# Clonal heterogeneity of endocrine therapy resistance in breast cancer

Lukas Beumers

Oral examination: 10.05.2023

Inaugural dissertation

for

obtaining the doctoral degree

of the

Combined Faculty of Mathematics, Engineering and Natural Sciences

of the

Ruprecht - Karls - University

Heidelberg

Presented by

M.Sc. Lukas Beumers

Born in: Heinsberg, Germany

Oral examination: 10.05.2023

# Clonal heterogeneity of endocrine therapy resistance in breast cancer

Referees: Prof. Dr. Stefan Wiemann

Prof. Dr. Peter Angel

*Für Andrea und Elsbeth*

## Declaration of Authorship

I hereby declare that the work presented in my dissertation was carried out between January 2020 and December 2022 under the supervision of Prof. Dr. Stefan Wiemann in the group of Molecular Genome Analysis at the German Cancer Research Center (DKFZ, Heidelberg, Germany).

If not stated differently and referenced within the text, the data described in my dissertation is original, has been gathered by myself and has not yet been presented as a part of a university examination. All main sources, as well as the work of joint cooperation have been referenced appropriately. I, as author, hereby declare no potential conflict of interest.

Heidelberg, ___________________________________

Lukas Beumers

## Table of Contents

**Summary** 1

**Zusammenfassung** 3

**1 Introduction** 5

1.1 Breast cancer 5

1.1.1 Molecular subtypes 5

1.1.2 Endocrine therapy 7

1.1.3 Acquired endocrine therapy resistance 8

1.2 Tumor heterogeneity 12

1.2.1 Clonal evolution and cancer stem cells 12

1.2.2 Selection of pre-existing clones and resistance acquisition 13

1.2.3 Tumor heterogeneity in breast cancer and endocrine therapy resistance 15

**2 Aims of the thesis** 19

**3 Materials and Methods** 21

3.1 Materials 21

3.1.1 Instruments 21

3.1.2 Consumables 21

3.1.3 Chemicals and reagents 22

3.1.4 Solutions 23

3.1.5 Commercial kits 24

3.1.6 siRNAs 24

3.1.7 RT-qPCR primers 25

3.1.8 Antibodies 25

3.1.9 Software 25

3.2 Methods 26

3.2.1 Cell culture 26

3.2.2 Sanger sequencing-based barcode analysis 33

3.2.3 NGS-based barcode analysis 35

3.2.4 Barcode integration site analysis 37

3.2.5 Sequencing of *ESR1* hotspot mutation codons 38

3.2.6 Analysis of RNA expression 38

3.2.7 Analysis of protein expression and phosphorylation 41

3.2.8 Mass Spectrometry-based phospho-proteomics 42

3.2.9 Estimation of kinase activities 44

3.2.10 Analysis of CPTAC-BRCA dataset ........ 44
3.2.11 Statistical analysis ........ 45
**4 Results ........ 47**
4.1 Cellular barcoding reveals distinct clonal complexities in endocrine therapy resistance ........ 47
4.2 Barcode composition analysis of final cell pools ........ 49
4.3 Project overview ........ 51
4.4 Deconvoluting the clonal complexity of barcoded cell pools ........ 52
4.5 Deconvoluting endocrine therapy resistant cell pools ........ 53
4.6 Endocrine therapy resistance is not driven by *ESR1* hotspot mutations ........ 57
4.7 Barcoded clonal populations can be robustly distinguished ........ 58
4.8 Endocrine therapy resistant clones display persisting or fast proliferating phenotypes ........ 59
4.9 Differential pathway activation in endocrine therapy resistant clones ........ 62
4.10 Differential transcription factor activity in endocrine therapy resistant clones ........ 64
4.11 Differential kinase activity in endocrine therapy resistant clones ........ 65
4.12 T47D *orange* clones model cycling persister cells and depend on AKT signaling ........ 67
4.13 MCF7 LTED clones upregulate JAK-STAT signaling through STAT2 ........ 71
4.14 MCF7 TAMR and LTED clones upregulate NFκB signaling in part through p65 ........ 75
4.15 UPR activation reveals clonal vulnerability to Bortezomib ........ 76
4.16 PKC activation as shared clonal driver of endocrine therapy resistance ........ 78
4.17 Clinical significance of *in vitro* findings ........ 80
**5 Discussion ........ 85**
5.1 Endocrine therapy resistant clones display distinct phenotypes ........ 85
5.2 Divergent transcriptional re-wiring in endocrine therapy resistant clones ........ 86
5.3 Kinome-wide alterations in endocrine therapy resistant clones ........ 88
5.4 AKT signaling is important for T47D *orange* clones ........ 89
5.5 STAT2 mediated JAK-STAT signaling ........ 90
5.6 Private UPR activation and vulnerability to Bortezomib ........ 90
5.7 PKC activation as shared clonal driver of endocrine therapy resistance ........ 91
5.8 Clinical significance of *in vitro* findings ........ 93
5.9 Limitations of the study ........ 93
**6 Conclusion ........ 97**

7 Supplementary Figures ........................................................................................ 99
8 Supplementary Tables ......................................................................................... 107
9 Abbreviations ........................................................................................................ 113
10 References ........................................................................................................... 117
11 Contributions ....................................................................................................... 131
12 Acknowledgements ............................................................................................. 133

## List of Figures

**Figure 1:** Molecular subtypes of breast cancer. .................... 5
**Figure 2:** Estrogen signaling and endocrine therapy. .................... 7
**Figure 3:** Examples of endocrine therapy resistance mechanisms. .................... 9
**Figure 4:** Tumor heterogeneity and evolution. .................... 12
**Figure 5:** Clonal evolution and cancer stem cell model. .................... 13
**Figure 6:** Selection of pre-existing clones and acquired changes by persisting cells. .................... 14
**Figure 7:** Schematic overview of cellular barcoding and resistance acquisition. .................... 47
**Figure 8:** Proliferation of barcoded cell pools. .................... 48
**Figure 9:** Number of reads and identified barcodes. .................... 48
**Figure 10:** Barcode composition analysis of final cell pools. .................... 49
**Figure 11:** Project overview. .................... 51
**Figure 12:** Deconvolution overview. .................... 52
**Figure 13:** Barcode PCR and distribution of identified barcodes. .................... 53
**Figure 14:** Overview of isolated clones from selected T47D and MCF7 replicates. .................... 54
**Figure 15:** Barcode composition analysis of MCF7 *blue* TAMR_1 clones. .................... 55
**Figure 16:** Frequency of barcode reads in T47D and MCF7 replicates. .................... 56
**Figure 17:** Sequencing of *ESR1*. .................... 57
**Figure 18:** Clustering of clones based on the top 200 most variable genes. .................... 58
**Figure 19:** Proliferation of T47D and MCF7 clones. .................... 60
**Figure 20:** Pathway activity analysis of T47D clones. .................... 63
**Figure 21:** Pathway activity analysis of MCF7 clones. .................... 64
**Figure 22:** Activated and repressed transcription factor activities in a comparison of T47D and MCF7 clones. .................... 65
**Figure 23:** Clustering of clones based on the consistently identified phosphosites. .................... 66
**Figure 24:** Activated and repressed kinase and phosphatase activities in a comparison of T47D and MCF7 clones. .................... 67
**Figure 25:** Cycling and dying/dead cells for T47D *orange* clones. .................... 68
**Figure 26:** Short- and long-term persister cells. .................... 69
**Figure 27:** Activation of AKT1-3 in short-term estrogen deprived T47D *orange* clones. .................... 70
**Figure 28:** Phenotypic effects of AKT inhibition. .................... 70
**Figure 29:** Activation of STAT transcription factors in MCF7 LTED clones. .................... 72
**Figure 30:** *STAT* knockdown efficiencies. .................... 72
**Figure 31:** Phenotypic effects of *STAT2* knockdown. .................... 73
**Figure 32:** Phenotypic effects of *STAT1* knockdown. .................... 74
**Figure 33:** Ruxolitinib response curve. .................... 74
**Figure 34:** p65 activation in MCF7 LTED clones. .................... 75

**Figure 35:** *RELA* knockdown efficiency. ........................................................................... 75
**Figure 36:** Phenotypic effects of *RELA* knockdown. ............................................................ 76
**Figure 37:** UPR activation and sensitivity to Bortezomib. ................................................... 77
**Figure 38:** Phenotypic effects of Bortezomib treatment. ..................................................... 78
**Figure 39:** PKC activation and sensitivity to Sotrastaurin. .................................................. 79
**Figure 40:** Quantification of PKC inhibition by Sotrastaurin. ............................................... 79
**Figure 41:** Phenotypic effects of Sotrastaurin treatment. .................................................... 80
**Figure 42:** Pathway activities in breast cancer patients. ..................................................... 81
**Figure 43:** TF, UPR and PKC activities in breast cancer patients. ..................................... 82
**Figure 44:** Pathway activity profiles of individual patients with ER- disease from the CPTAC cohort. ............................................................................................................... 82
**Figure 45:** TF, UPR and PKC activity profiles of individual patients with ER- disease from the CPTAC cohort. ................................................................................................... 83

## List of Supplementary Figures

**Supplementary Figure 1:** Clonality dynamics in T47D LTED_2 cell pool. .......................... 99
**Supplementary Figure 2:** Integration site verification for T47D and MCF7 clones. .......... 100
**Supplementary Figure 3:** Proliferation of T47D *orange* and *green* +E2_5 clones under endocrine therapy. ........................................................................................... 101
**Supplementary Figure 4:** Overlap of detected phosphosites in biological replicates. ...... 103
**Supplementary Figure 5:** Quantification of AKT inhibition by MK-2206. .......................... 103
**Supplementary Figure 6:** IKK activities in MCF7 TAMR and LTED clones. ..................... 104
**Supplementary Figure 7:** Pathway and PKC activity profiles of individual patients with ER+ disease from the CPTAC cohort. ......................................................................... 104
**Supplementary Figure 8:** HER2 protein overexpression in T47D and MCF7 TAMR and LTED clones. .......................................................................................................... 105
**Supplementary Figure 9:** *IRF9* gene and IRF9 protein overexpression in MCF7 TAMR and LTED clones. ................................................................................................. 105

## List of Tables

**Table 1:** Description of used human breast cancer cell lines and culture media. ................ 26
**Table 2:** Generation of single cell clones used for this thesis. ............................................ 28
**Table 3:** Plated cell numbers. .............................................................................................. 29
**Table 4:** Click-iT reaction mix. ............................................................................................. 32
**Table 5:** PCR Master mix. .................................................................................................... 34
**Table 6:** Thermocycler program for barcode amplification. ................................................. 34

**Table 7:** Thermocycler program for barcode amplification to be sequenced on MiSeq. ...... 35
**Table 8:** Utilized NGS primers adapted from Bhang *et al.* 2015. ........................................ 35
**Table 9:** Primers used for integration site validation. .......................................................... 38
**Table 10:** Thermocycler program for integration site validation. ......................................... 38
**Table 11:** Thermocycler program for *ESR1* amplification. ................................................... 38
**Table 12:** Reverse transcription reaction mix. ...................................................................... 40
**Table 13:** Reverse transcription thermocycler program. ..................................................... 40
**Table 14:** RT-qPCR master mix. .......................................................................................... 40
**Table 15:** RT-qPCR thermocycler program. ......................................................................... 41

## List of Supplementary Tables

**Supplementary Table 1:** Sequences and color coding of barcodes representing isolated clones. ........................................................................................................................ 107
**Supplementary Table 2:** Genomic integration sites of individual barcodes. .................... 108
**Supplementary Table 3:** Gene set enrichment analysis of top 200 most variable genes for T47D. ....................................................................................................................... 108
**Supplementary Table 4:** Gene set enrichment analysis of top 200 most variable genes for MCF7. ...................................................................................................................... 109
**Supplementary Table 5:** Transcription factor activities for T47D TAMR and LTED clones. ........................................................................................................................ 109
**Supplementary Table 6:** Transcription factor activities for MCF7 TAMR and LTED clones. ........................................................................................................................ 109
**Supplementary Table 7:** Kinase and phosphatase activities for T47D TAMR and LTED clones. ........................................................................................................................ 110
**Supplementary Table 8:** Kinase and phosphatase activities for MCF7 TAMR and LTED clones. ........................................................................................................................ 111

Author: Lukas Beumers
Title: Clonal heterogeneity of endocrine therapy resistance in breast cancer
Faculty supervisor: Prof. Dr. Stefan Wiemann

## Summary

Breast cancer is the most common malignancy in women worldwide and roughly two-thirds of breast tumors are characterized by estrogen receptor (ER) α expression allowing for a targeted treatment approach by suppressing estrogen signaling. Clinically, suppression of estrogen signaling is achieved by Tamoxifen or Fulvestrant which prevent the receptor's activation or aromatase inhibitors which prevent the formation of (peripheral) estrogen. Unfortunately, relapses are observed in up to 41% of cases and tumor heterogeneity has recently been implicated in therapeutic resistance. To analyze endocrine therapy resistance on a clonal level, I used a previously developed barcoded *in vitro* model. There, two endocrine therapy sensitive ER+ breast cancer cell lines had been transduced with a barcode library. These cells had then been rendered resistant to Tamoxifen treatment (TAMR) or long-term estrogen deprivation (LTED), the latter to mimic clinically used aromatase inhibitors. Barcode analysis of complex cell pools suggested that endocrine therapy resistance arose either due to the selection of pre-existing clones or the rewiring of initially treatment-persisting cells.

In the next step, I isolated and characterized endocrine therapy sensitive and resistant clones from biological replicates using phenotypic assays, RNA-Sequencing and Mass spectrometry-based (phospho-)proteomics. Phenotypically, endocrine therapy resistant clones showed either weak or strong proliferative capacity. On pathway, transcription factor and kinase activity level, I observed heterogeneity between the cell lines utilized resembling inter-tumor heterogeneity and between endocrine therapy resistant clones isolated from each cell line resembling intra-tumor heterogeneity. Activation of the unfolded protein response (UPR) was a private event, which was only observed in a single TAMR population, and correlated with sensitivity to the proteasome inhibitor Bortezomib. Conversely, TAMR and LTED populations shared the activation of multiple protein kinase C (PKC) isoforms, however to different degrees. Treatment with the pan-PKC inhibitor Sotrastaurin preferentially reduced cellular viability of endocrine therapy resistant populations showing stronger PKC activation.

Finally, my *in vitro* findings were supported by clinical findings from the CPTAC-BRCA cohort. Generally, strong heterogeneity between individual patients was evident on pathway, transcription factor and kinase activity levels. After I had deconvoluted the cohort on a per patient basis, I identified patients with estrogen independent tumors showing UPR and PKC activation, closely resembling my *in vitro* models. Taken together, in the presented PhD thesis I could identify private and shared clonal endocrine therapy resistance drivers with clinical importance.

Autor: Lukas Beumers
Titel: Klonale Heterogenität der endokrinen Therapieresistenz bei Brustkrebs
Fakultätsgutachter: Prof. Dr. Stefan Wiemann

## Zusammenfassung

Brustkrebs ist die am häufigsten diagnostizierte Krebserkrankung bei Frauen weltweit. Rund zwei Drittel aller Brusttumore sind durch die Expression des Östrogenrezeptor (ER) α charakterisiert und können gezielt durch die Unterdrückung der Östrogen-Signalübertragung behandelt werden. Klinisch wird dafür die Rezeptoraktivierung durch Tamoxifen oder Fulvestrant blockiert, oder die Bildung von (peripherem) Östrogen wird durch Aromatasehemmer verhindert. Leider erleiden bis zu 41% der Patientinnen ein Rezidiv und Tumorheterogeneität kann die Resistenzentwicklung begünstigen. Um die endokrine Therapieresistenz auf klonaler Ebene zu untersuchen, habe ich zuvor entwickelte *in vitro* Modelle verwendet. Dort wurden zwei bisher unbehandelte ER+ Brustkrebszelllinien mit einer Barcode-Bibliothek transduziert. Diese Zellen wurden dann mit Tamoxifen oder Östrogenentzug behandelt, sodass die Zelllinien Resistenzen gegen die Tamoxifenbehandlung (TAMR) und den Östrogenentzug (LTED) entwickeln. Der Östrogenentzug *in vitro* spiegelte die Verwendung von Aromatasehemmern in der Klinik wider. Die Barcodeanalyse der komplexen Zellpools deutete darauf hin, dass Resistenz gegen die endokrine Therapie entweder durch die Selektion bereits vorhandener Klone zustande kam oder, dass Zellen, die ursprünglich die Therapie nur überlebt hatten, weitere Veränderungen zur Therapieresistenz erworben haben.

Im nächsten Schritt habe ich aus den komplexen Zellpools einzelne Klone isoliert und charakterisiert. Hierzu habe ich phänotypische Tests, RNA-Sequenzierung und Massenspektrometrie basierte (Phospho-)Proteomik verwendet. Generell beobachtete ich große Unterschiede auf phänotypischer und molekularer Ebene zwischen den verwendeten Zelllinien und einzelnen Klonen, welche von der gleichen therapieresistenten Zelllinie isoliert wurden. Respektiver Weise spiegeln diese Unterschiede Heterogenität zwischen Patientinnen (inter-tumor) und zwischen einzelnen Klonen von der gleichen Patientin (intra-tumor) wider.

Lediglich eine einzige TAMR-Population wies eine signifikante Aktivierung der ungefalteten Protein Antwort (UPR) auf. Diese UPR-Aktivierung korrelierte mit einer besonders starken Sensibilität gegenüber dem Proteasom-Inhibitor Bortezomib. Somit stellte die UPR-Aktivierung und Sensibilität gegenüber Bortezomib ein privates Ereignis dar. Andererseits waren mehrere TAMR und LTED-Populationen durch die signifikante Aktivierung verschiedener Isoformen der Proteinkinase C (PKC) charakterisiert. Allerdings war diese Aktivierung unterschiedlich stark ausgeprägt. Die Reduktion der zellulären Viabilität nach Behandlung mit dem PKC Inhibitor Sotrastaurin korrelierte hier mit der Stärke der PKC-Aktivierung in einzelnen therapieresistenten Klonen.

Abschließend konnte ich meine *in vitro* Ergebnisse durch klinische Ergebnisse aus der CPTAC-BRCA Studie unterstützen. Generell konnte ich starke Unterschiede zwischen den einzelnen Patientinnen auf molekularer Ebene sehen. Außerdem identifizierte ich einzelne Patientinnen, deren Tumorwachstum nicht von Östrogen angetrieben wurde, und deren Tumore eine signifikante Aktivierung der UPR und einzelner PKC Isoformen aufwiesen. Dadurch ähnelten diese Patientinnen sehr stark meinen *in vitro* Resistenz-Modellen.

Zusammengefasst konnte ich in meiner Promotion Treiber der endokrinen Therapieresistenz identifizieren, welche entweder privat und auf einzelne Klone beschränkt waren oder zwischen mehreren Klonen geteilt waren. Außerdem waren diese Veränderungen in einzelnen Patientinnen mit Östrogen-unabhängigen Tumoren stark ausgeprägt, was die klinische Relevanz meiner *in vitro* Ergebnisse unterstreicht.

# 1 Introduction

## 1.1 Breast cancer

Breast cancer is the most commonly diagnosed cancer and the leading cause of cancer death in women worldwide accounting for around 2.3 million new cancer cases and around 700,000 cancer deaths in 2020 [1]. Survival after breast cancer diagnosis depends on the molecular subtype of the tumor and the amount to which the tumor has spread through the body. Overall, the five-year relative survival rate is around 90%. For localized disease, the five-year relative survival rate is as high as 99% whereas only 28% of patients with distant metastatic disease survive past the five year mark [2].

### 1.1.1 Molecular subtypes

Breast cancer comprises a heterogeneous group of cancerous diseases of the breast with distinct subgroup-specific clinical behavior and histopathological characteristics [3, 4]. Subgrouping can be performed based on immunohistochemistry (IHC) stainings for three receptors and the proliferative marker Ki-67 [5]. Luminal tumors are characterized by expression of the estrogen (ER) and progesterone (PR) receptors, whereas HER2+ tumors show HER2 protein overexpression. Tumors lacking any of the receptors as determined by IHC staining are defined as triple negative breast cancer (Figure 1).

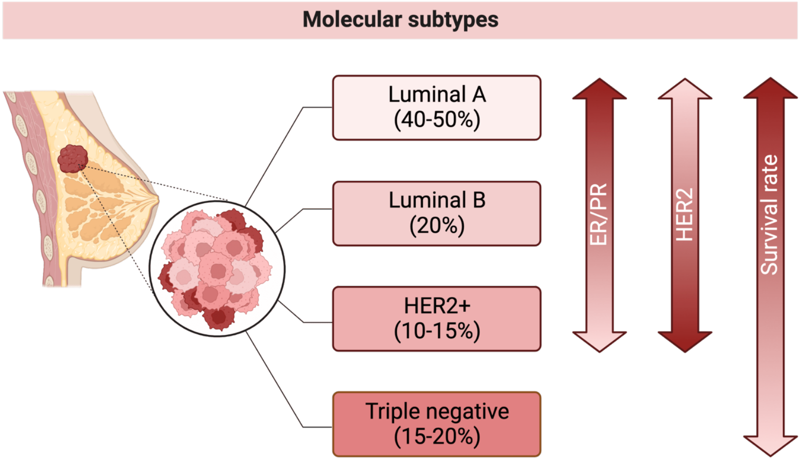


**Figure 1: Molecular subtypes of breast cancer.** Darker color denotes higher (ER/PR and HER2) receptor expression or higher survival rate. Adapted from Wong and Rebelo, 2012 [6] and 'Intrinsic and Molecular Subtypes of Breast Cancer' by BioRender.com (accessed: 09.05.2022). Created with BioRender.com.

Up to 70% of breast tumors stain positive for ER and PR by IHC [7-9]. These tumors can be further divided into luminal A and luminal B tumors. The majority of ER+ tumors belong to the luminal A subtype accounting for 40-50% of all breast cancer diagnoses. Around 20% of all breast tumors belong to the luminal B subtype [10-12]. While luminal A tumors are characterized by high PR positivity and low Ki-67 staining, luminal B tumors are characterized by low PR positivity and high Ki-67 staining. Luminal B tumors may additionally stain positive for HER2 [13, 14] and show a higher prevalence of *TP53* mutations compared to luminal A tumors [15]. In line, luminal A tumors have a better prognosis with relative five-year survival rates of about 95% while luminal B tumors have relative five-year survival rates of about 90% [16, 17]. The backbone for treating either luminal subtype is endocrine therapy impinging on estrogen signaling. Chemotherapy is additionally given to patients with luminal B tumors because of the increased proliferative rate of their tumors [18-20].

*ERRB2* gene (as determined by fluorescence in situ hybridization [FISH] [21]) and/or HER2 protein overexpression is characteristic for around 10-15% of breast tumors [10-12]. The backbone of clinically targeting HER2+ tumors has been Trastuzumab [22-25], a humanized monoclonal antibody. Trastuzumab may be used as single agent, in combination with chemotherapy or in a triple therapy containing a second humanized monoclonal antibody, Pertuzumab [26]. Cytotoxic payloads may additionally be coupled by a linker to Trastuzumab. The antibody-drug conjugates (ADCs) Trastuzumab Emtansine and Trastuzumab Deruxtecan combine the HER2 targeting antibody with a tubulin and topoisomerase I inhibitor, respectively. Both are suitable therapeutic options in patients with advanced HER2+ breast tumors showing activity after previous Trastuzumab treatment [27-30]. Besides monoclonal antibodies with(out) cytotoxic payload, HER2+ tumors can also be effectively therapeutically challenged by HER2-targeting tyrosine kinase inhibitors (TKIs). Lapatinib [31, 32], Neratinib [33] and Tucatinib [34] are notable examples of HER2-targeting TKIs showing improved patient outcome in combination or sequence with Trastuzumab and/or chemotherapy. Overall, targeted therapeutic approaches translate to five-year survival rates of around 85% [16, 35].

Breast tumors lacking any actionable target in form of ER (and PR) or HER2 are labeled as triple negative (TNBC) and account for 15-20% of breast tumors [36, 37]. Further classification into four subtypes can be performed according to Lehmann *et al.* [38]. Surgery and chemotherapy remain the main therapeutic options [39] resulting in five-year survival rates of about 77% [16, 35]. Immunotherapy with PD-L1 or PD-1 targeting antibodies alone or in combination with standard of care chemotherapy for metastatic TNBC showed improvements in clinical outcome [40]. The membrane antigen TROP-2 has attracted attention in different tumor types including TNBC as higher expression has been shown to be significantly associated with worse overall survival in solid tumors [41]. Sacituzumab Govitecan, an ADC combining a humanized anti-Trop-2 monoclonal antibody linked with a topoisomerase I inhibitor showed significantly improved overall survival in patients with metastatic TNBC over physician’s choice of single agent chemotherapy [42]. Utilizing TROP-2 as target to deliver cytotoxic payloads may potentially improve treatment options for patients with TNBC.

### 1.1.2 Endocrine therapy

Around 60-70% of breast tumors are characterized by the expression of the hormone receptors ER and PR. In these tumors, estrogen signaling drives the uncontrolled proliferation of the cells. Briefly, estrogen binds to its receptor, which then translocates into the nucleus and forms dimers. The ER dimers then bind to estrogen response elements (ERE) and recruit coactivator complexes ultimately activating the expression of estrogen target genes [43, 44]. The backbone of targeting estrogen signaling are aromatase inhibitors (AIs) such as Letrozole, selective ER modulators (SERMs) such as Tamoxifen and selective ER degraders (SERDs) such as Fulvestrant (Figure 2). Tamoxifen and AIs were applied as first line therapy for five years. Different trials have by now underlined clinical benefits when continuing Tamoxifen therapy for ten instead of five years [45, 46] and five years of either Tamoxifen or aromatase inhibitor followed by five years of aromatase inhibitor over treatment stop after five years [47].

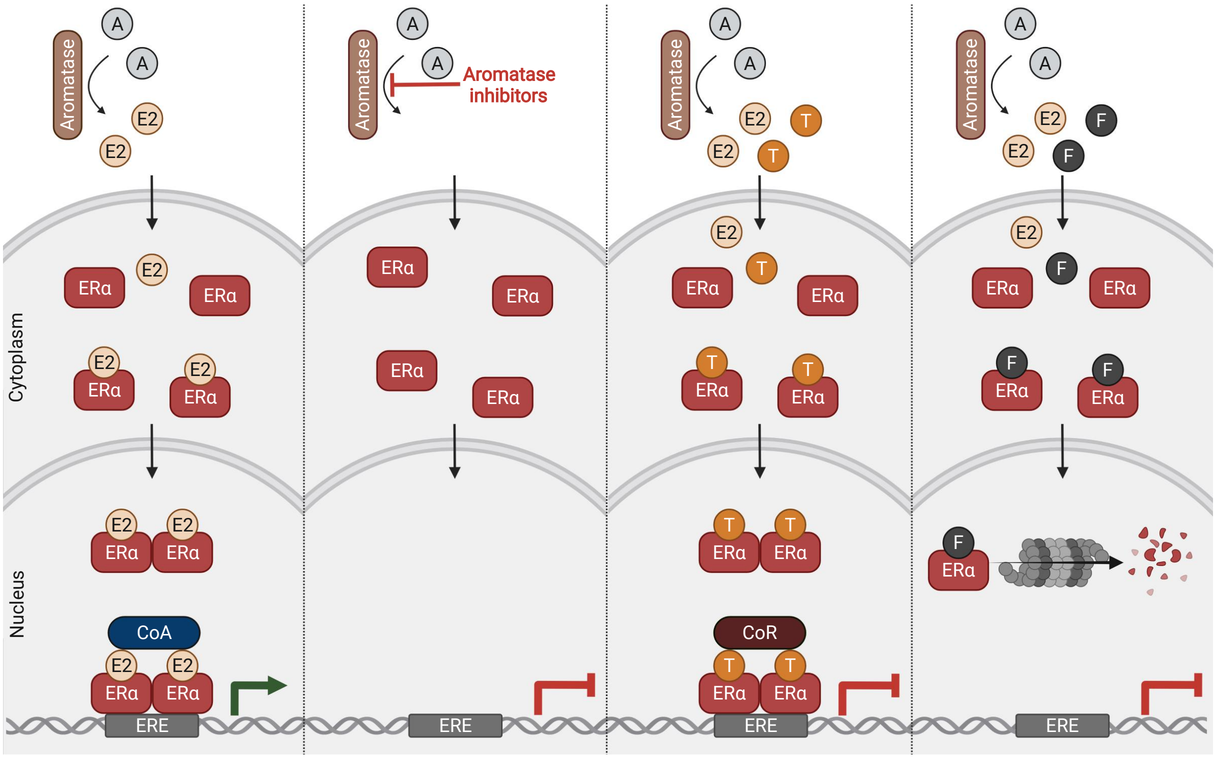


**Figure 2: Estrogen signaling and endocrine therapy.** A: androgen. E2: estrogen. ERα: estrogen receptor α. CoA: co-activator complex. ERE: estrogen response element. T: Tamoxifen. CoR: co-repressor complex. F: Fulvestrant. Adapted from Hanker, *Cancer Cell* 2020 [48] and 'Estrogen Receptor Signaling' by BioRender.com (accessed: 09.05.2022). Created with BioRender.com.

The aromatase enzyme catalyzes the last and rate-limiting step in the synthesis of estrogens from androgens (androstenedione and testosterone) to estrogens (estrone and estradiol). This enzyme is expressed in the granulosa cells of the ovaries and peripheral tissues such as fatty tissue. The majority of estrogen is produced in the ovaries in premenopausal women whereas postmenopausal estrogen production exclusively relies on the peripheral aromatase enzyme [49-51]. Three highly specific third-generation aromatase inhibitors (AIs) are clinically used to treat postmenopausal patients with ER+ disease. Anastrozole and Letrozole are non-steroidal reversible inhibitors while Exemestane is a steroidal irreversible inhibitor inducing suicide inhibition [52]. Respectively, Anastrozole [53], Letrozole [54] and Exemestane [55] reduce the peripheral aromatase's enzymatic activity by 97%, 99% and 98% thereby preventing the downstream activation of ER.

As a SERM, Tamoxifen possesses estrogenic and anti-estrogenic effects. Tamoxifen has estrogenic effects in the bone where it increases bone mineral density [56, 57]. Given its anti-estrogenic effects in the breast, Tamoxifen was approved in the 1970s for the treatment of ER+ breast cancer [58]. Other SERMs have been developed but Tamoxifen remains the first choice in the clinic [59]. Tamoxifen itself is a prodrug with limited affinity to the ER [60]. Cytochrome P450 enzymes metabolize Tamoxifen to biologically more active compounds such as 4-hydroxytamoxifen (4-OHT) [61-63] which has an 80% higher affinity for ER than estradiol itself [60]. Mechanistically, Tamoxifen-bound ER recruits co-repressor complexes which leads to the transcriptional inhibition of ER-dependent genes [64, 65].

The only SERD with current approval for clinical use is Fulvestrant but other SERDs are being developed as novel therapeutics. Fulvestrant has a similar affinity for the ER as estradiol [66, 67]. In contrast to estradiol, treatment with Fulvestrant abrogates the expression of ER-regulated genes [68-72]. Fulvestrant displaces estrogen in a competitive manner from the ER and induces a conformation of the receptor preventing transcriptional activation [73]. Binding of this drug with its bulky sidechain leads to the surface exposure of a large hydrophobic patch from the ligand-binding domain (LBD) and prevents the association between the transactivation helix 12 and the rest of the LBD thereby mimicking misfolded or denatured proteins [74]. Another proposed mechanism of Fulvestrant's action ascribes increased ER turnover as consequence of low chromatin accessibility and decreased intra-nuclear mobility [75]. Both proposed mechanisms result in the proteasomal degradation of ER in the nucleus [76-78]. Fulvestrant is very effective in patients, also after progression on prior aromatase inhibitor or Tamoxifen treatment [79-82], but Fulvestrant's limited bioavailability requires it to be given by injection. To circumvent the necessity of injection, newer generation bioavailable SERDs are being developed [83].

### 1.1.3 Acquired endocrine therapy resistance

Endocrine therapy has undoubtedly transformed and saved the lives of countless patients. Unfortunately, up to 41% of patients which originally presented with late stage disease relapse within 10 years of treatment start [84]. Therapy resistance and tumor progression remain a major clinical challenge. Endocrine therapy resistance may arise due to alterations in ER signaling, cell cycle deregulation and upregulation of other growth promoting pathways as shown in Figure 3 [85-92] and described in detail below. Further resistance mechanisms include interactions with the tumor microenvironment, epigenetics and cancer stem cells as reviewed elsewhere [48].

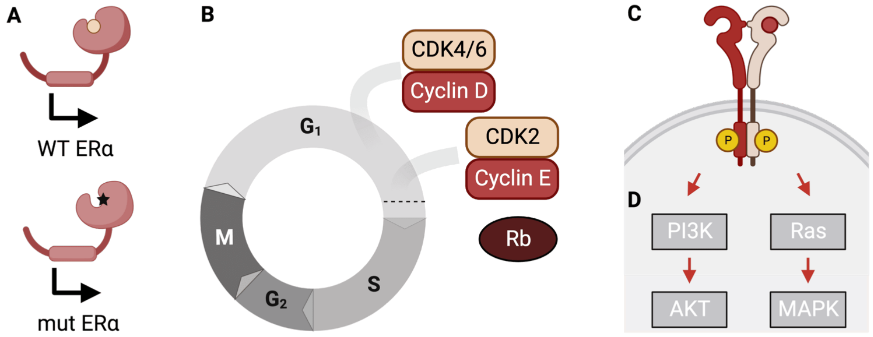


**Figure 3: Examples of endocrine therapy resistance mechanisms. A** *ESR1* mutations (star) commonly affect the ligand-binding domain rendering ER active in the absence of estrogen (beige circle) and drive gene expression (indicated by arrow). Adapted from Katzenellenbogen, *Nature Reviews Cancer* 2018 [93]. **B** Common cell cycle alterations in cancer. Rb: Retinoblastoma protein. Adapted from Asghar, *Nature Reviews Drug Discovery* 2015 [94] and 'Cell Cycle Deregulation in Cancer' by BioRender.com (accessed: 09.05.2022). **C** and **D** Upregulation of RTKs (**C**) and downstream pathways (**D**). Adapted from Hanker, *Cancer Cell* 2020 [48]. Created with BioRender.com.

### 1.1.3.1 *ESR1* mutations

ER is the effector of estrogen signaling and is the main therapeutic target in most breast tumors as illustrated above. Mutations in the *ESR1* LBD are rarely found in primary treatment-naïve tumors (0-7%) but are strongly enriched in advanced/metastatic tumors (12-55%) [95-100], especially after aromatase inhibitor treatment [101]. Several hotspot mutations affecting codons 536 to 538 stabilize helix 12 in the receptor's active conformation without requiring the binding of estrogen and drive ligand-independent gene transcription [93, 102]. A recent clinical meta-study could highlight *ESR1* mutations as predictor of resistance to aromatase inhibitor containing treatment regimens but failed to implicate *ESR1* mutations as driver of resistance to Fulvestrant containing regiments [103].

In addition to *ESR1* mutations, genetic alterations of other transcription factors such as *MYC* were found in endocrine therapy resistant tumors. These alterations were mutually exclusive with *ESR1* and MAPK alterations pointing to an alternative route of resistance [104]. In accordance with the implication of different transcription factors in therapy resistance, we recently identified Activating Transcription Factor 3 (ATF3) as novel driver of endocrine therapy resistance [105].

### 1.1.3.2 Cell cycle deregulation

Cell cycle progression is tightly controlled in normal cells whereas abnormal cell cycle progression is a cancerous hallmark [106-108]. Deregulation of key players involved in regulating the progression through the restriction point are commonly found in (breast) cancers. In ER+

disease, *CCND1* (gene encoding for Cyclin D1) amplification and Cyclin D1 overexpression correlated with worse prognosis such as increased risk of relapse and shortened overall survival [109, 110]. *RB1* was also frequently found to be mutated in metastatic ER+ breast tumors [111]. For patients receiving Tamoxifen, loss of Rb protein was additionally associated with worse overall survival [112]. Upregulation of either CDK6 or cyclin E2 in patients treated with Fulvestrant correlated with worse progression free survival [113, 114]. Further, estrogen-independent derepression of E2F can be achieved through CDK4 upregulation and an E2F activation gene signature correlated with reduced response to aromatase inhibitors [115]. Combination of endocrine therapy, mainly the non-steroidal aromatase inhibitors Letrozole and Anastrozole or Fulvestrant, and CDK4/6 inhibition by Ribociclib, Abemaciclib or Palbociclib, respectively, has been investigated in the MONALEESA-2/3/7 [116-120], MONARCH-2/3 [121-124] and PALOMA-1/2/3 trials [125-131]. Combination of endocrine therapy and CDK4/6 inhibition showed significantly improved progression-free as well as overall survival and a longer time to chemotherapy in patients with advanced ER+ HER2- disease [132].

### 1.1.3.3 Alterations in receptor tyrosine kinases

Receptor tyrosine kinases (RTKs) translate extracellular signals in form of growth factors intracellularly. Aberrant RTK signaling is among common alterations observed in different cancers entities and can be exploited by therapeutically targeting e.g. EGFR and HER2 [133, 134].

*Fibroblast Growth Factor Receptor 1* (*FGFR1*) amplification has been demonstrated as predictor of poor outcome in patients with ER+ disease [135] and correlates with higher Ki-67 staining during therapy [136, 137]. Its gene amplification further mediates resistance to endocrine treatment alone and/or in combination with CDK4/6 inhibition [136, 138]. In endocrine therapy resistant metastatic breast tumors, *FGFR4*, a gene encoding for another member of the FGFR family, was additionally found to be overexpressed or mutated [139]. An early clinical trial showed encouraging results using a FGFR inhibitor in combination with Anastrozole or Letrozole [140].

