Functional investigation of *ESR1* fusions identified in endocrine therapy refractory estrogen receptor positive breast cancer

by

Zeynep Erdoğan-Yıldırım

MD, Medical University Vienna, Austria, 2008

Submitted to the Graduate Faculty of

Department of Human Genetics

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2019

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This thesis was presented

by

Zeynep Erdoğan-Yıldırım

It was defended on

February 13, 2019

and approved by

Committee members:

Ryan Minster, PhD
Assistant Professor
Department of Human Genetics
Graduate School of Public Health
University of Pittsburgh

Zsolt Urban, PhD
Associate Professor
Department of Human Genetics
Graduate School of Public Health
University of Pittsburgh

Thesis Advisor:

Adrian Lee, PhD
Professor
Department of Pharmacology and Chemical Biology
School of Medicine
Department of Human Genetics
Graduate School of Public Health
University of Pittsburgh

Adrian Lee, PhD

Functional investigation of *ESR1* fusions identified in endocrine therapy refractory estrogen receptor positive breast cancer

Zeynep Erdoğan-Yıldırım, MS

University of Pittsburgh, 2019

Abstract

Metastatic estrogen receptor (ER)-positive breast cancer is an incurable disease that remains a clinical challenge and a public health burden. Over 40,000 women die each year from breast cancer and over 90% of these are due to metastatic disease. Although there has been great success with anti-endocrine treatment, most patients with metastatic disease develop resistance during the course of therapy. Loss of ER or missense mutations in the ligand-binding domain (LBD) of ER are reported as causal mechanisms of resistance. Recently, our laboratory showed that *ESR1* fusions involving loss of LBD, play a role in lack of response to therapy. Only limited knowledge exists on the actual frequency and functional role of *ESR1* fusions. Hence, I aimed to expand the search for further *ESR1* fusions in advanced breast cancer disease and to characterize the functional role of *ESR1* fusions (*ESR1-DAB2, ESR1-GYG1* and *ESR1-SOX9*) our lab published earlier along with the *ESR1-LPP* fusion found in a PDX-model.

Screening of RNA-seq data of primary-metastatic paired breast tumors (n=45) in University of Pittsburgh cohort and metastatic BrCa (n=91) in the MET500 cohort revealed a total of five *ESR1* fusions sharing identical breakpoint with *ESR1-YAP1* fusion. *ESR1-GYG1* was the only *ESR1* fusion identified in the Pitt-cohort. For functional assessment, the *ESR1* fusions were transiently and stably transfected into cell lines. The immunofluorescence staining confirmed the predominant nuclear localization of the *ESR1* fusions, while *ESR1-LPP* and *ESR1-GYG1* fusions additionally displayed cytoplasmic localization. ER activity assays via luciferase assay and qRT-

PCR demonstrated estrogen-independent constitutive activity of *ESR1* fusions that is unresponsive to anti-endocrine treatment. While *ESR1-DAB2* and *ESR1-SOX9* fusions induced the transcription of estrogen-responsive genes, only *ESR1-SOX9* demonstrated statistically significant estrogen-independent proliferation in stable expressing T47D cells. Overall, active *ESR1* fusions may have a critical role in developing anti-endocrine resistance and promoting tumor progression. Since *ESR1* fusions with loss of LBD are recurrent in therapy-refractory ER-positive breast cancer, further comprehensive studies are needed (1) to determine their true frequency, (2) to understand their mechanism of action and (3) to determine their value as prognostic and therapeutic biomarkers.

Table of Contents

List of Tables

List of Figures

Acknowledgement

I want to thank everyone for their valuable involvement throughout my project and finalization as a thesis. First of all, I want to thank Dr. **Adrian Lee** for being my Thesis Advisor and for introducing me to the intriguing field of cancer genetics and sharing his immense expertise. I am grateful for your unwavering support and encouragement. Many thanks to Dr. **Steffi Oesterreich** for providing her insights and expertise to the project that have been invaluable. I want to particularly thank Dr. **Jennifer Xavier**, Research Advisor in Lee/Oesterreich Lab, for helping me plan out the experiments and troubleshoot technical problems. I greatly appreciate that you always made the time and had an open door. I would like to express my deepest appreciation to Dr. **Ryan Minster** and Dr. **Zsolt Urban** for serving on my committee. I owe Dr. **Yesim Demirci**, who is an inspirational person and professor, deepest gratitude. She heartened me to go after my dream to explore the role of human genetics in diseases and supported my application to the program. I want to express my sincere gratitude to my school advisor Dr. **Candace Kammerer** for the excellent mentorship she provided. Thank you very much for your continued support and always finding the time to reach out to help and share your wisdom.

I want to thank Dr. **Ryan Hartmaier** who built the foundation of my research project for his hard work and allowing me to contribute as a co-author to the published research article "Recurrent hyperactive *ESR1* fusion proteins in endocrine therapy-resistant breast cancer." I owe deepest gratitude to **Nolan, Nick** and **Hannah**. Many thanks to Nick Smith for introducing me to the computational work during my rotation and for his continued support and patience throughout my time in the lab. Endless gratitude to Nolan Priedigkeit for introducing me to the wet lab, teaching me how to selectively approach the numerous algorithms available for detecting fusions

and to keep records of the overwhelming data in an organized manner. I am particularly grateful for Hannah Guzolik's assistance in performing the ERE-assays. Special thanks to Dr. **Nilgun Tasdemir** and Dr. **Susan Farabaugh** for their instrumental input and practical suggestions. I want to also thank my lab mates **Ahmed, Ashu, Beth, John, Kai, Kevin, Lan, Nick, Nolan, Rekha, Tian and Vaciry (Zheqi)** for their companionship. It has been a joyful experience to work with such incredible colleagues.

Moreover, I want to humbly thank the many patients who have contributed samples used in this research. None of this work would have been possible without you.

Finally, I want to thank my friends and family. I am extremely grateful to my remarkable friends, especially **Rehab Sherlala** for her unconditional friendship, support and motivation. I want to extend my deepest gratitude to my family, especially my parents, **Fatma** and **Yusuf**, for their selfless love, and for sacrifices they made to provide me my education and a future, I dreamed of. Most importantly, my heartfelt appreciation and gratitude goes to my husband **Hilmi** for being my partner in life, for his everlasting love, continuous support, encouragement and humor. Without you and your peaceful spirit making me see the beauty in everything, I couldn't have accomplished this journey. I am very thankful to my beloved and adorable daughters **Hatice Vera** and **Betül Neva** for their pure love, utmost patience and sacrificing irreplaceable time. Your endless curiosity to learn fuels my thirst for knowledge.

Scientific Publications

Results belonging to the early phase of this project have been published elsewhere.

Peer-reviewed research article:

Hartmaier, R.J., Trabucco, S.E., Priedigkeit, N., Chung, J.H., Parachoniak, C.A., Vanden Borre, P., Morley, S., Rosenzweig, M., Gay, L.M., Goldberg, M.E., Suh, J., Ali, S.M., Ross, J., Leyland-Jones, B., Young, B., Williams, C., Park, B., Tsai, M., Haley, B., Peguero, J., Callahan, R.D., Sachelarie, I., Cho, J., Atkinson, J.M., Bahreini, A., Nagle, A.M., Puhalla, S.L., Watters, R.J., **Erdogan-Yildirim, Z.**, Cao, L., Oesterreich, S., Mathew, A., Lucas, P.C., Davidson, N.E., Brufsky, A.M., Frampton, G.M., Stephens, P.J., Chmielecki, J., Lee, A.V., 2018. "*Recurrent hyperactive ESR1 fusion proteins in endocrine therapy-resistant breast cancer.*" Ann. Oncol. 29, 872–880. https://doi.org/10.1093/annonc/mdy025

Poster:

Erdogan-Yildirim Z, Guzolik H, Hartmaier R, Priedigkeit N, Oesterreich S, Xavier J, Lee AV (2018). *"ESR1-fusions - A recurrent mechanism of endocrine treatment resistance in metastatic ER+ breast cancer"*, **Poster**, 20th Annual Dean's Day of Graduate School of Public Health at University of Pittsburgh, Pittsburgh, USA. 5-9 April 2018

1.0 Background and Significance

1.1 Public health problem - Breast cancer

Breast cancer is the most frequent cancer in females and is the second leading cause of cancer death (Siegel, Miller, & Jemal, 2018). Worldwide, breast cancer is diagnosed in around 1.7 million women annually and was responsible for the death of 522,000 women in 2012 (Ginsburg et al., 2017). In US, the American Cancer Society estimates 266,120 new cases and 40,920 related deaths for 2018 (Siegel et al., 2018). Advances in understanding the disease pathology, early detection and availability of targeted therapies have significantly improved the overall survival in early-stage breast cancers and led to an increase in the 5-year survival rate to 99% for localized and 85% for regional disease (Siegel et al., 2018). Yet, survival rate for recurrent/distant metastatic disease has remained steady at approximately 25% for more than a decade (Siegel et al., 2018). Since vast majority of patients with metastatic disease acquire resistance to the state-of-art therapies, metastatic breast cancer is considered incurable. About 10% of patients are predicted to present with metastatic/advanced breast cancer at the initial diagnosis and 20-30% of patients diagnosed with early-stage cancer are estimated to relapse with metastasis (Early Breast Cancer Trialists' Collaborative, 2005; Siegel et al., 2018). It is important to note that more than 3.4 million females in US live with history of breast cancer according to the report of the Surveillance, Epidemiology, and End Results (SEER) program of National Cancer Institute (Noone et al., 2017) and are at potential risk of developing metastasis. Since metastatic disease is fatal and is

responsible for 90% of breast cancer related deaths, it represents an important public health burden in women not only in the developed countries but also worldwide (Ginsburg et al., 2017). For successful prevention and/or better management of metastatic disease, there is a need to better understand what mechanisms drive a tumor to metastasize, if/how metastatic disease is distinct from the primary tumor and which particular molecular features in the metastasis cause unresponsiveness to currently available therapies.

Although breast cancer is a disease confined to the breast, it is a complex and extremely diverse disease. In its most recent classification, the World Health Organization (WHO) divides invasive breast cancers based on their histological characteristics into two major groups "invasive carcinoma of no special type (NST)", also commonly known as "invasive ductal carcinoma" (IDC), and "special subtypes", which includes more than 20 different entities that differ in their morphology and growth (Lakhani SR, 2012). Invasive carcinoma of NST accounts for the majority (75%), with invasive lobular cancer (ILC) being the most frequent special subtype representing 5-15% of all invasive breast cancers (Lakhani SR, 2012). In general, the histological grading based on the Nottingham Histologic Score combined with the clinical tumor staging by the TNM system (tumor size, number of lymph nodes and presence of distant metastasis) continues to be an important key component for clinical decision-making and prognostic evaluation (Edge & Compton, 2010; Elston & Ellis, 1991; Lakhani SR, 2012). Although this classification efficiently predicts the prognosis for a large group of patients, the substantial heterogeneity even within the same histologic subtype limits its prognostic value/predictions for an individual patient. Additional assessment of the expression of the human epidermal growth factor receptor 2 (*HER2*) and the hormone receptors (HR) - estrogen receptor (ER) and progesterone receptor (PR) - by

immunohistochemistry (IHC) represents a further important prognostic determinant and is pivotal for the selection of adequate adjuvant systemic therapy.

Taking these measures into account, the current management of invasive breast cancer broadly consists of surgical removal of the tumor (lumpectomy or mastectomy) and subsequent sequential treatment with one or more of the available adjuvant therapy modalities that include radiotherapy, chemotherapy, endocrine therapy and targeted therapy, which depend mostly upon the size of the tumor and the expression of key biomarkers. Systemic therapy can also be applied preoperatively as neoadjuvant therapy. In breast tumors that lack the expression of all three receptors (HRs and HER2) - the so-called **triple negative breast cancers (TNBC)** - the mainstay therapy consists of sequentially applied chemotherapeutics containing taxanes and anthracyclines (Andreopoulou, Kelly, & McDaid, 2017; Bines, Earl, Buzaid, & Saad, 2014; Stover, Bell, & Tolaney, 2016). TNBCs, which represent 12.2% of invasive breast tumors, have the poorest prognosis (Andreopoulou et al., 2017; Howlader et al., 2014). About 15-20% of the invasive breast tumors overexpress **HER2** or have HER2 amplification (Howlader et al., 2014; Wilson et al., 2017). HER2 amplification in breast cancer and its association with poor outcome was shown for the first time in 1987 (Lukong, 2017; Slamon et al., 1987). HER2 is a ligand orphan receptor tyrosine kinase that belongs to the epidermal growth factor receptor (EGFR) family (Lukong, 2017; Roskoski, 2014). Its activation upon homo- or heterodimerization with other members of the EGFR/ErbB family triggers the downstream pathways PI3K/AKT and Raf/MAPK, which are very well studied pathways known to induce cell-cycle progression, proliferation and survival (Iqbal & Iqbal, 2014; Lukong, 2017; Reynolds, Sarangi, Bardia, & Dizon, 2014; Roskoski, 2014). The use of monoclonal antibodies targeting the extracellular domain of the HER2 called trastuzumab (Herceptin®) was initially approved in metastatic HER2 positive patients in a fast-

track process in 1998 by the FDA (Lukong, 2017). In fact, its discovery was a breakthrough because it was the first therapeutic antibody approved for the treatment of a solid tumor that by targeting an oncogene blocked its downstream signaling. Thereby, HER2 positive primary tumors, once a type that used to be associated with very poor prognosis, became better managed using the adjuvant treatment with trastuzumab, and its risk of recurrence decreased by more than 50% (Lukong, 2017). Lastly, the most prevalent subtype of invasive breast cancers is **ER-positive** and accounts for 75% of the breast cancer tumors. ER is a nuclear receptor that has been described and characterized in the 1960s (Jensen, Jacobson, Walf, & Frye, 2010; Lukong, 2017; Toft & Gorski, 1966). The long suspected relationship between estrogen and accelerated growth of breast cancers was demonstrated, which paved the way for development of drugs antagonizing the effects of estrogens (Jensen et al., 2010; Lukong, 2017; Toft & Gorski, 1966). Tamoxifen, which was initially designed as a contraceptive in 1962, was later shown to block the estrogen receptor from binding its ligand estrogen and to lead the tumor to remission similarly to the outcomes when adrenalectomy or ovarectomy were performed (Jensen et al., 2010; Jordan, 2003; Lukong, 2017). Tamoxifen received FDA approval for treatment of advanced ER-positive breast cancer in post-menopausal women in 1977 and from 1984 was regarded as the preferred adjuvant treatment for ER-positive breast cancer by the National Cancer Institute (Jordan, 2003; Lukong, 2017). Currently, additional anti-estrogenic treatments are available and include selective ER modulators (SERM) that competitively binds to ER, and selective ER degraders (SERD), which when bound to ER induces its rapid degradation via the proteasome. The first-in-class SERD is fulvestrant, which was initially approved by the FDA as a second-line treatment in postmenopausal patients with progressive ER-positive metastatic disease in 2002 (Boer, 2017; Lukong, 2017). Another way to reduce the effects of estrogen on the tumor is through aromatase inhibitors (AI), which inhibit

the enzyme aromatase that converts androgens into estrogen in peripheral tissues. The non-steroidal inhibitors (letrozole and anastrazole) bind to the aromatase enzyme reversibly, whereas steroidal inhibitors (exemestane) bind irreversibly and may cause androgenic effects. AIs serve as standard of care for breast cancer in postmenopausal women and are the first-line treatment for metastatic disease (Chumsri, Howes, Bao, Sabnis, & Brodie, 2011; Lukong, 2017).

Overall, patients with primary ER-positive breast cancer greatly benefit from endocrine therapy, which reduces the 10-year disease recurrence risk by around 50% and the mortality rate by around 30% (Clarke, Tyson, & Dixon, 2015; De Marchi, Foekens, Umar, & Martens, 2016; Early Breast Cancer Trialists' Collaborative et al., 2011; Pan et al., 2017). However, approximately 30% of patients initially diagnosed with early-stage breast cancer will experience a local recurrence and/or metastasis to the predilection sites of bone, brain, liver, lung and distant lymph nodes (Early Breast Cancer Trialists' Collaborative, 2005; Rabbani & Mazar, 2007). Moreover, around 40% of patients with recurrent disease do not respond to initial therapy and those that initially respond develop resistance during therapy (Angus, Beije, Jager, Martens, & Sleijfer, 2017). Since therapy failure adversely affects disease progression and outcome, metastatic ER-positive breast cancer remains a great challenge.

Efforts to stratify patients and tailor treatment based upon the expression of key proteins in breast cancer, such as estrogen receptor and tamoxifen therapy, was a breakthrough and served as a vanguard for the successful application of precision medicine, as it increased the quality of life and the survival of many patients. Yet, the classical characterization of breast cancer by receptor expression using IHC has its limitations not only due to lack of uniformity in testing but moreover due to the wide heterogeneity even within the same tumor. With advances in technology and broad application of next generation sequencing (NGS) methodologies, genome-wide

profiling of breast cancer by gene expression led to its classification into five major intrinsic molecular subtypes (Perou et al., 2000; Sorlie et al., 2001). These five subtypes are luminal A, luminal B, human epidermal growth factor receptor 2 (*HER2*) overexpression, basal-like and normal-like tumors (Perou et al., 2000; Sorlie et al., 2001). Remarkably, these subtypes clustered hierarchically first by the presence of ER and second by the presence of HER2 expression in ER-negative subtypes (Russnes, Lingjaerde, Borresen-Dale, & Caldas, 2017). Currently, the intrinsic subtyping along with the traditional pathological assessment allows the most comprehensive characterization and stratification of patients in terms of prognosis and therapeutic decision making. Furthermore, the application of NGS also sheds light on the complex molecular features of primary tumors including intra- and inter-tumor heterogeneity and allows us to monitor genomic evolution by identifying the acquisition of recurrence/metastasis-specific driver mutations and aberrations. This allows the unraveling of pathways involved in both cancer pathogenesis and progression and identification of known and unknown mutations that confer metastatic potential and resistance to treatment which may be clinically actionable. Since the main reason for breast cancer mortality is a result from complications of recurrence or metastasis, understanding these tumor dynamics is of critical importance. This will serve as a guide for the development of further targeted treatments that will hopefully overcome resistance and halt the fatal tumor advancement.

This thesis will mainly address endocrine treatment resistance in metastatic ER-positive breast cancer, which is regarded as an incurable disease that affects the major subgroup of patients with breast cancer who sorely need redefinition of treatment guidelines and additional treatment options in the metastatic setting.

1.2 Genetics of ER-positive breast cancer

NGS paved the way for large-scale genomic analysis that augmented our understanding of the genetics, biology and underlying mechanisms involved in pathogenesis of various diseases. The Cancer Genome Atlas (TCGA), The International Cancer Genome Consortium (ICGC) and the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) represent the three most important and comprehensive large-scale studies for breast cancer research in elucidating the molecular basis of breast cancer, focusing mostly on primary tumors. In these studies, *PIK3CA*, PTEN*, AKT1, TP53*, GATA3, CDH1, RB1, MLL3, MAP3K1* and *CDKN1B* were identified as commonly mutated genes in early breast cancer pathogenesis, where the genes marked with asterisk are found in at least 10% of tumors (Cancer Genome Atlas, 2012; Curtis et al., 2012; Nik-Zainal et al., 2016).

In luminal tumors, which are ER-positive, *MAP3K1, PIK3CA* and *GATA3* are the predominantly mutated genes in primary tumors (Cancer Genome Atlas, 2012). Another frequently observed alteration is Cyclin D1 amplifications (Cancer Genome Atlas, 2012; Courjal et al., 1996) and recurrent deletions of the putative tumor suppressor *MAP2K4* (Curtis et al., 2012). The pathways involving the tumor suppressors *RB1* and *TP53* are largely intact in ER-positive tumors and bestow better prognosis (Cancer Genome Atlas, 2012). Although *ESR1* and *XBP1* are highly expressed, they are found to be infrequently mutated in primary tumors. However, *ESR1* mutations arise in metastatic samples which relapse under endocrine treatment (Toy et al., 2013; Yates et al., 2017). Additionally, mutations involving actionable genes *RPTOR*, *ERBB3, RB1, TSC1/2, PALB2, NOTCH3, ALK* are also found to be enriched in metastatic lesions compared to early breast cancer (Lefebvre et al., 2016; Toy et al., 2013; Toy et al., 2017). Of note, the latter conclusion should be regarded with caution, since the metastatic samples are matched to a mix of early breast cancer

cases in the TCGA analysis rather than their own paired primary tissue. A primary-metastatic matched analysis of tumor evolution revealed the occurrence of *TP53* mutations (40-50%) and loss of ER and PgR expression (17%) in metastatic ER-positive tumors, where the loss of ER expression is concomitant mostly with the driver mutations in *TP53* and less frequently with mutations in *ARID1A (Yates et al., 2017)*.

Overall, the mutational profile of primary and metastatic tumors, especially synchronous cases, show high concordance in matched studies (Meric-Bernstam et al., 2014; Yates et al., 2017). Distant metastasis (via hematogenous spread) and relapses that spread late from the primary tumor acquire further somatic as well as driver mutations that are distinct from the primary as well as between different metastatic sites within the same patient indicating the continuous genetic evolution of tumor cells (Yates et al., 2017). The genetic evolution of a tumor results from an interplay of a) continuous acquisition of mutations due to increased proliferation and impaired DNA repair pathways, b) intra-tumor genetic heterogeneity and c) clonal selection due to fitness and due to resistance to therapies via intrinsic or acquired mutations.

1.3 The estrogen receptor and its role in breast cancer

Estrogens play an important role in the development and regulation of physiological functions in the reproductive system as well as in the regulation of various other tissues within the cardiovascular, skeletal and neuroendocrine system. They exert their effects primarily by interacting with two different estrogen receptors ERα and ERβ, which belong to the members of the nuclear receptor superfamily (Hall, Couse, & Korach, 2001). Although ERα and ERβ show a homology in the DNA-binding (97%) and in the ligand-binding domain (60%), they are two

distinct receptors encoded by separate genes - *ESR1* on chromosome 6 and *ESR2* on chromosome 14 - and have different expression patterns in different tissues (Hall et al., 2001). ERα (generally referred as the ER) is the main estrogen receptor found in breast tissues and is highly expressed in 75% of breast cancers.

The *ESR1* gene consists of ten exons, from which exon 3 through exon 10 contains the coding sequence for ER. The ER is a protein of 595 amino acids length that is formed of six functional regions A through F (**Figure 1**). Starting at the N-terminus, the domains are the ligand-independent Activation Function-1 (AF-1, region A/B), the DNA-binding domain (DBD, region C), the Hinge domain (region D), the Ligand-binding domain (LBD) with the ligand-dependent Activation Function-2 (AF-2, region E) and the multifunctional carboxyl terminal domain F (region F) (Kocanova, Mazaheri, Caze-Subra, & Bystricky, 2010; R. Kumar et al., 2011; V. Kumar et al., 1987; Yasar, Ayaz, User, Gupur, & Muyan, 2017). The ER is predominantly found in the nucleus regardless of the presence or absence of its ligand and exhibits nucleocytoplasmic shuttling in unbound conditions (Kocanova et al., 2010).

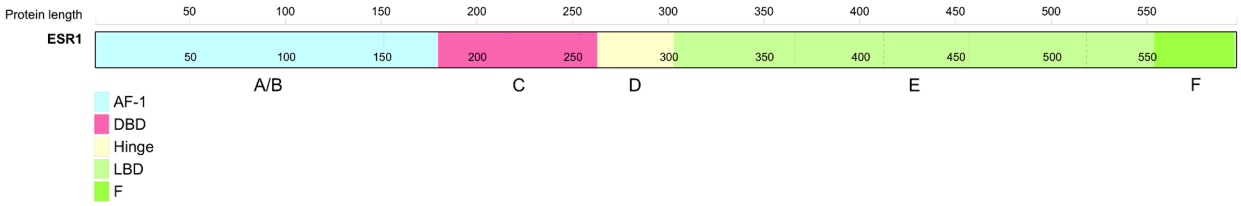

Figure 1 Functional structure and domains of the ERα

The domains from N-terminal (left) to C-terminal (right) end are: N-terminal domain (NTD) and Transactivation domain AF-1 (region A/B), DNA binding domain (DBD, region C), Hinge region (region D), Ligand-binding domain (LBD) with the Transactivation domain AF-2 (region E) and the C-terminal F-domain. The letters A-F below the graph labels the six regions. The dashed lines show the exon boundaries, starting with exon 3 (= the first exon in the coding sequence) up to the last exon 10. The protein length represents the codon numbers. Graph was designed with ProteinPaint (https://proteinpaint.stjude.org/)

Estrogenic signaling via ER is classically initiated by the binding of estradiol (E2) to the LBD that shelters the AF-2 and the dimerization interface. Consequent conformational changes in

the LBD enable the interaction of the AF-2 domain with co-activators at its interaction surface created by a positional shift of H12, as well as allow homo- and hetero-dimerization. Adjacent to the LBD is the hinge domain that serves as a linker between the LBD and the DBD and harbors the nuclear localization signal (NLS) that directs the receptor to the nucleus upon its activation (Kocanova et al., 2010; R. Kumar et al., 2011). The DBD binds to estrogen response elements (EREs) in DNA, which are palindromic sequences located proximally to the promoter regions or to the distant enhancer regions linked to the transcription initiation sites of ER-regulated genes. The EREs, for which consensus and non-consensus sequences exist, influence not only the affinity of the estrogen-ER complex to the DNA but also the recruitment of basal transcriptional machinery and co-regulatory proteins (including both activators and repressors) (Gruber, Gruber, Gruber, Wieser, & Huber, 2004; R. Kumar et al., 2011; Rosenfeld & Glass, 2001; Schwabe, Chapman, Finch, & Rhodes, 1993). The most well described activators are SRC-1 (NCOA1), TIF2 and AIB1, which belong to the p160 family of nuclear receptor co-activators (Rosenfeld & Glass, 2001; Yasar et al., 2017). The level of transcriptional activity of the ER is determined by the interaction of the ligand-independent AF-1 and the ligand-dependent AF-2 domain, which exert synergistic effect on the transcription and trigger full activity when their transactivation is harmonized. The two transactivation domains exert promoter and cell specific activities (R. Kumar et al., 2011).

In addition to the "classical" estrogen signaling known as **genotropic ERE-dependent ER signaling pathway**, there are three known "non-classical" pathways (McDevitt et al., 2008). Ligand/estradiol-stimulated ER can also mediate the transcription of target genes indirectly through association with other transcription factors (e.g. AP1, Sp1, NFkB, GATA1, ATF/CREB and STAT5) (Marino, Galluzzo, & Ascenzi, 2006; McDevitt et al., 2008; O'Lone, Frith, Karlsson, & Hansen, 2004). The so called **genotropic ERE-independent ("tethered") ER signaling**

activates approximately 35% of the estrogen-responsive genes (O'Lone et al., 2004). ER signaling can be also achieved via **non-genomic signaling by the membrane associated ER** (e.g. G protein-coupled estrogen receptor 1, GPER1). Upon E2-stimulation, GPER1 - genetically and structurally distinct from the ERα and ERβ – increases intracellular calcium and cyclic AMP (adenosine monophosphate) levels and triggers various protein-kinase cascades involved in EGFR, MAPK, PI3K and phospholipase C signaling pathways that control transcriptional activities (McDevitt et al., 2008; Pupo, Maggiolini, & Musti, 2016; Vrtacnik, Ostanek, Mencej-Bedrac, & Marc, 2014). The ER can further elicit **ligand-independent signaling** in the absence of estrogen and initiate transcription of its target genes, when phosphorylation of Ser118 in the AF-1 via growth factor signaling pathways MAPK, AKT and JNK induces ER activation (Lee, Cui, & Oesterreich, 2001; Marino et al., 2006; McDevitt et al., 2008).

