Queen Mary
University of London
Barts and The London

Thesis submitted in partial fulfilment of the requirements for the

Degree of Doctor of Philosophy

Delineating early transformational events in HER2 positive breast cancer using an inducible MCF10A cell line

---

August 2019

Ateequllah Hayat

Statement of originality

I, Ateequllah Hayat, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis. I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Signature:

Date: 08.08.2019

The work presented in figures 6.9A, 6.11, 6.12, and 6.13 were performed by Ana-Alexandra Greere as part of her Queen Mary University of London's Master's Project under my supervision. Figure 5.2A was plotted by Gabriella Ficz. The phosphoproteomic sample preparations were carried out by Saul Alvarez Teijeiro.

Acknowledgments

I would like to thank my supervisor, Gabriella Ficz, for her encouragement, guidance and support throughout my PhD.

I owe my thanks to the wonderful members of Ficz lab for their support, assistance and the laughter we have shared together. It's been incredible having you alongside. To Emily, Hemalvi, Michael, and Lily - thank you for everything from beginning to end. I would like to thank to Kriszta and Alexandra whom I have supervised, for their help with this project.

My thanks go to my colleagues and friends in the office. Federico, Arran, Maru, Dave, Ryan, Pedro, Henry, Emma, Faith, and Mariette.

I would like to thank our collaborators Pedro Cutillas and Edward Carter and for the funding I have received from Leverhulme Trust.

I am grateful to my family for their continued support and patience, especially to my dad who motivated me all the way. I would like to thank Qudsia for her love and support throughout and to baby Musa, who made the end of this PhD difficult yet a very special experience.

Abstract

HER2 protein overexpression in breast cancer patients is a predictor of poor prognosis and resistance to therapies. Despite significant advances in the development of targeted therapies and improvements in the 5-year survival rate of metastatic HER2 positive breast cancer patients, new approaches are needed to better understand the disease at an early stage in order to identify means to inhibit its progression. An inducible breast cancer transformation system allows examination of early molecular changes at high temporal resolution. Here, we show that HER2 overexpression to similar levels as those observed in a subtype of HER2 breast cancer patients is sufficient to induce transformation of MCF10A cells. We found that HER2 activation generated gross morphological changes in 3D cell culture, increased anchorage-independent growth of cells and altered the transcriptional programme of various genes associated with oncogenic transformation. Global phosphoproteomic analysis during early transformation uncovered numerous signalling changes associated with cancer upon HER2 overexpression. Candidate pathways included chromatin regulators, in addition to known cascades such as MAPK, focal adhesion, mTOR, and HER signalling pathways. To understand the effect of kinase signalling on chromatin accessibility landscape, we performed ATAC-seq on acini isolated from 3D cell culture. This enables elucidation of HER2 induced signalling effects on chromatin architecture and its contribution to transformation at temporal resolution. Uniquely, we identify that HER2 overexpression promotes reprogramming-associated heterogeneity, with a subset of cells acquiring a stem-like phenotype, expressing breast stem and cancer stem cell markers, making them likely targets for malignant transformation. Our preliminary data show that this population of cells, which counterintuitively enriches for relatively low HER2 protein abundance, possesses transformational drive, resulting in increased anchorage-independent growth *in vitro* compared to cells not enriching for stem markers. Our data provide a discovery platform for signalling to chromatin pathways in HER2-

driven cancers, offering an opportunity for biomarker discovery and identification of novel drug targets.

List of Figures

List of tables

Contents

Acronyms

AKT/PKB	Protein kinase B
ATAC-seq	Assay for transposase-accessible chromatin using sequencing
BMP6	Bone Morphogenetic Protein 6
BMPR2	Bone Morphogenetic Protein Receptor Type 2
ECL	Enhanced chemiluminescence
cDNA	Complementary DNA
CRE	Cyclisation recombination
DAPI	4',6-diamidino-2-phenylindole
DAVID	Functional Annotation Bioinformatics Microarray Analysis
DCIS	Ductal carcinoma in-situ
DMEM	Dulbecco's Modified Eagle Medium
Dox	Doxycycline hyclate
EMT	Epithelial to mesenchymal
EF1α	Elongation factor 1-alpha promoter
EGF	Epidermal growth factor
EGFR/HER1	Epidermal growth factor receptor/ Human epidermal receptor 1
ELK1	ETS Like-1 protein
ERK	extracellular signal-regulated kinase
FACS	Fluorescence-activated cell sorting
FDR	False discovery rate
FISH	Fluorescence in-situ hybridisation
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GFP	Green fluorescent protein
HER2	Human epidermal receptor 2

HRG	Heregulin
IDC	Invasive ductal carcinoma
IGV	Integrative Genomics Viewer
IF	Immunofluorescence analysis
IHC	Immunohistochemistry
ILK	Integrin-linked kinase
LOX	Lysyl oxidase
LOXL2	Lysyl Oxidase Homolog 2
MCF10A	Human normal immortalised mammary epithelial cell line
NFkB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NRG	Neuregulin
pINDUCER21	Inducible vector
RT-PCR	Real-time polymerase chain reaction
rtTA	Reverse transcription transactivator
Tet-On	Tetracycline On system
TGFα	Transforming growth factor alpha
TiO_2	Titanium dioxide
Tn5	Hyperactive transposase enzyme
TRE	Tetracycline response element
VEGFC	Vascular endothelial growth factor C
3D	3-dimensational

Chapter 1

1. Introduction

1.1 Neoplastic transformation

1.1.1 The process of transformation

Transformation takes place when cells acquire the key hallmarks of cancer (1). These include; morphological changes, increased migration and invasion potential, anchorage-independent growth, foci formation, as well as differences in the genetic and epigenetic landscape between normal and transformed cells (2-4). One of the methods to achieve neoplastic transformation of cells is the introduction of cancer associated oncogenic lesion(s). Transformation of normal human cells has been achieved by a step wise process of immortalisation and then conversion of the immortalised cells to metastatic transformation (5). Studies have shown that transformation of normal rodent cells can be achieved by the activation of a single oncogene in immortalised rodent cell lines, as they have already undergone genetic and/or epigenetic changes (6, 7). However, primary rodent cells are transformed by the co-expression of two distinct co-operating oncogenes or in combination with mutation or inactivation of a tumour suppressor gene (8).

Similar strategies have been used to convert normal primary human cells to tumourigenic state. It is suggested that three distinct oncogenic “hits” may be required, which lead to growth-regulating alterations to transform primary human cells (9). The foreskin fibroblasts (BJ), human mammary epithelial cells (HMECs) and human embryonic kidney cells (HEK) were transformed by genomic versions of H-Ras, hTERT and SV-40 LT genes (9-11). This suggests that there are fundamental differences for transformation in rodent versus human cells. An explanation has emerged that may elucidate such differences. The primary rodent cells are

easier to immortalise compared to the primary human cells (12). The latter rarely undergo spontaneous, immortalisation, whereas the rodent cells can be spontaneously immortalised, which indicates that the control of cellular lifespan between the two different cells is very different (13, 14). This change could partially be attributed to telomere biology. Unlike rodent cells, the human cells lack detectable telomerase activity and have relatively shorter telomeres, which erodes and triggers cellular senescence (15, 16). Interestingly, both inbred and wild type mice have telomerase activity, with wild type mice having shorter telomere length as similarly observed in humans, however, the growth characteristics between wild type and inbred animals are similar (17). However, it could be that inbred mice have a more "permissive" genetics, which make them more prone to immortalisation.

The vast majority of the *in vivo* and *in vitro* transformation models have been able to study the events occurring between normal and already transformed cells. This has made it impossible to track the early aberrant events taking place during the process of transformation upon an oncogene induction. Many experimental models both *in vivo* and *in vitro* have implicated a variety of mechanisms involved in oncogene-mediated transformation, but a unified mechanistic system cannot yet be proposed, in part due the lack of understanding of the early events in oncogene mediated tumourigenesis. To overcome the challenge in characterising the earliest changes, such as those in the signalling network and chromatin dynamics during transformation upon oncogene induction, an inducible model could be utilised.

1.1.2 Signalling by HER proteins

The human epidermal receptor (HER) family of proteins belong to the type I transmembrane growth factor receptors that function to activate a rich network of intracellular signalling pathways in response to extracellular signals (18, 19). The HER family has four members that are structurally and functionally very similar; HER1 (EGFR, or ErbB1), HER2 (ErbB2 or neu), HER3 (ErbB3), and HER4 (ErbB4). Their structure consists of an extracellular ligand-binding

domain, a transmembrane domain and an intracellular tyrosine kinase domain. This group of receptors have individual features which include the ligand-deprived HER2 receptor and a kinase-inactive HER3 receptor (20). In mammalian cells, at least 12 ligands are known to induce dimerisation, with each ligand favouring a specific combination of receptor dimerisation in a specific hierarchical order. However, there is a marked preference of HER2 as a dimerising partner of the three other partners (21, 22). The HER2 heterodimer with HER3 generates the most potent intracellular signal compared to those originating from other combinations, because HER2 contains the strongest catalytic kinase activity (23). In addition, HER2 heterodimers have slow ligand dissociation, prolonged firing, rapid recycling, slow endocytosis, slow ligand dissociation and internalisation (24). The HER proteins are normally widely expressed in numerous non-haematopoietic cells and are functionally important (25). The receptors are essential in tissue growth and development and knock out models have shown that they are critical for the development of organs such as lung, brain, gastrointestinal tract and skin (26-28).

The extracellular binding domain of the receptors except HER2 can be in active (open) or inactive (closed) conformation. Upon ligand binding, the extracellular binding domain of the HER protein undergoes structural change to an active conformation, which promotes dimerisation of the receptors. This leads to auto-phosphorylation of the intracellular tyrosine

kinase domain, which initiates a plethora of downstream signalling pathways and cross talks with other signalling proteins, leading to the regulation of numerous cellular activities (29, 30).

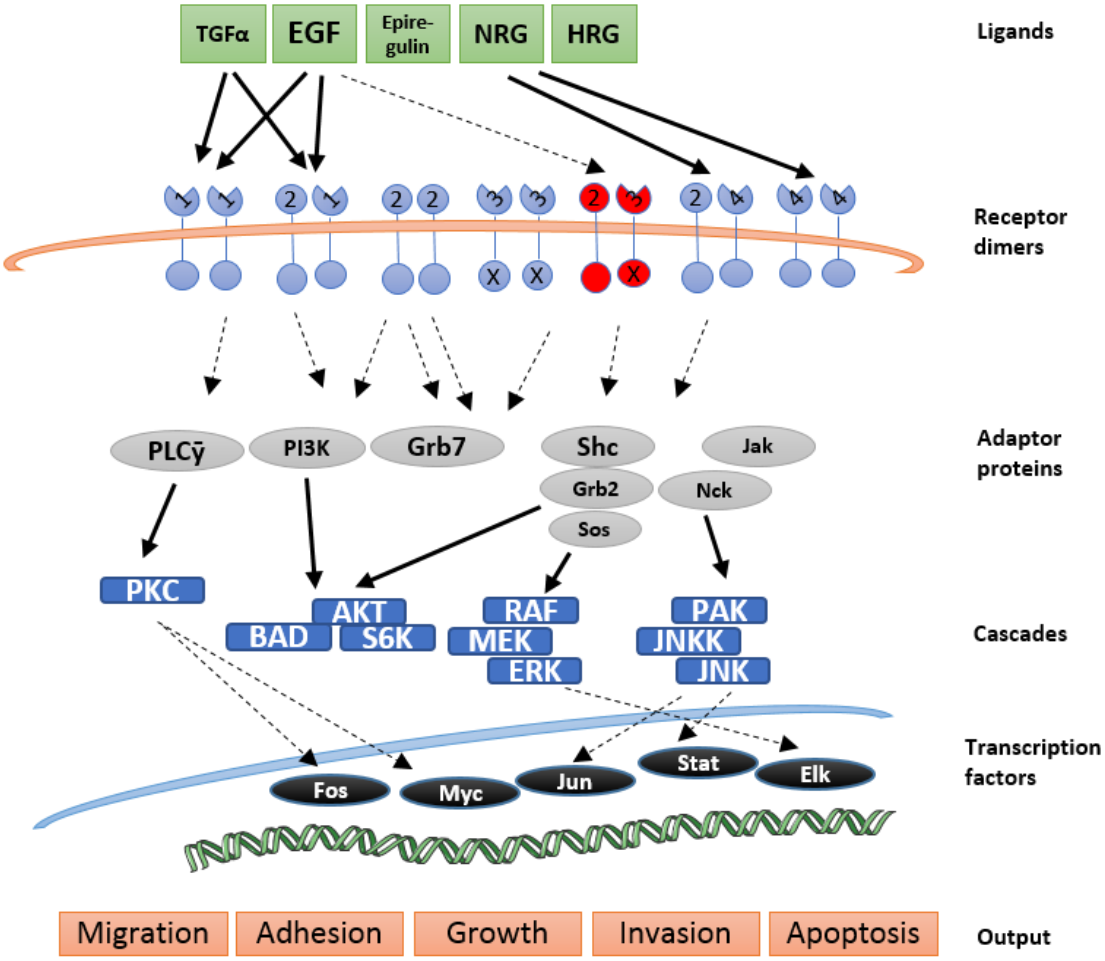


Figure 1.1: A simplified outline of HER signalling network. For simplicity only 5 ligands are shown out of the 12 that have been identified in mammalian cells. Number in each receptor circle or semi-circle indicate the respective receptor of the HER family. HER2 does not bind to any ligand hence a closed (circular) conformation, other HER receptors have an open conformation and the green blocks indicate the respective ligand that induces the dimerisation. HER3 has an inactive catalytic intracellular tyrosine kinase domain indicated by a cross. HER2-HER3 heterodimer is coloured in red because they generate the most potent signals. Signalling is transmitted to the adaptor proteins and enzymes, which activate a large network of signalling cascades of which only some of them are shown here. Signalling pathways

activate various transcription factors but at present time, their translation to a specific type of output is not fully understood.

HER2 does not require a specific ligand as its extracellular binding domain is always in a constitutively open (active) conformation, unlike its family members (Figure 1.1). In normal cells with endogenous levels of HER2 expression, the activation of pathways carefully regulates normal cell growth, adhesion, survival, and differentiation and other biological processes as the dimerisation of receptors and the ensuing activation are temporary and spatially controlled. Furthermore, in normal cells the excess signalling induces apoptosis due to the presence of a wild type p53 and other tumour suppressor genes (31). Expectedly, p53 inactivation is associated with HER2 induced tumours (32).

1.1.3 Transformation potential of HER2

The data supporting the ability of HER2 to transform human cells is compelling. HER2 protein over expression or gene amplification in breast epithelial cells has been shown to cause morphological alterations in the mammary acini and induce proliferation (33). HER2 over expression alone in NIH-3T3 cells is sufficient to transform cells *in vitro* and its over expression in invasive breast cancer cell line (MCF-7) is known to enhance tumourigenicity (34, 35). The evidence showing the transformation potential of neu (nomenclature of HER2 for rodent counterparts) in rodents is also robust and rodent cells are simpler to transform compared to the human cells (36). Transgenic mice with active neu developed mammary adenocarcinomas in a step-wise progression and neu was sufficient to induce transformation (37). Wild type (WT) neu over expression in the basal layer of mouse epidermis allowed for proliferation and tumour formation as early as six weeks (38). Numerous other studies have shown the potent transforming potential of neu inducing malignant transformation in a variety of organs and model systems (39-45).

1.1.4 Transformation of MCF10A cell line

To study the transformational events upon oncogene expression, the focus has historically been on the immortalised yet non-tumourigenic cell lines as a starting model. MCF10A cell line, the human mammary epithelial cells have been extensively used for this purpose. Forced ectopic over expression of constitutively active and inducible oncogenes such as Ha-Ras (46), HER1 (47), B-Raf (48), MYC (BHLH Transcription Factor) (49) NCT (Nicastrin) (50), RANK (Receptor activator of nuclear factor κ B) (51) and HER2 (3, 52) in the mammary epithelial cells produced many transformation associated phenotypic and transcriptional changes. These alterations include the morphological disruption in 3-dimensaional (3D) cell cultures and transcriptomic differences between the oncogene induced transformation relative to control cells (3).

Furthermore, HER2 gene in breast cancer appears to hold the transformational potential through its amplification alone. Therefore, to investigate the effects of HER2 over expression in tumourigenesis, a cell line that contains either “low” or endogenous levels of HER2 would be an ideal starting model to appropriately quantify the impact of HER2 and changes that occur consequently. MCF10A cell line is thought to have very low levels of endogenous HER2. Moreover, since HER receptors work closely with each other the expression of HER2 family members are also of importance in breast cancer. MCF10A cells express “normal” levels of EGFR and very low endogenous levels of HER3 (53, 54). This is essential to understand the first steps of HER2 over expression to dissect its effects and to reliably attribute the changes to the HER2 levels alone without ambiguity from other factors.

MCF10A cells expresses markers associated with basal/myoepithelial and luminal phenotype as is seen in the normal breast (55). When grown in 3D cell culture of matrigel and collagen mixture, MCF10A cells form a lumen as a result of apoptotic (e.g. anoikis) conditions in the centre of the acini (56, 57). This resembles the acini of the normal breast tissue with clustered

lobules, connecting the interlobular ductules with each other (58). This shows that overall MCF10A is a good initial model to study the transformational changes in the context of human mammary breast cancer, and analysis from this model could be further extended.

1.1.5 Epithelial to mesenchymal transition in transformation

Epithelial to mesenchymal transition (EMT) is a reversible cellular process that is known to have important roles in morphogenesis, wound healing, embryogenesis, development, tumour invasiveness and malignant transformation (59-61). During the EMT process, epithelial cells progressively lose their phenotype, which involves remodelling of the cell-extracellular matrix and cell-cell interactions. This results in the detachment of the epithelial cells from each other and the underlying base membrane, resulting in the activation of a new transcriptional programme that encourages the mesenchymal state (62). A widely studied phenotype of cells that have undergone EMT is the transformation of their normal compact – epithelial-cell-like morphology to a more elongated, spindle-like - mesenchymal morphology (63-65). Since EMT is a reversible process, mesenchymal cells can revert back to epithelial cells, known as the mesenchymal to epithelial transition (MET).

Moreover, normal epithelial cells are held together by tight junctions, gap junctions, desmosomes and adherens junctions, which consist of cell surface epithelial cadherin (E-cadherin) genes. This structure is critical for the integrity of epithelial cells. Upon EMT induction, the E-cadherins are downregulated alongside the repression and activation of other markers, leading to the arising of mesenchymal cells. This involves the breakdown of normal morphology of cells and acquisition of a more fibroblastic mesenchymal phenotype (60). The malignant transformation of many different tumours is dependent on EMT activation (66, 67). In transformation, the consequences of EMT activation are the degradation of the underlying

basement membrane, disruption to cell-cell interactions and cell polarity, as well as the abnormal reorganisation of the extracellular matrix (68) (Figure 1.2).

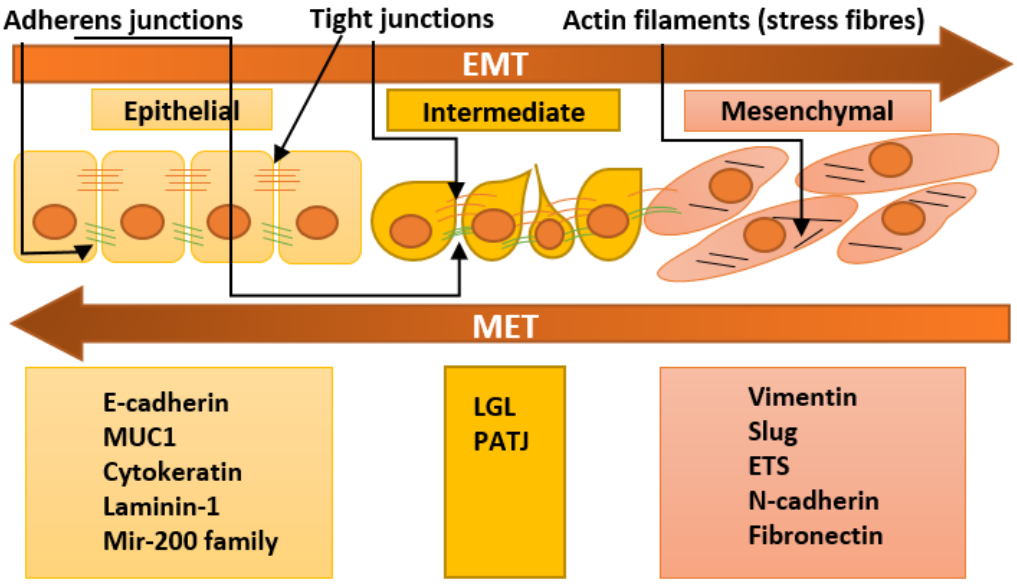


Figure 1.2: Schematic of EMT programme. Epithelial cells are held together by adherens junction, tight junctions, and are linked to the basement membrane by hemi-desmosomes. These cells express genes that are associated with the epithelial state and sustain polarity of cells (list of genes in light yellow box). The epithelial state has downregulation of molecules associated with mesenchymal state. EMT induction results in the expression of genes associated with mesenchymal state (listed in the orange box) and the concomitant down regulation of the epithelial genes. The alterations in gene expression in epithelial state leads to disruption of tight junctions, adherens junctions and the disassembly of cell-cell and cell-basement membrane attachments. Epithelial cells progressively lose their features by the acquisition of intermediate stage and associated gene expression. In certain circumstances, full EMT features occur but cells rarely advance to complete mesenchymal state. EMT is a reversible programme, and cells can revert back by undergoing MET.

In addition, prominent genes that are associated with the epithelial state, such as cytokeratin and E-cadherin are repressed, whilst at the same time, expression of genes that are linked to the mesenchymal state are activated. These include fibronectin, N-cadherin and vimentin (69).

Furthermore, in clinical setting, the protein markers that are associated with EMT activation could be used as specific indicators of high grade malignant transformation by pathologists (70). However, the transition from epithelial to mesenchymal state does not work as a binary switch and cancer cells do not always execute the complete EMT reprogramming to drive cells to an unequivocal mesenchymal state. The process appears to be more dynamic, and that is crucial for driving tumourigenesis, which contributes to full malignant transformation (62, 71).

Likewise, in cancer progression it has been widely known that during early carcinomas, cells are in the epithelial-cell-like state, and as the transformation progresses, cells gradually gain more mesenchymal features. The EMT activation in cancer cells has been associated with higher resistance to several therapies (68). Additionally, in breast cancer cells, the EMT programme is known to associate with more cancer stem-like phenotype, which in turn has a higher transformational potential (72).

1.1.6 Breast cancer progression – the role of HER2 over expression

It has been documented that upregulated levels of HER2 expression can be detected in mammary tissues that show features of partial transformation, but are not yet completely transformed. Generally, HER2 is expressed at low levels or is absent in benign breast lesions (73, 74). For example, HER2 is almost undetectable in terminal ductal lobular units (TDLUs), and has been detected at very low levels (0-9%) in atypical ductal hyperplasia (ADH) (75). In contrast, HER2 protein over expression and gene amplification are readily detected in the pre-invasive stage, in ductal carcinoma in situ (DCIS), with approximately 70% of patients exhibiting HER2 over expression (76-78). The progression from low HER2 expression levels – or its absence – in the benign breast biopsies to high incidence of HER2 over expression in the pre-invasive stage of the disease suggests that HER2 over expression is an early lesion in breast tumourigenesis. However, not all of the DCIS cases possess the ability to invade and metastasise, since about 20%-30% of the invasive breast cancers have HER2 over

expression/amplification. Thus, this points to the possibility of there being underlying differences causing some of the pre-invasive breast disease to either remain stable or progress to an invasive stage. It has been known that minor aberrations in HER2 over expressions cases are sufficient to induce transformation (79). Additional alterations alongside HER2 over expression may also play an important role in the progression of HER2 positive breast cancers from benign to invasive disease, such as associated abnormalities in p53 and E-cadherin genes (80, 81).

On the other hand, there is evidence to suggest that HER2 over expression does not change between primary tumours and those that metastasise. For instance, there was no drastic change in HER2 expression between primary tumours and lymph node metastatic cases, as HER2 over expression was found in 55% of primary disease, but also in metastatic cases at the same incidence rate (82). Indeed, there are many studies that have shown very little to no difference in HER2 over expression status between primary tumours and the corresponding metastatic breast cancer stage (83-87).

1.2 HER2 in cancer

1.2.1 HER2 over expression in cancer

The data from the experimental models is well supported by a significant body of clinical data from patients. HER2 is over expressed in approximately 20-30% of breast and ovarian cancers and is correlated with worse prognosis (88, 89). In addition, over expression of HER2 is observed in lung, head and neck, endometrial, oesophageal and kidney cancers and is also associated with worse prognosis (90) (Figure 1.3). HER2 over expression is a significant and early event in breast tumourigenesis and its expression is sustained through the different stages of breast cancer, from early detection, to invasive disease, to node and finally distal metastasis (91, 92). However, despite HER2 being maintained throughout disease progression, its over expression in early stage defines a sub type of breast cancer (HER2 positive), notwithstanding its expression at later stages (92-95).

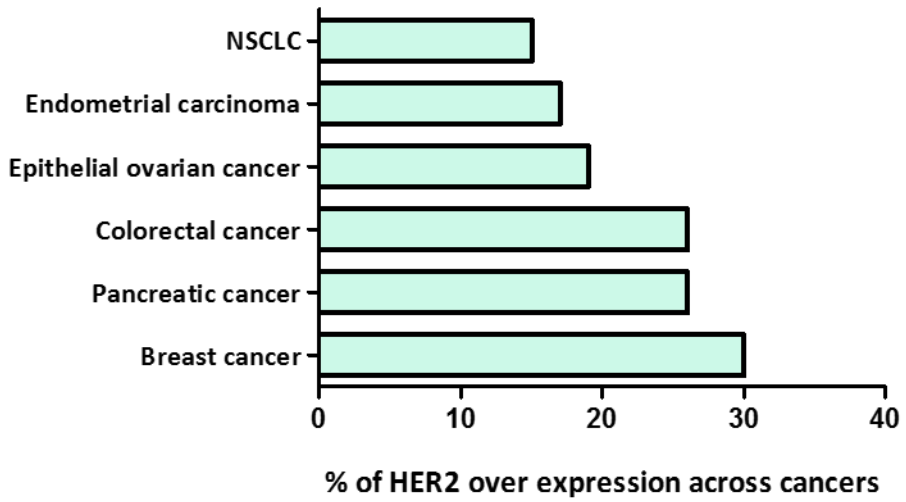


Figure 1.3: HER2 protein over expression in various malignancies. HER2 gene amplification and protein over expression has been identified in many cancer types. Only a few different types of cancers are

selected here out of the many other HER2 over expressing cancers. HER2 over expression shown here is determined by IHC and/or FISH.

Different studies have reported that within the same cancer type there is a wide range of variation/heterogeneity in the pattern of HER2 over expression, despite the same standardised fluorescence in situ hybridisation (FISH) and immunohistochemistry (IHC) analysis being used for detection (96). The source(s) of variation in HER2 expression is not yet elucidated, it could however be the intra-laboratory techniques, or that this subtype of breast cancer is extremely heterogeneous. The effect of HER2 over expression in breast cancer is well characterised. However, the clinical behaviour of HER2 in patients displaying varying levels of heterogeneity require much additional study. The intra-heterogeneity of HER2 expression within the same patient requires additional investigation to dissect if different levels of HER2 expression have different transformation potential.

1.2.2 HER2 positive breast cancer

HER2 over expression is the result of HER2 gene amplification and/or increased transcription. The extent of HER2 over expression can be evaluated at mRNA level by Real-Time PCR (RT-PCR) and FISH, or by IHC to quantify the protein levels. Currently, breast cancer patients undergo testing to check for HER2 positivity but the ideal way to evaluate the HER2 positive status remains unclear and controversial, because there is no standardised criteria for assessing HER2 as a prognostic marker (97). However, the guidelines have been updated in 2013 and more recently focused updated in 2018 by the American Society of Clinical Oncology (ASCO)/College of American Pathologists (CAP) (98, 99). The guidelines are based on HER2 gene and/or HER2 protein assessment and recommend the use of FISH and IHC assays to inform diagnoses of HER2 positivity in breast cancer. The utility of RT-PCR to diagnose or serve as a substitute for either FISH or IHC remains unclear because of high rates of false negative results and insufficient evidence to support its use as it is not fully validated in diagnostic settings (100).

The extent of HER2 protein over expression is firstly diagnosed by IHC to grade the tumour. If the protein staining results in unequivocal 0 or 1+ grade, the cancer is considered HER2 negative. If the results are unequivocal 3+ grade, the cancer is considered HER2 positive. If it is graded equivocal 2+, subsequent FISH analysis is used to determine the positivity of HER2 gene amplification (101).

In a phase III clinical trial (CLEOPATRA), women who have higher HER2 mRNA or protein over expression corresponds to a higher magnitude of benefit from Trastuzumab (Herceptin) (102). However, this is not true in all clinical cases as Trastuzumab treatments of lower HER2 expressing tumours are still associated with clinical benefit (103). This might account for the spatial heterogeneity and variation of HER2 expression which under appreciates the bona fide percentage of HER2 positivity in cells.

1.2.3 Clinical evidence of anti-HER2 therapies

Over the past 20 years, there have been significant advances in the therapeutic strategies employed for the treatment of HER2 positive breast cancer. The commonly recommended anti-HER2 therapies include trastuzumab, lapatinib, ado-trastuzumab emtansine (T-DM1), and pertuzumab (104). Trastuzumab was approved as a first-line treatment alongside paclitaxel for metastatic HER2 positive breast cancer after it was approved in 1998 (105). The benefit of trastuzumab in treating patients with metastatic disease has been well documented in clinical trials led by the North Central Cancer Treatment Group (NCCTG) and the National Surgical Adjuvant Breast and Bowel Project (NSABP). The results of the trials compared chemotherapy with or without trastuzumab. They found, after a follow up of two years, that there were 133 events in the trastuzumab group compared to 261 events in the chemotherapy treated patients without trastuzumab. The percentages of patients alive in the trastuzumab treated group were 87.1% compared to 75.4% in the control group in the medial follow-up of two years. At four years, the percentage of patients alive with trastuzumab were 85.3%, compared

to 67.1% in the control group (106). Therefore, trastuzumab in combination with paclitaxel in adjuvant setting significantly improved patients' disease-free survival (DFS) and overall survival (OS).

Furthermore, lapatinib – which is known to target the intracellular tyrosine kinase domain of the HER2 receptor (107) – was shown to be effective in treating HER2 positive breast cancer tumours that were resistant to trastuzumab (108, 109). Several clinical trials have shown that the combination of lapatinib with trastuzumab had significantly better progression-free survival (PFS) than lapatinib treatment alone (110). The median survival for the combination treatment was 12 weeks compared to 8.1 weeks with lapatinib alone (111).

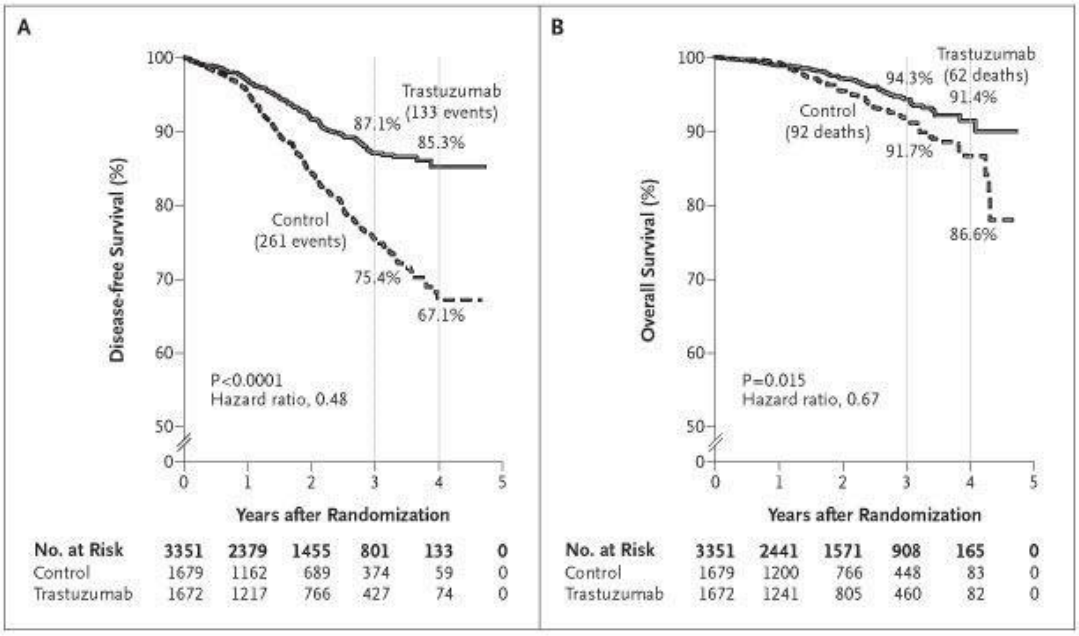


Figure 1.4: Kaplan-Meier curves depicting disease-free survival (DFS) (A) and overall survival (OS) (B) of patient treated with chemotherapy alone (control) vs Trastuzumab and chemotherapy.

1.2.4 Inducible transformation models in cancer

To achieve dose-dependent, reversible and uniform temporal control of a gene of interest, an inducible system has obvious advantages in many experimental settings. Numerous inducible

systems have been developed to understand the function of specific lesions in different diseases. For instance, a mouse embryonic cell line called C3H/10T1/2 with *ras* oncogene under the transcriptional control of the inducible mouse metallothionein-I promoter induced by heavy meal (zinc) ions induced conditional and reversible transformation (112). In addition, it has been shown that mutations in the DNA methyltransferase (DNMT3a) using Cre-inducible (cyclisation recombination) system and nucleophosmin (NMP1) enhanced clonogenic potential and eventually induced transformation in experimental mice (113). Furthermore, a doxycycline (dox) inducible H-RAS V12G mutation induced in the melanocytes of mice resulted in them developing spontaneous melanomas that eventually regressed upon withdrawal of doxycycline (114). Inducible HER2 over expression in primary human mammary cells induced various tumourigenic alterations to the ductal bilayer observed in early breast tumorigenesis (115).

The inducible models mentioned above and others provide advantages over conventional non-inducible systems. Firstly, in some cases the expression of constitutively active gene could be toxic to the cells, therefore the ability to control the timing and levels of ectopically expressed transgenes is extremely valuable. Secondly, inducibility grants the ability to track and characterise the very early molecular changes that occur upon gene induction, which would be impossible to capture otherwise. Thirdly, the reversibility of gene expression and phenotype upon withdrawal of the inducing agent can be useful to investigate because once the stimulus is removed, an inducible gene returns to inactive, basal level.

The inducible systems in published works have been valuable to show the combination of oncogenes required for transformation. However, in many cases they do not reflect the endogenous expression of oncogenes or the inactivation of tumour suppressor genes as presented clinically in patients. They are based on the forced ectopic expression of genes that are not normally seen in tumours. Therefore, there is a need for an inducible *in vitro* system to model transformation in a way that better reflects the early progression of HER2 breast

cancer, with endogenous levels of expression of oncogene(s) comparable to those observed in physiological conditions.

1.2.5 Inducible transformational vectors

1.2.5.1 Tetracycline inducible system

Conditional gene expression tools that can control the induction and reversibility of a gene are essential tools in research with broad applications. The tetracycline inducible vector is a responsive and tightly regulated system that produces robust expression of a gene of interest in the target cells (116). There are two subtypes of tetracycline inducible system.

Firstly the Tet-On, which is based on the reverse tetracycline controlled transactivator (rtTA). The rtTA consists of VP16 transactivation domain and the TetR repressor. The tetracycline response element (TRE), which is the inducible element (promoter) contains the Tet operator (TetO) sequence can bind the rtTA in the TRE of the target transgene in the presence of doxycycline (dox). Thus, the addition of dox regulates the expression of the gene of interest quantitatively and temporally (Figure 1.4). The Tet-Off system functions in the opposite manner, in the presence of dox, expression from the TRE is reduced, resulting in blocking of transcription (116, 117).

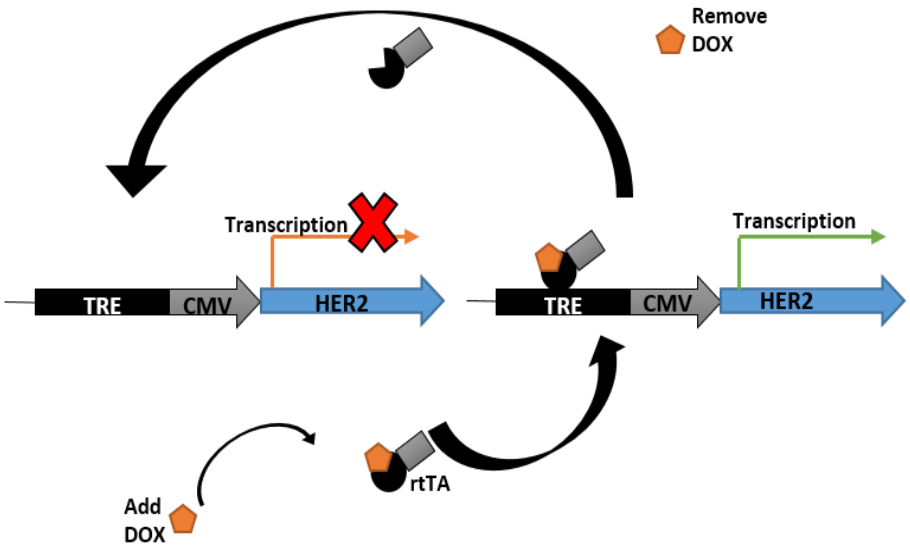


Figure 1.5: Schematic of tetracycline (ON) inducible system. In the presence of an inducing agent (dox), rtTA is bound by it, which binds to the TRE and induces the expression of HER2. Withdrawal of dox leaves the rtTA empty and therefore, transcription of HER2 is blocked.

1.2.5.2 Cre-Lox Inducible technology

Cre-Lox inducible model derived from the P1 bacteriophage is a specific and potent system for conditional control of gene expression. The inactivation of the allele is maintained by an inhibitory cassette called the lox-STOP-lox or LSL. The cre recombinase enzyme recognises the loxP sites (34bp recognition site), which results in the recombination reaction and the removal of one loxP site and the STOP cassette making the LSL cassette dysfunctional and thus permits the activation of a target gene (6, 118).

These two inducible systems are the most widely used and reported in the literature out of the many other inducible models that exists. In our study, we have used the Tet-On system because it offers tight control of gene expression and is reversible upon dox withdrawal and as we require the activation of gene occasionally upon dox treatment, using the Tet-On system is the most appropriate model.

1.3 Chromatin and transcriptional regulation

1.3.1 Mechanisms of transcriptional regulation

1.3.1.1 The structure of chromatin

In eukaryotic cells, the genomic DNA is not found naked but is bound by proteins which is tightly and efficiently packaged. The combination of the compacted proteins and DNA is known as chromatin. The canonical nucleosome, which is the repeating unit of chromatin, is formed by wrapping approximately 145-147 base pairs of DNA around the histone octamer (H2A, H2B, H3, H4 - two molecules of each histone) (119, 120). Nucleosomes are connected to each other by linker DNA to form nucleosomal arrays, also known as the beads-on-a-string structure (10 nm fibre), where each nucleosome is linearly and individually organised. Fibre-fibre interactions can contribute to higher order conformations and cause chromatin to become condensed (121). This generates the secondary chromatin structure (a 30 nm fibre) and eventually produces the high-order chromatin known as the tertiary structure, which can compact the original DNA by an extraordinary 10,000-fold of its original length (122).

The chromatin structure is dependent on environmental cues and stimuli, which can make chromatin highly accessible or inaccessible. Therefore, chromatin structure has a significant impact on transcriptional regulation. Chromatin is classified into two states: heterochromatin and euchromatin. Heterochromatin is highly compact and condensed ("inactive") chromatin and covers approximately 96% of the mammalian genome. Euchromatin refers to decondensed or open ("active") chromatin and comprises approximately 2-3% of the entire DNA sequence but captures over 90% of transcription factors (TFs) bound to it (123) (Figure 1.5). There are ever increasing numbers of post-translational modifications (PTMs) that are being identified, alongside nucleosome-binding proteins, architectural chromatin proteins (ACPs) and ATP-hydrolysis dependent chromatin re-modellers (such as the SN1/SWF family re-

modellers) impacting the conformation, and essentially the active and inactive states of chromatin at all levels.

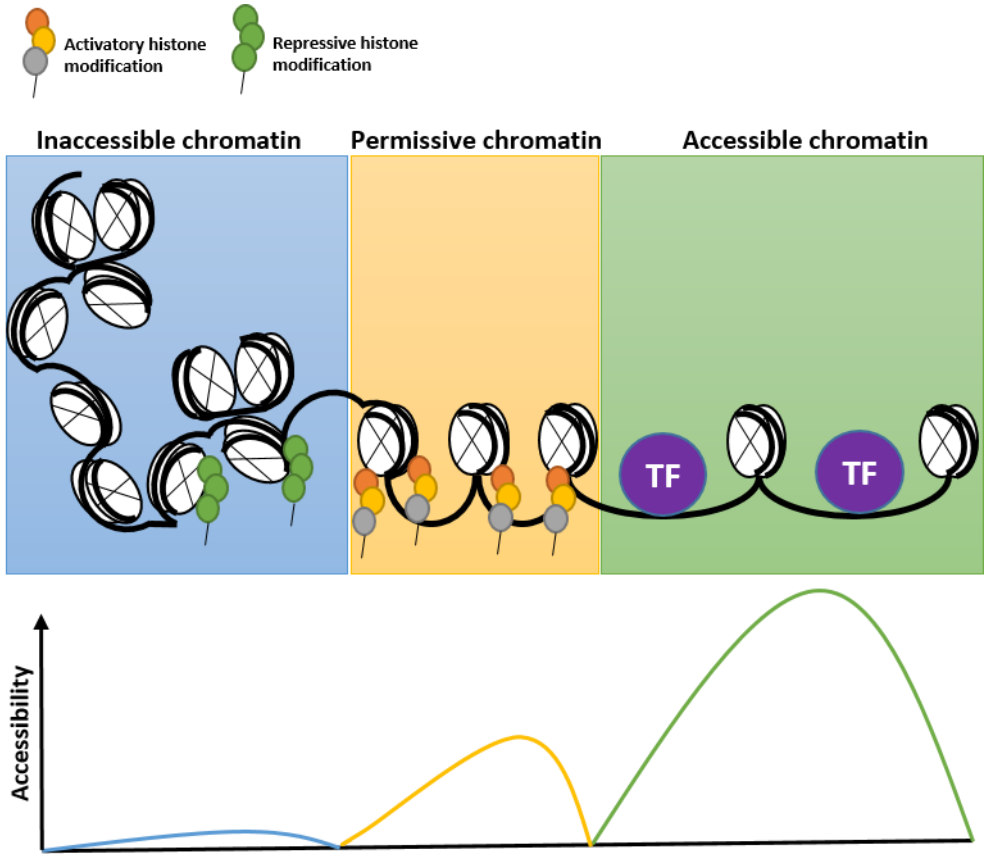


Figure 1.6: Simplified model of chromatin states. Majority of DNA is packaged into inactive (closed) chromatin marked by the repressive histone modifications and methylated CpG islands. Permissive chromatin is sufficiently dynamic to be modified by active histone modifications, which mediate remodelling and establishes an open chromatin state. The topological organisation of nucleosomes regulate chromatin accessibility through various distinct mechanisms such as altering the transcription factor binding to the DNA. The graph shows closed chromatin (blue) indicated by a lower peak, lower accessibility. Permissive chromatin (yellow), which is an intermediate stage has open chromatin with the nucleosomes arranged linearly and individually, shows an increase in chromatin accessibility. The open chromatin (green), is below sub-nucleosomal level and has a higher peak indicating chromatin is open and accessible.

1.3.2 Transcription by epigenetic regulation

DNA methylation, catalysed by one of the DNA methyltransferases (DNMTs), is an epigenetic modification and has been associated with both activation and repression of genes (124). In cancer cells, some CpG islands in promoter sequences become highly methylated, resulting in transcriptional repression of tumour suppressor genes. The gene bodies are generally methylated in normal cells, and this pattern is reversed in cancer cells (125). Histone methyltransferases (HMTs) are specific enzymes catalysing the methylation of histone tails. The lysine methylation marks are both linked to activation and inactivation. For example, HEK9me3 and H3K27me3 are both repressive methylation marks (126). The H3K79me, K3K4me3 and HEK36me3 are associated with active transcription (127).

Acetylation of histone residues is generally associated with transcriptional activation. The histone acetyl-transferases (HATs) are recruited to the histone tails to catalyse the addition of an acetyl group, which promotes transcription. The histone de-acetyl-transferases (HDACs) are repressors and reverse this modification (128). Several activatory acetylation marks include H4K16ac and H3K14ac (129, 130).

Furthermore, the activity of kinases associated with intracellular signalling pathways have been linked to changes in gene expression. For instance, MAPK, c-Jun, and PKC can directly catalyse the phosphorylation of various histones and have been correlated with gene activation (131).

1.3.2.1 Chromatin accessibility in cancer

Recent technological developments have dramatically improved our ability to measure chromatin accessibility by decreasing the amount of biological material required to levels that

are clinically achievable. This has made it possible to catalogue chromatin architectural changes between normal and transformed cells (132, 133). The phenotypic changes observed in tumour progression would most likely require transcriptional and/or epigenetic changes that drive migration, invasion, and metastasis (134). At present time, it appears that there is no universal signature of chromatin accessibility of normal versus cancer cells.

However, it has been shown that the over expression of a transcription factor known as Nfib (nuclear factor I B) is sufficient to globally alter the chromatin state (135). Nfib was shown to transform cells *in vivo* and induce widespread increase in the chromatin accessibility. In addition, there was a dramatic increase in the chromatin accessibility between primary tumours and metastatic cancer (135). Furthermore, SETD2 mutation was found to alter the chromatin organisation in primary human kidney tumours (97). It has been found that there was widespread decompaction of heterochromatin in actively transcribed genes of cancer cells compared to normal cells. These chromatin accessibility changes were associated with defects in RNA processing (136).

Additionally, it is known that, as cells progress from an embryonic stem cell state to a more differentiated state, the proportion of accessible chromatin regions is reduced. In transformed cells, the accessible chromatin landscape, which is normally repressed in the developmental programme is re-activated. It has been shown that, whilst the chromatin accessible landscape of normal cells is clearly distinct, cancer chromatin accessible regions resemble those found in embryonic stem cells (132). Accessible regions in pancreatic, prostate and lung adenocarcinoma cells coincide with endodermal stage of development, whereas malignant melanoma and mammary ductal carcinoma open chromatin loci converge with ectodermal stage of development. Overall, the majority (88-97%) of the accessible chromatin regions found in tumourigenesis of 21 different cancer cell lines were also found in normal foetal and adult cells or tissues (132). Furthermore, ATAC-seq (Assay for Transposase-Accessible

Chromatin using sequencing) analysis of cutaneous T cell lymphoma (CTCL) displayed distinct open chromatin signatures as patient samples aligned very well with H3K27ac (active histone modification mark), showing that the detected regions of the DNA are accessible and open. They found that the sites that were highly accessible in normal cells were less accessible in cancer cells, indicating disease-specific signatures between normal and cancer cells (137).

Interestingly, the acidosis-adapted colorectal cancer cell line (SW60-AA), showing enhanced invasion and metastasis *in vivo*, had 12,010 fewer ATAC-peaks, indicating a reduction in the overall accessibility compared to non-acidosis-adapted SW60 cell line (138). Furthermore, knockdown of ARID1A and ARID1B in colorectal carcinoma cells resulted in decreased ATAC-chromatin accessibility at 112,623 sites (12.5%) but showed increase in chromatin accessibility at 5264 sites (5.2%). The effect of decreased accessibility by ARID1B was only possible when ARID1A was not present, as ARID1B knock down had no effect (139). Moreover, over expression of nuclear auto-antigenic sperm protein (NASP) induces *in vitro* transformation in hepatocellular carcinoma and forms tumours *in vivo.* NASP over expression is also known to decrease chromatin accessibility, as its knock down leads to enhanced chromatin accessibility and transcription (140).

1.3.3 Signalling to chromatin

Accessibility of DNA within chromatin is an important feature that impacts DNA-dependent functions such as replication, transcription and repair. The structure of chromatin can be locally and globally altered by interactions with architectural proteins such as High-Mobility Group (HMG) proteins that influences chromatin accessibility (141). The activation of MAPK signalling pathway by the addition of a stimulus such as EGF, propagates signalling from the cellular membrane through to the nucleus, resulting in histone tail modification and induction of transcription (142). The induction of MAPK pathway leads to the activation of nuclear kinases such MSK1 and MSK2 that phosphorylate histone H3 on serine 10 and serine 28. These

phosphorylation events are rapid, occurring within minutes of stimulation with factors such as UV irradiation, Anisomycin, and EGF. This leads to the activation of immediate-early genes (e.g. junB, c-myc, c-fos, junD, fosB) regulating chromatin accessibility (143).

Furthermore, the phosphorylation and mutations in transcription factors such STAT5 and STAT3 can activate the JAK-STAT pathway. It is known that STAT3 acetylation by histone acetyltransferases can promote transcriptional activation as a result of chromatin remodelling (144). The silencing of the JAK-STAT pathway can globally affect the heterochromatin through the disruption of HP1 binding. This is especially important in differentiation, as the formation of heterochromatin leads to silencing of genes whose inactivation is required during differentiation (145, 146).

1.3.4 Cellular hierarchy in the breast tissue

Breast cancer is extremely heterogeneous and has been categorised into at least five different subclasses (147-149). These are luminal A, luminal B, basal, normal-like, and HER2 over expressing breast cancers. It has been widely recognised that the mammary compartment is made up of the inner luminal cells, which is covered by the outer layer of myoepithelial cells. Nevertheless, there is growing evidence which suggests that the mammary epithelium compartment exists as a cellular hierarchy spanning from stem cells, to biprogenitor cells, to the fully differentiated cells (150-152). The mammary stem cells, also known as MaSCs, have the self-renewal ability and organise the development of the breast gland during embryonic development. In the stem cell hierarchy model, stem and progenitor cells are of great interest, as they are possible targets for initial transformational events and cancer cells are generated from the stem cell population (153, 154).

Evidence has shown the existence of breast cancer stem cells that express surface stem proteins, such as CD44 +ve and CD24 -ve phenotype, exhibit increased tumour formation ability compared to other breast cancer tumours (155). The markers of cancer stem cells in the

human mammary gland are inferred from *in vitro* assays, flow cytometry and xenotransplantation. Generally, existing data has shown that the MaSCs are enriched for CD44 +ve, ALDH1 +ve, CD49F +ve, EpCAM -ve, and MUC1 -ve, with a more basal-like phenotype in the mammary compartment (154-156). The second most abundant cell type in the mammary epithelial hierarchy are the bipotent progenitors that have MUC1 -ve, EpCAM -ve and CD49F +ve phenotype, which are characterised as being more luminal-like. These cells can diverge into ductal epithelial cells or ductal myoepithelial cells, which enrich for the CD49 +ve or EpCAM -ve phenotype.

It appears that there is no universal breast cancer stem cell set of markers, since combinations of different stem markers have been associated with different breast tumours. For example, the CD44 +ve, EpCAM +ve, CD24 -ve phenotype was found in more than 80% of tumours analysed in a study (155). Furthermore, mammospheres generated from CD44 +ve and CD24 -ve cells resulted in tumours in immunodeficient mice (157). In other cases, the expression of ALDH1 protein is a predictor of poor outcome in patients, and its expression alongside CD44 +ve and CD24 -ve phenotype is associated with heightened tumourigenicity (158, 159).

1.4 Tumour heterogeneity

1.4.1 Intra- and inter- tumoural heterogeneity in cancer

Intra-tumoural heterogeneity, which has long been recognised, refers to the existence of distinct cellular populations with specific phenotypic features/markers within a tumour (160, 161). This phenomenon has been well-characterised in many different types of cancers including breast cancer (162), colorectal cancer (163), ovarian cancer (164, 165), brain cancer (166), and kidney cancer (167). Within cancers, variations can occur by multiple biological processes. These could be alterations in the genetic code or in the epigenome between single cells, or macroscopic heterogeneity involving changes in the morphology between regions of the same tumour. There is significant evidence of intra-tumoural heterogeneity shown in the early breast cancer, in ductal carcinoma in-situ (DCIS) stage of the disease (168). The evidence for heterogeneity is provided by traditional histopathology, biomarker expression (169), genetic signature (170) and non-genomic lesions such epigenomics (171), metabolomics (172, 173), and transcriptomics (174, 175). The histopathological intra-tumoural heterogeneity in DCIS include mitotic features, chromatin rearrangements, nuclear size and nucleolar prominence (176-178). About 50% of DCIS cases exhibit multiple architectural characteristics such as concurrent cribriform and solid and micro-papillary features, concurrent cribriform and micro-papillary features, and concurrent cribriform and solid features and so on (179-182). Furthermore, most cases of DCIS present some degree of heterogeneity when evaluated for biomarker expression. Approximately 70% of DCIS cases are oestrogen receptor (ER) positive (183). Similarly, HER2 over expression is observed heterogeneously in DCIS, with clusters of spatially intense regions to adjacent unamplified regions (184). Other markers such as p16, COX-2, p53, and ki67 also exhibit heterogeneous expression (185-187).

One of the mechanism for generating intra-tumoural heterogeneity is the presence of stem-like phenotype in tumours (“stemness”) (188). It is known that a subset of stem cells within a

tumour can self-renew and differentiate into many other types of cells, each type having its own capabilities and phenotypes (189-191). As the process of differentiation takes place, tumours are organised into a hierarchy of distinct cell types, including tumourigenic cancer stem cells, which can give rise to intermediate progenitors and differentiated cells (189). Therefore, these cancer stem cells are a source of intra- and inter- heterogeneity as well as being drivers of tumour initiation (190).

Other types of heterogeneity have also been described in cancer biology. The most well-known is inter-patient heterogeneity, which suggests that any two patients carrying the same sub-type of tumour are not the same and will have distinct clinical behaviour before and/or after treatment. This could be due to a variety of factors such as differences in the epigenome, mutations that arise within the tumour of individual patients, germline alterations and the tumour microenvironment (192). The resulting metastasis from primary tumours can give rise to distinct cellular populations, which consequently gives rise to heterogeneity in metastases from the same subtype of tumour, known as inter-metastatic heterogeneity (192). Moreover, each heterogeneous metastatic cancer can independently evolve and acquire different genetic mutations and/or epigenetic changes, which results in intra-metastatic heterogeneity (192). The intertumour heterogeneity in early breast carcinoma is illustrated by disease stage based on imaging and physical examinations.