Acquired antiestrogen resistance has been shown to be mediated through the upregulation of HER2 [141-143]. *ERBB2* mutations were further found in around 5% of endocrine therapy resistant, non-HER2 amplified tumors [104]. HER2 upregulation and mutations in the (transmembrane and) kinase domain(s) lead to the hyperactivity of the PI3K/AKT/mTOR axis or MAPK pathway and ultimately antiestrogen resistance [144, 145]. Hotspot mutations in *ESR1* and *ERBB2* were mutually exclusive pointing to alternative routes of acquired resistance [104, 145]. Dual blockage by Fulvestrant (targeting ER) and Neratinib (targeting HER2 and EGFR) appeared to be a suitable therapeutic option in ER+/HER2$^{\text{mutated}}$ disease [146].

#### 1.1.3.4 Upregulation of growth promoting pathways

*PIK3CA*, *PTEN* and *AKT1* encode for key players of the PI3K/AKT/mTOR pathway and are among the most frequently mutated genes in ER+ breast cancer [10, 104, 111]. Aberrant activation of phosphoinositide 3-kinase (PI3K) [147-149] and downstream protein kinase B (AKT) [150-152] induce estrogen independence and resistance towards Tamoxifen and Fulvestrant. The SOLAR-1 phase trial showed significantly improved progression free survival when adding Alpelisib, an α subunit specific PI3K inhibitor [153], to Fulvestrant for patients with *PIK3CA* mutated tumors progressing on or after a non-steroidal aromatase inhibitor [154]. According to the BYLieve study [155], Alpelisib is also a suitable therapeutic option for patients with *PIK3CA* mutated tumors progressing on or after endocrine therapy paired with a CDK4/6 inhibitor. In accordance with mammalian target of rapamycin (mTOR) activation downstream of AKT, the BOLERO-2 [156, 157], GINECO [158] and MANTA [159] trials showed improved progression free survival for the addition of the mTOR inhibitor Everolimus to the steroidal aromatase inhibitor Exemestane, Tamoxifen and Fulvestrant for patients progressing on or after a prior non-steroidal aromatase inhibitor.

Besides the PI3K/AKT/mTOR axis, the mitogen-activated protein kinase (MAPK) pathway is a growth promoting pathway commonly deregulated in malignant diseases [160]. Different alterations in this phosphorylation cascade have been identified in patients with endocrine therapy resistant tumors and were associated with worse outcome. Loss of neurofibromin 1 (NF1) activity resulted in constitutive RAS activity and was associated with endocrine therapy resistance [161-163]. Different members of the *RAS* superfamily as well as *ERBB3*, *BRAF* and *MAP2K1* were found to be frequently mutated in patients relapsing or progressing on endocrine therapy [104, 164]. Mechanistically, MAP kinases phosphorylate Ser104, Ser106 or Ser118 in the AF-1 region of ER leading to the receptor's ligand-independent activation and agonistic effects of Tamoxifen [165-167]. Steroidal antiestrogens such as Fulvestrant also inhibit the receptor's AF-1 region preventing Ser118 phosphorylation through the MAPK cascade and resulting ER activation [165]. In Tamoxifen-resistant tumors with sustained estrogen signaling as result of MAPK induced ER phosphorylation, Fulvestrant may be a suitable choice for second-line endocrine therapy. In conclusion, patients with ER+ disease harboring activating alterations in the MAPK axis may benefit from combining endocrine therapy with MAPK cascade inhibitors which are routinely used in treating other malignancies [168-170].

## 1.2 Tumor heterogeneity

Different tumor types including breast tumors demonstrate a variety of histopathological, genetic and epigenetic differences. This intertumor heterogeneity (Figure 4 A), describing such variations between tumors from different patients, results in very different clinical outcomes [171-173]. Individual tumors may also contain genetically, epigenetically or phenotypically distinct tumor populations and the tumor composition may change over time (Figure 4 B and C). This phenomenon is described as intratumor heterogeneity [171-173]. Nowell's clonal evolution and the cancer stem cell (CSC) model offer explanations for how heterogeneous tumors arise and progress.

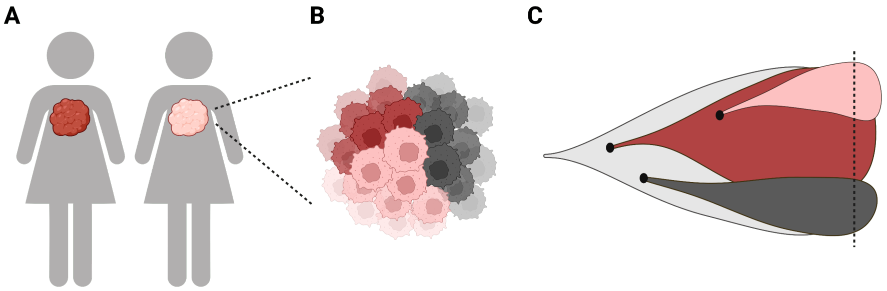


**Figure 4: Tumor heterogeneity and evolution. A** and **B** Inter- (**A**) and intra- (**B**) tumor heterogeneity. Adapted from Burrell, *Nature* 2013 [174]. **C** Tumors may evolve over time. New clones can emerge after a fitness advantage was gained exemplarily through a mutation (black dots). The dashed line indicates the time point at which the heterogeneity of the tumor in **B** was investigated. Adapted from Koren, *Molecular Cell* 2015 [175] and 'Cancer Evolution' by BioRender.com (accessed: 09.05.2022). Created with BioRender.com.

### 1.2.1 Clonal evolution and cancer stem cells

Intra- and intertumor heterogeneity may arise due to clonal evolution, from a subset of cancer stem cells or a combination of clonal evolution and (de)differentiation (Figure 5). According to Nowell's hypothesis [176], (pre-)cancerous cells acquire random mutations which are subsequently selected for increased fitness. Consequently, tumor development and progression are driven by Darwinian evolution. Cancer cells with similar fitness may give rise to populations acquiring mutations independently of each other. These populations then mutually form the tumor mass. The cancer stem cell model attributes tumor development, its progression and arising populations to a small subset of stem cell-like cells. Similar to stem cells, their cancerous equivalents are postulated to be capable of self-renewal and differentiation whereby both, cancer and normal stem cells, utilize similar signaling pathways [177, 178]. These models are not mutually exclusive. Depending on the tumor entity, stronger influence of one mechanism over the other however is likely [179].

Evidence for clonal evolution as driver of tumor heterogeneity comes from regional sequencing of tumors. Intratumor heterogeneity in non-small-cell lung cancer (NSCLC) was discovered in copy number alterations, translocations and APOBEC-associated mutations [180]. Mutations were additionally found to be present in all clones or only detected in a subset of clones. Exemplarily, mutations in *EGFR* were almost exclusively early and detected in all clones whereas mutations in *PIK3CA* were detected in a subset of clones and occurred later during tumor progression [181].

Besides clonal evolution, heterogeneity may be explained by the stem cell hypothesis. This model proposes CSCs on top of a cellular hierarchy which gives rise to different populations forming the final tumor mass [182-184]. Breast cancer stem cells (BCSCs) can be characterized by a $CD44^{+}/CD24^{-/low}/Lin^{-}$ phenotype. As little as 100 $CD44^{+}/CD24^{-/low}/Lin^{-}$ cells were able to form tumors in immunodeficient mice whereas 10,000 cells with reciprocal markers failed to do so. The tumors formed by the $CD44^{+}/CD24^{-/low}/Lin^{-}$ cells contained BCSCs and other phenotypic populations [185]. The cancer stem cell model however does not imply an exclusively unidirectional differentiation of CSCs since various findings show that CSCs can arise from non-CSCs through epithelial-mesenchymal transition [186, 187] or spontaneous conversion of non-CSCs to CSCs [188, 189].

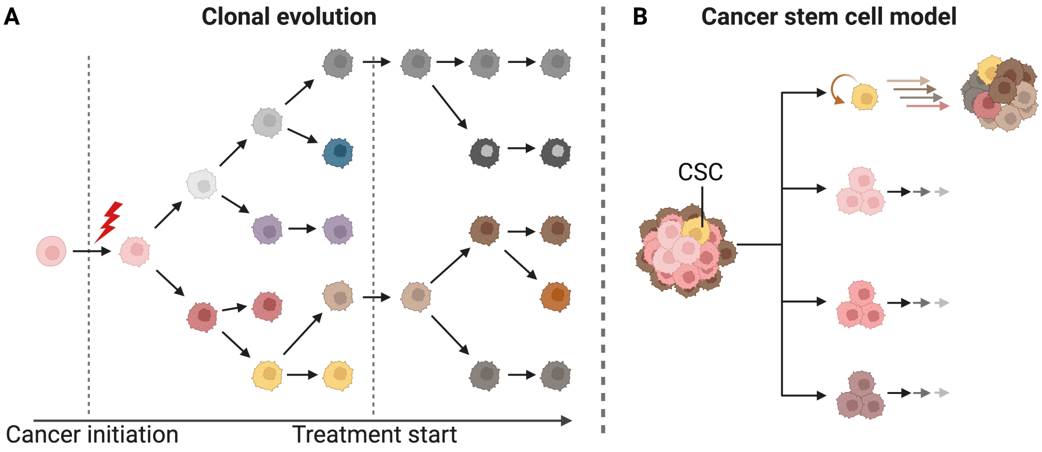


**Figure 5: Clonal evolution and cancer stem cell model. A** Adapted from Nowell, *Science* 1976 [176] and 'Cancer Evolution with Treatment' by BioRender.com (accessed: 09.05.2022). **B** Adapted from Wang, *EMBO Reports* 2015 [190] and 'Stochastic vs Cancer Stem Cell Models' by BioRender.com (accessed: 09.05.2022). Created with BioRender.com.

### 1.2.2 Selection of pre-existing clones and resistance acquisition

Tumor heterogeneity has been associated with worse disease outcome [172] and allows for treatment resistance by two general routes. In the initial treatment sensitive tumor bulk, a pre-existing minor clone may be intrinsically resistant to treatment. Upon treatment, most treatment sensitive tumor cells are eliminated, resulting in a drastic overall response. The treatment resistant clone is not affected by the therapy and may grow out to eventually form a treatment resistant refractory tumor (Figure 6 A). Alternatively, some clones in the tumor bulk may survive/persist the initial treatment in the absence of evident proliferation. These cells then acquire *de novo* genetic or epigenetic alterations mediating treatment resistance and finally the outgrowth of a refractory tumor (Figure 6 B). These two phenomena appear to not be mutually exclusive but can be (to some extend) implicated in response to the same treatment. An illustrative example comes from a study investigating resistance against the EGFR inhibitor Gefitinib in a NSCLC model [191]. Resistance either arose quickly due to the selection of a population harboring the EGFR p.T790M gatekeeper mutation or later due to resistance acquisition after a drug-induced dormant state. In the later stage, only 5 out of 16 clones harbored the EGFR p.T790M gatekeeper mutation [191].

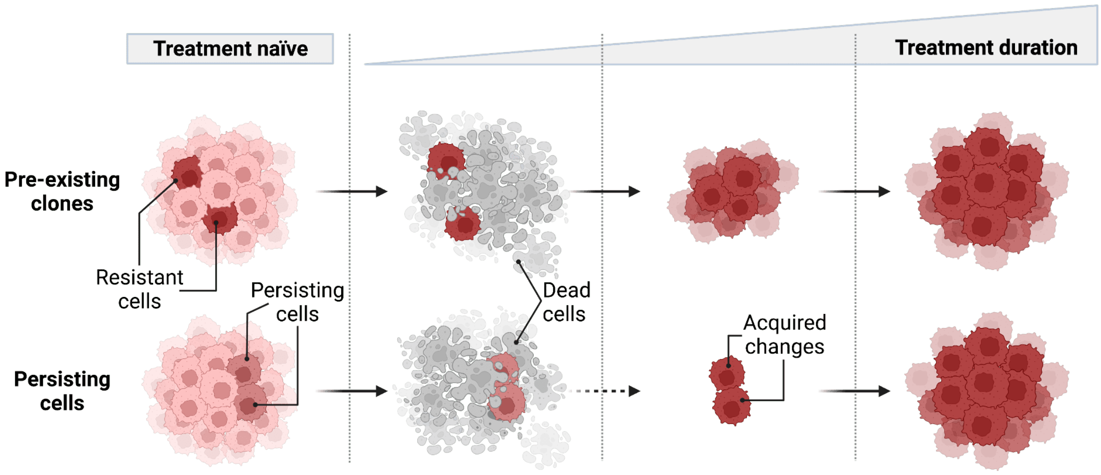


**Figure 6: Selection of pre-existing clones and acquired changes by persisting cells.** Adapted from Dagogo-Jack and Shaw, *Nature Reviews Clinical Oncology* 2018 [172] and 'Intrinsic and Acquired Drug Resistance' by BioRender.com (accessed: 09.05.2022). Created with BioRender.com.

Two case reports of NSCLC patients have explicitly tied tumor relapse to the selection of a minor treatment-resistant population already present in the treatment-naïve tumor. Shaw *et al.* [192] presented a patient relapsing on the ALK/ROS1 inhibitor Crizotinib due to the selection of a minor clone present in the treatment-naïve tumor harboring the ALK p.C1156Y Crizotinib resistance mutation [193]. In a second case report by Li *et al.* [194], relapsed disease was due to selection of a clone with BRAF p.V600E mutation. *In vitro*, populations emerging from as little as a single cell can be traced by cellular barcoding approaches. Barcoding generally refers to the lentiviral integration of a cassette containing a unique DNA-sequence ('barcode'), a marker to select for successful integration events and can contain additional elements such as fluorescent labels. An *in vitro* barcoding study revealed that resistance to EGFR inhibition in NSCLC cells and resistance to ABL1 inhibition in chronic myeloid leukemia (CML) cells was largely resulting from the selection of resistant pre-existent clones [195].

Accumulating evidence highlights that a different population of tumor cells may enter a dormant or persistent cell state escaping treatment stress [196]. This dormant state is characterized by negligible growth and has been associated with stem cell-like features in different cancer entities, such as gastric [197], breast [198], colorectal [199], ovarian [200] and lung cancer [201]. A pool of dormant cells may represent a reservoir for the emergence of heterogeneous acquired resistance mechanisms. A dormant cell state with >100-fold reduced sensitivity to Erlotinib has been shown to be maintained by insulin-like growth factor 1 receptor signaling and an altered chromatin state induced by the histone demethylase KDM5A [202].

Under prolonged treatment, initially persisting cells developed full resistance via heterogeneous routes, including the acquisition of the EGFR p.T790M mutation in some cases [203]. Drug persisting cells have not only been identified in NSCLC but in a multitude of cancer entities, exemplarily melanoma [204-206], breast [204, 206] and ovarian cancer [204].

### 1.2.3 Tumor heterogeneity in breast cancer and endocrine therapy resistance

Breast cancer is a heterogeneous disease and clinical subtyping is performed according to the ER, PR, *ERRB2* (by FISH)/HER2 and Ki-67 status by IHC as described above. More in-depth analyses of intra- and inter-tumor heterogeneity has recently been enabled by the broad-based adoption of next-generation sequencing (NGS). A series of analyses revealed complex genetic intra- and inter-tumor heterogeneity in primary breast tumors [207-210] as well as between primary tumors and disseminated cells in form of circulating tumor cells (CTCs) [211, 212] and metastases [213, 214].

Tumor heterogeneity has been shown to be involved in the progression of premalignant to malignant conditions as well [215]. Ductal carcinoma *in situ* (DCIS) is considered a pre-cancerous condition of the breast, in which the basal membrane surrounding the milk ducts has not yet been penetrated. According to a study by Casasent *et al.* [210], most mutations and copy number alterations (CNA) were already present prior to invasion in the pre-cancerous DCIS part and in at least half of all cases, two or more clones co-migrated into adjacent tissue establishing the invasive carcinomas. Invasion of the basal membrane is a very early form of cancerous progression, whereas in later stages tumor cells disseminate also to other parts of the body. Disseminated tumor cells or CTCs embody a further step in tumor progression and showed clear differences to their primary matched tumors on CNA [212] and mutational level [211]. The final and deadliest stage of tumor evolution is metastatic disease [216]. Two studies involving patients with metastatic ER+ [213] and triple negative disease [214], respectively, found individual mutational allele frequencies to be decreased, constant, enriched or newly detected in the metastatic sample compared to the primary tumor pointing to large changes in clonal tumor makeup with tumor progression. Multiple metastases could be seeded either by a single common metastatic precursor or arise from different clones present within the primary tumor [217]. Individual metastatic sites within the same patient were additionally shown to contain a different makeup of genetically defined clones and single metastatic lesions were not able to capture the breadth of the disease [218, 219]. A study by De Mattos-Arruda *et al.* [220] further pointed to the propagation and evolution of metastases as communities of different clones. One patient with ER+ disease relapsed within three years after diagnosis on subsequent chemo- and endocrine therapy with all metastases sharing *ESR1* hotspot activating and *PTEN* truncating mutations. The metastases further differentiated into more diverse subgroups, two of which were characterized by *BRCA2* and *PIK3CA* mutations, respectively. In a second patient with ER+ positive disease initially treated with subsequent chemo- and endocrine therapy, coexisting clusters of ER+ and ER- lesions were detected. The ER+ subgroup of metastases was characterized by an *ESR1* activating hotspot mutation. Metastases from these two patients clearly highlight the genomic heterogeneity of endocrine therapy resistant disease. Genomic heterogeneity may further be accompanied by transcriptional heterogeneity. Within a single patient with TNBC disease, over 50 pathologically diverse cell clusters from the primary tumor and lymph node metastasis have been identified by laser capture microdissection and subsequent DNA- and RNA-seq [221]. This study further highlighted transcriptionally defined clones to be more complex than genomically defined clones.

Formation of metastases generally is a lengthy process and tumor heterogeneity may change during the process. Previous studies have already highlighted profound changes to also be induced by short-term neoadjuvant endocrine therapy [222-224] but considered the bulk of tumor masses without assessing the tumors' clonal makeup. Miller *et al.* [225] utilized WGS to interfere this clonal heterogeneity in response to four months neoadjuvant aromatase inhibitor treatment in 18 patients. Overall, four genomic patterns were observed: (i) 'collision' tumors consisting of two intermingled but genomically separate tumors, treatment stable tumors with (ii) simple or (iii) complex genomic makeup and (iv) treatment dynamic tumors with complex clonal makeup. Most tumors were treatment dynamic with a complex clonal makeup (15/19, 79%). The tumors in this group were clonally heterogeneous and their clonal makeup changed during the four months treatment period.

Lower genetic tumor heterogeneity before applied chemotherapy has further been shown to be significantly associated with pathological clinical response across ER+, HER2+ and TNBC breast tumors. Conversely, intra-tumor genetic diversity did not change by the applied therapy in tumors showing only partial or no response [226]. Transcriptional heterogeneity in treatment-naïve ER+ breast tumors [227] may additionally prime the tumor for therapy resistance. Mechanistically, transcriptional tumor heterogeneity may induce fluctuations in gene expression and accordingly protein levels as well as phenotypes on population level [228]. It has been postulated that this gene expression 'noise' prepares the cancer population but not individual cells to defend against therapies [229, 230]. Exemplarily, Griffiths *et al.* [231] analyzed the evolutionary trajectories of early-stage breast tumors receiving Letrozole alone or in combination with the CDK4/6 inhibitor Ribociclib utilizing scRNA-Seq. Across patients and treatments, differences in clonal diversity were evident. However, also similarities between different patients were observed, such as estrogen-independent growth mediated through increased JNK signaling in patients treated with Letrozole and Ribociclib.

In concordance with findings for resistance acquisition of NSCLC to EGFR inhibitor treatment presented above, *in vitro* models for breast cancer also suggest involvement of pre-existing clones [232] and re-wiring of persisting cells [233] in endocrine therapy resistance.

Utilizing a barcoding approach, Hinohara *et al.* [232] on the one hand highlighted endocrine therapy resistance to arise from selecting pre-existing clones. Hong *et al.* [233] on the other hand identified a rare population of pre-adapted treatment-naïve cells mimicking a cellular state generally observed during short-term estrogen deprivation. These pre-adapted cells were poised to survive short-term estrogen deprivation and required further genetic and epigenetic alterations to become fully resistant.

Taken together, these findings suggest that different mechanisms drive endocrine therapy resistance and that resistant clones can either be initially present in the treatment naïve populations or that resistance is acquired by clones during the course of the applied therapy. Common genetic determinants of endocrine therapy resistance such as *ESR1* mutations can be found in individual tumor clones.

Studies focusing on transcriptional clonal alterations are increasing while phosphoproteomic analysis of whole tumor samples is gaining strong attention and may help to improve precision cancer therapy [234, 235]. Further studies are required to improve our understanding of how transcriptional and phosphoproteomic alterations drive clonal endocrine therapy resistance and how these alterations can be therapeutically exploited.

## 2 Aims of the thesis

Endocrine therapy resistance remains an urgent clinical problem affecting up to 41% of patients diagnosed with late stage disease [84]. Recent findings highlighted the importance of tumor heterogeneity in therapy resistance. To disentangle clonal endocrine therapy resistance *in vitro*, Dr. Simone Borgoni, a former PhD student in our lab, generated barcoded endocrine therapy resistant cell line models. Therefore, the ER+ breast cancer cell lines T47D and MCF7 were transduced with the ClonTracer [195] library to mark individual cells with a barcode. The cells were then rendered resistant to Tamoxifen treatment (TAMR) and long-term estrogen deprivation (LTED), the latter mimicking clinically used aromatase inhibitors. As comparison, cells were also kept under normal growth condition (+E2). Barcode analysis of the bulk treatment sensitive and resistant replicates highlighted that Tamoxifen resistance was achieved either through the selection of pre-existing populations (T47D) or through resistance acquisition of initially treatment persisting cells (MCF7) [236].

Based on these initial findings, the aims of my PhD thesis were:

(i) to illustrate the intra- and inter-tumor heterogeneity in endocrine therapy resistant breast cancer cells on a clonal level and
(ii) to validate private and shared clonal endocrine therapy resistance drivers with clinical importance.

Given the breadth of scientific literature on genetically defined heterogeneity and resistance mechanisms, the focus of this thesis was on non-genetically defined resistance mechanisms. To (i) illustrate tumor heterogeneity and (ii) unravel resistance drivers, I applied the following approaches:

- Isolation of multiple single cell clones from the +E2, TAMR and LTED replicates with distinct barcodes as foundation for further experiments
- Profiling of isolated clones for stable lineages or potential lineage change
- Analysis of phenotypes (mainly by proliferation assays), gene expression profiles and activation of kinases (through phospho-proteomics analysis)
- Identification of private (only for a single clone) and shared (across multiple clones) potential resistance drivers
- Experimental testing of candidates for their potential to drive endocrine therapy resistance
- Assessment of molecular mechanisms (proliferation, cell viability, percentage of cycling and dying/dead cells) after gene knockdown and inhibitor treatment
- Correlation of *in vitro* data with data from the CPTAC-BRCA cohort to add clinical relevance to the *in vitro* findings

# 3 Materials and Methods

## 3.1 Materials

### 3.1.1 Instruments

| Instrument | Company |
|---|---|
| 2200 TapeStation System | Agilent Technologies |
| Axiovert 40 CFL | Zeiss |
| Bacterial incubator (37°C) | Memmert |
| Bacterial shaking incubator (37°C) | INFORS HAT |
| Biohit Proline multichannel pipette | Sartorius |
| Cell culture hood HERA Safe | Thermo Fisher Scientific |
| Cell culture incubator | Heraeus |
| cellenONE X1 | Cellenion |
| Centrifuges | Eppendorf |
| F.SIGHT | Cytena |
| Freezer (-20°C) | Liebherr |
| Freezer (-80°C) | Sanyo |
| Fridge (4°C) | Liebherr |
| Gel documentation | Herolab |
| GloMax Discover System | Promega |
| ImageXpress Micro XLS | Molecular Devices |
| ImageXpress Micro Confocal | Molecular Devices |
| Nanodrop ND-1000 spectrophotometer | Thermo Fisher Scientific |
| Neubauer cell counting chamber | BRAND |
| NYONE | SYNENTEC |
| Odyssey Infrared Imaging System | LI-COR Biosciences |
| Pipetboy acu pipette | INTEGRA Biosciences |
| Protein Gel Apparatus MiniProtean II | Bio-Rad |
| QuantStudio 5 real-time PCR machine | Thermo Fisher Scientific |
| Thermocycler | Applied Biosystems |
| Titramax 100 rocking platform | Heidolph |
| Trans-Blot Turbo Transfer System | Bio-Rad |
| Vacuboy aspiration device | INTEGRA Biosciences |
| VIPS | Solentim |
| Vortex mixer | Heideloph |
| Water bath | JULABO |

### 3.1.2 Consumables

| Consumable | Company |
|---|---|
| 384-well plates for TaqMan | Applied Biosystems |
| 4-15% Mini-PROTEAN TGX™ Precast Protein Gels (10, 12 and 15-well) | Bio-Rad |
| 96-well plate (black) | Greiner Bio-one |
| Adhesive Optically Clear Plate Seal | Thermo Fisher Scientific |
| Canonical tubes (15 and 50 ml) | BD Falcon |
| Cell culture dishes (100 mm and 150 mm) | Greiner Bio-one |
| Cell scraper | Corning |

| | |
|---|---|
| Cryogenic vials | Thermo Fisher Scientific |
| Disposable filtertips for micropipettes | Starlab |
| Disposable tips for micropipettes | Steinbrenner |
| Microcentrifuge tubes (1.5, 2 and 5 ml) | Eppendorf AG |
| Mini-PROTEAN TGX Stain-Free$^{TM}$ Gels | Bio-Rad |
| Multi-well cell culture plates (6, 12, 24, 48, 96-well) | Greiner Bio-one |
| PCR strips | Steinbrenner Laborsysteme |
| PVDF blotting membrane | Merck |
| Reservoirs (50 ml) | Corning |
| Serological pipettes (5 ml, 10 ml, 25 ml, 50 ml) | BD Falcon |
| Syringes | Sigma-Aldrich |
| Syringe filters (0.22 µm) | Sigma-Aldrich |
| Trans-Blot Turbo™ Mini PVDF Transfer Packs | Bio-Rad |
| Whatman 3 MM filter paper | GE Healthcare |

### 3.1.3 Chemicals and reagents

| Chemical/reagent | Company |
|---|---|
| 0.25% Trypsin EDTA Solution | Gibco |
| 17-ß-estradiol | Sigma-Aldrich |
| 5xHF buffer | Thermo Fisher Scientific |
| 6x Orange Loading Dye | Fermentas |
| Agarose | Carl Roth |
| Benzonase | Merck |
| Bortezomib | MedChemExpress |
| Bovine serum albumin (BSA) | Sigma-Aldrich |
| Charcoal stripped fetal bovine serum (CSFBS) | Sigma-Aldrich |
| cOmplete EDTA-free protease inhibitor | Roche |
| DAPI | Thermo Fisher Scientific |
| DMEM | Gibco |
| DMSO | Sigma-Aldrich |
| dNTPs | Thermo Fisher Scientific |
| Essential 8 $^{TM}$ media | Thermo Fisher Scientific |
| Ethanol | Sigma-Aldrich |
| Ethidium bromide | Sigma-Aldrich |
| Fetal Bovine Serum (FBS) | Gibco |
| Glycine | Sigma-Aldrich |
| Hoechst 33258 | Sigma-Aldrich |
| Isopropanol | Greiner Bio-one |
| L-glutamine | Gibco |
| Lipofectamine RNAiMAX | Thermo Fischer Scientific |
| MassRuler DNA Ladder Low Range | Fermentas |
| Methanol | Sigma-Aldrich |
| $MgCl_2$ | Thermo Fisher Scientific |
| MK-2206 dihydrochloride | MedChemExpress |
| Nuclease-free $H_2O$ | Thermo Fisher Scientific |
| Opti-MEM | Gibco |
| PageRuler$^{TM}$ Prestained Protein Ladder | Thermo Fisher Scientific |

| | |
|---|---|
| Paraformaldehyde 16% | Thermo Fisher Scientific |
| Penicillin/Streptomycin (P/S) | Gibco |
| Phosphate-Buffered Saline (PBS) | Gibco |
| PhosSTOP Phosphatase Inhibitor Cocktail | Roche |
| Phusion Hot Start II DNA-Polymerase (2 U/µl) | Thermo Fisher Scientific |
| Power SYBR Green PCR Master Mix (2x) | Thermo Fisher Scientific |
| Proteinase K | Thermo Fisher Scientific |
| Puromycin dihydrochloride | Sigma-Aldrich |
| RIPA Lysis and Extraction buffer | Thermo Fischer Scientific |
| RNase A | Qiagen |
| RNase-free DNase | Qiagen |
| Rockland Blocking Buffer | Rockland Immunochemicals |
| Roti-Load 1, 4x sample loading buffer | Carl Roth |
| Ruxolitinib | MedChemExpress |
| Sodium dodecyl sulfate (SDS) | Carl Roth |
| Sodium fluoride (NaF) | Bernd Kraft |
| Sodium orthovanadate ($Na_3VO_4$) | Sigma-Aldrich |
| Sodium pyruvate | Gibco |
| Sotrastaurin | MedChemExpress |
| TPA | Sigma-Aldrich |
| Tris-base | Sigma-Aldrich |
| Triton X-100 | Sigma-Aldrich |
| Trypsin 0.05%/EDTA 0.1% in PBS, w/o: Ca and Mg | PAN-Biotech GmbH |
| Tween 20 | Sigma-Aldrich |
| (Z)-4-hydroxyTamoxifen ≥98% | Sigma-Aldrich |

### 3.1.4 Solutions

| **Barcode amplification** | |
|---|---|
| 50x TAE (Tris-acetate-EDTA) | 242 g Tris-base<br>57.1 ml Acetic acid<br>100 ml 0.5 M EDTA<br>Add 1 l with $ddH_2O$<br>pH 8.0 |
| Lysis buffer | per 200 µl:<br>40 µl 5x HF buffer<br>10 µl Proteinase K<br>1 µl RNase A<br>149 µl Nuclease-free $H_2O$ |
| **Western Blotting** | |
| 10x TBS | 1.37 M NaCl<br>200 mM Tris<br>pH 7.6 |
| Blocking buffer | 1:1 Rockland blocking buffer:TBS<br>10 mM NaF<br>1 mM $Na_3VO_4$ |

| | |
|---|---|
| Lysis buffer | 10 ml RIPA buffer<br>1x cOmplete EDTA-free protease inhibitor<br>1x PhosSTOP phosphatase inhibitor<br><br>Additionally for Mass Spectrometry:<br>10 mM NaF<br>1 mM $Na_3VO_4$<br>250 U/ml Benzonase<br>10 U/ml RNase-Free DNase |
| SDS running buffer | 192 mM glycine<br>25 mM Tris<br>0.1% SDS |
| Transfer buffer | 20% Trans-Blot Turbo™ 5x Transfer Buffer<br>20% EtOH<br>60% $H_2O$ |

### 3.1.5 Commercial kits

| Kit | Company |
|---|---|
| AllPrep DNA/RNA Micro Kit | Qiagen |
| CellTiter-Glo® Luminescent Cell Viability Assay | Promega |
| Click-iT™ EdU Cell Proliferation Kit for Imaging, Alexa Fluor™ 594 dye | Thermo Fisher Scientific |
| DNeasy Blood & Tissue Kit | Qiagen |
| Pierce™ BCA Protein Assay Kit | Thermo Fisher Scientific |
| RevertAid™ H Minus First Strand cDNA synthesis kit | Thermo Fisher Scientific |
| RNeasy Mini Kit | Qiagen |
| Trans-Blot Turbo mini PVDF Transfer Kit | Bio-Rad |
| Wizard SV Gel and PCR Clean-up System | Promega |

### 3.1.6 siRNAs

Four individual siRNAs for *RELA*, *STAT1* and *STAT2* were each ordered from Dharmacon. The individual siRNAs were pooled for each target.

| Gene | Catalogue number | Annotation | Target sequences |
|---|---|---|---|
| ON-TARGET control | D-001810-10-20 | siONT+ | 1 – UGGUUUACAUGUCGACUAA<br>2 – UGGUUUACAUGUUGUGUGA<br>3 – UGGUUUACAUGUUUUCUGA<br>4 – UGGUUUACAUGUUUUCCUA |
| *RELA* | LQ-003533-00-0002 | siRELA | 1 – GGAUUGAGGAGAAACGUAA<br>2 – CCCACGAGCUUGUAGGAAA<br>3 – GGCUAUAACUCGCCUAGUG<br>4 – CCACACAACUGAGCCCAUG |
| *STAT1* | LQ-003543-00-0002 | siSTAT1 | 1 – GCACGAUGGGCUCAGCUUU<br>2 – CUACGAACAUGACCCUAUC<br>3 – GAACCUGACUUCCAUGCGG<br>4 – AGAAAGAGCUUGACAGUAA |
| *STAT2* | LQ-012064-00-0002 | siSTAT2 | 1 – GGACUGAGUUGCCUGGUUA<br>2 – GGACUGAGGAUCCAUUAUU<br>3 – GACCCCUCCUGGCAAGUUA<br>4 – GAUUUGCCCUGUGAUCUGA |

### 3.1.7 RT-qPCR primers

| Gene | Primer left | Primer right |
|---|---|---|
| *ACTB* | ATTGGCAATGAGCGGTTC | GGATGCCACAGGACTCCA |
| *GAPDH* | GAGTCCACTGGCGTCTTCAC | GTTCACACCCATGACGAACA |
| *RELA* | GCTTGTAGGAAAGGACTGCC | GCTGCTCTTCTATAGGAACTTGG |
| *STAT1* | GAAGTTCACCCTTCTAGACTTCAG | AGAGCCCACTATCCGAGAC |
| *STAT2* | GCTAGGCCGATTAACTACCC | CAGCTGTGAACCATGTCTCC |

### 3.1.8 Antibodies

| **Primary antibodies** | | | | |
|---|---|---|---|---|
| **Target** | **Antibody ID** | **Company** | **Species** | **Dilution** |
| ß-Actin | 69100 | MP Biologicals | Mouse | 1:1,250 |
| p(Ser) PKC Substrate | CST2261 | Cell Signaling Technology | Rabbit | 1:1,000 |
| p65 | CST8242 | Cell Signaling Technology | Rabbit | 1:1,000 |
| PRAS40 | CST2691 | Cell Signaling Technology | Rabbit | 1:1,000 |
| PRAS40 (Thr246p) | CST2997 | Cell Signaling Technology | Rabbit | 1:1,000 |
| STAT1 | CST9172 | Cell Signaling Technology | Rabbit | 1:1,000 |
| STAT2 | CST72604 | Cell Signaling Technology | Rabbit | 1:1,000 |
| **Secondary antibodies** | | | | |
| **Antibody** | **Antibody ID** | **Company** | **Species** | **Dilution** |
| F(ab')2-Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor™ 680 | A-21077 | Thermo Fisher Scientific | Goat | 1:10,000 |
| Goat anti-Mouse IgG (H+L) Secondary Antibody, DyLight 800 4X PEG | SA5-35521 | Thermo Fisher Scientific | Goat | 1:10,000 |

### 3.1.9 Software

| Software | Company |
|---|---|
| Image Studio Lite v5.2 | LI-COR Biosciences |
| Molecular Devices Analysis Software | Molecular Devices |
| Odyssey 2.1 | LI-COR Biosciences |
| Prism 9.4.1 | GraphPad Software, Inc. |
| QuantSudio Analysis and Design v1.5.1 | Thermo Fisher Scientific |
| SnapGene software 6.1.2 | Insightful Science |

## 3.2 Methods

### 3.2.1 Cell culture

#### 3.2.1.1 Cultivation of luminal A breast cancer cell lines

Parental T47D and MCF7 breast cancer cell lines were obtained from ATCC. These cell lines had been used by Dr. Simone Borgoni to generate barcoded treatment naïve as well as TAMR and LTED cell lines [236]. Cell lines were regularly authenticated by Multiplexion GmbH and tested for potential mycoplasma contamination. All cells were cultured in the appropriate growth media as in Table 1 within a humified incubator at 37°C and 5% $CO_2$. Cells were subcultured when they reached a confluency of 70-80%. Therefore, the growth media was aspirated, the cells washed once with PBS and detached from the surface with 0.25% Trypsin-EDTA. Fresh growth media was added to neutralize the trypsin and cells were reseeded in an appropriate dilution. To eliminate estrogenic effects of the pH-indicator phenol red [237], LTED cells were trypsinized in phenol red-free Trypsin and resuspended in phenol red-free growth media. T47D and MCF7 cells transduced with the ClonTracer library were additionally cultured in the presence of 1 µg/ml Puromycin.