Estrogen signaling, as previously mentioned, has important regulatory effects in multiple organ systems in females. Most importantly, it is critical for the development and differentiation of the mammary gland (JavanMoghadam, Weihua, Hunt, & Keyomarsi, 2016). When the control mechanisms of ER signaling are disturbed, the uninhibited transcriptional activity of ER can initiate tumor development through uncontrolled growth (Carroll, 2016). The relationship between estrogen and breast cancer development has been well characterized (Russo & Russo, 2006). In approximately 75% of breast cancers, the expression of ER is the key driver that can be effectively and safely targeted by anti-endocrine treatments such as SERDs, SERMs and AIs (described in previous section). However, according to the review from Angus et al. 40% of patients do not respond to initial treatment due to intrinsic resistance and the remainder acquire resistance during therapy (Angus et al., 2017). By definition, intrinsic resistance, also known as primary resistance, is outlined as a progression of the metastatic disease within 6 months under first-line treatment or

as a relapse within 2 years while on adjuvant endocrine therapy. Secondary resistance is defined as a relapse that either occurs after 2 years while on ongoing or within 12 months after the completion of the adjuvant endocrine therapy or is a progression that occurs at least 6 months after starting the endocrine therapy for metastatic disease (Cardoso et al., 2017).

ER-positive breast cancers can gain resistance to anti-endocrine treatment through multiple mechanisms. The leading mechanism in both *de novo* and intrinsic resistance is attributed to lack of ER expression (Musgrove & Sutherland, 2009). Further, structural and functional changes of ER, epigenetic remodeling along with shifts in expression of its co-regulators (overexpression and phosphorylation of co-activators as AIB1 or down-regulation of co-repressors as NCoR) confer unresponsiveness to therapy (Musgrove & Sutherland, 2009; Osborne & Schiff, 2011). Changes in the ER pathway, deregulation of the cell cycle by activating cyclin-dependent kinases (CDKs) and blocking the CDK inhibitors p21 or p27, loss of estrogen dependence and/or up-regulation of and crosstalk with growth factor signaling pathways are further mechanisms contributing to acquired resistance to endocrine treatment (Musgrove & Sutherland, 2009; Osborne & Schiff, 2011). Alterations in the ER-regulated genes such as loss of *PTEN* expression or activating mutations in *PIK3CA* represent additional mechanisms of opposing the effects of ER inhibition (Shoman et al., 2005).

Altogether, therapy failure in patients especially with recurrent and metastatic disease is a major problem, that adversely affects disease progression and renders metastatic disease incurable.

1.3.1 Alterations in *ESR1* leading to endocrine treatment resistance

ER is encoded by the *ESR1* gene and is the main target of endocrine treatment in ER-positive breast cancer, where ER is directly targeted by SERDs and SERMS and indirectly by

depletion of its main ligand/estrogen through aromatase inhibitors (AIs) or oophorectomy. Alterations in the structure and/or function of ER are commonly acquired in metastatic disease and lead to insensitivity to treatment. The complete loss of ER is attributed to only 10-20% of patients with a complete lack of therapy response (Li et al., 2013). Alterations involving the *ESR1* gene leading to resistance include recurrent point mutations, copy number amplifications and fusions resulting from genomic rearrangements (Jeselsohn, De Angelis, Brown, & Schiff, 2017).

1.3.2 *ESR1* mutations

In a substantial group (18-55%) of metastatic patients with ER-positive breast cancer, recurrent activating mutations in the LBD domain of ER are associated with resistance to anti-endocrine treatment (Carroll, 2016). These missense mutations are almost exclusively metastasis-specific and occur in patients with a prior history of anti-hormonal therapy, especially with aromatase inhibitors (Jeselsohn, De Angelis, et al., 2017).

One of the most common missense mutations in the LBD involves codon 537, which is a tyrosine kinase phosphorylation site. The mutation Y537N identified in a metastatic sample in 1997 was the first *ESR1* mutation found to be ligand-independently active and insensitive to tamoxifen and ICI (Zhang, Borg, Wolf, Oesterreich, & Fuqua, 1997). D538G (Merenbakh-Lamin et al., 2013), Y537S/C (Robinson et al., 2013; Toy et al., 2013) and L536Q/R (Robinson et al., 2013; Toy et al., 2013) are further commonly detected mutations in metastatic patients, whereas S463P and E380Q are less frequent (Angus et al., 2017; Toy et al., 2017). All the above mentioned mutations cause higher S118 phosphorylation and higher stabilization of the receptor compared to the wildtype (wt)-ER, however the greatest stabilization and phosphorylation is observed in the mutants with Y537S and S463P mutations (Toy et al., 2013; Toy et al., 2017). Although, there is

no correlation between the level of constitutive activity and the stabilizing effect of the mutation on the ER, the level of phosphorylation correlates with its activity in ERE assays (Toy et al., 2013; Toy et al., 2017). Fulvestrant can inhibit the activity of most of the above-mentioned ER mutants, however, the Y537S mutant requires significantly higher levels of drug that exceeds achievable dosage in patients (Jeselsohn et al., 2014; Toy et al., 2017). In general, SERDs and SERMs have only limited activity in tumors with *ESR1* mutations involving the LBD (Merenbakh-Lamin et al., 2013; Robinson et al., 2013; Toy et al., 2013). In addition to the acquisition of ligand-independence and endocrine resistance, genomic analyses also reveal mutation- and context-specific transcriptional effects of the mutant ERs which cannot be provoked even under long-term E2 stimulation of the wt-ER suggesting that mutant ERs gain distinct neomorphic changes (Bahreini et al., 2017; Blanchard, Vahrenkamp, Berret, Arnesen, & Gertz, 2018).

The presence of *ESR1* mutations has a negative impact upon progression-free survival of patients with advanced disease who are under treatment with aromatase inhibitors (Fribbens et al., 2016; Schiavon et al., 2015). Multiple studies show that *ESR1* mutations can be identified by droplet digital PCR (ddPCR) with high sensitivity and accuracy in cell-free DNA (cfDNA) or circulating tumor DNA of patients, which is often referred to as a "liquid biopsy" and serves as a surrogate for a metastatic tumor biopsy (Guttery et al., 2015; Schiavon et al., 2015; P. Wang et al., 2016). Our laboratory previously reported that *ESR1* mutations in cfDNA are detected in only a small fraction of patients with primary tumors, and the variant allele frequency (VAF) is as low as 0.001%, however ER mutations are enriched in metastatic tumors with a VAF of 34.3-44.9% in brain metastases and 1.4% in bone metastases (P. Wang et al., 2016). Moreover, these studies demonstrate the practicability and feasibility of liquid biopsy to detect *ESR1* mutations and its advantageous application for monitoring resistance development through serial blood draws. Thus,

patients with ER-positive breast cancer can be screened for presence of *ESR1* mutations at the time of initial diagnosis and during treatment, which confers the advantage of early detection of resistance before the establishment of metastatic disease, and allows adjustment of therapy based on mutation status.

In recent years, CDK4/6 inhibitors have been approved in combination with letrozole (based on phase II of PALOMA-1 trial) as first-line treatment in hormone-receptor (HR) positive HER2-negative advanced breast cancer or in combination with fulvestrant (based on PALOMA-3) for treatment of women with HR-positive, HER2-negative advanced or metastatic breast cancer with disease progression following endocrine treatment (Walker et al., 2016). The progression-free survival significantly increased with both combination therapies in the indicated patient groups with advanced disease. Since one of the pathways activating the CDKs is through ER-induced expression of cyclin D, the use of CDK4/6-inhibitors to prevent the G1-to-S phase transition in tumor cells with active ER mutants is appealing. Preliminary studies suggest that patients with *ESR1* mutations may benefit from a combined therapy of palbociclib with fulvestrant and warrants further studies for confirmation of its efficacy (Fribbens et al., 2016).

1.3.3 *ESR1* copy number (CN) amplifications

ESR1 amplifications have been reported in ER-positive primary and metastatic breast cancer (Basudan et al., 2018; Holst et al., 2007; Nembrot, Quintana, & Mordoh, 1990; Tomita et al., 2009). Earlier studies using fluorescence *in situ* hybridization (FISH) or array comparative genomic hybridization (aCGH) estimated the frequency of copy number amplifications as 20.6% to 22.6% and gains as 11.3% to 15.3% (Holst et al., 2007; Tomita et al., 2009). However, these frequencies were contradicted by other studies, which found *ESR1* amplifications to be a rare event

both in primary and metastatic samples (1%-1.7% and 1.3% respectively) (Brown et al., 2008; Jeselsohn et al., 2014; Reis-Filho et al., 2008). These conflicting results have been attributed to the different methods used to detect CN changes, sizes of applied amplicons, different signal counting guidelines, selection of breast cancer cases and contamination of tumor cells with normal cells (Horlings et al., 2008; Moelans et al., 2011; Nielsen et al., 2011). A recent study from our laboratory investigating the presence of amplifications in primary-metastatic paired patient-matched breast cancer samples using nanoString identified that the *ESR1* amplifications are specific to ER-positive breast cancers and were enriched in metastatic samples (14.9%) compared to primary lesions (6.6%) (Basudan et al., 2018), which aligns with earlier observations that amplifications are frequently found alterations in *ESR1* in ER-positive breast cancer. Importantly, this study also revealed that *ESR1* amplifications were accompanied by enrichment of *MYC* (45%) or *CCND1* (36%). The level of *ESR1* amplification correlates with the protein expression levels of ER in most studies, however, *ESR1* amplification was not the underlying mechanism in all samples that that showed high ER expression. Although ambiguity exists over its influence on outcome in patients, a great number of studies show that *ESR1* amplifications are associated with poor disease-free and overall survival and are found more frequently in advanced disease (Basudan et al., 2018; Brown et al., 2008; Ejlertsen et al., 2012; Holst et al., 2007; Li et al., 2013; Markiewicz et al., 2013; Moelans et al., 2011; Nielsen et al., 2011; Tomita et al., 2009). According to Li et al., *ESR1* gene amplification in an ER-positive patient-derived xenograft (PDX) stemming from a patient with endocrine-resistant metastasis sustained tumor growth in the absence of estrogen (Li et al., 2013), which is consistent with the observation that ER-positive breast cancer cell lines adapt in response to estrogen deprivation by overexpressing *ESR1*. Certainly, the increase in expression of a gene that is associated with growth and survival advantage is a mechanism for cancer cells to

thrive (Holst et al., 2007). Moreover, the detection of such amplifications in premalignant lesions (36% of benign papillomas and 8.3% of ductal hyperplasias) suggests that ER amplifications may also play a role as a tumor-initiating event (Holst et al., 2007). However, the clinical significance of *ESR1* amplifications and its contribution to endocrine resistance needs to be further investigated.

1.3.4 Fusion formation in cancer and its detection

An emerging and important mechanism leading to somatic genetic changes that drive cancer are genomic rearrangements, which are caused by the defective repair of DNA double strand breaks (DSB) and by erroneous DNA replication (Hasty & Montagna, 2014). DSBs can be repaired via three different pathways, which include non-homologous end joining (NHEJ), homologous recombination (HR) and break-induced replication (BIR). The classical NHEJ repairs the broken ends of the DNA by directly ligating them without the need for a template of a homologous sequence (Rodgers & McVey, 2016). Aberrations in this pathway are associated with fusion of chromosomes via telomere joining events as well as with chromosomal translocations, which include among others more complex rearrangements as seen in chromothripsis (Bailey et al., 1999; Bunting & Nussenzweig, 2013; Marcand, 2014). If defects are present in classical NHEJ, alternative pathways of NHEJ use microhomology-based mechanisms and end up repairing the DSB imprecisely. As a consequence, translocations are generated, which harbor those microhomologies at the fusion points and can be found in cancer cells (e.g. present in BRCA1/2 positive breast cancer) (Bunting & Nussenzweig, 2013; Hasty & Montagna, 2014). The less error-prone mechanism of HR corrects DNA damage with better specificity by using information from a homologous template (Hasty & Montagna, 2014).

Structural rearrangements include deletions, inversions, tandem duplications, translocations and amplifications. As a consequence, two or more unrelated genes conjoin and form a gene fusion (Edwards & Howarth, 2012; Stephens et al., 2009). These can also arise through non-structural rearrangements such as transcription read-through of neighboring genes (*cis*-transcription-induced gene fusions (TIGF)) or splicing mRNAs of genes located far from another or on different chromosomes (*trans*-TIGF) (Latysheva & Babu, 2016; Mertens, Johansson, Fioretos, & Mitelman, 2015). Depending on the site of breakpoint within each gene, the 3'-fusion gene may stay in-frame or become out-of-frame. Although the consequences of out-of-frame fusions are unknown, most of these fusions are potentially removed via nonsense-mediated decay or if the decay is escaped, a truncated form of the 5'-partner may exhibit its function (Mittal & McDonald, 2017). In fusions which sustain an intact open reading frame, a chimeric protein may arise through a merge of two known protein segments or a non-canonical chimeric transcript with novel function may occur. Furthermore, a gene can inherit a new 5' untranslated region (UTR) or 3'UTR upon fusion. These UTRs harbor promoter sites and/or binding sites for microRNAs and thus can change gene regulation and expression (Mittal & McDonald, 2017). In summary, formation of fusions may mutate the normal genes into tumor genes or interfere with the activity of oncogenes and/or tumor suppressor genes by altering their gene expression, regulation, dimerization, subcellular localization or by adding or removing functional/binding domains (Edwards, 2010; Hasty & Montagna, 2014). The occurrence of such variations in a tumor cell may lead to a survival benefit as well as may alter the drug targets causing drug resistance. This may give rise to tumor heterogeneity and development of metastasis through clonal evolution (Hasty & Montagna, 2014). However, it is thought that the majority of detected fusions are byproducts or "passengers" which do not exhibit tumorigenic features and are a sign of chromosomal instability

in cancer cells (Vogelstein et al., 2013). It is important to determine which gene fusions recurrently found in tumors play an active role in tumorigenesis as this may reveal novel drug targets.

In general, genomic rearrangements in hematologic tumors have been well known to alter normal genes into cancer genes. The most prominent example is the Philadelphia chromosome, which was the first described gene fusion (*BCR-ABL1*) associated with various forms of leukemia, particularly the chronic myeloid leukemia (CML) (Hasty & Montagna, 2014). More importantly, the ability to directly target the cancer promoting fusion protein - through which a tyrosine kinase becomes constitutively activated- revolutionarily transformed a once aggressive disease (CML) into a chronic disease with significantly improved prognosis. In contrast to hematological malignancies, the role of fusion genes in solid tumors was underappreciated due to their "low prevalence" that is implied by the ineffective detection of fusions by cytogenetic techniques (Giemsa-banding, FISH and array comparative genomic hybridization (CGH)) (Kumar-Sinha, Tomlins, & Chinnaiyan, 2008; Paratala et al., 2016; Stephens et al., 2009). The most clinically important fusions in solid tumors detected by guided methods are *EWSR1-FLI1* fusions in Ewing's sarcoma (via cytogenetic analysis), *TMPRSS2-ERG* fusions in prostate cancer (via microarray technology) and *EML4-ALK* fusions in non-small cell lung cancer (NSCLC) (via PCR with subsequent Sanger sequencing) (Parker & Zhang, 2013). Besides the limitations of these techniques, biological diversity of fusions, unfeasibility of guided methods to screen for fusions with unknown fusion partner genes and tumor heterogeneity in solid tumors rendered fusion detection even more challenging. Owing to the emergence of NGS methods and the development of computational tools which allow for comprehensive genomic analyses, many recurrent fusions with functional importance and high frequency have been detected in solid tumors, like the fusions *FGFR3-TACC3* in glioblastoma multiforme, *ESRRA-C11orf20* in ovarian cancer or *RET* fusions,

to give just a few examples (Kumar-Sinha et al., 2008; Paratala et al., 2016; Parker & Zhang, 2013; Stephens et al., 2009). With the application of novel technologies, not only the numbers of detected fusions have intensified dramatically but also the time to detect decreased significantly, in contrast to the guided methods (directed by cytogenetic information) which yielded less than 50 fusions, unbiased sequencing techniques detected roughly 500 to 8000 fusions per year (Mertens et al., 2015).

Paired-end whole-genome sequencing (WGS) and RNA-seq are currently the two predominantly used sequencing technologies for detecting gene fusions. Both deep-sequencing methods are unbiased (WGS more than RNA-seq) and accurate approaches providing a far superior resolution compared to the guided methods, such as the ability to detect intrachromosomal rearrangements even when involved genes are in close proximity. However, the whole process demands extensive time from sample preparation to completing compute-intensive analysis. Furthermore, since not all fusion genes are transcribed or even translated, validation of potential fusion gene candidates detected by WGS costs additional time and resources and is a concerning disadvantage (Mertens et al., 2015). Although, RNA-seq favorably allows the detection of the cis-/trans-TIGFs, it may yet suffer on the one hand from suboptimal RNA quality leading to increased rate of false negative results and on the other hand from incapability to detect chimeric transcripts when present in low levels (Schram, Chang, Jonsson, & Drilon, 2017).

A great number of computational tools/algorithms have recently been developed to detect fusions from RNA-seq +/- WGS results. In brief, to detect fusions, most of the current tools align the paired-end reads to the reference genome and identify the fusions through two types of supporting reads, which can be named differently by each developer. The first type of supporting reads are the *spanning reads*, where each read of the pair aligns to a different gene. The second

type of reads are the *split reads,* where one of the read pair contains sequences from both genes involved in fusion, and have the advantage to provide the exact junction point of the fusion (Liu et al., 2016). While most of the tools (e.g. ChimeraScan, FusionCatcher, JAFFA, TopHat-Fusion) use only RNA-seq input data, some tools (FusionMap, INTEGRATE, nFuse) are able to use both WGS and RNA-seq as input (S. Kumar, Razzaq, Vo, Gautam, & Li, 2016). Other distinguishing aspects of these numerous tools include format of reads (single- vs. paired end reads), alignment method (against reference genome vs. *de novo* assembly) as well as performance in terms of number of detected fusions, sensitivity, specificity, computational time consumption and memory usage (S. Kumar, Razzaq, et al., 2016; S. Kumar, Vo, Qin, & Li, 2016). The calling of fusions also depends on the filtering based on various parameters that may be unique to each tool, such as the required minimum size of the spanning and split reads, trimming options and other criteria (Liu et al., 2016), in addition to filters applied by the end user. Many comparative assessments on the fusion callers have been performed and all agree that no tool outperforms in all evaluated aspects and that there is very low overlap of called fusions between different tools (S. Kumar, Vo, et al., 2016; Latysheva & Babu, 2016; Liu et al., 2016). However, these assessments represent a base for researchers to select the best algorithm(s) suited to their needs and available resources.

A newer method to screen for fusions is anchored multiplex PCR (AMP) for targeted NGS that is commercialized as the FusionPlex assay (ArcherDx, Boulder, CO). AMP is a rapid and cost-effective assay that uses unidirectional one-sided nested primers and can be employed for custom targeted sequencing (Zheng et al., 2014). Although, it has some limitations with detection of reciprocal fusion transcripts, it has a practical use to screen for fusions with one of the fusion partner being consistently involved (promiscuous genes) and having altering breakpoints or fusion partners (Afrin et al., 2018; Zheng et al., 2014).

Overall, fusions which are unique to cancer cells may play an important role as biomarkers for disease prognosis and therapy response to support therapeutic decision making and serve as unique targets for developing new targeted therapies. More importantly, detection and characterization of fusions involving drug target proteins may help unravel acquired resistance mechanisms to current state-of-art treatment. With specific relation to ER-positive breast cancers, in my thesis I am investigating the role *ESR1* fusions as a mechanism of endocrine resistance.

1.3.5 *ESR1* fusions

ESR1 fusions - similar to *ESR1* hotspot mutations - are almost exclusively found in metastatic ER-positive breast cancers and are a novel mechanism associated with endocrine resistance that needs to be elucidated. Currently, only a few research groups, including our lab, are investigating the frequency and functional role in ER-positive breast cancer.

The first described *ESR1* fusion was *ESR1-YAP1,* which was detected in a patient-derived human breast cancer xenograft (PDX) model in 2013 (Li et al., 2013). This fusion proved enticing particularly because the translocation took place at exon 6 (E6), at AA sequence 365, causing loss of the ER LBD. While retaining the expression of N-terminal *ESR1*, it acquires the C-terminal end of the *YAP1* gene that harbors the WW protein-interacting and the transactivation domains. In *in vitro* functional assays, the *ESR1-YAP1* fusion provokes estrogen-independent cell growth and as expected, renders insensitivity to fulvestrant, which exerts its effect by binding to the LBD of ER. In addition, *ESR1* fusions that partner with multiple truncated forms of its adjacent gene *CCDC170* have been identified, which occur relatively frequently not only in breast cancer but also in endometrial carcinoma (Holst et al., 2016; Hu et al., 2018; Veeraraghavan et al., 2014). In these fusions, *ESR1* endows its promoter to the truncated *CCDC170,* which is described as non-coding

promoter switching. By conjoining with the *ESR1* promoter, the truncated *CCDC170* gains higher expression and increases the tumorigenic, metastatic and endocrine-resistant potential of tumor cells (Veeraraghavan et al., 2014).

In a comprehensive sequencing effort of six patients with therapy-refractory advanced ER-positive breast cancer enrolled in the University of Pittsburgh rapid autopsy program, our laboratory analyzed the genetic evolution of normal-primary-metastasis matched samples. In a nodal recurrence of a patient, Dr. Hartmaier, a previous postdoctoral fellow, discovered a metastasis-specific *ESR1-DAB2* fusion that not only possessed the same *ESR1* breakpoint as the previously described *ESR1-YAP1* fusion but also was detected in the RNA and protein level (Hartmaier et al., 2018). When our lab screened for further *ESR1* fusions in breast cancer samples sequenced via exome–capture RNA-seq on 59 primary-metastasis paired as well as 13 metastatic and 10 primary unmatched breast cancer samples, we found an *ESR1-GYG1* fusion in a bone metastasis sample that again revealed the identical breakpoint as the previous *ESR1* fusions with *YAP1* and *DAB2*. I validated the RNA and protein expression for *ESR1-GYG1* fusion via RT-PCR and immunoblot, respectively (Hartmaier et al., 2018). To estimate the frequency of such fusions, Dr. Hartmaier expanded the exploration to hybrid-capture-based target sequencing in a large cohort of 9,542 breast tumors and 254 ctDNA samples from patients with advanced breast cancer collected by Foundation Medicine Inc. (FMI, Cambridge, MA). A total of 7 *ESR1* fusions (**Figure 2**), showing the identical junction breakpoint as *ESR1-YAP1* were detected in four solid tumors (*ESR1-SOX9, ESR1-MTHFD1L, ESR1-PLKHG1* and *ESR1-TFG*) and in three ctDNA samples (*ESR1-NKAIN2, ESR1-AKAP12* and *ESR1-CDK13*). Subsequent *in vitro* characterization of the fusions *ESR1-DAB2, ESR1-GYG1* and *ESR1-SOX9* in 293T cells via ERE-Tk-luciferase assay demonstrated the ligand-independent activity and resistance to ER-targeting drugs, where the level

of activity was influenced by its 3'-partner. Additionally, Hartmaier et al. revealed the presence of frequent copyshifts, which are indicative of intragenic rearrangements, in the *ESR1* gene upstream of exon 6 in ER-positive metastatic samples. The successful application of the *copyshift* method as an initial screening tool was demonstrated to uncover rearrangements within genes that harbor fusions. Particularly, the four *ESR1* fusions with *DAB2*, *SOX9, MTHFD1L* and *PLEKHG1* exhibited *copyshift*-positivity (Hartmaier et al., 2018).

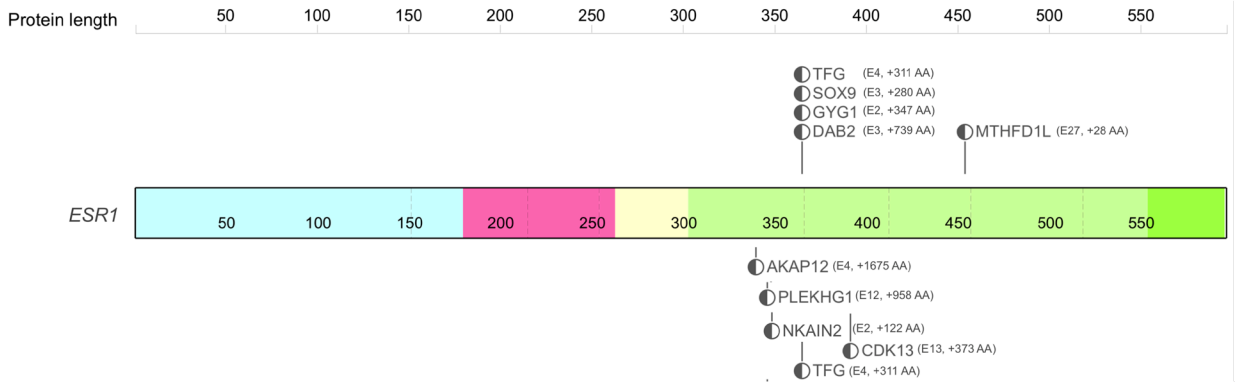

Figure 2 *ESR1* fusions with a breakpoint clustering at exon 6

The *ESR1* reference transcript with color coded functional domains is shown above. The dashed lines show the exon boundaries, starting with exon 3 (= the first exon in the coding sequence) up to the last exon 10. All fusions, except *ESR1-DAB2* and *ESR1-GYG1*, are detected in the patient samples collected by Foundation Medicine (Hartmaier et al., 2018). The numbers next to the 3'-partner gene refer to the breakpoint at exon level and additional amino acids acquired from the fusion partner, respectively. Domains from left to right: Activation Function-1; DNA-binding domain; Hinge domain; Activation function-2; and F domain.

Another important study highlighting the functional role of *ESR1* fusions (E6) in therapy-refractory advanced breast cancer was recently published, where biological insights about three in-frame *ESR1* fusions with a truncation at E6 were described (Lei, Shao, et al., 2018). The two fusions detected in therapy-refractory advanced breast cancer, *ESR1-YAP1* (earlier published by the same lab) and *ESR1-PCDH11X* were confirmed to be ligand-independently active and resistant to fulvestrant, whereas the fusion *ESR1-NOP2* found in a treatment-naïve primary tumor was inactive despite being highly expressed. This trend in activity was also reflected in *in vitro* and *in*

vivo proliferation assays, where fusions with *YAP1* and *PCDH11X* induced growth in low-estrogen conditions that was higher than the growth promotion in parental cells when stimulated with E2. Additionally, the two active *ESR1* fusions together share 445 DNA binding sites with activated ER and induce the expression of estrogen-responsive genes in the absence of estrogen. Furthermore, the authors suggest that activating *ESR1* fusions stimulate cell division by an increased phosphorylation of Rb (retinoblastoma associated protein). Thus, palbociclib, a CDK4/6 inhibitor was shown to work successfully against the fusion driven growth in *in vitro* and *in vivo* models in a dose-dependent manner. Based upon their observations, Lei et al. divide the *ESR1* fusions, where *ESR1* acts as the 5'-partner, into 3 groups: (1) *ESR1*-e2 fusions as promoter trap, (2) *ESR1*-e6 in-frame fusions as driver of metastasis and therapy resistance and (3) *ESR1*-e7 in-frame or out-of-frame *ESR1* fusions with truncation at any exon as inactive in E2-absence and thus non-proliferative (Lei, Gou, & Ellis, 2018).

In summary, *ESR1* fusions are predominantly distinct private fusions unique to each breast cancer. Unlike the prototype fusion *BCR-ABL* found in acute lymphoblastic leukemia, *ESR1* partners with various genes and thus, is classified as a promiscuous fusion gene partner. *ESR1* has may act as a 5'- fusion partner retaining the N-terminal domains or as 3'-fusion partners retaining the C-terminal domains (**Figure 3**). Although *ESR1* has been found to break at any exon, *ESR1* fusions with functional consequences occur more likely at exon 2 and exon 6. Since *ESR1* fusions affect the main driver and main drug target in ER-positive breast cancers, it is important to determine the phenotype these fusions confer and how they relate to endocrine treatment resistance.