1.4.2 How cancer heterogeneity arises

Tumour heterogeneity can be the consequence of genetic and non-genetic sources. The latter include epigenetic alterations or concerted or stochastic biological and biochemical processes within each cell and heterogeneous cancer cell microenvironment (193, 194). The epigenetic factors can include upregulation of polycomb group proteins of transcriptional repressors such as EZH2 and BMI-1, which are associated with normal stem cell self-renewal. These can have heterogeneous expression levels in tumours and contribute to tumourigenesis (195). The

genetic causes include cancers that spontaneously arise through clonal evolution and acquire "driver" mutations, which impact the cancer cell survival and proliferation, alongside "passenger" events that are believed to be phenotypically repressed and do not grant the tumour a selective fitness advantage (196).

One model for clonal evolution in cancer is that most cancers arise from a single previously normal cell, which gives it a sequential selective advantage over the adjacent normal cell triggering many other clonal expansions and the acquisition of driver aberrations, which will eventually outgrow and outcompete the normal cell in a typical Darwinian-like clonal evolution. This model does not suggest that a single mutation cannot affect other cells in the tissue, but suggests that the tumour results in linear steps and that the developing tumour evolves from the progeny of a single cell (197). Nevertheless, evidence is increasingly showing that cancer populations have multiple separate subpopulations that have distinct genetic make-up, at different locations that co-exist within the same tumour, rather than being the consequence of a series of gradual intermediates (198) (Figure 1.6).

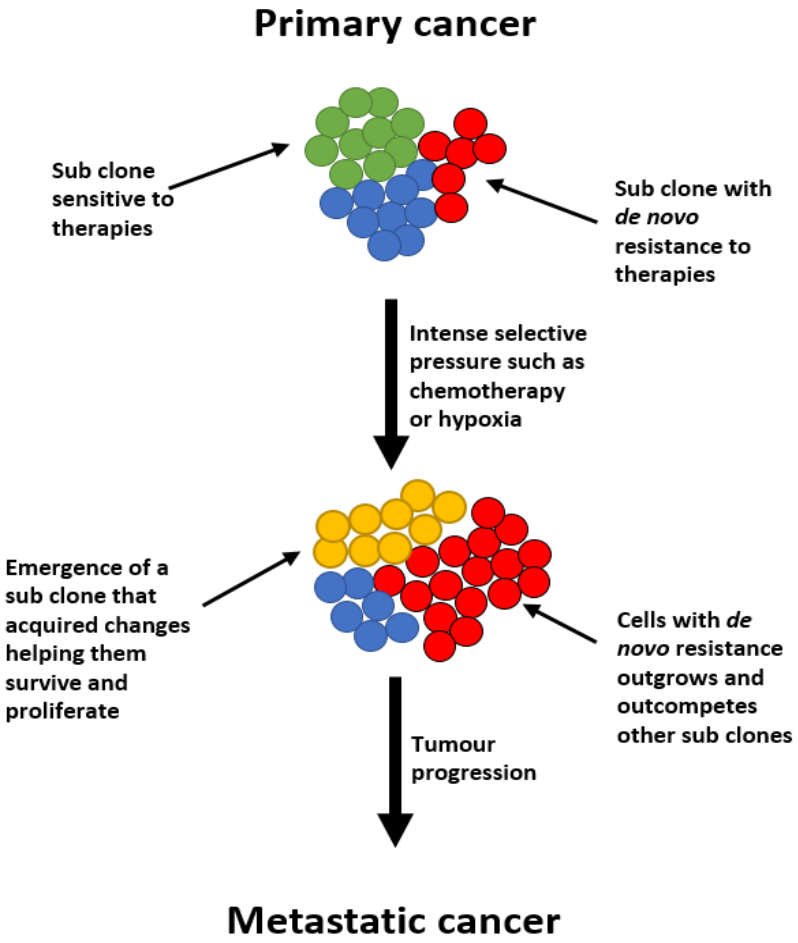


Figure 1.7: Cancer metastasis and Intra-tumoural heterogeneity. Primary cancers comprises many distinct types of sub clones which may be subjected to a variety of selection pressures such as chemotherapy. Under such type of selection pressures, sub clones (green) that are sensitive to therapies are diminished as a result of therapy. Sub clones (red) with *de novo* resistance outgrow and dominate the tumour mass, contributing to cancer progression. Other sub clones (yellow) may also emerge as the tumour acquires secondary mutations, which could potentially lead to cancer metastasis.

1.4.3 Evidence of heterogeneity in HER2 positive breast cancer

HER2 positive breast cancers exhibit cell to cell, temporal, and spatial heterogeneity both at inter- and intra-tumoural levels, as has been acknowledged for some time. The heterogeneous nature of this cancer might explain why it remains a challenging task to treat it, despite having well established treatments such as Trastuzumab and Lapatinib. The HER2 protein staining and

gene amplification can be highly heterogeneous (199, 200), and can ultimately impact disease-free survival (DFS) (201). Some cases of HER2 positive cancers can have gene amplifications by FISH without protein over expression, or protein over expression by IHC without gene amplification, or substantial intra-tumoural heterogeneity (202). The amplification of HER2 gene in a single location of a tumour is sufficient to categorise a tumour as HER2 amplified. This maximises patient eligibility for personalised medicine without consideration of clinical implications of intra-tumoural heterogeneity (203). Heterogeneous expression of other markers in HER2 positive cancer has been noted and these include HER1 (EGFR) (204), c-myc (205), p53 (199), PCNA (Proliferating cell nuclear antigen) (206), and cyclin D1 amongst other proteins (205). Epigenetic silencing of RASSF1A (Ras Association Domain Family Member 1) (207) and p16 (208) has also been recorded.

Interestingly, the borderline equivocal (2+) cases of HER2 positive cancers tend to have a higher HER2 biomarker heterogeneity than the unequivocal (3+ or 0/1+) cases, which tend to have a more homogenous HER2 expression. This is clinically relevant to Trastuzumab response, as the unequivocal cases respond better to Trastuzumab therapy compared to the borderline cases, indicating challenges to overcome HER2 biomarker variations (209).

Various stem cell markers have been proposed to identify cancer stem cells in HER2 positive breast cancer patients. Breast cancer stem cells express cell surface markers *in vitro* such as high levels of CD44, ALDH1, and low levels of CD24 (210, 211). High expression of CD44 and low expression of CD24 are also associated with EMT. The expression of stem like markers is a possible mechanism of Trastuzumab resistance (212). HER2 interaction with other signalling pathways is involved in the regulation of cancer stem cells through the Wnt, PI3 kinase, and AKT signalling pathways (213). For example, in HER2 positive breast cancers, HER2 has been shown to interact with CXCR1, and the blockade of CXCR1 leads to apoptosis of CSCs via the FAK/AKT/FOXO3A axis (214, 215).

1.18 Reprogramming-associated heterogeneity in transformation

The acquisition of stem like phenotype has been associated with human neoplastic transformation. It has been shown that DU145 prostate cancer cells were activated by heregulin growth factor (HGF) through Notch signalling, which induces a molecular signature associated with stem cells. This consists of upregulation of CD49f, CD49b, SOX9, and CD44 and downregulation of CD24 (216). Furthermore, loss of the transcription factor ETS is known to determine EMT and transformation in prostate epithelial cells. The knockdown of ETS also increased several genes associated with stem-like phenotype, which include NANOG, POU5F1, STAT3, and BMI-1 (217). In U251 glioma cells, tumour-like characteristics such as migration, invasion and proliferation were enhanced by exosome induction. This was also associated with the upregulation of markers associated with "stemness" such as Nestin and CD133 (218). Moreover, Scaffidi et al have shown that fibroblasts transformed by stable ectopic expression of H-Ras-V12, h-TERT, and SV40 LT and Small ST antigens exhibited differential expression of a stem marker known as stage-specific embryonic antigen (SSEA-1) in approximately 1% of transformed cells, but which was absent in the control cells (219).

During oncogene-induced transformation, cells reprogramme from a differentiated state to a more primitive, stem-like state that has high degree of plasticity, which gives the cells the ability to self-renew and differentiate into multiple lineages. It is interesting to note that various genes implicated in normal reprogramming from stem cell stage to differentiation are also involved in transformation, such as SOX2 in breast cancer (220), and the expression of KLF4 in human gastrointestinal cancer (221). This indicates that normal reprogramming and transformation occur through similar pathways/processes.

Furthermore, cancer by and large arises due the combination of genetic aberrations and epigenetic lesions that induces growth advantage in afflicted cells (222). It has been shown that histone modifications, DNA methylation, and chromatin remodelling can have a profound

influence in cellular transformation (223). Chromatin becomes condensed as differentiation proceeds, this process is reversed by cellular reprogramming. Cellular reprogramming involves local and genome-wide changes to the chromatin architecture as cells enter into a state of plasticity during reprograming. The chromatin of embryonic stem cells (ESCs) is open and accessible, which is reflected in the elevated activity of transcriptional programme as it is associated with enrichment of active histone marks such as H3K14ac, H3K9ac, H3K36me, H3K4me3, and H3K36me2 (224). Pioneering work by Yamanaka showed that differentiated cells can be reprogrammed back to more primitive or 'induced' pluripotent cells (iPS) by the addition of four transcription factors; SOX2, KLF4, OCT4, and c-Myc (225). During the transition to iPS, transcription factor mediated chromatin activation and associated transcriptional dynamics occur rapidly and early as is shown by the increase levels of euchromatin mark, H3K4me2 (226). Transformation gives rise to distinct cell types establishing subclones with heterogeneous genetic profile that has an epigenetic hierarchy, which may include aberrant chromatin state and DNA methylation changes (227). Cellular reprogramming involves the acquisition of epigenetic changes similar to those observed in cellular transformation such as promoter-specific DNA hyper-methylation and the inactivation of DNA methyltransferase enzymes (227).

1.19 Aims and Objectives:

Cancer cells display profound rearrangements of the signalling and epigenetic landscape but how such changes unfold is not fully understood. A limited number of studies have focused on the very early transformational events in transition from normal to cancer cells but rarely in the context of the chromatin. More specifically, how re-wiring of the signalling events can impact the epigenetic landscape, which can pave the way to fully transformed cells is not yet elucidated. To understand this, we used a relatively simple experimental *in vitro* system to characterise the events that enable emergence of transformed cells. The aims of the project were to:

- Establish and characterise the HER2 inducible transformation in breast epithelial cells (MCF10A cell line).
- Investigate the dynamics of global early signalling changes upon HER2 over expression.
- Assess the genome-wide chromatin accessibility alterations in HER2 induced transformation.
- Investigate HER2 induced reprogramming-associated heterogeneity.

Chapter 2

2. Materials and Methods

2.1 Monolayer cell culture

MCF10A cells were examined using a light microscope at 4X or 10X magnifications, and were passaged before they could reach 70% confluency. These cells were plated in either a 6-well plate, T25 cm^2, or a T75 cm^2 flask depending on the experimental setting. To split the cells, medium was aspirated, and cells were washed using phosphate-buffered saline (PBS) (1X) (GIBCO #14190-094). Cells were then incubated for 15 minutes with Trypsin-ethylenediaminetetraacetic acid (EDTA) (GIBCO #R-001-100) at 37 °C. Flask or plates were then gently tapped to detach adhering cells attached to the plastic, and trypsin was immediately inactivated using full growth medium. Cell suspension was gently pipetted upon and down to create single cell suspension and remove any formed clumps and directly added to a 15 mL falcon tubes. Cells were then centrifuged at 1200 RMP for 3 minutes at room temperature. The supernatants were discarded and cells were resuspended in fresh growth medium. Cells were then seeded into an appropriate new flask depending on the experimental requirements. The flask/plate was gently swirled in a figure of 8 to distribute the cell content evenly in the plate. MCF10A cell medium consists of Dulbecco's Modified Eagle's Medium (DMEM/F12) (SIGMA #D8347) supplemented with 5% Horse Serum (SIGMA #H1138), 0.5 µg/mL Hydrocortisone (SIGMA #H0888), 20 ng/mL Epidermal Growth Factor (EGF) (SIGMA #E4127), 100 ng/mL Cholera Toxin (SIGMA #C8052), 10 µg/mL Insulin (SIGMA #i9278) and 1X Pen/Strep. To induce the overexpression of HER2, 1 µg/mL of Doxycycline (SIGMA #DN891) was added to the media.

HEK293T cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) (SIGMA#D5796) in 10% foetal bovine serum (FBS) with 1X Pen/Strep. Cells were cultured in appropriate sized

sterile cell culture flask/plate depending on the experimental requirements. To detach cells from the flask/plate, cells were flashed with full growth media to remove any adherent cells attached to the plastic.

2.1.1 Freezing

Cells were cultured as above (see section 2.1) and centrifuged to obtain a cell pellet. The cell pellet was resuspended in full fresh growth medium containing 10% dimethyl sulfoxide (DMSO) (FISHER CHEMICAL #D/412/PB08), and aliquoted in 1 mL in cryovials and transferred to a Mr. Frosty freezing containers and stored in a -80 °C freezer for 24 hours. Cells were then transferred to liquid nitrogen for long term storage.

2.1.2 Thawing

Cells were retrieved from the liquid nitrogen in dry ice to prevent defrosting prematurely. The cryovial containing cells were placed in the 37 °C water bath for approximately 2-3 minutes. Cells were immediately transferred to a 15ml falcon tube containing 5ml full growth medium and resuspended. Cells were centrifuged at 1200 RPM for 3 minutes at room temperature. The supernatants were removed and cells were resuspended in 1ml of full growth medium, and plated in an appropriate cell culture dish.

2.1.3 Cell counting

To count a specific number of cells, a haemocytometer (BRIGHT LINE #520188) or automated cell counting device such as the Luna cell counter (LOGOBIO #L20001). For counting with haemocytometer 10 µL cells were resuspended in 90 µL of 0.4% trypan blue solution (GIBCO #15250-061) in a 96-well plate and added to the counting chamber to be counted under a light microscope (LIFE TECHNOLOGIES, EVOS XL CORE) using 10X magnification. For the Luna cell counter, 10 µL of cells were mixed with 10 µL of trypan blue and added to the counting slide

(LUNA #10182907) and inserted into the instrument, to quantify cell number. Readings were generally taken twice and averaged to get the correct number of cells.

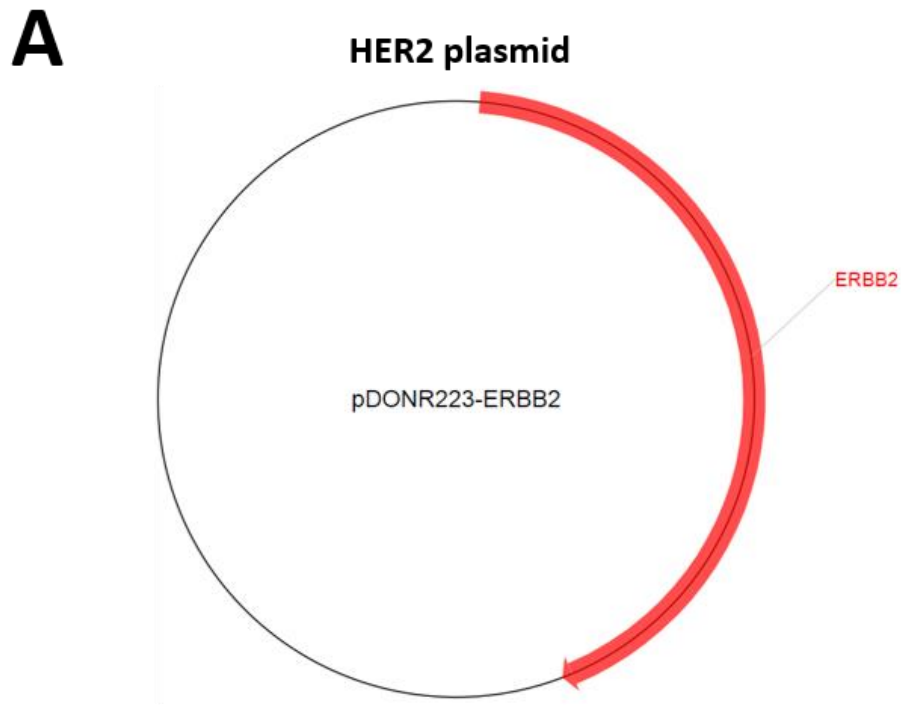
2.1.4 3D cell culture

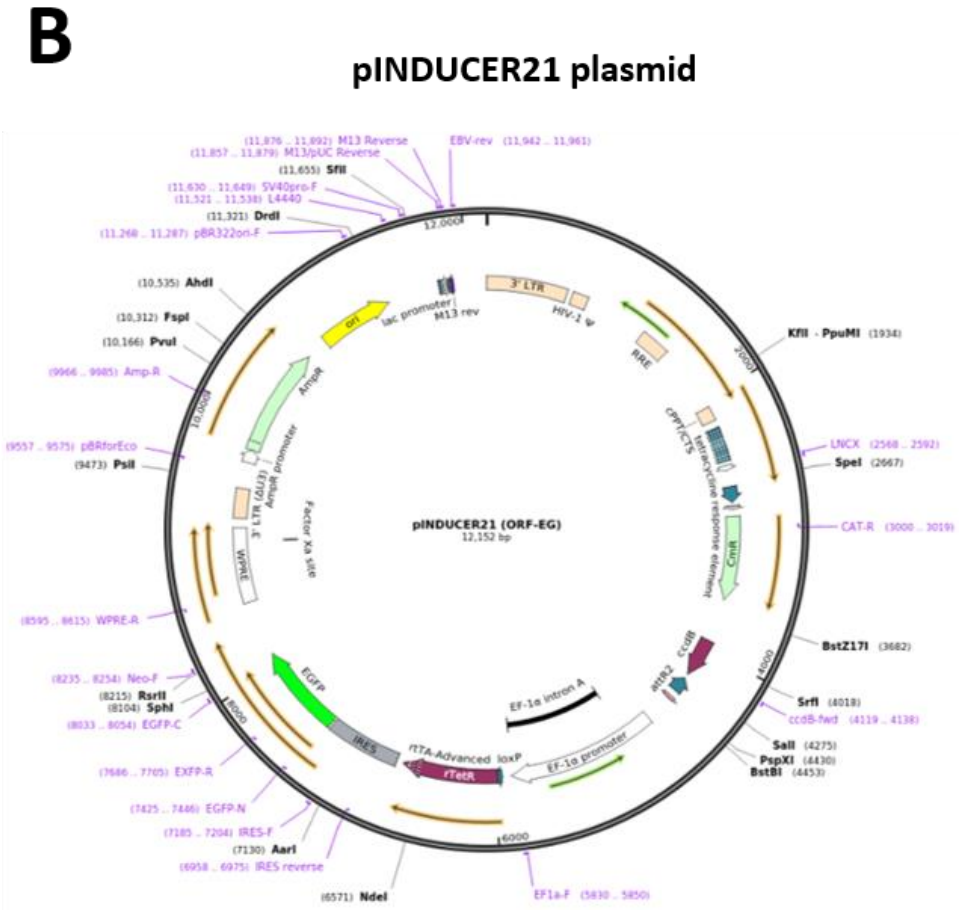
Matrigel (CORNING #356230) was thawed on ice at 4 °C overnight and aliquoted into 2 mL eppendorf tubes followed by freezing at -20 °C until required. Pipette tips were kept in -20 °C for 30 minutes and then used to prevent matrigel from solidifying whilst pipetting. To neutralise the acidic pH of collagen type I (CORNING #11563550) , it was required to add 62.5 µL of 10X PBS (THERMOFISHER # 70011044) and 62.5 µL of 0.1M NaOH (SIGMA ALDRICH #43617) to 500 µL of collagen. This helped neutralise the pH and allow cells to proliferate in the medium. To prepare one 8-well chamber slide 300 µL of matrigel and 200 µL of collagen mixture is required. They are both mixed whilst on ice to prevent it from solidifying and is pipetted up and down gently until a homogenous mixture is formed. 42 µL of this mixture is added to the centre of the well and a 10 µL pipette tip is used to spread the mixture evenly to create a layer of the mixture covering the entire well, without overspreading to the edges. The chamber is then placed in the incubator for 30-45 minutes to solidify. Meanwhile, cells are trypsinised and counted using a haemocytometer (see section 2.1.3). Cell mixture is resuspended thoroughly to avoid cell clamping and to make a single cell suspension before plating on to the wells. A cell suspension of 10,000 cells per mL was made and in each well 400 µL cell suspension containing 4000 cells was plated. 2% of matrigel mixture is added to the 400 µL cell suspension and carefully added to the wells by pipetting evenly into the well. Cells are re-fed with the 2% matrigel containing medium the next day and then medium is replaced as normal thereafter, until experimental endpoint.

2.2 Lentiviral transduction and generation of HER2 inducible cell line

HEK293T cells were used for the production of lentiviral particles due to their high transfectibilty. HEK293T cells were plated in a 6-well plate in full growth media until they were approximately 90% confluent the next day. For transfections, jetPRIME transfection reagent (POLYPLUS #114-15). The following plasmids were prepared: 5 µg of the HER2 (ADDGENE #23888) plasmid, which was sub-cloned into pINDUCER21 (ADDGENE #46948) plasmid as described in (116), 1.75µg pMD2.G (ADDGENE #12259) [envelope plasmid], and 3.25µg of pCMV delta R8.2 (ADDGENE #12263) [packaging plasmid]. The appropriate amount of jetPRIME buffer was added to the plasmid DNA and a ratio of 1:2 of DNA to jetPRIME reagent was used, briefly vortexed and incubated for 10 minutes at room temperature. The transfection mix was then added to the cells in a 6-well plate and incubated for 24 hours at 37 °C. The next day, lentiviral particles were harvested from HEK293T cells by collecting the media from cells and transferring it to a 15 mL falcon tube and centrifuging it for 1 hour, at 1500 RPM at 4 °C. The freshly produced lentiviral particles were added to the MCF10A cells to infect them, which were approximately 30% confluent, for an additional 48 hours.

2.2.1 Plasmid maps





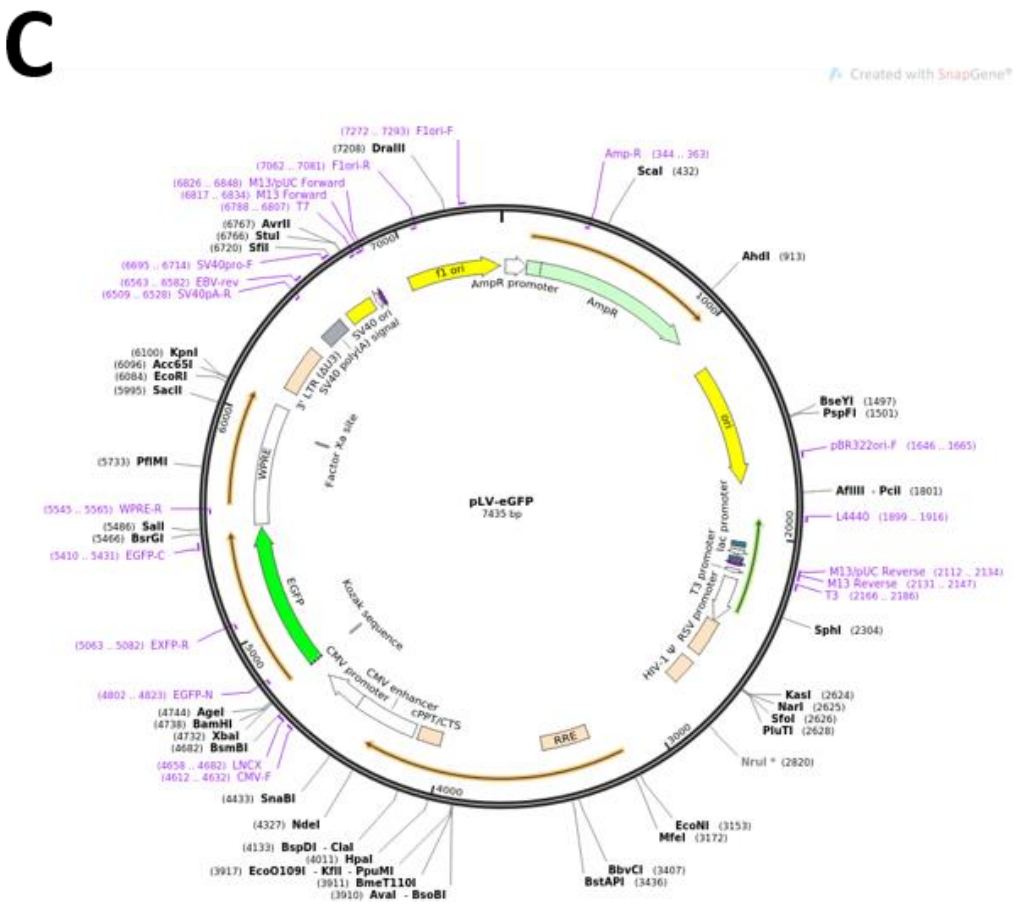


Figure 2.1. Plasmid construction. (A) The HER2 plasmid containing WT human HER2 sequence was cloned into the (B) inducible vector (pINDUCER21) plasmid to construct an inducible HER2 plasmid. The inducible plasmid map shows GFP gene which was used to select for HER2 positive cells. (C) plasmid map for the non-inducible GFP vector.

2.3 Preparation of protein lysates

Cells were seeded in a 6-well plate in full growth medium before protein analysis. For identification of phosphoproteins cells were seeded in serum starved medium (without horse serum and EGF) overnight. The following day, medium was removed and cells were stimulated with full medium for a desired time point depending on the experimental requirement. For the

analysis of total proteins, medium was removed and cells were washed with PBS. Cell lysis buffer, 200µl of either NP40 (ABCAM #142227) or RIPA buffer (THERMOFISHER #89900) containing complete cocktail of the protease inhibitors (SIGMA #P38340) or phosphatase inhibitors (SIGMA #P8340) when probing for phosphorylated proteins were added onto the cells. Cells were scraped off using a cell scraper and added to a 1 mL labelled eppendorf tubes on ice, for 30 minutes, with occasional vortexing for 5 seconds every 10 minutes in between. The suspension was then centrifuged at 4 °C, for 10 minutes at 10,000 *g*. The supernatant containing the protein was transferred to newly labelled eppendorf tubes and kept on ice until BCA assay (THERMOFISHER #23225).

2.4 BCA (Bicinchoninic Acid) Protein Assay

A BCA assay was used to determine the protein concentration of samples. In a 96-well plate, 200 µL of BCA reagent A and BCA reagent B (ratio 50:1) was added followed by the BCA protein standards and each sample in duplicates. The absorbance of each well was quantified using a plate reader (DYNEX TECHNOLOGIES OPSYS MR #CG34328) at an excitation of 562 nm. The plate reader automatically generates a standard curve and an equation using linear regression, whilst also giving us the concentrations of proteins in µg/µL. The required, but equal amount of protein (in concentration and in volume – equalised by adding some lysis buffer) was added to new 1 mL eppendorf tubes with the sample buffer (INVTROGEN #2020067) to a final concentration 1X and proteins denatured by placing the samples in a heat block (EPPENDORF THERMOSTAT PLUS) at 95 °C for 5 minutes.

2.5 SDS-PAGE

Proteins samples were resolved using 4-10% Bis-Tris mini gels (THERMOFISHER #NP0301). The tank was filled with 1X MOPS running buffer (THERMOFISHER # NP000102), and equal amounts of protein were loaded in wells ,alongside a colour pre-stained protein ladder (NEW ENGLAND

BIO-LABS #P7712). The gel was run at 70V for the first 20 minutes and then at 150V for another 50-60 minutes making sure the proteins have resolved to a sufficient degree, or until the blue down has reached the bottom of the running tank.

2.6 Protein transfer and antibody incubation

After successful running of the gel, proteins were transferred to a PVDF membrane (IMMOBILON #IPVH15150) by either wet transfer or semi-dry transfer (TRANSFER STACKS #AB401002) using i-Blot. For the wet transfer, 1 litre of 1X transfer buffer (THERMOFISHER #NP0006) containing 20% methanol and milli-Q water. The gel was removed from the tank and a transfer 'sandwich' was made. This was done by placing a sponge, followed by filter paper, and the gel. The PVDF membrane was activated by placing it in methanol for 1 minute and placed on top of the gel. This was followed by placing another filter paper and a sponge, the cassette was closed and placed in the chamber, in the transfer tank. A cold pack was placed in the side of the tank and the tank was filled to the top with the transfer buffer. The tank was placed at 4 °C overnight and 20V was applied to allow the negatively charged proteins to transfer to the membrane.

For the semi-dry transfer, gels were carefully removed from the plastic cassette and placed onto the "bottom" transfer stack so that the gel is facing the PVDF membrane. A filtered paper was paced on the back of the gel following by placing the "top" transfer stack. For smaller proteins ranging from 20-50 kDa, the transfer time was set up to 6 minutes, for larger proteins (approximately 180 kDa), transfer time was increased to 11 minutes.

This was followed by incubation of the PVDF membrane in 5% semi-skimmed milk (SIGMA ALDRICH 70166) for blocking to avoid non-specific antibody binding, for 1 hour at room temperature. Membranes were cut to size and appropriate antibody (see table 2.1) was added in 5 mL of BSA solution and incubated overnight at 4 °C, with gentle rocking or rolling in a 50 mL

falcon tube. Membranes were washed 3 times, 15 minutes each in 0.5% PBS-Tween (SIGMA ALDRICH P1379) followed by the incubation of the appropriate, species-specific secondary (see table 2.1) antibody diluted in 5 mL of BSA solution for 1 hour, at room temperature, gently rolling in a 50 mL falcon tube.

2.7 Detection of proteins

A 1:1 mixture of SuperSignal™ West Pico PLUS Chemiluminescent Substrate (THERMOFISHER #34580) was added to a 15 mL falcon tube and briefly vortexed. An appropriate volume (usually 1 mL) of the mixture was added to the membrane making sure that the entire membrane is covered and incubated at room temperature for 3-5 minutes. ECL was removed and the membrane was placed in a clear plastic film and exposed using Chemidoc (AMERSHAM IMAGER 600 #56930330) for an appropriate length of time.

Protein	**Antibody**	**Source**	**Dilution**
HER2	HER2/ERBB2 Rabbit mAb	CELLSIGNALLING #2165	1:5000
pAKT	Phospho-AKT (Ser473)	CELLSIGNALLING #9271	1:5000
tAKT	total AKT	CELLSIGNALLING #9272	1:5000
GAPDH	GAPDH	CELLSIGNALLING #2118	1:2500
Alpha-Tubulin	Anti-alpha tubulin antibody (DM1A)	ABCAM #7291	1:5000
p53	p53 (7F5) Rabbit mAb #2527	CELLSIGNALLING #2527	1:1000
Anti-rabbit secondary	GE HEALTH CARE LIFE SCIENCES	Amersham ECL Rabbit IgG, HRP-linked whole Ab #NA934	1:5000

p21	p21 Waf1/Cip1 (12D1) Rabbit mAb	CELLSIGNALLING #2947	1:1000
p27	p27 Kip1 (D69C12) XP® Rabbit mAb	CELLSIGNALLING #3836	1:1000
ACTIN (dye)	Rhodamine Phalloidin	THERMOFISHER # R415	1:200 to 1:500

Table 2.1: Antibody list with name, dilution and source.

2.8 Soft agar colony formation assay

The ultra-pure culture grade agarose (THERMOFISHER #16500500) were first diluted down to 1% in PBS and placed in a microwave to melt the agarose and then autoclaved. Soft agar assays were performed in either 12 well tissue culture plates or 24 well plates. Firstly, 0.8% of ultra-pure agarose layer (mixed with an appropriate medium) was made at the base of the wells and allowed to settle for 30 minutes at room temperature. Secondly, 10,000 cells for 12-well plates or 5000 cells for 24-well plates were mixed with 0.3% agarose and plated evenly, drop-wise, on top of the base layer and incubated for 21 days, with medium changed every 2 days. This was performed with three technical triplicates. After 21 days, medium was aspirated and cells washed with PBS. Colonies were fixed using 4% formaldehyde (PFA) at room temperature for 20-30 minutes. PFA was removed and colonies were washed with PBS and permeabalised by adding 100% methanol for 2 minutes at room temperature. Methanol was removed and colonies were washed by PBS. Colonies were stained by adding 0.05% of crystal violet dye diluted in PBS for 1 hour at room temperature. Crystal violet was removed and added to a 15 mL falcon tube to be used again. Colonies were washed with PBS, 3 times to make sure no dye remains. Images were taken of nearly the entire well using a dissecting microscope. Images were then quantified using imageJ.

2.9 Immunofluorescence

Autoclaved glass coverslips were placed in a 12-well tissue culture plates and appropriate number of cells seeded on the coverslips one day before immunofluorescence assay. The following day, media was removed and cells were washed with PBS 3 times, and an immuno-pen was used to draw a barrier around the glass cover to prevent spill over of buffers and antibodies. Cells were fixed by 4% PFA at room temperature, for 15 minutes and then washed in PBS 3 times. Cells were blocked in blocking solution (2% FBS/PBS) for 1 hour at room temperature. The blocking solution was removed, and appropriate antibodies were added onto the cells for 1 hour at room temperature. The antibodies were removed, and coverslips washed by PBS 3 times, 5 minutes each. The appropriate secondary antibodies were added to the cells for 1 hour at room temperature in the dark. Cells were washed 3 times in PBS 5 minutes each. A drop of mounting media either Glass anti-fade reagent (INVITROGEN #B36982) was added to the coverslips and were inverted into the glass sides and allowed to settle in dark for 30 minutes at room temperature. Excess mounting media was removed using tissue and a nail varnish was used to draw around coverslips to make sure they stay unmoved. Cells were imaged using the fluorescence microscope.

2.9 Immunofluorescence of acini in 3D cell culture

Media was aspirated from each well of the chamber and wells are washed with PBS carefully not to detach the layer of matrigel from the wells. Acini were fixed with 4% PFA for 30 minutes at room temperature. PFA was removed and acini washed with PBS 1 time. Acini were permeabalised with 0.5% Triton-X for 10 minutes at room temperature. Acini are then blocked in 10% goat serum in PBS-Tween, for 1 hour at room temperature. Acini are stained with Phalloidin dye over night at 4°C. Phalloidin dye was removed and acini washed with PBS 3

times, 10 minutes each at room temperature. At this point, the detachable chambers are removed and acini mounted in mounting media reagent and allowed to dry in the dark at room temperature for 4 hours. Once dried, slides are visualised using a confocal or a fluorescence microscope.

2.10 Transwell migration/invasion assay

Matrigel or collagen was diluted 1:5 with chilled growth factor reduced medium and pipetted up and down slowly to generate a homogenous mixture. 90µl of chilled diluted matrigel or collagen mixture was directly pipetted on the centre of an 8 µm pore size transwell inserts (MILLICELL #MCEP12H48) that was placed onto a 12-well plate. No matrix was placed onto the transwell insert if migration was measured. The 12-well plate was placed into an incubator for 30 minutes to allow the matrix and collagen to solidify. Meanwhile, 500µl of full medium containing growth factors (chemoattractant) was added to the wells in the 12-well plate. Cells were detached by trypsinisation and 150,000 cells were added in 200µl reduced growth factor medium, which were pipetted onto the transwell insert either coated with a matrix or the uncoated inserts. Plates were placed in the incubator for 16 hours. Highly invasive cells had invaded towards the chemoattractant, which were then stained with 0.05% of crystal violet dye. Images of random regions are taken using a standard light microscope and quantified using imageJ.

2.11 Sample preparation for flow cytometry and flow sorting

Cells were trypsinised and 500,000 cells were added to 1 mL of 2% horse serum/PBS in a polystyrene round bottomed tubes. Cells were centrifuged 5 minutes, 1200 RPM, at room temperature. Whilst cells were centrifuging, the lights in the cell culture hood were turned off and the antibody master mix was prepared in 1.5 mL eppendorf tubes. Cells were retained from the centrifuge and supernatants discarded. Antibodies were added to the polystyrene

tubes containing cells and thoroughly resuspended. Cells with antibodies were incubated for 20 minutes at room temperature, covered with kitchen foil in the dark. After 20 minutes, cells were resuspended in 2% horse serum/PBS and centrifuged for 5 minutes, 1200 RPM, at room temperature. Whilst centrifuging, DAPI suspension was made in 2% horse serum/PBS in 1:2000 dilution. Cells were retained from the centrifuge, and supernatants were discarded. Cells were washed again in 1 mL 2% horse serum/PBS by centrifugation for 5 minutes. DAPI suspension was added to the cells or just the staining buffer for unstained controls.

Fluorescence minus-one-controls (FMOs) were made for appropriate interpretation of the flow cytometry data, to make sure that the gating is based on the context of data spread in a panel with multiple fluorochromes. To do this, the FMO control contains all the antibodies except one in the designed panel with the same dilutions as shown in table 2.2.

For compensation, AbC™ Total Antibody Compensation Bead Kit (INVITOGEN #A10513) was used. The total compensation capture beads (component A) and negative beads (component B) were vortexed for 10 seconds before use. Flow cytometry tubes were labelled with the respective antibody name and 1 drop of component A was added to each tube. Pre-titrated amount of each antibody was directly added to the bead suspension and mixed well and incubated for 15 minutes at room temperature, protected from light. The beads/antibody mixture were washed by adding 3 mL of PBS by centrifugation at 250 x *g*, for 5 minutes. Supernatants were removed and the bead pellet was resuspended by adding 500 µL of PBS to the tubes. 1 drop of component B was added to the tubes and mixed well. The samples and bead pellets were kept on ice, protected from light and proceeded to flow cytometry analysis.

For flow sorting, same protocol as above was employed. Additionally, the required number of 15 mL falcon tubes or polystyrene tubes containing the appropriate medium was taken to obtain the sorted cells for further propagation in cell culture.

Protein	Antibody	Source	Dilution
HER2	BV650 Mouse Anti-Human Her2/Neu Clone NEU 24.7 (RUO)	BD Biosciences	1:100
EpCAM	APC Mouse Anti-Human EpCAM Clone EBA-1 (RUO (GMP)	BD Biosciences	2:100
MUC1	BV786 Mouse Anti-Human MUC1 (CD227) Clone HMPV (RUO)	BD Biosciences	2:100
CD44	PE Mouse Anti-Human CD44 Clone 515 (RUO)	BD Biosciences	1:100
CD24	Brilliant Violet 711™ anti-human CD24 Antibody	BD Biosciences	2:100
CD49F	BV650 Rat Anti-Human CD49f Clone GoH3 (RUO)	BD Biosciences	1:100

Table 2.2: List of antibodies used for flow cytometry or flow sorting with source and dilutions.

2.12 qRT PCR

Cells were grown and passaged in 6-well plates as previously described and resuspended in 1 mL of medium. The cell suspension was transferred to a 1.5 mL eppendorf tubes and centrifuged at 250 *g* for 5 minutes to obtain a cell pellet. The supernatants were discarded and cells were lysed in an appropriate volume (for 1 million cells use 300 µL) of TRI reagent (ZYMO (R2050-1-200) (kept in 4 °C) by pipetting up and down thoroughly. An equal volume (to the TRI reagent) of 100% ethanol and was added and mixed. RNA extraction was performed using Zymo kit (#R2050). The mixture was transferred to a Zymo-Spin Column placed in a collection tube and centrifuged at 10,000 *g* for 30 seconds. The column was transferred to a new

collection tube and the flow through was discarded. Next, 400 µL of Direct-zol RNA PreWash was added to the column and centrifuged at 10,000 *g* for 30 seconds. Flow through was discarded and this step was repeated again. 700 µL of RNA wash Buffer was added to the column and 10,000 *g* for 2 minutes. The column was carefully transferred into a labelled RNase-free tube. RNA was eluted by adding 50 µL of DNA/RNase-free water directly onto the column matrix and centrifuged at 10,000 *g* for 30 seconds. The extracted RNA was then subjected to DNase treatment using DNA-free kit (INVITROGEN #AM1906). This reaction was performed in 10 µL. Firstly, 0.1 volume (e.g. 1 µL in a 10 µL reaction) 10X DNase I buffer and 1 µL rDNase I was added to the RNA and gently mixed. This was incubated at 37 °C for 20-30 minutes. Then the resuspended DNase Inactivation Reagent (0.1 volume) was added to the mixture and mixed well by pipetting up and down. Tubes were incubated at room temperature for 2 minutes. Samples were centrifuged at 10,000 *g* for 90 seconds and RNA transferred to a clean 1.5 mL labelled eppendorf tubes. RNA was diluted in RNase free water to a concentration of 200 ng/µl.

RNA was reverse transcribed into a cDNA using the high capacity cDNA reverse transcription kit (APPLIED BIOSYSTEMS #4368814).The master mix consisted of the following components:

Component	**Volume/Reaction**
10 RT Buffer	2 µL
25X dNTP Mix (100 mM)	0.8 µL
10X RT Random Primers	2 µL
MultiScribe Reverse Transcriptase	1 µL
Nuclease-free water	4.2 µL
Total per reaction	**10 µL**

Table 2.3: Reagents and volume of the cDNA master mix to convert RNA to cDNA.

The master mix was placed on ice and gently mixed. Per each reaction, 10 µL of RNA was mixed added to 10 µL RT (reverse transcriptase) master mix in PCR tubes. Tubes were briefly centrifuged to spin down the contents and eliminate any existing air bubbles. Tubes were then placed into a thermal cyclers under the following conditions:

	Step 1	**Step 2**	**Step 3**	**Step 4**
Temperature	25 °C	37 °C	85 °C	4 °C
Time	10 minutes	120 minutes	5 minutes	-

Table 2.4: PCR conditions required cDNA synthesis.

After the reaction was completed, cDNA was then analysed by qRT PCR in three technical replicates using SsoAdvanced™ Universal SYBR® Green Supermix, (#1725274). 1 µL of cDNA was added per 10 µL reaction. The master mix contained 0.275 µL forward and reverse primers, 5.5 µL SYBR probe (Bio-Rad kit), and 3.95 µL nuclease-free water).

At the endpoint of qPCR Ct values are generated, which were used to analyse expression levels using the (2-ΔΔCt) [delta-delta Ct) method. 18S was used as a house keep gene. The primers used for qPCR are listen in the table below:

Gene	**Forward Primer**	**Reverse primer**
HER2	TGACACCTAGCGGAGCGA	GGGGATGTGTTTTCCCTCAA
BMP6	ACATGGTCATGAGCTTTGTGA	ACTCTTTGTGGTGTCGCTGA
BMPR2	GCCCAGGGGAGGAAGATA	TGGTGCCATATATCTGATAGTGC
LOX	GGGAATGGCACAGTTGTCA	ACTTGCTTTGTGGCCTTCAG

VEGFC	TGCCAGCAACACTACCACAG	GTGATTATTCCACATGTAATTGGTG
ILK	AACACGGAGAACGACCTCAA	CATCTCAACCACAGCAGAGC
18S	AAACGGCTACCACATCCAAG	CCTCCAATGGATCCTCGTTA

Table 2.5: List of forward and reverse primers used for RT-PCR.

2.13 ATAC-seq library preparation

300,000 cells per condition were grown in chamber wells as per the 3D cell culture overlay method described in section 2.1.4. Cells were isolated from the matrigel/collagen mixture using the cell recovery solution (Corning™ Cell Recovery Solution #354253). The cell recovery was done by removing the medium from the cells and washing cells with cold PBS. The removable chambers were detached from the slides and 2 mL of recovery solution was added. The matrix (matrigel/collagen mixture) was gently scraped using a sterile cell scraper onto an ice cold 15 mL falcon tube. The slides were rinsed again with 1mL of recovery solution onto the falcon tube to make sure all of the matrix and cells are recovered. The falcon tube is inverted a few times and placed on ice for 30 minutes until the matrix has been completely dissolved. The falcon tube is flicked with the finger tips back and forth to speed up the procedure. After about 15 minutes, cells begin to settle at the bottom of the falcon tube, indicating that the matrigel/collagen is dissolving. After 30 minutes, the matrix would have completely dissolved and cells are then pelleted to the bottom of the falcon tube by centrifugation at 200-300 *g*, for 5 minutes at 4 °C. The supernatants are discarded and cells were washed with PBS and centrifuged again for 5 minutes at 4 °C. Finally, cells were resuspended in 1 mL PBS for counting.

Cells were counted using the Luna counting device. 50,000 cells were used from each condition and time point to perform ATAC seq library preparation. In this experiment, we have used the OMNI-ATAC protocol with some optimisations (228). 50µl of cold ATAC-Resuspension

Buffer (RSB) containing 0.1% NP40, 0.1% Tween-20 AND 0.01% Digitonin was added to the cell pellet (in 1.5 mL eppendorf tube) and pipetted up and down 3 times. Cell pellet was incubated on ice for 3 minutes. The lysis was washed with 1 mL of cold ATAC-RSB containing 0.1% Tween-20 but no digitonin or NP40 and the eppendorf tube was inverted 3 times to mix. Nuclei were pelleted at 500 RCF for 10 minutes, at 4 °C. The tubes were retained and supernatants discarded using two separate pipetting steps, to be careful not to touch the almost visible cell pellet. To do this, remove 900 µL of the supernatant first with a p1000 pipette and use a p200 pipette to aspirate the remaining 100 µL supernatants. The cell pellet was then resuspended in 50 µL of the transposition mixture by pipetting up and down 6 times. The transposition mixture consisted of: 25 µL 2x TD buffer, 2.5 µL transposase (100 nM final), 16.5 µL PBS, 0.5 µL digitonin, 0.5 µL of 10% Tween-20, and 5 µL of water. The reaction was incubated at 37 °C for 30 minutes in a thermomixer with 1000 RMP mixing.

The reaction was cleaned up with a Zymo DNA Clean and Concentrator-5 kit (ZYMO #D4014). To do this, 250 µL of the DNA binding buffer was added to the DNA samples and DNA was transferred to a Zymo-Spin columns in a collection tubes. The column was centrifuged at 1000 *g*, for 30 seconds and flow through was discarded. 200 µL of DNA wash buffer was added to the columns, centrifuged for 30 seconds. This step was repeated 1 more time. Finally, DNA was eluted in 21 µL sterile water.

The ultra-pure DNA was now subjected to amplification by PCR. For amplification conditions see table below:

Lastly, the PCR samples were cleaned up using the Zymo DNA Clean and Concentrator-5 kit (ZYMO # D4014) as described above.

The DNA library profile was viewed using the automated electrophoresis tool, the Agilent TatpeStation System (serial number DEDAA01244). All the reagents were equilibrated to room temperature for 30 minutes. 1 µl of DNA sample was mixed with 1 µ1 of high sensitivity

(D1000) sample buffer (AGILENT #5067-5585) in strips (AGILENT #401428) and closed with caps (AGILENT #401425). Samples were vortexed for 60 seconds and span down and were analysed by the TapeSation.

2.14 Phosphoproteomic sample preparation

Cell medium was aspirated from cells in 6-well plates and 1 mL of ice cold PBS containing phosphatase and protease inhibitors (Add 20 µL NaF and 100 µL Na3VO4 to 10 mL of PBS) were added onto the wells whilst keeping the flask on ice. PBS was aspirated and this step was repeated again. 500 µL of lysis buffer was added to each well, cells were scraped off and transferred to a 1.5 mL eppendorf tubes. The cell suspension were sonicated at 50% intensity for 15 seconds, then rested for 10 seconds. This step was repeat two further times. Cell suspension was centrifuged at 20,000 *g* for 10 minutes, at 4 °C. Supernatant were recovered to a 1.5 mL eppendorf protein Lo-bind tube.

2.14.1 In Solution Tryptic digestion

Protein quantification was performed by BCA assay as described in 2.1.7. All samples were normalised to 250 µg concentration of total protein in a final volume of 300 µL. An appropriate volume of 1 M DTT to a final concentration of 10 mM (e.g. 3 µL in 300 µL) was added and incubated at room temperature, for 30 minutes with agitation (in the dark). Then, 415 mM iodoacetamide (IAM) was added to a final concentration of 16.6 mM (e.g. 12 uL in 300 µL). This was incubated at room temperature for 30 minutes with agitation (in the dark). Tubes are retained and 0.04 µL beads/µL of lysate containing 250 µg protein is added for the digest (e.g. 0.02 µl beads/ µg of protein). Appropriate volume of beads from stock beads container and aliquoted into a 1.5 mL Lo-bind Eppendorf tubes and centrifuged at 2,000 g for 5 min; 4 °C. HEPES buffer was added in equal volume to that of the beads (i.e. 1:1) and centrifuged at 2,000 g for 5 min at 4 °C. Supernatants was removed and replaced with fresh HEPES buffer

(1:1). The last two steps were repeated two further times. Samples were diluted 4X with HEPES buffer after the IAM incubation (e.g. 900 µL HEPES buffer to 300 µL lysate; 1200 µL total). The appropriate amount of conditioned beads (48 µL of beads for a 1200 µL digest containing 250 µg of protein) and incubated overnight at 37 °C with agitation. The next day, samples were transferred onto ice and centrifuged at 2,000 g for 5 min at 4 °C. Supernatants were transferred to a Lo-bind protein eppendorf tubes on ice. Meanwhile, vacuum manifold was setup to ~5 inHg. Samples were equilibrated at room temperature and loaded onto the vacuum manifold using the lowest flow rate possible. Samples were washed with 1 mL desalting loading buffer. Samples were retained and eluted with 0.5 mL Elution buffer A.

2.14.2 Phosphopeptide enrichment

An appropriate amount of TiO_2 beads from stock vial (50 µg beads/1 µg protein) were re-suspended in 1% TFA and vortexed. This was kept at 4 °C, when not in use. All the OASIS eluted fraction(s) volumes were adjusted to 500 µL with 1 M Glycolic acid in 80% ACN/ 5% TFA. 25 µL (i.e. 12.5 mg) of re-suspended TiO_2 beads were added to the OASIS eluted fraction(s) and vortexed. The TiO_2 beads were resuspended between samples before adding them, this is to ensure the same quantity of TiO_2 beads is added. Samples were incubated for 5 minutes with rotation/agitation. For spintips equilibration: the spintip(s) were placed in normal 2 mL eppendorfs and 200 µL 100% ACN applied to spintip(s), followed by centrifugation for 3 min at 1,500 *g* and flow through was discarded. Samples were incubated with TiO_2 for 5 minutes and then span down for 30 seconds, at 1500 *g*. The supernatants were transferred to protein Lo-bind 1.5 mL eppendorf tubes on ice. The TiO_2 beads were resuspended in the remaining 100 µL of solution and vortexed. 100 µL of re-suspended samples were applied to the empty spintip(s) and centrifuged for 2 min at 1,500 *g*. 100 µL 1 M Glycolic acid in 80% ACN/ 5% TFA was added to the sample tubes and the remaining TiO_2 beads were resuspended. Vortexed and span for 10 seconds. The remaining TiO_2 beads were applied to the spintip(s) and centrifuged for 2 min

at 1,500 *g*. Flow through was discarded. The 400 µL remaining aliquots was removed from ice to equilibrate at room temperature. The 400 µL of remaining sample was applied to the TiO_2-filled spintip(s); 2 x 200 µL batches – centrifuged for 3 min at 1,500 *g*. Flow through was discarded. 100 µL 1M Glycolic acid in 80% ACN/ 5% TFA was applied to the spintip(s). Centrifuged for 2 minutes at 1,500 *g*. Flow through was discarded to remove non-phosphorylated peptides. 100 µL 100 mM Ammonium Acetate (25% ACN) was applied to the spintip(s). Centrifuged for 2 minutes at 1,500 *g*. Flow through was discarded to remove acidic non-phosphorylated peptides. 100 µL 90/10 H_2O/CAN was applied to spintip(s). Centrifuged for 2 min at 1,500 g. Flow through was discarded. The last step was repeated twice to remove any salts and HILIC-mode bound non-phosphorylated peptides from the TiO_2 layer before the elution step. The spintip(s) were transferred to fresh 2 mL protein Lo-bind eppendorf tubes. 50 µL 5% NH4OH (10% ACN) was applied to the spintip(s) and centrifuged for 2 min at 1,500 *g*. The flow-through(s) were kept and pooled to elute phosphopeptides from the TiO_2 layer. This step was repeated 3 more times. Samples were snap-frozen and placed in speed-vac to dry overnight. Samples were then subjected to mass spectrometry analysis.

2.15 Image J quantification

Image J software was used to perform densitometry analysis on the western blots. A rectangular area around the first band was drawn using the “rectangular select” tool. Sequentially, a rectangle is drawn and selected for all of the bands of interest. Additionally, to compensate for the background noise, five random representative regions of the same size as the bands of interest were also selected. Once all the bands were selected, CTRL 3 was pressed and another image with histograms appear for each selected region. To obtain the results, we selected the “magic wand” button and clicked in each histogram. The average of five different random regions were subtracted from the band of interest to compensate for the background and results were plotted using prism as shown by bar graphs.

