variants in invasive breast cancer cells. These new alternative splice variants of *HER2* have also been detected in *HER2*-positive breast cancer samples. Furthermore, the splice factor *SRPK1* and *SRSF1* have shown regulatory effects on the expression of *HER2* in *HER2*-positive cell line MDA-MB-453.

This study identifies for the first time two novel splice variants with deletions in the transmembrane and kinase domains of the *HER2* gene, both with very distinct functional and structural differences. These findings conclude that alternative splicing plays a crucial role in the regulation of *HER2* expression, and possibly in the response of breast cancer patients to current targeted *HER2* therapies.

CHAPTER 1. GENERAL INTRODUCTION

1.1 Breast Cancer

Breast cancer is the most common form of cancer in women worldwide, accounting for approximately 16% of all cancers in women (World Health Organisation, 2010). The WHO statistics for breast cancer estimate that nearly 1.7 million women were diagnosed of breast cancer in 2012 and over 400,000 women died from breast cancer worldwide. In the UK, over 49,936 new cases of breast cancer were diagnosed in 2010, with approximately 11,600 patients dying from the disease (Cancer Research UK, 2010). A study of breast cancer cases reported between 2002 and 2006 by Cancer Research, UK showed that based on ethnicity, the incidence of breast cancer in women in the UK is significantly higher in white women (71% incidence rate). This rate is much lower in Asian, Black, and mixed ethnicities (1.4%, 0.14%, and 0.2%, respectively). Studies also show that breast cancer in black and Asian women tend to be have more aggressive phenotypes, and are mostly triple negative (ER-, PR- and *HER2*-), while white women tend to present a less aggressive phenotype. These differences may be due to differences in genetic predispositions or variations in lifestyle patterns (Anon, 2013).

Breast cancer is caused by malignant tumours arising within the breast epithelia. The progression of breast cancer is believed to be a result of various aberrant transformations which abnormally lead to changes in the breast epithelial cells (Sasso *et al.*, 2011). A number of risk factors have been implicated in the incidence of breast

cancer. Research has shown that the risk of developing breast cancer increases significantly with increase in age (Key *et al.*, 2011; Parkin, Boyd & Walker, 2011). Other risk factors include the use of oestrogen-progesterone contraceptives, excessive alcohol intake, tobacco smoking, diet, and mutations in the BRCA1/BRCA2 genes (Eliassen *et al.*, 2006; Key *et al.*, 2011). Breast cancer is not a single disease, but comprises various subtypes which vary in morphology, prognostic profiles, molecular entities and clinical outcomes (Jackson *et al.*, 2013; Eroles *et al.*, 2012). Prognostic factors for breast cancer include tumour grade, histological type, tumour size, lymph node involvement, distant metastasis, expression of steroid and growth factor receptors, oestrogen-inducible genes e.g Cathepsin-D, proto-oncogenes e.g. *HER2*, mutations in certain genes, e.g. *TP53* (Sørlie *et al.*, 2001). Atypical hyperplasia in breast epithelia may lead to ductal or lobular carcinomas in situ; where malignant cells remain in the ducts or lobules respectively, or may lead to invasive ductal or lobular carcinomas; where they become invasive and the carcinoma cells penetrate the basement membrane of the breast epithelia, and spread into surrounding stroma, skin or muscles, or metastasize to surrounding lymph nodes or distant tissues such as the brain, liver or bone (Jackson *et al.*, 2013). In addition, the overexpression or amplification of certain biomarkers such as the Oestrogen Receptor (ER), the Progesterone Receptor (PR), and the Human Epidermal Growth Factor Receptor 2 (*HER2/neu/c-erbB-2*) is also important in determining the prognosis and treatment of breast cancer (Ciocca *et al.*, 2006).

1.2 Classification of breast cancers

In the past few decades, research into the biology of breast cancer has revolutionised with the focus on new methods of understanding the expression, regulation and function of critical signalling pathways active in the incidence and progression of breast cancers (Eroles *et al.*, 2012; Alvarez & Hortobagyi, 2013). The cellular and molecular heterogeneity of breast cancer contributes to it being a highly complex disease; certain prognostic values are limited in the information that can be obtained about the biology of the disease. Therefore the parameters used to provide prognostic profiles are not self-sufficient in adequately predicting patient outcome in many cases of breast cancer. The model of individualised treatment (based on individual molecular profiles of different patients) has gained much support from the cancer research community, and has led to the identification of different subgroups of patients and the development of various targeted therapies for breast cancer patients (Zagozdzon, Gallagher & Crown, 2011). Examples of these targeted therapies include the successful use of hormonal therapy for women with hormone-sensitive tumour subtypes, and the use of anti-human epidermal growth factor receptor 2 therapies for women with Human Epidermal Growth Factor 2 (*HER2*)- overexpressing tumours (Alvarez & Hortobagyi, 2013).

New classifications of breast cancers have been proposed with the aim of giving potentially more significant prognostic information and providing guides for treatment options for individual subtypes (Boyle, 2012). Studies have focused on classifying breast cancers based on a combination of changes in gene expression microarrays and

immunohistochemical subtypes, resulting in breast cancers being classified into six distinct phenotypical subtypes (Table 1.1)

1.2.1 Luminal A

The luminal A subtype represents 50-60% of breast cancer subtypes, and based on its histological profile, tumours in this subtype are mostly lobular carcinomas in-situ and infiltrating lobular carcinomas (Eroles *et al.*, 2012). The luminal A subtype is immunohistochemically ER positive and/or PR positive, and *HER2* negative, and is known to have a good prognostic outcome (Eroles *et al.*, 2012; Boyle, 2012; Li *et al.*, 2013). Treatment is mainly based on hormonal receptor modulators such as Tamoxifen (Eroles *et al.*, 2012).

1.2.2 Luminal B

The luminal B subtype represents 10-20% of all breast cancers and is a more aggressive phenotype compared to the luminal A subtype (Eroles *et al.*, 2012). Like the Luminal A subgroup, Luminal B breast cancers are ER positive and/or PR positive. However, this subtype is identified as a tumour subgroup with poorer patient outcome compared to the luminal A subtype, due to the increased expression of proliferation genes such as *MKI67*, *Cyclin B1* and oncogenes such as *HER2* (Boyle, 2012; Eroles *et al.*, 2012). The luminal B subtype is also characterised by higher histological grade. In combination with tamoxifen treatment, patients in this subgroup respond well to neoadjuvant chemotherapy (Eroles *et al.*, 2012).

1.2.3 *HER2* overexpressing

This molecular subtype accounts for 10-25% of all breast cancers (Eroles *et al.*, 2012). *HER2* overexpressing cancers are characterised by the absence of ER and PR, and a high expression of the *HER2* gene and also exhibit an overexpression of genes associated with cellular proliferation (Boyle, 2012; Eroles *et al.*, 2012; Li *et al.*, 2013). *HER2* overexpressing tumours are histologically high grade tumours and have unfavourable prognostic implications. Treatment is mainly with anti-*HER2* therapies and neoadjuvant chemotherapy (Eroles *et al.*, 2012).

1.2.4 Basal-like

This subtype represents 10-20% of all breast cancers. Basal-like tumours are mainly infiltrating ductal carcinomas with high lymph node involvement and metastatic potential (Eroles *et al.*, 2012). The most important characteristic of this tumour subtype is the absence of all three breast cancer biomarkers *ER, PR* and *HER2* (Boyle, 2012; Eroles *et al.*, 2012). This is why they are also often referred to as triple negative breast cancers (TNBC). Triple-negative breast basal-like cancers are also positive for EGFR and CK5/6 (Li *et al.*, 2013). Basal-like cancers are usually histologically high grade tumours, and treatment is usually with chemotherapy. Poly-ADP ribose- polymerase-1 (PARP-1) inhibitors are currently being developed to treat this subset of tumours, but are currently only being used in clinical trials (Eroles *et al.*, 2012).

1.2.5 Normal breast-like

About 5-10% of breast carcinomas have been classified under the normal breast-like subtype (Eroles *et al.*, 2012). Normal breast-like breast cancers are negative for *ER*, *PR* and *HER2*, but unlike the triple-negative breast cancer subtype, are also negative for *CK5/6* and *EGFR* (Eroles *et al.*, 2012). Normal breast-like breast cancers are poorly characterised and are very rare (Sørlie *et al.*, 2001; Eroles *et al.*, 2012). Some studies hypothesize that this subtype is only a technical artefact derived from a high contamination of normal tissue during the preparation of tissue microarrays. It is difficult to establish a clinical significance of the normal breast-like breast cancer subtype due to their rarity (Eroles *et al.*, 2012).

1.2.6 Claudin-low

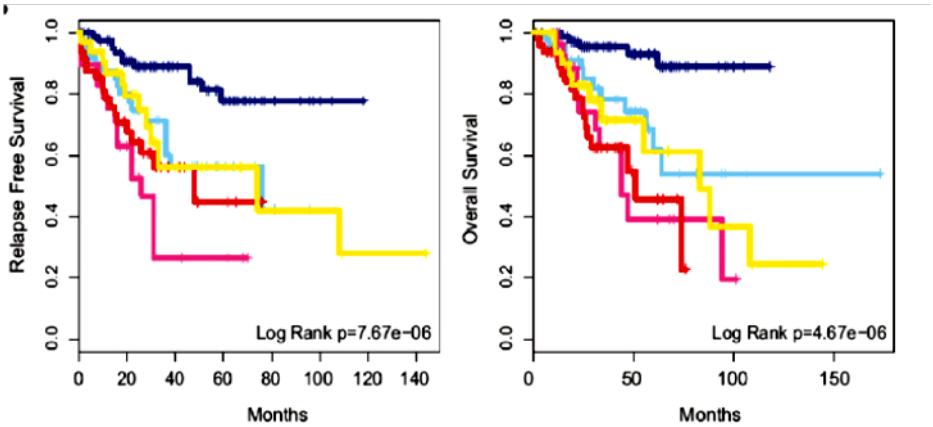
About 12-14% of breast cancers are found under this subtype (Eroles *et al.*, 2012). The claudin-low breast cancer subtype is characterized by a low expression of certain genes which play a crucial role in the formation of tight junctions and in cell adhesion. Examples of these genes are claudin -3, -4, -7 cingulin, occludin and E-cadherin, hence the name claudin-low (Perou, 2011). This subtype is characterised by a high expression of immune cells. Claudin-low tumours also express a low level of genes which are related to cell proliferation, yet this subtype belongs to a poor prognosis group (Eroles *et al.*, 2012; Perou, 2011). Claudin-low breast tumours are relatively high grade infiltrating ductal carcinomas. Immunohistochemically, Claudin-low tumours are mostly triple negative, and treatment of patients is mainly with neoadjuvant chemotherapy. This new classification of breast cancers is still being researched and

has not yet been fully applied in clinical settings due to a lack of standardization for testing of individual tumours (Eroles *et al.*, 2012).

MOLECULAR SUBTYPE	**FREQUENCY**	**ER/PR/ *HER2* STATUS**	**HISTOLOGIC GRADE**	**PROGNOSTIC VALUE**
Basal-like	10-20%	ER-, PR-, *HER2*-	High	Poor
HER2 overexpressing	10-15%	ER-, PR-, *HER2*+	High	Poor
Normal breast-like	5-10%	ER-/+, *HER2*-	Low	Intermediate
Luminal A	50-60%	ER+, PR+, *HER2*-	Low	Excellent
Luminal B	10-20%	ER+/-, PR+/-, *HER2*-/+	Intermediate/ high	Intermediate/ Poor
Claudin-low	12-14%	ER-, PR-, *HER2*-	High	Poor

Table 1.1: Molecular subtypes of *HER2* and their characteristic features. Adapted and used with permission from (Eroles *et al.*, 2012).

A



B

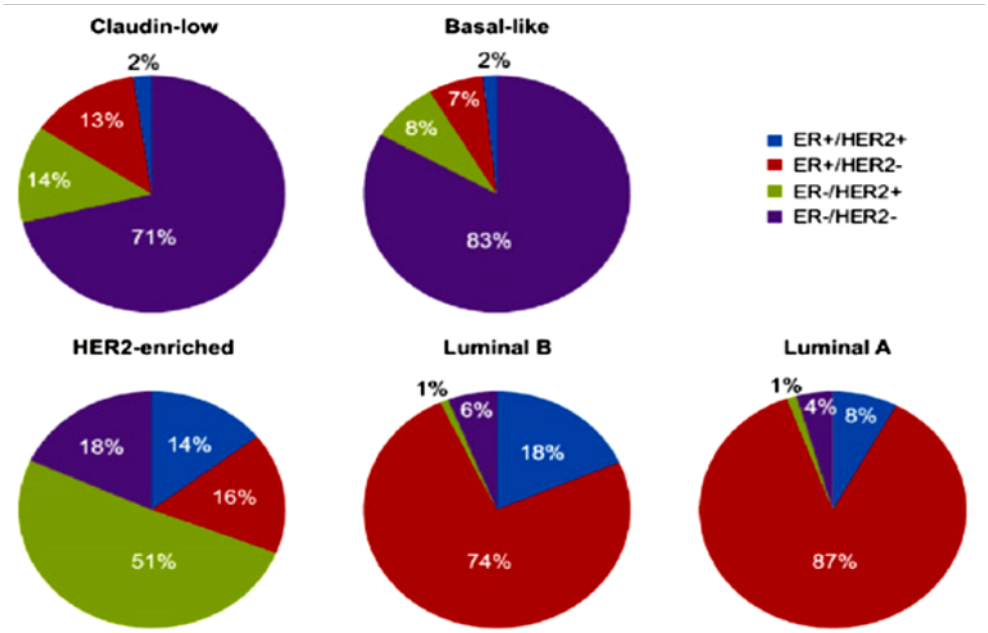


Figure 1.1 (A) Kaplan–Meier curves of disease-free survival and overall survival based on UNC337 database. Dark blue, luminal A; light blue, luminal B; red, basal-like; pink, *HER2*-enriched; yellow, Claudin-low. (B) Distribution of ER and *HER2* in the different subtypes of breast cancer based on mRNA expression. [Source and permissions: (Eroles *et al.*, 2012)].

1.3 Discovery of the Human Epidermal Growth Factor Receptor 2 (*HER2*)

The Human Epidermal Growth Factor Receptor 2 (*HER2*) is a proto-oncogene which belongs to the Human Epidermal Growth Factor Receptor (HER/EGFR) family of transmembrane receptor tyrosine kinases (Dean-Colomb & Esteva, 2008; Normanno *et al.*, 2006). The EGFR family of genes consists of HER1 (EGFR / c-erbB-1), *HER2* (c-erbB-2 / verb-B2 / Neu), HER3 (c-erbB-3), and HER4 (c-erbB-4) (Shah & Chen, 2010; Wan, Sazani & Kole, 2009). Aberrant expression or functioning of the epidermal growth factor family has been implicated in the development and evolution of various cancers (Baselga & Swain, 2009). The *HER2*/*Neu* oncogene was first characterised in 1981 in experiments using genetically modified mice (Sińczak-Kuta *et al.*, 2007; Siegel *et al.*, 1999). *Neu*, the rat homologue of *HER2*, was identified as a transforming gene in transfection experiments using genomic DNA isolated from chemically induced neuroblastoma models (Marchini *et al.*, 2011; Sińczak-Kuta *et al.*, 2007; Siegel *et al.*, 1999; Jackson *et al.*, 2013). Single point mutations in the transmembrane region activate the neu oncogene by converting a valine residue to glutamic acid (Siegel *et al.*, 1999). This mutation in the transmembrane domain results in increased ligand-independent dimerization causing tyrosine kinase activity. The human homolog of the Neu gene (*HER2*) was identified and isolated due to its homology with the Neu oncogene. *ERBB2* is the official name for *HER2* provided by the HUGO Gene Nomenclature Committee for the v-erb-b2 erythroblastic leukemia viral oncogene homolog 2 gene (Wolff *et al.*, 2007).

The *HER2* gene is located on the long arm of human chromosome 17 (17 q21-q22) and encodes a 185kDa tyrosine kinase receptor (Freudenberg *et al.*, 2009; Nuti *et al.*, 2011) that is constitutively active as a dimer and shares extensive homology with the other three members of the EGFR family (Castiglioni *et al.*, 2006); all Human Epidermal Growth factor receptors have an extracellular ligand-binding domain, a short hydrophobic transmembrane region and a cytoplasmic domain with intrinsic tyrosine kinase catalytic activity (Sińczak-Kuta *et al.*, 2007; Hynes & Lane, 2005) (Figure 1.2).

About 10-25% of human breast carcinomas are positive for *HER2* (Koletsa *et al.*, 2008; Scaltriti *et al.*, 2007). The overexpression or gene amplification of *HER2* has been observed in many other human epithelial cancers including colorectal, ovarian, endometrial, prostate, pancreatic, oral, lung, and gastric carcinomas (Freudenberg *et al.*, 2009; Gebhardt, Zänker & Brandt, 1998; Nuti *et al.*, 2011; Blok *et al.*, 2013; Fuse, 2011), and is known to be associated with an unfavourable prognosis. *HER2* positivity in breast cancers is associated with a more aggressive tumour phenotype with increased cell proliferation and metastatic potential (Dean-Colomb & Esteva, 2008), earlier recurrence, significantly lower disease-free and overall survival rates, shorter time to relapse, and overall poor prognosis (Doherty *et al.*, 1999a; Freudenberg *et al.*, 2009; Shah & Chen, 2010).

1.4 *HER2* signalling pathways

The Human Epidermal Growth Factor Receptor (HER/EGFR) tyrosine kinase family of transmembrane proteins are activated following binding with peptide growth factors

of the EGF-family of proteins (Normanno *et al.*, 2006). The EGFRs play a crucial role in normal physiology and evidence also suggests that the EGFR family of receptors is involved in the pathogenesis and progression of different carcinoma types (Normanno *et al.*, 2006; Hynes & MacDonald, 2009). Epidermal growth factor receptors also play a role in embryogenesis, and are important factors in tissue remodelling and renewal throughout adult life. Epidermal Growth Factor Receptors 1 and 2 (EGFRs 1 & 2) in particular are mutated in many epithelial cancers, and clinical studies suggest that they play roles in the development and progression of various cancer types (Hynes & MacDonald, 2009). Thirteen cognate ligands have been characterized that bind to the HER receptors, with the exception of *HER2* which has no known ligand. The cellular mechanism of *HER2* activation is therefore not completely understood (Shah & Chen, 2010). The HER proteins remain in an inactive form, assuming a tethered conformation until they are activated on ligand binding, but *HER2* remains constitutively active and can therefore induce transformation in a ligand-independent way (Alvarez & Hortobagyi, 2013; Nuti *et al.*, 2011). Upon ligand binding to the extracellular domain, homo- or heterodimerization of the *HER2* receptor with itself or other members of the EGFR family leads to phosphorylation of residues from the intracellular domain of the *HER2* receptor and consequent activation of the *HER2* protein resulting in the recruitment of signalling molecules from the cytoplasm and the activation of several signalling pathways (Alvarez & Hortobagyi, 2013). Downstream of *HER2*, phosphorylation results in the recruitment of signalling molecules from the cytoplasm and the induction of several potent intracellular signalling pathways which result in cell differentiation, cell migration, signal transduction, cell motility, cell adhesion, increased cell proliferation, protease expression and activation, and a cascade of other

events which lead to functional changes in both embryonic and adult tissues (Zagozdzon, Gallagher & Crown, 2011; Gebhardt, Zänker & Brandt, 1998; Nuti *et al.*, 2011; Shah & Chen, 2010). The downstream signalling pathway which is activated following *HER2* dimerization is significantly influenced by the pattern of dimerization (Tai, Mahato & Cheng, 2010). Different signalling cascades can potentially be initiated depending on the dimeric combination of *HER2* with itself or other members of the HER family.

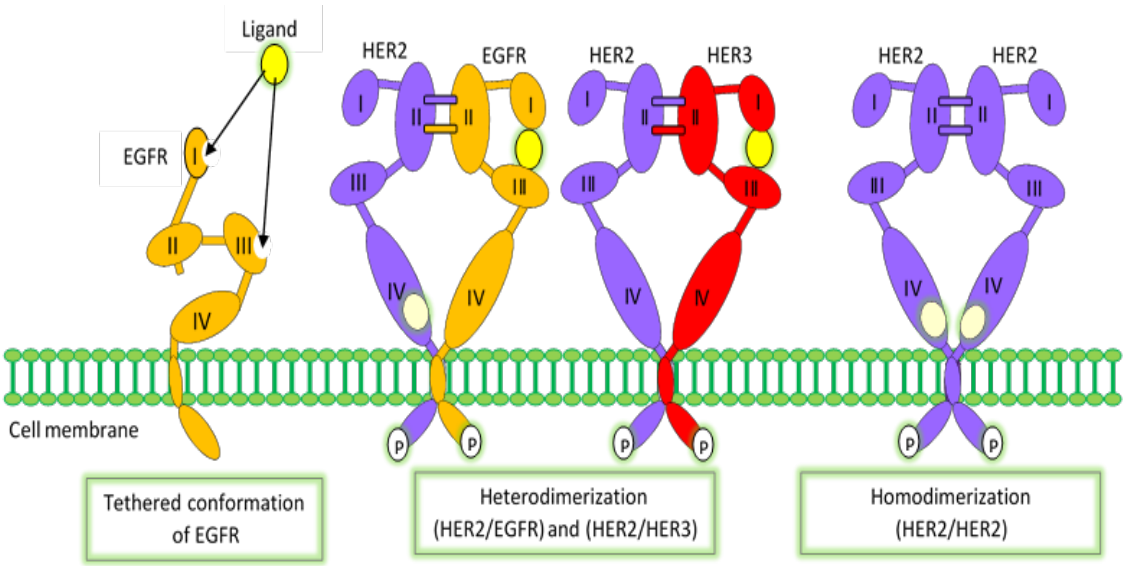


Figure 1.2 Schematic representation of *HER2*. *HER2* (purple) can form heterodimers as pictured with EGFR (orange), HER3 (red), as well as HER4 (not pictured). EGFR (orange) assumes a tethered conformation in the absence of a ligand. *HER2* remains active and is naturally ready for ligand dimerization.

Two of the most studied downstream signalling pathways in EGFR signalling are the phosphoinositide-3-kinase (PI3K/AKT) and the mitogen activated protein kinase (MAPK) cascades.

1.4.1. Phosphoinositide-3-kinase (PI3K/AKT) cascade

The PI3K/AKT pathway is arguably the most significant pathway activated downstream of *HER2* phosphorylation in cancer. The PI3K/AKT lipid kinase activity is stimulated when *HER2* signals in conjunction with HER3. Apart from the *HER2* + HER3 heterodimer, it is also possible to induce the PI3K/AKT cascade by tyrosine-phosphorylated HER3 homodimerization. This is because HER3 possesses numerous binding domains which can interact with the regulatory subunit p85 of PI3K (Tai, Mahato & Cheng, 2010; Hynes & MacDonald, 2009). Homodimers containing HER3 also have the potential to activate the AKT kinase via the PI3K lipid kinase. At the cell membrane, Phosphatidylinositol (3,4,5)-trisphosphate (PIP_3)-bound AKT becomes phosphorylated, resulting in the activation of the mechanistic target of rapamycin (mTOR) (Emde, Köstler & Yarden, 2012; Pohlmann, Mayer & Mernaugh, 2009). The activation of mTOR induces several intracellular functions, interactions with transcription factors, activation of metabolic pathways, apoptosis and angiogenesis, which result in cell proliferation, invasion and survival (Pohlmann, Mayer & Mernaugh, 2009). Following the activation of AKT, PIP_3 is dephosphorylated by the tumour suppressor gene phosphatase and tensin homolog (PTEN) to PIP_2, which makes PTEN a negative regulator of the PI3K/AKT signalling cascade (Figure 1.3). PTEN is a protein encoded by the PTEN gene, and functions as a tumour suppressor gene (Nahta & O'Regan, 2010). PTEN is important in the PI3K/AKT pathway because it inhibits the downstream signalling of P13K (Saal *et al.*, 2008). In normal signalling of the EGFR receptors, the PI3K/AKT pathway induces cell survival and inhibits apoptosis (Jackson *et al.*, 2013).

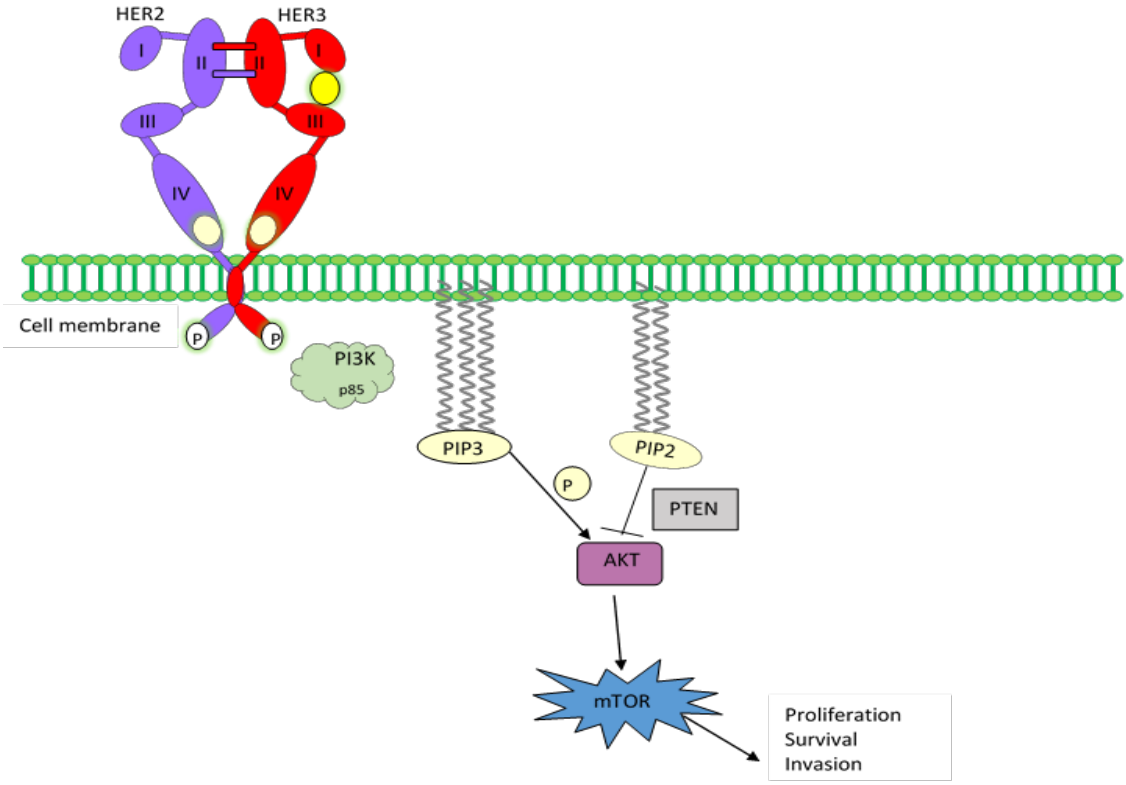


Figure 1.3 Schematic representation of the PI3K/AKT/mTOR pathway. Following ligand binding, *HER2* dimerizes with activated HER3 resulting in PI3K/AKT/mTOR signalling.

1.4.2. Mitogen Activated Protein Kinase (MAPK) cascade

This is also known as the RAS/RAF/MEK/MAPK pathway. Unlike the PI3K/AKT pathway, the MAPK pathway can be activated by all dimerizations involving *HER2* (HER1/*HER2*, *HER2*/*HER2*, *HER2*/HER3, and *HER2*/HER4) (Tai, Mahato & Cheng, 2010). In this cascade, the adaptor protein known as Growth Factor Receptor Bound Protein 2 (GRB2) which recognises tyrosine-phosphorylated sites on the activated receptor, binds to the guanine nucleotide exchange factor Son of Sevenless (SOS). The GRB2/SOS complex binding to the receptor activates SOS, resulting in a loss of Guanosine Diphosphate (GPD) from inactive RAS. Free RAS is then activated by binding to Guanosine-5'-Triphosphate (GTP). RAS/GTP complex then binds to activate Raf-1

(MAP3K). Raf-1 can then activate MEK1 (MAP2K1) and MEK2 (MAP2K2) which are essential for downstream signalling of RAS and Raf-1. The activation of MEK results in the phosphorylation of ERK (Figure 1.4). ERK activation is crucial to certain physiological functions in the cell, including cell cycle control, differentiation, migration, apoptosis and angiogenesis (Pohlmann, Mayer & Mernaugh, 2009). The pathway also serves in stimulating cell proliferation (Jackson *et al.*, 2013).

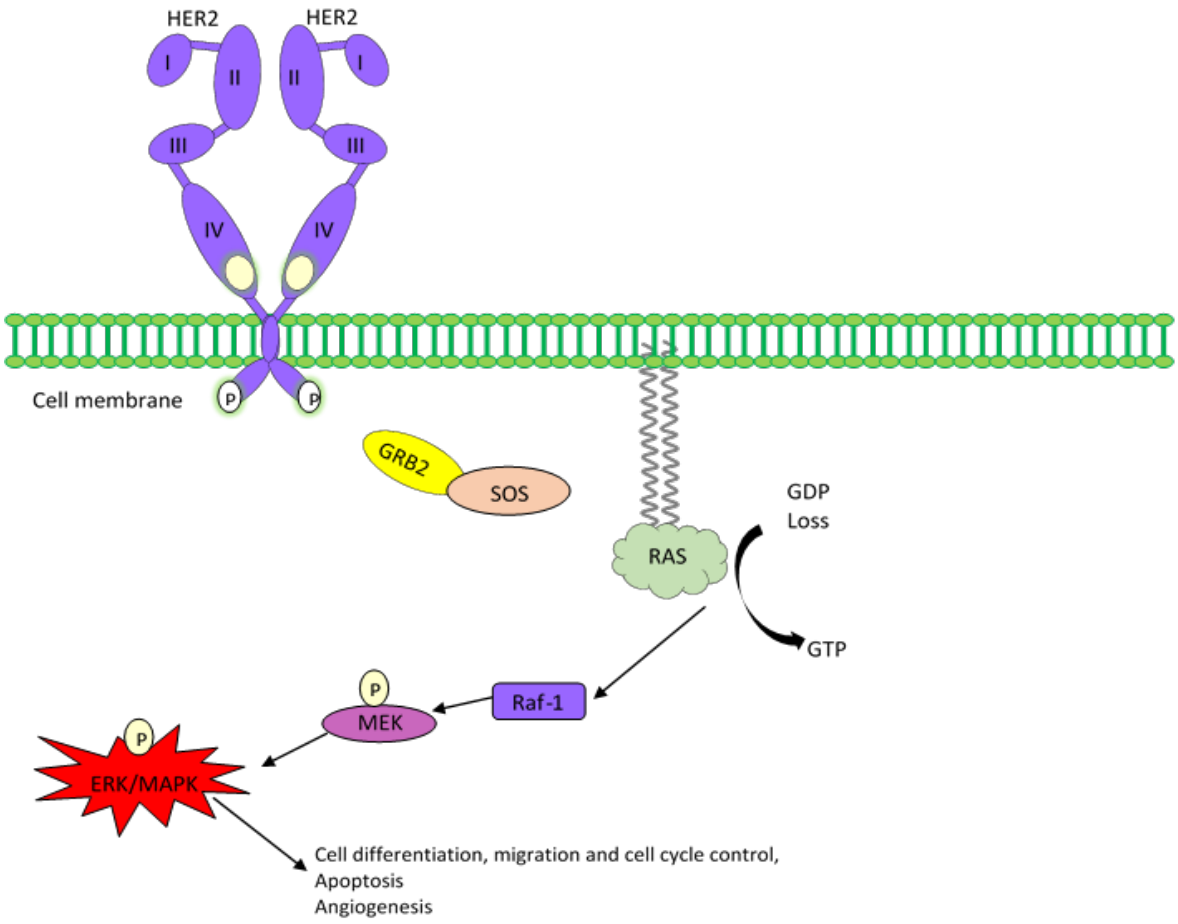


Figure 1.4 Schematic representation of the RAS/RAF/MEK/MAPK cascade. The MAPK pathway can be activated by dimerization of *HER2* with HER1, HER3 or HER4.

1.5 *HER2* as a prognostic factor in breast cancer

Gene amplification of *HER2* is the main mechanism by which abnormally high levels of the 185kDa glycoprotein are found in *HER2* overexpressed tumours (Wolff *et al.*, 2007). It is now recommended that the *HER2* status is determined for all invasive breast cancer patients (Dean-Colomb & Esteva, 2008). Determination of *HER2* status is important for a number of reasons. *HER2* positivity is associated with poor prognosis in patients with early breast cancer who do not receive any adjuvant systemic therapy. This makes *HER2* status a crucial factor in determining the type of treatment to administer to patients, particularly in cases where adjuvant therapy might be beneficial (Wolff *et al.*, 2007). *HER2* positivity is known to be associated with relative resistance to endocrine therapy, e.g. *tamoxifen*. *HER2* status has also been recently associated with resistance or sensitivity to various chemotherapeutic agents such as *nonathracycline* and *nontaxane*. Conversely, studies have shown that *HER2* positivity predicts a favourable response to *anthracycline*-containing chemotherapy regimens. *HER2* status may also play a role in predicting response to paclitaxel (Wolff *et al.*, 2007). More importantly, *HER2* positivity is known to predict a favourable response to *Trastuzumab* (Herceptin®, Genentech, Inc., South San Francisco, CA, USA; Hoffmann-La Roche Ltd., Basel, Switzerland), a humanised monoclonal antibody that targets the extracellular domain of *HER2* and inhibits tumour growth in vitro and in vivo (Saini *et al.*, 2011).

1.6 Testing for *HER2* status

Studies have shown a good correlation in the protein overexpression, gene amplification and mRNA levels of *HER2*. This allows for reliable testing for *HER2* status in cancer cells by immunistochemistry and FISH. These two methods are currently recommended for testing *HER2* status in breast cancer patients (Vogel, 2010; Wolff *et al.*, 2007).

1.6.1. Testing for *HER2* status at the protein level

1.6.1.1. Immunohistochemistry

Immunohistochemistry is currently the most widely used primary technique in the determination of *HER2* status (Moelans *et al.*, 2011). The use of immunohistochemical methods in testing for *HER2* in breast cancer allows tumours to be classified based on semi-quantitative methods of evaluation of *HER2* protein expression, such as the mouse monoclonal antibody clone CB11 (Novocastra Laboratories, Newcastle upon Tyne, England), the Rabbit monoclonal antibody clone SP3 (LabVision corporation, Runcorn, England), the Rabbit monoclonal antibody clone 4D5 (Ventana Medical Systems, Tuscon, Arizona), and the HercepTest (Dako, Glostrup, Denmark) which is based on the DAKO A0485 rabbit polyclonal antibody (Rhodes *et al.*, 2010; Vogel, 2010; Wolff *et al.*, 2007; van der Vegt *et al.*, 2009). Tumours are graded 0 to 3+ using the HercepTest score, an algorithm by which tumours are scored based on staining intensity and percentage of tumour cells stained. The HercepTest guidelines are used

to grade tumours as 0 (negative), 1+ (weak and incomplete membrane staining), 2+ (complete non-uniform membrane staining or with a weak intensity in more than 10% of the tumour cells) or 3+ (complete strong-positive membrane staining in more than 30% of tumour cells) (Moelans *et al.*, 2011). Patients with IHC scores of 0 and 1+ are considered negative for *HER2* and therefore not eligible for *Trastuzumab* therapy. Tumours with IHC scores of 2+ are considered to be equivocal cases, as several studies have shown that such tumours can be found to have *HER2* gene amplification. Patients with an IHC score of 3+ are considered to be a definite *HER2* positivity, and patients in this group are considered eligible for *Trastuzumab* treatment (Moelans *et al.*, 2011). The use of immunohistochemistry in testing for *HER2* status in patients is however laden with a number of drawbacks. Quality control and standardization of tests is critical in eliminating inter-observer variability. In contrast to FISH, IHC is prone to staining artefacts which may be caused by inappropriate tissue handling. Subjective interpretation of the intensity of membrane staining may also pose a major obstacle in the interpretation of IHC results (Moelans *et al.*, 2011; Vogel, 2010). In cases with equivocal IHC 2+ score, determination of treatment approach is usually after further analysis of the tumour by other methods of testing such as Fluorescence i*n situ* Hybridization (FISH) and chromogenic *in situ* hybridization (CISH).

1.6.1.2. The Enzyme-Linked Immunosorbant Assay (ELISA)

The Enzyme-Linked Immunosorbant Assay (ELISA) is used to measure protein levels in serum. The use of ELISA in testing for *HER2* status measures the amount of *HER2* in serum extracellular domain. Matrix metalloproteinases cleaved to the *HER2* receptor

protein can be detected in the extracellular domain that is released into the circulation. Elevated serum ECD levels (≥ 15ng/ml) is associated with poor prognosis, progressive metastasis, and a poor response to treatment. However, not all patients with *HER2* positive breast cancers appear to have elevated serum ECD levels, therefore the use of this method of testing in determination of treatment is not recommended in clinical settings (Moelans *et al.*, 2011).

1.6.2. Testing for *HER2* status at the DNA level

1.6.2.1. Fluorescence *In Situ* Hybridization (FISH)

Fluorescence *in situ* hybridization is a cytogenetic technique in which fluorescently labelled DNA probes are used to detect and visualise a specific DNA sequence with which they share a high degree of complementarity. Dual-colour FISH (Vysis Pathvysion, DAKO PharmDx) is the FDA-approved FISH testing kit, which hybridizes complementary *HER2* DNA on slides, and the resulting probes are visualised using a fluorescence microscope. In the current guidelines for *HER2* testing by FISH, a *HER2* gene signals to chromosome 17 signals ratio (FISH ratio) of ≤2:2 is considered a normal *HER2* expression (Wolff *et al.*, 2007). A FISH ratio of ≥2.2 counted in a minimum of 20 tumour cells, and in at least 2 invasive tumour areas, is considered positive for gene amplification. In equivocal cases of FISH assay, it is recommended that additional cells are counted or the test is repeated. In some cases, an IHC test may be required to confirm true equivocality. Due to the complexity of the FISH technique, and the fact that the scoring process is very time consuming, it is not a very practical primary

screening tool. In many clinical settings, the FISH technique is only used to determine treatment decisions in patients with equivocal IHC 2+ score (Moelans *et al.*, 2011).

1.6.2.2. Chromogenic *In Situ* Hybridization (CISH)

The Chromogenic in situ hybridization technique is a method of *HER2* testing developed by Tanner et al. as an alternative to FISH. CISH was approved by the FDA in 2008. CISH detects the *HER2* gene copies using a conventional immunoperoxidase reaction. This method of visualization of *HER2* allows scoring with a conventional light microscope. Using this method, a scoring system has been established where an average copy number of >10 or with big clusters in more than 50% of the tumour nuclei, is considered *HER2* amplification (Tanner *et al.*, 2000). A minimum of 30 tumour cells are counted. The CISH method has also been shown to correlate well with FISH and IHC methods (Moelans *et al.*, 2011). Most available CISH assays only score the *HER2* copy number, but recently, a new dual-colour CISH (Dako duo CISH kit) assay has been developed which allows for the detection of a *HER2* probe (red) and a chromosome 17 probe (blue), therefore allowing for the assessment of the *HER2* gene ratio relative to chromosome 17 signals (Moelans *et al.*, 2011).

1.6.2.3. Silver *In Situ* Hybridization (SISH)

The SISH technique is used for determination of *HER2* gene expression (Jacquemier *et al.*, 2013). SISH combines the accuracy of the FISH technique and the morphological control of the IHC technique, with the use of opaque silver, instead of fluorescent spot-

like signals. The resulting signal is a permanent result that can be visualised by an ordinary light microscope. The SISH technique is relatively new, and though studies have shown a high concordance between FISH and CISH techniques, further independent validation is recommended before the SISH test can be used in routine clinical settings (Gómez-Martin *et al.*, 2012; Jacquemier *et al.*, 2013; Shousha *et al.*, 2009).

1.6.3. Testing for *HER2* status at the RNA level

1.6.3.1. Quantitative Real-Time Reverse Transcription Polymerase Chain Reaction (qRT-PCR)

Due to the relative instability of RNA compared to DNA, and the severe degradation caused by cross-linking during fixation, RNA samples isolated from Formalin-fixed paraffin-embedded (FFPE) samples, the qRT-PCR method is not often used in *HER2* testing in clinical settings. This method, however, has been shown to correlate well with IHC and FISH techniques, even in RNA templates isolated from FFPE material. A very important use of qRT-PCR in cancer testing is the Oncotype DX assay (Genomic health, USA), a method by which the likelihood of disease recurrence in women with breast cancer is predicted. A recurrence score is obtained after analysing tumour cells using a panel of 21 genes, which include *HER2*, ER and PR (Moelans *et al.*, 2011).

The optimal method of testing for *HER2* in breast cancer remains controversial despite the development of various reliable tests. Most organisations rely on *HER2* 3+

positivity by immunohistochemistry as the preferred method for selecting patients for *Trastuzumab* therapy. (Vogel, 2010). The IHC technique has been shown to correlate well with the FISH technique, with only equivocal cases needing confirmatory tests, usually carried out by FISH (Meijer *et al.*, 2011). Therefore screening of newly diagnosed breast cancers is mostly performed by immunohistochemistry (Vogel, 2010). Accurate testing for *HER2* is important as false positive tests could lead to the administration of an expensive (£25,000 - £35,000 per patient per year) and ineffective treatment with very serious side effects; due to the cardiotoxicity of the drug, some patients who undergo *Trastuzumab* treatment are likely to develop cardiac dysfunction (Minami, Matsumoto & Horiuchi, 2010; Moelans *et al.*, 2011). A false negative test on the other hand, would deprive the patient of an important therapeutic option (Moelans *et al.*, 2011).

1.7 Therapies in *HER2* positive cancer

The current therapy for breast cancer patients with early and metastatic disease, involves the use of multiple agents. Endocrine therapies are administered to patients with hormone receptor-positive disease, anti-*HER2* therapies for *HER2*-overespressing tumours, and the poly-ADP-ribose polymerase (PARP) inhibitors are currently being developed in clinical trials, to treat patients with breast cancer gene 1 or 2 (BRCA1/2)-mutated tumours (Awada, Bozovic-Spasojevic & Chow, 2012; Eroles *et al.*, 2012). Accurate assessment of the *HER2* status of all invasive ductal breast carcinomas (IDC) is essential in determining the appropriate treatment regimen for individual cases. The inhibition of growth factor receptors can be achieved either by the use of monoclonal

antibodies which bind to the extracellular epitopes found in tumour cells, or the use of small tyrosine kinase inhibitors (TKIs) directed to extracellular epitopes and intracellular signalling pathways (Alvarez & Hortobagyi, 2013). These two complementary approaches have become part of the standard of care for patients with *HER2* positive breast cancer. A number of chemotherapeutic targets have been approved by the US Food and Drug agency (FDA) for use in the treatment of *HER2*-positive breast cancer.

1.7.1. *Trastuzumab* treatment for patients with *HER2* positive invasive breast cancer

Perhaps the most important reason for determining *HER2* status in patients is for the appropriate administration of the humanized anti-*HER2* monoclonal antibody therapy *Trastuzumab* to patients with invasive breast cancer that overexpress *HER2* (Bartlett *et al.*, 2001; Vogel, 2010). *Trastuzumab* was approved for use in September 1998, in the treatment of metastatic *HER2*-positive breast cancer. In January 2008, *Trastuzumab* was approved for use as first-line treatment for patients with early stage *HER2*-positive breast cancer (Piccart-Gebhart *et al.*, 2005; Romond *et al.*, 2005). More recently, *Trastuzumab* is also used in in the treatment of metastatic, unresectable *HER2*-positive gastric and gastro-oesophageal junction cancer (Bang *et al.*, 2010). *Trastuzumab* is also administered in both metastatic and adjuvant settings and is currently administered with chemotherapies involving paclitaxel and docetaxel which results in increased time to disease progression, and overall survival compared with Herceptin therapy alone (Kang *et al.*, 2008).

The *Trastuzumab* antibody consists of two antigen-specific sites that bind to the extracellular domain of *HER2*, resulting in the inhibition of *HER2* and apoptosis of tumour cells overexpressing *HER2* (Vogel, 2010; Geyer *et al.*, 2006), and the inhibition of the intracellular pathways involved in *HER2* activation (Stern, 2012) (Figure 1.5). The exact mechanism by which *Trastuzumab* exerts its antitumor activity has still not been fully elucidated. Several possible mechanisms have been proposed; these include the activation of antibody-dependent cellular cytotoxicity, inhibition of angiogenesis, blockage of proteolytic cleavage of the *HER2* extracellular domain and consequent downregulation of *HER2* receptors, disruption of downstream proliferative pathways, inhibition of cell cycle progression, inhibition of signal transduction in the intracellular domain of *HER2*, and the inhibition of repair of DNA damage caused by cancer treatment (Spector & Blackwell, 2009; Awada, Bozovic-Spasojevic & Chow, 2012).

The Initial approval of *Trastuzumab* in the treatment of *HER2* overexpressing breast cancers, was based on studies in patients with metastatic breast cancer (Dean-Colomb & Esteva, 2008). A phase III clinical trial by Slamon *et al.* (2001) compared the use of *Trastuzumab* plus various chemotherapeutic agents as a first-line treatment, with the use of chemotherapy alone in metastatic breast cancer. This study found that there was a significant improvement in overall survival (25.1 vs. 20.3 months) and overall response rate (50 vs. 32%), and a significantly longer median progression-free survival (6.9 vs. 3.0 months). These findings resulted in a 62% reduction in the risk of disease progression (Slamon *et al.*, 2001). Based on the results of the study, *Trastuzumab* was approved with paclitaxel as a first-line treatment of *HER2*-overexpressing metastatic breast cancer (Alvarez & Hortobagyi, 2013). Several independent randomized studies

have also shown that the addition of *Trastuzumab* to chemotherapy reduced the rate of recurrence by 50% in women with *HER2*-positive early breast cancer. In 2005, four clinical trials in patients with *HER2* overexpressing early breast cancer compared the effects administration of *Trastuzumab* as an adjuvant versus observation. The results showed that the recurrence rate of patients was reduced by a third, and the mortality rate was reduced by half (Wolff *et al.*, 2007). On the basis of these results, the UK National Institute for Health and Care Excellence (NICE) and the US Food and Drug Administration (FDA) and the European Medicines Agency (EMA) approved the adjuvant use of *Trastuzumab* with chemotherapy in patients with *HER2* overexpressing early breast cancer (Pivot *et al.*, 2013; Wolff *et al.*, 2007; Emde, Köstler & Yarden, 2012; Goddard *et al.*, 2012).

1.7.2. *Lapatinib*

In March 2007, *Lapatinib* (Tykerb, GlaxoSmithKline, Philadelphia, PA, USA) was approved for use in the metastatic *HER2*-positive breast cancers. *Lapatinib* is a small molecule reversible dual EGFR/HER1 and *HER2* tyrosine kinase inhibitor administered in combination with capecitabine for the treatment of advanced metastatic *HER2*-positive breast cancers where the patients had been given prior therapy including anthracycine, a taxane, and *Trastuzumab*, and where these therapies have failed (Geyer *et al.*, 2006; Awada, Bozovic-Spasojevic & Chow, 2012; Wolff *et al.*, 2007). In January 2010, *Lapatinib*, in combination with Letrozole (Femara®), was also approved for the treatment of *HER2*-positive breast cancer patients whose tumours were also hormone receptor-positive, and for whom hormonal therapy is indicated (Alvarez & Hortobagyi, 2013). *Lapatinib* binds to the intracellular tyrosine kinase domains of HER1

and *HER2*, and selectively inhibits HER1 or *HER2* overexpressing tumour cells, resulting in the inhibition of phosphorylation, and the inhibition of downstream pathways which lead to cell proliferation and cell survival (Awada, Bozovic-Spasojevic & Chow, 2012) (figure 1.5).

1.7.3. Pertuzumab

In June 2012, pertuzumab was approved for use in combination with *Trastuzumab* and docetaxel, in the treatment of patients with *HER2* positive breast cancer who had not received prior *HER2* therapy or chemotherapy for metastatic disease (Baselga *et al.*, 2012).

Pertuzumab is an anti *HER2* antibody that binds to subdomain II of the *HER2* extracellular domain (Franklin *et al.*, 2004). Pertuzumab functions in the prevention of *HER2* dimerization with ligand-activated *HER2* receptors, mostly HER3 (Baselga & Swain, 2009; Agus *et al.*, 2002). *Trastuzumab* is directed towards ligand-independent *HER2* signalling (*HER2*/*HER2* interactions), while Pertuzumab interferes with activation of *HER2* via ligand-dependent HER3-mediated signalling (Scheuer *et al.*, 2009) (figure 1.5). A combinatorial therapy involving *Trastuzumab* and pertuzumab in *HER2* overexpressed breast cancer has been associated with significant antitumor activity (Scheuer *et al.*, 2009; Nahta, Hung & Esteva, 2004).

Other chemotherapeutic agents including *Trastuzumab* Emtansine (T-DM1), and Neratinib, have shown significant activity in the inhibition of cell membrane receptors

in clinical trials, but have not yet been approved for clinical practice (Alvarez & Hortobagyi, 2013).

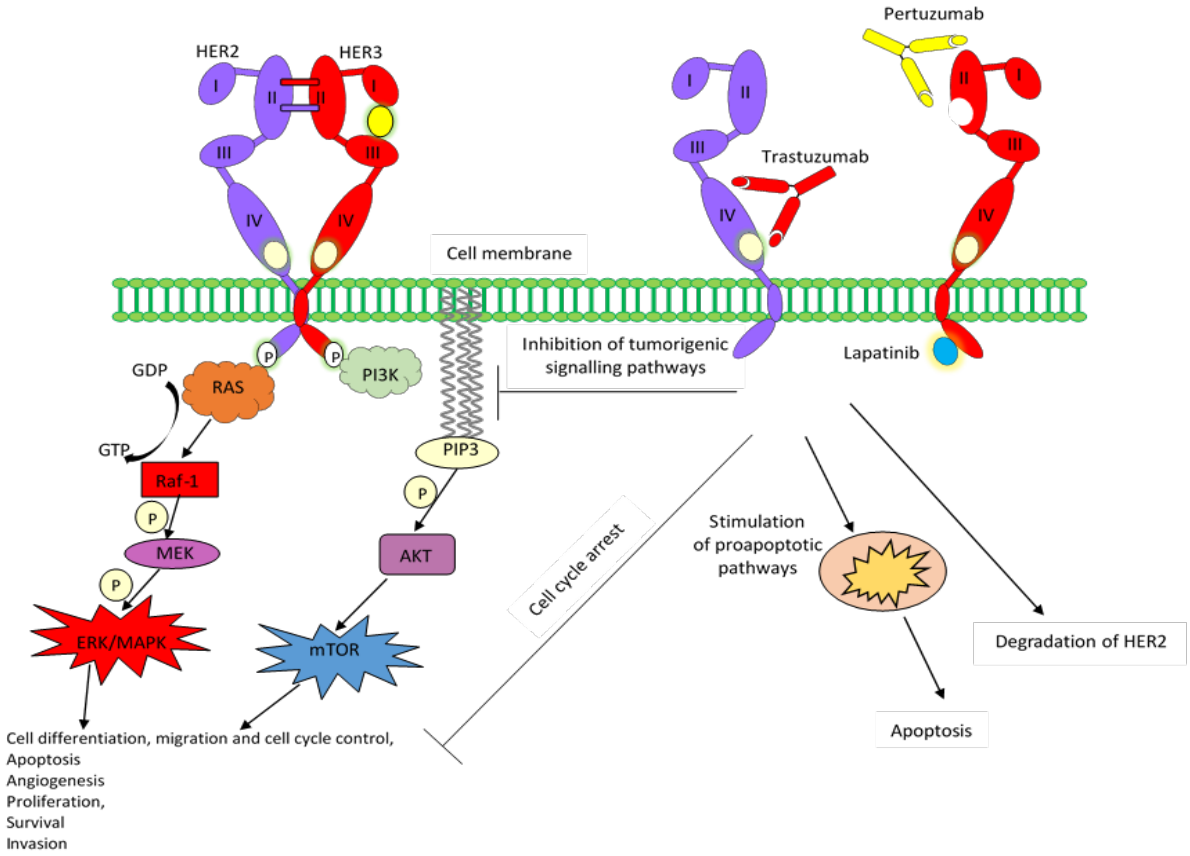


Figure 1.5 Schematic representation of the mechanisms of action of current therapies for *HER2* overexpressing breast cancer. Downstream signalling is inhibited by monoclonal antibodies *Trastuzumab* and pertuzumab, which inhibit dimerization by binding to subdomains IV and II of HER receptors respectively. *Lapatinib*, a small-molecule tyrosine kinase inhibitor, inhibits phosphorylation by binding to the intracellular kinase domain of *HER2* receptors.

1.8 Challenges and unmet needs in the treatment of *HER2* positive breast cancer

Although the administration of *Trastuzumab*-based treatment has become the standard of care for patients with *HER2* positive metastatic breast cancer, for reasons that are unclear, many patients will either not respond to *Trastuzumab* treatment, or eventually progress despite treatment (Awada, Bozovic-Spasojevic & Chow, 2012; Koletsa *et al.*, 2008). It is estimated that up to 40% of patients with metastatic breast cancer do not respond to *Trastuzumab* or combinatorial therapies which include *Trastuzumab* (Awada, Bozovic-Spasojevic & Chow, 2012). Patients who achieve an initial response to Herceptin plus chemotherapy generally acquire resistance within 1 year (Freudenberg *et al.*, 2009; Awada, Bozovic-Spasojevic & Chow, 2012; Kang *et al.*, 2008). As a result, significant efforts have been applied to elucidating the mechanisms underlying Herceptin resistance, and finding other therapies besides Herceptin for the treatment of *HER2* positive breast cancer.

Development of resistance to *HER2* therapies poses a serious concern, and ultimately results in shorter time to tumour progression, and limited overall survival. There are several mechanisms by which *HER2* overexpressing tumour cells may develop resistance to *HER2* therapies. Cross-talk between intracellular signalling pathways and redundancy in growth factor receptors have been implicated in the development of resistance in most patients with *HER2*-positive breast cancer. For example, *Trastuzumab* may be ineffective in inhibiting PI3K due to lateral activation of the pathway by other members of the HER family (i.e. HER1 and HER3), thereby leading to

continuing cell proliferation (Kerbel, 2009). Alterations in receptor-antibody interactions may lead to resistance, either through mutations in *HER2* that disrupt binding, the masking of antigens on the tumour cell surface through glycoproteins such as MUC-4, or the overexpression of truncated isoforms of *HER2*, e.g. p95 that lacks the extracellular domain, and therefore does not have a *Trastuzumab* binding site (Sperinde *et al.*, 2010). In addition, mutations in PTEN which result in loss of function, or activating mutations in PI3K, may lead to enhanced phosphorylation and signalling of AKT which is downstream of *HER2*, therefore bypassing any *HER2*-directed therapy, and resulting in cellular proliferation of *HER2* (Coughlin *et al.*, 2010). Mutations or conditions that lead to the loss of tumour suppressor gene Cyclin-dependent kinase inhibitor 1B (*p27*) may contribute to *Trastuzumab* resistance. The phosphorylation of the *p27* gene prevents the degradation of *Trastuzumab* and leads to cell cycle arrest (Bedard, de Azambuja & Cardoso, 2009). The *p27* gene is therefore crucial in the efficiency of *Trastuzumab* via the inhibition of *HER2*. In addition, the amplification or overexpression of cyclin E will result in increased proliferation and increased tumour growth, which may subsequently result in resistance and decreased sensitivity to *Trastuzumab* (Scaltriti *et al.*, 2011).

Due to shortcomings with current targeted therapies such as *Trastuzumab* and *Lapatinib*, a variety of novel and improved targets are being investigated for the treatment of *HER2* positive breast cancer, many of which have the potential to address the unmet needs of current treatment regimens (Awada, Bozovic-Spasojevic & Chow, 2012). Some characteristics of ideal novel targets would include significant antitumor activity, good tolerability, a limited propensity for the development of drug resistance,

good selectivity for the chosen therapeutic targets, potent inhibition of commonly expressed molecular targets, and such treatments would display irreversible binding to its molecular targets, thereby producing longer-lasting effects (Awada, Bozovic-Spasojevic & Chow, 2012). Recently, several agents have been developed which have the potential to inhibit *HER2* in breast tumours, either as monotherapy or in combination with other therapies. These novel therapies include the tyrosine kinase inhibitors neratinib (HKI-272) and afatinib (bibw-2992), and the anti-*HER2* monoclonal antibodies pertuzumab and *Trastuzumab*-MCC-DM1 (T-DM1)(Jones & Buzdar, 2009). The use of agents that target molecular pathways such as the Vascular Endothelial Growth Factor (VEGF) receptor, mammalian target of rapamycin (mTOR), PI3 kinase, insulin growth factor receptors (IGFRs), HSP-90 and other important kinases (Awada, Bozovic-Spasojevic & Chow, 2012) may also be useful as alternative targets for *HER2* therapy. However, innovative clinical studies using well characterised clinical subjects will be required in order to establish the true clinical value of these potential novel anti-*HER2* targets (Awada, Bozovic-Spasojevic & Chow, 2012).

1.9 Alternative Splicing

RNA splicing is a process in the nucleus by which introns are removed from pre-mRNA and exons are ligated together, converting pre-mRNA to mature mRNA (Watson et al., 2008). Some pre-mRNAs in individual genes may be differentially spliced, generating multiple alternative mature mRNA products from a common mRNA precursor, known as isoforms. Alternative splicing is the process by which more than one mRNA is produced by a single pre-mRNA, leading to the production of several structurally

distinct protein isoforms, which can have diverse functions (Garcia-Blanco, 2005). Alternative splicing is a key post-transcriptional mechanism by which the expression of multiple protein products from a single gene can be controlled (Li *et al.,* 2006) and is considered as one of the key generators of proteomic diversity (Ladomery, Harper & Bates, 2007); alternative splicing increases the protein diversity without increasing genome size, and therefore forms one of the most significant components of the complexity of the human genome (Modrek & Lee, 2002). It is estimated that alternative splicing can occur in up to 75% of all human genes (Watson et al., 2008). The number of different variants a given gene can encode by alternative splicing varies from two to thousands. For example, the rat *Slo* gene encodes a potassium channel which is expressed in neurons, and has the capacity to encode up to 500 alternative transcripts of the gene, and one *Drosphilia melanogaster* gene has the potential to encode up to 38,000 products as a result of alternative splicing (Watson, et al., 2008).

In the majority of human genes, alternative splicing may give rise to transcript variants and/or protein isoforms that can vary distinctly in structure and functional properties (Figure 1.6). For example, an alternative splice variant of *VEGF* gives rise to multiple protein isoforms, which display either pro-angiogenic or anti-angiogenic activities (Nowak *et al.*, 2008; Ladomery, Harper & Bates, 2007).

Alternative splicing is the most plausible solution for the miscorrelation between the number of genes transcribed in eukaryotic cells and the number of proteins translated in the same signalling pathway (Kim, Goren & Ast, 2008; Pal, Gupta & Davuluri, 2012).

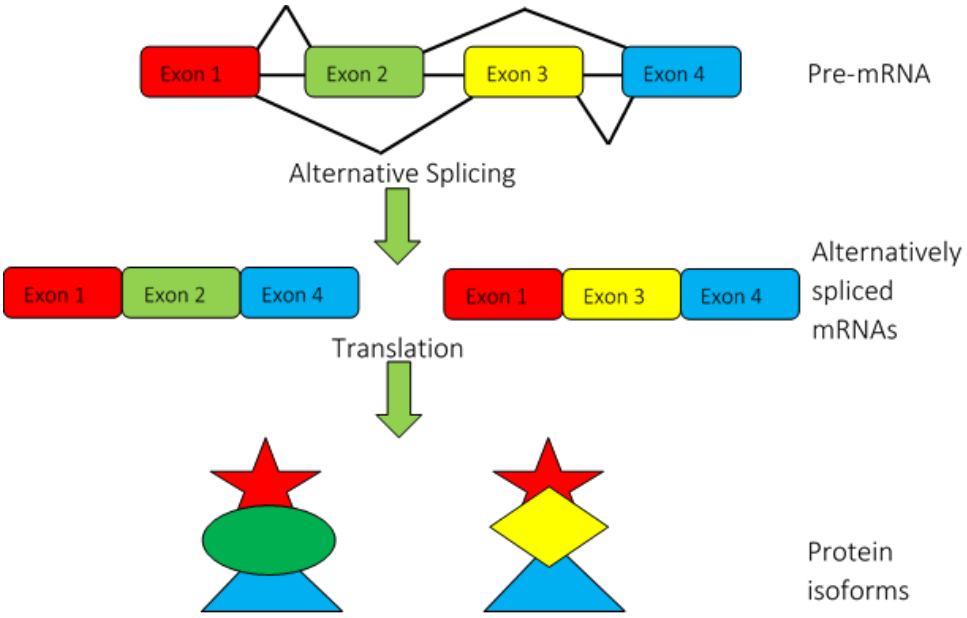


Figure 1.6 Schematic representation of alternative splicing. Due to alternative splicing, a single gene can produce more than one transcript, which can be translated into different protein isoforms.

Alternative splicing plays various important roles in the cell, chief of which is to increase the diversity of the proteome and transcriptome by allowing the generation and expression of multiple mRNA products from a single gene. Due to alternative splicing, the coding capacity of the human genome is greatly increased (Tazi, Bakkour & Stamm, 2009), leading to an increased complexity of the transcriptome and proteome (Stamm *et al.*, 2005). This increased complexity of the human genome has clear repercussions for the regulation of gene expression in many organisms and for the balance between human health and disease (Garcia-Blanco, 2005; Stamm *et al.*, 2005). Today, alternative splicing is increasingly linked with the aetiology of cancer (Ladomery, Harper & Bates, 2007).

Regulation of alternative splicing

A number of factors are known to play a crucial role in alternative splicing. Of these known factors, the most studied are RNA-binding proteins and transcription factors. RNA-binding proteins, also known as regulatory proteins, play a wide role in mRNA biogenesis (David & Manley, 2008). When bound to exons, Serine and Arginine-rich (SR) proteins tend to promote exon inclusion, while heterogeneous nuclear ribonuclear proteins (hnRNPs) modulate exon skipping (David & Manley, 2008). Pre-mRNA splicing is catalysed by the spliceosome, a ribonuclearprotein complex composed of five small nuclear ribonuclearprotein particles (hnRNPs), and a number of accessory polypeptides, in a sequential and highly coordinated pathway (Smith & Valcárcel, 2000).

The spliceosome begins its function by recognising the consensus elements on both ends of the intron; the U1 Small nuclear ribonucleoprotein (snRNP) recognises the 5' splice site, while the 65 kDa subunit of the U2 snRNP (U2AF65) binds to the polypyrimidine tract, and the 35kDa subunit of the U2 (U2AF35) binds to the 3' splice site. Bridging interactions between U1 snRNP bound to the 5' splice site and U2AF bound to the 3' splice-site region are known to be modulated by SR proteins (Smith & Valcárcel, 2000). The SR proteins are involved in both constitutive and regulated splicing (Ladomery, Harper & Bates, 2007). SR proteins contain N-terminal RNA recognition motifs, which mediate binding to the pre-mRNA. Their C-terminal recognise exonic splice enhancers (ESEs) and through protein-protein interactions, bind to the snRNP U2AF35 and U1 snRNP, therefore promoting U1 and U2AF binding

to splice sites (Smith & Valcárcel, 2000; Ladomery, Harper & Bates, 2007). Smaller regulatory complexes also play a role in achieving cell-type-specific splicing; exonic splice enhancers (ESEs) or silencers (ESSs) promote or inhibit exon inclusion of proximal exons, while intronic splice enhancers (ISEs) or silencers (ISSs) enhance or inhibit the use of exon from an intronic location, respectively (Wang & Burge, 2008; Smith & Valcárcel, 2000). Apart from the functional role of regulatory proteins, changes in their expression levels or post-transcriptional changes may alter their activities, providing a means for the regulation of alternative splicing (David & Manley, 2008). For example, hnRNPA1 which functions in the inclusion of many alternative exons, may become phosphorylated upon osmotic shock, resulting in its cytoplasmic accumulation, which may in turn lead to changes in alternative splicing (Allemand *et al.*, 2005).

Transcription factors have also been studied, and may have a potential influence on alternative splicing, either by influencing the concentration of direct regulators of alternative splicing (e.g. SR proteins and snRNPs), or by altering the rate of RNA polymerase II elongation, leading to indirect effects on alternative splicing. It is therefore important to identify the tissue-specific changes in transcription factor expression during splicing, and the instances in which these changes alter patterns of alternative splicing.

1.9.1. The role of alternative splicing in the development of cancer

The abnormal generation of mRNA splice variants has been implicated in the oncogenic tendencies of several important biomarkers to include vascular endothelial growth factor (*VEGF*) and Wilms Tumour 1 (*WT1*); with evidence accumulating to support the theory that some splice variants are more oncogenic than others. This evidence is used to explain why in some studies *WT1* appears to act as a tumour suppressor, whilst in others it appears to be operating as an oncogene. Many protein isoforms produced via alternative splicing are tightly regulated during normal development, but may be misregulated in cancer cells. Aberrant expression of alternative splice variants in many genes has been linked with disease progression and prognosis, and cancer cells may manipulate the mechanisms that regulate drug resistance and patient survival (Pal, Gupta & Davuluri, 2012). Recent studies have also demonstrated that the modulation of alternative transcript expression in various genes may impede tumour growth and act as a model for targeting disease at the isoform level (Pal, Gupta & Davuluri, 2012). There are various modes of alternative splicing. The most common types include competing 5' splice sites, competing 3' splice sites, cassette exons, mutually exclusive exons, and retained introns (Ladomery, Harper & Bates, 2007; Scaltriti *et al.*, 2007) (Figure 1.7).

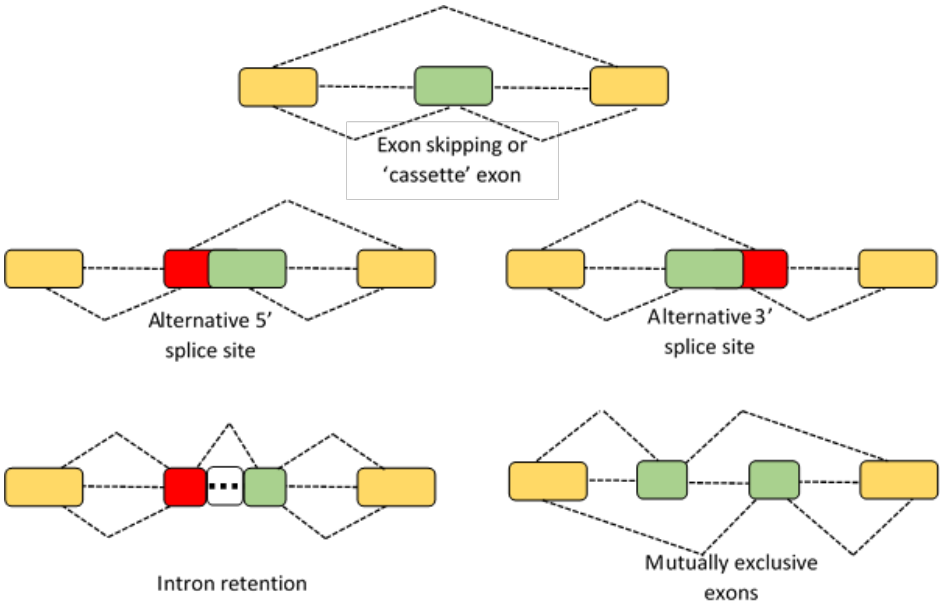


Figure 1.7 Modes of alternative splicing showing exon skipping, alternative 5' and 3' splice sites, intron retention and mutually exclusive exons (adapted from Ladomery, Harper & Bates, 2007 and used with permission of the author).

1.10 Alternative splicing of *HER2* and *HER2* Splice Isoforms

1.10.1. Herstatin

1n 1999, *Doherty et al*. described a truncated alternative *HER2* transcript which is a secreted protein of ≈68kDa with growth inhibitory properties. Herstatin is a naturally occurring *HER2* protein and is generated from alternative *HER2* mRNA transcripts that retain the intron 8 (Doherty *et al.*, 1999b). Herstatin is a soluble protein, which can be secreted from cells it has been produced by, and lacks a transmembrane domain (Koletsa *et al.*, 2008). Herstatin, also described as p68 *HER2*, and dimercept, delineates 340 amino acid residues identical to subdomain I and II of the extracellular domain of p185 *HER2*. The extracellular domain precedes a unique C-terminal sequence of 79 amino acids encoded by intron 8, which functions as a receptor binding domain (Figure 1.8).

Herstatin is expressed in foetal kidney and liver, and in normal tissues, and is considered as a growth regulatory factor during normal development (Kolesta *et al.,* 2008). Herstatin mRNA and protein have been shown to be expressed in the noncancerous breast tissue, in areas adjacent to breast carcinomas (Kolesta *et al.,* 2008). Recent evidence suggests that Herstatin has a potential significance in the regulation of the wild type (p185) *HER2* in normal and malignant development, due to its specific inhibitory effect (Doherty *et al.,* 1999). When bound to p185*HER2*, Herstatin disrupts the dimerization of *HER2* with itself and other *HER2* full length receptor homologues, and results in a noticeably reduced tyrosine phosphorylation of *HER2* (Wang *et al.*, 2013). Herstatin may therefore have tumour-supressing activity of down-regulating the expression of p185*HER2* through the inhibition of receptor dimerization and the consequent inhibition of tumour formation (Jackson *et al.*, 2013).

Herstatin has also been shown to bind to EGFR, HER3 and HER4, and in a way similar to pertuzumab, blocks homomeric and heteromeric receptor interactions (Justman & Clinton, 2002; Koletsa *et al.*, 2008). More so, In a similar manner to *Trastuzumab*, binding of Herstatin to the extracellular domain of *HER2* leads to degradation of *HER2* receptor and endocytosis (Wang *et al.*, 2013). Herstatin specifically blocks dimer phosphorylation by disrupting *HER2*/HER3 and EGFR/*HER2* dimers. Although the majority of tumours express Herstatin mRNA, the Herstatin protein is absent in 75% of breast cancers. This may be because the cancer cells are protected by an intrinsic mechanism against the growth-inhibitory effects of Herstatin (Koletsa *et al.*, 2008).

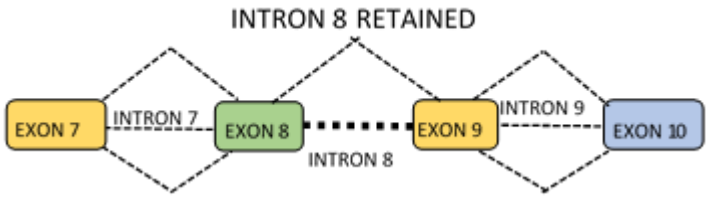


Figure 1.8 Schematic representation of Herstatin showing the retention of intron 8.

1.10.2. *HER2*Δ16

An alternative splice form of the human *HER2* containing an in-frame skipping of exon 16, a 48bp cassette exon, has been detected in human breast cancers (Kwong & Hung, 1998; Castiglioni *et al.*, 2006). This splice variant, designated *HER2*Δ16encodes a receptor that lacks the amino-acids encoded by exon 16, which is a small region of the extracellular domain of *HER2* (Jackson *et al.*, 2013; Mitra *et al.*, 2009) (Figure 1.9). Exon 16 immediately precedes the transmembrane domain and contains two cysteine residues, and lacks the amino acids 634 – 649 in domain IV of the *HER2* extracellular domain (Wang *et al.*, 2013; Kwong & Hung, 1998). The resultant loss of these cysteine residues in the extracellular domain of *HER2* leads to a change in the conformation of the *HER2* receptor extracellular domain promoting the homodimerization of stable receptors capable of transforming cells, via the formation of disulphide bonds (Jackson *et al.*, 2013; Castiglioni *et al.*, 2006). *HER2*Δ16 is purported to constitute a more aggressive *HER2* variant compared to the wild type, p185 *HER2*, and is said to play a crucial role in the malignant transformation and disease progression of *HER2* positive breast cancers. This suggests that patients expressing *HER2*Δ16 may benefit from more aggressive therapy. *HER2*Δ16 has also been linked to *Trastuzumab* resistance, advocating the use of tyrosine kinase inhibitors as an alternative therapy. Studies by

Castiglioni *et al*. show that the *HER2*Δ16 splice variant comprises about 9% of the total *HER2* mRNA of a collection of 46 breast carcinomas tested, with *HER2* expression levels ranging from 0 to 3+, as determined by the HercepTest. These studies showed the exon 16 skipped *HER2* variant to have much stronger transformation activity than the wild type *HER2* (Castiglioni *et al.*, 2006). Mitra *et al.* report that *HER2*Δ16 expression promotes cell invasion and *Trastuzumab* resistance through direct coupling of *HER2*Δ16 to Src kinase (Mitra *et al.*, 2009). There is evidence that alternative splicing leading to cassette exons within the extracellular domain of some growth factor receptors provides a unique mechanism for the generation of novel isoforms which may encode potentially active molecules (Baek *et al.*, 2004; Collesi *et al.*, 1996; Li *et al.*, 1995). For this reason, the *HER2*Δ16 variant represents a constitutively active form of *HER2* similar to the mutated gene (Marchini *et al.*, 2011). This provides further evidence to the theory that mutations in splice variants affect the progression from normal breast cells to invasive cells, rather than increase in receptor numbers alone (Castiglioni *et al.,* 2006). Therefore, a better understanding of the involvement of *HER2* splice variants in the response or resistance to current therapies targeting the *HER2* receptor might be crucial in improving response rates in patients with *HER2* overexpressing cancers (Castiglioni *et al.*, 2006).

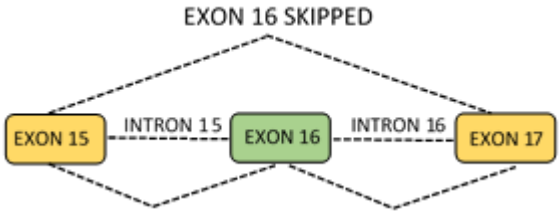


Figure 1.9 Schematic representation of *HER2*Δ16 showing the cassette exon on exon 16.

1.10.3. P100 *HER2*

The p100 *HER2* transcript was first described by Scott et al. as a spliced variant of *HER2* which encodes a 2.3kb protein constituting only the extracellular domain of the full length protein (Scott *et al.*, 1993; Jackson *et al.*, 2013). P100 *HER2* arises via an in-frame stop codon which results from the retention of intron 15, and has been found to interfere with oncogenic activity, with the capacity to inhibit cell proliferation (Jackson *et al.*, 2013) (Figure 1.10). The 5' end of the p100 *HER2* gene is a 2.1kb segment of the truncated *HER2* transcript homologous to the 5' end of the full length *HER2* transcript, while the 3' end diverges to reveal an exonic extension with an in-frame stop codon and a poly(A) addition site (Scott *et al.*, 1993). The alternative polyadenylaton signal permits reading into the intron which results in the translation of a 100kDa isoform of *HER2*. The truncated *HER2* transcript encodes a 100-kDa variant of the full length 185-kDa transmembrane receptor that, with the exception of 20 C-terminal amino acids, contains the entire *HER2* extracellular domain (Scott *et al.*, 1993). Further exploration into the role of p100 *HER2* revealed that this secreted truncated form of *HER2* gives rise to a decrease in the downstream signal transduction pathway such as the MAPK cascade. Being a secreted protein, the p100 *HER2* isoform may also serve as a serum biomarker, and may be crucial in making treatment decisions.

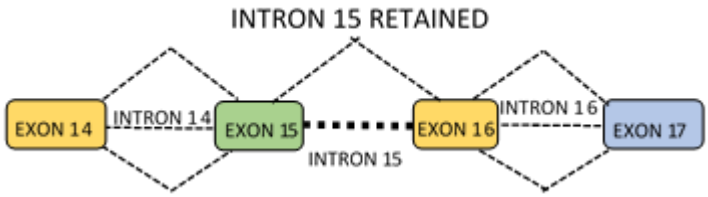


Figure 1.10 Schematic representation of p100 *HER2* showing the retention of intron 15.

1.11 Hypothesis and objectives of this Study

1.11.1. Hypothesis

Studies have shown the existence of alternative splice variants of *HER2*, and in addition, that *HER2* splice variants may serve various crucial functions in continuing disease progression of *HER2* overexpressed or amplified breast carcinomas despite current therapies. The lack of activating mutations in *HER2* positive breast cancers and the high proportion of patients who do not respond to current therapies, suggests that it is not only the total number of *HER2* receptors that is responsible for malignant transformations in the cell. The expression of alternative splice variants of *HER2* resulting in protein isoforms with potent cellular functions, may lead to the development of specific assays to include splice-variant status of certain important isoforms, in the aim of improving diagnosis and treatment. Mutations in the oncoprotein, in particular small deletions in the extracellular domain, may also result in the formation of disulphide bonds. Current investigations into the three identified naturally occurring *HER2* splice variants (*HER2*Δ16, Herstatin and p100 *HER2*) has been geared on developing new therapeutic strategies to tackle issues with treatment

resistance and to further understand impact of *HER2* overexpression with particular focus on these splice variants (Jackson *et al.*, 2013).

The hypothesis of this study is that in addition to the previously identified roles of *HER2* alternative splice variants in the progression and continued transformation of invasive breast cancer, currently unidentified alternative splice variants may be a contributing factor in the poor response of patients to current treatment regimes. The discovery of novel alternative splice variants may lead to identification of *HER2* isoforms which play crucial roles in the induction or inhibition of *HER2* signalling. The use of these splice variants in the manipulation of *HER2* signalling may be of therapeutic benefit to patients diagnosed with *HER2* positive breast cancers.

1.11.2. Objectives

1. To develop specific standard RT-PCR primers for the detection of novel splice isoforms of *HER2* in invasive breast cancer cell lines.
2. To verify the identity of amplified PCR products by DNA sequencing.
3. To identify the functions of novel splice variants, relating splice variant expression to clinical parameters, using bioinformatics analysis.
4. To identify the expression of novel splice variants of *HER2* in clinical breast cancer samples categorised with 0 to 3+ by immunohistochemistry.
5. To develop double dye (Taqman) probes for the detection of novel alternative *HER2* splice variants by quantitative real-time PCR.
6. To study the regulation of *HER2* and *HER2* splice variants by manipulating cell culture conditions, e.g. induction of hypoxia.

7. To study the effects of tyrosine kinase inhibitors *SRPIN340*, *TG003* and *INDY* on the regulation of *HER2* and *HER2* alternative splice variants.
8. To develop siRNAs to knockdown certain splice factors which may be known to play a role in *HER2* splicing, and to relate the effects of gene silencing on the expression and regulation of *HER2* and *HER2* splice variants.

CHAPTER 2. GENERAL MATERIALS AND METHODS

2.1. Antigen Retrieval Immunohistochemistry

2.1.1. Buffers used in Immunohistochemistry

Sodium Citrate buffer (pH 6.0) was used as the antigen retrieval buffer. 0.1M Sodium Citrate buffer was freshly prepared by adding 29.4g of Sodium Citrate and 54ml of 1M HCl to 10L of distilled water and stirring vigorously. The pH value of the buffer solution was adjusted using a pH meter by adding a few drops of freshly prepared 1M HCl or 0.1M NaOH solution to obtain the desired value. The buffer was then stored at 4°C (Kong *et al.*, 2006).

TRIS Buffered Saline (TBS) pH7.4 was used as the washing buffer. 0.005M TBS was prepared by adding 80g of Sodium Chloride, 6.05g of TRIS and 44ml of 1M HCl to 10L distilled water. All antibodies were diluted in TBS (Kong *et al.*, 2006) (Table 2.1).

ANTIBODY CLONE	SPECIFICITY	SOURCE	CONCENTRATION
SP3	*HER2*	Labvision Corporation, UK	0.6 µg/mL
6F11	ER	Novocastra Laboratories, UK	0.8µg/mL
CB11	*HER2*	Novocastra Laboratories, UK	0.5µg/mL
PGR636	PR	Dako Uk Ltd	0.5 µg/mL

Table 2.1 Antibodies used in immunohistochemistry, their specificities and concentrations.

2.1.2. Antigen retrieval microwave heating technique

Tissue sections 5 microns thick mounted on slides were prepared for antigen retrieval by deparaffinising in two changes of Histoclear (5 minutes each), rehydrating in two changes of 100% Industrial methylated spirit (3 minutes each) and rinsing in tap water for a few minutes. The sections were treated with 3% hydrogen peroxide to block endogenous peroxidise for 5 minutes, and rinsed thoroughly in distilled water. The slides were then placed in a plastic slide rack and transferred into a plastic jar containing the antigen retrieval buffer solution and the slide rack was completely covered in antigen retrieval buffer solution. The plastic jar was covered with a loose-fitting cap and the slides were heated in a microwave oven at a temperature of 100°C.

The duration of heating was 25 minutes in Sodium Citrate buffer (pH 6.0). After heating the sections were washed and cooled to room temperature in running tap water and rinsed in TBS before immunostaining (Hayat, 2002).

2.1.3. Immunohistochemical staining methods

All immunohistochemical assays were performed at room temperature (20-25°C). Although most automated immonhistochemical systems suggest that conducting immunohistochemical staining at 4°C may improve antigenicity, in most instances, the use of room temperature is generally acceptable and has been known to give highly reproducible results (Vogel, 1901). Antibodies used for immunohistochemical staining are listed in table 2.1. The Vector Elite ABC system for immunohistochemistry, with a peroxidase label (Vector Labs, Uk) was used as the detection kit. Diaminobenzidine - DAB (Dako Uk Ltd) was used to visualise the labelled antibodies.

The sections were incubated in normal goat serum for 5 minutes to prevent non-specific binding of the antibodies, and then incubated in primary antibody for 60 minutes. After staining with the primary antibody, sections were then incubated with a linking secondary biotinyilated antibody for 30 minutes. Sections were washed after each stage with TBS at pH 7.4. The slides were then incubated in the enzyme-conjugated ABC label for 30 minutes, rinsed in TBS and washed in distilled water. Conjugation of the primary antibody with biotin is a method of labelling which, due to the high affinity of avidin to biotin, enables the visualization of proteins when bound to fluorescent or enzyme-labelled avidin. The resulting avidin-biotin (ABC) complex forms a stable complex, and is a highly sensitive method for detection of the primary

antibody. Finally the sections were treated with the chromogen 3,3' diaminobenzidine (DAB) for 10 minutes, washed in running tap water, counterstained in Harris' haematoxylin for 2 minutes, differentiated in 1% acid alcohol for a few seconds, left to blue in running tap water and dehydrated in alcohol, cleared and mounted in DPX for examination under a microscope (Hayat, 2002).

2.1.4. Evaluation of Immunohistochemistry results

Slides were examined using Nikon Eclipse 50i microscope. Photomicrographs were taken with a Nikon Digital Sight DS-UI camera.

2.2. Cell culture

Human breast and ovarian cancer cell lines were obtained from the European Collection of Cell Cultures (ECACC; Salisbury, UK) and the American Type Culture Collections (ATCC; Middlesex, UK). In order to eliminate contaminants arising from high passage numbers, all six cell line stocks were frozen down at passages 5 to 10, and cells were cultured up to a maximum of 10 passages before being replaced by frozen stocks (van Staveren *et al.*, 2009). The cell lines were cultured in their respective growth media in cell culture flasks according to their manufacturers' recommendations, in a humidified atmosphere.

All cell culture media were supplemented with 10% Foetal Bovine Serum (Biosera, UK), 2mM L-Glutamine (Lonza, UK) and 100u/ml penicillin-streptomycin solution (Lonza, UK). All cells lines were adhering cells. Upon reaching confluence, cells were washed in

Phosphate Buffered Saline (PBS) and removed from the bottom of the flasks by trypsinization in 1x trypsin/EDTA solution (Lonza, UK) (Goren *et al.*, 2010).

CELL LINE	**ORIGIN**	***HER2* STATUS**	**CULTURE MEDIA**	**CULTURE CONDITIONS**
SKOV3	Ovarian Cancer	*HER2* 3+	McCoy's 5A Medium (Lonza, UK)	37°C, 5% CO_2
SKBR3	Breast Cancer	*HER2* 3+	McCoy's 5A Medium (Lonza, UK)	37°C, 5% CO_2
BT-20	Breast Cancer	*HER2* -	DMEM (Lonza, UK)	37°C, 5% CO_2
MCF-7	Breast Cancer	*HER2* -	DMEM (Lonza, UK)	37°C, 5% CO_2
MDA-MB-453	Breast Cancer	*HER2* 2+	DMEM (Lonza, UK)	37°C, 5% CO_2
MDA-MB-361	Breast Cancer	*HER2* 2+	DMEM (Lonza, UK)	37°C, 5% CO_2

Table 2.2 Cell line models used in cell culture studies and their culture conditions.