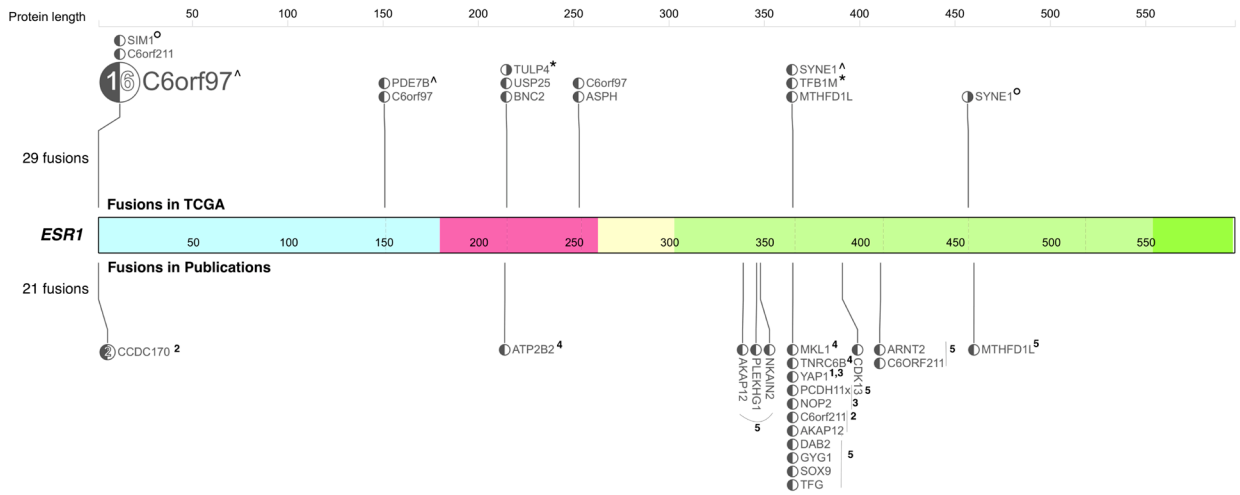

Figure 3 Panorama of identified *ESR1* fusions in publications

The structure of *ESR1* is depicted along with the exon boundaries and amino acid sequence length. The functional domains are color coded. The fusion partners of *ESR1* are graphed/mapped on the *ESR1* structure, where they fuse. The symbols are ◐ for *ESR1* as 5'-partner; ◑ for *ESR1* as 3'-partner. The fusions depicted above the structure are found in the TCGA database. The fusions annotated with ^ are found in uterine corpus endometrial cancer (UCEC). * in lung squamous cell carcinoma (LUSC) and ° in sarcoma (SARC). *C6orf97* is also known as *CCDC170* and 4 of these are found in UCEC. All other fusions stem from breast carcinoma. The fusions below the structure are found in publications. *C6orf211* is outdated name for *ARMT1*. 1) (Li et al., 2013), 2) (Giltnane et al., 2017), 3) (Lei, Shao, et al., 2018), 4)Tan et al.(2016) SABCS Poster: P3-04-01, 5)(Hartmaier et al., 2018); Most common breakpoints are bimodally distributed on the *ESR1* with a peak at 5' UTR promoter site and the other at the end of exon 6. Graph was designed with ProteinPaint (https://proteinpaint.stjude.org/)

1.4 Study objectives

Based upon observations from our laboratory and others, we hypothesize that *ESR1* fusions (E6) involving the loss of the LBD domain lead to tumor progression via unresponsiveness to anti-endocrine treatment as well as through acquired ligand-independent constitutive activation of ER. Thus, the main goal of this project is to understand the biological function and significance of the *ESR1* (E6) fusions identified in metastatic breast cancer patient samples and a patient derived xenograft (PDX) breast cancer model harboring an *ESR1-LPP* fusion (identified in collaboration with Champions Oncology). Thus, the objectives to fulfill the main goal are comprised of:

(1) assessing the ligand-independence of ER fusions in hormone-deprived conditions by employing ERE-Tk luciferase reporter assays

(2) determining the effect of *ESR1* fusions on subcellular localization

(3) determining if the *ESR1* fusions confer a growth/proliferative advantage to tumor cells with/out endocrine treatment

(4) evaluating the transactivation properties of *ESR1* fusions on known ER-regulated downstream genes.

2.0 Methods

2.1 Screening for *ESR1* fusions and *in silico* functional prediction

To identify fusion transcripts in ER-positive breast cancer, RNA-seq was performed on six patients from our rapid autopsy program and exome–capture RNA-seq on 45 patient-matched primary-metastasis paired breast cancer samples. The process of how these samples were obtained and sequenced have been previously described in detail (Hartmaier et al., 2018; Priedigkeit et al., 2017; Vareslija et al., 2018). After receiving permission from University of Michigan Clinical Sequencing Exploratory Research (CSER) to access their "MET500" dataset (dbGaP Study Accession number phs000673.v3.p1), which is a large dataset containing WES and RNA-seq information of metastatic tumor samples, a further screening of *ESR1* fusions in the metastatic breast cancer samples were performed (Robinson et al., 2017). Fusion transcripts were discovered by FusionCatcher version 99.7b (Nicorici et al., 2014) and STAR-Fusion version v1.1.0 (Haas et al., 2017). For both algorithms the default parameters were used. The code used for FusionCatcher is *fusioncatcher -d $data_reference -i $fastq_folder -o $out -p $ppn --visualization-sam*, where the current data reference is downloaded by using the built-in tool *fusioncatcher-build* (for more details, please refer to the online available manual). To identify fusions with STAR-Fusion, first the raw sequences were aligned with STAR version 2.5.3a using the code *STAR --genomeDir $star_index --readFilesIn $sampleR1 $sampleR2 --readFilesCommand zcat --twopassMode Basic --chimOutType SeparateSAMold --outReadsUnmapped None --chimSegmentMin 12 --chimJunctionOverhangMin 12 --alignSJDBoverhangMin 10 --alignMatesGapMax 200000 --alignIntronMax 200000 --chimSegmentReadGapMax parameter 3 --alignSJstitchMismatchNmax*

5 -1 5 5 --runThreadN $ppn --outSAMtype BAM SortedByCoordinate --outFileNamePrefix $sample.star followed by the command *STAR-Fusion --left_fq $sampleR1 --right_fq $sampleR2 --genome_lib_dir $starFusion_library -J $sample.starChimeric.out.junction --output_dir $out.* The genome reference GRCh38_gencode_v24_CTAT_lib_Mar292017_prebuilt and the index STAR-index-gencode24-starFusion-74 for STAR-Fusion is downloaded and installed per the descriptions in the online manual. Fusions in tumor samples from rapid autopsy had been also detected with the fusion calling algorithm ChimeraScan (Iyer, Chinnaiyan, & Maher, 2011) by Dr. Hartmaier and were available for a reference (Hartmaier et al., 2018).

The functional annotation of fusion genes was done by the two prediction tools called Oncofuse-1.1.1 (Shugay, Ortiz de Mendibil, Vizmanos, & Novo, 2013) and Pegasus (Abate et al., 2014). Since the set of fusions called by ChimeraScan were already available, first those outputs were annotated using both algorithms to evaluate if they produce equal predictions. The command used for Oncofuse is *java -Xmx2G -jar Oncofuse.jar (-a hg38) $input_file $input_type EPI $output_file* and uses by default the hg19 as genome assembly unless specified otherwise (see online manual). Pegasus, which annotates only based on genome assembly hg19, was run with the command *pegasus.pl -c config.txt -d data_spec.txt -l log_folder -o output_folder* (see online manual for further details).

After verifying that the results obtained by both tools were similar, the outputs generated by FusionCatcher and STAR-Fusion were annotated only with Oncofuse in order to avoid any errors that would occur while transforming the data into Pegasus-conform general input file. The *ESR1* fusions with significant driver scores were depicted for functional analysis.

2.2 Cell culture

All cell lines were obtained from ATCC, authenticated at the University of Arizona Genetics Core and tested negative for mycoplasma with a MycoAlert Mycoplasma Detection Kit (Lonza, Basel, Switzerland). The cells were maintained at 37°C in 5% CO_2. All culture media and media supplements were obtained from Life Technologies. HEK293T, MCF7 and MDA-MB-231 (MM231) cells were cultured in Dulbecco's modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (FBS). T47D cells were cultured in Roswell Park Memorial Institute (RPMI) medium 1640 supplemented with 10% FBS. MCF7/C4-12 (C4-12) cells, an estrogen-independent derivative of MCF7 cells generated after 9 months of long-term hormone deprivation were cultured in phenol red free minimum essential medium (MEM) Alpha supplemented with 5% charcoal-stripped FBS (CSS) (Oesterreich et al., 2001).

2.3 Plasmids

HA-tagged wild-type (wt) *ESR1* in pcDNA3.1(+) vector was purchased from Addgene (#49498). All the other *ESR1* constructs (truncated *ESR1* at exon 6 (E6), *ESR1-DAB2*, *ESR1-GYG1*, *ESR1-SOX9*, *ESR1-LPP* and *ESR1*(E7)-*LATS1*), once they were computationally identified, were *de novo* synthesized with an N-terminal HA-tag into pcDNA3.1(+) between the NheI and KpnI restriction cutting sites by Gene Universal Inc. (Newark, DE), who confirmed the sequences of all constructs by Sanger sequencing and diagnostic digestion.

Scheme of the constructs:

NheI – Kozak sequence – ATG –HA-tag – **sequence of interest** – Stop codon - KpnI

2.4 Subcloning into lentiviral backbone

To generate lentiviral plasmids, the above described *ESR1* constructs in pcDNA3.1 were cut first with the restriction enzyme KpnI (New England Biolabs, NEB), the plasmid with wt-*ESR1* was cut with ApaI (NEB) and the empty lentiviral vector pCDH-MSCV-MCS-EF1α -GFP-T2A-Puro (#CD713B-1, System Biosciences, **Figure 4**) was cut with NotI (NEB). Because the 3'-ends of the *ESR1* constructs and the 3' insertion site of the lentiviral vector are incompatible, the ends were blunted with T4 DNA Polymerase (NEB) at 12°C for 15 minutes (min). The blunting reaction was inactivated by adding EDTA (UltraPure™ 0.5M EDTA, pH 8.0, Invitrogen) and by heat treatment at 75°C for 20 min. After column purification with Qiaquick PCR purification kit (Qiagen), all plasmids were cut with NheI (NEB). All digested plasmid products and only 5ul of the lentiviral backbone were run through the agarose gel and visualized by ChemiDoc™ XRS+ System (BioRad). The bands in accordance with the expected base pair lengths for the inserts were cut out from the gel and extracted using the Qiaquick gel extraction kit (Qiagen). The remaining 45 ul of lentiviral backbone was treated with the calf-intestinal alkaline phosphatase (CIP) (NEB) to prevent re-ligation and column purified using Qiaquick PCR purification kit (Qiagen).

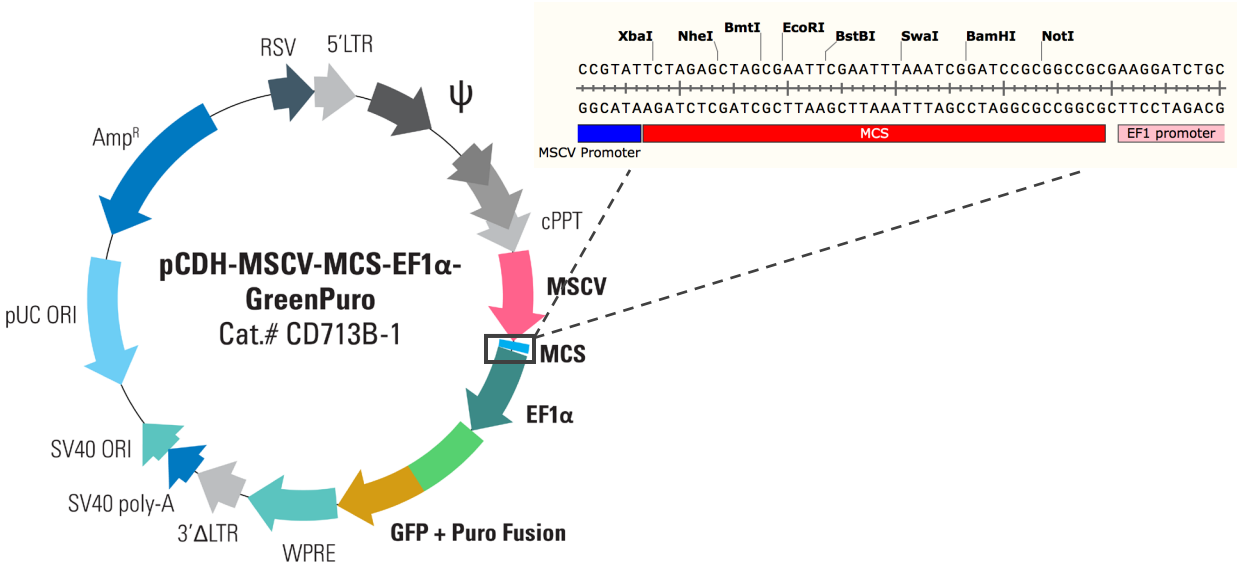

Figure 4 The lentiviral vector pCDH-MSCV-MCS-EF1α-GFP-T2A-Puro

The map of the vector (left) and its multi cloning site sequence showing the cutting sites by the restriction enzymes (right) are depicted in the graph above.

Next, the *ESR1* constructs were ligated in a 1:3 insert to vector ratio using T4 DNA ligase (#M0202T, NEB) into the target lentiviral backbone. After transformation of the Stbl-3 cells (One Shot® Stbl3™ Chemically Competent E. coli, Thermo Fisher Scientific) with each *ESR1*-construct, the selection followed with ampicillin (VWR) on agar plates at 37°C for 16 hours (h) and up to five clones per construct were picked. After obtaining pure plasmids using the Qiagen Plasmid Mini Kit (Qiagen), a first diagnostic digest was performed with the restriction cutting enzyme SpeI (NEB). Second, the sequences of the plasmids, which passed the diagnostic digestion, were further confirmed through sequencing by GENEWIZ (South Plainfield, NJ). For sequencing, the primers (**Table 1**) recommended by System Biosciences and purchased from IDT (Integrated DNA Technologies) were used.

Table 1 PCR primer sequences for sequence analysis of the insert in pCDH backbone

Primer name	Description	Primer Sequence
pCDH-MSCV_F	is a forward primer starting at MSCV promoter, immediately before the insert	GGGGTACAGTGCAGGGGAAAGAAT
pCDH-EF1_R	is a reverse primer that primes at EF1 promoter, which is located right after the insert site	CTCGACCTGAGCTTTAAACTTACC

2.5 Lentiviral production

For generating cells stably expressing the *ESR1*-constructs, the previously subcloned *ESR1*-constructs (lentiviral vector or empty vector; pCDH-MSCV-MCS-EF1-GFP-T2A-Puro) along with the packaging plasmids pEZ-DNA, pMDL, pRev and pVSVG were introduced into the packaging cells HEK293T by transfection with PEI (polyethylenimine, Polysciences) in a 3:1 PEI to DNA ratio. In the following days, conditioned medium (RPMI for T47D cells and MEM alpha for MCF7/C4-12 cells), which was collected after incubation with the packaging cells, was filtered through a 0.45 μm syringe filter (Merck, MilliporeSigma) and used to infect the target cells in the presence of polybrene (4 μg/ml, Sigma Aldrich) 36h, 48h and 60h post-transfection, respectively. Screening using the GFP channel with an Olympus IX83 inverted microscope ascertained that the plasmid was transfected into the target cells expressing GFP protein. However, the efficiency of the infection was relative to the size of the insert. One day after the recovery period, the target cells were selected with 1μg/μl puromycin (Life Technologies) for 7 days. The complete elimination of uninfected non-resistant parental cells under puromycin selection in 4 days and the presence of only GFP-expressing cells under microscope were reassuring for a successful selection. Subsequently, cells were collected for protein extraction to confirm the expression of interest fusion proteins by immunoblot.

2.6 Protein extraction and quantification

After washing the cells with ice-cold PBS, the cell suspension was centrifuged at 1,000*g* for 4 min. The cell pellet was lysed for 20 min with RIPA buffer containing 1x Halt protease and phosphatase inhibitor (Thermo Scientific) while keeping it on ice with occasional pipetting up and down. The cells were further disrupted by 2 cycles of ultra-sonication (Ultrasonic Processor GEX130, Cole-Parmer) for 5 seconds with 10 seconds cooling down on ice in between. Then, the samples were centrifuged at 14000 rpm at 4°C for 10 min. The supernatant was aliquoted and stored at -80°C.

For estimating the total protein concentration, the bicinchoninic acid assay (BCA) was performed. Bovine serum albumin (BSA) standards ranging from 0 mg/ml to 1.5 mg/ml were used to generate a standard curve. The samples were diluted 1:25 before adding to the wells of a 96 well plate. Next, a working reagent was prepared with 1 part of BCA reagent A and 50 parts of BCA reagent B (Pierce BCA Protein Assay Kit, Thermo Scientific) and added to both the samples and the standards in a 1:20 ratio. After 30 min of incubation at 37°C, the absorbance of the samples was read at 562 nm wavelength with a spectrophotometer (GLOMAX®-Multi+ Microplate Multimode Reader, Promega). The total protein concentration in the samples was determined with reference to the standard curve generated by the correlation of the BSA concentrations to the BSA absorption.

2.7 Immunoblot

To validate the expression of the *ESR1* fusions in the stably expressing cells, 60 µg whole cell protein lysates were separated on 10% SDS-PAGE gel at 180V after the proteins ran at 80V until stacking. After the completion of the gel separation, the proteins were transferred onto PVDF membrane at 80V at 4°C for 90 minutes. Post-transfer, the membrane was dried at room temperature for 30 minutes, reactivated in methanol and then blocked in Odyssey PBS blocking buffer (LI-COR) at room temperature for 1h. Subsequently, the membrane was incubated with primary antibody (anti-*ESR1*, see specifics in **Table 2**) overnight at 4°C. After washing the membrane, the blot was incubated with the adequate secondary antibody for 60 min at room temperature. After removal of the excess secondary antibody, the blot was imaged with the Odyssey Infrared Imager System (LI-COR). After stripping the membrane with 1x NewBlot PVDF 5X Stripping Buffer (LI-COR) for 20 min and ensuring of the complete removal of fluorescence, the membrane was re-blotted with the anti-HA-tag and anti-β-actin primary antibodies overnight at 4°C (specifics in **Table 2**).

Table 2 The specifics of primary and secondary antibodies used in Immunoblot

Primary Antibodies	**Species**	**Dilution ratio**	**Vendor**	**Catalog #**
HA-tag	rabbit	1:1500	Cell signaling	3724
ESR1 (N-terminal)	rabbit	1:2000	Millipore	04-820
β-Actin	mouse	1:15000	Sigma-Aldrich	A5441
Secondary Antibodies	**Target Ab**	**Dilution ratio**		
IRDye® 680LT Goat anti-Mouse IgG	Anti-β-actin	1:15000	LI-COR	925-68020
IRDye® 800CW Goat anti-Rabbit IgG	Anti-HA Anti-ER	1:15000	LI-COR	926-32211

2.8 ERE-Assays

Cells were seeded in 24-well-plates and kept in regular growth medium until they became 80% confluent. For each cell line, the seeding density was adjusted to reach the target confluency level in 24h. Cells were then serum starved in modified improved MEM media enriched with 5% CSS for 24 hours until their transfection using the Lipofectamine 3000 (Invitrogen) method following the manufacturer's protocol. The ERE-assay was performed in both parental cells that were transiently transfected with *ESR1* plasmids or stably expressing the *ESR1* constructs of interest. For employing the ERE-assay on transiently transfected cells, in addition to the vectors containing the firefly luciferase that is controlled by estrogen-response-element (ERE-Tk-Luc) and the Renilla luciferase that is constitutively expressed under null-promoter as an internal control, the cells were also transfected with one of the vectors containing the *ESR1* constructs including *ESR1* fusions, wild-type or truncated forms of *ESR1* or the empty vector (pcDNA3.1+). The ERE-assay in stably expressing cells required only the co-transfection of firefly luciferase with Renilla luciferase in 5:1 ratio (250ng and 50ng, respectively). 24 hours after transfection, the cells were exposed to treatment with 100 pM estradiol only or a combination of 100 pM estradiol with one of the anti-estrogen drugs ICI182,780 (ICI) or 4-hydroxy-tamoxifen (4-OHT) to a 100 nM final concentration. 48 hours post-transfection (and there for 24 hours post E2 and drug treatment), the cells were harvested and assay was performed with Dual-Luciferase Reporter Assay kit (Promega) following the manufacturer's protocol. The luminescence was measured with GLOMAX®-Multi+ Microplate Multimode Reader (Promega) and analyzed by normalizing the firefly luciferase values to the values for Renilla luciferase, which served as an internal control for expression. The fold change in luciferase activity for any transiently transfected or stable expressing cell was compared to the vehicle treated empty vector containing cells and was presented in bar graphs with mean +/-

SD. Statistical analysis was done by employing one-way ANOVA with Dunnett's post hoc test for multiple comparison in GraphPad Prism 7. P-values less than 0.05 were considered statistically significant and were annotated as follows * $p<0.5$, ** $p<0.01$. *** $p<0.001$ and **** $p<0.0001$.

2.9 Cell proliferation assay

Proliferation assays were performed in stably expressing T47D cells. Cells were grown to a confluency level of 70 – 80% in regular medium. On the day of the start of the experiment, the cells were washed twice with PBS and then detached from the plate with the help of phenol-red free 0.5% trypsin-EDTA (Invitrogen). After washing off the trypsin and resuspending the cells in IMEM with 5% CSS, cells were counted via automated cell counter (Countess II, Invitrogen). 5000 cells/well were seeded into 96 well plates in IMEM with 5% CSS. Additionally, for the controls (parental cells, empty vector and wt-*ESR1*-expressing cells), a 1 nM E2 and a vehicle treated condition was included in the experiment. The cells were followed for 14 days in the incubator and an additional set was monitored for morphological assessment using IncuCyte Zoom Live Cell Imaging System (Essen Bioscience).

On Day 0, 1, 3, 7 and 14, the cells were incubated with the cell viability indicator PrestoBlue (Thermofisher Scientific) for 2h and the fluorescence was measured by the Victor X4 Plate Reader (Perkin-Elmer). After subtracting the blank values, all time point values were normalized to Day 0 value in each construct. The mean of the 6 replicates +/- its SD was calculated. The mean ratios +/- SEM of growth comparing Day 7/Day 0 and Day 14/Day 0 of two independent proliferation assays were presented as bar graphs. Statistical significance was determined by one-way ANOVA with Sidak's post hoc test for multiple comparison using GraphPad Prism 7. P-

values less than 0.05 were considered statistically significant and were annotated as follows * $p<0.5$, ** $p<0.01$. *** $p<0.001$ and **** $p<0.0001$.

2.10 RNA extraction and quantitative PCR

For evaluation of the expression of genes regulated by wildtype estrogen receptor (ER), cells were hormone-deprived for 3 consecutive days. After washing the cells daily twice with PBS and growing them in IMEM + 5% CSS, 700,000 cells/well were seeded into 6 well plates. 24 hours post-plating, cells were exposed to one of the 4 different treatment conditions (vehicle, 1nM E2 and/or 100nM ICI) for 24 hours. After the cells were washed with PBS and media was fully aspirated, the plates were put in -80°C.

Cells were lysed according to the Qiagen protocol for RNA-extraction and treated with RNase-free DNase (Qiagen) for the elimination of the DNA. Subsequently, employing the RNeasy Mini Kit (Qiagen) as per manufacturer's protocol, RNA was extracted from the cell lysates and the concentration was quantified via the spectrophotometer Nanodrop 2000 (Thermo Scientific).

In a 10µl reverse transcription reaction at 37°C for 15 min, 500 ng RNA was converted to cDNA (complementary DNA) with PrimeScript RT Master Mix (Takara Bio). The ER-regulated genes *GREB1, TFF1* and *IGFBP4* and the housekeeping gene *RPLP0* as the reference gene were employed to assess the transcriptional effect of the *ESR1* fusions through quantitative real-time PCR. The cDNA was diluted 1:9 in RNA-free water and mixed with the SsoAdvanced SYBR Gren Supermix (Bio-Rad) along with the forward and reverse primers (the sequences can be found in the **Table 3)**. A 12ul reaction was performed in 384 well plates in the CFX384 thermocycler (Bio-Rad) according to the manufacturer's protocol. Each samples was normalized to its *RPLP0*

housekeeper control and assessed using the $2^{-\Delta\Delta C_T}$ method as previously described (Livak & Schmittgen, 2001; Schmittgen & Livak, 2008).

Table 3 The qPCR primer sequences

Primer name	Primer Sequence	Company
RPLP0-F	TAAACCCTGCGTGGCAATC	IDT
RPLP0-R	TTGTCTGCTCCCACAATGAAA	
IGFBP4-F	ACGAGGACCTCTACATCATCC	IDT
IGFBP4-R	GTCCACACACCAGCACTTG	
GREB1-F	GGTTCTTGCCAGATGACAATGG	IDT
GREB1-R	CTTGGGTTGAGTGGTCAGTTTC	
PGR-F	TCGCCTTAGAAAGTGCTGTC	IDT
PGR-R	GCTTGGCTTTCATTTGGAACG	
ESR1-F (3'-end)	GAGTATGATCCTACCAGACCCTTC	IDT
ESR1-R (3'-end)	CCTGATCATGGAGGGTCAAATC	

The fold change in gene expression for any stable expressing cell was compared to the untreated parental cells and was presented in bar graphs with mean +/- SD. For assessing its statistical significance, two-way ANOVA with Dunnett's post hoc test for multiple comparison was employed using GraphPad Prism 7 (Goni, Garcia, & Foissac, 2009). P-values less than 0.05 were considered statistically significant and were annotated as follows * $p<0.5$, ** $p<0.01$, *** $p<0.001$ and **** $p<0.0001$.

2.11 Immunocytochemistry of the *ESR1* fusions in transiently transfected cells

MCF7, T47D and 293T cells were plated onto coverslips in 24-well-plate in regular media. When cells reached a confluency level of at least 50%, they were transfected with 500 ng of the *ESR1* constructs (wt-*ESR1*, trunc*ESR1*, *ESR1-DAB2*, *ESR1-GYG1*, *ESR1-SOX9* and *ESR1-LPP*)

using Lipofectamine 3000 as per manufacturer's protocol. 48 hours post-transfection, the cells were fixed with 4% paraformaldehyde (PFA, Electron Microscope Sciences) dissolved in 1x PBS (phosphate buffered saline, Fisher Scientific) for 10 min at room temperature (RT) and blocked for 30 min in blocking buffer containing 0.3% Triton X-100 (Sigma) and 5% BSA (Bovine Serum Albumin (A9647), Sigma). Subsequently, the cells were first exposed to the primary mouse antibody against ER (ER-6F11, Leica) diluted 1:50 in blocking buffer for 60 min at RT, washed and then incubated with the 1:200 diluted secondary anti-mouse antibody Alexa Fluor 546 (Invitrogen) another 45 min before further washing. After the completion of immunostaining of ER, the cells were treated with a 1:500 diluted rabbit antibody against HA-tag (3724, Cell Signaling) for 60 min, then washed, followed by secondary antibody treatment with anti-rabbit Alexa Fluor 488 (Invitrogen) for 45 min. After the antibody labeling, the cover slips with the cells were transferred onto slides and mounted with a drop of Prolong Diamond Antifade mountant with DAPI (Invitrogen). Coverslips were imaged using both the Olympus IX83 inverted microscope and the Nikon A1 advanced confocal system.