2.16 Bioinformatics analysis

The ATAC-seq data was provided as FASTQ files. The initial quality control checks were performed on each sample using the FastQC tool. The adapter sequences were removed with cutadapt using:

Cutadapt -a CTGTCTCTTATACACATCT -A CTGTCTCTTATACACATCT -o out.1.fastq -p out.2.fastq *infastqfile1 infastqfile2*

Samples were aligned to the human genome, Genome Reference Consortium Human Build 38 patch release 13 (GRCh38.p13), using bowtie2, and a SAM file was obtained.

bowtie2 index -1 trimmed FASTQ file -2 trimmed FASTQ file –S 1.sam

SAM files were converted to BAM files (binary files) using the following command:

Samtools view –Sb in.samfile > out.bamfile

Bam files were sorted using:

Samtools sort in.bamfile -o out.bamfile

The sorted files were then indexed using:

Samtools index in.bamfile

The ATAC-seq files can have a large number of reads (40-60%) that align to the mitochondrial DNA which should be removed using:

Samtools view –h in.bamfile | removeChrom - - chrM | Samtools view - b - > out.bamfile

PCR duplicates were removed from the files using Picard tools:

Java -jar picard.jar MarkDuplicates I=in.bamfile O= out.bamfile M=dups.txt REMOVE_DUPLICATES=true VALIDATION_STRIGENCY=LENIENT

Samples were downsampled to 25 million reads by working out the ratio (which was done by dividing 25 million by the total number of reads were in that specific file. The reason for down sampling/normalising all reads in each sample to 25 million was because there were drastically uneven number of reads (ranging from 27-55 million reads) obtained from our ATAC-seq data. To make sure we account for the differences in sequencing depth between the different samples, all samples were trimmed to 25 million reads prior to analysis to bring different samples onto a common scale. The logic behind down sampling to the lowest number of sequences produced from any sample is generally unreported, but presumably it is to compromise between data set balance and information loss. This down sampling is only performed when visualising and checking the quality of the data. The differential analysis would be performed by the raw BAM files, and programmes such as DiffBind would have an internal control to compensate for the differences in coverage.

samtools view -b -s *0.5* in.bam > out.Downsampled.bam

Peaks were called using the MACS2 tool for each bam file separately using:

MACS2 callpeak -t inbamfile -f BAMPE -n in.bamfile -g ce –keep-dup all

The two biological replicates were intersected using bedtools with the following script:

Bedtools intersect -a peakfile.1 -b peakfile.2 -f 0.50 -r > out.bedfile

To report unique entries we used:

Bedtools intersect -a bed.file -b bedfile.2 -v > 1.bed

To report overlap entries we used:

Bedtools intersect -a bed.file -b bed.file2 -u > 1.bed

To create a matrix to then generate a heatmap, we first converted the bam files to bigwif files using:

bamCoverage -b in.bam -o coverage.bw

computeMatrix reference-point –referencePoint center –S in.bigwigfile.1 bigwigfile.2 –R bedfile.1 –a 1000 –b 1000 –o matrix.1

To plot a heatmap:

plotHeatmap –m marix.1 –o Heatmap.png

To plot correlation we first produced a multiBamSummary and and multibigWigSummary using:

multiBamSummary bins --bamfiles file1.bam file2.bam -o results.npz

multiBigwigSummary bins -b file1.bw file2.bw -o results.npz

Then the correlation were plotted using:

plotCorrelation –n result.npz – corMethod Pearson –skipZeros – plotTitle “Pearson Correlation” --whatToPlot heatmap --colorMap RdYlBu--plotNumbers--o heatmappearson.png

To plot profile we used the following script:

plotProfile -m matrix.mat.gz --perGroup --kmeans 2 -plotType heatmap -outExampleProfile1.png

Number of peaks were counted using:

samtools view -c in.bam

2.17 Statistics

The appropriate statistics were performed using either GraphPad Prism 5.4 or Microsoft Excel 2013. For each experiment see accompanying figure legend.

Chapter 3

Establishment and characterisation of HER2 inducible transformation model

3.1 Introduction

To overcome significant challenges in delineating early transformational events of normal cells progressing towards cancer, we took advantage of the tetracycline (Tet-On) inducible system. HER2 over expression is observed in up to 30% of breast cancers and have been found to promote tumourigenesis. However, the early changes occurring upon HER2 over expression, particularly those regarding intracellular signalling, chromatin architecture and cell physiology need further investigation. The main advantage of this model lies in the ability to control the levels and timing of expression of the gene of interest. Furthermore, it provides us with a platform to investigate the events leading to oncogenic transformation in a reproducible experimental setting, which to date have not been fully exploited. The over expression of HER2 using Tet-On system in human mammary epithelial (MCF10A) cells represents a simple, yet versatile model for the study of early changes in the process of transformation. This system post establishment and characterisation will provide novel insights into the effects of HER2 over expression on signalling and chromatin conformation changes.

3.2 Generation of HER2-MCF10A cell line using tetracycline inducible system

To generate a stable inducible HER2 over expressing cell line we used a third generation stable lentiviral transduction system. We transiently co-transfected the human embryonic kidney epithelial (HEK) 293T cell line with pCMV delta R8.2 (packaging vector), pMD2.G (envelope vector) and the Tet-On inducible pINDUCER21 HER2 construct containing a surrogate marker, a

constitutively active GFP (green fluorescent protein) gene for tracking and selection purposes (Figure 3.1). This plasmid has been generated in a series of inducible vectors for inducible control of gene function (116). It has been previously used to investigate the relationship between myoepithelial-luminal cells in progression of breast cancer (115). In parallel, as a control for our subsequent experiments, 293T cells were co-transfected with an empty constitutively expressed pLV-eGFP vector along with the same packaging and envelope vectors. Lentiviral particles were harvested 24 hours post transfection and MCF10A cells were infected with the viral particles for an additional 48 hours.

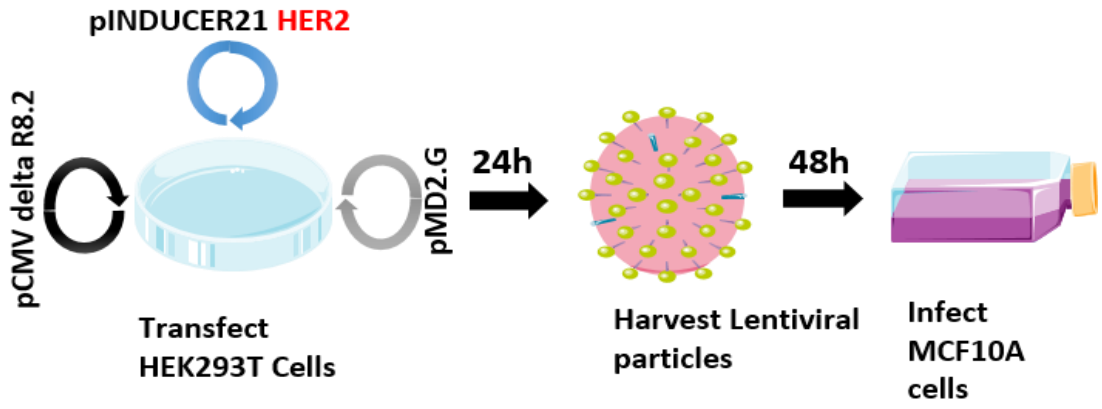


Figure 3.1: Schematic of third generation lentiviral transduction. pINDUCER21 vector containing wildtype HER2 insert was co-transfected with the packaging (pCMV delta R8.2) and envelope (pMD2.G) vectors in HEK293T cells. Viral particles were purified with centrifugation 24 hours after transient transfection by the labelled vectors. MCF10A cells were infected with the freshly produced virus for 48 hours.

In control cells, expression of GFP, which is under the control of the constitutively active cytomegalovirus immediate-early (CMV) promoter, was readily detected using the fluorescence imaging microscope in the HEK293T cells and subsequently in the infected control GFP-MCF10A cells (Figure 3.2A.). However, barely detectable levels of GFP were observed in HEK293T and MCF10A cells transduced with the inducible pINDUCER21 HER2 construct. GFP in

the pINDUCER21 HER2 construct is expressed from a weak constitutively active human elongation factor 1α (EF1α) promoter. Therefore, to adequately quantify the percentage and relative intensity of the GFP positive cells, flow cytometry was performed. It showed that GFP expression is notably lower in the pINDUCER21 HER2 transduced cells, with a 2.27% transduction efficiency. This is in comparison to GFP-MCF10A cells, which exhibited a 35.6% transduction efficiency (Figure 3.2B). The successfully transduced cells were flow sorted at high purity (approximately 90%) based on GFP expression from the non-transduced background population. The purity check is routinely performed directly after FACS sorting has been completed and we found that the GFP positive cells were purified at 90%. Subsequently, from the pINDUCER21 HER2 transduced cells, the GFP positive cells were separated into two populations for propagation with either doxycycline hyclate (dox), to induce HER2 expression (DOX +ve cells), or without dox as a parental control (DOX -ve cells).

Furthermore, to investigate if GFP efficiency remains stable over time. We passaged cells for an additional 6 times and 8 times respectively and measured the GFP expression of cells by flow cytometry. We found that that there was a decrease in GFP expression from the original 90% of pure GFP population to approximately 65% GFP expression (Figure 3.2C).

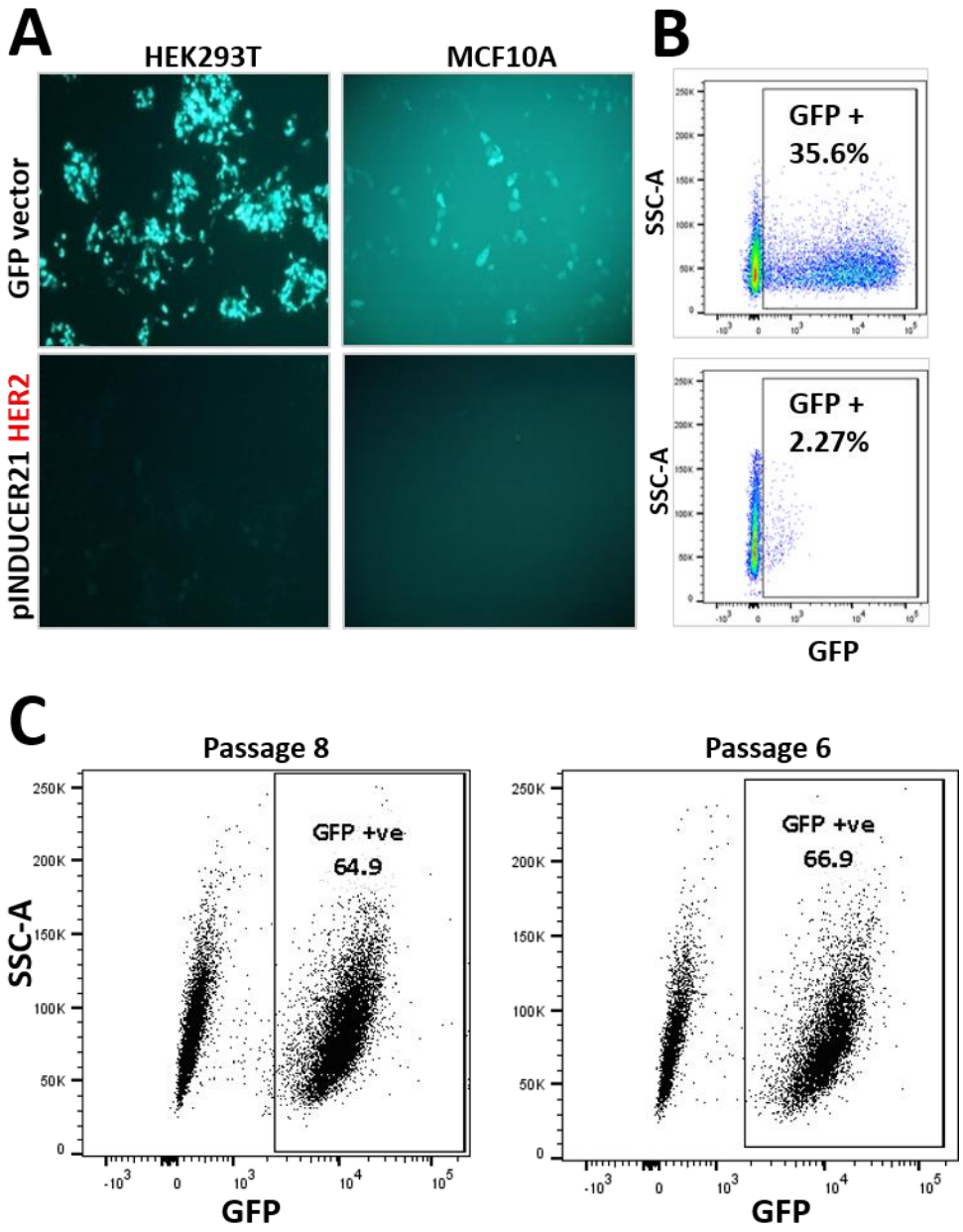


Figure 3.2: Generating a HER2 inducible MCF10A cell line using lentiviral transduction. (A) Fluorescence images of HEK293T cells transfected with pINDUCER21 HER2 and control empty GFP plasmids 24 hours post transfection. MCF10A cells were infected with the virus for 48 hours resulting in fluorescence in GFP transduced cells but not in the pINDUCER21 HER2 transduced cells (B) Scatter plots of flow cytometric analysis of MCF10A cells transduced with pINDUCER21 HER2 and control GFP expressing virus to check for transduction efficiency. (C) Flow cytometry analysis of cells cultured for 6 and 8 passages after the initial transduction and GFP expression measured by flow cytometry.

3.3 Dose and time dependent HER2 expression

To investigate inducibility of HER2 expression at the protein level, we selected five different concentrations of dox and cultured cells for 24 hours. Some of these concentrations selected here have been previously used to induce the expression of gene of interest (115, 229). HER2 over expression was readily and efficiently induced in the infected MCF10A cells upon exposure of cells to dox. Typically, HER2 expression was induced by dox in a dose-dependent manner. It also shows that protein expression is tightly regulated, as there is no "leaky" expression of HER2 in the DOX -ve cells (Figure 3.3A). We wished to examine whether this model could be used to express similar levels of HER2 protein as seen in a subset of HER2 over expressing breast cancer patients, with the aim to generate a more physiologically relevant human context (230). Using Real-Time Polymerase Chain Reaction (RT-PCR), we verified the increase in HER2 expression at gene expression level with increasing concentration of dox. In addition, we determined that there is approximately 18-fold more HER2 mRNA transcripts in the DOX +ve cells when cells are exposed to 1 μg/ml of dox relative to normal MCF10A cells (Figure 3.3B). We found that the HER2 gene expression is relatively similar to levels observed in the 2+ grade tumours in some HER2 positive breast cancer patients (231, 232)

Immunofluorescence (IF) staining was performed to confirm the cellular localisation of HER2 protein in DOX +ve cells, after cells were cultured in 1 μg/ml of dox for 24 hours. As a cell surface receptor, HER2 localised around the plasma membrane as expected, whereas in normal MCF10A cells, HER2 levels were negative or below the detection threshold of IF (Figure 3.3C). Additionally, we noted that HER2 was expressed heterogeneously, with some cells appearing brighter than others and a fraction of cells not exhibiting any fluorescence. Next, we sought to quantify the levels of HER2 in DOX +ve cells in a time-dependent manner. We selected 12 different time points, ranging from 0 hours to 17 hours, with samples collected every 1.5 hours. As shown in Figure 3.3D, HER2 expression increased in a time-dependent

manner until the 10.5-hour time point, after which it remained constant. This shows that 10.5 hours of dox induction is sufficient to cause saturation of the cell receptor (HER2).

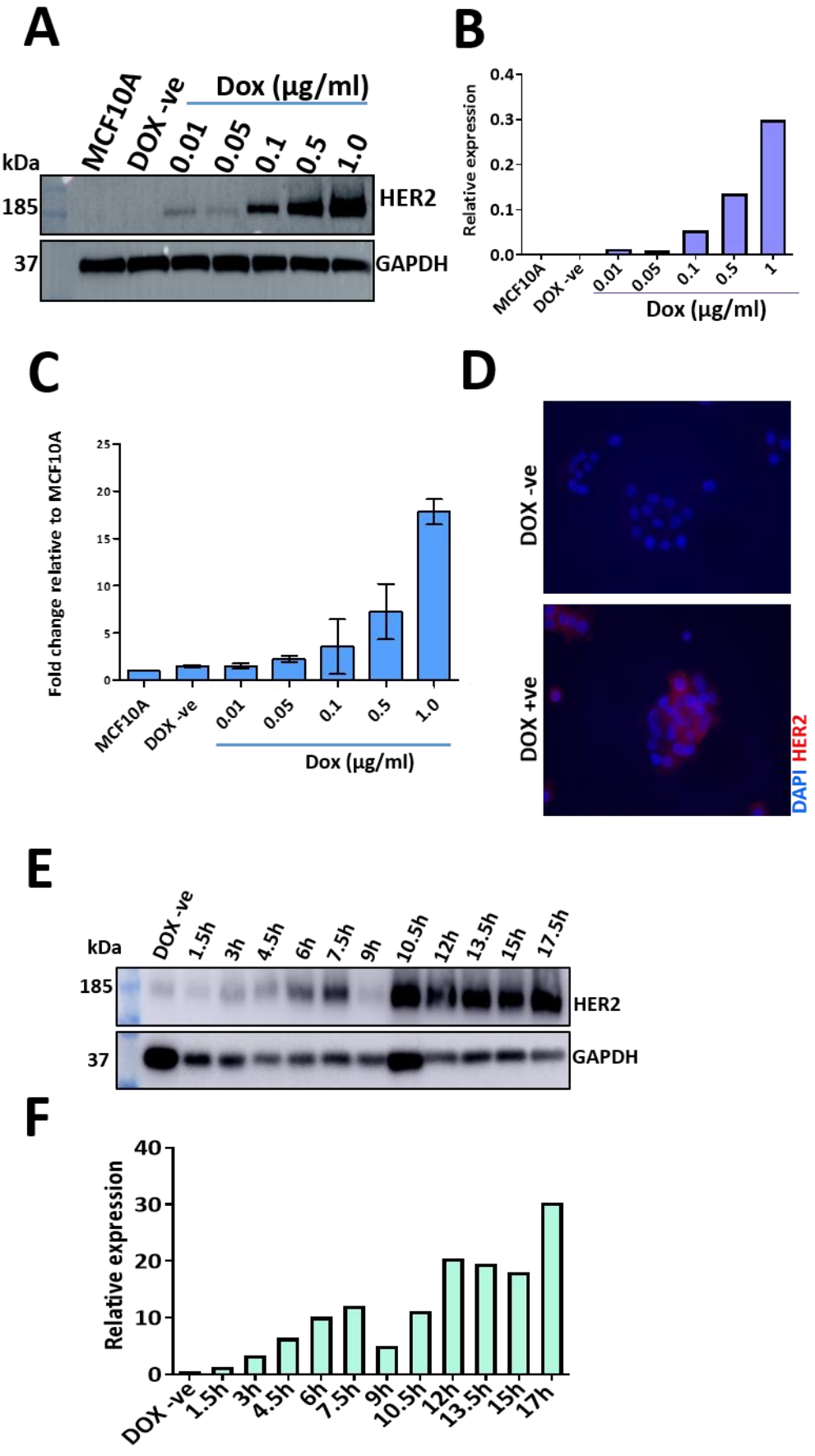

A
MCF10A
DOX -ve
Dox (µg/ml)
0.01
0.05
0.1
0.5
1.0
kDa
185
HER2
37
GAPDH
B
Relative expression
0.4
0.3
0.2
0.1
0.0
MCF10A
DOX -ve
0.01
0.05
0.1
0.5
1
Dox (µg/ml)
C
Fold change relative to MCF10A
25
20
15
10
5
0
MCF10A
DOX -ve
0.01
0.05
0.1
0.5
1.0
Dox (µg/ml)
D
DOX -ve
DOX +ve
DAPI HER2
E
kDa
DOX -ve
1.5h
3h
4.5h
6h
7.5h
9h
10.5h
12h
13.5h
15h
17.5h
185
HER2
37
GAPDH
F
Relative expression
40
30
20
10
0
DOX -ve
1.5h
3h
4.5h
6h
7.5h
9h
10.5h
12h
13.5h
15h
17h

Figure 3.3: Doxycycline induces HER2 over expression. (A and B) HER2 expression analysis by western blot and the densitometry analysis of HER2 expression normalised to GAPDH (loading control) using Image J (n=1). (C) RT-PCR in MCF10A cells infected with inducible HER2 lentiviral particles and cultured in various concentration of dox (0.01, 0.05, 1.0, 0.5, and 1.0 µg/ml) for 24 hours. 18S was used as internal control for RT-PCR (n=2). (D) Fluorescence microscopy performed 24 hours after inducing HER2 expression by dox. Cells were stained with DAPI and HER2 antibody for nuclear and protein visualisation respectively. Alexafluor 555 was used as a secondary antibody. Scale bars represent 100µm. (E and F) Western blot analysis of time-dependent HER2 expression for the indicated time points and the densitometry analysis of HER2 expression normalised to GAPDH control using Image J (n=1).

3.4 HER2 over expression induces morphological changes

A key hallmark of transformed cells is the loss of cell organisation, and proliferation, as well as cell to cell membrane contact and cell to cell adhesion to their control counterparts (56). In monolayer cell culture, un-transduced MCF10A and DOX -ve cells grew in expanding colonies with the cobblestone-like structure characteristic of epithelial cells. However, DOX +ve cells exhibited a more fibroblastic and spindle-like shape after being in 2-dimensional (2D) culture for 7 days (Figure 3.4A). We extended our analysis by studying the morphology of cells in 3-dimensional (3D) basement membrane cell culture (rBM) in overlay or “on-top” matrigel/collagen assay over a period of 9 days. The MCF10A cell morphology progression series grown in 3D rBM cultures is a powerful system to study human mammary transformation and is simple to track morphological changes compared to 2D cultures. At day 0, both DOX -ve and DOX +ve cells anchored into the matrix and formed similarly-sized spherical masses of cells termed “acini”. After day 3 and until day 9, the acini of DOX -ve cells continued to grow in size while retaining their overall spherical structure with smoother outer edges. During the same time, the DOX +ve cells became easily distinguishable, as they appeared flat and lacked even edges. They not only grew in size, but appeared to have

produced a multi-acinar conformation with branched connections. The cells exhibited a more irregular, disorganised, arm-protruding, and invasive-like structure as they became denser and darker (Figure 3.4B). Moreover, we questioned whether the morphological alterations could have been due to the addition of dox and not HER2 over expression. To answer this, we added the same amount (1 μg/ml) of dox to an empty vector (GFP) transduced MCF10A cells and DOX +ve cells. Indeed, the addition of dox had no impact on GFP-MCF10A cell morphology, as they retained circular organised acini conformation similar to DOX -ve cells, but the DOX +ve did not (Figure 3.4C). This shows that dox addition is not the cause of morphological changes but HER2 over expression.

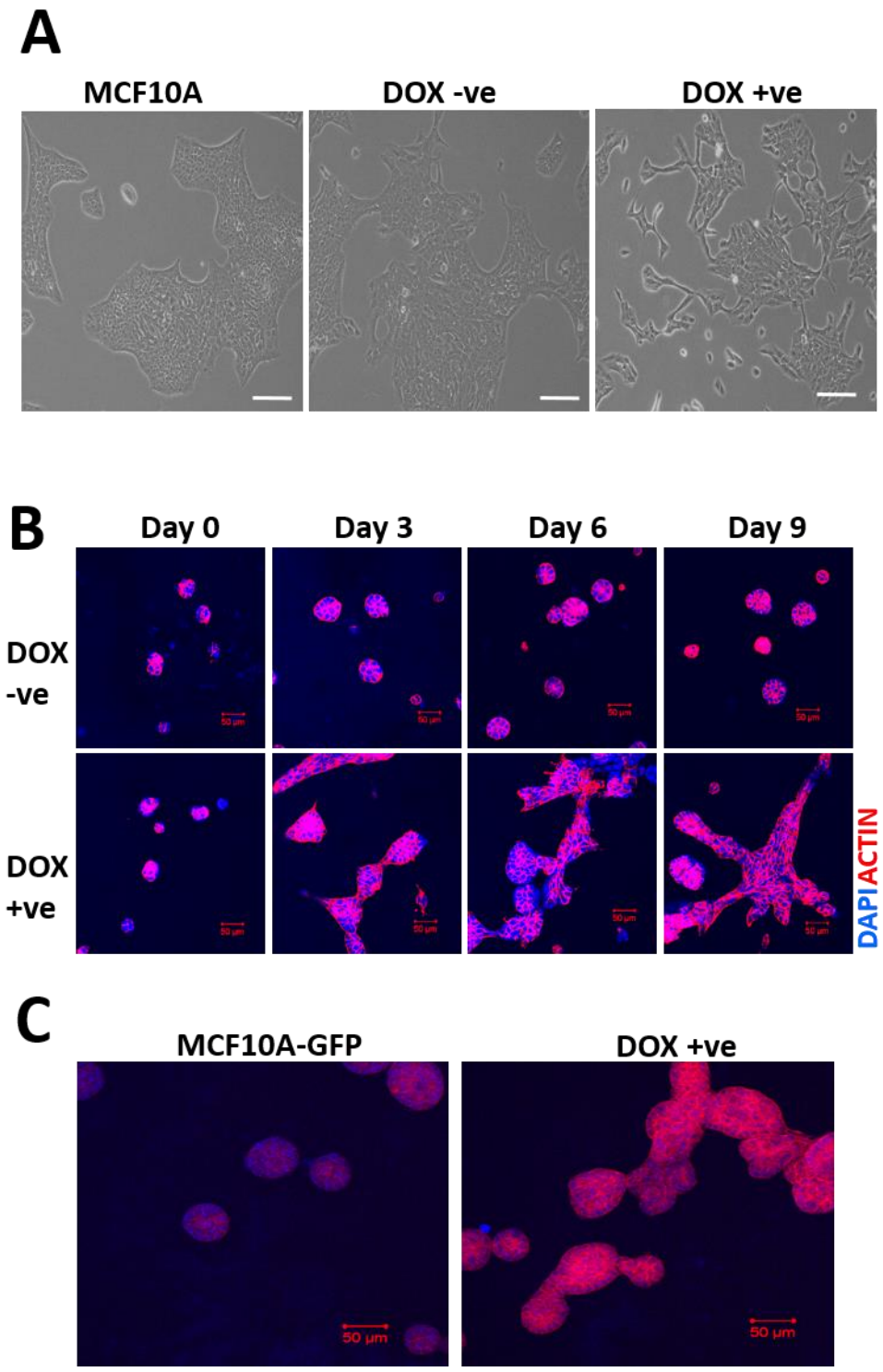


Figure 3.4: HER2 over expression disrupts normal MCF10A morphology. (A) HER2 expression was maintained for 7 days in monolayer cell culture. Bright-field images of MCF10A and DOX -ve cells show normal cobble-stone like morphology. DOX +ve cells show a more spindle-like appearance as a result of HER2 expression. Scale bars represent 250µm. (B) DOX -ve and DOX +ve cells were cultured in 3D matrix

over 9 days. DOX -ve cells formed spherical acini which increased in size over time. DOX +ve cells formed flat projecting cells of complex masses, typical of transformed cells. (C) 1µg/ml dox was added to GFP-transduced MCF10A and DOX +ve cells. Dox had no effect on MCF10A cell morphology without HER2 expression. Scale bars represent 50µm.

3.5 Effects of lapatinib and dox absence on HER2 expressing cells

Lapatinib is a dual HER1 and HER2 kinase inhibitor and thus is effective in inhibiting downstream signals through the MAPK signalling pathway (233, 234). To explore the effects of lapatinib on phosphorylated ERK (phospho-p44/42 MAPK (Erk1/2) (Thr202/Tyr204) in our system, cells were acutely treated with four different concentrations of lapatinib for 3 hours. Such short-term treatment times have been previously reported and it has been shown that even one hour of lapatinib treatment of HER2 over expressing cells may be sufficient to decrease HER2 signalling (234, 235). Indeed, lapatinib treatment decreased pERK abundance in DOX +ve cells in a dose-dependent manner (Figure 3.5A). Furthermore, to determine if HER2 activity is required to sustain the aberrant morphological phenotype of DOX +ve acini in 3D, we investigated the effect that long-term treatment with lapatinib has on the morphology of MCF10A cells. Therefore, we maintained lapatinib treatment of HER2 over-expressing acini for 5 days. Indeed, cells did not progress to the aberrant morphology, as observed in the untreated DOX +ve cells (positive control). At the same time, established aberrant acini formed after 4 days were treated with 5µM of lapatinib for 24 hours (Figure 3.5B). This confirmed that inhibiting HER2 activity after the formation of invasive morphology induced a significant reversal of the aberrant morphology of DOX +ve acini.

To test whether the induction of HER2 protein expression from the DOX +ve cells is reversible, HER2 expression was induced by the addition of dox for 24 hours, followed by the removal of dox by three consecutive PBS washes. Western blot analysis at 24, 48 and 72 hours after removal of dox showed that HER2 expression decreased over time, but did not reach the basal

levels of un-transduced MCF10A or DOX -ve cells within 72 hours (Figure 3.5C). This shows that, as expected, this system provides reversible and temporal control of protein expression, but may require more time for reversal of HER2 expression to the uninduced state. The HER2 expression remains at low levels despite dox removal and the half-time of HER2 protein is known to be 19.6 hours (236).

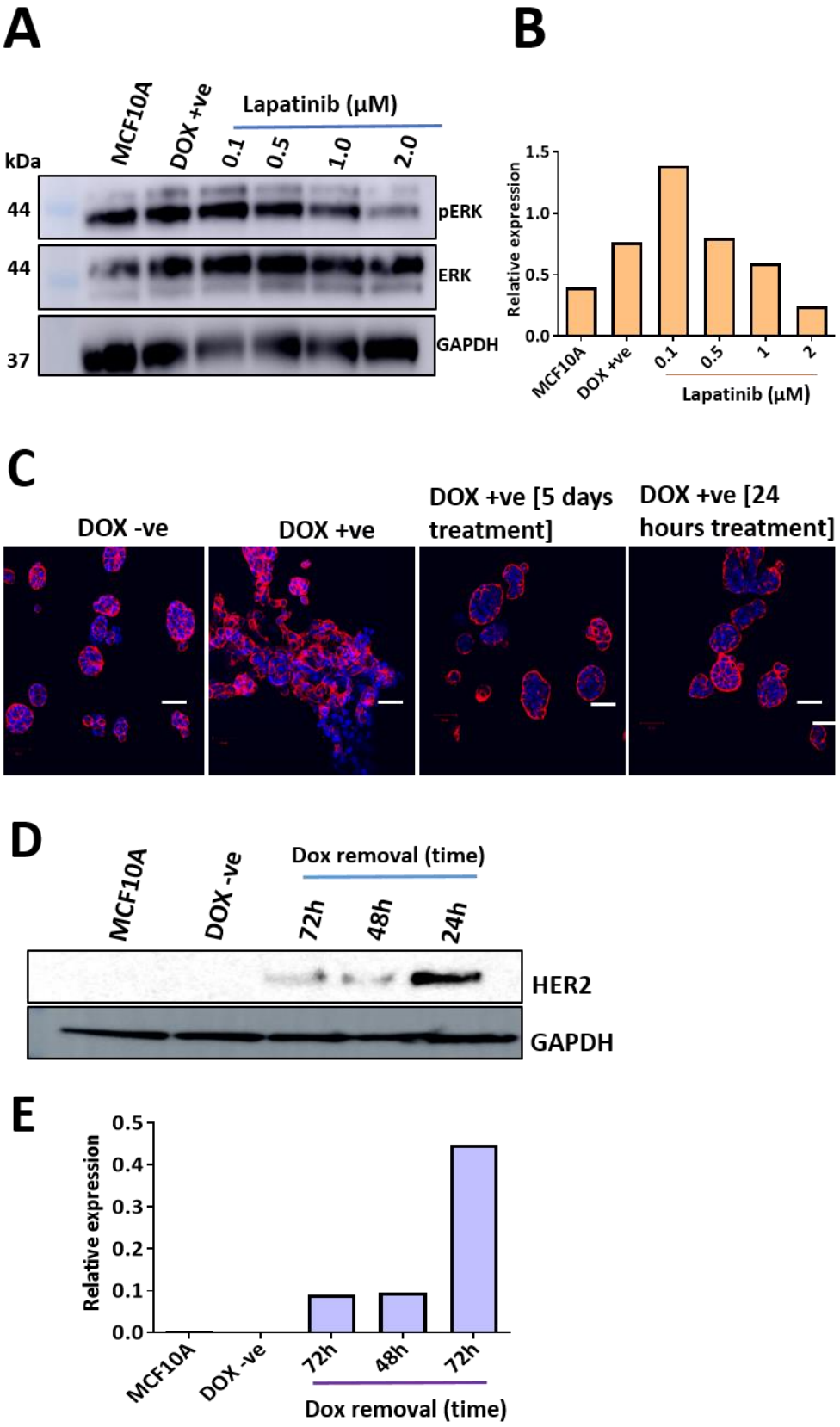

A
MCF10A
DOX +ve
Lapatinib (µM)
0.1
0.5
1.0
2.0
kDa
44
pERK
44
ERK
37
GAPDH
B
Relative expression
1.5
1.0
0.5
0.0
MCF10A
DOX +ve
0.1
0.5
1
2
Lapatinib (µM)
C
DOX -ve
DOX +ve
DOX +ve [5 days treatment]
DOX +ve [24 hours treatment]
D
MCF10A
DOX -ve
Dox removal (time)
72h
48h
24h
HER2
GAPDH
E
Relative expression
0.5
0.4
0.3
0.2
0.1
0.0
MCF10A
DOX -ve
72h
48h
72h
Dox removal (time)

Figure 3.5: Inhibition of HER2 and its signalling. (A and B) Cells were serum starved for 24 hours and treated with various concentrations (0.1, 0.5, 1.0, and 2.0 µM) of lapatinib for three hours. Cells were then stimulated with full serum media for 10 minutes before harvesting protein lysates. Western blot analysis shows dose-dependent reduction of pERK signalling and the associated densitometry analysis normalised to GAPDH control (n=1). (C) DOX +ve cells were treated with 5 µM of lapatinib for 5 days or treated on day 4 for 24 hours. DOX -ve and DOX +ve cells served as controls. Scales bars represent 50 µM (n=1). (D and E) HER2 was induced for 24 hours and cells were washed with PBS three times consecutively. Western blot analysis was performed at 24, 48, 72 hours. HER2 levels decreased over time but did not reach the basal levels observed in 72 hours as is also shown in the densitometry analysis (n=1).

3.6 HER2 induces invasion of cells and activates associated pathway

An indication of cells progressing towards transformation is their ability to invade the surrounding tissue (3, 237). To test the effect of HER2 over expression on the ability of cells to invade through matrix barriers, we performed *in vitro* transwell migration and invasion assays. We maintained HER2 over expression in DOX +ve cells for 5 days and plated 150,000 cells in parallel to DOX -ve cells on 8µm transwell filters for 16 hours in a low serum media. The transwell filters were either matrigel-coated, collagen-coated, or left uncoated to estimate migration through the transwell filters. We demonstrated that the DOX +ve cells exhibit considerably higher migration and invasion capacity towards the full serum containing media compared to the DOX -ve cells (Figure 3.6A). This shows that within 5 days of HER2 over expression, cells have acquired an invasive phenotype.

Previously, microRNAs (miRNAs) such as miR-21 has been shown to enhance invasion and metastatic potential in HER2 over expressing cells (238). To confirm the invasive phenotype observed through the transwells, we investigated and confirmed a known HER2 signalling pathway that induces the expression of miR-21 through the MAPK signalling pathway, which is known to contribute to the increased invasive phenotype observed in DOX +ve cells compared

to DOX -ve cells. We grew cells in 3D cell culture in a mixture of matrigel/collagen assay for 5 days to be comparable to the functional invasion assay performed. We determined that HER2 over expression dysregulates this pathway at gene expression level as it increased the levels of a known transcription factor (ETS-1). This upregulates the expression of primary transcript coding for miR-21 (pri-miR-21), which is further processed into miR-21. The Pri-miR21 subsequently decrease the expression of PDCD4 (programmed cell death 4) and/or other

unidentified genes allowing active cell invasion relative to control cells (Figure 3.6B) (238).

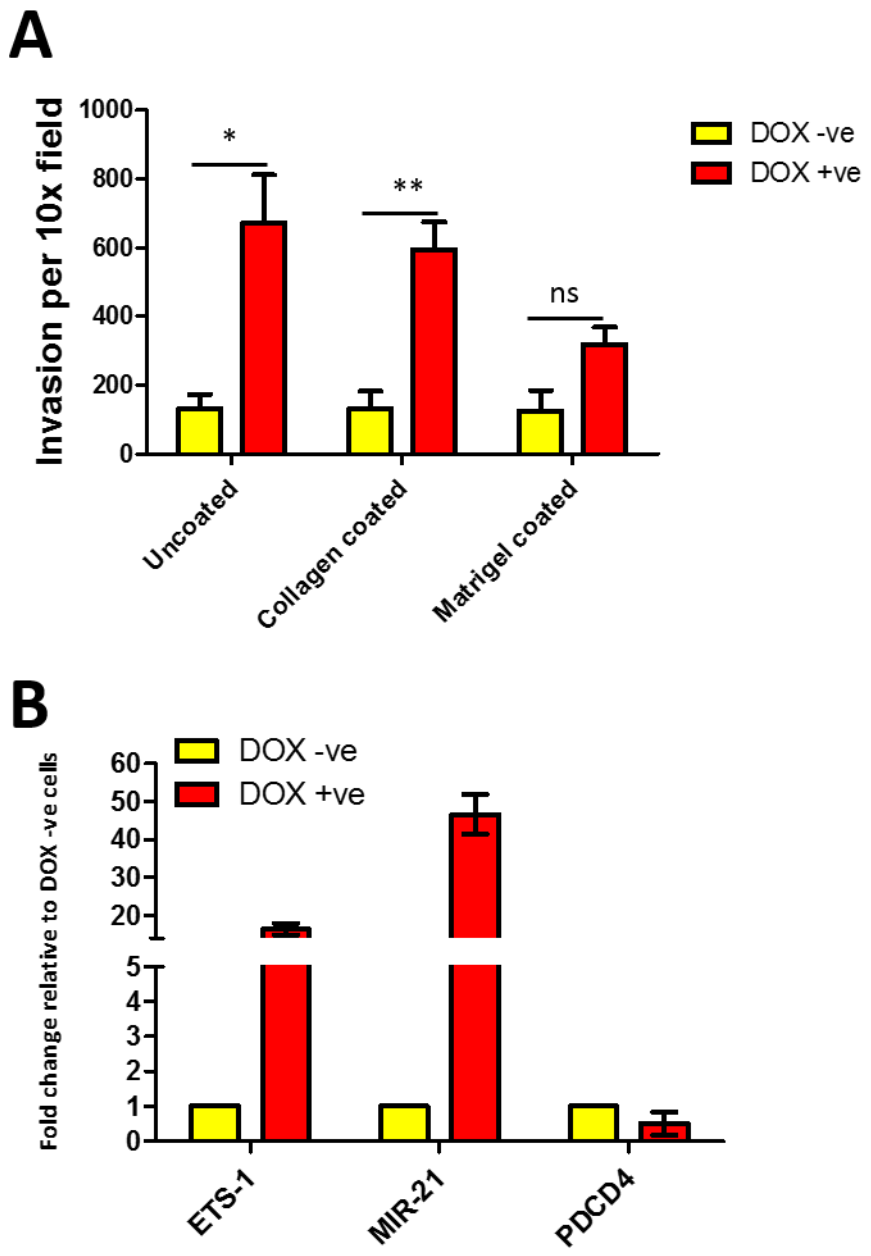


Figure 3.6: HER2-associated migration and invasion. (A) Cell migration was analysed through the transwell membranes over 16-hour period of chemotactic migration towards full serum media. The ability of cell invasion was measured in collagen or matrigel coated transwells (n=3). Student's t-test was performed and statistical significance is shown as * for p-value < 0.05, ** for p-value < 0.01. (B) Gene expression analysis by RT-PCR of a known pathway associated with HER2-induced invasion. Cells were grown in 3D cell culture for 5 days and acini recovered. PDCD4 levels decreases via upregulation of ETS1, and primary mir-21 (n=2).

3.7 HER2 over expression induces anchorage-independent growth

Transformed cells have the ability to grow large colonies compared to normal cells in the soft agar, a characteristic known as anchorage-independent growth, and a hallmark for *in vitro* transformation (239, 240). The formation of domed-shaped colonies are strongly correlated with formation of tumours in experimental mice (241, 242). We tested the anchorage-independent growth capability of DOX +ve and DOX -ve cells by growing 5000 cells mixed with a low percentage (0.3%) of ultra-pure agarose on top of a 0.8% layer of ultra-pure agarose. Cells were replenished with fresh media containing dox to maintain HER2 over expression over the course of 21 days and colonies were quantified after imaging by a dissecting microscope. HER2 over expression in DOX +ve cells induced anchorage-independent growth in the soft agar, but the control cells did not (Figure 3.7A). Some DOX +ve cells grew large colonies (above 100µm perimeter) but DOX -ve cells did not aggregate or form larger colonies after 21 days of

being in the soft agar. The DOX +ve cells increased in their sizes and became more distinguishable after 10 days, some of which were very large, round structures (Figure 3.7B).

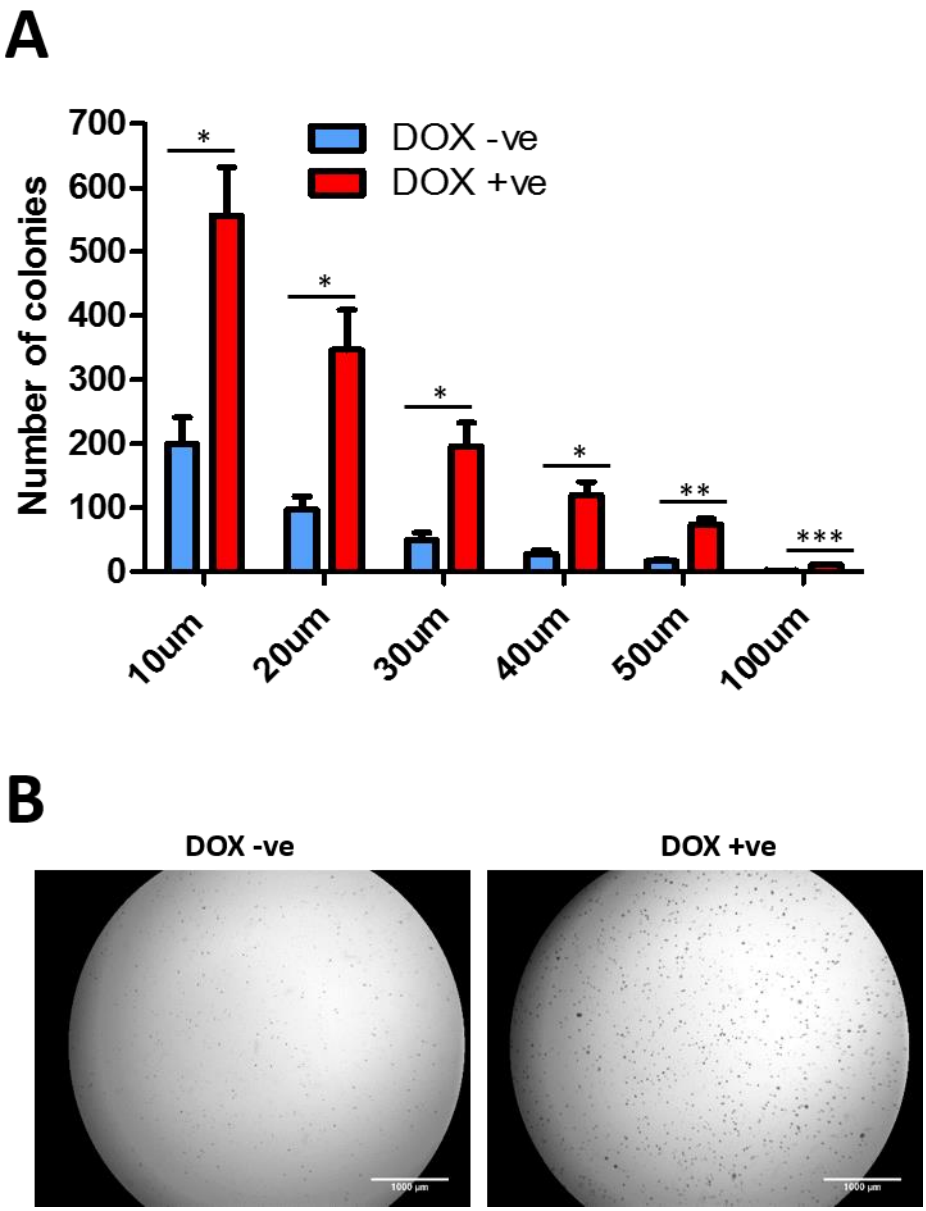


Figure 3.7: HER2 induces MCF10A cell anchorage-independent growth. (A) Colony growth of DOX -ve and DOX +ve cells in 0.3% ultra-pure agarose. ImageJ analysis of six different size colonies were quantified. Student's t-test was performed and statistical significance is shown as * for p-value < 0.05, ** for p-value < 0.01, *** for p-value <0.001. n=3 (B) Representative microscopic images of colonies stained with crystal violet after three weeks. Images are at 1.6x magnification. Scale bars represent 1000µm.

3.8 Discussion

Previously, inducible transformation models have studied various differences occurring between normal and transformed cells, for example by inducing over expression of HER2/neu in mouse models (243, 244) and primary luminal cells to characterise a phenotype of transformation, namely the filling of the lumen (115). Whilst this has been valuable to understand the role of HER2 in driving transformation, it has an important limitation in that is not suitable to track the alterations that occur at the very outset of transformation with physiological levels of HER2 expression. Here, we ectopically generated a stable MCF10A cell line transduced with inducible HER2 gene using the Tet-On system that allows for characterisation of the earliest changes. The wild type HER2 over expression alone is sufficient to transform the immortalised, yet normal MCF10A cell line (245). We anticipate to fully exploit this system in understanding the genome-wide early signalling and chromatin structure changes upon HER2 induction.

To work with the successfully transduced HER2 MCF10A cells, we FACS selected only the very high 2.3% of cells based on GFP expression, which is expressed from the weak EF1α promoter, despite the HER2 being driven by a distinct strong inducible TRE (tetracycline response element) promoter. The FACS selection of GFP cells with high fluorescence intensity would have ensured that the majority of the transduced cells contained more than one viral integration per cell. However, there would still be variation in the HER2 levels among different MCF10A cells within the same population as is shown by the IF. This variation in HER2 levels could present key caveats due to the rapidly evolving heterogeneity in HER2 expression or due to outgrowth of one clone over others in cell culture affecting the inducibility and the overall expression of HER2 over time. This limitation of the inducible system cannot be differentiated from the response of cells to dox, as it is already known that there is variation in inducibility

over long periods of time, with cells having reduce response to dox, resulting in significantly decreased inducibility (246).

An essential feature of an inducible system is its high inducibility in the presence of an inducing agent (dox) and its low background or leakiness in the transduced cells in the absence of dox. Our results show tightness of dox-regulation as the DOX -ve cells are identical to the untransduced MCF10A cells exhibiting no “leakiness” of HER2 protein in absence of dox. In addition, the system displays that HER2 expression is strictly dependent on dox treatment, as the plasmid enables the expression of HER2 in a graded fashion by titrating the dox concentration. Furthermore, the plasmid encodes GFP as a surrogate marker, which can only help to monitor the successful delivery of the plasmid into the target cells. This marker cannot be used to monitor the expression of gene of interest (HER2), as both of the genes are driven by two different prompters and are therefore, transcribed and expressed at different levels.

The 3-dimensional basement membrane cell culture offers significant insights compared to 2D cultures of normal cells progressing to cancer, making it an ideal system for us to study the morphogenesis in a more physiological relevant context. We monitored the morphological transformation of cells over a period of 9 days. DOX +ve clearly showed an aberrant morphology whereas the DOX -ve did not. The normal MCF10A and DOX -ve cells are known to exhibit growth arrest, which appear to be delayed or does not occur in the transformed cells. This growth arrest in the DOX -ve cells results in a notable lumen formation after 10 days of being in 3D cell culture. Interestingly, the transformed cells demonstrate lack of apoptosis and hence the lack of lumen formation. We have not been able to observe this phenotype in our experiments since our experimental end point was 9 days, with the lumen formation appearing after 10 days. Nevertheless, we show that HER2 over expression disrupts the normal MCF10A morphology within 3 days and this invasive morphology is maintained by HER2 expression, which was the main objective of our experiment.

Our results show a marked increase in the ability of DOX +ve cells to form colonies compared to DOX -ve cells. Intriguingly, some of the DOX -ve cells also have the ability of colony formation. The colony forming ability of normal MCF10A cells to this extent has not been reported before. Previous reports show only some colony formation or the complete inability of MCF10A cells to grow in the soft agar compared to oncogene-induced transformed MCF10A cells (247-250). However, these reports do not take into consideration the sizes of the colonies, whereas we show that approximately 100 colonies were detected in DOX -ve cells compared to the 350 colonies in DOX +ve cells when 10µm perimeter size was set as a threshold. Moreover, there is an incremental reduction in the number of colonies detected in the DOX -ve cells as the size threshold was increased. For example, no colonies were detected in the DOX -ve cells when the threshold was increased to 100um perimeter compared to 10 large colonies detected in the DOX +ve cells. Another explanation that may explain this discrepancy is that MCF10A cells are extremely adherent, form round domed-shaped colonies, and generally aggregate together and therefore, the colonies observed in the DOX -ve cells are a mere characteristic of these cells exhibited in the ultra-pure agarose and are not relevant to transformation.

Chapter 4

Investigating the dynamics of early signalling changes upon HER2 over expression

4.1 Introduction

Signalling pathways convey cellular information into the nucleus in response to external stimuli by various post translational modifications (PTMs) to proteins (251). The PTMs contribute to vital roles by regulating biological processes such as cell growth, survival, invasion, differentiation, and protein turnover. Importantly, reversible phosphorylation events play a central role in the growth of tumours. For example, the HER receptor family is activated by various ligands, which in turn can initiate a cascade of widespread phosphorylation in downstream signalling pathways to promote tumour development (251, 252). Indeed, phosphorylation in cancer cell signalling has been actively studied in various biological contexts, but there is a need for network-wide analysis of each signalling dynamics to define the signalling machinery at the system level.

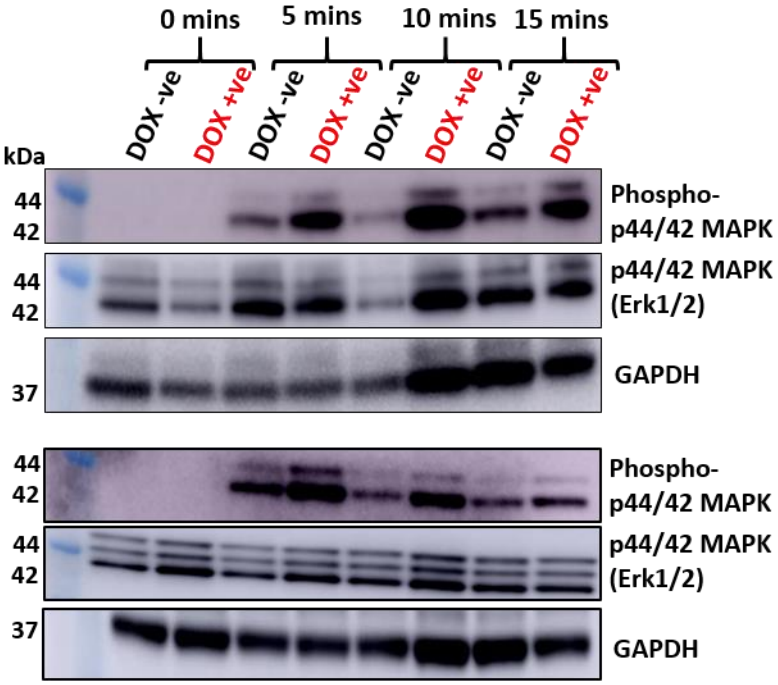
To better characterise the phenotypic consequences we have observed through HER2 mediated cellular neoplastic transformation, we undertook a detailed global study to investigate the molecular signalling events in the phosphoproteome driven by HER2 expression by an unbiased and comprehensive phosphoproteomic approach. We were particularly interested in determining the molecular changes that take place at the very outset of HER2 protein induction and transformation. We aimed to assess the effect of HER2 over expression at short time points after protein induction, in the absence of any other genetic or epigenetic alterations except from those already occurring in the immortalised MCF10A cell line. In this system, HER2 protein levels increase in a time depend manner upon doxycycline addition.

Therefore, the low-level sequential activation of HER2 in early time points may mimic physiological signalling events as observed in cancer cells. Indeed, in many different cancers, wild type HER2 is over expressed as is the case in this system.

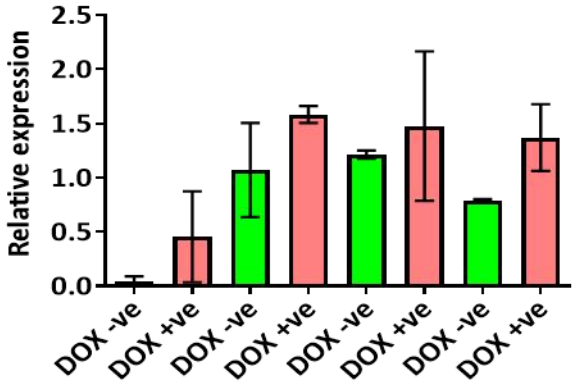
4.2 Detection of downstream signalling events upon HER2 induction

To establish if the HER2 inducible construct transduced in the MCF10A cells is functioning, we first assessed the phosphorylation status of known proteins activated by HER2 protein induction. The DOX -ve and DOX +ve cells were cultured in serum free media (by removing EGF and horse serum) for 24 hours to lower pathway activities close to basal levels. Cells were then stimulated with full media (containing EGF and horse serum) for either 5, 10, 15 minutes or left untreated in the serum free media, as a negative control. Western blotting analysis revealed notably higher phosphorylation of ERK in DOX +ve cells compared to DOX -ve cells, as is shown by the increase in phosphorylation of the activatory modification (ERK[1/2] Thr/202/Tyr204) levels upon stimulation with full media, in two biological replicates (Figure 4.1A). Moreover, to determine the phosphorylation changes occurring upon the activation of the PI3K-AKT signalling pathway, we performed western blotting for AKT activatory modification at serine 473 (S473). Cells were serum starved for 24 hours and then stimulated with full media for either 15, 30, 45 minutes or left untreated in the serum free media. We chose longer time-points of stimulation compared to ERK phosphorylation as it is known that AKT phosphorylation is slower compared to rapid transmission of signalling through the MAPK signalling pathway. Interestingly, HER2 over expression did not have an effect on the activation of AKT at this specific residue (Figure 4.1B).