The cell lines used in this study represent samples with varying levels of *HER2* expression, and were therefore considered a good model for *HER2* expression analysis, based on the repository of cells available for selection.

2.3. RNA extraction

2.3.1 RNA extraction from cell cultures

Total RNA was extracted from cells using TRI reagent (Ambion, UK), a monophasic solution of phenol, guanidine isothiocyanate, and other proprietary components which

facilitate the isolation of a variety of RNA species of large or small molecular size (Mita *et al.*, 2007). Cells were washed twice in PBS, and 1ml TRI reagent was added to the cells. The cells were removed from the flasks or plates by scraping to one side; the resulting lysate was then removed and placed in a 1.5ml eppendorf and incubated for 5 minutes at room temperature, 0.2ml chloroform (Sigma, UK) was added to the lysates and the suspension was mixed by inversion and incubated at room temperature for 15 minutes. The suspension was then centrifuged at 12,000xg for 15 minutes at 4°C, after which the aqueous phase was transferred into fresh micro centrifuge tubes. To pellet the RNA 0.5ml of isopropanol was added to each micro centrifuge tube and mixed thoroughly by vortexing. The mixture was incubated at room temperature for 10 minutes and then centrifuged at 12,000xg for 8 minutes at room temperature. The resulting supernatant was carefully removed and the RNA pellet was washed by adding 1ml of 75% ethanol and centrifuged at 7,500xg for 5 minutes (Doherty *et al.*, 1999b). After removal of the supernatant, the RNA was air dried and resuspended in 20-50µl of nuclease free water, depending on the size of the pellet. The RNA suspension was incubated for 5 minutes at 55°C on a heating block to completely dissolve the RNA pellet. The resulting RNA was then stored at -80°C.

2.3.2 RNA extraction from Formalin Fixed, Paraffin Embedded (FFPE) samples

Fourteen FFPE samples were kindly supplied by Dr Muhammed Sohail (Department of Histopathology, United Hospitals Bristol Foundation Trust, Bristol Royal Infirmary, UK). The department of histopathology holds a general pathology accreditation awarded by the Clinical Pathology Accreditation (UK) Ltd. The criteria for selection of breast cancer tissue were based on the tissues being *HER2* 2+ or 3+ as determined by HercepTest.

10µm thick sections were cut from selected blocks of breast cancer which were confirmed to be *HER2* 2+ and 3+ by immunohistochemistry, as determined by HercepTest carried out by the Bristol Royal infirmary laboratory. Total RNA was extracted from FFPE samples using the RNEasy FFPE kit (Qiagen, UK) according to manufacturer's protocols, with the following modifications: Four 10µm thick sections were placed in 1.5ml eppendorf tubes and deparaffinised by adding 1ml of Histoclear, vortexing for 10 seconds and centrifuging at full speed for 2 minutes to bring the samples to the bottom of the tubes. The supernatant was then removed and discarded. To remove the Histoclear, 1ml of 100% ethanol was added to the pellets, mixed by vortexing and centrifuged at full speed for 2 minutes. The supernatant was removed and discarded, and the pellets were incubated at room temperature for 10-15 minutes to completely dry the pellet, thereby removing residual ethanol. 150µL of buffer PKD was added to the tubes; the solution was mixed by vortexing, and centrifuged for 1 minute at 11,000xg. 10µl of Proteinase K was then added to the lower, clear phase containing the RNA, mixed by gently pipetting, incubated overnight at 15°C, and then for 15 minutes at 80°C. This step is crucial for reversal of cross links to ensure optimal performance in downstream applications such as qRT-PCR. The lower, clear phase was then transferred to new 2ml microcentrifuge tubes, incubated on ice for 3 minutes, and then centrifuged for 15 minutes at 20,000xg. The resulting supernatant was transferred into fresh microcentrifuge tubes. To treat for genomic DNA contamination, 16µL DNAse booster and 10µL DNAse I solution were added to the tubes and incubated for 15minutes at room temperature. 320ml of buffer RBC was added to the tubes to adjust the binding conditions, and mixed by vortexing. 700ml of ethanol was added to the tubes, and mixed by pipetting. The sample was then passed

through an RNeasy MinElute spin column placed in a 2 ml collection tube. The spin column was then centrifuged at 8,000xg for 15 seconds and the flow-through was discarded. This step was repeated until the entire sample was passed through the MinElute spin column. 500ml of buffer RPE was then added to the spin columns and centrifuged at 8,000xg for 15 seconds and the flow-through was discarded. 500ml of buffer RPE was again added to the spin columns to wash the spin column membrane, centrifuged at 8,000xg for 2 minutes, and the flow-through discarded along with the collection tube. The spin column was then transferred into fresh collection tube and centrifuged at full speed for 5 minutes. This step is necessary to ensure that the residual ethanol on the membranes of the spin columns is removed. The spin columns were then placed in fresh 1.5ml microcentrifuge tubes and 20µL of RNase-free water was added to the column membrane. After incubating for 5 minutes at room temperature, the spin columns were centrifuged for 1 minute at full speed to elute the RNA (Abramovitz *et al.*, 2008).

2.3.3 Assessment of RNA yield and quality

The yield and quality of isolated RNA was determined using a Nanodrop 1000A UV spectrophotometer (Thermo Fisher Scientific, USA). RNA concenttrations for the cell lines used in this study can be found in Table 3.1. The Nanodrop UV spectrophotometer is used for analysing nucleic acid concentrations and the purity of nucleic acid samples with high accuracy using as small as 1µl of sample volume. A ratio of ~1.8 of sample absorbance at 260/280 nm was considered pure for RNA samples.

2.4. Reverse Transcription Polymerase Chain Reaction (RT-PCR)

2.4.1. First-Strand Synthesis of cDNA

Prior to first strand synthesis total RNA was treated for genomic DNA contamination with RQ1 RNAse-free DNAse (Promega, UK). 1µg of RNA was diluted up to 8µl in nuclease free water and incubated with 1µl RQ1 RNAse-free DNAse 10x reaction buffer and 1µl RQ1 RNAse-free DNAse for 30 minutes at 37°C. The DNAse reaction was terminated by the addition of RQ1 DNAse stop solution and heating to 65°C for 10 minutes. First-strand cDNA synthesis was then performed using equal amounts of total RNA by incubating with random hexamers (Qiagen UK) at 65°C for 5 minutes and then cooling immediately on ice. The reverse transcriptase buffer, dNTP mix, RNase inhibitor and Omniscript reverse transcriptase (Qiagen, UK) were then added to the reaction and incubated at 37°C for 60 minutes. To ensure that subsequent PCR amplification was derived from RNA and not genomic DNA or other contaminants, a no-RT control was included in every reverse transcription experiment (Bustin *et al.*, 2009). The yield and quality of cDNA synthesized was determined using a Nanodrop 1000A UV spectrophotometer (Thermo Fisher Scientific, USA). A ratio of ~1.8 of sample absorbance at 260/280nm was considered pure for cDNA samples.

2.4.2. Standard Reverse Transcription Polymerase Chain Reaction (RT-PCR)

Standard PCR reactions of 20µl volumes consisted of 1x green GoTaq® Flexi Buffer; 2µM $MgCL_2$; 0.2µM of each dNTP; 0.5µM of each primer; 1.25U GoTaq® Hotstart DNA polymerase and ~200ng template cDNA. All standard PCR reagents were supplied by Promega UK. For control amplifications cDNA was replaced by nuclease free water.

Cycling was performed using a PTC200 Peltier Thermal Cycler (MJ research, USA) using the following conditions: hotstart at 95°C for 2 minutes followed by 35 cycles of 95°C for 1 minute (denaturing), 54-64°C for 1 minute (annealing), 72°C for 1 minute (extending), and a final extension of 5 minutes at 72°C. A soaking cycle of 4°C was included to hold the tubes after amplification, prior to storage at -20°C (Hillig *et al.*, 2012).

2.4.3. Agarose gel electrophoresis

RT-PCR products were separated using molecular grade Agarose (Bioline UK). A 2% Agarose gel was prepared by dissolving Agarose powder in 1x TAE buffer and heating in a microwave until the Agarose powder was dissolved in the buffer. Ethidium bromide (10mg/ml; Sigma-Aldrich, UK) was added to a final concentration of 0.5µg/ml and the molten Agarose solution was poured into a gel tray and left to set at room temperature for 15 minutes. DNA samples were loaded onto the gel and subjected to electrophoresis at 120V for 1 hour 20 minutes. DNA hyperladder (Bioline UK) was run simultaneously to allow estimation of DNA fragment sizes. DNA was visualised under UV light on a MiniBis gel documentation system (Berthold Technologies, Germany).

2.5. Cloning of RT-PCR products for sequencing

2.5.1. Gel extraction and purification of PCR products

100µl reactions of PCR products destined for cloning were separated on a 3% Agarose gel and the bands excised from the gel prior to purification using a sterile scalpel. PCR products were purified using Qiaquick Gel Extraction Kit (Qiagen UK) (Lukas *et al.*, 2001). The excised gel bands were transferred to micro centrifuge tubes and weighed.

Three volumes of buffer QG were added to one gel volume, incubated at 50°C for 10 minutes and vortexed every 2-3 minutes until the gel was completely dissolved. To increase the DNA yield, 1 gel volume of isopropanol was added to the sample and mixed by inverting the tube several times. To bind the DNA, the sample was then transferred to a Qiaquick spin column, placed in a collection tube and centrifuged at 10,000xg for 1 minute and the flow-through was discarded, 500µl of buffer QG was added to the spin column and centrifuged at 10,000 x g for 1 minute and the flow-through discarded. To wash the DNA, 750µl of buffer PE was added to the column and centrifuged at 10,000 x g for 1 minute. To remove residual ethanol, the flow-through was discarded and the spin column was centrifuged for an additional 1 minute at 10,000 x g. To elute the DNA the Qiaquick column was transferred to a fresh 1.5ml micro centrifuge tube and 10µl of buffer EB was added to the column and incubated for about 1 minute and centrifuged at 10,000 x g for 1 minute. The purified DNA was stored at 4°C.

2.5.2. Preparation of LB broth and LB/Agar plates with ampicillin/IPTG/X-GAL

LB broth medium was prepared by suspending 20g of LB power (Sigma UK) in 1L of distilled water and sterilized by autoclaving. LB/Agar plates were prepared by suspending 20g of LB powder (Sigma UK) and 15g of Agar (Sigma UK) in 1L of distilled water and sterilized by autoclaving. The LB + Agar mixture was cooled to approximately 50°C and stock solutions of IPTG, X-GAL and ampicillin (bioline, UK) were added to the LB + Agar mixture to a final concentration of 100µg/ml each, and then mixed by inversion (Doherty *et al.*, 1999b). The mixture was then poured into petri dishes and cooled until solid.

2.5.3. Ligation into pGEM-T Easy vector

The pGEM®-T Easy Vector system (Promega, UK) is a convenient system for cloning of PCR products. The pGEM-T Easy vector is prepared by digestion with EcoRV and adding a 3' terminal thymidine terminal overhang to both ends to prevent recircularisation of the vector and enable insertion of PCR products generated using thermostable polymerases that add a deoxyadenosine to the 3' end of amplified PCR products. The high copy number pGEM®-T Easy Vector contains T7 and SP6 RNA polymerase promoters flanking an alpha peptide multiple cloning region, which when inactivated, allows recombinant clones to be identified by blue/white screening on indicator plates.

The ligation reaction consisted of 1µl (50ng) pGEM-T® Easy Vector, 5µl 2x T4 DNA Ligase rapid ligation buffer, 1µl (3 Weiss units/µl) T4 DNA Ligase and 3µl gel purified PCR product. The reactions were incubated overnight at 4°C to produce a maximum number of transformants.

2.5.4. Transformation into JM109 High Efficiency Competent E. coli cells

PCR ligation reactions were inserted into JM109 High Efficiency Competent E. Coli cells (Promega, UK). 50µl of cells were added to fresh microcentrifuge tubes and incubated on ice for 20 minutes with 2µl of each ligation reaction. Cells were heat-shocked in a water bath at 42°C for 50 seconds and then placed on ice for 2 minutes. 950µl of LB broth was added to the cells then incubated at 37°C for 1.5 hours with shaking. Transformed cells were then grown overnight on LB plates with ampicillin/IPTG/X-GAL at 37°C. Single white E. coli colonies were selected and used to inoculate 5ml LB medium containing 100µg/ml ampicillin and incubated overnight at 37°C with shaking.

2.5.5. Extraction of plasmid DNA

Plasmids were extracted from E. coli cells using The PureYield™ Plasmid miniprep system (Promega UK). The PureYield™ Plasmid Miniprep System is designed to purify plasmid DNA from an overnight culture of bacteria transformed with a high-copy number plasmid. Cells were pelleted by centrifugation at 10,000xg for 5 minutes and the supernatant discarded. Cells were resuspended in 600µl of nuclease free water, lysed by adding 100µl of cell lysis buffer, and mixed by inverting the tube several times. Lysis was terminated by adding 350µl of neutralization solution and the solution was thoroughly mixed by inversion. The reaction was then centrifuged at 10,000xg for 3 minutes to pellet the debris resulting from lysis. The supernatant was then transferred to a PureYield™ Mini column and centrifuged at 10,000xg for 15 seconds and the flow through discarded. The column was then washed by adding 200µl of Endotoxin Removal Wash (ERB) to the mini column and centrifuged at 10,000xg for 15 seconds, and then adding 400µl of column wash solution (CWS) to the mini column and centrifuging at 10,000xg for 30 seconds. The resulting DNA was eluted by adding 30µl of Elution Buffer directly to the mini column matrix and incubating for 1 minute. The column was then transferred to a fresh 1.5ml tube and centrifuged at 10,000xg for 30 seconds. The eluted purified plasmid DNA was quantified using a Nanodrop 1000A UV spectrophotometer (Thermo Fisher Scientific, USA) and then stored at -20°C.

2.6. Quantitative real-time PCR (qRT-PCR)

2.6.1. Quantitative real-time PCR amplifications

Primers and probes for qRT-PCR were designed by Primer Deisgn UK. The amplicon lengths for all primer sets were 90-150 base pairs. All PCR reactions were performed using a Step One Plus™ Real-Time PCR system (Life Technologies, UK). Real-time PCR reactions were set up using Sensifast Probe Hi-Rox and Sensifast SYBR Hi-Rox kits (Bioline, UK). 25ng cDNA was used for each reaction containing 1x Sensifast master mix, primers (400nM each), probe (100nM) and nuclease free water (up to 20ul). All reactions were set up in triplicate in a 96 well plate. The cycling conditions comprised polymerase activation for 2 minutes at 95°C, followed by 40 cycles of denaturation at 95°C for 5 minutes, and annealing at 60°C for 15 seconds. For SYBR green chemistry, the specificity of the primers was determined using a melt curve analysis by increasing the cycling temperature in one degree increments, starting at 50°C, for 40 cycles of 10 seconds each. The efficiency of the primers was also determined using a standard curve. Six serial doubling dilutions of pooled cDNA samples were run in triplicate alongside the main reaction. C_T values were plotted against arbitrary log values on an excel spreadsheet using a scatter plot. R^2 values of ≥0.96 were accepted for correlation efficiency (Derveaux, Vandesompele & Hellemans, 2010).

2.6.2. Calculations

C_T values were automatically determined by the Step One Plus™ Real-Time PCR system (Life Technologies, UK). Relative expression of target mRNA in different samples were normalised to a set of reference genes using the Relative Standard Curve method. A

standard curve was set up for each primer set by running six serial doubling dilutions of pooled cDNA in triplicate alongside the main reactions. The standard curve was determined using the equation:

$$y = mx+c.$$

The C_T (y) values generated from the real-time PCR reaction was then used to derive the following equation:

$$x = [(y-c)/-m)]$$

Once the x values were derived, they were then converted to real numbers by anti-logging, normalised against a set of reference genes, and the fold difference between control and experimental samples were calculated (Emir, 2011). The statistical significance of the fold difference between two samples was determined by student's t-test.

2.6.3. Normalisation of real-time qRT-PCR

The use of quantitative real-time PCR in gene expression analysis has become increasingly important in biological research. Due to the sensitivity of the qRT-PCR assay, accurate normalization of data is necessary to ensure that changes in gene expression from sample to sample are not due to errors like variations in sample size, pipetting errors or errors in the reverse transcription step.

The use of reference genes as endogenous controls in qRT-PCR is one of the most common methods used in the normalization of qRT-PCR data. However, the expression of reference genes may vary considerably between experiments as no single gene is stably expressed from sample to sample or under all experimental conditions. Randomly selecting a reference gene of choice may lead to large errors in data analysis; therefore it is essential to identify a reference gene or a set of genes from a set of candidate genes, that are most stably expressed in every experimental model (Derveaux, Vandesompele & Hellemans, 2010; Vandesompele *et al.*, 2002; Romanowski *et al.*, 2007; Andersen, Jensen & Ørntoft, 2004; Bustin *et al.*, 2009).

Two different methods of reference gene selection were used in this project:

- Genorm^{PLUS} (Ghent, Belgium): the Genorm^{PLUS} algorithm calculates the most stably expressed reference gene or set of reference genes from a set of genes based on the geometric mean of a number of housekeeping genes (Vandesompele *et al.*, 2002).

- Normfinder (Molecular Diagnostics Laboratory, Denmark): The normfinder algorithm ranks a set of candidate reference genes according to their stability in a given experimental design (Andersen, Jensen & Ørntoft, 2004).

A set of 12 candidate reference genes were chosen for use in the selection of optimal reference genes for the normalisation of qRT-PCR (Table 2.3).

GENE NAME	DESIGNATION	ACCESSION	FUNCTION
18s Ribosomal RNA	18S	NM_10098	Protein translation
Beta actin	ACTB	NM_001101	Cytoskeletal protein involved in cell motility, structure, and integrity.
Homo sapiens ATP synthase, H+ transporting, mitochondrial F1 complex, beta polypeptide	ATP5B	NM_001686	ATP synthesis
Beta-2-microglobulin	B2M	NM_004048	Beta-chain of major histocompatibility complex class I molecules involved in the presentation of peptide antigens to the immune system.
Cytochrome c-1	CYC1	NM_001916	Electron transport in mitochondrial respiratory chain
Eukaryotic translation initiation factor 4A2	E1F4A2	NM_001967	Initiation of translation

GENE NAME	DESIGNATION	ACCESSION	FUNCTION
Glyceraldehyde-3-phosphate dehydrogenase	GAPDH	NM_002046	Oxidoreductase in glycolysis
Ribosomal protein L13A	RPL13A	NM_012423	Protein synthesis
Succinate dehydrogenase complex, subunit A, flavoprotein	SDHA	NM_004168	Mitochondrial respiration
Topoisomerase DNA I	TOP1	NM_003286	Control and alteration of DNA topology during transcription
Ubiquitin C	UBC	NM_021009	Ubiquitination
Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta	YWHAZ	NM_145690	Mediation of signal transduction

Table 2.3: Reference genes used for qRT-PCR and their functions in normal physiology.

2.7. Protein analysis

2.7.1. Protein extraction

Protein extraction was carried out using Radioimmunoprecipitation assay (RIPA) lysis buffer (Fisher, UK). Cells grown in monolayer were washed twice with cold PBS. 500µL of RIPA buffer, supplemented with protease and phosphatase inhibitors (sigma, UK) was added to a 75cm^2 flask or 5x10^6 cells and kept on ice for 5 minutes with occasional

swirling of the flask for uniform spreading. The lysate was then removed from the flask by scraping to gather to one side, and transferred into a microcentrifuge tube. To increase yield, the lysates ware homogenised by taking the homogenate up and down 20 times using a 21G needle. The homogenate was then centrifuged at 13,000 rpm for 15 minutes at 4°C to pellet the cell debris. The supernatant was then transferred to a fresh microcentrifuge tube and stored at -80°C.

2.7.2. Protein quantification

Protein concentration was determined using the Cooomassie Blue dye binding method for protein estimation (Bradford Method) (Siegel *et al.*, 1999). This assay is based on an absorbance shift of the coomassie blue G dye when it binds to proteins. The increase in absorbance at 595nm is proportional to the amount of bound dye and thus, the concentration of protein present in the sample.

A stock dye solution was prepared dissolving 330mg of coomassie blue G dye in 100ml phosphoric acid/ethanol (2:1) mixture. The working dye solution was prepared by mixing 1.5ml of stock dye reagent, 4ml of phosphoric acid, 1.9ml of ethanol and making up to 50ml in distilled water. A protein standard was prepared by dissolving albumin in deionised water to a final concentration of 1µg/µL. Reactions were carried out in triplicate in a 96 well plate and the absorbance was measured at 560nm.

2.7.3. SDS PAGE

Proteins were separated using a criterion XT system (Bio-rad, UK). Samples were thawed on ice from frozen and 20µg of protein lysates were mixed with 4x NuPAGE LDS sample buffer diluted to 1x in distilled water (Life Technologies, UK), and made up

to 20µl volume in distilled water. The concentration of protein for SDS was determined following the general protocols for western blotting by the Bio-rad systems. The mixture was heated to 95°C for 5 minutes prior to being loaded onto a criterion XT 4%-12% Bis-Tris pre-cast gel (Bio-rad, UK). The Criterion gel system (Bio-rad, UK) was assembled according to manufacturer's instructions and filled with 1 litre of 1x XT-MES running buffer (Life Technologies, UK). The criterion XT pre-cast gel was inserted into the gel tank and 10µl HyperPAGE II broad range pre stained protein marker (Bioline, UK) was loaded to the first well. Samples were loaded to subsequent wells and run at 120 volts for 1-2 hours.

2.7.4. Western blot analysis

Following SDS PAGE, Polyvinylidene fluoride (PVDF) membrane (GE Healthcare, UK) was cut to appropriate size and activated by immersing in methanol for 2 minutes. The membrane was then rinsed in distilled water and then left to equilibrate in 1x transfer buffer (20x Nupage transfer buffer, 5% methanol in distilled water). The gel and membrane were then transferred onto a Criterion blotter (Bio-Rad, UK) and the proteins were transferred at 50 volts at 4°C for 1 hour. Following transfer the membranes were transferred to blocking solution [5% (w/v) non-fat dry milk solution, 0.02% (v/v) Tween 20 (Life Technologies, UK) in PBS] for 1 hour at room temperature with rocking. After blocking, the membranes were incubated in primary antibody (in 5% non-fat dry milk solution) at 4°C overnight. Following incubation in primary antibody, the membrane was washed 3 times with PBS-T at room temperature for 15 minutes each with rocking. The PBS-T was then removed and the membranes were incubated in HRP-linked secondary antibody (Santa Cruz Biotechnology, UK) for 45

minutes at room temperature. The membranes were then washed 3 times with PBS-T at room temperature for 15 minutes each with rocking, and then rinsed in distilled water (Al Okail, 2010). Table 6.1 shows a list of antibodies used in western blotting and their specificities.

After washing the membranes, HRP-chemiluminescent substrate was applied to the membranes and the bands were visualised in a dark room by developing on a film (Amersham Hyperfilm ECL; GE Healthcare).

2.8. RNAi methods

2.8.1. siRNA transfection of cells

Cells were prepared for transfection by seeding in 2ml antibiotic free media in 6-well plates at a density of $0.6x10^6$ for 24 hours. On the day of transfection, cells were serum starved in Opti-MEM® I Reduced Serum Media (Life Technologies, UK) for 2 hours. siRNA transfection was then carried out using siGENOME siRNA smartpool *SRPK1* and *SRSF1* (Thermoscientific, UK) and siGENOME Non-Targeting siRNA Pool #2 as a negative control. Dharmafect 2 transfection reagent (Thermoscientific, UK) was used (Sahlberg *et al.*, 2013). siRNA transfection was carried out according to the manufacturer's guidelines with the following modifications:

A 2mM siRNA solution was prepared by diluting in 1X siRNA Buffer (Thermoscientific, UK). In separate tubes, 100nM siRNA and Dharmafect 2 transfection reagent were diluted using Opti-Mem 1 medium. Contents of each tube were carefully mixed by pipetting up and down and incubated for 5 minutes at room temperature. The content

of both tubes were then carefully mixed together by carefully pipetting up and down. The final mixture was then incubated for 20 minutes at room temperature to allow the formation of transfection complexes. Prior to adding the transfection reagents to the cells, cell culture medium was removed from the cells and fresh opti-Mem I was added to the cells, and then the appropriate transfection mix was added to the cells at a final concentration of 100nM siRNA and 6µl Dharmafect 2 reagent per well. Cells were then incubated for 6 hours to allow the transfection reagents to permeate the cell membranes. After 6 hours of transfection, the transfection reagent was removed from the cells and replaced with antibiotic free medium, and the cells were then incubated for 24-48 hours before protein or RNA extraction. All knockdowns were carried out in triplicate, with each replicate performed on a different passage of cells. A mock transfection and a non-treated control was added to each experiment.

CHAPTER 3. DISCOVERY OF *HER2* AND *HER2* ALTERNATIVE SPLICE VARIANTS IN BREAST AND OVARIAN CANCER CELL LINES

3.1 Introduction

Amplification of the *HER2* oncogene is one of the genetic abnormalities in breast tissue associated with the progression from normal breast epithelia to invasive cancer cells (Castiglioni *et al.*, 2006). *HER2* testing is essential for the appropriate administration of the humanized anti-*HER2* monoclonal antibody therapy *Trastuzumab* (Herceptin ®, Genetech, South San Francisco, CA) to invasive breast cancer patients with *HER2* overexpression or gene amplification (Bartlett *et al.*, 2001; Dean-Colomb & Esteva, 2008). However, the discovery of alternative splice variants of *HER2* potentially adds extra complexity in mediating patient response to *HER2* therapies as different *HER2* splice variants may have differing biological properties (Mitra *et al.*, 2009; Marchini *et al.*, 2011). Research studies suggest that expression and secretion of aberrant *HER2* splice variants can interfere with the oncogenic *HER2* activity (Aigner *et al.*, 2001). Whereas the wild type p185 *HER2* is associated with pro-oncogenic receptor activity, resulting in poor prognosis and disease progression if overexpressed or over amplified in breast cancer, the *HER2* isoforms Herstatin and p100 *HER2* are thought to have anti-oncogenic function by inhibiting receptor dimerization and subsequently inactivating signal transduction pathways (Wang *et al.*, 2013). Also, the *HER2* exon 16 immediately precedes the transmembrane domain and contains two cysteine residues. The loss of exon 16 in *HER2*Δ16 therefore leads to a change in the conformation of *HER2* receptor extracellular domain that promotes intermolecular disulphide bonding, thereby

promoting the formation of constitutively activated *HER2* homodimers capable of transforming cells (Castiglioni *et al.*, 2006). Therefore it may be of significant prognostic value to determine the *HER2* variant status of patients with invasive breast cancer. The recommended methods for testing *HER2* status in breast cancer patients are Immunohistochemistry and Fluorescence In-situ Hybridization (FISH) (Wolff *et al.*, 2007), and more recently, Silver in-situ hybridization (SISH) is also being used as an alternative to the FISH technique (Moelans *et al.*, 2011). With growing evidence of the potential involvement of alternative splice variants as biomarker candidates in cancer diagnosis (Lukas *et al.*, 2001; Marchini *et al.*, 2011; Omenn, Yocum & Menon, 2010), the discovery of novel alternative splice variants of *HER2* and their potential role in disease progression or resistance to *Trastuzumab* may be crucial in developing potential novel methods of testing and treatment of *HER2* in patients with invasive breast cancer.

Objectives

1. To detect *HER2* protein expression in breast cancer cell lines by immunohistochemistry.
2. To detect *HER2*Δ16 in *HER2* overexpressing breast and ovarian cancer cell lines.
3. To design PCR primers for the detection of *HER2* and possibly novel *HER2* splice variants in cell lines by RT-PCR.
4. To sequence the *HER2* cDNA and *HER2* splice variants following RT-PCR amplification.

3.2 Methods

3.2.1 Antigen Retrieval for Immunohistochemistry

Antigen retrieval was performed prior to Immunohistochemical staining in SKBR3 (*HER2* 3+; ER-; PR-), BT-20 (*HER2*-; ER-; PR-) and MCF-7 (*HER2*-, ER+; PR+) cell lines. Breast cancer cells mounted on slides were stained with monoclonal antibodies to *HER2* (SP3; CB11), Oestrogen Receptors (6F11), and Progesterone Receptors (PGR636) (Table 3.1).

3.2.2 *HER2* primer design for RT-PCR

The polymerase chain reaction (PCR) is a reliable method used to amplify specific genes of interest, or sections of a gene (Schrader *et al.*, 2012). Specific primer sequences corresponding to a specific gene of interest are used to amplify sections of DNA, with the use of DNA polymerase, enabling the generation of unlimited copies of a small fragment of DNA (Joshi & Deshpande, 2011; Schrader *et al.*, 2012).

The open reading frame of *HER2* mRNA as predicted by the ExPASY translate tool was amplified by PCR using 12 sets of sequence-specific primers (Figure 3.1). The primers were designed using OligoPerfect™ Designer (Invitrogen, UK). The nucleotide sequences used for the primer design were based on *HER2* ncbi GenBank® accession number NM_004448 (NCBI, 2010). NP1, NP2, NP5 and NP6 primer sequences for the amplification of *HER2*Δ16 were obtained from Kwong & Hung (1998). Primers were

synthesized by Eurofins MWG Operon (Ebesberg, Germany). Primer sequences for all primers used in *HER2* amplification are listed in Appendix A.

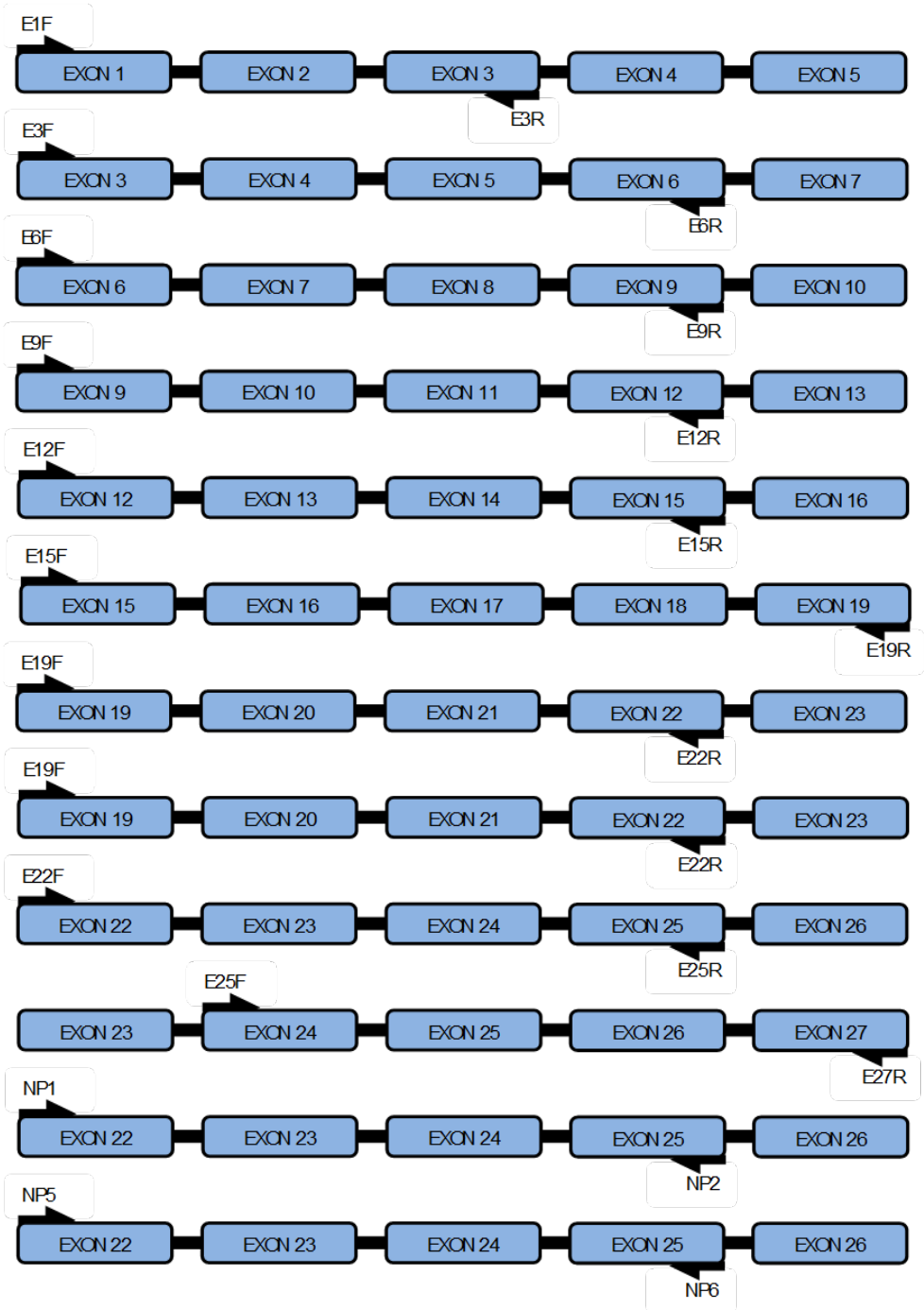


Figure 3.1: Design of *HER2*-specific RT-PCR primers for used to amplify *HER2* cDNA. Arrows indicate positions of primers in target exons.

3.2.3 DNA sequencing

All DNA products isolated and purified from RT-PCR analysis were sequenced using the Sanger method of DNA sequencing to confirm that the insert sequences were from *HER2* mRNA. DNA sequencing was performed at three independent locations:

- University of the West of England: Plasmid DNA templates were supplied at a concentration of 200ng/µl in a total volume of 10µl. The sequencing reactions consisted of 8 µl of 1 x BigDye Terminator v. 3.1 ready reaction mix (Applied Biosystems, UK), 3 µl of dilution buffer (Applied Biosystems), 3.2 pmol of primer, and 0.2 µg of template DNA in a final reaction volume of 20 µl. Cycling conditions included an initial one minute denaturation step at 96°C, followed by 25 cycles at 96°C for 10 seconds, 50°C for 5 seconds, and 60°C for 4 minutes, and a final extension at 15°C for 10 minutes. Samples were electrophoresed on an Applied Biosystems 3730xl automated DNA sequencing instrument, using 37 cm capillary arrays and POP-7 polymer. Sequencing of plasmids was performed using T7 and SP6 promoter primers. Data were analyzed using PE-Biosystems version 3.7 of Sequencing Analysis.
- MWG operon: 50-100ng/µl of plasmid DNA in a final volume of 15µl was sent to MWG operon for commercial single strand sequencing using T7 and SP6 promoter primers.
- University of Exeter: following agarose gel electrophoresis, DNA was obtained directly from the agarose gel by stabbing the bands with a p20 pipette tip while being visualised with the UV transilluminator. The DNA was then re-amplified

by RT-PCR, using the same primers used for the initial PCR amplification, tagged with M13 primers (primer sequences are listed in appendix A), under the same cycling conditions as the initial PCR reaction. The resulting re-PCR products were then sequenced at the University of Exeter.

3.2.4 Analysis of sequencing results

All sequenced plasmid DNA templates were analysed by first identifying SP6 or T7 promoter primer sequences, and M13 primer sequences, and then identifying the insert sequences. The plasmid insert sequences were entered into NCBI Nucleotide Basic Local Alignment Search Tool (BLASTN) to confirm that they were actual *HER2* mRNA sequences (NCBI, 2010). After sequences were confirmed to be *HER2*, insert sequences were then aligned using Clustal Omega Multiple Sequence Alignment Tool (European Bioinformatics Institute, 2010) with the reference *HER2* sequence to determine homology between sequences obtained from GenBank and sequences obtained from cloned inserts of PCR amplified products.

3.3 Results

3.3.1 Detection of *HER2* protein in cell lines by Immunohistochemistry

Immunohistochemistry was performed to give a general overview of *HER2* expression in cell lines. Three breast cancer cell lines; SKBR3 (*HER2* 3+; ER-; PR-), BT-20 (*HER2*-; ER-; PR-) and MCF-7 (*HER2*-, ER+; PR+) were used for immunohistochemical analysis. The

antibodies used for detection of *HER2* bind to the extracellular domain while the ER and PR antibodies bind to the nucleus. *HER2* antibodies are not known to bind to a specific *HER2* isoform, and may therefore be positive to the generic wild-type *HER2*, regardless of any isoforms which may be co expressed in the same cells. Monoclonal antibodies used in this study and their specificities are listed in Table 3.1.

SKBR-3 cell line: Immunohistochemical analysis of SKBR3 cell line showed the presence of membrane staining with SP3 and CB11 monoclonal antibodies (Figures 3.2 and 3.3), which is indicative of *HER2* positivity. The absence of nuclear staining of SKBR3 cells with 6F11 and PGR636 antibodies (Figures 3.4 and 3.5) show negative results for Oestrogen and Progesterone receptors respectively.

BT-20 cell line: Immunohistochemical analysis of BT-20 cell line showed the absence of membrane staining with SP3 and CB11 antibodies (Figures 3.2 and 3.3), which is indicative of *HER2* negativity. The absence of nuclear staining of BT-20 cell line with 6F11 and PGR636 antibodies (Figures 3.4 and 3.5) is also indicative of negativity of oestrogen and progesterone receptors respectively.

MCF-7 cell line: Immunohistochemical analysis of MCF-7 cell line showed the absence of membrane staining with SP3 and CB11 antibodies (Figures 3.2 and 3.3), which is indicative of *HER2* negativity. The presence of nuclear staining in MCF-7 cell line with CB11 and PGR636 antibodies (Figures 3.4 and 3.5) is indicative of oestrogen and progesterone receptor positivity, respectively. Negative controls used in immunohistochemistry, stained negative for SP3 antibody (Figure 3.6).

The results obtained from immunohistochemistry correlate well with previous *HER2* studies on the cell lines used in this study (Rhodes *et al.*, 2010) However, overall protein expression as determined by IHC may not be conclusive in determining the patients' splice variant status, and therefore may not be sufficient in predicting patients' response to treatment.

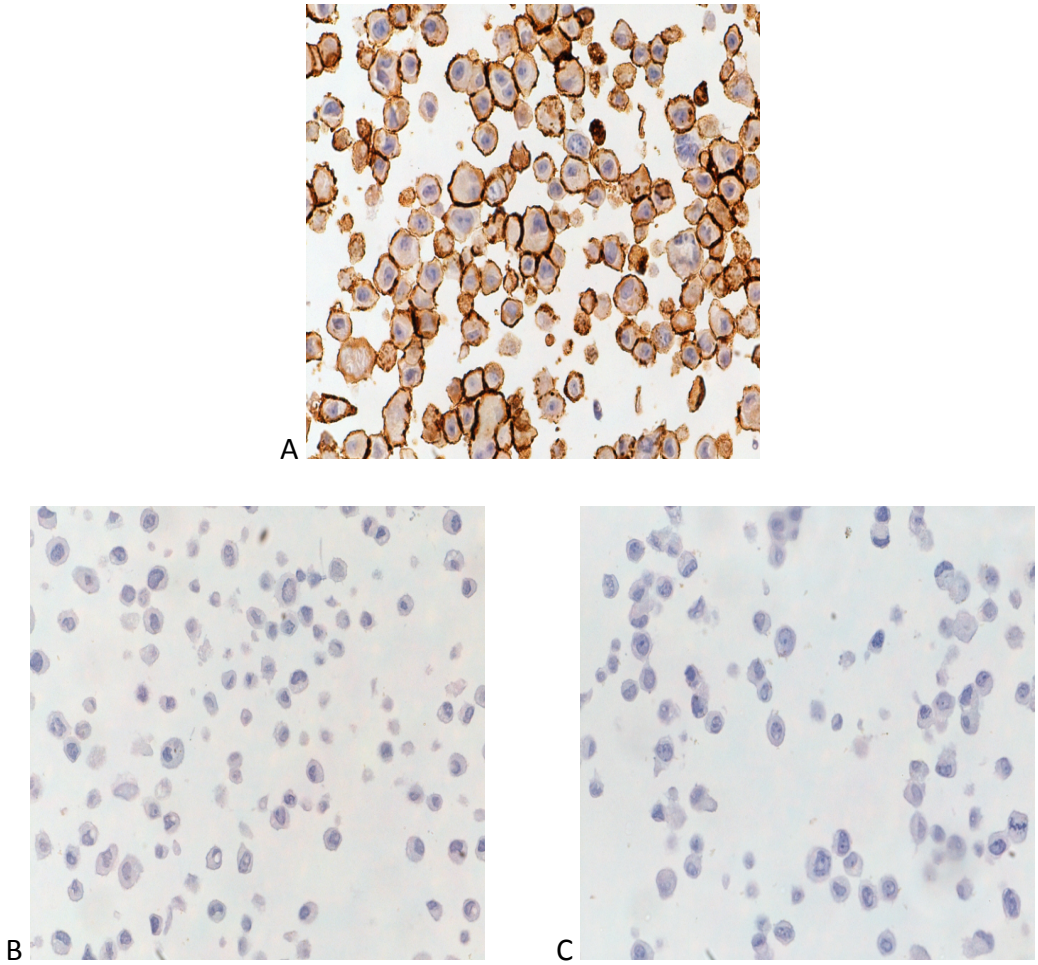


Figure 3.2: Immunohistochemical staining of cell lines SKBR3 (A), BT-20 (B) and MCF-7 (C) using SP3 monoclonal antibody. Membrane staining of SKBR3 indicates *HER2* positivity (magnification: x 40).

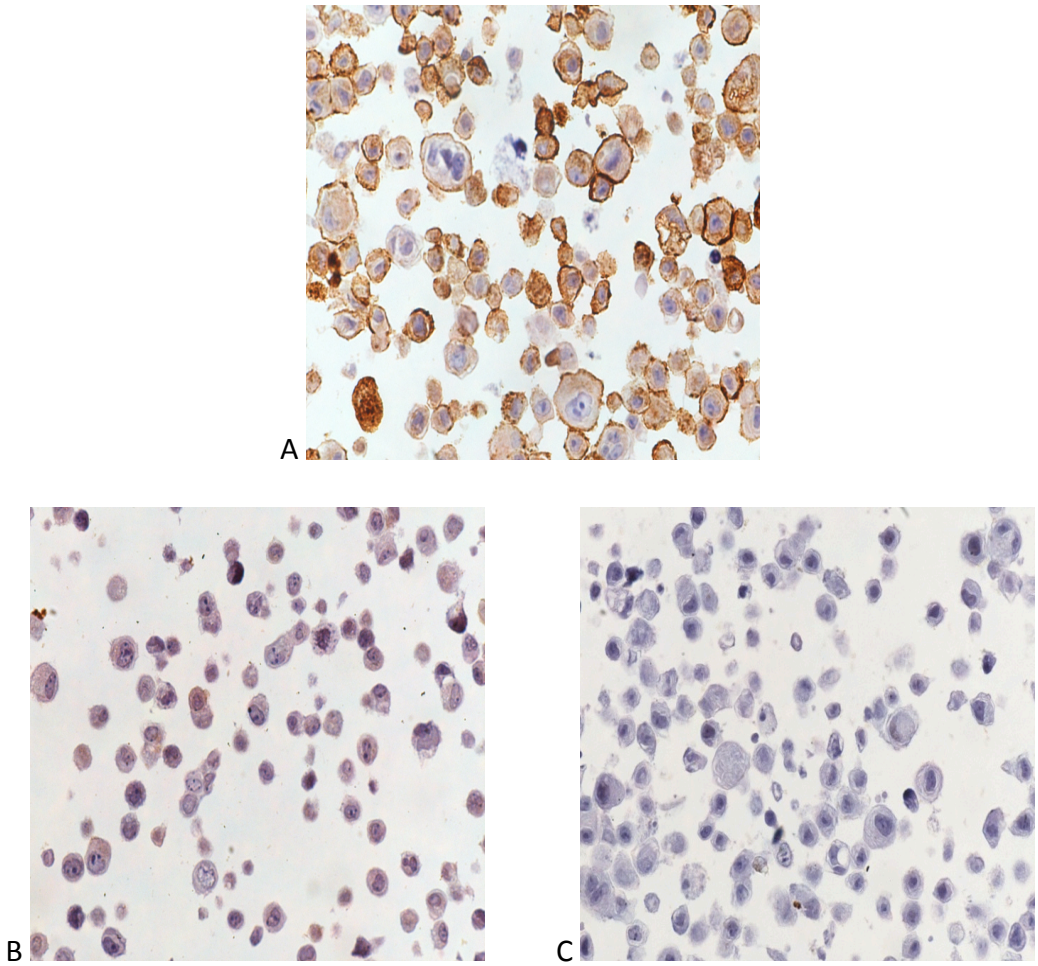


Figure 3.3: Immunohistochemical staining of cell lines SKBR3 (A), BT-20 (B) and MCF-7 (C) using CB11 monoclonal antibody. Membrane staining of SKBR3 indicates *HER2* positivity (magnification: x 40).

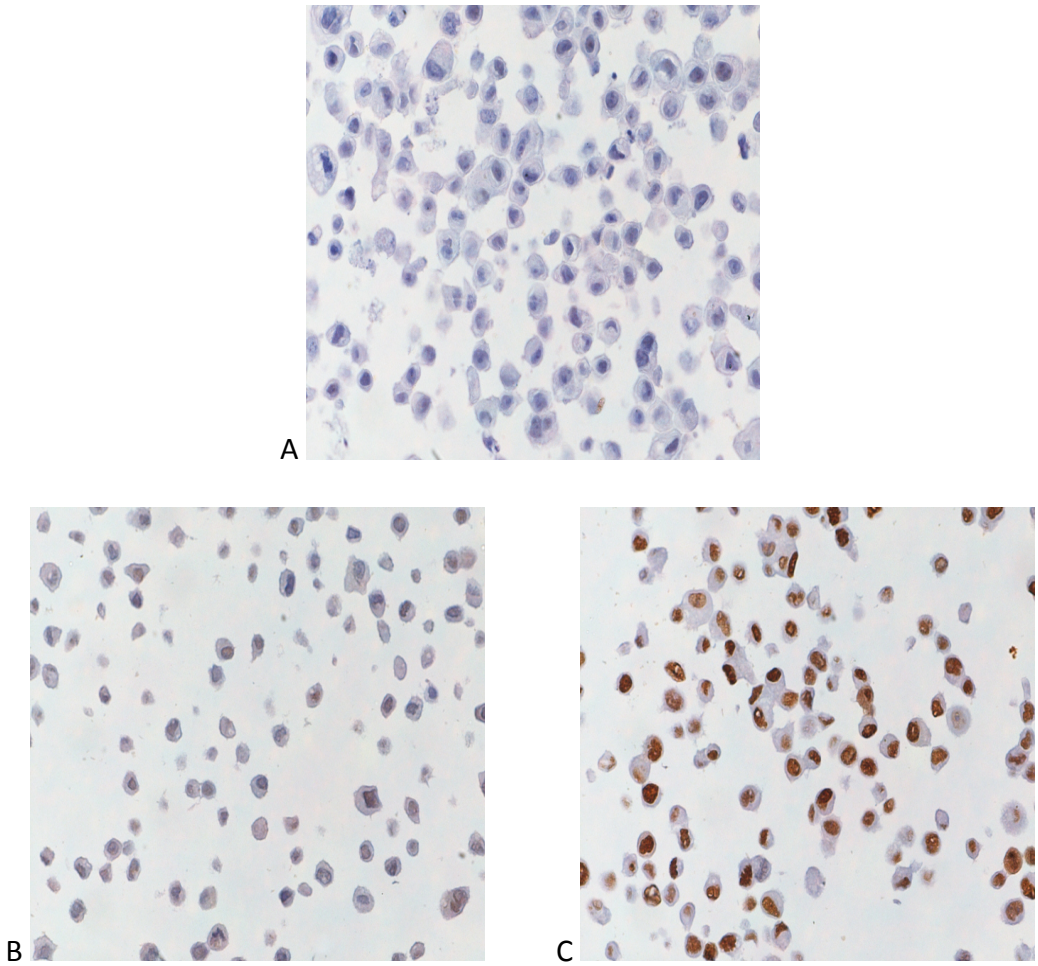


Figure 3.4: Immunohistochemical staining of cell lines SKBR3 (A), BT-20 (B) and MCF-7 (C) using 6F11 monoclonal antibody. Nuclear staining of MCF-7 indicates ER positivity (magnification: x 40).

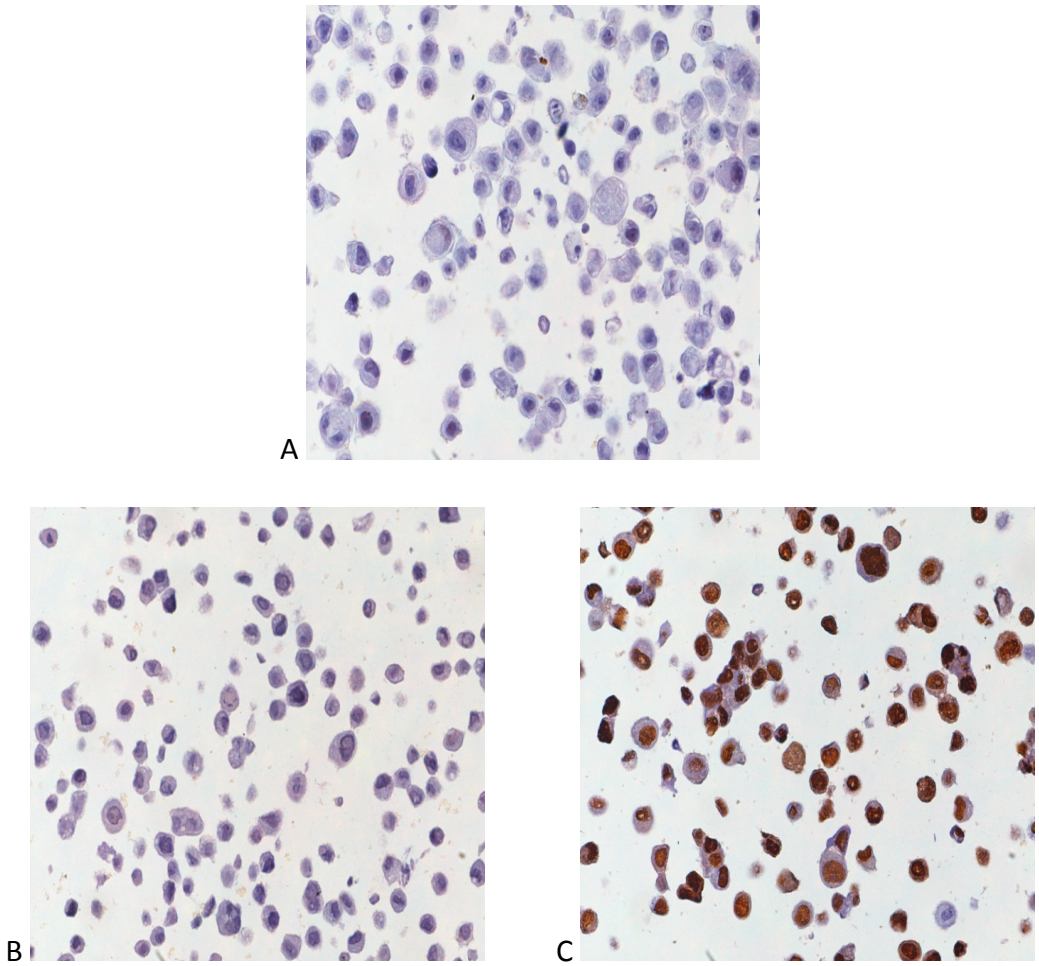


Figure 3.5: Immunohistochemical staining of cell lines SKBR3 (A), BT-20 (B) and MCF-7 (C) using PGR636 monoclonal antibody. Nuclear staining of MCF-7 indicates PR positivity (magnification: x 40).

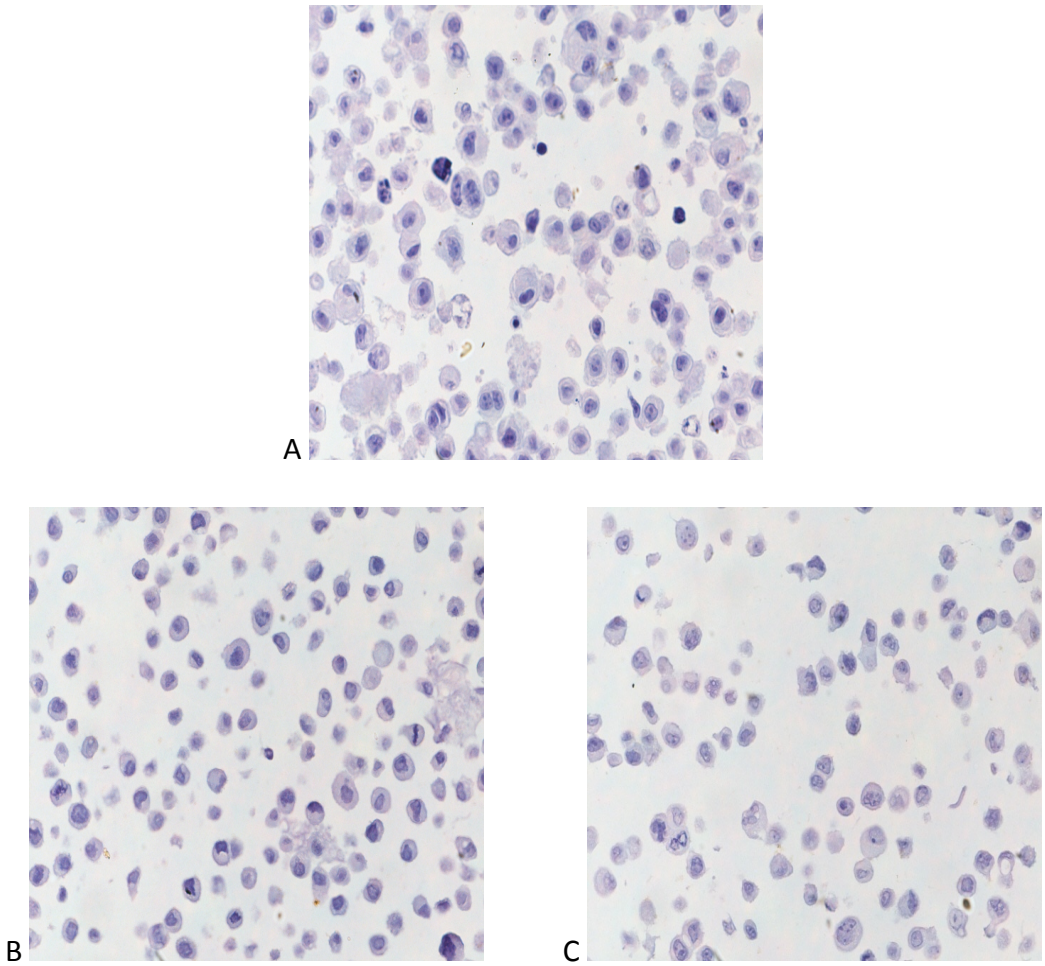


Figure 3.6: Negative controls used in immunohistochemistry showing cell lines SKBR3 (A), BT-20 (B) and MCF-7 (C) using SP3, CB11 and 6F11 monoclonal antibodies respectively (magnification: x 40).

3.3.2 Detection of *HER2* mRNA expression in cell lines by RT- PCR

The Polymerase Chain Reaction provides a rapid and sensitive method for amplifying a specific segment of complementary DNA (cDNA) produced by reverse-transcription of RNA extracted from cells or tissues, making it possible to delineate a template sequence or a specific region of a gene of interest. The specified sequence corresponds to the amplicon size of the PCR product. PCR has also been used to identify alternative

splice isoforms in genes where products which do not correspond to the expected amplicon size have been obtained.

RT-PCR analysis of cell lines shows that there are numerous potential alternative splice variants in *HER2*. In addition to the already published exon 16 deleted *HER2* isoform using primer pairs NP1/NP2 and NP5/NP6 (Figures 3.15 and 3.16) (Kwong & Hung, 1998), multiple bands were observed in exons 12-15 with primers E12F/E15R (Figure 3.10), exons 15-19 with primers E15F/E19R, and exons 19-22 with primers E19F/E22R (Figure 3.12). To ensure that the RT-PCR results were from actual *HER2* mRNA and not genomic DNA, all RNA templates were treated with RQ1 RNAse-Free DNAse (Promega, UK). As an added control, a no RT template was added to each reverse transcription reaction, and the template was PCR amplified along with the experimental cDNA samples. RT-PCR results are representative of three biological repeats.

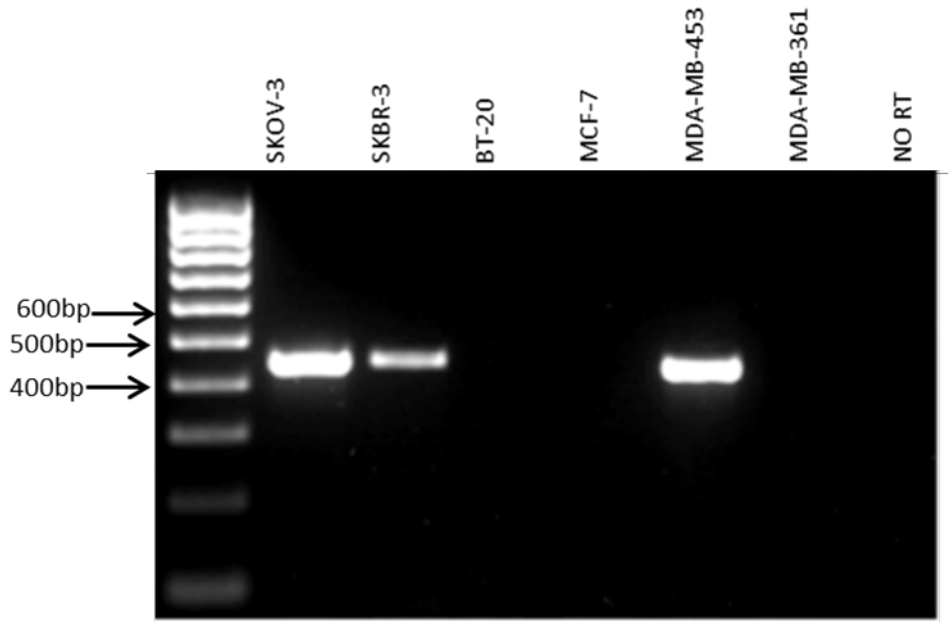


Figure 3.7: RT-PCR amplification of *HER2* exons 3-6 (primer pair E3F+E6R) using all six cell lines, and a negative (no RT) control. SKOV3, SKBR3 and MDA-MB-453 show expected band sizes at 478 base pairs. Each rung of the hyperladder IV represents 100bp.

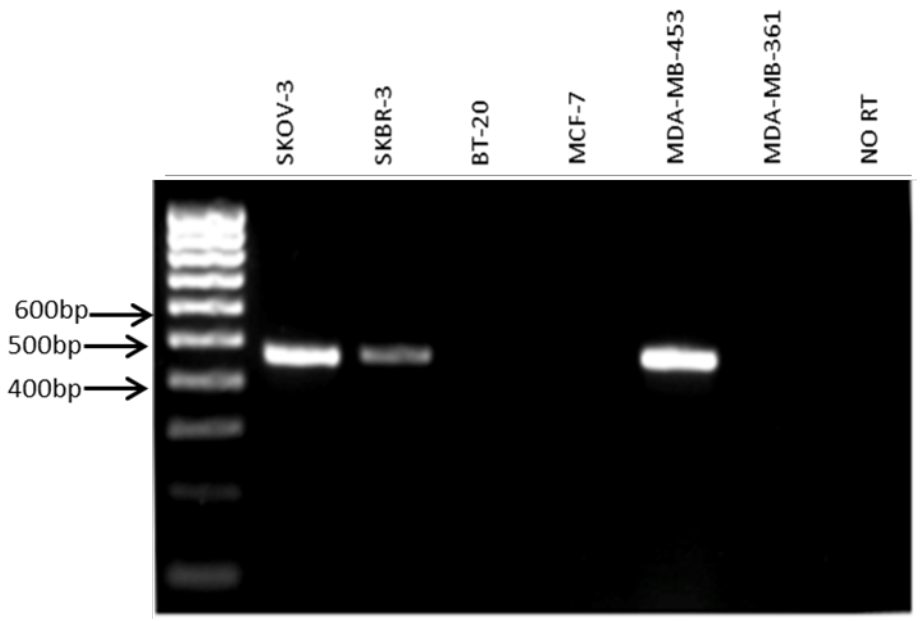


Figure 3.8: RT-PCR amplification of *HER2* exons 6-9 (primer pair E6F+E9R) using all six cell lines, and a negative (no RT) control. SKOV3, SKBR3 and MDA-MB-453 show expected band sizes at 441bp base pairs. Each rung of the hyperladder IV represents 100bp.

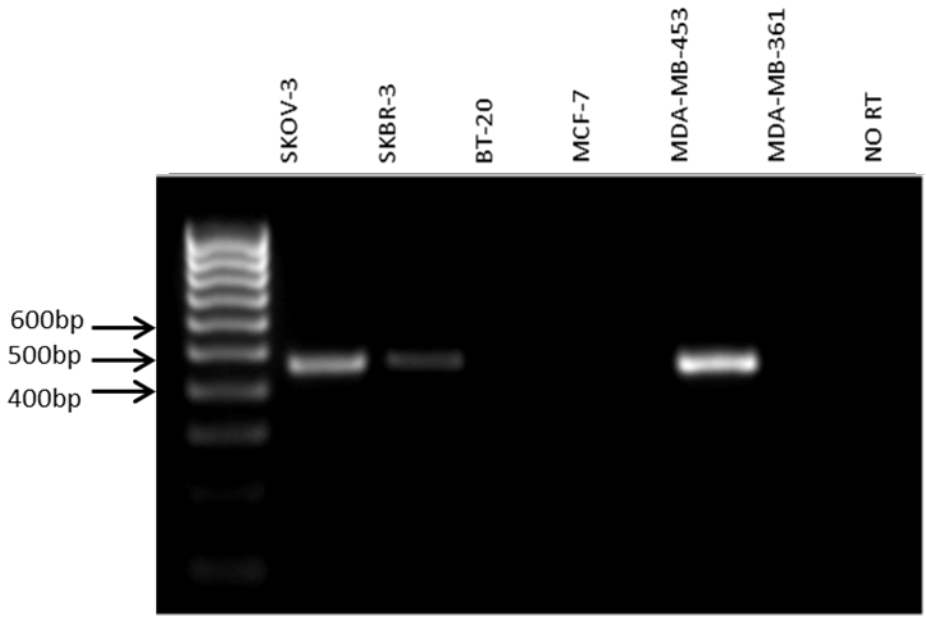


Figure 3.9: RT-PCR amplification of *HER2* exons 9-12 (primer pair E9F+E12R) using all six cell lines, and a negative (no RT) control. SKOV3, SKBR3 and MDA-MB-453 show expected band sizes at 454 base pairs. Each rung of the hyperladder IV represents 100bp.

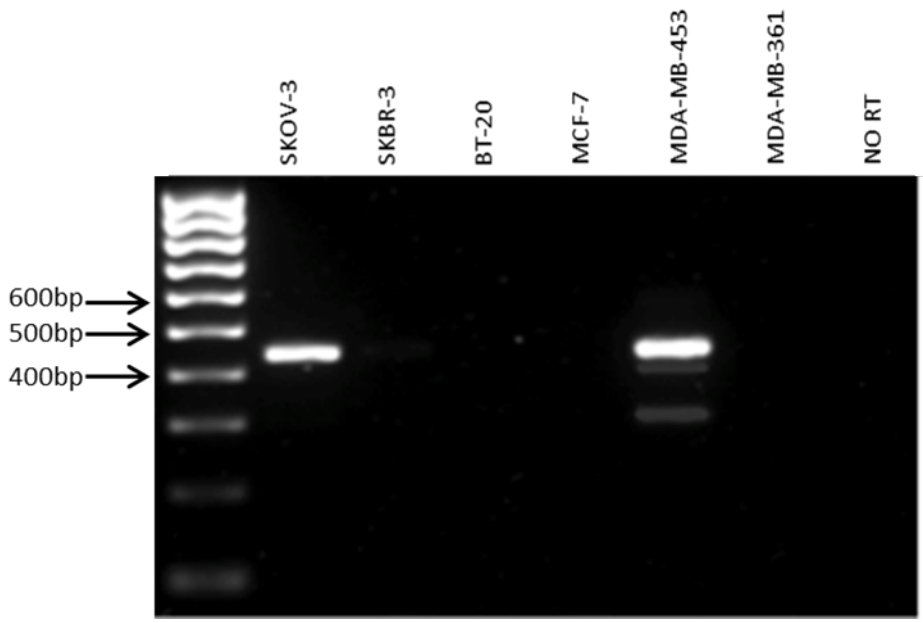


Figure 3.10: RT-PCR amplification of *HER2* exons 12-15 (primer pair E12F+E15R) using all six cell lines, and a negative (no RT) control. SKOV3, SKBR3 and MDA-MB-453 show expected band sizes at 432 base pairs, and MDA-MB-453 shows two unexpected additional bands, indicative of potential novel alternative splice variants. Each rung of the hyperladder IV represents 100bp.

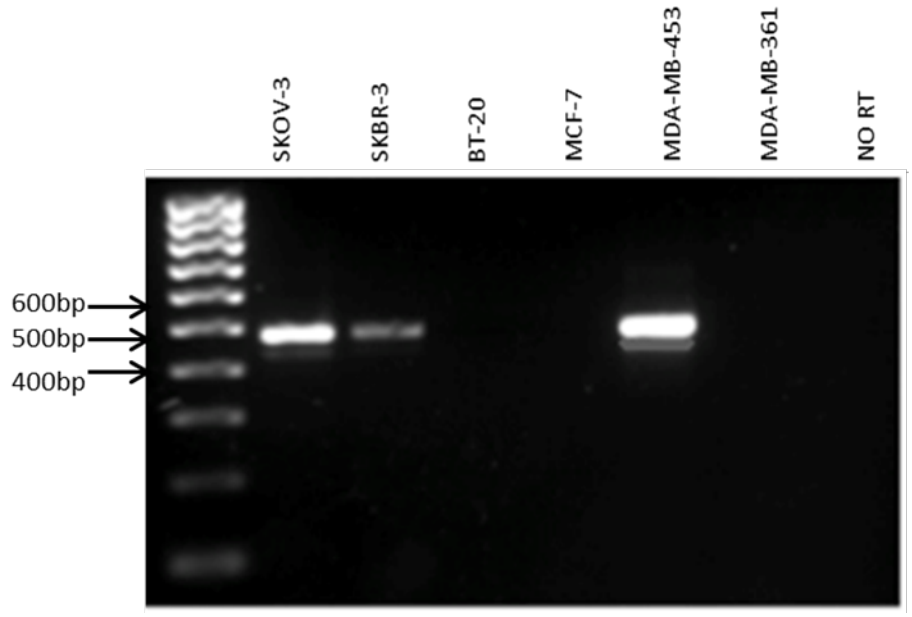


Figure 3.11: RT-PCR amplification of *HER2* exons 15-19 (primer pair E15F+E19R) using all six cell lines, and a negative (no RT) control. SKOV3, SKBR3 and MDA-MB-453 show expected band sizes at 480 base pairs; SKOV-3, SKBR-3, and MDA-MB-453 show one unexpected additional band, indicative of potential novel alternative splice variants. Each rung of the hyperladder IV represents 100bp.

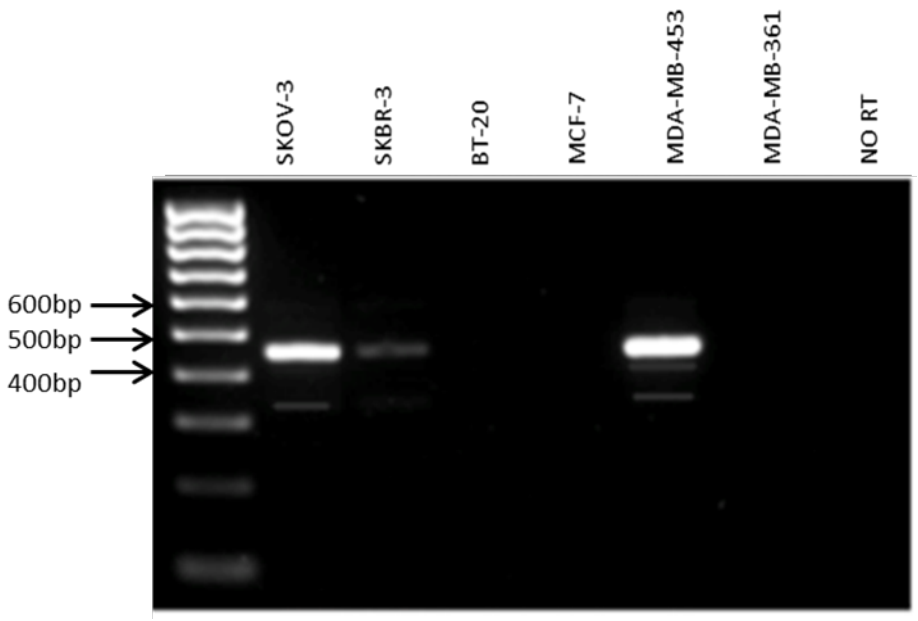


Figure 3.12: RT-PCR amplification of *HER2* exons 19-22 (primer pair E19F+E22R) using all six cell lines, and a negative (no RT) control. SKOV3, SKBR3 and MDA-MB-453 show expected band sizes at 450 base pairs; SKOV3 and MDA-MB-453 shows two unexpected additional bands, indicative of potential novel alternative splice variants. Each rung of the hyperladder IV represents 100bp.

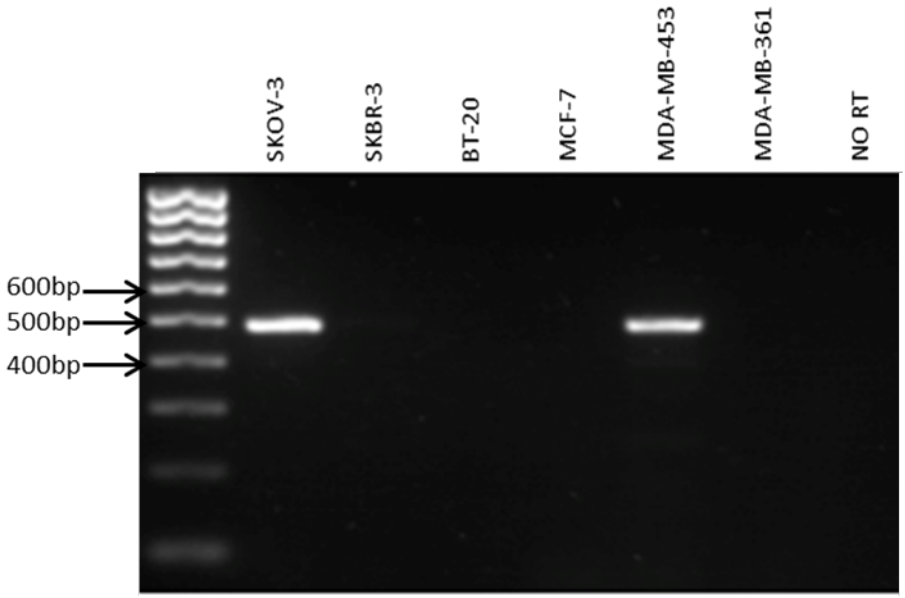


Figure 3.13: RT-PCR amplification of *HER2* exons 22-25 (primer pair E22F+E25R) using all six cell lines, and a negative (no RT) control. SKOV3 and MDA-MB-453 show expected band sizes at 489 base pairs. Each rung of the hyperladder IV represents 100bp.

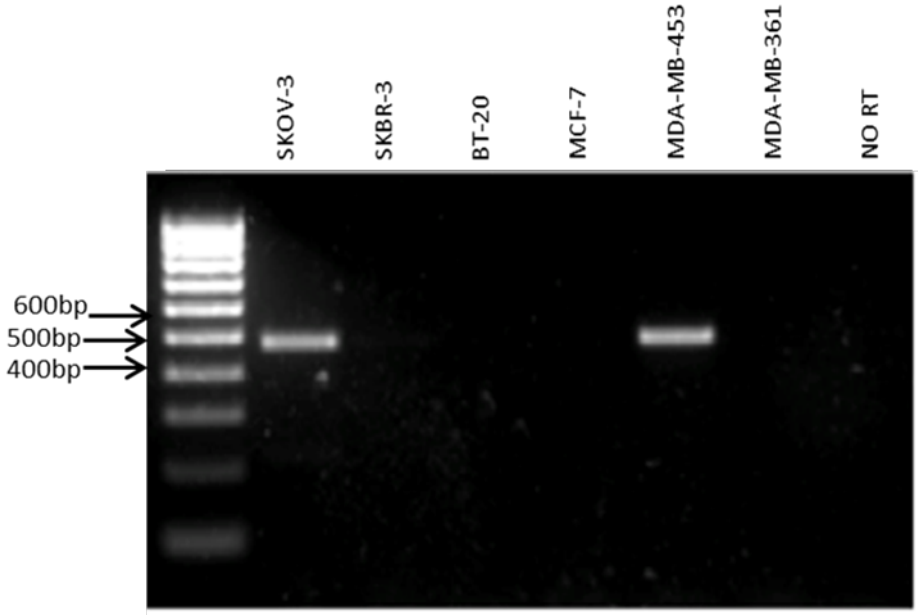


Figure 3.14: RT-PCR amplification of *HER2* exons 25-27 (primer pair E25F+E27R) using all six cell lines, and a negative (no RT) control. SKOV3 and MDA-MB-453 show expected band sizes at 450 base pairs. Each rung of the hyperladder IV represents 100bp.

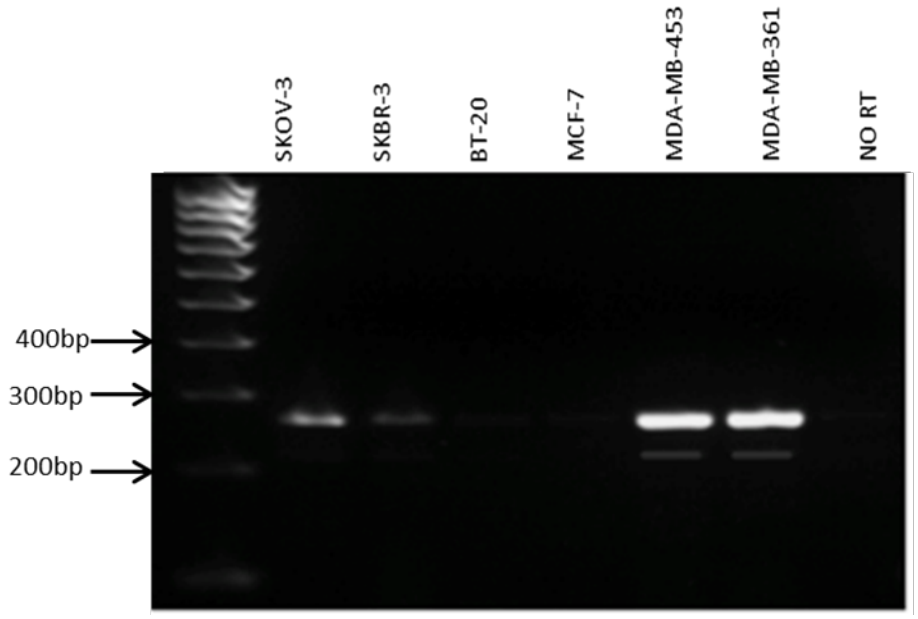


Figure 3.15: RT-PCR amplification of *HER2* exons 16-18 (primer pair NP1+NP2) using all six cell lines, and a negative (no RT) control. SKOV3, SKBR3, MDA-MB-453 and MDA-MB-361 show expected band sizes at 266 base pairs; MDA-MB-453 and MDA-MB-361 show one expected additional band at approx. 224bp. Each rung of the hyperladder IV represents 100bp.

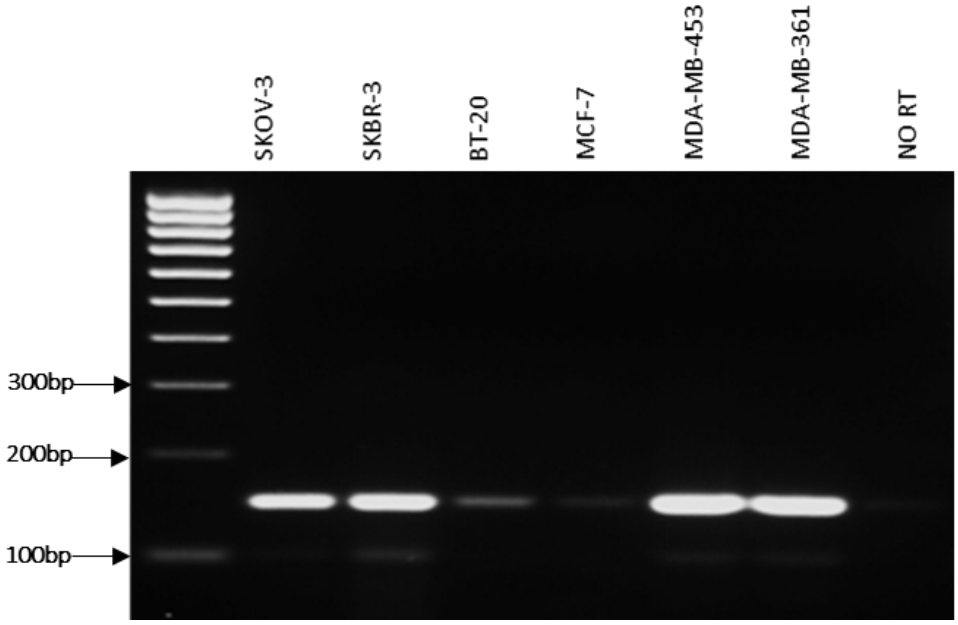


Figure 3.16: RT-PCR amplification of *HER2* exons 16-18 (primer pair NP5+NP6) using all six cell lines, and a negative (no RT) control. SKOV3, SKBR3, MDA-MB-453 and MDA-MB-361 show expected band sizes at 146 base pairs; SKBR3 shows one additional band at approximately 104bp. Each rung of the hyperladder IV represents 100bp.

Quality control results for cDNA and RNA templates used for the assessment of *HER2* mRNA expression are shown in tables 3.1 and 3.2.

SAMPLE	ng/µl	260/280
SKOV3	1019.6	1.99
SKBR3	696.8	1.95
BT-20	441.2	1.90
MCF-7	501.7	1.93
MDA-MB-453	494.3	1.94
MDA-MD-361	432.2	1.92

Table 3.1: RNA concentrations and absorbance at 260/280 for each cell line. The ratio of absorbance at 260/280 is used to assess the purity of RNA and DNA. A ratio of ~2.0 and ~1.8 is generally accepted as 'pure' for RNA and DNA respectively.

SAMPLE	ng/µL	260/280
SKOV3	1718.3	1.8
SKBR3	1709.9	1.8
BT-20	2036.4	1.77
MCF-7	1764.7	1.78
MDA-MB-453	1666.8	1.79
MDA-MD-361	1661.3	1.78

Table 3.2: cDNA concentrations and absorbance at 260/280 for each cell line. The ratio of absorbance at 260/280 is used to assess the purity of RNA and DNA. A ratio of ~2.0 and ~1.8 is generally accepted as 'pure' for RNA and DNA respectively.

3.3.3 Analysis of *HER2* cDNA amplicon sequences

To ensure that the results obtained from the sequences were actual *HER2* mRNA, the sequences were entered into NCBI Nucleotide Basic Local Alignment Search Tool (BLASTN) after which the insert sequences were then aligned to the full *HER2* sequence using Clustal Omega to determine homology between sequences obtained from GenBank and sequences obtained from cloned inserts of PCR amplified products. Bioinformatic analysis was then carried out on the regions of the *HER2* gene with potential alternative splice variants observed as loss of entire exons (cassette exons) or loss of part of an exon (alternative 3' or 5' splice sites). Full sequence alignment of the rest of the *HER2* exons resulting from the RT-PCR products which did not show multiple bands, or which did not have alternative splice variants, are found in Appendix A.

HER2 Exons 12-15; top band (Primers E12F + E15R)

```
HER2FL      CTTCGTGCACACGGTGCC-CTGGGACCAGCTCTTTCGGAACCCGCACCAAGCTCTGCTCC
INSERT      -----TGCAGGCGGCCGCGAATTCACTAGT--GATTGGAACCCGCACCAAGCTCTGCTCC
                 **** .***   * .:   ** **    :* ************************

HER2FL      ACACTGCCAACCGGCCAGAGGACGAGTGTGTGGGCGAGGGCCTGGCCTGCCACCAGCTGT
INSERT      ACACTGCCAACCGGCCAGAGGACGAGTGTGTGGGCGAGGGCCTGGCCTGCCACCAGCTGT
            ************************************************************

HER2FL      GCGCCCGAGGGCACTGCTGGGGTCCAGGGCCCACCCAGTGTGTCAACTGCAGCCAGTTCC
INSERT      GCGCCCGAGGGCACTGCTGGGGTCCAGGGCCCACCCAGTGTGTCAACTGCAGCCAGTTCC
            ************************************************************

HER2FL      TTCGGGGCCAGGAGTGCGTGGAGGAATGCCGAGTACTGCAGGGGCTCCCCAGGGAGTATG
INSERT      TTCGGGGCCAGGAGTGCGTGGAGGAATGCCGAGTACTGCAGGGGCTCCCCAGGGAGTATG
            ************************************************************

HER2FL      TGAATGCCAGGCACTGTTTGCCGTGCCACCCTGAGTGTCAGCCCCAGAATGGCTCAGTGA
INSERT      TGAATGCCAGGCACTGTTTGCCGTGCCACCCTGAGTGTCAGCCCCAGAATGGCTCAGTGA
            ************************************************************

HER2FL      CCTGTTTTGGACCGGAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCT
INSERT      CCTGTTTTGGACCGGAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCT
            ************************************************************

HER2FL      TCTGCGTGGCCCGCTGCCCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGA
INSERT      TCTGCGTGGCCCGCTGCCCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGA
            ************************************************************

HER2FL      AGTTTCCAGATGAGGAGGGCGCATGCCAGCCTTGCCCCATCAACTGCACCC---ACTCCT
INSERT      AGTTTCCAGATGAGGAGGGCGCATGCCAGCCTTGCCCCATCAACTGCAAATCGAATTCCC
            ************************************************..    * ***
```

Figure 3.17: Sequence alignment of *HER2* insert with the reference *HER2* exons 12-15 using Clustal Omega. The alignment shows the expected gene sequence for the region amplified in the top band using primer pairs E12F + E15R.

HER2 Exons 12-15; bottom band (Primers E12F + E15R)

```
HER2FL      TGGCGCCTACTCGCTGACCCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACT
INSERT      TGGCGCCTACTCGCTGACCCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACT
            ************************************************************

HER2FL      GAGGGAACTGGGCAGTGGACTGGCCCTCATCCACCATAACACCCACCTCTGCTTCGTGCA
INSERT      GAGGGAACTGGGCAGTGGACTGGCCCTCATCCACCATAACACCCACCTCTGCTTCGTGCA
            ************************************************************

HER2FL      CACGGTGCCCTGGGACCAGCTCTTTCGGAACCCGCACCAAGCTCTGCTCCACACTGCCAA
INSERT      CACGGTGCCCTGGGACCAGCTCTTTCGGAACCCGCACCAAGCTCTGCTCCACACTGCCAA
            ************************************************************

HER2FL      CCGGCCAGAGGACGAGTGTGTGGGCGAGGGCCTGGCCTGCCACCAGCTGTGCGCCCGAGG
INSERT      CCGGCCAGAGGACGAGTGTG----------------------------------------
            ********************

HER2FL      GCACTGCTGGGGTCCAGGGCCCACCCAGTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCA
INSERT      ------------------------------------------------------------


HER2FL      GGAGTGCGTGGAGGAATGCCGAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAG
INSERT      ---------------------------------GCTCCCCAGGGAGTATGTGAATGCCAG
                                             ****************************

HER2FL      GCACTGTTTGCCGTGCCACCCTGAGTGTCAGCCCCAGAATGGCTCAGTGACCTGTTTTGG
INSERT      GCACTGTTTGCCGTGCCACCCTGAGTGTCAGCCCCAGAATGGCTCAGTGACCTGTTTTGG
            ************************************************************

HER2FL      ACCGGAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGC
INSERT      ACCGGAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGC
            ************************************************************

HER2FL      CCGCTGCCCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGA
INSERT      CCGCTGCCCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGA
            ************************************************************

HER2FL      TGAGGAGGGCGCATGCCAGCCTTGCCCCATCAACTGCACCCACTC
INSERT      TGAGGAGGGCGCATGCCAGCCTTGCCCCATCAACTGCACCCACTC
            **********************************************
```

Figure 3.18: Sequence alignment of *HER2* insert with the wild type *HER2* exons 12-15 using Clustal Omega. The alignment shows the deletions in the gene sequence for the region amplified in the top band using primer pairs E12F + E15R. The missing sequence corresponds with the skipping of exon 13.