**Table 1: Description of used human breast cancer cell lines and culture media.**

| Cell line | Description | Medium | |
|---|---|---|---|
| T47D wildtype (WT)<br>MCF7 wildtype (WT) | Luminal A breast cancer cell lines and clones | ‘+E2’ | DMEM supplemented with 10% FBS, 1% P/S, $10^{-8}$ M E2<br>1 µg/ml Puromycin was added for barcoded cell line pools and clones only |
| T47D +E2_5 pool and clones<br>Clones:<br>B5, B8, C12, D3, D11, E2, F6, H9<br>MCF7 +E2_1, +E2_3, +E2_5 pools | | | |
| T47D TAMR_2 and TAMR_4 pools and clones<br>Clones:<br>TAMR_2: B3, B5, C4, D8, D12, H2<br>TAMR_4: B2, C3, D8, E2<br>MCF7 TAMR_A pool and clones<br>Clones: B3, B6, E10 | Luminal A breast cancer cell lines and clones resistant to treatment with 100 nM 4-OHT | ’+4-OHT’ | DMEM supplemented with 10% FBS, 1% P/S, 100 nM 4-OHT<br>1 µg/ml Puromycin was added for barcoded cell line pools and clones only |
| T47D LTED_1 and LTED_2 pools and clones<br>Clones<br>LTED_1: B10, C4, C6, F10<br>LTED_2: C2, C11, F5<br>MCF7 LTED_2 and LTED_5 pools and clones<br>Clones:<br>LTED_2: B5, C4, D6, E11, G2<br>LTED_5: A9, C9, E4 | Luminal A breast cancer cell lines and clones resistant to estrogen deprivation, mimicking aromatase inhibition | ‘-E2’ | DMEM phenol red free supplemented with 10% CSFBS, 1% P/S 1% sodium pyruvate, 1% L-glutamine<br>1 µg/ml Puromycin was added for barcoded cell line pools and clones only |

Frozen cell stocks were generated by pelleting the cells (1,200 rpm for 5 min) and resuspending the cell pellet in freeze media consisting of 70% +E2 media, 20% FBS and 10% DMSO. $1x10^6$ cells in 1 ml freeze media were transferred to cryogenic vials and cooled down in an isopropanol bath at -80°C. Frozen cells were kept in liquid nitrogen for long-term storage.

Recovery of frozen cell vials was performed by thawing at 37°C, pelleting and aspiration of the DMSO containing media. Cells were resuspended and seeded in their respective growth media. The next day, the media was changed to remove dead cells.

#### 3.2.1.2 Generation and deconvolution of barcoded T47D control cells

T47D cells were transduced with individual barcodes obtained from the ClonTracer library [195] (Addgene #67267) and deconvoluted utilizing the F.SIGHT$^{TM}$ dispenser in collaboration with the DKFZ Cellular Tools Core Facility as described in the corresponding Application Note [238]. In brief, *Escherichia coli* cells were transformed with the ClonTracer library and single colonies obtained. From individual colonies, plasmids were isolated and the barcode sequences analyzed. Lentiviral barcode plasmids were produced in HEK293FT cells and T47D cells were transduced with eight different barcode plasmids. 1 µg/ml Puromycin was then used to select for successful integration events. Afterwards, single cells were spotted with the F.SIGHT$^{TM}$ dispenser in poly-L-lysine coated 96-well plates and grown out for five weeks in +E2 media. Clonal outgrowth was monitored with the automated cell culture microscope NYONE. Media was aspirated and the cells lysed as described below. The barcode region was PCR-amplified and subjected to Sanger sequencing as described below. Generation of stable cells, cell spotting, monitoring of clonal outgrowth and sequence analysis were performed by the Cellular Tools Core Facility.

#### 3.2.1.3 Generation of barcoded endocrine therapy resistant *in vitro* models

The barcoded T47D and MCF7 resistance models were generated by Dr. Simone Borgoni as previously explained [236]. In brief, $1x10^5$ treatment-naïve T47D and MCF7 cells were each lentivirally transduced with the ClonTracer library [195] at an MOI of 0.05 to likely ensure the delivery of a single barcode per cell. Successfully transduced cells were selected and grown out in the presence of 1 µg/ml Puromycin. Once enough cells were obtained, the cells were split into 20 replicates of $5x10^6$ cells each. Five replicates were immediately frozen down ('Initial'). Five biological replicates were each treated for 8 months with either 100 nM 4-OHT to generate Tamoxifen resistant (TAMR) cell lines or deprived of estrogen (-E2) mimicking clinically used aromatase inhibitors to generate long-term estrogen deprived (LTED) cell lines. Finally, five biological replicates were kept for 8 months under control conditions with E2 supplementation (+E2).

To determine the barcode and therefore cell composition of the initial, +E2, TAMR and LTED cell pools, all replicates were subjected to NGS of the barcode region by Dr. Simone Borgoni [236]. Barcode composition analysis was then performed by Dr. Luca Penso Dolfin as described previously (https://github.com/luca8651/Barcode_analyses-python, accessed: 11.11.20).

#### 3.2.1.4 Generation of single cell clones

To delineate clonal endocrine therapy resistance drivers, the previously established barcoded replicates were deconvoluted. Single cell suspensions were plated manually or by using a cell printer with the help of the DKFZ single-cell Open Lab, the DKFZ Cellular Tools Core Facility, Solentim and Cytena.

Table 2 shows an overview of all generated single cell clones including the spotting method and outgrow conditions. For better attachment of single cells, some 96 well plates used for cell spotting were coated with poly-L-lysine. Single cells were grown out for four weeks in 100 µl +E2 media, preconditioned +4-OHT, +4-OHT media without drug ('TAM media without 4-OHT') or Essential 8$^{TM}$ media. Wells were topped up with fresh media on a weekly basis. For T47D TAMR_2 and TAMR_4 clones, pre-conditioned media was used. Therefore, TAMR_2 and TAMR_4 cell pools were kept in +4-OHT media for 3-4 days. Afterwards, the media was sterile filtered and diluted with freshly prepared +4-OHT media before the media was added to the single cell clones. No individual clones from the T47D LTED_1 cell pool could be established using spotting in 96 well format. Therefore, 100 single cells were plated in poly-L-lysine coated 6 well plates and individual cell clones were grown out.

**Table 2: Generation of single cell clones used for this thesis.**

| **Cell line** | **Clones** | **Printer** | **Initial (4 weeks) growth media (without Puromycin)** | **Coating** |
|---|---|---|---|---|
| T47D +E2_5 | B5, B8, C12, D3, D11, E2, F6, H9 | Cellenion cellenONE X1 | +E2 | none |
| T47D TAMR_2 | C4 | Cellenion cellenONE X1 | Pre-conditioned +4-OHT media | poly-L-lysine |
| | B3, B5, D8, D12, H2 | Solentim VIPS | TAM media without 4-OHT | none |
| T47D TAMR_4 | B2, C3, D8, E2 | Cellenion cellenONE X1 | Pre-conditioned +4-OHT media | poly-L-lysine |
| T47D LTED_1 | B10, C4, C6, F10 | manually plated in 6 well plates | TAM media without 4-OHT | poly-L-lysine |
| T47D LTED_2 | C2, C11, F5 | Solentim VIPS | Essential 8™ Medium | none |
| MCF7 LTED_2 | B5, E11, G2 | Cellenion cellenONE X1 | TAM media without 4-OHT | poly-L-lysine |
| | C4, D6 | Cytena F.SIGHT | TAM media without 4-OHT | poly-L-lysine |
| MCF7 LTED_5 | A9, C9, E4 | Cytena F.SIGHT | TAM media without 4-OHT | poly-L-lysine |

### 3.2.1.5 Cell plating conditions

Isolated T47D and MCF7 clones, T47D and MCF7 WT cells without barcode, and MCF7 +E2_1, +E2_3 and +E2_5 pools were used for all further experiments and assays as described in detail below. All plating conditions are summarized in Table 3. Cell counting was performed manually using a Neubauer chamber.

**Table 3: Plated cell numbers.**

| Cell line(s) or clone(s) | Assay | Media | Culture dish | Cell number plated |
|---|---|---|---|---|
| T47D WT | 14 d pre-treatment | +4-OHT or -E2 | 100 mm | $3x10^5$ |
| | Cell viability | +E2 + DMSO or 100 nM MK-2206 | 96 well plate | 2,500 |
| | Proliferation, EdU, DAPI | +E2 | 96 well plate | 500 |
| T47D WT pre-treated | Cell viability | -E2 + DMSO or 100 nM MK-2206 | 96 well plate | 5,000 |
| | Proliferation | +4-OHT | 96 well plate | 500 |
| | Proliferation, EdU, DAPI | -E2 | 96 well plate | 2,000 |
| T47D +E2_5 B8, C12, E2, F6, H9 | 14 d pre-treatment | +4-OHT or -E2 | 100 mm | $2x10^5$ |
| | Cell viability | +E2 + DMSO or 100 nM MK-2206 | 96 well plate | 2,500 |
| | Proliferation, EdU, DAPI | +E2 | 96 well plate | 500 |
| | Proliferation | +4-OHT | 96 well plate | 500 |
| T47D +E2_5 E2 | AKT inhibition | +E2 + DMSO or MK-2206 | 6 well plate | $2x10^5$ |
| T47D +E2_5 B5, D3, D11 | 14 d pre-treatment | +4-OHT | 100 mm | $2x10^5$ |
| | 14 d pre-treatment | -E2 | 100 mm | $1.2x10^5$ |
| | Proliferation | +E2 or +4-OHT | 96 well plate | 500 |
| T47D +E2_5 B5, D3, D11 pre-treated | Proliferation | +4-OHT | 96 well plate | 500 |
| | Proliferation | -E2 | 96 well plate | 2,000 |
| T47D LTED_1 B10, C4, C6, F10<br>T47D LTED_2 C2, C11, F5 | 14 d pre-treatment | +E2 | 100 mm | $5x10^4$ |
| | Cell viability | -E2 + DMSO or 100 nM MK-2206 | 96 well plate | 5,000 |
| | Proliferation, EdU, DAPI | -E2 | 96 well plate | 2,000 |
| T47D LTED_1 C6 | AKT inhibition | +E2 + DMSO or MK-2206 | 6 well plate | $4x10^5$ |
| T47D LTED clones pre-treated | Proliferation, EdU, DAPI | +E2 | 96 well plate | 500 |
| T47D TAMR_2 B3, B5, C4, D8, D12, H2<br>T47D TAMR_4 B2, C3, D8, E2 | Proliferation | +4-OHT | 96 well plate | 500 |
| MCF7 WT | 14 d pre-treatment | +4-OHT or -E2 | 100 mm | $5x10^4$ |

| | | | | |
|---|---|---|---|---|
| | Proliferation | +E2 | 96 well plate | 500 |
| MCF7 WT pre-treated | Proliferation | +4-OHT or -E2 | 96 well plate | 500 |
| MCF7 +E2_1, +E2_3, +E2_5 | 14 d pre-treatment | +4-OHT or -E2 | 100 mm | 6x10$^4$ |
| | Proliferation | +E2 | 96 well plate | 500 |
| MCF7 +E2_5 | Proliferation, cell viability, EdU, DAPI | +E2 + DMSO or Bortezomib, Ruxolitinib, Sotrastaurin | 96 well plate | 500 |
| | Inhibitor treatment | +E2 + DMSO or Sotrastaurin | 6 well plate | 1.5x10$^5$ |
| | Transfection | +E2 | 96 well plate | 800 |
| | Transfection | +E2 | 6 well plate | 3x10$^4$ |
| MCF7 +E2_1, +E2_3, +E2_5 pre-treated | Proliferation | +4-OHT or -E2 | 96 well plate | 500 |
| MCF7 TAMR_1 B3, B6, E10 | Proliferation | +4-OHT | 96 well plate | 500 |
| MCF7 TAMR_1 B6 | Proliferation, cell viability, EdU, DAPI | +4-OHT + DMSO or Bortezomib, Ruxolitinib, Sotrastaurin | 96 well plate | 500 |
| | Inhibitor treatment | +4-OHT + DMSO or Sotrastaurin | 6 well plate | 1.25x10$^5$ |
| | Transfection | +4-OHT | 96 well plate | 600 |
| | Transfection | +4-OHT | 6 well plate | 2x10$^4$ |
| MCF7 LTED_2 B5, C4, D6, E11, G2 | Proliferation | -E2 | 96 well plate | 500 |
| MCF7 LTED_2 C4, E11 | Proliferation, cell viability, EdU, DAPI | -E2 + DMSO or Bortezomib, Ruxolitinib, Sotrastaurin | 96 well plate | 750 |
| | Inhibitor treatment | -E2 + DMSO or Sotrastaurin | 6 well plate | 2x10$^5$ |
| | Transfection | -E2 | 96 well plate | 1,600 |
| | Transfection | -E2 | 6 well plate | 6x10$^4$ |
| MCF7 LTED_5 A9, C9, E4 | Proliferation | -E2 | 96 well plate | 500 |
| MCF7 LTED_5 C9 | Proliferation, cell viability, EdU, DAPI | -E2 + DMSO or Bortezomib, Ruxolitinib, Sotrastaurin | 96 well plate | 750 |
| | Inhibitor treatment | -E2 + DMSO or Sotrastaurin | 6 well plate | 2x10$^5$ |
| | Transfection | -E2 | 96 well plate | 1,200 |
| | Transfection | -E2 | 6 well plate | 6x10$^4$ |

#### 3.2.1.6 Pretreatment of cells

Besides the long-term adaptations, also short-term responses of WT and control (+E2) cell lines and clones to 4-OHT treatment and estrogen deprivation (-E2) were assessed. Further, isolated T47D LTED clones were grown in +E2 media to assess their short-term adaptation to the drug removal. Cells were plated as in Table 3 in their respective growth media and allowed to attach overnight. The next day, the media was aspirated, the cells were washed with PBS and the pre-treatment media was added to the cells. Cells pre-treated for 14 days with +E2, +4-OHT or -E2 media are designated with an additional apostrophe to their name.

#### 3.2.1.7 Cell proliferation assay

Cells were plated in 100 µl +E2, +4-OHT or -E2 media per well into 96 well plates according to Table 3. To assess the proliferation of cells normally kept in +E2 media under 4-OHT treatment, cells were detached, resuspended in +E2 media and pelleted. The cell pellet was washed once with PBS before reconstitution, counting and plating in +4-OHT media. Pre-treated cells were detached and resuspended in the respective pre-treatment media. Cells pre-treated by estrogen deprivation were additionally detached using phenol-red free trypsin. Cells were allowed to attach overnight, grown for another 7 days and stained with 20 mM Hoechst 33342 (final dilution: 1:5,000) for 30 min. Cell numbers were determined by microscopy-based nuclei counting using the ImageXpress Micro XLS or ImageXpress Micro Confocal microscope and the Molecular Devices Software. Final cell numbers were normalized to the seeding controls.

#### 3.2.1.8 Cell viability assay

Cell viability was estimated from adenosine triphosphate (ATP) levels [239-241]. On the assays' final day, total cell numbers were first determined by Hoechst staining and the cell viability was determined with the CellTiter-Glo Luminescent Cell Viability Assay according to the manufacturer's instructions. In short, freshly prepared CellTiter-Glo Reagent was added to the cells, the plate briefly placed on an orbital shaker and the luminescence signal measured after 10 min using the GloMax Explorer Multimode Microplate Reader.

#### 3.2.1.9 Determination of cells in S-phase

Cells progressing through the cell cycle's S-phase were detected by 5-ethynyl-2'-deoxyuridine (EdU) incorporation and microscopic detection of the attached Alexa Fluor using the Click-iT™ EdU Cell Proliferation Kit for Imaging, Alexa Fluor™ 594 dye. 72 hours after measuring the plating controls by Hoechst staining, cells were pulsed with EdU for 21 h by replacing half of the media with fresh media containing 20 µM EdU (final concentration: 10 µM

EdU). The next day, media was aspirated and the cells were fixed with 4% paraformaldehyde (PFA) for 15 min. PFA was removed, cells were washed twice with 3% bovine serum albumin (BSA) in PBS and permeabilized with 0.5% Triton X-100 in PBS for 30 min. Afterwards, cells were washed twice with 3% BSA in PBS and freshly prepared Click-iT reaction mix (Table 4) was added to the cells. After 30 min incubation, cells were washed once with 3% BSA in PBS and once with PBS. Then, 100 µl Hoechst-33258 (1:5,000 diluted in PBS) was added. The cells were incubated for 30 min protected from light and cells were imaged with the ImageXpress Micro XLS or ImageXpress Micro Confocal microscope and the Molecular Devices Software. Cells going through S-phase were determined as Alexa Fluor 594 positive whereas total cell numbers were determined by Hoechst 33342 staining. The percentage of cells going through S-phase was calculated as ratio of Alexa Fluor 594 to Hoechst 33342 positive cells.

**Table 4: Click-iT reaction mix.**

| Component | Per reaction |
|---|---|
| Click-iT EdU reaction buffer | 43 µl |
| $CuSO_4$ | 2 µl |
| Alexa Fluor azide | 0.12 µl |
| Click-iT EdU buffer additive | 5 µl |
| Total volume | 50.12 µl |

### 3.2.1.10 Determination of dying/dead cells

Dying/dead cells were determined by using the exclusion dye 4′,6-diamidino-2-phenylindole (DAPI) [242] and Hoechst 33342 for total cell quantification. 96 h after measuring the plating controls with Hoechst staining, DAPI diluted in PBS was added to the cells with a final concentration of 1 µg/ml and after a 5 min incubation period, plates were imaged with the ImageXpress Micro XLS or ImageXpress Micro Confocal microscope and the Molecular Devices Software. Afterwards, Hoechst 33342 (final dilution: 1:1,000) was added to the cells and after 30 min incubation, the cells were imaged again. The percentage of dying/dead cells was determined as ratio of DAPI positive to Hoechst positive cells.

### 3.2.1.11 Inhibitor treatments

AKT signaling was inhibited in T47D WT and T47D *orange* clones with the pan-AKT inhibitor MK-2206 [243]. Cells were plated according to Table 3 in poly-L-lysine coated black 96 well plates. The next day, media was aspirated and replaced with media containing 100 nM MK-2206 or the corresponding volume of DMSO. 72 h after treatment start, total cell numbers and cell viability were determined. 100 nM MK-2260 was determined by a dilution experiment as previously described [244]. T47D +E2_5 E2 and LTED_1 C6 cells were plated in poly-L-lysine coated 6 well plates. 48 h later, media was replaced with media containing increasing concentrations of MK-2206. After 1 h, protein lysates were prepared and subjected to Western Blotting.

MCF7 +E2_5, TAMR_1 B6, LTED_2 C4 and E11, and LTED_5 C9 cells were treated with Bortezomib (proteasome inhibitor [245]), Ruxolitinib (JAK1/2 inhibitor [246]) or Sotrastaurin (pan-PKC inhibitor [247]). Cells were plated in black 96 well plates. The next day, media was aspirated and the cells treated with increasing concentrations of either inhibitor. Total cell numbers were determined after 7 days.

For the assessment of PKC inhibition by Sotrastaurin, cells were plated in 6 well plates and allowed to attach overnight. Cells were then pre-treated for 24 h with 2.50 µM Sotrastaurin before being stimulated for 30 min with 200 nM 2-O-Tetra-decanoylphorbol 13-acetate (TPA) [248, 249]. Phosphorylation of PKC substrates was then assessed by Western Blotting.

#### 3.2.1.12 Transfections

MCF7 +E2_5, TAMR_1 B6, LTED_2 C4 and E11, and LTED_5 C9 cells were transfected with siRNA targeting *RELA*, *STAT1* and *STAT2*. Cells were plated in 96 well plates in P/S-free media and transfected the next day using Lipofectamine RNAimax and siRNAs were used at a final concentration of 30 nM. Briefly, siRNAs and the RNAimax were diluted in Opti-MEM, the transfection mix vortexed and incubated for 5 min. The transfection mix was then added dropwise to the cells. In parallel, cells were each plated in 6 well plates in P/S-free media and transfected the next day in the same manner. Media was changed 6 h after the transfection.

### 3.2.2 Sanger sequencing-based barcode analysis

#### 3.2.2.1 Cell lysis for barcode amplification

During the outgrowth of single cell clones, residual cells were replated in 96 well plates and the media was aspirated when wells reached a confluency of at least 10%. 60 µl lysis buffer were added to each well, the lysates transferred to 1.5 ml Eppendorf tubes and afterwards, samples were incubated at 56 °C for 30 min and then at 96°C for 5 min. Samples were stored at 4°C until further processing.

#### 3.2.2.2 DNA and RNA isolation

DNA and RNA were simultaneously isolated using the AllPrep DNA/RNA Micro Kit according to manufacturer’s instructions. For the isolation of either DNA or RNA alone, the DNeasy Blood & Tissue Kit or RNeasy Mini Kit were used according to the manufacturer’s instructions, respectively. DNA and RNA were eluted in nuclease-free $H_2O$. Additionally, DNA digestion was performed for isolated RNA using RNase-Free DNase. Concentrations were determined with the NanoDrop ND-1000 UV-Vis Spectrophotometer.

#### 3.2.2.3 Amplification of barcode region and visualization of PCR products

To identify the barcodes in the generated single cell clones, the lentivirally integrated region containing the barcode sequence was amplified using Phusion Hot Start II DNA Polymerase and the primer pair: Fw: GCTGTGCCTTGGAATGCTAGTTGG, Rev: TCTGCTG-TCCCTGTAATAAACCCG. 10 µl lysis or 200 ng DNA were used as input and the PCR reactions were run as described in Table 5 and Table 6, respectively. PCR products were run in 1% Agarose gels in 1x TAE buffer at 100 V for 1 h and visualized after ethidium bromide staining using the Herolab Gel documentation system. 5 µl MassRuler DNA Ladder Low Range were used as size standard.

**Table 5: PCR Master mix.**

| Component | Amount |
|---|---|
| 5xHF buffer | 10 µl |
| dNTP Mix (10mM each) | 1 µl |
| Forward primer (10 µM) | 2 µl |
| Reverse primer (10 µM) | 2 µl |
| $MgCl_2$ | 0.5 µl |
| Polymerase | 0.5 µl |
| Lysis/DNA | 10 µl/200 ng |
| $H_2O$ | up to 50 µl |

**Table 6: Thermocycler program for barcode amplification.**

| | | |
|---|---|---|
| 98°C | 2 min | |
| 98°C | 10 s | 32x |
| 71°C | 20 s | |
| 72°C | 1 min 10 s | |
| 72°C | 10 min | |
| 4°C | ∞ | |

#### 3.2.2.4 PCR purification

The PCR reactions were purified using the Wizard SV Gel and PCR Clean-up System according to manufacturer’s instructions. PCR products were eluted in nuclease-free $H_2O$ and concentrations were determined with the NanoDrop ND-1000 UV-Vis Spectrophotometer.

#### 3.2.2.5 Sanger sequencing

Purified PCR products were subjected to Sanger sequencing [250] performed by Eurofins Genomics with the Fw primer utilized as sequencing primer and the barcode sequences were analyzed with the SnapGene software.

### 3.2.3 NGS-based barcode analysis

#### 3.2.3.1 PCR amplification and sequencing

The Sanger chromatograms for isolated MCF7 TAMR_1 and LTED_2 clones were too complex to be deconvoluted. Hence, Illumina indices modified from Bhang *et al.* 2015 [195] (https://www.addgene.org/pooled-library/clontracer/, assessed: 06/01/2023) were added by PCR and 48 clones selected for sequencing. 200 ng DNA were used as input and a PCR reaction with Rev_Index_049 and Rev_Index_050 primers without DNA input were each run as non-template controls. The PCR was performed according to Table 5 and Table 7 with the utilized primers listed in Table 8.

**Table 7: Thermocycler program for barcode amplification to be sequenced on MiSeq.**

| | | |
|---|---|---|
| 98°C | 2 min | |
| 98°C | 10 sec | 32x |
| 69°C | 20 sec | |
| 72°C | 12 sec | |
| 72°C | 10 min | |
| 4°C | ∞ | |

**Table 8: Utilized NGS primers adapted from Bhang *et al.* 2015 [195].** The three bases CTG were added to each reverse primer and can be distinguished from the original primers by small letters.

| | |
|---|---|
| WS PCR Forward Primer | AATGATACGGCGACCACCGAGATCTACACACTGACTGCAGTCTGAGTCTGACAG |
| WS_Rev_Index_001 | CAAGCAGAAGACGGCATACGAGATACGATCGTGAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_002 | CAAGCAGAAGACGGCATACGAGATCTAGATCGTGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_003 | CAAGCAGAAGACGGCATACGAGATGACTCGATCAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_004 | CAAGCAGAAGACGGCATACGAGATTGACTAGCTCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_005 | CAAGCAGAAGACGGCATACGAGATATGCTCAGCAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_006 | CAAGCAGAAGACGGCATACGAGATCGATCTGCATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_007 | CAAGCAGAAGACGGCATACGAGATGATAGCTGACGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_008 | CAAGCAGAAGACGGCATACGAGATTCAGCTACGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_009 | CAAGCAGAAGACGGCATACGAGATAGTACGCATGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_010 | CAAGCAGAAGACGGCATACGAGATCACGTCGATAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |

| | |
|---|---|
| WS_Rev_Index_011 | CAAGCAGAAGACGGCATACGAGATGTATCACGACGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_012 | CAAGCAGAAGACGGCATACGAGATTCGCAGTACTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_013 | CAAGCAGAAGACGGCATACGAGATAGCGTCTGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_014 | CAAGCAGAAGACGGCATACGAGATCAGCATGTCTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_015 | CAAGCAGAAGACGGCATACGAGATGTACTCATCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_016 | CAAGCAGAAGACGGCATACGAGATTCTGCAGCTAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_017 | CAAGCAGAAGACGGCATACGAGATACTGTACTCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_018 | CAAGCAGAAGACGGCATACGAGATCGACAGCTATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_019 | CAAGCAGAAGACGGCATACGAGATGTCATGCGTAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_020 | CAAGCAGAAGACGGCATACGAGATTAGTCGCATGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_021 | CAAGCAGAAGACGGCATACGAGATATCGATGACGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_022 | CAAGCAGAAGACGGCATACGAGATCGATAGTCGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_023 | CAAGCAGAAGACGGCATACGAGATGAGCTGTATCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_024 | CAAGCAGAAGACGGCATACGAGATTCTGATCGCAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_025 | CAAGCAGAAGACGGCATACGAGATAGCATCGTCTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_026 | CAAGCAGAAGACGGCATACGAGATCTACGTCTAGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_027 | CAAGCAGAAGACGGCATACGAGATGCTAGATGCTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_028 | CAAGCAGAAGACGGCATACGAGATTCGAGTGCATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_029 | CAAGCAGAAGACGGCATACGAGATACGCTGACATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_030 | CAAGCAGAAGACGGCATACGAGATCATACAGTGCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_031 | CAAGCAGAAGACGGCATACGAGATGAGCACTAGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_032 | CAAGCAGAAGACGGCATACGAGATTGCATGTAGCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_033 | CAAGCAGAAGACGGCATACGAGATAGTGATCGACGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_034 | CAAGCAGAAGACGGCATACGAGATCTGACATGCAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_035 | CAAGCAGAAGACGGCATACGAGATGTAGCAGATCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_036 | CAAGCAGAAGACGGCATACGAGATTCACTATGCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_037 | CAAGCAGAAGACGGCATACGAGATACTCGATACGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_038 | CAAGCAGAAGACGGCATACGAGATCGCATGATCAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_039 | CAAGCAGAAGACGGCATACGAGATGCAGATCACTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_040 | CAAGCAGAAGACGGCATACGAGATTCGACTAGTGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |

| | |
|---|---|
| WS_Rev_Index_041 | CAAGCAGAAGACGGCATACGAGATATCAGCGATGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_042 | CAAGCAGAAGACGGCATACGAGATCTGTATGAGCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_043 | CAAGCAGAAGACGGCATACGAGATGTGACTGTCAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_044 | CAAGCAGAAGACGGCATACGAGATTACGCTGCATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_045 | CAAGCAGAAGACGGCATACGAGATAGCTGATGCAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_046 | CAAGCAGAAGACGGCATACGAGATCTATGCACTGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_047 | CAAGCAGAAGACGGCATACGAGATGCTCATGTCAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_048 | CAAGCAGAAGACGGCATACGAGATTAGCGATCTGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_049 | CAAGCAGAAGACGGCATACGAGATACGTACTGCTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |
| WS_Rev_Index_050 | CAAGCAGAAGACGGCATACGAGATCATAGCATCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGCACTAGCATAGAGTGCGTAGCTctg |

PCR products were cleaned and the final concentration and size of the PCR products were evaluated with the 2200 TapeStation System. All samples were pooled and sequenced in the NGS Core Facility at the DKFZ on a MiSeq V2 300 Nano. A custom sequencing primer with the following sequence was used: TCTACACACTGACTGCAGTCTGAGTCTGACAG.

#### 3.2.3.2 Barcode composition analysis

Fastq files from the NGS of the barcode region were analyzed for individual barcode sequences using a custom script. The ShortRead package [251] was used in R (version 4.0.3) [252] and R studio [253] to read fastq files individually. In the next step, the respective barcodes were counted using the 'grepl' command. Analysis was performed by Dr. Birgitta Michels.

### 3.2.4 Barcode integration site analysis

Integration site analysis of barcoded complex cell pools was performed by Genewerk GmbH (Heidelberg, Germany) to rule out disruption of coding or regulatory elements as drivers of therapy resistance in the *in vitro* models as previously described by Dr. Simone Borgoni [236]. To verify that isolated T47D and MCF7 clones indeed came from the outgrowth of the same initially barcoded cell, the integration sites for all isolated T47D TAMR and LTED, and the MCF7 LTED_5 clones were verified by a PCR assay based on the integration site analysis of complex barcoded cell pools [236]. A specific primer upstream or downstream of each integration site to be investigated (Table 9) was combined with a primer specific for the viral 3' long term repeat (LTR) with the following sequence: CCCAACGAAGATAAGATCTGC. Amplification was performed using Phusion Hot Start II DNA Polymerase as described above and the thermocycling program is shown in Table 10. PCR product visualization and sequencing with the respective Fw primer of each reaction was performed as described above. For the MCF7 TAMR_1 and LTED_2 clones, no additional integration site validation was performed by this PCR-based assay given the multitude of barcodes integrated and lacking pool integration site analysis data.

**Table 9: Primers used for integration site validation.**

| Integration locus | Strand | Mainly found in (cell pools) | Primer sequence |
|---|---|---|---|
| Chr3:189,684,206 | + | T47D TAMR_2 | TCCCTTGACATCACTAATAGGC |
| Chr4:110,457,931 | + | MCF7 LTED_5 | GTTACCTCTGCTTGGCGAAC |
| Chr4:112,000,728 | + | T47D +E2_2 and LTED_2 | ACTCACATTGAGGCTAAAGGG |
| Chr6:120,131,260 | - | T47D TAMR_2 | CTTGCTGAAATTAAGGAGAGACG |
| Chr7:110,418,502 | + | T47D +E2_2 and LTED_2 | ACCAAAATCCTCTCGGGCTG |
| Chr8:66,103,520 | - | T47D TAMR_2 | CTGAGGCCACAATATTAACAGC |
| Chr8:91,606,636 | + | T47D TAMR_2 | GAGATTCAGCAGCAATAATCAGG |
| Chr11:31,426,888 | + | T47D +E2_2 and LTED_2 | GCTTGTGACCCACGTCTTAG |

**Table 10: Thermocycler program for integration site validation.**

| | | |
|---|---|---|
| 98°C | 2 min | |
| 98°C | 10 sec | 32x |
| 65°C | 20 sec | |
| 72°C | 60 sec | |
| 72°C | 10 min | |
| 4°C | ∞ | |

### 3.2.5 Sequencing of *ESR1* hotspot mutation codons

The genomic region covering the *ESR1* hotspot codons 536 to 538 was amplified using Phusion Hot Start II DNA Polymerase as described above and the primer pair: Fw: AATACCCACTCCTGCTTGGC, Rev: TATCTGAACCGTGTGGGAGC. The thermocycling program is shown in Table 11. PCR product visualization and sequencing with the Fw primer were performed as described above.

**Table 11: Thermocycler program for *ESR1* amplification.**

| | | |
|---|---|---|
| 98°C | 2 min | |
| 98°C | 10 sec | 32x |
| 66°C | 20 sec | |
| 72°C | 60 sec | |
| 72°C | 10 min | |
| 4°C | ∞ | |

### 3.2.6 Analysis of RNA expression

#### 3.2.6.1 RNA Sequencing

RNA sequencing (RNA-Seq) was performed in the NGS Core Facility at the DKFZ using the NovaSeq 6K PE 50 S1 for T47D samples or the NovaSeq 6K PE 50 SP for MCF7 samples. Further downstream processing and data analysis (3.2.6.2 to 3.2.6.4) was performed by Dr. Efstathios-Iason Vlachavas.

Bioinformatic analyses were generally performed using R (version 4.1.0) and the Bioconductor software [254]. To determine the read quality of the raw gene expression data, the FASTQC tool (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/, accessed: 07.04.22) was used. Next, the Rsubread R package [255] (version 2.6.4) was used to align the raw paired-end reads to the selected human reference genome version "GRCh38.p13. primary_assembly_gencode" (https://www.gencodegenes.org/human/release_39.html). Gene expression counts were then determined using the featureCounts pipeline [256] run with default settings. Processing, alignment and quantification were performed on the Omics IT and Data Management Core Facility (ODCF) DKFZ Compute Cluster, running under CentOS Linux 7 (Core) (https://odcf.dkfz.de/, accessed: 07.04.22) [257].

### 3.2.6.2 Pathway activity estimation

Downstream analysis was performed for each cell line individually. Genes with limited evidence of expression were removed by a non-specific intensity procedure (function filterByExpr::edgeR R package). Here, genes that did not have at least ten counts in at least one of the biological barcoding conditions were removed. Next, Trimmed Mean of M-values (TMM) normalization from edgeR R package (version 3.36.0) [258] along with the voom function from the limma R package (version 3.50.0) [259] was applied. These allowed to adjust for library composition bias, to compensate for non-biological variability and to increase statistical power for biological discoveries [260]. The limma R package was then used to perform differential expression analysis between the different cell lines and respective clones.

Next, Pathway RespOnsive GENes (PROGENy) [261, 262] was used to estimate the activity of 14 major pathways. Pathway activity estimation was based on previous high-throughput perturbation experiments and a consensus gene signature for each pathway. The tool (R package, version 1.16.0) was run using default settings and gene-level statistics such as limma t-values were used as input since these represent the respective differential expression comparisons between the distinct clonal conditions.

### 3.2.6.3 Transcription factor activity estimation

Transcription factor activities were determined using the DoRothEA-decouple R framework [263]. DoRothEA (R package, version 1.6.0) contains information about interactions between TFs and their targets which were obtained from different types of evidence resources (A to C confidence levels). Simplified, the TF activity was estimated from the (deregulated) expression of its target genes (molecular footprint). The decouple R package (version 2.1.6, weighted mean as selected statistic) [264] was then used for the statistical evaluation of the TF activity significance.

### 3.2.6.4 Gene Set Enrichment Analysis (GSEA)

The GSEA module implemented in the R package clusterProfiler (version 4.2.2, fgsea algorithm) [265] and the Hallmarks(H) gene sets from the msigdbr R package (version 7.5.1) (https://github.com/DavisLaboratory/msigdb, accessed: 07.11.22) were used for functional enrichment analysis.

### 3.2.6.5 Reverse transcription

Isolated mRNA was reverse transcribed using the RevertAid RT Reverse Transcription Kit according to the manufacturer's instructions. Briefly, a 12 µl mix of 1 µg RNA, 1 µl oligo-dT primer and nuclease-free $H_2O$ was incubated at 70°C for 5 min. Afterwards, the reverse transcription mix (Table 12) was added and the reaction incubated according to Table 13.

**Table 12: Reverse transcription reaction mix.**

| Reagent | Volume/reaction |
|---|---|
| 5x Reaction Buffer | 4 µl |
| 10 mM dNTP Mix | 2 µl |
| RiboLock RNase Inhibitor (20 U/µl) | 1 µl |
| RevertAid Reverse Transcriptase (200 U/µl) | 1 µl |

**Table 13: Reverse transcription thermocycler program.**

| | |
|---|---|
| 37°C | 5 min |
| 42°C | 1 h |
| 70°C | 10 min |
| 4°C | ∞ |

### 3.2.6.6 Quantitative reverse transcription polymerase chain reaction (RT-qPCR)

10 ng of total cDNA in 5 µl were added per well in a 384-well plate. Afterwards, 6 µl qPCR master mix as in Table 14 were added and the reaction incubated as in Table 10 on a QuantStudio 5. The QuantStudio™ Design & Analysis Software v1.5.0 was used for the analysis. *ACTB* and *GAPDH* were used as reference genes [266]. Relative changes were calculated by the comparative Ct method (ΔΔCt method) [267] and relative abundances were visualized as $2^{-\Delta\Delta Ct}$.

**Table 14: RT-qPCR master mix.**

| Reagent | Volume/reaction |
|---|---|
| SYBR *Green* PCR Master Mix | 5.5 µL |
| Left primer (20 µM) | 0.2 µl |
| Right primer (20 µM) | 0.2 µl |
| Nuclease-free $H_2O$ | 0.1 µl |

**Table 15: RT-qPCR thermocycler program.**

| | | |
|---|---|---|
| 50°C | 2 min | |
| 95°C | 15 min | |
| 95°C | 15 s | 45x |
| 60°C | 1 min | |
| 95°C | 15 s | |
| 60°C | 1 min | 0.075°C/s |
| 95°C | 15 s | |

### 3.2.7 Analysis of protein expression and phosphorylation

#### 3.2.7.1 Protein isolation

Cell culture plates were taken from the incubator and immediately placed on ice. Cells were washed twice with ice-cold PBS and lysis buffer was added to the cells (6 well plate: 40 µl, 10 cm dishes: 200 µl, 15 cm dishes: 400 µl). Cells were then detached using a cell scraper. The cell lysates were incubated on ice for 30 min and centrifuged for 20 min at 15,000xg and 4°C. Subsequently, the supernatant was transferred to a fresh tube. The protein content was determined by BCA assay and the lysates were stored at -80°C.

#### 3.2.7.2 BCA assay

Protein concentrations were determined by the Pierce™ BCA Protein Assay Kit according to the manufacturer's instructions. A serial dilution of BSA in PBS was used as reference and the colorimetric readout at 562 nm was performed using the GloMax Discover System.