A


0 mins
5 mins
10 mins
15 mins
DOX -ve
DOX +ve
kDa
44
42
Phospho-
p44/42 MAPK
p44/42 MAPK
(Erk1/2)
37
GAPDH

B


Relative expression
2.5
2.0
1.5
1.0
0.5
0.0
DOX -ve
DOX +ve

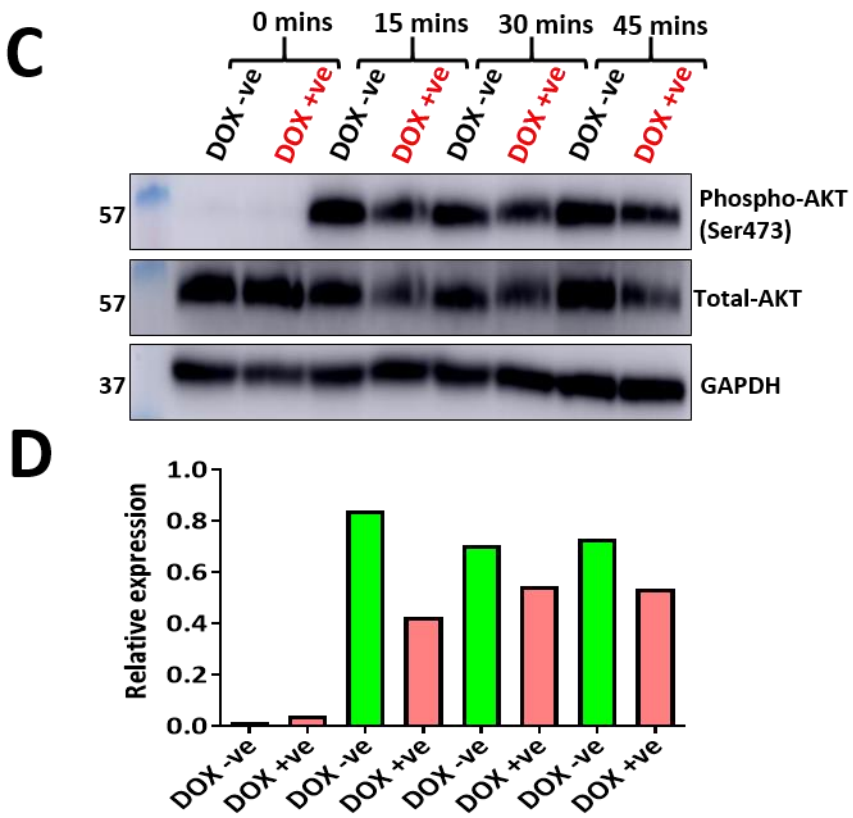


Figure 4.1: Detection of ERK and AKT activation upon HER2 protein over expression. (A) To detect ERK activation DOX -ve and DOX +ve cells were gown in serum starved cell media for 24 hours and then stimulated with full media for the indicated time points or left in the serum starved media as a negative control. Detection of phospho-ERK (Thr202/Tyr204) is shown. After stripping, the same membrane was blotted for total-ERK (ERK 1/2) and GAPDH was used a loading control. n=2. (B) Densitometry analysis of pERK expression normalised to GAPDH (loading control) was performed using Image J (n=2). (C) For AKT activation DOX -ve and DOX +ve cells were gown in serum starved cell media for 24 hours and then stimulated with full media for the indicated time points or left in the serum starved media as a negative control. Detection of phospho-AKT (S473) is shown. After stripping, the same membrane was probed with a total-AKT antibody and GAPDH was used as a control. n=1. (D) Densitometry analysis of pAKT expression normalised to GAPDH (loading control) was performed using Image J (n=1).

4.3 HER2 over expression increases expression of genes related to angiogenesis and adhesion mediators

To validate the expression of genes that are known to be enhanced upon HER2 over expression and in the presence of exogenous EGF, we performed RT-PCR for 6 different genes known to be involved in transcriptional induction of adhesion, morphogenesis and angiogenesis (3). HER2 over expression was maintained in DOX +ve cells for 5 days in 2D cell culture and RT-PCR was carried out on total RNA from both DOX -ve and DOX +ve cells. We found that there was no significant change in the expression of angiogenic and adhesion factors including LOX, BMPR2, ILK, VEGFC, BMP6, and LOXL2 when cells are grown in 2D cell culture (Figure 4.2A). We thought that this because these genes are relevant to the processes of cell adhesion and angiogenesis and there expression may not be directly significant in 2D cell culture. Therefore, we extended our analysis to 3-dimensional (3D) cell culture and plated DOX -ve and DOX +ve cells in a mixture of matrigel and collagen overlay ("on top") 3D cell culture method for the same number of days (5 days) as the 2D cell culture. Acini from 3D cell culture were recovered and expression of the same genes as above were validated by RT-PCR. Intriguingly, HER2 over expression increased the transcription of adhesion and angiogenic molecules in the acini of the MCF10A cells (Figure 4.2B).

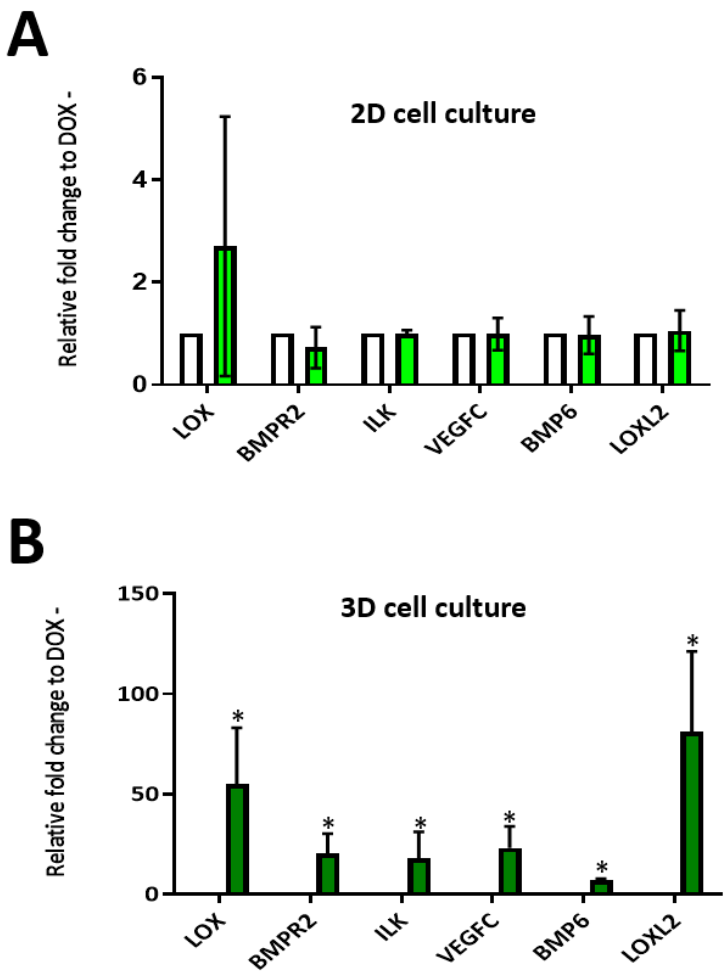


Figure 4.2: Transcriptional effects of HER2 over expression. mRNA expression of genes known to be upregulated upon HER2 over expression. (A) Total RNA was isolated after 5 days of HER2 over expression in MCF10A cells in 2D cell culture, subjected to reverse transcription and analysed by RT-PCR. Light green and bars represent DOX +ve cells and white bars show DOX -ve cells (n=2). (B) HER2 was induced in 3D cell culture for 5 days, acini recovered and RT-PCR was performed. Green bars represent relative expression of DOX +ve cells compared to DOX -ve cells (n=2). Two-tailed student t-test was performed and is depicted as significant [* < 0.05 *p-value*].

4.4 Phospho-proteomic analysis of HER2 activation – an overview of experimental design

In order to map the early molecular signalling events induced by HER2 protein over expression and cellular transformation, we performed liquid chromatography tandem-mass spectrometry

(LC-MS/MS) based phospho-proteomic analysis. To ensure reproducibility of minor quantitative changes, the experiment was repeated in 3 biological replicates and each sample was analysed twice by mass spectrometry. We selected 4 different time points for DOX -ve cells and added doxycycline to induce HER2 protein expression at 0 hours, 0.5 hours, 4 hours, and 7 hours to capture signalling dynamics at early and early-immediate points. The decision of selecting these time points was based on the western blotting of HER2 over expression in a time-dependent manner, and we saw that HER2 is expressed early upon dox addition, we wanted to study signalling changes at the very outset of HER2 protein expression (Figure 4.3).

As a control, we added doxycycline to an empty GFP vector transduced in MCF10A cells at the same time points. The analysis compared the signalling changes in a time-dependent manner by comparing each time point to the 0 hour time point (0.5 hours vs 0 hours, 4 hours vs 0 hours, and 7 hours vs 0 hours) to capture the earliest changes during the process of transformation (Figure 4.3). To obtain differentially regulated phospho-peptides from our dataset, we filtered out background phosphorylation events occurring natively and by the addition of dox in the GFP transduced MCF10A cells. We defined a phospho-peptide to be significantly differentially regulated if changes in phosphorylation intensity, such as increases or decreases in expression had a False Discovery Rate (FDR) corrected *p*-value of less < 0.05.

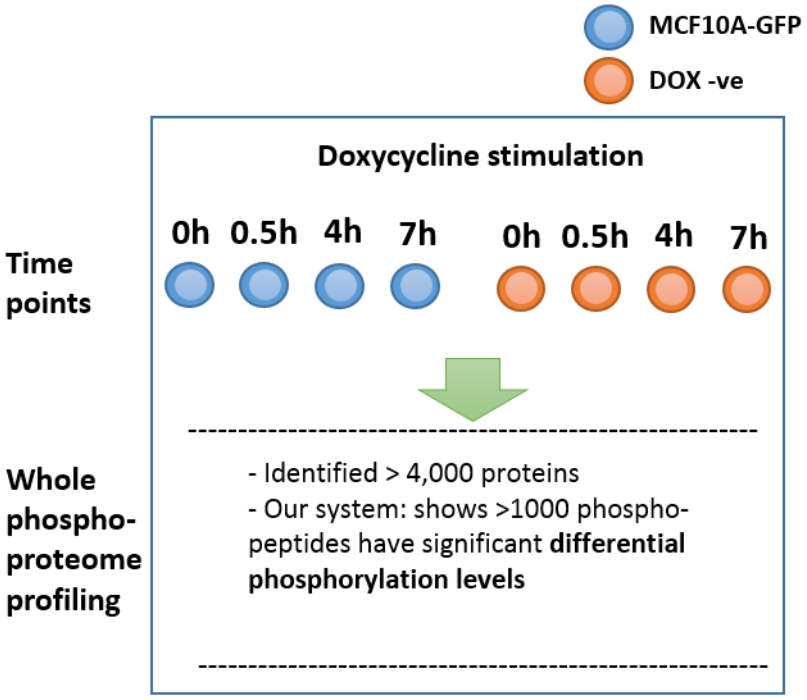


Figure 4.3: Schematic of experimental outline and phospho-proteome dataset overview. GFP transduced MCF10A cells and DOX -ve cells were stimulated by dox and collected at time points labelled. Samples were then subjected to mass spectrometry analysis. Collectively, over 4000 unique proteins were found to be modified and post our analysis pipeline, more than 1000 phospho-peptides were found to have differential phosphorylation levels.

4.5 Overview of the phosphorylation changes upon HER2 activation

In total, our data analysis workflow revealed changes in 4089 proteins containing one or more phosphopeptide. The differentially regulated phosphorylation changes were observed in 800 proteins, which equalled to 1004 phosphopeptides. From this, 383 proteins were enriched in phosphorylation and 417 proteins were depleted of phosphorylation. We also quantified phosphorylation changes occurring at each time point upon HER2 activation. We found that there were 310 differentially regulated phospho-peptides at 0.5 hours (that may or may not also be significantly changing at other time points), 701 at 4 hours and 663 at 7 hours upon HER2 induction. The effects of HER2 on all proteins in our experimental setting was also quantified (Figure 4.4). Of those proteins that showed increase in phosphorylation, less than

15 proteins that contained a phospho-site that showed higher than 5 fold increase in phosphorylation abundance at any time point. Fewer than 100 proteins across all time points exhibited increase in phosphorylation between 1 to 5 fold intensity. Interestingly, most of the affected proteins exhibited a decrease in abundance of phosphorylation because at every time point the downregulated phosphorylation sites outnumbered those that were upregulated.

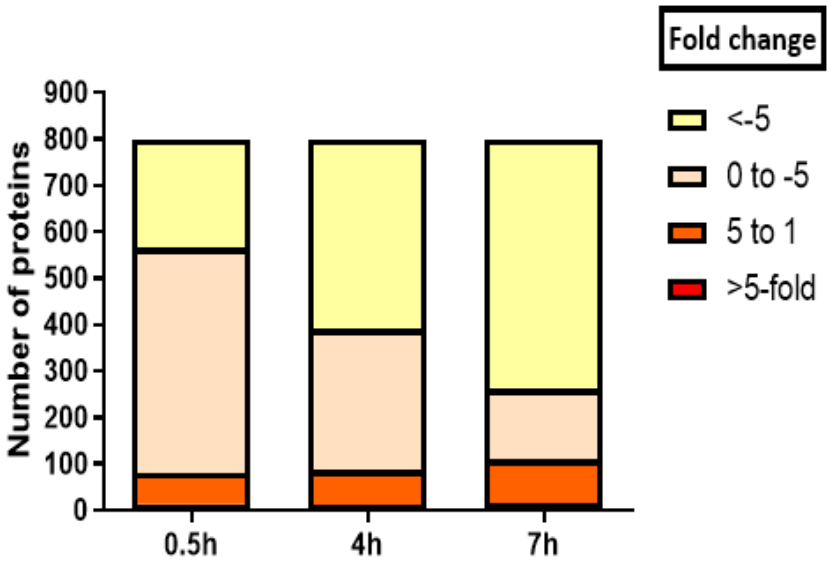


Figure 4.4: Quantification of protein phosphorylation. HER2 effect on phosphorylation of proteins that show varying fold changes of phosphorylation.

To ensure that the experimental design has been correctly executed we first checked if known proteins were activated upon HER2 induction. The proteins that should be phosphorylated in this model are HER2 and its family member HER1 (EGFR). As expected we observed an increase in both the HER2 and HER1 phosphorylation levels at sites T701 and Y1110 in a time-dependent manner (Figure 4.5).

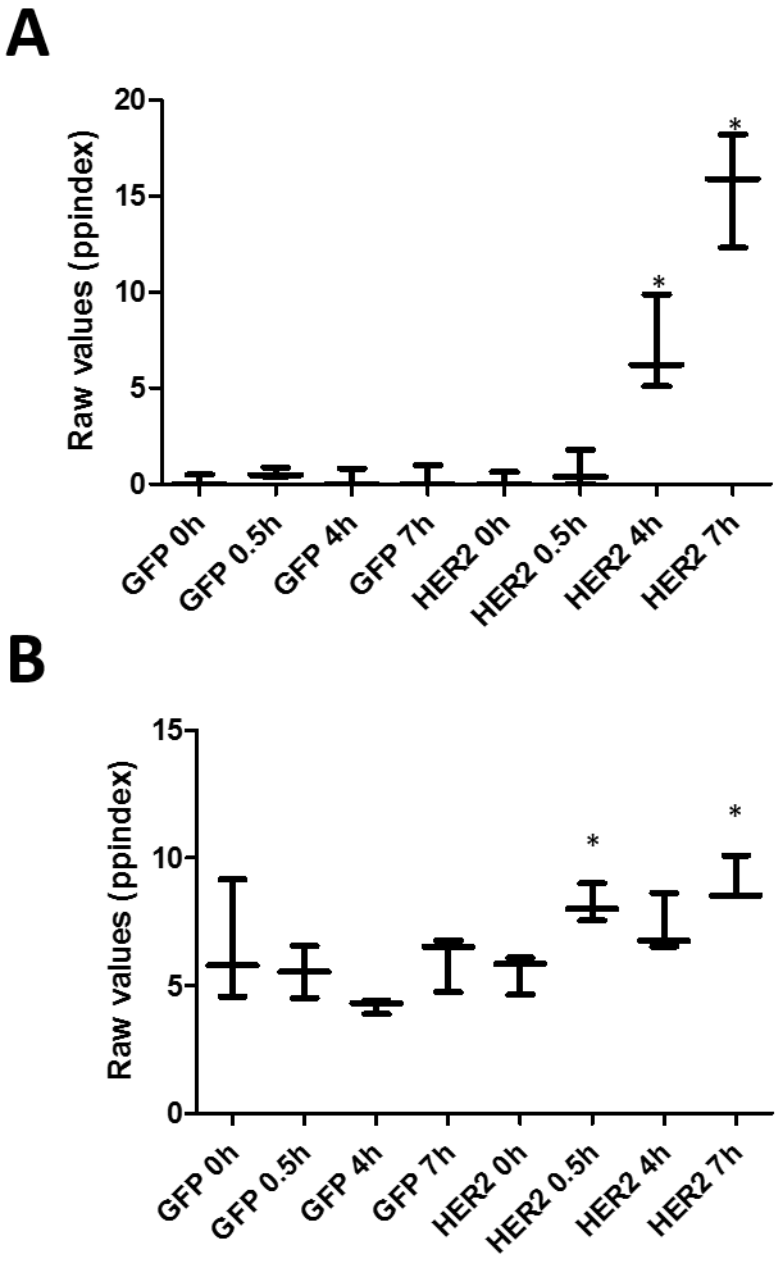


Figure 4.5: An internal quality control (QC) for phospho-proteomic analysis. (A) HER2 phosphorylation modification (T701) increases in a time dependent manner. (B) EGFR (Y1110) also becomes marginally activated in a time dependent manner compared to control cells. [* FDR corrected p-value of < 0.05.

4.6 HER2 induced time-dependent differentially regulated phosphorylation events

Having screened the phospho-proteome of MCF10A cells upon HER2 activation, we next analysed how this phospho-proteome is impacted by cellular transformation induced by HER2

protein expression. A volcano plot for the 0.5 hours' time point is shown in the Figure 4.6A. This shows the immediate early phosphorylation events of 310 proteins, which include the upregulation of the HER1 (EGFR), the transcription factor JUN, the activation of PAK2, and NKTR, but also the downregulation of a novel DNMT1 phospho-peptide amongst many other changes. More specifically, of the 310 differentially and significantly changing phospho-peptides, 153 were significantly depleted (log2 fold change < -0.5, FDR corrected *p*-value of < 0.05), whereas 94 were significantly enriched (log2 fold change > 0.5 fold, FDR corrected *p*-value < of 0.05). We used the 0.5 log2 fold cut off for upregulation and downregulation (represented by the dotted vertical line), since most of these peptides exhibited only marginal phosphorylation change. To understand how many phospho-sites are significantly increasing with a higher fold change, we picked a 2 log2 fold cut off and found that only 39 phospho-peptides were changing significantly and 4 phospho-peptides were significantly down-regulated with a cut off of log2 fold change <-2. The phospho-peptides that were changing (upregulated or downregulated) at 0.5 hours were subjected to ontology enrichment analysis (Figure 4.6B). Using the DAVID bioinformatics functional annotation, we identified the biological processes that are significantly altered upon HER2 over expression in 0.5 hours. The clusters consisted of establishment of RNA localisation (cluster 1), cell-cell adhesion (cluster 4), chromosome organisation (cluster 5), and gene silencing (cluster 6). KEGG PATHWAY analysis and DISEASE annotations did not reveal any significant terms in any pathway or disease clusters, respectively, at the 0.5 hours' time point.

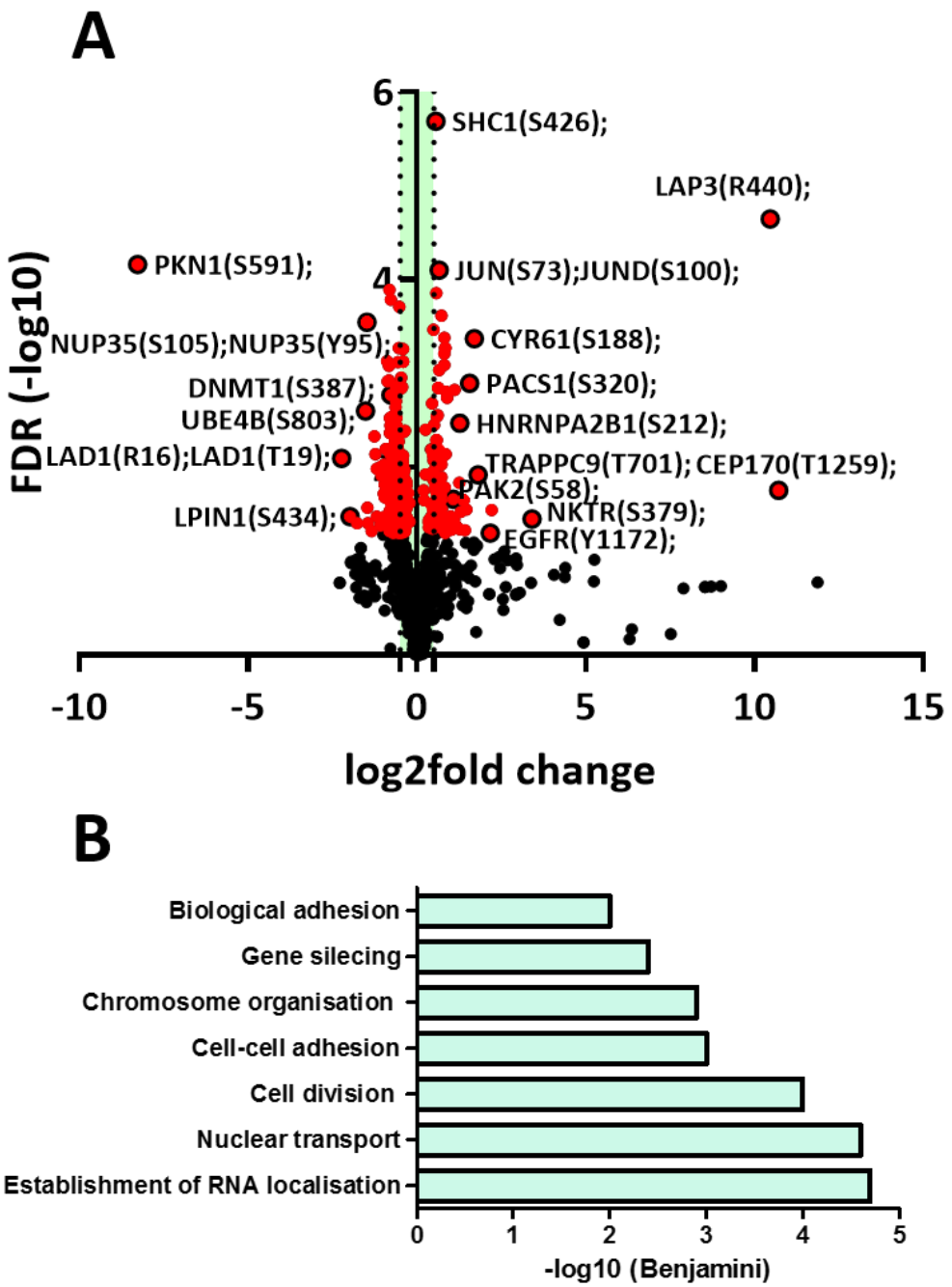
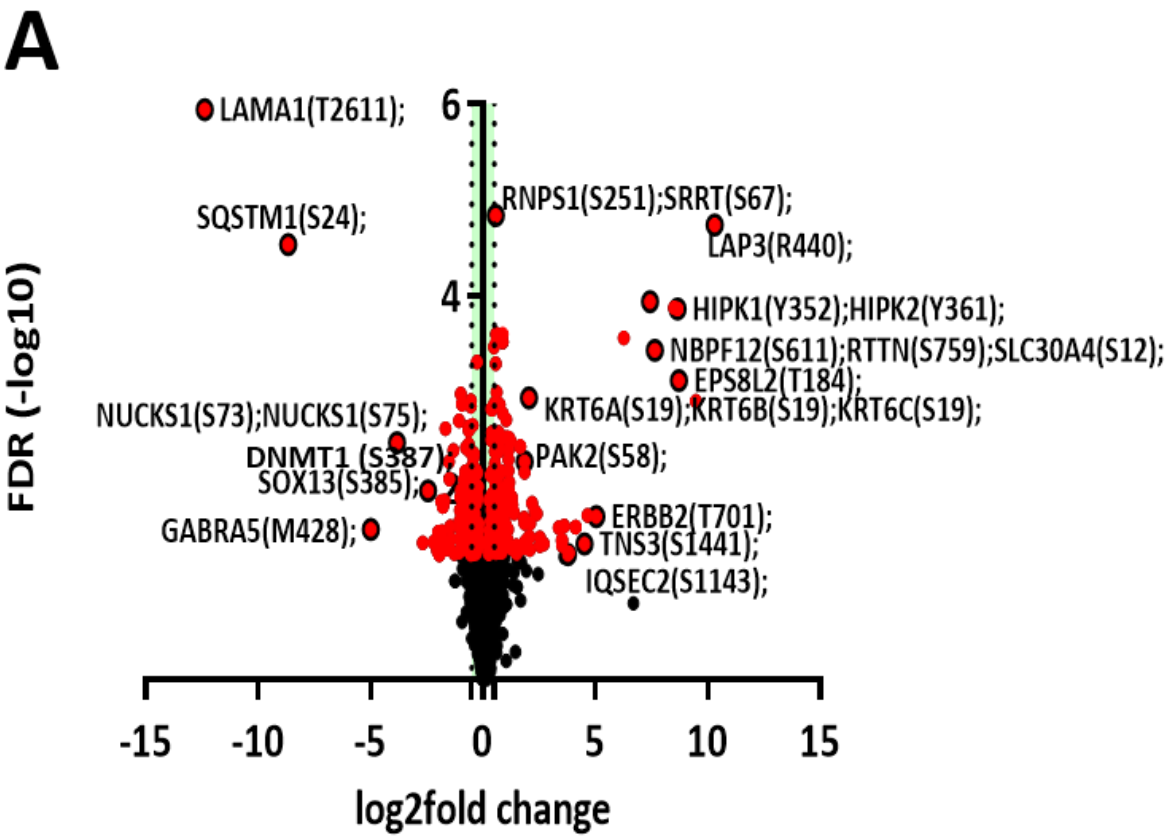


Figure 4.6: Volcano plot of showing the phospho-proteome of HER2 induced changes in MCF10A cells at 0.5 hours. The red circles show the significant differential phosphorylation changes and the black circles show non-significant changes. The labelled phospho-peptides are indicated by red/black circles. The statistical significance was –log10 of the FDR corrected p-values (y axis) and the fold change is shown on the x axis. The vertical dotted line indicates a 0.5-fold change. (B) Gene ontology analysis of biological processes using DAVID of all the changes occurring at 0.5 hour upon HER2 induction and

transformation. The resulting Benjamini p-values for each term were –log10 transformed with a threshold of 0.05.

Next, we assessed the differential phosphorylation changes of 390 phospho-peptides upon 4 hours of HER2 induction and transformation. These alterations are visualised in a volcano plot (Figure 4.7A). There were 125 phospho-peptides significantly depleted (log2 fold change < -0.5, FDR corrected *p*-value of < 0.05) and 168 were significantly upregulated (log2 fold change > 0.5 fold, FDR corrected *p*-value < of 0.05). We find the emergence of HER2 phosphorylation, and the activation of PAK2 alongside the hyper phosphorylation of a known HER2 interactor, EPS8L2. Interestingly, the downregulation of p53 binding protein was noted, and the sustained downregulation of the same phospho-site of DNMT1 observed in the earlier 0.5 hours' time point. To investigate how many phospho-sites are significantly increasing with a higher fold change, we picked a higher cut off threshold of 2 log2 fold cut off and found that only 35 phospho-sites were upregulated and 10 were significantly down-regulated with a cut off of log2 fold change <-2. Ontology analysis of these phosphorylated proteins identified biological processes enriched for various terms, such as Ras protein signal transduction (cluster 1), cell-cell adhesion (cluster 5), and the regulation of signalling (cluster 7) (Figure 4.7B). Similar to the

0.5 hours' time point, KEGG PATHWAY and DISEASE annotation analysis did not enrich for any significant terms.



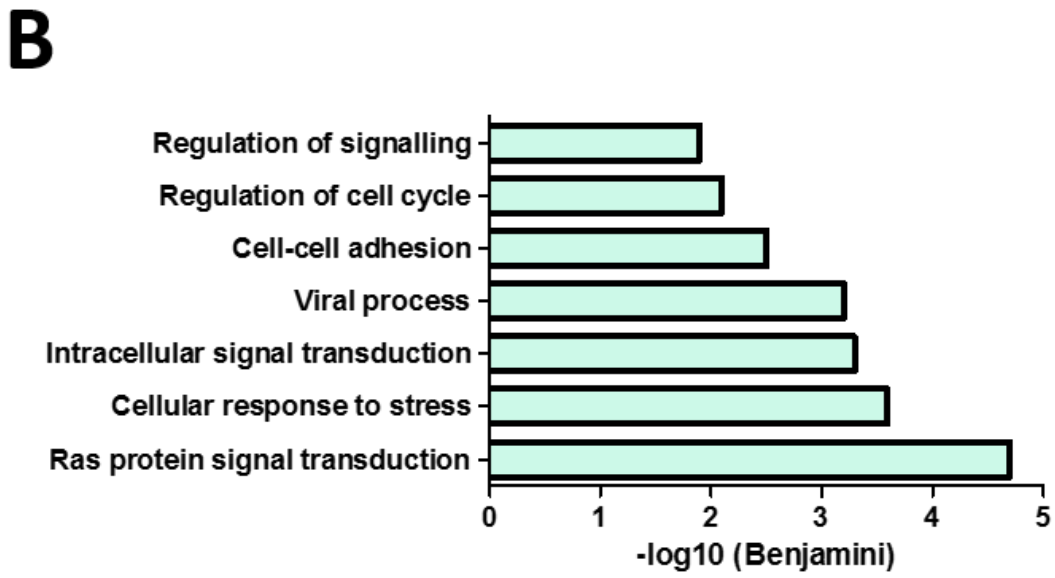


Figure 4.7: Volcano plot of showing the phospho-proteome of HER2 induced changes in MCF10A cells at 4 hours. The red circles show the significant differential phosphorylation changes and the black circles show non-significant changes. The statistical significance was –log10 of the FDR corrected *p*-values (y axis) and the fold change is shown on the x axis. The vertical dotted line indicates a 0.5-fold change. (B) Gene ontology analysis of biological processes using DAVID of all the changes occurring at 4 hours upon HER2 induction and transformation. The resulting Benjamini *p*-values for each term were –log10 transformed with a threshold of 0.05.

We next examined the differential phosphorylation changes upon 7 hours of HER2 induction and cellular transformation, which resulted in 455 differentially regulated phospho-peptides, visualised by a volcano plot (Figure 4.8A). Of these, 157 phospho-peptides were significantly depleted (log2 fold change < -0.5, FDR corrected *p*-value of < 0.05) and 213 were significantly enriched (log2 fold change > 0.5 fold, FDR corrected *p*-value < of 0.05). Indeed, there is an overlap of phospho-peptides that were observed in the 0.5 hours or 4 hours' time points, as the HER2 (T701), DNMT1 (S487), TP53BP1 (S1067), PAK2 (S58), are all maintained indicating that these changes are not transient. However, novel changes that were not observed in the previous two time points also appear. These include a second HER2 phospho-peptide (HER2 T1060) as well as the activation of NKTR and PTK2 (FAK2) amongst many other alterations. To investigate how many phospho-peptides are significantly upregulated with a higher fold change, we picked a higher cut off threshold of 2 log2 fold cut off and found that only 45 phospho-sites were upregulated and 14 were significantly down-regulated with a cut off of log2 fold change <-2. To understand the biological significance of these alterations, we performed ontology analysis of the changes occurring at the 7 hours' time point. There were overlapping enrichment of biological processes such as cell-cell adhesion (cluster 1), nuclear chromosome segregation (cluster 3), cell projection organisation (cluster 4) and cell development (cluster 6), amongst others (Figure 4.8B). Interestingly, KEGG PATHWAY analysis

revealed significant changes in 3 different pathways, including the ErbB (HER) signalling pathway (Figure 4.8C).

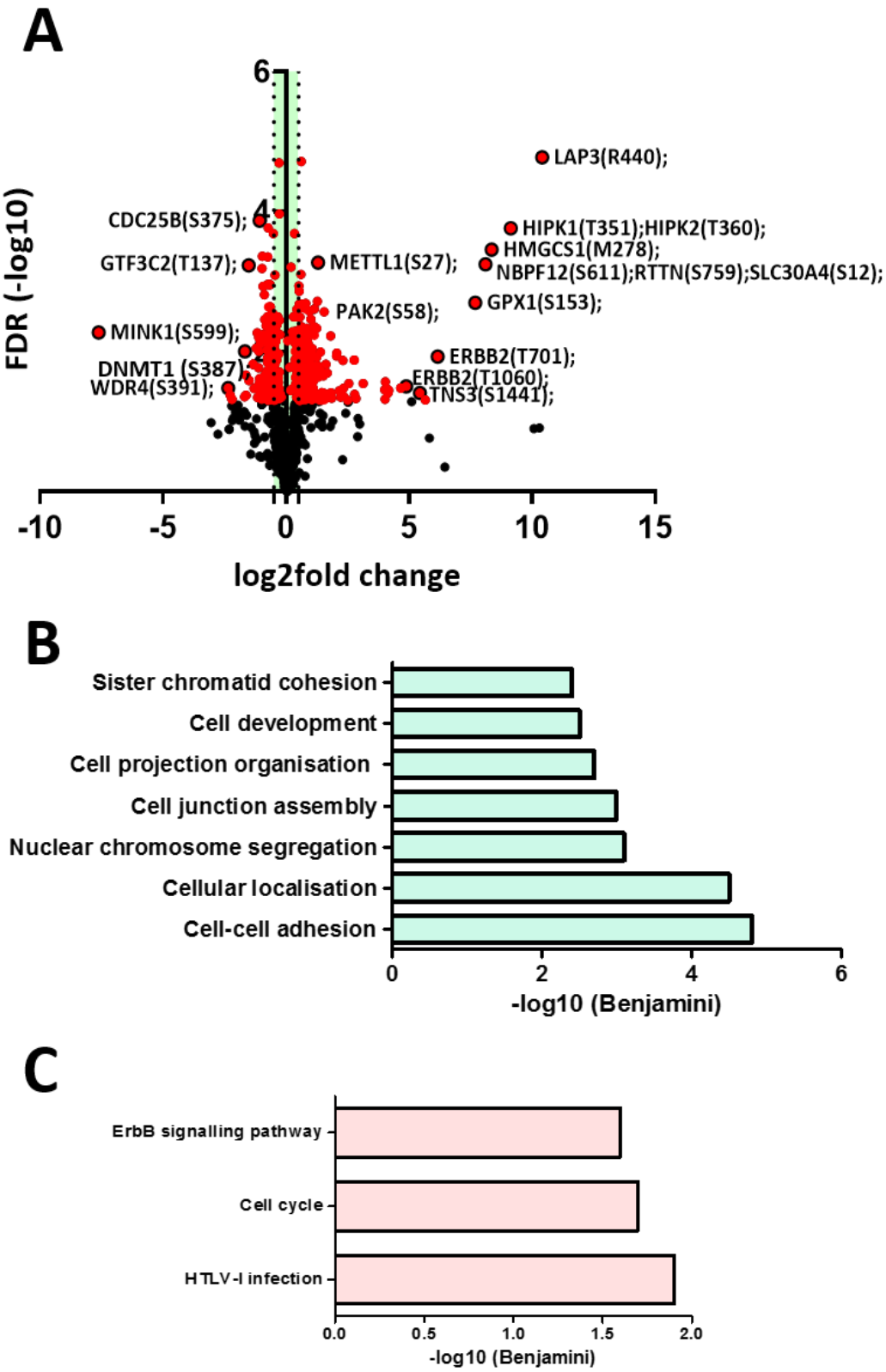


Figure 4.8: Volcano plot of showing the phospho-proteome of HER2 induced changes in MCF10A cells at 7 hours. The red circles show the significant differential phosphorylation changes and the black circles

show non-significant changes. The statistical significance was –log10 of the FDR corrected p-values (y axis) and the fold change is shown on the x axis. The vertical dotted line indicates a 0.5-fold change. (B) Gene ontology analysis of biological processes using DAVID of all the changes occurring at 7 hours upon HER2 induction and transformation. (C) Pathway enrichment analysis using DAVID. The resulting Benjamini p-values for each term were –log10 transformed with a threshold of 0.05.

4.7 Multisite protein phosphorylation

Multisite protein phosphorylation is a major mechanism of regulating the activity of proteins (253). We found that across all time points in the upregulated or the downregulated phospho-peptides, single phosphorylation sites per protein were strongly represented compared to multi-sites phospho-peptides (Figure 4.9A). Approximately 80% of the identified peptides were phosphorylated on just one residue, whereas the remaining 20% were phosphorylated at multiple sites of 2 or more. As we observed that many proteins could potentially have multi-site protein phosphorylation, we next asked if they were activatory or inhibitory. We detected changes in the phosphorylation status of several regulators and kinases activated upon HER2 expression with a multitude of phosphorylation modifications. These included HER2, SRC substrate cortactin, EGFR, FAK1, and P63 amongst many others (Figure 3.9B). We then manually inspected the multiplicity of phospho-peptides of several proteins to assess if they are associated with the activation or the inhibition of that protein. Interestingly, out of the 12 proteins we searched for on PhosphoSitePlus, 5 of them had an activating (inducing) function (coloured red) or the function is not yet known. The other 7 phospho-peptide function (activatory or inhibitory) is yet to be elucidated (white bars) and no protein was found to have an inhibitory effect (Figure 4.9B). Compared to single site phosphorylation, multi-site protein phosphorylation maybe considered as an on/off switch for protein function and it increases the possibilities for protein regulation, with each phospho-site have a distinct characteristic (254).

Furthermore, of the 800 differentially regulated proteins we investigated a possible correlation with HER2 activity by calculating the Pearson's correlation coefficient (R^2) between the HER2 phospho-peptide (HER2 T701) and the abundance of other phosphorylated peptides following similar intensity as the HER2 T701 modification. By applying a cut off value of 0.8, we identified 148 proteins strongly following the trend of HER2 fold increase pattern (Figure 4.9C).

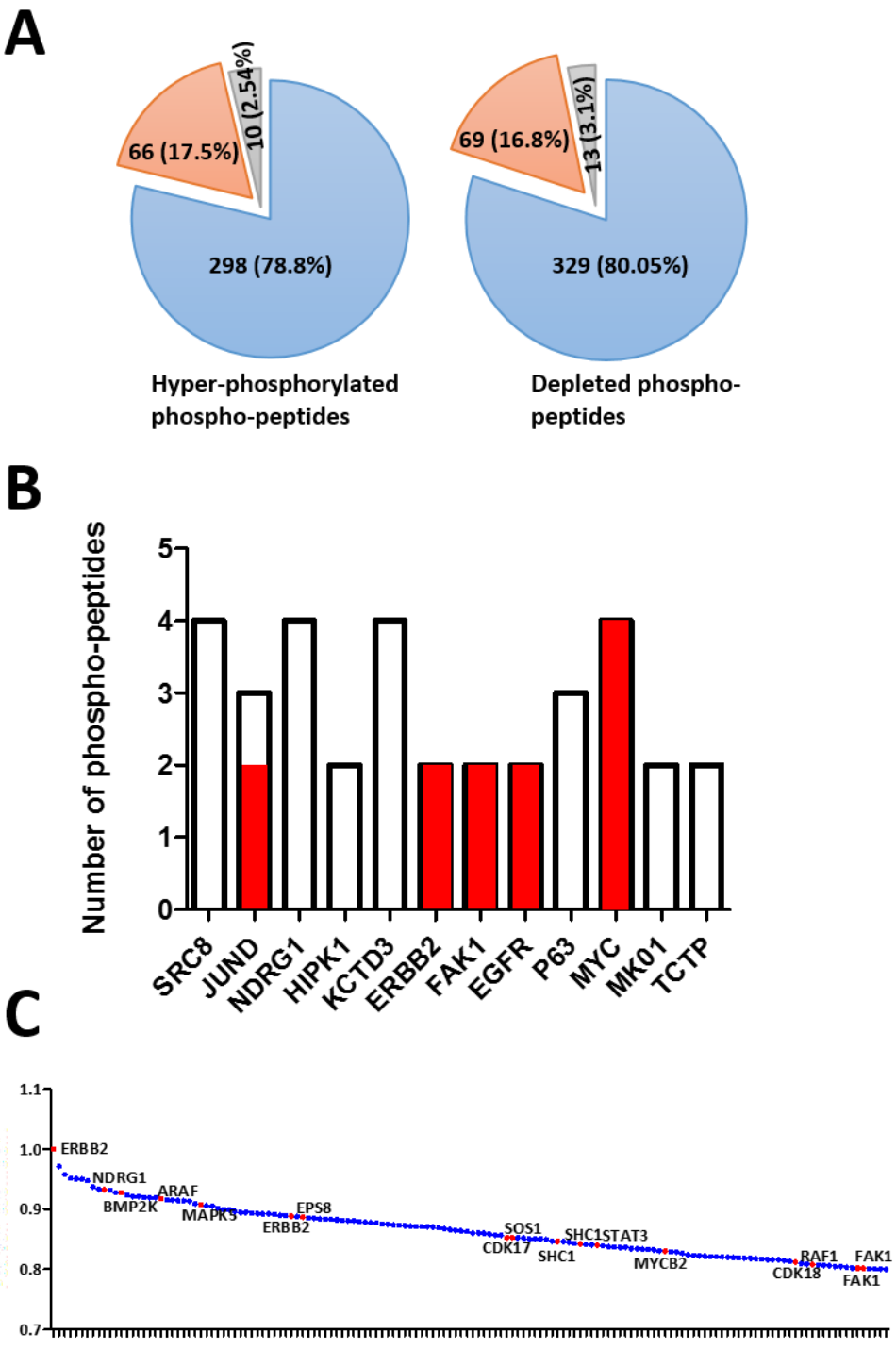


Figure 4.9: Phospho-proteome identification. (A) Pie charts represent the multi-site protein phosphorylation of the identified phospho-peptides. Percentage of phospho-peptides carrying either a single (blue), double (orange) or more than three residues (grey) are indicated. (B) Bar chart showing the multiplicity of the phosphorylation sites of some proteins. The red bars show the already known activating phospho-peptides. The white bars represent the activatory or inhibitory effects that are not

yet known. No inhibitory effect was found. (C) Pearson's ranked phosphorylation changes of phospho-peptides following the intensity of HER2 T701 residue.

4.8 Quantitative phospho-proteomic analysis of HER2 induced changes

To determine the signalling pathways activated upon HER2 induction at all time points (from 0 to 7 hours) in the neoplastic transformation, we interrogated our dataset of 1004 differentially phosphorylated peptides using DAVID bioinformatics. We found at least 13 terms significantly enriched (Benjamini corrected *p*-value of below 0.05). These included enrichment for ErbB (HER) signalling pathway, mTOR signalling pathway, endometrial cancer, and MAPK signalling pathway amongst others (Figure 4.10A). To check which components of the HER signalling pathway are enriched, a schematic of the canonical KEGG pathway is shown in figure 4.10B. The red stars represent the proteins either activated or depleted in our system upon HER2 over expression at all time points. The data shows the homodimerisation between HER1-HER1 and HER2-HER2 partners, but also the heterodimerisation between HER1-HER2 family members. Interestingly, HER3 and HER4 remain inactive. A novel observation exhibits the neuregulin (NRG4) ligand itself is being activated. The predominant pathway that showed phosphorylation events was the MAPK signalling pathway, which enriched for SHC, SOS, RAF, ERK, and MYC. However, some proteins of other pathways were also enriched, including FAK, PAK, and the activation of AKT and p21, indicating HER2 is able to induce phosphorylation changes through many distinct pathways. These changes were only exhibited in the HER2 transduced cells as none of these changes were significant in the GFP-transduced MCF10A cells (Figure 4.10B).

Furthermore, to understand the biological significance of these results, the same phospho-peptides were subjected to ontology analysis (Figure 4.10C). Phospho-peptides associated with cell-cell adhesion were enriched similarly to the gene ontology (GO) terms we determined for each time point, indicating that this cellular process is very sensitive to alterations in the HER

signalling. The analysis also revealed the enrichment of a number of other biological processes GO terms associated with chromatin organisation, cell projection organisation, cell ageing, and regulation of signal transduction pathways amongst other terms. Lastly, the only disease significantly enriched was breast cancer.

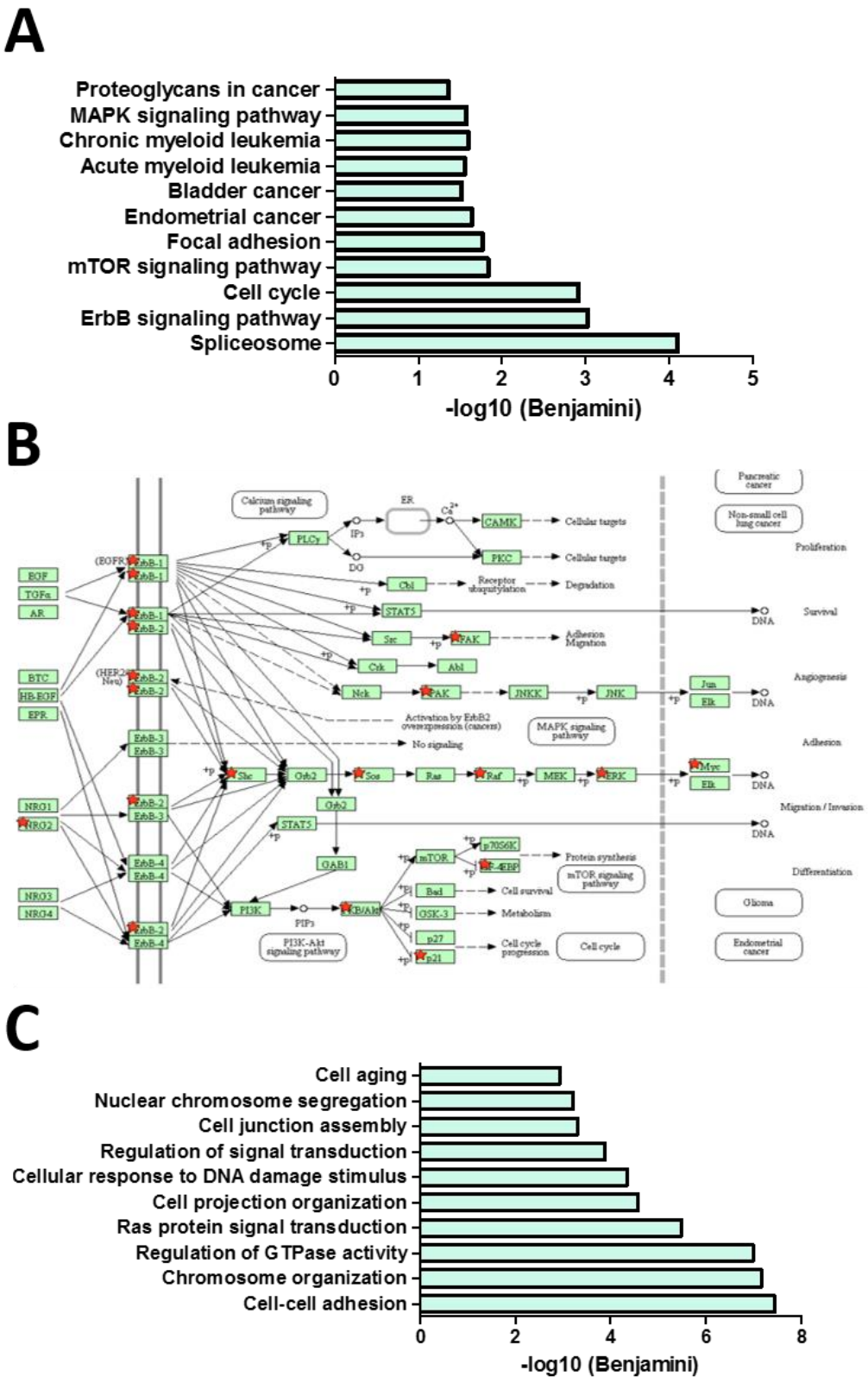


Figure 4.10: Signalling and biological function analysis of the early immediate changes in transformation. (A) Signalling pathway analysis using the DAVID KEGG PATHWAY tools of the

differentially phosphorylated events at all time points upon the HER2 protein induction is shown. (B) Selection of the canonical ErbB signalling pathway, shown the changes in our system indicated by the red stars. (C) Gene ontology analysis of the enrichment biological processes.

4.9 Time dependent changes upon HER2 over expression

To gain a better understanding of HER2 regulated time-dependent changes, we selected only those phospho-peptides that are changing in the 0.5 hour time point and then continuously maintained until the final 7 hour time point. We first applied a FDR corrected *p*-value of at least < 0.05 to focus on phospho-peptides that showed statistically significant differential regulation (up or down regulation) compared to DOX -ve cells. We found that by applying such a stringent threshold it would likely represent fewer but genuine phosphorylation events, thus overall only 57 phosphopeptides were differentially regulated, of which 32 phospho-peptides were marked as being down regulated and 25 hyper phosphorylated or up regulated. Some of the significant down regulated phospho-peptides included DNA methyltransferase (DNMT1 S387), AKT2 [T451], and ELF4 [S186]. One of the hyper phosphorylated peptides was LAP3 [R440], which is involved in the turnover of intracellular proteins, but the specific function for this phospho-site is not yet known and it has not been associated with breast lesions or HER2 positive breast cancer.

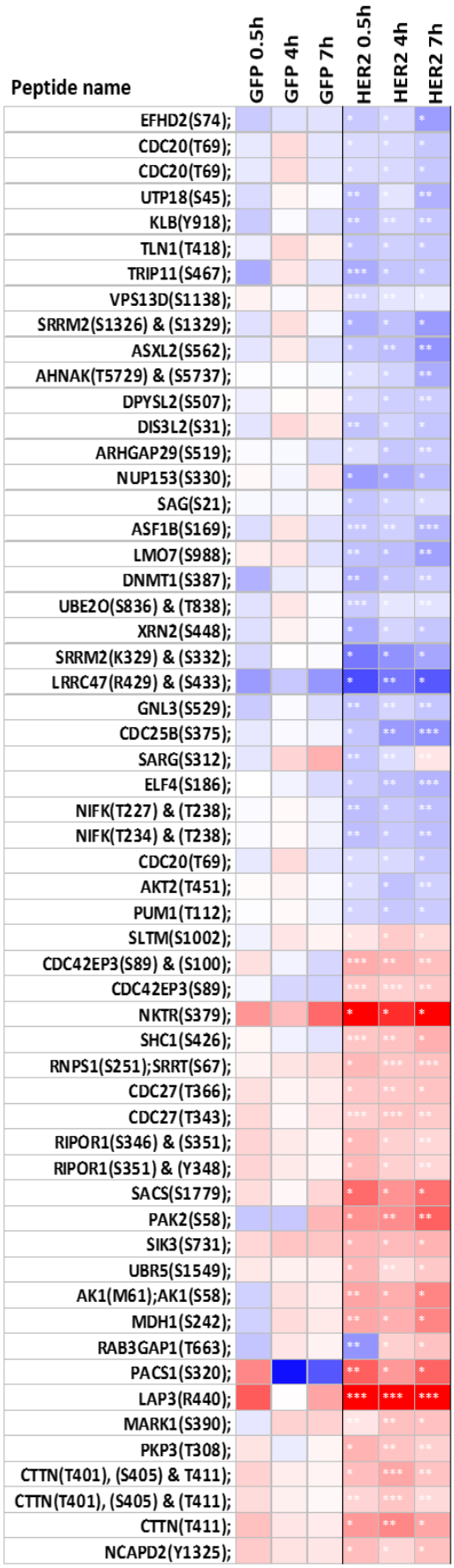

Peptide name
GFP 0.5h
GFP 4h
GFP 7h
HER2 0.5h
HER2 4h
HER2 7h
EFHD2(S74);
CDC20(T69);
CDC20(T69);
UTP18(S45);
KLB(Y918);
TLN1(T418);
TRIP11(S467);
VPS13D(S1138);
SRRM2(S1326) & (S1329);
ASXL2(S562);
AHNAK(T5729) & (S5737);
DPYSL2(S507);
DIS3L2(S31);
ARHGAP29(S519);
NUP153(S330);
SAG(S21);
ASF1B(S169);
LMO7(S988);
DNMT1(S387);
UBE2O(S836) & (T838);
XRN2(S448);
SRRM2(K329) & (S332);
LRRC47(R429) & (S433);
GNL3(S529);
CDC25B(S375);
SARG(S312);
ELF4(S186);
NIFK(T227) & (T238);
NIFK(T234) & (T238);
CDC20(T69);
AKT2(T451);
PUM1(T112);
SLTM(S1002);
CDC42EP3(S89) & (S100);
CDC42EP3(S89);
NKTR(S379);
SHC1(S426);
RNPS1(S251);SRRT(S67);
CDC27(T366);
CDC27(T343);
RIPOR1(S346) & (S351);
RIPOR1(S351) & (Y348);
SACS(S1779);
PAK2(S58);
SIK3(S731);
UBR5(S1549);
AK1(M61);AK1(S58);
MDH1(S242);
RAB3GAP1(T663);
PACS1(S320);
LAP3(R440);
MARK1(S390);
PKP3(T308);
CTTN(T401), (S405) & T411);
CTTN(T401), (S405) & (T411);
CTTN(T411);
NCAPD2(Y1325);

Figure 4.11. Time dependent phosphorylation events. Heatmap displaying differentially regulated phosphorylation changes that are time-dependent (significant) that occur in all time points analysed upon HER2 over expression but none of these changes are significant in the control cells. [* FDR corrected p-value of < 0.05, **FDR corrected p-value of < 0.001, *** FDR corrected p-value of < 0.001].

4.10 Activation of chromatin regulators and transcription factors

As we identified various biological processes enriched for proteins associated with chromosome organisation, nuclear chromosome segregation, sister chromatid segregation and other processes related chromatin, we wanted to assess if there were any transcription factors or regulators of chromatin in our dataset that have a molecular effect on transcription of genes that are significantly changing upon HER2 over expression that do not alter significantly in the GFP transduced cells. We identified 29 phospho-peptides that satisfy those conditions by checking their molecular function on PhosphoSitePlus. The alterations included the activation of NFkB, JUN, SIRT1, and SOX13 amongst other changes (Figure 4.11). It is interesting to note that the majority (72%) of these changes affecting the transcription factors/chromatin regulators incidentally occur at the later time points of 4 hours and 7 hours and the remaining 28% at 0.5 hours. This is in contrast to the activation of for example, the HER signalling pathway, in which the majority (70%) of the proteins become active at either 0.5 or 4 hours' time points. These changes were not significant in the GFP transduced cells.