3.0 Results

3.1 Screening for *ESR1* fusions

A comprehensive sequencing of a series of breast cancers in six patients from our rapid autopsy program revealed an *ESR1* fusion (*ESR1-DAB2*) in a nodal recurrence of a patient (Hartmaier et al., 2018). The supraclavicular lymph node metastasis had developed while the patient was under treatment with Letrozole, following a switch from long-term treatment with Tamoxifen. Interestingly, the *ESR1-DAB2* rearrangement occurred at the end of E6 (AA365) and led to a loss of the ligand-binding domain of the ER which is the target domain where anti-estrogen treatments bind. This *ESR1* fusion presented with the same breakpoint as the previously reported *ESR1-YAP1* fusion that was found in a PDX model, and was found to show not only unresponsiveness to endocrine treatment but also hormone-independent activity, which was further reflected in the proliferation assays under hormone-deprived conditions (Li et al., 2013). The co-incidence of the *ESR1* fusion gene found in our lab with an identical breakpoint at AA365 as the *ESR1-YAP1* and the shared history of fulvestrant unresponsiveness led us to screen for further *ESR1* fusions in our cohort of exome-capture RNA sequencing of 45 patient-matched primary-metastasis paired breast cancer samples, where metastatic samples originated from one of 4 different recurrence sites (21 brain, 11 bone, 10 ovary and 3 gastrointestinal) (Priedigkeit et al., 2017) (Vareslija et al., 2018).

Using two different fusion calling algorithms (FusionCatcher and STAR-Fusion) (Haas et al., 2017; Nicorici et al., 2014), I detected fusion transcripts from the raw sequences. With FusionCatcher, a total of nine *ESR1* fusions were detected – two of them having *ESR1* as 3'-partner

and seven of them as 5'-partner. With STAR-Fusion, a total of 14 fusions were called, in six of which the *ESR1* was as 5'- and in eight of them as 3'-partner (**Figure 5**). Both algorithms shared only 2 identical *ESR1* fusions, the first being *ESR1-DAB2* that had been previously found by Dr. Hartmaier, and the second is an *ESR1-GYG1* fusion that I detected (Hartmaier et al., 2018). Notably, these two fusions were also the only *ESR1* fusions detected in all samples that had a breakpoint at exon 6. The fusion with *GYG1* was found in a bone metastasis and its expression both as a transcript and protein was confirmed. However, it is unknown if this fusion is metastasis-specific similarly as the other described *ESR1* fusions (E6), since no RNA-seq from the primary tumor was available.

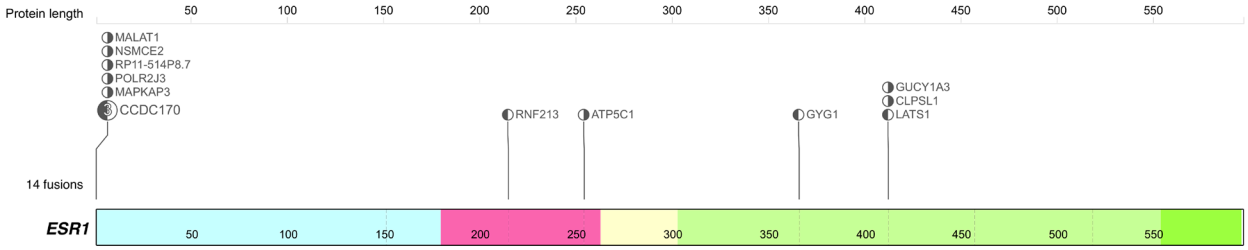

Figure 5 Breakpoint and fusion site of *ESR1* fusions detected in our samples

The structure of *ESR1* gene is depicted with its domains: Transactivation domain (AF1) (teal), DNA binding site (pink), Hinge region (yellow) and Transactivation domain (AF2) with Ligand-binding domain (green). The dashed lines show the exon boundaries, starting with exon 3 (= the first exon in the coding sequence) up to the last exon 10. The numbers stand for the amino acid position. ◐ *ESR1* as 5'-partner; ◑ *ESR1* as 3'-partner

To assess the oncogenic potential of the identified fusions and prioritize them by their predicted tumor driver score, I initially employed the two available algorithms Oncofuse and Pegasus in the readily available fusion calls of rapid autopsy samples by ChimeraScan. Both *in silico* prediction tools yielded for the *ESR1* fusion with *DAB2* a significant driver score of higher than 0.99 (1 being maximum). After verifying that the results produced by both tools were similar,

the outputs generated by FusionCatcher and STAR-Fusion were annotated with Oncofuse only to avoid any errors that would occur while transforming the data into a Pegasus-compatible input file.

The *ESR1-GYG1* fusion was the only *ESR1* fusion in our dataset that revealed a breakpoint at exon 6 with a significant driver probability of 0.9993 (**Appendix A, Table 4**). Oncofuse predicted a driver score of greater than 0.986 for three other *ESR1* fusions (*ESR1-LATS1, ATP5C1-ESR1, NSMCE2-ESR1*). Since these fusions are structurally different (*ESR1(E7)-LATS1* is an out-of-frame fusion) from the *ESR1* fusions (E6), we did not further investigate these.

Since *ESR1* fusions are unique to metastatic sites, I searched for further *ESR1* fusions in MET500 cohort, which is a public dataset of RNA-seq data from 91 patients with metastatic breast cancer. A total 46 *ESR1* fusions were detected in 12 patient samples (**Appendix A, Table 5**). As shown in **Figure 6**, consistent with prior observations, *ESR1* fusions clustered at 5' promoter site and the end of exon 6. At the 5'-promoter site, 19 *ESR1* fusions are detected. In 14 of these fusions, the total *ESR1* coding sequence is retained as the 3'-fusion partner. This may positively or negatively affect the expression of ER. The driver probability predicted for most of these type of *ESR1* fusions is significantly high with a value greater than 0.99 (**Appendix A, Table 4-5**). In the remaining 5 fusions, *ESR1* endows its promoter to the 3'-fusion partner. In summary, 4 further *ESR1* fusions with identical breakpoints at exon 6 (*ESR1-ARNT2, ESR1- RIMS2, ESR1-ARID1B* and *ESR1-ARMT1*) were found, which also revealed a driver score of above 0.99. Importantly, while most *ESR1* fusions at E6 reveal distinct 3'-fusion gene partners, *ESR1-ARMT1* (*ARMT1* is also known as *C6orf211*) is the first example, which reoccurs and demonstrates identical breakpoint for both genes (Giltnane et al., 2017). This is probably attributed to the proximal localization of *ARMT1* to *ESR1* (similarly as with *CCDC170*) and due to its tight co-expression with *ESR1 (Perry et al., 2015)*.

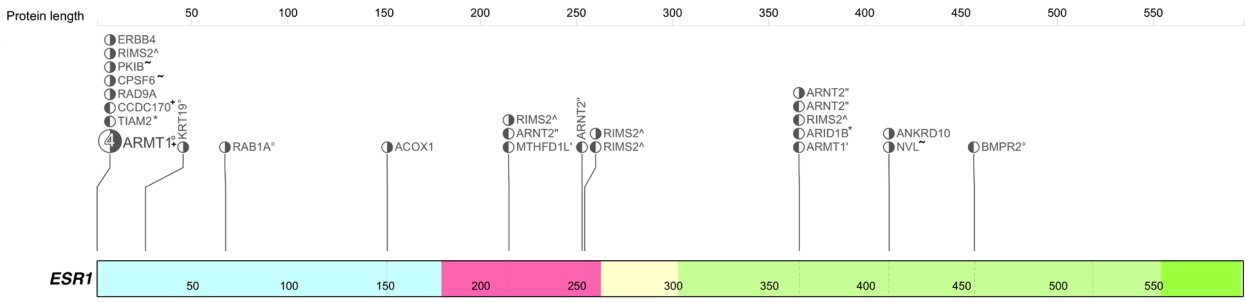

Figure 6 Breakpoint and fusion site of *ESR1* fusions detected in MET500 samples

The junction sites for *ESR1* fusions detected in MET500 dataset cluster in 5'-UTR and at exon 6 consistent with other datasets. Multiple *ESR1* fusions are predicted within a single sample, which are marked with a distinct symbol. Details including genomic breakpoint and driver probability are listed in **Appendix A, Table 5**. ◐ *ESR1* as 5'-partner; ◑ *ESR1* as 3'-partner

3.2 Ligand-independent activity and treatment-resistance of *ESR1* fusions

To understand the transcriptional effect of *ESR1* fusions on ER-regulated gene expression, and the impact of different ER-targeting compounds (E2, tamoxifen or fulvestrant (ICI182-780)) on their activity, ERE-Dual Luciferase assays were performed first in human embryonic kidney 293T cells, which show high transfection efficiency, and then in commonly used ER-positive and –negative breast cancer cell lines. Cells were transiently co-transfected with either one of the *ESR1* fusions (*ESR1-DAB2*, *ESR1-GYG1*, *ESR1-SOX9* and *ESR1-LPP*) or one of the controls (truncated *ESR1* (E6), wt-*ESR1* or empty vector), and the reporter plasmid *Firefly* luciferase that is under an estrogen response element (ERE) promoter along with the constitutively expressed transfection control vector *Renilla* luciferase.

In cells that lack ER expression (293T, MM231, C4-12), the basal signal is very low and can not be induced upon E2 treatment. The overexpression of wt-*ESR1* in those cells results in a pattern that we observe for endogenous ER in ER-positive cells, where there is a baseline activity that after addition of E2 induces transcriptional activation and returns to/below baseline upon

treatment with ER inhibitors. The activity induced by truncated *ESR1* (E6) is mostly comparable to the cell's baseline activity or in most cases even lower (considered inactive or potentially dominant-negative) and has lack of the above described pattern of induction and reversal upon various treatments. In general, all *ESR1* fusions show a ligand-independent activation and unresponsiveness to both treatments, consistent with our hypothesis.

In 293T and MM231 (**Figure 7A-B**), the two fusions *ESR1-SOX9* and *ESR1-LPP* show more than around 20x higher ER activity compared to basal ER activity level. The *ESR1-DAB2* fusion exhibited approximately 12-fold and 5-fold higher activity compared to the basal ER activity in 293T and MM231 cells, respectively. The *ESR1-GYG1* fusion has a 5-fold activity in 293T, but is inactive in MM231. In C4-12 cells (**Figure 7C**), a derivative of MCF7 cells, where the ER-expression is silenced, the activity level of all constructs are similar to the MM231 cells.

We further investigated the transcriptional effect of *ESR1* fusions in ER-dependent cells. ER-positive cell lines MCF7 and T47D (**Figure 7D-E**) have endogenous ER signaling that can be triggered upon addition of E2 and reversed on exposure to anti-endocrine treatment. The pattern of activity in cells overexpressing the wt-*ESR1* is similar to the endogenous ER activity although the signal level is greater for all four conditions. The overexpression of *ESR1* fusions in both cell lines again demonstrated hormone-independent activity that remained constant even after addition of E2. However, a slight decrease of activity was observed upon treatment with tamoxifen or fulvestrant that is likely due to the background signaling through endogenous ER present in cells. Even with the anti-endocrine therapies, ERE activity remained high. The *ESR1* fusion with *DAB2* and *LPP* has 4-5-fold ERE activity in both cell lines. The activity of the *ESR1-SOX9* fusion in MCF7 cells is around 4-fold higher than the endogenous, however in T47D, around 15-fold change is observed. The *ESR1-GYG1* fusion is inactive in both T47D cells and MCF7 cells.

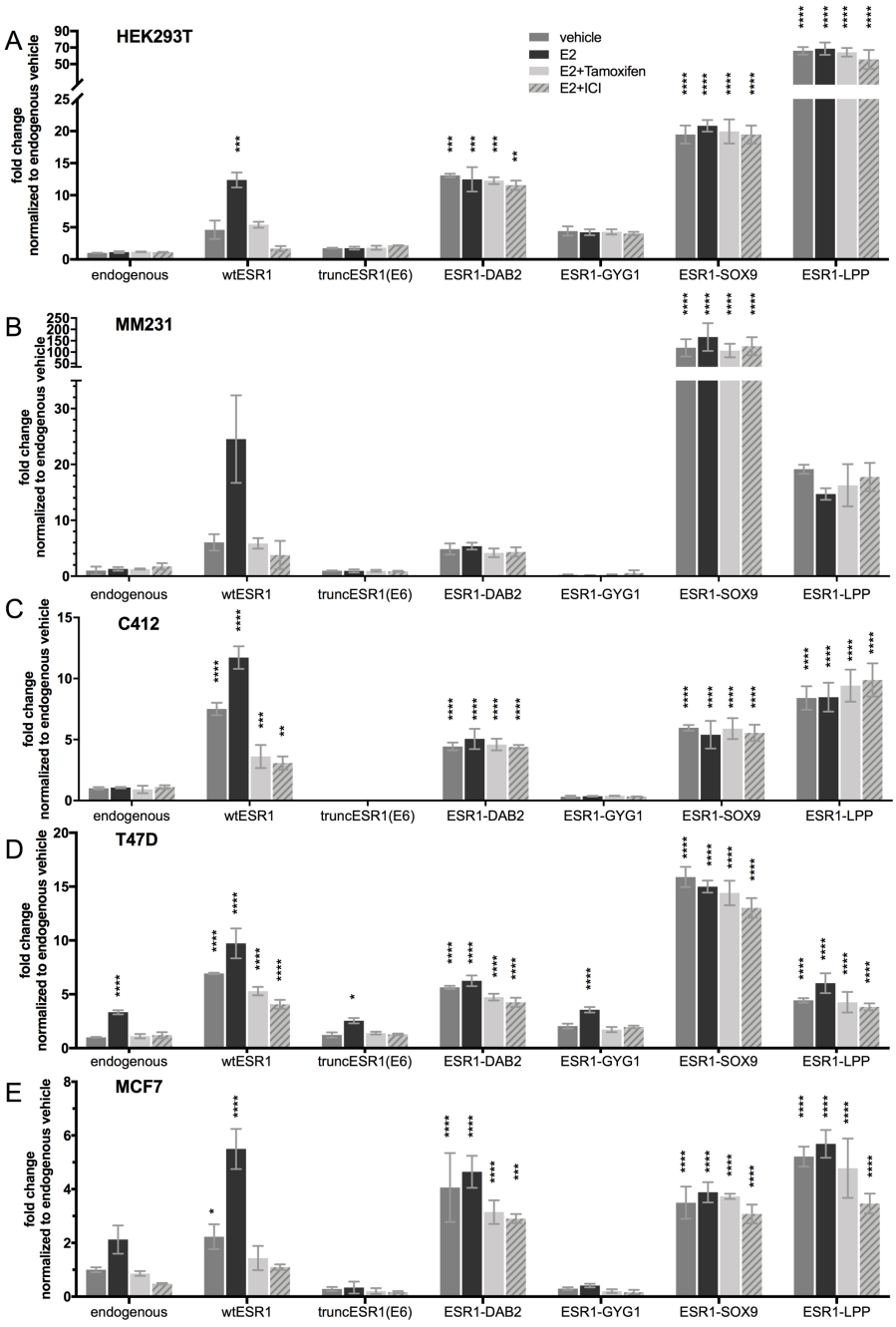

Figure 7 ERE-Tk-Luciferase assays reveal context-dependent ligand independence and endocrine resistance conferred by *ESR1* fusions

Cells were treated with 100 pM E2 or combined with 100 nM tamoxifen or 100 nM ICI. The bar graphs show the mean fold-change of activity compared to the untreated endogenous ER with the error bar representing 3 technical replicates. For independent repeats, n=2 for 293T, MM231 and C4-12, and n=3 for T47D and MCF7 cells.

In summary, our data suggest that truncated *ESR1* has a dominant-negative effect on ER, and that a fusion with a 3'-partner causes constitutive activity of the chimeric ER-protein. More importantly, the domains added by the 3'-fusion partner influences the activity and the function of the chimeric ER-protein. Furthermore, the results presented above also highlights the cell-dependent signaling behavior of each *ESR1* fusion, where the level of transcriptional activity can be more or less prominent.

3.3 *ESR1* fusions display predominantly nuclear localization

To investigate the subcellular localization of the chimeric fusion ER and the endogenous ER protein, cells were stained with fluorochrome-labeled antibodies targeting the N-terminal end of endogenous and chimeric ER as well as the HA-tag that is present in the overexpressed constructs only. Immunofluorescence staining was performed in 293T and the two ER-positive breast cancer cell lines T47D and MCF7 after a day of hormone-deprivation. The untransfected cells within the same slide served as an internal control to the cells in which the *ESR1* construct integrated successfully. The endogenous wt-ER protein was enriched in the nucleus of untransformed T47D and MCF7 cells, which could be visualized only in higher exposures, than used for transfected cells due to low endogenous signal. In 293T cells, which are known to be negative for ER expression, there was no signal for endogenous ER. Although the truncated and the chimeric ER proteins in all cell lines expressed well and were mainly detected in the nucleus the *ESR1* fusions with *GYG1* and *LPP* demonstrated in some cells both nuclear and cytoplasmic localization (**Figure 8; Appendix B Figure 13-15**).

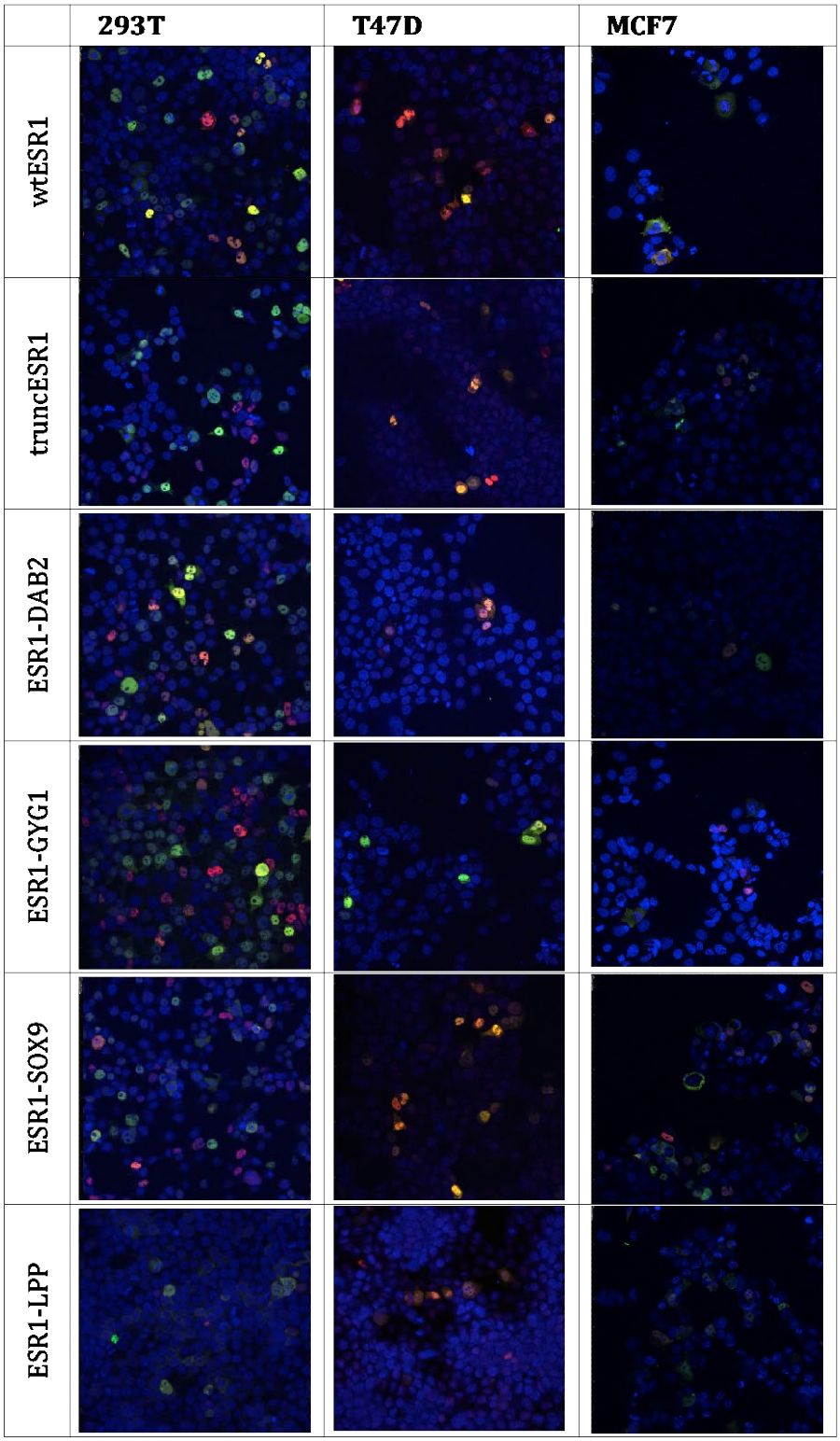

Figure 8 *ESR1* fusions display predominant nuclear localization in HEK293T, T47D and MCF7 cells following transient transfection of expression plasmids

Images were taken on a 40x objective following staining with DAPI (blue), *ESR1* (red) and HA (green). Representative composite images shown are generated by confocal microscopy.

3.4 Stable expression of *ESR1* fusions in T47D cells

To examine the effect of the *ESR1* fusions on the proliferation of breast cancer cells, we aimed to stably express them in T47D, MCF7 and in C4-12 cells (Oesterreich et al., 2001). Our initial effort to develop single clones of stably expressing cells by transfection and selection with constructs in the pcDNA3.1(+) backbone using Lipofectamine yielded very low efficiency. The majority of the single clones selected with neomycin did not produce any protein that was detectable by immunoblot. Remarkable was that none of the MCF7 cells expressing the truncated *ESR1* (E6) survived. Thus, we switched to a lentiviral infection method to induce T47D to stably express wt-*ESR1*, truncated *ESR1* (E6) and the *ESR1* fusions with *DAB2, SOX9* and *LPP*. Since the *ESR1-GYG1* fusion did not significantly induce transcriptional activity, it was excluded from generating stably expressing cells.

The expression of the *ESR1* constructs and fusions in T47D cells was confirmed by blotting with both the anti-ER and anti-HA antibodies (**Figure 9**). The cells stably expressing the *ESR1-DAB2* fusion yielded very low fusion protein levels, which was only seen in higher exposure times. One possible reason for the low protein expression is that the *ESR1-DAB2* fusion is the longest fusion consisting of 3285 base pairs (bp) and it is known that larger constructs are harder to express. After insertion into the lentiviral plasmid, the total vector size increased to 11546 bp and thus exceeded the size for optimal packaging and most probably produced low titers. However, because the cells with the *ESR1-DAB2* fusion survived antibiotic selection and express the *GFP* marker, we continued to perform experiments with them.

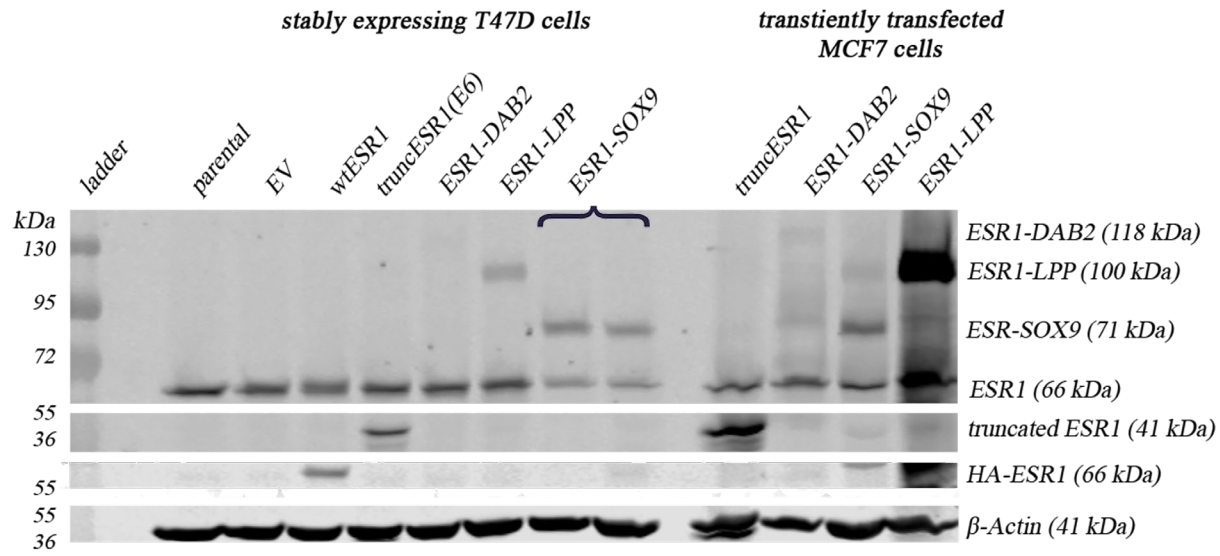

Figure 9 Immunoblot reveals successful expression of wt-*ESR1*, trunc*ESR1* and the *ESR1* fusions with *DAB2*, *SOX9* and *LPP* in T47D cells

Cell lysates of transiently transfected MCF7 cells served as a positive control. The membrane was first blotted with anti-ER antibody (1:2000) and post-stripping with anti-HA (1:1500) and anti-β-Actin (1:15000) antibodies.

3.5 Estrogen-independent proliferation of cells expressing *ESR1* fusions

To test our hypothesis that *ESR1* fusions (retaining AA 1-365) promote hormone-independent cell growth, T47D cells stably expressing *ESR1* fusions were seeded into hormone-deprived media. Additionally, the control cells - parental T47D, and EV and wt-*ESR1* stably expressing T47D cells – were exposed to either 1 nM E2 or vehicle treatment. The growth of cells were monitored for 14 days and the relative growth of cells comparing Day 7 to Day 0 and Day 14 to Day 0 represented in the graph below (**Figure 10**). As expected, the addition of E2 to control cells significantly promoted growth compared to untreated controls. In comparison to the untreated parental T47D, cells stably expressing wt-*ESR1* showed as expected a slightly greater relative growth in untreated condition and significantly higher relative growth upon E2 treatment. While

the *ESR1-SOX9* fusion promoted estrogen-independent growth, the *ESR1* fusions with *DAB2* and *LPP* and truncated *ESR1* were inactive and had similar growth pattern as the untreated parental T47D cells. Based on our transient ERE activity in T47D cells, we anticipated that both *SOX9* and *LPP* fusions may confer ligand-independent growth. Notably, both treated and untreated control cells stopped growing 7 days post-plating, but the *ESR1-SOX9* fusions kept proliferating. The proliferative properties of T47D cells with the *ESR1-SOX9* fusion under hormone-deprived condition was similar to the growth activity of parental T47D cells in E2-containing medium. Overall, the growth promoting activity of *SOX9* fusion is consistent with our transient ERE data. For the *DAB2* and *LPP* fusions, we did not observe the anticipated growth phenotype, however this could be due to poor fusion expression, especially in the case of *DAB2*.