#### 3.2.7.3 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins were electrophoretically separated according to their molecular weight by discontinuous sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) [268]. Proteins lysates were diluted with $H_2O$ and 4x Roti®-Load 1, and incubated at 95°C for 5 min. 5 µl of the size marker PageRuler™ Prestained Protein Ladder or up to 25 µg of total protein were loaded per lane onto Mini-PROTEAN TGX Stain-Free™ Gels. Electrophoresis was performed in 1x Running buffer at 140 V for 1 h.

#### 3.2.7.4 Western Blotting

After electrophoretic separation, proteins were transferred onto polyvinylidene fluoride (PVDF) membranes using the Trans-Blot Turbo™ Mini PVDF Transfer Packs and the Trans-Blot Turbo™ Transfer System according to the manufacturer's instructions. Unspecific binding sites were blocked in blocking buffer for 1 h. Subsequently, the membrane was incubated with the primary antibody diluted in blocking buffer overnight at 4°C. The membrane was washed three times for 5 min each with TBS containing 0.1% Tween 20 (TBS-T) and then incubated with the corresponding secondary antibody for 1 h diluted in TBS-T. After three washing steps as before with TBS-T, the membrane was imaged using the Odyssey Infrared Imaging System and quantification was performed using Image Studio Lite.

### 3.2.8 Mass Spectrometry-based phospho-proteomics

Protein lysates were prepared as described above. Consecutive sample preparation, mass spectrometry measurement and data analysis (3.2.8.1-3.2.8.4) were performed by Luisa Schwarzmüller.

#### 3.2.8.1 Sample preparation workflow for phospho-proteomics

320 µg protein were precipitated according to Wessel and Flügge [269]. Four volumes of methanol and one volume of chloroform were added to the protein extract. In the next step, three volumes of $H_2O$ were added and the mixture centrifuged for 2 min at 18,000xg leading to the formation of three phase-separated layers. Here, the protein containing phase was between a chloroform phase and a $H_2O$/methanol phase. The upper phase was discarded and the protein pellet was washed with three volumes of methanol and centrifuged. The supernatant was discarded and the protein pellet was dried and stored at -20°C.

Tryptic digestion was performed next. Therefore, the protein pellet was reconstituted in 8 M urea which was supplemented with 100 mM NaCl, 50 mM Tetraethylammonium bromide, cOmplete EDTA-free protease inhibitor and PhosSTOP phosphatase inhibitor. For the reduction of disulfide bridges, 10 mM Dithiothreitol (DTT) was added and the samples were incubated for 1 h at 27°C before alkylation with 30 mM iodoacetamide was performed for 30 min. The reaction was quenched with 10 mM DTT for 15 min. Proteins were then digested with lysyl endopeptidase (Fujifilm, #125-05061) (enzyme-protein ratio: 1:100) for 4 h at 30°C. Afterwards, the urea concentration was diluted to 1.6 M and the proteins were digested with trypsin (enzyme-protein ratio: 1:50) for 16 h at 37°C. Formic acid with a final concentration of 2% (v/v) was added to stop the digestion.

The peptides were afterwards desalted using Sep-Pak C18 cartridges (Waters, #WAT054955). Previously, the cartridges had been conditioned with 100% acetonitrile. These were then washed with 0.6% acetic acid in 80% acetonitrile and equilibrated twice with 2.5% formic acid. Proteins were loaded onto the cartridges and the flow-through was collected. Loading and collection of flow-through was performed twice. Peptides were then washed four times with 2.5% formic acid before being eluted in two steps using 0.6% acetic acid in 80% acetonitrile. The eluate was vacuum centrifuged to dryness.

Phosphorylated peptides contain a negative charge and were enriched using an Immobilized Metal Affinity Chromatography (IMAC) column (Thermo Fisher Scientific,

#063276). The column was prepared as follows: Iron Chloride ($FeCl_3$) from previous enrichments was removed with 50 mM EDTA (pH 8) at a flow rate of 1 ml/min for 24 min and the column was flushed with $H_2O$. Charging with 25 mM $FeCl_3$ in 100 mM acetic acid was performed at a flow rate of 0.2 ml/min for 30 min. The column was then rinsed with 0.1% formic acid at a flow rate of 4 ml/min for 4 h and stored at 4°C until further use.

To load the peptides onto the IMAC column, dried peptides were first reconstituted in 30% acetonitrile with 0.07% trifluoroacetic acid (TFA). Loading onto the IMAC column was performed with a flow rate of 0.2 ml/min and elution was achieved using an ammonium gradient. The collected phospho-fraction was vacuum centrifuged to dryness.

Self-made Stop and Go Extraction (STAGE) tips were used to desalt the enriched phospho-peptides. Therefore, three layers of C18 material (Supelco, #66883-U) were placed in a pipette tip. The C18 material was then pre-wetted with methanol and washed with 80% acetonitrile in 0.1% TFA. Equilibration was done with 0.1% TFA. Dried phospho-peptides were reconstituted in 0.1% TFA and loaded onto the STAGE tips. Peptides were then washed with 0.1% TFA and eluted in two steps using 60% and 80% acetonitrile, respectively. The eluate was vacuum centrifuged to dryness and stored at -20°C until further analysis.

#### 3.2.8.2 Liquid Chromatography

Peptides were dissolved in ULC/MS grade $H_2O$ containing 0.1% TFA and 2.5% 1,1,1,3,3,3-Hexafluoro-2-propanol, sonicated for 5 min and transferred to autosamplers vials. Autosamplers vials were then placed in the autosampler module of the Ultimate 3000 liquid chromatography system. The LC was operated at a flow rate of 300 nl/min and the columns were heated to 35°C. First, peptides were loaded onto a trapping column (Thermo Fisher Scientific, 160454) using a mixture of 98% loading buffer A (0.1% TFA in $H_2O$) and 2% loading buffer B (0.1% TFA in acetonitrile). Peptides were then eluted from the trapping column and loaded onto the analytical column (Waters, 186008795, BEH C18 130Å 1.7 µm 75x250 mm). Using a linear gradient of 2-28% acetonitrile, peptides were separated according to their hydrophobicity. Separated peptides were ionized by electrospray ionization with an applied voltage of 2,200 V.

#### 3.2.8.3 Mass Spectrometry in DIA mode

Mass spectrometry was performed in data-independent acquisition (DIA) mode. MS1 scans were acquired at a resolution of 120K covering the range from 350-1400 m/z. Maximum injection time was 45 ms and the automated gain control (AGC) target was set to 3e6. MS2 scans were acquired in 48 precursor isolation windows of variable width and 1 m/z overlap that covered the range from 400-1200 m/z. The orbitrap was operated at a resolution of 30K and a normalized collision energy of 26% was applied. Maximum injection time was 54 ms and the AGC target was set to 1e6. The total cycle time equaled 3.6 s.

#### 3.2.8.4 Data analysis

Biognosys software Spectronaut (version 15) was used for protein identification and quantification form the obtained DIA raw data. For phosphopeptides, the newly developed identification and localization algorithm implemented in Spectronaut was used [270]. Identified phosphorylated peptides were site-collapsed using the Perseus [271] (version 1.6.2.3) plug-in PeptideCollapse [270] with a localization cut-off at 0.95.

### 3.2.9 Estimation of kinase activities

From the phospho-proteomics data, kinase activities were estimated by Dr. Efstathios-Iason Vlachavas. The downstream analysis was performed using R/Bioconductor software (R version 4.1.0). Raw measurements were processed with the PhosR R package [272] (version 1.4.0) and the intensities were normalized using the VSN R package (version 3.62.0) [273]. Phosphosites with limited evidence of quantification were excluded. Limited evidence of quantification was defined as not quantified in at least 2/3 samples in at least one biological condition. Differential analysis of phosphosites was performed with the limma R package (version 3.50.0). Kinase activities were determined in the next step. Simplified, the kinase activity was estimated from the phosphorylation of its substrates. The Omnipath R package (version 3.2.5) was used as database for prior knowledge as it contains information about signaling network interactions, enzyme-post translation modifications relationships, protein complexes, protein annotations and intracellular communication roles [274]. Normalized weighted mean was selected as statistical method for the decouple R package.

### 3.2.10 Analysis of CPTAC-BRCA dataset

The CPTAC-BRCA [234] dataset was analyzed to correlate *in vitro* and clinical findings. This dataset comprises 122 treatment-naïve breast cancer patients and contains clinical, genomic aberrations, gene expression, protein abundance and (phospho-)proteomics data. Two patients were removed because no information about their ER status was available. The cohort was initially classified into patients with ER+ (81 patients) and ER- disease (39 patients). Raw data from the patient dataset was obtained from the cBioPortal [275, 276]. Analysis of gene expression data and phospho-proteomics data were performed as described above by Dr. Efstathios-Iason Vlachavas. In the next step, the cohort was analyzed again on a per patient basis by Dr. Efstathios-Iason Vlachavas. Gene expression data was z-score transformed to compare per patient expression data to the rest of the profiled cohort. Pathway and transcriptional factor activities were estimated using the scaled and normalized gene expression values as input. Single sample gene set enrichment analysis (ssGSEA) [277] was performed using the available script “ssgsea-gui.R” (https://github.com/broadinstitute/ssGSEA2.0, accessed: 20.12.22) with default settings (sample.norm.type: none, weight: 0.75, statistic: area.under.RES, nperm: 1000, output.score.type: NES and correl.type: z.score). The hallmark signatures gmt file (version 7.5.0) (https://github.com/broadinstitute/ssGSEA2.0/tree/master/db/msigdb, accessed: 20.12.22) and the scaled gene expression matrix were used as input.

### 3.2.11 Statistical analysis

Statistical analysis was performed using GraphPad Prism 9.4.1 as individually indicated. P values <0.05 were considered statistically significant and p values <0.05, <0.01, <0.001 and <0.0001 are indicated in figures with one, two, three and four asterisks respectively.

# 4 Results

## 4.1 Cellular barcoding reveals distinct clonal complexities in endocrine therapy resistance

In a previous PhD project [236], barcoded endocrine therapy resistant ER+ breast cancer cell lines were generated to study the clonal complexity of endocrine therapy resistance. There, Dr. Simone Borgoni had lentivirally transduced treatment-naïve T47D and MCF7 cell lines with the ClonTracer library [236]. Following an initial expansion, four different conditions were generated from the starting population with five biological replicates each. One set of replicates was immediately frozen down ('Initial') to assess the complexity of the transduced library. Of the other three sets of replicates, one set was kept under normal growth conditions (+E2), one set was treated with 4-OHT, the active metabolite of Tamoxifen, and one set was deprived of estrogen (-E2) to mimic clinically used aromatase inhibitors. Cells were chronically exposed to either growth condition for eight months (Figure 7).

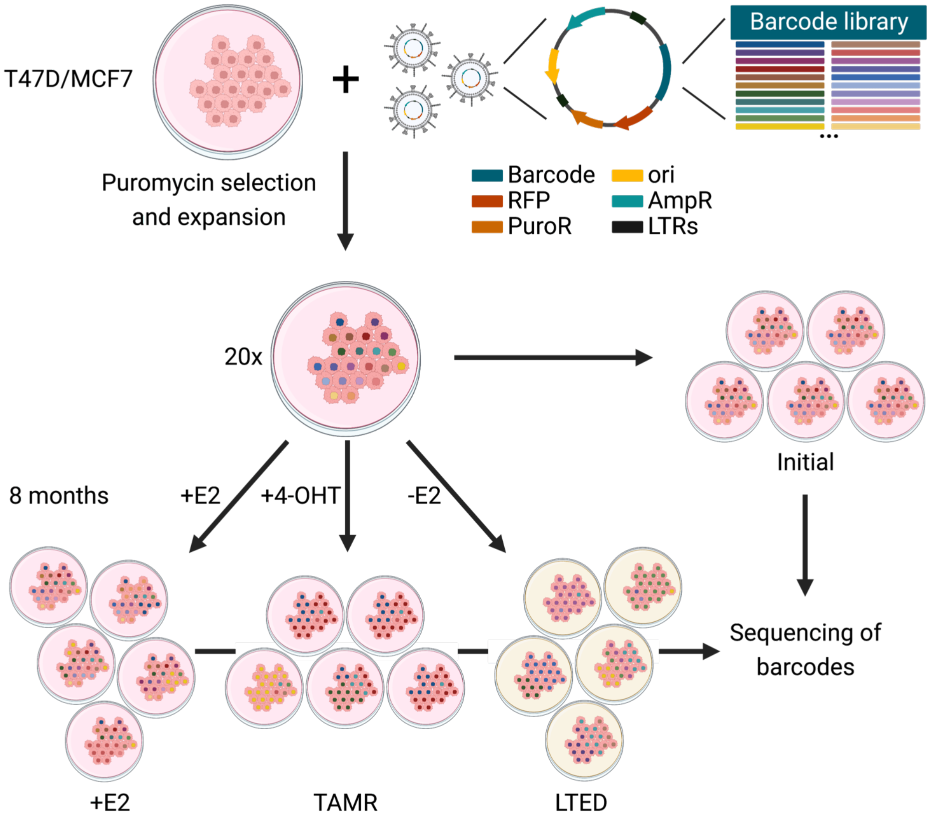


**Figure 7: Schematic overview of cellular barcoding and resistance acquisition.** T47D and MCF7 cells were transduced with the ClonTracer barcode library at a very low MOI and selected for successful integration events by Puromycin. For eight months, one set of replicates each was kept under normal growth conditions (+E2) or treated with either 4-OHT or estrogen deprivation (-E2) yielding TAMR and LTED cell lines (see below). RFP: red fluorescent protein. PuroR: Puromycin resistance gene. ori: origin of replication. Amp: Ampicillin resistance gene. LTRs: Long terminal repeats. Adapted from Dr. Simone Borgoni [236] and Bhang, *Nature Medicine* 2015 [195]. Created with BioRender.com.

After eight months, cells chronically treated with 4-OHT or estrogen deprivation (-E2) had regained a proliferative phenotype indicative of endocrine therapy resistance. It should be noted however, that T47D LTED cells showed only slightly increased proliferation compared to control (+E2) cells treated by estrogen deprivation (-E2). The cells rendered resistant to 4-OHT and estrogen deprivation (-E2) are accordingly referred to as TAMR and LTED cells, respectively (Figure 8).

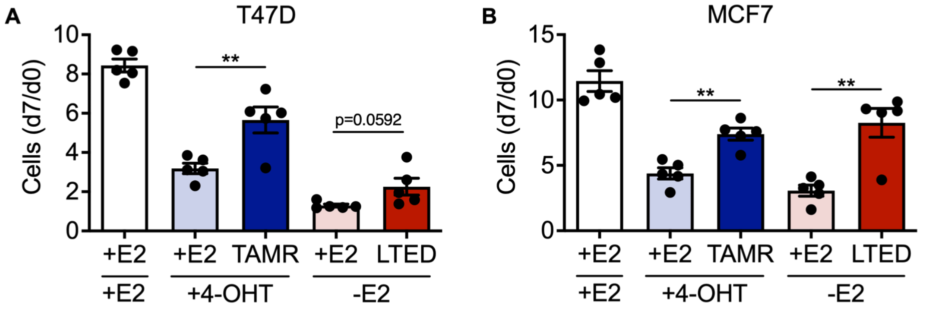


**Figure 8: Proliferation of barcoded cell pools.** Cells (+E2, TAMR, LTED) were kept in or treated with the indicated media (+E2, +4-OHT, -E2). Control cells (+E2) were pre-treated for 14 days by estrogen deprivation (-E2). Cell numbers were determined by microscopy-based nuclei counting at day 0 (d0) and day 7 (d7). Individual barcoded replicates were combined and the average (mean) ± SEM is shown. n = 1. * represents p<0.05 and ** represents p<0.01 as determined by unpaired two-tailed t-tests. Data generated by and figure adapted from Dr. Simone Borgoni [236].

After having confirmed the resistance of the TAMR and LTED cell lines, the barcode complexity of all initial, +E2, TAMR and LTED cell line pools was assessed by Dr. Simone Borgoni [236]. In brief, barcode regions of complex cell pools were deeply sequenced by Dr. Simone Borgoni and deconvoluted by Dr. Luca Penso Dolfin. Figure 9 highlights the number of reads per sample and the detected barcodes in the T47D and MCF7 cell line models.

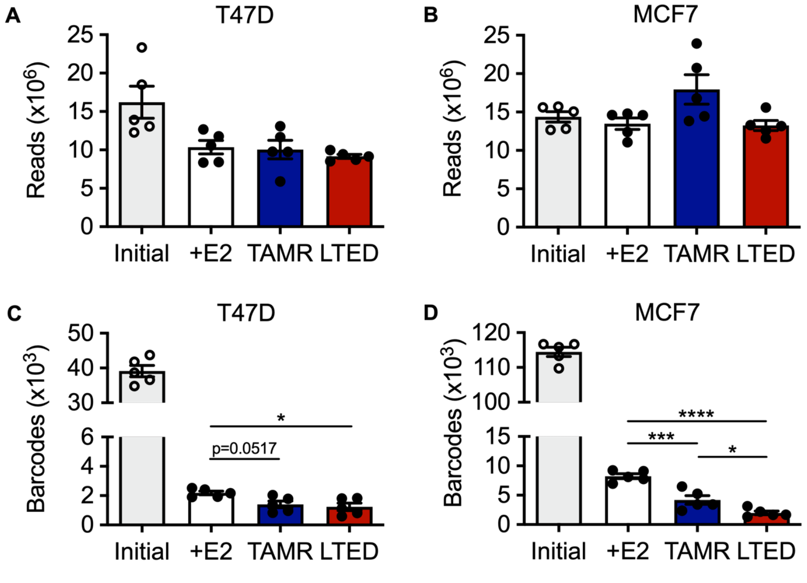


**Figure 9: Number of reads and identified barcodes.** Number of total reads for T47D (**A**) and MCF7 (**B**) as well as number of identified barcodes for T47D (**C**) and MCF7 (**D**) are shown. **C** and **D** * represents p<0.05, *** represents p<0.001 and **** represents p<0.0001 as determined by ordinary one-way ANOVA with Tukey multiple comparisons test. Initial cell populations were not included in the statistical comparisons. Data generated by and figure adapted from Dr. Simone Borgoni [236].

Sequencing of individual samples generated on average 9-16x$10^6$ reads for T47D and 13-18x$10^6$ reads for MCF7 replicates. Additionally, no sample dropped out during the sequencing pointing to similar coverage across all replicates assessed (Figure 9 A and B). For

T47D and MCF7, in the initial cell populations a median of $38.6*10^3$ and $115.8*10^3$ individual barcodes were detected, respectively. For MCF7, the obtained number of detected barcodes was close to the theoretically delivered $100.0*10^3$ individual barcodes, whereby only half of the expected barcodes were detected for the T47D. Continuous subculturing under control (+E2) conditions led to a strong decrease in the amount of detected barcode for both cell lines. A further reduction in detected barcodes was evident in the TAMR and LTED conditions (Figure 9 C and D). These findings point to a random dropout of clones during subculturing and further enrichment of clones in response to the applied endocrine therapies.

## 4.2 Barcode composition analysis of final cell pools

Dr. Simone Borgoni was then further interested in the frequency and distribution of specific barcodes across replicates to see if different or recurrent barcodes would be enriched under control (+E2) or treatment (TAMR, LTED) conditions [236]. The overall barcode distribution is shown in Figure 10.

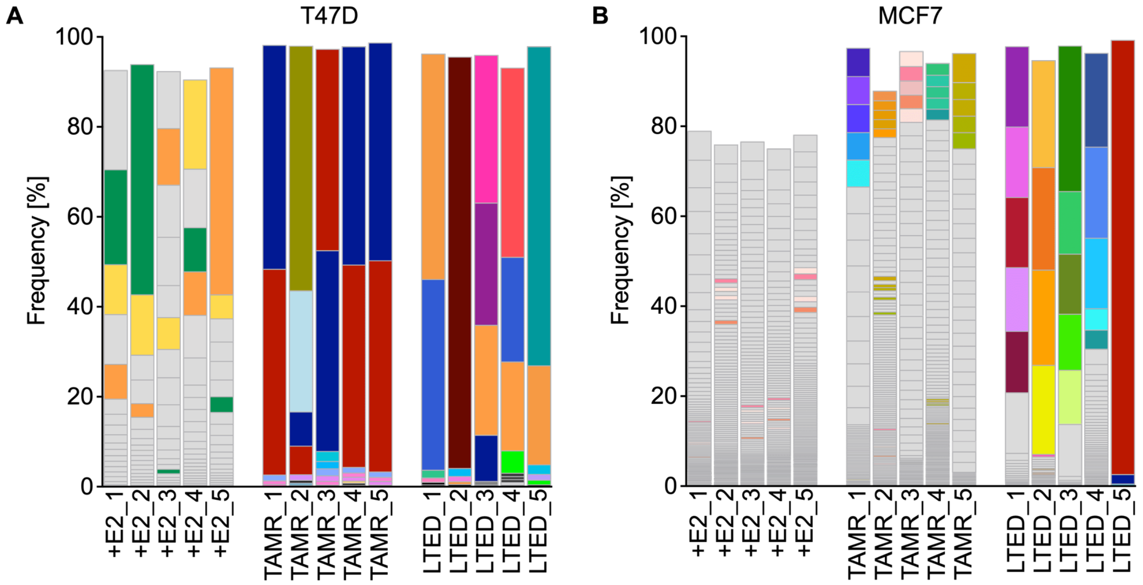


**Figure 10: Barcode composition analysis of final cell pools.** The detected barcodes in the control (+E2), 4-OHT treated (TAMR) and estrogen deprived (LTED) cell pools are shown for T47D (**A**) and MCF7 (**B**). For T47D TAMR and LTED (**A**): the overall top 25 barcodes are differentially color-coded. An additional clonal population is highlighted in the +E2 replicates in *green*. For MCF7 TAMR and LTED replicates (**B**): the top five detected barcodes in each TAMR and LTED replicate are differentially color-coded. Data generated by and figure adapted from Dr. Simone Borgoni [236].

For T47D TAMR replicates, the *dark red* and *blue* barcodes were strongly enriched in four out of five replicates pointing to the selection of pre-existing clones. Conversely, the detection of multiple and distinct barcodes in the MCF7 TAMR replicates pointed to resistance acquisition. It should be highlighted as well that under LTED conditions, in one T47D and MCF7 replicate each, a single barcoded clonal population almost exclusively dominated the respective replicates (T47D LTED_2 and MCF7 LTED_5), pointing to a strong selective advantage of these clonal populations. In the T47D model, the *orange* clonal population was of key interest as it was enriched under control (+E2) and treatment (LTED) conditions pointing to selective advantages under normal growth as well as treatment conditions. The frequency of the *orange* clonal populations in individual +E2 and LTED replicates ranged from less than 5% (+E2_2, LTED_2) to up to 50% (+E2_5, LTED_1, Figure 10 A) pointing to a recessive or dominant character compared other barcoded clonal populations, respectively.

Dr. Borgoni next wanted to identify chromosomal loci where the lentivirus had integrated, to rule out that viral integration into gene coding or regulatory regions could have been drivers of the clonal selection. To this end, five barcoded replicates were analyzed by LAM-PCR at Genewerk GmbH followed by Illumina sequencing [236]. Dr. Simone Borgoni then correlated the barcode read data (as analyzed by Dr. Luca Penso Dolfin) with the integration site data to match individual barcodes and their likely integration sites. Exemplarily, the *brown* clonal population made up >90% of the T47D LTED_2 replicate (Figure 10 A). >99% of viral DNA for this replicate were found to be integrated into Chr4:112,000,728 (GRCh38/hg38) suggesting that the *brown* barcode had been integrated into this genomic region encoding for intron 2 of *LINC02945*, a gene with unknown significance. Based on these findings, Dr. Simone Borgoni concluded for selected individual replicates that the respective viral integration sites did not disrupt any coding or regulatory regions and, therefore, were not drivers of clonal selection *per se*.

## 4.3 Project overview

The results I have described so far were generated by Dr. Simone Borgoni and are described in more detail in his PhD thesis [236]. Based on these tools and findings, I took over the project starting with the barcoded complex control (+E2) and endocrine therapy resistant (TAMR and LTED) cell pools. The focus of my PhD project was then to identify and perturb individual clonal resistance drivers in selected replicates of interest. To tackle this challenge, I subdivided my project into three major parts (Figure 11).

Initially, I performed single cell plating to deconvolute complex cell pools and obtain clones originating from a single cell. Multiple clones coming from the same progenitor cell as determined by the same barcode were isolated for selected complex cell pools. Afterwards, I expanded the clones and characterized them on phenotypic level mainly by proliferation assays, on transcriptomic level by RNA-Seq, and by Mass spectrometry-based (phospho-) proteomics. Based on the gene expression and phospho-proteomics data, the activities of pathway, transcription factors and kinases were determined. Treatment- and clone-specific alterations and similarities between different endocrine therapy resistant clones were identified and highlighted. I then perturbed individual private (limited to a single clonal population) as well as shared (by different clones) resistance drivers by knockdown or inhibitor treatments. Finally, I highlighted the clinical significance of individual *in vitro* resistance drivers.

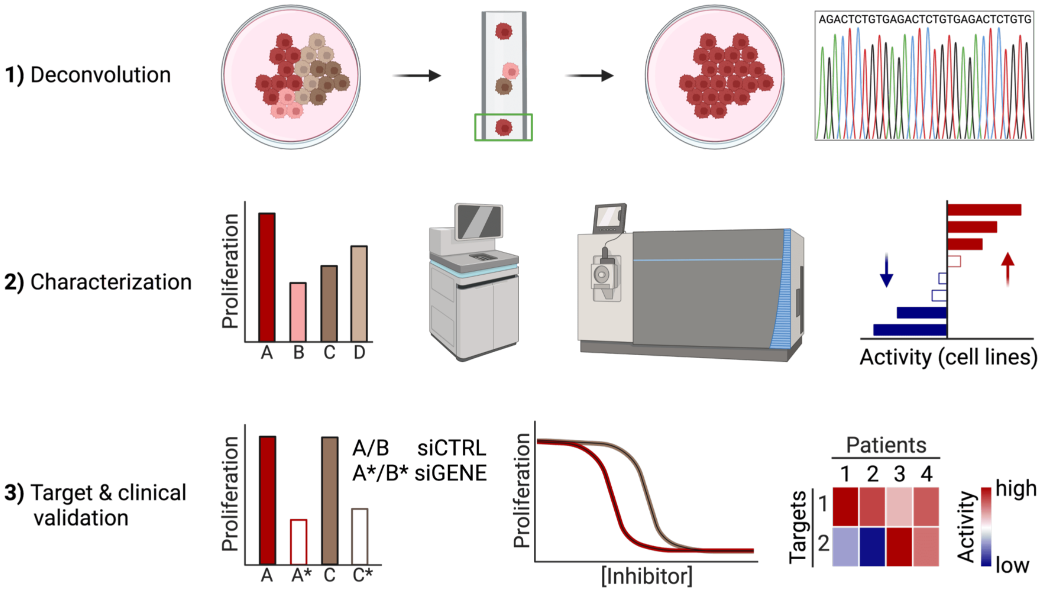

1) Deconvolution
AGACTCTGTGAGACTCTGTGAGACTCTGTG
2) Characterization
Proliferation
A
B
C
D
Activity (cell lines)
3) Target & clinical validation
Proliferation
A
A*
C
C*
A/B siCTRL
A*/B* siGENE
Proliferation
[Inhibitor]
Patients
1
2
3
4
Targets
1
2
Activity
high
low

**Figure 11: Project overview. 1)** Deconvolution of complex cell pools to obtain clonal populations originating from a single cell with known barcode. Green frame indicates the detection window for a single cell to be dispersed **2)** Characterization by proliferation assays, expression analysis by RNA-Seq and (phospho-) proteomics (left to middle right). Based on the obtained data, activities of signaling pathways, transcription factors and kinases were determined (right). **3)** Targeting of potential resistance drivers with siRNA (left) or inhibitors (middle). Correlation of in *in vitro* identified resistance drivers with clinical findings (right). Created with BioRender.com.

## 4.4 Deconvoluting the clonal complexity of barcoded cell pools

Deconvolution of complex cell pools had not been established in the lab. Therefore, I first did a proof-of-principle experiment to validate the applicability of the available experimental setup. This work was done with help of Dr. Rainer Will (Cellular Tools Core Facility, DKFZ). T47D cells were initially transduced with eight different lentiviral particles containing known barcodes from the ClonTracer library [195]. Cells were then grown out, mixed in equimolar ratios and spotted using the Cytena's F.SIGHT spotter. Clones originating from single cells were grown out and the outgrowth monitored with the NYONE microscope. Barcodes were then analyzed by Sanger sequencing (Figure 12).

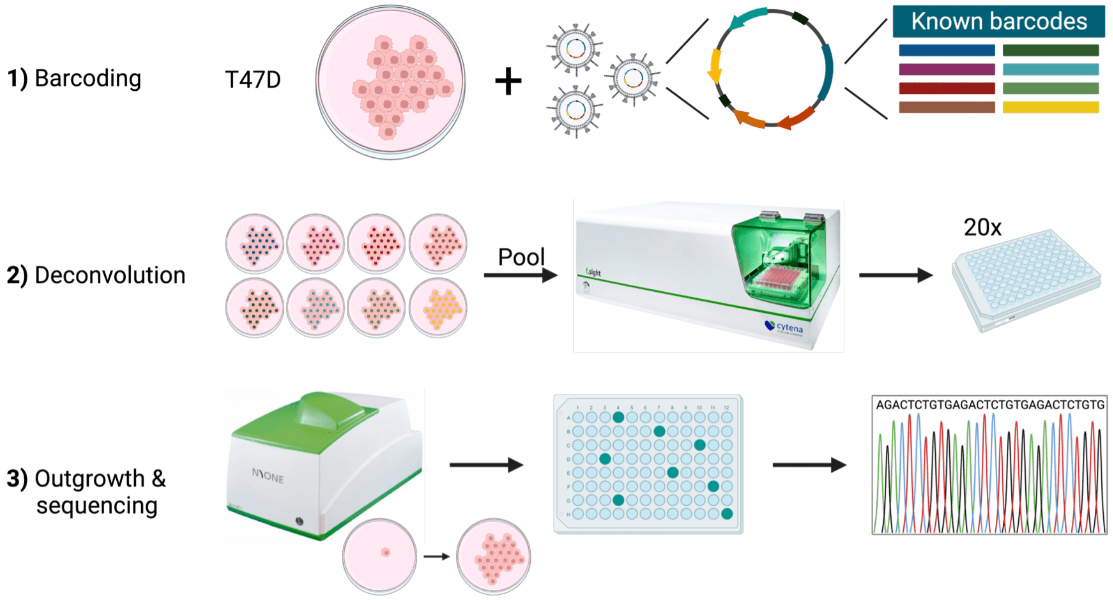


**Figure 12: Deconvolution overview.** Figure adapted from our Application Note [238]. Created with BioRender.com.

After five weeks, 279 clones were obtained with at least 20 cells resulting in a clonal outgrowth rate of 15% (279 clones obtained/20*94 cells spotted) under control (+E2) growth conditions. PCR products were generated and analyzed by Sanger sequencing for 94 clones as exemplarily shown in Figure 13 A and B. Each barcode was on average detected 11.75 times (12.50%), which represents the expected absolute frequency of 94 analyzed clones characterized by eight different barcodes as indicated by the dotted line in Figure 13 C. The details of this experiment have been published in an Application Note [238].

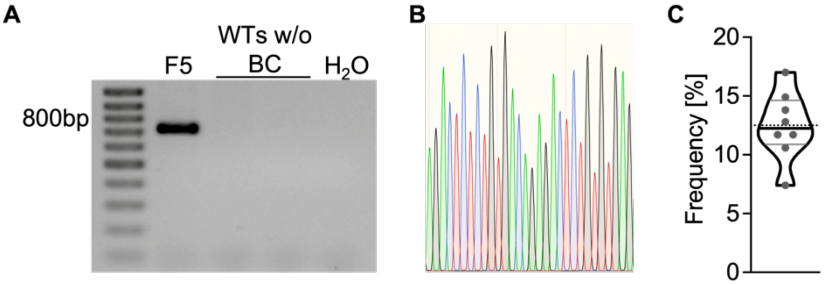


**Figure 13: Barcode PCR and distribution of identified barcodes. A** and **B** EtBR stained agarose gel and Sanger chromatogram of amplified barcode sequence. **A** F5 represents T47D LTED_2 F5 clone (see below) and 'WTs w/o BC' refers to non-barcoded T47D and MCF7 cell pools used as additional controls. **B** Barcode (AGACTCTCTGTGACAGAGACTCTGTGTGAG) sequence represents *orange* barcode. **C** 94 clones were analyzed by sequencing. Expected frequency per barcode: 12.50 (dotted line). Measured mean: 12.49. Subfigure **C** was adapted from our Application Note [238].

## 4.5 Deconvoluting endocrine therapy resistant cell pools

Having proven the reliability of the deconvolution protocol and technology, I next wanted to go ahead with the deconvolution of individual therapy-resistant clones. I selected different replicates for deconvolution as I next wanted to identify and then validate clonal resistance drivers (Figure 14). In the T47D cell line model, clones characterized by the *orange* barcode were of strong interest since cells carrying the same barcode had been enriched under +E2 conditions and also under estrogen deprivation. I decided for the +E2_5 and LTED_1 pools since both contained the recurrently enriched *orange* clonal population with a share of roughly 50%. The LTED_2 pool was chosen for deconvolution as well since the *brown* clonal population dominated this replicate pointing to a strong selective advantage of this clonal population. Given a temporary shift in the clonal makeup of this pool (Supplementary Figure 1), I was not able to isolate *brown* clones but rather obtained *orange* clones also from the LTED_2 replicate. Besides the +E2 and LTED replicates, two TAMR replicates were additionally chosen for deconvolution. Four out of five replicates (TAMR_1, 3, 4 and 5) were mainly dominated by two clonal populations characterized by a *dark red* or *blue* barcode indicative of selection of pre-existing clones. These two clonal populations were also present in the fifth replicate (TAMR_2), however seemed to have a selective disadvantage compared to two other clonal populations indicated by the *olive* and *turquoise* colors (Figure 10 A). In line, I chose the TAMR_2 and TAMR_4 replicates for further deconvolution. Sanger sequencing the barcode region of single cell clones obtained from the T47D TAMR_2 and TAMR_4 replicates revealed the integration of both *(dark red* and *dark blue*) barcode into the same cell of origin. Accordingly, the *dark blue* and *dark red* barcodes represented the same clonal population and not two distinct clonal populations as originally assumed. Accordingly, they are described as *purple* TAMR henceforth and shaded colors in Figure 14 highlight clones that are characterized by more than one barcode.

For MCF7, TAMR clones had shown up more heterogeneous compared to their T47D counterparts in the initial sequencing experiment (compare Figure 10), as more and different barcodes were enriched in the final cell pools. This pointed to a rewiring of persisting cells as predominant resistance mechanism. I selected only one TAMR replicate, namely TAMR_1, for deconvolution as I did not expect to recover several clones sharing the same barcodes from any replicate. In the LTED setting, I chose the LTED_2 and LTED_5 replicates as the LTED_2 replicate was made up of different clonal populations while the LTED_5 replicate was dominated by a single *red* clonal population pointing to a strong selective advantage of this clonal population.

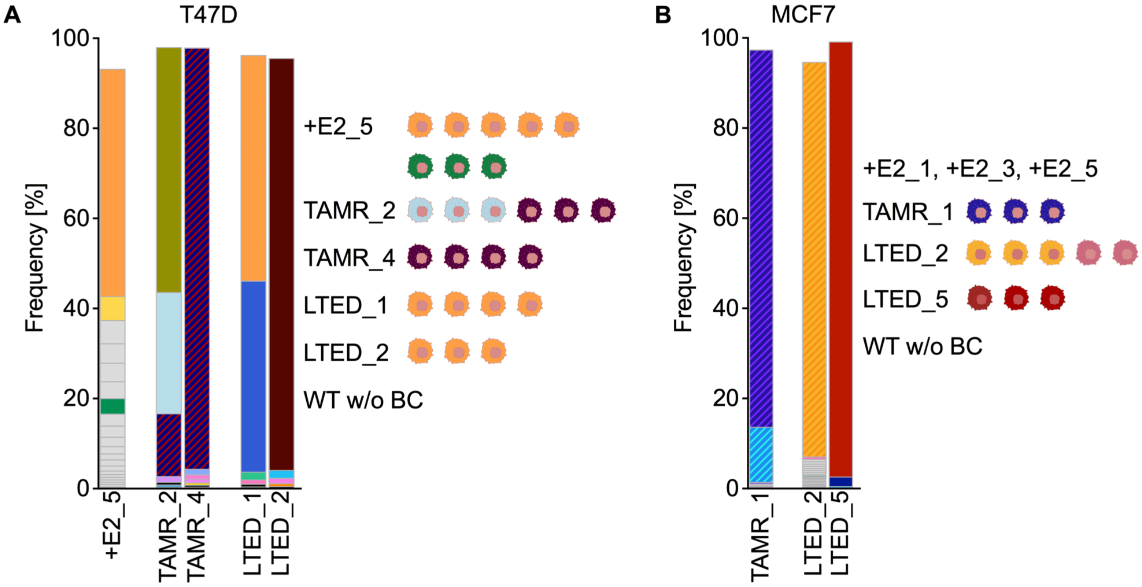


**Figure 14: Overview of isolated clones from selected T47D and MCF7 replicates.** Single cell clones are color coded according to the clonal population of the complex T47D (**A**) and MCF7 (**B**) cell pools they originated from. Color description as in Figure 10. Dashes: Clones characterized by at least two distinct barcodes. 'WT w/o BC' indicates non-barcoded T47D and MCF7 WT cell lines that were used as additional controls throughout the study. Barcode sequences, clone names and the color coding of all isolated clones presented here are shown in Supplementary Table 1. Cancer cells were created with BioRender.com.