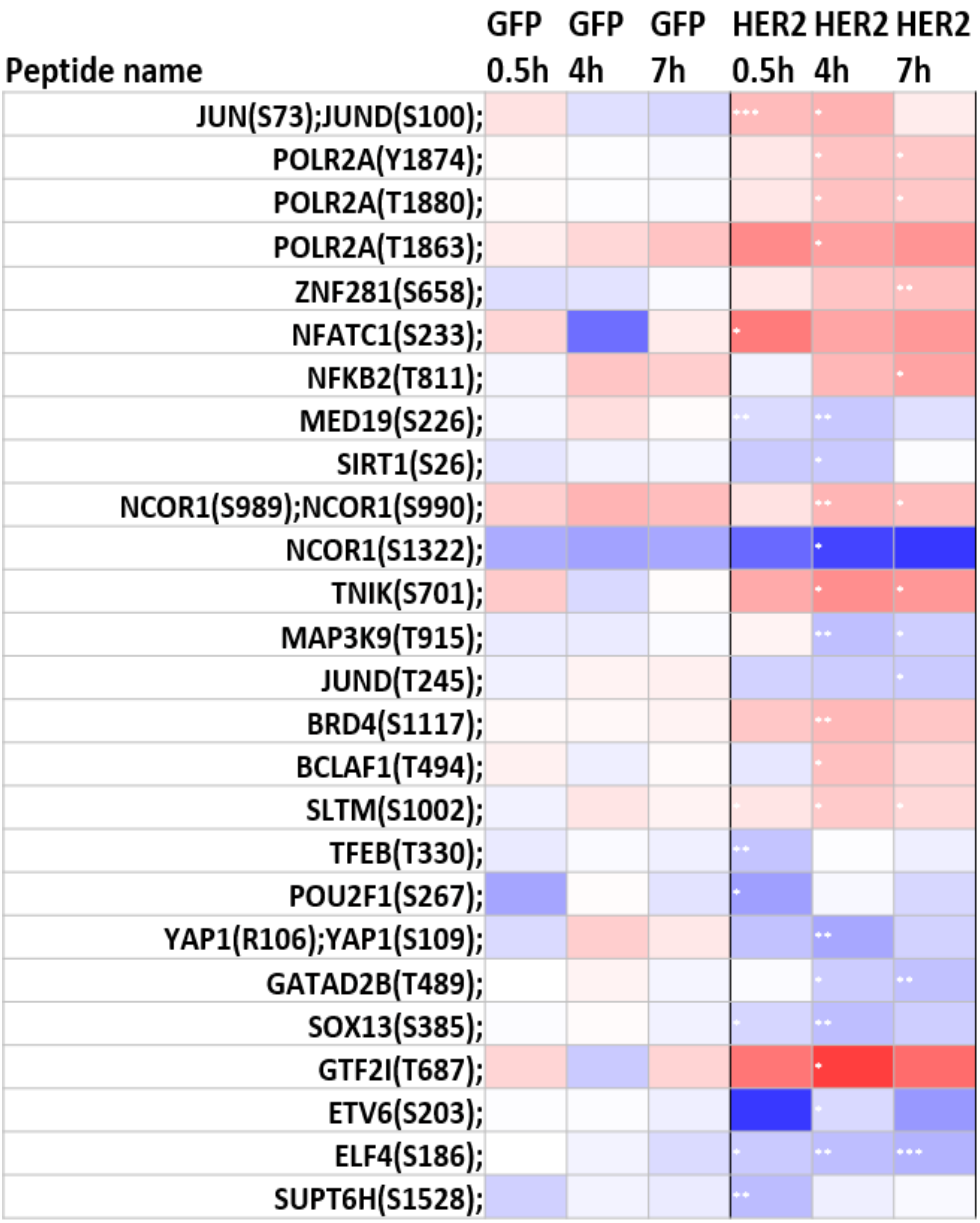


Figure 4.12: Identification of transcription factors and chromatin regulators. A list of transcription factors and chromatin regulators becoming differentially regulated upon HER2 expression in at least one time point is shown. [* FDR corrected p-value of < 0.05, **FDR corrected p-value of < 0.001, *** FDR corrected p-value of < 0.001].

Many proteins of the HER signalling pathway and others we have identified here are known to directly impact transcription factors, which can ultimately alter chromatin architecture (i.e. its accessibility or inaccessibility (255, 256)). Therefore, it would be extremely valuable to

understand the association between HER2 induced signalling changes and its effects on chromatin organisation, which ultimately plays an important role in transcription. To achieve this, we have performed ATAC-seq (Assay for Transposase-Accessible Chromatin using sequencing) analysis to interrogate the architectural chromatin alterations upon HER2 over expression and during the early stages of cellular transformation. Indeed, we have already shown that the components of the HER/MAPK signalling pathway are activated rapidly at 0.5 and 4 hours' time points, but the various chromatin regulators become activated at the later time points (which can impact chromatin organisation), indicating a series of events in a time-dependent manner that can ultimately alter chromatin state and contribute to transformation. Therefore, performing ATAC-seq alongside our phospho-proteomic data set will help us dissect the mechanism(s) by which HER2 induces transformation in an experimental setting and help understand the contribution of signalling and chromatin structural changes to transformation (Figure 4.12).

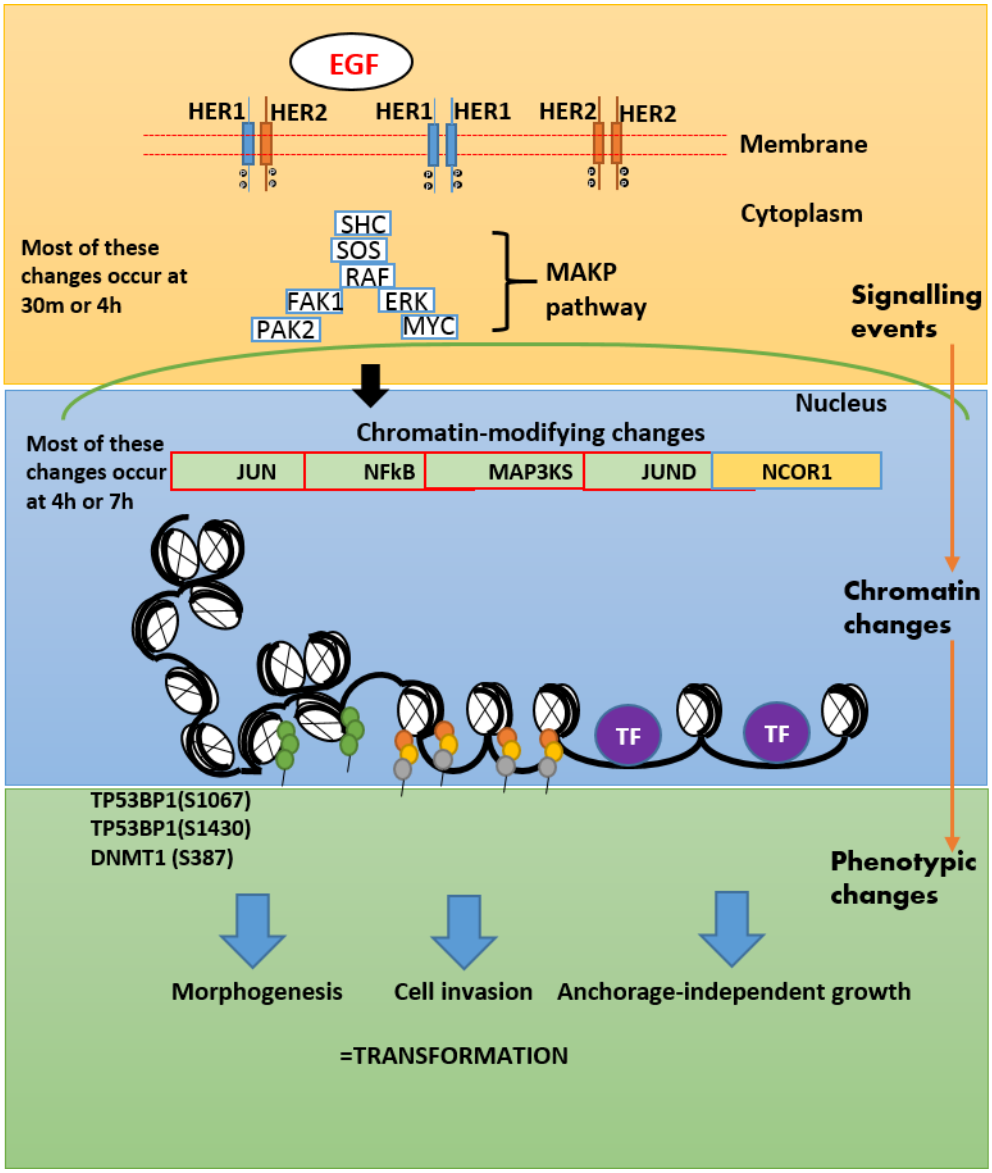


Figure 4.13: Potential mechanism of HER2 induced transformation in MCF10A cells. The phenotypic changes such as morphological alterations, high invasion potential, and anchorage-independent growth of HER2 over expressing MCF10A cells will most likely accompany changes at the molecular level. Here, we hypothesise that the signalling changes induced by HER2 over expression will result in gross chromatin organisational changes. Those changes may include accessible regions of the chromatin at where various proto-oncogene reside (potentially activating them) and inaccessible regions of chromatin may be enriched where tumour suppressor genes reside (potentially inactivating them), contributing the transformative phenotypes we have observed with various functional assays performed.

4.11 Discussion

In this investigation, we have carried out an in-depth characterisation of the phospho-proteome of early-immediate signalling changes in the process of cellular transformation. This study provides a detailed picture of the downstream consequences (at the phospho-peptide level), of neoplastic transformation induced by the activation of a proto-oncogene. The phospho-proteomic changes in MCF10A cells upon HER2 protein induction and neoplastic transformation of our dataset is in contrast with the other studies that have examined the effects of HER2 activation in transformed cells. This is because those systems study already transformed cells, or examine the effects of mutations at long time points, when presumably other genetic and/or epigenetic aberrations have taken place. To achieve the aim of dissecting signalling changes at the very outset of transformation and HER2 expression, our HER2 inducible MCF10A system provides obvious advantages. The low levels of HER2 activation at early time points may closely mimic, to a partial extent, the early signalling changes occurring in HER2 positive breast cancer patients. The signalling changes at global scale of low level HER2 induction has not been performed to date.

We have previously shown by western blotting that HER2 protein levels increase in a time-dependent manner by the addition of 1µg/ml of doxycycline, and that the protein levels fully saturate after 12 hours in doxycycline containing media. However, our phospho-proteomic analysis was performed at the final time point of 7 hours, by which the HER2 expression would not be fully induced. Therefore, our phospho-proteomic screen is constrained to the acute effects of HER2 activation, since HER2 is not fully expressed; as a result, we have not measured the signalling activity of a fully induced HER2 protein.

Furthermore, MCF10A cells require the addition of ligands to survive, as they induce signalling to allow the cells to divide and proliferate. Our simple model requires the exogenous addition of a single ligand, which is the epidermal growth factor (EGF). This causes the

heterodimerisation between HER1-HER2 or homodimerisation between HER1-HER1, and the non-ligand independent homodimerisation between HER2-HER2 receptors. The deprivation of additional available ligands in our model results in lack of dimerisation between the other HER2 binding partners and family members such as HER3 and HER4. This system is therefore restricted in characterising the signalling changes upon just three combinations of dimerisation. However, in mammalian cells, 12 different ligands have been identified that can induce signalling which would not be reproduced by this system, suggesting the lack of complexity in this model to recapitulate the phosphorylation events occurring in HER2 positive breast cancer patients.

Moreover, despite the ectopic over expression of HER2 in cells at early time points, we identified less than 2% of phosphotyrosine peptides even though a large number of tyrosine kinases are present in the genome and it is known that tyrosine phosphorylation occurs earlier on compared to phoshotheronine and phosphoserine. This may be attributed to the technical aspect of the experimental setting, such as the use of titanium dioxide (TiO_2), which is known to bind to tyrosine phosphorylations less favourably compared to serine and threonine modifications, which may explain the lower enrichment of phosphotyrosines (257). There are several known biological reasons for the relatively low phosphotyrosine sites identified. Firstly, phosphotyrosines become activated only during specific circumstances (258). Secondly, phosphotyrosines have a short half-life, due to high levels of activity of phosphotyrosine phosphatases or PTPs, unless the phosphotyrosines are protected by the PTP and SH2 domains (259). Lastly, since the number of phosphopeptides observed in our system is not high, this may correlate to the fewer phosphotyrosines identified. This is because it is known that tyrosine phosphorylation occurs on proteins with high abundance. Nevertheless, it appears that there is an inherent bias due the method employed for identifying fewer phosphotyrosines compared to threonine and serine modifications, but is difficult to dissect if that is due to a biological effect, which might be vital or a technical caveat. If technical, then it

means that many important phosphopeptides were not identified that may be critical for the process of transformation (257). However, since the focus of our study was not only to identify phosphotyrosines, but also serine and threonine phosphorylations, the relatively low enrichment of tyrosine phosphorylations did not pose a major concern.

Chapter 5

Assessing global chromatin accessibility alterations in HER2 induced transformation

5.1 Introduction

The opening of chromatin that is accessible for binding by transcription factors is correlated with biological activity at a specific genomic region (133). The phenotypic changes induced upon HER2 over expression in our model during cellular transformation are likely to be driven by alterations in the gene expression, which are themselves governed by the accessibility and inaccessibility of chromatin architecture. There are reports that have documented the chromatin landscape differences between normal and transformed cells, and have begun to define the chromatin state of cancer cell lines (132, 136). However, the specific changes in chromatin state driving the transition from normal to transformed cells are still remaining to be explored. More specifically, the over expression of a cell surface receptor and the ensuing activation of a plethora of signalling networks and its effects on chromatin landscape is not yet elucidated. Here, we attempt to understand the impact of signalling events on the chromatin state, and how that contributes to cellular transformation.

5.2 ATAC-seq library preparation – attacking the chromatin

To probe for DNA accessibility with a sensitive and fast alternative to other methods such as DNase-seq or Mnase-seq, we employed ATAC-seq with next generation sequencing. This method uses a hyperactive Tn5 transposase enzyme that inserts sequencing adapters to random DNA sequences, but only in accessible regions of the chromatin (260). In an attempt to be partially physiologically relevant to the *in vivo* microenvironment, we prepared DNA libraries for ATAC-seq from 50,000 cells from acini grown in 3D cell cultures in contrast to cells growing in 2D cell culture. However, the sensitivity of library preparation for ATAC seq from acini recovered from 3D cell culture proved challenging initially, because of the difficulty in isolating cells that are homogenous population (intact and free from any debris) and also that they are the correct number, as the ratio of cell number to transposase enzyme is critical for success library preparation.

50,000 cells were thus recovered from 3D cell culture and libraries prepared as per protocol (see 2.13). DNA library profiles were analysed by a bio-analyser, and a representative profile is shown in figure 5.1A. At the beginning, it appeared that we were transposing fewer than 50,000 cells despite counting with a haemocytometer twice and using the average of those values, as the fragments appears to be “over-transposed” with a preponderance of shorter fragments lacking the periodicity as is generally expected in ATAC-seq libraries. To overcome this, we switched to counting the cells with Luna automated cell counter and performed the library preparation. The library profile was then as expected for ATAC-seq, and the rest of the libraries were prepared in the same manner (Figure 5.1B).

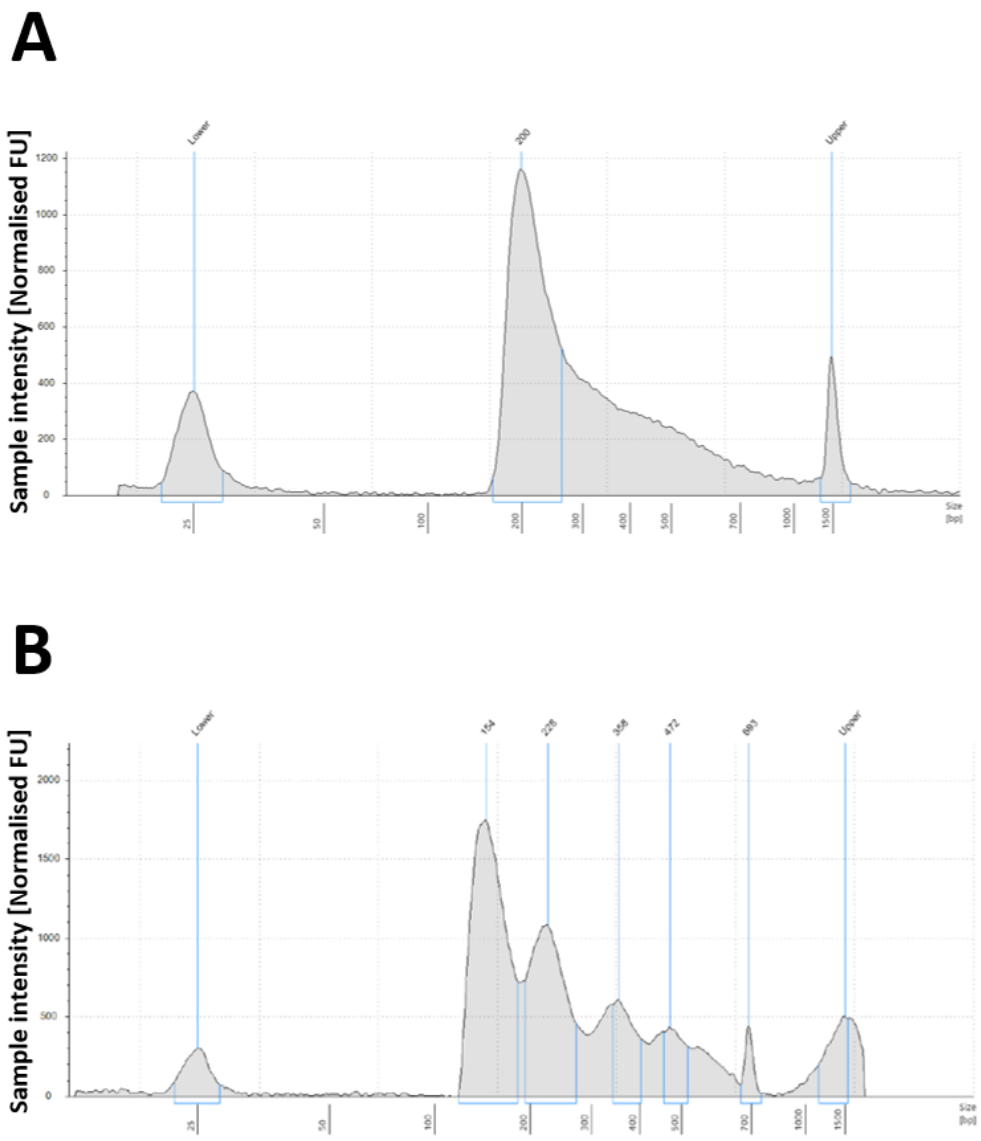


Figure 5.1: ATAC-seq library profiles. (A) Library profile for incorrect number of cells, which has been over-transposed. (B) The correct library profile, with periodicity of Tn5 cutting chromatin at different fragment lengths.

5.3 Quality metrics and validation

To identify genome-wide changes in the chromatin accessibility during cellular transformation, we analysed 6 different time points. These were 0 hour, 1 hour, 4 hours, 7 hours, 24 hours, and 48 hours in two biological replicates, encompassing early and late time points upon HER2 induction (DOX +ve cells) and their control counterparts, DOX -ve cells. Here, we will confine

our analysis to the early time points of transformation only, taking into consideration just the first three time points. These are 0 hours DOX -ve, 1 hour DOX +ve, and 4 hours DOX +ve time points. For these sample, we obtained on average 57.94% mappability to the human genome. We first assessed the fragment length distribution which has 124 base pairs (bp) adapter sequence removed, a representative plot is shown in figure 5.2A. This shows more than half of the reads tend to be shorter than 150 bp, which are sub-nucleosomal and approximately half of the reads appear to be larger than 150 bp. This as it has been previously shown, is an expected profile of ATAC-seq library (260).

To ensure that the biological replicates are reproducible we clustered the samples based on Pearson correlation coefficient. The correlation coefficient indicates how strong the relationship between two samples is, which consists of numbers from -1 to 1 (where 1 indicates perfect correlation and -1 indicates perfect anti-correlation). This method is used to determine if different samples can be separated. For example, generally it would be expected that samples from two biological replicates of the same condition would have greater similarity between them, compared to samples from two different conditions. In our case, it appears that the biological replicates are more similar to each other than samples collected at different times within the same condition (Figure 5.2B). The PCA plot shows that DOX +ve sample cluster together and broadly there is a clear separation between the DOX -ve and DOX +ve samples (Figure 5.2C).

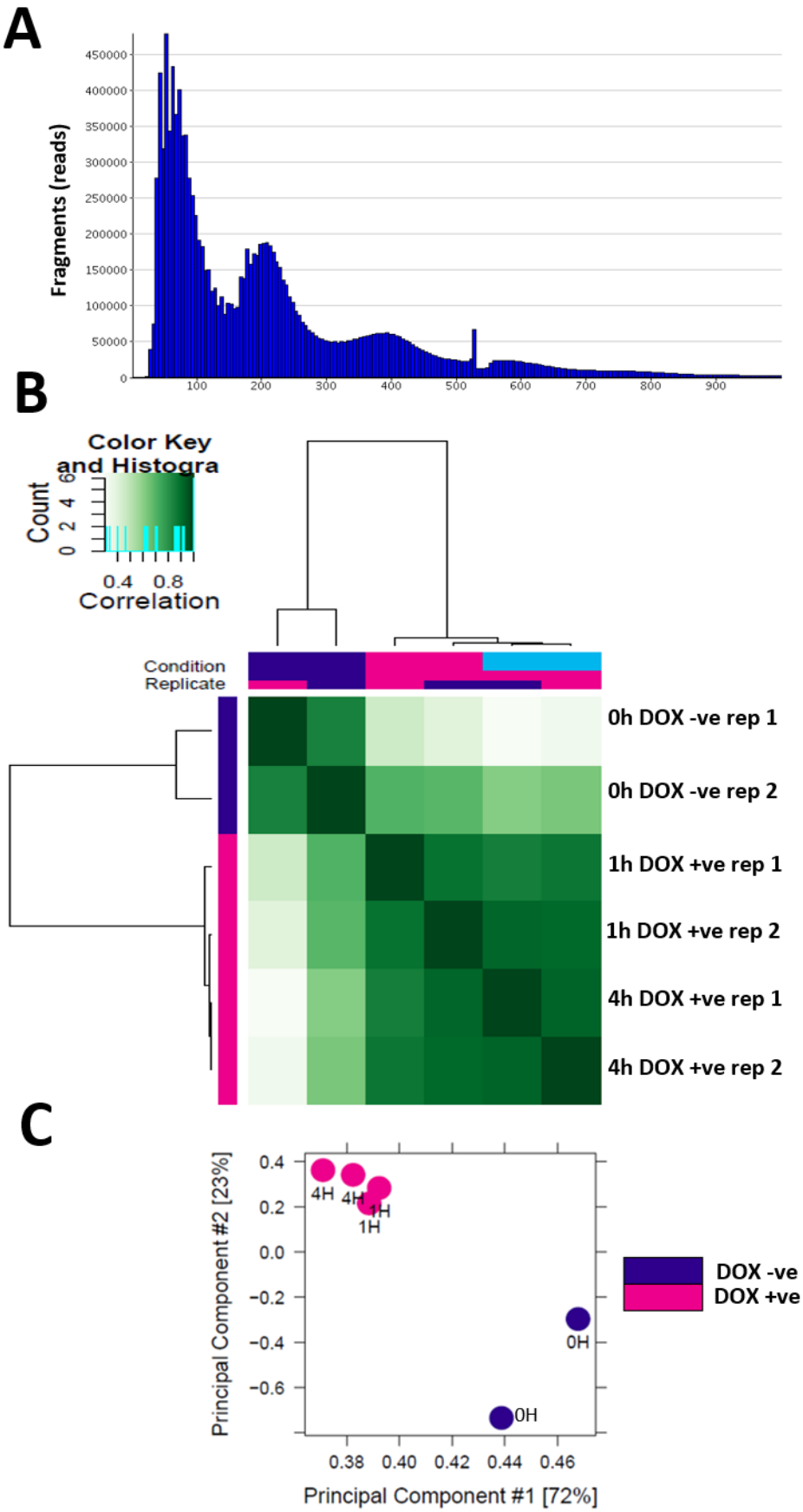

A
Fragments (reads)
450000
400000
350000
300000
250000
200000
150000
100000
50000
0
100
200
300
400
500
600
700
800
900
B
Color Key
and Histogra
Count
0 2 4 6
0.4 0.8
Correlation
Condition
Replicate
0h DOX -ve rep 1
0h DOX -ve rep 2
1h DOX +ve rep 1
1h DOX +ve rep 2
4h DOX +ve rep 1
4h DOX +ve rep 2
C
Principal Component #2 [23%]
0.4
0.2
0.0
−0.2
−0.4
−0.6
4H
4H
1H
1H
0H
0H
0.38 0.40 0.42 0.44 0.46
Principal Component #1 [72%]
DOX -ve
DOX +ve

Figure 5.2: Fragment size and evaluation of reproducibility. (A) Insert size as determined by high throughput sequencing, adapter sequences are an additional ~124 base pairs. (B) Correlation heatmap using peak caller score data across all the time points in biological replicates. (C) PCA plot showing the clustering between DOX -ve and DOX +ve samples and their biological replicates.

To visualise the enrichment ATAC-seq signal over specific target regions, we plotted heatmaps of the signal coverage between the two biological replicates. The y-axis of the heatmap shows the regions of accessible chromatin i.e. peaks. The x-axis shows the read counts were "centered" on the center of each peak region, which were extended to include 1000 bp of upstream of each peak start and 1000 bp downstream of each peak end. This simple peak calling with default parameters generated consistent regions between the biological replicates (Figure 5.3).

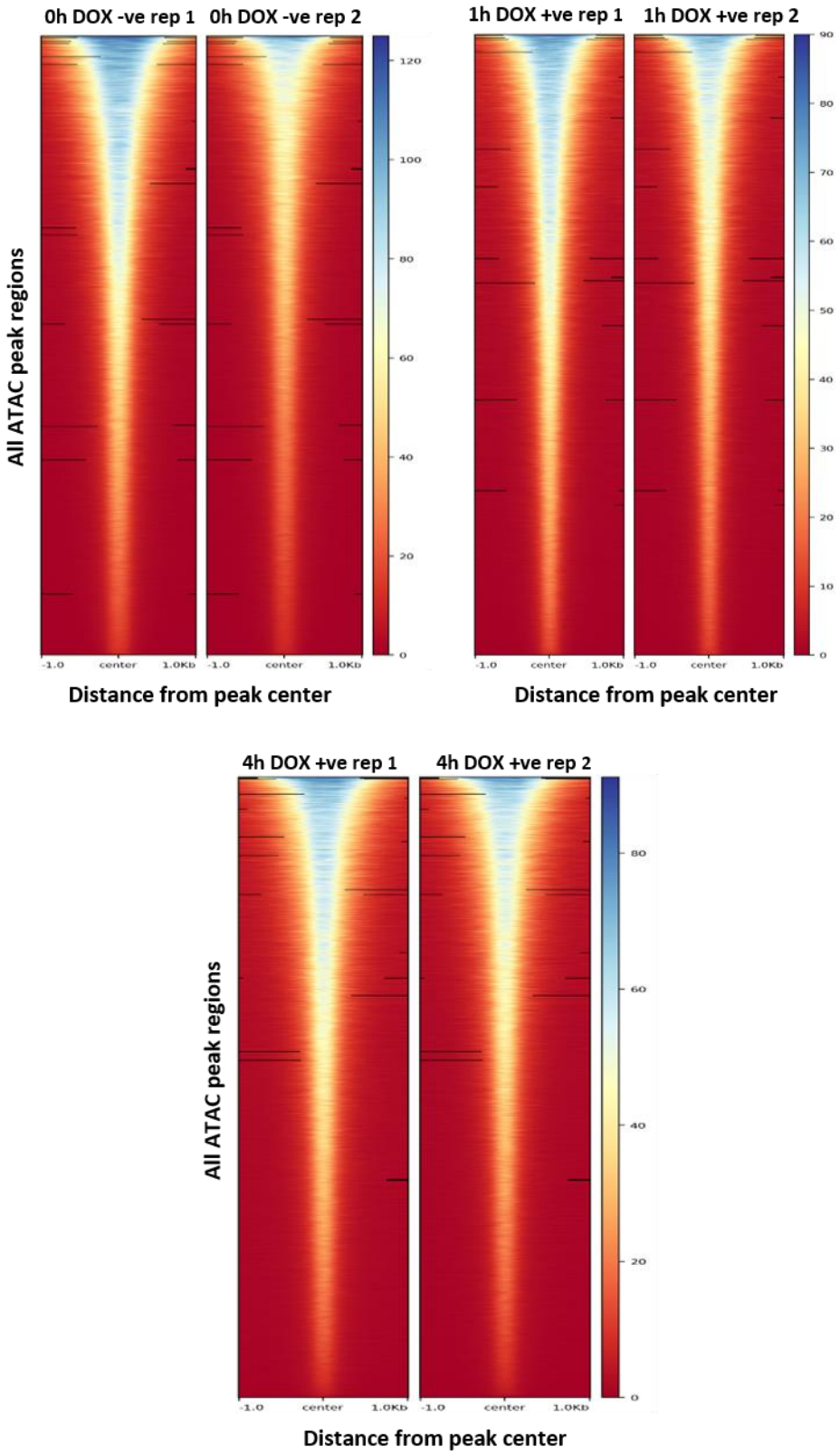


Figure 5.3: ATAC-seq quality metrics. Heatmaps showing normalised read coverage for ATAC-seq enrichment signal ± 1000 base pairs from the center of the peak for the biological replicates. The scale shows highly accessible regions in blue and inaccessible regions in red, based on the fold-change value

from each peak. Each row represents one peak. The heatmaps were created using a matrix, which requires BigWig files and a BED file. The BigWig file is an indexed, compressed and binary file of the genome-wide signal data for various types of calculations. The BED file is a text file format, containing the chromosome name, the chromosome start position and end position. Therefore, the matrix used to create the heatmaps are all ordered in the same way for the different samples. They rows are ordered by the chromosome name, and the start position.

5.4 Overview of chromatin accessibility landscape

Next, we intersected the two biological replicates and measured the total number of peaks (open chromatin regions) in each time point using default MACS2 settings and without applying filters or any statistical power. In total we identified dynamic DNA access (71,699 peaks) at the 0 hour DOX -ve time point, 73,457 peaks in the 1 hour DOX +ve time point, and 74,375 peaks in the 4 hours DOX +ve point. The majority of the peaks were identified across the samples, representing a total of 61,162 shared peaks. However, a number of them were also unique to each point (Figure 5.4). It appears that chromatin accessibility between the three samples is approximately the same, potentially reflecting that HER2 overexpression does not cause large scale changes in chromatin accessibility.

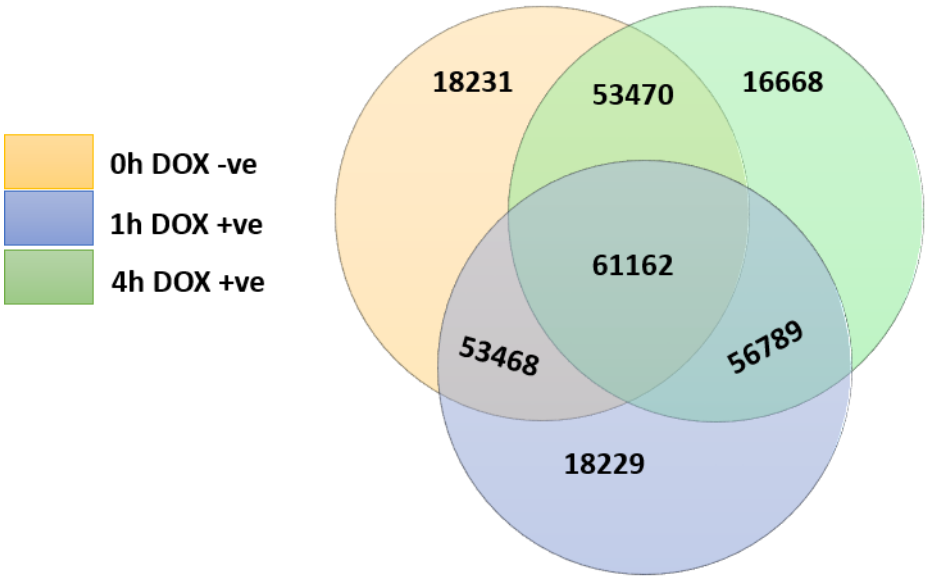


Figure 5.4: Quantification of accessible chromatin. Venn diagram shows the peaks that overlap and those that are unique to the specific time point. All the samples were downsampled (normalised) to 25 million reads. Peaks were called by MACS2 and the different number of peaks were counted by samtools.

The data revealed categories of peaks that are either unique to, or are overlapping between the different time points. For instance, a peak associated with SHNG5 is overlapping between all the time points and is enriched in all the biological replicates, whereas a peaks associated with FAM19A2 and KRT17 are only unique to the time points with HER2 expression, in 1 and 4 hours DOX +ve time points, with background noise peaks for the DOX -ve replicates (Figure 5.5).

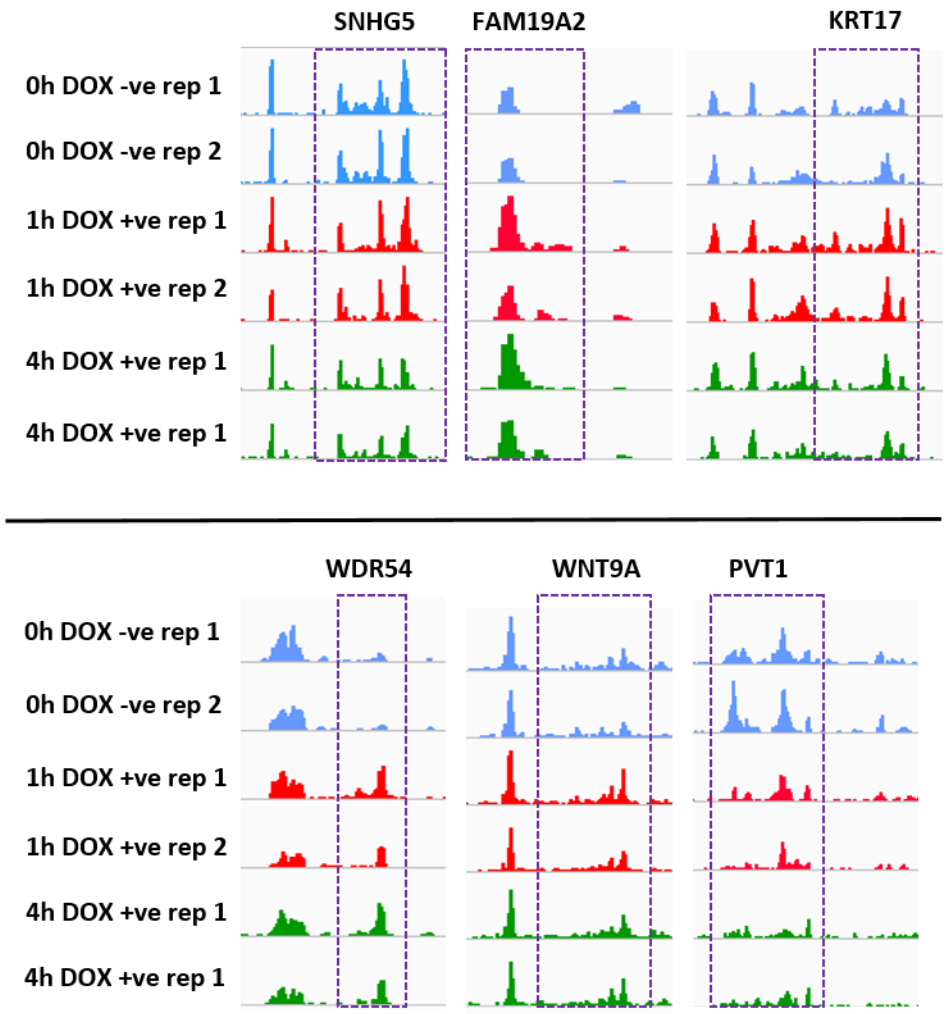
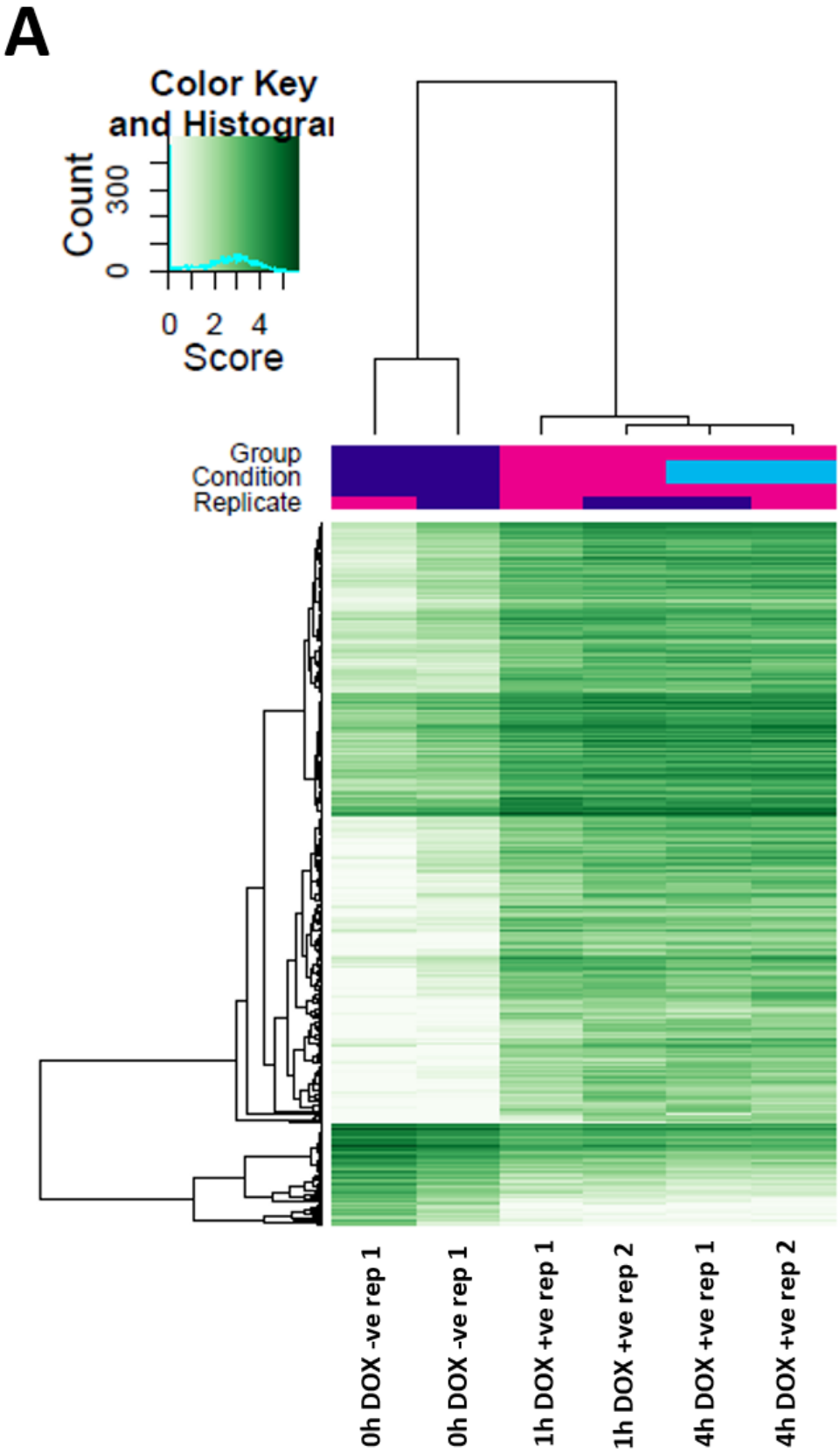


Figure 5.5: Visualising the peaks and associated genes. Integrative Genomics Viewer (IGV) for ATAC-seq signal for the indicated genes across the biological replicates and time points.

To visualise the significant and differential chromatin accessible peaks between the different time points and the biological replicates, we generated a heatmap with all the regions that are changing (Figure 5.6A). These plots were made after statistically significant peaks were selected by setting the threshold with an FDR corrected p-value of 9 and fold-change of at least 7. There is a high degree of similarity between the biological replicates and there is a distinct pattern in the accessible chromatin from 0h time point to 1 hour time point. The heatmap also shows the clustering between the DOX -ve samples and the DOX +ve samples. To identify which data points are identified as being differentially chromatin accessible regions, we plotted an MA plot (log_2 fold change vs. mean average) to visualise changes in chromatin

accessibility for all peaks. In total, we identified 22,296 differentially accessible changes (Figure 5.6B). The MA plot takes into account all the changes between DOX -ve and DOX +ve samples at all the time points (0 hour, 1 hour, and 4 hours). Without taking into account the time-specific changes, there appears to be a decrease in the global chromatin accessibility, since more of accessible regions have decrease intensities (< -0.5 fold change) (Figure 5.6B).


A
Color Key
and Histogra
Count
300
0
0 2 4
Score
Group
Condition
Replicate
0h DOX -ve rep 1
0h DOX -ve rep 1
1h DOX +ve rep 1
1h DOX +ve rep 2
4h DOX +ve rep 1
4h DOX +ve rep 2

B

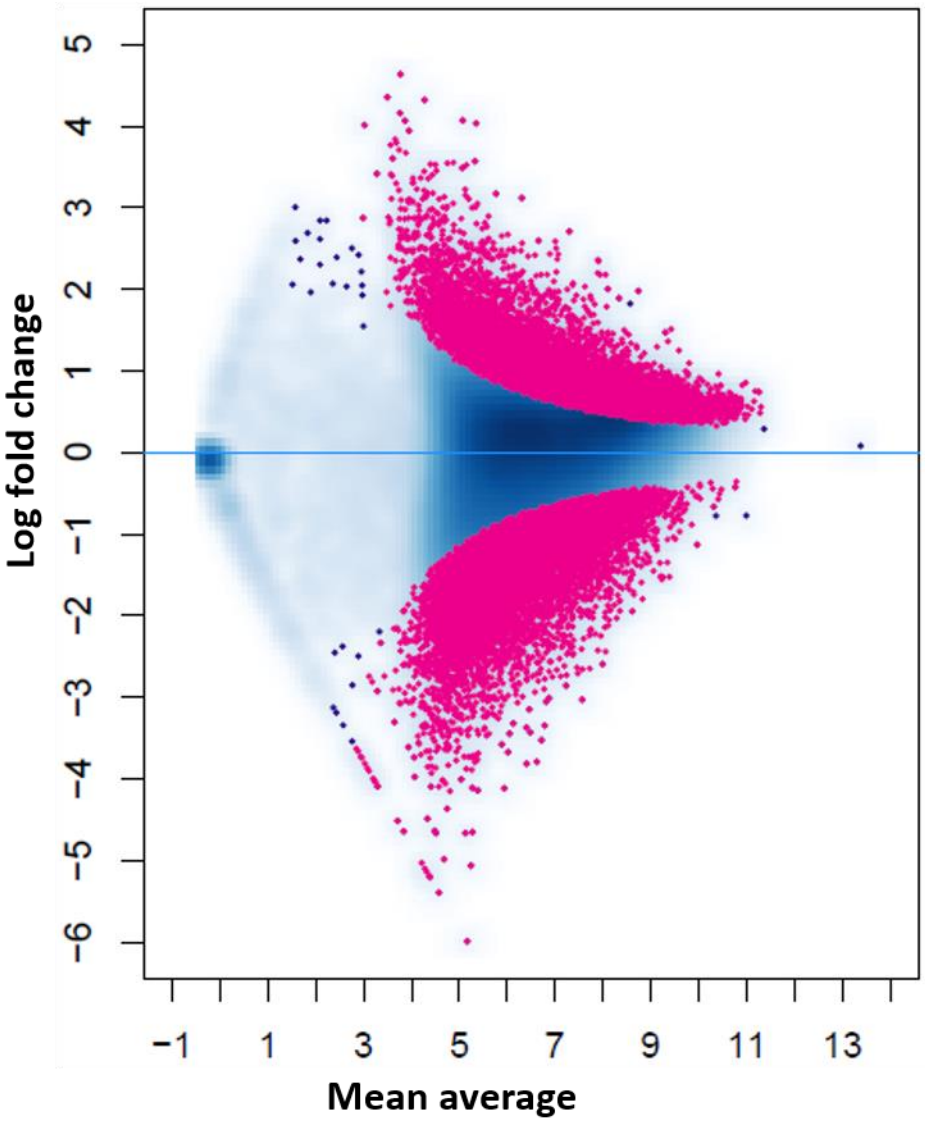


Figure 5.6: Profiling of chromatin accessibility at early time points upon HER2 over expression. (A) Heatmap displaying relative chromatin accessibility. (B) Sites that have differentially accessibility are coloured in pink. The differentially accessible regions have an absolute log fold difference of at least 0.5.

5.5 Discussion

We provide a non-comprehensive and simple initial quality metrics and visualisation of our ATAC-seq data for some of our samples. We performed some very basic tests to check if our dataset that has been aligned to human genome meet our expectations.

A technical aspect of the this ATAC-seq experimental setting included the isolation of acini from a 3D matrix at 4°C using the cell recovery solution, to depolymerise the matrigel/collagen mixture. Although there is no direct evidence to suggest that such a recovery method would impact the chromatin dynamics, the cellular microenvironment between the physiological growth conditions at 37 °C and the conditions during the detachment of cells from the matrix are quite different. It has been previously shown that the phosphoproteome of cells recovered with the recovery solution has significant impact on the phosphoproteomic status of cells (261). Therefore, it is conceivable to think that the use of recovery solution may also have an effect on the chromatin landscape.

Furthermore, we used DOX -ve cells as a parental control for ATAC-seq analysis, which does not have the addition of dox that is added to the DOX +ve cells for HER2 protein induction. It is worth bearing in mind that addition of dox may induce chromatin changes that are not associated with HER2 expression. However, there are reports of using dox as an inducing agent where a separate control for dox was not performed (262, 263), whereas others have included a dox control (264).

Finally, there is ongoing comprehensive analysis of this dataset across all the time points to help understand the impact of HER2 induced transformation on the chromatin architecture, with a more specific aim of understanding the effect of HER2 signalling on DNA accessibility.

To achieve this we anticipate to address the following:

1. It will be important to address whether HER2 overexpression creates a chromatin accessibility pattern early on after induction, a pattern that is maintained throughout the subsequent time points; or is it the case that differential chromatin regions (DCRs) are dynamic and time-dependent (i.e specific to time points).
2. One way of exploring the implication of signalling on chromatin changes is to integrate the phosphoproteomic dataset with ATAC-seq data. Accessible chromatin have peaks at specific genomic regions and these can be used to identify motifs for transcription factor binding. Transcription factors found in the phosphoproteomic data that correspond to these genomic motifs could be targets for further investigation as they may be involved in regulating chromatin architecture at these regions.
3. Since transcriptomic changes (chromatin changes) are very closely related to epigenetic changes, it would be useful to perform RNA-seq and DNA methylation analysis (or use available datasets) to identify chromatin accessibility in differentially methylated regions and to correlate the accessible peaks with gene expression by RNA-seq. This will give us a combined dataset that could be explored from different angles at high temporal resolution. In fact, our collaborators are performing single-cell RNA-seq experiments upon HER2 induction and during the transformation process with the aim of mapping the transcriptional process of HER2 induced transformation and the heterogeneity of the process.
4. Since the morphological changes take place early upon HER2 induction in our model system, it would be interesting to see if the chromatin accessibility changes that occur in the early time points play a driving role in the morphological changes we observed, or whether these chromatin changes are independent, or facultative, of the process of transformation.

Chapter 6

Investigating HER2 induced reprogramming associated heterogeneity

6.1 Introduction

Breast cancer can originate from different cells in the differentiation hierarchy, and can present different survival outcomes, mutational landscapes, and have distinct biological and clinical phenotypes (265-267); hence it can be categorised into several distinct subtypes based on the genetic and histopathological signatures. An example of this is the classification proposed by Perou *et al*. using microarrays (148), which has led to the formulation of five defined intrinsic subtypes, namely luminal A, luminal B, normal-like, HER2 enriched, and basal-like. The newly diagnosed breast cancers can now be designated to one of these subtypes based on the gene expression patterns of the PAM50, which are the 50 informative genes (148). However, the model of somatic cells acquiring mutations sequentially may be overly simplistic, and the concept of breast cancer stem cells has gained significant attention recently (268). These are thought to reside within the basal compartment of the gland because they share gene expression profiles and cell surface with the basal cells (269). The heterogeneity in cancer incidence, patient prognosis and patient response to therapies can also be ascribed to committed cells in the mammary compartment acquiring a stem cell-like phenotype during breast tumourigenesis (270-272).

Aberrant signalling events (273), induction of EMT transition (274), mutations in genes (275), and oncogene over expression (276) can induce cells to undergo tumour-reprogramming processes and enrich for markers known to be active in stem cells. The acquisition of the stem-like phenotype is associated with higher transformational potential (277, 278), leading to more aggressive forms of cancers because of their ability to self-renew (279, 280). The reason for

this is that it allows stem cells to produce a large number of progeny cells, thus increasing the probability of cells to acquire further genetic and/or epigenetic aberrations.

Our first observation was that in the *in vitro* transformation assays (measuring anchorage-independent growth of cells) of HER2-induced MCF10A cells, only a small fraction of cells plated out of the total population were able to form colonies. This could have been caused by cell death that can occur when cells are placed in relatively harsh conditions in this assay. Nevertheless, we also considered that perhaps not all cells have the potential to form colonies, and that upon induction of HER2 overexpression, a subset of cells would acquire markers associated with breast stem cells, which would increase their ability to form colonies. We therefore investigated the "stemness" of MCF10A cells upon HER2 protein induction and its control counterparts and hypothesised that the sub-population of cells with enrichment of stem-like markers will exhibit a higher transformative potential compared to bulk population or those that have non-stem like markers.

6.2 Identification of stem cell markers upon HER2 protein induction

To investigate if HER2 protein over expression induces reprogramming-associated heterogeneity in early cellular transformation, we tested the expression of proteins associated with breast stem-like phenotype. As a starting reference into identifying possible stem cell proteins that may be differentially expressed in DOX +ve cells compared to DOX -ve cells, we explored existing literature and investigated the cell surface markers proposed in the mammary epithelial cell hierarchy (154), as well as markers associated with embryonic stem cells, and cancer stem cells. We induced HER2 over expression for 72 hours and used DOX -ve parental population as control. Firstly, we began by performing immunofluorescence analysis to check for the expression of stem markers such as SOX2 (SRY-Box 2) and KLF4 (kruppel-like factor 4), which are enriched in pluripotent stem-like cells, and the expression of MUC1 (CD227), which is depleted in breast stem cells (154). We found that there was no difference in

the protein expression of SOX2 and KLF4 between DOX -ve and DOX +ve cells, with both cell types exhibiting homogenous and depleted levels of the cell surface proteins (Figure 6.1). Therefore, SOX2 and KLF4 were disqualified from our panel of markers for identifying the stemness of cells due the lack of differential expression between DOX -ve and DOX +ve cells. However, we observed heterogeneous expression of the MUC1 protein (Figure 6.1). MUC1 is a type I transmembrane, which is normally expressed at low levels in the luminal epithelial cells of the mammary gland, and its expression is low or negative in normal breast stem cells (154, 281, 282). DOX -ve cells showed no variability in its protein expression as most cells were expressing similar levels of MUC1. On the other hand, DOX +ve cells (72 hours after induction of HER2) exhibited heterogeneous expression of MUC1, with some cells being negative for MUC1 expression or below the detection threshold of immunofluorescence analysis.