HER2 Exons 15-19; top band (Primers E15F + E19R)

```
HER2        GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGC
INSERT      GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGC
            ************************************************************

HER2        TGCCCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAG
INSERT      TGCCCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAG
            ************************************************************

HER2        GAGGGCGCATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGAC
INSERT      GAGGGCGCATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGAC
            ************************************************************

HER2        AAGGGCTGCCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTT
INSERT      AAGGGCTGCCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTT
            ************************************************************

HER2        GGCATTCTGCTGGTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAG
INSERT      GGCATTCTGCTGGTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAG
            ************************************************************

HER2        CAGAAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCG
INSERT      CAGAAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCG
            ************************************************************

HER2        CTGACACCTAGCGGAGCGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGAG
INSERT      CTGACACCTAGCGGAGCGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGAG
            ************************************************************

HER2        CTGAGGAAGGTGAAGGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGG
INSERT      CTGAGGAAGGTGAAGGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGG
            ************************************************************

HER2        ATCCCTGATGGGGAGAATGTGAAAATTCCAGTGGCCATCAAAGTGTTGAGGGAAAACACA
INSERT      ATCCCTGATGGGGAGAATGTGAAAATTCCAGTGGCCATCAAAGTGTTGAGGGAAAACACA
            ************************************************************

HER2        TCCCCCAAAGCCAACAAAGAAATCTTAGAC
INSERT      TCCCCCAAAGCCAACAAAGAAATCTTAGAC
            ******************************
```

Figure 3.19: Sequence alignment of *HER2* insert with the wild type *HER2* exons 15-19 using Clustal Omega. The alignment shows the expected gene sequence for the region amplified in the top band using primer pairs E15F + E19R.

HER2 Exons 15-19; lower band (Primers E15F + E19R)

```
HER2FL          GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGC 60
INSERT          GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGC 60
                ************************************************************

HER2FL          TGCCCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAG 120
INSERT          TGCCCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAG 120
                ************************************************************

HER2FL          GAGGGCGCATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGAC 180
INSERT          GAGGGCGCATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGAC 180
                ************************************************************

HER2FL          AAGGGCTGCCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTT 240
INSERT          AAGGGCTGCCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTT 240
                ************************************************************

HER2FL          GGCATTCTGCTGGTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAG 300
INSERT          GGCATTCTGCTGGTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAG 300
                ************************************************************

HER2FL          CAGAAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCG 360
INSERT          CAGAAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCG 360
                ************************************************************

HER2FL          CTGACACCTAGCGGAGCGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGAG 420
INSERT          CTGACACCTAGCGGAGCGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGAG 420
                ************************************************************

HER2FL          CTGAGGAAGGTGAAGGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGG 480
INSERT          CTGAGGAAGG-----------------------------------------GCATCTGG 438
                **********                                         ********

HER2FL          ATCCCTGATGGGGAGAATGTGAAAATTCCAGTGGCCATCAAAGTGTTGAGGGAAAACACA 540
INSERT          ATCCCTGATGGGGAGAATGTGAAAATTCCAGTGGCCATCAAAGTGTTGAGGGAAAACACA 498
                ************************************************************

HER2FL          TCCCCCAAAGCCAACAAAGAAATCTTAGAC 570
INSERT          TCCCCCAAAGCCAACAAAGAAATCTTAGAC 528
                ******************************
```

Figure 3.20: Sequence alignment of *HER2* insert with the wild type *HER2* exons 15-19 using Clustal Omega. The alignment shows the deletions in the gene sequence for the region amplified in the top band using primer pairs E15F + E19R. This deletion corresponds with the use of an alternative 3' splice site in exon 18.

HER2 Exons 15-18; top band (Primers NP1 + NP2)

```
HER2            GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGC 60
INSERT          GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGC 60
                ************************************************************

HER2            TGCCCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAG 120
INSERT          TGCCCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAG 120
                ************************************************************

HER2            GAGGGCGCATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGAC 180
INSERT          GAGGGCGCATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGAC 180
                ************************************************************

HER2            AAGGGCTGCCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTT 240
INSERT          AAGGGCTGCCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTT 240
                ************************************************************

HER2            GGCATTCTGCTGGTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAG 300
INSERT          GGCATTCTGCTGGTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAG 300
                ************************************************************

HER2            CAGAAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCG 360
INSERT          CAGAAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCG 360
                ************************************************************

HER2            CTGACACCTAGCGGAGCGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGAG 420
INSERT          CTGACACCTAGCGGAGCGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGAG 420
                ************************************************************

HER2            CTGAGGAAGGTGAAGGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAG 471
INSERT          CTGAGGAAGGTGAAGGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAG 471
                ***************************************************
```

Figure 3.21: Sequence alignment of *HER2* exons 15-18 using Clustal Omega. The alignment shows the expected gene sequence for the region amplified in the top band using primer pairs NP1 + NP2.

HER2 Exons 15-18; bottom band (Primers NP1 + NP2)

```
HER2            GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGC 60
INSERT          GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGC 60
                ************************************************************

HER2            TGCCCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAG 120
INSERT          TGCCCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAG 120
                ************************************************************

HER2            GAGGGCGCATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGAC 180
INSERT          GAGGGCGCATGCCAGCCTTGCCCCATCAACTGCACCCACTCC------------------ 162
                ****************************************

HER2            AAGGGCTGCCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTT 240
INSERT          ------------------------------CCTCTGACGTCCATCATCTCTGCGGTGGTT 192
                                              ******************************

HER2            GGCATTCTGCTGGTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAG 300
INSERT          GGCATTCTGCTGGTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAG 252
                ************************************************************

HER2            CAGAAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCG 360
INSERT          CAGAAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCG 312
                ************************************************************

HER2            CTGACACCTAGCGGAGCGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGAG 420
INSERT          CTGACACCTAGCGGAGCGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGAG 372
                ************************************************************

HER2            CTGAGGAAGGTGAAGGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAG 471
INSERT          CTGAGGAAGGTGAAGGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAG 423
                ***************************************************
```

Figure 3.22: Sequence alignment of *HER2* exons 15-18 using Clustal Omega. The alignment shows deletions in the gene sequence for the region amplified in the bottom band using primer pairs NP1 + NP2. This deletion corresponds to the skipping of exon 16.

HER2 Exons 15-17; Top band (Primers NP5 + NP6)

```
HER2        GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGC
INSERT      GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGC
            ************************************************************

HER2        TGCCCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAG
INSERT      TGCCCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAG
            ************************************************************

HER2        GAGGGCGCATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGAC
INSERT      GAGGGCGCATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGAC
            ************************************************************

HER2        AAGGGCTGCCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTT
INSERT      AAGGGCTGCCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTT
            ************************************************************

HER2        GGCATTCTGCTGGTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAG
INSERT      GGCATTCTGCTGGTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAG
            ************************************************************

HER2        CAGAAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAG
INSERT      CAGAAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAG
            ************************************************
```

Figure 3.23: Sequence alignment of *HER2* exons 15-17 using Clustal Omega. The alignment shows the expected gene sequence for the region amplified in the top band using primer pairs NP5 + NP6.

HER2 Exons 15-17; Bottom band (Primers NP5 + NP6)

```
HER2        GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGC
INSERT      GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGC
            ************************************************************

HER2        TGCCCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAG
INSERT      TGCCCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAG
            ************************************************************

HER2        GAGGGCGCATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGAC
INSERT      GAGGGCGCATGCCAGCCTTGCCCCATCAACTGCACCCACTC-------------------
            ****************************************

HER2        AAGGGCTGCCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTT
INSERT      ------------------------------CCTCTGACGTCCATCATCTCTGCGGTGGTT
                                          ******************************

HER2        GGCATTCTGCTGGTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAG
INSERT      GGCATTCTGCTGGTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAG
            ************************************************************

HER2        CAGAAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAG
INSERT      CAGAAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAG
            ************************************************
```

Figure 3.24: Sequence alignment of *HER2* exons 15-17 using Clustal Omega. The alignment shows deletions in the gene sequence for the region amplified in the bottom band using primer pairs NP5 + NP6. This deletion corresponds to the skipping of exon 16.

Of all the additional bands observed in RT-PCR experiments, only the bands obtained with E12F+E15R, E15F+E19R, NP1+NP2 and NP5+NP6 were successfully sequenced. Additional bands were observed with primer pairs E19F+E22R, and E12F+E15R. Sequencing of these bands, however, did not produce reliable results. This may have been as a result of inadequate template concentrations required for the reliable sequencing of plasmids or artefacts produced in the cDNA amplification process. RT-PCR amplification of exons 1 and 2 did not produce bands after agarose gel electrophoresis. Sequence alignments for all other primer pairs for the *HER2* gene where additional bands were not observed, can be found in Appendix A.

3.4 Summary

Analyses of SKBR3, BT-20 and MCF-7 cell lines by immunohistochemistry show that *HER2* is expressed in SKBR-3 cell lines, and not in BT-20 and MCF-7 cell lines. This is consistent with published studies on the *HER2* expression in these cell lines (Figures 3.2-3.6). Immunohistochemical testing of breast tumours is a reliable method of testing for *HER2* status in breast cancer patients. However, overall protein expression as determined by IHC may not be conclusive in determining a patient's response to treatment.

RT-PCR analysis of cell lines shows that there are novel alternative splice variants in *HER2*. In addition to the already published exon 16 deleted *HER2* isoform, two novel splice variants of *HER2* have been successfully characterised; a cassette exon in exon 13, in which the entire exon is skipped, and an alternative 5' splice site in exon 18, in which the 42 bases corresponding with 14 amino acids at the 3' end of the exon have been skipped.

CHAPTER 4. BIOINFORMATIC ANALYSIS OF *HER2* AND *HER2* ALTERNATIVE SPLICE VARIANTS

4.1 Introduction

The Human Epidermal Growth Factor Receptor 2 (*HER2*) gene (accession number NM_004448) was first isolated in 1985 (Coussens *et al.*, 1985; Semba *et al.*, 1985; Bargmann, Hung & Weinberg, 1986) in rat NIH3T3 cells as an oncogene called *neu,* and shown to reside on chromosome 17 (Figure 4.1). *HER2* is classified as one of the most important genes in human cancer, because of its frequent amplification in cancers such as breast carcinomas (Shih *et al.*, 1981). Schechter et al (1985) showed that the *neu* gene encoded a protein of relative molecular mass 185,000 (p185) (Schechter *et al.*, 1985). The *neu* gene was found to share significant similarity to the avian erythroblastosis virus (*v-erbB*), and was homologous with the cellular gene (*c-erbB*) which encodes *EGFR* (Downward *et al.*, n.d.; Vennström & Bishop, 1982; Yamamoto *et al.*, 1983). Schechter et al (1985) also demonstrated that the homology between the proteins encoded by the *neu* and *c-erbB* genes were limited to the kinase domain of the *EGFR* protein, and the human *v-erbB*-related sequence was then identified, and shown to be distinct from *EGFR* (Yamamoto *et al.*, 2011). The *HER2* nucleic acid sequence is approximately 4.8kb long, the open reading frame encodes a 1255 amino acid protein approximately 185kDa in mass. An extensive bioinformatic analysis of the *HER2* gene and protein structure is available on databases such as ScanProsite, UniprotKB and NCBI. With the discovery of novel splice variants of *HER2* as described in chapter 2, the use of bioinformatics was interrogated to investigate the potential

functional and structural differences between alternative splice variants of *HER2*, and the regulation of splicing in the *HER2* gene.

4.2 Objectives

1. To review the literature and databases on the bioinformatics of *HER2* gene and protein.
2. To use bioinformatics to analyse the functional properties of the wild-type *HER2* gene.
3. To use bioinformatics to understand the role and function of *HER2* alternative splice variants in comparison to the wild-type *HER2*.

4.3. Methods

4.3.1 *HER2* sequence retrieval

The *HER2* mRNA sequence was obtained using NCBI Refseq®, a genetic sequence database which contains an annotated collection of all publicly available DNA sequences. *HER2* exons were configured using the Friendly Alternative Splicing and Transcripts DataBase (FastDB; GenoSplice technology, Paris). The *HER2* protein sequence was obtained using the UniprotKB blast tool, a database which consists of high quality and freely accessible resource of protein sequences and their functional information. The *HER2* protein sequence was also derived from the *HER2* RNA sequence using the ExPASy translate tool (Swiss Institute of Bioinformatics). The *HER2*

open reading frame was predicted using ExPASy Translate Tool and the NCBI ORF finder.

4.3.2 Alignment of *HER2* transcript variants and isoforms

HER2 transcript variants and *HER2* isoforms obtained from the NCBI database were aligned using the Clustal Omega multiple sequence alignment tool (Appendix B). Novel *HER2* splice variants and their isoforms characterised during this study were aligned in comparison with the *HER2* transcript variant 1 mRNA (Accession number NM_004448) using the Clustal Omega multiple sequence alignment tool (Appendix B).

4.3.3 Analysis of potential splice factor binding sites in *HER2* alternative splice variants.

SpliceAid, a database of experimentally assessed human RNA target sequences (introni.it, 2013), was used to identify motifs that may predict the pattern of RNA splicing by identifying splice factors which are involved in *HER2* splicing. The SpliceAid database was used to predict binding motifs for exonic splice enhancers (ESE), Intronic splice silencers (ISS), exonic splice silencers (ESS) and Intronic splice enhancers (ISE) within the alternatively spliced exons and their flanking introns.

4.3.4 Structural and functional characterisation of the wild-type *HER2* protein

Structural and functional characterisation of the wild-type *HER2* protein was carried out by entering the amino acid sequence of *HER2* isoform 1 (accession number P02464) into various bioinformatic databases; this variant encodes the longest protein isoform, and has been chosen as the 'canonical' *HER2* sequence. All comparative analyses in this study have been made in reference to it. The *HER2* protein sequence analysis was carried out using UniprotKB, and the prediction of *HER2* domain structure and function was obtained using ScanProsite. The Protein Families (Pfam) database (Wellcome trust, Sanger Institute, Cambridge) was used to predict the functional domains of *HER2* and their amino acid sequences. The secondary structure of *HER2* was predicted using the PSIPRED programme. The *HER2* protein sequence was then entered into the Protein Homology/Analogy Recognition Engine V 2.0 (Phyre[2]) programme to predict its tertiary structure. ProtParam was used to predict the molecular weight, half-life and amino acid composition of *HER2.*

4.4. Results

4.4.1 *HER2* RNA sequence analysis

The NCBI Genbank® nucleotide search using the search term '*HER2*' returned the following results:

- Definition*: V-erb-b2* avian erythroblastic leukemia viral oncogene homolog 2 (*ERBB2*), transcript variant 1, mRNA.
- Accession: NM_004448.
- Source organism: Homo sapiens.
- Location: 17q12.
- Genomic sequence: 17; NC_000017.11 (39688140...39728662) reference GRCH38 p13 primary assembly.

HER2 is flanked upstream by post-*GPI* attachment to proteins 3 (*PGAP3*) and migration and invasion enhancer 1 (*MIEN1*), and downstream by titin-cap (*TCAP*), Phenylethanolamine N-methyltransferase (*PNMT*), microRNA 4728 (*MIR4728*), and growth factor receptor-bound protein 7 (*GRB7*) (Figure 4.1).

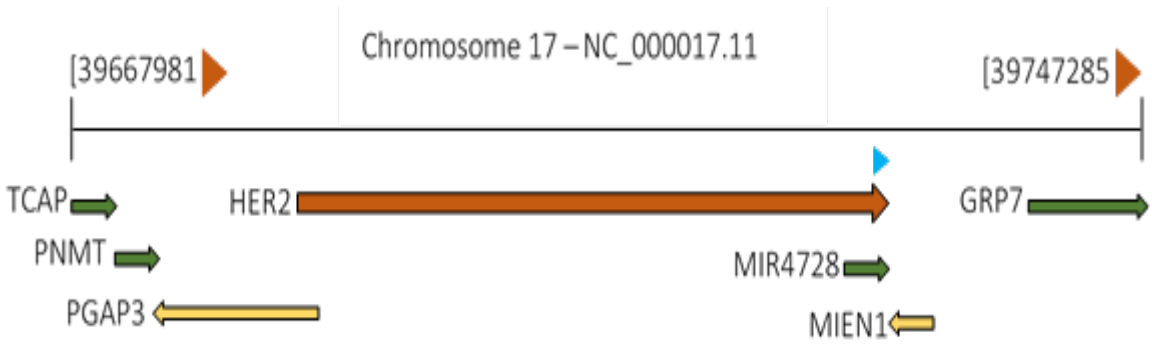


Figure 4.1. Schematic representation of the location of *HER2* gene on chromosome 17, and flanking genes.

Six transcript variants of *HER2* mRNA were obtained from the NCBI database;

1. ***Homo sapiens v-erb-b2 Avian Erythroblastic Leukaemia Viral Oncogene Homolog 2 (ERBB2), Transcript Variant 1, mRNA (Accession number NM_004448):*** This 4,664 base pair transcript encodes a protein known as *HER2* isoform a, which represents the longest *HER2* protein isoform.
2. ***Homo sapiens v-erb-b2 Avian Erythroblastic Leukaemia Viral Oncogene Homolog 2 (ERBB2), Transcript Variant 2, mRNA (Accession number NM_001005862):*** The *HER2* transcript variant 2 is 4,889 base pairs long and encodes a protein known as *HER2* isoform b. The *HER2* transcript variant 2 lacks a portion of the 5' coding region, and initiates translation at a downstream start codon, resulting in a shorter N terminus, compared to isoform a.
3. ***Homo sapiens v-erb-b2 Avian Erythroblastic Leukaemia Viral Oncogene Homolog 2 (ERBB2), Transcript Variant 3, mRNA (Accession number NM_001289936; XM_006721766):*** The *HER2* transcript variant 3 is 4,940 base pairs long and encodes a protein known as *HER2* isoform c. Unlike the transcript variant 1, the *HER2* transcript variant 3 has an alternative 5' UTR and 5' coding region, resulting in an isoform with a shorter N-terminus compared to isoform a.
4. ***Homo sapiens v-erb-b2 Avian Erythroblastic Leukaemia Viral Oncogene Homolog 2 (ERBB2), Transcript Variant 4, mRNA (Accession number NM_001289937):*** The *HER2* transcript variant 4 is 4,411 base pairs long and encodes a protein known as the *HER2* isoform d. this variant lacks an exon in

the 3' coding region, resulting in a translational frame shift. The resulting protein isoform has a distinct and shorter c-terminus compared to isoform a.

5. ***Homo sapiens v-erb-b2 Avian Erythroblastic Leukaemia Viral Oncogene Homolog 2 (ERBB2), Transcript Variant 5, mRNA (Accession number NM_001289938):*** The *HER2* transcript variant 5 is 2,590 base pairs long and encodes a protein known as the *HER2* isoform e. this variant has multiple coding differences, and differs in the 5' and 3' UTRs, compared to variant 1. The resulting isoform has a shorter N-terminus and a truncated C-terminus, compared to isoform a.
6. ***Homo sapiens v-erb-b2 Avian Erythroblastic Leukaemia Viral Oncogene Homolog 2 (ERBB2), Transcript Variant 6, long non-coding RNA (Accession number NR_110535):*** This 4,998 base pair transcript variant has an alternative 5' splice site compared to the *HER2* transcript variant 1. The *HER2* variant 6 is designated as a non-coding RNA because use of the 5'-most expected translation start codon renders the transcript a candidate for nonsense-mediated mRNA decay (NMD).

The RNA sequences of all six *HER2* transcript variants were compared using Clustal Omega multiple sequence alignment tool (Appendix B).

4.4.2 *HER2* protein sequence analysis

Four isoforms of *HER2* were obtained from the NCBI database. The protein sequences of the above *HER2* transcripts were also derived from their mRNA sequences using the Expasy Translate tool.

1. ***HER2* isoform 1** (accession number P042626; P04626-1) is a 1255 amino acid protein, encoded by Homo sapiens v-erb-b2 Avian Erythroblastic Leukaemia Viral Oncogene Homolog 2 (*ERBB2*), Transcript Variant 1, mRNA (Accession number NM_004448). This isoform is chosen as the putative (wild-type) *HER2* sequence. All positional and comparative analyses of *HER2* and *HER2* splice isoforms are made with reference to this isoform.
2. ***HER2* isoform 2** (accession number P04626-2) is 611 amino acids long, and is also known as CTF-611. This isoform is produced by alternative initiation at Met-611 of isoform 1, and is missing amino acids 1-610 of the canonical *HER2* isoform 1 sequence.
3. ***HER2* isoform 3** (accession number P04626-3) is 569 amino acids long, and is also known as CTF-687. This isoform is produced by alternative initiation at Met-687 of isoform 1, and is missing amino acids 1-687 compared to *HER2* isoform 1.
4. ***HER2* isoform 4** (accession number P04626-4) is a 1240 amino acid protein produced by alternative splicing of the 5' end of isoform 1. The alternative splicing of this *HER2* isoform produces replaces amino acids 1-23 (MELAALCRWGLLLALLPPGAAST…) with a shorter, 8-amino acid sequence (MPRGSWKP…).

The protein sequences of all *HER2* isoforms obtained from NCBI, as well as sequences derived from *HER2* transcript variants, were compared using the Clustal Omega multiple sequence alignment tool (Appendix B).

4.4.3 Structural and functional characterisation of the wild-type *HER2* (isoform 1)

The analyses of *HER2* isoform 1 (P04626) sequence for structural and functional properties predicted a 1255 amino acid protein with a protein kinase region at residues 720-987aa; 2417-2438nt, an ATP nucleotide binding region at residues 726-734; 2438-2461nt, an ATP binding site at 735aa; 2462-2465nt, and an active (proton acceptor) site at 845aa; 2794-2797nt. Further analysis of *HER2* protein structure using the UniprotKB analysis tool also predicted the signal peptide at 1-22aa; 260-325nt, the receptor tyrosine kinase domain at 23-1255aa; 326-4664nt, the transmembrane domain at 653-675aa; 2215-2284nt, and the cytoplasmic domain at 676-1255aa, 2285-4664nt. The nuclear localisation signalling region is predicted at 676-689aa; 2285-2326nt (KPNB1 and EEA1 activation site), and a PIK3C2B activation site at 1195-1197aa; 3842-3851nt.

Inputting the *HER2* RNA sequence into the Pfam programme returned four distinct domains in the extracellular region of *HER2* and one domain in the tyrosine kinase region (Figure 4.2).

- Receptor L domain: This domain constitutes subdomains I and III of the *HER2* ECD (residues 52-173; 336-468, respectively). The *HER2* receptor L domain makes up the bilobal ligand binding site, each domain consisting of a single-stranded right hand β-helix (Garrett *et al.*, 1998).
- Furin-like cysteine rich domain: This domain constitutes subdomain III of the *HER2* ECD (residues 183-343), and is usually found in eukaryotic proteins that

are involved in signal transduction by receptor tyrosine kinases (Raz, Schejter & Shilo, 1991).

- Growth factor receptor domain: This domain constitutes subdomain IV of the *HER2* ECD (residues 510-643). Interaction between the growth factor receptor domain and the furin-like domain regulates the binding of ligands to the receptor L domains (Cho & Leahy, 2002).
- Protein tyrosine kinase domain: Tyrosine kinases are a subclass of protein kinases. This domain (residues 720-976) constitutes an enzyme that can transfer a phosphate group from ATP to a protein in the cell, functioning as an 'on' or 'off switch in a variety of cellular functions (Hanks & Quinn, 1991; Hanks & Hunter, 1995; Hunter & Plowman, 1997).

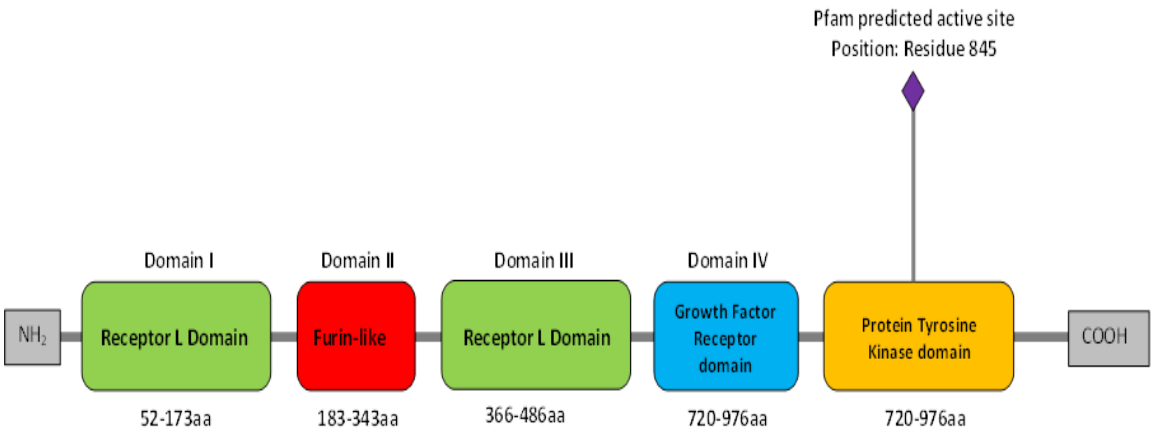


Figure 4.2 Schematic of Pfam output showing *HER2* functional domains and their positions.

The PSIPRED programme predicted the following *HER2* secondary structure: α-helix (2-14aa), β-strand (26-27aa), α-helix (39-48aa), β-strand (54-57aa, 60-64aa, 78-81aa, 83-88aa, 100-103aa, 112-117aa, 139-140aa), α-helix (147-148aa), β-strand (153-155aa, 181-184aa), α-helix (237-241aa), β-strand (272-273aa, 279-281aa, 304-305aa, 312-313aa, 321-323aa, 354-356aa), α-helix (360-362aa), β-strand (368-371aa, 374-377aa,

404-414aa, 429-432aa, 434-437aa, 440-446aa, 451-453aa, 460-461aa, 466-468aa, 493-495aa, 655-660aa, 663-667aa), α-helix (671-687aa, 709-714aa), β-strand (720-726aa, 732-741aa, 747-755aa), α-helix (761-775aa), β-strand (781-789aa, 794-799aa), α-helix (805-812aa, 819-832aa, 842-849aa), β-strand (858-861aa, 877-879aa,884-885aa, 887-888aa), α-helix (889-895aa), β-strand (911-916), α-helix (933-936aa, 949-959aa, 969-980aa), β-strand (987-988aa). The intervening sequences are random coils (appendix A).

The Phyre[2] programme predicted the 3D structure of the *HER2* protein, showing the positions of the α-helices and the β-strands (Figure 4.3).

ProtParam predicted *HER2* to have a molecular weight of 137910.5Da and an estimated half-life of 30 hours. The amino acid composition of the *HER2* protein as predicted by ProtParam is detailed in Table 4.1.

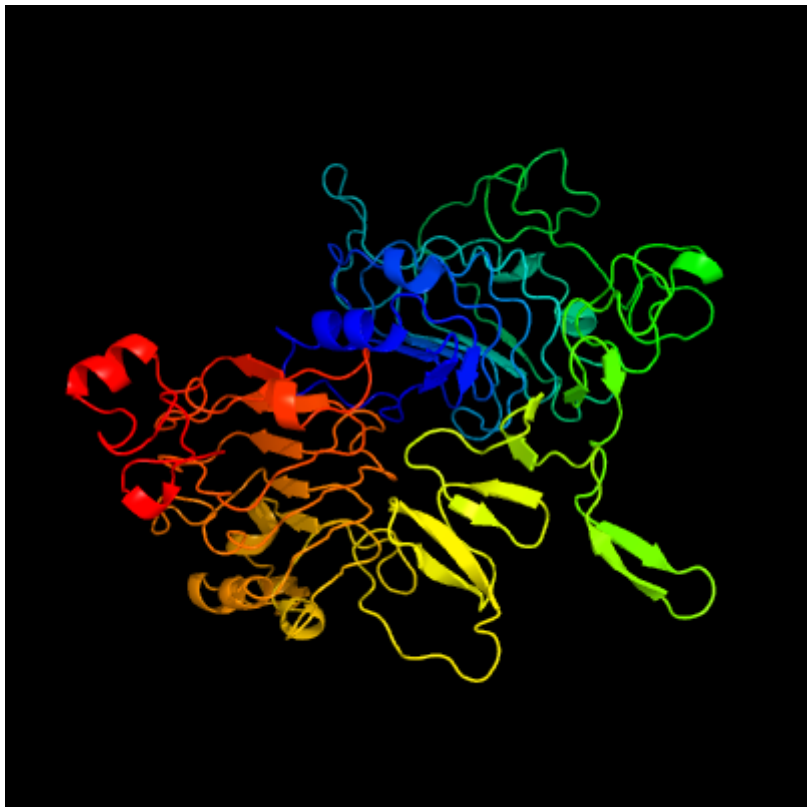


Figure 4.3 Phyre[2] output showing the 3D structure of *HER2*. The image is coloured by rainbow from N→C terminus. α-helices are represented by coils and β-strands are represented by arrowed regions.

AMINO ACID	OCCURRENCE	PERCENTAGE
Ala (A)	83	6.60%
Arg (R)	71	5.70%
Asn (N)	41	3.30%
Asp (D)	65	5.20%
Cys (C)	59	4.70%
Gln (Q)	62	4.90%
Glu (E)	77	6.10%
Gly (G)	101	8.00%
His (H)	35	2.80%
Ile (I)	44	3.50%
Leu (L)	138	11.00%
Lys (K)	39	3.10%
Met (M)	23	1.80%
Phe (F)	35	2.80%
Pro (P)	109	8.70%
Ser (S)	73	5.80%
Thr (T)	67	5.30%
Trp (W)	15	1.20%
Tyr (Y)	35	2.80%
Val (V)	83	6.60%

Table 4.1 *HER2* amino acid composition as predicted by ProtParam.

4.4.4 Analysis of potential splice factor binding sites

HER2 DNA sequences representing alternative spliced exons and their flanking introns were analysed for potential splice factor binding motifs. The results are shown in Figures 4.5-4.7.

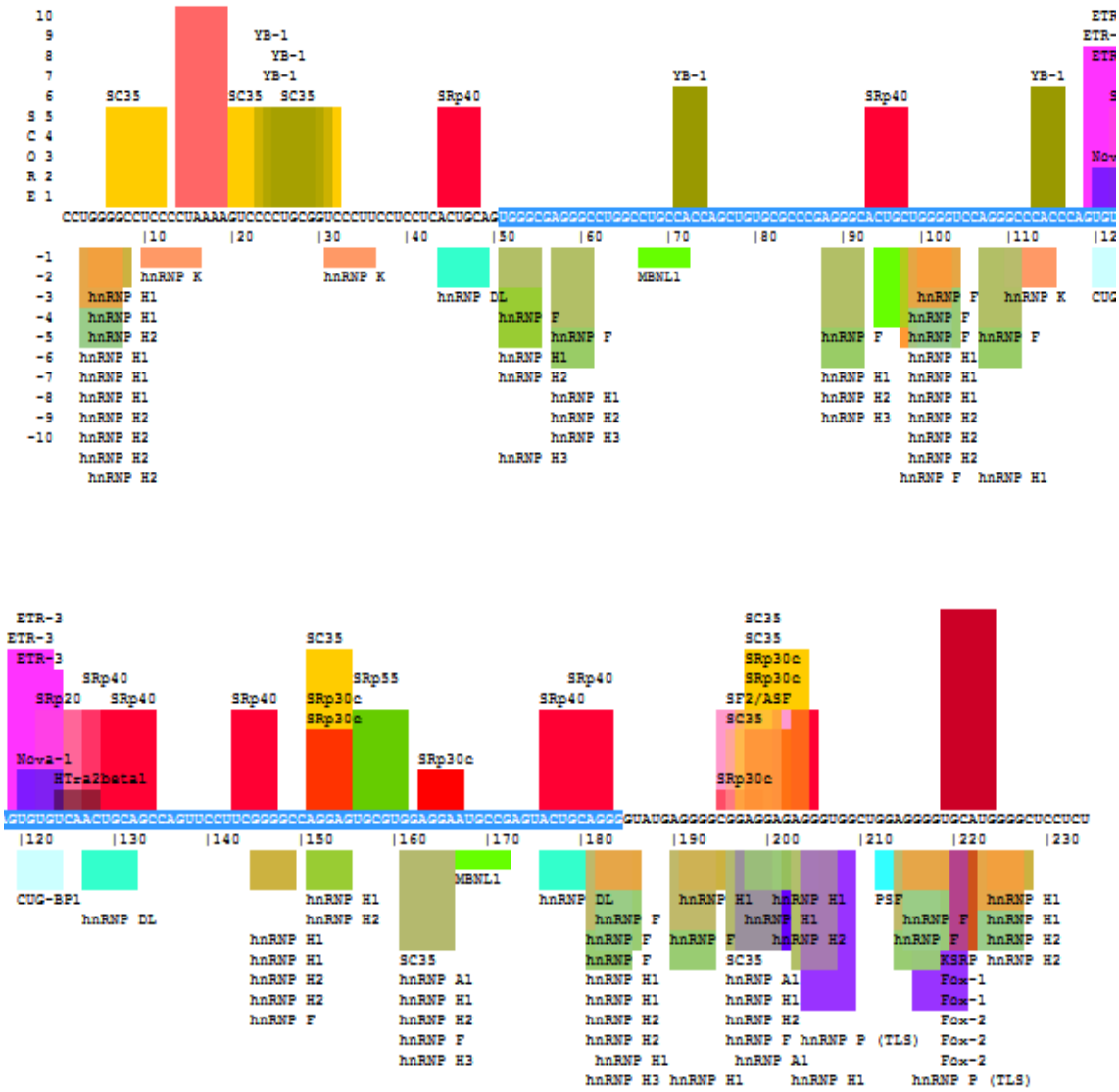


Figure 4.4 SpliceAid output for the analysis of splice factor binding motifs in exon 13 and 50 base pairs into the flanking introns. In exon 13 skipping, binding motifs that facilitate exon skipping are considerably more than those which facilitate exon inclusion. Positive scores represent target sequences that facilitate exon definition; exonic splice enhancer (ESE) and intronic splice silencer (ISS) motifs, and negative scores represent target sequences that facilitate intron definition; exonic splice silencers (ESS) and intronic splice enhancer (ISE) motifs. Target RNA sequences for splice factors are represented by histogram. Bars have variable heights and widths related to the binding affinity. The missing exon is highlighted in the DNA sequence.

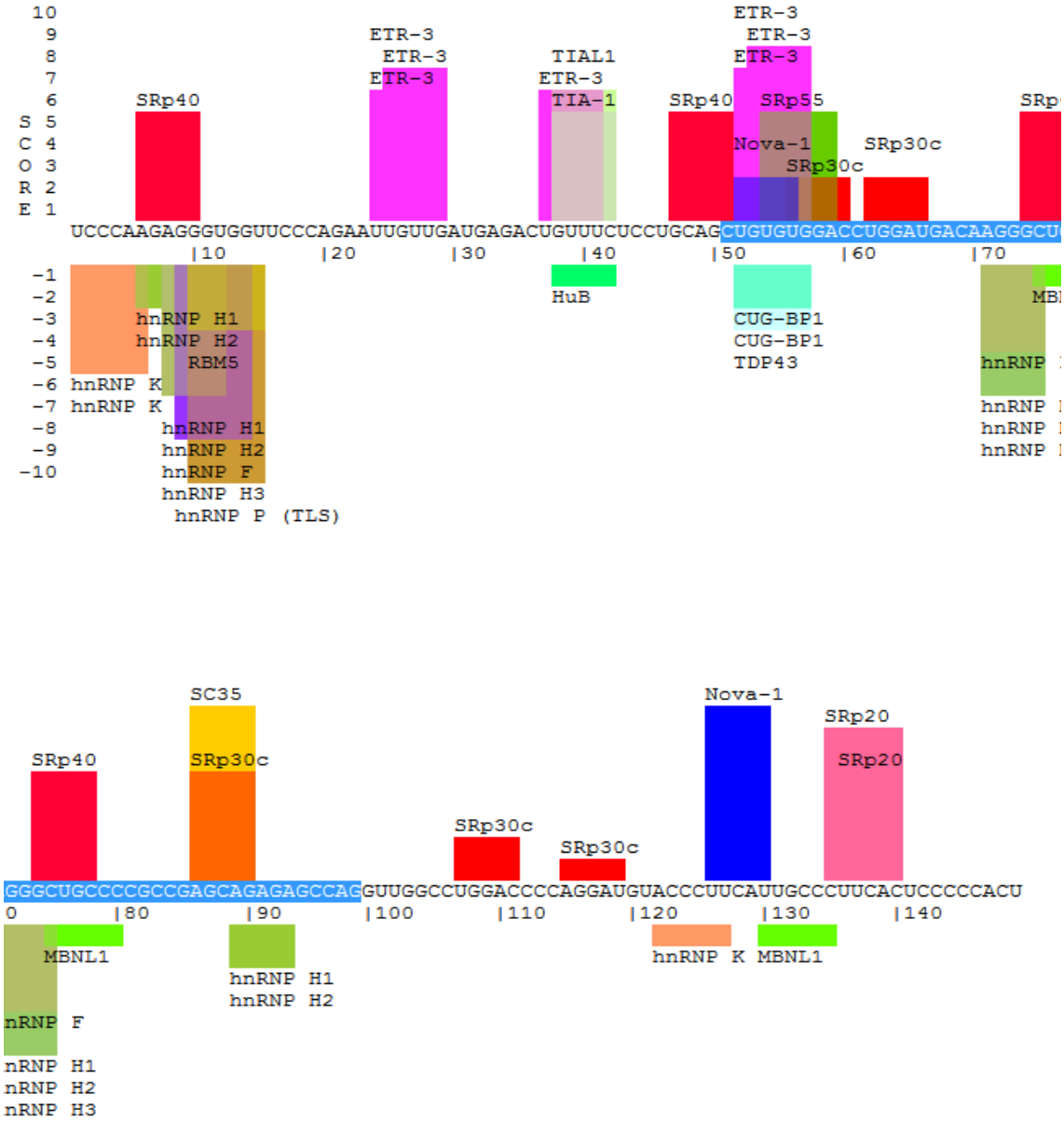


Figure 4.5 SpliceAid output for the analysis of splice factor binding motifs in exon 16 and 50 base pairs into the flanking introns. In exon 16 skipping, binding motifs that facilitate exon skipping are present in equal numbers as those which facilitate exon inclusion. Positive scores represent target sequences that facilitate exon definition; exonic splice enhancer (ESE) and intronic splice silencer (ISS) motifs, and negative scores represent target sequences that facilitate intron definition; exonic splice silencers (ESS) and intronic splice enhancer (ISE) motifs. Target RNA sequences for splice factors are represented by histogram. Bars have variable heights and widths related to the binding affinity. The missing exon is highlighted in the DNA sequence.

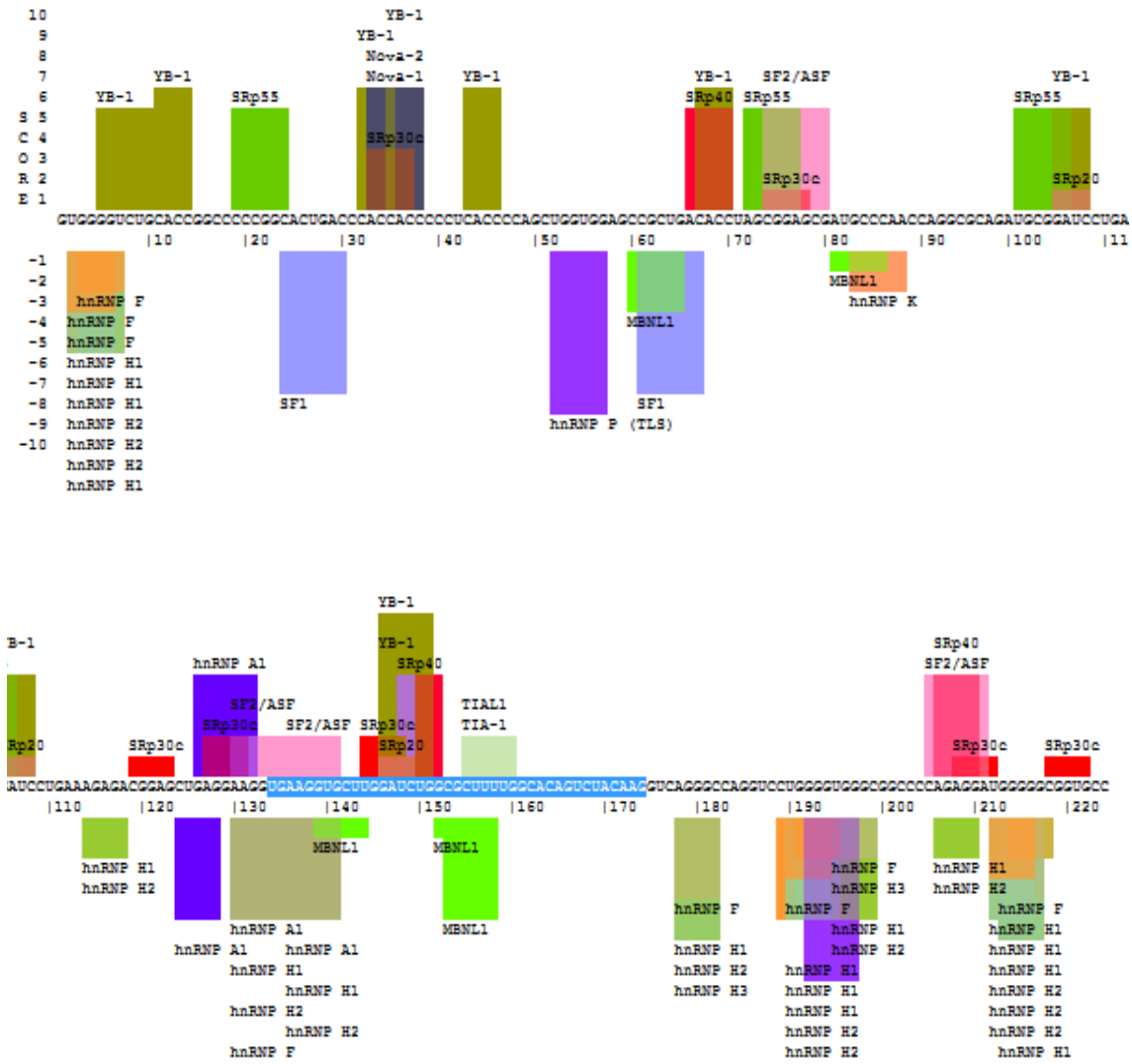


Figure 4.6 SpliceAid output for the analysis of splice factor binding motifs in exon 16 and 50 base pairs into the flanking introns. In exon 13 skipping, binding motifs that facilitate exon skipping are relatively more than those which facilitate exon inclusion. Positive scores represent target sequences that facilitate exon definition; exonic splice enhancer (ESE) and intronic splice silencer (ISS) motifs, and negative scores represent target sequences that facilitate intron definition; exonic splice silencers (ESS) and intronic splice enhancer (ISE) motifs. Target RNA sequences for splice factors are represented by histogram. Bars have variable heights and widths related to the binding affinity. The missing part of exon 18 is highlighted in the DNA sequence.

4.4.5 Post-translational modification of *HER2* protein

Post-translational modification is an enzymatic process of covalently altering one or more amino acids in a protein by either addition of functional groups or proteins, proteolytic cleavage of regulatory subunits or degradation of entire proteins. Post-translational modifications occur after translation from mRNA, and after the protein has been released from the ribosome. Post-translational modifications increase the functional diversity of the proteome and are therefore critical in cell biology. Various post-translational modifications include phosphorylation/autophosphorylation, glycosylation, ubiquitination, nitrosylation, methylation, acetylation, lipidation and proteolysis. The only researched post-translational modification of *HER2* that was found at this time was Autophosphorylation. In *HER2*, phosphorylation increases on the tyrosine residues following dimerisation. Autophosphorylation of *HER2* occurs in trans; receptor dimerisation occurs when one subunit of the dimeric receptor phosphorylates tyrosine residues on the other subunit (Deng *et al.*, 2007; Li *et al.*, 2007).

4.4.6 Structural and functional characterisation of novel *HER2* isoforms

In order to predict the structural and functional differences between the wild-type *HER2* and novel *HER2* transcript variants identified in this study, a comparative bioinformatic analysis was carried out on individual transcripts and their resulting protein isoforms. The previously identified *HER2*Δ16 transcript (Kwong & Hung, 1998)

was also analysed. All three splice variants of *HER2* are similar to the full length *HER2* transcript except for the skipping of exons 13 and 16, and an alternative 5' splice site in exon 18, respectively. Analysis of the protein isoforms of these splice variants will predict structural differences which may lead to functional changes in the *HER2* isoforms.

4.4.6.1. Additional band produced by primers E15F/E19R give rise to a loss of the *HER2* ATP binding pocket, and a novel *HER2* splice variant *HER2*ΔATP

The cDNA sequence of the multiple bands obtained using primer pairs E15F/E19R (Figures 3.22 and 3.24) were aligned using Clustal Omega multiple sequence alignment tool, and revealed a deletion of 42 base pairs (lower band) compared to the wild-type *HER2* (top band) (Figure 3.11). The amino acid sequences of both bands were obtained using ExPASy translate tool, and revealed an in-frame deletion of 14 amino acids. The structural and functional changes were compared to the *HER2* isoform 1 (P04626), and revealed the deletions in the lower *HER2* amplicon to be the loss of amino acids 724-737 in the kinase domain of the *HER2* protein, which corresponds to the ATP binding domain (the ATP binding domain is represented by amino acids 726-734). This analysis revealed a novel splice isoform containing a deletion of the 3' end of exon 18, and more specifically, the deletion of the entire ATP binding pocket (Figure 4.8). This novel alternative splice variant of *HER2* has been designated *HER2*ΔATP.

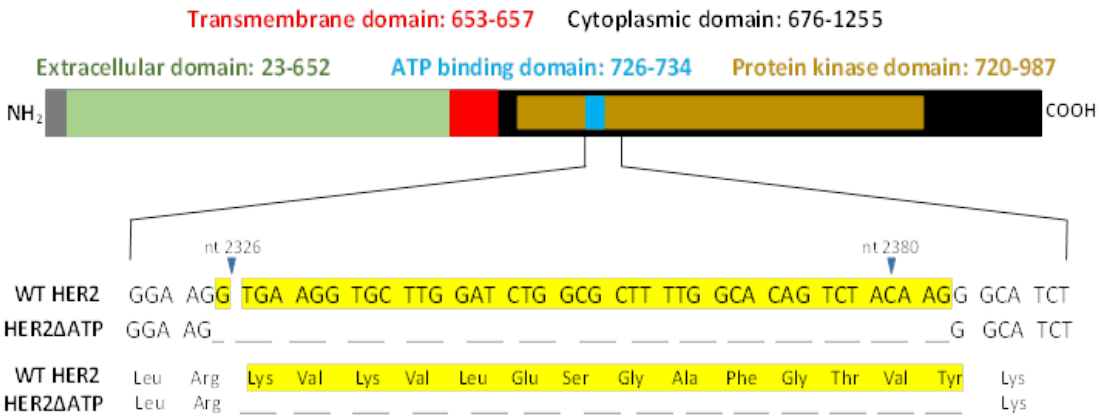


Figure 4.7 Analysis of cDNA and amino acid sequences of multiple bands obtained using primer pair E15F/E19R.

4.4.6.2. Additional band produced by primers E12F/E15R gives rise to the loss of the *HER2* extracellular domain, and a novel *HER2* splice variant *HER2ΔECD* .

The cDNA sequence of the multiple bands obtained using primer pairs E12F/E15R (Figures 3.18 and 3.20) were aligned using Clustal Omega multiple sequence alignment tool, and revealed a deletion of 133 base pairs (lower band) compared to the wild-type *HER2* (top band) (Figure 3.10). The amino acid sequence of the additional band was obtained using ExPASy, and revealed a truncated 645 amino acid *HER2* isoform. This truncated *HER2* shows a loss in amino acids 1-610 in the extracellular domain of *HER2*. Further analysis using ScanProsite and uniprotKB revealed that this novel alternative splice variant encodes a *HER2* protein with conserved active binding sites in the transmembrane domain of *HER2*, but with a loss of 644 amino acids upstream of *HER2* (*HER2ΔECD* consists 652 amino acids). This novel *HER2* alternative splice variant has been designated *HER2ΔECD* (Figure 4.9).

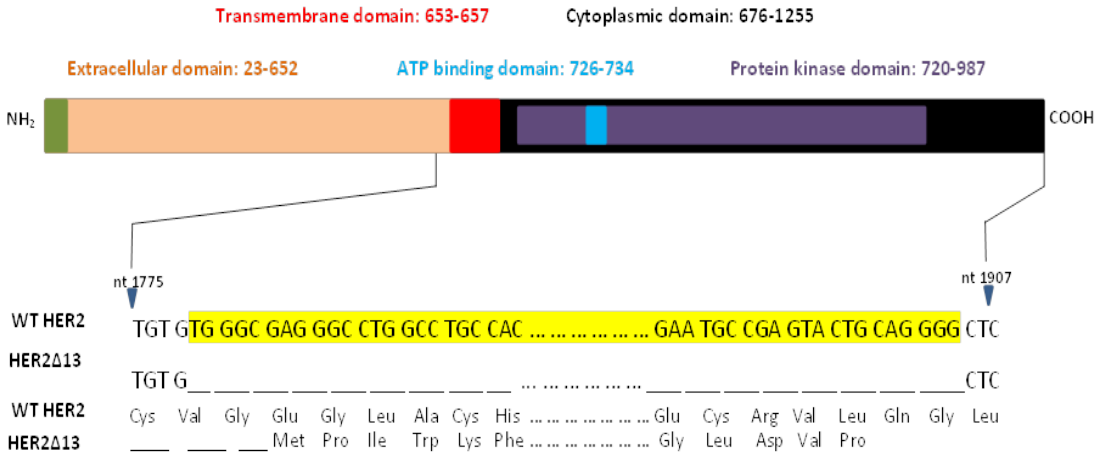


Figure 4.8 Analysis of cDNA and amino acid sequences of multiple bands obtained using primer pair E12F/E15R.

4.4.6.3. Additional bands produced using primer pairs NP1/NP2 and NP5/NP6 give rise to the HERΔ16 isoform corresponding to the loss of subdomain IV of the *HER2* extracellular domain

Both NP1/NP2 and NP5/NP6 primer sets have been used previously to identify the *HER2*Δ16 splice variant (Kwong & Hung, 1998). Although this isoform has been identified in previous studies, the present study identifies the expression of this isoform in SKOV3, SKBR3, and MDA-MB-453 and MDA-MB-361 cell lines. Expression in these cell lines has not been previously documented. Also, the splicing mechanisms underlying the deletion of exon 16 have not been elucidated. The alignment of cDNA sequences of both primer pairs confirms the expression of an alternative *HER2* isoform in addition to the wild-type *HER2*. This isoform shows a loss of exon 16, and has been shown to have increased transformation activity when expressed in *HER2* positive breast cancers (Kwong & Hung, 1998). Bioinformatics analysis using scanprosite and

uniprotkb revealed a 1239 conserved peptide with active structural functions (ATB binding domain, and tyrosine kinase domain). The cassette exon occurs in amino acids 634-649, which constitute a portion of subdomain IV of the extracellular domain of the wild-type *HER2* protein (Figure 4.10).

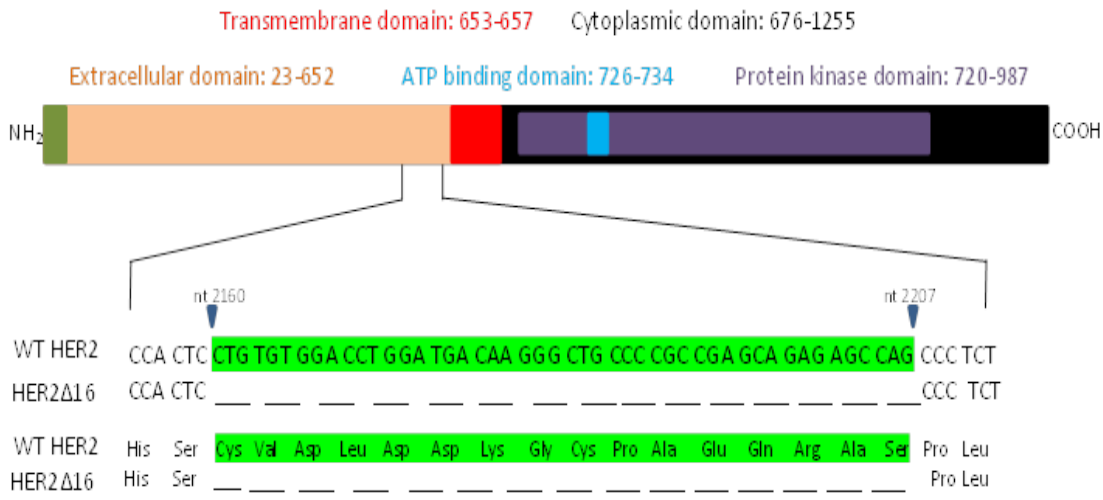


Figure 4.9 Analysis of cDNA and amino acid sequences of multiple bands obtained using primer pairs NP1/NP2 and NP5/NP6.

Using the information generated from the analyses of the *HER2* splice variants, a schematic was drawn comparing cDNA and protein sequences generated by the different splice variants.

4.5. Analysis of new 5' splice site boundaries for *HER2*ΔATP

The loss of 42 nucleotides gives rise to the loss of exactly 14 amino acids. The resulting new 5' splice site boundaries for *HER2*ΔATP were analysed in relation to the vertebrate splice site consensus.

In the full length *HER2* mRNA, exon 18 splices into exon 19 in this order:

…TCTACAAG**GGCATCTGG**…

In the new *HER2*ΔATP mRNA, exon 18 splices into exon 19 in this order:

…GAGGAAG**GGCATCTGG**…

The vertebrate 5’ splice site consensus is:

…MAG**GURAGU**…

Where:

M=A or C

R=A or G

U=C or U (T) (Elliot & Ladomery, 2011)

Therefore, the alternative 5’ splice site AAG fits the consensus MAG, which indicates that the new 5’ splice site in *HER2*ΔATP is a true isoform with conserved vertebrate 5’ splice site.

4.6. Summary

The structural and functional characterisation of *HER2* gives a better understanding of the function of the *HER2* protein, as well as the effects of alternative splicing in changing the function of the protein through change in structure.

- The loss of exon 13 gives rise to *HER2ΔECD* , a 645 amino acid protein with a loss of the entire *HER2* extracellular binding domain, and consequently a

potential loss of all signalling properties, as the *HER2* signalling domains have been lost in translation. This loss may potentially result in the loss of the binding sites of *HER2*, therefore conferring resistance of *HER2ΔECD* to *HER2* therapies targeted at the *HER2* extracellular domain.

- The loss of exon 16 constitutes the loss of *HER2* to translate amino acids 634 to 649, which is predicted to be within the region containing domain IV of the *HER2* extracellular domain. The *HER2* ECD domain IV is predicted to start and end at amino acids 510 and 643, respectively, and is designated as the growth factor receptor binding domain of *HER2* isoform 1 (P04626). The loss of exon 16 is therefore likely to alter the binding sites of *HER2*, conferring resistance to *HER2* ECD-targeted therapies.
- The utilization of the alternative 5' splice site of *HER2*ΔATP results in the loss of amino acids 722-735 in the protein tyrosine kinase domain of *HER2* (the *HER2* protein tyrosine kinase domain spans amino acids 720-976). As the *HER2* ECD remains intact, *HER2*ΔATP would still be capable of dimerization; however, the loss of amino acids in the tyrosine kinase region may inhibit phosphorylation and subsequent activation of downstream signalling pathways.
- *HER2*ΔATP is identified here as a true isoform with an alternative 5' splice site in exon 18 with a conserved 5' splice site.

CHAPTER 5. EXPRESSION OF *HER2* AND *HER2* ALTERNATIVE SPLICE VARIANTS IN NORMAL HUMAN TISSUES AND HUMAN BREAST TUMOURS

5.1 Introduction

Established human cancer cell lines derived from tumours are frequently used as *in vitro* tumour models for human cancers, and have been used to significantly advance the understanding of cancer biology (Domcke *et al.*, 2013). Human lesions obtained at surgery represent the real state of the tumour *in vivo*, and can be used to derive certain useful information such as their pathology, gene or biomarker expression, and metabolism. However, they only represent one time point in the evolution of the tumour, and therefore do not lend themselves to much experimentation (van Staveren *et al.*, 2009). Human cell lines are an example of good experimental models as they are known to retain the hallmarks of cancer cells, are easy to propagate and genetically manipulate, and can produce reproducible results when used under well-defined experimental conditions, even after numerous passages (van Staveren *et al.*, 2009). The use of cell lines in breast cancer studies has resulted in a wealth of information about deregulation of proliferation, migration and apoptosis, as well as the genes and signalling pathways that regulate these processes (Vargo-Gogola & Rosen, 2007; Neve *et al.*, 2006). However, gene expression profiles may sometimes be altered by activating mutations of kinases in cell lines which may not be present in primary breast tumours (van Staveren *et al.*, 2009).

The discovery of new alternative splice variants of *HER2* in *HER2* positive breast cancer cell lines in this study gives rise to a need for further exploration of these splice variants in human samples from a normal tissue panel and clinical cases of *HER2* positive breast cancer.

The use of a normal tissue panel and *HER2* positive breast tumour samples to test for *HER2* expression in this study was in order to investigate tumour-specificity of *HER2* alternative splice variants, particularly *HER2*ΔATP and *HER2*ΔECD. In a study by Mitra *et al* (2009), a panel of 18 normal tissues showed no expression of *HER2*Δ16. However *HER2*Δ16 was detected in 51% of a cohort of 85 primary invasive breast tumours. *HER2*Δ16 is therefore said to be a tumour-specific *HER2* oncogene (Mitra *et al.*, 2009).

For the benefit of this chapter, only the exons which have been confirmed to have alternative splice variants were analysed for the expression of *HER2* and *HER2* splice variant expression in human samples. qPCR probes were designed to target the wild-type *HER2* gene and to detect expression of *HER2*ΔECD, *HER2*Δ16 and *HER2*ΔATP .

5.2 Objectives

1. To investigate the expression of *HER2* and *HER2* alternative splice variants in normal tissues.
2. To investigate the expression of *HER2* and novel *HER2* alternative splice variants in *HER2* positive human breast cancer tissues which have been processed by freezing, and samples which have been processed by formalin fixation and embedded in paraffin wax (FFPE).

3. To compare the expression of *HER2* and novel *HER2* splice variants in normal tissues and human breast cancer tissues.

5.3 Methods

Standard RT-PCR primers listed in Table A 1 were used in the amplification of exons 12-15, exons 15-19, and exons 16-18. Taqman probes for the detection of *HER2* and *HER2* splice variants by quantitative real-time PCR were designed by Primer Design (Southampton,UK). Primer sequences are listed in Table 5.1.

PRIMER NAME	SENSE PRIMER	ANTISENSE PRIMER	PRODUCT LENGTH
ERBB2 (Global)	ACCTTCCTTCCTGCTTGAGT	GCCTCAGAATCCACAAAGACT	94
ERBB2_ex13del	CCAGAGGACGAGTGTGGAG	CGGTCCAAAACAGGTCACTG	120
ERBB2_ex16del	CAACTGCACCCACTCCCCT	CCAAGACCACGACCAGCAG	71
ERBB2_ex18del	GGAGCTGAGGAAGGGCAT	GGCTTTGGGGGATGTGTTTT	94

Table 5.1 Primer sequences for the detection of *HER2* and *HER2* splice variants by qRT-PCR.

5.3.1 Analysis of cDNA samples from a normal tissue panel for the expression of *HER2* and *HER2* alternative splice variants

A panel of ten BioBank human cDNA samples consisting of normal tissues was obtained from Primer Design (Southampton, UK). The BioBank is a high quality source of cDNA validated for use in real-time PCR experiments. The cDNA is reverse transcribed from high quality, DNAse treated RNA, from a variety of tissues or cell

cultures, using an optimised blend of oligo-dT and random nonamer primers. BioBank cDNA is therefore free of genomic DNA and PCR inhibitors and covers the widest possible range of RNA and mRNA transcripts in the specified tissue or cell line. BioBank cDNA is useful for expression profiling of newly identified genes, and also as a positive control for real-time PCR. The normal tissue panel consisted of the following tissues:

- Adipose
- Cervix
- Colon
- Kidney
- Liver
- Lung
- Ovary
- Placenta
- Prostate
- Spleen

The panel of normal tissues for used *HER2* testing was based on the repository of cDNA samples available for selection. To eliminate variations in results, tissues used were all treated from RNA extraction to reverse transcription, using the same protocol and processed at the same time. Positive control primers were also supplied with the tissue samples, which detect 18s ribosomal RNA. The tissue samples from the normal tissue panel were tested for the expression of *HER2* and *HER2* splice variants by standard PCR.

5.3.2 Analysis of frozen clinical samples from *HER2* positive breast tumours for the expression of *HER2* and *HER2* alternative splice variants

RNA samples from three matched invasive ductal carcinomas and adjacent normal tissues were obtained from the Wales cancer bank (Cardiff, UK). RNA extracted from frozen blocks is of high quality for use in techniques such as expression microarray systems. The RNA extraction was carried out by the Wales cancer bank using a Qiagen kit or Trizol® method. RNA was supplied as 5µg in 50µl aliquots. RNA quality was assessed by 260/230 and 260/280 ratio using a nanospectrophotometer, and then subjected to quality assurance by Agilent Bioanalyser. Table 5.2 shows a minimum data set for all three frozen samples, which includes the age of the patient, tumour type and grade, size of tumour and *HER2* status. For the purpose of this study, the samples were designated 01A, 02A and 03A for the breast tumours and 01B, 02B and 02B for the respective normal breast tissue obtained from each patient.

WCB No.	Age	Gender	Tumour Type	Tumour IHC Grade	Max Diameter of Invasive Tumour	Whole Size of Tumour	Receptor Status	
RR6BL0 000141 (01)	48	F	Invasive Ductal Carcinoma	3+	60	80	ER-/PGR-	*HER2* +
RR6BL0 000198 (02)	51	M	Invasive Ductal Carcinoma	3+	25	300	ER+/PGR-	*HER2* +
RR6BL0 000409 (03)	43	F	Invasive Ductal Carcinoma	3+	50	54	ER+/PGR +	*HER2* +

Table 5.2 Minimum data set for frozen samples from invasive ductal carcinomas obtained from the Wales cancer bank (Cardiff, UK). For anonymity and data protection, samples were designated numeric codes for the purpose of identification.

Following gDNA treatment and reverse transcription, the RNA samples obtained from the frozen breast tumours were tested for the expression of *HER2* and *HER2* splice variants by standard PCR and real-time PCR.

5.3.3 Analysis of formalin fixed and paraffin embedded (FFPE) clinical samples from *HER2* positive breast tumours for the expression of *HER2* and *HER2* alternative splice variants.

Total RNA was extracted from FFPE samples using the RNEasy FFPE kit (Qiagen, UK) according to manufacturer's protocols. After RNA extraction and quantification, 700ng of RNA was reverse transcribed to cDNA using Maxima H Minus Reverse Transcriptase (ThermoScientific, UK) according to the manufacturer's guidelines. The resulting cDNA was then diluted 1:10 and RT-PCR was performed using GoTaq Hotstart Taq Polymerase (Promega, UK) using the following thermal cycler program: hotstart at 95°C for 2 minutes followed by 39 cycles of 95°C for 1 minute (denaturing), 56°C for 1 minute (annealing), 72°C for 30 seconds (extending), and a final extension of 5 minutes at 72°C. A soaking cycle of 4°C was included to hold the tubes after amplification, prior to agarose gel electrophoresis, or storage at -20°C.

5.4 Results

5.4.1 Expression of *HER2* and *HER2* alternative splice variants in cDNA samples from a normal tissue panel

Amplification of *HER2*ΔECD: Figure 5.1 shows RT-PCR amplification of *HER2*ΔECD in normal human tissues. All tissue types (1-4, 6-10) except liver (5), express the wild-type *HER2*, but not *HER2*ΔECD. Liver tissue (5) does not appear to express either the wild-type *HER2* or *HER2*ΔECD.

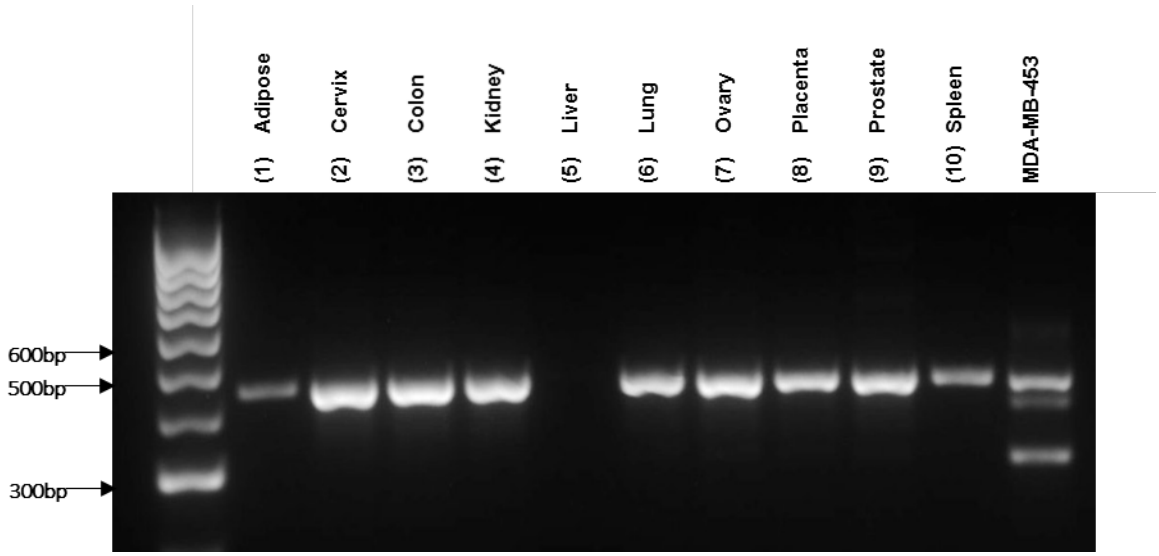


Figure 5.1: RT-PCR amplification of wild type *HER2* and *HER2*ΔECD (primer pair E12F+E15R) in normal human tissue cDNA (1-10), using MDA-MB-453 cell line as a positive control. Hyperladder IV was used as the DNA marker. The expected amplicon sizes for the wild-type *HER2* and *HER2*ΔECD were 432 and 299 base pairs respectively. Each rung of the hyperladder IV represents 100bp.

Amplification of *HER2*ΔATP: Figure 5.2 shows RT-PCR amplification of *HER2*ΔATP in normal tissues. All tissue types (1-4, 6-10) except liver (5), appear to express the wild type *HER2*, but not *HER2*ΔATP. In addition to the top band which represents the wild type *HER2*, Adipose, cervix, colon, kidney, ovary and placenta also show the smaller, lower bands which represent *HER2*ΔATP. Liver tissue (5) does not appear to express either the wild type *HER2* or *HER2*ΔATP.

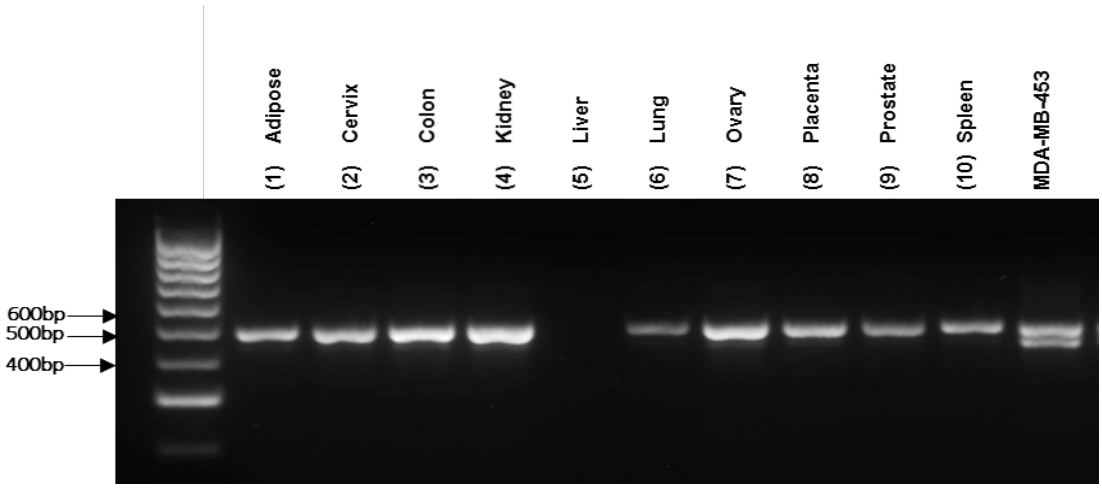


Figure 5.2: RT-PCR amplification of wild type *HER2* and *HER2*ΔATP (primer pair E15F+E19R) in normal human tissue RNA (1-10), using MDA-MB-453 cell line as a positive control. Hyperladder IV was used as the DNA marker. The expected amplicon sizes for the wild-type *HER2* and *HER2*ΔATP were 480 and 438 base pairs respectively. Each rung of the hyperladder IV represents 100bp.

Amplification of *HER2*Δ16: Figure 5.3 shows RT-PCR amplification of *HER2*Δ16 in normal tissues. All tissue types (1-4, 6-10) except liver (5), express the wild type *HER2*, but not *HER2*Δ16. Liver tissue (5) does not appear to express either the wild type *HER2* or *HER2*Δ16.

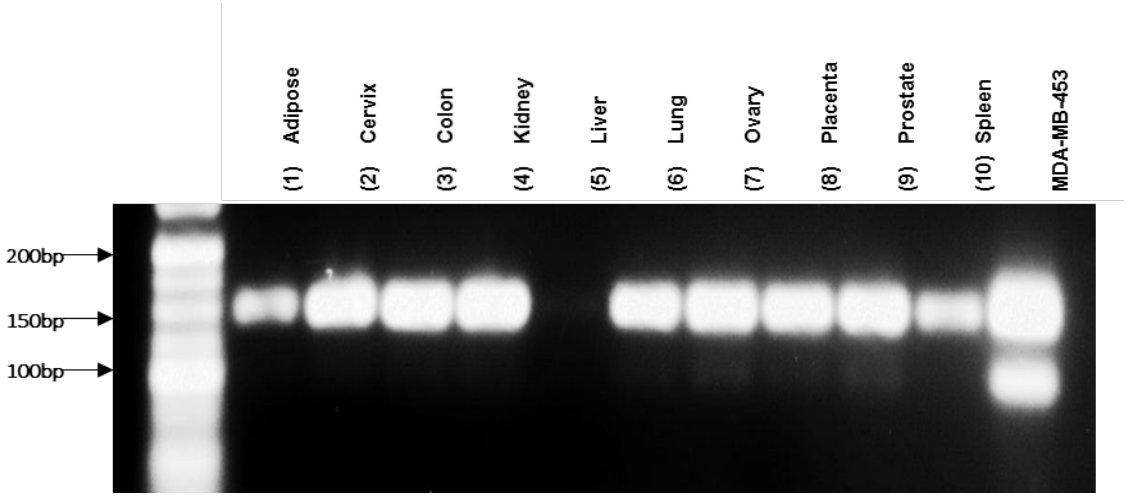


Figure 5.3: RT-PCR amplification of wild type *HER2* and *HER2*Δ16 (primer pair NP5+NP6) in normal human tissue RNA (1-10), using MDA-MB-453 cell line as a positive control. Hyperladder V was used as the DNA marker. The expected amplicon sizes for the wild-type *HER2* and *HER2*Δ16 were 146 and 104 base pairs respectively.

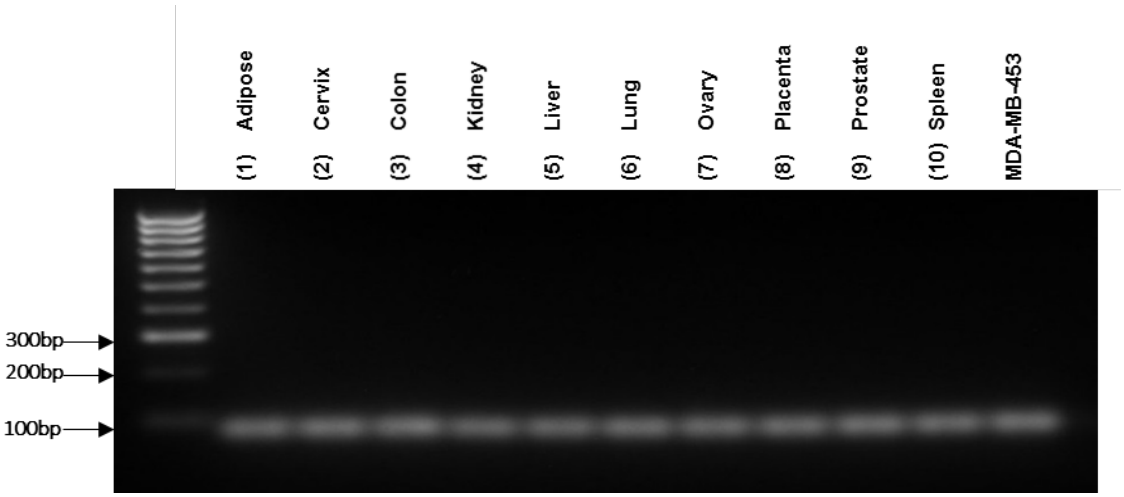


Figure 5.4: RT-PCR amplification of 18s in normal human tissue cDNA and MDA-MB-453 cell line. Amplicon size is approximately 88bp. Each rung of the hyperladder IV represents 100bp.

A negative (no RT) control was also included in all PCR experiments to ensure that PCR amplification was derived from RNA and not genomic DNA or other contaminants. All experiments were run in triplicate to ensure reproducibility.

5.4.2 Expression of *HER2* and *HER2* alternative splice variants in cDNA obtained from frozen clinical samples

Amplification of wild-type *HER2*:

Figure 5.5 shows the relative expression of the wild-type *HER2* in frozen clinical samples. qPCR analysis shows *HER2* expression to be generally higher in the tumours than in the matched normal tissues.

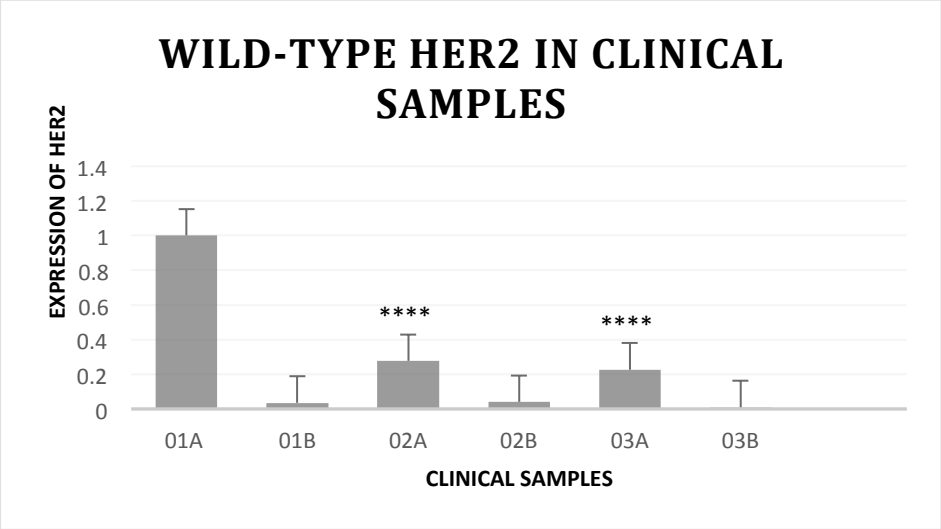


Figure 5.5: qPCR analysis of the expression of wild-type *HER2* cDNA in clinical samples. Each histogram bar is representative of one sample and three replicates (n=3). The x axis represents the individual samples (01A, 02A and 03A), and their matched normal breast tissue (01B, 02B and 03B). The error bars represent the standard deviations of the C_T values.

Amplification of *HER2*ΔECD: Figure 5.6 shows RT-PCR amplification of the wild-type *HER2* and *HER2*ΔECD in cDNA obtained from frozen clinical samples. All three tumours samples express the wild type *HER2*, but not *HER2*ΔECD. Figure 5.7 shows the relative expression of *HER2*ΔECD in the clinical samples by qRT-PCR.

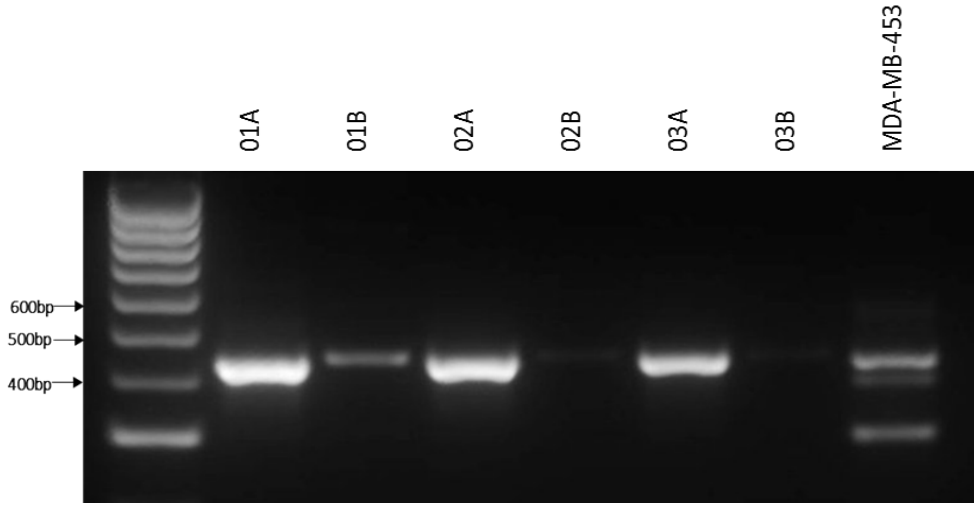


Figure 5.6: RT-PCR amplification of wild type *HER2* and *HER2*ΔECD (primer pair E12F+E15R) in cDNA obtained from frozen tumours, using MDA-MB-453 cell line as a positive control. The samples named 01A, 02A and 03A represent breast tumours, while 01B, 02B and 03B represent the matched normal breast tissue from the same patient, respectively. Hyperladder IV was used as the DNA marker. The expected amplicon sizes for the wild-type *HER2* and *HER2*ΔECD were 432 and 299 base pairs respectively. Each rung of the hyperladder IV represents 100bp.

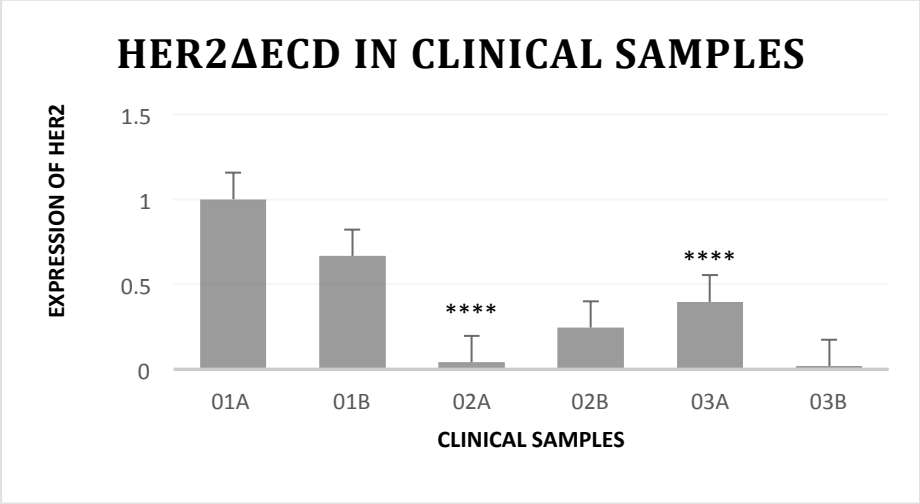


Figure 5.7: qPCR analysis of the expression of *HER2*ΔECD in clinical samples. Each histogram bar is representative of one sample and three replicates (n=3). The x axis represents the individual samples (01A, 02A and 03A), and their matched normal breast tissue (01B, 02B and 03B). The error bars represent the standard deviations of the C_T values.

Amplification of *HER2*ΔATP: Figure 5.8 shows RT-PCR amplification of the wild-type *HER2* and *HER2*ΔATP in cDNA obtained from frozen clinical samples. All three tumours samples express the wild type *HER2*, but not *HER2*ΔATP. Figure 5.9 shows the relative expression of *HER2*ΔATP in the clinical samples by qRT-PCR.

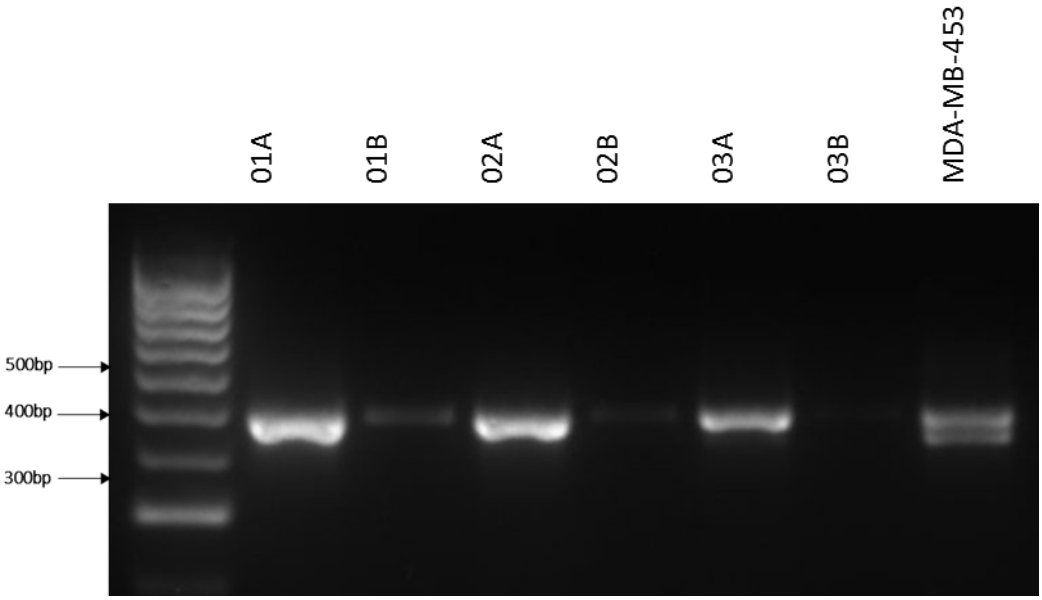


Figure 5.8: RT-PCR amplification of wild type *HER2* and *HER2*ΔATP (primer pair E15F+E19R) in RNA samples obtained from frozen tumours, using MDA-MB-453 cell line as a positive control. The samples named 01A, 02A and 03A represent breast tumours, while 01B, 02B and 03B represent the matched normal breast tissue from the same patient, respectively. Hyperladder IV was used as the DNA marker. The expected amplicon sizes for the wild-type *HER2* and *HER2*ΔATP were 480 and 438 base pairs respectively. Each rung of the hyperladder IV represents 100bp.

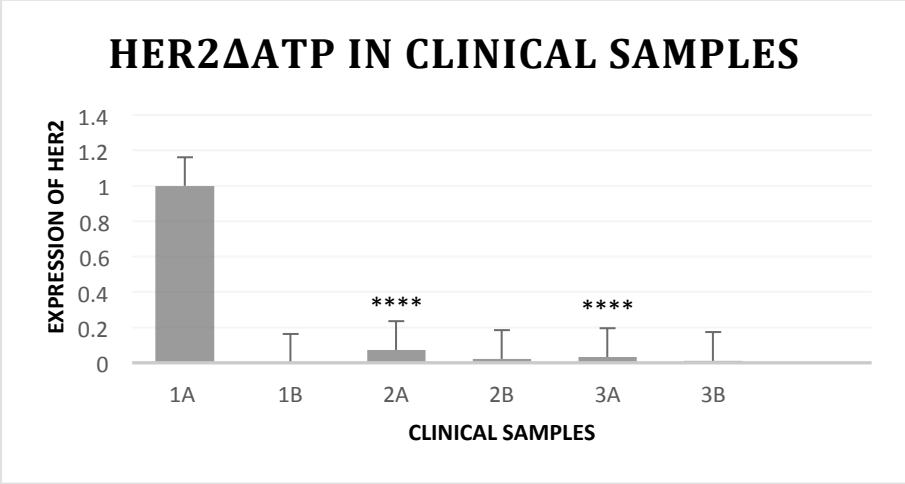


Figure 5.9 qPCR analysis of the expression of *HER2*ΔATP in clinical samples. Each histogram bar is representative of one sample and three replicates (n=3). The x axis represents the individual samples (01A, 02A and 03A), and their matched normal breast tissue (01B, 02B and 03B). The error bars represent the standard deviations of the C_T values.