From the original barcode analysis, MCF7 TAMR_1 was thought to be made up of at least 15 different clonal populations (Figure 10 B). After deconvolution and isolating single cells, the Sanger chromatograms were too complex to be analyzed and single clones likely contained multiple different barcode sequences (data not shown). Also, the Sanger chromatograms of MCF7 LTED_2 clones were too complex to be analyzed by Sanger sequencing, which again pointed to the presence of multiple barcodes having integrated into single cells as indicated by the shaded color coding in Figure 14. To determine the barcode composition of the individual MCF7 TAMR_1 and LTED_2 clones, I deeply sequenced their barcode region. To this end, I adjusted the NGS primers used by Bhang *et al.* [195] and PCR-amplified the barcode region to be sequenced on the MiSeq V2 300 Nano. In total, I subjected 41 MCF7 TAMR_1 and seven MCF7 LTED_2 clones for NGS of the barcode region. 37 of the isolated MCF7 TAMR_1 clones were characterized by the same 17 barcodes (Figure 15) and three clones were randomly chosen for further analysis. The two clonal populations isolated from the LTED_2 replicate were characterized by five and six integrated barcodes, respectively (data not shown). The isolated clones are accordingly referred to as *blue* TAMR_1 as well as *yellow* and *pink* LTED_2, in line with their color coding in Figure 14 B.

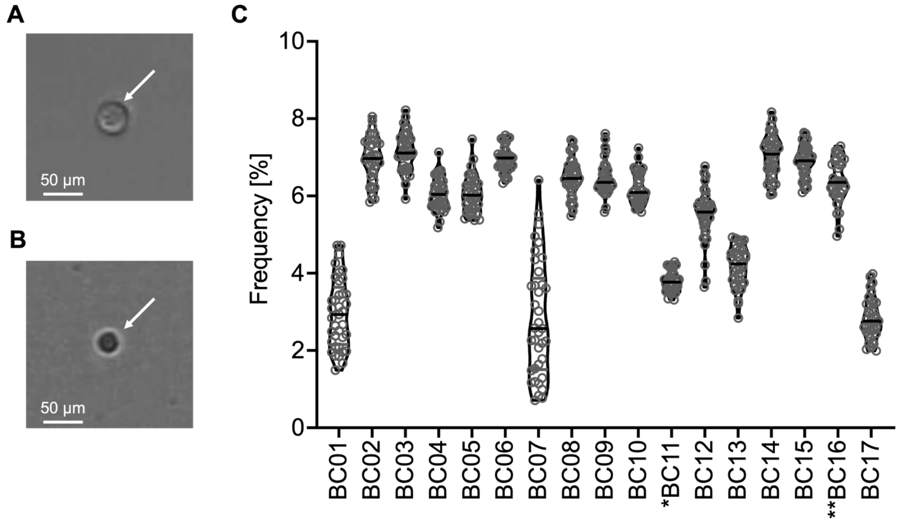


**Figure 15: Barcode composition analysis of MCF7 *blue* TAMR_1 clones**. **A** and **B** show a picture from the cell spotting and the attachment, respectively, highlighting the single cell origin of the B6 clone. **C** The percentual reads in the 37 individual clones included in the analysis are depicted for each detected barcode (BC) as violin plots. * Clustered reads of highly similar reads. ** One outlier was identified by the ROUT method (Q = 0.1%) and removed from the data. Barcode reads were determined from the sequencing data by Dr. Birgitta Michels.

In total, I established 28 resistant (T47D: 17, MCF7: 11) and 8 control (all T47D) clonal cell lines. Barcode sequences and color coding for all isolated clones are shown in Supplementary Table 1. MCF7 +E2 pools were by far more heterogeneous than T47D +E2 pools as evident from the barcode sequencing (Figure 10) and therefore I did not try to deconvolute the MCF7 +E2 pools but included the barcoded cell pools +E2_1, +E2_3 and +E2_5 as controls. Besides barcoded cell pools which had undergone clonal restriction during frequent passaging in limited culture for eight months (compare Figure 9 C and D), I included non-barcoded T47D and MCF7 WT cell pools as their cellular complexity was presumably larger compared to the barcoded control cell lines. Non-barcoded WT cell lines are indicated by ‘WT w/o BC’ (Figure 14).

In the next step, I was interested whether the enrichment of the isolated clones in the cell pools (Figure 14) was driven by biology or by an artifact. Specifically, enrichment of clones in the initial cell pools would have given individual clones a higher probability of survival compared to lowly abundant clones. To assess the potential enrichment of individual clones characterized by their barcodes, we plotted all barcode reads and specifically highlighted barcodes representing clones isolated from the individual replicates (Figure 16). Here, enrichment of isolated clones in the final T47D and MCF7 TAMR and LTED pools was not arbitrarily due to their enrichment during the initial expansion of untreated cells as neither of the barcodes characterizing the isolated clones was among the top enriched barcodes in the initial replicates. Exemplarily for T47D clones, the *turquoise* barcode characterizing a clonal population from the TAMR_2 replicate was among less abundant barcodes (exemplified by barcode ‘c’ in Initial cell pools in Figure 16 A). For MCF7, barcodes integrated into the *blue* TAMR_1 clones were among the least abundant barcodes (exemplified by barcode ‘m’ in Initial_1 and barcode ‘a’ in Initial_3 in Figure 16 B).

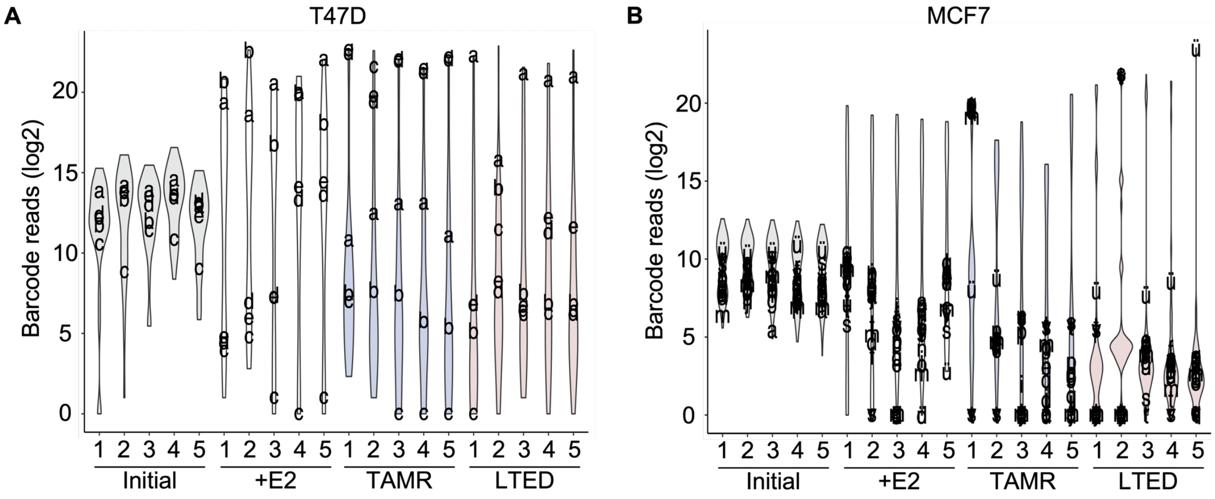
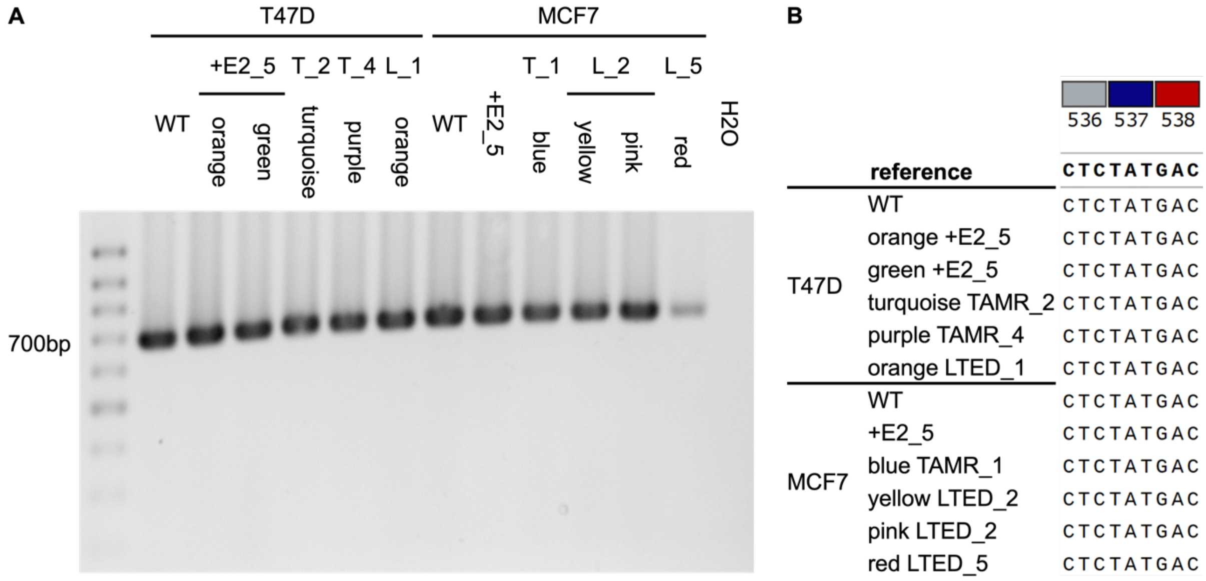


**Figure 16: Frequency of barcode reads in T47D and MCF7 replicates.** Barcodes whose sum were at least 5x$10^5$ for all T47D (**A**) or MCF7 (**B**) replicates were log2 normalized and are depicted as violin plots using ggplot2 in R [252, 278]. Barcodes from isolated TAMR and LTED clones are labelled alphabetically. Barcode labeling for T47D (**A**): a: *orange* +E2_5 and LTED barcode; b: *green* +E2_5 barcode; c: *turquoise* TAMR_2 barcode; d: *dark red* TAMR_2 and TAMR_4 barcode; e: *dark blue* TAMR_2 and TAMR_4 barcode. Barcode labeling for MCF7 (**B**): a-q: *blue* TAMR_1 barcodes; r-v: *yellow* LTED_2 barcodes; w-ö: *pink* LTED_2 barcodes; ü: *red* LTED_5 barcode. Analyses were performed by Dr. Birgitta Michels. The sequences of barcodes denoted here by letter coding are shown in Supplementary Table 1.

Next, I investigated if the integration of the viral sequences had disrupted gene coding or regulatory regions in the isolated clones. Therefore, I utilized the integration site analysis data from the T47D +E2_2, TAMR_2 and LTED_2, and MCF7 +E2_3 and LTED_5 cell line pools which was generated by Genewerk GmbH on behalf of Dr. Simone Borgoni [236]. Here, I correlated barcode frequencies in individual replicates, which were generated by Dr. Luca Penso Dolfin, and the frequency of individual integration sites in the respective cell lines as provided by Genewerk GmbH. Predicted integration sites were then validated by a PCR-based assay for isolated clones if possible. Specifically, I could show for the isolated MCF7 *red* LTED_5 clones that all clones were represented by the same barcode (Supplementary Table 1) and that they contained the same viral integration side in Chr4:110,457,931 (Supplementary Figure 2 B). Similarly, I could show that all clones isolated from T47D *turquoise* TAMR, *purple* TAMR, *orange* +E2_5 as well as *orange* LTED_1 and LTED_2 clonal populations contained the same barcodes (Supplementary Table 1) and the same viral integration sites characteristic for their respective clonal populations (Supplementary Figure 2 C and D). Two barcodes each integrated into intergenic sequences (> 15 kb distance to nearest gene) and two barcodes integrated into intronic genomic regions of unknown significance (*ENSG00000226965* and *ENSG00000253901*). To rule out potential effects of viral integrations on *ENSG00000226965* and *ENSG00000253901* expression, I utilized our RNA-Seq data. Here, no evidence of expression was found for either gene. One barcode characterizing the T47D *purple* TAMR clones integrated into intron 1 of *TP63*. Gene expression analysis showed no significant deregulation of *TP63* expression in T47D *purple* TAMR clones compared to the *orange* and *green* +E2_5 control cells. One viral particle carrying the *orange* barcode integrated into intron 4 of *DNAJC24*. The gene expression data showed significantly higher expression in *orange* LTED clones compared to *orange* +E2_5 clones. Since both clonal populations are characterized by the same integration site, *DNAJC24* upregulation appeared to be a treatment (-E2) specific effect. Taken together, these findings strongly suggest that a potential role of the viral integration sites in the selection of specific clones and in resistance acquisition at least for these clones can be ruled out (Supplementary Table 2).

### 4.6 Endocrine therapy resistance is not driven by *ESR1* hotspot mutations

After having isolated single clones and verified their clonality, I was interested whether resistance to Tamoxifen and/or estrogen deprivation was the result of a commonly observed resistance mechanism, namely *ESR1* hotspot mutations affecting codons 536 to 538 (RefSeq ID: NM_000125.4). I PCR-amplified and sequenced the genomic region covering these mutational hotspots in a representative clone from every barcode, replicate and condition. Additionally, I included non-barcoded T47D and MCF7 WT in the analysis. All utilized cell lines and clones showed wildtype sequence for this *ESR1* hotspot region and no mutations (Figure 17). Accordingly, endocrine therapy resistance in isolated clones was not driven by *ESR1* hotspot mutations covering codons 536 to 538.



| | reference | CTCTATGAC |
|---|---|---|
| T47D | WT | CTCTATGAC |
| | orange +E2_5 | CTCTATGAC |
| | green +E2_5 | CTCTATGAC |
| | turquoise TAMR_2 | CTCTATGAC |
| | purple TAMR_4 | CTCTATGAC |
| | orange LTED_1 | CTCTATGAC |
| MCF7 | WT | CTCTATGAC |
| | +E2_5 | CTCTATGAC |
| | blue TAMR_1 | CTCTATGAC |
| | yellow LTED_2 | CTCTATGAC |
| | pink LTED_2 | CTCTATGAC |
| | red LTED_5 | CTCTATGAC |

**Figure 17: Sequencing of *ESR1*.** PCR amplification (**A**) and sequence alignment using Snapgene (**B**) of genomic *ESR1* region covering the hotspot codons 536 to 538 for non-barcoded control cell lines (WT) as well as endocrine therapy sensitive and resistant cell lines and clones. A single representative clone was used for each cell

line or isolated clonal population. T_2: TAMR_2. T_4: TAMR_4. L_1: LTED_1. T_1: TAMR_1. L_2: LTED_2. L_5: LTED_5. Reference sequence was obtained from RefSeq ID: NM_000125.4 (accessed 13.12.22).

## 4.7 Barcoded clonal populations can be robustly distinguished

So far, I could show that isolated clones were enriched in the final cell pools because of biology rather than stochasticity and that clones characterized by the same barcode were of clonal origin. Further, resistance to Tamoxifen and estrogen deprivation was not driven by *ESR1* hotspot mutations covering codons 536-538, a commonly observed endocrine therapy resistance mechanism. I subsequently grew out the single cell clones for further analysis.

All single cell pools had originated from an individual cell dispensed either manually or by a cell spotter. Therefore, the single starting cell had to undergo many cell divisions and this process took between three to seven months for individual clones. As described in the Materials and Method section (Table 2), treatment was additionally withdrawn from some clones for an initial four-week period since these clones had previously not formed colonies after singularization in the presence of the applied therapy. In total, individual clones were passaged for over twelve months between barcoding of cell pools and final analysis of individual clones. Accordingly, I was interested whether clonal cell populations with the same barcode behaved similarly and could be distinguished from clonal cell populations characterized by different barcodes or if evolution on cellular level led to lineage convergence of different clonal populations. To answer this question, we utilized the RNA-Seq data and a hierarchical clustering approach of the 200 most variable genes (Figure 18).

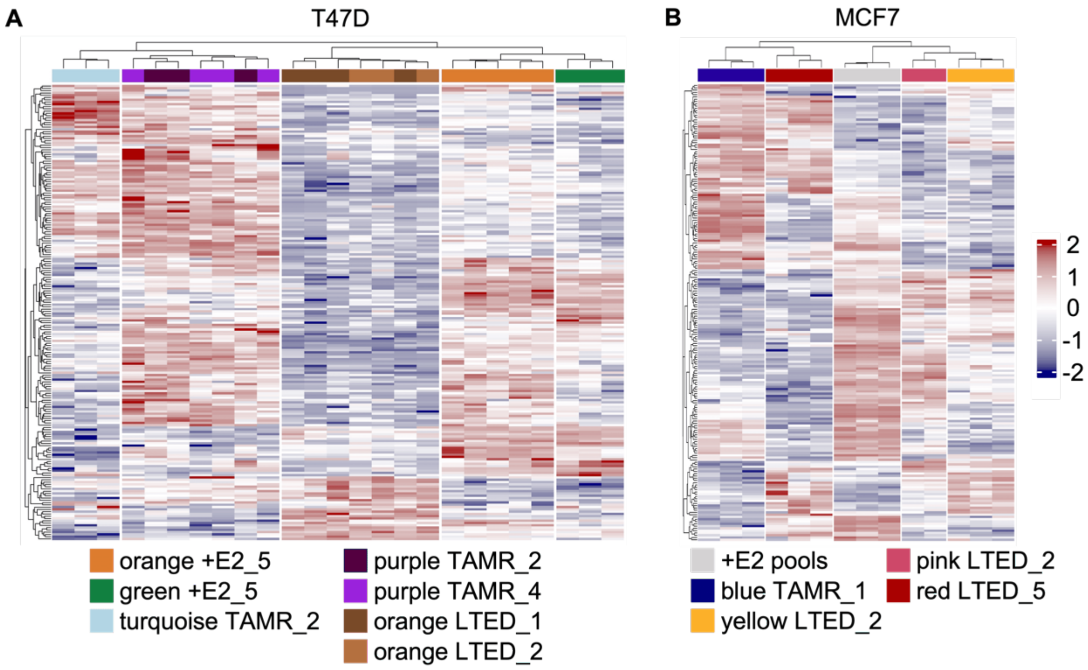


**Figure 18: Clustering of clones based on the top 200 most variable genes.** The R package ComplexHeatmap (version 2.10.0) was used to generate heatmaps of the top 200 most variable genes in the T47D (**A**) and MCF7 (**B**) cell lines. For genes and columns, hierarchical clustering was applied (Euclidean distance,

complete linkage). The illustrated values are scaled log2-counts per million (CPM) (TMM normalized) values. Analyses were performed by Dr. Efstathios-Iason Vlachavas.

Based on these most variable genes, clonal populations with the same barcode clustered together and could clearly be distinguished from clonal populations characterized by different barcodes. For the T47D cell line models, *orange* and *green* +E2_5 clones originated from the same parental replicate (+E2_5) but could be clearly separated. Even more interestingly, *turquoise* TAMR_2 clones clustered separately from *purple* clones isolated from TAMR_2 and TAMR_4 replicates. Accordingly, the clonal identity as defined by the respective barcode (in this case *turquoise* and *purple*) was more discriminatory than the complex cell pool (in this case TAMR_2 and TAMR_4), from which cells had been isolated. *Orange* clones formed two clusters based on the treatment (+E2, LTED) they had originated from. Of note, *orange* clones isolated from LTED_1 and LTED_2 replicates clustered together highlighting their similarity across biological replicates (Figure 18 A). Two distinct barcodes/clonal populations (*yellow* and *pink*) were also isolated from the MCF7 LTED_2 replicate and these clustered separately from each other (Figure 18 B). Given the separation of distinct clonal populations isolated from control (+E2) and treatment conditions (TAMR, LTED) in both cell line models, it is unlikely that the clustering approach was affected by the different media conditions of the clones. In conclusion, the isolated clonal populations were stable in their trajectories over time and no signs of lineage change were detected.

## 4.8 Endocrine therapy resistant clones display persisting or fast proliferating phenotypes

After having validated the lineage stability of the isolated clones, I assessed the proliferation of all 28 resistant (T47D: 17, MCF7: 11) and 13 control (T47D: 9, MCF7: 4) clones and cell pools. As performed previously for the barcoded +E2 pools (Figure 8), control (+E2) clones and pools were pre-treated for 14 days by estrogen deprivation. Additionally, these cell lines were also pre-treated with 4-OHT as further comparison (Figure 19).

The barcode analysis of the T47D +E2_5 cell line pool (Figure 10 A) revealed that the *orange* clonal population made up roughly 50% of the whole replicate while the *green* clonal population made up less than 5%. Hence, a stronger proliferation of the *orange* clones compared to the *green* clones had been expected. However, *green* clones isolated from the T47D +E2_5 replicate proliferated significantly stronger than the *orange* clones originating from the same parental replicate. Moreover, *orange* clones roughly proliferated at least as much as non-barcoded heterogeneous T47D WT cells (Figure 19 A). These findings point to selective advantage of *orange* over *green* clones in the T47D +E2_5 replicate that is not based on cellular proliferation.

21 days of treatment with 4-OHT led to a complete proliferative block of non-barcoded T47D WT cells. Most clones isolated from the two distinct (*turquoise* and *purple*) T47D TAMR clonal populations proliferated stronger than the pre-treated T47D WT cells. Specifically, T47D *turquoise* TAMR_2 B3 and *purple* T47D TAMR_2 C4 and H2 however showed weak proliferation while the other two *turquoise* and four *purple* T47D TAMR clones showed significantly stronger proliferation compared to the pre-treated T47D WT cells (Figure 19 B).

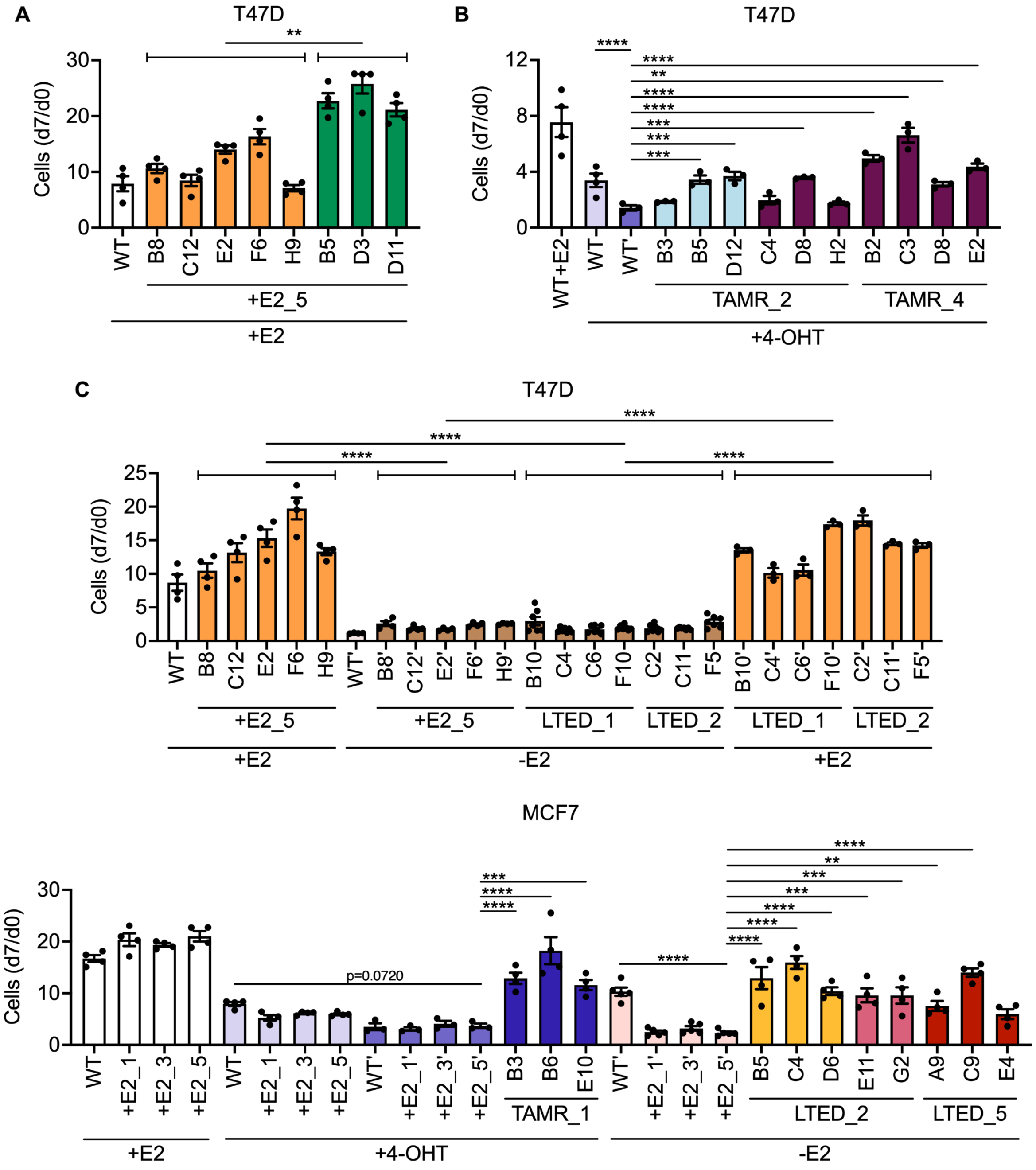


**Figure 19: Proliferation of T47D and MCF7 clones. A** Proliferation of T47D *orange* and *green* +E2_5 clones. **B** Proliferation of T47D *turquoise* and *purple* TAMR clones. **C** Proliferation of T47D *orange* +E2_5 and LTED clones. Plates were coated with poly-L-lysine for the proliferation assay. **D** Proliferation of MCF7 *blue* TAMR_1, *yellow* and *pink* LTED_2 and *red* LTED_5 clones. **A**-**D** Cells were kept in or treated with the indicated media (+E2, +4-OHT, -E2). Cell numbers were determined by microscopy-based nuclei counting at day 0 (d0) and day 7 (d7). WT indicates non barcoded cells. ' indicates 14 day pre-treatment with estrogen supplementation (+E2), +4-OHT or by estrogen deprivation (-E2) and then continued treatment for the duration (7 days) of the proliferation assay as indicated. n ≥ 3 with ≥ 5 technical replicates for all assays. Shown are mean ± SEM. **A** and **C** Clones with

the same barcode originating from the same treatment were grouped and the grouped clones compared. **A** ** represents $p<0.01$ by unpaired two-tailed t-tests. **B** to **D** * represents $p<0.05$, ** represents $p<0.01$, *** represents $p<0.001$ and **** represents $p<0.0001$ as determined by one-way ANOVA with Tukey or Dunnett multiple comparisons test. **D** Statistical comparison is shown against +E2_5' in the +4-OHT and -E2 treated comparisons. The trend is the same for +E2_1', +E2_3' and E2_5' (data not shown). **A**-**D** Labeling such as 'B8' designates individual clones as in Supplementary Table 1.

For the T47D cell line, *orange* clones deprived of estrogen for over one year (LTED) showed an almost non-proliferative phenotype resembling dormant/persister cells. A similar phenotype could be induced in T47D WT cells and *orange* clones isolated from the +E2_5 replicate after a total of 21 days of estrogen deprivation (-E2). In accordance with a possible persister phenotype, T47D *orange* LTED clones regained a strong proliferative phenotype within 21 days of estrogen supplementation and proliferated similarly to *orange* clones isolated from the +E2_5 pool and kept with estrogen supplementation at all times (+E2, Figure 19 C).

In the MCF7 model, pretreatment with 4-OHT or estrogen deprivation led to a strong and uniform reduction of proliferation in the barcoded +E2 pools. Non-barcoded MCF7 WT cells however were hardly affected by prolonged estrogen deprivation, which might be due to the much higher complexity of this cell population compared to the +E2 pools having passed clonal restriction during frequent passaging in limited culture. All isolated *blue* TAMR_1 as well as *yellow* and *pink* LTED_2 clones proliferated significantly stronger under the applied endocrine therapy compared to the respective +E2 cell pools treated with either 4-OHT or estrogen deprivation. Similarly, two out of three isolated *red* LTED_5 clones proliferated significantly stronger than the pre-treated +E2 cell pools, while the third clone (E4) showed a stronger, but statistically not significant, proliferation (Figure 19 D).

Given the divergence in response to estrogen deprivation between non-barcoded (WT) and barcoded MCF7 control cell lines (+E2_1, +E2_3, +E2_5), I additionally tested all *orange* and *green* T47D +E2_5 clones for their respective responses to 4-OHT treatment and estrogen deprivation (Supplementary Figure 3). All T47D *orange* and *green* +E2_5 clones except for a single clone showed almost complete abrogation of cellular proliferation under short-term 4-OHT treatment as observed also for pre-treated non-barcoded T47D WT cells (shown in Figure 19 B). Only T47D *orange* +E2_5 B8 showed minor but still significantly stronger proliferation under short-term 4-OHT treatment compared to non-barcoded T47D WT cells (Supplementary Figure 3 A). Additionally, 21 day estrogen deprivation uniformly abrogated cellular proliferation in *orange* and *green* T47D +E2_5 clones (Supplementary Figure 3 B). Given the uniform responses of T47D WT cells as well as T47D *orange* and *green* +E2_5 clones to short-term 4-OHT treatment and estrogen deprivation (-E2), I kept T47D WT as control in Figure 19 B.

Taken together, MCF7 TAMR and LTED clones showed stronger proliferation than their respective T47D counterparts. Moreover, individual clones showed some variation in their proliferative responses. Despite their variation on phenotypic levels, the unsupervised hierarchical clustering approach robustly distinguished different populations based on their clonal origin (Figure 18). Hence, the separation did also not represent an artefact originating from different levels of proliferation further highlighting the biological differences of the individual clonal populations. In line, gene set enrichment analysis of the top 200 most variable genes showed no enrichment of Hallmark gene sets related to proliferation or cell cycle progression (Supplementary Table 3 and Supplementary Table 4).

### 4.9 Differential pathway activation in endocrine therapy resistant clones

After I confirmed stable clonal trajectories and highlighted the phenotypic heterogeneity of the T47D and MCF7 TAMR and LTED clones, I next wanted to uncover pathway-wide alterations. Therefore, the RNA-Seq data was further analyzed by Dr. Efstathios-Iason Vlachavas using PROGENy [261, 262] to elucidate differential activities of 14 major pathways in the TAMR and LTED clones. The pathway activation pattern of T47D and MCF7 TAMR and LTED clones are shown in Figure 20 and Figure 21, respectively.

There, pathway activation or repression in each comparison does not mean an absolute activation or repression in either condition but rather elevated or reduced activities relative to the other condition in the binary comparison. Along these lines, *orange* and *green* +E2_5 clones were used as reference to find deregulated pathways in T47D TAMR and LTED clones. Two reference clones (*orange* and *green* +E2_5) were employed to correct for potential clone-specific effects. *Orange* LTED clones were compared also to *orange* +E2_5 clones since both clonal populations descended from the same cell of origin and this comparison might highlight clonal adaptations to estrogen deprivation. As stated above, the MCF7 +E2 replicates were considered too complex to be deconvoluted (Figure 10 B), and the three control cell lines +E2_1, +E2_3 and +E2_3 were used as controls.

*Turquoise* and *purple* T47D TAMR clones showed shared activation of androgen, hypoxia, transforming growth factor (TGF) ß, and vascular endothelial growth factor (VEGF) signaling. For *turquoise* T47D TAMR_2 clones, their activation scores ranging from 1.69-1.87 marginally missed statistical significance. Activation of p53 signaling was the only pathway to be upregulated in T47D *orange* LTED_1 and LTED_2 clones compared to (*green* and) *orange* +E2_5 clones. All T47D TAMR and LTED clones showed significant downregulation of estrogen, MAPK and Janus kinases-signal transducer and activator of transcription (JAK-STAT) signaling compared to the *orange* and *green* +E2_5 clones (Figure 20).

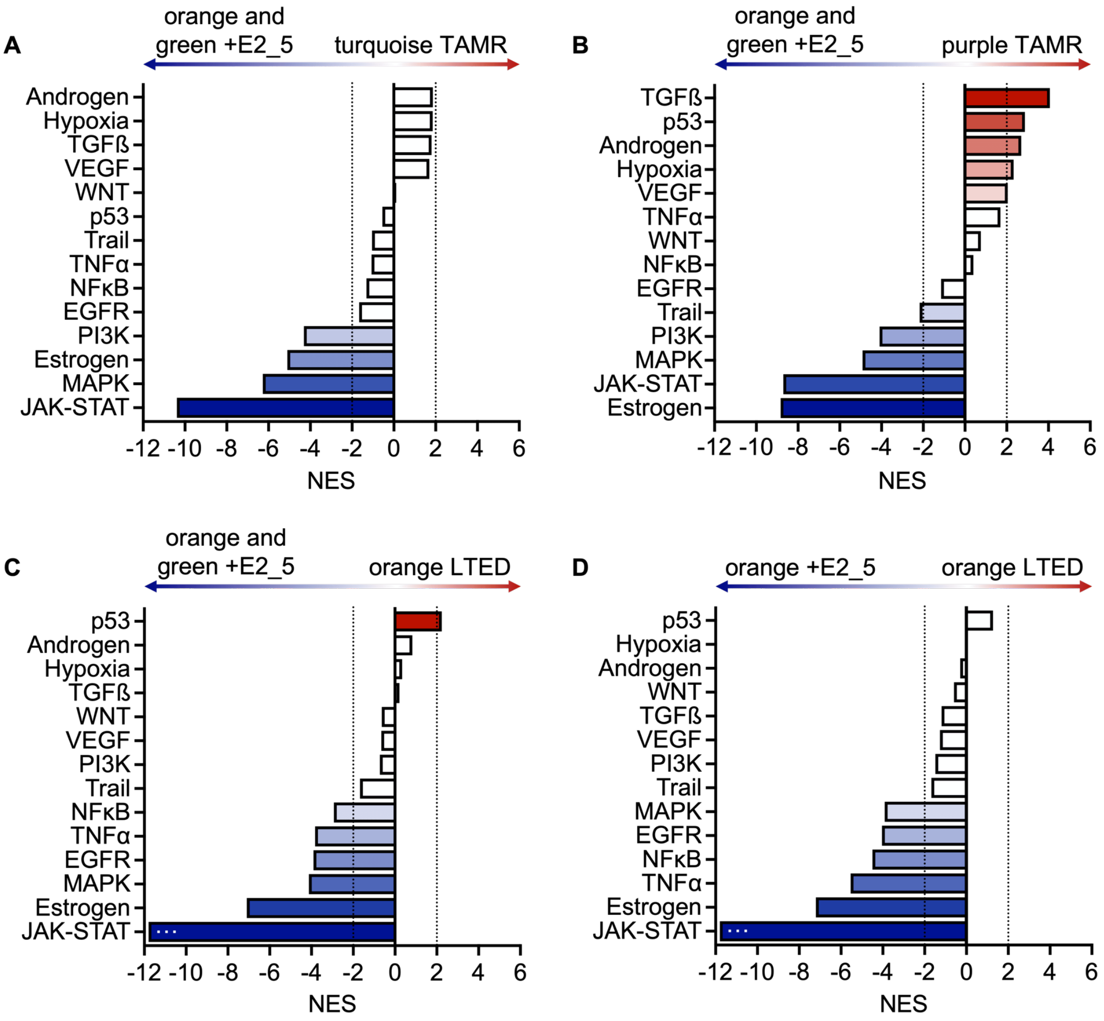


**Figure 20: Pathway activity analysis of T47D clones.** Absolute NES scores ≥ 2 were considered significant as indicated by dashed lines. Significantly activated pathways in TAMR and LTED clones are highlighted in red whereas significantly downregulated pathways in TAMR and LTED clones are highlighted in blue. Analyses were performed by Dr. Efstathios-Iason Vlachavas.

In the MCF7 LTED clones, JAK-STAT was the most prominently upregulated pathway while no activation was observed for MCF7 *blue* TAMR_1 clones. NFκB signaling on the one hand was among the strongest activated pathways compared to the +E2 control pools in all MCF7 TAMR and LTED clones. Estrogen signaling on the other hand was the only pathway to be significantly downregulated in all MCF7 TAMR and LTED clones (Figure 21). The significant repression of estrogen signaling in all T47D and MCF7 TAMR and LTED clones was further in accordance with the absence of *ESR1* hotspot mutations covering codon 536 to 538 (Figure 17), which are known to induce estrogen independent signaling [95-97, 279-281]. This limited overlap of commonly deregulated pathways highlighted the strong heterogeneity in the different T47D and MCF7 TAMR and LTED clones.