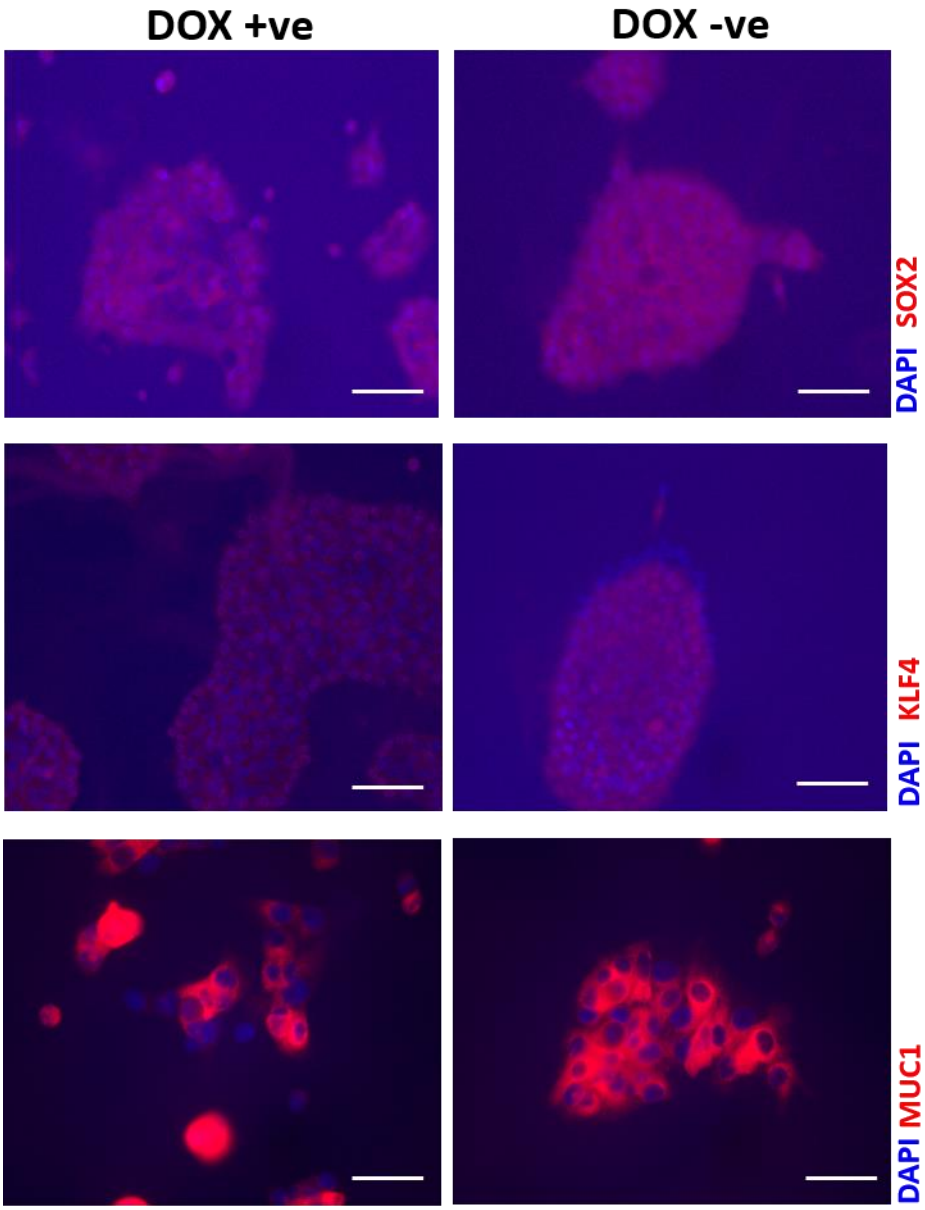
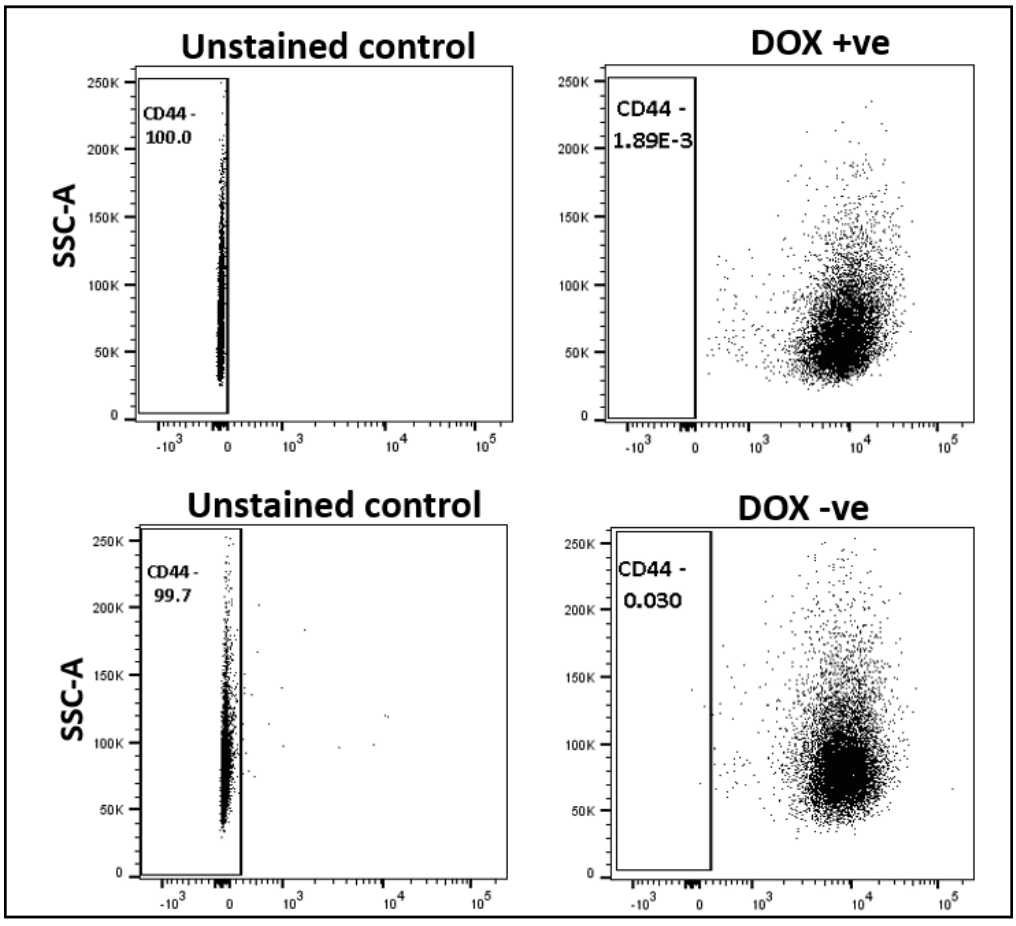


Figure 6.1: Investigating the acquisition of stem-like phenotypic features. DOX +ve and DOX -ve cells were grown for 3 days, fixed and subjected to staining by the indicated antibodies to stain for stem/progenitor cells using immunofluorescence assay. DAPI was used as a nuclear stain. Magnification: 20X for SOX2 and KLF4 images and 40X for MUC1 images. Scale bars represent 50µm in MUC1 images and 100µm for KLF4 and SOX2 images.

We moved on from using immunofluorescence analysis, which gives us a static image of protein expression, to flow cytometry to quantitatively measure the protein abundance. We

further investigated the protein expression of CD44 (cluster of differentiation 44) and CD49F (α6-Integrin subunit), both of which are highly expressed in mammary stem cells (283). Mammary cancer stem cells have been previously isolated by high expression of CD44 alongside CD24 -ve and Lin -ve markers (155). The co-expression of CD49F +ve with EpCAM -ve (epithelial cellular adhesion molecule) expression have also been used as prognostic markers for breast cancer (284). As with the previous experiment, HER2 expression was maintained for 72 hours, and flow cytometry analysis for the two proteins performed. Interestingly, flow cytometry analysis confirmed the high expression of CD44 and CD49F in both DOX -ve and DOX +ve cells in two independent biological replicates (Figure 6.2). For this reason, we also disqualified these markers of heterogeneity from our system, despite them being detected at high levels in both cell types as is seen in breast stem cells.


Unstained control
DOX +ve
SSC-A
CD44 -
100.0
CD44 -
1.89E-3
Unstained control
DOX -ve
SSC-A
CD44 -
99.7
CD44 -
0.030

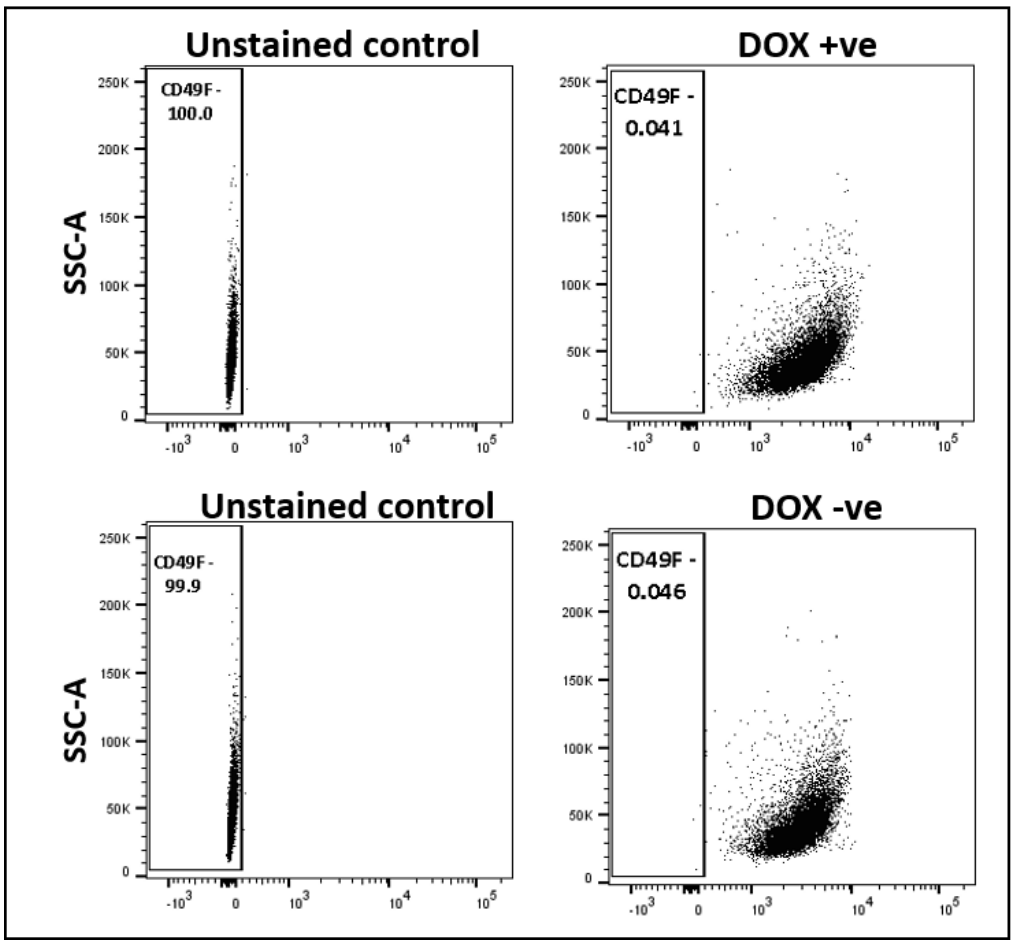
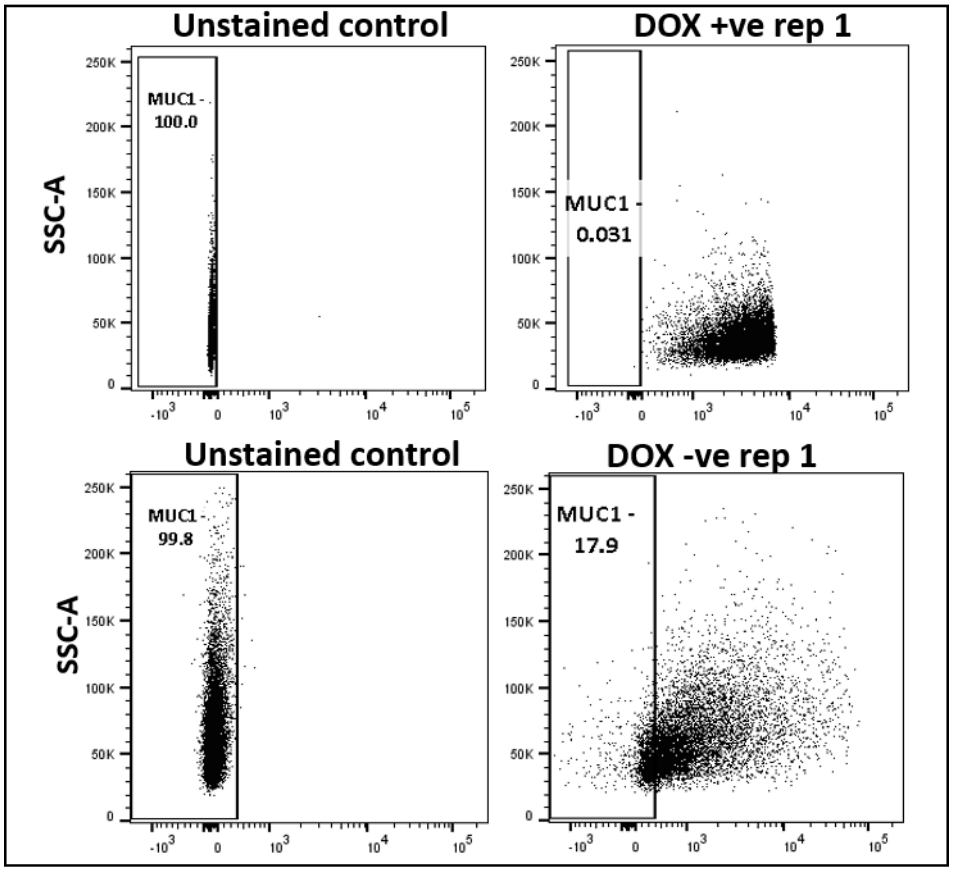

Unstained control
DOX +ve
SSC-A
CD49F -
100.0
CD49F -
0.041
Unstained control
DOX -ve
SSC-A
CD49F -
99.9
CD49F -
0.046

Figure 6.2: Determining the expression of stem markers. Flow cytometric analysis of single stains of CD44 and CD49F in in two independent biological replicates of DOX -ve and DOX +ve cells. The gating was based on the negative control.

6.3 Characterising HER2 induced MCF10A cells for stemness

We continued our investigation to identify stem markers that may be heterogeneously expressed upon HER2-induced transformation, and have differential expression between DOX +ve and DOX -ve cells. It has been shown that decreased expression of MUC1 and the EpCAM is associated with the most primitive cells in the mammary epithelial stem cell hierarchy (154).

We confirmed our previous observation by flow cytometry that MUC1 has decreased expression in DOX +ve cells compared to DOX -ve cells (Figure 6.3).


Unstained control
DOX +ve rep 1
SSC-A
MUC1 - 100.0
MUC1 - 0.031
Unstained control
DOX -ve rep 1
SSC-A
MUC1 - 99.8
MUC1 - 17.9

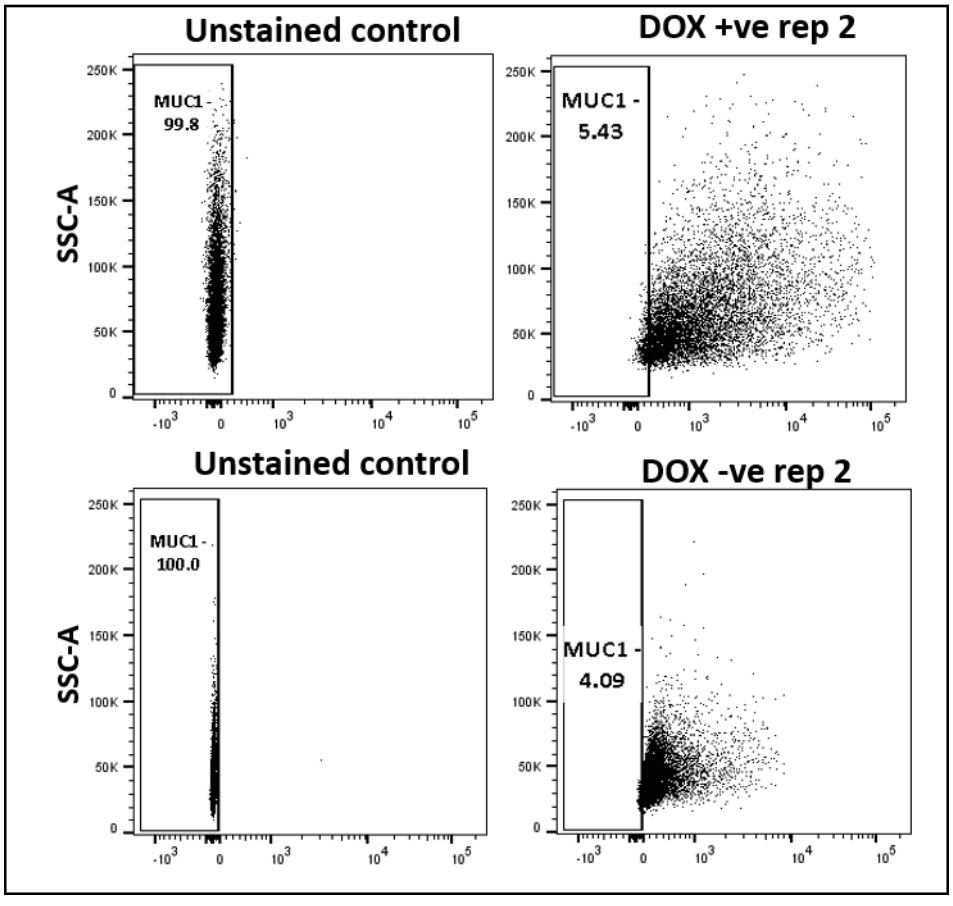
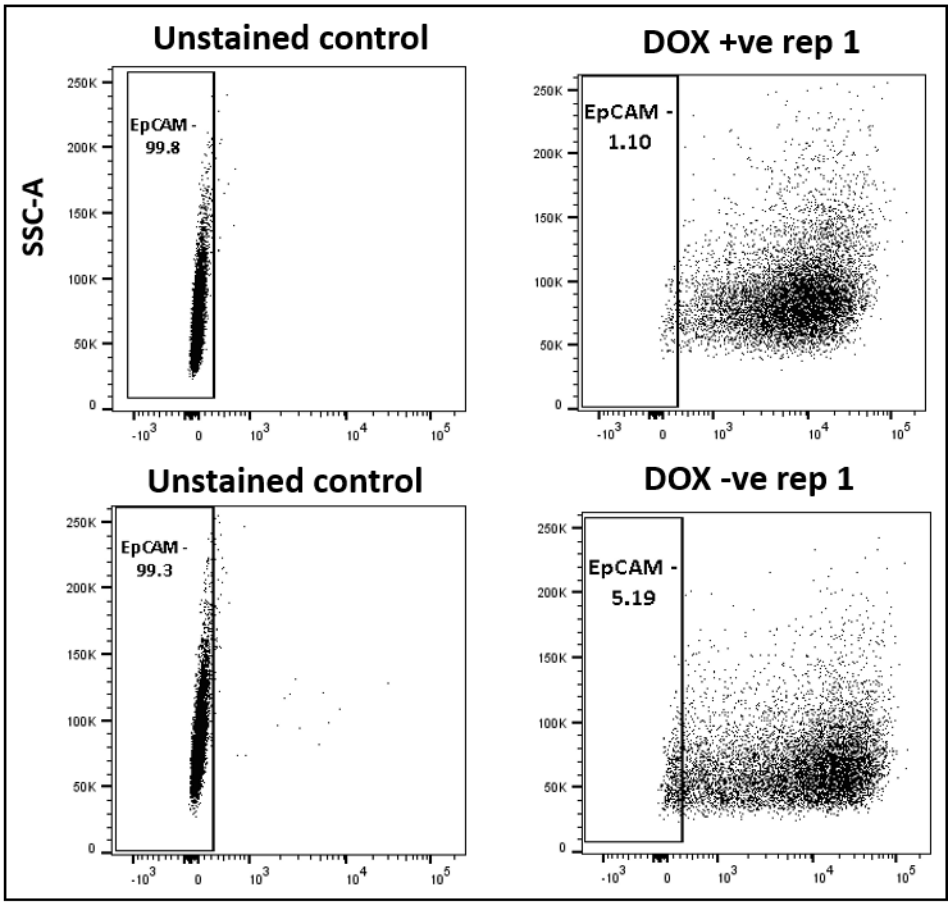

Unstained control
DOX +ve rep 2
SSC-A
MUC1 - 99.8
MUC1 - 5.43
Unstained control
DOX -ve rep 2
SSC-A
MUC1 - 100.0
MUC1 - 4.09

Figure 6.3: Relative abundance of MUC1 protein expression. Flow cytometry analysis of single stains of MUC1 in two independent biological replicates of DOX -ve and DOX +ve cells. The negative gating was based on unstained cells.

Additionally, we found that EpCAM has decreased expression upon HER2 over expression, resulting in a subpopulation of cells exhibiting a stem-like phenotype ("stemness"). We verified the expression of MUC1 and EpCAM in two independent biological replicates and found variable percentage of MUC1 -ve and EpCAM -ve cells in both DOX +ve and DOX -ve cells, with consistently higher enrichment of MUC1 -ve and EpCAM -ve population in the DOX +ve cells relative to DOX -ve cells. The variability in MUC1 and EpCAM expression may show that

acquisition of the stem-like phenotype is a stochastic process or is a result of technical aspects of the experiment (Figures 6.3 and 6.4).


Unstained control
DOX +ve rep 1
SSC-A
EpCAM - 99.8
EpCAM - 1.10
Unstained control
DOX -ve rep 1
EpCAM - 99.3
EpCAM - 5.19

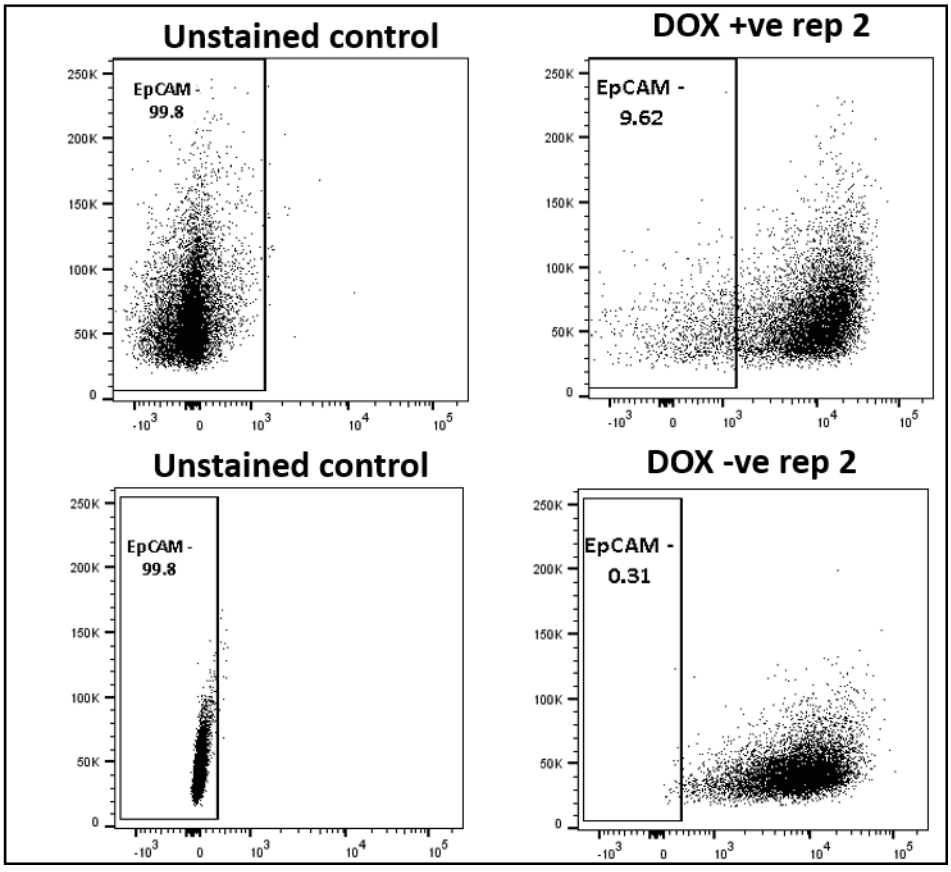

Unstained control
DOX +ve rep 2
EpCAM - 99.8
EpCAM - 9.62
Unstained control
DOX -ve rep 2
EpCAM - 99.8
EpCAM - 0.31

Figure 6.4: Relative abundance of EpCAM protein expression. Flow cytometric analysis of single stains of EpCAM in two independent biological replicates of DOX -ve and DOX +ve cells. The gating strategy was based on unstained cells.

To find out if the identified stem markers are co-expressed and co-localised as a result of HER2 over expression we carried out flow cytometry for the expression of MUC1, EpCAM, and added CD24 (cluster of differentiation 24), which is absent in breast cancer stem cells (285). The lack of CD24 expression alongside CD44 +ve expression in breast cells have been associated with enhanced tumourigenicity, and the conclusions from several investigations have shown a role in cancer initiation and metastasis (285-287). We identified that in the DOX +ve cells there were approximately 60% of cells expressing HER2 protein. Of these, we found 19.4% MUC1 -ve cells, of which 26.2% were EpCAM -ve cells. All of the MUC1 -ve/EpCAM -ve cells were also CD24 -ve (Figure 6.5). This was in contrast to the DOX -ve cells, which had depleted levels of stem markers. We identified 5.45% MUC1 -ve cells, of which 2.93% were EpCAM -ve. All of the MUC1 -ve/EpCAM -ve cells were also CD24 -ve (Figure 6.5). This suggests that *in vitro* transformation of MCF10A cells upon HER2 protein over expression favours/selects a subpopulation of cells enriched for cells with proteins expressed in stem cells based on the MUC1/EpCAM/CD24 -ve phenotype.

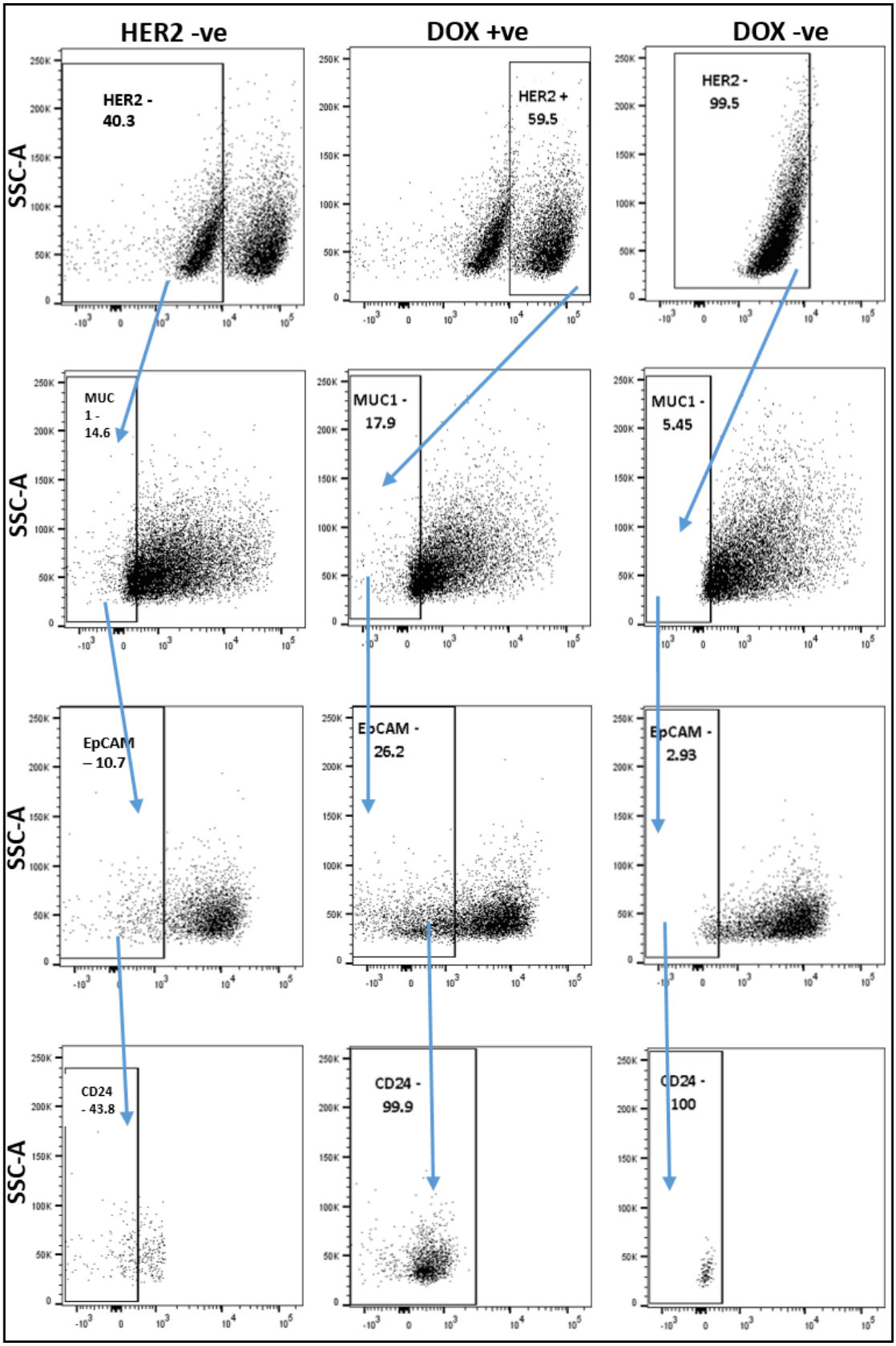


Figure 6.5: Outline of flow cytometry strategy of identifying the MUC1/EpCAM/CD24 co-expression in DOX +ve and DOX -ve cells. HER2 was induced for 72 hours and cells were treated with the combination

of antibodies as labelled. Cells were gated on HER2 expressing cells, which were further gated for the absence of MUC1, and thereafter for the absence of EpCAM and CD24. The negative gates were set on relative fluorescence minus one (FMO) controls whereby at least 99% of cells were selected.

We further wanted to know if variable HER2 expression (HER2 biomarker heterogeneity) induces differential expression levels of stem markers. At this point, we hypothesised that increased HER2 expression, more specifically the highest HER2 expressing cells in this experiment, would drive a more rapid acquisition of the stem state. We therefore selected the top 20% of HER2 expressing cells, and surprisingly, the stem-like markers were less enriched compared to the bulk HER2 positive population (6.6). This is because the enrichment of the stemness MUC1 -ve (11.6%) and EpCAM -ve (16.1%) had decreased compared to the bulk HER2 over expressing cells. Next, we selected the lowest 20% of HER2 expressing cells, and unexpectedly, we found that the stem cell markers have enriched in this population (Figure 6.6). There was an enrichment of stem markers, as MUC1 -ve (24.4%) and EpCAM -ve (30.1%) expression was higher compared to the high HER2 expressing cells, but also higher than the bulk HER2 positive cells (Figure 6.6).

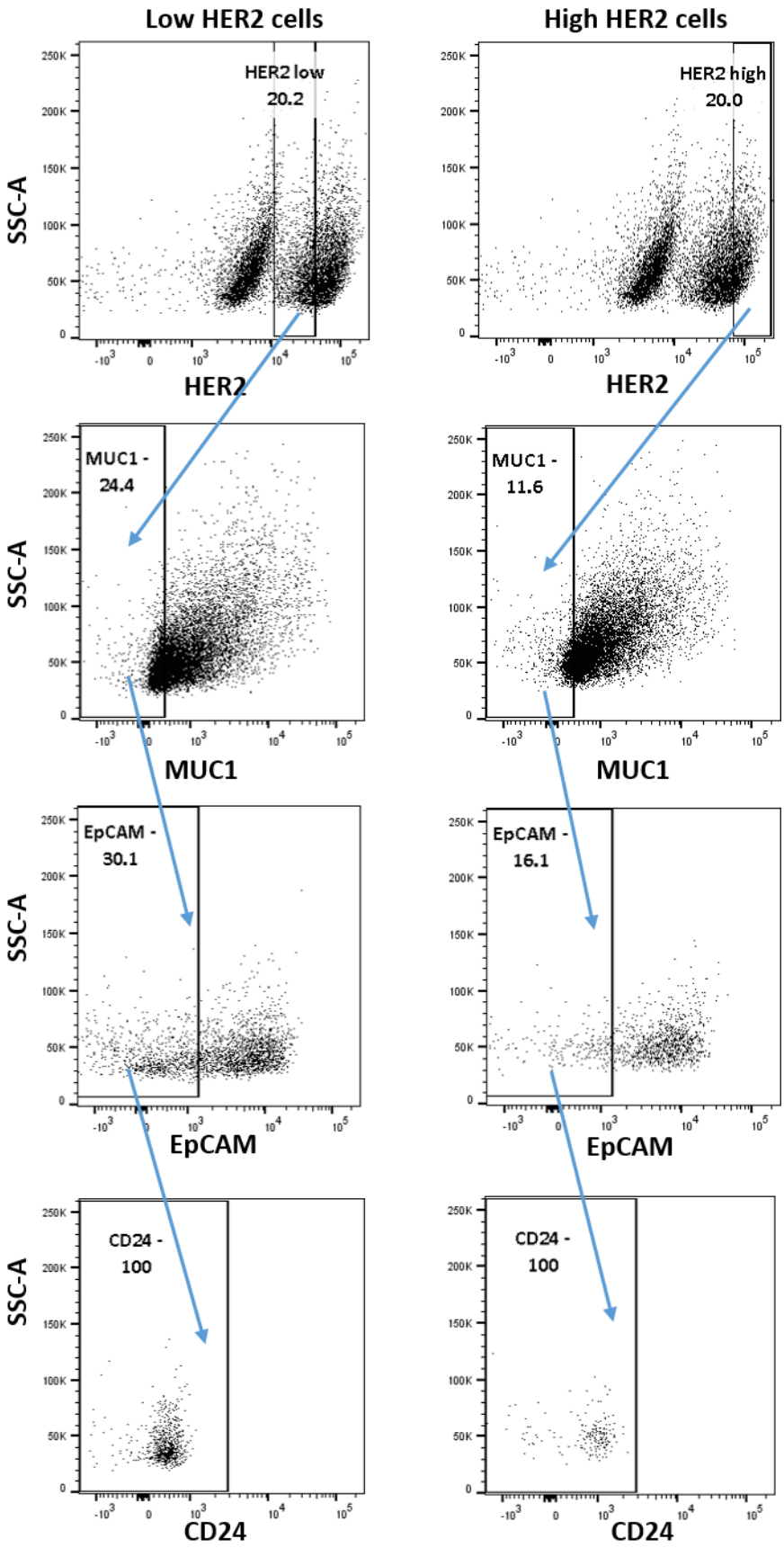

Low HER2 cells
High HER2 cells
HER2 low
20.2
HER2 high
20.0
SSC-A
HER2
MUC1 -
24.4
MUC1 -
11.6
MUC1
EpCAM -
30.1
EpCAM -
16.1
EpCAM
CD24 -
100
CD24

Figure 6.6: Identification of stemness based on HER2 biomarker heterogeneity. Of the DOX +ve cells, cells were gated based on 20% highest HER2 expression and 20% lowest HER2 expression. Cells were thereafter gated on the absence of MUC1, followed by the absence of EpCAM and CD24. The negative gates were set on relative fluorescence minus one (FMO) controls whereby at least 99% of cells were selected.

To enquire the stem-like phenotype of the medium HER2 expressing cells, we selected the middle 30% of HER2 expressing cells juxtaposed between the high and low HER2 positive cells. The resulting marker enrichment was the intermediate of the high HER2 and low HER2 expressing cells, with MUC1 -ve (23.4%) and EpCAM -ve (28.8%) (Figure 6.7).

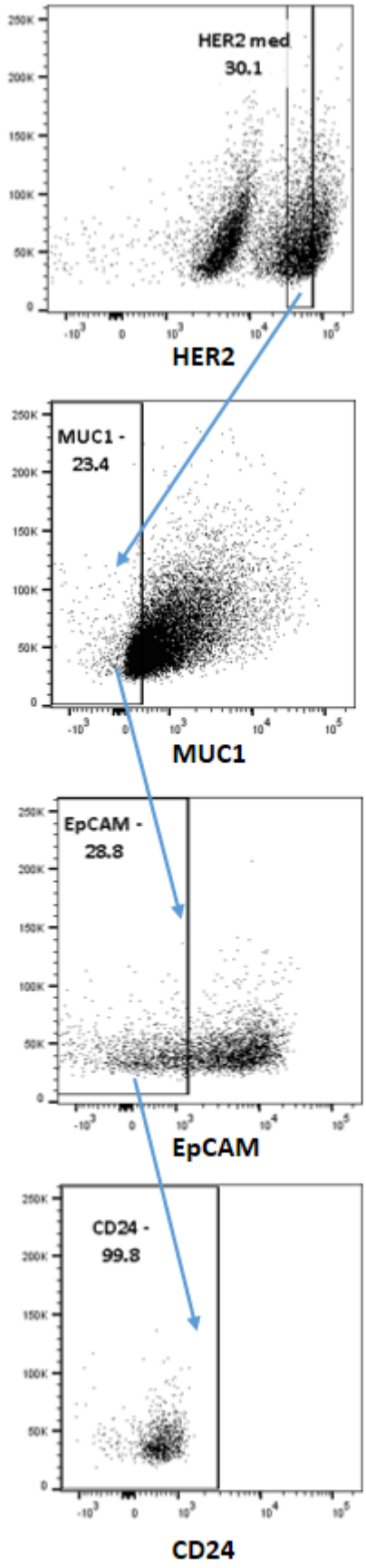

HER2 med
30.1
HER2
MUC1 -
23.4
MUC1
EpCAM -
28.8
EpCAM
CD24 -
99.8
CD24

Figure 6.7: Enrichment of stem markers based on "intermediate" expression of HER2. ~About 30% of HER2 medium expressing cells were first gated and thereafter on MUC1 –ve cells, which were further gated for EpCAM -ve and CD24 -ve cells. The negative gates were set on relative fluorescence minus one (FMO) controls whereby at least 99% of cells were selected.

To find out the percentage of stem cell markers based on the co-expression of MUC1/EpCAM/CD24 -ve proteins in the different subtypes of cells present in the total population, we plotted the enrichment of stem cell markers as a percentage value for simple visualisation. We see as previously shown that the low HER2 expressing cells have the most pronounced stem-like phenotype, followed by cells expressing "medium" HER2 expression.

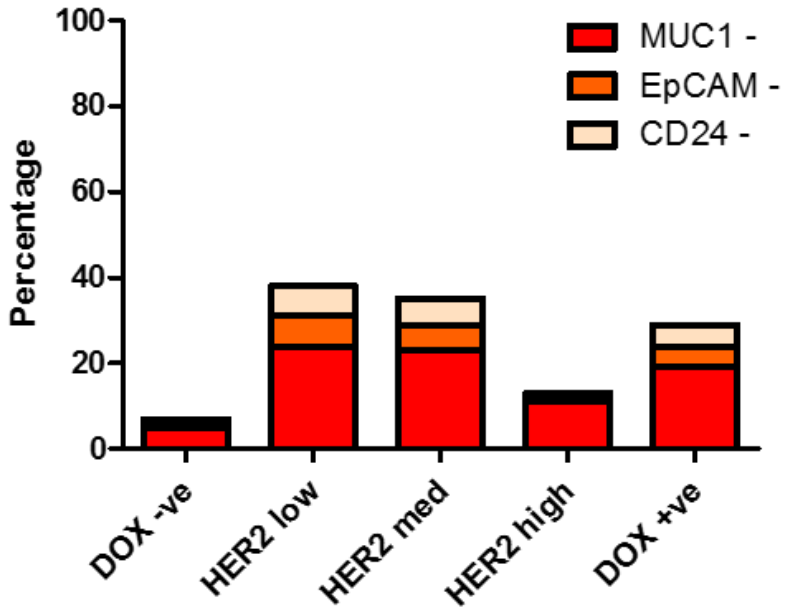


Figure 6.8: Determining enrichment of stem cell markers in subpopulations of HER2 positive cells. Cells were analysed by flow cytometry and HER2 positive cells were divided into three subpopulations of low, medium and higher HER2 expression as described above. The enrichment of stem markers is shown as a proportion of the total number of cells exhibiting MUC1 -ve, EpCAM -ve, and CD24 -ve phenotype.

6.4 Cells enriched for breast stem cell markers are associated with increased colony formation *in vitro*

As different HER2 expression levels correlate with distinct stem cell markers in our model based on the three proteins (MUC1, EpCAM, and CD24), where the expression of stem markers is especially enriched in the low HER2 expressing cells, we investigated the differences in the transformational potential between cells based on HER2 biomarker heterogeneity. The expression of HER2 was maintained for 72 hours in the DOX +ve cells and then cells were flow sorted based on the expression of HER2 protein. We sorted cells into three groups: the highest ~20% of HER2 expressing cells, the lowest ~20% of HER2 expressing cells, and the intermediate ~35% of cells, whilst using DOX -ve and DOX +ve cells as negative and positive controls (Figure 6.9A). To assess the ability to form colonies of the three sorted cell populations and associated controls, we performed soft agar colony formation assay by plating 5000 cells from each group in each well containing ultra-pure agarose. Interestingly, we found that the low HER2 expressing cells had a greater anchorage-independent growth capacity relative to the high or medium HER2 protein expressing cells, as they grew more colonies in the semi-solid media (Figure 6.9B).

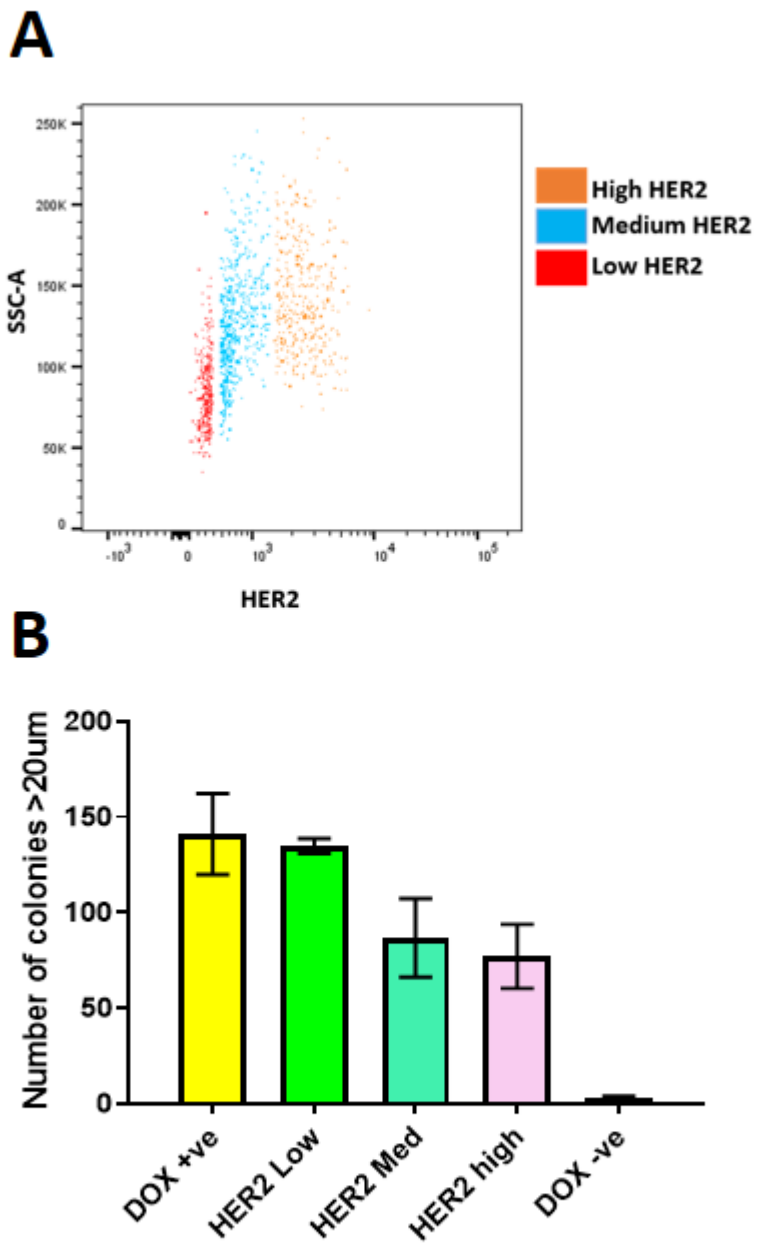

A
250K
200K
150K
100K
50K
0
SSC-A
HER2
High HER2
Medium HER2
Low HER2
B
Number of colonies >20um
200
150
100
50
0
DOX +ve
HER2 Low
HER2 Med
HER2 high
DOX -ve

Figure 6.9. Anchorage-independent growth of cells based on stem-like phenotype. (A) HER2 expression was induced for 3 days and cells were sorted based on HER2 expression into low, medium and high HER2 expression. (A) 5000 cells from each condition were plated into ultra-pure agarose to investigate their transformational potential over 21 days. N=2.

In the above experiment, we can see that DOX +ve cells have an enhanced colony formation ability compared to the low HER2 expressing cells or the others. However, the DOX +ve and DOX -ve cells were not flow sorted again in this experiment. We only FACS separated the DOX +ve cells into low, medium and high HER2 expressing cells. However, to make appropriate comparisons between the different types of cells, all of them must be subjected to the same procedures. To satisfy this, we FACS selected the cells into low, medium and high HER2 expressing cells, but also sorted the DOX +ve and DOX -ve cells. As previously, 5000 cells were then plated onto ultra-pure agarose to measure the anchorage-independency. Similar results to the previous experiments were observed. This is because the DOX +ve cells formed the highest number of colonies followed by the cells expressing low levels of HER2 (Figure 6.10).

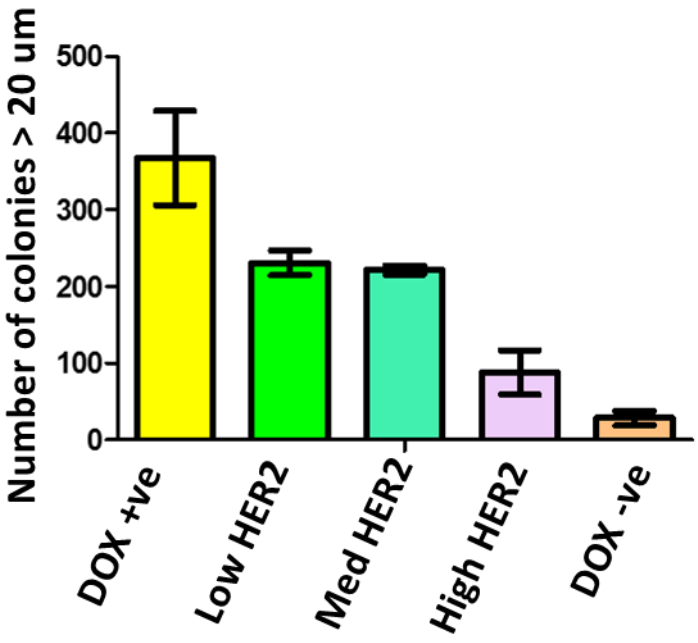


Figure 6.10: Measuring transformational potential of cells based on stem-like phenotype. As above, HER2 expression was induced for 3 days and cells were separated based on HER2 expression and then subjected to soft agar colony formation assay for 21 days. This experiment was performed in 3 technical replicates.

6.5 Investigating oncogene-induced senescence

Our results thus far indicate that high HER2 expressing cells form fewer colonies relative to the low and medium HER2 expressing cells. To understand why this was the case, we hypothesised that the cells with high HER2 expression undergo senescence, due to a phenomenon known as oncogene-induced senescence or OIS. Indeed, OIS has been previously observed with other oncogenes, such as Ras (288, 289). The high expression of HER2 is known to drive tumourigenesis, but paradoxically can also induce senescence (290). It has been found to induce senescence by upregulating various tumour suppressor proteins such as p16 (291). OIS is known to upregulate other tumour suppressor proteins such as p53, p27, and p21 (292). To test if high HER2 over expression leads to senescence, we carried out western blot analysis on proteins known to be upregulated in senescence. We investigated the senescence protein

expression in DOX -ve, low HER2, medium HER2, high HER2 expressing and DOX +ve cells. As positive controls we used DLD1 (colorectal cell line), expressing p53 and p27 and HCT166 cell line (human colon cancer cell line) expressing p53, p27, and p21. Another positive control was Naïve hESC (naïve human embryonic stem cells), which are also positive for p53, and NALM6 (acute lymphoblastic leukaemia (ALL) cell line) which is positive for p27 and p53. From this preliminary analysis, we concluded that there was no difference in the expression of proteins implicated in OIS. Therefore, high HER2 expressing cells do not induce OIS and so another mechanism may be responsible for the low colony growth in agarose. However, the loading controls (GAPDH and tubulin) are not equal as the high HER2 expressing have less protein loaded compared to other cell types. This analysis requires further attention to ensure appropriate conclusions are made (Figure 6.11).

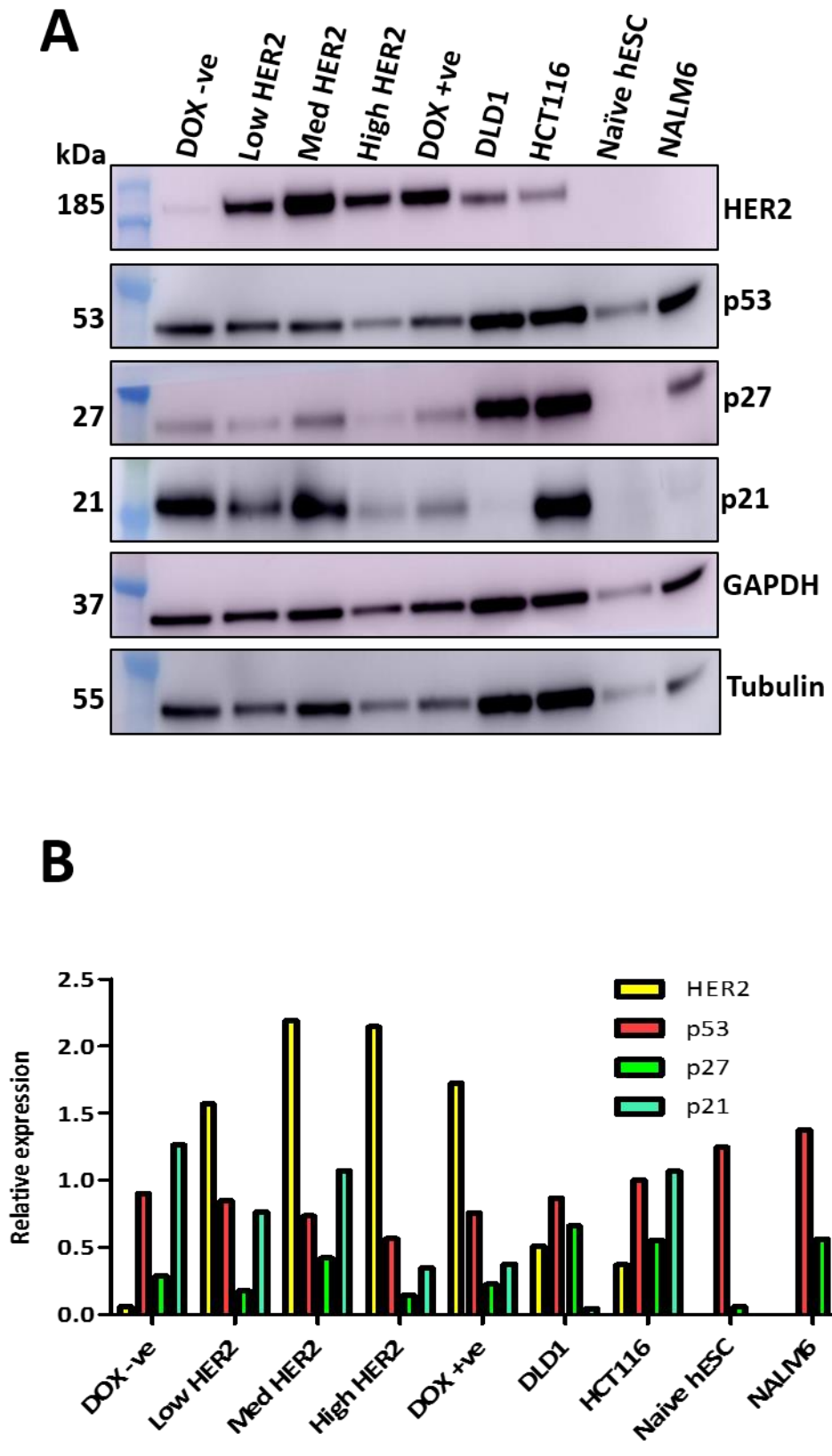


Figure 6.11: Investigating oncogene-induced senescence. (A and B) Western blot and densitometry analysis of the indicated proteins known to have higher expression in cells that have undergone OIS.

Protein lysates were prepared from cells sorted based on HER2 expression as previously (Figure 4.9). HER2 was induced in cells for 3 days (DOX +ve cells) and then FACS separated based on HER2 expression into three different subtype (low, medium, and high HER2 expressing cells). Cells were grown in culture and protein lysates were prepared. DOX -ve and DOX +ve cells were used as controls. DLD1, HCT116, Naïve hESC, and NALM6 protein lysates were used as positive controls (n=2).

As a way to explain why low HER2 expressing cells have a higher anchorage-independent growth compared to cells with medium or high HER2 expression, we hypothesised that medium and high HER2 expressing cells may have lower signalling activity of the MAPK signalling network relative to low HER2 expressing cells, which could be contributing to their weakened ability to form colonies. This could be caused by negative feedback loops acting to limit MAPK signalling in cells expressing very high levels of HER2 protein. To determine if there are differences in signalling between cells expressing varying levels of HER2, we as previously induced HER2 protein expression for 72 hours. Cells were serum starved for 24 hours (whilst maintaining HER2 expression). Cells were then stimulated for 5 minutes with full growth media to activate the signalling and protein lysates were prepared for western blot analysis. The preliminary western blot shows that there is no difference in the activation of ERK activity as the expression of phospho-ERK remained the same across the different populations. It appears that high HER2 expressing cells have lower ERK activation, but that is likely due to the lower protein loading as indicated by the lower total-ERK, tubulin and GAPDH protein expression (Figure 6.12). However, we consistently observed low levels of protein abundance in the high HER2 expressing cells based on the loading controls, despite multiple repeats of protein quantifications. This possibly indicates to a biological effect in high HER2 expressing cells, potentially during FACS selection, when cells are under stress, with high HER2 cells being more affected than other subpopulation of cells.

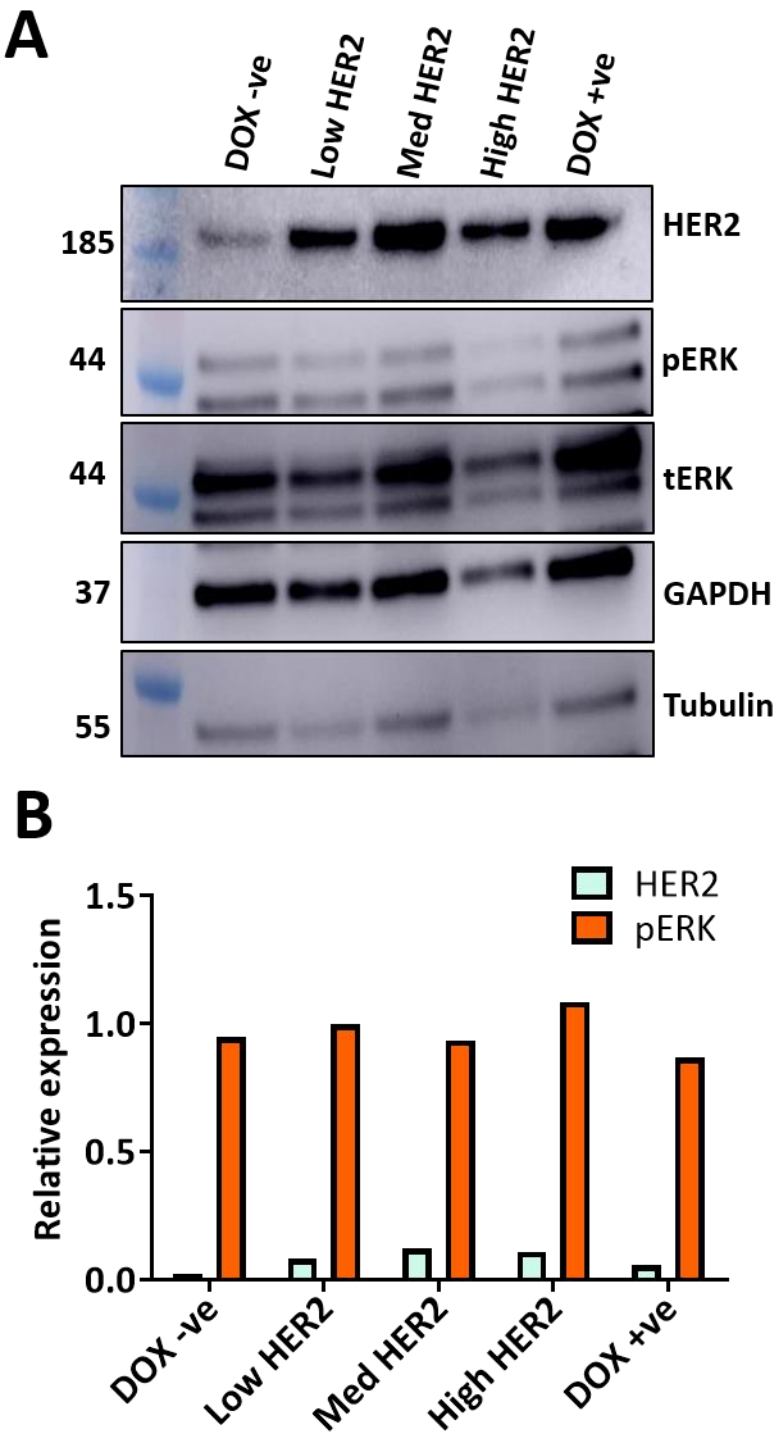


Figure 6.12: ERK phosphorylation in cells expressing differential levels of HER2 protein. (A and B) Western blot and densitometry analysis for HER2, phosphorylated-ERK, total-ERK, loading controls (GAPDH and tubulin) in cells expressing varying levels of HER2 protein (n=1).