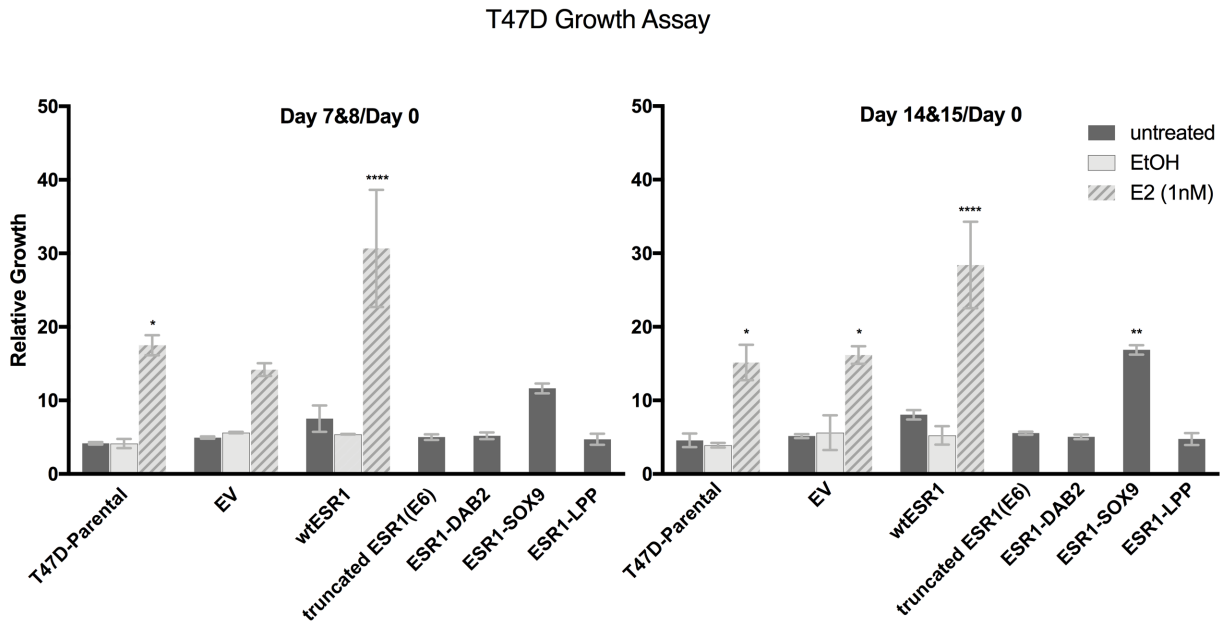

Figure 10 The *ESR1-SOX9* fusion leads to ligand independent cell proliferation

Cells were maintained in hormone-deprived medium for 14 days. For control cells, growth in 1nM E2 and vehicle treatment conditions were included. Bar graphs show the average of relative growth on Day 7 (left) and Day 14 (right) with the error bar representing the SEM of two independent experiments.

The above results examining proliferation were in concordance with observations via IncuCyte, where the morphology was captured in 8h intervals starting 24h post-plating. There were no significant morphological differences between the different stably expressing T47D cells. (**Figure 11**). Notable was that the areas of plate where originally >2 cells were attached to each other, became dense in shorter time (<7 days) compared to areas where only single cells were visible. Initially, most of the cells have a round shape that flattens as they grow and divide. However, if an area is highly proliferative, this monolayer appearance changes into multiple layers of smaller cells. Although all cells were prepared and dispensed the same way, the truncated *ESR1* and the *ESR1-DAB2* fusion expressing cells show clumped small cell colonies in the first day, which also tended to accumulate centrally in the plate. Yet, their overall growth was not significantly higher than the parental T47D cells in two biological replicates of proliferation assay with PrestoBlue. The *ESR1-SOX9* fusion has an increased proliferative property and tends to build very dense areas of cells.

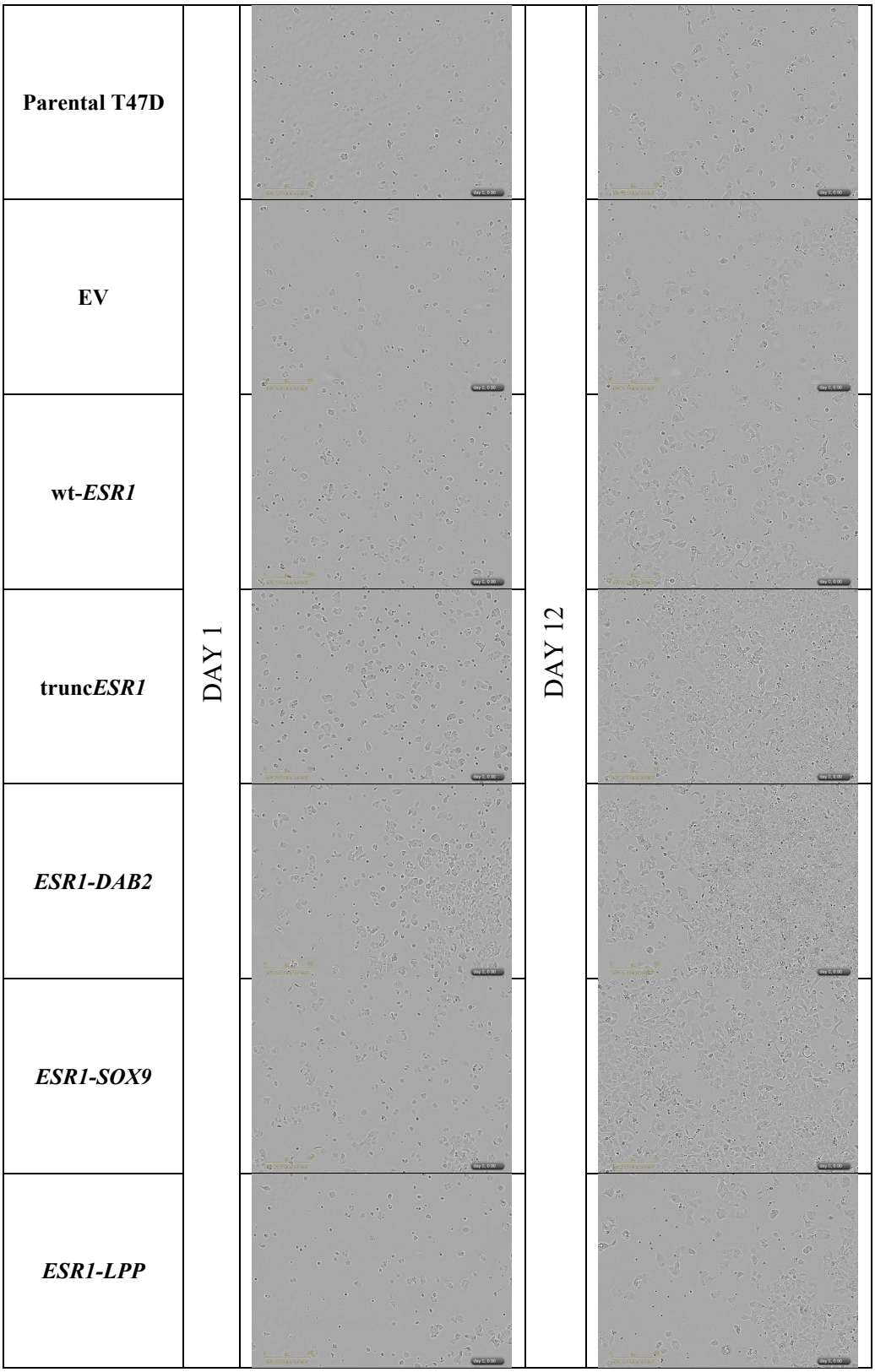

Figure 11 Morphology of stable T47D captured by IncuCyte on Day 1 and Day 12

The images are representative of the one quadrant of one replicate per stable expressing T47D cell. Images of whole single well are in **Appendix C**, **Figure 16-22**.

3.6 Estrogen-independent regulation of downstream genes by *ESR1* fusions

To examine the effect of the transcriptional properties of the *ESR1* fusions on messenger RNA (mRNA) expression of well characterized ER-regulated genes (*TFF1, GREB1* and *IGFBP4*), a qRT-PCR assay was performed in hormone-deprived stably expressing T47D cells that were cultivated 24h in the absence or presence of 1nM E2 and +/- 100nM ICI. The gene expression responses of downstream genes highlight that their stimulation and the level of their expression is dependent on the 3'-partner of the *ESR1* fusion.

In parental T47D cells, all downstream genes were stimulated when ER was activated by E2 and suppressed with ICI treatment, more so with ICI only. The expression observed with EV-inserted T47D cells were identical to the parental T47D. While the expression of *GREB1* and *IGFBP4* is only slightly more induced by cells overexpressing the wt-*ESR1*, the expression of *TFF1* **(Figure 12A**) is induced approximately 4-fold relative to the untreated state and increases even more with E2-stimulation. Interestingly, the truncated *ESR1*(E6) stimulates the expression of ER-regulated genes in all conditions significantly higher compared to the baseline activity in parental T47D. The *ESR1-SOX9* fusion promotes around 3-fold higher expression of *GREB1* (**Figure 12B**) and *IGFBP4* (**Figure 12C**) and around 75-fold for *TFF1,* and is fulvestrant-resistant. The *ESR1-LPP* fusion slightly increased the expression of *TFF1,* but did not reach statistical significance, and seems to have no stimulatory effect on the other ER-responsive genes. The *ESR1-DAB2* fusion shows constitutive transcriptionally active on all ER-regulated genes. However, it needs to be taken into account, that this experiment was performed only once and that the C_T levels of *RPLP0*, which serves as a reference gene for all values to be normalized to, for *ESR1-DAB2* was considerably higher (average C_T ~22) where all the other *ESR1* construct containing cells had an average C_T ranging from 16 to 18. Due to this reason, another statistical analysis was performed

where the *ESR1-DAB2* values were omitted from analysis, which consequently unmasked the statistical significance of other stably expressing cells (**Figure 12**).

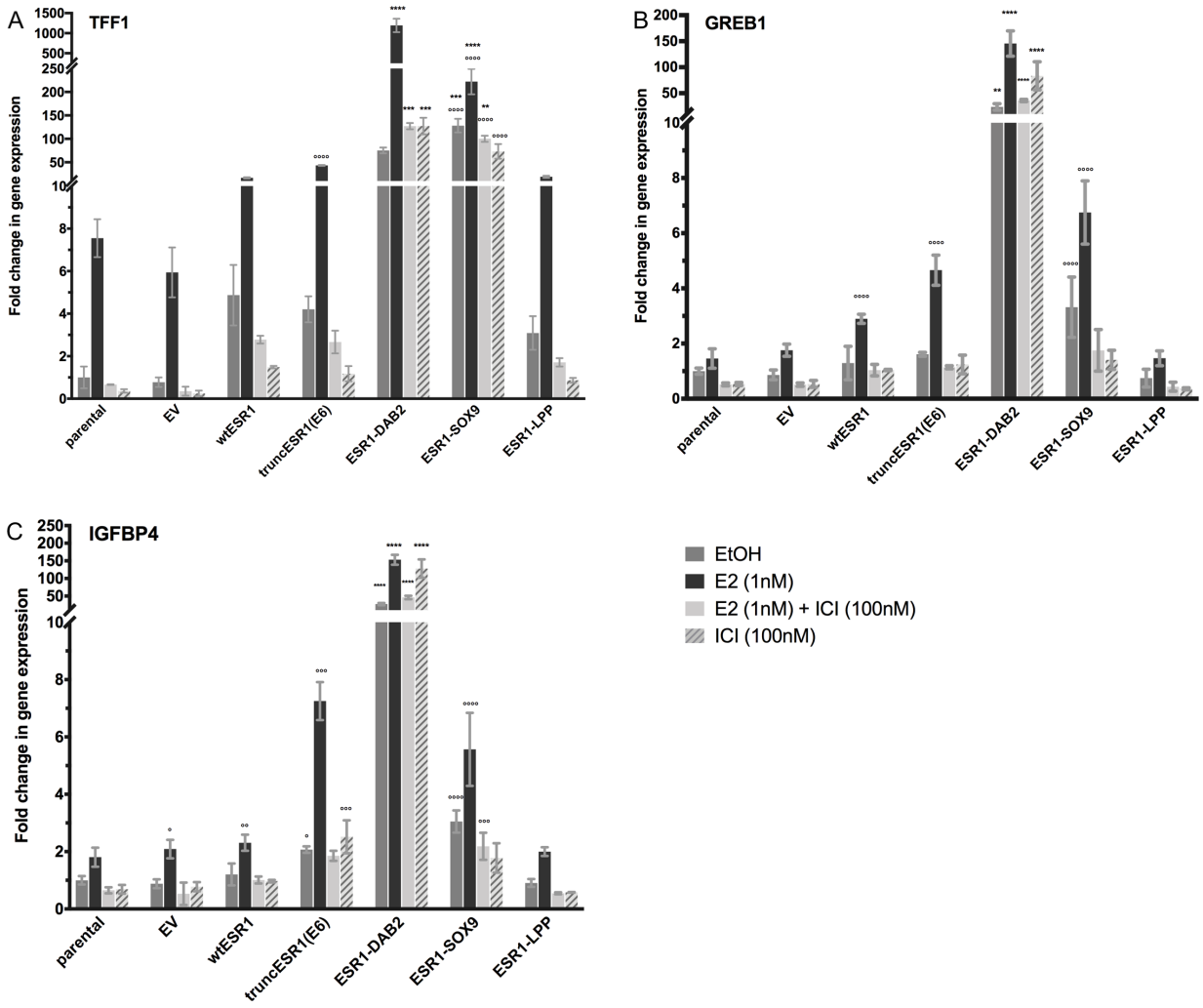

Figure 12 Expression of ER-regulated genes (*TFF1, GREB1* and *IGFBP4*) in T47D cells stably expressing the *ESR1* fusions

Bar graphs represent the mean fold change in gene expression normalized to the untreated parental T47D cells, for which the p-values are marked with asterisk(*). Data are from one experiment with the error bar representing the SD of 3 replicates. The ° marked p-values represent the statistical significance when data on *ESR1-DAB2* fusion is omitted from analysis.

In summary, the effect of *ESR1* fusions is variable on the downstream ER-regulated genes. While *ESR1* fusions with *DAB2* and *SOX9* stimulate the expression of all tested ER-responsive

genes, *ESR1-LPP* fusion only stimulates the transcription of *TFF1*. The stimulatory effect of *ESR1* fusions is estrogen-independent, which indicates the gain-of-function effect of *ESR1* fusions (E6). The transcriptional effect is maintained to a great extent even after degradation of wt-ER with ICI treatment, which further demonstrates that a dimerization of *ESR1* fusion with the wt-ER is probably not necessary for its activity.

4.0 Discussion and Conclusions

Metastatic ER-positive breast cancer is an incurable disease that remains a clinical challenge and a worldwide public health burden. The emergence of endocrine treatment resistance is a key factor in the lethality of breast cancer. Prior studies have demonstrated a now well accepted role for ER mutations in endocrine therapy resistance and the discovery of recurrent metastasis-specific *ESR1* fusions involving the loss of the LBD by our laboratory and collaborators suggests that *ESR1* fusions may represent an additional mechanism of therapy resistance in advanced ER-positive breast tumors (Hartmaier et al., 2018). Preliminary functional analysis of *ESR1* fusions in our laboratory has revealed that such *ESR1* fusions are not only unresponsive to anti-endocrine treatment, but also exert ligand-independent constitutive ER activity. Herein, I presented further *ESR1* fusion characterization that supports these findings and demonstrates the effect of such fusions on cell signaling and cell proliferation.

Owing to the promiscuous feature of partnering with various genes at its 3'-end, the *ESR1* fusions (E6) represent a heterogeneous group of fusions. However, these chimeric ERs share the loss of the LBD which yields them unresponsive to both its ligands (e.g. estrogens) and to anti-estrogen drugs as the target domain for drug interaction no longer exists. Similar observations have previously been reported for ER mRNA splicing variants lacking one or more of the exons encoding the LBD (E6-E8), which were found in normal as well as malignant breast and endometrial tissues (Holst et al., 2016; V. Kumar et al., 1987). Importantly, while truncated ERs have been consistently shown to be inactive in ERE assays by us and others (V. Kumar et al., 1987; Lei, Shao, et al., 2018), the gain of additional domains from the 3' fusion partner render the chimeric ERs ligand-independent. Our data and others' suggest that while some fusions are

hyperactive as determined by ERE assay (*ESR1* fusions with *DAB2, SOX9, LPP, PCDH11X* and *YAP1*), others demonstrate ligand independence but not hyperactivity (*ESR1* fusions with *GYG1* and *NOP2*) (Lei, Shao, et al., 2018). Employment of ERE assays in different breast cancer cell models has further shown that the activity level of the chimeric ERs are not only determined by the 3'-partner but are also cell line context-dependent (**Figure 7**). This could be attributed partially to the potentially different transfection rate of each construct in the transient setting as well as to differences in the molecular features of each cell line. The variability in the transfection efficiency after transient transfection and thus, the protein expression is reflected in the immunoblot (for MCF7, **Figure 9**) and immunofluorescence assay (**Figure 8**). It is important to note however that, high protein expression does not always mean higher activity and so the effects of the fusions are truly cell line and fusion partner dependent, and not only due to the expression levels.

Furthermore, I have provided provisional evidence that the chimeric ERs are predominantly localized in the nucleus, as is wild-type ER. However, fusions with *GYG1* and *LPP* demonstrate some cytosolic localization which is not evident in the other fusions or endogenous ER **(Figure 8**). *LPP* is a gene that encodes for the lipoma-preferred partner, which belongs to the zyxin family and harbors three LIM-domains at its C-terminal end that are retained in the fusion with *ESR1*. *LPP* is found transiently in nucleus as a transcription factor, but is mainly localized in the cytoplasm and is involved in the cell-cell adhesion and cell motility by interacting with α-actinin and VASP (vasomotor stimulated phosphoprotein) (Petit et al., 2005). Since *LPP* adds its nuclear export sequence (NES) located at AA 117-128 to the fusion (Petit et al., 2000), it is perhaps not surprising that the *ESR1-LPP* fusion additionally localizes to the cytoplasm. Hence, it can be concluded that depending on the 3'-partner and the domains afforded to the chimeric ER, the *ESR1* fusion can inherit new features that can cause neomorphic effects.

To understand the functional consequences of *ESR1* fusions, I examined the effect of fusions on proliferation in stably expressing T47D cells. The *ESR1-SOX9* fusion was the only *ESR1* fusion that conferred a statistically significant growth advantage to the cells in E2-deprived conditions and which showed a similar growth pattern as the parental T47D cells when stimulated by E2. Hence, the *ESR1-SOX9* fusion allows tumor cells to adapt to the hormone-depleted environment, conferring a ligand-independent cell proliferative phenotype. It should be noted however, that not all of the fusions were expressed as equal levels in the pools of overexpression cells, as indicated by immunoblot (**Figure 9**). Interestingly, *SOX9* (SRY-Box 9), which is a transcription factor critical for sex and skeletal development, also determines the stem cell fate in luminal progenitor cells in breast tissue. Moreover, the expression of *SOX9* is negatively regulated by ER upon competitively binding to the miR-190 against *ZEB1* with a subsequent reduction in the Wnt/β-catenin signaling (Domenici et al., 2019; Yu et al., 2019). Thus as expected, *SOX9* is highly expressed mainly in ER-negative and more aggressive triple negative breast tumor cells, rather than ER positive breast cancer (Domenici et al., 2019). Interestingly, the up-regulation of *SOX9* was additionally found to be associated with tamoxifen-/endocrine-resistant breast cancer and possessed metastatic properties, which could be reversed upon its down-regulation (Jeselsohn, Cornwell, et al., 2017; Yu et al., 2019). It will be interesting to study whether the *SOX9* fusion results in lower wild type *SOX9* in the cells baring the fusion, and how the *SOX9* part of the chimeric protein contributes to the cellular phenotype. In the *ESR1-SOX9* fusion, the PQS (proline, glutamine, serine)-rich domain and PQA (proline, glutamine, alanine) domain (residues 386-509 and 339-379), which are required for maximal transcriptional activation of *SOX9* are retained (McDowall et al., 1999). A main limitation of our studies is that it is unknown if the rearrangements lead to mono-allellic or bi-allellic disruption of the genes forming the fusion in single tumor cells.

Our cell models use overexpression of the fusion and it is unclear how this relates to the clinical situation in terms of expression level and allelic disruption. In my experiments, I used fulvestrant to deplete wt-ER from the cells so that we could more clearly study the effects of the fusion proteins. More comprehensive experiments with knocking down or downregulating the expression of *ESR1* and the fusion partner using shRNAs can be performed. Additionally, as mentioned, the fusions were overexpressed since the overexpression model is broadly used as an initial screening tool to roughly exploit the cellular and mechanistic effects of proteins. A controlled expression or expression under the endogenous promoter by generating the fusion via CRISPR would be superior methods to reflect a basal expression pattern and impact of the fusions.

As anticipated from the proliferation assays, there was heterogeneity in the transcriptional activation of ER-responsive downstream genes from each of the fusions. While *ESR1-SOX9* and *ESR1-DAB2* fusions induced transcription of the estrogen-responsive genes *TFF1, GREB1* and *IGFBP4*, *ESR1-LPP* stimulated the transcription of the *TFF1* gene only. This further proves that the 3'-fusion partner may add neomorphic and/or hypomorphic effects upon the chimeric ER. This is in line with the DNA-protein interaction studies, where active *ESR1* fusions gain estrogen-independent novel DNA-binding sites in contrast to loss of binding sites for inactive *ESR1* fusions (Lei, Shao, et al., 2018). A recent comparative examination of the known *ESR1* fusions revealed not only an independent replication of the results presented here for proliferation and signaling of *SOX9* and *DAB2* fusions, but also contrasted the action of the *SOX9* fusion to the *YAP1* fusion (Lei, Gou, Seker, et al., 2018). The tumor proliferative and transcriptional activation properties of the *SOX9* fusion is stronger than the *YAP1* fusion which was demonstrated to have comparable growth effects as the *ESR1* hotspot mutations (Y537S/N) (Li et al., 2013). Thus, *ESR1* fusions can be regarded as an important resistance mechanism to endocrine treatment in addition to the well-

known frequent hotspot mutations. Determining the true frequency of the *ESR1* fusions in endocrine resistance breast cancer will be an important step in fueling efforts to study these fusions and identify novel therapeutic agents targeting these fusions.

ESR1 fusions are unique to cancer cells and have so far been detected only in metastatic samples in therapy-refractory disease. Since 30% of patients diagnosed with early stage tumors will eventually experience recurrence and progression, patients may benefit from regular monitoring for *ESR1* related aberrations (mutations or fusions) via, for example, non-invasive liquid biopsy. Performed during both the disease-free period and during progression, early detection may indicate alternative treatment options before further disease progression. The use of CDK4/6 inhibitors which are approved as first-line treatment in advanced disease were also shown to effectively reverse the growth induced by *ESR1* fusions suggesting that patients with *ESR1* fusions may still respond to CDK4/6 inhibitors (Lei, Shao, et al., 2018; Walker et al., 2016). There is a lack of successful treatment options once patients develop therapy-refractory progressive disease. Although alternative compounds have been shown to successfully inhibit the association of ER with the ERE in *in vitro* assays, there are currently no approved drugs targeting the DNA-binding domain of ER or other domains (Mao et al., 2008; L. H. Wang et al., 2004). In addition to ongoing research seeking to further understand and target mutant ER, *ESR1* fusions and the resultant biology should be critically assessed for therapeutic vulnerabilities which could be exploited for patient benefit.

Overall, the comparative functional characterization of *ESR1* fusions published in the literature and presented here demonstrate that *ESR1* fusions, when active, have a critical role in developing endocrine resistance and in promoting tumor progression, comparable to the resistance observed with *ESR1* hotspot mutations. Since these fusions are observed to be metastasis-specific,

they may serve as biomarkers for diagnosis and for monitoring disease progression and most critically exploited to discover new drug targets. Although *ESR1* fusions involving the LBD are currently not as frequent in breast cancer as the hotspot mutations, they are also found in other gynecological tumors affecting the prognosis of patients. Therefore, a systematical analysis of the actual frequency of these fusions is warranted. Moreover, it is crucial to exploit more comprehensive functional characterization of the fusions including phenotypic assays (migration, invasion, colony formation), gene regulation and signaling (ChIP-seq, qPCR) and pharmacological assays in stably expressing or genome edited cell line models and in PDX-models harboring the fusions. Because the *ESR1* fusions have distinct partners and preferential breakpoints, characterization of these fusions may enable functional predictions based on the role and retained domains of the 3'-fusion partner. Furthermore, it is important to explore the feasibility to detect fusions in cfDNA, CTCs or exosomes via liquid biopsy and to assess if it provides prognostic value.

To summarize, my studies use transient and stable cell models to functionally characterize *ESR1* fusions. My data reveal that such fusions are ligand independent and may induce hyperactive signaling, the extent to which is dependent on the 3'-fusion partner. I further demonstrate for the *ESR1-SOX9* fusion that this elevated signaling results in ligand independent cellular proliferation and induction of ER target genes. Further studies are warranted to determine the true frequency of *ESR1* fusions in endocrine resistant breast cancer, and to understand their value as biomarkers of disease progression and as therapeutic targets.