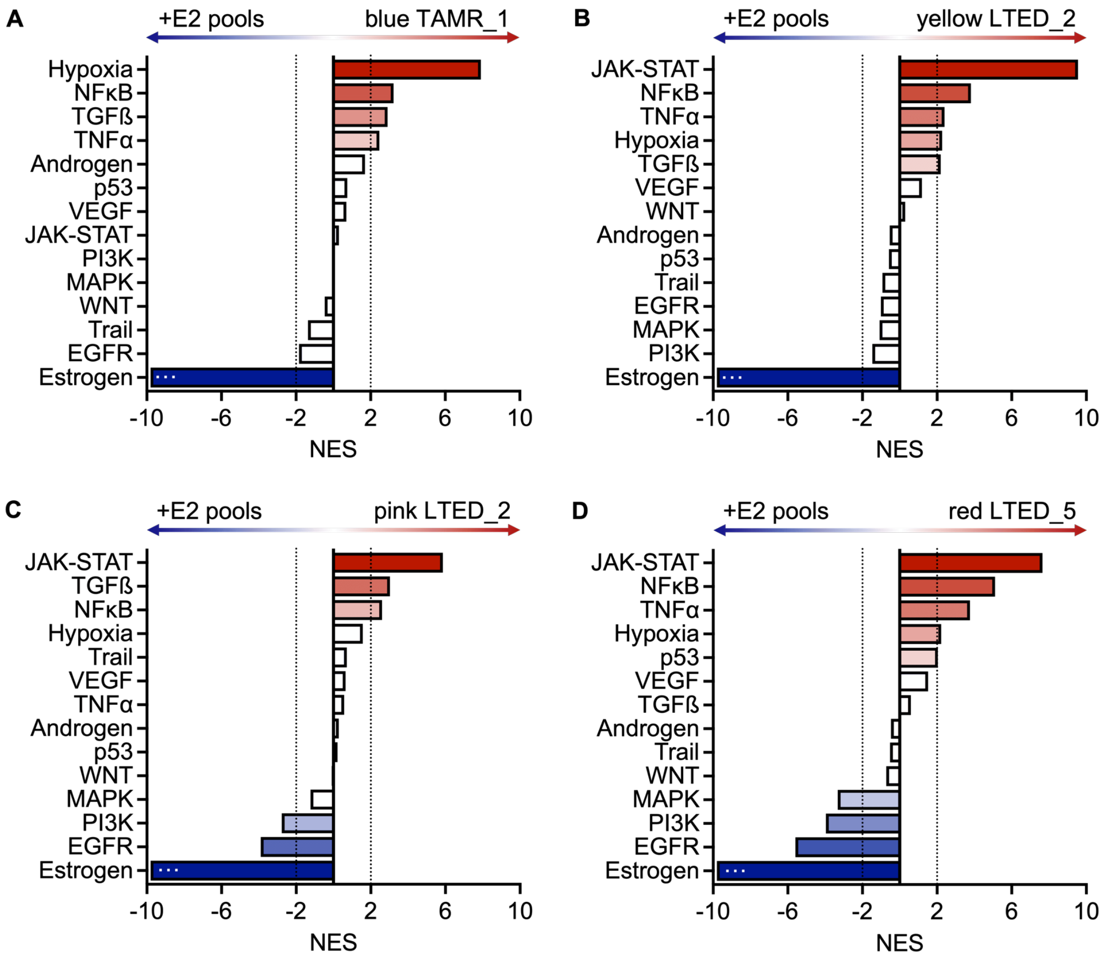


**Figure 21: Pathway activity analysis of MCF7 clones.** Absolute NES scores ≥ 2 were considered significant as indicated by dashed lines. Significantly activated pathways in TAMR and LTED clones are highlighted in red whereas significantly downregulated pathways in TAMR and LTED clones are highlighted in blue. Analyses were performed by Dr. Efstathios-Iason Vlachavas.

## 4.10 Differential transcription factor activity in endocrine therapy resistant clones

As PROGENy analysis [261, 262] had revealed differential pathway activations, I was next interested in differential activation or repression of TFs in the T47D and MCF7 TAMR and LTED clones. To this end, TF activities were estimated from the expression of their corresponding downstream targets using DoRothEA [263]. This analysis was performed by Dr. Efstathios-Iason Vlachavas. Tables showing the top 10 activated and repressed TFs in T47D and MCF7 TAMR and LTED clones compared to their respective control clones and cell lines are listed in the appendix (Supplementary Table 5 and Supplementary Table 6, respectively). Key findings from the analysis of TF activities are shown in Figure 22.

In line with the findings for the activation of different signaling pathways, heterogeneity was further evident from the activity of TFs in the T47D and MCF7 TAMR and LTED clones. On a global scale, significantly more TFs were activated in the MCF7 TAMR and LTED clones compared to T47D TAMR and LTED clones while conversely, the opposite was true for downregulated TFs between the two cell line models (Figure 22 A). Only three TFs, namely

ELF3, GRHL2 and KLF5, were significantly activated in all T47D and MCF7 TAMR and LTED clones (Figure 22 B). No TF was significantly repressed in all T47D and MCF7 TAMR and LTED clones (Figure 22 C).

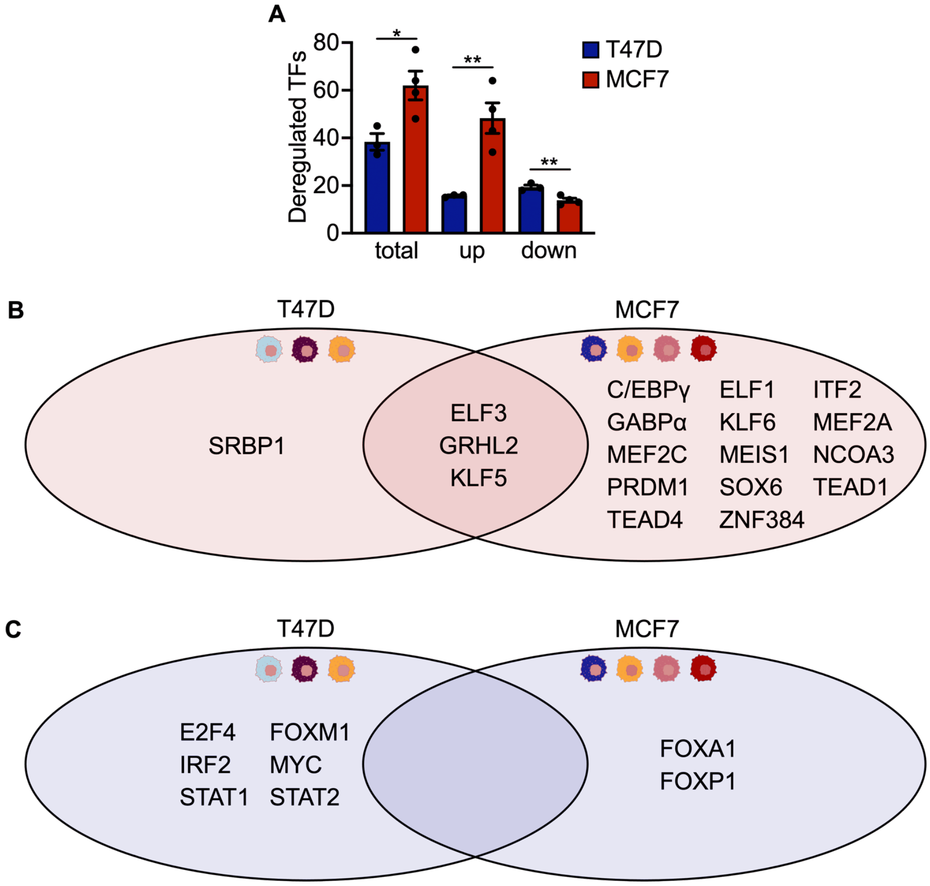


**Figure 22: Activated and repressed transcription factor activities in a comparison of T47D and MCF7 clones. A** Absolute numbers of significantly deregulated TFs in T47D and MCF7 TAMR and LTED clones. ‘up’ denotes activation while ‘down’ denotes repression. * represents $p<0.05$ and ** represents $p<0.01$ as determined by unpaired two-tailed t-tests. Commonly and specifically activated (**B**) and repressed (**C**) TFs in all T47D or MCF7 TAMR and LTED clones. Only significant hits (absolute activity score ≥ 2) are shown. T47D TAMR and LTED clones were compared to *orange* and *green* +E2_5 control clones. MCF7 TAMR and LTED clones were compared to the +E2 control cell lines. Analyses were performed by Dr. Efstathios-Iason Vlachavas. Cancer cells were created with BioRender.com.

## 4.11 Differential kinase activity in endocrine therapy resistant clones

RNA-Seq analysis highlighted tumor heterogeneity and allowed deep insights into molecular changes in the T47D and MCF7 TAMR and LTED clones on gene expression level. In the next step, I complemented the analysis of transcriptional changes with the analysis of differential activation or repression of kinases and phosphatases. Therefore, Mass spectrometry analysis of previously enriched phospho-proteins was performed by Luisa Schwarzmüller. Next, I assessed the reproducibility of detected phosphosites in the T47D and MCF7 replicates. Some 46-74% of all phosphosites detected were common in all biological replicates of clones carrying the same barcode for both cell lines (Supplementary Figure 4). Afterwards, we performed unsupervised hierarchical clustering as for the RNA-Seq shown above, but this time for all the phosphosites detected (Figure 23). Generally, the clustering on phosphosite level confirmed the clustering on RNA-Seq level (Figure 18) as clones with the same barcode could be robustly distinguished from clones with other barcodes based on their phosphorylation profile. Of note, T47D *purple* TAMR_2 C4 and H2 clones clustered separately from the other T47D *purple* TAMR clones (Figure 23 A).

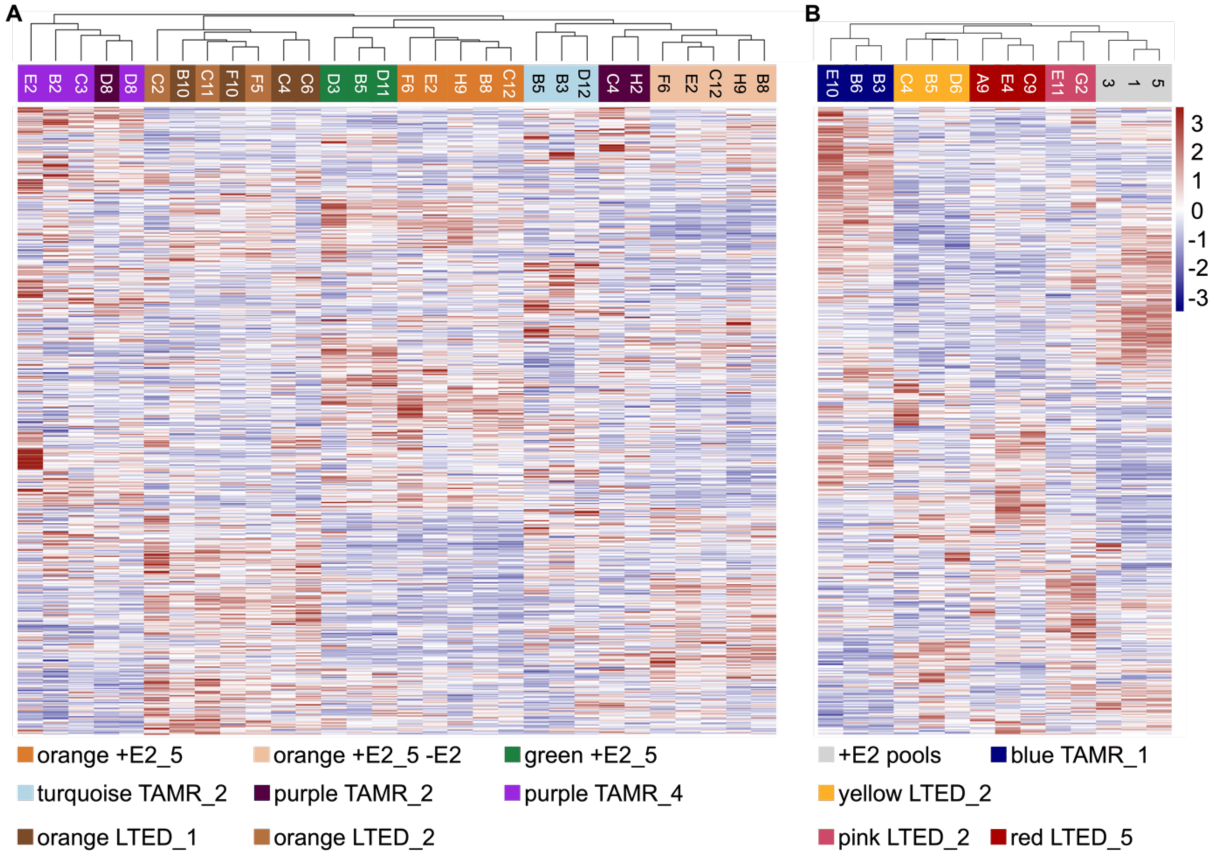


**Figure 23: Clustering of clones based on the consistently identified phosphosites.** The R package Pheatmap (version 1.0.12) was used to generate heatmaps of consistently identified phosphosites in the T47D (**A**) and MCF7 (**B**) cell lines. For phosphosites (rows) and cell lines (columns), hierarchical clustering was applied (Euclidean distance). The illustrated values are z-scored peptide intensities. Analyses were performed by Luisa Schwarzmüller.

After the initial quality control which highlighted the strong correlation between clones sharing the same barcode again, kinase activities were determined based on the phosphorylation of their substrates in the next step by Dr. Efstathios-Iason Vlachavas. Tables showing the top 10 activated and repressed kinases and phosphatases in T47D and MCF7 TAMR and LTED clones compared to their respective control clones and cell lines are listed in the appendix (Supplementary Table 7 and Supplementary Table 8, respectively).

Globally, significantly more kinases and phosphatases were activated in MCF7 TAMR and LTED clones compared to T47D TAMR and LTED clones (Figure 24 A). Similar to the results observed on transcriptional level (Figure 22 B), there was almost no overlap in significantly activated and repressed kinases and phosphatases between the T47D and MCF7 cell line models as CDK1 was the only kinase whose activity was commonly repressed in all T47D and MCF7 TAMR and LTED clones (Figure 24 C).

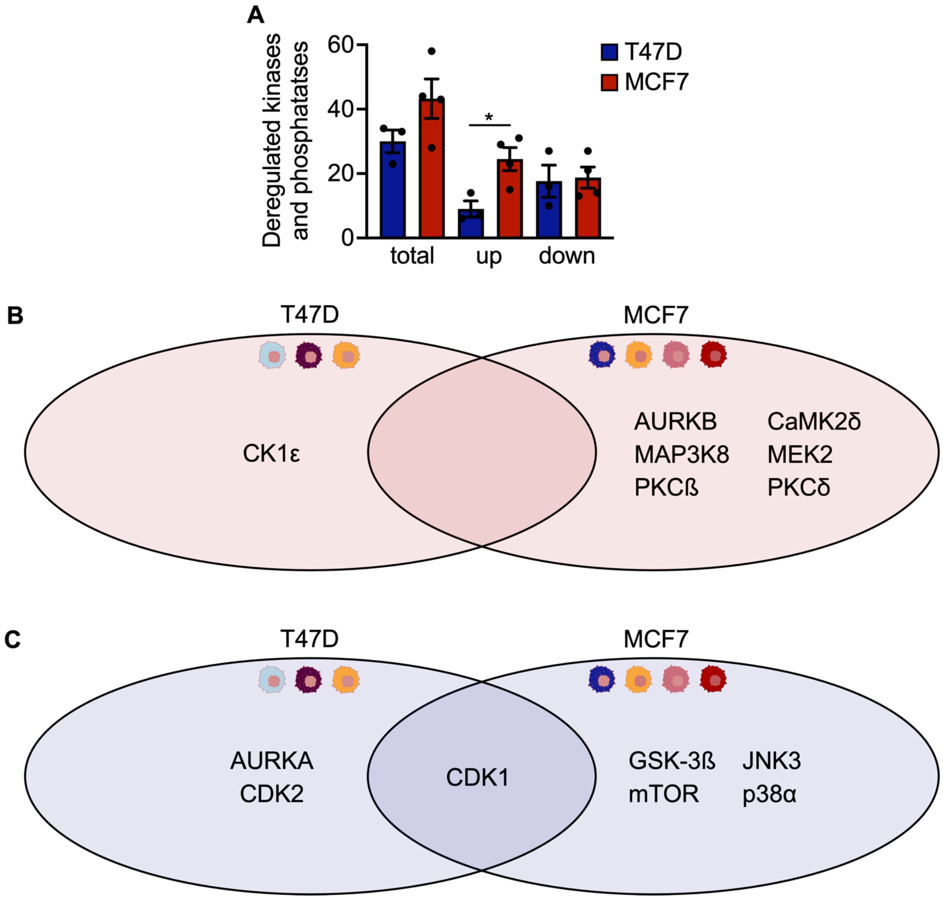


**Figure 24: Activated and repressed kinase and phosphatase activities in a comparison of T47D and MCF7 clones. A** Absolute numbers of significantly deregulated kinases and phosphatases in T47D and MCF7 TAMR and LTED clones. ‘up’ denotes activation while ‘down’ denotes repression. * represents $p<0.05$ as determined by unpaired two-tailed t-tests. Commonly and specifically activated (**B**) and repressed (**C**) kinases and phosphatases in all T47D or MCF7 TAMR and LTED clones. Only significant hits (absolute activity score ≥ 2) are shown. T47D TAMR and LTED clones were compared to *orange* and *green* +E2_5 control clones. MCF7 TAMR and LTED clones were compared to the +E2 control cell lines. Analyses were performed by Dr. Efstathios-Iason Vlachavas. Cancer cells were created with BioRender.com.

## 4.12 T47D *orange* clones model cycling persister cells and depend on AKT signaling

After I highlighted the inter- and intratumor heterogeneity on proliferation, pathway activation as well as TF and kinase activity level, I was interested in individual potential contributors to endocrine therapy resistance. In the T47D cell line model I was particularly interested in the *orange* clones. These cells were isolated from a treatment-naïve replicate (+E2_5) and from LTED replicates (LTED_1 and LTED_2). Since both isolated populations originated from the same cell of origin (Supplementary Figure 2 D), they allowed me to study the effect of estrogen supplementation and deprivation on the same clonal background.

As seen above (Figure 19 C), short-term estrogen deprivation of treatment-naïve *orange* +E2_5 clones resulted in a uniform and significant reduction of proliferation. Even after estrogen deprivation for months, T47D *orange* LTED clones still presented little proliferative capacity. When the therapeutic pressure was lifted however and cells were kept in estrogen supplemented (+E2) media, the LTED clones quickly regained a strong proliferative phenotype. In the next step, I was interested whether the strong reduction of total cells by estrogen withdrawal was the result of a decrease in cells going through S-phase or an increase in dying/dead cells. I therefore performed EdU and DAPI stainings to label cells going through S-phase and dying/dead cells, respectively (Figure 25).

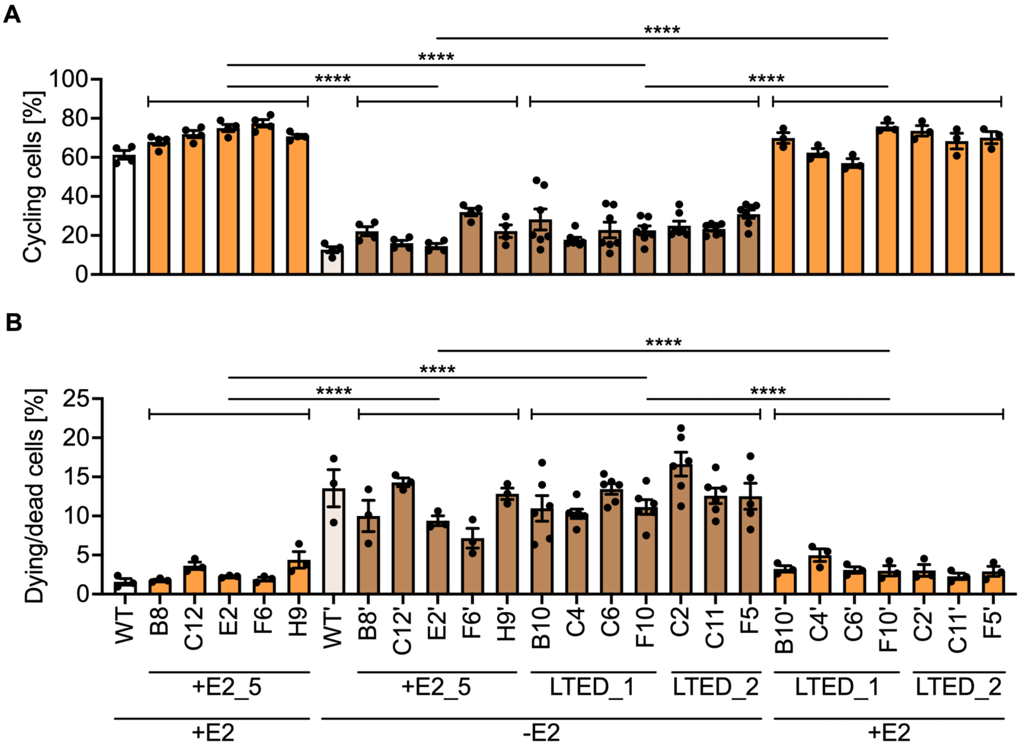


**Figure 25: Cycling and dying/dead cells for T47D *orange* clones. A** Percentage of cycling cells within 21 h were determined by EdU incorporation. **B** Dying/dead cells were determined by DAPI staining. **A** and **B** Cells were kept in or treated with the indicated media (+E2, -E2). T47D WT cells without barcode were used as additional controls. ' indicates 14 day pre-treatment by estrogen deprivation (-E2) or supplementation (+E2) and then continued treatment for the duration (4 days) of the assays. n ≥ 3 with ≥ 5 technical replicates for both assays. Highlighted are mean ± SEM. Clones with the same barcode originating from the same treatment were grouped and the grouped clones compared. **** represents p<0.0001 as determined by one-way ANOVA with Tukey multiple comparisons test.

Around 60-75% of T47D *orange* +E2_5 cells and non-barcoded T47D WT control cells went through S-phase under normal (+E2) growth conditions within the tested 21 h. Upon estrogen withdrawal for 14 days and further estrogen deprivation for four days during the assay, the percentage of cells going through S-phase was significantly reduced to around 10-30%. *Orange* LTED clones showed similar levels of cycling cells compared to the short-term deprived WT and *orange* +E2_5 cells (18-30%). Upon short-term estrogen supplementation (+E2) for a total of 19 days, the percentage of *orange* LTED cells actively going through S-phase was significantly increased to 58-75%. Accordingly, short-term estrogen deprivation in T47D WT and *orange* +E2_5 as well as short-term estrogen supplementation of *orange* LTED clones induced the same pattern of cells going through S-phase as their long-term counterparts (Figure 25 A). A similar pattern was also observed for the level of dying/dead cells. Under normal growth conditions with estrogen supplementation (+E2), around 1-5% of T47D WT and *orange* +E2_5 cells as well *orange* LTED cells supplemented with estrogen for a short-term period were dying/dead. Conversely, this percentage increased to 7-17% of short-term estrogen deprived T47D WT and *orange* +E2_5 clones as well as the *orange* LTED clones (Figure 25 B). Taken together, these data suggest that T47D *orange* clones can induce a reversible cycling persister phenotype. Clinically and as illustrated in Figure 26, tumors may initially respond very well to the applied therapy as illustrated by the strong reduction in cell numbers. However, few cells may survive and persist the treatment. Consequently, when the treatment is lifted, persisting cells may regain a proliferative phenotype again. Potentially, the regrown tumor may be sensitive to endocrine therapy again. The concept of persisting cells is supported by clinical evidence as distant relapses in breast cancer patients may recur up to 20 years after surgery [282-284].

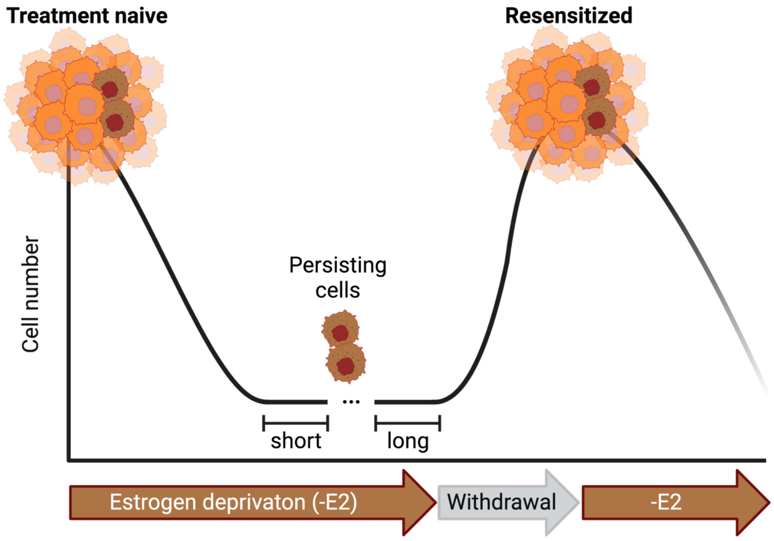


**Figure 26: Short- and long-term persister cells.** Adapted from Cabanos, *Cancers* 2021 [285]. Created with BioRender.com.

Consequently, I was interested in the differences between the short- and long-term estrogen deprived T47D *orange* clones as they may represent short- and long-term persisting clones as shown in Figure 26, respectively. Additionally, short-term estrogen deprived *orange* +E2_5 clones may represent the entry point of this persister phenotype. To this end, I further utilized the Mass spectrometry-based phospho-proteomics analysis for which I had also included T47D *orange* +E2_5 clones which were deprived of estrogen for 14 days.

Kinase and phosphatase activity analysis as described above revealed AKT1-3 as the strongest activated kinases in short-term estrogen deprived T47D *orange* +E2_5 clones compared to *orange* LTED clones (Figure 27). AKT1-3 were not altered (Kinase activities between -0.22 and -0.79) in T47D *orange* LTED clones compared to the *orange* +E2_5 clones when both cell lines were kept in their respective media (LTED: -E2, +E2_5: +E2) (data not shown).

Accordingly, AKT1-3 may have been crucial only in the short-term adaptation of *orange* clones to estrogen deprivation (Figure 27).

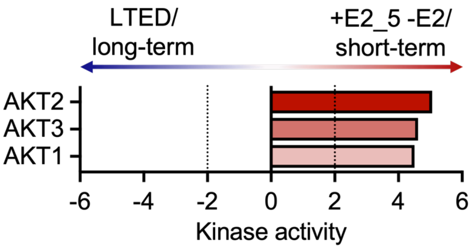


**Figure 27: Activation of AKT1-3 in short-term estrogen deprived T47D *orange* clones.** Absolute activity scores ≥ 2 were considered significant as indicated by dashed lines and significant activation in short-term estrogen deprived *orange* clones is highlighted in red. Analysis was performed by Dr. Efstathios-Iason Vlachavas.

Next, I challenged T47D *orange* clones with MK-2206, an AKT inhibitor currently investigated in clinical trials [286]. T47D +E2_5 E2 and LTED_1 C6 clones were treated with increasing concentration of MK-2206 for 1 h as previously described [244]. Inhibition of AKT signaling was assessed by quantification of Thr246 phosphorylation in the AKT downstream target PRAS40 [287]. Treatment with 100 nM MK-2206 reduced PRAS40 pThr246 levels by 80-90% (Supplementary Figure 5). In the next step, I treated T47D *orange* clones kept either with estrogen supplementation (+E2) or grown under short- or long-term estrogen deprivation (-E2) with 100 nM MK-2206, and measured cell viability and proliferation (Figure 28).

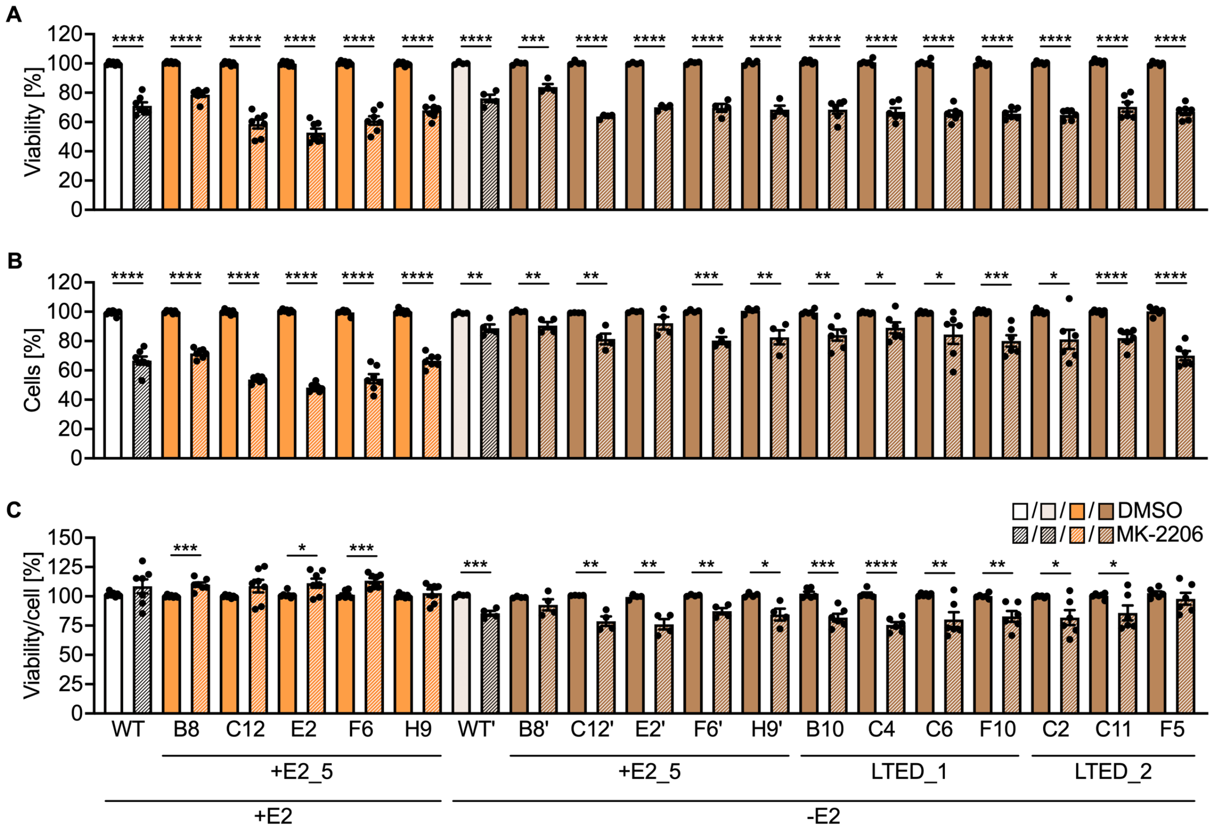


**Figure 28: Phenotypic effects of AKT inhibition.** T47D *orange* clones were kept in the indicated media (+E2, -E2) and were treated with DMSO (no filling) or 100 nM MK-2206 (shaded filling) for 72 h and cell viability (**A**) as well as cell numbers (**B**) were determined. The viability per counted cell is shown in **C**. **A-C** WT cells without barcode were used as additional controls. ‘ indicates 14 day pre-treatment by estrogen deprivation (-E2) and then

continued treatment for the duration (3 days) of the assay. n ≥ 3 with ≥ 5 technical replicates for all assays. Shown are mean ± SEM. * represents $p<0.05$, ** represents $p<0.01$, *** represents $p<0.001$ and **** represents $p<0.0001$ as determined by unpaired two-tailed t-tests.

AKT inhibition resulted in a significant reduction in observed cell viability by one-third on average (range: 17%-48%). The reduction in observed cellular viability was consistent for *orange* clones under normal (+E2) growth condition as well as during short-term (-E2) and long-term estrogen deprivation (LTED). T47D WT cells without barcode integration were used as additional controls and showed a similar response to AKT inhibition under normal (+E2) as well as short-term estrogen deprived (-E2) growth conditions. In my T47D model, AKT inhibition seemed to be a suitable pan-treatment approach for fast-proliferating (WT, +E2_5 clones) as well as dormant (short-term estrogen deprived WT and +E2_5 clones, LTED clones). It was further effective in decreasing observed cell viability on a clonal background (*orange* +E2_5 and LTED clones) as well as in a heterogeneous cell population (T47D WT). Despite the pronounced activation of AKT signaling in short-term estrogen deprived T47D *orange* +E2_5 clones however, AKT inhibition was not more effective in decreasing cellular viability under short-term estrogen deprivation (-E2) compared to normal (+E2) and LTED conditions (Figure 28 A). While the observed net-decrease in cellular viability was similar between treatment naïve (+E2), short-term (WT and +E2_5 -E2) and long-term estrogen deprived (LTED) cells, reduced proliferation appeared to be the driver for fast-proliferating T47D WT and +E2_5 clones. For short-term estrogen deprived *orange* +E2_5 clones and LTED clones, the observed reduction in cellular viability largely came as a combination of reduced proliferation and reduced viability per cell (Figure 28 B and C).

### 4.13 MCF7 LTED clones upregulate JAK-STAT signaling through STAT2

For all further validation experiments, I focused on the MCF7 TAMR and LTED models as they showed a more proliferative phenotype as well as more deregulated TFs and kinases than their T47D counterparts. Here, I used a selected control cell line (+E2_5) and a clone each representing the isolated clonal populations (*blue* TAMR_1: B6, *yellow* LTED_2: C4, *pink* LTED_2: E11 and *red* LTED_5: C9) for all experiments. I did not include non-barcoded MCF7 WT cells for further analyses as I had done for T47D since the MCF7 +E2_5 control cell line represented a complex barcoded cell line and not individual clones. Conversely, *orange* and *green* T47D +E2_5 clones originated from two single cells of origin, and I had therefore included the complex non-barcoded T47D WT cells for the prior analyses.

In the MCF7 models, JAK-STAT signaling was of keen interest as it was significantly upregulated in MCF7 LTED clones as inferred from the RNA-Seq data by PROGENy [261, 262] but not altered compared to +E2 control cell lines in the *blue* TAMR_1 clones (Figure 21). TF activity analysis by DoRothEA [263] further revealed that pathway activation was likely reflected by STAT2 activity and to a smaller extent by STAT1 activity (Figure 29).

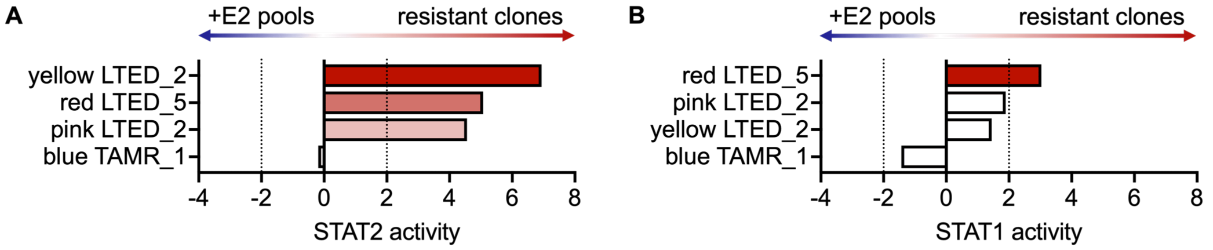


**Figure 29: Activation of STAT transcription factors in MCF7 LTED clones.** Absolute activity scores ≥ 2 were considered significant as indicated by dashed lines. Significant activation of STAT2 (**A**) and STAT1 (**B**) activity is highlighted in red. Analyses were performed by Dr. Efstathios-Iason Vlachavas.

In line with the specific activation of JAK-STAT signaling as well as STAT2 and STAT1, I investigated if knockdown of either STAT would specifically affect MCF7 LTED clones. Knockdown of *STAT2* and *STAT1* resulted in > 80% reduced mRNA and protein levels (Figure 30).

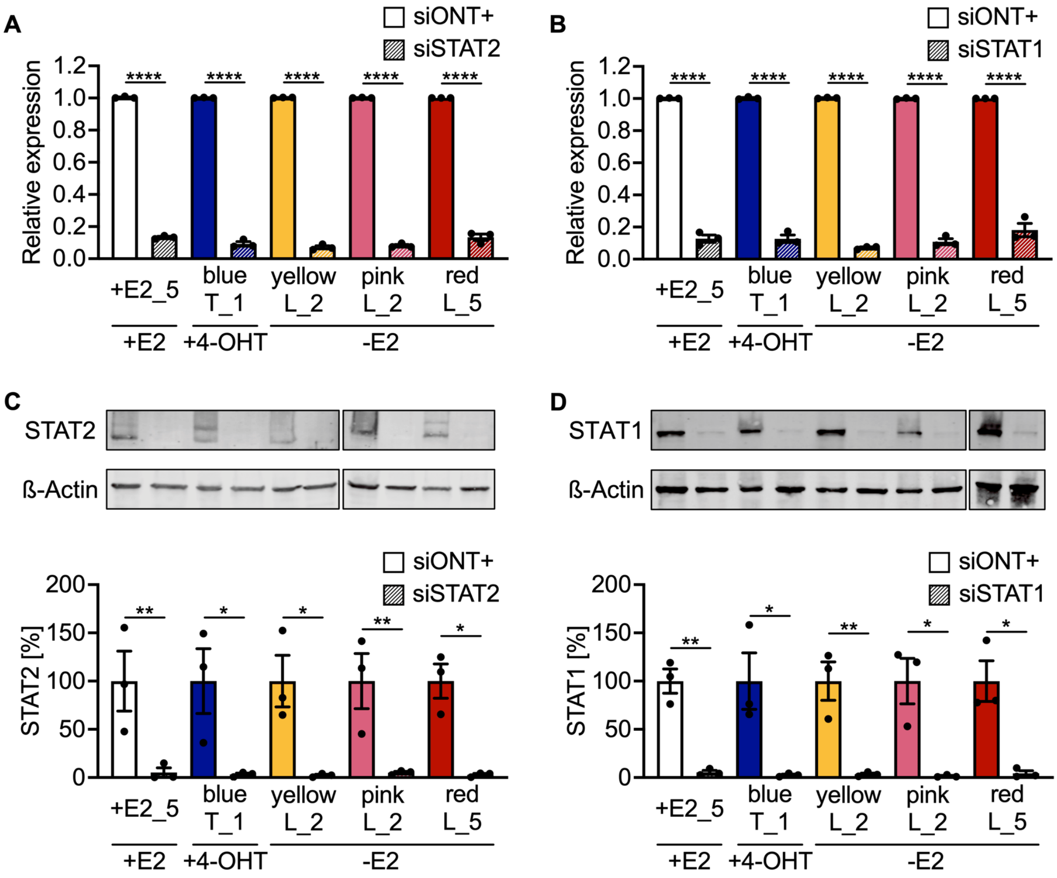


**Figure 30: *STAT* knockdown efficiencies.** MCF7 +E2_5 (☐), TAMR_1 B6 (*blue* T_1, ■), LTED_2 C4 (*yellow* L_2, ■), LTED_2 E11 (*pink* L_2, ■) and LTED_5 C9 (*red* L_5, ■) were kept in the indicated media (+E2, +4-OHT, -E2) and were transfected with control siRNA (siONT+, no filling) or siRNA targeting *STAT2* or *STAT1* (shaded filling). Expression on gene (**A** and **B**) and protein levels (**C** and **D**) were determined 6 days post transfection. **A** and **B** *STAT2* (**A**) and *STAT1* (**B**) abundances were normalized to *ACTB* and *GAPDH* abundances and to their respective siONT+ controls. n = 3 with 3 technical replicates. **** represents p<0.0001 as determined by unpaired two-tailed t-tests. **C** and **D** Representative Western Blots and quantification of *STAT2* (**C**) and *STAT1* (**D**) knockdowns on protein level. STAT2 (**C**) and STAT1 (**D**) signals were normalized to the respective ß-Actin loading controls and then normalized to cells transfected with control siRNA (siONT+). Shown are mean ± SEM. n = 3. * represents p<0.05 and ** represents p<0.01 as determined by unpaired two-tailed t-tests.