One of the earliest and consistent phenotypic alterations observed upon HER2 induction in MCF10A cells in this system is the aberrant morphological changes in 3D cell culture. To assess this phenotype we cultured the DOX -ve, DOX +ve low HER2, medium HER2, and high HER2 expressing cells in 3D culture for 3 days. We found that all cell populations were characterised by flattened morphology with protrusions, expect for DOX -ve cells that formed normal round conformation as previously shown. However, the extent of protrusions was variable between the cell types. The DOX +ve cells had the most pronounced invasive morphology, followed by the low HER2 expressing cells. To a large extent, the medium and high HER2 expressing cells also form aberrant structures, but there were also some normal, round acini as observed in the DOX -ve cells, which were absent in the DOX +ve and low HER2 expressing cells (6.13).

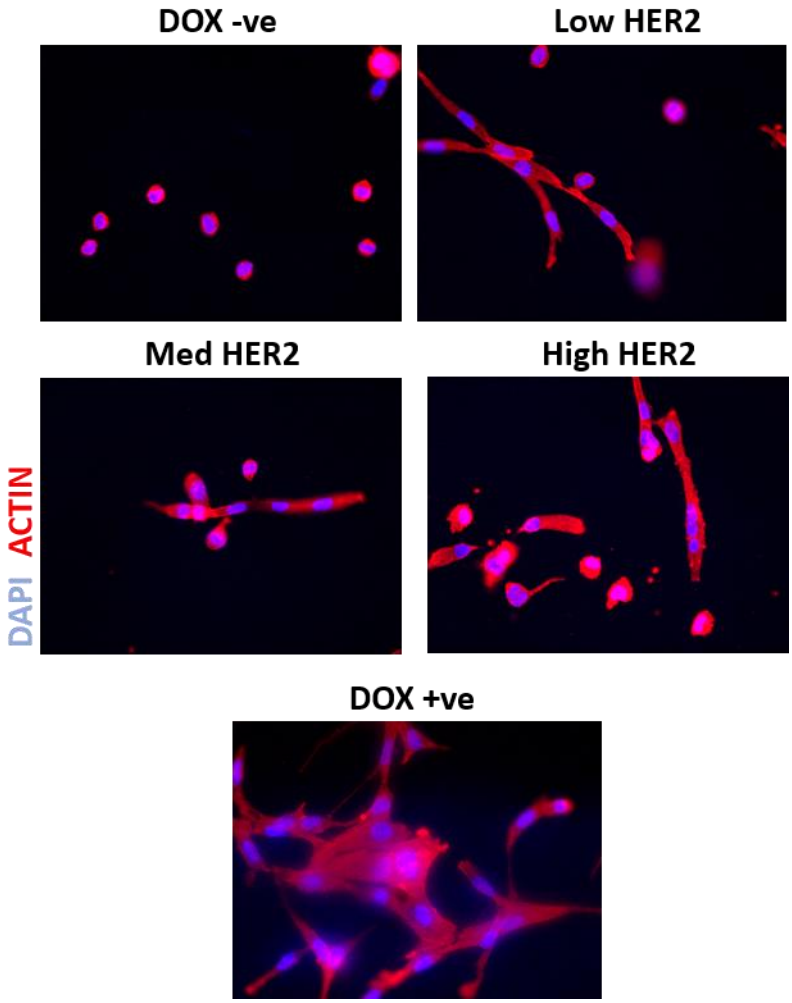


Figure 6.13: Morphology observed for the labelled cell populations grown in matrigel/collagen.

Representative fluorescence images of cells cultured in overlay 3D cell culture method for 3 days. Blue: nuclear staining with DAPI. Red: Actin staining by phalloidin dye. Magnification: 10X. n=1.

6.6 Discussion

Our results indicate that *in vitro* transformation of MCF10A cells as a result of HER2 protein over expression results in generation of markers present in stem cells. We find that there is an emergence of a subset of cells that have enrichment for markers of stem cells, which may have a higher transformational potential. Interestingly, these subpopulation of stem cells marked by the decreased levels of MUC1, EpCAM, and CD24 are counterintuitively enriched for low HER2 protein abundance compared to cells expressing high HER2 protein levels.

Since independent clones of HER2 transduced cells resulted in different percentages of cells giving rise to the stem-like subpopulation may suggest that the emergence of the stem-like phenotype in this model is a stochastic transition, indicating that the behaviours of the cells may partly be due the over expression of HER2 but also other unknown intrinsic and/or extrinsic factors. This results points to heterogeneous cancer stem cell population as distinct subset of cells acquires the capability to present the stem cell phenotype.

In this system, we have observed that the enrichment for stem-like cells arises three days after HER2 protein induction. However, this is an arbitrary time-point we had chosen and have continued our subsequent experiments at the three day time point, and we do not know precisely when exactly these stem-like cells are emerging. It may be that it arises much earlier than the three day time point such as at 24 hours or 48 hours and those time points may be associated with a more expanded stem-like phenotype. In that case, we would have missed the most critical stage of the transition to stem-like phenotype. Our work so far does not show the plasticity of these cells to reprogram back to dedifferentiated cells and cannot yet ascertain if these cells expand the stem cell population, or if the transition is static or is decreased as the HER2 over expression is maintained for longer.

As for the high HER2 expressing cells that grow fewer colonies in the ultra-pure agarose, we reject our hypothesis based on high HER2 cells inducing OIS. It is conceivable think that because there is lack of growth in high HER2 expressing cells, they might be undergoing apoptosis as high HER2 expressing cells in isolation may be toxic. It would be interesting to study the expression of markers associated with apoptosis such as caspase and PARP. Furthermore, it has been shown that non-malignant cells upregulates IRF6 (Interferon Regulatory Factor 6) (293), which leads to the blockage of anoikis. Since low HER2 expressing cells would -at first thought- be considered closer to normal cells than high HER2 expressing

cells, it could be that the low HER2 expressing cells are blocking cell death by the upregulation of IRF6 and eventually growing more colonies. However, this remains to be studied.

Furthermore, we have not yet elucidated the reason why the DOX +ve cells have an increased anchorage-independent growth compared to low HER2 expressing cells, despite having restricted stem-like phenotype. It may be that when there is a heterogeneous cell population (i.e. low, medium, high HER2 expressing cells), they conform to a more aggressive behaviour as compared to a more homogeneous cell population (low HER2 expressing cells only). The EMT transition between low, med, high HER2 expressing cells will also need to be compared to DOX +ve cells, since it could be that the DOX +ve cells undergo EMT to a higher extent than low HER2 expressing cells, making them more transformative. Indeed, other characteristics of heterogeneity would also need to be studied to understand the true nature of aggressiveness between low HER2 cells and the bulk (DOX +ve) cells. It could be that the DOX +ve despite forming more colonies would be easily eliminated because it has lower percentage of stem cells, whereas the low HER2 expressing cells have fewer colonies, but because of higher number of stem cells, would be resistant to therapies.

Chapter 7

Discussion

7.1 MCF10A cells – controversial model for breast cancer progression

The MCF10A human breast epithelial cell line provides an opportunity to investigate the initiation, development, and progression of breast cancer systematically. This cell line is arguably the most commonly used non-malignant breast cell model, as it exhibits characteristics of normal breast epithelium, such as dependency on growth factors for survival, lack of anchorage-independent growth and formation of acini in 3D cell culture (294, 295). These features render MCF10A cells a good model to study the effects of oncogene-induced transformation. However, despite not being transformed, the molecular features of MCF10A cells include the inactivation of p16 and p14ARF genes, which has allowed spontaneous immortalisation of these cells (294). The main objective of our system was to characterise the early events in transformation; however, several lines of evidence have shown that immortalisation is a prerequisite for transformation (296-298). Immortalisation of cells, which disrupts the physiological mechanisms regulating normal proliferation and cell growth, is a hallmark of cancer. To achieve the state of immortality, cells must gain additional genetic and/or epigenetic alterations, and since MCF10A cells are established to proliferate without limit *in vitro* (298) (by the inactivation of p16 and p14ARF), the initial phase of transformation has indeed already taken place before HER2 expression could be induced in our system.

As a consequence, our model is limited in that it is not possible to characterise the events at the very onset of transformation. To overcome this challenge, an alternative model may be proposed. Transformation of primary human breast cells can be generated by using an inducible oncogene in 3D cell culture to be more physiologically relevant to the human

context, which will allow the most appropriate characterisation *in vitro* of the early events upon an oncogenic insult. Indeed, primary breast epithelial cells have been previously oncogenically transformed (299-301). However, the accompanying molecular changes using an inducible oncogene that has been implemented in breast cancer to transform primary breast cells are not yet elucidated.

Nevertheless, it is recognised that HER2 over expression is not the only aberrant lesion in HER2 positive breast cancer and other changes such as p53 mutations are observed alongside HER2 over expression (91). However, p16 inactivation, as is the case in MCF10A cell line, is not an early event in HER2 positive breast cancer. This is because HER2 is over expressed in most cases of DCIS, but only about 20-30% of invasive ductal carcinomas (IDCs) exhibit over expression. These observations establish that HER2 over expression acts as an early event, or even as a first hit, which may be followed by a secondary hit – an invasion promoting hit – which impacts only a fraction of DCIS cases, and ultimately gives rise to IDCs. In the context of MCF10A cells, impaired HER2 expression cannot be the first hit as it is already established that MCF10A cell line's first hit is the inactivation of p16 and p14 ARF locus. Thus, this cell line does not follow the canonical progression of breast cancer as is seen in HER2 positive breast cancer patients and, therefore, many of the associated molecular events occurring in our model system cannot fully recapitulate those observed in patients.

7.2 Conditional oncogene expression – taking advantage of inducibility

By applying a tightly controlled doxycycline-inducible gene expression model to MCF10A cell line, we have further contributed to the improvement of this system, which is commonly used to characterise the early carcinogenic alterations and to understand the luminal epithelial cell biology. The main advantage of this model lies in the feature that it is inducible, allowing for high resolution analysis at the earliest time-points upon HER2 over expression. In our model, we introduced the pINDUCER21 (inducible) vector into MCF10A cells, which allow examination

of early transformational changes at a high temporal resolution. However, a key caveat when transducing cells with an inducible vector is the heterogeneous expression levels of transgene within the cell population, due to differences in the position of viral integration (116). Yet, in our model system the differences in gene expression levels were not an issue as the heterogeneity recapitulates the HER2 positive breast cancer patients better, since HER2 biomarker heterogeneity has been observed in clinical samples. Furthermore, this allows us to compare cells expressing high or low levels of HER2 protein with those cells that do not express the HER2 transgene at all, within the same genetic background.

The pINDUCER21 vector tightly controls HER2 expression under the control of TRE promoter, but the GFP is driven by a weak EF1α promoter, which we used to flow-sort cells to obtain only those cells that have the vector successfully transduced. However, since the GFP driven promoter is weak, we FACS-selected only the 2.3% cells based on GFP expression at high purity (approximately 90%), despite the fact that many cells would have had successful transduction. MCF10A cells are known to be heterogeneous, expressing various markers for breast stem cells, myoepithelial cells, and luminal cells (55), and the decision to select based on such a low percentage of cells poses the risk that the native heterogeneity may not be captured. This might have a profound effect on the subsequent experiments, making it difficult to make meaningful conclusions or comparisons with other systems employing MCF10A cells, especially comparisons to clinical samples as the heterogeneity observed in patients may not be replicated.

Although inducible systems allow for stringent control for characterising gene function, cells can lose over time the fraction of cells that have been successfully transduced. We found that when we cultured cells for an additional 8 passages, the GFP expression was reduced by up to 30 percent as measured by flow cytometry (Figure 3.2C). This would mean that the expression of HER2 is also reduced. Although we have not directly measured the inducibility of HER2

transgene over time, we can indirectly see from our experiments that not all cells express HER2 (Figure 3.3). It is documented that tetracycline inducible system gradually lose inducibility over time (246). Two possible explanations may be considered: firstly, it could be that the 10% non-transduced cells clonally outgrow and outcompete the transduced cells over time, and the cells with the successful transduction are progressively lost. Secondly, epigenetic silencing may act to inactivate the inducible promoter (tetracycline response element), resulting in decreased number of cells with the vector (246).

Finally, inducible systems require mediators such as tetracycline or its derivative doxycycline to induce the expression of the gene of interest. However, the use of these antibiotics can have confounding off targets effects at concentrations commonly used in inducible systems, from 100ng/ml to 5µg/ml. It has been identified that the use of dox in cell lines, including the MCF10A cell line, can decrease the proliferation of cells and induce metabolic gene expression alterations (302). This could have notable effects on the various phenotypic changes we have observed to characterise transformation. For instance, upon HER2 over expression, we have seen morphological changes in 3D cell culture, higher migration and invasion potential, and the formation of colonies in agarose, which all rely on proliferation to a partial extent. We would assume that the addition of dox may have significantly reduced the extent of the phenotypes observed. This is because these phenotypes – without dox addition and with sole HER2 over expression – would be more pronounced and with the introduction of dox have been decreased to a certain level. Therefore, the true extent of HER2 transformational drive may not be appropriately characterised.

7.3 HER2 induced phenotypic alterations

Using wild-type HER2 over expression as a model oncogene, we have further confirmed that aberrant ectopic expression of HER2 in MCF10A cells can alter the morphology, migration/invasion potential, and their ability to carry out metastatic properties by growing in

the semi-solid media (3, 239). These features can only be exhibited by cells that have undergone transformation. MCF10A cells cultured in matrigel/collagen develop important characteristics of normal mammary tissue, which can be manipulated by HER2 over expression. We provide evidence of the dramatic disruption of the normal architecture of MCF10A cells to produce morphological protrusions similar to the ones observed in early pre-malignant mammary lesions by means of HER2 over expression alone. However, 3D cell cultures are simplified microenvironments with reduced complexity compared to *in vivo* models, but they are still useful for mechanistic studies in transformation. Our observations are in full agreement with previous reports showing the elongated, larger and less cohesive features of MCF10A acini upon oncogene expression, whereas normal cells retain organised, spherical conformation (55, 295, 303-306). An essential feature of early breast cancer is the repopulation of the lumen with cancer cells (57). It would have been useful to allow the DOX -ve cells to form the lumen as it occurs in normal MCF10A cells after day 10. This could have been followed by the induction of HER2 to study if the lumen formation occurs as a result of HER2 expression. This phenotype would be useful to have been seen by a live imaging system such as an incuCYTE microscope to pinpoint the time it takes for HER2 to induce this phenotype.

Another neoplastic characteristic of transformed cells is the induction of migration and invasion of cells into the surrounding tissues. Likewise, the migratory and invasive features are in line with previous studies that show that constitutive expression of HER2 results in higher migration and invasion potential (48, 307, 308). We have shown that MCF10A cells with HER2 over expression (DOX +ve cells) are able to grow colonies in semi-solid media, which to an extent represents metastasis *in vivo* as cells are moved from their normal microenvironment to reside in an unsuitable one. However, the variation in performing the technical aspects of the assay and the individual quantification methods employed make it challenging to draw comparisons between our results and of the previous studies. For instance, in our experiments

we used 0.3% ultra-pure agarose to measure anchorage-independent growth of cells, whereas the type of matrix and its percentage can vary between different experiments, such as the use of noble agar (ultra-pure gelling agent) in this study (53). Another variation involves methods introduced to count the number of colonies. This could be overcome by the incorporation of fluorescent dye to enable high throughput counting. Furthermore, specialised soft agar or agarose solution could be used to facilitate the isolation of viable cells for easy counting after the assay end point to allow for protein, DNA and RNA samples to be prepared if required.

Although the assays we have performed show the transformative behaviour of cells as a result of HER2 over expression, there are other aspects of transformed cells that could also be investigated such as foci formation capability of cells and the ability of cells to survive and proliferate in reduced growth factor media (6).

One of the most fundamental and useful piece of information missing from transformational models in general is the question; how long does it take for cells to become fully transformed upon induction of an oncogene? The inability to answer this question may largely be attributed to the use of non-inducible systems and of normal versus cancer cells, because the timing of gene induction is not known. However, it appears that even inducible gene expression systems cannot answer this difficulty properly. For instance, we have seen in our model that the morphological alterations occur within three days of HER2 over expression, that the cells can migrate/invade after five days of HER2 induction, and that the anchorage-independency is acquired within the 21 days. However, this does not inform when cells attain full transformation *in vitro*, even though we know that by day 21 the cells have transformed relative to the normal cells according to the soft agar assay. It could be that cells gain migratory and invasive phenotype earlier than 5 days, it is just that we measured migration/invasion of cells at 5 day time point, and that the cells gain anchorage independency earlier than 21 days. However, we do not yet know if this phenotype is attained earlier in

transformation. The advantage of knowing when full transformation is achieved in cells could give us clues to perform the molecular analysis at the right and most relevant time points. For example, if complete transformation was reached by day 5, then the molecular analysis time points could be spread out to encompass full transformation without uncertainty, preventing the characterisation of molecular analysis from time points that are not in our objectives. Nevertheless, cellular transformation could sometimes be an ambiguous term and its proper definition is important in understanding what can be classified as transformed or not. For example, overexpression of cyclooxygenase 1 in spontaneously immortalized human umbilical vein endothelial cells were not anchorage-independent but grow tumours *in vivo* aggressively (309). These cells would be classified as transformed despite not growing colonies in soft agar because they are able to induce a more significant event, which is to grow tumours in mice. In another case, human primary foreskin fibroblasts attained anchorage-independent and grew tumours in vivo but were able to indefinitely proliferate in cell culture (310). Moreover, the human papilloma virus 16 E6 oncogene was sufficient to induce anchorage-independent growth but did not generate tumours when injected into mice (311). Therefore, transformation of cells is dependent on a number of factors and experimental settings, such as *in vivo* or *in vitro* work, the types of analysis performed, relative controls, and the types of analysis performed.

7.4 The signalling dynamics – taking a global approach

Our global phosphoproteomic study extends the knowledge of signalling induced by HER2 over expression by identifying previously uncharacterised downstream signalling proteins. In this experimental setting, we carried out a mass spectrometry screen under standard growth conditions as opposed to in response to acute external stimuli to faithfully mimic the physiological impact of HER2 expression at short time points after induction. We have also identified previously unknown phospho-peptides which include LAP3 (R440), HIPK1 (Y352), and

GPX1 (S153). Whilst it is extremely valuable to understand the early signalling changes upon HER2 protein over expression and cellular transformation, the screen is restricted to the short term effects of HER2 expression and may overlook the secondary effects which could rely on the transcription and translation of regulatory proteins.

A limitation of our investigation is the identification of a relatively modest number of phosphopeptides from our analysis. Our dataset shows approximately 4000 proteins with one or more phosphosite, which is lower than published reports of 7500 and 7214 phosphosites respectively (312-315). Furthermore, another limitation of this study is the reliance on the phosphoproteome of cells, without focusing on changes in protein abundance. In the absence of in-depth proteomic analysis, we cannot distinguish if the alterations in the phosphoproteome of our cells are a result of the protein phosphorylation stoichiometry or due to differential levels of total protein expression. However, it could be that it is primarily the activation of proteins (via phosphorylation) rather than total protein expression that may be vital for the regulation of molecular mechanisms involved in transformation (316).

Our phosphoproteomic analysis finds that, upon HER2 over expression in all the time points we have studied, approximately 20% of the phosphoproteome is significantly changed. Although not directly relevant, this is in contrast to the gastric cell line which displayed that 5% of the phosphoproteome was significantly altered compared to the parental cell line (317). In another case, GIST cell line showed approximately 75% of the phosphoproteome altered versus the parental cell line (318). One reason for these differences could be the underlying genomic drivers introduced between different cell lines. Another contributing factor may be that the depth of the phosphoproteome coverage is less comprehensive in our study and that we are only examining the phosphopeptides with the highest abundance in our cells. Finally, it could be that receptor tyrosine kinases (such as HER2), reprogramme signalling networks to achieve transformation using distinct set of mechanisms.

As expected, one the largest increases of phosphorylation was seen in the HER2 (at T701 and T1060) itself, because as we add dox, the levels of HER2 proteins increase and hence a higher fold change was observed. The fold change was some 6-fold more in the 7 hour time point compared to the 0.5h one. This is because the longer the cells were cultured in dox-containing media, the higher the time-dependent expression of HER2. This may mean that with higher protein abundance, HER2 can increasingly homodimerise and transmit potent signals downstream, as they do not rely on a ligand to induce active signalling due to their open extracellular conformation. Furthermore, as HER2 expression increased, counterintuitively there was a higher number of differentially regulated phosphopeptides observed. For instance, phosphopeptides that had differential levels of phosphorylation at the 0.5h time point were 310, at 4h time point they were 390, which increased to 455 at the 7h time point. This shows that higher HER2 expression is likely to change phosphorylation status of an increasing number of proteins. A rather simple observation maybe put forward: it is well known that the higher grade tumours (3+) of HER2 positive breast cancers are more aggressive due to the higher expression of HER2 protein expression as assessed by IHC. It is therefore conceivable to think that one of the reasons why they behave aggressively is the result of widespread activity in the signalling networks amongst other changes.

Amongst the earliest changes detected at the 0.5h time point upon HER2 protein over expression are the downregulation of phosphopeptides involved in cell-cell junction and adherens junctions, these phosphopeptides include: LMO7 (S988), which is downregulated at all time points, but also include CTND1 (T869), AKT2 (T451), and TLN1 (S488), amongst other changes. This is consistent with the observed phenotypic alterations, such as the morphological changes in 3D cell culture and anchorage-independent growth of cells.

To capture the dynamics and complexity of the signalling networks upon HER2 over expression, a cocktail of ligands should be used which could include heregulin (HRG), neregulin

(NRG), TGFα, EGF, and epiregulin. These would induce dimerisation of all possible combinations between the various family members of HER receptors. This is unlike the addition of EGF ligand alone, as in our case, which induces homodimerisation between HER1-HER1 and heterodimerisation between HER1-HER2. Notably, it is already known that stimulation by HRG activates a specific subset of the migration signalling network that is not induced by EGF (319). Nevertheless, the advantage of introducing a single ligand, such as EGF allows us to attribute the signalling changes to one factor without ambiguity from other ligands.

While the phosphoproteomic layer of protein regulation provides valuable and descriptive insight to the process of transformation, the challenge is that the results are not readily interpretable or actionable. For instance, we identified that a handful of signalling cascades are affected upon HER2 protein over expression, which may suggest that administration of specific kinase inhibitors could be used as a therapy, it does not reveal the complete mechanism of transformation. Nevertheless, there have been successful instances such as in Zeevi et al (320), where they employed an '-omics' dataset, patient data, and machine learning to implement a change in nutrition to regulate glucose levels, without deep insight of the mechanism. However, in the majority of cases, the absence of mechanistic information of disease progression makes it challenging to find targets for therapeutics with reliability. In order to move away from the 'big picture' provided by the phosphoproteomic data to investigate a testable hypothesis, it is documented that signal transduction pathways can modulate chromatin structure. To study the relationship between important signal transduction pathways and the chromatin architectural landscape in transformation, we have performed ATAC-seq analysis at similar time points to our phosphoproteomic study to analyse the link between cell signalling and chromatin structure. Interestingly, we have seen that in our system the MAPK signalling pathway, which is known to regulate gene expression at multiple levels, is the dominant cascade by which signalling is transduced. Among the downstream targets of the

MAPK signalling pathway, the MAPK5 and MAPK1 are of interest as they are able to directly target several transcription factors such as NFkB and ELK-1, which can in turn induce phosphorylation of Histone 3 and HMG-like proteins known to have an effect on chromatin accessibility (321, 322). We have observed from our dataset that various transcription factors and chromatin regulators become phosphorylated, and these include NFkB at (T811) and phosphorylation changes in various HMG phosphosites (such as HMG4BX [S497] and HMGA1 [T53 and S36]) amongst many others, which could potentially have an effect on the chromatin architecture.

7.5 Multiple layers of heterogeneity in breast cancer

We have shown that a sub-population of potentially cancer stem cells can emerge during the processes of cellular transformation by inducing the expression of HER2 protein. This subset of cells is uniquely marked by the absence of markers known to be either low or absent in breast stem cells, which include MUC1, EpCAM and CD24. Interestingly, we also identified that different HER2 expression levels coincide with distinct expression of stem markers, with the low HER2 expressing cells, unexpectedly, being the most enriched for stem cell markers compared to medium or high HER2 expressing cells.

We used a combination of well-known cell surface markers of MUC1/EpCAM/CD24 low/negative that is associated with stem cells (154, 323, 324). These markers individually have been implicated in stem and cancer stem cells, but their co-expression to identify stem-like phenotype to date has not yet been investigated. Nevertheless, there are other (cancer) stem markers known to be associated with stem-like phenotype, such as the high expression of ALDH1 (325, 326), high expression of CD44 and low expression of CD24, which together have been used as cancer stem cell markers in mammary cells (327-329).

Our findings are in line with previous investigations that have shown that EpCAM and CD24 negative or low-expressing cells are associated with mammary stem cells (330, 331). However,

it is the over expression of MUC1, rather than its decrease in protein abundance, that has been associated with worse prognosis in breast cancer (332). The upregulation of MUC1-C subunit is known to increase phospho-AKT and results in resistance to tamoxifen in breast cancer (333). Elevated expression of MUC1 has also been shown to be preserved in cancer stem cell population in luminal breast cancer cell lines (334). It seems that MUC1 has a multifaceted role in transformation, being associated with higher transformational potential when over expressed, but also, in our case, found to be associated with enhanced *in vitro* transformational properties when lowly expressed, jointly with EpCAM and CD24. Interestingly, reduced EpCAM expression is not only found in stem cells, but also in cells that display EMT phenotype (335).

In our experimental setting thus far, our approach to identifying stem-like phenotypic features was confined to considering the expression of markers enriched in stem cells. However, other potential mechanisms could be applied to identify, or at least confirm, the cells acquiring stemness. For instance, it is known that the rate of cell cycle of stem cells versus differentiated cells is different (336). It is identified that stem cells have a faster G1 phase of the cell cycle compared to differentiated cells (336). Furthermore, it would be useful to identify epigenetic signatures of stem cells compared to differentiated cells. It has already been found that the expression of EZH2, a core subunit of the PRC2 complex, can activate NOTCH1 signalling by binding to the NOTCH1 promoter and activating its signalling, which enhances the stem cell phenotype of cells (337). This is unprecedented, since EZH2 is normally known to have a suppressive role rather than activating one.

The identification of an expanded stem-like phenotype based on the MUC1/EpCAM/CD24 -ve expression and its association with low HER2 expressing cells in our system is a novel finding. This is because in HER2 positive cancers, high HER2 expression is associated with higher tumour grade and aggressive disease, and in turn worse prognosis and survival (338, 339).

Since it is known that the upregulation of stemness in cancer cells is associated with higher transformational potential (340), we would have expected that stem-like phenotype would be more highly enriched in the high HER2 expressing cells compared to the low or medium HER2 expressing cells. However, an observation may be made here; it is known that patients with high HER2 expressing cells (tumour grade 3+) tend to respond better to anti-HER2 therapy (341) compared to patients expressing borderline HER2 expression (tumour grade 2+) (342). This is partly attributed to HER2 regional biomarker heterogeneity. However, since cancers with upregulated stem-like phenotype are at the forefront of resistance to therapies, it is conceivable to think that part of the reason why the borderline HER2 positive breast cancer patients do not respond well to treatment is their expanded stemness, just as we have observed in our system compared to high expressing cells. To test this hypothesis, low, medium and high HER2 expressing cells could be separated and treated with trastuzumab or lapatinib to study what levels of HER2 confer higher levels of resistance to inhibition.

Furthermore, if the low HER2 expressing cells have more stemness – and considering that normal MCF10A cells have even lower levels of HER2 expression than the low HER2 expressing cells in the DOX +ve cells – it is logical to assume that MCF10A cells would have an even higher stem-like phenotype. However, the normal MCF10A cells (or the DOX -ve cells) did not have a MUC1/EpCAM/CD24 -ve phenotype. This shows that the low HER2 expressing cells have co-occurring aberrant alterations that make the cells acquire the stem-like phenotype.

7.6 HER2 over expression – what does it mean in the context of patients?

In the context of HER2 positive breast cancer patients, protein and gene expression levels provide critical information, as they act as predictive markers to diagnose patients based on biomarker expression. It is not clear, for example, whether the borderline (2+ grade) tumours have undergone complete neoplastic transformation or whether only the patients with 3+ grade tumours have full malignant transformation. Although not directly comparable to the

context and complexity observed in patients, a minimal and consistent increase in HER2 expression, as shown in our study, seems sufficient to induce transformation as measured by our *in vitro* assays. This raises the possibility of 3+ grade tumours undergoing additional changes (such as further HER2 amplification) that make them more aggressive. Nevertheless, the prognostic significance of the low-expressing HER2 positive cancers, such as 1+ grade tumours, which are generally regarded as HER2 negative alongside 0 grade tumours as assessed by IHC, have not yet been properly evaluated. One of the main reasons for not appropriately evaluating the prognostic value of low level HER2 expression is because the investigators generally group the 0+ and 1+ tumours categories together, assuming in advance that low level HER2 expression may not be clinically significant, despite systems such as ours showing that it may be sufficient to progress cancer. Additionally, many of these studies were published before the 0-3+ scoring system was clinically established by IHC. The HER2 positivity was defined by protein expression or by gene amplification above a given threshold by western blotting, or by immunostaining (343-349). Furthermore, it has been assumed that patients with low levels of HER2 expression many not benefit from targeted treatments such as trastuzumab, but existing data with regards to the low levels HER2 expression and their response to trastuzumab are contradictory and limited in number. For instance, in an evaluation by the National Surgical Breast and Bowel Project (NSABP) B-31, which looked at 161 patients found to be negative for HER2 expression by IHC and FISH. In this group of patients, the rate of relapse in patients treated with chemotherapy and trastuzumab versus chemotherapy alone was 8% and 21% respectively (350). In another similar study, in patients that were classified as HER2 negative both by IHC and FISH, the relapse rate of patients treated with chemotherapy and trastuzumab compared to chemotherapy alone was 15% and 30% respectively (348). However, the HER2 negative patients in the CALGB 9840 trial had a better response rate to chemotherapy and trastuzumab versus chemotherapy alone (35% versus 29%), but that was not significant (103). This points to the potential for low HER2 expressing

patients – generally thought as being negative – also benefitting from anti-HER2 therapy. Therefore, low levels of HER2 expression, in an *in vitro* system as we describe here, are critical in understanding how we define HER2 positivity and could provide us with a useful understanding of how low HER2 expressing breast cancer behaves. However, it should be emphasised that our investigation is at the hypothesis-generating stage, and should be extended further in order to aid our understanding of HER2 positive breast cancer.

Our model presented here is yet to be tested alongside primary HER2 positive breast cancer patient samples. Since the HER2 over expression in our system is low, it would valuable to test the HER2 gene and protein expression are similar to clinical samples from HER2 breast cancer patients with 0, 1+, 2+, and 3+ graded tumours. I would hypothesise that the HER2 expression levels would be similar to those observed in 1+ graded tumours. Based on the current HER2 assessments performed by IHC, the HER2 protein expression would be classified as normal in our system. However, as we have shown that such low levels of HER2 expression is sufficient to induce transformation and global changes in the signalling network as well as genome-wide changes in the epigenome. Another layer of complexity arises when such patients are not considered to be treated with HER2 targeted therapy, as the 1+ graded patients are seen as the "bystanders". If some of the work presented here could be replicated in a more physiologically relevant setting, such as *in vivo* work or the same levels of HER2 expression in primary breast cells, we could present a case for questioning the current practice of not treating 1+ graded tumours with anti-HER2 therapies. Especially when drug related toxicities of treating low HER2 expressing patients with anti-HER2 therapies are mild (351). This would be particularly useful for patients that present heterogeneous population of HER2 positivity. If potentially clear significant and compelling evidence is found that low HER2 expressing cells do not indeed benefit from anti-HER2 therapies, then at least an alternative method of therapeutics may be suggested for the 1+ scored tumours.

References

1. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646-74.
2. Woods Ignatoski KM, Grewal NK, Markwart S, Livant DL, Ethier SP. p38MAPK induces cell surface alpha4 integrin downregulation to facilitate erbB-2-mediated invasion. Neoplasia (New York, NY). 2003;5(2):128-34.
3. Pradeep CR, Zeisel A, Kostler WJ, Lauriola M, Jacob-Hirsch J, Haibe-Kains B, et al. Modeling invasive breast cancer: growth factors propel progression of HER2-positive premalignant lesions. Oncogene. 2012;31(31):3569-83.
4. Futscher BW. Epigenetic changes during cell transformation. Advances in experimental medicine and biology. 2013;754:179-94.
5. Rhim JS. Neoplastic transformation of human cells in vitro. Critical reviews in oncogenesis. 1993;4(3):313-35.
6. Tuveson DA, Shaw AT, Willis NA, Silver DP, Jackson EL, Chang S, et al. Endogenous oncogenic K-ras(G12D) stimulates proliferation and widespread neoplastic and developmental defects. Cancer cell. 2004;5(4):375-87.
7. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000;100(1):57-70.
8. Woo RA, Poon RYC. Activated oncogenes promote and cooperate with chromosomal instability for neoplastic transformation. Genes & development. 2004;18(11):1317-30.
9. Hahn WC, Counter CM, Lundberg AS, Beijersbergen RL, Brooks MW, Weinberg RA. Creation of human tumour cells with defined genetic elements. Nature. 1999;400(6743):464-8.
10. Elenbaas B, Spirio L, Koerner F, Fleming MD, Zimonjic DB, Donaher JL, et al. Human breast cancer cells generated by oncogenic transformation of primary mammary epithelial cells. Genes & development. 2001;15(1):50-65.
11. Lundberg AS, Randell SH, Stewart SA, Elenbaas B, Hartwell KA, Brooks MW, et al. Immortalization and transformation of primary human airway epithelial cells by gene transfer. Oncogene. 2002;21(29):4577-86.
12. Blasco MA, Lee HW, Rizen M, Hanahan D, DePinho R, Greider CW. Mouse models for the study of telomerase. Ciba Foundation symposium. 1997;211:160-70; discussion 70-6.
13. Chin L, Artandi SE, Shen Q, Tam A, Lee SL, Gottlieb GJ, et al. p53 deficiency rescues the adverse effects of telomere loss and cooperates with telomere dysfunction to accelerate carcinogenesis. Cell. 1999;97(4):527-38.
14. Imam SA, Kim MS, Anker L, Datar RH, Law RE, Taylor CR. Systematic determination of telomerase activity and telomerase length during the progression of human breast cancer in cell culture models. Anticancer research. 1997;17(6d):4435-41.
15. McEachern MJ, Krauskopf A, Blackburn EH. Telomeres and their control. Annual review of genetics. 2000;34:331-58.
16. Harley CB, Kim NW, Prowse KR, Weinrich SL, Hirsch KS, West MD, et al. Telomerase, cell immortality, and cancer. Cold Spring Harbor symposia on quantitative biology. 1994;59:307-15.
17. Seluanov A, Hine C, Bozzella M, Hall A, Sasahara TH, Ribeiro AA, et al. Distinct tumor suppressor mechanisms evolve in rodent species that differ in size and lifespan. Aging cell. 2008;7(6):813-23.
18. Olayioye MA. Intracellular signaling pathways of ErbB2/HER-2 and family members. Breast Cancer Research. 2001;3(6):385.
19. Elster N, Collins DM, Toomey S, Crown J, Eustace AJ, Hennessy BT. HER2-family signalling mechanisms, clinical implications and targeting in breast cancer. Breast cancer research and treatment. 2015;149(1):5-15.

20. Kraus MH, Issing W, Miki T, Popescu NC, Aaronson SA. Isolation and characterization of ERBB3, a third member of the ERBB/epidermal growth factor receptor family: evidence for overexpression in a subset of human mammary tumors. Proceedings of the National Academy of Sciences of the United States of America. 1989;86(23):9193-7.
21. Graus-Porta D, Beerli RR, Daly JM, Hynes NE. ErbB-2, the preferred heterodimerization partner of all ErbB receptors, is a mediator of lateral signaling. The EMBO journal. 1997;16(7):1647-55.
22. Tzahar E, Waterman H, Chen X, Levkowitz G, Karunagaran D, Lavi S, et al. A hierarchical network of interreceptor interactions determines signal transduction by Neu differentiation factor/neuregulin and epidermal growth factor. Molecular and cellular biology. 1996;16(10):5276-87.
23. Yarden Y. The EGFR family and its ligands in human cancer. signalling mechanisms and therapeutic opportunities. European journal of cancer (Oxford, England : 1990). 2001;37 Suppl 4:S3-8.
24. Yarden Y, Sliwkowski MX. Untangling the ErbB signalling network. Nature Reviews Molecular Cell Biology. 2001;2:127.
25. Press MF, Cordon-Cardo C, Slamon DJ. Expression of the HER-2/neu proto-oncogene in normal human adult and fetal tissues. Oncogene. 1990;5(7):953-62.
26. Morris JK, Lin W, Hauser C, Marchuk Y, Getman D, Lee KF. Rescue of the cardiac defect in ErbB2 mutant mice reveals essential roles of ErbB2 in peripheral nervous system development. Neuron. 1999;23(2):273-83.
27. Sibilia M, Wagner EF. Strain-dependent epithelial defects in mice lacking the EGF receptor. Science (New York, NY). 1995;269(5221):234-8.
28. Erickson SL, O'Shea KS, Ghaboosi N, Loverro L, Frantz G, Bauer M, et al. ErbB3 is required for normal cerebellar and cardiac development: a comparison with ErbB2-and heregulin-deficient mice. Development (Cambridge, England). 1997;124(24):4999-5011.
29. Prenzel N, Fischer OM, Streit S, Hart S, Ullrich A. The epidermal growth factor receptor family as a central element for cellular signal transduction and diversification. Endocrine-related cancer. 2001;8(1):11-31.
30. Bazley LA, Gullick WJ. The epidermal growth factor receptor family. Endocrine-related cancer. 2005;12 Suppl 1:S17-27.
31. Peles E, Bacus SS, Koski RA, Lu HS, Wen D, Ogden SG, et al. Isolation of the neu/HER-2 stimulatory ligand: a 44 kd glycoprotein that induces differentiation of mammary tumor cells. Cell. 1992;69(1):205-16.
32. Li B, Rosen JM, McMenamin-Balano J, Muller WJ, Perkins AS. neu/ERBB2 cooperates with p53-172H during mammary tumorigenesis in transgenic mice. Molecular and cellular biology. 1997;17(6):3155-63.
33. Muthuswamy SK, Li D, Lelievre S, Bissell MJ, Brugge JS. ErbB2, but not ErbB1, reinitiates proliferation and induces luminal repopulation in epithelial acini. Nature cell biology. 2001;3(9):785-92.
34. Hudziak RM, Schlessinger J, Ullrich A. Increased expression of the putative growth factor receptor p185HER2 causes transformation and tumorigenesis of NIH 3T3 cells. Proceedings of the National Academy of Sciences of the United States of America. 1987;84(20):7159-63.
35. Benz CC, Scott GK, Sarup JC, Johnson RM, Tripathy D, Coronado E, et al. Estrogen-dependent, tamoxifen-resistant tumorigenic growth of MCF-7 cells transfected with HER2/neu. Breast cancer research and treatment. 1992;24(2):85-95.
36. Holliday R. Neoplastic transformation: the contrasting stability of human and mouse cells. Cancer surveys. 1996;28:103-15.

37. Muller WJ, Sinn E, Pattengale PK, Wallace R, Leder P. Single-step induction of mammary adenocarcinoma in transgenic mice bearing the activated c-neu oncogene. Cell. 1988;54(1):105-15.
38. Kiguchi K, Bol D, Carbajal S, Beltran L, Moats S, Chan K, et al. Constitutive expression of erbB2 in epidermis of transgenic mice results in epidermal hyperproliferation and spontaneous skin tumor development. Oncogene. 2000;19(37):4243-54.
39. Xie W, Chow LT, Paterson AJ, Chin E, Kudlow JE. Conditional expression of the ErbB2 oncogene elicits reversible hyperplasia in stratified epithelia and up-regulation of TGFalpha expression in transgenic mice. Oncogene. 1999;18(24):3593-607.
40. Finkle D, Quan ZR, Asghari V, Kloss J, Ghaboosi N, Mai E, et al. HER2-targeted therapy reduces incidence and progression of midlife mammary tumors in female murine mammary tumor virus huHER2-transgenic mice. Clinical cancer research : an official journal of the American Association for Cancer Research. 2004;10(7):2499-511.
41. Guy CT, Webster MA, Schaller M, Parsons TJ, Cardiff RD, Muller WJ. Expression of the neu protooncogene in the mammary epithelium of transgenic mice induces metastatic disease. Proceedings of the National Academy of Sciences of the United States of America. 1992;89(22):10578-82.
42. Bol D, Kiguchi K, Beltran L, Rupp T, Moats S, Gimenez-Conti I, et al. Severe follicular hyperplasia and spontaneous papilloma formation in transgenic mice expressing the neu oncogene under the control of the bovine keratin 5 promoter. Molecular carcinogenesis. 1998;21(1):2-12.
43. Kiguchi K, Carbajal S, Chan K, Beltran L, Ruffino L, Shen J, et al. Constitutive expression of ErbB-2 in gallbladder epithelium results in development of adenocarcinoma. Cancer research. 2001;61(19):6971-6.
44. Weinstein EJ, Kitsberg DI, Leder P. A mouse model for breast cancer induced by amplification and overexpression of the neu promoter and transgene. Molecular medicine (Cambridge, Mass). 2000;6(1):4-16.
45. Andrechek ER, Hardy WR, Siegel PM, Rudnicki MA, Cardiff RD, Muller WJ. Amplification of the neu/erbB-2 oncogene in a mouse model of mammary tumorigenesis. Proceedings of the National Academy of Sciences of the United States of America. 2000;97(7):3444-9.
46. Gustafson TL, Wellberg E, Laffin B, Schilling L, Metz RP, Zahnow CA, et al. Ha-Ras transformation of MCF10A cells leads to repression of Singleminded-2s through NOTCH and C/EBPbeta. Oncogene. 2009;28(12):1561-8.
47. Bessette DC, Tilch E, Seidens T, Quinn MC, Wiegmans AP, Shi W, et al. Using the MCF10A/MCF10CA1a Breast Cancer Progression Cell Line Model to Investigate the Effect of Active, Mutant Forms of EGFR in Breast Cancer Development and Treatment Using Gefitinib. PloS one. 2015;10(5):e0125232.
48. Herr R, Wöhrle FU, Danke C, Berens C, Brummer T. A novel MCF-10A line allowing conditional oncogene expression in 3D culture. Cell communication and signaling : CCS. 2011;9:17-.
49. Wasylishen AR, Stojanova A, Oliveri S, Rust AC, Schimmer AD, Penn LZ. New model systems provide insights into Myc-induced transformation. Oncogene. 2011;30(34):3727-34.
50. Lombardo Y, Filipovic A, Molyneux G, Periyasamy M, Giamas G, Hu Y, et al. Nicastrin regulates breast cancer stem cell properties and tumor growth in vitro and in vivo. Proceedings of the National Academy of Sciences of the United States of America. 2012;109(41):16558-63.
51. Palafox M, Ferrer I, Pellegrini P, Vila S, Hernandez-Ortega S, Urruticoechea A, et al. RANK induces epithelial-mesenchymal transition and stemness in human mammary epithelial cells and promotes tumorigenesis and metastasis. Cancer research. 2012;72(11):2879-88.
52. Ward TM, Iorns E, Liu X, Hoe N, Kim P, Singh S, et al. Truncated p110 ERBB2 induces mammary epithelial cell migration, invasion and orthotopic xenograft formation, and is associated with loss of phosphorylated STAT5. Oncogene. 2013;32(19):2463-74.

53. Zhou X, Agazie YM. The signaling and transformation potency of the overexpressed HER2 protein is dependent on the normally-expressed EGFR. Cellular signalling. 2012;24(1):140-50.
54. Pasleau F, Grooteclaes M, Gol-Winkler R. Expression of the c-erbB2 gene in the BT474 human mammary tumor cell line: measurement of c-erbB2 mRNA half-life. Oncogene. 1993;8(4):849-54.
55. Qu Y, Han B, Yu Y, Yao W, Bose S, Karlan BY, et al. Evaluation of MCF10A as a Reliable Model for Normal Human Mammary Epithelial Cells. PloS one. 2015;10(7):e0131285.
56. Imbalzano KM, Tatarkova I, Imbalzano AN, Nickerson JA. Increasingly transformed MCF-10A cells have a progressively tumor-like phenotype in three-dimensional basement membrane culture. Cancer cell international. 2009;9:7.
57. Mailleux AA, Overholtzer M, Brugge JS. Lumen formation during mammary epithelial morphogenesis: insights from in vitro and in vivo models. Cell cycle (Georgetown, Tex). 2008;7(1):57-62.
58. Arnandis T, Godinho SA. Studying centrosome function using three-dimensional cell cultures. Methods in cell biology. 2015;129:37-50.
59. Nieto MA. Epithelial-Mesenchymal Transitions in development and disease: old views and new perspectives. The International journal of developmental biology. 2009;53(8-10):1541-7.
60. Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. The Journal of clinical investigation. 2009;119(6):1420-8.
61. Thiery JP, Acloque H, Huang RY, Nieto MA. Epithelial-mesenchymal transitions in development and disease. Cell. 2009;139(5):871-90.
62. Nieto MA, Huang RY, Jackson RA, Thiery JP. EMT: 2016. Cell. 2016;166(1):21-45.
63. Leggett SE, Sim JY, Rubins JE, Neronha ZJ, Williams EK, Wong IY. Morphological single cell profiling of the epithelial-mesenchymal transition. Integrative biology : quantitative biosciences from nano to macro. 2016;8(11):1133-44.
64. Moreno-Bueno G, Peinado H, Molina P, Olmeda D, Cubillo E, Santos V, et al. The morphological and molecular features of the epithelial-to-mesenchymal transition. Nature protocols. 2009;4(11):1591-613.
65. Lamouille S, Xu J, Derynck R. Molecular mechanisms of epithelial-mesenchymal transition. Nature reviews Molecular cell biology. 2014;15(3):178-96.
66. Krebs AM, Mitschke J, Lasierra Losada M, Schmalhofer O, Boerries M, Busch H, et al. The EMT-activator Zeb1 is a key factor for cell plasticity and promotes metastasis in pancreatic cancer. Nature cell biology. 2017;19(5):518-29.
67. Ye X, Tam WL, Shibue T, Kaygusuz Y, Reinhardt F, Ng Eaton E, et al. Distinct EMT programs control normal mammary stem cells and tumour-initiating cells. Nature. 2015;525(7568):256-60.
68. Shibue T, Weinberg RA. EMT, CSCs, and drug resistance: the mechanistic link and clinical implications. Nature reviews Clinical oncology. 2017;14(10):611-29.
69. Lamouille S, Xu J, Derynck R. Molecular mechanisms of epithelial-mesenchymal transition. Nature reviews Molecular cell biology. 2014;15(3):178-96.
70. Kalluri R. EMT: when epithelial cells decide to become mesenchymal-like cells. The Journal of clinical investigation. 2009;119(6):1417-9.
71. Grande MT, Sanchez-Laorden B, Lopez-Blau C, De Frutos CA, Boutet A, Arevalo M, et al. Snail1-induced partial epithelial-to-mesenchymal transition drives renal fibrosis in mice and can be targeted to reverse established disease. Nature medicine. 2015;21(9):989-97.
72. Morel A-P, Lièvre M, Thomas C, Hinkal G, Ansieau S, Puisieux A. Generation of Breast Cancer Stem Cells through Epithelial-Mesenchymal Transition. PloS one. 2008;3(8):e2888.

73. Rohan TE, Hartwick W, Miller AB, Kandel RA. Immunohistochemical detection of c-erbB-2 and p53 in benign breast disease and breast cancer risk. Journal of the National Cancer Institute. 1998;90(17):1262-9.
74. van de Vijver MJ, Peterse JL, Mooi WJ, Wisman P, Lomans J, Dalesio O, et al. Neu-protein overexpression in breast cancer. Association with comedo-type ductal carcinoma in situ and limited prognostic value in stage II breast cancer. The New England journal of medicine. 1988;319(19):1239-45.
75. Gusterson BA, Machin LG, Gullick WJ, Gibbs NM, Powles TJ, Price P, et al. Immunohistochemical distribution of c-erbB-2 in infiltrating and in situ breast cancer. International journal of cancer. 1988;42(6):842-5.
76. Collins LC, Schnitt SJ. HER2 protein overexpression in estrogen receptor-positive ductal carcinoma in situ of the breast: frequency and implications for tamoxifen therapy. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc. 2005;18(5):615-20.
77. Ridolfi RL, Jamehdor MR, Arber JM. HER-2/neu testing in breast carcinoma: a combined immunohistochemical and fluorescence in situ hybridization approach. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc. 2000;13(8):866-73.
78. Quenel N, Wafflart J, Bonichon F, de Mascarel I, Trojani M, Durand M, et al. The prognostic value of c-erbB2 in primary breast carcinomas: a study on 942 cases. Breast cancer research and treatment. 1995;35(3):283-91.
79. Weinstein EJ, Kitsberg DI, Leder P. A mouse model for breast cancer induced by amplification and overexpression of the neu promoter and transgene. Molecular medicine (Cambridge, Mass). 2000;6(1):4-16.
80. Bobrow LG, Happerfield LC, Gregory WM, Springall RD, Millis RR. The classification of ductal carcinoma in situ and its association with biological markers. Seminars in diagnostic pathology. 1994;11(3):199-207.
81. Soslow RA, Carlson DL, Horenstein MG, Osborne MP. A comparison of cell cycle markers in well-differentiated lobular and ductal carcinomas. Breast cancer research and treatment. 2000;61(2):161-70.
82. Carlsson J, Nordgren H, Sjöström J, Wester K, Villman K, Bengtsson NO, et al. HER2 expression in breast cancer primary tumours and corresponding metastases. Original data and literature review. British journal of cancer. 2004;90(12):2344-8.
83. Simon R, Nocito A, Hübscher T, Bucher C, Torhorst J, Schraml P, et al. Patterns of HER-2/neu Amplification and Overexpression in Primary and Metastatic Breast Cancer. JNCI: Journal of the National Cancer Institute. 2001;93(15):1141-6.
84. Gancberg D, Di Leo A, Cardoso F, Rouas G, Pedrocchi M, Paesmans M, et al. Comparison of HER-2 status between primary breast cancer and corresponding distant metastatic sites. Annals of Oncology. 2002;13(7):1036-43.
85. Masood S, Bui MM. Assessment of Her-2/neu overexpression in primary breast cancers and their metastatic lesions: an immunohistochemical study. Annals of clinical and laboratory science. 2000;30(3):259-65.
86. Tsutsui S, Ohno S, Murakami S, Kataoka A, Kinoshita J, Hachitanda Y. EGFR, c-erbB2 and p53 protein in the primary lesions and paired metastatic regional lymph nodes in breast cancer. European journal of surgical oncology : the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology. 2002;28(4):383-7.
87. Shimizu C, Fukutomi T, Tsuda H, Akashi-Tanaka S, Watanabe T, Nanasawa T, et al. c-erbB-2 protein overexpression and p53 immunoreaction in primary and recurrent breast cancer tissues. Journal of surgical oncology. 2000;73(1):17-20.