Appendix A Details of *ESR1* fusions

Table 4 Junctionpoint and Onfocuse Driver Score of *ESR1* fusions from our dataset

Sample_ID	LeftGene	RightGene	LeftBreakpoint	RightBreakpoint	Spanning|Split_Reads	DriverProbability	FrameStatus	DetectionMethod
0013-56M	ESR1	CCDC170	chr6:151702005:+	chr6:151544572:+	3 | 3	0.0218255726059303	UTR/CDS(truncated)	Fusioncatcher; Starfusion
	ESR1	CCDC170	chr6:151702005:+	chr6:151536318:+	3 | 1	0.0218255726059303		Starfusion
19M	ESR1	LATS1	chr6:152011794:+	chr6:149662238:-	0 | 2	0.996943475069151		Starfusion
22M	ESR1	CCDC170	chr6:151702005:+	chr6:151585889:+	2 | 2	0.0218255726059303	UTR/CDS(truncated)	Fusioncatcher; Starfusion
	ESR1	CCDC170	chr6:151702005:+	chr6:151573174:+	1 | 1	0.0218255726059303		Starfusion
A26M	MAPKAPK3	ESR1	chr3:50614018:+	chr6:151807843:+	0 | 2			Starfusion
BM08	ESR1	GYG1	chr6:151944508:+	chr3:148994142:+	50 | 23	0.999317091901761	in-frame	Fusioncatcher; Starfusion
BM24	ATP5C1	ESR1	chr10:7802854:+	chr6:151944173:+	0 | 2	0.986824001207693		Starfusion
BM74	NSMCE2	ESR1	chr8:125182256:+	chr6:151807843:+	0 | 2	0.999981401476875		Starfusion
BP72	MALAT1	ESR1	chr11:65500520:+	chr6:151802184:+	6 | 2		exonic(no-known-CDS)/intronic	Fusioncatcher
ERLR_08_N1	GUCY1A3	ESR1	chr4:155667419:+	chr6:152060991:+	0 | 2	0.365260651542769		Starfusion
GM3A	CLPSL1	ESR1	chr6:35787120:+	chr6:152060991:+	0 | 2	0.114316674719834		Starfusion
GP7	POLR2J3	ESR1	chr7:102569835:-	chr6:151807843:+	0 | 3			Starfusion
	RP11-514P8.7	ESR1	chr7:102569835:-	chr6:151807843:+	0 | 3			Starfusion
OP8	ESR1	RNF213	chr6:151842787:+	chr17:80325030:+	0 | 2			Starfusion

Table 5 Junctionpoint and Onfocuse Driver Score of *ESR1* fusions from MET500 dataset

Sample_ID	LeftGene	RightGene	LeftBreakpoint	RightBreakpoint	Spanning|Split_Reads	DriverProbability	FrameStatus	DetectionMethod
SRR4306274	ARMT1	ESR1	chr6:151452586:+	chr6:151807843:+	0 | 2	0.9999893		Starfusion
SRR4306580	ACOX1	ESR1	chr17:75978534:-	chr6:151842597:+	0 | 2	0.9999990		Starfusion
SRR4306748	ESR1	ARMT1	chr6:151944508:+	chr6:151454548:+	1 | 1	0.9993171		Starfusion
	ESR1	ARMT1	chr6:151808364:+	chr6:151454548:+	1 | 1	0.0111461		Starfusion
	ESR1	MTHFD1L	chr6:151842787:+	chr6:151092467:+	1 | 1	0.0643566		Starfusion
SRR4306857	ANKRD10	ESR1	chr13:110906033:-	chr6:152060991:+	5 | 1	0.7215169	in-frame	Fusioncatcher
SRR4306905	ESR1	ARID1B	chr6:151944508:+	chr6:157189642:+	63 | 1	0.9998880	in-frame	Fusioncatcher; Starfusion
	ESR1	TIAM2	chr6:151702005:+	chr6:155211204:+	3 | 1	0.0389343	UTR/CDS(truncated)	Fusioncatcher; Starfusion
SRR4306946	ARMT1	ESR1	chr6:151452586:+	chr6:151807843:+	1 | 1	0.9999893	CDS(truncated)/UTR	Fusioncatcher; Starfusion
	ESR1	BMPR2	chr6:152061124:+	chr2:202530794:+	11 | 1	0.9998880	in-frame	Fusioncatcher; Starfusion
	ESR1	KRT19	chr6:151808082:+	chr17:41528121:-	10 | 1	0.6625967	out-of-frame	Fusioncatcher
	ESR1	KRT19	chr6:151808088:+	chr17:41528056:-	10 | 1	0.6625967	in-frame	Fusioncatcher
	KRT19	ESR1	chr17:41528045:-	chr6:151807987:+	10 | 1	0.9987066	in-frame	Fusioncatcher
	KRT19	ESR1	chr17:41528044:-	chr6:151807988:+	10 | 2	0.9987066	in-frame	Fusioncatcher
	RAB1A	ESR1	chr2:65129983:-	chr6:151808113:+	3 | 2	0.9999735	UTR/CDS(truncated)	Fusioncatcher
SRR4306950	CPSF6	ESR1	chr12:69262562:+	chr6:151807843:+	1 | 2	0.9999897		Starfusion
	NVL	ESR1	chr1:224330071:-	chr6:152060991:+	1 | 2	0.0336325		Starfusion
	PKIB	ESR1	chr6:122633367:+	chr6:151807843:+	1 | 23	0.9999375		Starfusion
SRR4307279	ESR1	RIMS2	chr6:151944508:+	chr8:103885298:+	182 | 2	0.9998880	out-of-frame	Fusioncatcher; Starfusion
	ESR1	RIMS2	chr6:151983560:+	chr8:103885298:+	182 | 6	0.9998880	UTR/CDS(truncated)	Fusioncatcher
	ESR1	RIMS2	chr6:151880771:+	chr8:103912053:+	182 | 3	0.9998880	out-of-frame	Fusioncatcher
	RIMS2	ESR1	chr8:103886223:+	chr6:151807843:+	1 | 1	0.9999997		Starfusion
	RIMS2	ESR1	chr8:103942926:+	chr6:151807843:+	1 | 1	0.9999997		Starfusion
	RIMS2	ESR1	chr8:103886223:+	chr6:151842597:+	1 | 1	0.9999997		Starfusion
	RIMS2	ESR1	chr8:103921784:+	chr6:151807843:+	1 | 1	0.9999997		Starfusion
	RIMS2	ESR1	chr8:103989421:+	chr6:151807843:+	1 | 2	0.9999997		Starfusion
	RIMS2	ESR1	chr8:103886223:+	chr6:151944173:+	182 | 58	0.9999933	in-frame	Fusioncatcher; Starfusion
	RIMS2	ESR1	chr8:103886219:+	chr6:151807841:+	182 | 4	0.9999997	CDS(truncated)/UTR	Fusioncatcher
	RIMS2	ESR1	chr8:103886223:+	chr6:151880655:+	182 | 1	0.9999933	in-frame	Fusioncatcher
SRR4307330	ERBB4	ESR1	chr2:212538449:-	chr6:151807843:+	0 | 4	0.9999866		Starfusion
SRR4307578	ARNT2	ESR1	chr15:80581404:+	chr6:152011656:+	610 | 1	0.9999893	in-frame	Fusioncatcher; Starfusion
	ARNT2	ESR1	chr15:80581404:+	chr6:152011657:+	610 | 1	0.9999893	out-of-frame	Fusioncatcher
	ARNT2	ESR1	chr15:80581401:+	chr6:152011656:+	610 | 30	0.9999893	in-frame	Fusioncatcher
	ARNT2	ESR1	chr15:80581401:+	chr6:152011657:+	610 | 29	0.9999893	out-of-frame	Fusioncatcher
	ESR1	ARNT2	chr6:151944508:+	chr15:80591568:+	610 | 29	0.9998880	in-frame	Fusioncatcher; Starfusion
	ESR1	ARNT2	chr6:151944506:+	chr15:80591568:+	610 | 2	0.9998880	out-of-frame	Fusioncatcher
	ESR1	ARNT2	chr6:151944508:+	chr15:80591569:+	610 | 133	0.9998880	out-of-frame	Fusioncatcher
	ESR1	ARNT2	chr6:151880768:+	chr15:80591568:+	610 | 126	0.9998880	in-frame	Fusioncatcher
	ESR1	ARNT2	chr6:151880771:+	chr15:80591568:+	610 | 125	0.9998880	in-frame	Fusioncatcher
	ESR1	ARNT2	chr6:151983560:+	chr15:80591568:+	610 | 10	0.9998880	UTR/CDS(truncated)	Fusioncatcher
	ESR1	ARNT2	chr6:151944506:+	chr15:80591569:+	610 | 9	0.9998880	in-frame	Fusioncatcher
	ESR1	ARNT2	chr6:151944506:+	chr15:80591567:+	610 | 2	0.9998880	CDS(truncated)/intronic	Fusioncatcher
	ESR1	ARNT2	chr6:151842787:+	chr15:80591568:+	610 | 2	0.0643566	in-frame	Fusioncatcher
SRR4307807	ARMT1	ESR1	chr6:151452586:+	chr6:151807843:+	1 | 1	0.9999893	CDS(truncated)/UTR	Fusioncatcher; Starfusion
	ESR1	CCDC170	chr6:151702005:+	chr6:151536318:+	0 | 1	0.0218256		Starfusion
SRR4307851	ARMT1	ESR1	chr6:151452586:+	chr6:151807843:+	0 | 3	0.9999893		Starfusion

Appendix B Subcellular localization of *ESR1* fusions

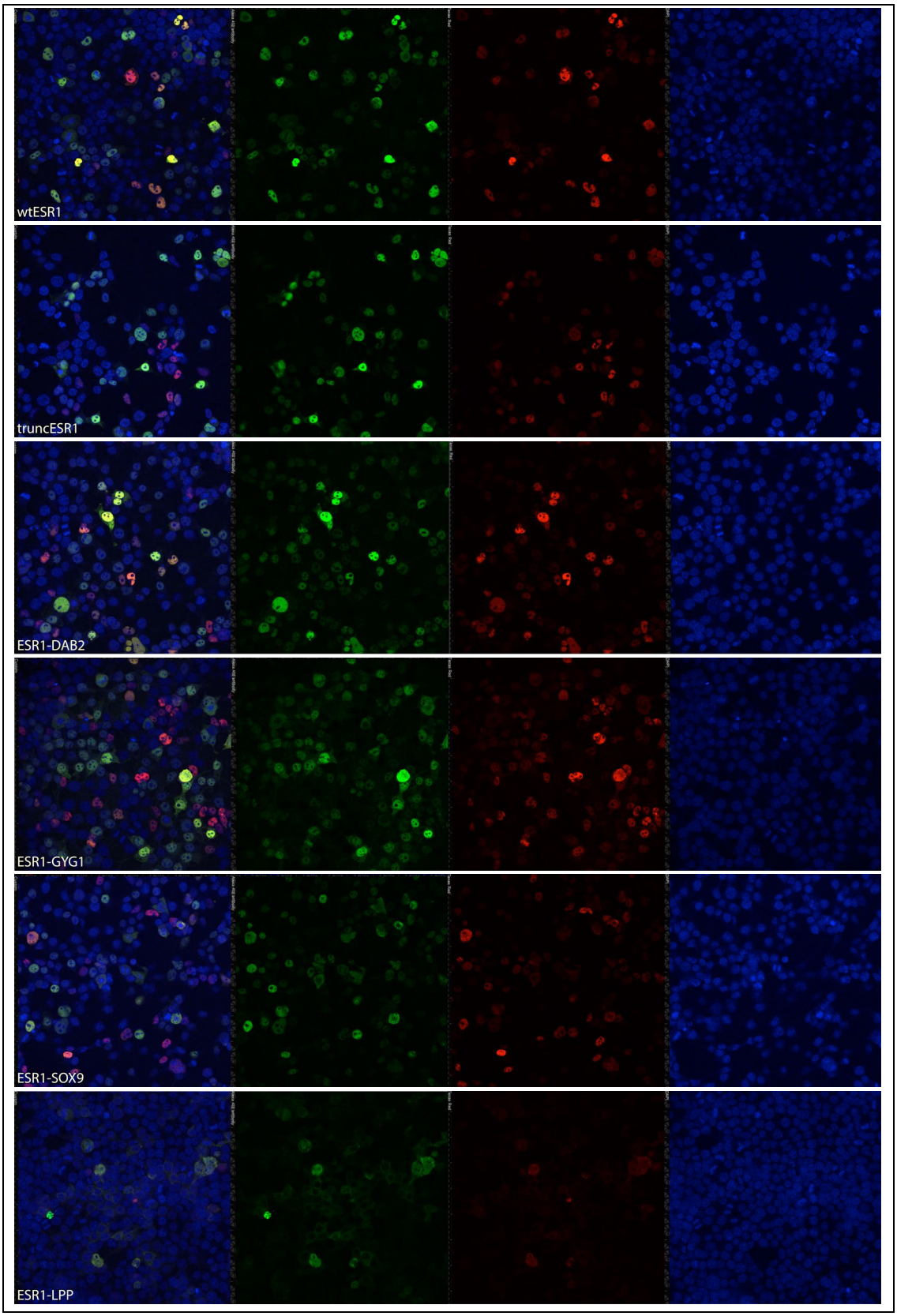

Figure 13 Subcellular localization of *ESR1* fusion in 293T cells via confocal microscopy with 40x objective
Composite and single channels: DAPI (blue), *ESR1* (red) and HA (green).

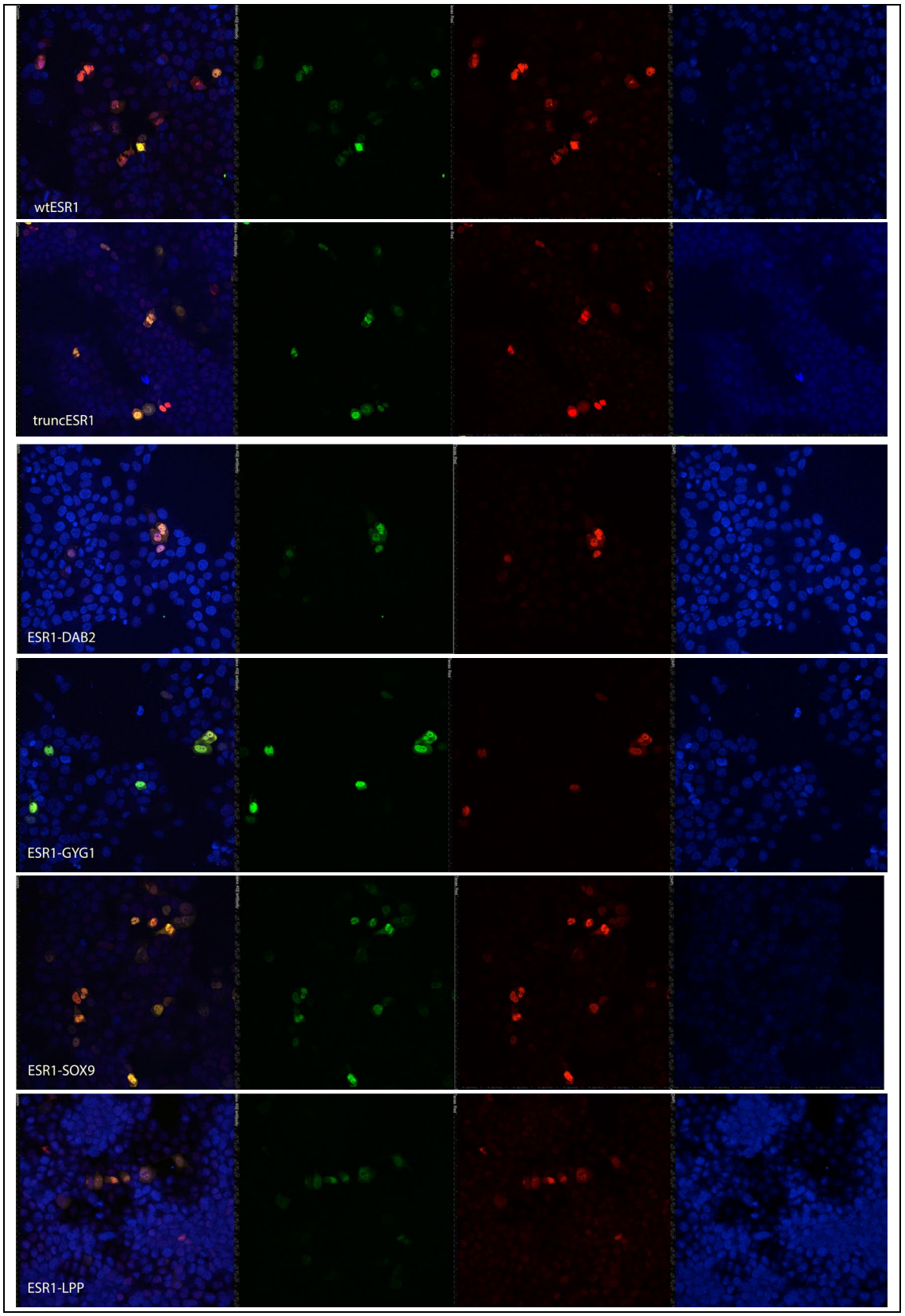

Figure 14 Subcellular localization of *ESR1* fusion in T47D cells via confocal microscopy with 40x objective
Composite and single channels: DAPI (blue), *ESR1* (red) and HA (green).

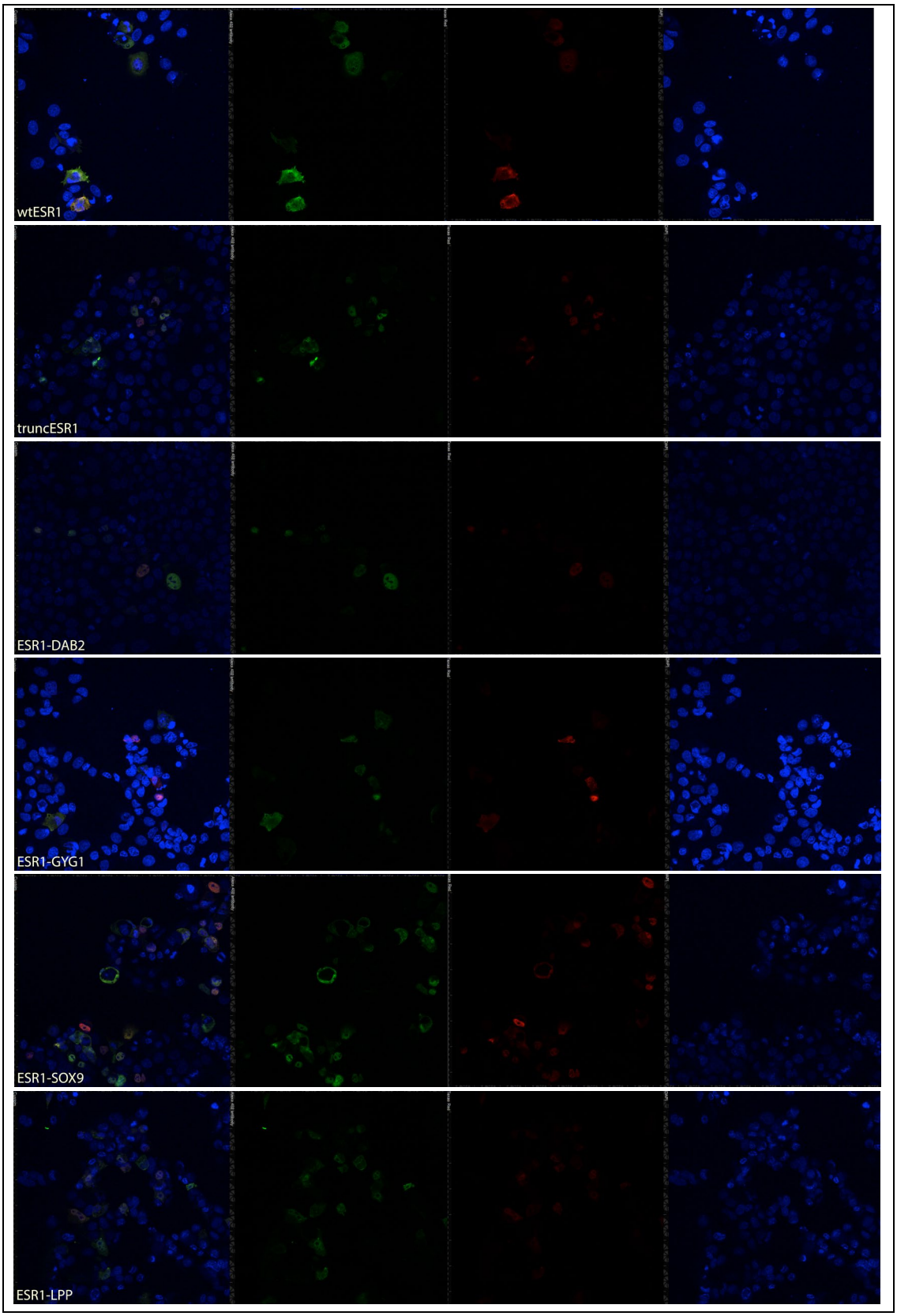

Figure 15 Subcellular localization of *ESR1* fusion in MCF7 cells via confocal microscopy with 40x objective
Composite and single channels: DAPI (blue), *ESR1* (red) and HA (green).

Appendix C Morphological assessment of stable T47D cells by IncuCyte

The morphology of cells were started to be monitored via IncuCyte 24h after plating. For each stable expressing cell had 6 replicates were plated. To visualize the total area of each well, images were captured dividing a well into 4 quadrants every 8h over the course of 12 days. The images below represent the total area of a single well from one replicate of each stable expressing T47D on Day 1 and Day 12.

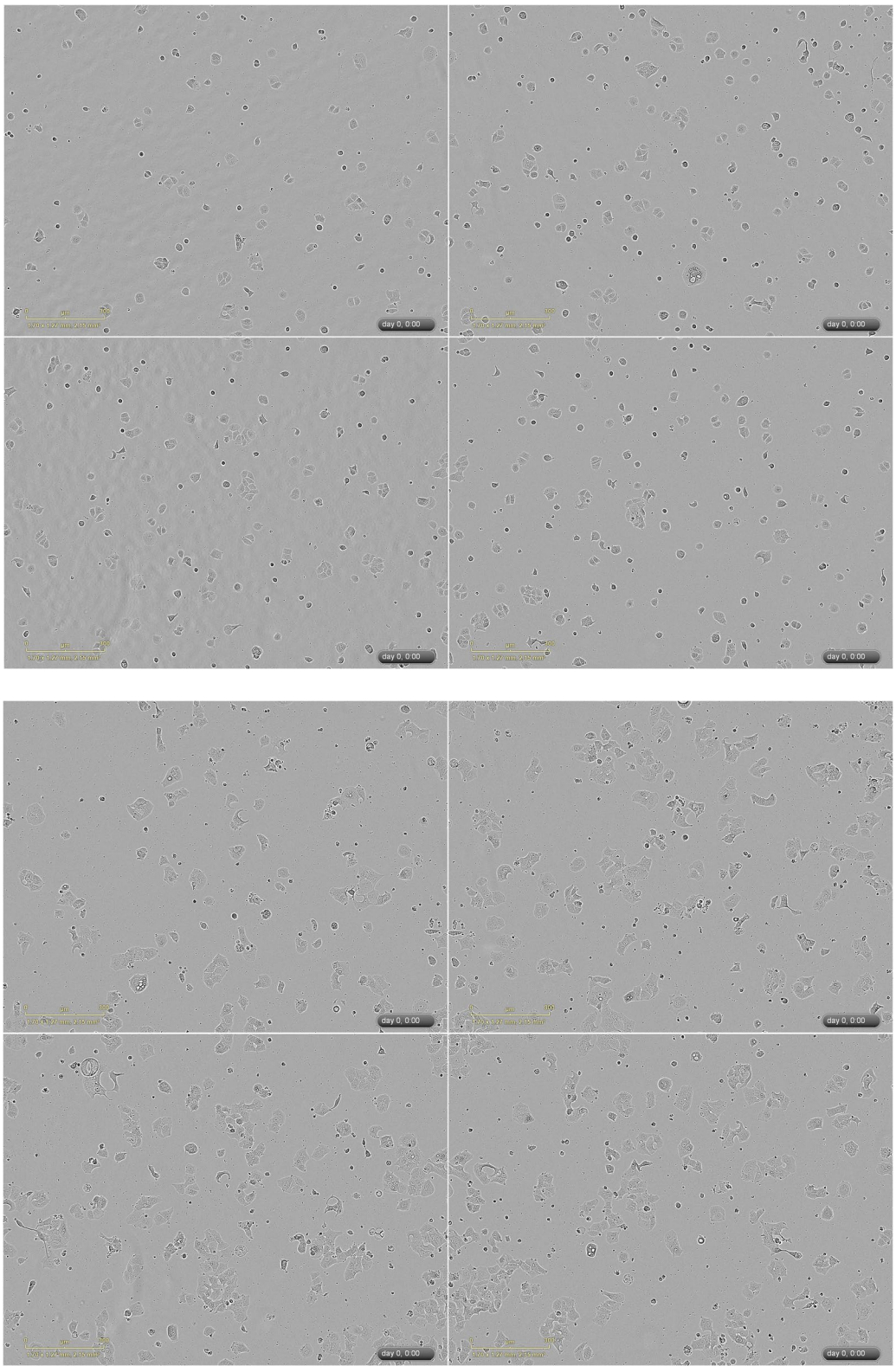

Figure 16 The cell viability of parental T47D cells on Day 1 (top) and Day 12 (bottom) in hormone-deprived medium

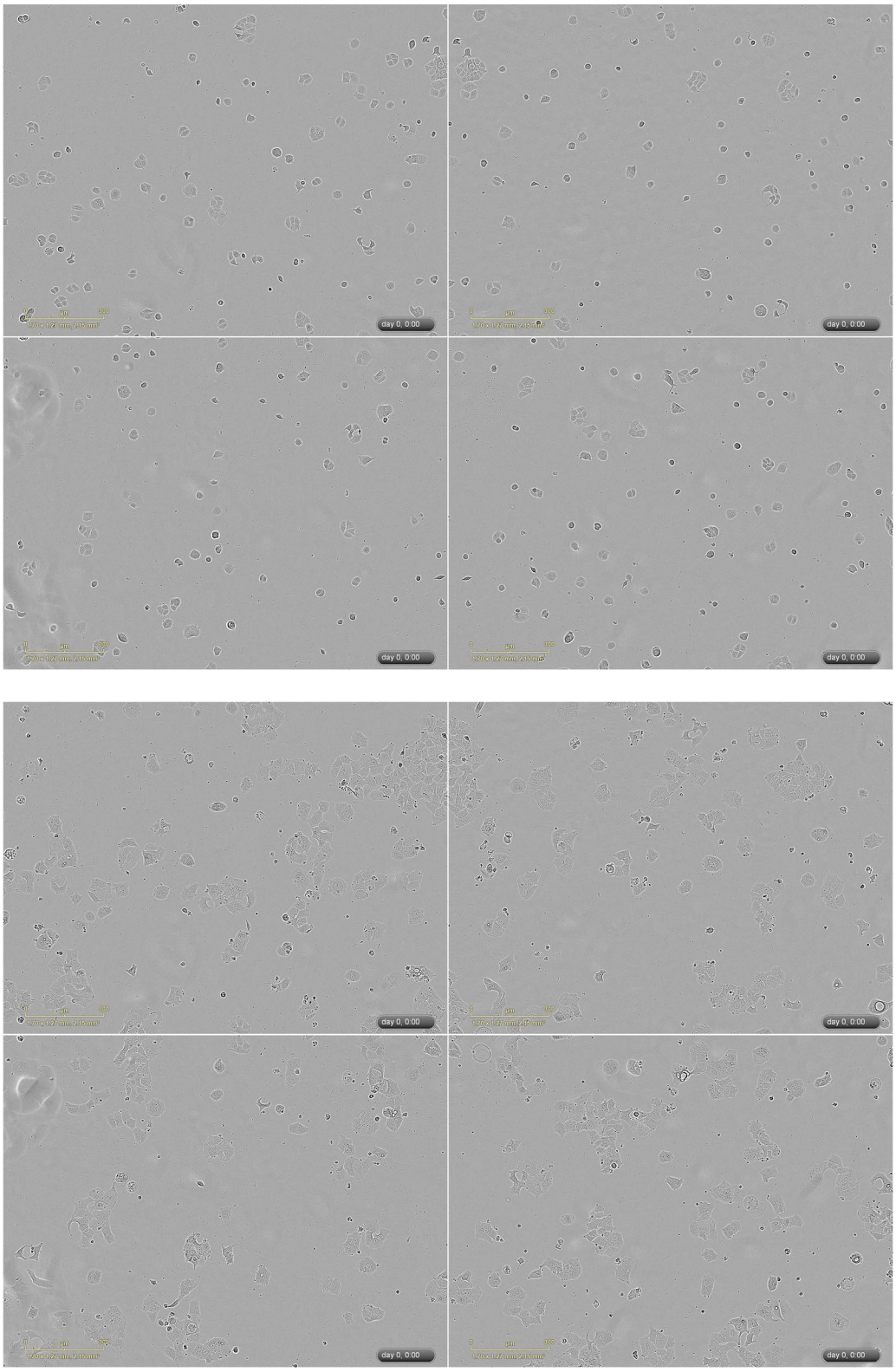

Figure 17 The cell viability of EV expressing T47D cells on Day 1 (top) and Day 12 (bottom) in hormone-deprived medium