On phenotypic level, I investigated the effect of either STAT knockdown on cellular proliferation, viability, cell cycle progression and cell death utilizing nuclear cell count, ATP-levels, EdU and DAPI incorporation, respectively. *STAT2* knockdown had highly similar effects on endocrine sensitive control cells (+E2_5) as well as TAMR and LTED clones irrespective of their STAT2 activation score. *STAT2* knockdown significantly reduced cellular proliferation, viability, and the percentage of cycling cells, while simultaneously inducing a significant increase in dying/dead cells (Figure 31).

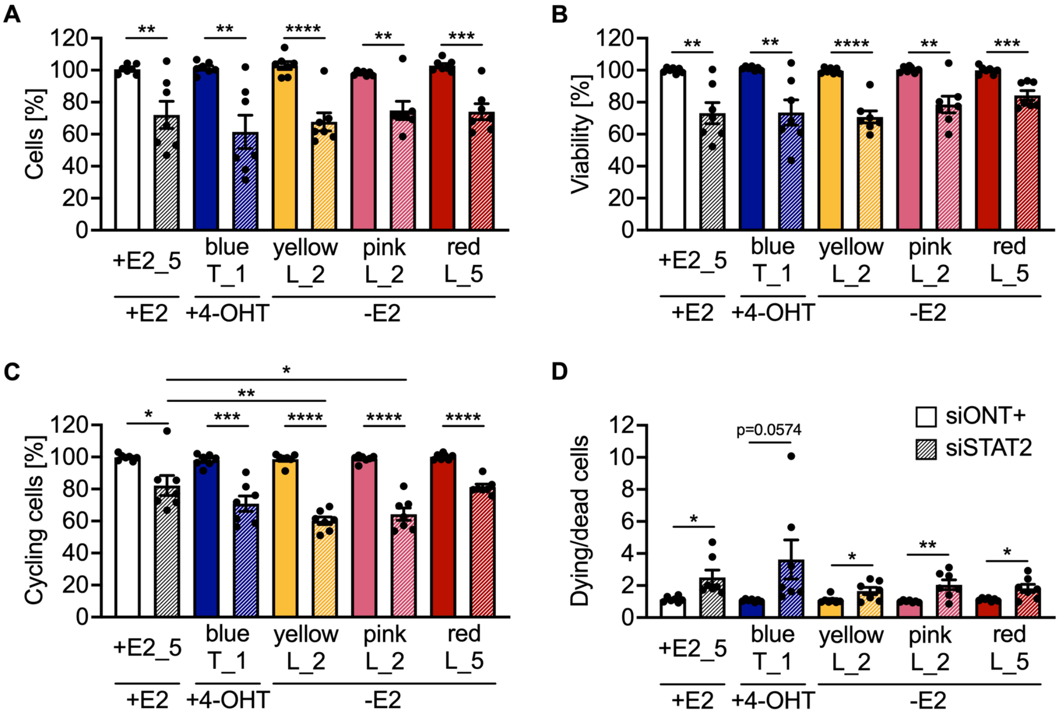


**Figure 31: Phenotypic effects of *STAT2* knockdown.** MCF7 +E2_5 (□), TAMR_1 B6 (*blue* T_1, ■), LTED_2 C4 (*yellow* L_2, ■), LTED_2 E11 (*pink* L_2, ■) and LTED_5 C9 (*red* L_5, ■) were kept in the indicated media (+E2, +4-OHT, -E2) and were transfected with control siRNA (siONT+, no filling) or siRNA targeting *STAT2* (shaded filling). Nuclear cell count (**A**) and ATP levels (**B**) were determined 6 days after transfection while EdU (**C**) and DAPI incorporation (**D**) were assessed 4 days after transfection. Results were normalized to non-targeting control (siONT+) siRNA. n = 7 with ≥ 4 technical replicates. Shown are mean ± SEM. * represents $p<0.05$, ** represents $p<0.01$, *** represents $p<0.001$ and **** represents $p<0.0001$ as determined by unpaired two-tailed t-tests or one-way ANOVA with Dunnett multiple comparisons test.

*STAT1* knockdown had similar effects as *STAT2* knockdown, meaning that it affected endocrine sensitive +E2_5 control cells as well as TAMR and LTED clones. In contrast to *STAT2* knockdown however, the effect of *STAT1* knockdown on different phenotypes was less pronounced for *yellow* and *pink* MCF7 LTED_2 clones (Figure 31).

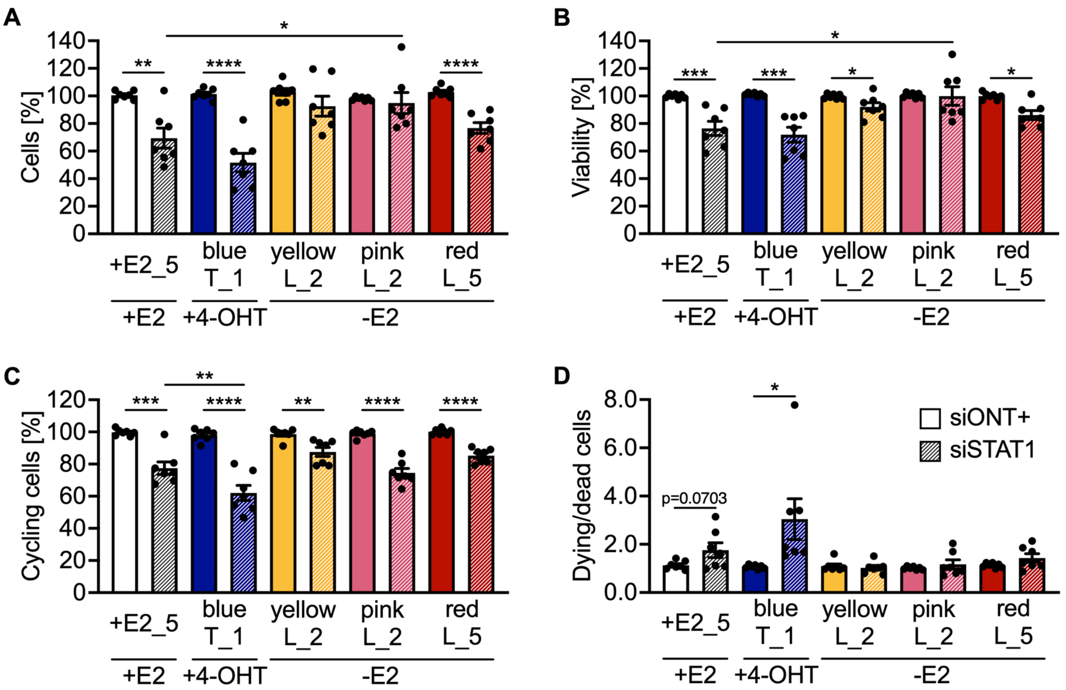


**Figure 32: Phenotypic effects of *STAT1* knockdown.** MCF7 +E2_5 (☐), TAMR_1 B6 (*blue* T_1, ■), LTED_2 C4 (*yellow* L_2, ■), LTED_2 E11 (*pink* L_2, ■) and LTED_5 C9 (*red* L_5, ■) were kept in the indicated media (+E2, +4-OHT, -E2) and were transfected with control siRNA (siONT+, no filling) or siRNA targeting *STAT1* (shaded filling). Nuclear cell count (**A**) and ATP levels (**B**) were determined 6 days after transfection while EdU (**C**) and DAPI incorporation (**D**) were assessed 4 days after transfection. Results were normalized to non-targeting control (siONT+) siRNA. n = 7 with ≥ 4 technical replicates. Shown are mean ± SEM. * represents p<0.05, ** represents p<0.01, *** represents p<0.001 and **** represents p<0.0001 as determined by unpaired two-tailed t-tests or one-way ANOVA with Dunnett multiple comparisons test.

Knockdown of either STAT highlighted their important role in cellular signaling. The knockdowns reduced STAT2 and STAT1 protein levels by > 90% (Figure 30 C and D) thereby potentially masking the impact of differential STAT2/STAT1 activities and their implications in resistance to aromatase inhibitor treatment. To titrate potential inhibitory effects, I utilized an inhibitor specific to Janus kinase (JAK) 1 and JAK2, which are upstream activators of different STAT transcription factors. However, all utilized cell lines and clones showed similar sensitivity to the JAK1/2 inhibitor Ruxolitinib [246], irrespective of their (in)sensitivity to endocrine therapy and JAK-STAT activation (Figure 33).

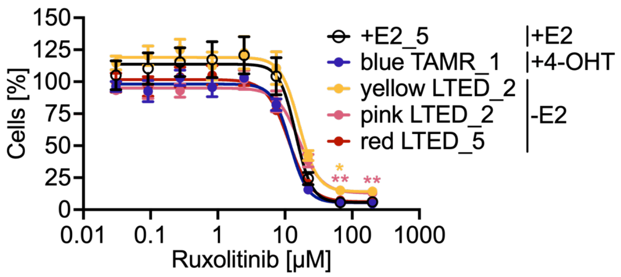


**Figure 33: Ruxolitinib response curve.** MCF7 +E2_5 (○), TAMR_1 B6 (*blue* TAMR_1, ●), LTED_2 C4 (*yellow* LTED_2, ●), LTED_2 E11 (*pink* LTED_2, ●) and LTED_5 C9 (*red* LTED_5, ●) were kept in the indicated media (+E2, +4-OHT, -E2) and were treated with increasing concentrations of Ruxolitinib. Cell numbers were determined by microscopy-based nuclei counting at day 6 and normalized to the DMSO controls. n = 3 with 3 technical replicates. * represents p<0.05 and ** represents p<0.01 as determined by two-way ANOVA with Dunnett multiple comparisons test.

## 4.14 MCF7 TAMR and LTED clones upregulate NFκB signaling in part through p65

Besides JAK-STAT, NFκB signaling was of strong interest to me, since all MCF7 TAMR and LTED clones showed significant activation of this pathway (Figure 21). The downstream TF p65 was significantly activated in *red* LTED_5 and *pink* LTED_2 clones, while *yellow* LTED_2 and *blue* TAMR_1 clones showed only a tendency of further activation over +E2 control cell lines with NES > 1.10 (Figure 34).

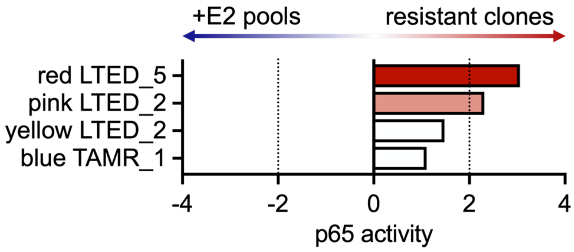


**Figure 34: p65 activation in MCF7 LTED clones.** Absolute activity scores ≥ 2 were considered significant as indicated by dashed lines. Significant activation in *red* LTED_5 and *pink* LTED_2 clones is highlighted in red. Analysis was performed by Dr. Efstathios-Iason Vlachavas.

Next, I targeted *RELA* with specific siRNAs and knockdown efficiency was > 90% at RNA and protein levels (Figure 35).

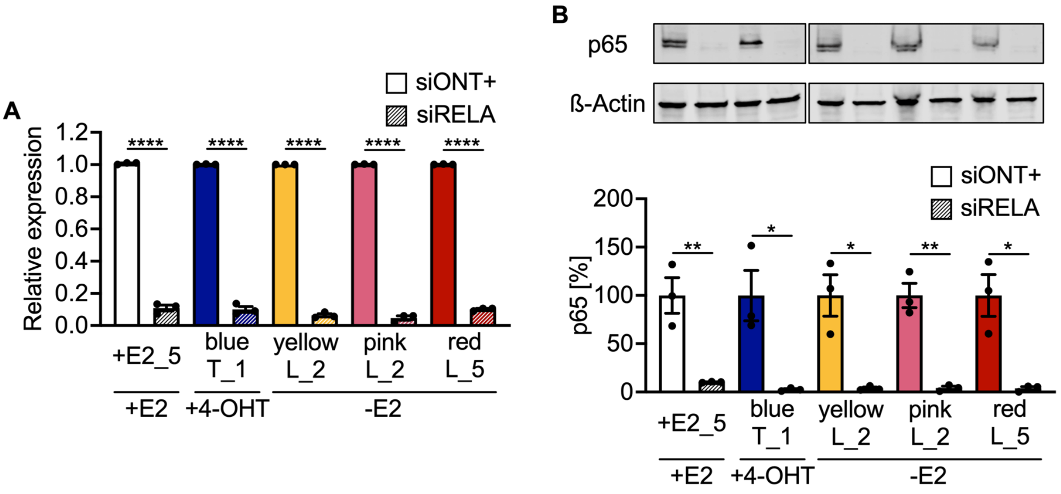


**Figure 35: *RELA* knockdown efficiency.** MCF7 +E2_5 (□), TAMR_1 B6 (*blue* T_1, ■), LTED_2 C4 (*yellow* L_2, ■), LTED_2 E11 (*pink* L_2, ■) and LTED_5 C9 (*red* L_5, ■) were kept in the indicated media (+E2, +4-OHT, -E2) and were transfected with control siRNA (siONT+, no filling) or siRNA targeting *RELA* (shaded filling). Expression on gene (**A**) and protein levels (**B**) were determined 6 days post transfection. **A** *RELA* abundance was normalized to *ACTB* and *GAPDH* abundances and to their respective siONT+ controls. n = 3 with 3 technical replicates. **** represents p<0.0001 as determined by unpaired two-tailed t-tests. **B** Representative Western Blot and quantification of *RELA* knockdown on protein level. p65 signals were normalized to the respective ß-Actin loading controls and then normalized to cells transfected with control siRNA (siONT+). Shown are mean ± SEM. n = 3. * represents p<0.05 and ** represents p<0.01 as determined by unpaired two-tailed t-tests.

Phenotypically, *RELA* knockdown induced similar phenotypes in MCF7 +E2_5 as well as TAMR and LTED clones. Proliferation, cell viability and the percentage of cycling cells were all significantly decreased while the fraction of dying/dead cells was significantly increased upon *RELA* knockdown irrespective of endocrine therapy sensitivity (Figure 36). Like STAT2 and STAT1, p65 appeared to play an important role in cell proliferation and survival in MCF7 TAMR and LTED clones but also in endocrine sensitive +E2_5 control cells.

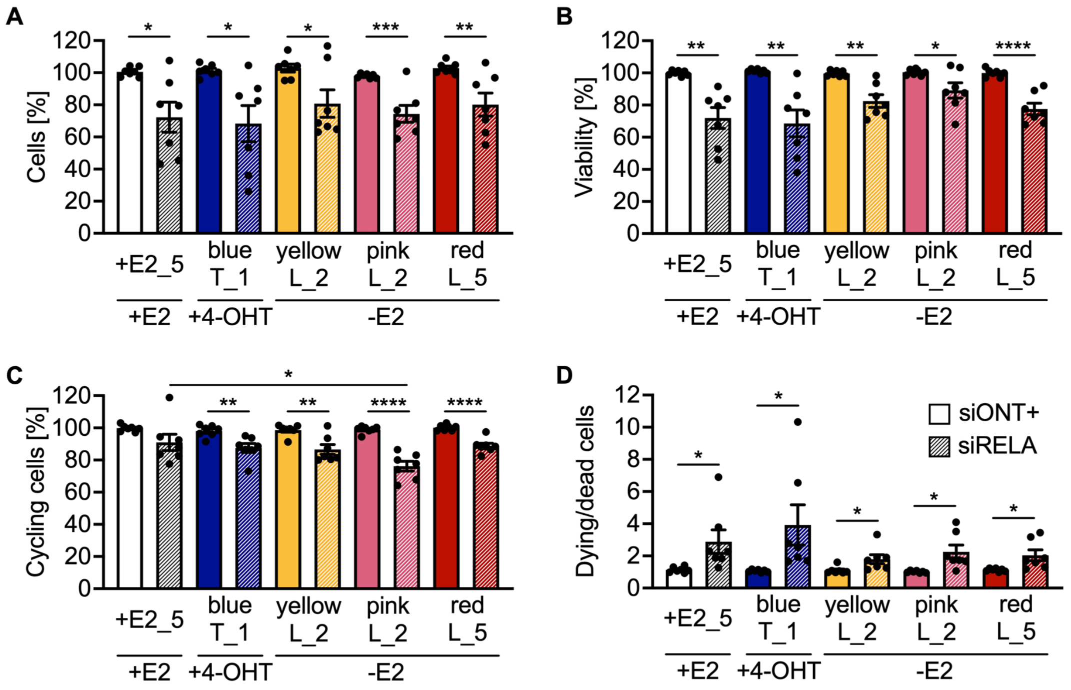


**Figure 36: Phenotypic effects of *RELA* knockdown.** MCF7 +E2_5 (□), TAMR_1 B6 (*blue* T_1, ■), LTED_2 C4 (*yellow* L_2, ■), LTED_2 E11 (*pink* L_2, ■) and LTED_5 C9 (*red* L_5, ■) were kept in the indicated media (+E2, +4-OHT, -E2) and were transfected with control siRNA (siONT+, no filling) or siRNA targeting *RELA* (shaded filling). Nuclear cell count (**A**) and ATP levels (**B**) were determined 6 days after transfection while EdU (**C**) and DAPI incorporation (**D**) were assessed 4 days after transfection. Results were normalized to non-targeting control (siONT+) siRNA. n = 7 with ≥ 4 technical replicates. Shown are mean ± SEM. * represents $p<0.05$, ** represents $p<0.01$, *** represents $p<0.001$ and **** represents $p<0.0001$ as determined by unpaired two-tailed t-tests or one-way ANOVA with Dunnett multiple comparisons test.

NFκB signaling is regulated and can be interfered with at different levels. In unstimulated conditions, NFκB dimers are sequestered in the cytosol by nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor (IκB) proteins. Upon stimulation, IκB is phosphorylated by IκB kinases (IKKs) marking it for ubiquitination and subsequent degradation by the proteasome, thereby freeing the NFκB dimer [288-290]. We utilized our phospho-proteomic data again to determine the activity of different IKKs and to investigate whether IKK activity could explain overall NFκB pathway activity. Neither of the four IKKs (IKKα, IKKß, IKKε and TBK1) showed a significant deregulation nor did their activity correlate with NFκB or p65 activity (Supplementary Figure 6). Taken together, targeting NFκB signaling by utilizing an IKK inhibitor seemed to not be a suitable option.

### 4.15 UPR activation reveals clonal vulnerability to Bortezomib

NFκB signaling appeared to be involved in endocrine therapy resistance in the MCF7 TAMR and LTED clones. However, targeting p65 signaling with specific siRNAs reduced p65 protein levels by > 90% and thereby did not allow to titrate a potential effect on NFκB signaling. Clinically, NFκB-driven Multiple Myeloma are commonly treated in the first line setting with a protease inhibitor such as Bortezomib [291], preventing the degradation of the negative regulator IκB and thereby blocking NFκB signaling. Previously, Bortezomib treatment has additionally been shown to induce a strong unfolded protein response (UPR) [292, 293]. Generally, the UPR serves to restore normal protein levels (protein homeostasis). Increases in protein secretion and/or disruptive protein folding in the endoplasmic reticulum can lead to the accumulation of un- or misfolded protein and associated stress of the endoplasmic reticulum. The UPR is then activated to adjust transcription and translation as well as protein modifications and secretion to finally restore protein homeostasis [294]. While mild UPR activation has been hypothesized to be protective and potentially drive proliferation in endocrine therapy treated cells [293], proteasome inhibitors such as Bortezomib can induce a terminal/detrimental UPR ultimately killing the cancer cells [292, 293]. Accordingly, we utilized the RNA-Seq data again for further analysis using gene set enrichment analysis (GSEA) [295-297] looking specifically for enrichment of the UPR in MCF7 TAMR and LTED clones. Here, exclusively MCF7 *blue* TAMR_1 clones revealed a significant enrichment of the UPR pointing to its activation (Figure 37 A). Accordingly, I speculated that Bortezomib might be a suitable therapeutic option, either as NFκB inhibitor for all TAMR and LTED clones or specifically for MCF7 *blue* TAMR_1 cells, which showed an activation of the UPR that would be potentiated detrimentally by Bortezomib. A titrating experiment revealed high sensitivity to treatment with 1.85 nM Bortezomib for MCF7 *blue* TAMR_1 (red arrow in Figure 37 B), in accordance with the private UPR activation. I was next interested in how treatment would affect cell viability, cell cycle progression and cell death.

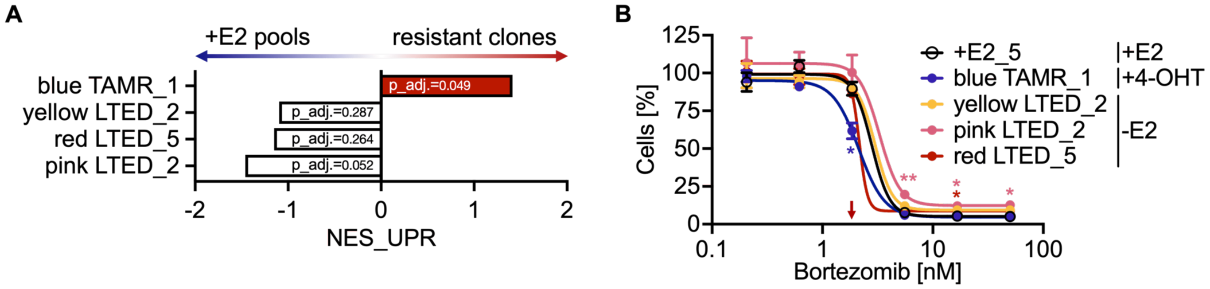


**Figure 37: UPR activation and sensitivity to Bortezomib. A** GSEA of unfolded protein response (UPR). Significant enrichment in MCF7 *blue* TAMR_1 is highlighted in red. Adjusted p-values are shown for each comparison. Analysis was performed by Dr. Efstathios-Iason Vlachavas. **B** MCF7 +E2_5 (○), TAMR_1 B6 (*blue* TAMR_1, ●), LTED_2 C4 (*yellow* LTED_2, ●), LTED_2 E11 (*pink* LTED_2, ●) and LTED_5 C9 (*red* LTED_5, ●) were kept in the indicated media (+E2, +4-OHT, -E2) and were treated with increasing concentrations of Bortezomib. Cell numbers were determined by microscopy-based nuclei counting at day 7 and normalized to the DMSO controls. n = 3 with 3 technical replicates. * represents $p<0.05$ and ** represents $p<0.01$ as determined by two-way ANOVA with Dunnett multiple comparisons test.

Bortezomib strongly reduced proliferation, cell viability and the percentage of cycling cells specifically for the utilized MCF7 *blue* TAMR_1 B6 clone, which was highly significant compared to the Bortezomib-treated +E2_5 control cells. Treatment also resulted in a significant increase in dying/dead TAMR_1 B6 cells. Additionally, Bortezomib had a similar but only minimally effect on the MCF7 *red* LTED_5 clone, significantly reducing cellular viability and the percentage of cycling cells compared to DMSO-treated *red* LTED_5 cells. However, none of these phenotypic effects were statistically significant compared to Bortezomib-treated +E2_5 cells indicating no increased sensitivity of *red* LTED_5 cells over control (+E2_5) cells

(Figure 38). Taken together, UPR activation and sensitivity to Bortezomib emerged as a private event specific for MCF7 *blue* TAMR_1.

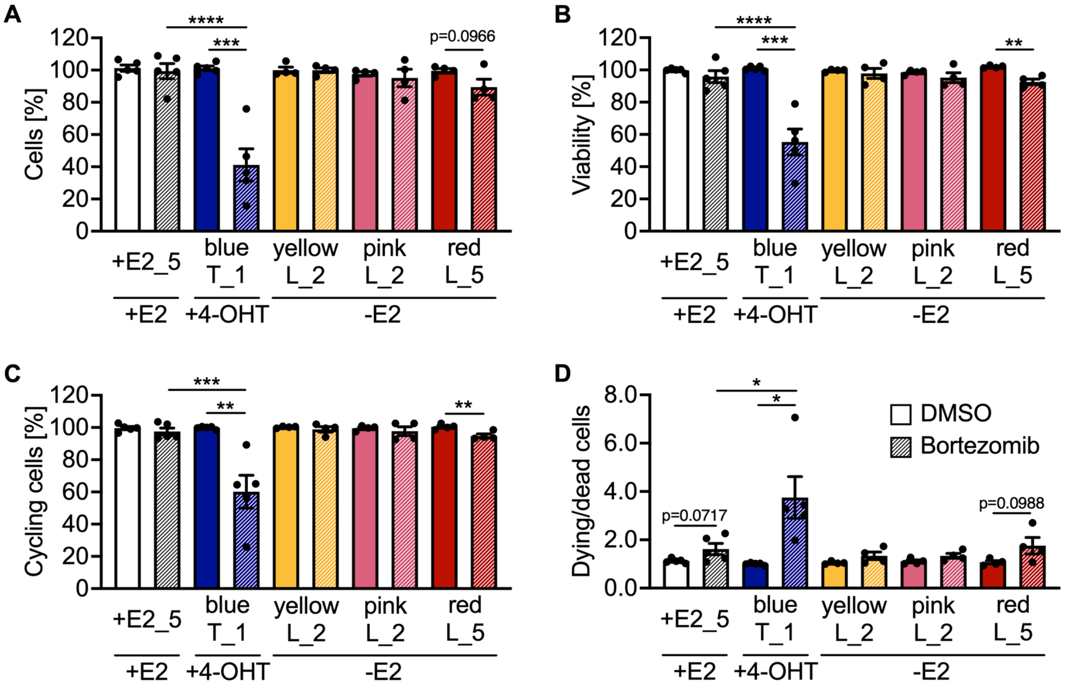


**Figure 38: Phenotypic effects of Bortezomib treatment.** MCF7 +E2_5 (□), TAMR_1 B6 (*blue* T_1, ■), LTED_2 C4 (*yellow* L_2, ■), LTED_2 E11 (*pink* L_2, ■) and LTED_5 C9 (*red* L_5, ■) were kept in the indicated media (+E2, +4-OHT, -E2) and were treated with DMSO (no filling) or 1.85 nM Bortezomib (shaded filling). Nuclear cell count (**A**) and ATP levels (**B**) were determined 6 days after treatment start while EdU (**C**) and DAPI incorporation (**D**) were assessed 4 days after treatment start. Results were normalized to DMSO controls. n ≥ 4 with ≥ 4 technical replicates. Shown are mean ± SEM. * represents $p<0.05$, ** represents $p<0.01$, *** represents $p<0.001$ and **** represents $p<0.0001$ as determined by unpaired two-tailed t-tests or one-way ANOVA with Dunnett multiple comparisons test.

## 4.16 PKC activation as shared clonal driver of endocrine therapy resistance

Besides NFκB activation, PKCß and PKCδ of the protein kinase C family were also significantly activated in all MCF7 TAMR and LTED clones compared to the +E2 control cell lines as determined from the phospho-proteomics data (Figure 24 B). PKCα, PKCγ and PKCε were further significantly activated in MCF7 *blue* TAMR_1, *yellow* LTED_2 and *red* LTED_5 clones (Figure 39 A). To investigate whether inhibition of the shared activation of different PKC isoforms would be a suitable therapeutic option, I performed an titration experiment using the pan-PKC inhibitor Sotrastaurin [247] which revealed strong inhibition of proliferation for *yellow* LTED_2 cells (Figure 39 B).

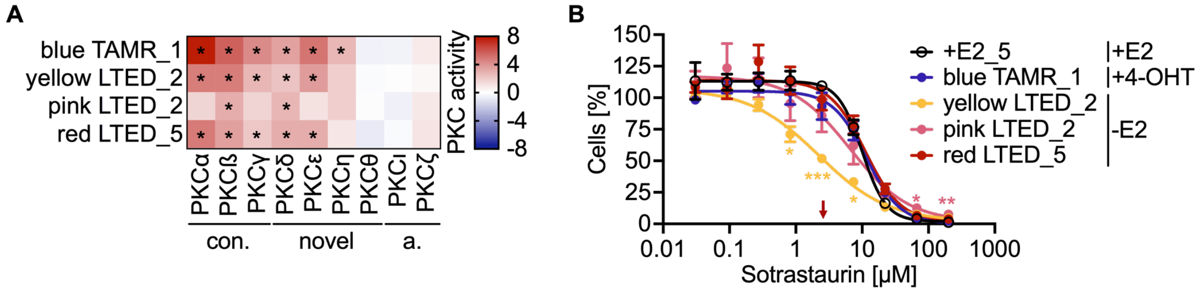


**Figure 39: PKC activation and sensitivity to Sotrastaurin. A** Activities of different PKC isoforms in MCF7 TAMR and LTED clones. con.: conventional. a.: atypical. Activation and repression of PKC isoforms in TAMR and LTED clones is indicated by red and blue, respectively. * indicates absolute kinase activity ≥ 2. Analysis was performed by Dr. Efstathios-Iason Vlachavas. **B** MCF7 +E2_5 (○), TAMR_1 B6 (*blue* TAMR_1, ●), LTED_2 C4 (*yellow* LTED_2, ●), LTED_2 E11 (*pink* LTED_2, ●) and LTED_5 C9 (*red* LTED_5, ●) were kept in the indicated media (+E2, +4-OHT, -E2) and were treated with increasing concentrations of Sotrastaurin. Cell numbers were determined by microscopy-based nuclei counting at day 7 and normalized to DMSO controls. n = 4 with 3 technical replicates. * represents p<0.05, ** represents p<0.01 and *** represents p<0.001 as determined by two-way ANOVA with Dunnett multiple comparisons test.

Next, I investigated the inhibition of TPA-induced PKC activation [248, 249] by 2.50 µM Sotrastaurin (red arrow in Figure 39 B). Here, I pre-treated the cells either with DMSO or Sotrastaurin for 24 h before stimulation with TPA for 30 min. As readout and for quantification, I utilized an antibody that recognizes phosphorylated PKC substrates. Treatment with Sotrastaurin significantly and reproducibly reduced PKC activation by 65-80% in all cell lines tested (Figure 40).

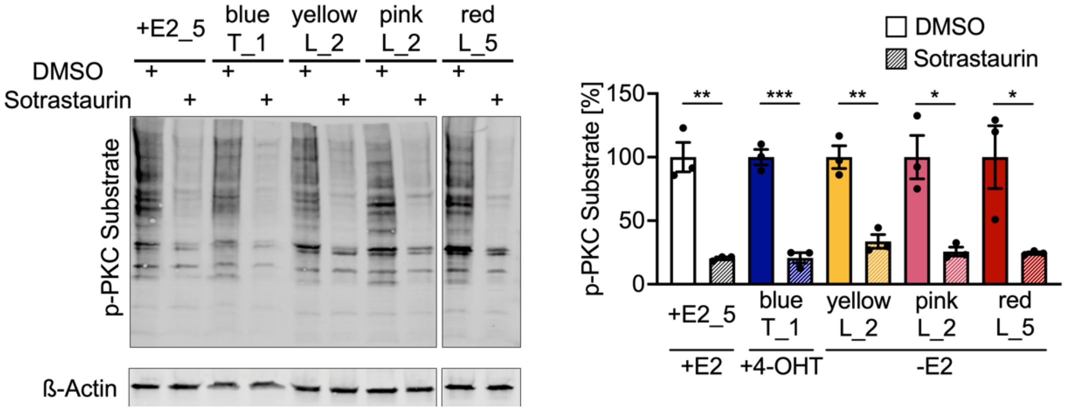


**Figure 40: Quantification of PKC inhibition by Sotrastaurin.** MCF7 +E2_5 (□), TAMR_1 B6 (*blue* T_1, ■), LTED_2 C4 (*yellow* L_2, ■), LTED_2 E11 (*pink* L_2, ■) and LTED_5 C9 (*red* L_5, ■) were treated for 24 h with DMSO (no filling) or 2.50 µM Sotrastaurin (shaded filling) before being stimulated with 200 nM TPA for 30 min. Cells were kept in the indicated media (+E2, +4-OHT, -E2). Representative Western Blot (left panel) and quantification of three biological replicates (right panel) of DMSO or Sotrastaurin treatment before TPA stimulation are depicted. p-PKC Substrate signals were normalized to the respective ß-Actin loading controls and then normalized to DMSO-treated cells. Shown are mean ± SEM. * represents p<0.05, ** represents p<0.01 and *** represents p<0.001 as determined by unpaired two-tailed t-tests.

Having confirmed significant PKC inhibition in all cell lines tested, I next assessed the effect of treatment with 2.50 µM Sotrastaurin on proliferation, cell viability, cycling and dying/dead cells (Figure 41). Fully in line with the sensitivity assay described in Figure 39 B, *yellow* LTED_2 cells showed the highest sensitivity to inhibition of PKC at the levels of cell numbers, viability, and cell death. In contrast, the *red* LTED_5 clone, which was suggested to be insensitive to this Sotrastaurin concentration showed a prominent reduction in proliferation and viability as well, and a small yet insignificant increase in dying/dead cells (Figure 41). Effects of Sotrastaurin on the other clones tested were mostly mild, in accordance with the drug sensitivity test. Of note, *blue* TAMR_1 clones showed reduced viability, which showed a trend towards significance compared to Sotrastaurin-treated +E2_5 control cell lines. This observation could be explained by neither a reduction in cycling cells nor an increase in cell death. Taken together, the reductions in cellular viability by Sotrastaurin were largely in line with the activation level of different PKC isoforms in the MCF7 TAMR and LTED clones.

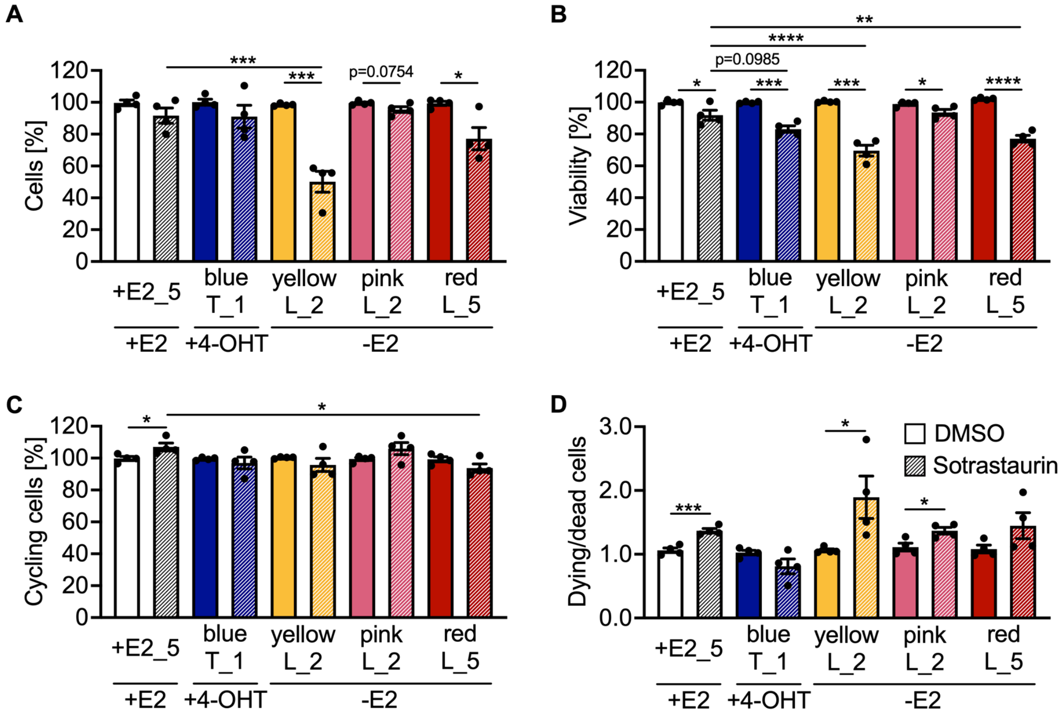


**Figure 41: Phenotypic effects of Sotrastaurin treatment.** MCF7 +E2_5 (□), TAMR_1 B6 (*blue* T_1, ■), LTED_2 C4 (*yellow* L_2, ■), LTED_2 E11 (*pink* L_2, ■) and LTED_5 C9 (*red* L_5, ■) were kept in the indicated media (+E2, +4-OHT, -E2) and were treated with DMSO (no filling) or 2.50 µM Sotrastaurin (shaded filling). Nuclear cell count (**A**) and ATP levels (**B**) were determined 6 days after treatment start while EdU (**C**) and DAPI incorporation (**D**) were assessed 4 days after treatment start. Results were normalized to DMSO controls. n = 4 with ≥ 4 technical replicates. Shown are mean ± SEM. * represents $p<0.05$, ** represents $p<0.01$, *** represents $p<0.001$ and **** represents $p<0.0001$ as determined by unpaired two-tailed t-tests or one-way ANOVA with Dunnett multiple comparisons test.