Amplification of *HER2*Δ16: Figure 5.10 shows RT-PCR amplification of the wild-type *HER2* and *HER2*Δ16 in cDNA obtained from frozen clinical samples. All three tumours samples express the wild type *HER2*, but not *HER2*Δ16. Figure 5.9 shows the relative expression of *HER2*ΔATP in the clinical samples.

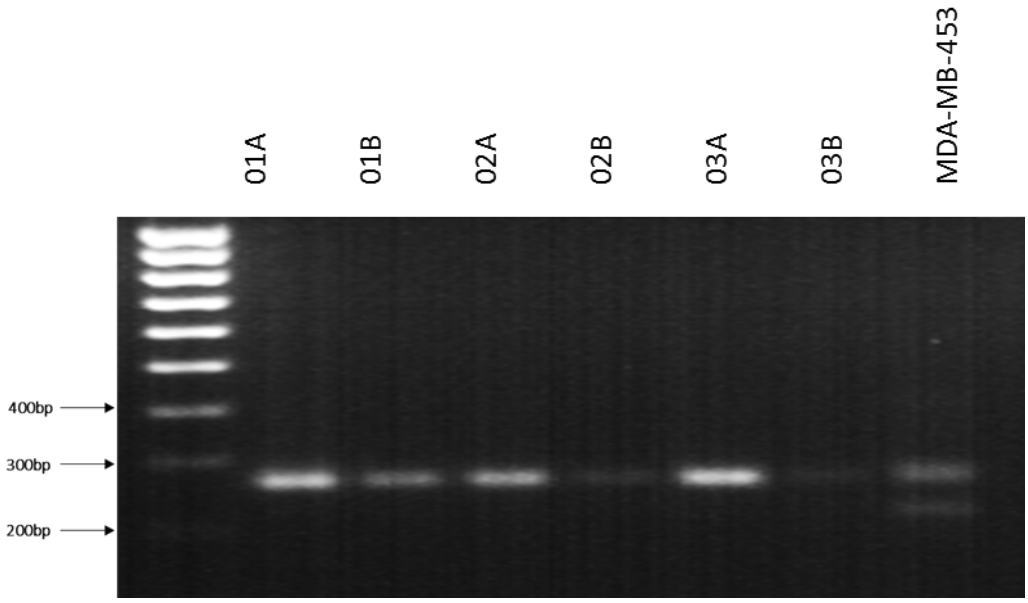


Figure 5.10: RT-PCR amplification of wild type *HER2* and *HER2*Δ16 (primer pair NP1 + NP2) in cDNA samples obtained from frozen tumours, using MDA-MB-453 cell line as a positive control. The samples named 01A, 02A and 03A represent breast tumours, while 01B, 02B and 03B represent the matched normal breast tissue from the same patient, respectively. Hyperladder IV was used as the DNA marker. The expected amplicon sizes for the wild-type *HER2* and *HER2*Δ16 were 266 and 218 base pairs respectively. Each rung of the hyperladder IV represents 100bp.

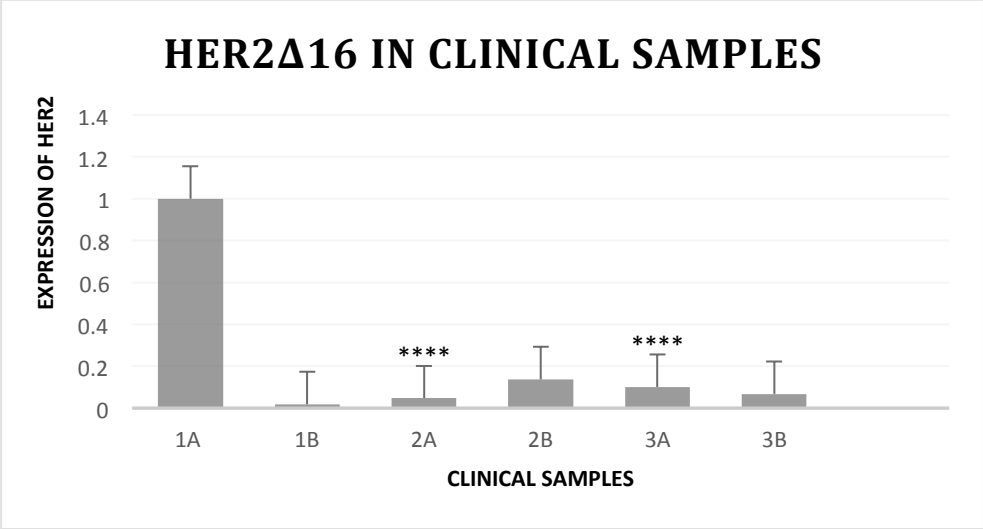


Figure 5.11: qPCR analysis of the expression of *HER2*Δ16 in clinical samples. Each histogram bar is representative of one sample and three replicates (n=3). The x axis represents the individual samples (01A, 02A and 03A), and their matched normal breast tissue (01B, 02B and 03B). The error bars represent the standard deviations of the C_T values.

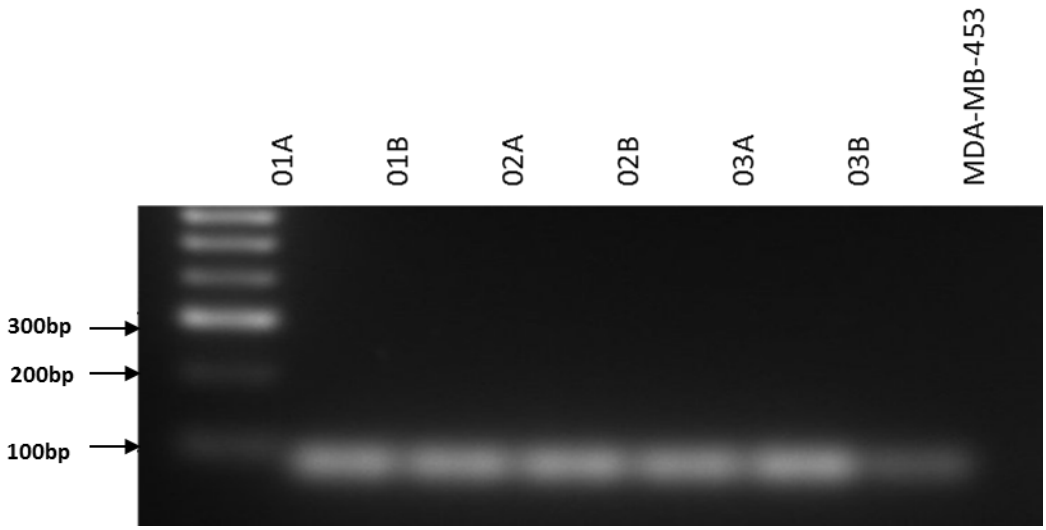


Figure 5.12: RT-PCR amplification of 18s in normal human tissue RNA and MDA-MB-453 cell line. Each rung of the hyperladder IV represents 100bp.

5.4.3 Expression of *HER2* and *HER2* splice variants in formalin fixed and paraffin embedded (FFPE) clinical samples.

SAMPLE	*HER2* STATUS	ng/µl	260/280
11	2+	19.09	1.79
15B	2+	284.52	1.94
21B	3+	86.69	1.93
24B	3+	140.99	1.89
27B	3+	367.41	1.93
28C	3+	359.44	1.95
31A	2+	179.75	1.98
34B	3+	135.91	1.94
35	2+	367.66	1.92
36B	3+	289.44	1.94
43A	2+	535.65	2
53B	3+	385.82	1.91
75A	2+	144.81	1.91
86	2+	111.84	1.93

Table 5.3: *HER2* status, quantification and integrity of RNA obtained from FFPE samples. For anonymity and data protection, samples were designated numeric codes for the purpose of identification.

Amplification of *HER2*ΔECD: Figure 5.13 shows RT-PCR amplification of the wild-type *HER2* and *HER2*ΔECD in cDNA obtained from FFPE clinical samples. When matched against MDA-MB-453 as a positive control, PCR amplification shows wild-type *HER2* expressed in samples 21B, 28C, 35 and 36B, and *HER2*ΔECD expressed in samples 15B, 21b, 28C and 36B.

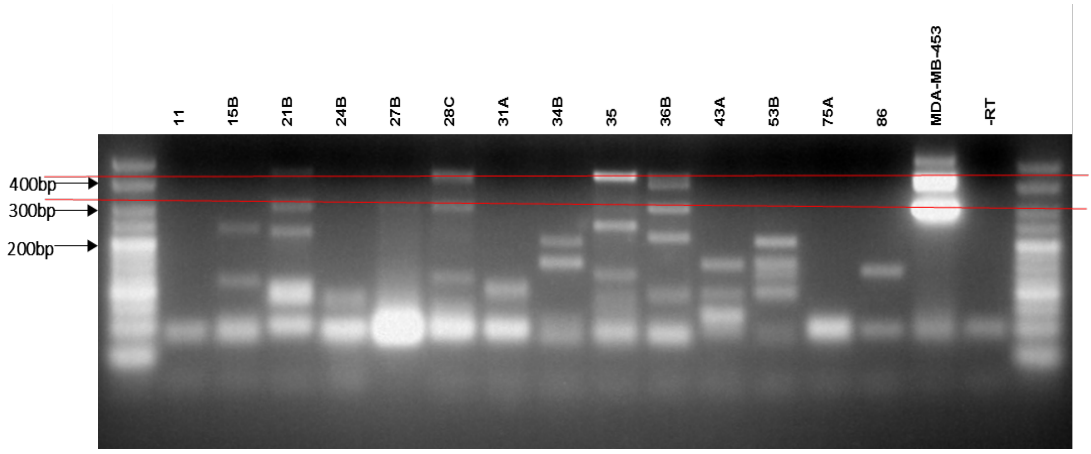


Figure 5.13: RT-PCR amplification of wild type *HER2* and *HER2*ΔECD (primer pair E12F+E15R) in cDNA obtained from FFPE clinical samples, using MDA-MB-453 cell line as a positive control. The numbers above the lanes (11, 15B, 21B, 24B, 27B, 28C, 31A, 34B, 35, 36B, 43A, 53B, 75A, and 86) represent individual patient samples. The red lines are to aid in identification of expected amplicons. Hyperladder V was used as the DNA marker. The expected amplicon sizes for the wild-type *HER2* and *HER2*ΔECD were 432 and 299 base pairs respectively.

Amplification of *HER2*ΔATP: Figure 5.14 shows RT-PCR amplification of the wild-type *HER2* and *HER2*ΔATP in cDNA obtained from FFPE clinical samples. When matched against MDA-MB-453 as a positive control, PCR amplification shows wild-type *HER2* expressed in samples 28C, 34B, 35 and 36B, and *HER2*ΔATP expressed in sample 36B.

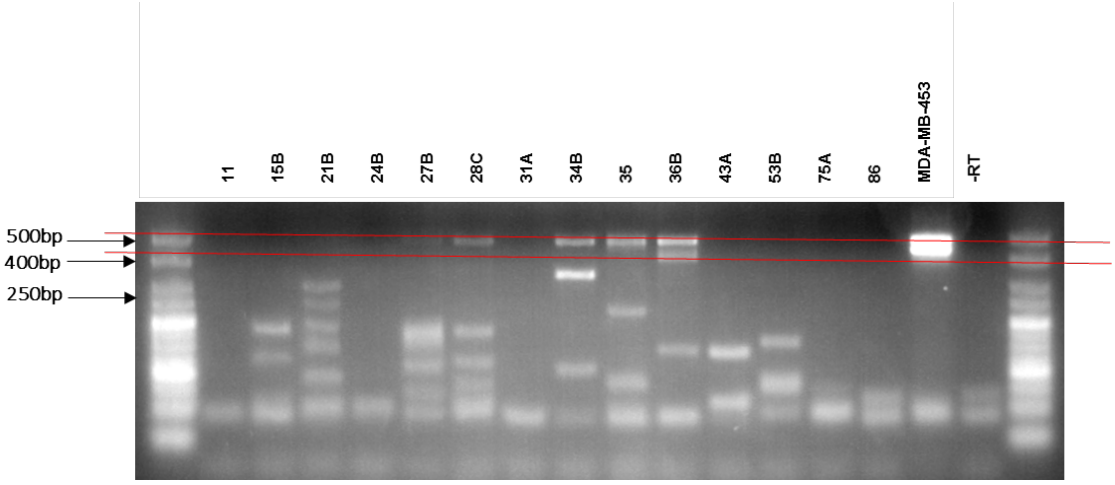


Figure 5.14: RT-PCR amplification of wild type *HER2* and *HER2*ΔATP (primer pair E15F+E19R) in cDNA obtained from FFPE clinical samples, using MDA-MB-453 cell line as a positive control. The numbers above the lanes (11, 15B, 21B, 24B, 27B, 28C, 31A, 34B, 35, 36B, 43A, 53B, 75A, and 86) represent individual patient samples. The red lines are to aid in identification of expected amplicons. Hyperladder V was used as the DNA marker. The expected amplicon sizes for the wild-type *HER2* and *HER2*ΔATP were 480 and 438 base pairs respectively.

Amplification of *HER2*Δ16: Figure 5.15 shows RT-PCR amplification of the wild-type *HER2* and *HER2*Δ16 in cDNA obtained from FFPE clinical samples. When matched against MDA-MB-453 as a positive control, PCR amplification shows wild-type *HER2* expressed in samples 15B, 21B, 27B, 28C, 34B, 35, 36B, 43A, 53B and 75A, and *HER2*Δ16 expressed in samples 15B, 21B, 28C, 34B, 35, 36B, 43A and 53B.

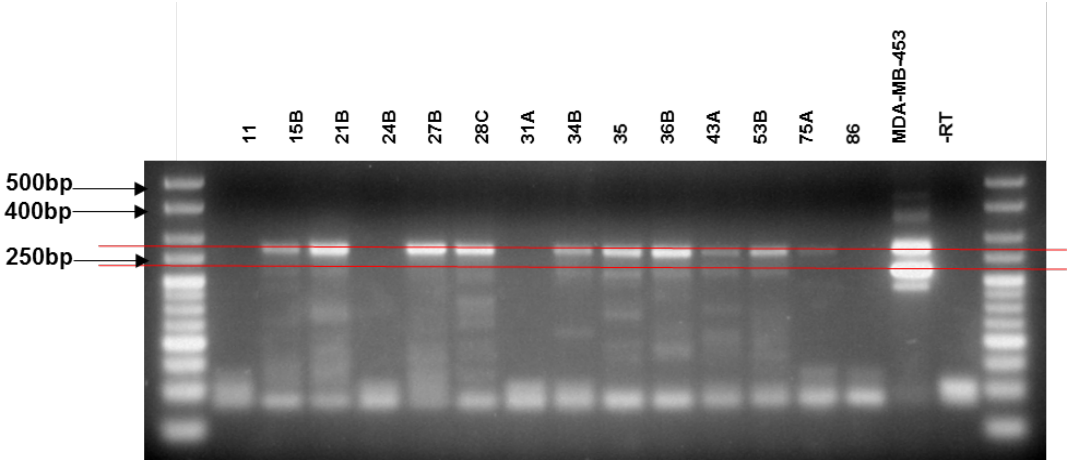


Figure 5.15: RT-PCR amplification of wild type *HER2* and *HER2*Δ16 (primer pair NP1 and NP2) in cDNA obtained from FFPE clinical samples, using MDA-MB-453 cell line as a positive control. The numbers above the lanes (11, 15B, 21B, 24B, 27B, 28C, 31A, 34B, 35, 36B, 43A, 53B, 75A, and 86) represent individual patient samples. The red lines are to aid in identification of expected amplicons. Hyperladder V was used as the DNA marker. The expected amplicon sizes for the wild-type *HER2* and *HER2*Δ16 were 266 and 218 base pairs respectively.

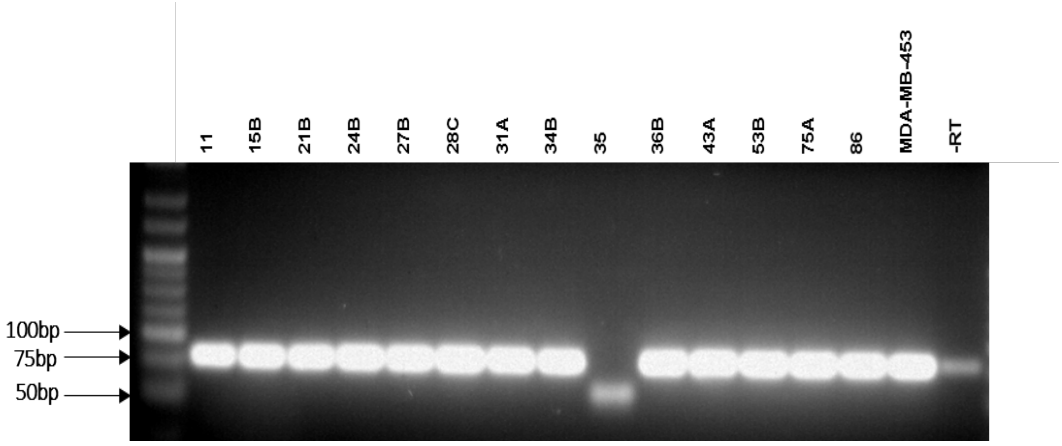


Figure 5.16: RT-PCR amplification of 18s in cDNA obtained from FFPE clinical samples, using MDA-MB-453 cell line as a positive control.

5.5 Summary

The use of cell line models in the investigation of alternative splicing in invasive breast and ovarian cancer cell lines in this study has led to the discovery of novel splice variants in invasive cancer. However, it is also important to correlate these findings with expression in human samples. The discovery of *HER2* alternative splice variants in human tissues particularly *HER2* positive cases of breast cancer, may lead to establishing a potential clinical significance of these new alternative splice variants.

The use of qRT-PCR in this study was in order to give a general overview of the expression of *HER2* and *HER2* alternative splice variants in tissue samples, alongside the standard RT-PCR. Due to the absence of a control sample, it is not conclusive in this study the tissue-specific changes in the expression levels of *HER2* and *HER2* alternative splice variants. It is also important to note that the detection by qRT-PCR of *HER2* expression and tissue-specific changes in *HER2* expression, though they may seem statistically significant, may only be minute changes which can only be verified by the use of a tissue control and absolute quantification of *HER2* expression in the controls to compare to the expression in the tissues tested. To summarise:

- In this study, alternative splice variant *HER2*Δ16 was not detected in normal human tissues (Figure 5.3). This finding is in accordance with current reports (Mitra *et al.*, 2009). The newly identified *HER2*Δ13 splice variant was also not detected in normal human tissues. However, the *HER2*ΔATP splice variant appears to be expressed in normal tissues.

- The newly identified alternative splice variants *HER2*ΔECD and *HER2*ΔATP were also found to be expressed in clinical samples of breast carcinomas which were processed by formalin fixation (Figures 5.13 and 5.14).
- In accordance with findings by Mitra *et al* (2009), this study has identified the expression of the *HER2*Δ16 splice variant in *HER2* positive breast cancer, and for the first time, in FFPE tissues (Figure 5.15).

CHAPTER 6. REGULATION OF *HER2* AND *HER2* SPLICE VARIANTS IN CELL LINE MODELS

6.1. Introduction

The regulation of normal mRNA splicing is dependent on the recognition of intron-exon boundaries, the removal of intervening introns, and the ligation of exons by the spliceosome (Fackenthal & Godley, 2008). In cancer cells, the fidelity of this process may be altered by defects in several splicing mechanisms (Skotheim & Nees, 2007). In some cases of misregulation of splicing, the aberrant mRNAs and their encoded proteins confer unique functions to the expressing cancer cells, and have unique properties that alter the growth and differentiation properties, and other cellular characteristics of the cell (Kim, Goren & Ast, 2008) . The regulation of alternative splicing is still being widely investigated and remains incompletely understood. It is suggested that the disruption of splicing regulatory proteins may play a role in alternative splicing (Kim, Goren & Ast, 2008). Serine and arginine-rich (SR) proteins control the functions of exonic splice enhancers (ESEs), which are purine-rich cis-acting elements that promote splicing of nearby sequences (Fackenthal & Godley, 2008). High levels of phosphorylation of SR proteins are also thought to play a role in inhibiting splicing (Gui *et al.*, 1994). Bioinformatic analysis of binding sites for novel *HER2* splice variants and HERΔ16 in chapter 4 revealed potential splice factor binding motifs for *ASF/SF2* (SRSF2), which is a splice factor in the cytoplasm phosphorylated by the SR protein *SRPK1*. The *SRPK1* protein kinase is also involved in RNA transcript processing, and increased levels of *SRPK1* have been known to play a role in the development of certain cancers such as chronic myelogenous leukemia (CML), colonic and pancreatic

carcinomas (Salesse, Dylla & Verfaillie, 2004; Hayes, Carrigan & Miller, 2007). There is also evidence that *SRPK1* plays an important role in the regulation and of Vascular Endothelial Growth Factor Receptor (VEGF), and is one of the factors responsible for the balance between the pro- and anti- angiogenic isoforms of VEGF (Nowak *et al.*, 2008). In this chapter the SR protein *SRSF1* (*ASF/SF2*) and it phosphorylating protein kinase *SRPK1* were investigated by siRNA knockdown, to determine their involvement in the regulation of *HER2* and *HER2* alternative splice variants using cell line models. Other factors were also investigated, including the role of Hypoxia Inducible Factor 1α (HIF1-α), and the role of Splice factor kinases. HIF1-α is known to function as a tumour suppressor in breast cancer cells (Chiavarina *et al.*, 2010), and can be induced by the use of hypoxia mimetic factors such as Cobalt Chloride (Vengellur & LaPres, 2004). The HIF1-α protein is a transcription factor subunit with intrinsic cellular response to hypoxia. HIF1-α is known to be upregulated by hypoxia, and is known as a gold standard in the detection of hypoxia (Vordermark, Brown & Phil, 2003; Lekas *et al.*, n.d.; Ke & Costa, 2006).

The role of splice factor kinases was investigated by the use of protein kinase inhibitors to the protein kinase inhibitors of interest. The selection protein kinases tested in this study were based on an available repository of protein kinases which have been shown to be related to certain mammalian cancers, or have been shown to regulate certain factors which play a role in cancer. Current group research interests have been focused on investigating the roles of these protein kinase inhibitors in various cancer types including prostate cancer and leukaemia. The kinase inhibitors investigated in this study include *SRPIN340*, *TG003* and *INDY*. *SRPIN340* is a selective inhibitor of *SRPK1*.

TG003 is a CDC2-like kinase inhibitor. The CDC2-like kinase is a member of an evolutionary conserved family of dual-specificity kinases belonging to the Cyclin-dependent (CDK), Mitogen-activated (MAPK), Glycogen synthase (GSK) and CDK-like kinases (CMGC) (Jain *et al.*, 2014; Rodgers *et al.*, 2010). CDC2-like kinases have been shown to alter the regulation of SR proteins both *in vitro* and *in vivo* (Rodgers *et al.*, 2010). Though a high level regulator of alternative splicing via phosphorylation of SR domains on splice factors, the connection between CDC2-like kinases and breast cancer has not been previously studied.

INDY is an inhibitor of the Dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 1A (DYRK1A), a protein kinase that is a member of the highly conserved dual-specificity tyrosine phosphorylation regulated kinase (DYRK) family (Courcet *et al.*, 2012). DYRK1A modulates alternative splicing by phosphorylating splice factor SRSF6. It is also known to phosphorylate serine, threonine and tyrosine residues in its sequence. In addition, DYRK1A participates in multiple biological pathways, including the phosphorylation of ASF and the splicing factor SF3b1/SAP155 (Courcet *et al.*, 2012).

6.2. Methods

6.2.1 Treatment of cells with protein kinase inhibitors

To study the effect of protein kinases on *HER2* and *HER2* alternative splice variants, MDA-MB-453, SKBR3 and BT-20 cells were used to investigate the effects of protein kinase inhibitors *SRPIN340*, *TG003* and *INDY*. Stock solutions were prepared by diluting in DMSO. 1µl of DMSO was also added to each well of the untreated cells to ensure that the changes in the cells after treatments are not due to DMSO in treated cells.

Protein kinase inhibitors were added to cells according to the manufacturer's guidelines at a final concentration of 10µM, in complete cell culture media, and incubated for 24 and 48 hours, prior to RNA extraction, reverse transcription, and qPCR analysis. In addition to untreated cells, two negative controls were also added to the treatment; *SRPIN349* was used as a negative control for *SRPIN340*, and *TG009* was used as a negative control for *TG003*.

6.2.2 Treatment of cells with hypoxia mimetic factor Cobalt Chloride ($CoCl_2$)

Cells were treated with $CoCl_2$ to investigate the effects of hypoxia stimulated by $CoCl_2$ in breast cancer cell line MDA-MB-453 and SKBR3. A 100mM stock solution was prepared by dissolving $CoCl_2$ powder in sterile distilled water. $CoCl_2$ was added to complete cell culture media at final concentrations of 100µM, 200µM, 300µM, 400µM and 500µM. Cells were incubated for 24 hours under normal cell culture conditions before RNA extraction, reverse transcription and qPCR analysis. HIF1-α accumulation in cells following $CoCl_2$ treatment have been seen to increase within only 6-12 hours in western blot experiments (Ke & Costa, 2006; Al Okail, 2010). Moreover, treatment with $CoCl_2$ for longer than 24 hours in this study resulted in a high level of cell death which may be due to the high toxicity of $CoCl_2$ to the cells (Al Okail, 2010). The 24 hour time point was chosen for all $CoCl_2$ treatments. However the effects of different time points have not been carried out within the scope of this project.

6.2.3 siRNA silencing of *SRPK1* and *SRSF1* in MDA-MB-453 and SKBR3 cell lines

MDA-MB-453 and SKBR3 cell lines were used to investigate the effects of siRNA knockdown of *SRPK1* and *SRSF1* on the expression of *HER2* and the *HER2* alternative splice variants of interest. MDA-MB-453 cell line was used for *SRPK1* and *SRSF1* knockdown because it showed reliable expression of *HER2* and *HER2* splice variants and is a reliable source of DNA material for sequencing. SKBR3 cell line was used for *SRPK1* and *SRSF1* knockdowns because it is *HER2* 3+ by immunohistochemistry, and represents the highest clinical expression of *HER2* based on tests used in *HER2* diagnosis.

Cells were seeded onto 6-well tissue culture plates at a density of $0.6x10^6$ the day before transfection, and allowed to adhere overnight in their respective media + 10% Foetal Bovine Serum (FBS) + 2mM L-Glutamine, without antibiotics. On the day of transfection, cells were serum-starved for two hours in Opti-MEM I reduced-serum medium (Life Technologies, UK). Transfection was carried out using Dharmafect transfection reagent (Dharmacon, UK) at a volume of 6µl per well. Smartpool siGENOME siRNAs targeted against *SRPK1* and *SRSF1* were used at a final concentration of 100nM. Smartpool siGENOME siRNA consists of a set of 4 siRNAs provided as a single reagent, providing the advantaged of high specificity and potency. Total RNA was extracted from the cells 24 and 48 hours post transfection, followed by reverse transcription and qRT-PCR.

6.2.4 Western blot analysis

To confirm changes in protein expression following transfection, siRNA knockdowns of *SRPK1* and *SRSF1* in MDA-MB-453 cells were studied using western blot analysis. MDA-MB-453 cells were used for this assay due to its reliability in expressing *HER2* and *HER2* alternative splice variants earlier in this study. Cells were transfected for 24 and 48 hours prior to protein extraction. *SRPK1* mouse monoclonal antibody clone EE-13:sc100443 and *ASF/SF2* (*SRSF1*) mouse monoclonal antibody clone 96:sc33652 (Santa Cruz Biotechnology) were used as primary antibodies to *SRPK1* and *SRSF1*, respectively. β-actin was used as a loading control.

Antibody Clone	**Specificity**	**Source**	**Concentration**
EE-13:sc100443	*SRPK1*	Santa Cruz Biotechnology	1µg/mL
96:sc33652	*SRSF1*	Santa Cruz Biotechnology	1µg/mL
sc-47778	β-actin	Santa Cruz Biotechnology	0.1µg/mL

Table 6.1: Antibodies used in Western blotting and their specificities

6.2.5 Real-time qPCR analysis of *HER2*, *HER2* alternative splice variants, *SRPK1* and *SRSF1*.

Real-time qPCR analysis was performed following treatment of cells with protein kinase inhibitors and Cobalt Chloride, and following transfection of cells with *SRPK1* and *SRSF1* siRNAs. Double-dye probes were used for the detection of *HER2* and *HER2* alternative splice variant expression (Table 5.1). Primers for the detection of *SRPK1*, *SRSF1* and *HIF1-α* by sybr green chemistry were designed by primer design UK (Table 6.2). Real-time qPCR assays were normalised against the most stable of 12 reference genes, using the *Normfinder* algorithm.

Primer name	**5'-3' sequence**	**Product length**
SRSF1 sense	GATGGAATTGTGTTTTGCGTTTT	101bp
SRSF1 antisense	CATCTACTCGTGCTGAATCCTT	
SRPK1 sense	ACAAGCAAGAAGAATCAGAGAGT	124bp
SRPK1 antisense	CGTTCCATAAGCGTTTGATCC	
HIF1-α sense	TGCCACATCATCACCATATAGAG	132bp
HIF1-α antisense	TGACTCAAAGCGACAGATAACA	

Table 6.2: *SRPK1, SRSF1* and *HIF1-α* primer sequences.

6.3. Results

6.3.1 Inhibition of *SRPK1* by *SRPIN340* modulates the expression of *HER2* and *HER2* alternative splice variants in MDA-MB-453 cell line

Gene name	Stability value		
18S	0.072	**Best gene**	GAPDH
ACTB	0.079	**Stability value**	0.035
ATP5B	0.113		
B2M	0.138		
CYC1	0.119		
E1F4A2	0.155		
GAPDH	0.035		
RPLI3A	0.079		
SDHA	0.171		
TOP1	0.073		
UBC	0.119		
YWHAZ	0.153		

Table 6.3: *Normfinder* output for the selection of an optimal reference gene in MDA-MB-453 cells treated with protein kinase inhibitors *SRPIN340*, *TG003* and *INDY*.

6.3.1.1. MDA-MB-453 cell line

The untreated samples were used as a calibrator to measure the fold change at 24 and 48 hours post-treatment, using *GAPDH* as a normalisation factor (Table 6.3).

6.3.1.1.1. Changes in the expression of wild-type *HER2* following treatment with protein kinase inhibitors

There was no significant change in the expression wild-type *HER2* after treatment for 24 hours with *SRPIN340* (fold change = 0.94; $p \geq 0.05$), *TG003* (fold change = 1.07; $p \geq 0.05$) and *INDY* (fold change = 1.07; $p \geq 0.05$). Negative controls *SRPIN340* and *TG009* also showed no change in expression after 24 hours (fold change = 0.97; $p \geq 0.05$; and 1.12; $p \geq 0.05$, respectively) (Figure 6.1). After 48 hours, a significant increase was observed in the expression of the wild-type *HER2* following *SRPIN340* treatment (fold change = 0.26; $p < 0.05$). No significant change was observed in the wild-type *HER2* after treatment with *TG003* and *INDY* (fold change = 0.93; $p \geq 0.05$; and 1.28; $p \geq 0.05$, respectively). Negative controls *SRPIN349* and *TG009* also showed no significant change in the expression of wild-type *HER2* following treatment for 48 hours of MDA-MB-453 cells (fold change = 1.22; $p \geq 0.05$; and 0.70; $p \geq 0.05$, respectively) (Figure 6.2).

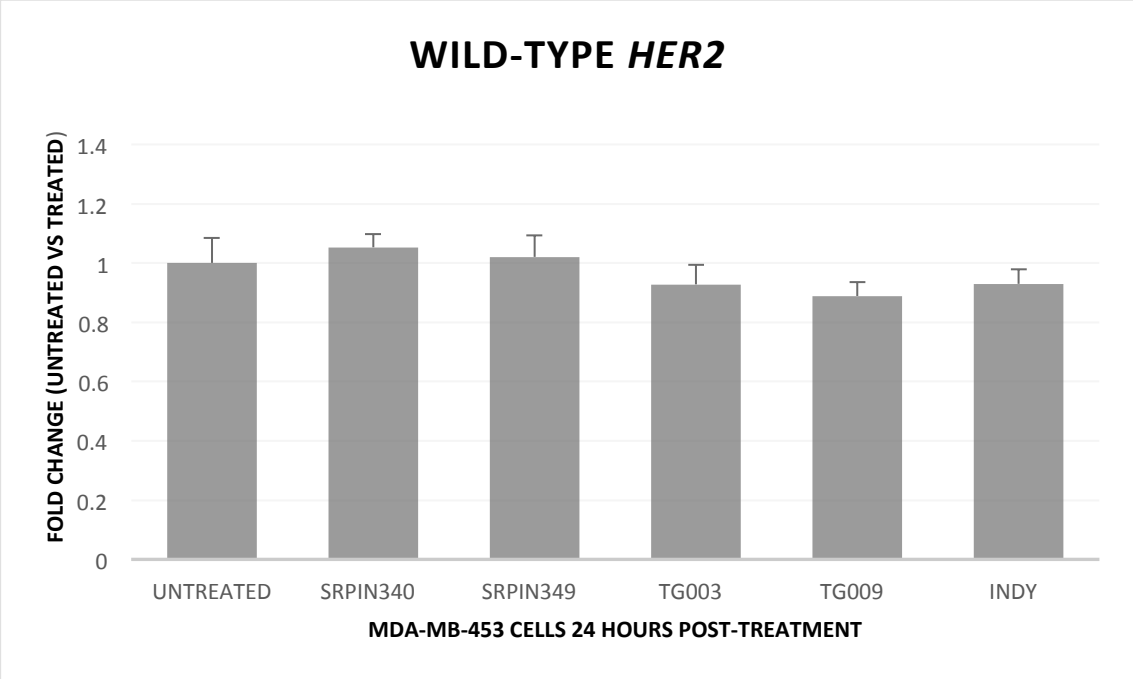


Figure 6.1: Effect of protein kinase inhibitors *SRPIN340*, *TG003* and *INDY* on the wild-type *HER2* in MDA-MB-453 cells 24 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

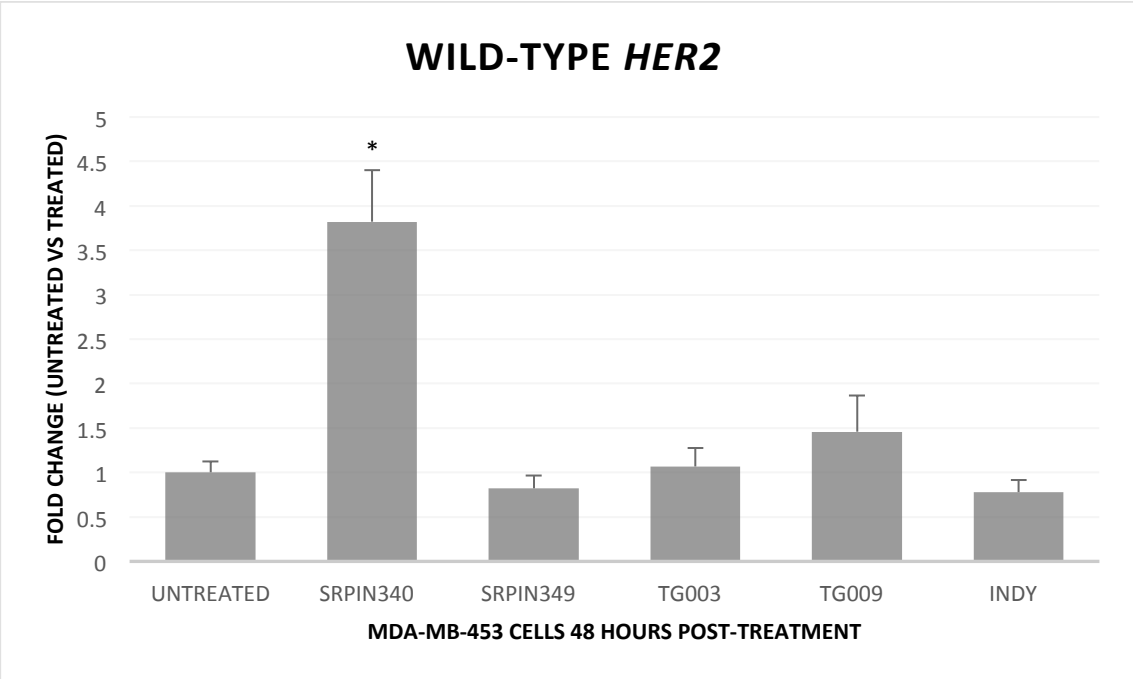


Figure 6.2: Effect of protein kinase inhibitors *SRPIN340*, *TG003* and *INDY* on the wild-type *HER2* in MDA-MB-453 cells 24 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

6.3.1.1.2. Changes in the expression of *HER2ΔECD* following treatment with protein kinase inhibitors

There was no significant change in the expression *HER2ΔECD* after treatment for 24 hours with *SRPIN340* (fold change = 1.0; $p \geq 0.05$), *TG003* (fold change = 1.0; $p \geq 0.05$) and *INDY* (fold change = 1.21; $p \geq 0.05$). Negative controls *SRPIN340* and *TG009* also showed no change in expression after 24 hours (fold change = 0.92; $p \geq 0.05$; and 1.14; $p \geq 0.05$, respectively) (Figure 6.3). After 48 hours, a significant increase was observed in the expression of the *HER2ΔECD* following *SRPIN340* treatment (fold change = 0.13; $p < 0.05$). A slight reduction in the expression of *HER2ΔECD* was observed after treatment with *TG003* and *INDY* (fold change = 0.4; $p \geq 0.05$; and 0.42; $p \geq 0.05$, respectively). These changes, however, are not statistically significant. Negative controls *SRPIN349* and *TG009* also showed no significant change in the expression of wild-type *HER2* following treatment for 48 hours of MDA-MB-453 cells (fold change = 1.06; $p \geq 0.05$; and 0.7; $p \geq 0.05$, respectively) (Figure 6.4).

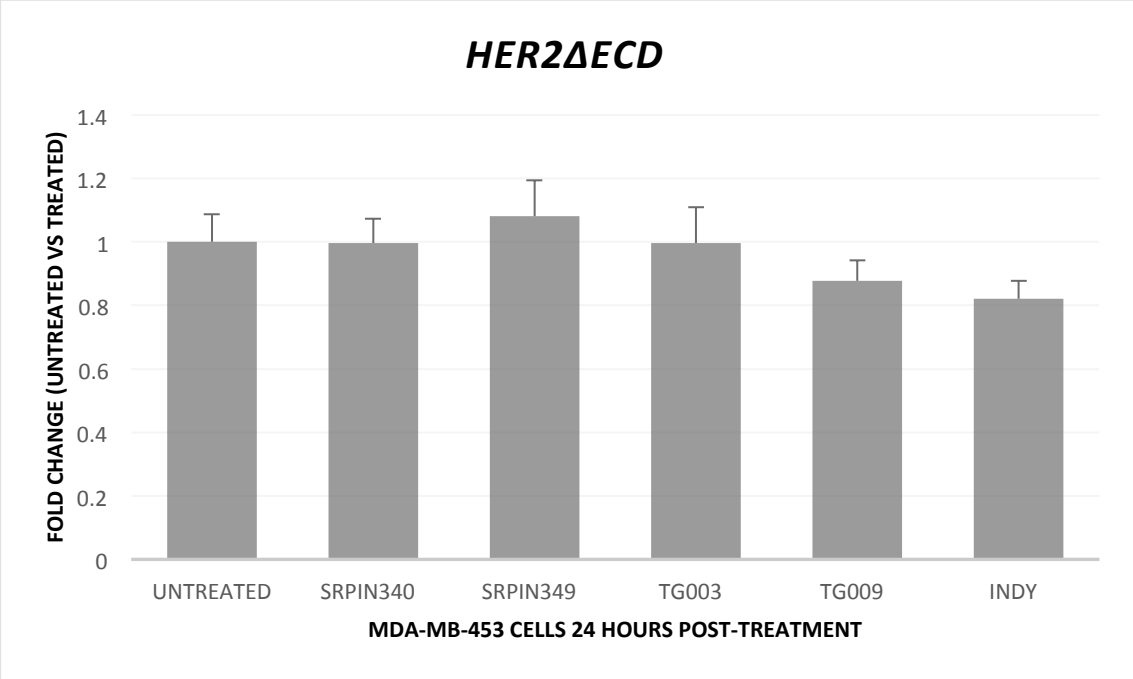


Figure 6.3: Effect of protein kinase inhibitors *SRPIN340*, *TG003* and *INDY* on the *HER2ΔECD* alternative splice variant in MDA-MB-453 cells 24 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

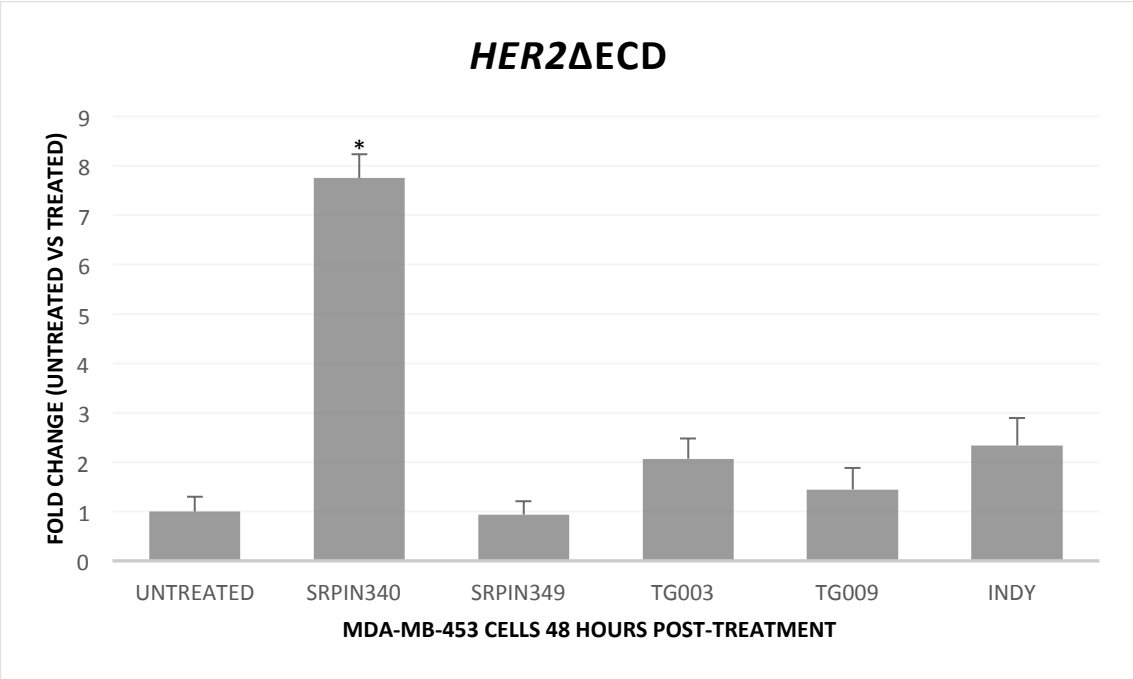


Figure 6.4: Effect of protein kinase inhibitors *SRPIN340*, *TG003* and *INDY* on the *HER2ΔECD* alternative splice variant in MDA-MB-453 cells 48 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

6.3.1.1.3. Changes in the expression of *HER2*Δ16 following treatment with protein kinase inhibitors

There was no significant change in the expression *HER2*Δ16 after treatment for 24 hours with *SRPIN340* (fold change = 1.12; $p \geq 0.05$) and *TG003* (fold change = 1.1; $p \geq 0.05$). A slight reduction was observed following treatment with *INDY* (fold change = 1.53; $p \geq 0.05$). Negative controls *SRPIN340* and *TG009* also showed no change in expression after 24 hours (fold change = 0.98; $p \geq 0.05$; and 1.23; $p \geq 0.05$, respectively) (Figure 6.5). After 48 hours, a significant increase was observed in the expression of the *HER2*Δ16 following *SRPIN340* treatment (fold change = 0.13; $p < 0.01$). A slight reduction in the expression of *HER2*Δ16 was observed after treatment with *TG003* and *INDY* (fold change = 0.5; $p \geq 0.05$; and 0.42; $p \geq 0.05$, respectively). These changes, however, are not statistically significant. Negative controls *SRPIN349* and *TG009* also showed no significant change in the expression of wild-type *HER2* following treatment for 48 hours of MDA-MB-453 cells (fold change = 1.0; $p \geq 0.05$; and 0.9; $p \geq 0.05$, respectively) (Figure 6.6).

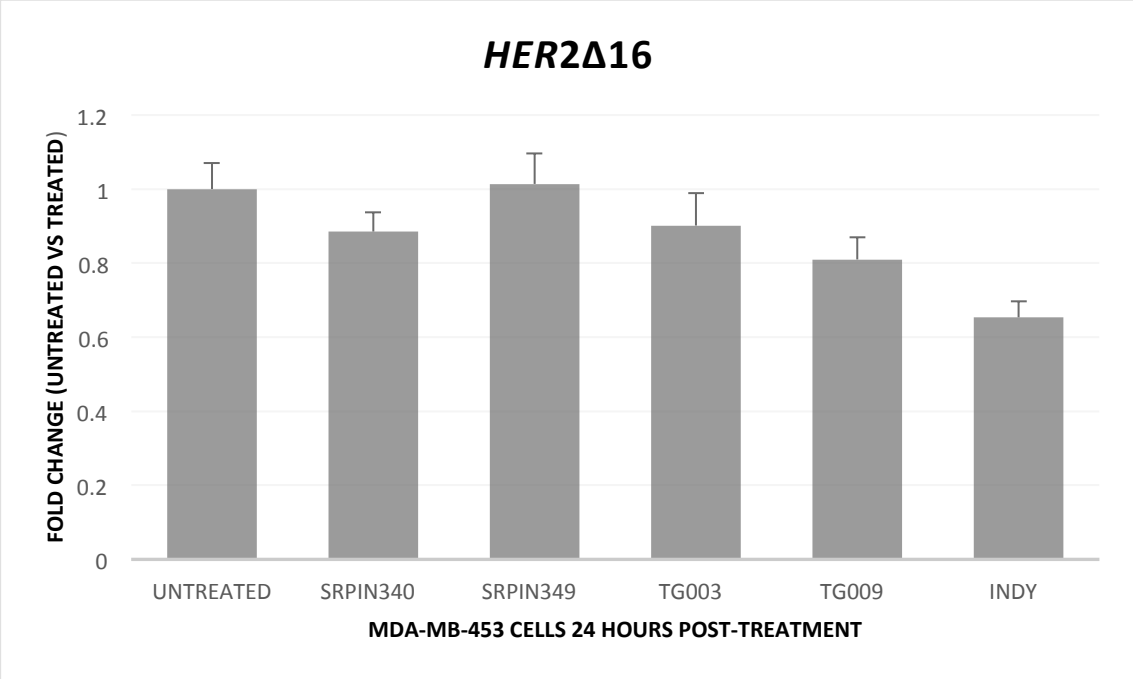


Figure 6.5: Effect of protein kinase inhibitors *SRPIN340*, *TG003* and *INDY* on the *HER2*Δ16 alternative splice variant in MDA-MB-453 cells 24 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

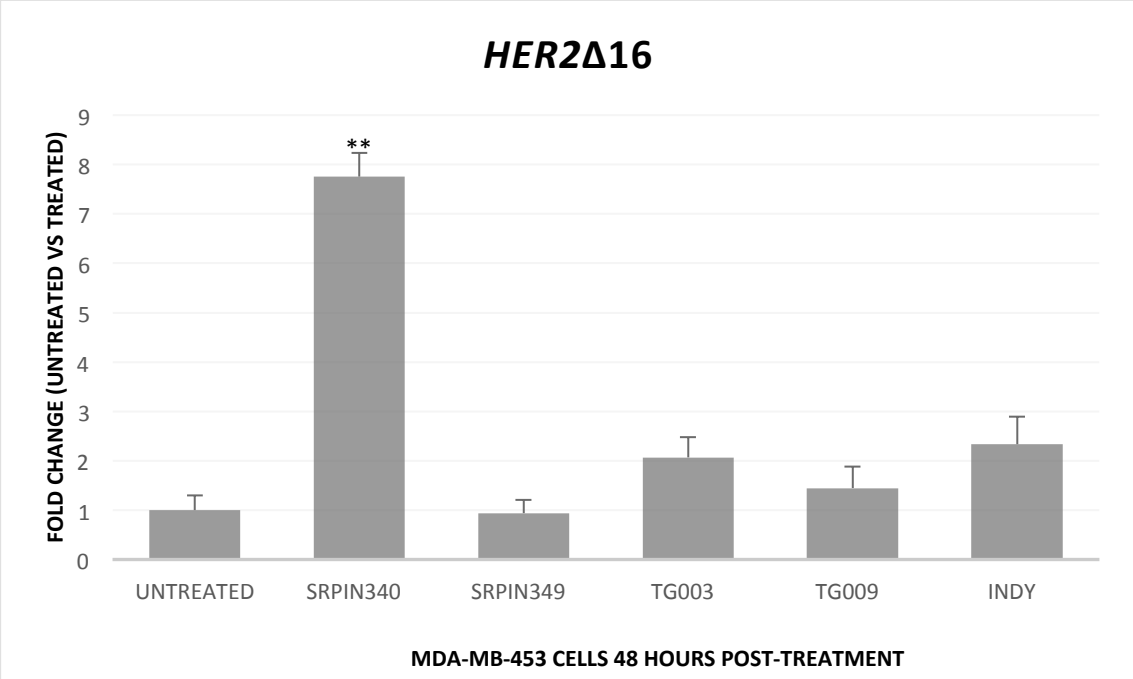


Figure 6.6: Effect of protein kinase inhibitors *SRPIN340*, *TG003* and *INDY* on the *HER2*Δ16 alternative splice variant in MDA-MB-453 cells 48 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

6.3.1.1.4. Changes in the expression of *HER2*ΔATP following treatment with protein kinase inhibitors

There was no significant change in the expression *HER2*ΔATP after treatment for 24 hours with *SRPIN340* (fold change = 1.0; $p \geq 0.05$) and *TG003* (fold change = 0.96; $p \geq 0.05$). A slight increase was observed following treatment with *INDY* (fold change = 1.31; $p \geq 0.05$). Negative controls *SRPIN340* and *TG009* also showed no change in expression after 24 hours (fold change = 0.96; $p \geq 0.05$; and 1.16; $p \geq 0.05$, respectively) (Figure 6.7). After 48 hours, a significant increase was observed in the expression of the *HER2*ΔATP following *SRPIN340* treatment (fold change = 0.13; $p < 0.01$). A slight increase in the expression of *HER2*ΔATP was observed after treatment with *TG003* and *INDY* (fold change = 0.4; $p \geq 0.05$; and 0.42; $p \geq 0.05$, respectively). These changes, however, are not statistically significant. Negative controls *SRPIN349* and *TG009* also showed no significant change in the expression of wild-type *HER2* following treatment for 48 hours of MDA-MB-453 cells (fold change = 1.0; $p \geq 0.05$; and 0.9; $p \geq 0.05$, respectively) (Figure 6.8).

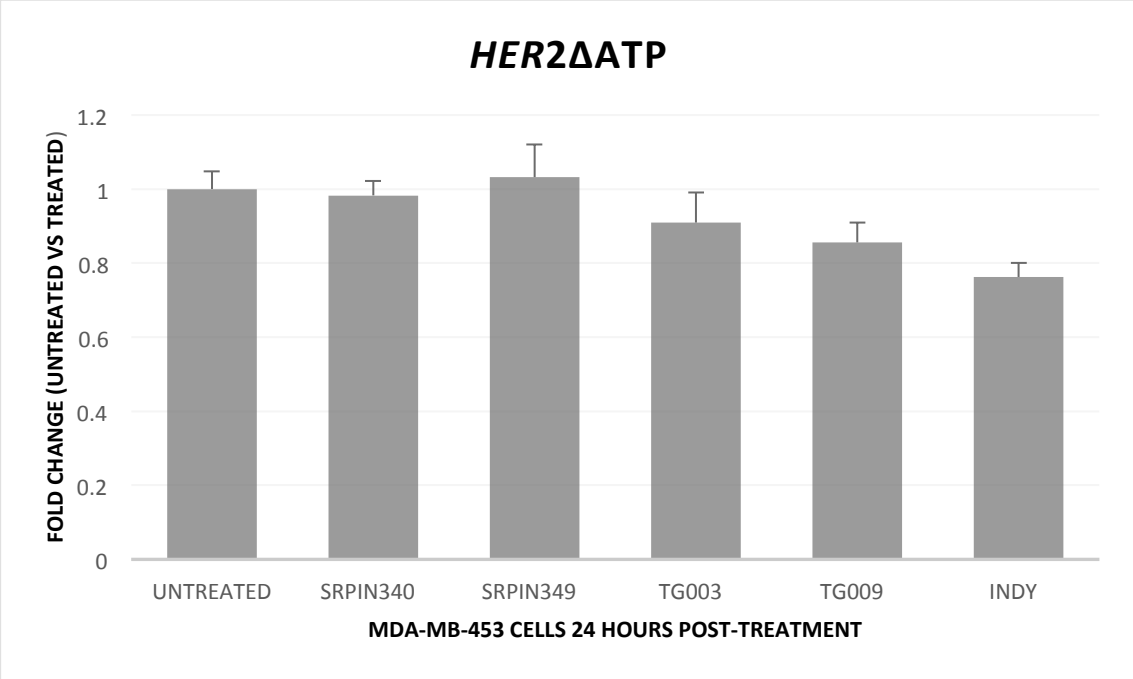


Figure 6.7: Effect of protein kinase inhibitors *SRPIN340*, *TG003* and *INDY* on the *HER2*ΔATP alternative splice variant in MDA-MB-453 cells 24 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

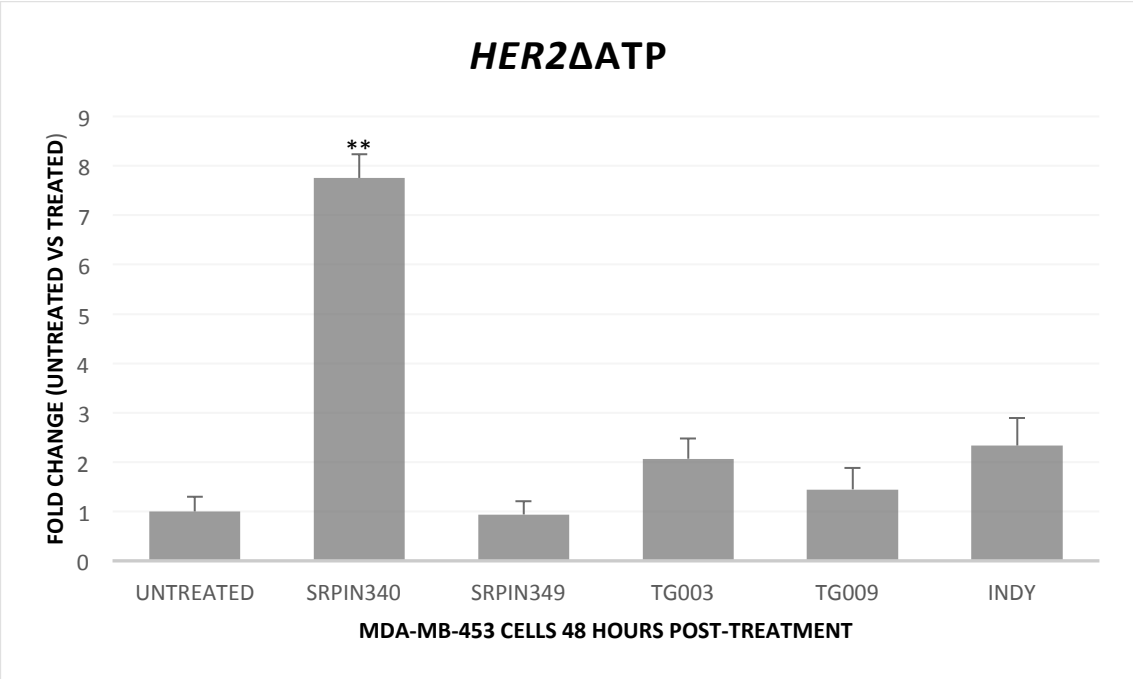


Figure 6.8: Effect of protein kinase inhibitors *SRPIN340*, *TG003* and *INDY* on the *HER2ΔECD* alternative splice variant in MDA-MB-453 cells 48 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

6.3.1.2 SKBR3 cell line

The untreated cells were used as a calibrator to measure the fold change at 24 and 48 hours post-treatment, using *ATP5B* as a normalisation factor (Table 6.4).

Gene name	**Stability value**		
18S	0.099	**Best gene**	ATP5B
ACTB	0.110	**Stability value**	0.067
ATP5B	0.067		
B2M	0.080		
CYC1	0.098		
E1F4A2	0.119		
GAPDH	0.102		
RPLI3A	0.151		
SDHA	0.178		
TOP1	0.085		
UBC	0.170		
YWHAZ	0.080		

Table 6.4: Normfinder output for the selection of an optimal reference gene in SKBR3 cells treated with protein kinase inhibitors *SRPIN340*, *TG003* and *INDY*.

6.3.1.2.1 The expression of wild-type *HER2* following treatment with protein kinase inhibitors

There was no significant change in the expression wild-type *HER2* after treatment for 24 hours with *SRPIN340* (fold change = 0.94; p≥0.05), *TG003* (fold change = 1.09; p≥0.05) and *INDY* (fold change = 0.87; p≥0.05). Negative controls *SRPIN340* and *TG009* also showed no change in expression after 24 hours (fold change = 0.9; p≥0.05; and 0.91; p≥0.05, respectively) (Figure 6.9). After 48 hours, there was also no significant change in the expression of the wild-type *HER2* following treatment with *SRPIN340* (fold change = 0.8; p≥0.05), *TG003* (fold change = 1.1; p≥0.05) and *INDY* (fold change = 1.0; p≥0.05). Negative controls *SRPIN349* and *TG009* also showed no significant change in the expression of wild-type *HER2* following treatment for 48 hours of MDA-MB-453 cells (fold change = 0.8; p≥0.05; and 0.9; p≥0.05, respectively) (Figure 6.10).

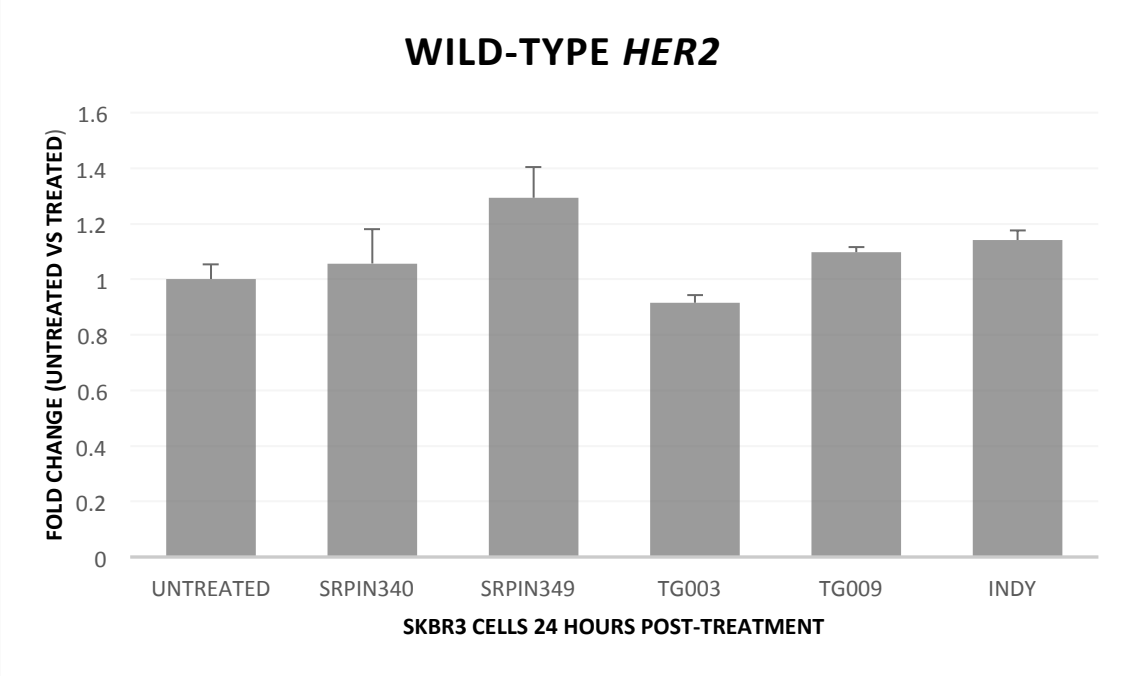


Figure 6.9: Effect of protein kinase inhibitors *SRPIN340*, *TG003* and *INDY* on the wild-type *HER2* in SKBR3 cells 24 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

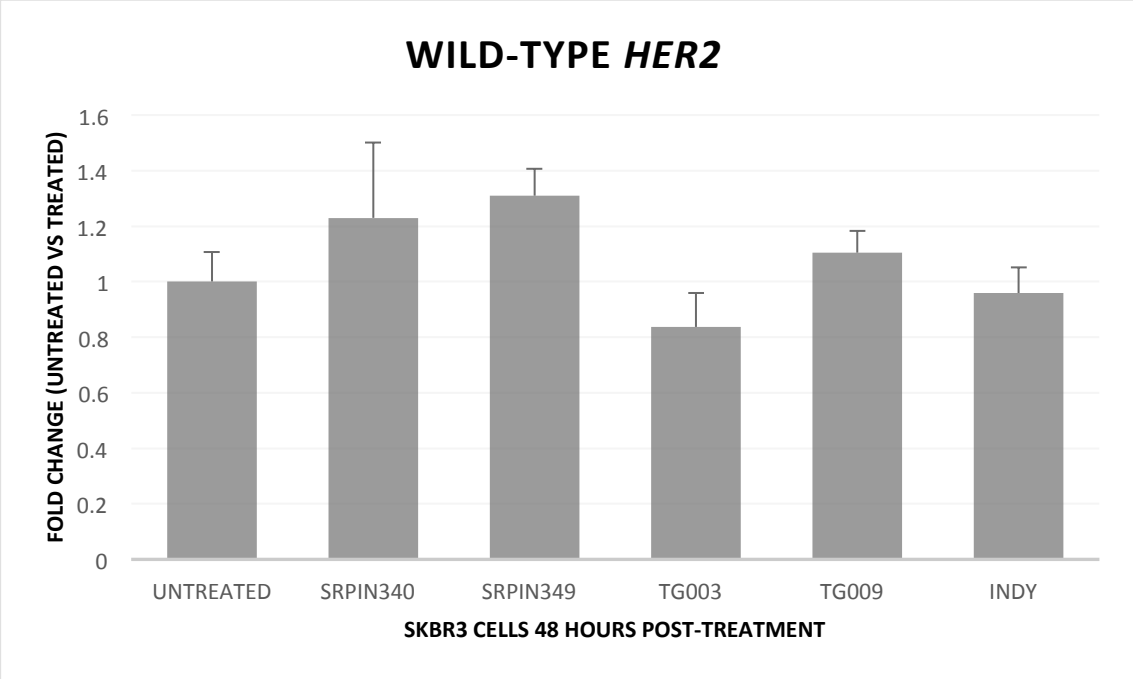


Figure 6.10: Effect of protein kinase inhibitors *SRPIN340*, *TG003* and *INDY* on the wild-type *HER2* in SKBR3 cells 48 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

6.3.1.2.2 The expression of *HER2ΔECD* following treatment with protein kinase inhibitors

There was no significant change in the expression *HER2ΔECD* after treatment for 24 hours with *SRPIN340* (fold change = 1.0; p≥0.05), *TG003* (fold change = 1.2; p≥0.05) and *INDY* (fold change = 1.2; p≥0.05). Negative controls *SRPIN340* and *TG009* also showed no change in expression after 24 hours (fold change = 1.0; p≥0.05; and 1.2; p≥0.05, respectively) (Figure 6.11). After 48 hours, there was also no significant change in the expression of *HER2ΔECD* following treatment with *SRPIN340* (fold change = 0.75; p≥0.05), *TG003* (fold change = 0.75; p≥0.05) and *INDY* (fold change = 1.0; p≥0.05). Negative controls *SRPIN349* and *TG009* also showed no significant change in the

expression of *HER2ΔECD* following treatment for 48 hours of MDA-MB-453 cells (fold change = 0.8; p≥0.05; and 1.1; p≥0.05, respectively) (Figure 6.12).

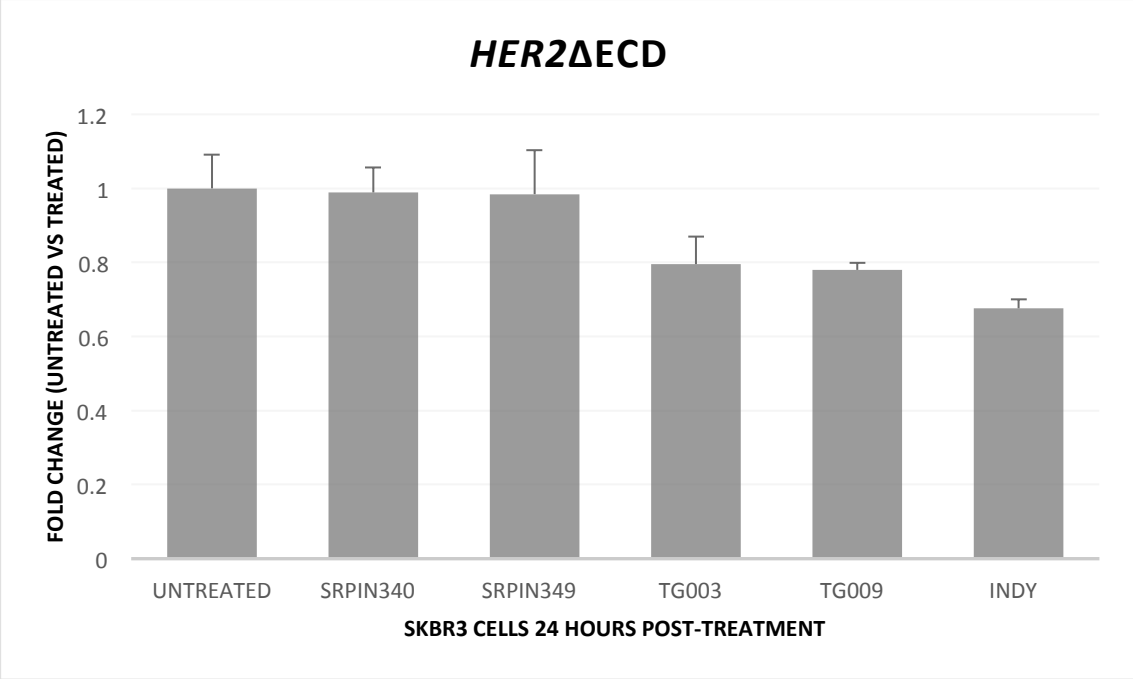


Figure 6.11: Effect of protein kinase inhibitors *SRPIN340*, *TG003* and *INDY* on the *HER2ΔECD* alternative splice variant in SKBR3 cells 24 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

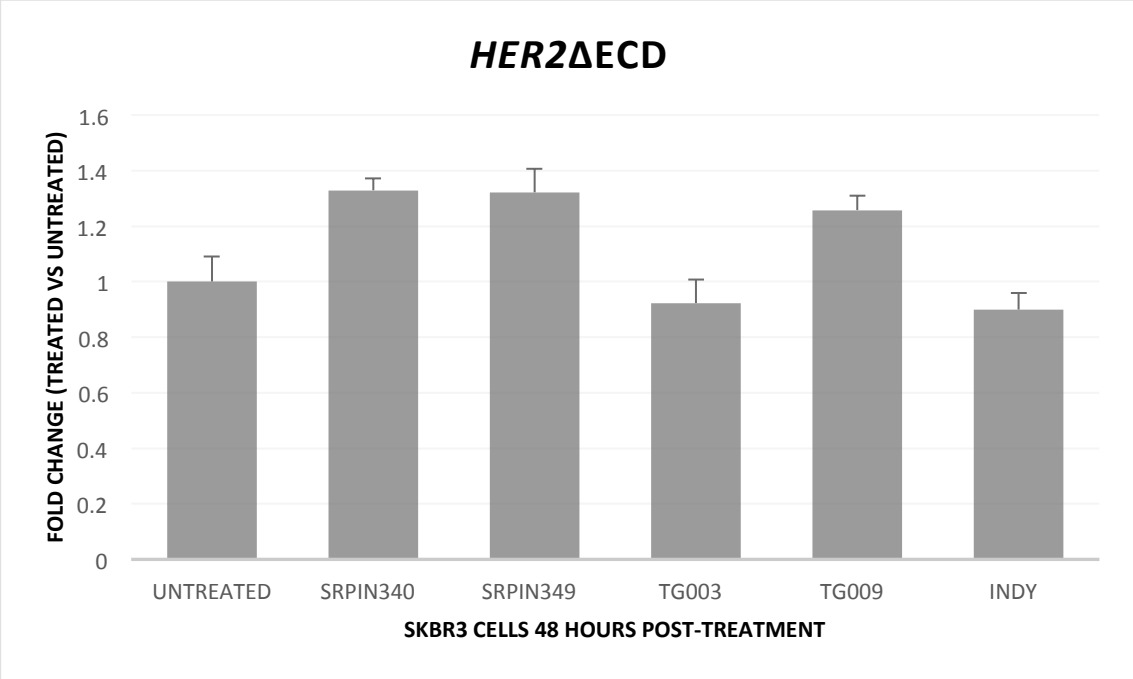


Figure 6.12: Effect of protein kinase inhibitors *SRPIN340*, *TG003* and *INDY* on the *HER2ΔECD* alternative splice variant in SKBR3 cells 48 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

6.3.1.2.3 The expression of *HER2*Δ16 following treatment with protein kinase inhibitors

There was no significant change in the expression *HER2*Δ16 after treatment for 24 hours with *SRPIN340* (fold change = 0.8; p≥0.05), *TG003* (fold change = 1.0; p≥0.05) and *INDY* (fold change = 0.9; p≥0.05). Negative controls *SRPIN340* and *TG009* also showed no change in expression after 24 hours (fold change = 0.8; p≥0.05; and 0.9; p≥0.05, respectively) (Figure 6.13). After 48 hours, there was also no significant change in the expression of *HER2*Δ16 following treatment with *SRPIN340* (fold change = 1.0; p≥0.05), *TG003* (fold change = 1.1; p≥0.05) and *INDY* (fold change = 0.8; p≥0.05). Negative controls *SRPIN349* and *TG009* also showed no significant change in the

expression of *HER2*Δ16 following treatment for 48 hours of MDA-MB-453 cells (fold change = 0.8; p≥0.05; and 0.8; p≥0.05, respectively) (Figure 6.14).

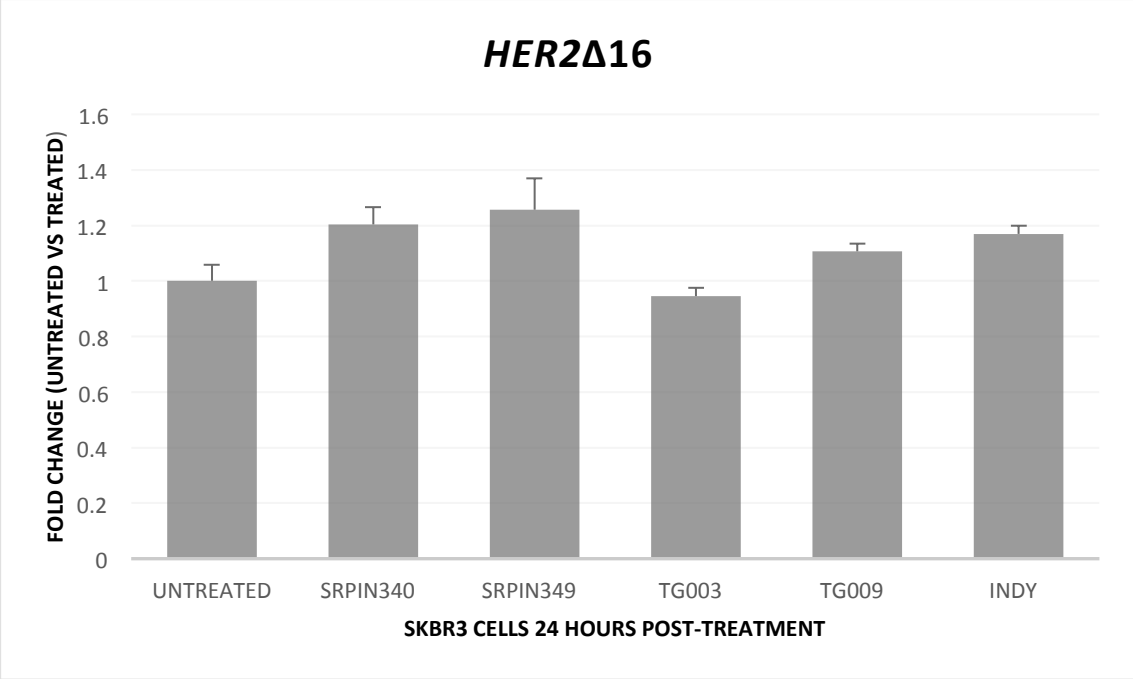


Figure 6.13: Effect of protein kinase inhibitors *SRPIN340*, *TG003* and *INDY* on the *HER2*Δ16 alternative splice variant in SKBR3 cells 24 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

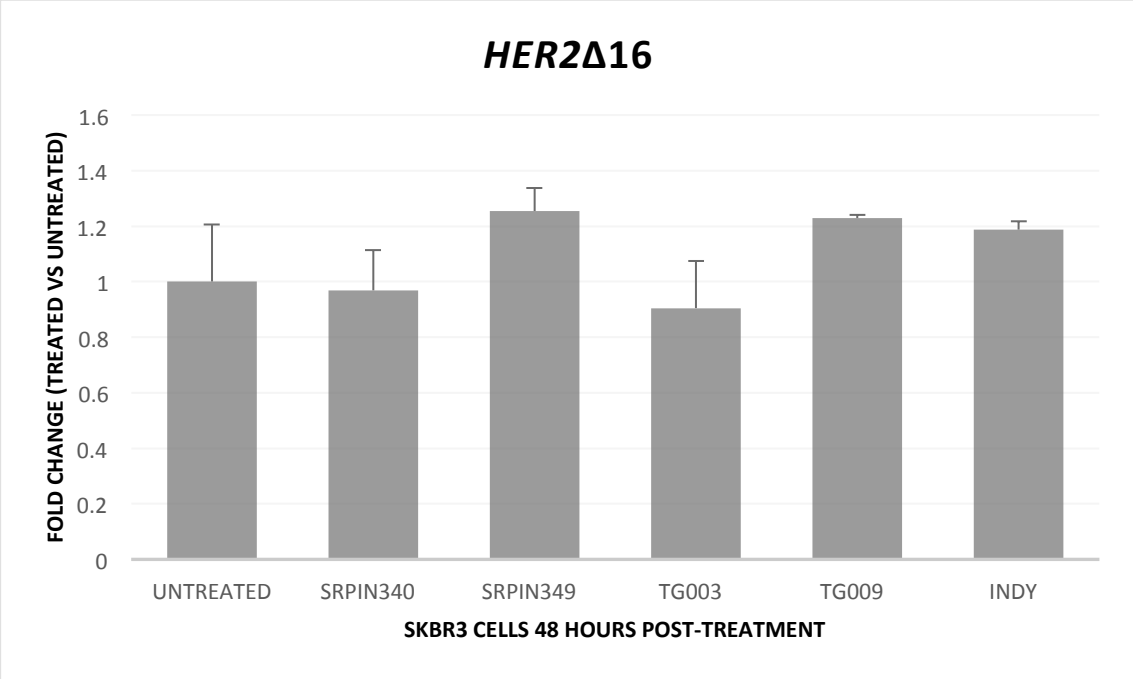


Figure 6.14: Effect of protein kinase inhibitors *SRPIN340*, *TG003* and *INDY* on the *HER2*Δ16 alternative splice variant in SKBR3 cells 48 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

6.3.1.2.4 The expression of *HER2*ΔATP following treatment with protein kinase inhibitors

There was no significant change in the expression *HER2*ΔATP after treatment for 24 hours with *SRPIN340* (fold change = 0.96; p≥0.05), *TG003* (fold change = 1.1; p≥0.05) and *INDY* (fold change = 1.1; p≥0.05). Negative controls *SRPIN340* and *TG009* also showed no change in expression after 24 hours (fold change = 0.9; p≥0.05; and 1.0; p≥0.05, respectively) (Figure 6.15). After 48 hours, there was also no significant change in the expression of *HER2*ΔATP following treatment with *SRPIN340* (fold change = 1.0; p≥0.05), *TG003* (fold change = 1.0; p≥0.05) and *INDY* (fold change = 0.9; p≥0.05). Negative controls *SRPIN349* and *TG009* also showed no significant change in the

expression of *HER2*ΔATP following treatment for 48 hours of MDA-MB-453 cells (fold change = 1.0; p≥0.05; and 1.0; p≥1.0, respectively) (Figure 6.16).

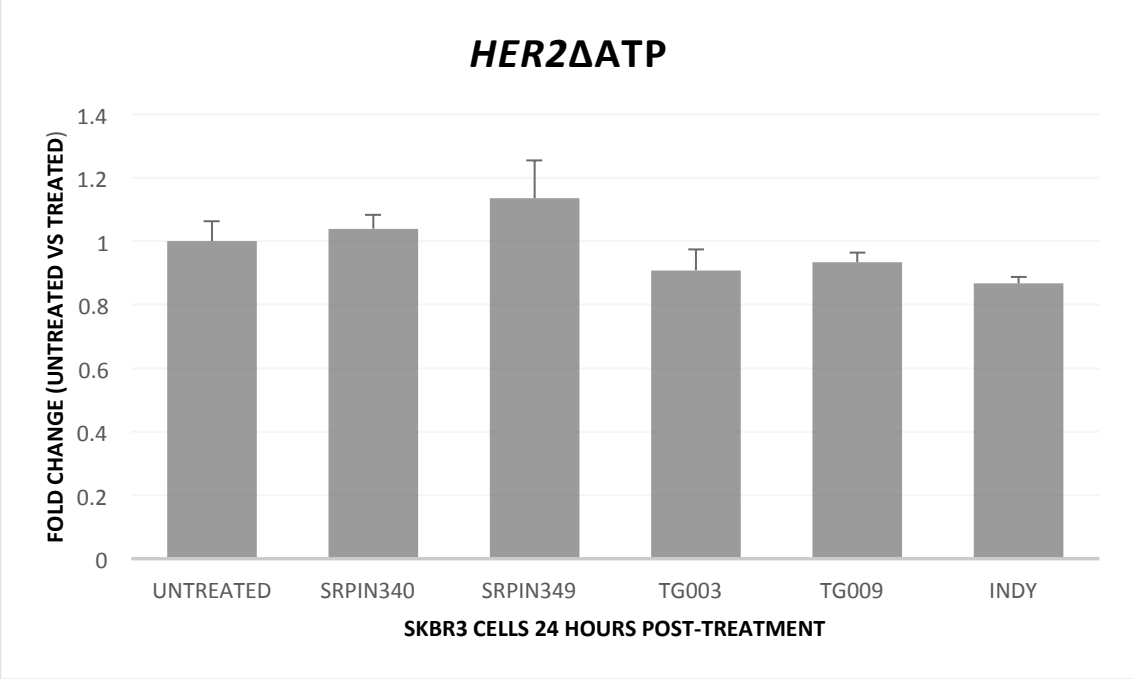


Figure 6.15: Effect of protein kinase inhibitors *SRPIN340*, *TG003* and *INDY* on the *HER2*ΔATP alternative splice variant in SKBR3 cells 24 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

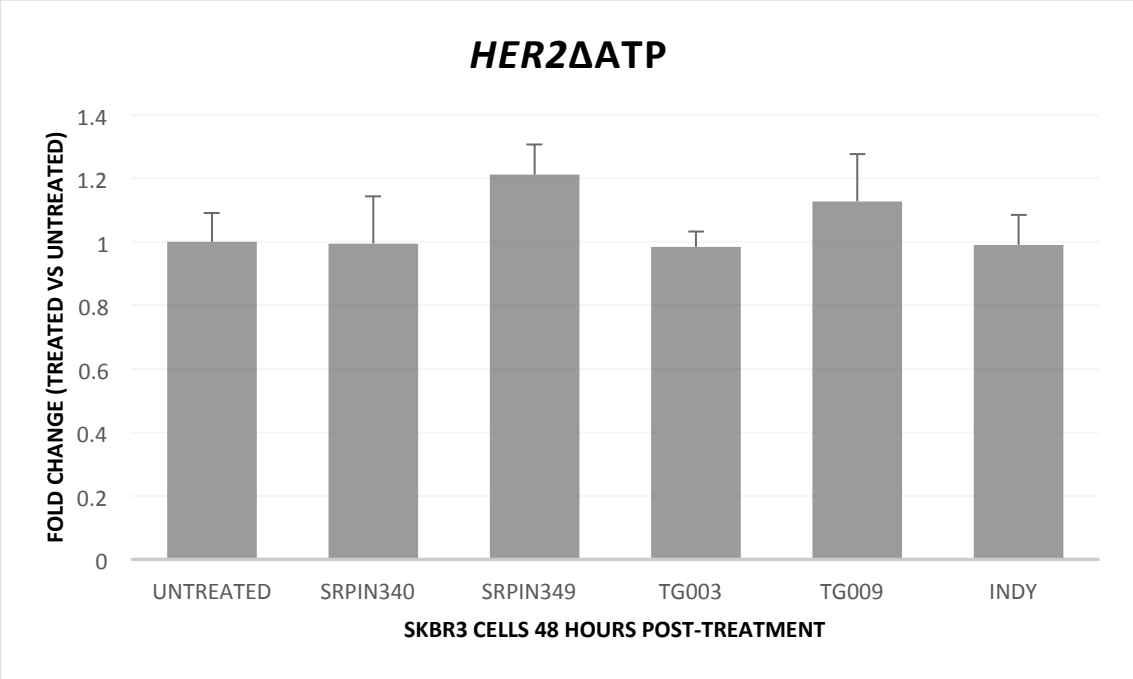


Figure 6.16: Effect of protein kinase inhibitors *SRPIN340*, *TG003* and *INDY* on the *HER2*ΔATP alternative splice variant in SKBR3 cells 48 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

6.3.1.3 BT-20 cell line

The untreated cells were used as a calibrator to measure the fold change at 24 and 48 hours post-treatment, using *B2M* as a normalisation factor (Table 6.5). Due to the significant changes observed in MDA-MB-453 cells, triple-negative BT-20 cells were treated with protein kinase inhibitors to investigate whether the changes in expression of *HER2* and *HER2* alternative splice variants were specific to MDA-MB-453 cells or breast cancer cells in general.

Gene name	Stability value		
18S	0.160	**Best gene**	B2M
ACTB	0.141	**Stability value**	0.100
ATP5B	0.160		
B2M	0.100		
CYC1	0.176		
E1F4A2	0.118		
GAPDH	0.171		
RPLI3A	0.151		
TOP1	0.346		
UBC	0.214		
YWHAZ	0.115		

Table 6.5: *Normfinder* output for the selection of an optimal reference gene in BT-20 cells treated with protein kinase inhibitors *SRPIN340*, *TG003* and *INDY*.

6.3.1.3.1 The expression of wild-type *HER2* following treatment with protein kinase inhibitors

There was no significant change in the expression wild-type *HER2* after treatment for 24 hours with *SRPIN340* (fold change = 0.82; $p \geq 0.05$). An increased expression in wild-type *HER2* expression was observed following treatment with *TG003* and *INDY* (fold change = 0.04; $p<0.01$; 0.03; $p<0.01$, respectively). Negative controls *SRPIN340* and

TG009 also showed no change in expression after 24 hours (fold change = 0.8; p≥0.05; and 0.7; p≥0.05, respectively) (Figure 6.17). After 48 hours, there was no significant change in the expression of the wild-type *HER2* following treatment with *SRPIN340* (fold change = 1.1; p≥0.05), *TG003* (fold change = 1.1; p≥0.05) and *INDY* (fold change = 0.9; p≥0.05). Negative controls *SRPIN349* and *TG009* also showed no significant change in the expression of wild-type *HER2* following treatment for 48 hours of MDA-MB-453 cells (fold change = 1.2; p≥0.05; and 0.9; p≥0.05, respectively) (Figure 6.18).