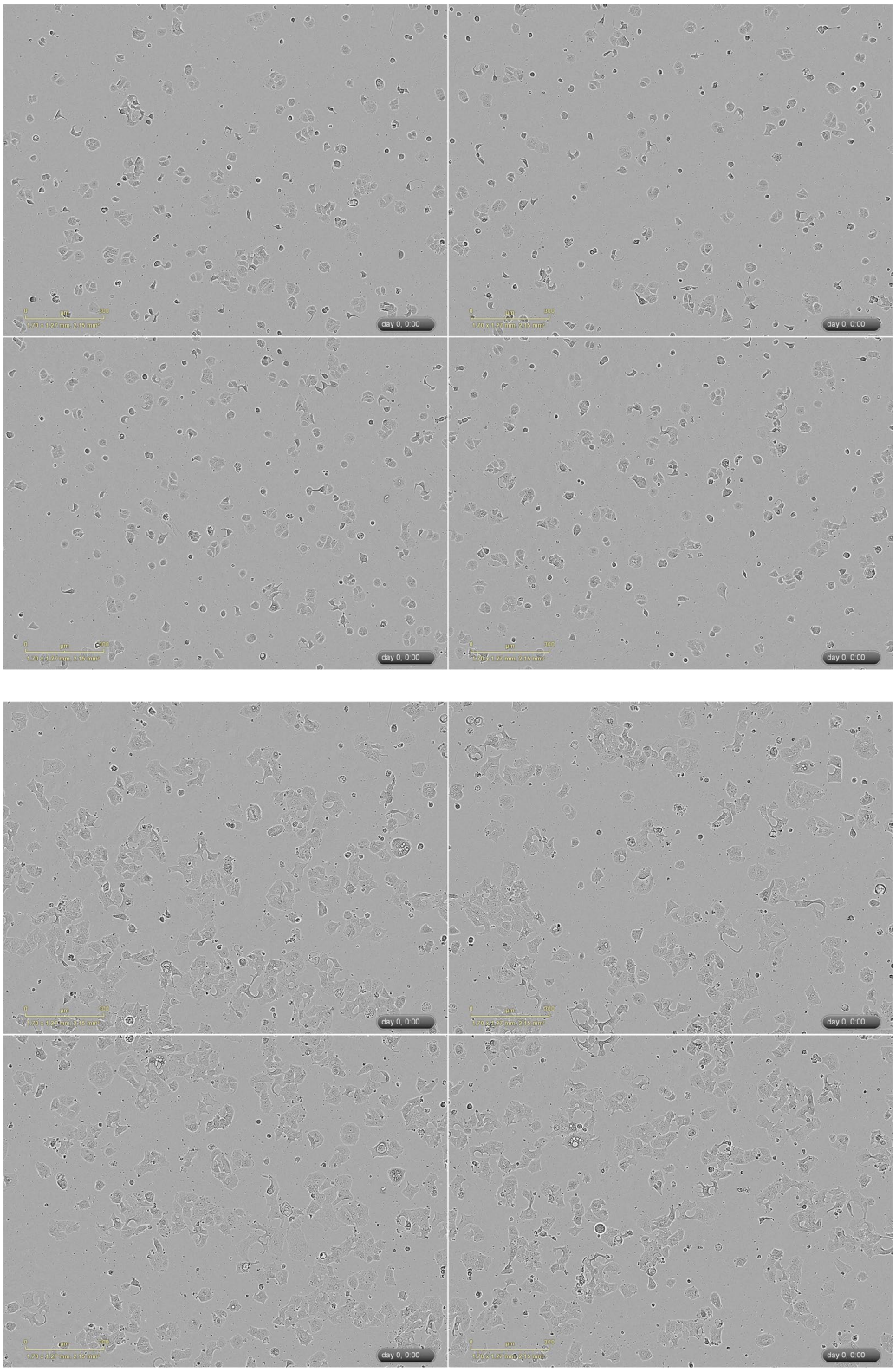

Figure 18 The cell viability of wt-*ESR1* expressing T47D cells on Day 1 (top) and Day 12 (bottom) in hormone-deprived medium

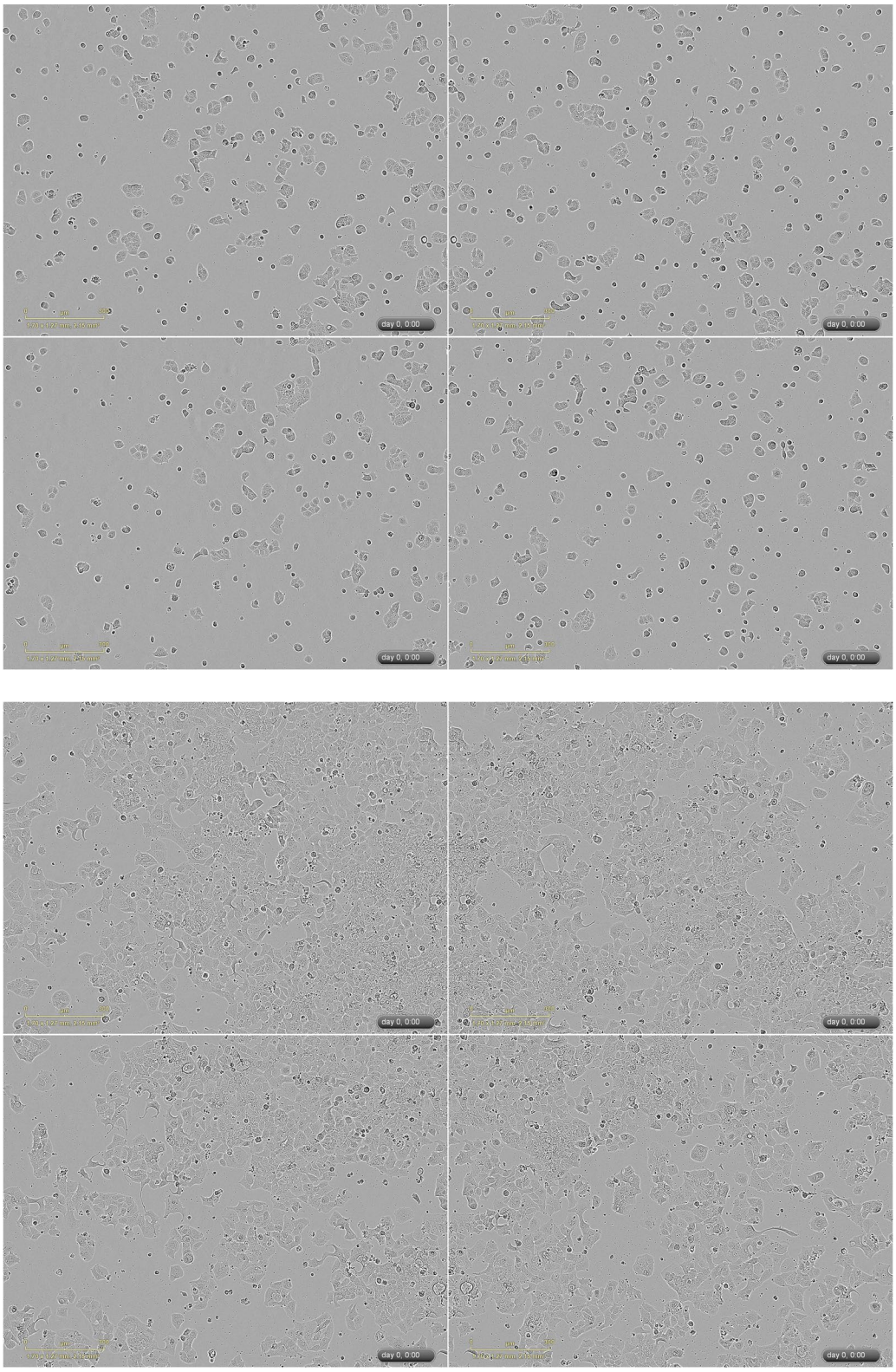

Figure 19 The cell viability of *truncESR1* expressing T47D cells on Day 1 (top) and Day 12 (bottom) in hormone-deprived medium

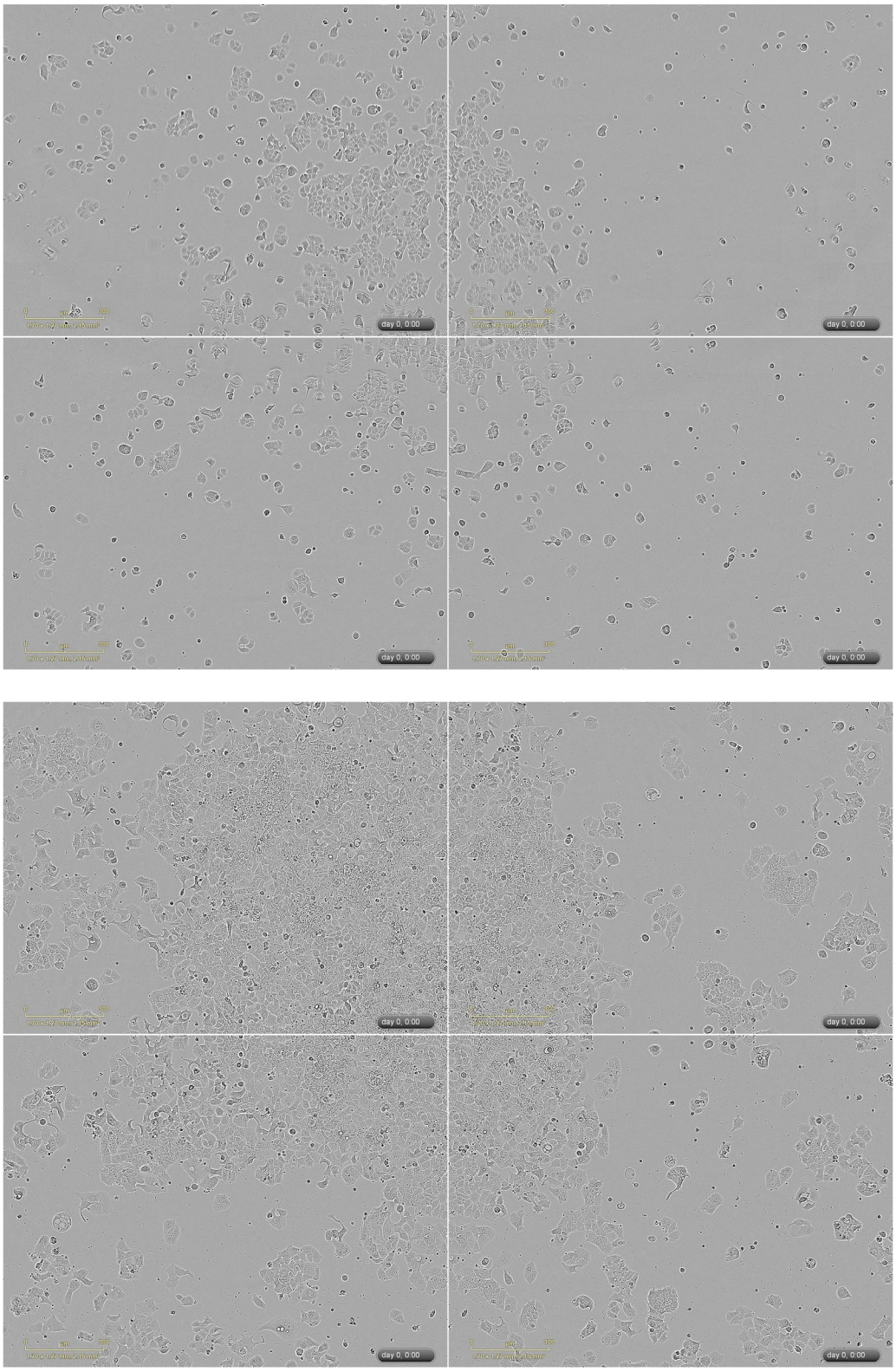

Figure 20 The cell viability of *ESR1-DAB2* expressing T47D cells on Day 1 (top) and Day 12 (bottom) in hormone-deprived medium

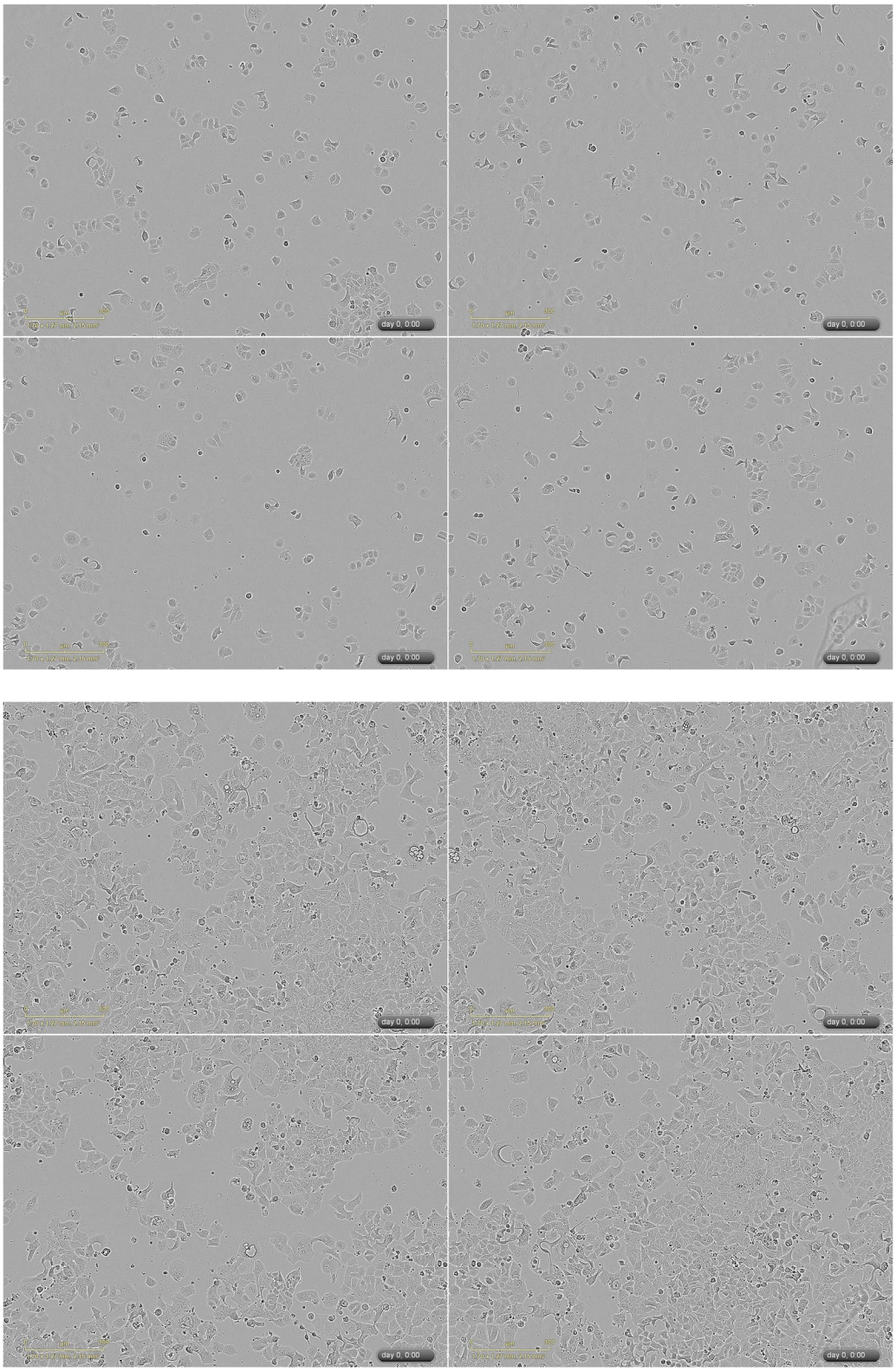

Figure 21 The cell viability of *ESR1-SOX9* expressing T47D cells on Day 1 (top) and Day 12 (bottom) in hormone-deprived medium

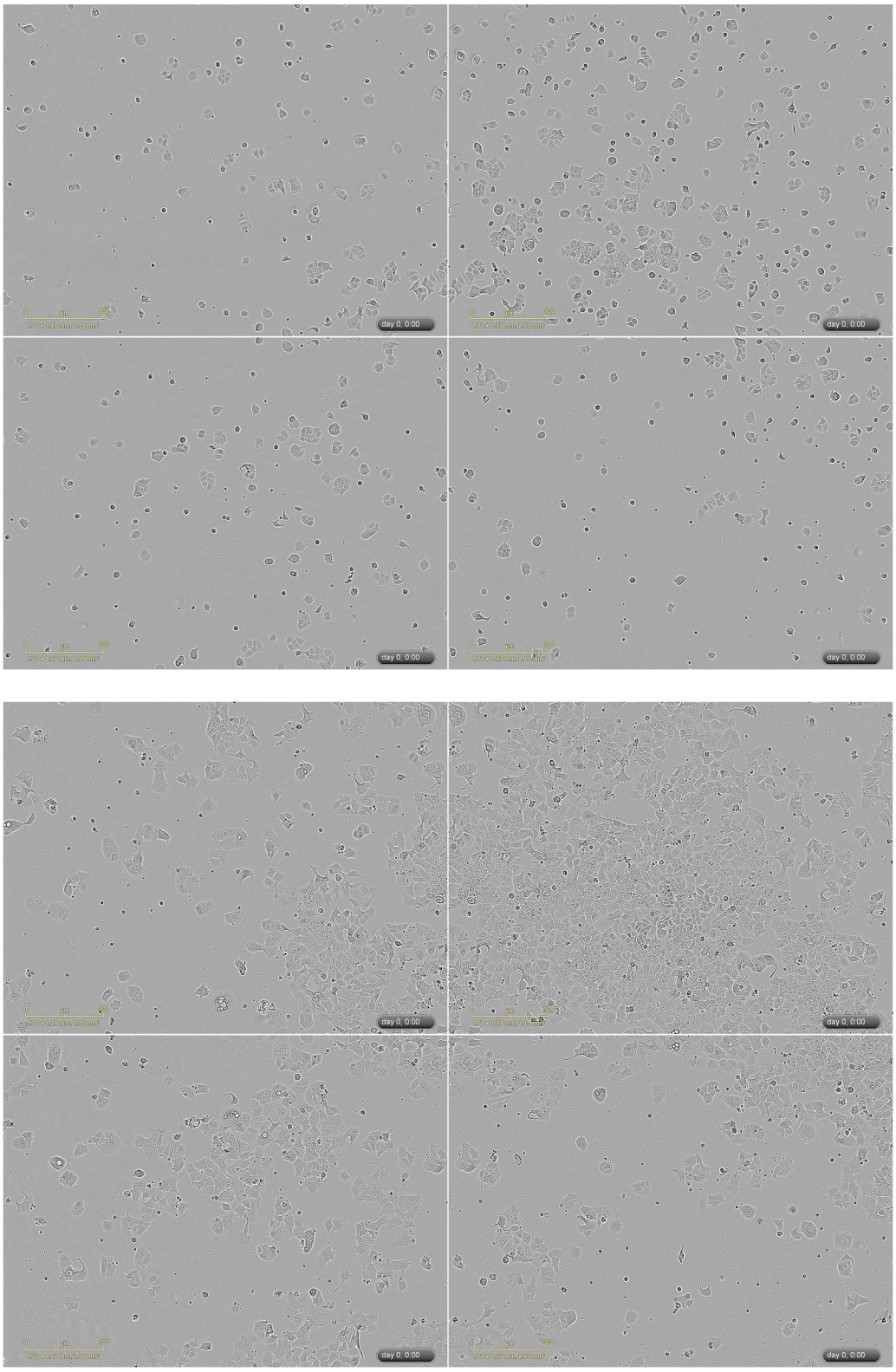

Figure 22 The cell viability of *ESR1-LPP* expressing T47D cells on Day 1 (top) and Day 12 (bottom) in hormone-deprived medium

Bibliography

Abate, F., Zairis, S., Ficarra, E., Acquaviva, A., Wiggins, C. H., Frattini, V., . . . Rabadan, R. (2014). Pegasus: a comprehensive annotation and prediction tool for detection of driver gene fusions in cancer. *BMC Syst Biol, 8*, 97. doi:10.1186/s12918-014-0097-z

Afrin, S., Zhang, C. R. C., Meyer, C., Stinson, C. L., Pham, T., Bruxner, T. J. C., . . . Moore, A. S. (2018). Targeted Next-Generation Sequencing for Detecting MLL Gene Fusions in Leukemia. *Mol Cancer Res, 16*(2), 279-285. doi:10.1158/1541-7786.MCR-17-0569

Andreopoulou, E., Kelly, C. M., & McDaid, H. M. (2017). Therapeutic Advances and New Directions for Triple-Negative Breast Cancer. *Breast Care (Basel), 12*(1), 21-28. doi:10.1159/000455821

Angus, L., Beije, N., Jager, A., Martens, J. W., & Sleijfer, S. (2017). ESR1 mutations: Moving towards guiding treatment decision-making in metastatic breast cancer patients. *Cancer Treat Rev, 52*, 33-40. doi:10.1016/j.ctrv.2016.11.001

Bahreini, A., Li, Z., Wang, P., Levine, K. M., Tasdemir, N., Cao, L., . . . Oesterreich, S. (2017). Mutation site and context dependent effects of ESR1 mutation in genome-edited breast cancer cell models. *Breast Cancer Res, 19*(1), 60. doi:10.1186/s13058-017-0851-4

Bailey, S. M., Meyne, J., Chen, D. J., Kurimasa, A., Li, G. C., Lehnert, B. E., & Goodwin, E. H. (1999). DNA double-strand break repair proteins are required to cap the ends of mammalian chromosomes. *Proc Natl Acad Sci U S A, 96*(26), 14899-14904.

Basudan, A., Priedigkeit, N., Hartmaier, R. J., Sokol, E. S., Bahreini, A., Watters, R. J., . . . Oesterreich, S. (2018). Frequent ESR1 and CDK Pathway Copy-Number Alterations in Metastatic Breast Cancer. *Mol Cancer Res*. doi:10.1158/1541-7786.MCR-18-0946

Bines, J., Earl, H., Buzaid, A. C., & Saad, E. D. (2014). Anthracyclines and taxanes in the neo/adjuvant treatment of breast cancer: does the sequence matter? *Ann Oncol, 25*(6), 1079-1085. doi:10.1093/annonc/mdu007

Blanchard, Z., Vahrenkamp, J. M., Berret, K. C., Arnesen, S., & Gertz, J. (2018). Estrogen-independent molecular actions of mutant estrogen receptor alpha in endometrial cancer. *BioRxiv*. doi:10.1101/369165

Boer, K. (2017). Fulvestrant in advanced breast cancer: evidence to date and place in therapy. *Ther Adv Med Oncol, 9*(7), 465-479. doi:10.1177/1758834017711097

Brown, L. A., Hoog, J., Chin, S. F., Tao, Y., Zayed, A. A., Chin, K., . . . Caldas, C. (2008). ESR1 gene amplification in breast cancer: a common phenomenon? *Nat Genet, 40*(7), 806-807; author reply 810-802. doi:10.1038/ng0708-806

Bunting, S. F., & Nussenzweig, A. (2013). End-joining, translocations and cancer. *Nat Rev Cancer, 13*(7), 443-454. doi:10.1038/nrc3537

Cancer Genome Atlas, N. (2012). Comprehensive molecular portraits of human breast tumours. *Nature, 490*(7418), 61-70. doi:10.1038/nature11412

Cardoso, F., Costa, A., Senkus, E., Aapro, M., Andre, F., Barrios, C. H., . . . Winer, E. (2017). 3rd ESO-ESMO International Consensus Guidelines for Advanced Breast Cancer (ABC 3). *Ann Oncol, 28*(1), 16-33. doi:10.1093/annonc/mdw544

Carroll, J. S. (2016). Mechanisms of oestrogen receptor (ER) gene regulation in breast cancer. *Eur J Endocrinol, 175*(1), R41-49. doi:10.1530/EJE-16-0124

Chumsri, S., Howes, T., Bao, T., Sabnis, G., & Brodie, A. (2011). Aromatase, aromatase inhibitors, and breast cancer. *J Steroid Biochem Mol Biol, 125*(1-2), 13-22. doi:10.1016/j.jsbmb.2011.02.001

Clarke, R., Tyson, J. J., & Dixon, J. M. (2015). Endocrine resistance in breast cancer--An overview and update. *Mol Cell Endocrinol, 418 Pt 3*, 220-234. doi:10.1016/j.mce.2015.09.035

Courjal, F., Louason, G., Speiser, P., Katsaros, D., Zeillinger, R., & Theillet, C. (1996). Cyclin gene amplification and overexpression in breast and ovarian cancers: evidence for the selection of cyclin D1 in breast and cyclin E in ovarian tumors. *Int J Cancer, 69*(4), 247-253. doi:10.1002/(SICI)1097-0215(19960822)69:4<247::AID-IJC1>3.0.CO;2-X

Curtis, C., Shah, S. P., Chin, S. F., Turashvili, G., Rueda, O. M., Dunning, M. J., . . . Aparicio, S. (2012). The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature, 486*(7403), 346-352. doi:10.1038/nature10983

De Marchi, T., Foekens, J. A., Umar, A., & Martens, J. W. (2016). Endocrine therapy resistance in estrogen receptor (ER)-positive breast cancer. *Drug Discov Today, 21*(7), 1181-1188. doi:10.1016/j.drudis.2016.05.012

Domenici, G., Aurrekoetxea-Rodriguez, I., Simoes, B. M., Rabano, M., Lee, S. Y., Millan, J. S., . . . Vivanco, M. D. (2019). A Sox2-Sox9 signalling axis maintains human breast luminal progenitor and breast cancer stem cells. *Oncogene*. doi:10.1038/s41388-018-0656-7

Early Breast Cancer Trialists' Collaborative, G. (2005). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet, 365*(9472), 1687-1717. doi:10.1016/S0140-6736(05)66544-0

Early Breast Cancer Trialists' Collaborative, G., Davies, C., Godwin, J., Gray, R., Clarke, M., Cutter, D., . . . Peto, R. (2011). Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. *Lancet, 378*(9793), 771-784. doi:10.1016/S0140-6736(11)60993-8

Edge, S. B., & Compton, C. C. (2010). The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. *Ann Surg Oncol, 17*(6), 1471-1474. doi:10.1245/s10434-010-0985-4

Edwards, P. A. (2010). Fusion genes and chromosome translocations in the common epithelial cancers. *J Pathol, 220*(2), 244-254. doi:10.1002/path.2632

Edwards, P. A., & Howarth, K. D. (2012). Are breast cancers driven by fusion genes? *Breast Cancer Res, 14*(2), 303. doi:10.1186/bcr3122

Ejlertsen, B., Aldridge, J., Nielsen, K. V., Regan, M. M., Henriksen, K. L., Lykkesfeldt, A. E., . . . International Breast Cancer Study, G. (2012). Prognostic and predictive role of ESR1 status for postmenopausal patients with endocrine-responsive early breast cancer in the Danish cohort of the BIG 1-98 trial. *Ann Oncol, 23*(5), 1138-1144. doi:10.1093/annonc/mdr438

Elston, C. W., & Ellis, I. O. (1991). Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology, 19*(5), 403-410.