88. Mitri Z, Constantine T, O'Regan R. The HER2 Receptor in Breast Cancer: Pathophysiology, Clinical Use, and New Advances in Therapy. Chemotherapy research and practice. 2012;2012:743193-.
89. Luo H, Xu X, Ye M, Sheng B, Zhu X. The prognostic value of HER2 in ovarian cancer: A meta-analysis of observational studies. PloS one. 2018;13(1):e0191972-e.
90. Yan M, Schwaederle M, Arguello D, Millis SZ, Gatalica Z, Kurzrock R. HER2 expression status in diverse cancers: review of results from 37,992 patients. Cancer metastasis reviews. 2015;34(1):157-64.
91. Freudenberg JA, Wang Q, Katsumata M, Drebin J, Nagatomo I, Greene MI. The role of HER2 in early breast cancer metastasis and the origins of resistance to HER2-targeted therapies. Experimental and molecular pathology. 2009;87(1):1-11.
92. Latta EK, Tjan S, Parkes RK, O'Malley FP. The role of HER2/neu overexpression/amplification in the progression of ductal carcinoma in situ to invasive carcinoma of the breast. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc. 2002;15(12):1318-25.
93. Tsuda H, Akiyama F, Terasaki H, Hasegawa T, Kurosumi M, Shimadzu M, et al. Detection of HER-2/neu (c-erb B-2) DNA amplification in primary breast carcinoma. Interobserver reproducibility and correlation with immunohistochemical HER-2 overexpression. Cancer. 2001;92(12):2965-74.
94. Carlsson J, Nordgren H, Sjostrom J, Wester K, Villman K, Bengtsson NO, et al. HER2 expression in breast cancer primary tumours and corresponding metastases. Original data and literature review. British journal of cancer. 2004;90(12):2344-8.
95. Park K, Han S, Kim HJ, Kim J, Shin E. HER2 status in pure ductal carcinoma in situ and in the intraductal and invasive components of invasive ductal carcinoma determined by fluorescence in situ hybridization and immunohistochemistry. Histopathology. 2006;48(6):702-7.
96. Gancberg D, Jarvinen T, di Leo A, Rouas G, Cardoso F, Paesmans M, et al. Evaluation of HER-2/NEU protein expression in breast cancer by immunohistochemistry: an interlaboratory study assessing the reproducibility of HER-2/NEU testing. Breast cancer research and treatment. 2002;74(2):113-20.
97. Gilcrease MZ, Woodward WA, Nicolas MM, Corley LJ, Fuller GN, Esteva FJ, et al. Even low-level HER2 expression may be associated with worse outcome in node-positive breast cancer. Am J Surg Pathol. 2009;33(5):759-67.
98. Wolff AC, Hammond ME, Hicks DG, Dowsett M, McShane LM, Allison KH, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2013;31(31):3997-4013.
99. Wolff AC, Hammond MEH, Allison KH, Harvey BE, Mangu PB, Bartlett JMS, et al. Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer: American Society of Clinical Oncology/College of American Pathologists Clinical Practice Guideline Focused Update. Archives of pathology & laboratory medicine. 2018;142(11):1364-82.
100. Bartlett JM, Starczynski J. Quantitative reverse transcriptase polymerase chain reaction and the Oncotype DX test for assessment of human epidermal growth factor receptor 2 status: time to reflect again? Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2011;29(32):4219-21.
101. Murphy CG, Modi S. HER2 breast cancer therapies: a review. Biologics : targets & therapy. 2009;3:289-301.
102. Baselga J, Cortes J, Im SA, Clark E, Ross G, Kiermaier A, et al. Biomarker analyses in CLEOPATRA: a phase III, placebo-controlled study of pertuzumab in human epidermal growth

factor receptor 2-positive, first-line metastatic breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2014;32(33):3753-61.
103. Kaufman PA, Broadwater G, Lezon-Geyda K, Dressler LG, Berry D, Friedman P, et al. CALGB 150002: Correlation of HER2 and chromosome 17 (ch17) copy number with trastuzumab (T) efficacy in CALGB 9840, paclitaxel (P) with or without T in HER2+ and HER2- metastatic breast cancer (MBC). 2007;25(18_suppl):1009-.
104. Wang J, Xu B. Targeted therapeutic options and future perspectives for HER2-positive breast cancer. Signal transduction and targeted therapy. 2019;4:34.
105. Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, et al. Use of Chemotherapy plus a Monoclonal Antibody against HER2 for Metastatic Breast Cancer That Overexpresses HER2. 2001;344(11):783-92.
106. Romond EH, Perez EA, Bryant J, Suman VJ, Geyer CE, Davidson NE, et al. Trastuzumab plus Adjuvant Chemotherapy for Operable HER2-Positive Breast Cancer. 2005;353(16):1673-84.
107. Burris HA, 3rd. Dual kinase inhibition in the treatment of breast cancer: initial experience with the EGFR/ErbB-2 inhibitor lapatinib. The oncologist. 2004;9 Suppl 3:10-5.
108. O'Donovan N, Byrne AT, O'Connor AE, McGee S, Gallagher WM, Crown JJIND. Synergistic interaction between trastuzumab and EGFR/HER-2 tyrosine kinase inhibitors in HER-2 positive breast cancer cells. 2011;29(5):752-9.
109. Burstein HJ, Storniolo AM, Franco S, Forster J, Stein S, Rubin S, et al. A phase II study of lapatinib monotherapy in chemotherapy-refractory HER2-positive and HER2-negative advanced or metastatic breast cancer. Annals of oncology : official journal of the European Society for Medical Oncology. 2008;19(6):1068-74.
110. Blackwell KL, Burstein HJ, Storniolo AM, Rugo H, Sledge G, Koehler M, et al. Randomized study of Lapatinib alone or in combination with trastuzumab in women with ErbB2-positive, trastuzumab-refractory metastatic breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2010;28(7):1124-30.
111. Blackwell KL, Burstein HJ, Storniolo AM, Rugo H, Sledge G, Koehler M, et al. Randomized Study of Lapatinib Alone or in Combination With Trastuzumab in Women With ErbB2-Positive, Trastuzumab-Refractory Metastatic Breast Cancer. 2010;28(7):1124-30.
112. Trimble WS, Boulianne GL, Hozumi N. Morphological transformation and tumorigenicity in C3H/10T1/2 cells transformed with an inducible c-Ha-ras oncogene. Bioscience reports. 1987;7(7):579-85.
113. Loberg MA, Bell RK, Goodwin LO, Eudy E, Miles LA, SanMiguel JM, et al. Sequentially inducible mouse models reveal that Npm1 mutation causes malignant transformation of Dnmt3a-mutant clonal hematopoiesis. Leukemia. 2019.
114. Wong AK, Chin L. An inducible melanoma model implicates a role for RAS in tumor maintenance and angiogenesis. Cancer Metastasis Rev. 2000;19(1-2):121-9.
115. Carter EP, Gopsill JA, Gomm JJ, Jones JL, Grose RP. A 3D in vitro model of the human breast duct: a method to unravel myoepithelial-luminal interactions in the progression of breast cancer. Breast cancer research : BCR. 2017;19(1):50.
116. Meerbrey KL, Hu G, Kessler JD, Roarty K, Li MZ, Fang JE, et al. The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo. Proceedings of the National Academy of Sciences of the United States of America. 2011;108(9):3665-70.
117. Baron U, Freundlieb S, Gossen M, Bujard H. Co-regulation of two gene activities by tetracycline via a bidirectional promoter. Nucleic acids research. 1995;23(17):3605-6.
118. Le Y, Sauer B. Conditional gene knockout using Cre recombinase. Molecular biotechnology. 2001;17(3):269-75.
119. Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature. 1997;389(6648):251-60.

120. Venkatesh S, Workman JL. Histone exchange, chromatin structure and the regulation of transcription. Nature reviews Molecular cell biology. 2015;16(3):178-89.
121. Tremethick DJ. Higher-order structures of chromatin: the elusive 30 nm fiber. Cell. 2007;128(4):651-4.
122. Woodcock CL, Dimitrov S. Higher-order structure of chromatin and chromosomes. Current opinion in genetics & development. 2001;11(2):130-5.
123. Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, et al. The accessible chromatin landscape of the human genome. Nature. 2012;489(7414):75-82.
124. Jin B, Li Y, Robertson KD. DNA methylation: superior or subordinate in the epigenetic hierarchy? Genes & cancer. 2011;2(6):607-17.
125. Kulis M, Esteller M. DNA methylation and cancer. Advances in genetics. 2010;70:27-56.
126. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, et al. High-resolution profiling of histone methylations in the human genome. Cell. 2007;129(4):823-37.
127. Carrozza MJ, Li B, Florens L, Suganuma T, Swanson SK, Lee KK, et al. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell. 2005;123(4):581-92.
128. Graff J, Tsai LH. Histone acetylation: molecular mnemonics on the chromatin. Nature reviews Neuroscience. 2013;14(2):97-111.
129. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. Cell research. 2011;21(3):381-95.
130. Eberharter A, Becker PB. Histone acetylation: a switch between repressive and permissive chromatin. Second in review series on chromatin dynamics. EMBO reports. 2002;3(3):224-9.
131. Miotto B. Kinases and chromatin structure: who regulates whom? Epigenetics. 2013;8(10):1008-12.
132. Stergachis AB, Neph S, Reynolds A, Humbert R, Miller B, Paige SL, et al. Developmental fate and cellular maturity encoded in human regulatory DNA landscapes. Cell. 2013;154(4):888-903.
133. Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, et al. The accessible chromatin landscape of the human genome. Nature. 2012;489(7414):75-82.
134. Sethi N, Kang Y. Unravelling the complexity of metastasis - molecular understanding and targeted therapies. Nature reviews Cancer. 2011;11(10):735-48.
135. Denny SK, Yang D, Chuang CH, Brady JJ, Lim JS, Gruner BM, et al. Nfib Promotes Metastasis through a Widespread Increase in Chromatin Accessibility. Cell. 2016;166(2):328-42.
136. Simon JM, Hacker KE, Singh D, Brannon AR, Parker JS, Weiser M, et al. Variation in chromatin accessibility in human kidney cancer links H3K36 methyltransferase loss with widespread RNA processing defects. Genome research. 2014;24(2):241-50.
137. Qu K, Zaba LC, Satpathy AT, Giresi PG, Li R, Jin Y, et al. Chromatin Accessibility Landscape of Cutaneous T Cell Lymphoma and Dynamic Response to HDAC Inhibitors. Cancer cell. 2017;32(1):27-41.e4.
138. Zhou ZH, Wang QL, Mao LH, Li XQ, Liu P, Song JW, et al. Chromatin accessibility changes are associated with enhanced growth and liver metastasis capacity of acid-adapted colorectal cancer cells. Cell cycle (Georgetown, Tex). 2019;18(4):511-22.
139. Kelso TWR, Porter DK, Amaral ML, Shokhirev MN, Benner C, Hargreaves DC. Chromatin accessibility underlies synthetic lethality of SWI/SNF subunits in ARID1A-mutant cancers. eLife. 2017;6.
140. Kang X, Feng Y, Gan Z, Zeng S, Guo X, Chen X, et al. NASP antagonize chromatin accessibility through maintaining histone H3K9me1 in hepatocellular carcinoma. Biochimica et biophysica acta Molecular basis of disease. 2018;1864(10):3438-48.

141. Pogna EA, Clayton AL, Mahadevan LC. Signalling to chromatin through post-translational modifications of HMGN. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms. 2010;1799(1):93-100.
142. Clayton AL, Mahadevan LC. MAP kinase-mediated phosphoacetylation of histone H3 and inducible gene regulation. FEBS Letters. 2003;546(1):51-8.
143. Hazzalin CA, Mahadevan LC. MAPK-Regulated transcription: A continuously variable gene switch? Nature Reviews Molecular Cell Biology. 2002;3(1):30-40.
144. Wingelhofer B, Neubauer HA, Valent P, Han X, Constantinescu SN, Gunning PT, et al. Implications of STAT3 and STAT5 signaling on gene regulation and chromatin remodeling in hematopoietic cancer. Leukemia. 2018;32(8):1713-26.
145. Wagatsuma K, Tani-ichi S, Liang B, Shitara S, Ishihara K, Abe M, et al. STAT5 Orchestrates Local Epigenetic Changes for Chromatin Accessibility and Rearrangements by Direct Binding to the TCRγ Locus. 2015;195(4):1804-14.
146. Arzate-Mejia RG, Valle-Garcia D, Recillas-Targa F. Signaling epigenetics: novel insights on cell signaling and epigenetic regulation. IUBMB life. 2011;63(10):881-95.
147. Sotiriou C, Neo SY, McShane LM, Korn EL, Long PM, Jazaeri A, et al. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. Proceedings of the National Academy of Sciences of the United States of America. 2003;100(18):10393-8.
148. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature. 2000;406(6797):747-52.
149. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proceedings of the National Academy of Sciences of the United States of America. 2001;98(19):10869-74.
150. Pechoux C, Gudjonsson T, Ronnov-Jessen L, Bissell MJ, Petersen OW. Human mammary luminal epithelial cells contain progenitors to myoepithelial cells. Developmental biology. 1999;206(1):88-99.
151. Smith GH. Experimental mammary epithelial morphogenesis in an in vivo model: evidence for distinct cellular progenitors of the ductal and lobular phenotype. Breast cancer research and treatment. 1996;39(1):21-31.
152. Stingl J, Eaves CJ, Zandieh I, Emerman JT. Characterization of bipotent mammary epithelial progenitor cells in normal adult human breast tissue. Breast cancer research and treatment. 2001;67(2):93-109.
153. Shackleton M, Quintana E, Fearon ER, Morrison SJ. Heterogeneity in cancer: cancer stem cells versus clonal evolution. Cell. 2009;138(5):822-9.
154. Stingl J. Detection and analysis of mammary gland stem cells. The Journal of pathology. 2009;217(2):229-41.
155. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF. Prospective identification of tumorigenic breast cancer cells. Proceedings of the National Academy of Sciences of the United States of America. 2003;100(7):3983-8.
156. Ginestier C, Hur MH, Charafe-Jauffret E, Monville F, Dutcher J, Brown M, et al. ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. Cell stem cell. 2007;1(5):555-67.
157. Fillmore CM, Kuperwasser C. Human breast cancer cell lines contain stem-like cells that self-renew, give rise to phenotypically diverse progeny and survive chemotherapy. Breast cancer research : BCR. 2008;10(2):R25-R.
158. Tomita H, Tanaka K, Tanaka T, Hara A. Aldehyde dehydrogenase 1A1 in stem cells and cancer. Oncotarget. 2016;7(10):11018-32.
159. Morimoto K, Kim SJ, Tanei T, Shimazu K, Tanji Y, Taguchi T, et al. Stem cell marker aldehyde dehydrogenase 1-positive breast cancers are characterized by negative estrogen

receptor, positive human epidermal growth factor receptor type 2, and high Ki67 expression. Cancer science. 2009;100(6):1062-8.
160. Bloom HJ, Richardson WW. Histological grading and prognosis in breast cancer; a study of 1409 cases of which 359 have been followed for 15 years. British journal of cancer. 1957;11(3):359-77.
161. Heppner GH, Miller BE. Tumor heterogeneity: biological implications and therapeutic consequences. Cancer Metastasis Rev. 1983;2(1):5-23.
162. Yates LR, Gerstung M, Knappskog S, Desmedt C, Gundem G, Van Loo P, et al. Subclonal diversification of primary breast cancer revealed by multiregion sequencing. Nature medicine. 2015;21(7):751-9.
163. Sottoriva A, Kang H, Ma Z, Graham TA, Salomon MP, Zhao J, et al. A Big Bang model of human colorectal tumor growth. Nature genetics. 2015;47(3):209-16.
164. Schwarz RF, Ng CK, Cooke SL, Newman S, Temple J, Piskorz AM, et al. Spatial and temporal heterogeneity in high-grade serous ovarian cancer: a phylogenetic analysis. PLoS medicine. 2015;12(2):e1001789.
165. Bashashati A, Ha G, Tone A, Ding J, Prentice LM, Roth A, et al. Distinct evolutionary trajectories of primary high-grade serous ovarian cancers revealed through spatial mutational profiling. The Journal of pathology. 2013;231(1):21-34.
166. Kumar A, Boyle EA, Tokita M, Mikheev AM, Sanger MC, Girard E, et al. Deep sequencing of multiple regions of glial tumors reveals spatial heterogeneity for mutations in clinically relevant genes. Genome biology. 2014;15(12):530.
167. Gerlinger M, Horswell S, Larkin J, Rowan AJ, Salm MP, Varela I, et al. Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. Nature genetics. 2014;46(3):225-33.
168. Sinha VC, Piwnica-Worms HJJoMGB, Neoplasia. Intratumoral Heterogeneity in Ductal Carcinoma In Situ: Chaos and Consequence. 2018;23(4):191-205.
169. Allred DC, Wu Y, Mao S, Nagtegaal ID, Lee S, Perou CM, et al. Ductal Carcinoma In situ and the Emergence of Diversity during Breast Cancer Evolution. Clinical Cancer Research. 2008;14(2):370.
170. Sun X-x, Yu Q. Intra-tumor heterogeneity of cancer cells and its implications for cancer treatment. Acta Pharmacol Sin. 2015;36(10):1219-27.
171. Locke WJ, Clark SJ. Epigenome remodelling in breast cancer: insights from an early in vitro model of carcinogenesis. Breast cancer research : BCR. 2012;14(6):215-.
172. Park SY, Kwon HJ, Lee HE, Ryu HS, Kim S-W, Kim JH, et al. Promoter CpG island hypermethylation during breast cancer progression. 2011;458(1):73-84.
173. Johnson KC, Koestler DC, Fleischer T, Chen P, Jenson EG, Marotti JD, et al. DNA methylation in ductal carcinoma in situ related with future development of invasive breast cancer. Clin Epigenetics. 2015;7(1):75-.
174. Kaur H, Mao S, Li Q, Sameni M, Krawetz SA, Sloane BF, et al. RNA-Seq of human breast ductal carcinoma in situ models reveals aldehyde dehydrogenase isoform 5A1 as a novel potential target. PloS one. 2012;7(12):e50249-e.
175. Abba MC, Gong T, Lu Y, Lee J, Zhong Y, Lacunza E, et al. A Molecular Portrait of High-Grade Ductal Carcinoma In Situ. Cancer research. 2015;75(18):3980-90.
176. Allred DC. Ductal carcinoma in situ: terminology, classification, and natural history. J Natl Cancer Inst Monogr. 2010;2010(41):134-8.
177. Pinder SE. Ductal carcinoma in situ (DCIS): pathological features, differential diagnosis, prognostic factors and specimen evaluation. Modern Pathology. 2010;23(2):S8-S13.
178. Gorringe KL, Fox SB. Ductal Carcinoma In Situ Biology, Biomarkers, and Diagnosis. Frontiers in oncology. 2017;7:248-.

179. Ko EC, Jhala NC, Shultz JJ, Chhieng DC. Use of a Panel of Markers in the Differential Diagnosis of Adenocarcinoma and Reactive Mesothelial Cells in Fluid Cytology. American journal of clinical pathology. 2001;116(5):709-15.
180. Lennington WJ, Jensen RA, Dalton LW, Page DL. Ductal carcinoma in situ of the breast. Heterogeneity of individual lesions. Cancer. 1994;73(1):118-24.
181. Perez AA, Balabram D, Salles MdA, Gobbi H. Ductal carcinoma in situ of the breast: correlation between histopathological features and age of patients. Diagn Pathol. 2014;9:227-.
182. Scripcaru G, Zardawi IM. Mammary ductal carcinoma in situ: a fresh look at architectural patterns. Int J Surg Oncol. 2012;2012:979521-.
183. Lari SA, Kuerer HM. Biological Markers in DCIS and Risk of Breast Recurrence: A Systematic Review. J Cancer. 2011;2:232-61.
184. Latta EK, Tjan S, Parkes RK, O'Malley FP. The Role of HER2/neu Overexpression/Amplification in the Progression of Ductal Carcinoma In Situ to Invasive Carcinoma of the Breast. Modern Pathology. 2002;15(12):1318-25.
185. Lebeau A, Unholzer A, Amann G, Kronawitter M, Bauerfeind I, Sendelhofert A, et al. EGFR, HER-2/neu, Cyclin D1, p21 and p53 in Correlation to Cell Proliferation and Steroid Hormone Receptor Status in Ductal Carcinoma in situ of the Breast. 2003;79(2):187-98.
186. Rakovitch E, Nofech-Mozes S, Hanna W, Narod S, Thiruchelvam D, Saskin R, et al. HER2/neu and Ki-67 expression predict non-invasive recurrence following breast-conserving therapy for ductal carcinoma in situ. British journal of cancer. 2012;106(6):1160-5.
187. Done SJ, Arneson NCR, Özçelik H, Redston M, Andrulis IL. p53 Mutations in Mammary Ductal Carcinoma in Situ but not in Epithelial Hyperplasias. Cancer research. 1998;58(4):785.
188. Miranda A, Hamilton PT, Zhang AW, Pattnaik S, Becht E, Mezheyeuski A, et al. Cancer stemness, intratumoral heterogeneity, and immune response across cancers. Proceedings of the National Academy of Sciences. 2019;116(18):9020.
189. Gerdes MJ, Sood A, Sevinsky C, Pris AD, Zavodszky MI, Ginty F. Emerging understanding of multiscale tumor heterogeneity. Frontiers in oncology. 2014;4:366.
190. Michor F, Polyak K. The origins and implications of intratumor heterogeneity. Cancer prevention research (Philadelphia, Pa). 2010;3(11):1361-4.
191. Plaks V, Kong N, Werb Z. The cancer stem cell niche: how essential is the niche in regulating stemness of tumor cells? Cell stem cell. 2015;16(3):225-38.
192. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer Genome Landscapes. Science (New York, NY). 2013;339(6127):1546.
193. Dick JE. Stem cell concepts renew cancer research. Blood. 2008;112(13):4793-807.
194. Marusyk A, Tabassum DP, Altrock PM, Almendro V, Michor F, Polyak K. Non-cell-autonomous driving of tumour growth supports sub-clonal heterogeneity. Nature. 2014;514(7520):54-8.
195. Sparmann A, van Lohuizen M. Polycomb silencers control cell fate, development and cancer. Nature reviews Cancer. 2006;6(11):846-56.
196. Abramyuk A, Wolf G, Hietschold V, Haberland U, van den Hoff J, Abolmaali N. Comment on "Developing DCE-CT to quantify intra-tumor heterogeneity in breast tumors with differing angiogenic phenotype". IEEE transactions on medical imaging. 2010;29(4):1088-9; author reply 9-92.
197. Nowell PC. The clonal evolution of tumor cell populations. Science (New York, NY). 1976;194(4260):23-8.
198. Navin N, Krasnitz A, Rodgers L, Cook K, Meth J, Kendall J, et al. Inferring tumor progression from genomic heterogeneity. Genome research. 2010;20(1):68-80.
199. Nassar A, Radhakrishnan A, Cabrero IA, Cotsonis GA, Cohen C. Intratumoral heterogeneity of immunohistochemical marker expression in breast carcinoma: a tissue

microarray-based study. Applied immunohistochemistry & molecular morphology : AIMM. 2010;18(5):433-41.
200. Allison KH, Dintzis SM, Schmidt RA. Frequency of HER2 heterogeneity by fluorescence in situ hybridization according to CAP expert panel recommendations: time for a new look at how to report heterogeneity. American journal of clinical pathology. 2011;136(6):864-71.
201. Seol H, Lee HJ, Choi Y, Lee HE, Kim YJ, Kim JH, et al. Intratumoral heterogeneity of HER2 gene amplification in breast cancer: its clinicopathological significance. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc. 2012;25(7):938-48.
202. Vance GH, Barry TS, Bloom KJ, Fitzgibbons PL, Hicks DG, Jenkins RB, et al. Genetic heterogeneity in HER2 testing in breast cancer: panel summary and guidelines. Archives of pathology & laboratory medicine. 2009;133(4):611-2.
203. Beca F, Polyak K. Intratumor Heterogeneity in Breast Cancer. Adv Exp Med Biol. 2016;882:169-89.
204. Chhieng DC, Frost AR, Niwas S, Weiss H, Grizzle WE, Beeken S. Intratumor heterogeneity of biomarker expression in breast carcinomas. Biotechnic & histochemistry : official publication of the Biological Stain Commission. 2004;79(1):25-36.
205. Glockner S, Buurman H, Kleeberger W, Lehmann U, Kreipe H. Marked intratumoral heterogeneity of c-myc and cyclinD1 but not of c-erbB2 amplification in breast cancer. Laboratory investigation; a journal of technical methods and pathology. 2002;82(10):1419-26.
206. Siitonen SM, Isola JJ, Rantala IS, Helin HJ. Intratumor variation in cell proliferation in breast carcinoma as determined by antiproliferating cell nuclear antigen monoclonal antibody and automated image analysis. American journal of clinical pathology. 1993;99(3):226-31.
207. Pasquali L, Bedeir A, Ringquist S, Styche A, Bhargava R, Trucco G. Quantification of CpG island methylation in progressive breast lesions from normal to invasive carcinoma. Cancer letters. 2007;257(1):136-44.
208. Reynolds PA, Sigaroudinia M, Zardo G, Wilson MB, Benton GM, Miller CJ, et al. Tumor suppressor p16INK4A regulates polycomb-mediated DNA hypermethylation in human mammary epithelial cells. The Journal of biological chemistry. 2006;281(34):24790-802.
209. Kim EJ, Seo AN, Jang MH, Park SY, Lee HJ, Kim JH, et al. HER2 Heterogeneity Affects Trastuzumab Responses and Survival in Patients With HER2-Positive Metastatic Breast Cancer. American journal of clinical pathology. 2014;142(6):755-66.
210. Liu S, Cong Y, Wang D, Sun Y, Deng L, Liu Y, et al. Breast cancer stem cells transition between epithelial and mesenchymal states reflective of their normal counterparts. Stem cell reports. 2014;2(1):78-91.
211. Diessner J, Bruttel V, Stein RG, Horn E, Hausler SF, Dietl J, et al. Targeting of preexisting and induced breast cancer stem cells with trastuzumab and trastuzumab emtansine (T-DM1). Cell death & disease. 2014;5:e1149.
212. Shah D, Osipo C. Cancer stem cells and HER2 positive breast cancer: The story so far. Genes & diseases. 2016;3(2):114-23.
213. Korkaya H, Paulson A, Charafe-Jauffret E, Ginestier C, Brown M, Dutcher J, et al. Regulation of mammary stem/progenitor cells by PTEN/Akt/beta-catenin signaling. PLoS biology. 2009;7(6):e1000121.
214. Ginestier C, Liu S, Diebel ME, Korkaya H, Luo M, Brown M, et al. CXCR1 blockade selectively targets human breast cancer stem cells in vitro and in xenografts. The Journal of clinical investigation. 2010;120(2):485-97.
215. Singh JK, Farnie G, Bundred NJ, Simões BM, Shergill A, Landberg G, et al. Targeting CXCR1/2 Significantly Reduces Breast Cancer Stem Cell Activity and Increases the Efficacy of Inhibiting HER2 via HER2-Dependent and -Independent Mechanisms. 2013;19(3):643-56.

216. van Leenders GJ, Sookhlall R, Teubel WJ, de Ridder CM, Reneman S, Sacchetti A, et al. Activation of c-MET induces a stem-like phenotype in human prostate cancer. PloS one. 2011;6(11):e26753.
217. Albino D, Civenni G, Rossi S, Mitra A, Catapano CV, Carbone GM. The ETS factor ESE3/EHF represses IL-6 preventing STAT3 activation and expansion of the prostate cancer stem-like compartment. Oncotarget. 2016;7(47):76756-68.
218. Ma Z, Cui X, Lu L, Chen G, Yang Y, Hu Y, et al. Exosomes from glioma cells induce a tumor-like phenotype in mesenchymal stem cells by activating glycolysis. Stem cell research & therapy. 2019;10(1):60.
219. Scaffidi P, Misteli T. In vitro generation of human cells with cancer stem cell properties. Nature cell biology. 2011;13(9):1051-61.
220. Chen Y, Shi L, Zhang L, Li R, Liang J, Yu W, et al. The molecular mechanism governing the oncogenic potential of SOX2 in breast cancer. The Journal of biological chemistry. 2008;283(26):17969-78.
221. Wei D, Kanai M, Huang S, Xie K. Emerging role of KLF4 in human gastrointestinal cancer. Carcinogenesis. 2006;27(1):23-31.
222. You JS, Jones PA. Cancer genetics and epigenetics: two sides of the same coin? Cancer cell. 2012;22(1):9-20.
223. Chatterjee A, Rodger EJ, Eccles MR. Epigenetic drivers of tumourigenesis and cancer metastasis. Seminars in cancer biology. 2018;51:149-59.
224. Efroni S, Duttagupta R, Cheng J, Dehghani H, Hoeppner DJ, Dash C, et al. Global transcription in pluripotent embryonic stem cells. Cell stem cell. 2008;2(5):437-47.
225. Yamanaka S. Induced pluripotent stem cells: past, present, and future. Cell stem cell. 2012;10(6):678-84.
226. Koche RP, Smith ZD, Adli M, Gu H, Ku M, Gnirke A, et al. Reprogramming factor expression initiates widespread targeted chromatin remodeling. Cell stem cell. 2011;8(1):96-105.
227. Suvà ML, Riggi N, Bernstein BE. Epigenetic reprogramming in cancer. Science (New York, NY). 2013;339(6127):1567-70.
228. Corces MR, Trevino AE, Hamilton EG, Greenside PG, Sinnott-Armstrong NA, Vesuna S, et al. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nature methods. 2017;14(10):959-62.
229. Leitner NR, Lassnig C, Rom R, Heider S, Bago-Horvath Z, Eferl R, et al. Inducible, Dose-Adjustable and Time-Restricted Reconstitution of Stat1 Deficiency In Vivo. PloS one. 2014;9(1):e86608.
230. Koudelakova V, Berkovcova J, Trojanec R, Vrbkova J, Radova L, Ehrmann J, et al. Evaluation of HER2 Gene Status in Breast Cancer Samples with Indeterminate Fluorescence in Situ Hybridization by Quantitative Real-Time PCR. The Journal of molecular diagnostics : JMD. 2015;17(4):446-55.
231. Park S, Wang H-Y, Kim S, Ahn S, Lee D, Cho Y, et al. Quantitative RT-PCR assay of HER2 mRNA expression in formalin-fixed and paraffin-embedded breast cancer tissues. Int J Clin Exp Pathol. 2014;7(10):6752-9.
232. Wang H-y, Kim S, Park S, Kim S, Jung D, Park KH, et al. Evaluation of a quantitative RT-PCR assay to detect HER2 mRNA overexpression for diagnosis and selection of trastuzumab therapy in breast cancer tissue samples. Experimental and Molecular Pathology. 2014;97(3):368-74.
233. Chen CT, Kim H, Liska D, Gao S, Christensen JG, Weiser MR. MET activation mediates resistance to lapatinib inhibition of HER2-amplified gastric cancer cells. Molecular cancer therapeutics. 2012;11(3):660-9.

234. Nagaria TS, Shi C, Leduc C, Hoskin V, Sikdar S, Sangrar W, et al. Combined targeting of Raf and Mek synergistically inhibits tumorigenesis in triple negative breast cancer model systems. Oncotarget. 2017;8(46):80804-19.
235. Konecny GE, Pegram MD, Venkatesan N, Finn R, Yang G, Rahmeh M, et al. Activity of the dual kinase inhibitor lapatinib (GW572016) against HER-2-overexpressing and trastuzumab-treated breast cancer cells. Cancer research. 2006;66(3):1630-9.
236. Dong H, Ma L, Gan J, Lin W, Chen C, Yao Z, et al. PTPRO represses ERBB2-driven breast oncogenesis by dephosphorylation and endosomal internalization of ERBB2. Oncogene. 2017;36(3):410-22.
237. Benlimame N, He Q, Jie S, Xiao D, Xu YJ, Loignon M, et al. FAK signaling is critical for ErbB-2/ErbB-3 receptor cooperation for oncogenic transformation and invasion. The Journal of cell biology. 2005;171(3):505-16.
238. Huang TH, Wu F, Loeb GB, Hsu R, Heidersbach A, Brincat A, et al. Up-regulation of miR-21 by HER2/neu signaling promotes cell invasion. The Journal of biological chemistry. 2009;284(27):18515-24.
239. Ueda Y, Wang S, Dumont N, Yi JY, Koh Y, Arteaga CL. Overexpression of HER2 (erbB2) in human breast epithelial cells unmasks transforming growth factor beta-induced cell motility. The Journal of biological chemistry. 2004;279(23):24505-13.
240. Kang JS, Krauss RS. Ras induces anchorage-independent growth by subverting multiple adhesion-regulated cell cycle events. Molecular and cellular biology. 1996;16(7):3370-80.
241. Colburn NH, Bruegge WFV, Bates JR, Gray RH, Rossen JD, Kelsey WH, et al. Correlation of Anchorage-independent Growth with Tumorigenicity of Chemically Transformed Mouse Epidermal Cells. Cancer research. 1978;38(3):624.
242. Cifone MA, Fidler IJ. Correlation of patterns of anchorage-independent growth with in vivo behavior of cells from a murine fibrosarcoma. Proceedings of the National Academy of Sciences of the United States of America. 1980;77(2):1039-43.
243. Moody SE, Sarkisian CJ, Hahn KT, Gunther EJ, Pickup S, Dugan KD, et al. Conditional activation of Neu in the mammary epithelium of transgenic mice results in reversible pulmonary metastasis. Cancer cell. 2002;2(6):451-61.
244. Moody SE, Perez D, Pan TC, Sarkisian CJ, Portocarrero CP, Sterner CJ, et al. The transcriptional repressor Snail promotes mammary tumor recurrence. Cancer cell. 2005;8(3):197-209.
245. Pradeep CR, Zeisel A, Köstler WJ, Lauriola M, Jacob-Hirsch J, Haibe-Kains B, et al. Modeling invasive breast cancer: growth factors propel progression of HER2-positive premalignant lesions. Oncogene. 2012;31(31):3569-83.
246. Yu Y, Lowy MM, Elble RC. Tet-On lentiviral transductants lose inducibility when silenced for extended intervals in mammary epithelial cells. Metab Eng Commun. 2016;3:64-7.
247. Sung YM, Xu X, Sun J, Mueller D, Sentissi K, Johnson P, et al. Tumor suppressor function of Syk in human MCF10A in vitro and normal mouse mammary epithelium in vivo. PloS one. 2009;4(10):e7445.
248. Zhu T, Starling-Emerald B, Zhang X, Lee KO, Gluckman PD, Mertani HC, et al. Oncogenic transformation of human mammary epithelial cells by autocrine human growth hormone. Cancer research. 2005;65(1):317-24.
249. Becker LE, Takwi AA, Lu Z, Li Y. The role of miR-200a in mammalian epithelial cell transformation. Carcinogenesis. 2015;36(1):2-12.
250. Yuen HF, Gunasekharan VK, Chan KK, Zhang SD, Platt-Higgins A, Gately K, et al. RanGTPase: a candidate for Myc-mediated cancer progression. Journal of the National Cancer Institute. 2013;105(7):475-88.
251. Schlessinger J. Cell Signaling by Receptor Tyrosine Kinases. Cell. 2000;103(2):211-25.
252. Blume-Jensen P, Hunter T. Oncogenic kinase signalling. Nature. 2001;411(6835):355-65.

253. Whitmarsh AJ, Davis RJ. Multisite phosphorylation by MAPK. Science (New York, NY). 2016;354(6309):179.
254. Salazar C, Hofer T. Multisite protein phosphorylation--from molecular mechanisms to kinetic models. The FEBS journal. 2009;276(12):3177-98.
255. Klein AM, Zaganjor E, Cobb MH. Chromatin-tethered MAPKs. Curr Opin Cell Biol. 2013;25(2):272-7.
256. Plotnikov A, Zehorai E, Procaccia S, Seger R. The MAPK cascades: signaling components, nuclear roles and mechanisms of nuclear translocation. Biochimica et biophysica acta. 2011;1813(9):1619-33.
257. Sharma K, D'Souza RC, Tyanova S, Schaab C, Wisniewski JR, Cox J, et al. Ultradeep human phosphoproteome reveals a distinct regulatory nature of Tyr and Ser/Thr-based signaling. Cell reports. 2014;8(5):1583-94.
258. Hunter T. Tyrosine phosphorylation: thirty years and counting. Curr Opin Cell Biol. 2009;21(2):140-6.
259. Sadowski I, Stone JC, Pawson T. A noncatalytic domain conserved among cytoplasmic protein-tyrosine kinases modifies the kinase function and transforming activity of Fujinami sarcoma virus P130gag-fps. Molecular and cellular biology. 1986;6(12):4396-408.
260. Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. Curr Protoc Mol Biol. 2015;109:21.9.1-.9.9.
261. Abe Y, Tada A, Isoyama J, Nagayama S, Yao R, Adachi J, et al. Improved phosphoproteomic analysis for phosphosignaling and active-kinome profiling in Matrigel-embedded spheroids and patient-derived organoids. Scientific reports. 2018;8(1):11401-.
262. Magli A, Baik J, Mills LJ, Kwak I-Y, Dillon BS, Mondragon Gonzalez R, et al. Time-dependent Pax3-mediated chromatin remodeling and cooperation with Six4 and Tead2 specify the skeletal myogenic lineage in developing mesoderm. PLoS biology. 2019;17(2):e3000153.
263. Tripodi IJ, Allen MA, Dowell RD. Detecting Differential Transcription Factor Activity from ATAC-Seq Data. Molecules (Basel, Switzerland). 2018;23(5).
264. Rogerson C, Britton E, Withey S, Hanley N, Ang YS, Sharrocks AD. Identification of a primitive intestinal transcription factor network shared between esophageal adenocarcinoma and its precancerous precursor state. Genome Res. 2019;29(5):723-36.
265. Polyak K. Heterogeneity in breast cancer. The Journal of clinical investigation. 2011;121(10):3786-8.
266. Koren S, Bentires-Alj M. Breast Tumor Heterogeneity: Source of Fitness, Hurdle for Therapy. Molecular cell. 2015;60(4):537-46.
267. Zardavas D, Irrthum A, Swanton C, Piccart M. Clinical management of breast cancer heterogeneity. Nature reviews Clinical oncology. 2015;12(7):381-94.
268. Crabtree JS, Miele L. Breast Cancer Stem Cells. Biomedicines. 2018;6(3):77.
269. Owens T, Naylor M. Breast cancer stem cells. 2013;4(225).
270. Phi LTH, Sari IN, Yang Y-G, Lee S-H, Jun N, Kim KS, et al. Cancer Stem Cells (CSCs) in Drug Resistance and their Therapeutic Implications in Cancer Treatment. Stem Cells Int. 2018;2018:5416923-.
271. Palomeras S, Ruiz-Martinez S, Puig T. Targeting Breast Cancer Stem Cells to Overcome Treatment Resistance. Molecules (Basel, Switzerland). 2018;23(9).
272. Chu P-Y, Hou M-F, Lai J-C, Chen L-F, Lin C-S. Cell Reprogramming in Tumorigenesis and Its Therapeutic Implications for Breast Cancer. Int J Mol Sci. 2019;20(8):1827.
273. Noh KH, Kim BW, Song KH, Cho H, Lee YH, Kim JH, et al. Nanog signaling in cancer promotes stem-like phenotype and immune evasion. The Journal of clinical investigation. 2012;122(11):4077-93.
274. Singla M, Kumar A, Bal A, Sarkar S, Bhattacharyya S. Epithelial to mesenchymal transition induces stem cell like phenotype in renal cell carcinoma cells. Cancer cell international. 2018;18(1):57.

275. Solomon H, Dinowitz N, Pateras IS, Cooks T, Shetzer Y, Molchadsky A, et al. Mutant p53 gain of function underlies high expression levels of colorectal cancer stem cells markers. Oncogene. 2018;37(12):1669-84.
276. Monterisi S, Lo Riso P, Russo K, Bertalot G, Vecchi M, Testa G, et al. HOXB7 overexpression in lung cancer is a hallmark of acquired stem-like phenotype. Oncogene. 2018;37(26):3575-88.
277. Wang X, Yang S, Zhao X, Guo H, Ling X, Wang L, et al. OCT3 and SOX2 promote the transformation of Barrett's esophagus to adenocarcinoma by regulating the formation of tumor stem cells. Oncology reports. 2014;31(4):1745-53.
278. Yu W, Ma Y, Ochoa AC, Shankar S, Srivastava RK. Cellular transformation of human mammary epithelial cells by SATB2. Stem cell research. 2017;19:139-47.
279. Shackleton M, Vaillant F, Simpson KJ, Stingl J, Smyth GK, Asselin-Labat M-L, et al. Generation of a functional mammary gland from a single stem cell. Nature. 2006;439(7072):84-8.
280. Ayob AZ, Ramasamy TS. Cancer stem cells as key drivers of tumour progression. Journal of Biomedical Science. 2018;25(1):20.
281. Gendler SJ. MUC1, The Renaissance Molecule. Journal of Mammary Gland Biology and Neoplasia. 2001;6(3):339-53.
282. Jing X, Liang H, Hao C, Yang X, Cui X. Overexpression of MUC1 predicts poor prognosis in patients with breast cancer. Oncology reports. 2019;41(2):801-10.
283. Eirew P, Stingl J, Raouf A, Turashvili G, Aparicio S, Emerman JT, et al. A method for quantifying normal human mammary epithelial stem cells with in vivo regenerative ability. Nature medicine. 2008;14(12):1384-9.
284. Ye F, Qiu Y, Li L, Yang L, Cheng F, Zhang H, et al. The Presence of EpCAM(-)/CD49f(+) Cells in Breast Cancer Is Associated with a Poor Clinical Outcome. Journal of breast cancer. 2015;18(3):242-8.
285. Sheridan C, Kishimoto H, Fuchs RK, Mehrotra S, Bhat-Nakshatri P, Turner CH, et al. CD44+/CD24- breast cancer cells exhibit enhanced invasive properties: an early step necessary for metastasis. Breast cancer research : BCR. 2006;8(5):R59.
286. Honeth G, Bendahl P-O, Ringnér M, Saal LH, Gruvberger-Saal SK, Lövgren K, et al. The CD44+/CD24-phenotype is enriched in basal-like breast tumors. Breast Cancer Research. 2008;10(3):R53.
287. Mylona E, Giannopoulou I, Fasomytakis E, Nomikos A, Magkou C, Bakarakos P, et al. The clinicopathologic and prognostic significance of CD44+/CD24(-/low) and CD44-/CD24+ tumor cells in invasive breast carcinomas. Human pathology. 2008;39(7):1096-102.
288. Ferbeyre G, de Stanchina E, Lin AW, Querido E, McCurrach ME, Hannon GJ, et al. Oncogenic ras and p53 Cooperate To Induce Cellular Senescence. 2002;22(10):3497-508.
289. Serrano M, Lin AW, McCurrach ME, Beach D, Lowe SW. Oncogenic ras Provokes Premature Cell Senescence Associated with Accumulation of p53 and p16INK4a. Cell. 1997;88(5):593-602.
290. Zacarias-Fluck MF, Morancho B, Vicario R, Luque Garcia A, Escorihuela M, Villanueva J, et al. Effect of Cellular Senescence on the Growth of HER2-Positive Breast Cancers. JNCI: Journal of the National Cancer Institute. 2015;107(5).
291. Trost TM, Lausch EU, Fees SA, Schmitt S, Enklaar T, Reutzel D, et al. Premature senescence is a primary fail-safe mechanism of ERBB2-driven tumorigenesis in breast carcinoma cells. Cancer research. 2005;65(3):840-9.
292. Lin AW, Barradas M, Stone JC, van Aelst L, Serrano M, Lowe SW. Premature senescence involving p53 and p16 is activated in response to constitutive MEK/MAPK mitogenic signaling. Genes & development. 1998;12(19):3008-19.

293. Khan IA, Yoo BH, McPhee M, Masson O, Surette A, Dakin-Hache K, et al. ErbB2-driven downregulation of the transcription factor Irf6 in breast epithelial cells is required for their 3D growth. Breast Cancer Research. 2018;20(1):151.
294. Soule HD, Maloney TM, Wolman SR, Peterson WD, Jr., Brenz R, McGrath CM, et al. Isolation and characterization of a spontaneously immortalized human breast epithelial cell line, MCF-10. Cancer research. 1990;50(18):6075-86.
295. Debnath J, Mills KR, Collins NL, Reginato MJ, Muthuswamy SK, Brugge JS. The role of apoptosis in creating and maintaining luminal space within normal and oncogene-expressing mammary acini. Cell. 2002;111(1):29-40.
296. Trott DA, Cuthbert AP, Overell RW, Russo I, Newbold RF. Mechanisms involved in the immortalization of mammalian cells by ionizing radiation and chemical carcinogens. Carcinogenesis. 1995;16(2):193-204.
297. Reddel RR. The role of senescence and immortalization in carcinogenesis. Carcinogenesis. 2000;21(3):477-84.
298. Hahn WC. Immortalization and transformation of human cells2002. 351-61 p.
299. Elenbaas B, Spirio L, Koerner F, Fleming MD, Zimonjic DB, Donaher JL, et al. Human breast cancer cells generated by oncogenic transformation of primary mammary epithelial cells. Genes & development. 2001;15(1):50-65.
300. Dimri G, Band H, Band V. Mammary epithelial cell transformation: insights from cell culture and mouse models. Breast Cancer Research. 2005;7(4):171.
301. Hahn WC, Counter CM, Lundberg AS, Beijersbergen RL, Brooks MW, Weinberg RA. Creation of human tumour cells with defined genetic elements. Nature. 1999;400(6743):464-8.
302. Ahler E, Sullivan WJ, Cass A, Braas D, York AG, Bensinger SJ, et al. Doxycycline Alters Metabolism and Proliferation of Human Cell Lines. PloS one. 2013;8(5):e64561.
303. Nash CE, Mavria G, Baxter EW, Holliday DL, Tomlinson DC, Treanor D, et al. Development and characterisation of a 3D multi-cellular in vitro model of normal human breast: a tool for cancer initiation studies. Oncotarget. 2015;6(15):13731-41.
304. Imbalzano KM, Tatarkova I, Imbalzano AN, Nickerson JA. Increasingly transformed MCF-10A cells have a progressively tumor-like phenotype in three-dimensional basement membrane culture. Cancer cell international. 2009;9(1):7.
305. Debnath J, Muthuswamy SK, Brugge JS. Morphogenesis and oncogenesis of MCF-10A mammary epithelial acini grown in three-dimensional basement membrane cultures. Methods (San Diego, Calif). 2003;30(3):256-68.
306. Underwood JM, Imbalzano KM, Weaver VM, Fischer AH, Imbalzano AN, Nickerson JA. The ultrastructure of MCF-10A acini. Journal of cellular physiology. 2006;208(1):141-8.
307. Pradeep CR, Zeisel A, Köstler WJ, Lauriola M, Jacob-Hirsch J, Haibe-Kains B, et al. Modeling invasive breast cancer: growth factors propel progression of HER2-positive premalignant lesions. Oncogene. 2011;31:3569.
308. Kim IY, Yong HY, Kang KW, Moon A. Overexpression of ErbB2 induces invasion of MCF10A human breast epithelial cells via MMP-9. Cancer letters. 2009;275(2):227-33.
309. Narko K, Ristimaki A, MacPhee M, Smith E, Haudenschild CC, Hla T. Tumorigenic transformation of immortalized ECV endothelial cells by cyclooxygenase-1 overexpression. The Journal of biological chemistry. 1997;272(34):21455-60.
310. Boehm JS, Hession MT, Bulmer SE, Hahn WC. Transformation of human and murine fibroblasts without viral oncoproteins. Molecular and cellular biology. 2005;25(15):6464-74.
311. Hoover AC, Spanos WC, Harris GF, Anderson ME, Klingelhutz AJ, Lee JH. The role of human papillomavirus 16 E6 in anchorage-independent and invasive growth of mouse tonsil epithelium. Arch Otolaryngol Head Neck Surg. 2007;133(5):495-502.
312. Lawrence RT, Searle BC, Llovet A, Villen J. Plug-and-play analysis of the human phosphoproteome by targeted high-resolution mass spectrometry. Nature methods. 2016;13(5):431-4.

313. Matheron L, van den Toorn H, Heck AJ, Mohammed S. Characterization of biases in phosphopeptide enrichment by Ti(4+)-immobilized metal affinity chromatography and TiO2 using a massive synthetic library and human cell digests. Analytical chemistry. 2014;86(16):8312-20.
314. Ruprecht B, Koch H, Medard G, Mundt M, Kuster B, Lemeer S. Comprehensive and reproducible phosphopeptide enrichment using iron immobilized metal ion affinity chromatography (Fe-IMAC) columns. Molecular & cellular proteomics : MCP. 2015;14(1):205-15.
315. Vyse S, McCarthy F, Broncel M, Paul A, Wong JP, Bhamra A, et al. Quantitative phosphoproteomic analysis of acquired cancer drug resistance to pazopanib and dasatinib. J Proteomics. 2018;170:130-40.
316. Elsberger B, Fullerton R, Zino S, Jordan F, Mitchell TJ, Brunton VG, et al. Breast cancer patients' clinical outcome measures are associated with Src kinase family member expression. British journal of cancer. 2010;103(6):899-909.
317. Lee YY, Kim HP, Kang MJ, Cho BK, Han SW, Kim TY, et al. Phosphoproteomic analysis identifies activated MET-axis PI3K/AKT and MAPK/ERK in lapatinib-resistant cancer cell line. Experimental & molecular medicine. 2013;45:e64.
318. Nagata K, Kawakami T, Kurata Y, Kimura Y, Suzuki Y, Nagata T, et al. Augmentation of multiple protein kinase activities associated with secondary imatinib resistance in gastrointestinal stromal tumors as revealed by quantitative phosphoproteome analysis. J Proteomics. 2015;115:132-42.
319. Wolf-Yadlin A, Kumar N, Zhang Y, Hautaniemi S, Zaman M, Kim H-D, et al. Effects of HER2 overexpression on cell signaling networks governing proliferation and migration. 2006;2(1):54.
320. Zeevi D, Korem T, Zmora N, Israeli D, Rothschild D, Weinberger A, et al. Personalized Nutrition by Prediction of Glycemic Responses. Cell. 2015;163(5):1079-94.
321. Dunn KL, Davie JR. Stimulation of the Ras-MAPK pathway leads to independent phosphorylation of histone H3 on serine 10 and 28. Oncogene. 2005;24(21):3492-502.
322. Edmunds JW, Mahadevan LC. MAP kinases as structural adaptors and enzymatic activators in transcription complexes. Journal of Cell Science. 2004;117(17):3715.
323. Wang R, Yang L, Li S, Ye D, Yang L, Liu Q, et al. Quercetin Inhibits Breast Cancer Stem Cells via Downregulation of Aldehyde Dehydrogenase 1A1 (ALDH1A1), Chemokine Receptor Type 4 (CXCR4), Mucin 1 (MUC1), and Epithelial Cell Adhesion Molecule (EpCAM). Med Sci Monit. 2018;24:412-20.
324. Agnoletto C, Corrà F, Minotti L, Baldassari F, Crudele F, Cook WJJ, et al. Heterogeneity in Circulating Tumor Cells: The Relevance of the Stem-Cell Subset. Cancers (Basel). 2019;11(4):483.
325. Douville J, Beaulieu R, Balicki D. ALDH1 as a functional marker of cancer stem and progenitor cells. Stem cells and development. 2009;18(1):17-25.
326. Nakahata K, Uehara S, Nishikawa S, Kawatsu M, Zenitani M, Oue T, et al. Aldehyde Dehydrogenase 1 (ALDH1) Is a Potential Marker for Cancer Stem Cells in Embryonal Rhabdomyosarcoma. PloS one. 2015;10(4):e0125454-e.
327. Li W, Ma H, Zhang J, Zhu L, Wang C, Yang Y. Unraveling the roles of CD44/CD24 and ALDH1 as cancer stem cell markers in tumorigenesis and metastasis. Scientific Reports. 2017;7(1):13856.
328. Ponti D, Costa A, Zaffaroni N, Pratesi G, Petrangolini G, Coradini D, et al. Isolation and in vitro propagation of tumorigenic breast cancer cells with stem/progenitor cell properties. Cancer research. 2005;65(13):5506-11.
329. Horimoto Y, Arakawa A, Sasahara N, Tanabe M, Sai S, Himuro T, et al. Combination of Cancer Stem Cell Markers CD44 and CD24 Is Superior to ALDH1 as a Prognostic Indicator in Breast Cancer Patients with Distant Metastases. PloS one. 2016;11(10):e0165253.

330. Sun Q, Lesperance J, Wettersten H, Luterstein E, DeRose YS, Welm A, et al. Proapoptotic PUMA targets stem-like breast cancer cells to suppress metastasis. The Journal of clinical investigation. 2018;128(1):531-44.
331. Lee E, Piranlioglu R, Wicha MS, Korkaya H. Plasticity and Potency of Mammary Stem Cell Subsets During Mammary Gland Development. 2019;20(9):2357.
332. Jing X, Liang H, Hao C, Yang X, Cui X. Overexpression of MUC1 predicts poor prognosis in patients with breast cancer. Oncology reports. 2019;41(2):801-10.
333. Kharbanda A, Rajabi H, Jin C, Raina D, Kufe D. Oncogenic MUC1-C promotes tamoxifen resistance in human breast cancer. Molecular cancer research : MCR. 2013;11(7):714-23.
334. Kravchenko YE, Alibaeva RA, Frolova EI, Chumakov SP. MUC1 expression is upregulated in CSC-enriched populations of luminal breast cancer cell lines2017. 2539-43 p.
335. Hyun KA, Koo GB, Han H, Sohn J, Choi W, Kim SI, et al. Epithelial-to-mesenchymal transition leads to loss of EpCAM and different physical properties in circulating tumor cells from metastatic breast cancer. Oncotarget. 2016;7(17):24677-87.
336. Pauklin S, Vallier L. The cell-cycle state of stem cells determines cell fate propensity. Cell. 2013;155(1):135-47.
337. Gonzalez ME, Moore HM, Li X, Toy KA, Huang W, Sabel MS, et al. EZH2 expands breast stem cells through activation of NOTCH1 signaling. Proceedings of the National Academy of Sciences of the United States of America. 2014;111(8):3098-103.
338. Halle MK, Tangen IL, Berg HF, Hoivik EA, Mauland KK, Kusonmano K, et al. HER2 expression patterns in paired primary and metastatic endometrial cancer lesions. British journal of cancer. 2017;118:378.
339. Tong Z-J, Shi N-Y, Zhang Z-J, Yuan X-D, Hong X-M. Expression and prognostic value of HER-2/neu in primary breast cancer with sentinel lymph node metastasis. Bioscience reports. 2017;37(4):BSR20170121.
340. A DACP, Lopes C. Implications of Different Cancer Stem Cell Phenotypes in Breast Cancer. Anticancer research. 2017;37(5):2173-83.
341. Scaltriti M, Nuciforo P, Bradbury I, Sperinde J, Agbor-Tarh D, Campbell C, et al. High HER2 expression correlates with response to the combination of lapatinib and trastuzumab. Clinical cancer research : an official journal of the American Association for Cancer Research. 2015;21(3):569-76.
342. Lee HJ, Seo AN, Kim EJ, Jang MH, Suh KJ, Ryu HS, et al. HER2 Heterogeneity Affects Trastuzumab Responses and Survival in Patients With HER2-Positive Metastatic Breast Cancer. American journal of clinical pathology. 2014;142(6):755-66.
343. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science (New York, NY). 1987;235(4785):177-82.
344. Tandon AK, Clark GM, Chamness GC, Ullrich A, McGuire WL. HER-2/neu oncogene protein and prognosis in breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 1989;7(8):1120-8.
345. Slamon DJ, Godolphin W, Jones LA, Holt JA, Wong SG, Keith DE, et al. Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. Science (New York, NY). 1989;244(4905):707-12.
346. Paik S, Hazan R, Fisher ER, Sass RE, Fisher B, Redmond C, et al. Pathologic findings from the National Surgical Adjuvant Breast and Bowel Project: prognostic significance of erbB-2 protein overexpression in primary breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 1990;8(1):103-12.
347. Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, et al. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. The New England journal of medicine. 2001;344(11):783-92.

348. Thor AD, Liu S, Edgerton S, Moore D, 2nd, Kasowitz KM, Benz CC, et al. Activation (tyrosine phosphorylation) of ErbB-2 (HER-2/neu): a study of incidence and correlation with outcome in breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2000;18(18):3230-9.
349. Wright C, Angus B, Nicholson S, Sainsbury JR, Cairns J, Gullick WJ, et al. Expression of c-erbB-2 oncoprotein: a prognostic indicator in human breast cancer. Cancer research. 1989;49(8):2087-90.
350. Paik S, Kim C, Jeong J, Geyer CE, Romond EH, Mejia-Mejia O, et al. Benefit from adjuvant trastuzumab may not be confined to patients with IHC 3+ and/or FISH-positive tumors: Central testing results from NSABP B-31. 2007;25(18_suppl):511-.
351. Pernas S, Tolaney SM. HER2-positive breast cancer: new therapeutic frontiers and overcoming resistance. Ther Adv Med Oncol. 2019;11:1758835919833519-.