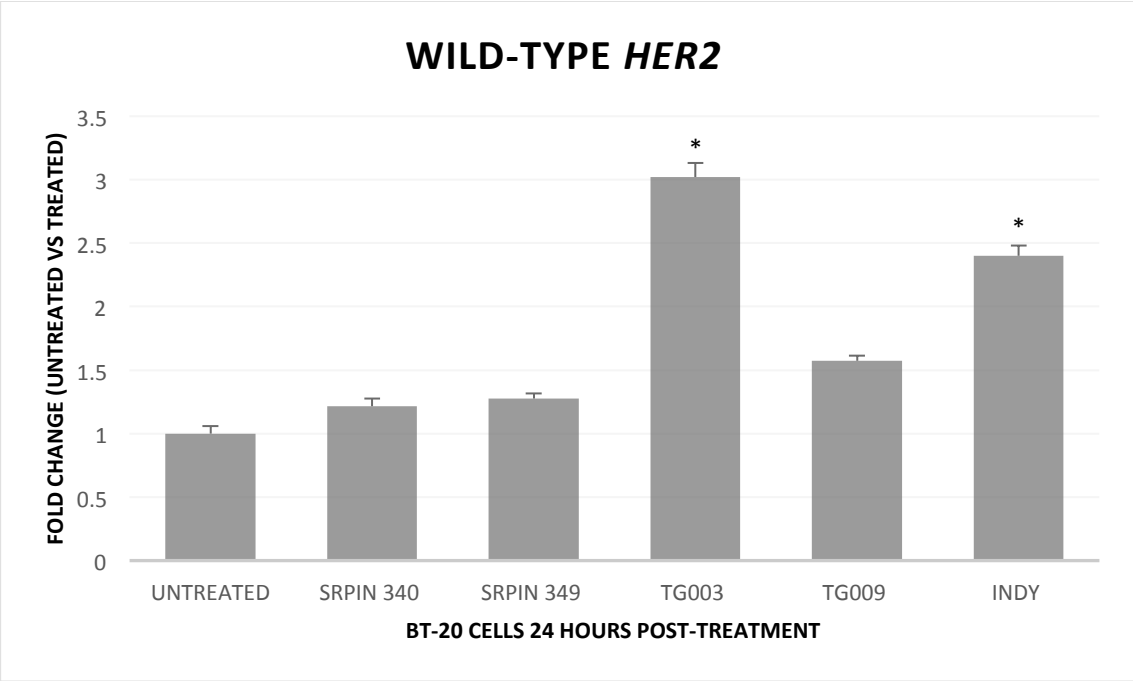


Figure 6.17: Effect of protein kinase inhibitors *SRPIN340*, *TG003* and *INDY* on the wild-type *HER2* in BT-20 cells 24 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

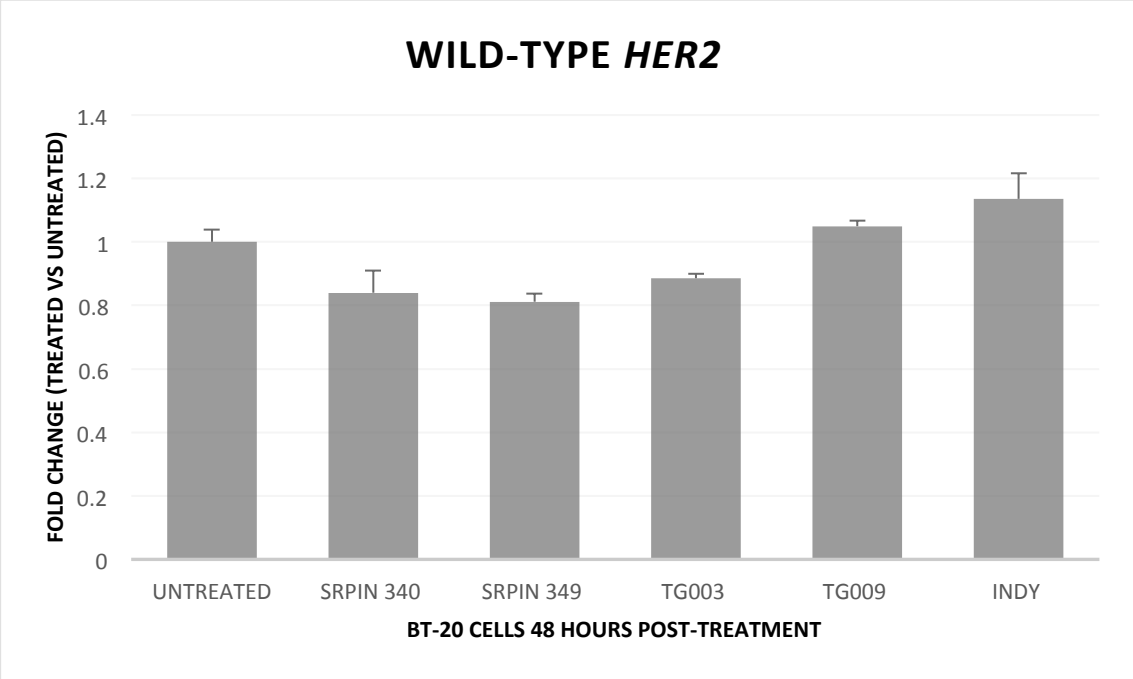


Figure 6.18: Effect of protein kinase inhibitors *SRPIN340*, *TG003* and *INDY* on the wild-type *HER2* in BT-20 cells 48 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

6.3.1.3.2 The expression of *HER2ΔECD* following treatment with protein kinase inhibitors

There was no significant change in the expression *HER2ΔECD* after treatment for 24 hours with *SRPIN340* (fold change = 1.3; p≥0.05), *TG003* (fold change = 0.8; p≥0.05) and *INDY* (fold change = 0.7; p≥0.05). Negative controls *SRPIN340* and *TG009* also showed no change in expression after 24 hours (fold change = 1.0; p≥0.05; and 1.0; p≥0.05, respectively) (Figure 6.19). After 48 hours, there was also no significant change in the expression of *HER2ΔECD* following treatment with *SRPIN340* (fold change = 0.7; p≥0.05), *TG003* (fold change = 0.9; p≥0.05) and *INDY* (fold change = 0.8; p≥0.05). Negative controls *SRPIN349* and *TG009* also showed no significant change in the

expression of *HER2ΔECD* following treatment for 48 hours of MDA-MB-453 cells (fold change = 1.1; p≥0.05; and 1.0; p≥0.05, respectively) (Figure 6.20).

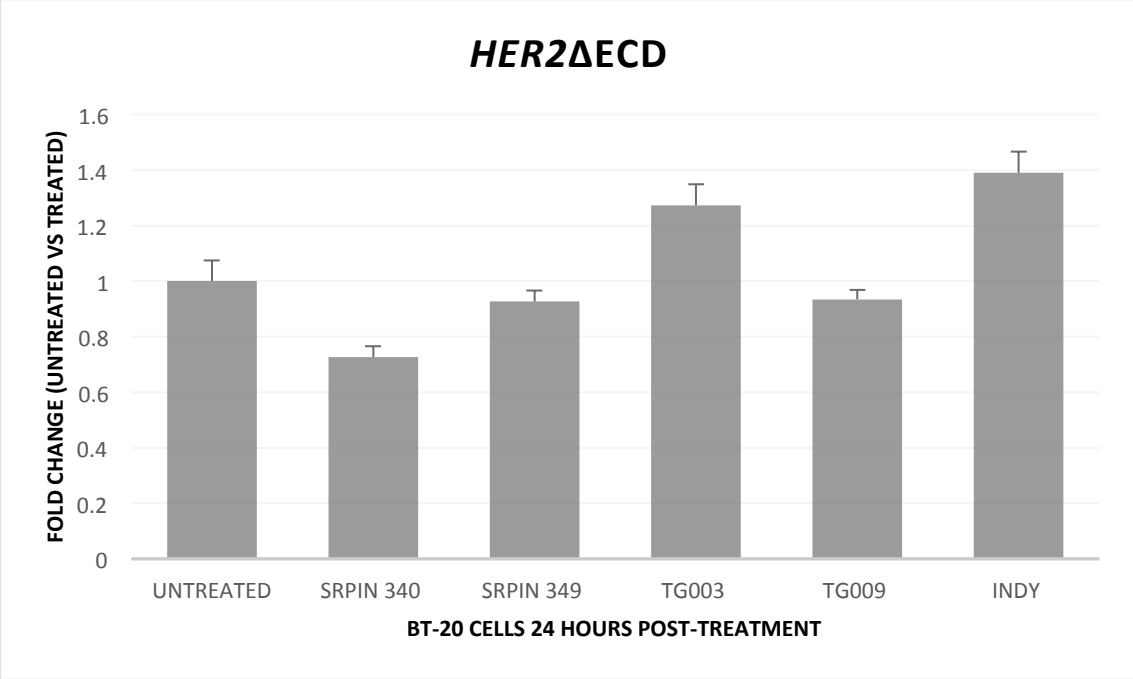


Figure 6.19: Effect of protein kinase inhibitors *SRPIN340*, *TG003* and *INDY* on the *HER2ΔECD* alternative splice variant in BT-20 cells 24 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

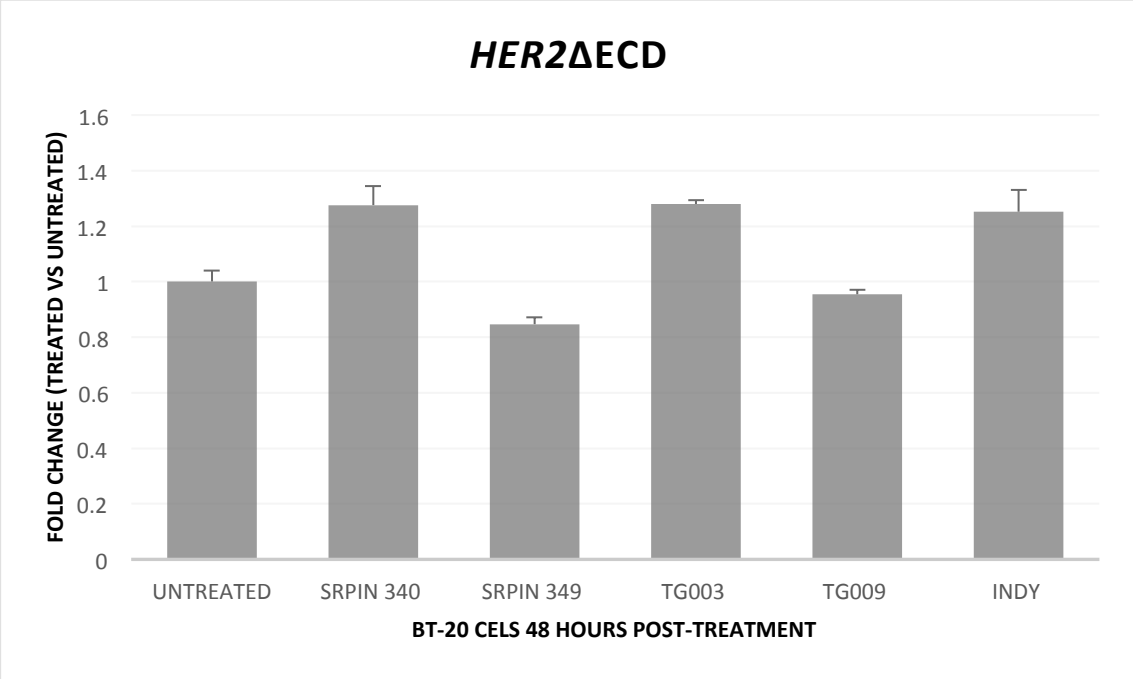


Figure 6.20: Effect of protein kinase inhibitors *SRPIN340*, *TG003* and *INDY* on the *HER2ΔECD* alternative splice variant in BT-20 cells 48 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

6.3.1.3.3 Changes in the expression of *HER2*Δ16 following treatment with protein kinase inhibitors

There was no significant change in the expression *HER2*Δ16 after treatment for 24 hours with *SRPIN340* (fold change = 1.3; p≥0.05) and *TG003* (fold change = 0.8; p≥0.05). An increase in the expression of *HER2*Δ16 was seen after 24 hour treatment with *INDY* (fold change = 0.6; p≥0.05). Negative controls *SRPIN340* and *TG009* showed no change in expression after 24 hours (fold change = 0.8; p≥0.05; and 0.9; p≥0.05, respectively) (Figure 6.21). After 48 hours, there was no significant change in the expression of *HER2*Δ16 following treatment with *SRPIN340* (fold change = 0.9; p≥0.05), *TG003* (fold change = 1.1; p≥0.05) and *INDY* (fold change = 1.0; p≥0.05). Negative controls *SRPIN349* and *TG009* also showed no significant change in the expression of

*HER2*Δ16 following treatment for 48 hours of MDA-MB-453 cells (fold change = 0.9; p≥0.05; and 1.0; p≥0.05, respectively) (Figure 6.22).

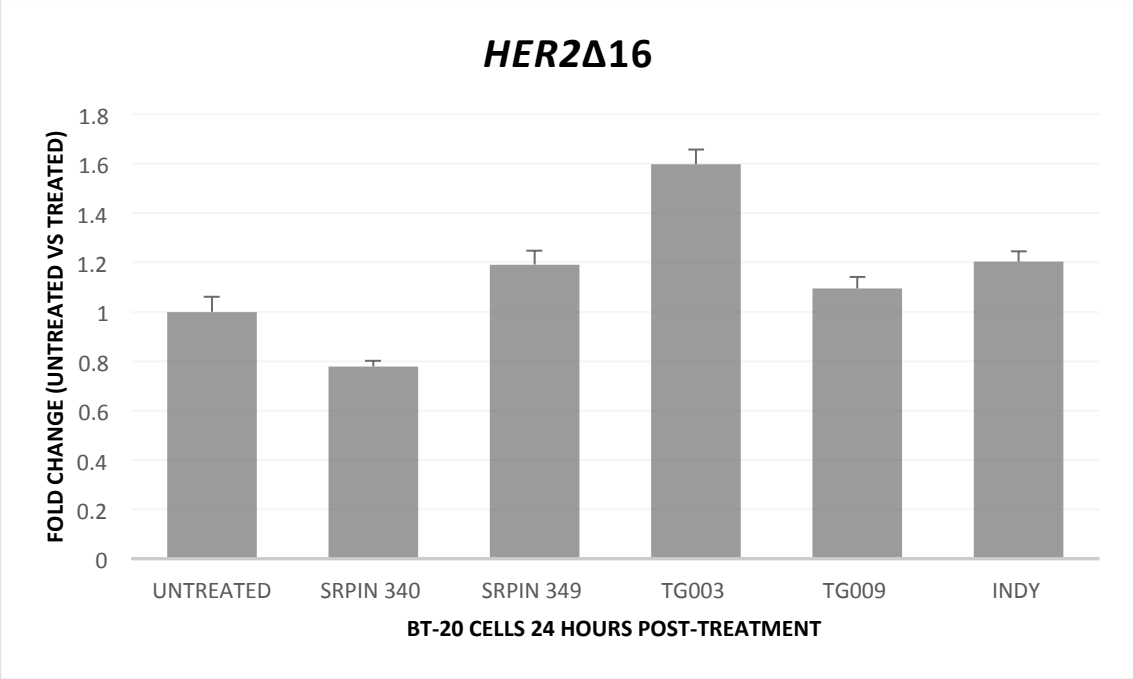


Figure 6.21: Effect of protein kinase inhibitors *SRPIN340*, *TG003* and *INDY* on the *HER2*Δ16 alternative splice variant in BT-20 cells 24 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

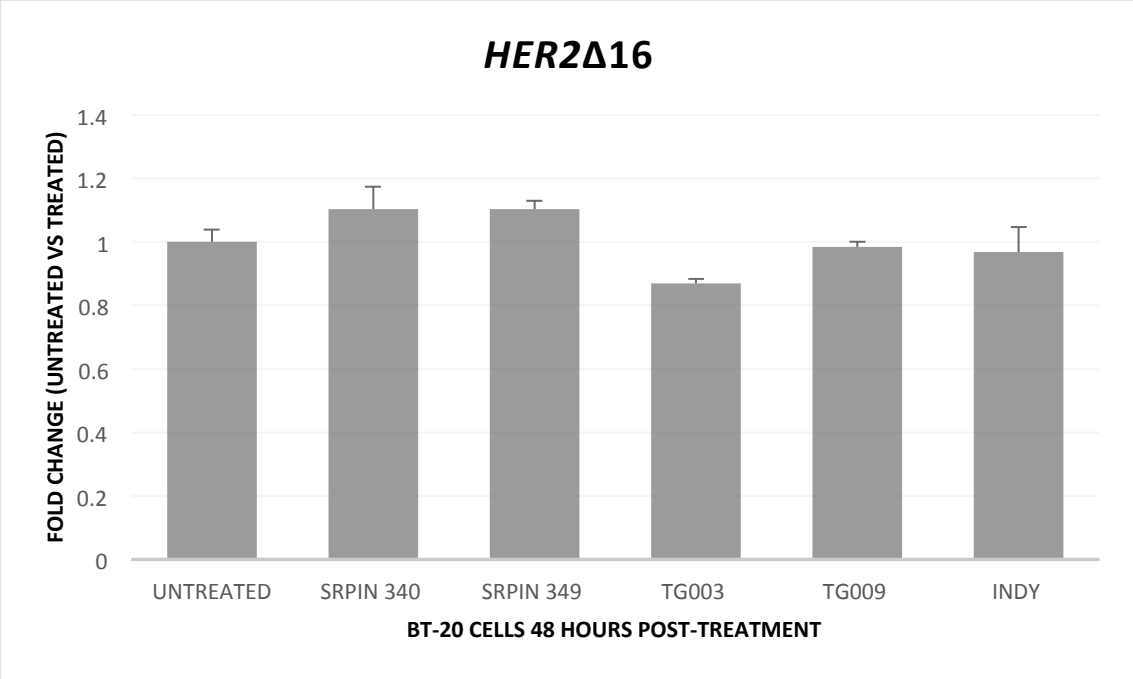


Figure 6.22: Effect of protein kinase inhibitors *SRPIN340*, *TG003* and *INDY* on the *HER2*Δ16 alternative splice variant in BT-20 cells 48 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

6.3.1.3.4 Changes in the expression of *HER2*ΔATP following treatment with protein kinase inhibitors

There was no significant change in the expression *HER2*ΔATP after treatment for 24 hours with *SRPIN340* (fold change = 1.0; p≥0.05) and *TG003* (fold change = 0.98; p≥0.05). An increase in the expression *HER2*ΔATP was observed after 24 hour treatment with *INDY* (fold change = 0.058; p≥0.05) Negative controls *SRPIN340* and *TG009* also showed no change in expression after 24 hours (fold change = 1.1; p≥0.05; and 0.9; p≥0.05, respectively) (Figure 6.23). After 48 hours, there was also no significant change in the expression of *HER2*ΔATP following treatment with *SRPIN340* (fold change = 1.0; p≥0.05), *TG003* (fold change = 0.98; p≥0.05) and *INDY* (fold change = 0.92; p≥0.05). Negative controls *SRPIN349* and *TG009* also showed no significant

change in the expression of *HER2*ΔATP following treatment for 48 hours of MDA-MB-453 cells (fold change = 0.97; p≥0.05; and 0.84; p≥1.0, respectively) (Figure 6.24).

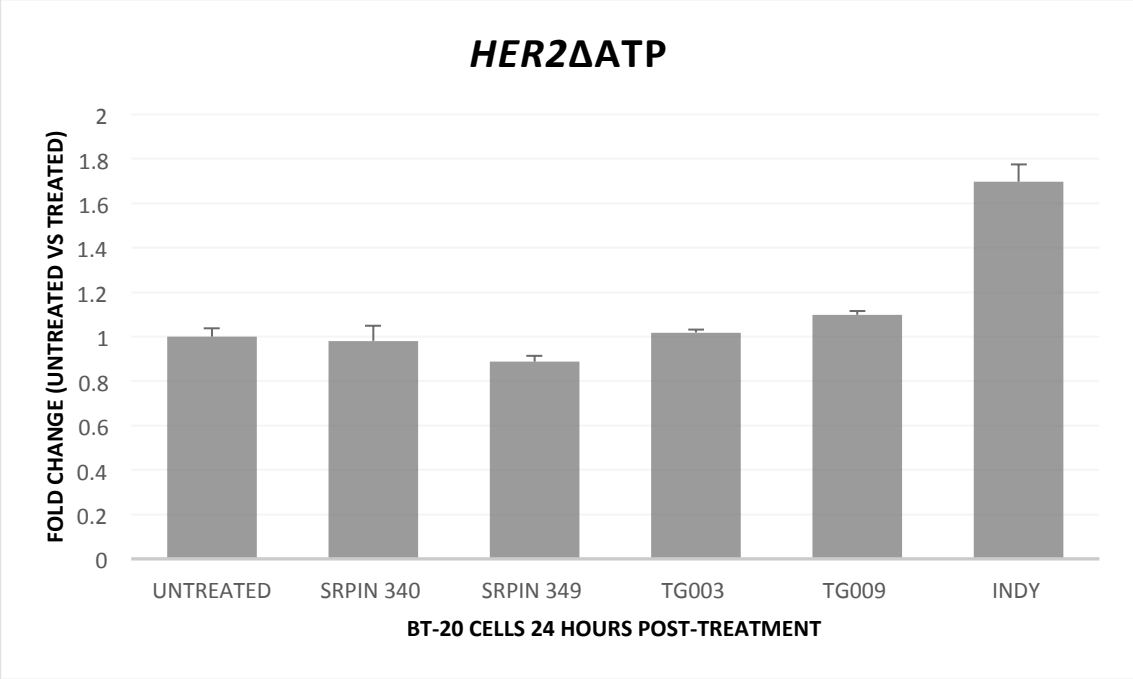


Figure 6.23: Effect of protein kinase inhibitors *SRPIN340*, *TG003* and *INDY* on the *HER2*ΔATP alternative splice variant in BT-20 cells 24 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

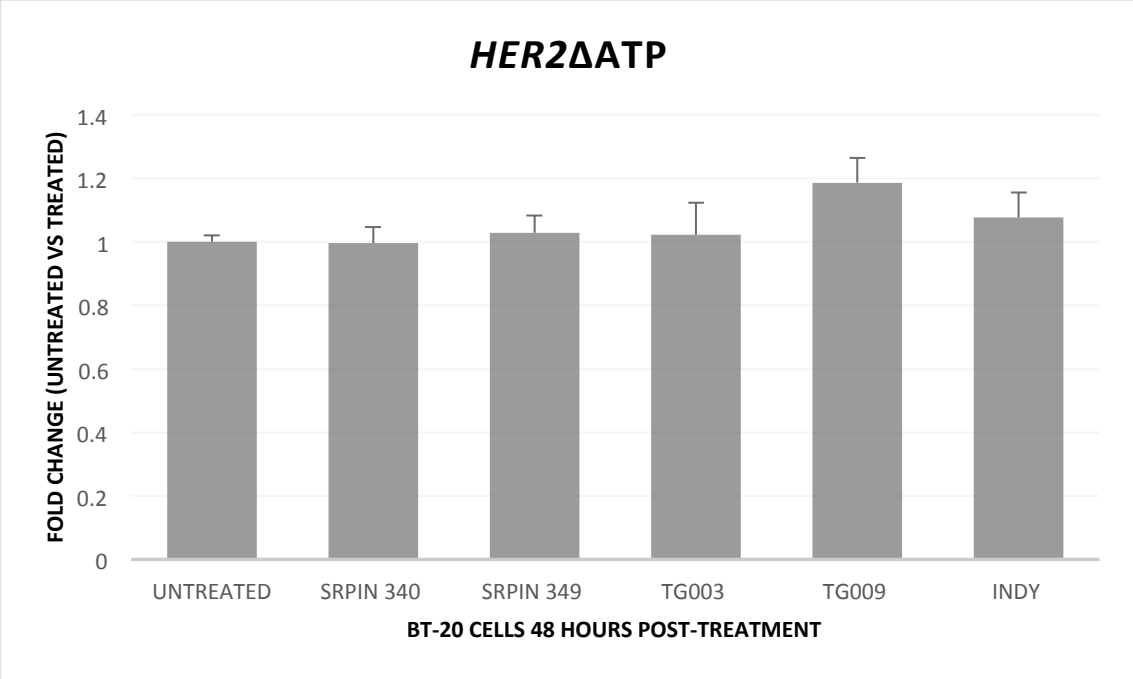


Figure 6.24: Effect of protein kinase inhibitors *SRPIN340*, *TG003* and *INDY* on the *HER2*ΔATP alternative splice variant in BT-20 cells 48 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

6.3.2 Induction of hypoxia by hypoxia mimetic factor Cobalt Chloride ($CoCl_2$) inhibits the expression of *HER2* and *HER2* alternative splice variants in SKBR3 cell line.

The untreated cells were used as a calibrator to measure the fold change at 24 and 48 hours post-treatment, using *ACTB* as a normalisation factor (Table 6.6). No results are shown for the $CoCl_2$ induction of hypoxia in MDA-MB-453 cell lines because the MDA-MB-453 cells did not thrive during this experiment.

Gene name	Stability value		
18S	0.355	**Best gene**	ACTB
ACTB	0.203	**Stability value**	0.203
ATP5B	0.398		
B2M	0.233		
CYC1	0.207		
E1F4A2	0.289		
GAPDH	0.438		
RPLI3A	0.284		
SDHA	0.494		
TOP1	0.258		
UBC	0.482		
YWHAZ	0.281		

Table 6.6: Normfinder output for the selection of an optimal reference gene in SKBR3 cells treated with Cobalt Chloride.

6.3.2.1. Changes in HIF1-α expression after treatment of SKBR3 cells with Cobalt Chloride for 24 and 48 hours

After 24 hours, and taking the untreated cells as 1, a significant increase was observed in the expression of HIF1-α with 100µM (fold change = 0.59; $p<0.001$), 200µM (fold change = 0.59; $p<0.001$), 300µM (fold change = 0.58; $p<0.001$), 400µM (fold change = 0.6; $p<0.01$), and 500µM (fold change = 0.7; $p<0.01$) concentrations of $CoCl_2$ (Figure 6.25).

After 48 hours, a significant reverse of this effect is observed; HIF1-α expression significantly reduces after treatment with 100µM (fold change = 2.15; p<0.0001), 200µM (fold change = 2.0; p<0.0001), 300µM (fold change 2.9; p<0.0001), 400µM (fold change = 4.9; p<0.0001), and 500µM (fold change = 5.3; p<0.0001) concentrations of $CoCl_2$ (Figure 6.26).

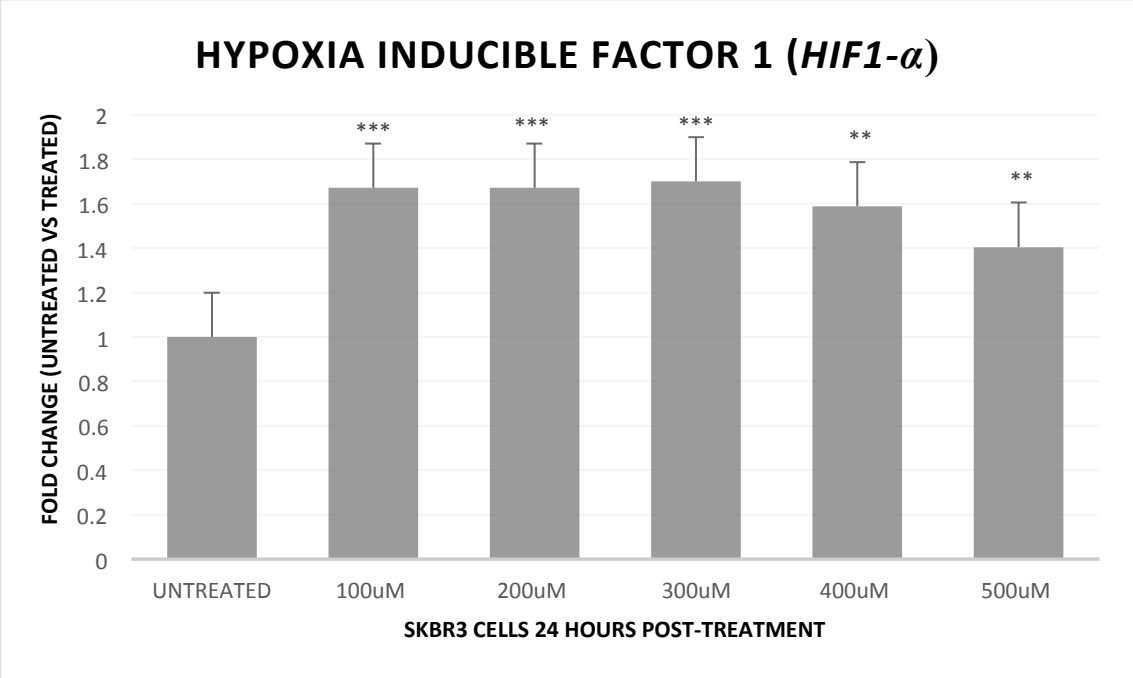


Figure 6.25: Effect of Cobalt chloride treatment on *HIF1-α* gene in SKBR3 cells 24 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

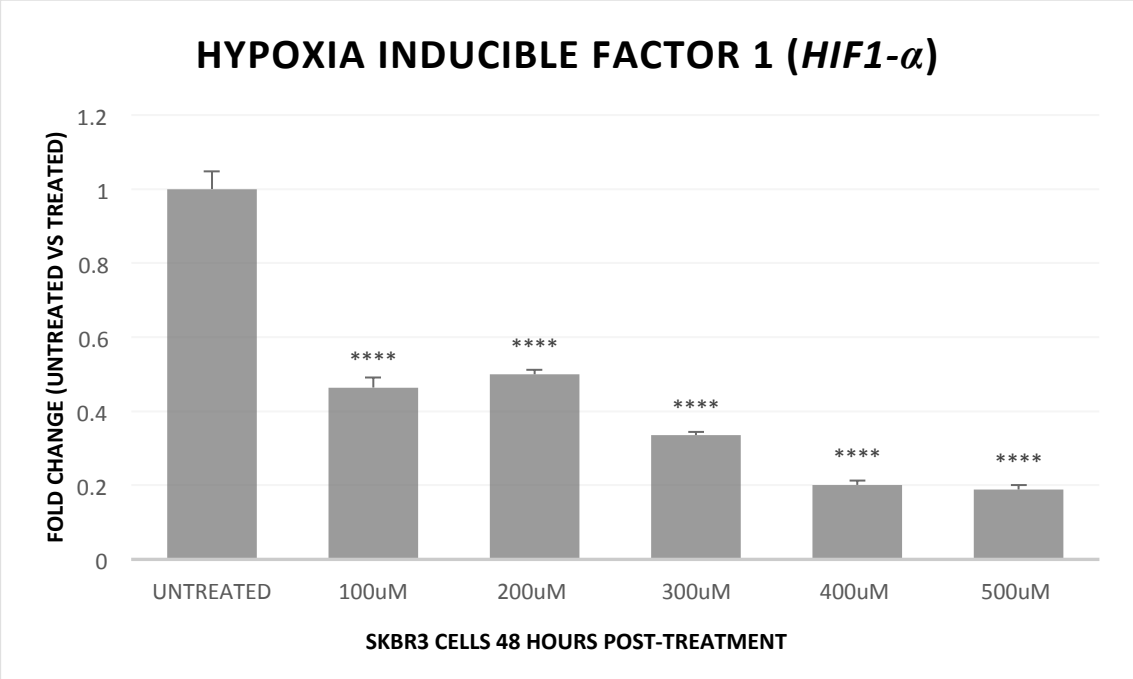


Figure 6.26: Effect of Cobalt chloride treatment on *HIF1-α* gene in SKBR3 cells 48 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

6.3.2.2. Changes in the expression of *HER2* and *HER2* alternative splice variants after treatment of SKBR3 cells with Cobalt Chloride for 24 and 48 hours

Treatment of *HER2*-positive SKBR3 cells with hypoxia mimetic $CoCl_2$ resulted in a significant reduction in the expression of *HER2* and *HER2* splice variants. At 24 hours post-treatment, a significant decrease in expression was observed in the wild-type *HER2* (fold change = 6.38; p<0.0001) (

Figure 6.27), *HER2ΔECD* (fold change = 4.39; p<0.001) (Figure 6.28), *HER2*Δ16 (fold change = 5.14; p<0.001) (Figure 6.29), and *HER2*ΔATP (fold change = 5.13; p<0.001) (Figure 6.30). At 48 hours post-treatment, the same significant decrease in expression

was observed in the wild-type *HER2* (fold change = 6.49; p<0.0001) (Figure 6.31), *HER2ΔECD* (fold change = 2.76; p<0.0001) (Figure 6.32), *HER2*Δ16 (fold change = 2.9; p<0.0001) (Figure 6.33), and *HER2*ΔATP (fold change = 2.9; p<0.0001) (Figure 6.34).

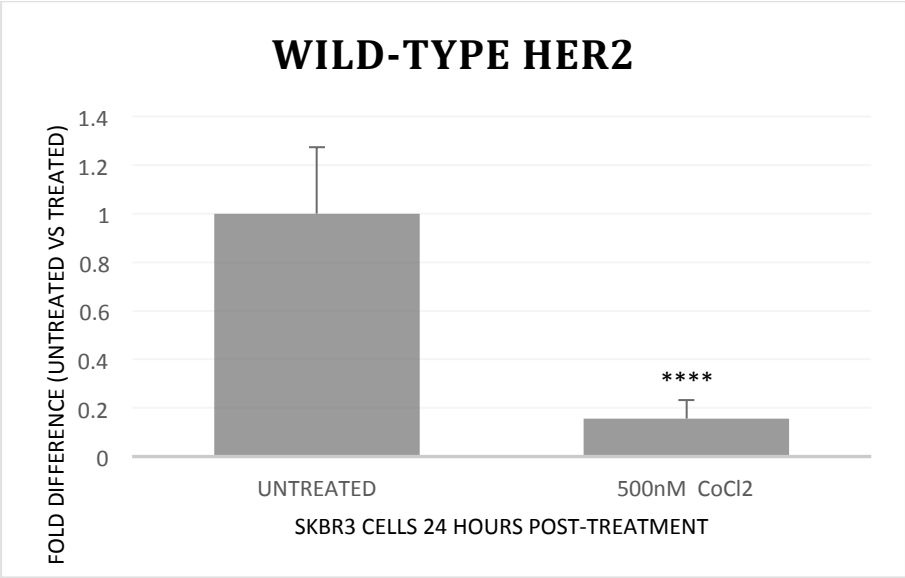


Figure 6.27: Effect of Cobalt Chloride treatment on the wild-type *HER2* in SKBR3 cells 24 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the Ct values.

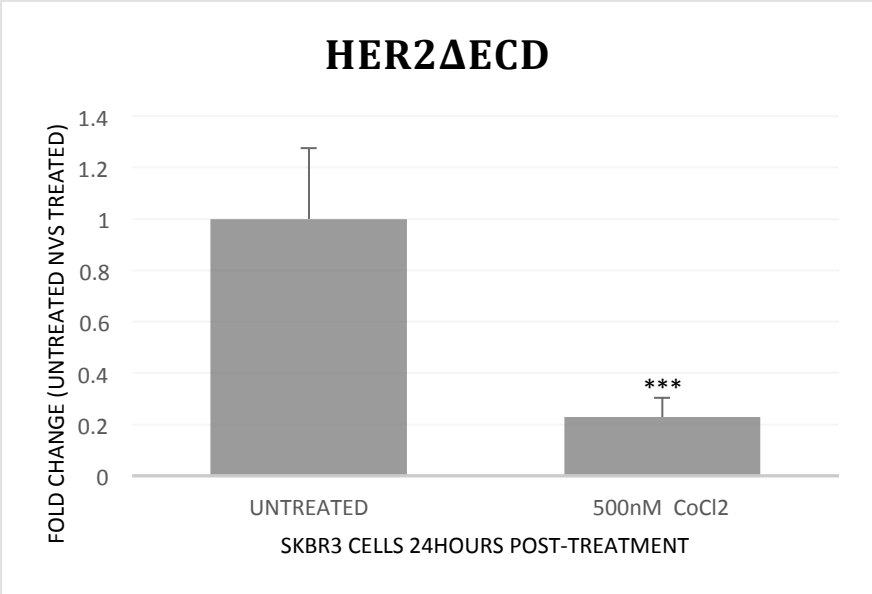


Figure 6.28: Effect of Cobalt Chloride treatment on the *HER2ΔECD* alternative splice variant in SKBR3 cells 24 hours after treatment. Each histogram bar is representative of an average of 9 samples

consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

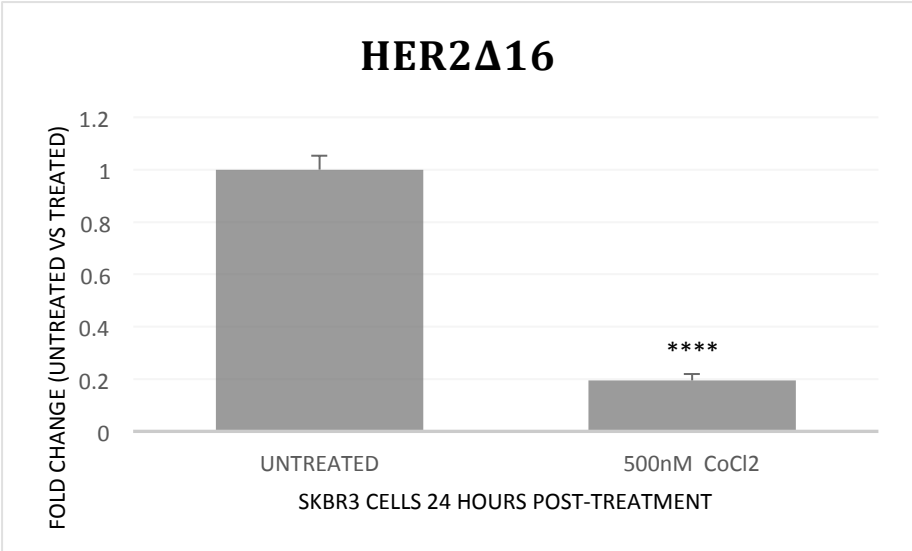


Figure 6.29: Effect of Cobalt Chloride treatment on the *HER2*Δ16 alternative splice variant in SKBR3 cells 24 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

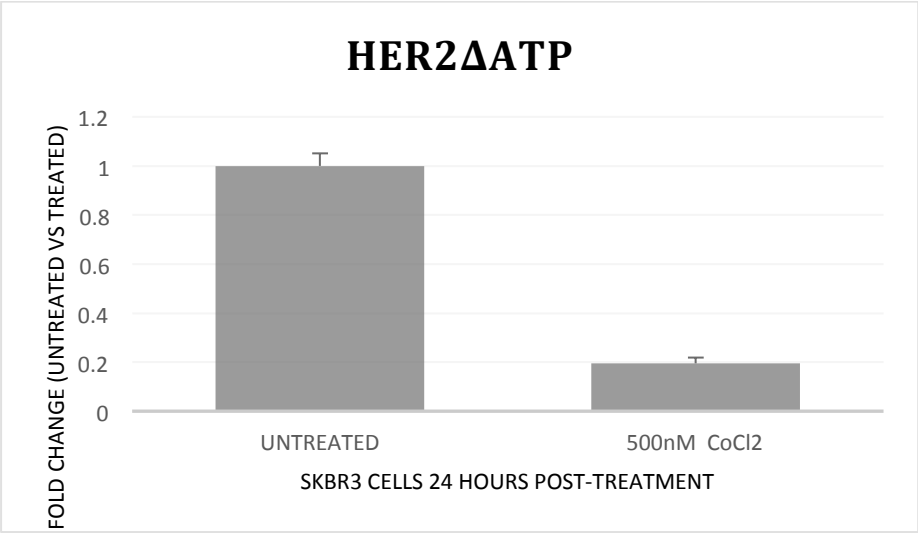


Figure 6.30: Effect of Cobalt Chloride treatment on the *HER2*ΔATP alternative splice variant in SKBR3 cells 24 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

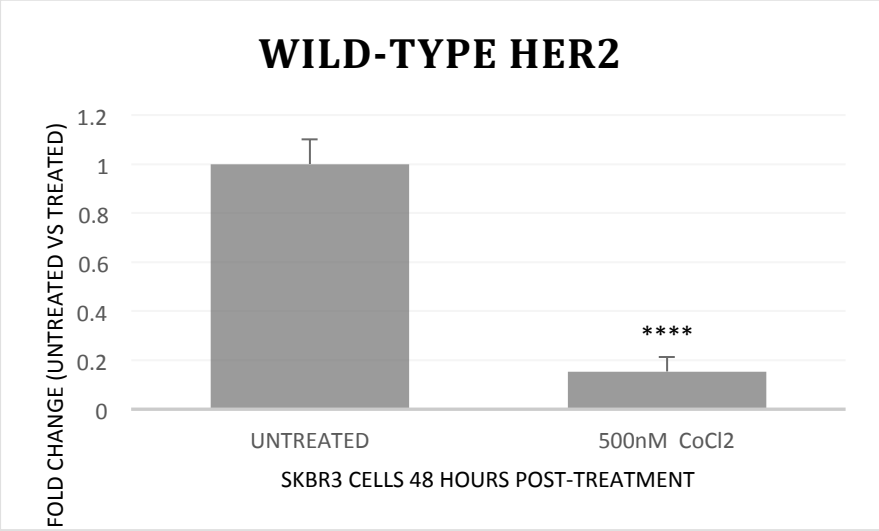


Figure 6.31: Effect of Cobalt Chloride treatment on the wild-type *HER2* in SKBR3 cells 48 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

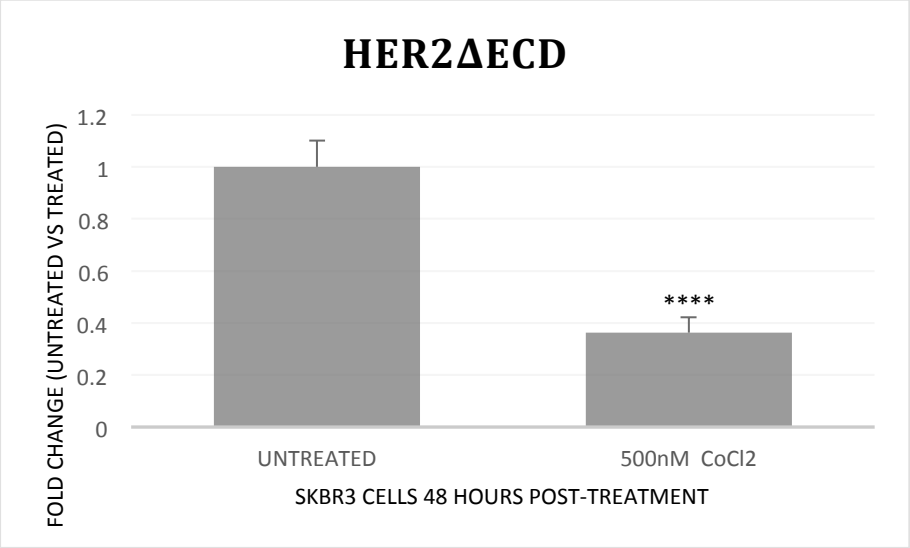


Figure 6.32: Effect of Cobalt Chloride treatment on the *HER2ΔECD* alternative splice variant in SKBR3 cells 48 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

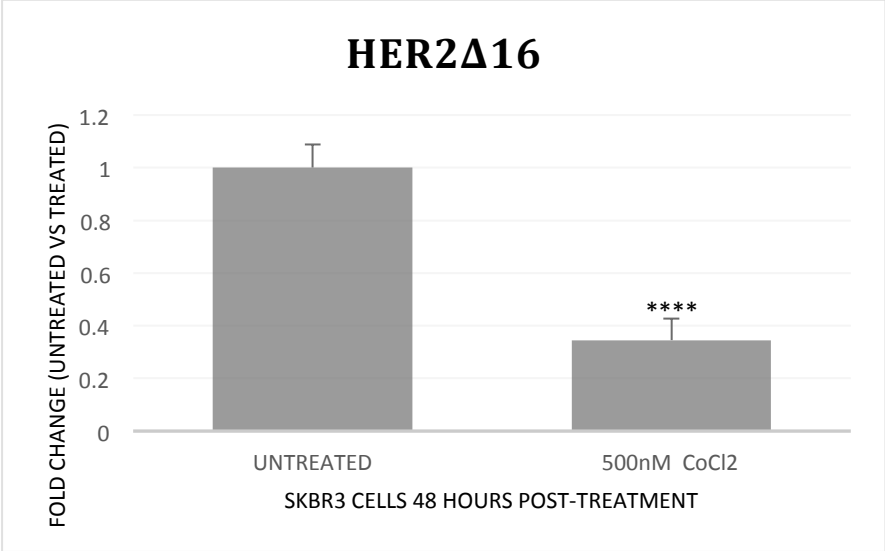


Figure 6.33: Effect of Cobalt Chloride treatment on the *HER2*Δ16 alternative splice variant in SKBR3 cells 48 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

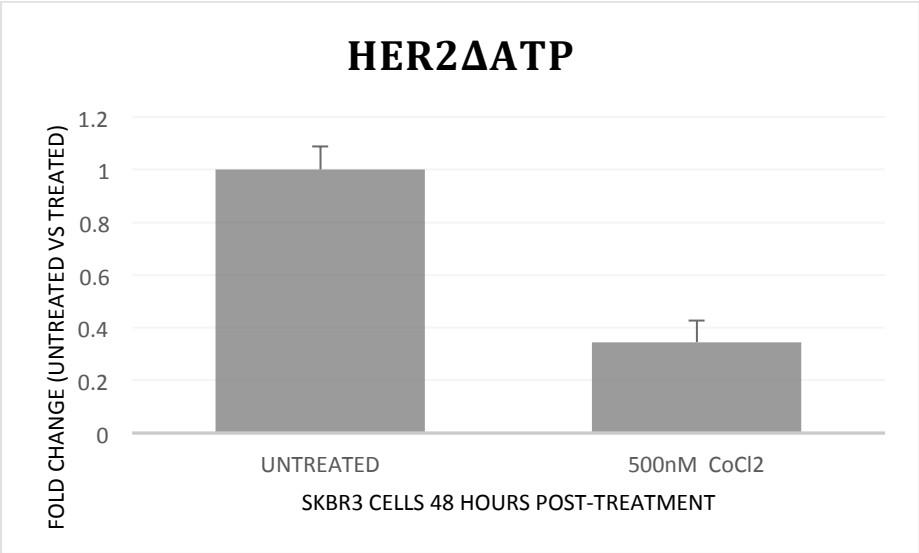


Figure 6.34: Effect of Cobalt Chloride treatment on the *HER2*ΔATP alternative splice variant in SKBR3 cells 48 hours after treatment. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

6.3.3. The effects of *SRPK1* and *SRSF1* knockdown on the expression of *HER2* and *HER2* alternative splice variants in *HER2*-positive MDA-MB-453 and SKBR3 breast cancer cell lines.

Cells transfected with a non-targeting siRNA were used as a calibrator to measure the fold change at 24 and 48 hours post-treatment, using *RPLI3A* as a normalisation factor (Table 6.7).

Gene name	Stability value		
18S	0.217	**Best gene**	RPLI3A
ACTB	0.119	**Stability value**	0.030
ATP5B	0.334		
B2M	0.269		
CYC1	0.184		
E1F4A2	0.168		
GAPDH	0.760		
RPLI3A	0.030		
SDHA	0.179		
TOP1	0.093		
UBC	0.048		
YWHAZ	0.089		

Table 6.7: Normfinder output for the selection of an optimal reference gene in MDA-MB-453 and SKBR3 cells after siRNA knockdown of *SRPK1* and *SRSF1* splice factors.

6.3.3.1 Confirmation of *SRPK1* and *SRSF1* knockdown in MDA-MB-453 cells

At both 24 and 48 hours post-transfection, a significant knockdown was observed for both *SRPK1* and *SRSF1* in MDA-MB-453 cells. Using non-targeting siRNA as a calibrator, a 24 hour transfection produced a reduction in *SRPK1* mRNA of 1.7fold (± 0.42; p<0.0001) (Figure 6.35), and a reduction in *SRSF1* mRNA of 0.76 fold (± 0.63; p<0.01) (Figure 6.37), and a 48 hour transfection produced a reduction in *SRPK1* mRNA of 2.0 fold (±0.05; p<0.0001) (Figure 6.36), and in *SRSF1* of 1.8 fold (±0.34; p<0.0001) (Figure 6.38).

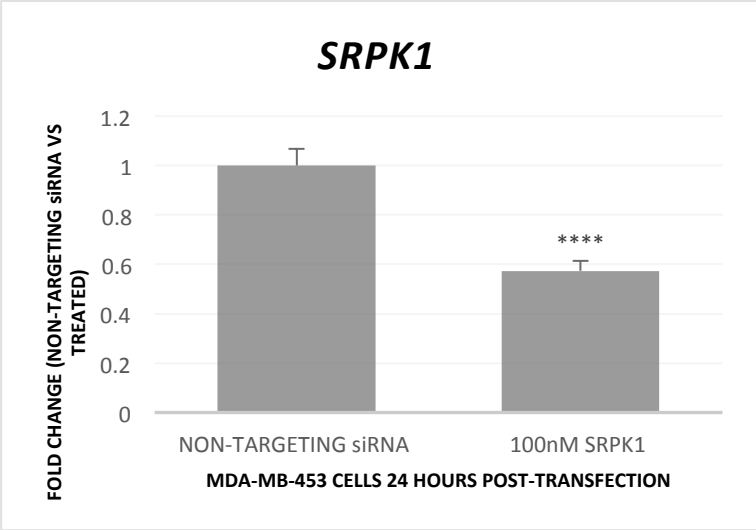


Figure 6.35: Knockdown of *SRPK1* mRNA in MDA-MB-453 cells after transfection with *SRPK1* smartpool siGENOME siRNA; a mixture of four separate siRNAs supplied in a single tube. MDA-MB-453 cells were transfected with 100nM of either a non-targeting siRNA or *SRPK1*-specific siRNA for 24 hours before RNA extraction, reverse transcription and qPCR analysis. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

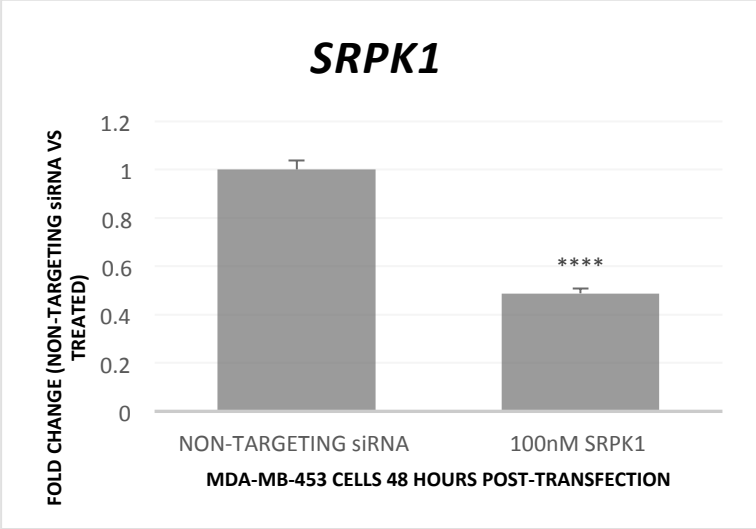


Figure 6.36: Knockdown of *SRPK1* mRNA in MDA-MB-453 cells after transfection with *SRPK1* smartpool siGENOME siRNA; a mixture of four separate siRNAs supplied in a single tube. MDA-MB-453 cells were transfected with 100nM of either a non-targeting siRNA or *SRPK1*-specific siRNA for 48 hours before RNA extraction, reverse transcription and qPCR analysis. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

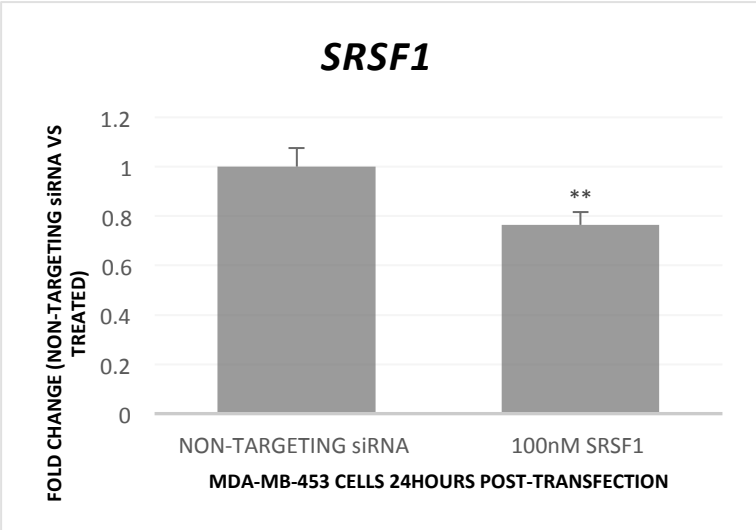


Figure 6.37: Knockdown of *SRSF1* mRNA in MDA-MB-453 cells after transfection with *SRSF1* smartpool siGENOME siRNA; a mixture of four separate siRNAs supplied in a single tube. MDA-MB-453 cells were transfected with 100nM of either a non-targeting siRNA or *SRSF1*-specific siRNA for 24 hours before RNA extraction, reverse transcription and qPCR analysis. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

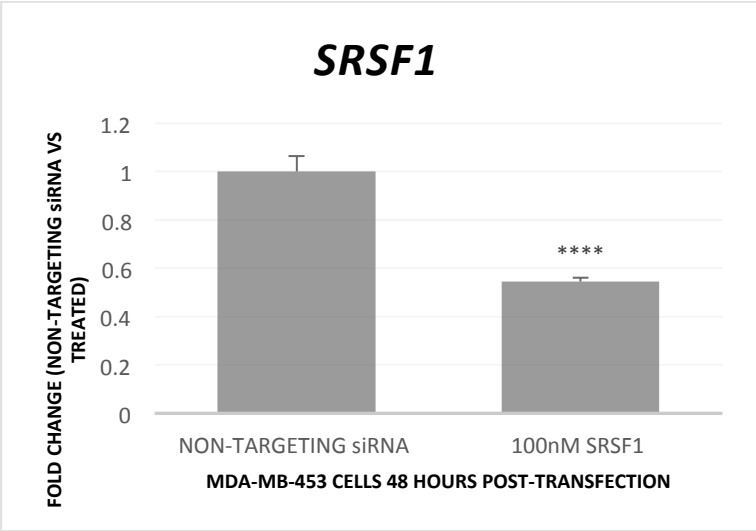


Figure 6.38: Knockdown of *SRSF1* mRNA in MDA-MB-453 cells after transfection with *SRSF1* smartpool siGENOME siRNA; a mixture of four separate siRNAs supplied in a single tube. MDA-MB-453 cells were transfected with 100nM of either a non-targeting siRNA or *SRSF1*-specific siRNA for 48 hours before RNA extraction, reverse transcription and qPCR analysis. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

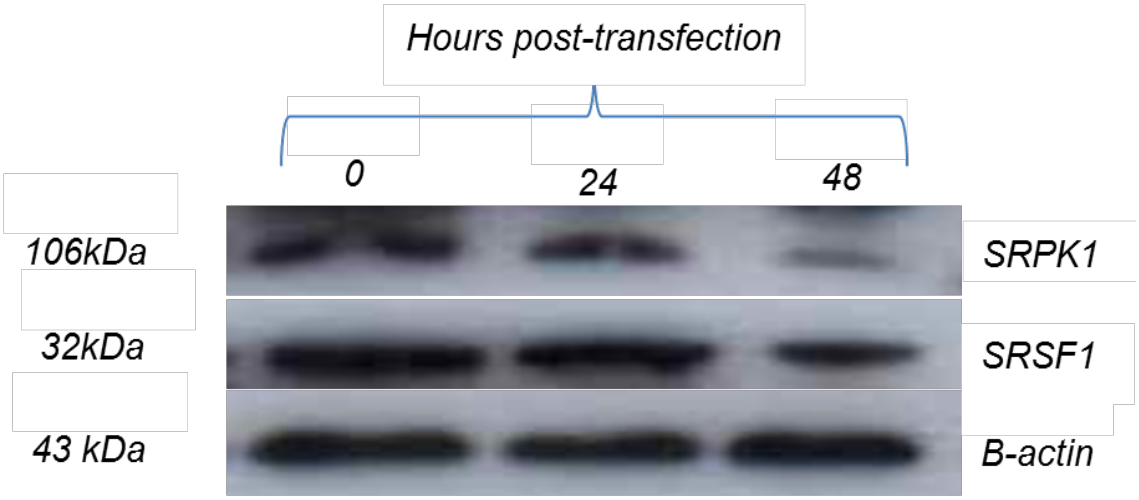


Figure 6.39: Western blot of *SRPK1* and *SRSF1* in MDA-MD-453 cells showing, 0, 24 and 48 hours post transfection. MDA-MB-453 cells were transfected with 100nM siGENOME siRNA targeting *SRPK1* or *SRSF1*. Western blot was performed using *SRPK1* mouse monoclonal antibody clone EE-13:sc100443 and *ASF/SF2* (*SRSF1*) mouse monoclonal antibody clone 96:sc33652 (Santa Cruz Biotechnology) and β-actin was used as a loading control Protein bands indicate that *SRPK1* and *SRSF1* expression is lowest 48 hours post-transfection.(representative of n-3).

6.3.3.2. Knockdown of *SRPK1* and *SFSF1* shows no significant effect on the expression of *HER2* and *HER2* alternative splice variants in MDA-MB-453 cells at mRNA level

In MDA-MB-453 cells, slight changes in the expression of *HER2* and *HER2* splice variants were observed after *SRPK1* and *SRSF1* transfection. After 24 hours of transfection with *SRPK1*, MDA-MB-453 cells showed a slight reduction in the expression of wild-type *HER2* (fold change = 1.23; p≥0.05) (Figure 6.40), *HER2ΔECD* (fold change = 1.28; p≥0.05) (Figure 6.41), *HER2*Δ16 (fold change = 1.5; p≥0.05) (Figure 6.42), and *HER2*ΔATP (fold change = 1.59; p≥0.05) (Figure 6.43); and a slight reduction in the expression of wild-type *HER2* (fold change = 1.45; p≥0.05) (Figure 6.40) and *HER2ΔECD* (fold change = 1.54; p≥0.05) (Figure 6.41) after 24 hours of transfection with *SRSF1*. These changes, however, are not statistically significant. *HER2*Δ16 and *HER2*ΔATP remain unchanged after siRNA transfection of *SRPK1* and *SRSF1* for 24 hours (Figure 6.42 and Figure 6.43). There were also no significant changes in the expression of wild-type *HER2*, *HER2ΔECD* , *HER2*Δ16 and *HER2*ΔATP after transfection with *SRPK1* for 48 hours [fold difference = (1.4; p≥0.05); (1.35; p≥0.05); (1.0; p≥0.05); (1.26; p≥0.05), respectively] (Figures 6.44-6.47), nor in the expression of wild-type *HER2*, *HER2ΔECD* , *HER2*Δ16 and *HER2*ΔATP after transfection with *SRSF1* for 48 hours [fold difference = (1.56; p≥0.05); (1.02; p≥0.05); (1.17; p≥0.05); (0.85; p≥0.05), respectively] (Figures 6.44-6.47).

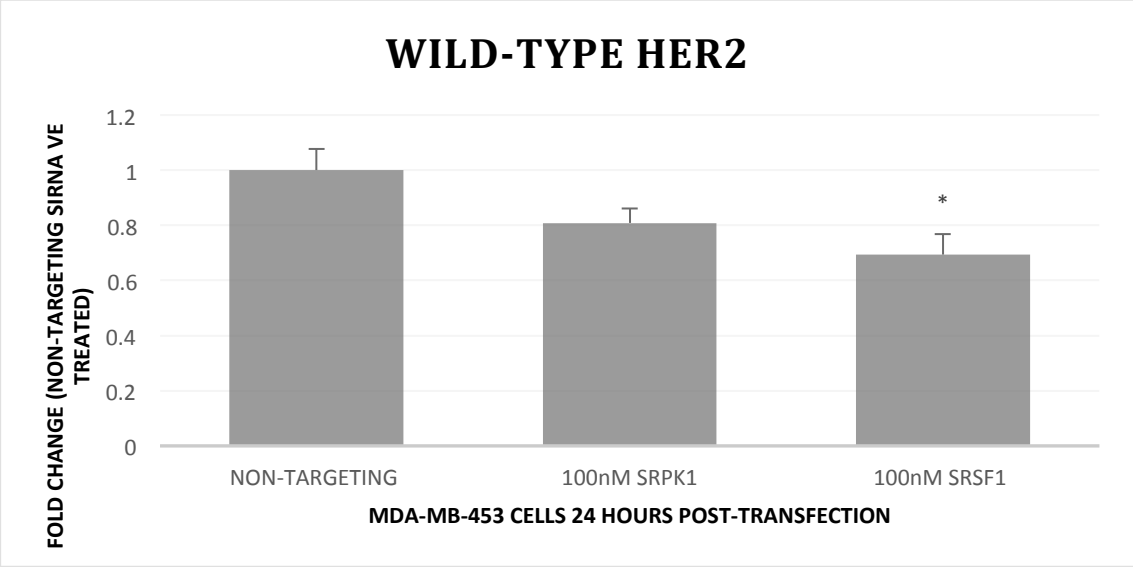


Figure 6.40: Effect of knockdown on wild-type *HER2* mRNA in MDA-MB-453 cells 24 hours after transfection with smartpool siGENOME siRNA specific to *SRPK1* and *SRSF1* splice factors. Non-targeting siRNA is used as a calibrator to measure the fold difference in *HER2* mRNA after transfection for 24 hours. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

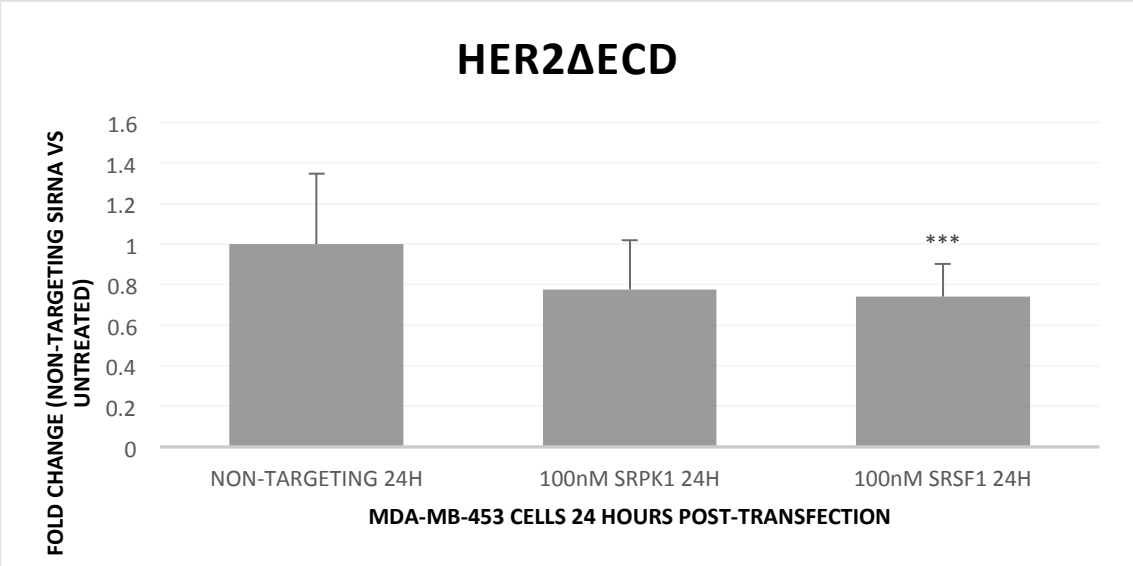


Figure 6.41: Effect of knockdown on *HER2ΔECD* mRNA in MDA-MB-453 cells 24 hours after transfection with smartpool siGENOME siRNA specific to *SRPK1* and *SRSF1* splice factors. Non-targeting siRNA is used as a calibrator to measure the fold difference in *HER2ΔECD* mRNA after transfection for 24 hours. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

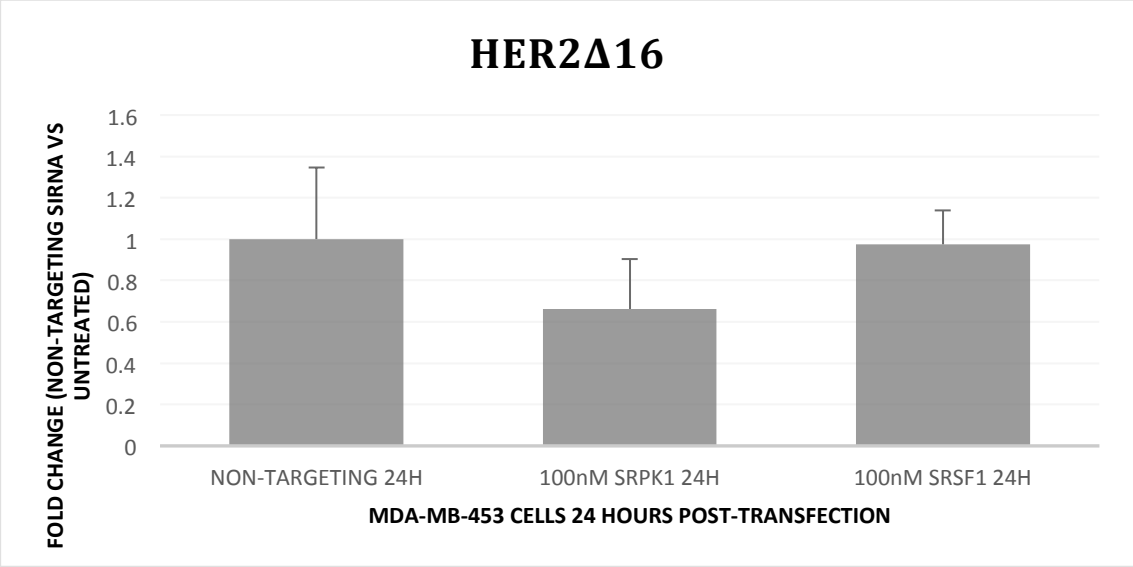


Figure 6.42: Effect of knockdown on *HER2*Δ16 mRNA in MDA-MB-453 cells 24 hours after transfection with smartpool siGENOME siRNA specific to *SRPK1* and *SRSF1* splice factors. Non-targeting siRNA is used as a calibrator to measure the fold difference in *HER2*Δ16 mRNA after transfection for 24 hours. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

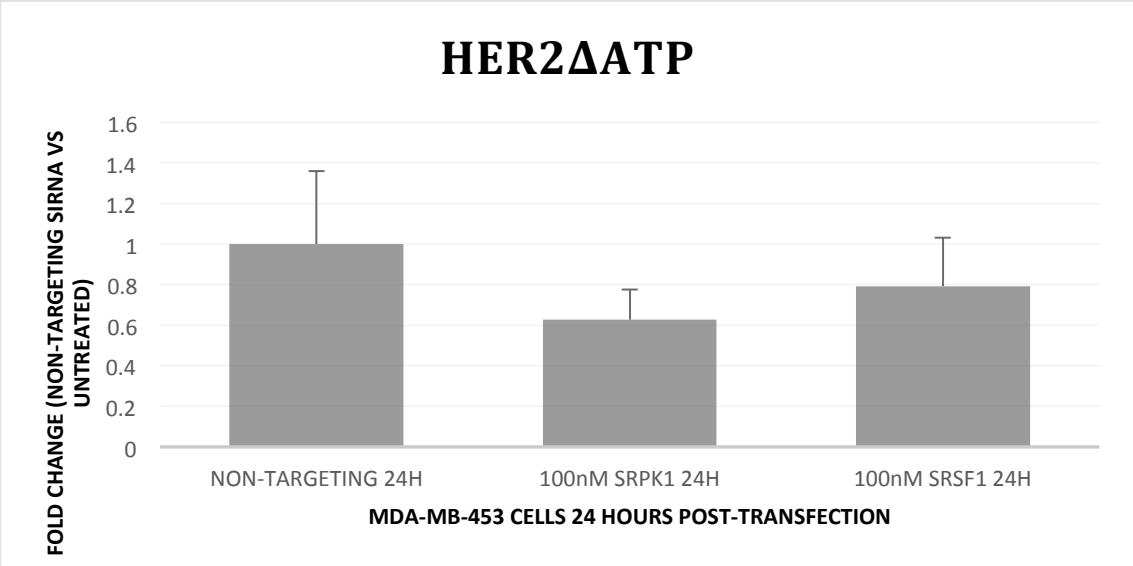


Figure 6.43: Effect of knockdown on *HER2*ΔATP mRNA in MDA-MB-453 cells 24 hours after transfection with smartpool siGENOME siRNA specific to *SRPK1* and *SRSF1* splice factors. Non-targeting siRNA is used as a calibrator to measure the fold difference in *HER2*ΔATP mRNA after transfection for 24 hours. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

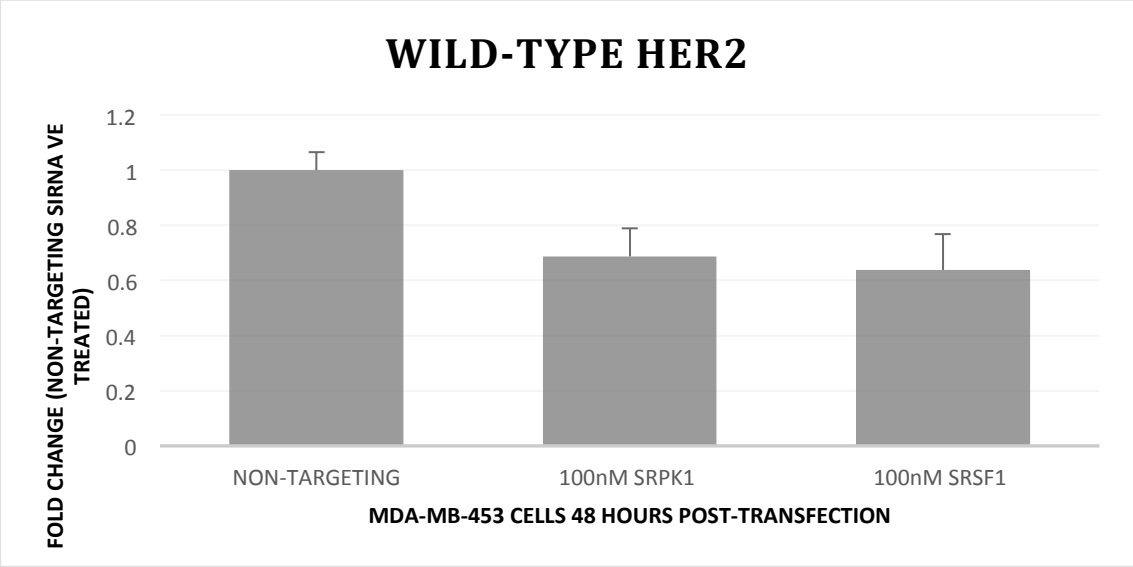


Figure 6.44: Effect of knockdown on wild-type *HER2* mRNA in MDA-MB-453 cells 48 hours after transfection with smartpool siGENOME siRNA specific to *SRPK1* and *SRSF1* splice factors. Non-targeting siRNA is used as a calibrator to measure the fold difference in wild-type *HER2* mRNA after transfection for 48 hours. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

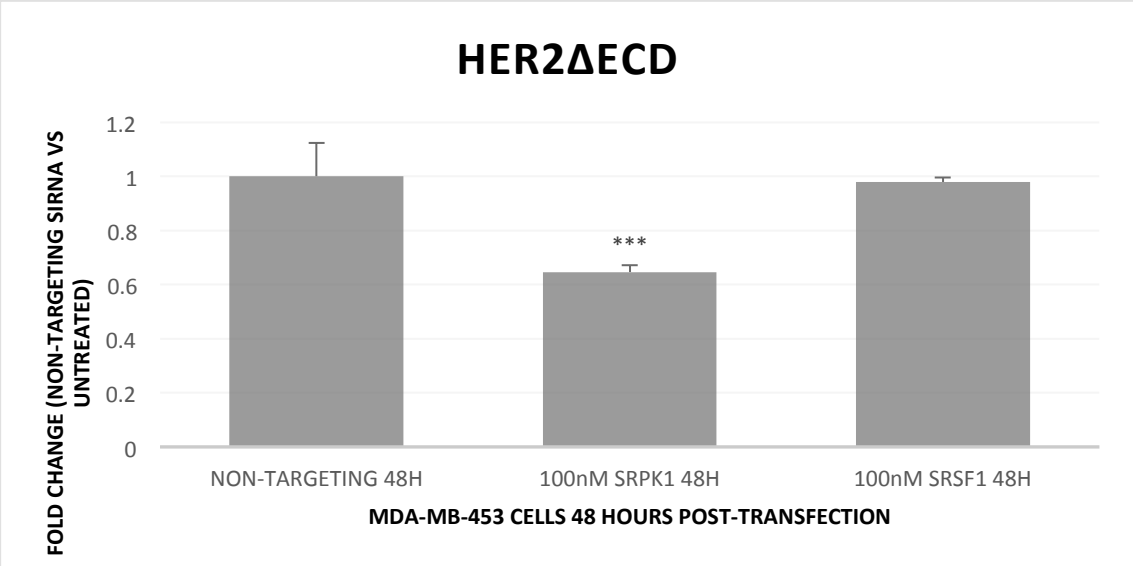


Figure 6.45: Effect of knockdown on *HER2ΔECD* mRNA in MDA-MB-453 cells 48 hours after transfection with smartpool siGENOME siRNA specific to *SRPK1* and *SRSF1* splice factors. Non-targeting siRNA is used as a calibrator to measure the fold difference in *HER2ΔECD* mRNA after transfection for 48 hours. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

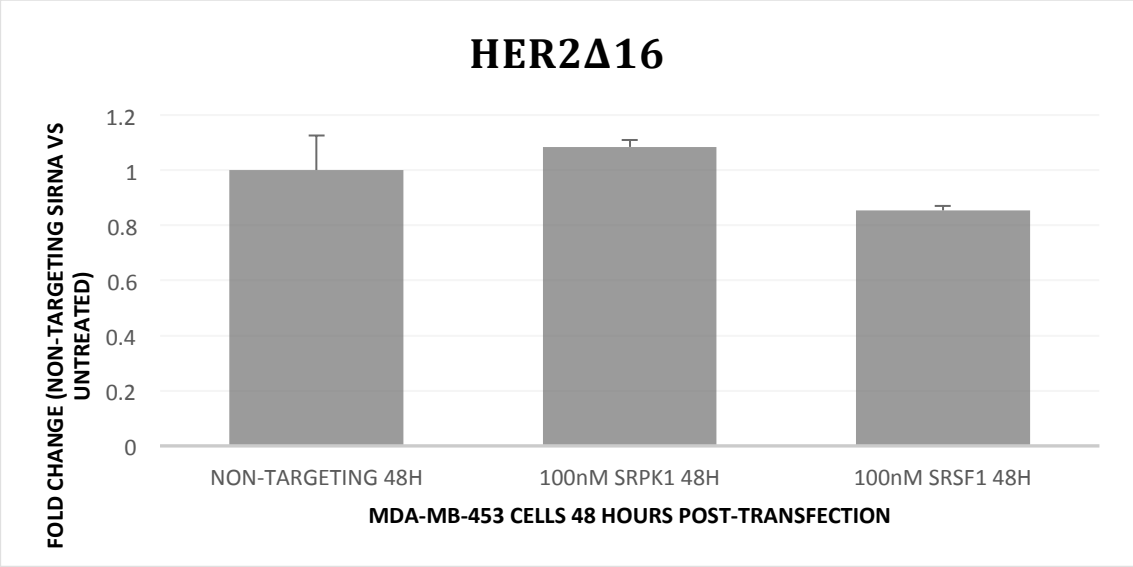


Figure 6.46: Effect of knockdown on *HER2*Δ16 mRNA in MDA-MB-453 cells 48 hours after transfection with smartpool siGENOME siRNA specific to *SRPK1* and *SRSF1* splice factors. Non-targeting siRNA is used as a calibrator to measure the fold difference in *HER2*Δ16 mRNA after transfection for 48 hours. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

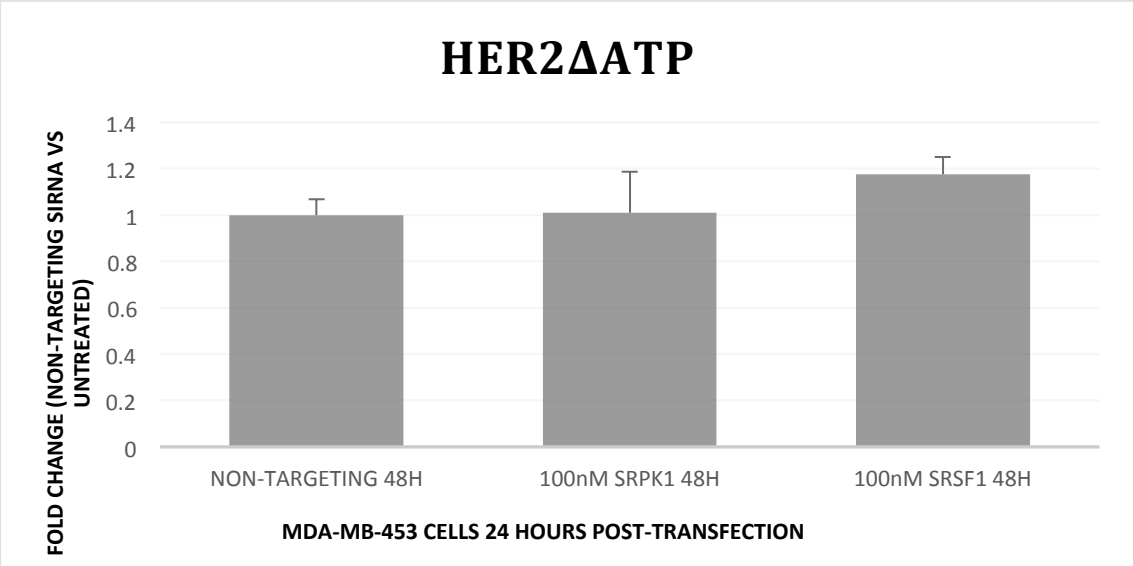


Figure 6.47: Effect of knockdown on *HER2*ΔATP mRNA in MDA-MB-453 cells 48 hours after transfection with smartpool siGENOME siRNA specific to *SRPK1* and *SRSF1* splice factors. Non-targeting siRNA is used as a calibrator to measure the fold difference in *HER2*ΔATP mRNA after transfection for 48 hours. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

6.3.3.3. Confirmation of *SRPK1* and *SRSF1* knockdown in SKBR3 cells

At both 24 and 48 hours post-transfection, a significant knockdown was observed for both *SRPK1* and *SRSF1* in SKBR3 cells. Using non-targeting siRNA as a calibrator, a 24 hour transfection produced a reduction in *SRPK1* mRNA of 4.1 fold (± 0.08; $p<0.0001$) (Figure 6.48), and a reduction in *SRSF1* mRNA of 4.8 fold (±0.34; $p<0.01$) (Figure 6.50), and a 48 hour transfection produced a reduction in *SRPK1* mRNA of 4.5 fold ($p<0.0001$) (Figure 6.49), and in *SRSF1* of 3.7 fold (±0.56; $p<0.0001$) (Figure 6.51).

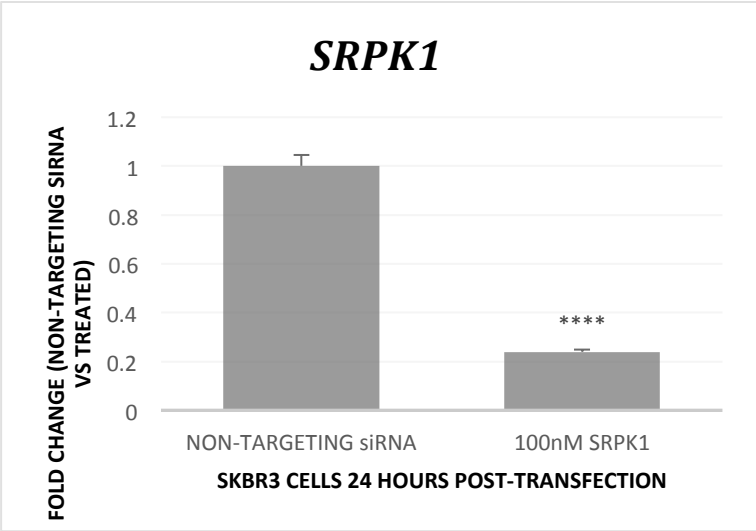


Figure 6.48: Knockdown of *SRPK1* mRNA in SKBR3 cell lines after transfection with *SRPK1* smartpool siGENOME siRNA; a mixture of four separate siRNAs supplied in a single tube. SKBR3 cells were transfected with 100nM of either a non-targeting siRNA or *SRPK1*-specific siRNA for 24 hours before RNA extraction, reverse transcription and qPCR analysis. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

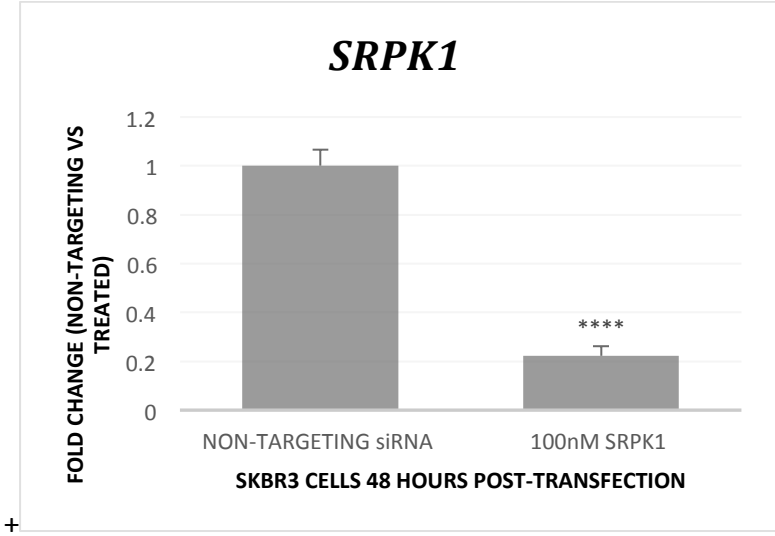


Figure 6.49: Knockdown of *SRPK1* mRNA in SKBR3 cell lines after transfection with *SRPK1* smartpool siGENOME siRNA; a mixture of four separate siRNAs supplied in a single tube. SKBR3 cells were transfected with 100nM of either a non-targeting siRNA or *SRPK1*-specific siRNA for 48 hours before RNA extraction, reverse transcription and qPCR analysis. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T Ct values.