Fribbens, C., O'Leary, B., Kilburn, L., Hrebien, S., Garcia-Murillas, I., Beaney, M., . . . Turner, N. C. (2016). Plasma ESR1 Mutations and the Treatment of Estrogen Receptor-Positive Advanced Breast Cancer. *J Clin Oncol, 34*(25), 2961-2968. doi:10.1200/JCO.2016.67.3061

Giltnane, J. M., Hutchinson, K. E., Stricker, T. P., Formisano, L., Young, C. D., Estrada, M. V., . . . Arteaga, C. L. (2017). Genomic profiling of ER(+) breast cancers after short-term estrogen suppression reveals alterations associated with endocrine resistance. *Sci Transl Med, 9*(402). doi:10.1126/scitranslmed.aai7993

Ginsburg, O., Bray, F., Coleman, M. P., Vanderpuye, V., Eniu, A., Kotha, S. R., . . . Conteh, L. (2017). The global burden of women's cancers: a grand challenge in global health. *Lancet, 389*(10071), 847-860. doi:10.1016/S0140-6736(16)31392-7

Goni, R., Garcia, P., & Foissac, S. (2009). The qPCR data statistical analysis *Integromics White Paper*. https://gene-quantification.de/integromics-qpcr-statistics-white-paper.pdf

Gruber, C. J., Gruber, D. M., Gruber, I. M., Wieser, F., & Huber, J. C. (2004). Anatomy of the estrogen response element. *Trends Endocrinol Metab, 15*(2), 73-78. doi:10.1016/j.tem.2004.01.008

Guttery, D. S., Page, K., Hills, A., Woodley, L., Marchese, S. D., Rghebi, B., . . . Shaw, J. A. (2015). Noninvasive detection of activating estrogen receptor 1 (ESR1) mutations in estrogen receptor-positive metastatic breast cancer. *Clin Chem, 61*(7), 974-982. doi:10.1373/clinchem.2015.238717

Haas, B., Dobin, A., Stransky, N., Li, B., Yang, X., Tickle, T., . . . Regev, A. (2017). STAR-Fusion: Fast and Accurate Fusion Transcript Detection from RNA-Seq. *BioRxiv*. doi:https://doi.org/10.1101/120295

Hall, J. M., Couse, J. F., & Korach, K. S. (2001). The multifaceted mechanisms of estradiol and estrogen receptor signaling. *J Biol Chem, 276*(40), 36869-36872. doi:10.1074/jbc.R100029200

Hartmaier, R. J., Trabucco, S. E., Priedigkeit, N., Chung, J. H., Parachoniak, C. A., Vanden Borre, P., . . . Lee, A. V. (2018). Recurrent hyperactive ESR1 fusion proteins in endocrine therapy-resistant breast cancer. *Ann Oncol, 29*(4), 872-880. doi:10.1093/annonc/mdy025

Hasty, P., & Montagna, C. (2014). Chromosomal Rearrangements in Cancer: Detection and potential causal mechanisms. *Mol Cell Oncol, 1*(1). doi:10.4161/mco.29904

Holst, F., Hoivik, E. A., Gibson, W. J., Taylor-Weiner, A., Schumacher, S. E., Asmann, Y. W., . . . Cherniack, A. D. (2016). Recurrent hormone-binding domain truncated ESR1 amplifications in primary endometrial cancers suggest their implication in hormone independent growth. *Sci Rep, 6*, 25521. doi:10.1038/srep25521

Holst, F., Stahl, P. R., Ruiz, C., Hellwinkel, O., Jehan, Z., Wendland, M., . . . Simon, R. (2007). Estrogen receptor alpha (ESR1) gene amplification is frequent in breast cancer. *Nat Genet, 39*(5), 655-660. doi:10.1038/ng2006

Horlings, H. M., Bergamaschi, A., Nordgard, S. H., Kim, Y. H., Han, W., Noh, D. Y., . . . van de Vijver, M. J. (2008). ESR1 gene amplification in breast cancer: a common phenomenon? *Nat Genet, 40*(7), 807-808; author reply 810-802. doi:10.1038/ng0708-807

Howlader, N., Altekruse, S. F., Li, C. I., Chen, V. W., Clarke, C. A., Ries, L. A., & Cronin, K. A. (2014). US incidence of breast cancer subtypes defined by joint hormone receptor and HER2 status. *J Natl Cancer Inst, 106*(5). doi:10.1093/jnci/dju055

Hu, X., Wang, Q., Tang, M., Barthel, F., Amin, S., Yoshihara, K., . . . Verhaak, R. G. W. (2018). TumorFusions: an integrative resource for cancer-associated transcript fusions. *Nucleic Acids Res, 46*(D1), D1144-D1149. doi:10.1093/nar/gkx1018

Iqbal, N., & Iqbal, N. (2014). Human Epidermal Growth Factor Receptor 2 (HER2) in Cancers: Overexpression and Therapeutic Implications. *Mol Biol Int, 2014*, 852748. doi:10.1155/2014/852748

Iyer, M. K., Chinnaiyan, A. M., & Maher, C. A. (2011). ChimeraScan: a tool for identifying chimeric transcription in sequencing data. *Bioinformatics, 27*(20), 2903-2904. doi:10.1093/bioinformatics/btr467

JavanMoghadam, S., Weihua, Z., Hunt, K. K., & Keyomarsi, K. (2016). Estrogen receptor alpha is cell cycle-regulated and regulates the cell cycle in a ligand-dependent fashion. *Cell Cycle, 15*(12), 1579-1590. doi:10.1080/15384101.2016.1166327

Jensen, E. V., Jacobson, H. I., Walf, A. A., & Frye, C. A. (2010). Estrogen action: a historic perspective on the implications of considering alternative approaches. *Physiol Behav, 99*(2), 151-162. doi:10.1016/j.physbeh.2009.08.013

Jeselsohn, R., Cornwell, M., Pun, M., Buchwalter, G., Nguyen, M., Bango, C., . . . Brown, M. (2017). Embryonic transcription factor SOX9 drives breast cancer endocrine resistance. *Proc Natl Acad Sci U S A, 114*(22), E4482-E4491. doi:10.1073/pnas.1620993114

Jeselsohn, R., De Angelis, C., Brown, M., & Schiff, R. (2017). The Evolving Role of the Estrogen Receptor Mutations in Endocrine Therapy-Resistant Breast Cancer. *Curr Oncol Rep, 19*(5), 35. doi:10.1007/s11912-017-0591-8

Jeselsohn, R., Yelensky, R., Buchwalter, G., Frampton, G., Meric-Bernstam, F., Gonzalez-Angulo, A. M., . . . Miller, V. A. (2014). Emergence of constitutively active estrogen receptor-alpha mutations in pretreated advanced estrogen receptor-positive breast cancer. *Clin Cancer Res, 20*(7), 1757-1767. doi:10.1158/1078-0432.CCR-13-2332

Jordan, V. C. (2003). Tamoxifen: a most unlikely pioneering medicine. *Nat Rev Drug Discov, 2*(3), 205-213. doi:10.1038/nrd1031

Kocanova, S., Mazaheri, M., Caze-Subra, S., & Bystricky, K. (2010). Ligands specify estrogen receptor alpha nuclear localization and degradation. *BMC Cell Biol, 11*, 98. doi:10.1186/1471-2121-11-98

Kumar, R., Zakharov, M. N., Khan, S. H., Miki, R., Jang, H., Toraldo, G., . . . Jasuja, R. (2011). The dynamic structure of the estrogen receptor. *J Amino Acids, 2011*, 812540. doi:10.4061/2011/812540

Kumar, S., Razzaq, S. K., Vo, A. D., Gautam, M., & Li, H. (2016). Identifying fusion transcripts using next generation sequencing. *Wiley Interdiscip Rev RNA, 7*(6), 811-823. doi:10.1002/wrna.1382

Kumar, S., Vo, A. D., Qin, F., & Li, H. (2016). Comparative assessment of methods for the fusion transcripts detection from RNA-Seq data. *Sci Rep, 6*, 21597. doi:10.1038/srep21597

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

Kumar-Sinha, C., Tomlins, S. A., & Chinnaiyan, A. M. (2008). Recurrent gene fusions in prostate cancer. *Nat Rev Cancer, 8*(7), 497-511. doi:10.1038/nrc2402

Lakhani SR, E. I., Schnitt Sj, Tan PH, van de Vijver MJ. (2012). WHO Classification of Tumours of the Breast. In *WHO Classification of Tumours* (4th ed., Vol. 4th). Lyon: IARC Press.

Latysheva, N. S., & Babu, M. M. (2016). Discovering and understanding oncogenic gene fusions through data intensive computational approaches. *Nucleic Acids Res, 44*(10), 4487-4503. doi:10.1093/nar/gkw282

Lee, A. V., Cui, X., & Oesterreich, S. (2001). Cross-talk among estrogen receptor, epidermal growth factor, and insulin-like growth factor signaling in breast cancer. *Clin Cancer Res, 7*(12 Suppl), 4429s-4435s; discussion 4411s-4412s.

Lefebvre, C., Bachelot, T., Filleron, T., Pedrero, M., Campone, M., Soria, J. C., . . . Andre, F. (2016). Mutational Profile of Metastatic Breast Cancers: A Retrospective Analysis. *PLoS Med, 13*(12), e1002201. doi:10.1371/journal.pmed.1002201

Lei, J. T., Gou, X., & Ellis, M. J. (2018). ESR1 fusions drive endocrine therapy resistance and metastasis in breast cancer. *Mol Cell Oncol, 5*(6), e1526005. doi:10.1080/23723556.2018.1526005

Lei, J. T., Gou, X., Seker, S., Haricharan, S., Lee, A. V., Robinson, D. R., & Ellis, M. (2018). *Functional and therapeutic significance of ESR1 fusions in metastatic ER+ breast cancer*. Paper presented at the San Antonio Breast Cancer Symposium, San Antonio.

Lei, J. T., Shao, J., Zhang, J., Iglesia, M., Chan, D. W., Cao, J., . . . Ellis, M. J. (2018). Functional Annotation of ESR1 Gene Fusions in Estrogen Receptor-Positive Breast Cancer. *Cell Rep, 24*(6), 1434-1444 e1437. doi:10.1016/j.celrep.2018.07.009

Li, S., Shen, D., Shao, J., Crowder, R., Liu, W., Prat, A., . . . Ellis, M. J. (2013). Endocrine-therapy-resistant ESR1 variants revealed by genomic characterization of breast-cancer-derived xenografts. *Cell Rep, 4*(6), 1116-1130. doi:10.1016/j.celrep.2013.08.022

Liu, S., Tsai, W. H., Ding, Y., Chen, R., Fang, Z., Huo, Z., . . . Tseng, G. C. (2016). Comprehensive evaluation of fusion transcript detection algorithms and a meta-caller to combine top performing methods in paired-end RNA-seq data. *Nucleic Acids Res, 44*(5), e47. doi:10.1093/nar/gkv1234

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods, 25*(4), 402-408. doi:10.1006/meth.2001.1262

Lukong, K. E. (2017). Understanding breast cancer - The long and winding road. *BBA Clin, 7*, 64-77. doi:10.1016/j.bbacli.2017.01.001

Mao, C., Patterson, N. M., Cherian, M. T., Aninye, I. O., Zhang, C., Montoya, J. B., . . . Shapiro, D. J. (2008). A new small molecule inhibitor of estrogen receptor alpha binding to estrogen response elements blocks estrogen-dependent growth of cancer cells. *J Biol Chem, 283*(19), 12819-12830. doi:10.1074/jbc.M709936200

Marcand, S. (2014). How do telomeres and NHEJ coexist? *Mol Cell Oncol, 1*(3), e963438. doi:10.4161/23723548.2014.963438

Marino, M., Galluzzo, P., & Ascenzi, P. (2006). Estrogen signaling multiple pathways to impact gene transcription. *Curr Genomics, 7*(8), 497-508.

Markiewicz, A., Welnicka-Jaskiewicz, M., Skokowski, J., Jaskiewicz, J., Szade, J., Jassem, J., & Zaczek, A. J. (2013). Prognostic significance of ESR1 amplification and ESR1 PvuII, CYP2C19*2, UGT2B15*2 polymorphisms in breast cancer patients. *PLoS One, 8*(8), e72219. doi:10.1371/journal.pone.0072219

McDevitt, M. A., Glidewell-Kenney, C., Jimenez, M. A., Ahearn, P. C., Weiss, J., Jameson, J. L., & Levine, J. E. (2008). New insights into the classical and non-classical actions of estrogen: evidence from estrogen receptor knock-out and knock-in mice. *Mol Cell Endocrinol, 290*(1-2), 24-30. doi:10.1016/j.mce.2008.04.003

McDowall, S., Argentaro, A., Ranganathan, S., Weller, P., Mertin, S., Mansour, S., . . . Harley, V. (1999). Functional and structural studies of wild type SOX9 and mutations causing campomelic dysplasia. *J Biol Chem, 274*(34), 24023-24030.

Merenbakh-Lamin, K., Ben-Baruch, N., Yeheskel, A., Dvir, A., Soussan-Gutman, L., Jeselsohn, R., . . . Wolf, I. (2013). D538G mutation in estrogen receptor-alpha: A novel mechanism for acquired endocrine resistance in breast cancer. *Cancer Res, 73*(23), 6856-6864. doi:10.1158/0008-5472.CAN-13-1197

Meric-Bernstam, F., Frampton, G. M., Ferrer-Lozano, J., Yelensky, R., Perez-Fidalgo, J. A., Wang, Y., . . . Gonzalez-Angulo, A. M. (2014). Concordance of genomic alterations between primary and recurrent breast cancer. *Mol Cancer Ther, 13*(5), 1382-1389. doi:10.1158/1535-7163.MCT-13-0482

Mertens, F., Johansson, B., Fioretos, T., & Mitelman, F. (2015). The emerging complexity of gene fusions in cancer. *Nat Rev Cancer, 15*(6), 371-381. doi:10.1038/nrc3947

Mittal, V. K., & McDonald, J. F. (2017). De novo assembly and characterization of breast cancer transcriptomes identifies large numbers of novel fusion-gene transcripts of potential functional significance. *BMC Med Genomics, 10*(1), 53. doi:10.1186/s12920-017-0289-7

Moelans, C. B., Monsuur, H. N., de Pinth, J. H., Radersma, R. D., de Weger, R. A., & van Diest, P. J. (2011). ESR1 amplification is rare in breast cancer and is associated with high grade and high proliferation: a multiplex ligation-dependent probe amplification study. *Cell Oncol (Dordr), 34*(5), 489-494. doi:10.1007/s13402-011-0045-5

Musgrove, E. A., & Sutherland, R. L. (2009). Biological determinants of endocrine resistance in breast cancer. *Nat Rev Cancer, 9*(9), 631-643. doi:10.1038/nrc2713

Nembrot, M., Quintana, B., & Mordoh, J. (1990). Estrogen receptor gene amplification is found in some estrogen receptor-positive human breast tumors. *Biochem Biophys Res Commun, 166*(2), 601-607.

Nicorici, D., Satalan, M., Edgren, H., Kangaspeska, S., Murumagi, A., Kallioniemi, O., . . . Kilkku, O. (2014). FusionCatcher - a tool for finding somatic fusion genes in paired-end RNA-sequencing data. *BioRxiv*. doi:10.1101/011650

Nielsen, K. V., Ejlertsen, B., Muller, S., Moller, S., Rasmussen, B. B., Balslev, E., . . . Mouridsen, H. T. (2011). Amplification of ESR1 may predict resistance to adjuvant tamoxifen in postmenopausal patients with hormone receptor positive breast cancer. *Breast Cancer Res Treat, 127*(2), 345-355. doi:10.1007/s10549-010-0984-y

Nik-Zainal, S., Davies, H., Staaf, J., Ramakrishna, M., Glodzik, D., Zou, X., . . . Stratton, M. R. (2016). Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature, 534*(7605), 47-54. doi:10.1038/nature17676

Noone, A. M., Howlader, N., Krapcho, M., Miller, D., Brest, A., Yu, M., . . . Cronin, K. A. (2017). *SEER Cancer Statistics Review, 1975-2015*. CA: A Cancer Journal for Clinicians. Bethesda, MD.

O'Lone, R., Frith, M. C., Karlsson, E. K., & Hansen, U. (2004). Genomic targets of nuclear estrogen receptors. *Mol Endocrinol, 18*(8), 1859-1875. doi:10.1210/me.2003-0044

Oesterreich, S., Zhang, P., Guler, R. L., Sun, X., Curran, E. M., Welshons, W. V., . . . Lee, A. V. (2001). Re-expression of estrogen receptor alpha in estrogen receptor alpha-negative MCF-7 cells restores both estrogen and insulin-like growth factor-mediated signaling and growth. *Cancer Res, 61*(15), 5771-5777.

Osborne, C. K., & Schiff, R. (2011). Mechanisms of endocrine resistance in breast cancer. *Annu Rev Med, 62*, 233-247. doi:10.1146/annurev-med-070909-182917

Pan, H., Gray, R., Braybrooke, J., Davies, C., Taylor, C., McGale, P., . . . Ebctcg. (2017). 20-Year Risks of Breast-Cancer Recurrence after Stopping Endocrine Therapy at 5 Years. *N Engl J Med, 377*(19), 1836-1846. doi:10.1056/NEJMoa1701830

Paratala, B. S., Dolfi, S. C., Khiabanian, H., Rodriguez-Rodriguez, L., Ganesan, S., & Hirshfield, K. M. (2016). Emerging Role of Genomic Rearrangements in Breast Cancer: Applying Knowledge from Other Cancers. *Biomark Cancer, 8*(Supple 1), 1-14. doi:10.4137/BIC.S34417

Parker, B. C., & Zhang, W. (2013). Fusion genes in solid tumors: an emerging target for cancer diagnosis and treatment. *Chin J Cancer, 32*(11), 594-603. doi:10.5732/cjc.013.10178

Perou, C. M., Sorlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., . . . Botstein, D. (2000). Molecular portraits of human breast tumours. *Nature, 406*(6797), 747-752. doi:10.1038/35021093

Perry, J. J., Ballard, G. D., Albert, A. E., Dobrolecki, L. E., Malkas, L. H., & Hoelz, D. J. (2015). Human C6orf211 encodes Armt1, a protein carboxyl methyltransferase that targets PCNA and is linked to the DNA damage response. *Cell Rep, 10*(8), 1288-1296. doi:10.1016/j.celrep.2015.01.054

Petit, M. M., Fradelizi, J., Golsteyn, R. M., Ayoubi, T. A., Menichi, B., Louvard, D., . . . Friederich, E. (2000). LPP, an actin cytoskeleton protein related to zyxin, harbors a nuclear export

signal and transcriptional activation capacity. *Mol Biol Cell, 11*(1), 117-129. doi:10.1091/mbc.11.1.117

Petit, M. M., Meulemans, S. M., Alen, P., Ayoubi, T. A., Jansen, E., & Van de Ven, W. J. (2005). The tumor suppressor Scrib interacts with the zyxin-related protein LPP, which shuttles between cell adhesion sites and the nucleus. *BMC Cell Biol, 6*(1), 1. doi:10.1186/1471-2121-6-1

Priedigkeit, N., Watters, R. J., Lucas, P. C., Basudan, A., Bhargava, R., Horne, W., . . . Lee, A. V. (2017). Exome-capture RNA sequencing of decade-old breast cancers and matched decalcified bone metastases. *JCI Insight, 2*(17). doi:10.1172/jci.insight.95703

Pupo, M., Maggiolini, M., & Musti, A. M. (2016). GPER Mediates Non-Genomic Effects of Estrogen. *Methods Mol Biol, 1366*, 471-488. doi:10.1007/978-1-4939-3127-9_37

Rabbani, S. A., & Mazar, A. P. (2007). Evaluating distant metastases in breast cancer: from biology to outcomes. *Cancer Metastasis Rev, 26*(3-4), 663-674. doi:10.1007/s10555-007-9085-8

Reis-Filho, J. S., Drury, S., Lambros, M. B., Marchio, C., Johnson, N., Natrajan, R., . . . Dowsett, M. (2008). ESR1 gene amplification in breast cancer: a common phenomenon? *Nat Genet, 40*(7), 809-810; author reply 810-802. doi:10.1038/ng0708-809b

Reynolds, K., Sarangi, S., Bardia, A., & Dizon, D. S. (2014). Precision medicine and personalized breast cancer: combination pertuzumab therapy. *Pharmgenomics Pers Med, 7*, 95-105. doi:10.2147/PGPM.S37100

Robinson, D. R., Wu, Y. M., Lonigro, R. J., Vats, P., Cobain, E., Everett, J., . . . Chinnaiyan, A. M. (2017). Integrative clinical genomics of metastatic cancer. *Nature, 548*(7667), 297-303. doi:10.1038/nature23306

Robinson, D. R., Wu, Y. M., Vats, P., Su, F., Lonigro, R. J., Cao, X., . . . Chinnaiyan, A. M. (2013). Activating ESR1 mutations in hormone-resistant metastatic breast cancer. *Nat Genet, 45*(12), 1446-1451. doi:10.1038/ng.2823

Rodgers, K., & McVey, M. (2016). Error-Prone Repair of DNA Double-Strand Breaks. *J Cell Physiol, 231*(1), 15-24. doi:10.1002/jcp.25053

Rosenfeld, M. G., & Glass, C. K. (2001). Coregulator codes of transcriptional regulation by nuclear receptors. *J Biol Chem, 276*(40), 36865-36868. doi:10.1074/jbc.R100041200

Roskoski, R., Jr. (2014). The ErbB/HER family of protein-tyrosine kinases and cancer. *Pharmacol Res, 79*, 34-74. doi:10.1016/j.phrs.2013.11.002

Russnes, H. G., Lingjaerde, O. C., Borresen-Dale, A. L., & Caldas, C. (2017). Breast Cancer Molecular Stratification: From Intrinsic Subtypes to Integrative Clusters. *Am J Pathol, 187*(10), 2152-2162. doi:10.1016/j.ajpath.2017.04.022

Russo, J., & Russo, I. H. (2006). The role of estrogen in the initiation of breast cancer. *J Steroid Biochem Mol Biol, 102*(1-5), 89-96. doi:10.1016/j.jsbmb.2006.09.004

Schiavon, G., Hrebien, S., Garcia-Murillas, I., Cutts, R. J., Pearson, A., Tarazona, N., . . . Turner, N. C. (2015). Analysis of ESR1 mutation in circulating tumor DNA demonstrates evolution during therapy for metastatic breast cancer. *Sci Transl Med, 7*(313), 313ra182. doi:10.1126/scitranslmed.aac7551

Schmittgen, T. D., & Livak, K. J. (2008). Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc, 3*(6), 1101-1108.

Schram, A. M., Chang, M. T., Jonsson, P., & Drilon, A. (2017). Fusions in solid tumours: diagnostic strategies, targeted therapy, and acquired resistance. *Nat Rev Clin Oncol, 14*(12), 735-748. doi:10.1038/nrclinonc.2017.127

Schwabe, J. W., Chapman, L., Finch, J. T., & Rhodes, D. (1993). The crystal structure of the estrogen receptor DNA-binding domain bound to DNA: how receptors discriminate between their response elements. *Cell, 75*(3), 567-578.

Shoman, N., Klassen, S., McFadden, A., Bickis, M. G., Torlakovic, E., & Chibbar, R. (2005). Reduced PTEN expression predicts relapse in patients with breast carcinoma treated by tamoxifen. *Mod Pathol, 18*(2), 250-259. doi:10.1038/modpathol.3800296

Shugay, M., Ortiz de Mendibil, I., Vizmanos, J. L., & Novo, F. J. (2013). Oncofuse: a computational framework for the prediction of the oncogenic potential of gene fusions. *Bioinformatics, 29*(20), 2539-2546. doi:10.1093/bioinformatics/btt445

Siegel, R. L., Miller, K. D., & Jemal, A. (2018). Cancer statistics, 2018. *CA Cancer J Clin, 68*(1), 7-30. doi:10.3322/caac.21442

Slamon, D. J., Clark, G. M., Wong, S. G., Levin, W. J., Ullrich, A., & McGuire, W. L. (1987). Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science, 235*(4785), 177-182.

Sorlie, T., Perou, C. M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., . . . Borresen-Dale, A. L. (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A, 98*(19), 10869-10874. doi:10.1073/pnas.191367098

Stephens, P. J., McBride, D. J., Lin, M. L., Varela, I., Pleasance, E. D., Simpson, J. T., . . . Stratton, M. R. (2009). Complex landscapes of somatic rearrangement in human breast cancer genomes. *Nature, 462*(7276), 1005-1010. doi:10.1038/nature08645

Stover, D. G., Bell, C. F., & Tolaney, S. M. (2016). Neoadjuvant and Adjuvant Chemotherapy Considerations for Triple-Negative Breast Cancer. *American Journal of Hematology / Oncology, 12*(3).

Toft, D., & Gorski, J. (1966). A receptor molecule for estrogens: isolation from the rat uterus and preliminary characterization. *Proc Natl Acad Sci U S A, 55*(6), 1574-1581.

Tomita, S., Zhang, Z., Nakano, M., Ibusuki, M., Kawazoe, T., Yamamoto, Y., & Iwase, H. (2009). Estrogen receptor alpha gene ESR1 amplification may predict endocrine therapy responsiveness in breast cancer patients. *Cancer Sci, 100*(6), 1012-1017. doi:10.1111/j.1349-7006.2009.01145.x

Toy, W., Shen, Y., Won, H., Green, B., Sakr, R. A., Will, M., . . . Chandarlapaty, S. (2013). ESR1 ligand-binding domain mutations in hormone-resistant breast cancer. *Nat Genet, 45*(12), 1439-1445. doi:10.1038/ng.2822

Toy, W., Weir, H., Razavi, P., Lawson, M., Goeppert, A. U., Mazzola, A. M., . . . Chandarlapaty, S. (2017). Activating ESR1 Mutations Differentially Affect the Efficacy of ER Antagonists. *Cancer Discov, 7*(3), 277-287. doi:10.1158/2159-8290.CD-15-1523

Vareslija, D., Priedigkeit, N., Fagan, A., Purcell, S., Cosgrove, N., O'Halloran, P. J., . . . Young, L. S. (2018). Transcriptome Characterization of Matched Primary Breast and Brain Metastatic Tumors to Detect Novel Actionable Targets. *J Natl Cancer Inst*. doi:10.1093/jnci/djy110

Veeraraghavan, J., Tan, Y., Cao, X. X., Kim, J. A., Wang, X., Chamness, G. C., . . . Wang, X. S. (2014). Recurrent ESR1-CCDC170 rearrangements in an aggressive subset of oestrogen receptor-positive breast cancers. *Nat Commun, 5*, 4577. doi:10.1038/ncomms5577

Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. A., Jr., & Kinzler, K. W. (2013). Cancer genome landscapes. *Science, 339*(6127), 1546-1558. doi:10.1126/science.1235122

Vrtacnik, P., Ostanek, B., Mencej-Bedrac, S., & Marc, J. (2014). The many faces of estrogen signaling. *Biochem Med (Zagreb), 24*(3), 329-342. doi:10.11613/BM.2014.035

Walker, A. J., Wedam, S., Amiri-Kordestani, L., Bloomquist, E., Tang, S., Sridhara, R., . . . Pazdur, R. (2016). FDA Approval of Palbociclib in Combination with Fulvestrant for the Treatment of Hormone Receptor-Positive, HER2-Negative Metastatic Breast Cancer. *Clin Cancer Res, 22*(20), 4968-4972. doi:10.1158/1078-0432.CCR-16-0493

Wang, L. H., Yang, X. Y., Zhang, X., Mihalic, K., Fan, Y. X., Xiao, W., . . . Farrar, W. L. (2004). Suppression of breast cancer by chemical modulation of vulnerable zinc fingers in estrogen receptor. *Nat Med, 10*(1), 40-47. doi:10.1038/nm969

Wang, P., Bahreini, A., Gyanchandani, R., Lucas, P. C., Hartmaier, R. J., Watters, R. J., . . . Oesterreich, S. (2016). Sensitive Detection of Mono- and Polyclonal ESR1 Mutations in Primary Tumors, Metastatic Lesions, and Cell-Free DNA of Breast Cancer Patients. *Clin Cancer Res, 22*(5), 1130-1137. doi:10.1158/1078-0432.CCR-15-1534

Wilson, F. R., Coombes, M. E., Wylie, Q., Yurchenko, M., Brezden-Masley, C., Hutton, B., . . . Cameron, C. (2017). Herceptin(R) (trastuzumab) in HER2-positive early breast cancer:

protocol for a systematic review and cumulative network meta-analysis. *Syst Rev, 6*(1), 196. doi:10.1186/s13643-017-0588-2

Yasar, P., Ayaz, G., User, S. D., Gupur, G., & Muyan, M. (2017). Molecular mechanism of estrogen-estrogen receptor signaling. *Reprod Med Biol, 16*(1), 4-20. doi:10.1002/rmb2.12006

Yates, L. R., Knappskog, S., Wedge, D., Farmery, J. H. R., Gonzalez, S., Martincorena, I., . . . Campbell, P. J. (2017). Genomic Evolution of Breast Cancer Metastasis and Relapse. *Cancer Cell, 32*(2), 169-184 e167. doi:10.1016/j.ccell.2017.07.005

Yu, Y., Yin, W., Yu, Z. H., Zhou, Y. J., Chi, J. R., Ge, J., & Cao, X. C. (2019). miR-190 enhances endocrine therapy sensitivity by regulating SOX9 expression in breast cancer. *J Exp Clin Cancer Res, 38*(1), 22. doi:10.1186/s13046-019-1039-9

Zhang, Q. X., Borg, A., Wolf, D. M., Oesterreich, S., & Fuqua, S. A. (1997). An estrogen receptor mutant with strong hormone-independent activity from a metastatic breast cancer. *Cancer Res, 57*(7), 1244-1249.

Zheng, Z., Liebers, M., Zhelyazkova, B., Cao, Y., Panditi, D., Lynch, K. D., . . . Le, L. P. (2014). Anchored multiplex PCR for targeted next-generation sequencing. *Nat Med, 20*(12), 1479-1484. doi:10.1038/nm.3729