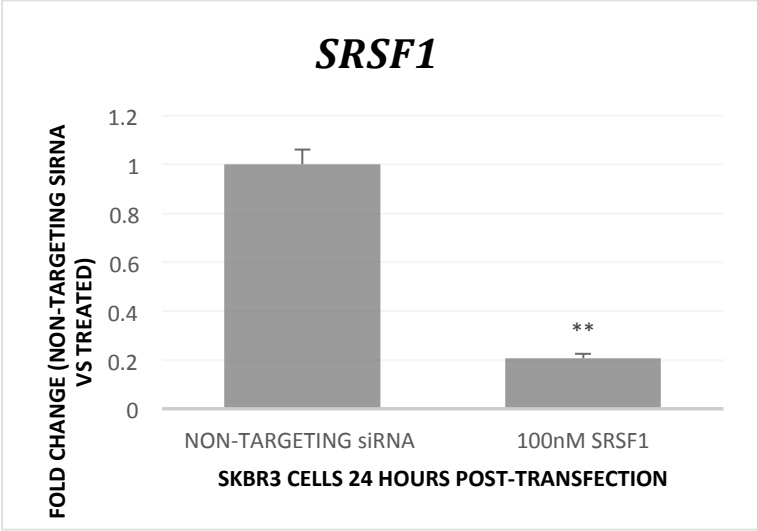


Figure 6.50: Knockdown of *SRSF1* mRNA in SKBR3 cell lines after transfection with *SRSF1* smartpool siGENOME siRNA; a mixture of four separate siRNAs supplied in a single tube. SKBR3 cells were transfected with 100nM of either a non-targeting siRNA or *SRSF1*-specific siRNA for 24 hours before RNA extraction, reverse transcription and qPCR analysis. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

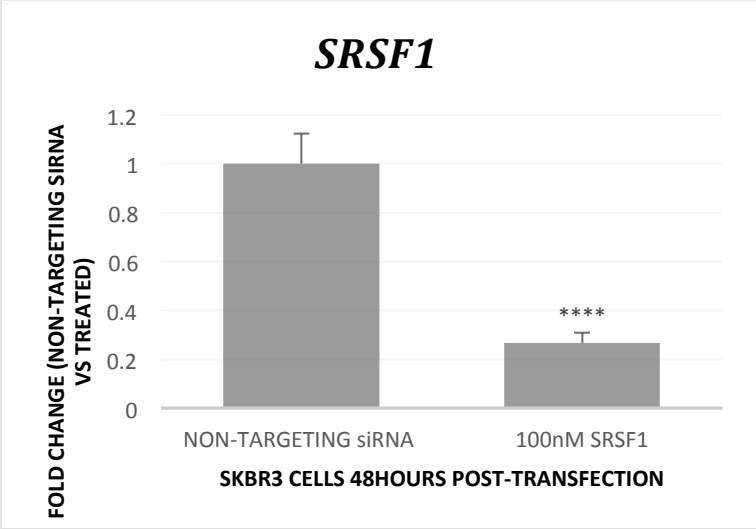


Figure 6.51: Knockdown of *SRSF1* mRNA in SKBR3 cell lines after transfection with *SRSF1* smartpool siGENOME siRNA; a mixture of four separate siRNAs supplied in a single tube. SKBR3 cells were transfected with 100nM of either a non-targeting siRNA or *SRSF1*-specific siRNA for 48 hours before RNA extraction, reverse transcription and qPCR analysis. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

6.3.3.4. Knockdown of *SRPK1* and SFSF1 affects the expression of *HER2* and *HER2* alternative splice variants in SKBR3 cells at mRNA level

Transfection of *SRPK1* and *SRSF1* siRNAs in SKBR3 cells shows a reduction in the expression of *HER2* and *HER2* splice variants. After 24 hours of transfection with *SRPK1*, SKBR3 cells showed a slight reduction in the expression of wild-type *HER2* (fold change = 1.02; p≥0.05) (Figure 6.52), *HER2ΔECD* (fold change = 1.16; p≥0.05) (Figure 6.53), *HER2*Δ16 (fold change = 1.13; p≥0.05) (Figure 6.54), and *HER2*ΔATP (fold change = 1.75; p≥0.05) (Figure 6.55). A slight reduction is observed in the expression of wild-type *HER2* (fold change = 1.47; p≥0.05) (Figure 6.52), *HER2ΔECD* (fold change = 1.26; p≥0.05) (Figure 6.53), and *HER2*ΔATP (fold change = 1.83; p≥0.05) (Figure 6.55) after 24 hour transfection with *SRSF1*. A more significant change is observed in the reduction of

expression of *HER2*Δ16 after transfection with *SRSF1* for 24 hours (fold change = 1.61; $p<0.05$) (Figure 6.54). Transfection of *SRPK1* and *SRSF1* siRNAs in SKBR3 cells shows a significant reduction in the expression of *HER2* and *HER2* splice variants after 48 hours. Transfection of *SRPK1* in SKBR3 cells showed a reduction in the expression of wild-type *HER2* (fold change = 1.81; $p<0.05$) (Figure 6.56), *HER2*Δ16 (fold change = 1.44; $p<0.05$) (Figure 6.58), and *HER2*ΔATP (fold change = 2.67; $p<0.05$) (Figure 6.59). At 48 hours post-transfection, there is no significant change in the expression of *HER2ΔECD* following *SRPK1* knockdown (fold change = 1.42; $p\geq0.05$) (Figure 6.57). A reduction is observed in the expression of wild-type *HER2* (fold change = 2.02; $p<0.05$) (Figure 6.56), *HER2*Δ16 (fold change = 1.9; $p<0.05$) (Figure 6.58), and *HER2*ΔATP (fold change = 3.44; $p<0.05$) (Figure 6.59) after 48 hour transfection with *SRSF1*. A less significant change is observed in the reduction of expression of *HER2*Δ16 after transfection with *SRSF1* and in the expression of *HER2ΔECD* after transfection for 48 hours with *SRPK1* (fold change = 1.61; $p\leq0.05$) (Figure 6.57).

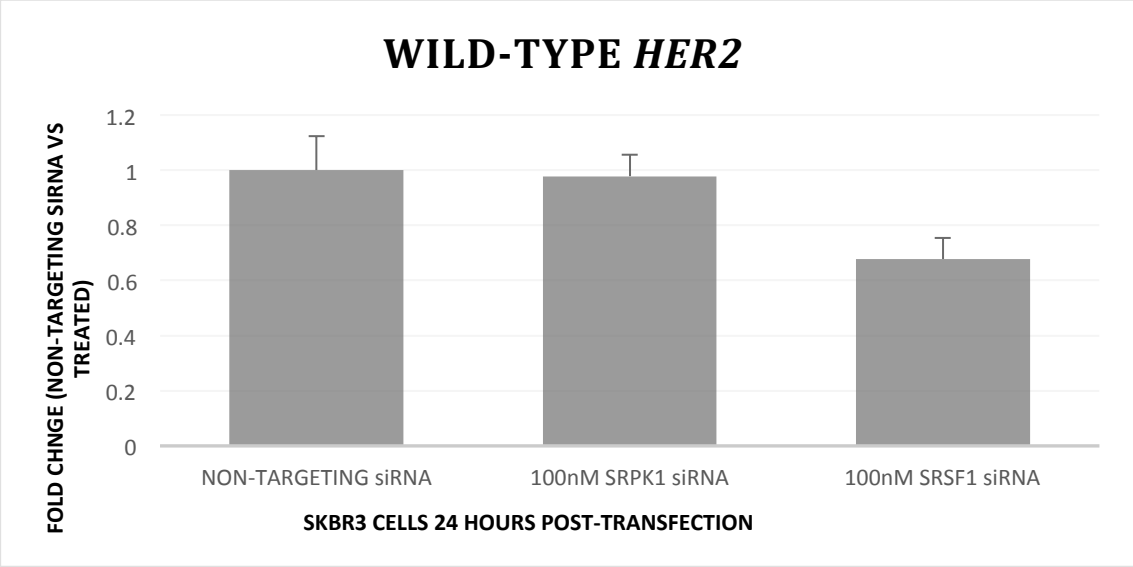


Figure 6.52: Effect of knockdown on wild-type *HER2* mRNA in SKBR3 cells 24 hours after transfection with smartpool siGENOME siRNA specific to *SRPK1* and *SRSF1* splice factors. Non-targeting siRNA is used as a calibrator to measure the fold difference in wild-type *HER2* mRNA after transfection for 24 hours. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

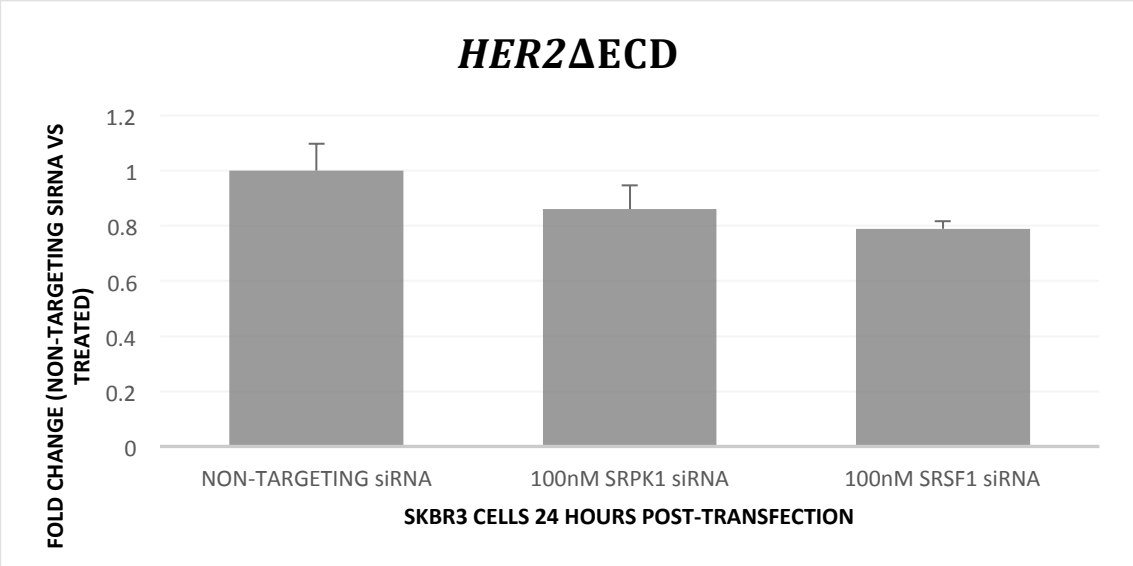


Figure 6.53: Effect of knockdown on *HER2ΔECD* mRNA in SKBR3 cells 24 hours after transfection with smartpool siGENOME siRNA specific to *SRPK1* and *SRSF1* splice factors. Non-targeting siRNA is used as a calibrator to measure the fold difference in *HER2ΔECD* mRNA after transfection for 24 hours. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

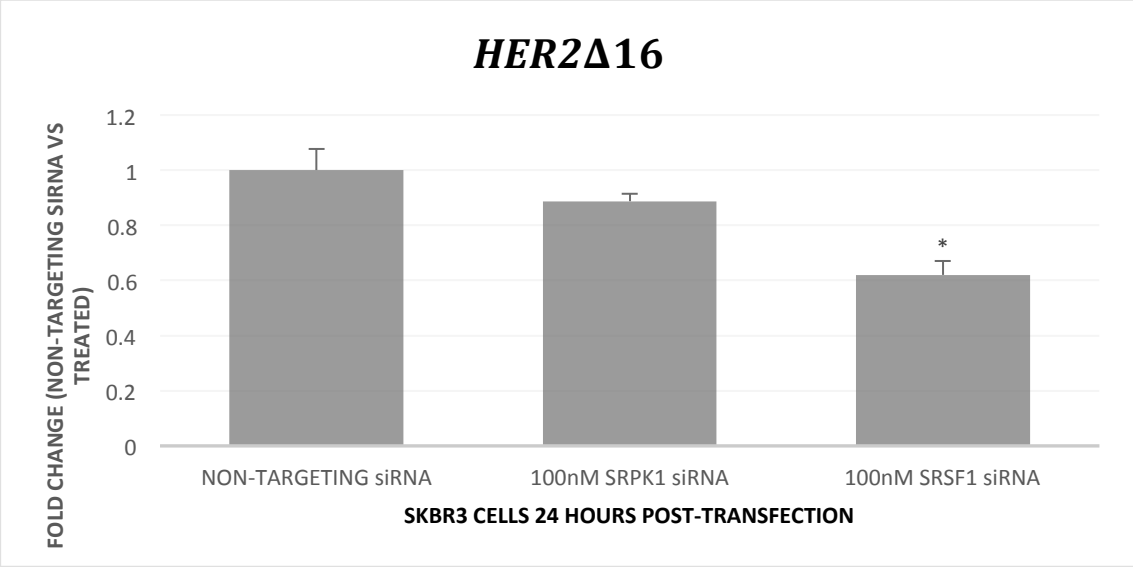


Figure 6.54: Effect of knockdown on *HER2*Δ16 mRNA in SKBR3 cells 24 hours after transfection with smartpool siGENOME siRNA specific to *SRPK1* and *SRSF1* splice factors. Non-targeting siRNA is used as a calibrator to measure the fold difference in *HER2*Δ16 mRNA after transfection for 24 hours. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

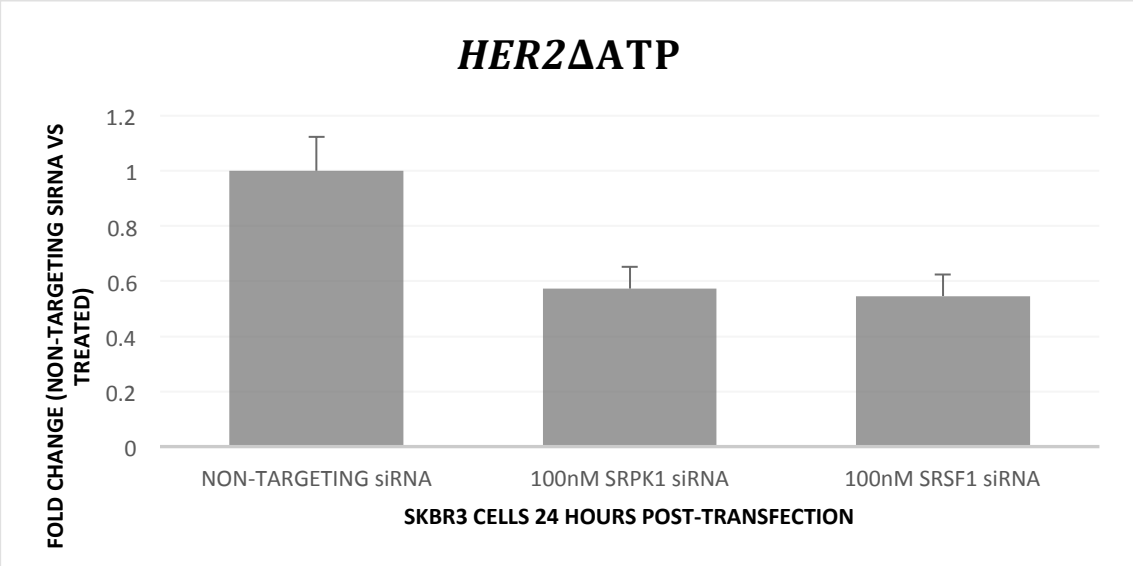


Figure 6.55: Effect of knockdown on *HER2*ΔATP mRNA in SKBR3 cells 24 hours after transfection with smartpool siGENOME siRNA specific to *SRPK1* and *SRSF1* splice factors. Non-targeting siRNA is used as a calibrator to measure the fold difference in *HER2*ΔATP mRNA after transfection for 24 hours. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

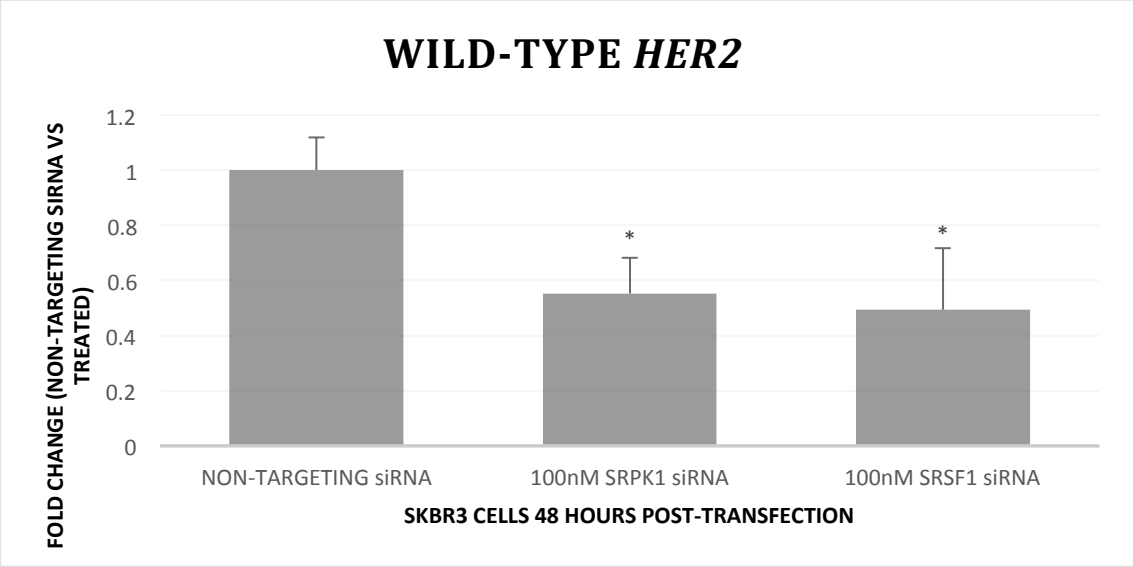


Figure 6.56: Effect of knockdown on wild-type *HER2* mRNA in SKBR3 cells 48 hours after transfection with smartpool siGENOME siRNA specific to *SRPK1* and *SRSF1* splice factors. Non-targeting siRNA is used as a calibrator to measure the fold difference in wild-type *HER2* mRNA after transfection for 48 hours. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

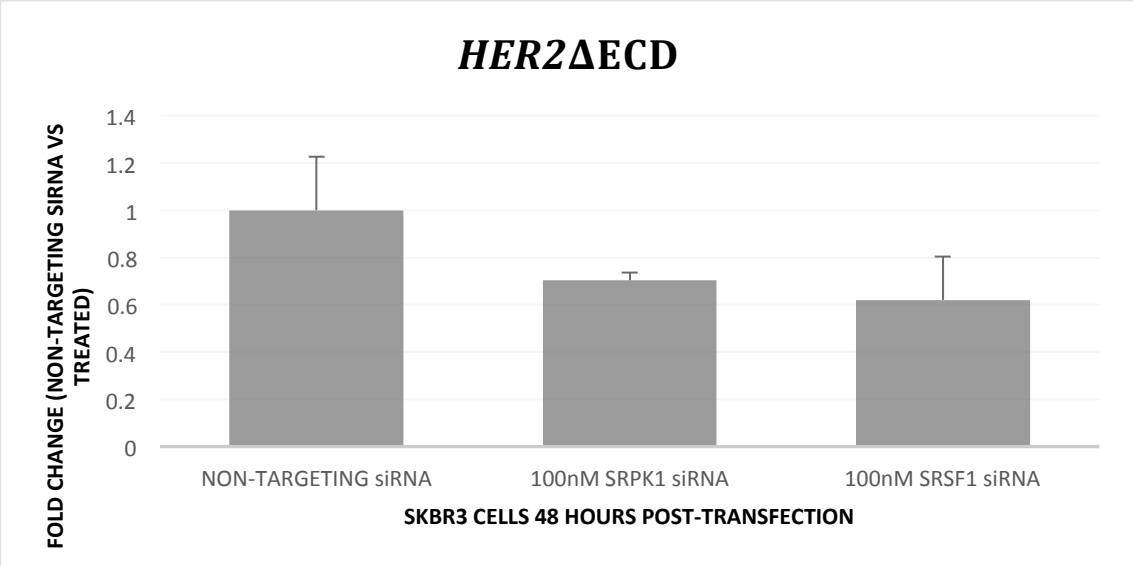


Figure 6.57: Effect of knockdown on *HER2ΔECD* mRNA in SKBR3 cells 48 hours after transfection with smartpool siGENOME siRNA specific to *SRPK1* and *SRSF1* splice factors. Non-targeting siRNA is used as a calibrator to measure the fold difference in *HER2ΔECD* mRNA after transfection for 48 hours. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

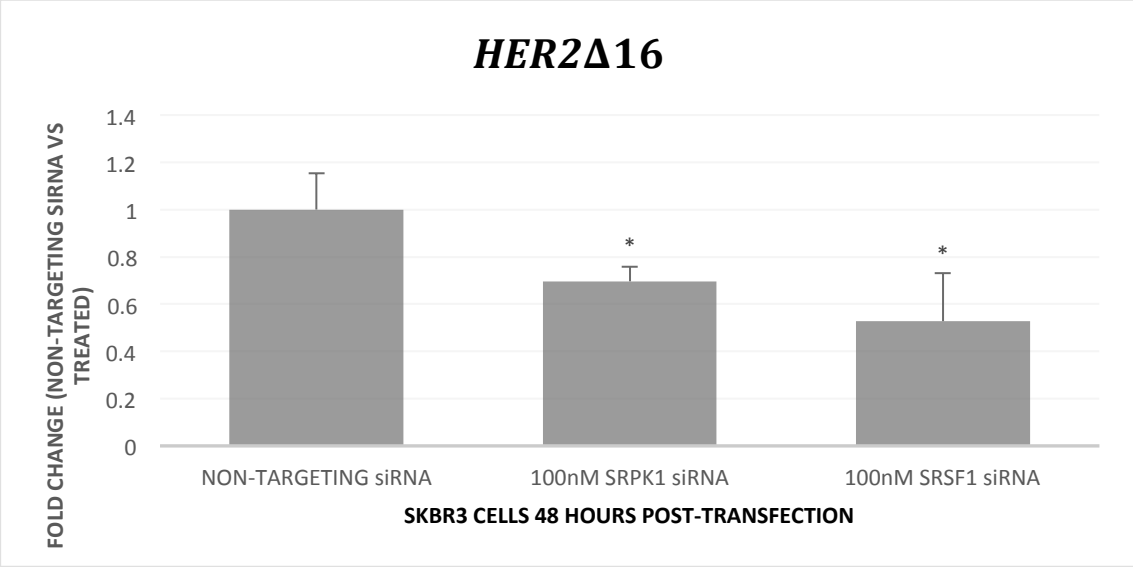


Figure 6.58: Effect of knockdown on *HER2*Δ16 mRNA in SKBR3 cells 48 hours after transfection with smartpool siGENOME siRNA specific to *SRPK1* and *SRSF1* splice factors. Non-targeting siRNA is used as a calibrator to measure the fold difference in *HER2*Δ16 mRNA after transfection for 48 hours. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

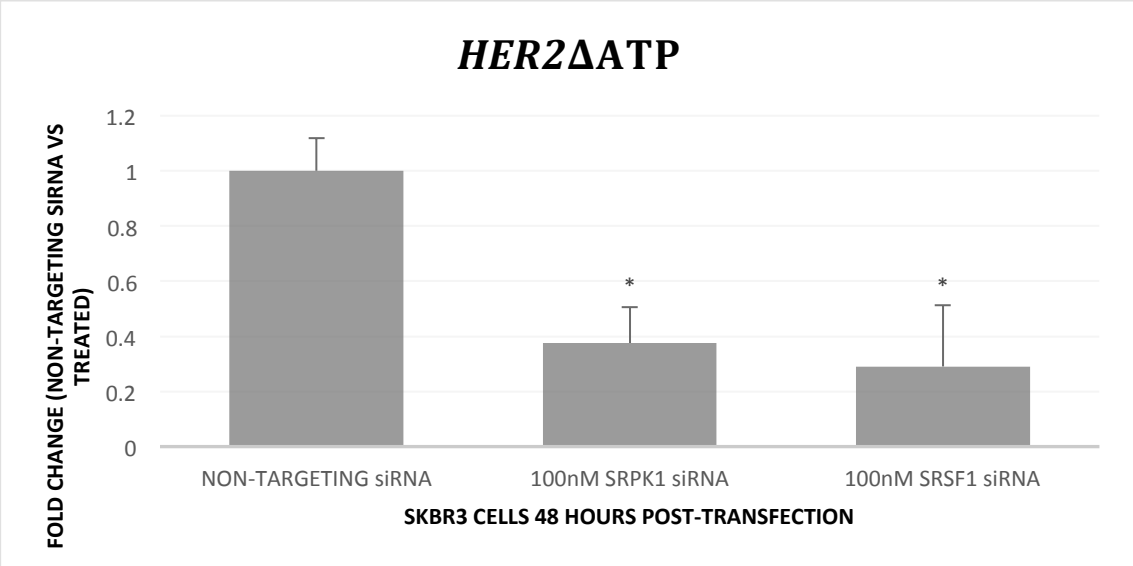


Figure 6.59: Effect of knockdown on *HER2*ΔATP mRNA in SKBR3 cells 48 hours after transfection with smartpool siGENOME siRNA specific to *SRPK1* and *SRSF1* splice factors. Non-targeting siRNA is used as a calibrator to measure the fold difference in *HER2*ΔATP mRNA after transfection for 48 hours. Each histogram bar is representative of an average of 9 samples consisting of three technical repeats of three biological replicates. Each biological replicate was run on a different passage of cells. Error bars represent the standard deviations of the C_T values.

6.4. Summary

The study of the regulation of splicing in *HER2* has not been well researched to date. With the identification of alternative splice variants of *HER2*, and the identification of splice factor binding motifs for *SRSF1* in chapter 4, the roles of certain potential regulatory factors was investigated. In summary:

- Inhibition of *SRPK1* via *SRPIN340* increases the expression of *HER2* and *HER2* alternative splice variants in MDA-MB-453 cells. This regulatory effect is not seen after treatment of SKBR3 and BT-20 cells with *SRPIN340*.
- Inhibition of CDC2-like kinase and DYRK1A via *TG003* and *INDY*, respectively, does not show any significant effects on the regulation of *HER2* and *HER2* alternative splice variants MDA-MB-453, SKBR3, and BT-20 cells.
- Induction of HIF1-α by Cobalt Chloride treatment has inhibitory effects on the expression of *HER2* and *HER2* alternative splice variants in SKBR3 cells.
- Knockdown of SR proteins *SRPK1* and *SRSF1* in MDA-MB-453 shows an inhibitory effect on the expression of *HER2* and *HER2* alternative splice variants. This inhibitory effect is not seen after *SRPK1* or *SRSF1* knockdown in SKBR3 cells.
- The MDA-MB-453 cell line is a good model for the investigation of *HER2* and *HER2* alternative splice variants in breast cancer.
- *SRSF1* and its phosphorylating protein *SRPK1* are potential regulators of *HER2* and *HER2* alternative splicing.

CHAPTER 7. DISCUSSION

Breast cancer is a heterogeneous disease, and with a wide range of histological and molecular subtypes, poses a challenge for diagnosis and prognosis in the clinical settings (Jackson *et al.*, 2013).

The *HER2* gene is an important biomarker in predicting the prognosis and outcome of patients with invasive breast cancer. *HER2* overexpression or gene amplification as determined by IHC or FISH is indicative of a more aggressive disease phenotype and low survival rates in patients with advanced breast cancer. It is well established that the *HER2* status in all breast cancer patients is crucial in identifying high risk groups and for the administration of appropriate therapies (Eroles *et al.*, 2012).

Trastuzumab, a recombinant monoclonal antibody targeted against the *HER2* transmembrane receptor binds with high affinity to the extracellular domain of *HER2*, providing substantial clinical benefits to patients with *HER2*-overexpressing advanced breast cancer, and improving overall survival when administered as an adjuvant in early breast cancer (Scaltriti *et al.,* 2007). *Trastuzumab* is currently the most important therapy in *HER2*-positive breast cancer. Several clinical trials have shown that *Trastuzumab* improves relapse-free survival and overall survival in patients with HER-2 overexpressing early breast cancer (Castiglioni *et al.,* 2006; Dean-Colomb & Esteva, 2008). However, due to findings that many patients with *HER2* positive breast cancer do not respond well to *Trastuzumab* therapy, recent studies have been aimed at elucidating and overcoming the mechanisms of *Trastuzumab* resistance. The

mechanisms underlying *Trastuzumab* resistance are not yet well understood, and elucidating the resistance of *HER2*-positive breast cancer to *Trastuzumab* therapy is critical in the diagnosis and treatment of patients whose cancers express this aggressive disease phenotype.

Several efforts have been made to overcome *Trastuzumab* resistance and to improve overall clinical outcome in targeted breast cancer treatment. These include the use of kinase antibodies such as *Pertuzumab* (Kolesta *et al.,* 2008). Pertuzumab binds the *HER2* extracellular domain in an area other than *Trastuzumab*, and interferes with the activation of *HER2*. An increasing body of evidence indicates that diverse adaptive and genetic changes within the transformed cell give rise to resistant phenotypes of the *HER2* protein, which allow them to survive in the presence of *HER2*-targeting antibodies (Freudenberg *et al.,* 2009). Khoury *et al.,* (2011) established that mutations in the juxtamembrane domain (*Trastuzumab* binding site) of *HER2* is a rare event, and would therefore not account for the relatively high frequency of *Trastuzumab* resistance encountered clinically (Khoury *et al.*, 2011).

This thesis establishes an overview of *HER2* expression in *HER2*-positive breast and ovarian cancer cell lines by immunohistochemistry. The results obtained correlate well with previous *HER2* studies on the cell lines used in this study (Wolff *et al.*, 2007). However, overall protein expression as determined by IHC may not be conclusive in determining a patient's true *HER2* status, and may therefore not be adequate as a single method of testing, to predict patients' response to *Trastuzumab* treatment. The antigen-antibody interaction of *Trastuzumab* and *HER2* is mediated by a portion of the

HER2 subdomain IV which consists of amino acid residues 510 – 643. The *Trastuzumab* antigen has affinity to amino acids 557 – 603. The coding regions of these amino acid residues are present in exons 17-18. The resistance of *HER2* breast cancer to *Trastuzumab* might be an indication that *Trastuzumab* binds differentially to different *HER2* splice isoforms. To be efficacious, the antibody must be able to bind to the *HER2* receptor, and consequently exact its growth inhibitory properties. This binding to the *HER2* receptor is entirely dependent on non-covalent interactions, therefore small changes in the antigen structure such as the loss of a single hydrogen bond can significantly reduce the strength of the antigen antibody interaction by over a 1000-fold (Khoury *et al.*, 2011).

This study hypothesized that factors other than the total copy number of *HER2* may be responsible for disease progression in *HER2*-overexpressing invasive breast cancers, and that currently unidentified alternative splice variants of *HER2* may give rise to protein isoforms with potent cellular functions, which can potentially alter the binding efficiency of *Trastuzumab* to the *HER2* receptor.

Bioinformatic analysis of the reference *HER2* gene revealed four binding sites on the cell surface receptor of the *HER2* protein which consists of two ligand binding sites, a signal transduction site, and a growth factor receptor domain; and an active binding site in the transmembrane domain which is responsible for tyrosine kinase phosphorylation, acting as an ‘on‘ or ‘off’ switch in a variety of cellular functions.

This thesis demonstrates for the first time, the expression of two novel *HER2* alternative splice variants, *HER2*ΔATP and *HER2ΔECD*, which, due to their very distinct structural and functional differences from the wild-type *HER2*, have the potential to alter the binding affinity of *Trastuzumab* to the *HER2* receptor, thereby contributing to the cellular changes that might affect patient response to *Trastuzumab* therapy.

Analysis of the sequences obtained from the smaller band produced by primer pair E15F/E19R revealed a loss of 42 base pairs downstream of the *HER2* mRNA sequence, indicative of an alternative 5' splice site in exon 18. This alternative 5' splice site resulted in the novel *HER2* isoform *HER2*ΔATP. The presence of a vertebrate splice site consensus sequence at the new splice boundaries suggest that this is not a splicing event mediated as a result of a mutation, but a true isoform resulting in an in-frame loss of 14 amino acids from the *HER2* protein.

Analysis of the resulting predicted protein revealed the deletions of this new *HER2* splice variant to be the loss of amino acids 724 – 737 in the tyrosine kinase domain of the *HER2* protein. The direct consequence of the loss of amino acids 724 – 737 is the loss of the entire ATP binding pocket of the *HER2* protein, which is represented by amino acids 726 – 734.

In the *HER2*ΔATP isoform, the predicted *HER2* extracellular domain remains unchanged. Since *HER2* dimerization occurs on subdomain II of the *HER2* extracellular domain, *HER2*ΔATP would still be capable of dimerization; however, the loss of amino acids in the tyrosine kinase region may inhibit phosphorylation and subsequent

activation of downstream signalling pathways. The activation of the ligand-binding domain of the *HER2* receptor triggers conformational changes within the *HER2* cytoplasmic domain, and the resulting phosphorylation of ATP to tyrosine residues on target substrates can only occur when the crucial key loops within the kinase domain are appropriately positioned (Telesco & Radhakrishnan, 2009). These functional loops form the molecular regulatory mechanisms in *HER2* phosphorylation; amino acid residues 884 – 850 form the catalytic loop (C-loop), which is crucial in the transfer of phosphoryl groups. Amino acid residues 761-775 and 727 – 732 form the αC helix and the nucleotide binding loop (N-loop), respectively, and these two loops are responsible for the coordination of the ATP substrate tyrosine. Amino acid residues 863 – 884 form the activation loop (A loop), which coordinates the activation of the tyrosine kinase by regulating accessibility of the target substrate to the C-loop. It is predicted that the *HER2*ΔATP splice variant translates a protein which is incapable of phosphorylation, as a portion of the nucleotide binding loop is lost in translation.

Analysis of the sequences obtained from the smaller band produced by primer pairs E12F/E15R revealed a loss of 133 nucleotides corresponding to the loss of the entire exon 13. Bioinformatic analysis of this novel alternative splice variant revealed a cassette exon in exon 13. Further analysis using exPASY translate tool revealed a truncated 645 amino acid protein which lacks amino acids 1-610 of the wild-type *HER2* isoform, but has conserved active binding sites in the transmembrane domain of the *HER2* protein.

Previous studies have demonstrated a *HER2* receptor termed p95*HER2* similar in structure to the *HER2ΔECD* alternative splice variant. P95*HER2* is an amino-terminally truncated carboxyl-terminal fragment (CTF) of *HER2* with kinase activity that lacks the *Trastuzumab* binding site, which is frequently found in *HER2* overexpressing breast cancer cell lines and tumours (Scaltriti *et al.*, 2007; Sperinde *et al.*, 2010; Sasso *et al.*, 2011). The p95*HER2* isoform is often cited as a contributing cause of *Trastuzumab* resistance. Studies have demonstrated that p95*HER2*-expressing tumours may be sensitive to the effects of *Lapatinib*, a small-molecule tyrosine kinase inhibitor which inhibits *HER2* phosphorylation by binding to the intracellular kinase domain of *HER2* (Awada, Bozovic-Spasojevic & Chow, 2012). The generation of p95*HER2* has been speculated to be by several mechanisms; the proteolytic shedding of the *HER2* extracellular domain at a site proximal to the transmembrane domain; alternative initiation of translation at the mRNA encoding *HER2* from methionine 611 located before the transmembrane domain, or alternative initiation of translation at the mRNA encoding *HER2* from methionine 687 located after the transmembrane domain (Scaltriti *et al.*, 2007; Sasso *et al.*, 2011). Interestingly, an mRNA precursor of the p95*HER2* has not been discovered to date.

This thesis predicts for the first time, a novel alternative splice variant of *HER2*, the *HER2ΔECD* splice variant, potentially corresponding to p95*HER2* protein.

In a previous study, Sperinde *et al*., (2010) demonstrated a novel antibody that measures p95*HER2 in vitro*. Due to the *HER2ΔECD* splice variant not being previously discovered, a method of detecting p95*HER2* in *HER2* mRNA has not been previously

demonstrated. In this thesis, for the first time, a method of detecting *HER2ΔECD* in genetic material has been demonstrated using RT-PCR priming, as well as a double-dye probe (Taqman) assay to detect the relative expression of *HER2ΔECD*, a predicted precursor of p95*HER2*.

Due to the significant findings leading to the discovery of novel *HER2* splice variants in this study, the expression of *HER2* and *HER2* alternative splice variants was investigated in a small cohort of human samples, to determine whether the splice variants identified are only tumour-specific. It has been shown here that despite the challenges of reversal of cross-linking in FFPE tissue samples, RNA can be successfully obtained from FFPE samples and for RT-PCR amplification of *HER2* and *HER2* splice variants. RNA from FFPE sample is known to be highly degraded, and results in RNA fragments approximately 200 nucleotides in length (Abramovitz *et al.*, 2008). It is reported here that *HER2* fragments with amplicon sizes of ~400 nucleotides in length and above were detected in the FFPE samples. Of the fourteen samples in the FFPE cohort, *HER2*ΔECD was detected in three samples, *HER2*ΔATP was detected in one sample, and *HER2*Δ16 was detected in six samples. Although the expression of *HER2*Δ16 has been previously identified in human samples (Castiglioni *et al.*, 2006), it is reported here for the first time in human tissue samples preserved by formalin fixation. P95*HER2*, which is predicted in this study to be the protein encoded by the novel *HER2* splice variant *HER2*ΔECD, has been previously identified in human samples as a protein. This *HER2* isoform is being identified in human tissue mRNA. In addition, the *HER2*ΔATP splice variant which corresponds with the loss of the ATP binding site of the *HER2* protein has also been detected here in FFPE samples.

The regulation of *HER2* and *HER2* splice variants was also investigated in cell line models. The role of splice factor kinases was investigated by the use of protein kinase inhibitors *SRPIN340*, *TG003* and *INDY* in *HER2* 3+ SKBR3, *HER2* 2+ MDA-MB-453, and *HER2*- BT-20 cell lines. Of all three cell lines, only MDA-MB-453 showed any changes in the expression of *HER2* and *HER2* alternative splice variants. The inhibition of *SRPK1* by the use of splice factor kinase *SRPIN340* greatly increased the expression of *HER2* and *HER2* alternative splice variants. Bioinformatic analysis of splice factor motifs in chapter 4 revealed potential binding sites for the SR kinase *SRSF1* (*ASF/SF2*), a splice factor potent regulated by the protein kinase *SRPK1*. *SRPIN340* is a selective inhibitor of *SRPK1*, and following *SRPIN340* treatment of MDA-MB-453 cells, levels of *HER2* were significantly elevated in the cells. Interestingly, direct siRNA knockdown of *SRPK1* and *SRSF1* and *SRSF1* were shown to have inhibitory effects on *HER2* and *HER2* alternative splice variants in MDA-MB-453 cells. Little is known about the regulation of *HER2* expression, nor the splice factors directly involved in *HER2* splicing. The up regulation of *HER2* following *SRPK1* inhibition via *SRPIN340*, and the down regulation of *HER2* following direct knockdown of *SRPK1* and *SRSF1* in MDA-MB-453 cells present a novel prediction that *SRPK1* and *SRSF1* may be potent regulators of *HER2* splicing.

Hypoxia Inducible Factor (HIF1-α), a potent physiologic marker for hypoxia in the cells, has been known to act as a tumour suppressor in certain cancers (Chiavarina *et al.*, 2010). The induction of HIF1-α in SKBR3 cells showed significantly reduced levels of *HER2* and *HER2* splice variants after as short as 24 hours following treatment with hypoxia mimetic factor $CoCl_2$. This finding confirms that HIF1-α may play also a crucial

role in driving down cancer cells, and reports for the first time that HIF1-α may be a positive factor in the regulation of *HER2* expression and splicing.

There are areas in this research which may benefit from further investigation. For example, there is not much information available on the exact protein sequences that *HER2* antibodies bind to. With findings of new alternative splice variants which are potentially capable of altering the conformation of the *HER2* protein with serious clinical implications, targeted treatment should focus on detecting more than one *HER2* protein, as a combination of different *HER2* isoforms may contribute in giving patients a *HER2* profile, thereby making it easier to determine what treatment will be beneficial to individual patients. For example, retrospective research has found that p95*HER2* is expressed in a large number of *HER2* positive breast cancers, and co expression of p95*HER2* with the wild-type *HER2* has been found to give rise to *Trastuzumab* resistant *HER2* tumours. These studies were carried out in retrospect, and thereforE did not directly benefit the patients at the time of diagnosis. Findings of this type would inform the use of other therapies like *Lapatinib* in combination with *trasuzumab* in clinical trials, with the aim of potentially inhibiting *HER2* at the extracellular level at the same time as the intracellular.

The discovery of *HER2* splice variants in breast tissues in this study leads to the recommendation that a larger cohort be tested, to give more significant information on the expression of *HER2* alternative splice variants in women who suffer from the disease.

It is also suggested that further work be carried out on the regulation of expression of *HER2* alternative splice variants, particularly the expression of *HER2* splice variants in non-*HER2* positive cell line models, to further the investigations into the role of these splice factors in breast cancer, and the role of SR proteins and splice factor kinases in the regulation of splicing in *HER2*. The wild-type *HER2* is often coexpressed with the known *HER2* splice variants, it may be beneficial to isolate the wild-type *HER2* in cell line models, and to investigate whether resistance to *Trastuzumab* may due to de novo resistance of *HER2*, or due to the coexpression of *HER2* splice variants which may have intrinsic resistance to *Trastuzumab*.

In conclusion this thesis demonstrates for the first time two novel alternative splice variants of *HER2*, and also has predicted potential regulators of splicing in *HER2* and *HER2* alternative splice variants, including the *HER2*Δ16 variant. The detection of these splice variants in human samples indicates a potential clinical significance, which can be established through further research and replication of these findings. The wild-type *HER2* and the *HER2* alternative splice variants appear to show similar changes in expression levels following treatments that have been administered to investigate their effects in cell line models. It might therefore be beneficial, for example, to look into the expression of HER2ΔATP *in vivo*. HER2ΔATP, being a potentially less active *HER2* isoform, may intrinsically inhibit phosphorylation with a higher potency than antibodies being administered. The findings of this study put alternative splicing of *HER2* at a focal point, in a hypothesis that the effects of alternative splicing in invasive breast cancer may be the continuing cause of resistance to existing therapies and

consequently disease progression in patients with *HER2* overexpressing or gene amplified invasive breast cancer.

REFERENCES

Abramovitz, M., Ordanic-Kodani, M., Wang, Y., Li, Z., Catzavelos, C., Bouzyk, M., Sledge, G.W., Moreno, C.S. & Leyland-Jones, B. (2008) Optimization of RNA extraction from FFPE tissues for expression profiling in the DASL assay. [online]. *BioTechniques*. 44 (3), pp. 417–423. [Accessed 23 October 2012].

Agus, D.B., Akita, R.W., Fox, W.D., Lewis, G.D., Higgins, B., Pisacane, P.I., Lofgren, J.A., Tindell, C., Evans, D.P., Maiese, K., Scher, H.I. & Sliwkowski, M.X. (2002) Targeting ligand-activated ErbB2 signaling inhibits breast and prostate tumor growth. [online]. *Cancer cell*. 2 (2), pp. 127–137. [Accessed 13 January 2014].

Aigner, A., Juhl, H., Malerczyk, C., Tkybusch, A., Benz, C.C. & Czubayko, F. (2001) Expression of a truncated 100 kDa HER2 splice variant acts as an endogenous inhibitor of tumour cell proliferation. [online]. *Oncogene*. 20 (17), pp. 2101–2111. [Accessed 16 October 2012].

Allemand, E., Guil, S., Myers, M., Moscat, J., Cáceres, J.F. & Krainer, A.R. (2005) Regulation of heterogenous nuclear ribonucleoprotein A1 transport by phosphorylation in cells stressed by osmotic shock. [online]. *Proceedings of the National Academy of Sciences of the United States of America*. 102 (10), pp. 3605–3610. [Accessed 15 January 2014].

Alvarez, R. & Hortobagyi, G. (2013) Dual human epidermal growth factor receptor 2 blockade for the treatment of HER2-positive breast cancer. [online]. *Breast cancer Tokyo Japan*. 20 (2), pp. 103–110.

Andersen, C.L., Jensen, J.L. & Ørntoft, T.F. (2004) Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. [online]. *Cancer research*. 64 (15), pp. 5245–5250. [Accessed 31 January 2013].

Anon (2013) Breast Cancer Facts and Figures 2013-2014. *American Cancer Society*. [online]. Available from: http://www.cancer.org/acs/groups/content/@research/documents/document/acspc-042725.pdfdoi:10.14496/dia.6104460190.12.

Awada, A., Bozovic-Spasojevic, I. & Chow, L. (2012) New therapies in HER2-positive breast cancer: a major step towards a cure of the disease? [online]. *Cancer treatment reviews*. 38 (5), pp. 494–504. [Accessed 20 September 2013].

Baek, C.-M., Jeon, S.-H., Jang, J.-J., Lee, B.S. & Lee, J.-H. (2004) Transforming variant of Met receptor confers serum independence and anti-apoptotic property and could be involved in the mouse thymic lymphomagenesis. [online]. *Experimental & molecular medicine*. 36 (4), pp. 283–291. [Accessed 17 January 2014].

Bang, Y.-J., Van Cutsem, E., Feyereislova, A., Chung, H.C., Shen, L., Sawaki, A., Lordick, F., Ohtsu, A., Omuro, Y., Satoh, T., Aprile, G., Kulikov, E., Hill, J., Lehle, M., et al. (2010) Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. [online]. *Lancet*. 376 (9742), pp. 687–697. [Accessed 13 January 2014].

Bargmann, C.I., Hung, M.C. & Weinberg, R.A. (1986) Multiple independent activations of the neu oncogene by a point mutation altering the transmembrane domain of p185. [online]. *Cell*. 45 (5), pp. 649–657. [Accessed 6 March 2014].

Bartlett, J.M., Going, J.J., Mallon, E.A., Watters, A.D., Reeves, J.R., Stanton, P., Richmond, J., Donald, B., Ferrier, R. & Cooke, T.G. (2001) Evaluating HER2 amplification and overexpression in breast cancer. [online]. *The Journal of pathology*. 195 (4), pp. 422–428. [Accessed 16 October 2012].

Baselga, J., Cortés, J., Kim, S.-B., Im, S.-A., Hegg, R., Im, Y.-H., Roman, L., Pedrini, J.L., Pienkowski, T., Knott, A., Clark, E., Benyunes, M.C., Ross, G. & Swain, S.M. (2012) Pertuzumab plus trastuzumab plus docetaxel for metastatic breast cancer. [online]. *The New England journal of medicine*. 366 (2), pp. 109–119. [Accessed 13 January 2014].

Baselga, J. & Swain, S.M. (2009) Novel anticancer targets: revisiting ERBB2 and discovering ERBB3. [online]. *Nature reviews. Cancer*. 9 (7), pp. 463–475. [Accessed 31 January 2013].

Bedard, P.L., de Azambuja, E. & Cardoso, F. (2009) Beyond trastuzumab: overcoming resistance to targeted HER-2 therapy in breast cancer. [online]. *Current cancer drug targets*. 9 (2), pp. 148–162. [Accessed 14 January 2014].

Blok, E.J., Kuppen, P.J., Van Leeuwen, J.E. & Sier, C.F. (2013) Cytoplasmic Overexpression of HER2: a Key Factor in Colorectal Cancer. [online]. *Clinical Medicine Insights. Oncology*. 7pp. 41–51. [Accessed 6 January 2014].

Boyle, P. (2012) Triple-negative breast cancer: epidemiological considerations and recommendations. [online]. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO*. 23 Suppl 6 (Supplement 6), pp. vi7–vi12. [Accessed 19 November 2012].

Bustin, S. a, Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J. & Wittwer, C.T. (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. [online]. *Clinical chemistry*. 55 (4), pp. 611–622. [Accessed 26 October 2012].

Castiglioni, F., Tagliabue, E., Campiglio, M., Pupa, S.M., Balsari, A. & Ménard, S. (2006) Role of exon-16-deleted HER2 in breast carcinomas. [online]. *Endocrine-related cancer*. 13 (1), pp. 221–232. [Accessed 16 October 2012].

Chiavarina, B., Whitaker-Menezes, D., Migneco, G., Martinez-Outschoorn, U.E., Pavlides, S., Howell, A., Tanowitz, H.B., Casimiro, M.C., Wang, C., Pestell, R.G., Grieshaber, P., Caro, J., Sotgia, F. & Lisanti, M.P. (2010) HIF1-alpha functions as a tumor promoter in cancer associated fibroblasts, and as a tumor suppressor in breast cancer cells: Autophagy drives compartment-specific oncogenesis. [online]. *Cell cycle (Georgetown, Tex.)*. 9 (17), pp. 3534–3551. [Accessed 28 May 2014].

Cho, H.-S. & Leahy, D.J. (2002) Structure of the extracellular region of HER3 reveals an interdomain tether. [online]. *Science (New York, N.Y.)*. 297 (5585), pp. 1330–1333. [Accessed 10 April 2014].

Ciocca, D.R., Gago, F.E., Fanelli, M.A. & Calderwood, S.K. (2006) Co-expression of steroid receptors (estrogen receptor alpha and/or progesterone receptors) and Her-2/neu: Clinical implications. [online]. *The Journal of steroid biochemistry and molecular biology*. 102 (1-5), pp. 32–40. [Accessed 16 October 2012].

Collesi, C., Santoro, M.M., Gaudino, G. & Comoglio, P.M. (1996) A splicing variant of the RON transcript induces constitutive tyrosine kinase activity and an invasive phenotype. [online]. *Molecular and cellular biology*. 16 (10), pp. 5518–5526. [Accessed 17 January 2014].

Coughlin, C.M., Johnston, D.S., Strahs, A., Burczynski, M.E., Bacus, S., Hill, J., Feingold, J.M., Zacharchuk, C. & Berkenblit, A. (2010) Approaches and limitations of phosphatidylinositol-3-kinase pathway activation status as a predictive biomarker in the clinical development of targeted therapy. [online]. *Breast cancer research and treatment*. 124 (1), pp. 1–11. [Accessed 14 January 2014].

Courcet, J.-B., Faivre, L., Malzac, P., Masurel-Paulet, A., Lopez, E., Callier, P., Lambert, L., Lemesle, M., Thevenon, J., Gigot, N., Duplomb, L., Ragon, C., Marle, N., Mosca-Boidron, A.-L., et al. (2012) The DYRK1A gene is a cause of syndromic intellectual disability with severe microcephaly and epilepsy. [online]. *Journal of medical genetics*. 49 (12), pp. 731–736. [Accessed 28 May 2014].

Coussens, L., Yang-Feng, T.L., Liao, Y.C., Chen, E., Gray, A., McGrath, J., Seeburg, P.H., Libermann, T.A., Schlessinger, J. & Francke, U. (1985) Tyrosine kinase receptor with extensive homology to EGF receptor shares chromosomal location with neu oncogene. [online]. *Science (New York, N.Y.)*. 230 (4730), pp. 1132–1139. [Accessed 4 March 2014].

David, C.J. & Manley, J.L. (2008) The search for alternative splicing regulators: new approaches offer a path to a splicing code. [online]. *Genes & development*. 22 (3), pp. 279–285. [Accessed 1 February 2013].

Dean-Colomb, W. & Esteva, F.J. (2008) Her2-positive breast cancer: herceptin and beyond. [online]. *European journal of cancer (Oxford, England : 1990)*. 44 (18), pp. 2806–2812. [Accessed 16 October 2012].

Deng, S., Hirschberg, A., Worzfeld, T., Penachioni, J.Y., Korostylev, A., Swiercz, J.M., Vodrazka, P., Mauti, O., Stoeckli, E.T., Tamagnone, L., Offermanns, S. & Kuner, R. (2007) Plexin-B2, but not Plexin-B1, critically modulates neuronal migration and patterning of the developing nervous system in vivo. [online]. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 27 (23), pp. 6333–6347. [Accessed 26 April 2014].

Derveaux, S., Vandesompele, J. & Hellemans, J. (2010) How to do successful gene expression analysis using real-time PCR. [online]. *Methods (San Diego, Calif.)*. 50 (4), pp. 227–230. [Accessed 31 January 2013].

Doherty, J.K., Bond, C., Jardim, A., Adelman, J.P. & Clinton, G.M. (1999a) The HER-2/neu receptor tyrosine kinase gene encodes a secreted autoinhibitor. [online]. *Proceedings of the National Academy of Sciences of the United States of America*. 96 (19), pp. 10869–10874. [Accessed 16 October 2012].

Doherty, J.K., Bond, C.T., Hua, W., Adelman, J.P. & Clinton, G.M. (1999b) An alternative HER-2/neu transcript of 8 kb has an extended 3'UTR and displays increased stability in SKOV-3 ovarian carcinoma cells. [online]. *Gynecologic oncology*. 74 (3), pp. 408–415. [Accessed 31 January 2013].

Domcke, S., Sinha, R., Levine, D.A., Sander, C. & Schultz, N. (2013) Evaluating cell lines as tumour models by comparison of genomic profiles. [online]. *Nature communications*. 4pp. 2126. [Accessed 2 May 2014].

Downward, J., Yarden, Y., Mayes, E., Scrace, G., Totty, N., Stockwell, P., Ullrich, A., Schlessinger, J. & Waterfield, M.D. (n.d.) Close similarity of epidermal growth factor receptor and v-erb-B oncogene protein sequences. [online]. *Nature*. 307 (5951), pp. 521–527. [Accessed 6 March 2014].

Eliassen, A.H., Missmer, S.A., Tworoger, S.S., Spiegelman, D., Barbieri, R.L., Dowsett, M. & Hankinson, S.E. (2006) Endogenous steroid hormone concentrations and risk of breast cancer among premenopausal women. *Journal of the National Cancer Institute*. 98 (19), pp. 1406–1415.

Elliot, D. & Ladomery, M. (2011) *Molecular Biology of RNA*. Oxford University Press.

Emde, A., Köstler, W.J. & Yarden, Y. (2012) Therapeutic strategies and mechanisms of tumorigenesis of HER2-overexpressing breast cancer. [online]. *Critical reviews in oncology/hematology*. 84 Suppl 1pp. e49–57. [Accessed 7 October 2013].

Emir, H.D.V.. (2011) *RT-qPCR Guidelines : From designing to publishing your data qPCR 2011*.

Eroles, P., Bosch, A., Pérez-Fidalgo, J.A. & Lluch, A. (2012) Molecular biology in breast cancer: intrinsic subtypes and signaling pathways. [online]. *Cancer treatment reviews*. 38 (6), pp. 698–707. [Accessed 22 September 2013].

Fackenthal, J.D. & Godley, L. a (2008) Aberrant RNA splicing and its functional consequences in cancer cells. [online]. *Disease models & mechanisms*. 1 (1), pp. 37–42. [Accessed 1 February 2013].

Franklin, M.C., Carey, K.D., Vajdos, F.F., Leahy, D.J., de Vos, A.M. & Sliwkowski, M.X. (2004) Insights into ErbB signaling from the structure of the ErbB2-pertuzumab complex. [online]. *Cancer cell*. 5 (4), pp. 317–328. [Accessed 13 January 2014].

Freudenberg, J.A., Wang, Q., Katsumata, M., Drebin, J., Nagatomo, I. & Greene, M.I. (2009) The role of HER2 in early breast cancer metastasis and the origins of resistance to HER2-targeted therapies. [online]. *Experimental and molecular pathology*. 87 (1), pp. 1–11. [Accessed 16 October 2012].

Fuse, N. (2011) Relation of HER2 status and prognosis in gastric cancer patients. *Gan to kagaku ryoho Cancer chemotherapy*. 38 (7), pp. 1073–1078.

Garcia-Blanco, M.A. (2005) Methods for the study of alternative splicing. [online]. *Methods (San Diego, Calif.)*. 37 (4), pp. 289–291. [Accessed 16 October 2012].

Garrett, T.P., McKern, N.M., Lou, M., Frenkel, M.J., Bentley, J.D., Lovrecz, G.O., Elleman, T.C., Cosgrove, L.J. & Ward, C.W. (1998) Crystal structure of the first three domains of the type-1 insulin-like growth factor receptor. [online]. *Nature*. 394 (6691), pp. 395–399. [Accessed 26 April 2014].

Gebhardt, F., Zänker, K.S. & Brandt, B. (1998) Differential expression of alternatively spliced c-erbB-2 mRNA in primary tumors, lymph node metastases, and bone marrow micrometastases from breast cancer patients. [online]. *Biochemical and biophysical research communications*. 247 (2), pp. 319–323. [Accessed 16 October 2012].

Geyer, C.E., Forster, J., Lindquist, D., Chan, S., Romieu, C.G., Pienkowski, T., Jagiello-Gruszfeld, A., Crown, J., Chan, A., Kaufman, B., Skarlos, D., Campone, M., Davidson, N., Berger, M., et al. (2006) Lapatinib plus capecitabine for HER2-positive advanced breast cancer. [online]. *The New England journal of medicine*. 355 (26), pp. 2733–2743. [Accessed 13 January 2014].

Goddard, K.A.B., Weinmann, S., Richert-Boe, K., Chen, C., Bulkley, J. & Wax, C. (2012) HER2 evaluation and its impact on breast cancer treatment decisions. [online]. *Public health genomics*. 15 (1), pp. 1–10.

Gómez-Martin, C., Garralda, E., Echarri, M.J., Ballesteros, A., Arcediano, A., Rodríguez-Peralto, J.L., Hidalgo, M. & López-Ríos, F. (2012) HER2/neu testing for anti-HER2-based therapies in patients with unresectable and/or metastatic gastric cancer. [online]. *Journal of clinical pathology*. 65 (8), pp. 751–757. [Accessed 8 May 2014].

Goren, A., Kim, E., Amit, M., Vaknin, K., Kfir, N., Ram, O. & Ast, G. (2010) Overlapping splicing regulatory motifs—combinatorial effects on splicing [online]. *Nucleic Acids Research*. 38 (10), pp. 3318–3327.

Gui, J.F., Tronchère, H., Chandler, S.D. & Fu, X.D. (1994) Purification and characterization of a kinase specific for the serine- and arginine-rich pre-mRNA splicing factors. [online]. *Proceedings of the National Academy of Sciences of the United States of America*. 91 (23), pp. 10824–10828. [Accessed 28 May 2014].

Hanks, S.K. & Hunter, T. (1995) Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. [online]. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*. 9 (8), pp. 576–596. [Accessed 26 April 2014].

Hanks, S.K. & Quinn, A.M. (1991) Protein kinase catalytic domain sequence database: identification of conserved features of primary structure and classification of family members. [online]. *Methods in enzymology*. 200pp. 38–62. [Accessed 10 April 2014].

Hayat, M.A. (2002) *Microscopy, Immunohistochemistry, and Antigen Retrieval Methods: For Light and Electron Microscopy*. 2nd edition. Springer US.

Hayes, G.M., Carrigan, P.E. & Miller, L.J. (2007) Serine-arginine protein kinase 1 overexpression is associated with tumorigenic imbalance in mitogen-activated protein kinase pathways in breast, colonic, and pancreatic carcinomas. [online]. *Cancer research*. 67 (5), pp. 2072–2080. [Accessed 28 May 2014].

Hillig, T., Thode, J., Breinholt, M.F., Franzmann, M.-B., Pedersen, C., Lund, F., Mygind, H., Sölétormos, G. & Rudnicki, M. (2012) Assessing HER2 amplification by IHC, FISH, and real-time polymerase chain reaction analysis (real-time PCR) following LCM in formalin-fixed paraffin embedded tissue from 40 women with ovarian cancer. [online]. *APMIS : acta pathologica, microbiologica, et immunologica Scandinavica*. 120 (12), pp. 1000–1007. [Accessed 26 October 2012].

Hunter, T. & Plowman, G.D. (1997) The protein kinases of budding yeast: six score and more. [online]. *Trends in biochemical sciences*. 22 (1), pp. 18–22. [Accessed 26 April 2014].

Hynes, N.E. & Lane, H.A. (2005) ERBB receptors and cancer: the complexity of targeted inhibitors. [online]. *Nature reviews. Cancer*. 5 (5), pp. 341–354. [Accessed 16 October 2012].

Hynes, N.E. & MacDonald, G. (2009) ErbB receptors and signaling pathways in cancer. [online]. *Current opinion in cell biology*. 21 (2), pp. 177–184. [Accessed 6 January 2014].

Jackson, C., Browell, D., Gautrey, H. & Tyson-Capper, A. (2013) Clinical Significance of HER-2 Splice Variants in Breast Cancer Progression and Drug Resistance. [online]. *International journal of cell biology*. 2013pp. 973584.

Jacquemier, J., Spyratos, F., Esterni, B., Mozziconacci, M.-J., Antoine, M., Arnould, L., Lizard, S., Bertheau, P., Lehmann-Che, J., Fournier, C.B., Krieger, S., Bibeau, F.,

Lamy, P.-J., Chenard, M.P., et al. (2013) SISH/CISH or qPCR as alternative techniques to FISH for determination of HER2 amplification status on breast tumors core needle biopsies: a multicenter experience based on 840 cases. [online]. *BMC cancer*. 13pp. 351. [Accessed 8 May 2014].

Jain, P., Karthikeyan, C., Moorthy, N.S.H.N., Waiker, D.K., Jain, A.K. & Trivedi, P. (2014) Human CDC2-like kinase 1 (CLK1): a novel target for Alzheimer's disease. [online]. *Current drug targets*. 15 (5), pp. 539–550. [Accessed 28 May 2014].

Jones, K.L. & Buzdar, A.U. (2009) Evolving novel anti-HER2 strategies. [online]. *The lancet oncology*. 10 (12), pp. 1179–1187. [Accessed 30 September 2013].

Joshi, M. & Deshpande, J.. (2011) *POLYMERASE CHAIN REACTION : METHODS , PRINCIPLES AND*. 2 (1), .

Justman, Q.A. & Clinton, G.M. (2002) Herstatin, an autoinhibitor of the human epidermal growth factor receptor 2 tyrosine kinase, modulates epidermal growth factor signaling pathways resulting in growth arrest. [online]. *The Journal of biological chemistry*. 277 (23), pp. 20618–20624. [Accessed 16 January 2014].

Kang, S.H., Kang, K.W., Kim, K.-H., Kwon, B., Kim, S.-K., Lee, H.-Y., Kong, S.-Y., Lee, E.S., Jang, S.-G. & Yoo, B.C. (2008) Upregulated HSP27 in human breast cancer cells reduces Herceptin susceptibility by increasing Her2 protein stability. [online]. *BMC cancer*. 8pp. 286. [Accessed 16 October 2012].

Ke, Q. & Costa, M. (2006) Hypoxia-inducible factor-1 (HIF-1). *Molecular pharmacology*. 70 (5), pp. 1469–1480.

Kerbel, R.S. (2009) Issues regarding improving the impact of antiangiogenic drugs for the treatment of breast cancer. [online]. *Breast (Edinburgh, Scotland)*. 18 Suppl 3pp. S41–7. [Accessed 14 January 2014].

Key, T.J., Appleby, P.N., Reeves, G.K., Roddam, A.W., Helzlsouer, K.J., Alberg, A.J., Rollison, D.E., Dorgan, J.F., Brinton, L.A., Overvad, K., Kaaks, R., Trichopoulou, A., Clavel-Chapelon, F., Panico, S., et al. (2011) Circulating sex hormones and breast cancer risk factors in postmenopausal women: reanalysis of 13 studies. *British journal of cancer*. 105 (5), pp. 709–722.

Khoury, T., Mojica, W., Hicks, D., Starostik, P., Ademuyiwa, F., Janarthanan, B. & Cheney, R.T. (2011) ERBB2 juxtamembrane domain (trastuzumab binding site) gene mutation is a rare event in invasive breast cancers overexpressing the ERBB2 gene. [online]. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc*. 24 (8), pp. 1055–1059. [Accessed 29 May 2014].

Kim, E., Goren, A. & Ast, G. (2008) Insights into the connection between cancer and alternative splicing. [online]. *Trends in genetics : TIG*. 24 (1), pp. 7–10. [Accessed 1 February 2013].

Koletsa, T., Kostopoulos, I., Charalambous, E., Christoforidou, B., Nenopoulou, E. & Kotoula, V. (2008) A Splice Variant of HER2 Corresponding to Herstatin Is Expressed in the Noncancerous Breast and in Breast Carcinomas1 [online]. *Neoplasia (New York, N.Y.)*. 10 (7), pp. 687–696. [Accessed 16 October 2012].

Kong, S.-Y., Nam, B.-H., Lee, K.S., Kwon, Y., Lee, E.S., Seong, M.-W., Lee, D.H. & Ro, J. (2006) Predicting tissue HER2 status using serum HER2 levels in patients with metastatic breast cancer. [online]. *Clinical Chemistry*. 52 (8), pp. 1510–1515.

Kwong, K.Y. & Hung, M.C. (1998) A novel splice variant of HER2 with increased transformation activity. [online]. *Molecular carcinogenesis*. 23 (2), pp. 62–68. [Accessed 16 October 2012].

Ladomery, M.R., Harper, S.J. & Bates, D.O. (2007) Alternative splicing in angiogenesis: the vascular endothelial growth factor paradigm. [online]. *Cancer letters*. 249 (2), pp. 133–142. [Accessed 16 October 2012].

Lekas, A., Lazaris, A.C., Deliveliotis, C., Chrisofos, M., Zoubouli, C., Lapas, D., Papathomas, T., Fokitis, I. & Nakopoulou, L. (n.d.) The expression of hypoxia-inducible factor-1alpha (HIF-1alpha) and angiogenesis markers in hyperplastic and malignant prostate tissue. *Anticancer research*. 26 (4B), pp. 2989–2993.

Li, X., Park, W.J., Pyeritz, R.E. & Jabs, E.W. (1995) Effect on splicing of a silent FGFR2 mutation in Crouzon syndrome. [online]. *Nature genetics*. 9 (3), pp. 232–233. [Accessed 17 January 2014].

Li, Y.-Q., Liu, F.-F., Zhang, X.-M., Guo, X.-J., Ren, M.-J. & Fu, L. (2013) Tumor secretion of CCL22 activates intratumoral Treg infiltration and is independent prognostic predictor of breast cancer. [online]. *PloS one*. 8 (10), pp. e76379. [Accessed 9 January 2014].

Li, Z., Mei, Y., Liu, X. & Zhou, M. (2007) Neuregulin-1 only induces trans-phosphorylation between ErbB receptor heterodimer partners. [online]. *Cellular signalling*. 19 (3), pp. 466–471. [Accessed 26 April 2014].

Lukas, J., Gao, D.Q., Keshmeshian, M., Wen, W.H., Tsao-Wei, D., Rosenberg, S. & Press, M.F. (2001) Alternative and aberrant messenger RNA splicing of the mdm2 oncogene in invasive breast cancer. [online]. *Cancer research*. 61 (7), pp. 3212–3219. [Accessed 26 October 2012].

Marchini, C., Gabrielli, F., Iezzi, M., Zenobi, S., Montani, M., Pietrella, L., Kalogris, C., Rossini, A., Ciravolo, V., Castagnoli, L., Tagliabue, E., Pupa, S.M., Musiani, P., Monaci, P., et al. (2011) The human splice variant Δ16HER2 induces rapid tumor onset in a reporter transgenic mouse. Maria G Castro (ed.) [online]. *PloS one*. 6 (4), pp. e18727. [Accessed 26 October 2012].

Meijer, S.L., Wesseling, J., Smit, V.T., Nederlof, P.M., Hooijer, G.K.J., Ruijter, H., Arends, J.W., Kliffen, M., van Gorp, J.M., Sterk, L. & van de Vijver, M.J. (2011) HER2 gene

amplification in patients with breast cancer with equivocal IHC results. [online]. *Journal of clinical pathology*. 64 (12), pp. 1069–1072. [Accessed 6 January 2014].

Minami, M., Matsumoto, S. & Horiuchi, H. (2010) Cardiovascular Side-Effects of Modern Cancer Therapy [online]. *Circulation Journal*. 74 (9), pp. 1779–1786. [Accessed 14 January 2014].

Mita, K., Zhang, Z., Ando, Y., Toyama, T., Hamaguchi, M., Kobayashi, S., Hayashi, S., Fujii, Y., Iwase, H. & Yamashita, H. (2007) Prognostic significance of insulin-like growth factor binding protein (IGFBP)-4 and IGFBP-5 expression in breast cancer. [online]. *Japanese journal of clinical oncology*. 37 (8), pp. 575–582. [Accessed 31 January 2013].

Mitra, D., Brumlik, M.J., Okamgba, S.U., Zhu, Y., Duplessis, T.T., Parvani, J.G., Lesko, S.M., Brogi, E. & Jones, F.E. (2009) An oncogenic isoform of HER2 associated with locally disseminated breast cancer and trastuzumab resistance. [online]. *Molecular cancer therapeutics*. 8 (8), pp. 2152–2162. [Accessed 16 October 2012].

Modrek, B. & Lee, C. (2002) A genomic view of alternative splicing. [online]. *Nature genetics*. 30 (1), pp. 13–19. [Accessed 16 October 2012].

Moelans, C.B., de Weger, R.A., Van der Wall, E. & van Diest, P.J. (2011) Current technologies for HER2 testing in breast cancer. [online]. *Critical reviews in oncology/hematology*. 80 (3), pp. 380–392. [Accessed 22 October 2012].

Nahta, R., Hung, M.-C. & Esteva, F.J. (2004) The HER-2-targeting antibodies trastuzumab and pertuzumab synergistically inhibit the survival of breast cancer cells. [online]. *Cancer research*. 64 (7), pp. 2343–2346. [Accessed 13 January 2014].

Nahta, R. & O'Regan, R.M. (2010) Evolving strategies for overcoming resistance to HER2-directed therapy: targeting the PI3K/Akt/mTOR pathway. [online]. *Clinical breast cancer*. 10 Suppl 3 (November), pp. S72–8. [Accessed 7 October 2013].

Neve, R.M., Chin, K., Fridlyand, J., Yeh, J., Baehner, F.L., Fevr, T., Clark, L., Bayani, N., Coppe, J.-P., Tong, F., Speed, T., Spellman, P.T., DeVries, S., Lapuk, A., et al. (2006) A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. [online]. *Cancer cell*. 10 (6), pp. 515–527. [Accessed 16 October 2012].

Normanno, N., De Luca, A., Bianco, C., Strizzi, L., Mancino, M., Maiello, M.R., Carotenuto, A., De Feo, G., Caponigro, F. & Salomon, D.S. (2006) Epidermal growth factor receptor (EGFR) signaling in cancer. [online]. *Gene*. 366 (1), pp. 2–16. [Accessed 17 December 2013].

Nowak, D.G., Woolard, J., Amin, E.M., Konopatskaya, O., Saleem, M.A., Churchill, A.J., Ladomery, M.R., Harper, S.J. & Bates, D.O. (2008) Expression of pro- and anti-angiogenic isoforms of VEGF is differentially regulated by splicing and growth

factors. [online]. *Journal of cell science*. 121 (Pt 20), pp. 3487–3495. [Accessed 31 January 2013].

Nuti, M., Bellati, F., Visconti, V., Napoletano, C., Domenici, L., Caccetta, J., Zizzari, I.G., Ruscito, I., Rahimi, H., Benedetti-Panici, P. & Rughetti, A. (2011) Immune Effects of Trastuzumab [online]. *Journal of Cancer*. 2pp. 317–323.

Al Okail, M.S. (2010) Cobalt chloride, a chemical inducer of hypoxia-inducible factor-1α in U251 human glioblastoma cell line [online]. *Journal of Saudi Chemical Society*. 14 (2), pp. 197–201. [Accessed 25 March 2015].

Omenn, G.S., Yocum, A.K. & Menon, R. (2010) Alternative splice variants, a new class of protein cancer biomarker candidates: findings in pancreatic cancer and breast cancer with systems biology implications. [online]. *Disease Markers*. 28 (4), pp. 241–251.

Pal, S., Gupta, R. & Davuluri, R. V (2012) Alternative transcription and alternative splicing in cancer. [online]. *Pharmacology & therapeutics*. [Accessed 22 October 2012].

Parkin, D.M., Boyd, L. & Walker, L.C. (2011) 16. The fraction of cancer attributable to lifestyle and environmental factors in the UK in 2010. *British journal of cancer*. 105 Supplpp. S77–81.

Perou, C.M. (2011) Molecular stratification of triple-negative breast cancers. [online]. *The oncologist*. 16 Suppl 1pp. 61–70. [Accessed 17 May 2015].

Piccart-Gebhart, M.J., Procter, M., Leyland-Jones, B., Goldhirsch, A., Untch, M., Smith, I., Gianni, L., Baselga, J., Bell, R., Jackisch, C., Cameron, D., Dowsett, M., Barrios, C.H., Steger, G., et al. (2005) Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. [online]. *The New England journal of medicine*. 353 (16), pp. 1659–1672. [Accessed 13 January 2014].

Pivot, X., Romieu, G., Debled, M., Pierga, J.-Y., Kerbrat, P., Bachelot, T., Lortholary, A., Espié, M., Fumoleau, P., Serin, D., Jacquin, J.-P., Jouannaud, C., Rios, M., Abadie-Lacourtoisie, S., et al. (2013) 6 months versus 12 months of adjuvant trastuzumab for patients with HER2-positive early breast cancer (PHARE): a randomised phase 3 trial. [online]. *The lancet oncology*. 14 (8), pp. 741–748. [Accessed 11 May 2014].

Pohlmann, P.R., Mayer, I. a & Mernaugh, R. (2009) Resistance to Trastuzumab in Breast Cancer. [online]. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 15 (24), pp. 7479–7491. [Accessed 5 October 2012].

Raz, E., Schejter, E.D. & Shilo, B.Z. (1991) Interallelic complementation among DER/flb alleles: implications for the mechanism of signal transduction by receptor-tyrosine kinases. [online]. *Genetics*. 129 (1), pp. 191–201. [Accessed 26 April 2014].

Rhodes, A., Sarson, J., Assam, E.E., Dean, S.J.R., Cribb, E.C. & Parker, A. (2010) The reliability of rabbit monoclonal antibodies in the immunohistochemical assessment of estrogen receptors, progesterone receptors, and HER2 in human breast carcinomas. [online]. *American journal of clinical pathology*. 134 (4), pp. 621–632. [Accessed 16 October 2012].

Rodgers, J.T., Haas, W., Gygi, S.P. & Puigserver, P. (2010) Cdc2-like kinase 2 is an insulin-regulated suppressor of hepatic gluconeogenesis. [online]. *Cell metabolism*. 11 (1), pp. 23–34. [Accessed 28 May 2014].

Romanowski, T., Markiewicz, A., Bednarz, N. & Bielawski, K.P. (2007) [Housekeeping genes as a reference in quantitative real-time RT-PCR]. [online]. *Postępy higieny i medycyny doświadczalnej (Online)*. 61pp. 500–510. [Accessed 25 November 2013].

Romond, E.H., Perez, E.A., Bryant, J., Suman, V.J., Geyer, C.E., Davidson, N.E., Tan-Chiu, E., Martino, S., Paik, S., Kaufman, P.A., Swain, S.M., Pisansky, T.M., Fehrenbacher, L., Kutteh, L.A., et al. (2005) Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. [online]. *The New England journal of medicine*. 353 (16), pp. 1673–1684. [Accessed 10 January 2014].

Saal, L.H., Gruvberger-Saal, S.K., Persson, C., Lövgren, K., Jumppanen, M., Staaf, J., Jönsson, G., Pires, M.M., Maurer, M., Holm, K., Koujak, S., Subramaniyam, S., Vallon-Christersson, J., Olsson, H., et al. (2008) Recurrent gross mutations of the PTEN tumor suppressor gene in breast cancers with deficient DSB repair. [online]. *Nature genetics*. 40 (1), pp. 102–107. [Accessed 24 October 2012].

Sahlberg, K.K., Hongisto, V., Edgren, H., Mäkelä, R., Hellström, K., Due, E.U., Moen Vollan, H.K., Sahlberg, N., Wolf, M., Børresen-Dale, A.-L., Perälä, M. & Kallioniemi, O. (2013) The HER2 amplicon includes several genes required for the growth and survival of HER2 positive breast cancer cells. [online]. *Molecular oncology*. 7 (3), pp. 392–401. [Accessed 9 January 2014].

Saini, K.S., Azim, H.A., Metzger-Filho, O., Loi, S., Sotiriou, C., de Azambuja, E. & Piccart, M. (2011) Beyond trastuzumab: new treatment options for HER2-positive breast cancer. [online]. *Breast (Edinburgh, Scotland)*. 20 Suppl 3pp. S20–7. [Accessed 22 October 2012].

Salesse, S., Dylla, S.J. & Verfaillie, C.M. (2004) p210BCR/ABL-induced alteration of pre-mRNA splicing in primary human CD34+ hematopoietic progenitor cells. [online]. *Leukemia*. 18 (4), pp. 727–733. [Accessed 28 May 2014].

Sasso, M., Bianchi, F., Ciravolo, V., Tagliabue, E. & Campiglio, M. (2011) HER2 splice variants and their relevance in breast cancer [online]. *Journal of Nucleic Acids Investigation*. 2 (1), . [Accessed 29 May 2014].

Scaltriti, M., Eichhorn, P.J., Cortés, J., Prudkin, L., Aura, C., Jiménez, J., Chandarlapaty, S., Serra, V., Prat, A., Ibrahim, Y.H., Guzmán, M., Gili, M., Rodríguez, O., Rodríguez,

S., et al. (2011) Cyclin E amplification/overexpression is a mechanism of trastuzumab resistance in HER2+ breast cancer patients. [online]. *Proceedings of the National Academy of Sciences of the United States of America*. 108 (9), pp. 3761–3766. [Accessed 14 January 2014].

Scaltriti, M., Rojo, F., Ocaña, A., Anido, J., Guzman, M., Cortes, J., Di Cosimo, S., Matias-Guiu, X., Ramon y Cajal, S., Arribas, J. & Baselga, J. (2007) Expression of p95HER2, a truncated form of the HER2 receptor, and response to anti-HER2 therapies in breast cancer. [online]. *Journal of the National Cancer Institute*. 99 (8), pp. 628–638. [Accessed 16 October 2012].

Schechter, A.L., Stern, D.F., Vaidyanathan, L., Decker, S.J., Drebin, J.A., Greene, M.I. & Weinberg, R.A. (1985) The neu oncogene: an erb-B-related gene encoding a 185,000-Mr tumour antigen. [online]. *Nature*. 312 (5994), pp. 513–516. [Accessed 6 March 2014].

Scheuer, W., Friess, T., Burtscher, H., Bossenmaier, B., Endl, J. & Hasmann, M. (2009) Strongly enhanced antitumor activity of trastuzumab and pertuzumab combination treatment on HER2-positive human xenograft tumor models. [online]. *Cancer research*. 69 (24), pp. 9330–9336. [Accessed 13 January 2014].

Schrader, C., Schielke, a, Ellerbroek, L. & Johne, R. (2012) PCR inhibitors - occurrence, properties and removal. [online]. *Journal of applied microbiology*. 113 (5), pp. 1014–1026. [Accessed 1 November 2013].

Scott, G.K., Robles, R., Park, J.W., Montgomery, P.A., Daniel, J., Holmes, W.E., Lee, J., Keller, G.A., Li, W.L. & Fendly, B.M. (1993) A truncated intracellular HER2/neu receptor produced by alternative RNA processing affects growth of human carcinoma cells. [online]. *Molecular and cellular biology*. 13 (4), pp. 2247–2257. [Accessed 26 October 2012].

Semba, K., Kamata, N., Toyoshima, K. & Yamamoto, T. (1985) A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epidermal growth factor-receptor gene and is amplified in a human salivary gland adenocarcinoma. [online]. *Proceedings of the National Academy of Sciences of the United States of America*. 82 (19), pp. 6497–6501. [Accessed 4 March 2014].

Shah, S. & Chen, B. (2010) Testing for HER2 in Breast Cancer: A Continuing Evolution [online]. *Pathology research international*. 2011pp. 903202.

Shih, C., Padhy, L.C., Murray, M. & Weinberg, R.A. (1981) Transforming genes of carcinomas and neuroblastomas introduced into mouse fibroblasts [online]. *Nature*. 290 (5803), pp. 261–264. [Accessed 6 March 2014].

Shousha, S., Peston, D., Amo-Takyi, B., Morgan, M. & Jasani, B. (2009) Evaluation of automated silver-enhanced in situ hybridization (SISH) for detection of HER2 gene amplification in breast carcinoma excision and core biopsy specimens. [online]. *Histopathology*. 54 (2), pp. 248–253. [Accessed 8 May 2014].

Siegel, P.M., Ryan, E.D., Cardiff, R.D. & Muller, W.J. (1999) Elevated expression of activated forms of Neu/ErbB-2 and ErbB-3 are involved in the induction of mammary tumors in transgenic mice: implications for human breast cancer. [online]. *The EMBO journal*. 18 (8), pp. 2149–2164.

Sińczak-Kuta, A., Tomaszewska, R., Rudnicka-Sosin, L., Okoń, K. & Stachura, J. (2007) Evaluation of HER2/neu gene amplification in patients with invasive breast carcinoma. Comparison of in situ hybridization methods. [online]. *Polish journal of pathology : official journal of the Polish Society of Pathologists*. 58 (1), pp. 41–50.

Skotheim, R.I. & Nees, M. (2007) Alternative splicing in cancer: noise, functional, or systematic? [online]. *The international journal of biochemistry & cell biology*. 39 (7-8), pp. 1432–1449. [Accessed 27 May 2014].

Slamon, D.J., Leyland-Jones, B., Shak, S., Fuchs, H., Paton, V., Bajamonde, A., Fleming, T., Eiermann, W., Wolter, J., Pegram, M., Baselga, J. & Norton, L. (2001) Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. [online]. *The New England journal of medicine*. 344 (11), pp. 783–792. [Accessed 13 January 2014].

Smith, C.W. & Valcárcel, J. (2000) Alternative pre-mRNA splicing: the logic of combinatorial control. [online]. *Trends in biochemical sciences*. 25 (8), pp. 381–388. [Accessed 16 October 2012].

Sørlie, T., Perou, C.M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M.B., Rijn, M. Van De, Jeffrey, S.S., Thorsen, T., Quist, H., Matese, J.C., Brown, P.O., et al. (2001) *Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications*. 98 (19), .

Spector, N.L. & Blackwell, K.L. (2009) Understanding the mechanisms behind trastuzumab therapy for human epidermal growth factor receptor 2-positive breast cancer. [online]. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 27 (34), pp. 5838–5847. [Accessed 13 January 2014].

Sperinde, J., Jin, X., Banerjee, J., Penuel, E., Saha, A., Diedrich, G., Huang, W., Leitzel, K., Weidler, J., Ali, S.M., Fuchs, E.-M., Singer, C.F., Köstler, W.J., Bates, M., et al. (2010) Quantitation of p95HER2 in paraffin sections by using a p95-specific antibody and correlation with outcome in a cohort of trastuzumab-treated breast cancer patients. [online]. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 16 (16), pp. 4226–4235. [Accessed 14 January 2014].

Stamm, S., Ben-Ari, S., Rafalska, I., Tang, Y., Zhang, Z., Toiber, D., Thanaraj, T.A. & Soreq, H. (2005) Function of alternative splicing. [online]. *Gene*. 344pp. 1–20. [Accessed 10 October 2012].

Van Staveren, W.C.G., Solís, D.Y.W., Hébrant, A., Detours, V., Dumont, J.E. & Maenhaut, C. (2009) Human cancer cell lines: Experimental models for cancer cells in situ? For cancer stem cells? [online]. *Biochimica et biophysica acta*. 1795 (2), pp. 92–103. [Accessed 2 May 2014].

Stern, H.M. (2012) Improving treatment of HER2-positive cancers: opportunities and challenges. [online]. *Science translational medicine*. 4 (127), pp. 127rv2. [Accessed 13 January 2014].

Tai, W., Mahato, R. & Cheng, K. (2010) The role of HER2 in cancer therapy and targeted drug delivery. [online]. *Journal of controlled release : official journal of the Controlled Release Society*. 146 (3), pp. 264–275. [Accessed 22 October 2012].

Tanner, M., Gancberg, D., Di Leo, A., Larsimont, D., Rouas, G., Piccart, M.J. & Isola, J. (2000) Chromogenic in situ hybridization: a practical alternative for fluorescence in situ hybridization to detect HER-2/neu oncogene amplification in archival breast cancer samples. [online]. *The American journal of pathology*. 157 (5), pp. 1467–1472. [Accessed 11 January 2014].

Tazi, J., Bakkour, N. & Stamm, S. (2009) Alternative splicing and disease. [online]. *Biochimica et biophysica acta*. 1792 (1), pp. 14–26. [Accessed 13 October 2012].

Telesco, S.E. & Radhakrishnan, R. (2009) Atomistic insights into regulatory mechanisms of the HER2 tyrosine kinase domain: a molecular dynamics study. [online]. *Biophysical journal*. 96 (6), pp. 2321–2334. [Accessed 29 May 2014].

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A. & Speleman, F. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. [online]. *Genome biology*. 3 (7), pp. RESEARCH0034. [Accessed 31 January 2013].

Vargo-Gogola, T. & Rosen, J.M. (2007) Modelling breast cancer: one size does not fit all. [online]. *Nature reviews. Cancer*. 7 (9), pp. 659–672. [Accessed 7 May 2014].

Van der Vegt, B., de Bock, G.H., Bart, J., Zwartjes, N.G. & Wesseling, J. (2009) Validation of the 4B5 rabbit monoclonal antibody in determining Her2/neu status in breast cancer. [online]. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc*. 22 (7), pp. 879–886. [Accessed 8 May 2014].

Vengellur, A. & LaPres, J.J. (2004) The role of hypoxia inducible factor 1alpha in cobalt chloride induced cell death in mouse embryonic fibroblasts. [online]. *Toxicological sciences : an official journal of the Society of Toxicology*. 82 (2), pp. 638–646. [Accessed 28 May 2014].

Vennström, B. & Bishop, J.M. (1982) Isolation and characterization of chicken DNA homologous to the two putative oncogenes of avian erythroblastosis virus. [online]. *Cell*. 28 (1), pp. 135–143. [Accessed 6 March 2014].

Vogel, J. (1901) *Histology and Immunohistochemistry: A Practical Guide* [online]. (October), pp. 1–11. [Accessed 17 May 2015].

Vogel, U.F. (2010) Confirmation of a low HER2 positivity rate of breast carcinomas - limitations of immunohistochemistry and in situ hybridization [online]. *Diagnostic Pathology*. 5pp. 50.

Vordermark, D., Brown, J.M. & Phil, D. (2003) Evaluation of hypoxia-inducible factor-1?? (HIF-1??) as an intrinsic marker of tumor hypoxia in U87 MG human glioblastoma: In vitro and xenograft studies. *International Journal of Radiation Oncology Biology Physics*. 56 (4), pp. 1184–1193.

Wan, J., Sazani, P. & Kole, R. (2009) Modification of HER2 pre-mRNA alternative splicing and its effects on breast cancer cells. [online]. *International journal of cancer. Journal international du cancer*. 124 (4), pp. 772–777. [Accessed 16 October 2012].

Wang, J., Willumsen, N., Zheng, Q., Xue, Y., Karsdal, M.A. & Bay-Jensen, A.C. (2013) Bringing cancer serological diagnosis to a new level: focusing on HER2, protein ectodomain shedding and neoepitope technology. [online]. *Future oncology (London, England)*. 9 (1), pp. 35–44. [Accessed 25 October 2013].

Wang, Z. & Burge, C.B. (2008) Splicing regulation: from a parts list of regulatory elements to an integrated splicing code. [online]. *RNA (New York, N.Y.)*. 14 (5), pp. 802–813. [Accessed 1 February 2013].

Wolff, A.C., Hammond, M.E.H., Schwartz, J.N., Hagerty, K.L., Allred, D.C., Cote, R.J., Dowsett, M., Fitzgibbons, P.L., Hanna, W.M., Langer, A., McShane, L.M., Paik, S., Pegram, M.D., Perez, E. a, et al. (2007) American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. [online]. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 25 (1), pp. 118–145. [Accessed 16 October 2012].

Yamamoto, T., Nishida, T., Miyajima, N., Kawai, S., Ooi, T. & Toyoshima, K. (1983) The erbB gene of avian erythroblastosis virus is a member of the src gene family. [online]. *Cell*. 35 (1), pp. 71–78. [Accessed 6 March 2014].

Yamamoto, T., Saito, M., Kumazawa, K., Doi, A., Matsui, A., Takebe, S., Amari, T., Oyama, M. & Semba, K. (2011) *ErbB2/HER2: Its Contribution to Basic Cancer Biology and the Development of Molecular Targeted Therapy, Breast Cancer - Carcinogenesis, Cell Growth and Signalling Pathways*. Mehmet Gunduz (ed.). [online]. InTech.

Zagozdzon, R., Gallagher, W.M. & Crown, J. (2011) Truncated HER2: implications for HER2-targeted therapeutics. [online]. *Drug discovery today*. 16 (17-18), pp. 810–816. [Accessed 22 October 2012].

APPENDICES

APPENDIX A:

HER2 sequence (Accession Number NM_004448)

EXON 1 (238bp)
GGGACCGGAGAAACCAGGGGAGCCCCCCGGGCAGCCGCGCGCCCCTTCCCACGGGGCCCTTT
ACTGCGCCGCGCGCCCGGCCCCCACCCCTCGCAGCACCCCGCGCCCCGCGCCCTCCCAGCCGG
GTCCAGCCGGAGCCATGGGGCCGGAGCCGCAGTGAGCACCATGGAGCTGGCGGCCTTGTGC
CGCTGGGGGCTCCTCCTCGCCCTCTTGCCCCCCGGAGCCGCGAGCACCCAAG
EXON 2 (152bp)
TGTGCACCGGCACAGACATGAAGCTGCGGCTCCCTGCCAGTCCCGAGACCCACCTGGACATGC
TCCGCCACCTCTACCAGGGCTGCCAGGTGGTGCAGGGAAACCTGGAACTCACCTACCTGCCCA
CCAATGCCAGCCTGTCCTTCCTGCAG
EXON 3 (214bp)
GATATCCAGGAGGTGCAGGGCTACGTGCTCATCGCTCACAACCAAGTGAGGCAGGTCCCACT
GCAGAGGCTGCGGATTGTGCGAGGCACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGC
TAGACAATGGAGACCCGCTGAACAATACCACCCCTGTCACAGGGGCCTCCCCAGGAGGCCTG
CGGGAGCTGCAGCTTCGAAGCCTCACAG
EXON 4 (135bp)
AGATCTTGAAAGGAGGGGTCTTGATCCAGCGGAACCCCCAGCTCTGCTACCAGGACACGATTT
TGTGGAAGGACATCTTCCACAAGAACAACCAGCTGGCTCTCACACTGATAGACACCAACCGCT
CTCGGGCCT
EXON 5 (69bp)
GCCACCCCTGTTCTCCGATGTGTAAGGGCTCCCGCTGCTGGGGAGAGAGTTCTGAGGATTGTC
AGAGCC
EXON 6 (116bp)
TGACGCGCACTGTCTGTGCCGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGACTGCT
GCCATGAGCAGTGTGCTGCCGGCTGCACGGGCCCCAAGCACTCTGACTGCCTG
EXON 7 (142bp)
GCCTGCCTCCACTTCAACCACAGTGGCATCTGTGAGCTGCACTGCCCAGCCCTGGTCACCTACA
ACACAGACACGTTTGAGTCCATGCCCAATCCCGAGGGCCGGTATACATTCGGCGCCAGCTGTG
TGACTGCCTGTCCCT
EXON 8 (120bp)
ACAACTACCTTTCTACGGACGTGGGATCCTGCACCCTCGTCTGCCCCCTGCACAACCAAGAGG
TGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGAG
EXON 9 (127bp)
TGTGCTATGGTCTGGGCATGGAGCACTTGCGAGAGGTGAGGGCAGTTACCAGTGCCAATATC
CAGGAGTTTGCTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGGAGAGCTTTGAT
GG
EXON 10 (74bp)
GGACCCAGCCTCCAACACTGCCCCGCTCCAGCCAGAGCAGCTCCAAGTGTTTGAGACTCTGGA
AGAGATCACAG
EXON 11 (91bp)
GTTACCTATACATCTCAGCATGGCCGGACAGCCTGCCTGACCTCAGCGTCTTCCAGAACCTGCA
AGTAATCCGGGGACGAATTCTGCACAA

EXON 12 (200bp)
TGGCGCCTACTCGCTGACCCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACTGA
GGGAACTGGGCAGTGGACTGGCCCTCATCCACCATAACACCCACCTCTGCTTCGTGCACACGG
TGCCCTGGGACCAGCTCTTTCGGAACCCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCAG
AGGACGAGTGTG
EXON 13 (133bp)
TGGGCGAGGGCCTGGCCTGCCACCAGCTGTGCGCCCGAGGGCACTGCTGGGGTCCAGGGCC
CACCCAGTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGCGTGGAGGAATGCCGAG
TACTGCAGGG
EXON 14 (91bp)
GCTCCCCAGGGAGTATGTGAATGCCAGGCACTGTTTGCCGTGCCACCCTGAGTGTCAGCCCCA
GAATGGCTCAGTGACCTGTTTTGGACCG
EXON 15 (161bp)
GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGCTGC
CCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAGGGC
GCATGCCAGCCTTGCCCCATCAACTGCACCCACTC
EXON 16 (48bp)
CTGTGTGGACCTGGATGACAAGGGCTGCCCCGCCGAGCAGAGAGCCAG
EXON 17 (139bp)
CCCTCTGACGTCCATCATCTCTGCGGTGGTTGGCATTCTGCTGGTCGTGGTCTTGGGGGTGGT
CTTTGGGATCCTCATCAAGCGACGGCAGCAGAAGATCCGGAAGTACACGATGCGGAGACTGC
TGCAGGAAACGGAG
EXON 18 (123bp)
CTGGTGGAGCCGCTGACACCTAGCGGAGCGATGCCCAACCAGGCGCAGATGCGGATCCTGAA
AGAGACGGAGCTGAGGAAGGTGAAGGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAG
EXON 19 (99bp)
GGCATCTGGATCCCTGATGGGGAGAATGTGAAAATTCCAGTGGCCATCAAAGTGTTGAGGGA
AAACACATCCCCCAAAGCCAACAAAGAAATCTTAGAC
EXON 20 (186bp)
GAAGCATACGTGATGGCTGGTGTGGGCTCCCCATATGTCTCCCGCCTTCTGGGCATCTGCCTG
ACATCCACGGTGCAGCTGGTGACACAGCTTATGCCCTATGGCTGCCTCTTAGACCATGTCCGG
GAAAACCGCGGACGCCTGGGCTCCCAGGACCTGCTGAACTGGTGTATGCAGATTGCCAAG
EXON 21 (156bp)
GGGATGAGCTACCTGGAGGATGTGCGGCTCGTACACAGGGACTTGGCCGCTCGGAACGTGCT
GGTCAAGAGTCCCAACCATGTCAAAATTACAGACTTCGGGCTGGCTCGGCTGCTGGACATTGA
CGAGACAGAGTACCATGCAGATGGGGGCAAG
EXON 22 (76bp)
GTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCCGGCGGTTCACCCACCAGAGTGAT
GTGTGGAGTTATG
EXON 23 (147bp)
GTGTGACTGTGTGGGAGCTGATGACTTTTGGGGCCAAACCTTACGATGGGATCCCAGCCCGG
GAGATCCCTGACCTGCTGGAAAAGGGGGAGCGGCTGCCCCAGCCCCCCATCTGCACCATTGA
TGTCTACATGATCATGGTCAAAT
EXON 24 (98bp)
GTTGGATGATTGACTCTGAATGTCGGCCAAGATTCCGGGAGTTGGTGTCTGAATTCTCCCGCA
TGGCCAGGGACCCCCAGCGCTTTGTGGTCATCCAG

EXON 25 (189bp)

AATGAGGACTTGGGCCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTGGAGGAC
GATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCCAGCAGGGCTTCTTCTG
TCCAGACCCTGCCCCGGGCGCTGGGGGCATGGTCCACCACAGGCACCGCAGCTCATCTACCA
GG

EXON 26 (253bp)

AGTGGCGGTGGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAGGCCCCCAGGTCTC
CACTGGCACCCTCCGAAGGGGCTGGCTCCGATGTATTTGATGGTGACCTGGGAATGGGGGCA
GCCAAGGGGCTGCAAAGCCTCCCCACACATGACCCCAGCCCTCTACAGCGGTACAGTGAGGA
CCCCACAGTACCCCTGCCCTCTGAGACTGATGGCTACGTTGCCCCCCTGACCTGCAGCCCCCAG
CCTG

EXON 27 (969bp)

AATATGTGAACCAGCCAGATGTTCGGCCCCAGCCCCCTTCGCCCCGAGAGGGCCCTCTGCCTG
CTGCCCGACCTGCTGGTGCCACTCTGGAAAGGCCCAAGACTCTCTCCCCAGGGAAGAATGGG
GTCGTCAAAGACGTTTTTGCCTTTGGGGGTGCCGTGGAGAACCCCGAGTACTTGACACCCCAG
GGAGGAGCTGCCCCTCAGCCCCACCCTCCTCCTGCCTTCAGCCCAGCCTTCGACAACCTCTATT
ACTGGGACCAGGACCCACCAGAGCGGGGGGCTCCACCCAGCACCTTCAAAGGGACACCTACG
GCAGAGAACCCAGAGTACCTGGGTCTGGACGTGCCAGTGTGAACCAGAAGGCCAAGTCCGCA
GAAGCCCTGATGTGTCCTCAGGGAGCAGGGAAGGCCTGACTTCTGCTGGCATCAAGAGGTGG
GAGGGCCCTCCGACCACTTCCAGGGGAACCTGCCATGCCAGGAACCTGTCCTAAGGAACCTTC
CTTCCTGCTTGAGTTCCCAGATGGCTGGAAGGGGTCCAGCCTCGTTGGAAGAGGAACAGCAC
TGGGGAGTCTTTGTGGATTCTGAGGCCCTGCCCAATGAGACTCTAGGGTCCAGTGGATGCCAC
AGCCCAGCTTGGCCCTTTCCTTCCAGATCCTGGGTACTGAAAGCCTTAGGGAAGCTGGCCTGA
GAGGGGAAGCGGCCCTAAGGGAGTGTCTAAGAACAAAAGCGACCCATTCAGAGACTGTCCCT
GAAACCTAGTACTGCCCCCCATGAGGAAGGAACAGCAATGGTGTCAGTATCCAGGCTTTGTAC
AGAGTGCTTTTCTGTTTAGTTTTTACTTTTTTTGTTTTGTTTTTTTAAAGATGAAATAAAGACCCA
GGGGGAGAATGGGTGTTGTATGGGGAGGCAAGTGTGGGGGGTCCTTCTCCACACCCACTTTG
TCCATTTGCAAATATATTTTGGAAAAC

Sequences of *HER2* cDNA obtained from RT-PCR assays which showed additional bands and subsequent differences in cDNA sequence

5'…TGGCGCCTACTCGCTGACCCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACT
GAGGGAACTGGGCAGTGGACTGGCCCTCATCCACCATAACACCCACCTCTGCTTCGTGCACAC
GGTGCCCTGGGACCAGCTCTTTC*ggaacccgcaccaagctct*GCTCCACACTGCCAACCGGCCAGA
GGACGAGTGTGTGGGCGAGGGCCTGGCCTGCCACCAGCTGTGCGCCCGAGGGCACTGCTGG
GGTCCAGGGCCCACCCAGTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGCGTGGA
GGAATGCCGAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGTTTGCCGT
GCCACCCTGAGTGTCAGCCCCAGAATGGCTCAGTGACCTGTTTTGGACCGGAGGCTGACCAGT
GTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGCTGCCCCAGCGGTGTGA
AACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAGGGCGCATGCCAG*ccttg*
*ccccatcaactgca*CCCACTC…3'

Figure A 1: Sequence analysis of the top band amplified using primers E12F + E 15R, showing exon 12, exon 13, exon 14, and exon 15. In Italics – forward and reverse primers. Underlined – actual cDNA sequence obtained from excised band.

5'…TGGCGCCTACTCGCTGACCCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACT
GAGGGAACTGGGCAGTGGACTGGCCCTCATCCACCATAACACCCACCTCTGCTTCGTGCACAC
GGTGCCCTGGGACCAGCTCTTTC*ggaacccgcaccaagctct*GCTCCACACTGCCAACCGGCCAGA
GGACGAGTGTGTGGGCGAGGGCCTGGCCTGCCACCAGCTGTGCGCCCGAGGGCACTGCTGG
GGTCCAGGGCCCACCCAGTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGCGTGGA
GGAATGCCGAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGTTTGCCGT
GCCACCCTGAGTGTCAGCCCCAGAATGGCTCAGTGACCTGTTTTGGACCGGAGGCTGACCAGT
GTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGCTGCCCCAGCGGTGTGA
AACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAGGGCGCATGCCAG*ccttg*
*ccccatcaactgca*CCCACTC…3'

Figure A 2: Sequence analysis of the middle band amplified using primers E12F + E 15R, showing exon 12, exon 13, exon 14, and exon 15. In Italics – forward and reverse primers. Underlined – actual gene sequence obtained from excised band.

5'…GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGC
TGCCCC*agcggtgtgaaacctgacctc*TCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAGGGCG
CATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAGGGCTGCC
CCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTTGGCATTCTGCTGG
TCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAGAAGATCCGGAAG
TACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCGCTGACACCTAGCGGAG
CGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGAGCTGAGGAAGGTGAAGGT
GCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATCCCTGATGGGGAGAATGT
GAAAATTCCAGTGGCCATCAAAGTGT*tgagggaaaacacatccccc*AAAGCCAACAAAGAAATCTT
AGAC…3'

Figure A 3: Sequence showing the top band amplified using primers E15F + E 19R, showing exon 15, exon 16, exon 17, exon 18, and exon 19. In Italics – forward and reverse primers. Underlined – actual gene sequence obtained from excised band.

5'…GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGC
TGCCCC*agcggtgtgaaacctgacctc*TCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAGGGCG
CATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAGGGCTGCC
CCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTTGGCATTCTGCTGG
TCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAGAAGATCCGGAAG
TACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCGCTGACACCTAGCGGAG
CGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGAGCTGAGGAAGGTGAAGGT
GCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATCCCTGATGGGGAGAATGT
GAAAATTCCAGTGGCCATCAAAGTGT*tgagggaaaacacatccccc*AAAGCCAACAAAGAAATCTT
AGAC…3'

Figure A 4: Sequence showing the lower band amplified using primers E15F + E 19R, showing exon 15, exon 16, exon 17, exon 18, and exon 19. In Italics – forward and reverse primers. Underlined – actual gene sequence obtained from excised band.

5'…GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGC
TGCCCCAGCGGTGTGAAACCTGACCTCTCCTA*catgcccatctggaagtttc*CAGATGAGGAGGGCG
CATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAGGGCTGCC
CCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTTGGCATTCTGCTGG
TCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAGAAGATCCGGAAG
TACACGATGCGGAGACTGCTG*caggaaacggagctggtggagc*CGCTGACACCTAGCGGAGCGAT
GCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGAGCTGAGGAAGGTGAAGGTGCTT
GGATCTGGCGCTTTTGGCACAGTCTACAAG…3'

Figure A 5: Sequence showing the top band amplified using primers NP1+NP2, showing exon 15, exon 16, exon 17, and exon 18. In Italics – forward and reverse primers. Underlined – actual gene sequence obtained from excised band.

5'…GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGC
TGCCCCAGCGGTGTGAAACCTGACCTCTCCTA*catgcccatctggaagtttc*CAGATGAGGAGGGCG
CATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAGGGCTGCC
CCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTTGGCATTCTGCTGG
TCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAGAAGATCCGGAAG
TACACGATGCGGAGACTGCTG*caggaaacggagctggtggagc*CGCTGACACCTAGCGGAGCGAT
GCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGAGCTGAGGAAGGTGAAGGTGCTT
GGATCTGGCGCTTTTGGCACAGTCTACAAG…3'

Figure A 6: Sequence showing the top band amplified using primers NP1+NP2, showing exon 15, exon 16, exon 17, and exon 18. In Italics – forward and reverse primers. Underlined – actual gene sequence obtained from excised band.

5'…GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGC
TGCCCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG
GGCGC*catgccagccttgccccatcaactgc*ACCCACTCCTGTGTGGACCTGGATGACAAGGGCTGCCC
CGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTTGGCATTCTGC*tggtc*
*gtggtcttggggtggtct*TTGGGATCCTCATCAAGCGACGGCAGCAGAAGATCCGGAAGTACACG
ATGCGGAGACTGCTGCAGGAAACGGAG…3'

Figure A 7: Sequence showing the top band amplified using primers NP5+NP6, showing exon 15, exon 16, and exon 17. In Italics – forward and reverse primers. Underlined – actual gene sequence obtained from excised band.

5'…GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGC
TGCCCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG
GGCGC*catgccagccttgccccatcaactgc*ACCCACTCCTGTGTGGACCTGGATGACAAGGGCTGCCC
CGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTTGGCATTCTGC*tggtc*
*gtggtcttggggtggtct*TTGGGATCCTCATCAAGCGACGGCAGCAGAAGATCCGGAAGTACACG
ATGCGGAGACTGCTGCAGGAAACGGAG…3'

Figure A 8: Sequence showing the top band amplified using primers NP5+NP6, showing exon 15, exon 16, and exon 17. In Italics – forward and reverse primers. Underlined – actual gene sequence obtained from excised band.

RT-PCR PRIMER SEQUENCES

PRIMER NAME	EXON TARGET	SENSE PRIMER	ANTI-SENSE PRIMER
E1F1 + E1R	EXON 1	5' CCAGTAGAATGGCCGGAGGA '3	5' TTTCTCCGGTCCCAATGGAG '3
E1F2 + E2R	EXONS 1-2	5' GTGGAGGAGGAGGGCTGCTT '3	5' GGAGCCGCAGCTTCATGTCT '3
E1F3 + E3R	EXONS 1-3	5' CTTCCCACGGGGCCCTTTAC '3	5' AGCACGTAGCCCTGCACCTC '3
E3F + E6R	EXONS 3-6	5' TCCAGGAGGTGCAGGGCTAC '3	5' CAGCAGTCAGTGGGCAGTGG '3
E6F + E9R	EXONS 6-9	5' CAGCAGTCAGTGGGCAGTGG '3	5' AAATGCCAGGCTCCCAAAGA '3
E9F + E12R	EXONS 9-12	5' TGGAGCACTTGCGAGAGGTG '3	5' TTGGCAGTGTGGAGCAGAGC '3
E12F + E15R	EXONS 12-15	5' GGAACCCGCACCAAGCTCT '3	5' TGCAGTTGATGGGGCAAGG '3
E15F + E19R	EXONS 15-19	5' AGCGGTGTGAAACCTGACCTC '3	5' GGGGGATGTGTTTTCCCTCA '3
E19F + E22R	EXONS 19-22	5' GAAAACACATCCCCCAAAGCCAAC '3	5' CCACACATCACTCTGGTGGGTGAA '3
E22F + E25R	EXONS 22-25	5' GATGGCGCTGGAGTCCATTC '3	5' GATGAGCTGCGGTGCCTGT '3
E25 + E27R	EXONS 25-27	5' CAGGCACCGCAGCTCATCTA '3	5' CTCCTCCCTGGGGTGTCAAG '3
E27F + E27R1	EXON 27	5' GAGAACCCCGAGTACTTGACAC '3	5' GCAAATGGACAAAGTGGGTGTG '3
NP1 + NP2	EXONS 16-18	5' CATGCCCATCTGGAAGTTTC '3	5' GCTCCACCAGCTCCGTTTCCTG '3
NP5 + NP6	EXONS 16-18	5' ATGCCAGCCTTGCCCCATCAACTGC '3	5' ACCACCCCCAAGACCACGACCAG '3

Table A 1: List of RT-PCR oligonucleotide sequences

RAW VALUES FOR RT-PCR SEQUENCES (INCLUDING SEQUENCES WHICH DID NOT PROCUDE MULTIPLE BANDS, AND FOR WHICH ALTERNATIVE SPLICING WAS NOT OBSERVED)

> 001_T7 -- 12..1051 of sequence

CATGCTCCGGCCGCCTGGCGGCCGCGGGAATTCGATTCACCGGCACAGACATGAAGCTGCGG
CTCCCTGCCAGTCCCGAGACCCACCTGGACATGCTCCGCCACCTCTACCAGGGCTGCCAGGTG
GTGCAGGGAAACCTGGAACTCACCTACCTGCCCACCAATGCCAGCCTGTCCTTCCTGCAGGAT
ATCCAGGAGGTGCAGGGCTACGTGCTCATCGCTCACAACCAAGTGAGGCAGGTCCCACTGCA
GAGGCTGCGGATTGTGCGAGGCACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGCTAG
ACAATGGAGACCCGCTGAACAATACAATCACTAGTGAATTCGCGGCCGCCTGCAGGTCGACC
ATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTGTCACCTAAATA
GCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACA
CAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCA
CATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTA
ATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCT
CACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGG
TAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCA
GCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCC
TGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAA
GATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTAC
CGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGC

EXON 1-3
GGAGGAGGTGGAGGAGGAGGGCTGCTTGAGGAAGTATAAGAATGAAGTTGTGAAGCTGAG
ATTCCCCTCCATTGGGACCGGAGAAACCAGGGGAGCCCCCCGGGCAGCCGCGCGCCCCTTCC
CACGGGGCCCTTTACTGCGCCGCGCGCCCGGCCCCCACCCCTCGCAGCACCCCGCGCCCCGCG
CCCTCCCAGCCGGGTCCAGCCGGAGCCATGGGGCCGGAGCCGCAGTGAGCACCATGGAGCT
GGCGGCCTTGTGCCGCTGGGGGCTCCTCCTCGCCCTCTTGCCCCCCGGAGCCGCGAGCACCCA
AGTGTGCACCGGCACAGACATGAAGCTGCGGCTCCCTGCCAGTCCCGAGACCCACCTGGACA
TGCTCCGCCACCTCTACCAGGGCTGCCAGGTGGTGCAGGGAAACCTGGAACTCACCTACCTGC
CCACCAATGCCAGCCTGTCCTTCCTGCAGGATATCCAGGAGGTGCAGGGCTACGTGCTCATCG
CTCACAACCAAGTGAGGCAGGTCCCACTGCAGAGGCTGCGGATTGTGCGAGGCACCCAGCTC
TTTGAGGACAACTATGCCCTGGCCGTGCTAGACAATGGAGACCCGCTGAACAATACCACCCCT
GTCACAGGGGCCTCCCCAGGAGGCCTGCGGGAGCTGCAGCTTCGAAGCCTCACAG

> 002_T7 -- 14..1010 of sequence
ATGCTCCGGCCGCCATGGCGGCCGCGGGAATTCGATTAGGGACAGGCAGTCACACAGCTGGC
GCCGAATGTATACCGGCCCTCGGGATTGGGCATGGACTCAAACGTGTCTGTGTTGTAGGTGA
CCAGGGCTGGGCAGTGCAGCTCACAGATGCCACTGTGGTTGAAGTGGAGGCAGGCCAGGCA
GTCAGAGTGCTTGGGGCCCGTGCAGCCGGCAGCACACTGCTCATGGCAGCAGTCAGTGGGCA

GTGGCCCCTTGCAGCGGGCACAGCCACCGGCACAGACAGTGCGCGTCAGGCTCTGACAATCC
TCAGAACTCTCTCCCCAGCAGAATCACTAGTGAATTCGCGGCCGCCTGCAGGTCGACCATATG
GGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTGTCACCTAAATAGCTTG
GCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACA
TACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTA
ATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGA
ATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACT
GACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAAT
ACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAA
AAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGAC
GAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATA
CCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTACCGG

EXON 3-7 (REVERSE)
GATATCCAGGAGGTGCAGGGCTACGTGCTCATCGCTCACAACCAAGTGAGGCAGGTCCCACT
GCAGAGGCTGCGGATTGTGCGAGGCACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGC
TAGACAATGGAGACCCGCTGAACAATACCACCCCTGTCACAGGGGCCTCCCCAGGAGGCCTG
CGGGAGCTGCAGCTTCGAAGCCTCACAGAGATCTTGAAAGGAGGGGTCTTGATCCAGCGGAA
CCCCCAGCTCTGCTACCAGGACACGATTTTGTGGAAGGACATCTTCCACAAGAACAACCAGCT
GGCTCTCACACTGATAGACACCAACCGCTCTCGGGCCTGCCACCCCTGTTCTCCGATGTGTAAG
GGCTCCCGCTGCTGGGGAGAGAGTTCTGAGGATTGTCAGAGCCTGACGCGCACTGTCTGTGC
CGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGTGCTG
CCGGCTGCACGGGCCCCAAGCACTCTGACTGCCTGGCCTGCCTCCACTTCAACCACAGTGGCA
TCTGTGAGCTGCACTGCCCAGCCCTGGTCACCTACAACACAGACACGTTTGAGTCCATGCCCA
ATCCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTGACTGCCTGTCCCTACAACTACCTTTC
TACGGACGTGGGATCCTGCACCCTCGTCTGCCCCCTGCACAACCAAGAGGTGACAGCAGAGG
ATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGAG

> 003_T7 -- 14..1007 of sequence
ATGCTCCGGCCGCCATGGCGGCCGCGGGAATTCGATTAGGGACAGGCAGTCACACAGCTGGC
GCCGAATGTATACCGGCCCTCGGGATTGGGCATGGACTCAAACGTGTCTGTGTTGTAGGTGA
CCAGGGCTGGGCAGTGCAGCTCACAGATGCCACTGTGGTTGAAGTGGAGGCAGGCCAGGCA
GTCAGAGTGCTTGGGGCCCGTGCAGCCGGCAGCACACTGCTCATGGCAGCAGTCAGTGGGCA
GTGGCCCCTTGCAGCGGGCACAGCCACCGGCACAGACAGTGCGCGTCAGGCTCTGACAATCC
TCAGAACTCTCTCCCCAGCAGAATCACTAGTGAATTCGCGGCCGCCTGCAGGTCGACCATATG
GGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTGTCACCTAAATAGCTTG
GCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACA
TACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTA
ATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGA
ATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACT
GACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAAT
ACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAA
AAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGAC
GAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATA
CCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTA

EXON 3-7 (REVERSE)
GATATCCAGGAGGTGCAGGGCTACGTGCTCATCGCTCACAACCAAGTGAGGCAGGTCCCACT
GCAGAGGCTGCGGATTGTGCGAGGCACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGC
TAGACAATGGAGACCCGCTGAACAATACCACCCCTGTCACAGGGGCCTCCCCAGGAGGCCTG
CGGGAGCTGCAGCTTCGAAGCCTCACAGAGATCTTGAAAGGAGGGGTCTTGATCCAGCGGAA
CCCCCAGCTCTGCTACCAGGACACGATTTTGTGGAAGGACATCTTCCACAAGAACAACCAGCT
GGCTCTCACACTGATAGACACCAACCGCTCTCGGGCCTGCCACCCCTGTTCTCCGATGTGTAAG
GGCTCCCGCTGCTGGGGAGAGAGTTCTGAGGATTGTCAGAGCCTGACGCGCACTGTCTGTGC
CGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGTGCTG
CCGGCTGCACGGGCCCCAAGCACTCTGACTGCCTGGCCTGCCTCCACTTCAACCACAGTGGCA
TCTGTGAGCTGCACTGCCCAGCCCTGGTCACCTACAACACAGACACGTTTGAGTCCATGCCCA
ATCCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTGACTGCCTGTCCCTACAACTACCTTTC
TACGGACGTGGGATCCTGCACCCTCGTCTGCCCCCTGCACAACCAAGAGGTGACAGCAGAGG
ATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGAG

> 004_T7 -- 12..954 of sequence
CGCATGCTCCGGCCGCCATGGCGGCCGCGGGAATTCGATTCCAGGAGGTGCAGGGCTACGTG
CTCATCGCTCACAACCAAGTGAGGCAGGTCCCACTGCAGAGGCTGCGGATTGTGCGAGGCAC
CCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGCTAGACAATGGAGACCCGCTGAACAATAC
CACCCCTGTCACAGGGGCCTCCCCAGGAGGCCTGCGGGAGCTGCAGCTTCGAAGCCTCACAG
AGATCTTGAAAGGAGGGGTCTTGATCCAGCGGAACCCCCAGCTCTGCTACCAGGACACGATTT
TGTGGAAGGACATCTTCCACAAGAACAACCAGCTGGCTCTCACACTGATAGACACCAACCGCT
CTCGGGCCTGCCACCCCTGTTCTCCGATGTGTAAGGGCTCCCGCTGCTGGGGAGAGAGTTCTG
AGGATTGTCAGAGCCTGACGCGCACTGTCTGTGCCGGTGGCTGTGCCCGCTGCAAGGGGCCA
CTGCCCACTGACTGCTGAATCACTAGTGAATTCGCGGCCGCCTGCAGGTCGACCATATGGGAG
AGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTGTCACCTAAATAGCTTGGCGT
AATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACG
AGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTG
CGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCG
GCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACT
CGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGG
TTA

EXON 3-6
GATATCCAGGAGGTGCAGGGCTACGTGCTCATCGCTCACAACCAAGTGAGGCAGGTCCCACT
GCAGAGGCTGCGGATTGTGCGAGGCACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGC
TAGACAATGGAGACCCGCTGAACAATACCACCCCTGTCACAGGGGCCTCCCCAGGAGGCCTG
CGGGAGCTGCAGCTTCGAAGCCTCACAGAGATCTTGAAAGGAGGGGTCTTGATCCAGCGGAA
CCCCCAGCTCTGCTACCAGGACACGATTTTGTGGAAGGACATCTTCCACAAGAACAACCAGCT
GGCTCTCACACTGATAGACACCAACCGCTCTCGGGCCTGCCACCCCTGTTCTCCGATGTGTAAG
GGCTCCCGCTGCTGGGGAGAGAGTTCTGAGGATTGTCAGAGCCTGACGCGCACTGTCTGTGC
CGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGTGCTG
CCGGCTGCACGGGCCCCAAGCACTCTGACTGCCTGGCCTGCCTCCACTTCAACCACAGTGGCA

TCTGTGAGCTGCACTGCCCAGCCCTGGTCACCTACAACACAGACACGTTTGAGTCCATGCCCA
ATCCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTGACTGCCTGTCCCT

> 005_T7 -- 14..1056 of sequence
TGCTCCGGCCGCCATGGCGGCCGCGGGAATTCGATTAAATGCCAGGCTCCCAAAGATCTTCTT
GCAGCCAGCAAACTCCTGGATATTGGCACTGGTAACTGCCCTCACCTCTCGCAAGTGCTCCAT
GCCCAGACCATAGCACACTCGGGCACAGGGCTTGCTGCACTTCTCACACCGCTGTGTTCCATC
CTCTGCTGTCACCTCTTGGTTGTGCAGGGGGCAGACGAGGGTGCAGGATCCCACGTCCGTAG
AAAGGTAGTTGTAGGGACAGGCAGTCACACAGCTGGCGCCGAATGTATACCGGCCCTCGGGA
TTGGGCATGGACTCAAACGTGTCTGTGTTGTAGGTGACCAGGGCTGGGCAGTGCAGCTCACA
GATGCCACTGTGGTTGAAGTGGAGGCAGGCCAGGCAGTCAGAGTGCTTGGGGCCCGTGCAG
CCGGCAGCACACTGCTCATGGCAGCAGTCAGTGGGCAGTGGAATCACTAGTGAATTCGCGGC
CGCCTGCAGGTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTA
TAGTGTCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATC
CGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAAT
GAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTC
GTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCT
CTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAG
CTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCACGGGATAACGCACGAAAGAACATG
TGAGCAAAAGGCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTCCATA
GGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAGT

EXON 5-8 (REVERSE)
GCCACCCCTGTTCTCCGATGTGTAAGGGCTCCCGCTGCTGGGGAGAGAGTTCTGAGGATTGTC
AGAGCCTGACGCGCACTGTCTGTGCCGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACT
GACTGCTGCCATGAGCAGTGTGCTGCCGGCTGCACGGGCCCCAAGCACTCTGACTGCCTGGC
CTGCCTCCACTTCAACCACAGTGGCATCTGTGAGCTGCACTGCCCAGCCCTGGTCACCTACAAC
ACAGACACGTTTGAGTCCATGCCCAATCCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTG
ACTGCCTGTCCCTACAACTACCTTTCTACGGACGTGGGATCCTGCACCCTCGTCTGCCCCCTGC
ACAACCAAGAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTG
TGCCCGAGTGTGCTATGGTCTGGGCATGGAGCACTTGCGAGAGGTGAGGGCAGTTACCAGTG
CCAATATCCAGGAGTTTGCTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGGAGA
GCTTTGATGG

> 006_T7 -- 15..1032 of sequence
ATGCTCCGGCCGCCATGGCGGCCGCGGGAATTCGATTGGAGCACTTGCGAGAGGTGAGGGC
AGTTACCAGTGCCAATATCCAGGAGTTTGCTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATT
TCTGCCGGAGAGCTTTGATGGGGACCCAGCCTCCAACACTGCCCCGCTCCAGCCAGAGCAGCT
CCAAGTGTTTGAGACTCTGGAAGAGATCACAGGTTACCTATACATCTCAGCATGGCCGGACAG
CCTGCCTGACCTCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGGACGAATTCTGCACAATGG
CGCCTACTCGCTGACCCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACTGAGGG
AACTGGGCAGTGGACTGGCCCTCATCCACCATAACACCCACCTCTGCTTCGTGCACACGGTGC
CCTGGGACCAGCTCTTTCGGAACCCGCACCAAGCTCTGCTCCACACTGCCAAAATCACTAGTG
AATTCGCGGCCGCCTGCAGGTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCT
TGAGTATTCTATAGTGTCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTG
AAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTG
GGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCG

GGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCG
TATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGA
GCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCACGGGATAACGCACG
AAAGAACATGTGAGCAAACGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGC
GTTTTCCATAGGCTC

EXONS 8-10
TGTGCTATGGTCTGGGCATGGAGCACTTGCGAGAGGTGAGGGCAGTTACCAGTGCCAATATC
CAGGAGTTTGCTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGGAGAGCTTTGAT
GGGGACCCAGCCTCCAACACTGCCCCGCTCCAGCCAGAGCAGCTCCAAGTGTTTGAGACTCTG
GAAGAGATCACAGGTTACCTATACATCTCAGCATGGCCGGACAGCCTGCCTGACCTCAGCGTC
TTCCAGAACCTGCAAGTAATCCGGGGACGAATTCTGCACAATGGCGCCTACTCGCTGACCCTG
CAAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACTGAGGGAACTGGGCAGTGGACTGG
CCCTCATCCACCATAACACCCACCTCTGCTTCGTGCACACGGTGCCCTGGGACCAGCTCTTTCG
GAACCCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCAGAGGACGAGTGTG

> 010_T7 -- 13..946 of sequence
ATGCTCCGGCCGCCATGGCGGCCGCGGGAATTCGATTGGGGGATGTGTTTTCCCTCAACACTT
TGATGGCCACTGGAATTTTCACATTCTCCCCATCAGGGATCCAGATGCCCTTGTAGACTGTGCC
AAAAGCGCCAGATCCAAGCACCTTCACCTTCCTCAGCTCCGTCTCTTTCAGGATCCGCATCTGC
GCCTGGTTGGGCATCGCTCCGCTAGGTGTCAGCGGCTCCACCAGCTCCGTTTCCTGCAGCAGT
CTCCGCATCGTGTACTTCCGGATCTTCTGCTGCCGTCGCTTGATGAGGATCCCAAAGACCACCC
CCAAGACCACGACCAGCAGAATGCCAACCACCGCAGAGACGATGGACGTCAGAGGGCTGGC
TCTCTGCTCGGCGGGGCAGCCCTTGTCATCCAGGTCCACACAGGAGTGGGTGCAGTTGATGG
GGCAAGGCTGGCATGCGCCCTCCTCATCTGGAAACCTCCAGATGGGCATGTAGGAGAGGTCA
GGTTTCACACCGCTAATCACTAGTGAATTCGCGGCCGCCTGCAGGTCGACCATATGGGAGAGC
TCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTGTCACCTAAATAGCTTGGCGTAATC
ATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCC
GGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTT
GCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCA
ACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCT
GCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTA

> 011b_T7 -- 12..918 of sequence
GCATGCTCCGGCCGCCATGGCGGCCGCGGGAATTCGATTGGGGGATGTGTTTTCCCTCAACAC
TTTGATGGCCACTGGAATTTTCACATTCTCCCCATCAGGGATCCAGATGCCCTTCCTCAGCTCC
GTCTCTTTCAGGATCCGCATCTGCGCCTGGTTGGGCATCGCTCCGCTAGGTGTCAGCGGCTCC
ACCAGCTCCGTTTCCTGCAGCAGTCTCCGCATCGTGTACTTCCGGATCTTCTGCTGCCGTCGCTT
GATGAGGATCCCAAAGACCACCCCCAAGACCACGACCAGCAGAATGCCAACCACCGCAGAGA
CGATGGACGTCAGAGGGCTGGCTCTCTGCTCGGCGGGGCAGCCCTTGTCATCCAGGTCCACA
CAGGAGTGGGTGCAGTTGATGGGGCAAGGCTGGCATGCGGCCTCCTCATCTGGAAACTTCCA
GATGGGCATGTAGGAGAGGTCAGGTTTCACACCGCTAATCACTAGTGAATTCGCGGCCGCCT
GCAGGTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTG
TCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTC
ACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGT
GAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGC
CAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTC

CGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCA
CTCAAAGGCGGTAATACGGTTATCCA

> 013_T7 -- 12..1020 of sequence
CATGCTCCGGCCGCCATGGCGGCCGCGGGAATTCGATTCATGCCCATCTGGAAGTTTCCAGAT
GAGGAGGGCGCATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGAC
TAGGGCTGCCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCGTCTCTGCGGTGGTTGG
CATTCTGCTGGTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAGAA
GATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCAATCACTA
GTGAATTCGCGGCCGCCTGCAGGTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATA
GCTTGAGTATTCTATAGTGTCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGT
GTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGC
CTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAG
TCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTT
GCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGG
CGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGC
AGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTG
CTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAG
AGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGT
GCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGC
GTGGCG

> 014_T7 -- 13..977 of sequence
CATGCTCCGGCCGCCATGGCGGCCGCGGGAATTCGATTCATGCCCATCTGGAAGTTTCCAGAT
AAGGAGGGTGCATGCCAGCCTTGCCCCATCAACTGCACCCACTCCCCTCTGACGTCCATCGTCT
CTGCGGTGGTTGGCATTCAGCTGGTCGTGGTCTTGGGGGTGGTCTTTGGAATCCTCATCAAGC
GACGGCAGCAGAAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGT
GGAGCAATCACTAGTGAATTCGCGGCCGCCTGCAGGTCGACCATATGGGAGAGCTCCCAACG
CGTTGGATGCATAGCTTGAGTATTCTATAGTGTCACCTAAATAGCTTGGCGTAATCATGGTCAT
AGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCAT
AAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACT
GCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGG
GGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGG
TCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAAT
CAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAA
AAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCG
ACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTG
GAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTC
CCTTCGGGAAGCGTGGCGCTTC

> 015_T7 -- 13..1039 of sequence
ATGCTCCGGCCGCCATGGCGGCCGCGGGAATTCGATTCATGCCCATCTGGAAGTTTCCAGATG
AGGAGGGCGCATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACA
AGGGCTGCCCCGCCGAGCAGAGAGCCAGCGACGGCAGCAGAAGATCCGGAAGTACACGATG
CGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCAATCACTAGTGAATTCGCGGCCGCCTGC
AGGTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTGTC
ACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCAC

AATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGA
GCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCA
GCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCG
CTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTC
AAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCA
AAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCT
CCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAG
GACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCT
GCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCA
CGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCC
CCGTTCAGCCCGACGCTGCGCCT

> 017_T7 -- 12..980 of sequence
CATGCTCCGGCCGCCATGGCCGCGGGATTAGACCACCCCCAAGACCACGACCAGCAGAATGC
CAACCACCGCAGAGACGATGGACGTCAGAGGGGAGTGGGTGCAGTTGATGGGGCAAGGCTG
GCATAATCACTAGTGCGGCCGCCTGCAGGTCGACCATATGGGAGAGCTCCCAACGCGTTGGA
TGCATAGCTTGAGTATTCTATAGTGTCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTT
TCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTG
TAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGC
TTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAG
GCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCG
GCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGG
ATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGC
CGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTC
AAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCT
CCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCG
GGAAGCGTGGCGCTCTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGC
TCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCCGGTAA
CTATCGTCTTGAGTCCAACCCGGTAGACA

> 018_T7 -- 14..1063 of sequence
ATGCTCCGGCCGCCATGGCGGCCGCGGGAATTCGATTCCTCGGGGGCAGTGAGTGGGTACCT
CGGGCACAGGGCTTGCTGCACTTCTCACACCGCTGTGTTCCATCCTCTGCTGTCACCTCTTGGT
TGTGCAGGGGGCAGACGAGGGTGCAGGATCCCACGTCCGTAGAAAGGTAGTTGTAGGGACA
GGCAGTCACACAGCTAATCACTAGTGAATTCGCGGCCGCCTGCAGGTCGACCATATGGGAGA
GCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTGTCACCTAAATAGCTTGGCGTA
ATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGA
GCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGC
GTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGG
CCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTC
GCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGT
TATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGC
CAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCA
TCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGG
CGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCT
GTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGT

TCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGC
TGCGCCTTATCCGGTAACTATCGTCTTGAGTCCACCCGGTAGACACG

> 019_T7 -- 14..1086 of sequence
ATGCTCCGGCCGCCATGGCGGCCGCGGGAATTCGATTAGCTGTGTCACTGCCTGTCCCTACAA
CTACCTTTCTACGGACGTGGGATCCTGCACCCTCGTCTGCCCCCTGCACAACCAAGAGGTGAC
AGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGAGGTACCCAC
TCACTGCCCCCGAGGAATCACTAGTGAATTCGCGGCCGCCTGCAGGTCGACCATATGGGAGA
GCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTGTCACCTAAATAGCTTGGCGTA
ATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGA
GCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGC
GTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGG
CCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTC
GCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGT
TATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGC
CAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCA
TCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGG
CGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCT
GTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGT
TCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGC
TGCGCCTTATCCGGTAACTATCGTCTTGAGTCCACCCGGTAGACACGACTTATCGCACTGGCA
GCAGCAC

> 019b_T7 -- 9..1063 of sequence
TCGCATGCTCCGGCCGCCATGGCGGCCGCGGGAATTCGATTCCTCGGGGGCAGTGAGTGGGT
ACCTCGGGCACAGGGCTTGCTGCACTTCTCACACCGCTGTGTTCCATCCTCTGCTGTCACCTCTT
GGTTGTGCAGGGGGCAGACGAGGGTGCAGGATCCCACGTCCGTAGAAAGGTAGTTGTAGGG
ACAGGCAGTGACACAGCTAATCACTAGTGAATTCGCGGCCGCCTGCAGGTCGACCATATGGG
AGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTGTCACCTAAATAGCTTGGC
GTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATA
CGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAAT
TGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATC
GGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGAC
TCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACG
GTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAG
GCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAG
CATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCA
GGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATAC
CTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCA
GTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACC
GCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACAC

> 020_T7 -- 16..1050 of sequence
TGCTCCGGCCGCCATGGCGGCCGCGGGAATTCGATTGATGGAACACAGCGGTGTGAGAAGT
GCAGCAAGCCCTGTGCCCGAGTGTGCTATGGTCTGGGCATGGAGCACTTGCGAGAGGTGAG
GGCAGTTACCAGTGCCAATATCCAGGAGTTTGCTGGCTGCAAGAAGATCTTTGGGAGCCTGG
CAAATCACTAGTGAATTCGCGGCCGCCTGCAGGTCGACCATATGGGAGAGCTCCCAACGCGTT

GGATGCATAGCTTGAGTATTCTATAGTGTCACCTAAATAGCTTGGCGTAATCATGGTCATAGCT
GTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAA
GTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCC
CGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGA
GAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGT
TCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAG
GGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAA
GGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACG
CTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAA
GCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCT
TCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTT
CGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGT
AACTATCGTCTTGAGTCCAACCCGGTAGACACG

> 021_T7 -- 15..987 of sequence
CATGCTCCGGCCGCCATGGCGGCCGCGGGAATTCGATTCACAGACACGTTTGAGTCCATGCCC
AATCCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTGACTGCCTGTCCCTACAACTACCTTT
CTACGGACGTGGGATCCTGCACCCTCGTCTGCCCCCTGCACAACCAAGAGGTGACAGCAGAG
GATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGAGGTACCCACTCACTGTC
CCCGAGGCCAGCTGCAGTTCCTGTCCCTCTGCGCATGCAGCCTGGCCCAGCCCACCCTGTCCTA
TCCTTCCTCAGACCCAATCACTAGTGAATTCGCGGCCGCCTGCAGGTCGACCATATGGGAGAG
CTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTGTCACCTAAATAGCTTGGCGTAAT
CATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGC
CGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTT
GCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCA
ACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCT
GCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTAT
CCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAG
GAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCA
CAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGT
TTCCCCTGGAAGCTCCCTCGTGCGCTTCTC

> 021b_T7 -- 14..1028 of sequence
CATGCTCCGGCCGCCATGGCGGCCGCGGGAATTCGATTCCTCGGGGGCAGTGAGTGGGTACC
TAGACGAATGCTAGATGGCCTGGATCCTCATCCCATGCTTCCAAACTACAAAAACTTTAAAGG
CACGATTCAAGAACTTGGTCAAAACCAGTATGCAGTCAGTGGTGAAATATTTGTAGGGACAG
GCAGTGACACAGCTAATCACTAGTGAATTCGCGGCCGCCTGCAGGTCGACCATATGGGAGAG
CTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTGTCACCTAAATAGCTTGGCGTAAT
CATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGC
CGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTT
GCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCA
ACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCT
GCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTAT
CCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAG
GAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCA
CAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGT
TTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCC

GCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGG
TGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCG
CCTTATCCG

> 024_T7 -- 14..1069 of sequence
ATGCTCCGGCCGCCATGGCGGCCGCGGGAATTCGATTCACAGACACGTTTGAGTCCATGCCCA
ATCCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTGACTGCCTGTCCCTACAACTACCTTTC
TACGGACGTGGGATCCTGCACCCTCGTCTGCCCCCTGCACAACCAAGAGGTGACAGCAGAGG
ATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGAGTGTGCTATGATCTGGGC
ATGGAGCACTTGCGAGAGGTGAGGGCAGTTACCAGTGCCAATATCCAGGAGTTTGCTGGCTG
CAAGAAGATCTTTGGGAAATCACTAGTGAATTCGCGGCCGCCTGCAGGTCGACCATATGGGA
GAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTGTCACCTAAATAGCTTGGCG
TAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATAC
GAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATT
GCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATC
GGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGAC
TCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACG
GTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAG
GCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAG
CATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCA
GGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATAC
CTGTCCGCCTTTCTCCCTTCGGGAAGCGTGCGCTTTCTCATAGCTCACGCTGTA

>4a_T7_G01_0266.ab1

TTCGATTAGCGGTGTGAACCTGACCTCTCCTACATGCCATCTGGAAGTTTCCAGATGAGGAGG
GCGCATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAGGGCT
GCCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCGTCTCTGCGGTGGTTGGCATTCTGC
TGGTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCTGAAGATCCGG
AAGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCGCTGACACCTAGCG
GAGCGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGAGCTGAGGAAGGTGAA
GGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATCCCTGATGGGGAGA
ATGTGAAAATTCCAGTGGCCATCAAAGTGTTGAGGGAAAACACATCCCCCAATCACTAGTGAA
TTCGCGGCCGCCTGCAGGTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTG
AGTATTCTATAGTGTCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAA
ATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCT

EXON2 15-19

GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGCTGC
CCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAGGGC
GCATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAGGGCTGC
CCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTTGGCATTCTGCTG
GTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAGAAGATCCGGAA
GTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCGCTGACACCTAGCGGA
GCGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGAGCTGAGGAAGGTGAAGG
TGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATCCCTGATGGGGAGAATG

TGAAAATTCCAGTGGCCATCAAAGTGTTGAGGGAAAACACATCCCCCAAAGCCAACAAAGAA
ATCTTAGAC

FOLDER NAME: 2010-02-01_ABI55_0266

FILE NAME: 6a_T7_A02_0266.ab1

GGCGGCCGCGGGATTCGATTCATGCCCATCTGGAAGTTTCCAGATGAGGAGGGCGCATGCCA
GCCTTGCCCCATCAACTGCACCCACTCCCCTCTGACGTCCATCGTCTCTGCGGTGGTTGGCATT
CTGCTGGTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAGAAGAT
CCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCAATCACTAGTG
AATTCGCGGCCGCCTGCAGGTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCT
TGAGTATTCTATAGTGTCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTG
AAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTG
GGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCG
GGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCG
TATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGA
GCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGG
AAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAANCCGTAAAAAGGCCGCGTTGCT
GGC

EXON 15-18

GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGCTGC
CCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAGGGC
GCATGCCAGCCTTGCCCCATCAACTGCACCCACTCC**TGTGTGGACCTGGATGACAAGGGCTGC**
CCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTTGGCATTCTGCTG
GTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAGAAGATCCGGAA
GTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCGCTGACACCTAGCGGA
GCGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGAGCTGAGGAAGGTGAAGG
TGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAG

FOLDER NAME: 2010-02-01_ABI55_0266

FILE NAME: 7a_T7_B02_0266.ab1

CAGCCTTGCCCCATCAACTGCACCCACTCCCCTCTGACGTCCATCGTCTCTGCGGTGGTTGGCA
TTCAGCTGGTCGTGGTCTTGGGGGTGGTCTTTGGAATCCTCATCAAGCGACGGCAGCAGAAG
ATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCAATCACTAG
TGCGGCCGCCTGCAGGTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAG
TATTCTATAGTGTCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAAT
TGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGT
GCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGA
AACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTAT
TGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGC
GGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAA

AGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGC
GTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAA

EXONS 15-18

GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGCTGC
CCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAGGGC
GCATGCCAGCCTTGCCCCATCAACTGCACCCACTC**CTGTGTGGACCTGGATGACAAGGGCTGC**
CCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTTGGCATTCTGCTG
GTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAGAAGATCCGGAA
GTACACGATGCGGAGACTGCTGCAGGAAACGGAG

FOLDER NAME: 2010-01-29_ABI53_0265

FILE NAME: 1_t7_E11_0265.ab1

CATCTGGAAACTTCCAGATGGGCATGTAGGAGAGGTCAGGTTTCACACCGCTGGGGCAGCGG
GCCACGCAGAAGGGAGGGTCCTTATAGTGGGCACAGGCCACACACTGGTCAGCCTCCGGTCC
AAAACAGGTCACTGAGCCATTCTGGGGCTGACACTCAGGGTGGCACGGCAAACAGTGCCTGG
CATTCACATACTCCCTGGGGAGCCCCTGCAGTACTCGGCATTCCTCCACGCACTCCTGGCCCCG
AAGGAACTGGCTGCAGTTGACACACTGGGTGGGCCCTGGACCCCAGCAGTGCCCTCGGGCGC
ACAGCTGGTGGCAGGCCAGGCCCTCGCCCACACACTCGTCCTCTGGCCGGTTGGCAGTGTGG
AGCAGAGCTTGGTGCGGGTTCCAATCACTAGTGAATTCGCGGCCGCCTGCAGGTCGACCATAT
GGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTGTCACCTAAATAGCTT
GGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAAC
ATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATT
AATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGG

EXONS 12-15

TGGCGCCTACTCGCTGACCCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACTGA
GGGAACTGGGCAGTGGACTGGCCCTCATCCACCATAACACCCACCTCTGCTTCGTGCACACGG
TGCCCTGGGACCAGCTCTTTCGGAACCCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCAG
AGGACGAGTGTGTGGGCGAGGGCCTGGCCTGCCACCAGCTGTGCGCCCGAGGGCACTGCTG
GGGTCCAGGGCCCACCCAGTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGCGTGG
AGGAATGCCGAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGTTTGCCG
TGCCACCCTGAGTGTCAGCCCCAGAATGGCTCAGTGACCTGTTTTGGACCGGAGGCTGACCAG
TGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGCTGCCCCAGCGGTGTG
AAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAGGGCGCATGCCAGCCT
TGCCCCATCAACTGCACCCACTC

FOLDER NAME: 2010-01-29_ABI53_0265

FILE NAME: 4_t7_H11_0265.ab1

CATGGCGGCCGCGGGATTCGATTGGGGGATGTGTTTTCCCTCAACACTTTGATGGCCACTGGA
ATTTTCACATTCTCCCCATCAGGGATCCAGATGCCCTTGTAGACTGTGCCAAAAGCGCCAGATC

CAAGCACCTTCACCTTCCTCAGCTCCGTCTCTTTCAGGATCCGCATCTGCGCCTGGTTGGGCAT
CGCTCCGTTAGGTGTCAGCGGCTCCACCAGCTCCGTTTCCTGCAGCAGTCTCCGCATCGTGTAC
TTCCGGATCTTCTGCTGCCGTCGCTTGATGAGGATCCCAAAGACCACCCCCAAGACCACGACC
AGCAGAATGCCAACCACCGCAGAGACGATGGACGTCAGAGGGCTGGCTCTCTGCTCGGCGG
GGCAGCCCTTGTCATCCAGGTCCACACAGGAGTGGGTGCAGTTGATGGGGCAAGGCTGGCAT
GCGCCCTCCTCATCTGGAAACTTCCAGATGGGCATGTAGGAGAGGTCAGGTTTCACACCGCTA
ATCACTAGTGAATTCGCGGCCGCCTGCAGGTCGACCATATGGGAGAGCTCCCAACGCGTTGG
ATGCATAGCTTGAGTATTCTATAGTGTCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGT
TTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGT
GTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCG
CTTTCCAGTCGG

EXONS 15-19

GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGCTGC
CCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAGGGC
GCATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAGGGCTGC
CCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTTGGCATTCTGCTG
GTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAGAAGATCCGGAA
GTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCGCTGACACCTAGCGGA
GCGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGAGCTGAGGAAGGTGAAGG
TGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATCCCTGATGGGGAGAATG
TGAAAATTCCAGTGGCCATCAAAGTGTTGAGGGAAAACACATCCCCCAAAGCCAACAAAGAA
ATCTTAGAC

FOLDER NAME: 2010-01-29_ABI53_0265

FILE NAME: 5_t7_A12_0265.ab1

CGCCATGGCGGCCGCGGGAATTCGATTAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATC
TGGAAGTTTCCAGATGAGGAGGGCGCATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGT
GTGGACCTGGATGACAAGGGCTGCCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCGT
CTCTGCGGTGGTTGGCATTCTGCTGGTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAA
GCGACGGCAGCAGAAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTG
GTGGAGCCGCTGACACCTAGCGGAGCGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAG
AGACGGAGCTGAGGAAGGGCATCTGGATCCCTGATGGGGAGAATGTGAAAATTCCAGTGGC
CATCAAAGTGTTGAGGGAAAACACATCCCCCAATCACTAGTGAATTCGCGGCCGCCTGCAGGT
CGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTGTCACCT
AAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATT
CCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTA
ACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTG
CATTAATGAATCGGCCAACGCGCGGGGAGAGGCGCGCTC

EXONS 15-19

GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGCTGC
CCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAGGGC
GCATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAGGGCTGC

CCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTTGGCATTCTGCTG
GTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAGAAGATCCGGAA
GTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCGCTGACACCTAGCGGA
GCGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGAGCTGAGGAAGG**TGAAG**
GTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATCCCTGATGGGGAGAA
TGTGAAAATTCCAGTGGCCATCAAAGTGTTGAGGGAAAACACATCCCCCAAAGCCAACAAAG
AAATCTTAGAC

FOLDER NAME: 2010-01-29_ABI53_0265

FILE NAME: 7_t7_C12_0265.ab1

CAGCCTTGCCCCATCAACTGCACCCACTCCCCTCTGACGTCCATCATCTCTGCGGTGGTTGGCA
TTCTGCTGGTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAGAAG
ATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCAATCACTAG
TGCGGCCGCCTGCAGGTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAG
TATTCTATAGTGTCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAAT
TGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGT
GCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGA
AACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTAT
TGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGC
GGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCC

EXONS 15-18

GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGCTGC
CCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAGGGC
GCATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAGGGCTGC
CCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTTGGCATTCTGCTG
GTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAGAAGATCCGGAA
GTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCGCTGACACCTAGCGGA
GCGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGAGCTGAGGAAGGTGAAGG
TGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAG

FOLDER NAME: 2010-01-29_ABI53_0265

FILE NAME: 8_t7_D12_0265.ab1

GGACCTGGATGACAAGGGCTGCCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCGTCT
CTGCGGTGGTTGGCATTCTGCTGGTCGTGGTCTTGGGGGTGGTCTAATCACTAGTGCGGCCGC
CTGCAGGTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAG

TGTCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGC
TCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGA
GTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGT
GCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCT
TCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCT
CACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTG
AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATA
GGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCG
ACAGGACTATAAAGATACCAGGCGTTTCCCCC

EXONS 15-18

GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGCTGC
CCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAGGGC
GCATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAGGGCTGC
CCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTTGGCATTCTGCTG
GTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAGAAGATCCGGAA
GTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCGCTGACACCTAGCGGA
GCGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGAGCTGAGGAAGGTGAAGG
TGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAG

FOLDER NAME: 2009-12-14_abi54_0258

FILE NAME: 4_C07_0258.ab1

GAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAGGGCGCATGCCAGCC
TTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAGGGCTGCCCCGCCGAGCA
GAGAGCCAGCCCTCTGACGTCCATCGTCTCTGCGGTGGTTGGCATTCTGCTGGTCGTGGTCTT
GGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAGAAGATCCGGAAGTACACGATGC
GGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCGCTGACACCTAACGGAGCGATGCCCAA
CCAGGCGCAGATGCGGATCCTGAAAGAGACGGAGCTGAGGAAGGTGAAGGTGCTTGGATCT
GGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATCCCTGATGGGGAGAATGTGAAAATTCC
AGTGGCCATCAAAGTGTTGAGGGAAAACACATCCCCCAATCGAATTCCCGCGGCCGCCATGN
CGGNCGGGAGCATGCGACGT

EXONS 15-18

GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGCTGC
CCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAGGGC
GCATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAGGGCTGC
CCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTTGGCATTCTGCTG
GTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAGAAGATCCGGAA
GTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCGCTGACACCTAGCGGA
GCGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGAGCTGAGGAAGGTGAAGG
TGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATCCCTGATGGGGAGAATG
TGAAAATTCCAGTGGCCATCAAAGTGTTGAGGGAAAACACATCCCCCAAAGCCAACAAAGAA
ATCTTAGAC

FOLDER NAME: 2009-12-14_abi54_0258

FILE NAME: 5_D07_0258.ab1

ATTCTCCCCATCAGGGATCCAGATGCCCTTCCTCAGCTCCGTCTCTTTCAGGATCCGCATCTGC
GCCTGGTTGGGCATCGCTCCGCTAGGTGTCAGCGGCTCCACCAGCTCCGTTTCCTGCAGCAGT
CTCCGCATCGTGTACTTCCGGATCTTCTGCTGCCGTCGCTTGATGAGGATCCCAAAGACCACCC
CCAAGACCACGACCAGCAGAATGCCAACCACCGCAGAGACGATGGACGTCAGAGGGCTGGC
TCTCTGCTCGGCGGGGCAGCCCTTGTCATCCAGGTCCACACAGGAGTGGGTGCAGTTGATGG
GGCAAGGCTGGCATGCGCCCTCCTCATCTGGAAACTTCCAGATGGGCATGTAGGAGAGGTCA
GGTTTCACACCGCTAATCGAATTCCCGCGGCCGCCATGGCGGCCGGGAGCATGCGACGTCGG
GCCCAATTCGCCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGAC
TGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCA

EXON 15-19

GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGCTGC
CCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAGGGC
GCATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAGGGCTGC
CCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTTGGCATTCTGCTG
GTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAGAAGATCCGGAA
GTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCGCTGACACCTAGCGGA
GCGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGAGCTGAGGAAGGTGAAGG
TGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATCCCTGATGGGGAGAATG
TGAAAATTCCAGTGGCCATCAAAGTGTTGAGGGAAAACACATCCCCCAAAGCCAACAAAGAA
ATCTTAGAC

FOLDER NAME: 2009-12-14_abi54_0258

FILE NAME: 7_F07_0258.ab1

TCCTGCAGCAGTCTCCGCATCGTGTACTTCCGGATCTTCTGCTGCCGTCGCTTGATGAGGATCC
CAAAGACCACCCCCAAGACCACGACCAGCAGAATGCCAACCACCGCAGAGATGATGGACGTC
AGAGGGGAGTGGGTGCAGTTGATGGGGCAAGGCTGGCATGCGCCCTCCTCATCTGGAAACTT
CCAGATGGGCATGAATCCCGCGGCCATGGCGGCCGGGAGCATGCGACGTCGGGCCCAATTC
GCCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAA
CCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGC
GAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCG
CCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACT
TG

EXON 15-18

GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGCTGC
CCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAGGGC
GCATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAGGGCTGC
CCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTTGGCATTCTGCTG

GTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAGAAGATCCGGAA
GTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCGCTGACACCTAGCGGA
GCGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGAGCTGAGGAAGGTGAAGG
TGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAG

FOLDER NAME: 2009-12-14_abi54_0258

FILE NAME: 8_G07_0258.ab1

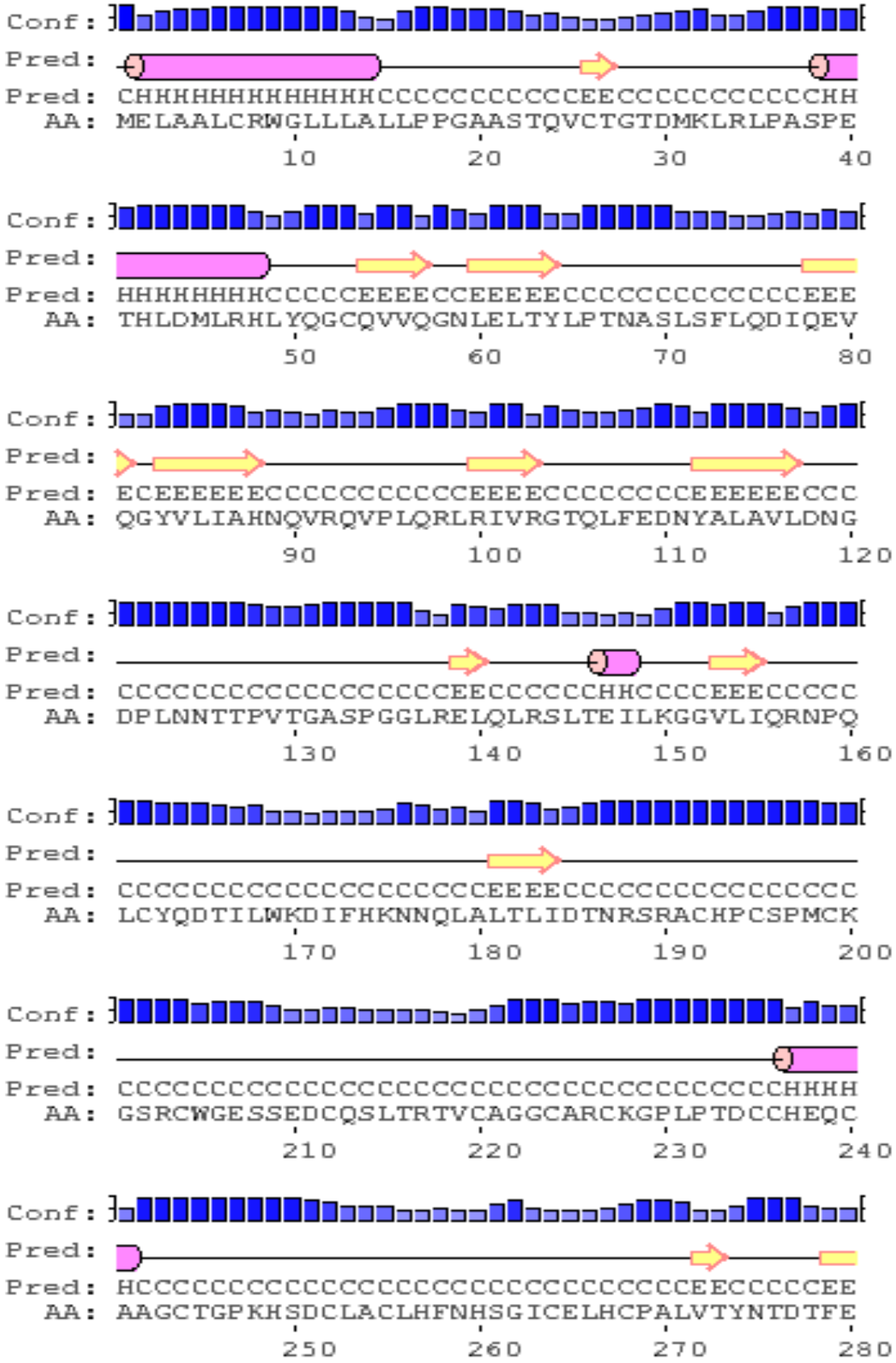
CAGAGGGCTGGCTCTCTGCTCGGCGGGGCAGCCCTTGTCATCCAGGTCCACACAGGAGTGGG
TGCAGTTGATGGGGCAAGGCTGGCATAATCCCGCGGCCATGGCGGCCGGGAGCATGCGACG
TCGGGCCCAATTCGCCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCG
TGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGC
TGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGG
CGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCG
TGACCGCTACACTTGCCAGCGCCCTA

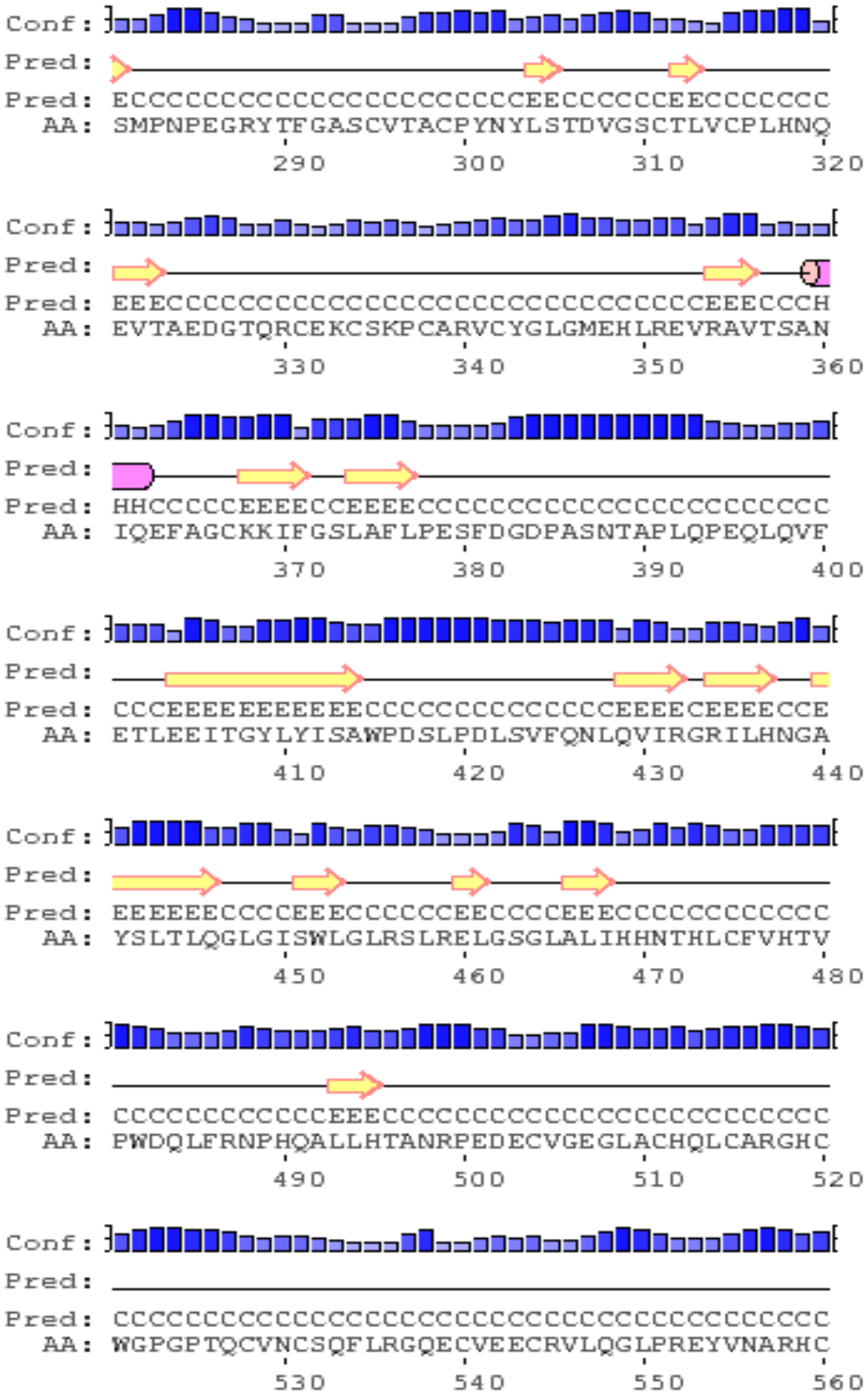
EXON 15-18

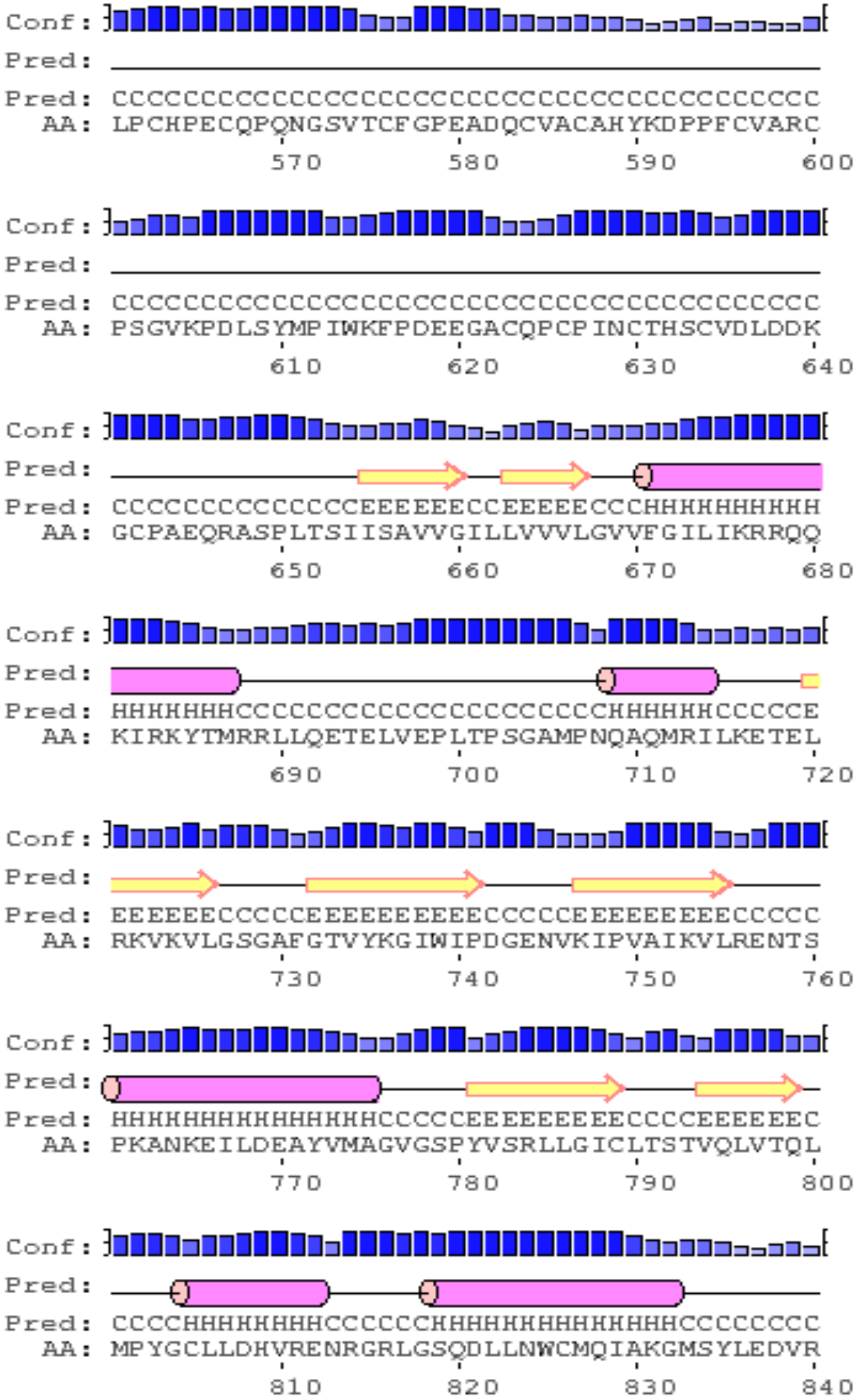
GAGGCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGCTGC
CCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAGGGC
GCATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAGGGCTGC
CCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTTGGCATTCTGCTG
GTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAGAAGATCCGGAA
GTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCGCTGACACCTAGCGGA
GCGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGAGCTGAGGAAGGTGAAGG
TGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAG

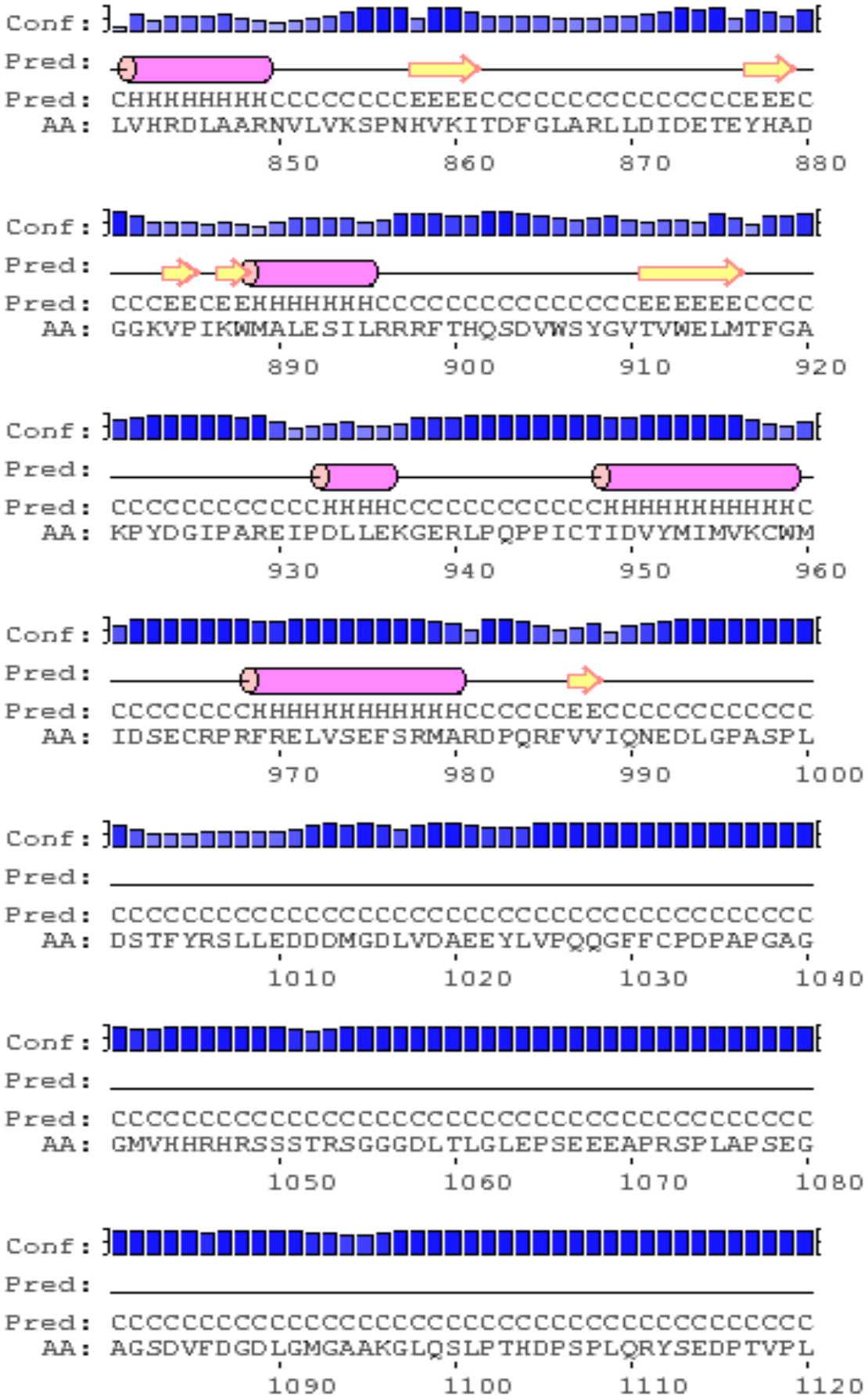
CCGCGGCCGCCATGGCGGCCGGGAGCATGCGACGTCGGGCCCAATTCGCCCTATAGTGAGTC
GTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCA
ACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCAC
CGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGC
ATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAG
CGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTC
TAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACT
TGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGAC
GTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATC
TCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGC
TGA

Secondary structure of *HER2* protein




Conf:
Pred:
Pred: ECCCCCCCCCCCCCCCCCCCCCCEECCCCCCEECCCCCCC
AA: SMPNPEGRYTFGASCVTACPYNYLSTDVGSCTLVCPLHNQ
290 300 310 320
Conf:
Pred:
Pred: EEECCCCCCCCCCCCCCCCCCCCCCCCCCCCCCEEECCCH
AA: EVTAEDGTQRCEKCSKPCARVCYGLGMEHLREVRAVTSAN
330 340 350 360
Conf:
Pred:
Pred: HHCCCCCEEEECCEEEECCCCCCCCCCCCCCCCCCCCCCC
AA: IQEFAGCKKIFGSLAFLPESFDGDPASNTAPLQPEQLQVF
370 380 390 400
Conf:
Pred:
Pred: CCCEEEEEEEEEEECCCCCCCCCCCCCCEEEECEEEECCE
AA: ETLEEITGYLYISAWPDSLPDLSVFQNLQVIRGRILHNGA
410 420 430 440
Conf:
Pred:
Pred: EEEEEECCCCEEECCCCCCEECCCCEEECCCCCCCCCCCC
AA: YSLTLQGLGISWLGLRSLRELGSGLALIHHNTHLCFVHTV
450 460 470 480
Conf:
Pred:
Pred: CCCCCCCCCCCCEEECCCCCCCCCCCCCCCCCCCCCCCCC
AA: PWDQLFRNPHQALLHTANRPEDECVGEGLACHQLCARGHC
490 500 510 520
Conf:
Pred:
Pred: CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
AA: WGPGPTQCVNCSQFLRGQECVEECRVLQGLPREYVNARHC
530 540 550 560


Conf:
Pred:
Pred: CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
AA: LPCHPECQPQNGSVTCFGPEADQCVACAHYKDPPFCVARC
570 580 590 600
Conf:
Pred:
Pred: CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
AA: PSGVKPDLSYMPIWKFPDEEGACQPCPINCTHSCVDLDDK
610 620 630 640
Conf:
Pred:
Pred: CCCCCCCCCCCCCCEEEEEECCEEEEECCCHHHHHHHHHH
AA: GCPAEQRASPLTSIISAVVGILLVVVLGVVFGILIKRRQQ
650 660 670 680
Conf:
Pred:
Pred: HHHHHHHCCCCCCCCCCCCCCCCCCCCHHHHHCCCCCE
AA: KIRKYTMRRLLQETELVEPLTPSGAMPNQAQMRILKETEL
690 700 710 720
Conf:
Pred:
Pred: EEEEEECCCCCEEEEEEEEECCCCCEEEEEEEECCCCC
AA: RKVKVLGSGAFGTVYKGIWIPDGENVKIPVAIKVLRENTS
730 740 750 760
Conf:
Pred:
Pred: HHHHHHHHHHHHHHCCCCCEEEEEEEECCCCEEEEEEC
AA: PKANKEILDEAYVMAGVGSPYVSRLLGICLTSTVQLVTQL
770 780 790 800
Conf:
Pred:
Pred: CCCCHHHHHHHHCCCCCCHHHHHHHHHHHHHCCCCCCCC
AA: MPYGCLLDHVRENRGRLGSQDLLNWCMQIAKGMSYLEDVR
810 820 830 840


Conf:
Pred:
Pred: CHHHHHHHHCCCCCCCCEEEECCCCCCCCCCCCCCCCEEEC
AA: LVHRDLAARNVLVKSPNHVKITDFGLARLLDIDETEYHAD
850 860 870 880
Conf:
Pred:
Pred: CCCEECEEHHHHHHHCCCCCCCCCCCCCCCEEEEEECCCC
AA: GGKVPIKWMALESILRRRFTHQSDVWSYGVTVWELMTFGA
890 900 910 920
Conf:
Pred:
Pred: CCCCCCCCCCCCHHHHCCCCCCCCCCCCHHHHHHHHHHHC
AA: KPYDGIPAREIPDLLEKGERLPQPPICTIDVYMIMVKCWM
930 940 950 960
Conf:
Pred:
Pred: CCCCCCCCHHHHHHHHHHHHCCCCCCEECCCCCCCCCCCC
AA: IDSECRPRFRELVSEFSRMARDPQRFVVIQNEDLGPASPL
970 980 990 1000
Conf:
Pred:
Pred: CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
AA: DSTFYRSLLEDDDMGDLVDAEEYLVPQQGFFCPDPAPGAG
1010 1020 1030 1040
Conf:
Pred:
Pred: CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
AA: GMVHHRHRSSSTRSGGGDLTLGLEPSEEEAPRSPLAPSEG
1050 1060 1070 1080
Conf:
Pred:
Pred: CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
AA: AGSDVFDGDLGMGAAKGLQSLPTHDPSPLQRYSEDPTVPL
1090 1100 1110 1120

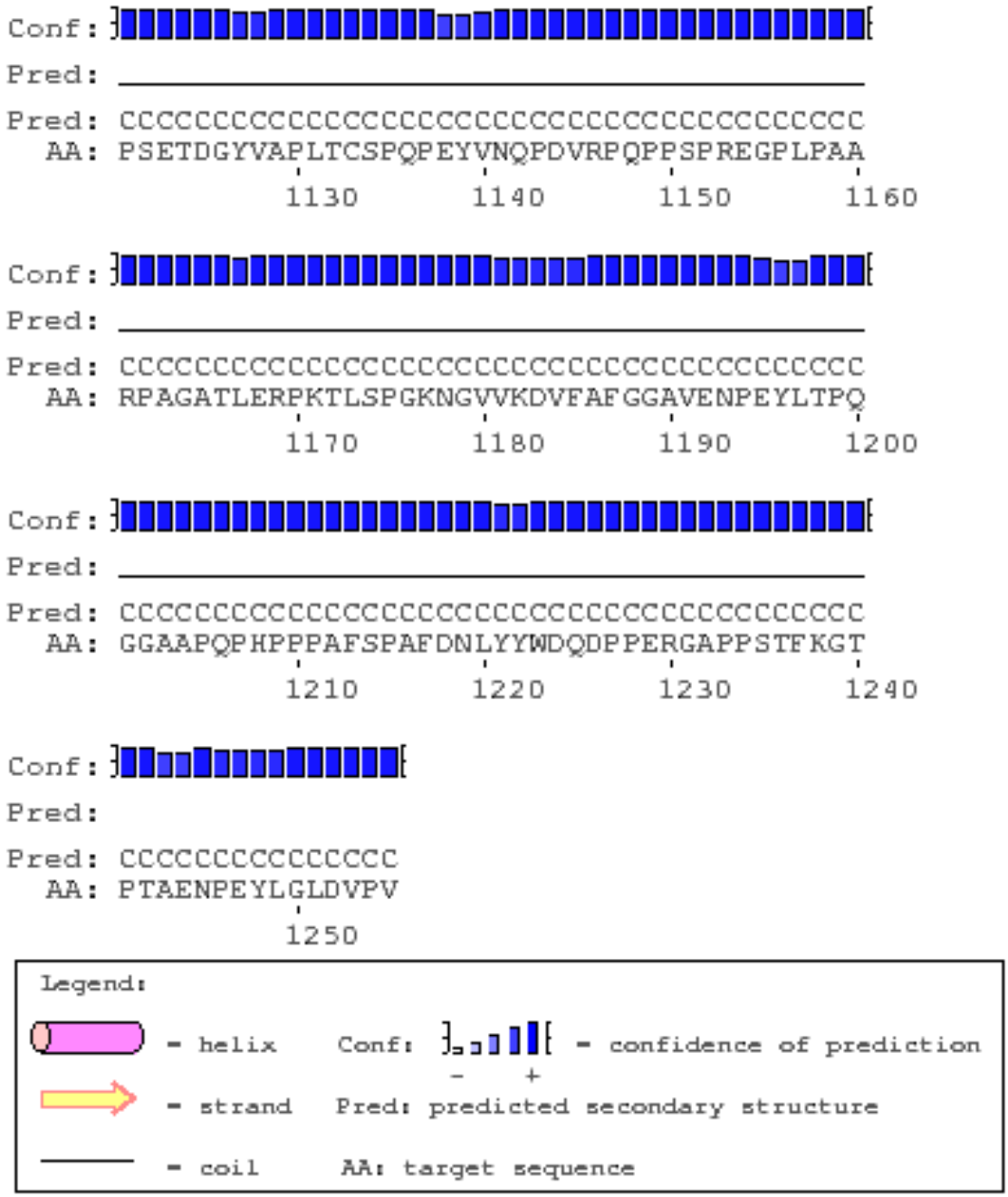


Figure A 9: PSIPRED output showing the secondary structure of the HER2 protein aligned with the amino acid sequence. α-helices are represented by pink cylinders, β-strands are represented by yellow arrows, and coils are represented by black lines. The predicted structures are also indicated by lettering; H – alpha helix, E -Beta sheet and C -coils.

pGEM-T Easy Vector Map

The pGEM-T Easy is a high copy number vector for cloning PCR products. It contains both T7 and SP6 RNA polymerase promoters flanking a multiple cloning region within the α-peptide coding region of the enzyme β-galactosidase to enable blue/white screening. It has multiple restriction enzyme sites and an ampicillin resistance gene to facilitate bacterial selection.

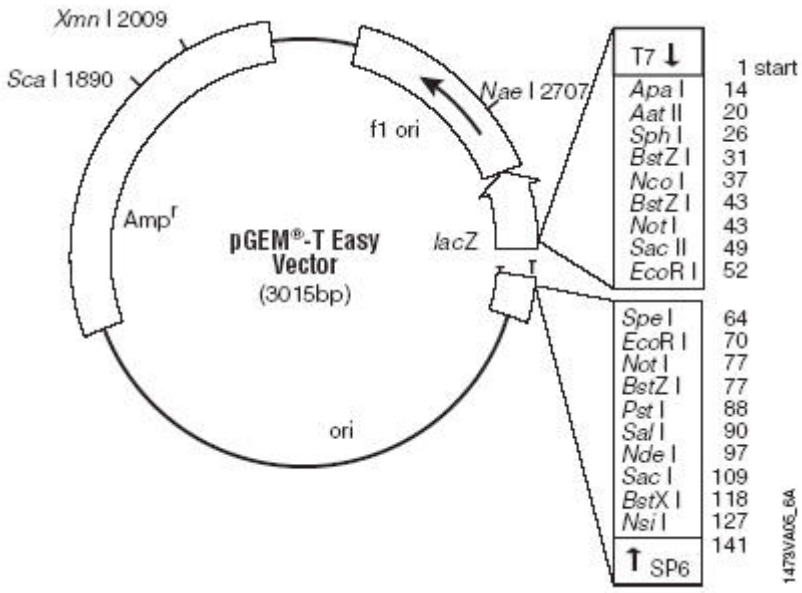


Figure A 10: pGEM-T Easy Vector map

APPENDIX B:

CLUSTAL OMEGA ALIGNMENTS OF *HER2* mRNA TRANSCRIPT VARIANTS OBTAINED FROM THE NCBI DATABASE:

```
NR_110535    ------------------------------------------------------------
NM_004448    ------------------------------------------------------------
NM_001289937 ------------------------------------------------------------
NM_001289936 AAGTTCCTGTGTTCTTTATTCTACTCTCCGCTGAAGTCCACACAGTTTAAATTAAAGTTC
NM_001005862 AAGTTCCTGTGTTCTTTATTCTACTCTCCGCTGAAGTCCACACAGTTTAAATTAAAGTTC
NM_001289938 AAGTTCCTGTGTTCTTTATTCTACTCTCCGCTGAAGTCCACACAGTTTAAATTAAAGTTC


NR_110535    ------------------------------------------------------------
NM_004448    ------------------------------------------------------------
NM_001289937 ------------------------------------------------------------
NM_001289936 CCGGATTTTTGTGGGCGCCTGCCCCGCCCCTCGTCCCCCTGCTGTGTCCATATATCGAGG
NM_001005862 CCGGATTTTTGTGGGCGCCTGCCCCGCCCCTCGTCCCCCTGCTGTGTCCATATATCGAGG
NM_001289938 CCGGATTTTTGTGGGCGCCTGCCCCGCCCCTCGTCCCCCTGCTGTGTCCATATATCGAGG

NR_110535    ------------------------------------------------------------
NM_004448    ------------------------------------------------------------
NM_001289937 ------------------------------------------------------------
NM_001289936 CGATAGGGTTAAGGGAAGGCGGACGCCTGATGGGTTAATGAGCAAACTGAAGTGTTTTCC
NM_001005862 CGATAGGGTTAAGGGAAGGCGGACGCCTGATGGGTTAATGAGCAAACTGAAGTGTTTTCC
NM_001289938 CGATAGGGTTAAGGGAAGGCGGACGCCTGATGGGTTAATGAGCAAACTGAAGTGTTTTCC

NR_110535    ------------------------------------------------------------
NM_004448    ------------------------------------------------------------
NM_001289937 ------------------------------------------------------------
NM_001289936 ATGATCTTTTTTGAGTCGCAATTGAAGTACCACCTCCCGAGGGTGATTGCTTCCCCATGC
NM_001005862 ATGATCTTTTTTGAGTCGCAATTGAAGTACCACCTCCCGAGGGTGATTGCTTCCCCATGC
NM_001289938 ATGATCTTTTTTGAGTCGCAATTGAAGTACCACCTCCCGAGGGTGATTGCTTCCCCATGC

NR_110535    -------------------------------------------------------GCTTGCT
NM_004448    -------------------------------------------------------GCTTGCT
NM_001289937 -------------------------------------------------------GCTTGCT
NM_001289936 GGGGTAGAACCTTTGCTGTCCTGTTCACCACTCTACCTCCAGCACAGAATTTGGCTTATG
NM_001005862 GGGGTAGAACCTTTGCTGTCCTGTTCACCACTCTACCTCCAGCACAGAATTTGGCTTATG
NM_001289938 GGGGTAGAACCTTTGCTGTCCTGTTCACCACTCTACCTCCAGCACAGAATTTGGCTTATG
                                                                   ****.

NR_110535    CCCAATCACAGGAGAAGGAGGAGGTGGAGGAGGAGGGCTGCTTGAGGAAGTATAAGAATG
NM_004448    CCCAATCACAGGAGAAGGAGGAGGTGGAGGAGGAGGGCTGCTTGAGGAAGTATAAGAATG
NM_00128993  CCCAATCACAGGAGAAGGAGGAGGTGGAGGAGGAGGGCTGCTTGAGGAAGTATAAGAATG
NM_001289936 CCTACTCAAT-GTGAAGA----TGATGAGGATGA---AAACC------------------
NM_001005862 CCTACTCAAT-GTGAAGA----TGATGAGGATGA---AAACC------------------
NM_001289938 CCTACTCAAT-GTGAAGA----TGATGAGGATGA---AAACC------------------
                  ** *.***.: *:****.     *: ***** **   .:.*

NR_110535    AAGTTGTGAAGCTGAGATTCCCCTCCATTGGGACCGGAGAAAC-----------------
NM_004448    AAGTTGTGAAGCTGAGATTCCCCTCCATTGGGACCGGAGAAAC-----------------
NM_001289937 AAGTTGTGAAGCTGAGATTCCCCTCCATTGGGACCGGAGA---AAC--------------
NM_001289936 --TTTGTGATGATCCACTTCCACTT---AATGAATGGTGGCAAAGCAAAGCTATATTCAA
NM_001005862 --TTTGTGATGATCCACTTCCACTT---AATGAATGGTGGCAAAGCAAAGCTATATTCAA
NM_001289938 --TTTGTGATGATCCACTTCCACTT---AATGAATGGTGGCAAAGCAAAGCTATATTCAA
                ******:*.* ...****.**    :. **. **:*.
```

```
NR_110535    ---------CAGGGGAGCCCCCCGGGCAGCCGCGCGCCCCTTCC--CA--CGGGGCCCTT
NM_004448    ---------CAGGGGAGCCCCCCGGGCAGCCGCGCGCCCCTT--CCCACGG--GGCCCTT
NM_001289937 ---------CAGGGGAGCCCCCCGGGCAGCCGCGCGCCCC--TTCCCACGGGG-CCC-TT
NM_001289936 GACCACATGCAAAGCTACTCCCTGAGCAA---AGAGTCACAGATAAAACGGGGGCACCAG
NM_001005862 GACCACATGCAAAGCTACTCCCTGAGCAA---AGAGTCACAGATAAAACGGGGGCACCAG
NM_001289938 GACCACATGCAAAGCTACTCCCTGAGCAA---AGAGTCACAGATAAAACGGGGGCACCAG
                     **..* :.* *** *.***.   .*.* *.*      .*       .* :

NR_110535    TACTGCGCCGCGCGCCCGGCCCCCACCCCTCGCAGCACCCCGCGCCCCGCGCCCTCCCAG
NM_004448    TACTGCGCCGCGCGCCCGGCCCCCACCCCTCGCAGCACCCCGCGCCCCGCGCCCTCCCAG
NM_001289937 TACTGCGCCGCGCGCCCGGCCCCCACCCCTCGCAGCACCCCGCGCCCCGCGCCCTCCCAG
NM_001289936 TAG-------AATGGCCAGGACAAACGCAGTGCAGCAC--AGAGACTCAGACCCTGGCAG
NM_001005862 TAG-------AATGGCCAGGACAAACGCAGTGCAGCAC--AGAGACTCAGACCCTGGCAG
NM_001289938 TAG-------AATGGCCAGGACAAACGCAGTGCAGCAC--AGAGACTCAGACCCTGGCAG
             **        .. * **.* .*..** *.  *******  .*.*.* *. .****  ***

NR_110535    CCGGGTCCAGCCGGAGCCATGGGGCCGGAGCCGCAGTGAGCACCATGGA-GCTGGCGGCC
NM_004448    CCGGGTCCAGCCGGAGCCATGGGGCCGGAGCCGCAGTGAGCACCATGGA-GCTGGCGGCC
NM_001289937 CCGGGTCCAGCCGGAGCCATGGGGCCGGAGCCGCAGTGAGCACCATGGA-GCTGGCGGCC
NM_001289936 CCAT-----GCCTGCGCA--------GGCAGTGATGAGAGTGACATGTACTGTTGTGGAC
NM_001005862 CCAT-----GCCTGCGCA--------GGCAGTGATGAGAGTGACATGTACTGTTGTGGAC
NM_001289938 CCAT-----GCCTGCGCA--------GGCAGTGATGAGAGTGACATGTACTGTTGTGGAC
             **.      *** *.**.        **..  *.:*:*** ..**** *   * * **.*

NR_110535    TTGTGCCGCTGGGGGCTCCTCCT--------CGCCCTCTTGCCCCCCGGAGCCGCGAGCA
NM_004448    TTGTGCCGCTGGGGGCTCCTCCT--------CGCCCTCTTGCCCCCCGGAGCCGCGAGCA
NM_001289937 TTGTGCCGCTGGGGGCTCCTCCT--------CGCCCTCTTGCCCCCCGGAGCCGCGAGCA
NM_001289936 ATGCACAAAAGTGAGATACTTCAAAGATTCCAGAAGATATGCCCCGGGGGTCCTGGAAGC
NM_001005862 ATGCACAAAAGTGA-----------------------------------------------
NM_001289938 ATGCACAAAAGTGA-----------------------------------------------
                  :** .*...:* *.

NR_110535    CCCAAGTGTGCACCGGCACAGACATGAAGCTGCGGCTCCCTGCCAGTCCCGAGACCCACC
NM_004448    CCCAAGTGTGCACCGGCACAGACATGAAGCTGCGGCTCCCTGCCAGTCCCGAGACCCACC
NM_001289937 CCCAAGTGTGCACCGGCACAGACATGAAGCTGCGGCTCCCTGCCAGTCCCGAGACCCACC
NM_001289936 CACAAGTGTGCACCGGCACAGACATGAAGCTGCGGCTCCCTGCCAGTCCCGAGACCCACC
NM_001005862 -----GTGTGCACCGGCACAGACATGAAGCTGCGGCTCCCTGCCAGTCCCGAGACCCACC
NM_001289938 -----GTGTGCACCGGCACAGACATGAAGCTGCGGCTCCCTGCCAGTCCCGAGACCCACC
                  *****************************************************

NR_110535    TGGACATGCTCCGCCACCTCTACCAGGGCTGCCAGGTGGTGCAGGGAAACCTGGAACTCA
NM_004448    TGGACATGCTCCGCCACCTCTACCAGGGCTGCCAGGTGGTGCAGGGAAACCTGGAACTCA
NM_001289937 TGGACATGCTCCGCCACCTCTACCAGGGCTGCCAGGTGGTGCAGGGAAACCTGGAACTCA
NM_001289936 TGGACATGCTCCGCCACCTCTACCAGGGCTGCCAGGTGGTGCAGGGAAACCTGGAACTCA
NM_001005862 TGGACATGCTCCGCCACCTCTACCAGGGCTGCCAGGTGGTGCAGGGAAACCTGGAACTCA
NM_001289938 TGGACATGCTCCGCCACCTCTACCAGGGCTGCCAGGTGGTGCAGGGAAACCTGGAACTCA
             ************************************************************

NR_110535    CCTACCTGCCCACCAATGCCAGCCTGTCCTTCCTGCAGGTGAGGCCCGTGGGCAACCCAG
NM_004448    CCTACCTGCCCACCAATGCCAGCCTGTCCTTCCTGCAG----------------------
NM_001289937 CCTACCTGCCCACCAATGCCAGCCTGTCCTTCCTGCAG----------------------
NM_001289936 CCTACCTGCCCACCAATGCCAGCCTGTCCTTCCTGCAG----------------------
NM_001005862 CCTACCTGCCCACCAATGCCAGCCTGTCCTTCCTGCAG----------------------
NM_001289938 CCTACCTGCCCACCAATGCCAGCCTGTCCTTCCTGCAG----------------------
             **************************************

NR_110535    CCAGGCCCTGCCTCCAGCTGGGCTGAGCCCTCTGTTTACAGGATATCCAGGAGGTGCAGG
NM_004448    ------------------------------------------GATATCCAGGAGGTGCAGG
NM_001289937 ------------------------------------------GATATCCAGGAGGTGCAGG
NM_001289936 ------------------------------------------GATATCCAGGAGGTGCAGG
NM_001005862 ------------------------------------------GATATCCAGGAGGTGCAGG
NM_001289938 ------------------------------------------GATATCCAGGAGGTGCAGG
                                                         *******************

NR_110535    GCTACGTGCTCATCGCTCACAACCAAGTGAGGCAGGTCCCACTGCAGAGGCTGCGGATTG
NM_004448    GCTACGTGCTCATCGCTCACAACCAAGTGAGGCAGGTCCCACTGCAGAGGCTGCGGATTG
NM_001289937 GCTACGTGCTCATCGCTCACAACCAAGTGAGGCAGGTCCCACTGCAGAGGCTGCGGATTG
NM_001289936 GCTACGTGCTCATCGCTCACAACCAAGTGAGGCAGGTCCCACTGCAGAGGCTGCGGATTG
```

```
NM_001005862 GCTACGTGCTCATCGCTCACAACCAAGTGAGGCAGGTCCCACTGCAGAGGCTGCGGATTG
NM_001289938 GCTACGTGCTCATCGCTCACAACCAAGTGAGGCAGGTCCCACTGCAGAGGCTGCGGATTG
             ************************************************************

NR_110535    TGCGAGGCACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGCTAGACAATGGAGACC
NM_004448    TGCGAGGCACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGCTAGACAATGGAGACC
NM_001289937 TGCGAGGCACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGCTAGACAATGGAGACC
NM_001289936 TGCGAGGCACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGCTAGACAATGGAGACC
NM_001005862 TGCGAGGCACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGCTAGACAATGGAGACC
NM_001289938 TGCGAGGCACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGCTAGACAATGGAGACC
             ************************************************************

NR_110535    CGCTGAACAATACCACCCCTGTCACAGGGGCCTCCCCAGGAGGCCTGCGGGAGCTGCAGC
NM_004448    CGCTGAACAATACCACCCCTGTCACAGGGGCCTCCCCAGGAGGCCTGCGGGAGCTGCAGC
NM_001289937 CGCTGAACAATACCACCCCTGTCACAGGGGCCTCCCCAGGAGGCCTGCGGGAGCTGCAGC
NM_001289936 CGCTGAACAATACCACCCCTGTCACAGGGGCCTCCCCAGGAGGCCTGCGGGAGCTGCAGC
NM_001005862 CGCTGAACAATACCACCCCTGTCACAGGGGCCTCCCCAGGAGGCCTGCGGGAGCTGCAGC
NM_001289938 CGCTGAACAATACCACCCCTGTCACAGGGGCCTCCCCAGGAGGCCTGCGGGAGCTGCAGC
             ************************************************************

NR_110535    TTCGAAGCCTCACAGAGATCTTGAAAGGAGGGGTCTTGATCCAGCGGAACCCCCAGCTCT
NM_004448    TTCGAAGCCTCACAGAGATCTTGAAAGGAGGGGTCTTGATCCAGCGGAACCCCCAGCTCT
NM_001289937 TTCGAAGCCTCACAGAGATCTTGAAAGGAGGGGTCTTGATCCAGCGGAACCCCCAGCTCT
NM_001289936 TTCGAAGCCTCACAGAGATCTTGAAAGGAGGGGTCTTGATCCAGCGGAACCCCCAGCTCT
NM_001005862 TTCGAAGCCTCACAGAGATCTTGAAAGGAGGGGTCTTGATCCAGCGGAACCCCCAGCTCT
NM_001289938 TTCGAAGCCTCACAGAGATCTTGAAAGGAGGGGTCTTGATCCAGCGGAACCCCCAGCTCT
             ************************************************************

NR_110535    GCTACCAGGACACGATTTTGTGGAAGGACATCTTCCACAAGAACAACCAGCTGGCTCTCA
NM_004448    GCTACCAGGACACGATTTTGTGGAAGGACATCTTCCACAAGAACAACCAGCTGGCTCTCA
NM_001289937 GCTACCAGGACACGATTTTGTGGAAGGACATCTTCCACAAGAACAACCAGCTGGCTCTCA
NM_001289936 GCTACCAGGACACGATTTTGTGGAAGGACATCTTCCACAAGAACAACCAGCTGGCTCTCA
NM_001005862 GCTACCAGGACACGATTTTGTGGAAGGACATCTTCCACAAGAACAACCAGCTGGCTCTCA
NM_001289938 GCTACCAGGACACGATTTTGTGGAAGGACATCTTCCACAAGAACAACCAGCTGGCTCTCA
             ************************************************************

NR_110535    CACTGATAGACACCAACCGCTCTCGGGCCTGCCACCCCTGTTCTCCGATGTGTAAGGGCT
NM_004448    CACTGATAGACACCAACCGCTCTCGGGCCTGCCACCCCTGTTCTCCGATGTGTAAGGGCT
NM_001289937 CACTGATAGACACCAACCGCTCTCGGGCCTGCCACCCCTGTTCTCCGATGTGTAAGGGCT
NM_001289936 CACTGATAGACACCAACCGCTCTCGGGCCTGCCACCCCTGTTCTCCGATGTGTAAGGGCT
NM_001005862 CACTGATAGACACCAACCGCTCTCGGGCCTGCCACCCCTGTTCTCCGATGTGTAAGGGCT
NM_001289938 CACTGATAGACACCAACCGCTCTCGGGCCTGCCACCCCTGTTCTCCGATGTGTAAGGGCT
             ************************************************************

NR_110535    CCCGCTGCTGGGGAGAGAGTTCTGAGGATTGTCAGAGCCTGACGCGCACTGTCTGTGCCG
NM_004448    CCCGCTGCTGGGGAGAGAGTTCTGAGGATTGTCAGAGCCTGACGCGCACTGTCTGTGCCG
NM_001289937 CCCGCTGCTGGGGAGAGAGTTCTGAGGATTGTCAGAGCCTGACGCGCACTGTCTGTGCCG
NM_001289936 CCCGCTGCTGGGGAGAGAGTTCTGAGGATTGTCAGAGCCTGACGCGCACTGTCTGTGCCG
NM_001005862 CCCGCTGCTGGGGAGAGAGTTCTGAGGATTGTCAGAGCCTGACGCGCACTGTCTGTGCCG
NM_001289938 CCCGCTGCTGGGGAGAGAGTTCTGAGGATTGTCAGAGCCTGACGCGCACTGTCTGTGCCG
             ************************************************************

NR_110535    GTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGTGCTG
NM_004448    GTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGTGCTG
NM_001289937 GTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGTGCTG
NM_001289936 GTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGTGCTG
NM_001005862 GTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGTGCTG
NM_001289938 GTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGTGCTG
             ************************************************************

NR_110535    CCGGCTGCACGGGCCCCAAGCACTCTGACTGCCTGGCCTGCCTCCACTTCAACCACAGTG
NM_004448    CCGGCTGCACGGGCCCCAAGCACTCTGACTGCCTGGCCTGCCTCCACTTCAACCACAGTG
NM_001289937 CCGGCTGCACGGGCCCCAAGCACTCTGACTGCCTGGCCTGCCTCCACTTCAACCACAGTG
NM_001289936 CCGGCTGCACGGGCCCCAAGCACTCTGACTGCCTGGCCTGCCTCCACTTCAACCACAGTG
NM_001005862 CCGGCTGCACGGGCCCCAAGCACTCTGACTGCCTGGCCTGCCTCCACTTCAACCACAGTG
NM_001289938 CCGGCTGCACGGGCCCCAAGCACTCTGACTGCCTGGCCTGCCTCCACTTCAACCACAGTG
             ************************************************************
```

```
NR_110535    GCATCTGTGAGCTGCACTGCCCAGCCCTGGTCACCTACAACACAGACACGTTTGAGTCCA
NM_004448    GCATCTGTGAGCTGCACTGCCCAGCCCTGGTCACCTACAACACAGACACGTTTGAGTCCA
NM_001289937 GCATCTGTGAGCTGCACTGCCCAGCCCTGGTCACCTACAACACAGACACGTTTGAGTCCA
NM_001289936 GCATCTGTGAGCTGCACTGCCCAGCCCTGGTCACCTACAACACAGACACGTTTGAGTCCA
NM_001005862 GCATCTGTGAGCTGCACTGCCCAGCCCTGGTCACCTACAACACAGACACGTTTGAGTCCA
NM_001289938 GCATCTGTGAGCTGCACTGCCCAGCCCTGGTCACCTACAACACAGACACGTTTGAGTCCA
             ************************************************************

NR_110535    TGCCCAATCCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTGACTGCCTGTCCCTACA
NM_004448    TGCCCAATCCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTGACTGCCTGTCCCTACA
NM_001289937 TGCCCAATCCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTGACTGCCTGTCCCTACA
NM_001289936 TGCCCAATCCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTGACTGCCTGTCCCTACA
NM_001005862 TGCCCAATCCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTGACTGCCTGTCCCTACA
NM_001289938 TGCCCAATCCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTGACTGCCTGTCCCTACA
             ************************************************************

NR_110535    ACTACCTTTCTACGGACGTGGGATCCTGCACCCTCGTCTGCCCCCTGCACAACCAAGAGG
NM_004448    ACTACCTTTCTACGGACGTGGGATCCTGCACCCTCGTCTGCCCCCTGCACAACCAAGAGG
NM_001289937 ACTACCTTTCTACGGACGTGGGATCCTGCACCCTCGTCTGCCCCCTGCACAACCAAGAGG
NM_001289936 ACTACCTTTCTACGGACGTGGGATCCTGCACCCTCGTCTGCCCCCTGCACAACCAAGAGG
NM_001005862 ACTACCTTTCTACGGACGTGGGATCCTGCACCCTCGTCTGCCCCCTGCACAACCAAGAGG
NM_001289938 ACTACCTTTCTACGGACGTGGGATCCTGCACCCTCGTCTGCCCCCTGCACAACCAAGAGG
             ************************************************************

NR_110535    TGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGAGTGT
NM_004448    TGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGAGTGT
NM_001289937 TGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGAGTGT
NM_001289936 TGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGAGTGT
NM_001005862 TGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGAGTGT
NM_001289938 TGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGAGTGT
             ************************************************************



NR_110535    GCTATGGTCTGGGCATGGAGCACTTGCGAGAGGTGAGGGCAGTTACCAGTGCCAATATCC
NM_004448    GCTATGGTCTGGGCATGGAGCACTTGCGAGAGGTGAGGGCAGTTACCAGTGCCAATATCC
NM_001289937 GCTATGGTCTGGGCATGGAGCACTTGCGAGAGGTGAGGGCAGTTACCAGTGCCAATATCC
NM_001289936 GCTATGGTCTGGGCATGGAGCACTTGCGAGAGGTGAGGGCAGTTACCAGTGCCAATATCC
NM_001005862 GCTATGGTCTGGGCATGGAGCACTTGCGAGAGGTGAGGGCAGTTACCAGTGCCAATATCC
NM_001289938 GCTATGGTCTGGGCATGGAGCACTTGCGAGAGGTGAGGGCAGTTACCAGTGCCAATATCC
             ************************************************************

NR_110535    AGGAGTTTGCTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGGAGAGCTTTG
NM_004448    AGGAGTTTGCTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGGAGAGCTTTG
NM_001289937 AGGAGTTTGCTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGGAGAGCTTTG
NM_001289936 AGGAGTTTGCTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGGAGAGCTTTG
NM_001005862 AGGAGTTTGCTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGGAGAGCTTTG
NM_001289938 AGGAGTTTGCTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGGAGAGCTTTG
             ************************************************************

NR_110535    ATGGGGACCCAGCCTCCAACACTGCCCCGCTCCAGCCAGAGCAGCTCCAAGTGTTTGAGA
NM_004448    ATGGGGACCCAGCCTCCAACACTGCCCCGCTCCAGCCAGAGCAGCTCCAAGTGTTTGAGA
NM_001289937 ATGGGGACCCAGCCTCCAACACTGCCCCGCTCCAGCCAGAGCAGCTCCAAGTGTTTGAGA
NM_001289936 ATGGGGACCCAGCCTCCAACACTGCCCCGCTCCAGCCAGAGCAGCTCCAAGTGTTTGAGA
NM_001005862 ATGGGGACCCAGCCTCCAACACTGCCCCGCTCCAGCCAGAGCAGCTCCAAGTGTTTGAGA
NM_001289938 ATGGGGACCCAGCCTCCAACACTGCCCCGCTCCAGCCAGAGCAGCTCCAAGTGTTTGAGA
             ************************************************************

NR_110535    CTCTGGAAGAGATCACAGGTTACCTATACATCTCAGCATGGCCGGACAGCCTGCCTGACC
NM_004448    CTCTGGAAGAGATCACAGGTTACCTATACATCTCAGCATGGCCGGACAGCCTGCCTGACC
NM_001289937 CTCTGGAAGAGATCACAGGTTACCTATACATCTCAGCATGGCCGGACAGCCTGCCTGACC
NM_001289936 CTCTGGAAGAGATCACAGGTTACCTATACATCTCAGCATGGCCGGACAGCCTGCCTGACC
NM_001005862 CTCTGGAAGAGATCACAGGTTACCTATACATCTCAGCATGGCCGGACAGCCTGCCTGACC
NM_001289938 CTCTGGAAGAGATCACAGGTTACCTATACATCTCAGCATGGCCGGACAGCCTGCCTGACC
             ************************************************************

NR_110535    TCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGGACGAATTCTGCACAATGGCGCCTACT
```

```
NM_004448    TCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGGACGAATTCTGCACAATGGCGCCTACT
NM_001289937 TCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGGACGAATTCTGCACAATGGCGCCTACT
NM_001289936 TCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGGACGAATTCTGCACAATGGCGCCTACT
NM_001005862 TCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGGACGAATTCTGCACAATGGCGCCTACT
NM_001289938 TCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGGACGAATTCTGCACAATGGCGCCTACT
             ************************************************************

NR_110535    CGCTGACCCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACTGAGGGAACTGG
NM_004448    CGCTGACCCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACTGAGGGAACTGG
NM_001289937 CGCTGACCCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACTGAGGGAACTGG
NM_001289936 CGCTGACCCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACTGAGGGAACTGG
NM_001005862 CGCTGACCCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACTGAGGGAACTGG
NM_001289938 CGCTGACCCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACTGAGGGAACTGG
             ************************************************************

NR_110535    GCAGTGGACTGGCCCTCATCCACCATAACACCCACCTCTGCTTCGTGCACACGGTGCCCT
NM_004448    GCAGTGGACTGGCCCTCATCCACCATAACACCCACCTCTGCTTCGTGCACACGGTGCCCT
NM_001289937 GCAGTGGACTGGCCCTCATCCACCATAACACCCACCTCTGCTTCGTGCACACGGTGCCCT
NM_001289936 GCAGTGGACTGGCCCTCATCCACCATAACACCCACCTCTGCTTCGTGCACACGGTGCCCT
NM_001005862 GCAGTGGACTGGCCCTCATCCACCATAACACCCACCTCTGCTTCGTGCACACGGTGCCCT
NM_001289938 GCAGTGGACTGGCCCTCATCCACCATAACACCCACCTCTGCTTCGTGCACACGGTGCCCT
             ************************************************************



NR_110535    GGGACCAGCTCTTTCGGAACCCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCAGAGG
NM_004448    GGGACCAGCTCTTTCGGAACCCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCAGAGG
NM_001289937 GGGACCAGCTCTTTCGGAACCCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCAGAGG
NM_001289936 GGGACCAGCTCTTTCGGAACCCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCAGAGG
NM_001005862 GGGACCAGCTCTTTCGGAACCCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCAGAGG
NM_001289938 GGGACCAGCTCTTTCGGAACCCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCAGAGG
             ************************************************************

NR_110535    ACGAGTGTGTGGGCGAGGGCCTGGCCTGCCACCAGCTGTGCGCCCGAGGGCACTGCTGGG
NM_004448    ACGAGTGTGTGGGCGAGGGCCTGGCCTGCCACCAGCTGTGCGCCCGAGGGCACTGCTGGG
NM_001289937 ACGAGTGTGTGGGCGAGGGCCTGGCCTGCCACCAGCTGTGCGCCCGAGGGCACTGCTGGG
NM_001289936 ACGAGTGTGTGGGCGAGGGCCTGGCCTGCCACCAGCTGTGCGCCCGAGGGCACTGCTGGG
NM_001005862 ACGAGTGTGTGGGCGAGGGCCTGGCCTGCCACCAGCTGTGCGCCCGAGGGCACTGCTGGG
NM_001289938 ACGAGTGTGTGGGCGAGGGCCTGGCCTGCCACCAGCTGTGCGCCCGAGGGCACTGCTGGG
             ************************************************************

NR_110535    GTCCAGGGCCCACCCAGTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGCGTGG
NM_004448    GTCCAGGGCCCACCCAGTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGCGTGG
NM_001289937 GTCCAGGGCCCACCCAGTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGCGTGG
NM_001289936 GTCCAGGGCCCACCCAGTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGCGTGG
NM_001005862 GTCCAGGGCCCACCCAGTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGCGTGG
NM_001289938 GTCCAGGGCCCACCCAGTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGCGTGG
             ************************************************************

NR_110535    AGGAATGCCGAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGTTTGC
NM_004448    AGGAATGCCGAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGTTTGC
NM_001289937 AGGAATGCCGAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGTTTGC
NM_001289936 AGGAATGCCGAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGTTTGC
NM_001005862 AGGAATGCCGAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGTTTGC
NM_001289938 AGGAATGCCGAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGTTTGC
             ************************************************************

NR_110535    CGTGCCACCCTGAGTGTCAGCCCCAGAATGGCTCAGTGACCTGTTTTGGACCGGAGGCTG
NM_004448    CGTGCCACCCTGAGTGTCAGCCCCAGAATGGCTCAGTGACCTGTTTTGGACCGGAGGCTG
NM_001289937 CGTGCCACCCTGAGTGTCAGCCCCAGAATGGCTCAGTGACCTGTTTTGGACCGGAGGCTG
NM_001289936 CGTGCCACCCTGAGTGTCAGCCCCAGAATGGCTCAGTGACCTGTTTTGGACCGGAGGCTG
NM_001005862 CGTGCCACCCTGAGTGTCAGCCCCAGAATGGCTCAGTGACCTGTTTTGGACCGGAGGCTG
NM_001289938 CGTGCCACCCTGAGTGTCAGCCCCAGAATGGCTCAGTGACCTGTTTTGGACCGGAGGCTG
             ************************************************************

NR_110535    ACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGCTGCCCCA
```

```
NM_004448    ACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGCTGCCCCA
NM_001289937 ACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGCTGCCCCA
NM_001289936 ACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGCTGCCCCA
NM_001005862 ACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGCTGCCCCA
NM_001289938 ACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGCTGCCCCA
             ************************************************************

NR_110535    GCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAGGGCG
NM_004448    GCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAGGGCG
NM_001289937 GCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAGGGCG
NM_001289936 GCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAGGGCG
NM_001005862 GCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAGGGCG
NM_001289938 GCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAGGGCG
             ************************************************************

NR_110535    CATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAGGGCT
NM_004448    CATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAGGGCT
NM_001289937 CATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAGGGCT
NM_001289936 CATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAGGGCT
NM_001005862 CATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAGGGCT
NM_001289938 CATGCCAGCCTTGCCCCATCAACTGCACCCACTCGTGAGTCCAACGGT----------CT
             ********************************** **:**  *.* *           **

NR_110535    GCCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTTGGCATTC
NM_004448    GCCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTTGGCATTC
NM_001289937 GCCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTTGGCATTC
NM_001289936 GCCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTTGGCATTC
NM_001005862 GCCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTTGGCATTC
NM_001289938 TTTCTGCAGAAAGGAGGACTTTC---------------CTTTCAGGGG------TCTTTC
                * **.**...***..* : *               .* **:* **       *:***

NR_110535    TGCTGGTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAGAAGA
NM_004448    TGCTGGTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAGAAGA
NM_001289937 TGCTGGTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAGAAGA
NM_001289936 TGCTGGTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAGAAGA
NM_001005862 TGCTGGTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAGAAGA
NM_001289938 TGGGGCTCTTACTATA----------AAAGGGGACCAACTCTCCCTTTGTCATATCTTGT
             **  * ** *. *.*:          :::***.:**:..**:. * : * ** .: ::*:

NR_110535    TCCGGA--------AGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCC
NM_004448    TCCGGA--------AGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCC
NM_001289937 TCCGGA--------AGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCC
NM_001289936 TCCGGA--------AGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCC
NM_001005862 TCCGGA--------AGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCC
NM_001289938 TTCTGATGACAAAAATAACACATTGTTAAAATTGTAAAATTA----AAACATGAA-----
             * * **        * :****.:**  .*.* ** :..*  *    .*.*: *:.

NR_110535    GCTGACACCTAGCGGAGCGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGA
NM_004448    GCTGACACCTAGCGGAGCGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGA
NM_001289937 GCTGACACCTAGCGGAGCGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGA
NM_001289936 GCTGACACCTAGCGGAGCGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGA
NM_001005862 GCTGACACCTAGCGGAGCGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGA
NM_001289938 ------AT---ATAAATTAATGCCCTAGCAAAAAAAAAAA----AAAAAAAAAAA-----
                   *    . ..*  .******:* **.....*.*:.    :..:.***.*.

NR_110535    GCTGAGGAAGGTGAAGGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTG
NM_004448    GCTGAGGAAGGTGAAGGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTG
NM_001289937 GCTGAGGAAGGTGAAGGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTG
NM_001289936 GCTGAGGAAGGTGAAGGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTG
NM_001005862 GCTGAGGAAGGTGAAGGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTG
NM_001289938 -------AAAAAAAA-------------------------------------------
                    **..:.**

NR_110535    GATCCCTGATGGGGAGAATGTGAAAATTCCAGTGGCCATCAAAGTGTTGAGGGAAAACAC
NM_004448    GATCCCTGATGGGGAGAATGTGAAAATTCCAGTGGCCATCAAAGTGTTGAGGGAAAACAC
NM_001289937 GATCCCTGATGGGGAGAATGTGAAAATTCCAGTGGCCATCAAAGTGTTGAGGGAAAACAC
NM_001289936 GATCCCTGATGGGGAGAATGTGAAAATTCCAGTGGCCATCAAAGTGTTGAGGGAAAACAC
NM_001005862 GATCCCTGATGGGGAGAATGTGAAAATTCCAGTGGCCATCAAAGTGTTGAGGGAAAACAC
```

```
NM_001289938 ------------------------------------------------------------

NR_110535    ATCCCCCAAAGCCAACAAAGAAATCTTAGACGAAGCATACGTGATGGCTGGTGTGGGCTC
NM_004448    ATCCCCCAAAGCCAACAAAGAAATCTTAGACGAAGCATACGTGATGGCTGGTGTGGGCTC
NM_001289937 ATCCCCCAAAGCCAACAAAGAAATCTTAGACGAAGCATACGTGATGGCTGGTGTGGGCTC
NM_001289936 ATCCCCCAAAGCCAACAAAGAAATCTTAGACGAAGCATACGTGATGGCTGGTGTGGGCTC
NM_001005862 ATCCCCCAAAGCCAACAAAGAAATCTTAGACGAAGCATACGTGATGGCTGGTGTGGGCTC
NM_001289938 ------------------------------------------------------------

NR_110535    CCCATATGTCTCCCGCCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACA
NM_004448    CCCATATGTCTCCCGCCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACA
NM_001289937 CCCATATGTCTCCCGCCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACA
NM_001289936 CCCATATGTCTCCCGCCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACA
NM_001005862 CCCATATGTCTCCCGCCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACA
NM_001289938 ------------------------------------------------------------

NR_110535    GCTTATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAAACCGCGGACGCCTGGGCTC
NM_004448    GCTTATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAAACCGCGGACGCCTGGGCTC
NM_001289937 GCTTATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAAACCGCGGACGCCTGGGCTC
NM_001289936 GCTTATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAAACCGCGGACGCCTGGGCTC
NM_001005862 GCTTATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAAACCGCGGACGCCTGGGCTC
NM_001289938 ------------------------------------------------------------

NR_110535    CCAGGACCTGCTGAACTGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGT
NM_004448    CCAGGACCTGCTGAACTGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGT
NM_001289937 CCAGGACCTGCTGAACTGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGT
NM_001289936 CCAGGACCTGCTGAACTGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGT
NM_001005862 CCAGGACCTGCTGAACTGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGT
NM_001289938 ------------------------------------------------------------

NR_110535    GCGGCTCGTACACAGGGACTTGGCCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGT
NM_004448    GCGGCTCGTACACAGGGACTTGGCCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGT
NM_001289937 GCGGCTCGTACACAGGGACTTGGCCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGT
NM_001289936 GCGGCTCGTACACAGGGACTTGGCCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGT
NM_001005862 GCGGCTCGTACACAGGGACTTGGCCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGT
NM_001289938 ------------------------------------------------------------

NR_110535    CAAAATTACAGACTTCGGGCTGGCTCGGCTGCTGGACATTGACGAGACAGAGTACCATGC
NM_004448    CAAAATTACAGACTTCGGGCTGGCTCGGCTGCTGGACATTGACGAGACAGAGTACCATGC
NM_001289937 CAAAATTACAGACTTCGGGCTGGCTCGGCTGCTGGACATTGACGAGACAGAGTACCATGC
NM_001289936 CAAAATTACAGACTTCGGGCTGGCTCGGCTGCTGGACATTGACGAGACAGAGTACCATGC
NM_001005862 CAAAATTACAGACTTCGGGCTGGCTCGGCTGCTGGACATTGACGAGACAGAGTACCATGC
NM_001289938 ------------------------------------------------------------

NR_110535    AGATGGGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCCGGCGGTT
NM_004448    AGATGGGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCCGGCGGTT
NM_001289937 AGATGGGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCCGGCGGTT
NM_001289936 AGATGGGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCCGGCGGTT
NM_001005862 AGATGGGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCCGGCGGTT
NM_001289938 ------------------------------------------------------------

NR_110535    CACCCACCAGAGTGATGTGTGGAGTTATGGTGTGACTGTGTGGGAGCTGATGACTTTTGG
NM_004448    CACCCACCAGAGTGATGTGTGGAGTTATGGTGTGACTGTGTGGGAGCTGATGACTTTTGG
NM_001289937 CACCCACCAGAGTGATGTGTGGAGTTATGGTGTGACTGTGTGGGAGCTGATGACTTTTGG
NM_001289936 CACCCACCAGAGTGATGTGTGGAGTTATGGTGTGACTGTGTGGGAGCTGATGACTTTTGG
NM_001005862 CACCCACCAGAGTGATGTGTGGAGTTATGGTGTGACTGTGTGGGAGCTGATGACTTTTGG
NM_001289938 ------------------------------------------------------------

NR_110535    GGCCAAACCTTACGATGGGATCCCAGCCCGGGAGATCCCTGACCTGCTGGAAAAGGGGGA
```

```
NM_004448    GGCCAAACCTTACGATGGGATCCCAGCCCGGGAGATCCCTGACCTGCTGGAAAAGGGGGA
NM_001289937 GGCCAAACCTTACGATGGGATCCCAGCCCGGGAGATCCCTGACCTGCTGGAAAAGGGGGA
NM_001289936 GGCCAAACCTTACGATGGGATCCCAGCCCGGGAGATCCCTGACCTGCTGGAAAAGGGGGA
NM_001005862 GGCCAAACCTTACGATGGGATCCCAGCCCGGGAGATCCCTGACCTGCTGGAAAAGGGGGA
NM_001289938 ------------------------------------------------------------

NR_110535    GCGGCTGCCCCAGCCCCCCATCTGCACCATTGATGTCTACATGATCATGGTCAAATGTTG
NM_004448    GCGGCTGCCCCAGCCCCCCATCTGCACCATTGATGTCTACATGATCATGGTCAAATGTTG
NM_001289937 GCGGCTGCCCCAGCCCCCCATCTGCACCATTGATGTCTACATGATCATGGTCAAATGTTG
NM_001289936 GCGGCTGCCCCAGCCCCCCATCTGCACCATTGATGTCTACATGATCATGGTCAAATGTTG
NM_001005862 GCGGCTGCCCCAGCCCCCCATCTGCACCATTGATGTCTACATGATCATGGTCAAATGTTG
NM_001289938 ------------------------------------------------------------

NR_110535    GATGATTGACTCTGAATGTCGGCCAAGATTCCGGGAGTTGGTGTCTGAATTCTCCCGCAT
NM_004448    GATGATTGACTCTGAATGTCGGCCAAGATTCCGGGAGTTGGTGTCTGAATTCTCCCGCAT
NM_001289937 GATGATTGACTCTGAATGTCGGCCAAGATTCCGGGAGTTGGTGTCTGAATTCTCCCGCAT
NM_001289936 GATGATTGACTCTGAATGTCGGCCAAGATTCCGGGAGTTGGTGTCTGAATTCTCCCGCAT
NM_001005862 GATGATTGACTCTGAATGTCGGCCAAGATTCCGGGAGTTGGTGTCTGAATTCTCCCGCAT
NM_001289938 ------------------------------------------------------------

NR_110535    GGCCAGGGACCCCCAGCGCTTTGTGGTCATCCAGAATGAGGACTTGGGCCCAGCCAGTCC
NM_004448    GGCCAGGGACCCCCAGCGCTTTGTGGTCATCCAGAATGAGGACTTGGGCCCAGCCAGTCC
NM_001289937 GGCCAGGGACCCCCAGCGCTTTGTGGTCATCCAGAATGAGGACTTGGGCCCAGCCAGTCC
NM_001289936 GGCCAGGGACCCCCAGCGCTTTGTGGTCATCCAGAATGAGGACTTGGGCCCAGCCAGTCC
NM_001005862 GGCCAGGGACCCCCAGCGCTTTGTGGTCATCCAGAATGAGGACTTGGGCCCAGCCAGTCC
NM_001289938 ------------------------------------------------------------

NR_110535    CTTGGACAGCACCTTCTACCGCTCACTGCTGGAGGACGATGACATGGGGGACCTGGTGGA
NM_004448    CTTGGACAGCACCTTCTACCGCTCACTGCTGGAGGACGATGACATGGGGGACCTGGTGGA
NM_001289937 CTTGGACAGCACCTTCTACCGCTCACTGCTGGAGGACGATGACATGGGGGACCTGGTGGA
NM_001289936 CTTGGACAGCACCTTCTACCGCTCACTGCTGGAGGACGATGACATGGGGGACCTGGTGGA
NM_001005862 CTTGGACAGCACCTTCTACCGCTCACTGCTGGAGGACGATGACATGGGGGACCTGGTGGA
NM_001289938 ------------------------------------------------------------

NR_110535    TGCTGAGGAGTATCTGGTACCCCAGCAGGGCTTCTTCTGTCCAGACCCTGCCCCGGGCGC
NM_004448    TGCTGAGGAGTATCTGGTACCCCAGCAGGGCTTCTTCTGTCCAGACCCTGCCCCGGGCGC
NM_001289937 TGCTGAGGAGTATCTGGTACCCCAGCAGGGCTTCTTCTGTCCAGACCCTGCCCCGGGCGC
NM_001289936 TGCTGAGGAGTATCTGGTACCCCAGCAGGGCTTCTTCTGTCCAGACCCTGCCCCGGGCGC
NM_001005862 TGCTGAGGAGTATCTGGTACCCCAGCAGGGCTTCTTCTGTCCAGACCCTGCCCCGGGCGC
NM_001289938 ------------------------------------------------------------

NR_110535    TGGGGGCATGGTCCACCACAGGCACCGCAGCTCATCTACCAGGAGTGGCGGTGGGGACCT
NM_004448    TGGGGGCATGGTCCACCACAGGCACCGCAGCTCATCTACCAGGAGTGGCGGTGGGGACCT
NM_001289937 TGGGGGCATGGTCCACCACAGGCACCGCAGCTCATCTACCAGGA----------------
NM_001289936 TGGGGGCATGGTCCACCACAGGCACCGCAGCTCATCTACCAGGAGTGGCGGTGGGGACCT
NM_001005862 TGGGGGCATGGTCCACCACAGGCACCGCAGCTCATCTACCAGGAGTGGCGGTGGGGACCT
NM_001289938 ------------------------------------------------------------

NR_110535    GACACTAGGGCTGGAGCCCTCTGAAGAGGAGGCCCCCAGGTCTCCACTGGCACCCTCCGA
NM_004448    GACACTAGGGCTGGAGCCCTCTGAAGAGGAGGCCCCCAGGTCTCCACTGGCACCCTCCGA
NM_001289937 ------------------------------------------------------------
NM_001289936 GACACTAGGGCTGGAGCCCTCTGAAGAGGAGGCCCCCAGGTCTCCACTGGCACCCTCCGA
NM_001005862 GACACTAGGGCTGGAGCCCTCTGAAGAGGAGGCCCCCAGGTCTCCACTGGCACCCTCCGA
NM_001289938 ------------------------------------------------------------

NR_110535    AGGGGCTGGCTCCGATGTATTTGATGGTGACCTGGGAATGGGGGCAGCCAAGGGGCTGCA
NM_004448    AGGGGCTGGCTCCGATGTATTTGATGGTGACCTGGGAATGGGGGCAGCCAAGGGGCTGCA
NM_001289937 ------------------------------------------------------------
NM_001289936 AGGGGCTGGCTCCGATGTATTTGATGGTGACCTGGGAATGGGGGCAGCCAAGGGGCTGCA
NM_001005862 AGGGGCTGGCTCCGATGTATTTGATGGTGACCTGGGAATGGGGGCAGCCAAGGGGCTGCA
```

```
NM_001289938 ------------------------------------------------------------

NR_110535    AAGCCTCCCCACACATGACCCCAGCCCTCTACAGCGGTACAGTGAGGACCCCACAGTACC
NM_004448    AAGCCTCCCCACACATGACCCCAGCCCTCTACAGCGGTACAGTGAGGACCCCACAGTACC
NM_001289937 ------------------------------------------------------------
NM_001289936 AAGCCTCCCCACACATGACCCCAGCCCTCTACAGCGGTACAGTGAGGACCCCACAGTACC
NM_001005862 AAGCCTCCCCACACATGACCCCAGCCCTCTACAGCGGTACAGTGAGGACCCCACAGTACC
NM_001289938 ------------------------------------------------------------

NR_110535    CCTGCCCTCTGAGACTGATGGCTACGTTGCCCCCCTGACCTGCAGCCCCCAGCCTGAATA
NM_004448    CCTGCCCTCTGAGACTGATGGCTACGTTGCCCCCCTGACCTGCAGCCCCCAGCCTGAATA
NM_001289937 ------------------------------------------------------------ATA
NM_001289936 CCTGCCCTCTGAGACTGATGGCTACGTTGCCCCCCTGACCTGCAGCCCCCAGCCTGAATA
NM_001005862 CCTGCCCTCTGAGACTGATGGCTACGTTGCCCCCCTGACCTGCAGCCCCCAGCCTGAATA
NM_001289938 ------------------------------------------------------------

NR_110535    TGTGAACCAGCCAGATGTTCGGCCCCAGCCCCCTTCGCCCCGAGAGGGCCCTCTGCCTGC
NM_004448    TGTGAACCAGCCAGATGTTCGGCCCCAGCCCCCTTCGCCCCGAGAGGGCCCTCTGCCTGC
NM_001289937 TGTGAACCAGCCAGATGTTCGGCCCCAGCCCCCTTCGCCCCGAGAGGGCCCTCTGCCTGC
NM_001289936 TGTGAACCAGCCAGATGTTCGGCCCCAGCCCCCTTCGCCCCGAGAGGGCCCTCTGCCTGC
NM_001005862 TGTGAACCAGCCAGATGTTCGGCCCCAGCCCCCTTCGCCCCGAGAGGGCCCTCTGCCTGC
NM_001289938 ------------------------------------------------------------

NR_110535    TGCCCGACCTGCTGGTGCCACTCTGGAAAGGCCCAAGACTCTCTCCCCAGGGAAGAATGG
NM_004448    TGCCCGACCTGCTGGTGCCACTCTGGAAAGGCCCAAGACTCTCTCCCCAGGGAAGAATGG
NM_001289937 TGCCCGACCTGCTGGTGCCACTCTGGAAAGGCCCAAGACTCTCTCCCCAGGGAAGAATGG
NM_001289936 TGCCCGACCTGCTGGTGCCACTCTGGAAAGGCCCAAGACTCTCTCCCCAGGGAAGAATGG
NM_001005862 TGCCCGACCTGCTGGTGCCACTCTGGAAAGGCCCAAGACTCTCTCCCCAGGGAAGAATGG
NM_001289938 ------------------------------------------------------------

NR_110535    GGTCGTCAAAGACGTTTTTGCCTTTGGGGGTGCCGTGGAGAACCCCGAGTACTTGACACC
NM_004448    GGTCGTCAAAGACGTTTTTGCCTTTGGGGGTGCCGTGGAGAACCCCGAGTACTTGACACC
NM_001289937 GGTCGTCAAAGACGTTTTTGCCTTTGGGGGTGCCGTGGAGAACCCCGAGTACTTGACACC
NM_001289936 GGTCGTCAAAGACGTTTTTGCCTTTGGGGGTGCCGTGGAGAACCCCGAGTACTTGACACC
NM_001005862 GGTCGTCAAAGACGTTTTTGCCTTTGGGGGTGCCGTGGAGAACCCCGAGTACTTGACACC
NM_001289938 ------------------------------------------------------------

NR_110535    CCAGGGAGGAGCTGCCCCTCAGCCCCACCCTCCTCCTGCCTTCAGCCCAGCCTTCGACAA
NM_004448    CCAGGGAGGAGCTGCCCCTCAGCCCCACCCTCCTCCTGCCTTCAGCCCAGCCTTCGACAA
NM_001289937 CCAGGGAGGAGCTGCCCCTCAGCCCCACCCTCCTCCTGCCTTCAGCCCAGCCTTCGACAA
NM_001289936 CCAGGGAGGAGCTGCCCCTCAGCCCCACCCTCCTCCTGCCTTCAGCCCAGCCTTCGACAA
NM_001005862 CCAGGGAGGAGCTGCCCCTCAGCCCCACCCTCCTCCTGCCTTCAGCCCAGCCTTCGACAA
NM_001289938 ------------------------------------------------------------

NR_110535    CCTCTATTACTGGGACCAGGACCCACCAGAGCGGGGGGCTCCACCCAGCACCTTCAAAGG
NM_004448    CCTCTATTACTGGGACCAGGACCCACCAGAGCGGGGGGCTCCACCCAGCACCTTCAAAGG
NM_001289937 CCTCTATTACTGGGACCAGGACCCACCAGAGCGGGGGGCTCCACCCAGCACCTTCAAAGG
NM_001289936 CCTCTATTACTGGGACCAGGACCCACCAGAGCGGGGGGCTCCACCCAGCACCTTCAAAGG
NM_001005862 CCTCTATTACTGGGACCAGGACCCACCAGAGCGGGGGGCTCCACCCAGCACCTTCAAAGG
NM_001289938 ------------------------------------------------------------

NR_110535    GACACCTACGGCAGAGAACCCAGAGTACCTGGGTCTGGACGTGCCAGTGTGAACCAGAAG
NM_004448    GACACCTACGGCAGAGAACCCAGAGTACCTGGGTCTGGACGTGCCAGTGTGAACCAGAAG
NM_001289937 GACACCTACGGCAGAGAACCCAGAGTACCTGGGTCTGGACGTGCCAGTGTGAACCAGAAG
NM_001289936 GACACCTACGGCAGAGAACCCAGAGTACCTGGGTCTGGACGTGCCAGTGTGAACCAGAAG
NM_001005862 GACACCTACGGCAGAGAACCCAGAGTACCTGGGTCTGGACGTGCCAGTGTGAACCAGAAG
NM_001289938 ------------------------------------------------------------

NR_110535    GCCAAGTCCGCAGAAGCCCTGATGTGTCCTCAGGGAGCAGGGAAGGCCTGACTTCTGCTG
```

```
NM_004448    GCCAAGTCCGCAGAAGCCCTGATGTGTCCTCAGGGAGCAGGGAAGGCCTGACTTCTGCTG
NM_001289937 GCCAAGTCCGCAGAAGCCCTGATGTGTCCTCAGGGAGCAGGGAAGGCCTGACTTCTGCTG
NM_001289936 GCCAAGTCCGCAGAAGCCCTGATGTGTCCTCAGGGAGCAGGGAAGGCCTGACTTCTGCTG
NM_001005862 GCCAAGTCCGCAGAAGCCCTGATGTGTCCTCAGGGAGCAGGGAAGGCCTGACTTCTGCTG
NM_001289938 ------------------------------------------------------------

NR_110535    GCATCAAGAGGTGGGAGGGCCCTCCGACCACTTCCAGGGGAACCTGCCATGCCAGGAACC
NM_004448    GCATCAAGAGGTGGGAGGGCCCTCCGACCACTTCCAGGGGAACCTGCCATGCCAGGAACC
NM_001289937 GCATCAAGAGGTGGGAGGGCCCTCCGACCACTTCCAGGGGAACCTGCCATGCCAGGAACC
NM_001289936 GCATCAAGAGGTGGGAGGGCCCTCCGACCACTTCCAGGGGAACCTGCCATGCCAGGAACC
NM_001005862 GCATCAAGAGGTGGGAGGGCCCTCCGACCACTTCCAGGGGAACCTGCCATGCCAGGAACC
NM_001289938 ------------------------------------------------------------

NR_110535    TGTCCTAAGGAACCTTCCTTCCTGCTTGAGTTCCCAGATGGCTGGAAGGGGTCCAGCCTC
NM_004448    TGTCCTAAGGAACCTTCCTTCCTGCTTGAGTTCCCAGATGGCTGGAAGGGGTCCAGCCTC
NM_001289937 TGTCCTAAGGAACCTTCCTTCCTGCTTGAGTTCCCAGATGGCTGGAAGGGGTCCAGCCTC
NM_001289936 TGTCCTAAGGAACCTTCCTTCCTGCTTGAGTTCCCAGATGGCTGGAAGGGGTCCAGCCTC
NM_001005862 TGTCCTAAGGAACCTTCCTTCCTGCTTGAGTTCCCAGATGGCTGGAAGGGGTCCAGCCTC
NM_001289938 ------------------------------------------------------------

NR_110535    GTTGGAAGAGGAACAGCACTGGGGAGTCTTTGTGGATTCTGAGGCCCTGCCCAATGAGAC
NM_004448    GTTGGAAGAGGAACAGCACTGGGGAGTCTTTGTGGATTCTGAGGCCCTGCCCAATGAGAC
NM_001289937 GTTGGAAGAGGAACAGCACTGGGGAGTCTTTGTGGATTCTGAGGCCCTGCCCAATGAGAC
NM_001289936 GTTGGAAGAGGAACAGCACTGGGGAGTCTTTGTGGATTCTGAGGCCCTGCCCAATGAGAC
NM_001005862 GTTGGAAGAGGAACAGCACTGGGGAGTCTTTGTGGATTCTGAGGCCCTGCCCAATGAGAC
NM_001289938 ------------------------------------------------------------

NR_110535    TCTAGGGTCCAGTGGATGCCACAGCCCAGCTTGGCCCTTTCCTTCCAGATCCTGGGTACT
NM_004448    TCTAGGGTCCAGTGGATGCCACAGCCCAGCTTGGCCCTTTCCTTCCAGATCCTGGGTACT
NM_001289937 TCTAGGGTCCAGTGGATGCCACAGCCCAGCTTGGCCCTTTCCTTCCAGATCCTGGGTACT
NM_001289936 TCTAGGGTCCAGTGGATGCCACAGCCCAGCTTGGCCCTTTCCTTCCAGATCCTGGGTACT
NM_001005862 TCTAGGGTCCAGTGGATGCCACAGCCCAGCTTGGCCCTTTCCTTCCAGATCCTGGGTACT
NM_001289938 ------------------------------------------------------------

NR_110535    GAAAGCCTTAGGGAAGCTGGCCTGAGAGGGGAAGCGGCCCTAAGGGAGTGTCTAAGAACA
NM_004448    GAAAGCCTTAGGGAAGCTGGCCTGAGAGGGGAAGCGGCCCTAAGGGAGTGTCTAAGAACA
NM_001289937 GAAAGCCTTAGGGAAGCTGGCCTGAGAGGGGAAGCGGCCCTAAGGGAGTGTCTAAGAACA
NM_001289936 GAAAGCCTTAGGGAAGCTGGCCTGAGAGGGGAAGCGGCCCTAAGGGAGTGTCTAAGAACA
NM_001005862 GAAAGCCTTAGGGAAGCTGGCCTGAGAGGGGAAGCGGCCCTAAGGGAGTGTCTAAGAACA
NM_001289938 ------------------------------------------------------------

NR_110535    AAAGCGACCCATTCAGAGACTGTCCCTGAAACCTAGTACTGCCCCCCATGAGGAAGGAAC
NM_004448    AAAGCGACCCATTCAGAGACTGTCCCTGAAACCTAGTACTGCCCCCCATGAGGAAGGAAC
NM_001289937 AAAGCGACCCATTCAGAGACTGTCCCTGAAACCTAGTACTGCCCCCCATGAGGAAGGAAC
NM_001289936 AAAGCGACCCATTCAGAGACTGTCCCTGAAACCTAGTACTGCCCCCCATGAGGAAGGAAC
NM_001005862 AAAGCGACCCATTCAGAGACTGTCCCTGAAACCTAGTACTGCCCCCCATGAGGAAGGAAC
NM_001289938 ------------------------------------------------------------

NR_110535    AGCAATGGTGTCAGTATCCAGGCTTTGTACAGAGTGCTTTTCTGTTTAGTTTTTACTTTT
NM_004448    AGCAATGGTGTCAGTATCCAGGCTTTGTACAGAGTGCTTTTCTGTTTAGTTTTTACTTTT
NM_001289937 AGCAATGGTGTCAGTATCCAGGCTTTGTACAGAGTGCTTTTCTGTTTAGTTTTTACTTTT
NM_001289936 AGCAATGGTGTCAGTATCCAGGCTTTGTACAGAGTGCTTTTCTGTTTAGTTTTTACTTTT
NM_001005862 AGCAATGGTGTCAGTATCCAGGCTTTGTACAGAGTGCTTTTCTGTTTAGTTTTTACTTTT
NM_001289938 ------------------------------------------------------------

NR_110535    TTTGTTTTGTTTTTTTAAAGATGAAATAAAGACCCAGGGGGAGAATGGGTGTTGTATGGG
NM_004448    TTTGTTTTGTTTTTTTAAAGATGAAATAAAGACCCAGGGGGAGAATGGGTGTTGTATGGG
NM_001289937 TTTGTTTTGTTTTTTTAAAGATGAAATAAAGACCCAGGGGGAGAATGGGTGTTGTATGGG
NM_001289936 TTTGTTTTGTTTTTTTAAAGATGAAATAAAGACCCAGGGGGAGAATGGGTGTTGTATGGG
NM_001005862 TTTGTTTTGTTTTTTTAAAGATGAAATAAAGACCCAGGGGGAGAATGGGTGTTGTATGGG
```

```
NM_001289938 ------------------------------------------------------------

NR_110535    GAGGCAAGTGTGGGGGGTCCTTCTCCACACCCACTTTGTCCATTTGCAAATATATTTTGG
NM_004448    GAGGCAAGTGTGGGGGGTCCTTCTCCACACCCACTTTGTCCATTTGCAAATATATTTTGG
NM_001289937 GAGGCAAGTGTGGGGGGTCCTTCTCCACACCCACTTTGTCCATTTGCAAATATATTTTGG
NM_001289936 GAGGCAAGTGTGGGGGGTCCTTCTCCACACCCACTTTGTCCATTTGCAAATATATTTTGG
NM_001005862 GAGGCAAGTGTGGGGGGTCCTTCTCCACACCCACTTTGTCCATTTGCAAATATATTTTGG
NM_001289938 ------------------------------------------------------------

NR_110535         AAAACAGCTAAAAAAAAAAAAAAAAAA
NM_004448         AAAACAGCTAAAAAAAAAAAAAAAAAA
NM_001289937      AAAACAGCTAAAAAAAAAAAAAAAAAA
NM_001289936      AAAACAGCTAAAAAAAAAAAAAAAAAA
NM_001005862      AAAACAGCTAAAAAAAAAAAAAAAAAA
NM_001289938      ---------------------------
```

CLUSTAL OMEGA ALIGNMENTS OF *HER2* ISOFORMS OBTAINED FROM THE NCBI DATABASE:

```
P04626        MELAALCRWGLLLALLPPGAASTQVCTGTDMKLRLPASPETHLDMLRHLYQGCQVVQGNL
P04626-1      ------------------------------------------------------------
P04626-3      ------------------------------------------------------------
P04626-4      ---------------MPRGSWKPQVCTGTDMKLRLPASPETHLDMLRHLYQGCQVVQGNL

P04626        ELTYLPTNASLSFLQDIQEVQGYVLIAHNQVRQVPLQRLRIVRGTQLFEDNYALAVLDNG
P04626-1      ------------------------------------------------------------
P04626-3      ------------------------------------------------------------
P04626-4      ELTYLPTNASLSFLQDIQEVQGYVLIAHNQVRQVPLQRLRIVRGTQLFEDNYALAVLDNG

P04626        DPLNNTTPVTGASPGGLRELQLRSLTEILKGGVLIQRNPQLCYQDTILWKDIFHKNNQLA
P04626-1      ------------------------------------------------------------
P04626-3      ------------------------------------------------------------
P04626-4      DPLNNTTPVTGASPGGLRELQLRSLTEILKGGVLIQRNPQLCYQDTILWKDIFHKNNQLA

P04626        LTLIDTNRSRACHPCSPMCKGSRCWGESSEDCQSLTRTVCAGGCARCKGPLPTDCCHEQC
P04626-1      ------------------------------------------------------------
P04626-3      ------------------------------------------------------------
P04626-4      LTLIDTNRSRACHPCSPMCKGSRCWGESSEDCQSLTRTVCAGGCARCKGPLPTDCCHEQC

P04626        AAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP
P04626-1      ------------------------------------------------------------
P04626-3      ------------------------------------------------------------
P04626-4      AAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP

P04626        YNYLSTDVGSCTLVCPLHNQEVTAEDGTQRCEKCSKPCARVCYGLGMEHLREVRAVTSAN
P04626-1      ------------------------------------------------------------
P04626-3      ------------------------------------------------------------
P04626-4      YNYLSTDVGSCTLVCPLHNQEVTAEDGTQRCEKCSKPCARVCYGLGMEHLREVRAVTSAN

P04626        IQEFAGCKKIFGSLAFLPESFDGDPASNTAPLQPEQLQVFETLEEITGYLYISAWPDSLP
P04626-1      ------------------------------------------------------------
P04626-3      ------------------------------------------------------------
P04626-4      IQEFAGCKKIFGSLAFLPESFDGDPASNTAPLQPEQLQVFETLEEITGYLYISAWPDSLP

P04626        ------------------------------------------------------------
P04626-1      ------------------------------------------------------------
P04626-3      ------------------------------------------------------------
P04626-4      DLSVFQNLQVIRGRILHNGAYSLTLQGLGISWLGLRSLRELGSGLALIHHNTHLCFVHTV

P04626        ------------------------------------------------------------
P04626-1      ------------------------------------------------------------
P04626-3      ------------------------------------------------------------
P04626-4      PWDQLFRNPHQALLHTANRPEDECVGEGLACHQLCARGHCWGPGPTQCVNCSQFLRGQEC

P04626        ------------------------------------------------------------
P04626-1      ------------------------------------------------------------
P04626-3      ------------------------------------------------------------
P04626-4      VEECRVLQGLPREYVNARHCLPCHPECQPQNGSVTCFGPEADQCVACAHYKDPPFCVARC

P04626        ------------------------------------------------------------
```

```
P04626-1      ----------MPIWKFPDEEGACQPCPINCTHSCVDLDDKGCPAEQRASPLTSIISAVVG
P04626-3      ------------------------------------------------------------
P04626-4      PSGVKPDLSYMPIWKFPDEEGACQPCPINCTHSCVDLDDKGCPAEQRASPLTSIISAVVG

P04626        ------------------------------------------------------------
P04626-1      ILLVVVLGVVFGILIKRRQQKIRKYTMRRLLQETELVEPLTPSGAMPNQAQMRILKETEL
P04626-3      --------------------------MRRLLQETELVEPLTPSGAMPNQAQMRILKETEL
P04626-4      ILLVVVLGVVFGILIKRRQQKIRKYTMRRLLQETELVEPLTPSGAMPNQAQMRILKETEL

P04626        ------------------------------------------------------------
P04626-1      RKVKVLGSGAFGTVYKGIWIPDGENVKIPVAIKVLRENTSPKANKEILDEAYVMAGVGSP
P04626-3      RKVKVLGSGAFGTVYKGIWIPDGENVKIPVAIKVLRENTSPKANKEILDEAYVMAGVGSP
P04626-4      RKVKVLGSGAFGTVYKGIWIPDGENVKIPVAIKVLRENTSPKANKEILDEAYVMAGVGSP

P04626        ------------------------------------------------------------
P04626-1      YVSRLLGICLTSTVQLVTQLMPYGCLLDHVRENRGRLGSQDLLNWCMQIAKGMSYLEDVR
P04626-3      YVSRLLGICLTSTVQLVTQLMPYGCLLDHVRENRGRLGSQDLLNWCMQIAKGMSYLEDVR
P04626-4      YVSRLLGICLTSTVQLVTQLMPYGCLLDHVRENRGRLGSQDLLNWCMQIAKGMSYLEDVR

P04626        ------------------------------------------------------------
P04626-1      LVHRDLAARNVLVKSPNHVKITDFGLARLLDIDETEYHADGGKVPIKWMALESILRRRFT
P04626-3      LVHRDLAARNVLVKSPNHVKITDFGLARLLDIDETEYHADGGKVPIKWMALESILRRRFT
P04626-4      LVHRDLAARNVLVKSPNHVKITDFGLARLLDIDETEYHADGGKVPIKWMALESILRRRFT

P04626        ------------------------------------------------------------
P04626-1      HQSDVWSYGVTVWELMTFGAKPYDGIPAREIPDLLEKGERLPQPPICTIDVYMIMVKCWM
P04626-3      HQSDVWSYGVTVWELMTFGAKPYDGIPAREIPDLLEKGERLPQPPICTIDVYMIMVKCWM
P04626-4      HQSDVWSYGVTVWELMTFGAKPYDGIPAREIPDLLEKGERLPQPPICTIDVYMIMVKCWM

P04626        ------------------------------------------------------------
P04626-1      IDSECRPRFRELVSEFSRMARDPQRFVVIQNEDLGPASPLDSTFYRSLLEDDDMGDLVDA
P04626-3      IDSECRPRFRELVSEFSRMARDPQRFVVIQNEDLGPASPLDSTFYRSLLEDDDMGDLVDA
P04626-4      IDSECRPRFRELVSEFSRMARDPQRFVVIQNEDLGPASPLDSTFYRSLLEDDDMGDLVDA

P04626        ------------------------------------------------------------
P04626-1      EEYLVPQQGFFCPDPAPGAGGMVHHRHRSSSTRSGGGDLTLGLEPSEEEAPRSPLAPSEG
P04626-3      EEYLVPQQGFFCPDPAPGAGGMVHHRHRSSSTRSGGGDLTLGLEPSEEEAPRSPLAPSEG
P04626-4      EEYLVPQQGFFCPDPAPGAGGMVHHRHRSSSTRSGGGDLTLGLEPSEEEAPRSPLAPSEG

P04626        ------------------------------------------------------------
P04626-1      AGSDVFDGDLGMGAAKGLQSLPTHDPSPLQRYSEDPTVPLPSETDGYVAPLTCSPQPEYV
P04626-3      AGSDVFDGDLGMGAAKGLQSLPTHDPSPLQRYSEDPTVPLPSETDGYVAPLTCSPQPEYV
P04626-4      AGSDVFDGDLGMGAAKGLQSLPTHDPSPLQRYSEDPTVPLPSETDGYVAPLTCSPQPEYV

P04626        ------------------------------------------------------------
P04626-1      NQPDVRPQPPSPREGPLPAARPAGATLERPKTLSPGKNGVVKDVFAFGGAVENPEYLTPQ
P04626-3      NQPDVRPQPPSPREGPLPAARPAGATLERPKTLSPGKNGVVKDVFAFGGAVENPEYLTPQ
P04626-4      NQPDVRPQPPSPREGPLPAARPAGATLERPKTLSPGKNGVVKDVFAFGGAVENPEYLTPQ

P04626        ------------------------------------------------------------
P04626-1      GGAAPQPHPPPAFSPAFDNLYYWDQDPPERGAPPSTFKGTPTAENPEYLGLDVPV-----
P04626-3      GGAAPQPHPPPAFSPAFDNLYYWDQDPPERGAPPSTFKGTPTAENPEYLGLDVPV-----
P04626-4      GGAAPQPHPPPAFSPAFDNLYYWDQDPPERGAPPSTFKGTPTAENPEYLGLDVPVDLSVF

P04626        ------------------------------------------------------------
P04626-1      ------------------------------------------------------------
P04626-3      ------------------------------------------------------------
```

```
P04626-4      QNLQVIRGRILHNGAYSLTLQGLGISWLGLRSLRELGSGLALIHHNTHLCFVHTVPWDQL

P04626        ------------------------------------------------------------
P04626-1      ------------------------------------------------------------
P04626-3      ------------------------------------------------------------
P04626-4      FRNPHQALLHTANRPEDECVGEGLACHQLCARGHCWGPGPTQCVNCSQFLRGQECVEECR

P04626        ------------------------------------------------------------
P04626-1      ------------------------------------------------------------
P04626-3      ------------------------------------------------------------
P04626-4      VLQGLPREYVNARHCLPCHPECQPQNGSVTCFGPEADQCVACAHYKDPPFCVARCPSGVK

P04626        ------------------------------------------------------------
P04626-1      ------------------------------------------------------------
P04626-3      ------------------------------------------------------------
P04626-4      PDLSYMPIWKFPDEEGACQPCPINCTHSCVDLDDKGCPAEQRASPLTSIISAVVGILLVV

P04626        ------------------------------------------------------------
P04626-1      ------------------------------------------------------------
P04626-3      ------------------------------------------------------------
P04626-4      VLGVVFGILIKRRQQKIRKYTMRRLLQETELVEPLTPSGAMPNQAQMRILKETELRKVKV

P04626        ------------------------------------------------------------
P04626-1      ------------------------------------------------------------
P04626-3      ------------------------------------------------------------
P04626-4      LGSGAFGTVYKGIWIPDGENVKIPVAIKVLRENTSPKANKEILDEAYVMAGVGSPYVSRL

P04626        ------------------------------------------------------------
P04626-1      ------------------------------------------------------------
P04626-3      ------------------------------------------------------------
P04626-4      LGICLTSTVQLVTQLMPYGCLLDHVRENRGRLGSQDLLNWCMQIAKGMSYLEDVRLVHRD

P04626        ------------------------------------------------------------
P04626-1      ------------------------------------------------------------
P04626-3      ------------------------------------------------------------
P04626-4      LAARNVLVKSPNHVKITDFGLARLLDIDETEYHADGGKVPIKWMALESILRRRFTHQSDV

P04626        ------------------------------------------------------------
P04626-1      ------------------------------------------------------------
P04626-3      ------------------------------------------------------------
P04626-4      WSYGVTVWELMTFGAKPYDGIPAREIPDLLEKGERLPQPPICTIDVYMIMVKCWMIDSEC

P04626        ------------------------------------------------------------
P04626-1      ------------------------------------------------------------
P04626-3      ------------------------------------------------------------
P04626-4      RPRFRELVSEFSRMARDPQRFVVIQNEDLGPASPLDSTFYRSLLEDDDMGDLVDAEEYLV

P04626        ------------------------------------------------------------
P04626-1      ------------------------------------------------------------
P04626-3      ------------------------------------------------------------
P04626-4      PQQGFFCPDPAPGAGGMVHHRHRSSSTRSGGGDLTLGLEPSEEEAPRSPLAPSEGAGSDV

P04626        ------------------------------------------------------------
P04626-1      ------------------------------------------------------------
P04626-3      ------------------------------------------------------------
P04626-4      FDGDLGMGAAKGLQSLPTHDPSPLQRYSEDPTVPLPSETDGYVAPLTCSPQPEYVNQPDV
```

```
P04626        ------------------------------------------------------------
P04626-1      ------------------------------------------------------------
P04626-3      ------------------------------------------------------------
P04626-4      RPQPPSPREGPLPAARPAGATLERPKTLSPGKNGVVKDVFAFGGAVENPEYLTPQGGAAP

P04626        --------------------------------------------------
P04626-1      --------------------------------------------------
P04626-3      --------------------------------------------------
P04626-4      QPHPPPAFSPAFDNLYYWDQDPPERGAPPSTFKGTPTAENPEYLGLDVPV
```

LIST OF PRESENTATIONS

1.

EVENT TITLE: SAN ANTONIO BREAST CANCER SYMPOSIUM

VENUE: HENRY B. GONZALEZ CONVENTION CENTER, SAN ANTONIO, TEXAS, USA

DATE: DECEMBER 6-11, 2011

AWARD RECEIVED: AACR scholar-in-training award by Susan G Komen for the Cure.

ABSTRACT

Background:

The Human Epidermal Growth Factor Receptor 2 (HER2) is an oncogene expressed in 25-30% of invasive breast cancers. HER2 shares extensive homology with other members of the HER family (HER1, HER3 and HER4), and is constitutively active as an homo-and heterodimer. The HER2 gene encodes an 185kDa transmembrane protein with tyrosine kinase activity. Gene amplification or protein expression of HER2 is a predictor of poor clinical outcome and decreased survival in women with breast cancer, and also indicates a favourable response to Trastuzumab (Herceptin) therapy, or a combinational therapy comprising Herceptin plus chemotherapy. However, resistance to Trastuzumab remains the case in approximately 50% of HER2 amplified/overexpressing tumours. Understanding the molecular mechanisms of Trastuzumab resistance and identifying more effective therapies, is critical in the treatment of patients whose breast cancers express this aggressive disease phenotype. In this study, it is postulated that the abnormal generation of mRNA splice variants may be responsible for the continued tumour growth and progression.

Aims:

The aim of this study is to increase our understanding of the role of HER2 splice variants in the development and progression of breast cancer. This will inform the development of more sophisticated and effective therapies that target specific HER2 isoforms, rather than Herceptin, which targets just the generic wild type HER2 protein.

Materials and Methods:

The coding region of HER2 cDNA was PCR-amplified using 12 sets of HER2 specific primers in HER2 positive cell lines (SKOV-3, SKBR-3, MDA-MB-453 and MDA-MB-361).

Results:

RT-PCR results showed multiple bands in various regions of the HER2 mRNA. Sequencing of these bands revealed novel alternative splice variants with deletions in exons 13 and 18 of the HER2 gene expressed in addition to the wild-type HER2. Bioinformatics analysis of the deletions revealed a cassette exon in exon 13, and a loss of 42 base pairs in the 3' end of exon 18 compared to the full length HER2. Both the full-length HER2 sequence and the sequence containing the deletions were translated using the ExPASy Translate tool. This revealed an in-frame deletion of 14 amino acids and a novel splice isoform with a deletion in the HER2 protein, which encompasses the entire ATP binding pocket. This was determined by analysis using UniProtKB to identify the composition of amino acids for each domain of HER2.

Discussion:

Our studies have identified novel splice variants in the tyrosine kinas domain of the HER2 gene in HER2-positive cell lines. The loss of an ATP binding site in the HER2 gene may lead to a less active HER2 isoform, which may play a significant role in prognosis. Current work is being carried out to study the regulation of these splice variants and to study the role of splice factor ASF/SF2 and its phosphorylating kinase SRPK1 in the regulation of HER2 splicing, and to elucidate any significant changes in the HER2 signalling pathways. In addition, the expression of these isoforms is currently being investigated in tissues from FFPE and frozen breast tumours.

2.

EVENT TITLE: POSTGRADUATE RESEARCH FORUM

VENUE: UNIVERSITY OF THE WEST OF ENGLAND

DATE: APRIL 12, 2011

TITLE: THE FUNCTION AND ROLE OF ALTERNATIVE SPLICE VARIANTS IN INVASIVE BREAST CANCER

ABSTRACT

Background:

The Human Epidermal Growth Factor Receptor 2 (HER2) is an oncogene expressed in 25-30% of invasive breast cancers. HER2 shares extensive homology with other members of the HER family (HER1, HER3 and HER4), and is constitutively active as an homo-and heterodimer. The HER2 gene encodes an 185kDa transmembrane protein with tyrosine kinase activity. Gene amplification or protein expression of HER2 is a predictor of poor clinical outcome and decreased survival in women with breast cancer, and also indicates a favourable response to Trastuzumab (Herceptin) therapy, or a combinational therapy comprising Herceptin plus chemotherapy. However, resistance to Trastuzumab remains the case in approximately 50% of HER2 amplified/overexpressing tumours. Understanding the molecular mechanisms of Trastuzumab resistance and identifying more effective therapies, is critical in the treatment of patients whose breast cancers express this aggressive disease phenotype. In this study, it is postulated that the abnormal generation of mRNA splice variants may be responsible for the continued tumour growth and progression.

Aims:

The aim of this study is to increase our understanding of the role of HER2 splice variants in the development and progression of breast cancer. This will inform the

development of more sophisticated and effective therapies that target specific HER2 isoforms, rather than Herceptin, which targets just the generic wild type HER2 protein.

Materials and Methods:

The entire coding region of HER2 cDNA was PCR amplified using 12 sets of HER2 specific primers in HER2 positive cell lines (SKOV-3, SKBR-3, MDA-MB-453 and MDA-MB-361), and HER2 negative cell lines (BT-20, MDA-MB-361 MCF-7). These cell lines were also tested for protein expression of HER2 using Immunohistochemistry.

Results:

RT-PCR results suggest that there are alternatively spliced variants of HER2 between exons 12-15, 15-19, and 19-22 in the SKOV-3 and MDA-MB-453cell lines, as well as potential exon deletions in the SKBR-3 cell line.

Conclusion:

In addition to a previously described exon 16 deleted HER2 transcript, our current findings have identified potential novel splice variants in the transmembrane and kinase domains of the HER2 gene in HER2 positive cell lines. Functional studies of the proteins encoded by these variants will be carried in order to elucidate any significant changes in the HER2 signalling pathways. In addition, the expression of these isoforms will be investigated in tissues from formalin fixed, paraffin embedded and frozen breast tumours.

3.

EVENT TITLE: CURRENT ISSUES IN BREAST CANCER CONFERENCE

VENUE: UNIVERSITY OF THE WEST OF ENGLAND

DATE: JULY 3, 2010

TITLE: BIOMEDICAL RESEARCH IN BREAST CANCER

4.

EVENT TITLE: CENTRE FOR RESEARCH IN BIOMEDICINE (CRIB) REVIEW DAY

VENUE: UNIVERSITY OF THE WEST OF ENGLAND

DATE: JULY 09, 2009

TITLE OF PRESENTATION: THE ROLE OF HER2 SPLICE VARIANTS IN INVASIVE BREAST CANCER

5.

EVENT TITLE: CENTRE FOR RESEARCH IN BIOMEDICINE (CRIB) REVIEW DAY

VENUE: UNIVERSITY OF THE WEST OF ENGLAND

DATE: MAY 26, 2009

TITLE OF PRESENTATION: THE FUNCTION AND ROLE OF HER2 SPLICE VARIANTS IN INVASIVE BREAST CANCER

6.

EVENT TITLE: CENTRE FOR RESEARCH IN BIOMEDICINE (CRIB) FORUM

VENUE: UNIVERSITY OF THE WEST OF ENGLAND

DATE: JANUARY 27, 2009

TITLE OF PRESENTATION: THE FUNCTION AND ROLE OF HER2 SPLICE VARIANTS IN INVASIVE BREAST CANCER