## 4.17 Clinical significance of *in vitro* findings

After having targeted and validated clonal resistance drivers in MCF7 TAMR and LTED clones, I next wanted to assess whether these *in vitro* findings have clinical significance. Therefore, we utilized and analyzed the CPTAC-BRCA cohort recently published by Krug *et al.* [234]. This data set contains genetic aberrations, gene and protein expression as well as phospho-proteomics data of 122 treatment-naïve patients. When analyzing this cohort, I was particularly interested in similarities between my MCF7 TAMR and LTED clones and patients with ER- disease, as the applied therapies in my *in vitro* models strongly suppressed estrogen signaling (Figure 20 and Figure 21), thereby rendering the TAMR and LTED clones 'quasi'

ER-. This observation was in accordance with the lack of *ESR1* hotspot mutations in any of the endocrine therapy resistant clones (Figure 17).

Of note, all downstream analyses were performed by Dr. Efstathios-Iason Vlachavas. First, we analyzed the different activation of the 14 major pathways assessed by PROGENy for patients with ER- as compared to ER+ disease. In patients with ER- disease, estrogen signaling was the strongest downregulated pathway (Figure 42), in line with my *in vitro* findings. Hypoxia, TNFα, NFκB and JAK-STAT signaling were further associated with ER- disease (Figure 42). These four pathways were also significantly activated in at least three out of four MCF7 TAMR and LTED clones (Figure 21), and together with the significant repression of estrogen signaling in patients with ER- disease highlighted the resemblance of my *in vitro* MCF7 endocrine therapy resistance models and patients with ER- disease in the CPTAC-BRCA cohort.

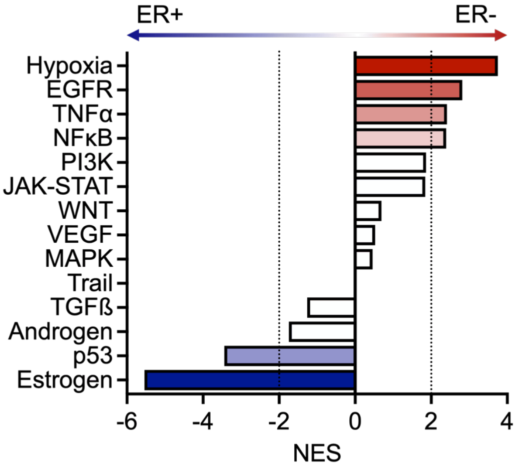


**Figure 42: Pathway activities in breast cancer patients.** The CPTAC-BRCA cohort was subgrouped into patients with ER- (n = 39) or ER+ (n = 81) disease. Absolute NES scores ≥ 2 were considered significant as indicated by dashed lines. Significantly activated pathways in patients with ER- disease are highlighted in red whereas significantly downregulated pathways are highlighted in blue. Analysis was performed by Dr. Efstathios-Iason Vlachavas.

In the next step, I was particularly interested if alterations which I had observed in my MCF7 *in vitro* models were also evident in patients with ER- disease. Specifically, I was interested in STAT2, STAT1 and p65, UPR and PKC activity patterns. To this end, we performed TF activity analysis, GSEA and kinase activity analysis of the stratified CPTAC-BRCA patients (Figure 43). p65 was significantly associated with ER- disease, while STAT2 and STAT1 activity did not correlate with either disease subtype. Further, UPR activation showed a trend towards association with ER- disease, however this association did not reach statistical significance. No patterns of PKC activation and ER-(in)dependent disease were evident and PKCδ was the only PKC isoform that showed a significant association, particularly with ER+ disease.

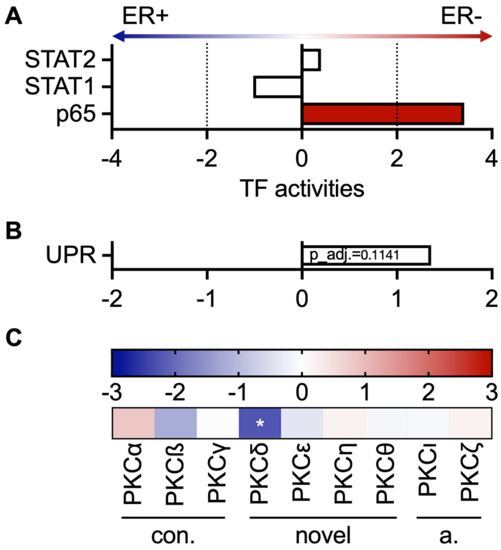


**Figure 43: TF, UPR and PKC activities in breast cancer patients.** The CPTAC-BRCA cohort was subgrouped into patients with ER- (n = 39) or ER+ (n = 81) disease. **A** and **C** Absolute NES scores ≥ 2 were considered significant. Significant activation of p65 (**A**) in patients with ER- disease is highlighted in red whereas significant downregulation of PKCδ (**C**) is highlighted in blue and by the asterisk. **B** UPR activation and adjusted p. value as determined by GSEA. Analyses were performed by Dr. Efstathios-Iason Vlachavas.

Accordingly, I speculated that specifically PKC activation of individual patients could be masked in a heterogenous cohort. To this end, we analyzed the cohort again, but this time on a per patient basis. And indeed, strong heterogeneity in the activation of the pathways estimated by PROGENy was evident across patients with ER- (Figure 44) and ER+ disease (Supplementary Figure 7 A).

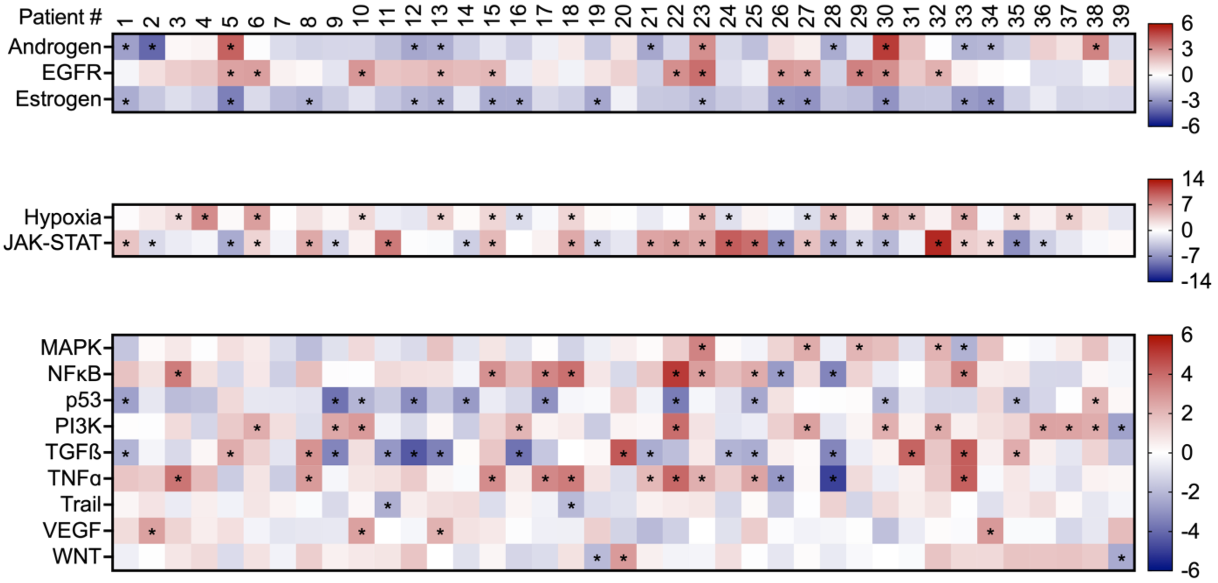


**Figure 44: Pathway activity profiles of individual patients with ER- disease from the CPTAC cohort.** Data from the RNA-Seq was z-scored prior to the analysis. Alphabetical ranking of the 14 major pathways assessed by PROGENy and their activities depicted as heatmaps. Activation and repression are visualized in red and blue, respectively, and significant deregulations are indicated by asterisks. Analysis was performed by Dr. Efstathios-Iason Vlachavas.

On one extreme, hypoxia signaling was significantly associated with ER- disease (Figure 43) and was significantly activated in 14 patients with ER- disease while it was significantly repressed in only 3 patients with ER- disease (Figure 44). Conversely, estrogen signaling was significantly associated with ER+ disease (Figure 43) and accordingly significantly downregulated in 14 patients with ER- disease, respectively. JAK-STAT signaling was significantly activated and repressed in 15 and 11 patients with ER- disease, respectively (Figure 44), potentially reflecting the heterogeneity observed in my MCF7 *in vitro* models (Figure 21). NFκB signaling was further significantly activated in eight patients and repressed in two patients with ER- disease, respectively, which was expected from its significant association with ER- disease (Figure 42).

Next, I was interested in deconvoluting the activity of potential endocrine therapy resistance drivers. Slightly more patients with ER- disease showed significant activation than repression of STAT2 and STAT1. For STAT2, 14 patients with ER- disease showed significant activation while 10 patients with ER- disease showed significant downregulation, respectively. For STAT1, 11 patients with ER- disease showed significant activation while 8 patients with ER- disease showed significant repression, respectively (Figure 45 A). These findings are in accordance with the observed activation of primarily STAT2 in my MCF7 LTED clones while no activation over the control (+E2) cell lines was observed for MCF7 *blue* TAMR_1 clones. Additionally, p65 activity was strongly and significantly associated with ER- disease on stratification level (Figure 43 A) and in line, more patients with ER- disease showed significant activation than repression of p65 activity on a per patient basis (16 vs. 2, Figure 45 A).

Similarly, UPR activity showed a trend towards association with ER- disease on stratification (ER- vs ER+) level (Figure 43 B) and was significantly activated in 16 patients with ER- disease while it was only significantly repressed in 2 patients with ER- disease, respectively (Figure 45 B).

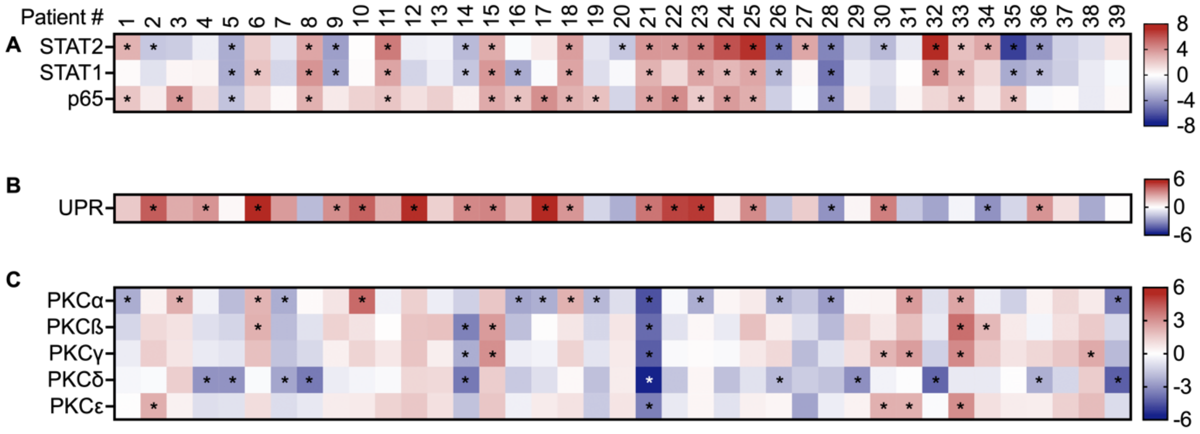


**Figure 45: TF, UPR and PKC activity profiles of individual patients with ER- disease from the CPTAC cohort. A** and **B** The data from the RNA-Seq was z-scored before further analysis of TF (**A**) and UPR activities (**B**). **C** PKC activities are in comparison to the reference sample as described by Krug *et al.* [234]. **A**-**C** Activation and repression are visualized in red and blue, respectively, and significant deregulations are indicated by asterisks. **B** UPR activation and repression were considered significant for absolute NES ≥ 2 and adj. p values < 0.05. **C** For better visibility, one asterisk (patient #21, PKCδ) is shown in white. Analyses were performed by Dr. Efstathios-Iason Vlachavas.

PKC activation in individual patients showed strong heterogeneity again. One common observation was the concurrent downregulation of PKCδ activity in 11 patients with ER- disease. Interestingly, PKCδ was hardly activated in the whole cohort as it was only significantly activated in 4 out of 120 analyzed patients (all ER+, #1, #28, #36, #48 and #70 in Supplementary Figure 7 B). In contrast, PKCß and PKCδ were significantly activated in all MCF7 TAMR and LTED clones (Figure 39 A), irrespective of their sensitivity to treatment with Sotrastaurin (Figure 41). Besides PKCß and PKCδ activation, two patients with ER- disease (#31 and #33) showed simultaneous activation of the other three PKC isoforms α, γ and ε (Figure 45 C). Accordingly, these two patients closely mimicked my MCF7 TAMR and LTED clones (Figure 39 A) and their tumors may be particularly sensitive to Sotrastaurin treatment. Of note, no kinase activity was predicted for PKCη, PKCθ, PKCι and PKCζ, but these isoforms did also not seem to be key players in the MCF7 TAMR and LTED clones (Figure 39 A).

Taken together, the clinical data provided strong support for the importance of my *in vitro* findings. Exemplarily, the per patient pathway activation pattern unraveled the strong inter-tumor heterogeneity of breast cancer patients. It is further tempting to speculate that individual tumors analyzed from the CPTAC-BRCA cohort additionally represented the assemble average measurements of multiple different clones within a single tumor. The need for per patient stratification became particularly evident for the activation of PKC isoforms. On patient level, I identified 16 patients with significant UPR activation closely resembling MCF7 *blue* TAMR_1 clones. Per patient analysis further revealed two patients with ER- disease whose PKC isoform activation pattern closely resembled the activation pattern of the MCF7 TAMR and LTED clones.

# 5 Discussion

Endocrine therapy resistance has remained an urgent clinical problem with relapse rates of around 20-25% within 10 years from diagnosis and treatment start [45, 298-300]. While ER+ patients whose tumors are diagnosed early (diameter < 2 cm and no lymph nodes involved) have 20-year relapse rates of as low as 13%, patients with advanced tumors (> 2 cm in diameter and 4-9 lymph nodes involved) experience relapses in as many as 41% of cases [84]. Many resistance mechanisms have been identified and for some known drivers, specific treatment options are routinely used in the clinics such as CDK4/6 inhibitors [85-92]. Tumor heterogeneity is increasingly appreciated to contribute to therapy resistance and recent years have shed further light on the clonal makeup of early and relapsing tumors across different entities [173]. The focus of my PhD thesis was to investigate the clonal heterogeneity and clonal drivers of endocrine therapy resistance.

## 5.1 Endocrine therapy resistant clones display distinct phenotypes

Dr. Simone Borgoni had established endocrine therapy sensitive and resistant barcoded cells lines in a previous PhD project [236]. I isolated single clones from both cell lines (T47D and MCF7) and both treatment arms (Tamoxifen treatment and estrogen deprivation) to uncover potential molecular and phenotypic heterogeneity in different clones and to identify individual resistance drivers. Previously, the cellular ability to change from one functional cell type to another type, also termed lineage plasticity, has been hypothesized to enable adaptation to targeted therapy and fuel intratumor heterogeneity [301, 302]. Therefore, I highlighted on gene expression (Figure 18) and phospho-proteomics level (Figure 23) that once clonal trajectories had been fixed as long-term adaptations to endocrine therapies, they were stable and no signs of further lineage change were detected.

On phenotypic level, I could further highlight the differences between the established cell line models. On the one hand, T47D *orange* LTED clones displayed a phenotype commonly referred to as ‘proliferative persister’ state [303]. Isolated MCF7 LTED clones on the other hand showed sustained proliferation under estrogen deprivation (Figure 19 D). Clinically, MCF7 LTED clones might be representative of earlier relapses while T47D *orange* LTED clones could be representative of dormant and late-relapsing diseases that could occur up to 20 years after surgery [282-284]. Further indication that T47D *orange* LTED clones may resemble late clinical relapses came from the fast resumption of a strong proliferative phenotype after short treatment withdrawal. In the clinics, patients receive endocrine therapy for ten years [45-47]. Around 10-40% of patients do not adhere to the prescribed endocrine therapy (defined by a medication possession rate of <80%) [304-307] potentially leading to a steady rate of relapse that could be attributed to treatment withdrawal and re-awakening of dormant cell populations.

## 5.2 Divergent transcriptional re-wiring in endocrine therapy resistant clones

After isolating and confirming the stable trajectories of +E2, TAMR and LTED clones, I was interested in the molecular changes underlying endocrine therapy resistance. Therefore, I mainly investigated differentially activated pathways, transcription factors and kinases.

Significant repression of estrogen signaling compared to treatment-naïve cell lines and clones was the only shared pathway deregulation between all T47D and MCF7 TAMR and LTED clones, highlighting the strong inter-tumor heterogeneity (Figure 20 and Figure 21). Additionally, no *ESR1* hotspot mutations covering codons 536 to 538 were detected (Figure 17) which might otherwise have rendered ER active independent of its ligand in these cells [95-97, 279-281]. Accordingly, estrogen signaling was effectively blocked in my *in vitro* models by Tamoxifen or estrogen deprivation mimicking clinically used aromatase inhibitors.

To complement the pathway activity analysis and to further identify putative drivers, TF activity was determined based on the RNA-Seq data. Here, the activity of three transcription factors (ELF3, GRHL2 and KLF5) was significantly upregulated in all T47D and MCF7 TAMR and LTED clones. ELF3 activation may have been the result of HER2 protein upregulation in my T47D and MCF7 TAMR and LTED clones (Supplementary Figure 8) as *ELF3* expression was previously found to be elevated in HER2+ tumors [308, 309] and associated with worse outcome of patients [310]. GRHL2 and KLF5 have both been implicated in therapy resistance of hormone dependent cancers. In the METABRIC cohort, the expression of GRHL2 target genes significantly correlated with worse overall patient survival only when patients were treated with endocrine therapy [311]. Higher GRHL2 protein expression has further been associated with Tamoxifen resistance [312]. KLF5 expression was elevated in a subset of prostate cancers resistance to androgen deprivation, the equivalent of endocrine therapy in breast cancer. Interestingly, KLF5 can also be activated through HER2 while HER2 inhibition was found to block KLF5-driven oncogenic phenotypes [313]. Taken together, these findings may point to an induction of ELF3 and KLF5 activities by HER2 in T47D and MCF7 TAMR and LTED clones.

SRBP1 was specifically activated in the T47D TAMR and LTED clones, while several TFs were significantly activated only in the MCF7 TAMR and LTED clones. SRBP1 and its paralog SRBP2 are required for lipid synthesis induced by oncogenic PI3K, such as in T47D cells harboring the activating *PIK3CA* p.H1047R mutation [314]. As lipogenesis and cholesterol synthesis have previously been implicated in resistance to aromatase inhibitor treatment [315], further SRBP1 activation in the T47D TAMR and LTED clones may have been a cell line specific adaptation to the therapeutic stress. Of note, MCF7 also carry an activating *PI3KCA* mutation [314] but SRBP1 was not significantly activated in all MCF7 TAMR and LTED clones, pointing to different rewiring between T47D and MCF7 endocrine therapy resistant clones.

Most TFs significantly activated in the MCF7 TAMR and LTED clones have previously been implicated in endocrine or other therapeutic resistance. Expression levels of *CEBPG* (encodes C/EBPγ) [316], *GABPA* (encodes GABPα) [317] and *TCF4* (encodes ITF2) [318] were elevated in endocrine therapy resistant cell lines or xenografts. KLF6 [319], MEF2A [320] and TEAD4 [321] have been implicated in (breast) cancer progression, while MEF2C [322] and MEIS1 [323] confer resistance to chemotherapeutics. *NCOA3* expression further correlated with worse recurrence-free survival for patients with invasive breast cancer [324] and melanoma [325]. *PRDM1* was found to be concurrently mutated with *PI3KCA* in a patient with endocrine therapy resistant breast cancer [326] and has been shown to be involved in chemoresistance of T-cell lymphoma [327].

No TF activity was significantly repressed in a shared manner between all T47D and MCF7 TAMR and LTED clones compared to their respective controls. A significant repression of ER activity however was expected because of the strong repression of estrogen signaling and the absence of activating *ESR1* mutations in the investigated hotspot region. Nevertheless, ER was among the top 10 most repressed TFs in MCF7 LTED clones (Supplementary Table 6). In the T47D TAMR and LTED clones, E2F4, FOXM1, IRF2, MYC, STAT1 and STAT2 were significantly less active whereas the activity of FOXA1 and FOXP1 was concurrently downregulated in MCF7 TAMR and LTED clones (Figure 22 C). E2F4 is important for a stable $G_2$ phase arrest in response to genotoxic stress in prostate cancer cells [328] and downregulation of E2F4 activity might thus have enabled the T47D TAMR and LTED clones to sustain a cycling phenotype under selective stress by the applied therapy. Reduced cell cycling activity as observed exemplarily in T47D *orange* LTED clones (Figure 25 A) may be reflected by the downregulation of transcriptional activity of MYC which generally promotes cell cycle progression [329]. FOXM1 is a transcriptional target of ER [330]. Hence, its significantly reduced activity may have been the result of reduced estrogen signaling in the T47D TAMR and LTED clones. For IRF2, tumor suppressive properties have recently been described in CRC [331, 332] and its downregulation may have promoted tumor progression to treatment resistant disease in the T47D cell line models. Reduced activity of STAT1 and STAT2 in isolated T47D TAMR and LTED clones likely represented the downregulation of JAK-STAT signaling in these clones. For the MCF7 TAMR and LTED clones, only the activity of the Forkhead box proteins FOXA1 and FOXP1 were repressed compared to the +E2 control cell lines. Clinically, dual suppression of FOXA1 and FOXP1 has been associated with worse relapse-free and overall survival in Tamoxifen-treated patients [333], suggesting a possible involvement in endocrine therapy resistance in the utilized MCF7 cell line model.

## 5.3 Kinome-wide alterations in endocrine therapy resistant clones

The RNA-Seq analysis highlighted prominent pathway alterations and the divergent re-wiring of important transcriptional networks in response to endocrine therapy. To complement this expression-based analysis, I was further interested in alterations of kinase activities in endocrine therapy resistant clones.

Like the TF activity analysis, the analysis of kinase activities underlined the heterogeneity of endocrine therapy resistance in my T47D and MCF7 cell line models with CDK1 as the only commonly repressed kinase compared to the control clones and cell lines (Figure 24). CK1ε was the only commonly activated kinase in all T47D TAMR and LTED clones while its isoform CK1δ [334] was not altered in any T47D TAMR or LTED clone. Accordingly, *CSNK1D* (encoding CK1δ) expression has been associated with longer overall survival while higher *CSNK1E* (encoding CK1ε) expression has been associated with shorter overall survival in patients with breast tumors [335]. These clinical findings support my *in vitro* observation, that CK1ε but not CK1δ may have been involved in tumor progression and therapy resistance.

Activation of AURKB, CaMK2δ, MAP3K8, MEK2, PKCß and of PKCδ was shared between all MCF7 TAMR and LTED clones. *AURKB* expression was previously found to be elevated in aromatase inhibitor-resistant MCF7 cell lines [336] and a high percentage of AURKB positive tumor cells was significantly associated with reduced disease-free and overall survival in Tamoxifen treated patients [337]. CaMK2δ is one of four calcium/calmodulin-dependent protein kinases II. This kinase family has been shown to modulate cell proliferation, cell cycle, invasion and metastasis, and therapy efficacy in different malignant diseases [338]. Suppression of estrogen signaling may have induced CaMK2δ activation as elevated CaMK2δ levels in ovariectomized rats was reduced by the addition of estrogen [339]. CaMK2δ might then have modulated important metabolic and signaling pathways with possible implications in resistance [340]. MAP3K8 and its downstream target MEK2 [341] are involved in MAPK signaling, which is commonly altered in endocrine therapy resistant disease. MAP3K8 has been identified as driver of androgen independent prostate cancer [342], the equivalent of endocrine therapy resistant breast cancer. In *in vitro* and *in vivo* models, combination of Anastrozole and the MEK specific inhibitor Selumetinib further resulted in significant growth reduction compared to Anastrozole treatment alone [343].

Conversely, AURKA, CDK1, CDK2, GSK-3ß, JNK3, mTOR and p38α were downregulated in all T47D or MCF7 TAMR and LTED clones (Figure 24 C). Downregulation of the cell cycle kinases AURKA, CDK1 and CDK2 likely was the result of reduced proliferation compared to treatment naïve control clones and cell lines, especially in the T47D TAMR and LTED clones (Figure 19). GSK-3ß has been considered either a tumor promoter or suppressor, depending on its cellular context [344]. In breast tumors, its downregulation has been shown to promote endocrine therapy resistance [345]. JNK3 [346] and p38α [347] have been described as tumor suppressor and promoter in HCC and ER- breast tumors, respectively. mTOR activation has been shown to induce endocrine therapy resistance [348], but appeared not to be a causal driver in my MCF7 *in vitro* resistance models.

Taken together, these findings on gene expression and phospho-proteomic level highlighted (i) the strong heterogeneity in my clonal endocrine therapy resistance *in vitro* models and (ii) the concordance between my findings and previously described findings. In the next step, I interfered with individual alterations in the T47D and MCF7 cell line models.

### 5.4 AKT signaling is important for T47D *orange* clones

Estrogen deprivation induced and maintained a reversible cell state commonly referred to as ‘cycling persister’ state [303] in T47D *orange* clones (Figure 19 C). Here, short-term estrogen deprived T47D *orange* +E2_5 clones may have represented the entry point to this cell state which was accompanied by significant upregulation of AKT1-3 activity (Figure 27). AKT inhibition led to a significant and similar reduction of observed cell viability by one-third on average in treatment-naïve, short-term estrogen deprived and LTED T47D *orange* clones (Figure 28 A), irrespective of their AKT1-3 activation. Clinically and in line with my *in vitro* findings, MK-2206 monotherapy has achieved only limited clinical activity in heavily pre-treated patients with advanced breast cancer carrying activating mutations in the PI3K/AKT/mTOR pathway or PTEN loss [286]. According to the authors of the study, tumor heterogeneity may have been one of the limiting factors. In my T47D cell line model however, MK-2206 was effective in reducing cell viability in heterogeneous (T47D WT) and clonal (T47D *orange* clones) populations, either as monotherapy or in combination with estrogen deprivation representing clinically utilized aromatase inhibitors. My experimental setting moreover closely mimicked a phase II clinical trial [349], where MK-2206 was added to the non-steroidal aromatase inhibitor Anastrozole after 28 days of previous treatment with Anastrozole alone. A strong reduction in proliferation as measured by Ki-67 staining was observed during the initial Anastrozole treatment but no further reduction was found upon additional MK-2206 treatment. Similarly, I hardly observed a reduction in total cell numbers when MK-2206 was additionally added to the pre-treated cells. Anastrozole treatment/short-term estrogen deprivation already led to an almost complete abrogation of cellular proliferation (Figure 19 C). Both clinical trials further confirmed additional adverse effects of MK-2206 treatment [286, 349]. Taken together, MK-2206 was effective in reducing the observed cellular viability on a (non-)clonal background in my *PI3KCA* mutated T47D *in vitro* model (Figure 28), but so far has had limited clinical efficacy when added to Anastrozole in the neoadjuvant setting or as monotherapy in heavily pre-treated patients with advance disease. The ability to target fast-proliferating and cycling persisting cells may not outweigh the lack of MK-2206’s added benefits in early clinical trials and in the light of its toxicity.

## 5.5 STAT2 mediated JAK-STAT signaling

The JAK-STAT pathway is an important pathway in immunity, cell division and death, and tumorigenesis [350]. Of the key downstream TFs, STAT2 and to a lesser extend STAT1 emerged as potential drivers of endocrine therapy resistance specifically in MCF7 LTED clones. Gene knockdown of either *STAT* and utilization of the JAK1/2 inhibitor Ruxolitinib however had similar effect on endocrine therapy sensitive control cells as well as TAMR and LTED clones (Figure 31, Figure 32 and Figure 33). Particularly the non-discriminatory effect of Ruxolitinib and the almost exclusive and strong activation of STAT2 point to activation and upregulation of the JAK-STAT pathway on transcriptional level by STAT2. After activation through type I interferon such as IFN-ß, phosphorylated STAT1 and STAT2 form a trimer with IFN-regulatory factor 9 (IRF9), known as IFN-stimulated gene (ISG) factor 3 (ISGF3) [351] leading to the increased expression of ISGs. Phosphorylated STAT2 can also form homodimers which interact with IRF9 and activate transcription without the involvement of STAT1 [352]. Unfortunately, no phosphorylated STAT2 could be detected by Mass Spectrometry, potentially due to its low abundance. Since MCF7 LTED clones with significantly activated STAT2 showed no increased sensitivity to Ruxolitinib, STAT2 activation may have been phosphorylation independent. Unphosphorylated STAT2 and IRF9 have previously been shown to drive *IL6* expression in the absence of STAT1 [353]. These findings highlight the potential of STAT2 (and IRF9) exclusive effects, irrespective of other STAT family members. I observed an upregulation of *IRF9* expression and IRF9 protein levels in MCF7 LTED clones whereas no upregulation of gene or protein levels was evident in MCF7 *blue* TAMR_1 clones (Supplementary Figure 9). Accordingly, higher STAT2 activity in MCF7 LTED clones may have been driven by higher abundancy of IRF9. Additionally, STAT2 activity can further be regulated by its protein stability. Exemplarily, different viruses have been shown to induce the proteasome-mediated degradation of STAT2 [354-356]. In melanoma, STAT2 stability has been shown to play an essential role in cell proliferation and colony growth [357]. Accordingly, STAT2 may have been more stable and therefore more active in MCF7 LTED clones compared to MCF7 +E2 control cell lines and *blue* TAMR_1 clones.

## 5.6 Private UPR activation and vulnerability to Bortezomib

In addition to NFκB activation that was observed in all MCF7 TAMR and LTED clones, MCF7 *blue* TAMR_1 clones showed an activation of the unfolded protein response (UPR) and accordingly a pronounced involvement of the proteasome (Figure 37 A). Therefore, I decided to use the proteasome inhibitor Bortezomib, which may (i) potentiate the UPR and/or (ii) inhibit NFκB signaling. Only MCF7 *blue* TAMR_1 cells were affected profoundly by Bortezomib (Figure 38), correlating with the prominent UPR activation seen in MCF7 *blue* TAMR (Figure 37 A), however not with the general activation of NFκB in all MCF7 TAMR and LTED clones

(Figure 34 A). It is tempting to speculate that the noxious effect of Bortezomib treatment was not mediated by inhibition of NFκB signaling but rather through the potentiation of the UPR activation in MCF7 *blue* TAMR_1 clones to a detrimental level [293]. UPR and NFkB signaling have previously been described to converge in endocrine therapy resistance [358]. The UPR comprises sophisticated mechanisms to restore protein homeostasis and protect cells against un- or misfolded proteins and the associated endoplasmic reticulum stress [294]. In brief and as described in detail exemplarily by Hetz *et al.* [294], protein kinase R (PKR)-like endoplasmic reticulum kinase (PERK, encoded by *EIF2AK3*), inositol-requiring enzyme 1 α (IRE1α, encoded by *ERN1*) and ATF6 are key effectors of the UPR. PERK phosphorylates eukaryotic translation initiation factor 2 subunit α (eIF2α) which ultimately leads to a decrease in mRNA translation initiation. Additionally, the transcription of further UPR target genes is activated through ATF4. IRE1α activates the expression of UPR target genes involved in endoplasmic reticulum proteostasis through splicing of *XBP1* mRNA and reduces protein load through cleavage of (non-coding) RNAs. ATF6 is activated by cleavage and then activates transcription of genes involved in protein folding homeostasis. Finally, un- or misfolded proteins are degraded by the endoplasmic-reticulum-associated protein degradation (ERAD) [294]. XBP1 was shown to be activated in breast cancer and specifically in TNBC, while ablation of *XBP1* inhibited tumor growth [359]. The PERK/ATF4 arm has been shown to be involved in the hypoxia-induced migration of breast cancer cell lines [360]. Specifically for ER+ breast cancer, Fulvestrant was shown to induce pro-survival UPR in a MCF7 model [361]. Elevated expression of a UPR gene signature was moreover a powerful prognostic marker of resistance to Tamoxifen, reduced time to recurrence and poor survival in breast cancer patients with ER+ disease [362]. This UPR activation in endocrine therapy treated patients may also be exploited clinically. Here, proteasome inhibition by Bortezomib may potentiate Fulvestrant's efficacy in the clinic as it enhances the aggregation of ER leading to the activation of a sustained UPR and ultimately apoptosis [363]. In accordance with this finding, a clinical phase II trial showed increased median progression free survival when adding Bortezomib to Fulvestrant for patients with AI-resistant metastatic breast tumors [364].

## 5.7 PKC activation as shared clonal driver of endocrine therapy resistance

The PKC family contains nine different isoforms which are grouped according to their mode of activation. 'Conventional' PKC isoforms α, ß and γ are activated by diacylglycerol (DAG) and Calcium ($Ca^{2+}$), while the 'novel' isoforms δ, ε, η and θ only require DAG for activation. 'Atypical' PKC isoforms ι and ζ are activated neither by DAG nor $Ca^{2+}$ [365-367]. PKC was exclusively been thought of as a tumor promoter since it was shown to be activated by TPA [368], the promoter in the classical two-step protocol for inducing skin cancer in mice utilizing 7,12-Dimethylbenz[a]anthracene (DMBA) as initiator [369]. Mechanistically, PKC binds to phorbol esters such as TPA with the same domain (C1) as PKC usually binds to DAG [370]. Decades of research have shown that different PKC isoforms can have tumor promoter and suppressor functions, also depending on the cellular context as extensively reviewed elsewhere [365-367, 371]. While all-*trans* retinoic acid (ATRA) induced growth arrest in ER+ and ER- cells, ATRA induced and decreased PKCα expression [372, 373], respectively, indicating opposing roles for regulating cell growth by PKCα. PKC isoforms may also compensate for the loss of another isoform, as exemplarily shown for the two typical isoforms α and ß [374]. Targeting individual isoforms by pharmacological inhibition may be challenging given the similarities between PKC isoforms. Staurosporine [375] and Sotrastaurin [247] are pan-PKC inhibitors and in the case of Staurosporine also inhibit other kinases with similar $IC_{50}$ values.

In accordance with potentially overlapping or compensatory modes of actions, different PKC isoforms were among the significantly activated kinases in my MCF7 TAMR and LTED clones (Figure 39 A) and I targeted PKC activation in all cell lines using the pan-PKC inhibitor Sotrastaurin. Here, reduction in cellular viability upon Sotrastaurin treatment largely correlated with the PKC activation of individual MCF7 TAMR and LTED clones (Figure 41). *Pink* LTED_2 clones only showed significant activation of PKCß and PKCδ and a weak response to Sotrastaurin treatment. PKCδ has been compared to *yin* and *yang* [371] as it may act tumor suppressive or promoting. In breast cancer, PKCδ is mainly considered tumor promoting [371] but has also been implicated in apoptosis induction and cell cycle arrest in $G_1$ phase in other cancer entities [376-379]. These findings could in part explain that activation of PKCδ alone may not predict sensitivity to Sotrastaurin treatment. Similarly, PKCη may generally promote the growth of breast cancer cells [380], but has also been shown to induce growth arrest and differentiation in other cell types [381-383]. The additional significant activation of PKCη in MCF7 *blue* TAMR_1 cells may explain their reduced sensitivity to Sotrastaurin compared to *yellow* LTED_2 and *red* LTED_5 clones which lacked significant activation of this isoform. Mechanistically, Sotrastaurin treatment had minor effects on the percentage of cycling cells but its growth inhibitory effect appeared to be driven at least in part through an increase in dying/dead cells (Figure 41 C and D). These findings are in accordance with previous studies indicating that PKCs predominantly affect $G_0/G_1$ and $G_2$ phases [384, 385].

Different PKC isoforms have previously been implicated in endocrine therapy resistance. *In vitro*, T47D and MCF7 TAMR cells expressed significantly higher levels of PKCα and PKCδ [386]. In patients, PKCδ expression was found to be associated with ER+ disease while PKCα expression was associated with ER- disease [386]. Double positivity of PKCα and PKCδ was associated with poor prognosis in patients receiving endocrine therapy [386]. PKCα expression was positively associated with TNBC [387] and HER2 overexpressing [388] tumors proposing a role of this PKC isoform in estrogen-independent breast cancers. In line, PKCα overexpression in MCF7 and T47D cell lines reduced ER expression [389, 390]. Overexpression of constitutively active PKCß moreover significantly increased the growth rate of MCF7 cells [391]. PKCε expression levels further correlated with estrogen receptor negative breast tumors as well as worse overall and disease-free survival [392]. Expression of the novel isoform *PRKCQ* (encoding PKCθ) has previously been described to be inversely correlated with ER positivity [393] while PKCη has been shown to contribute to breast cancer growth [380]. Taken together, these findings highlight the potentially important role of PKC isoforms in endocrine therapy resistance and estrogen-independent growth.

### 5.8 Clinical significance of *in vitro* findings

After I had analyzed clonal endocrine therapy resistance on different levels and validated individual resistance drivers *in vitro*, I further set out to validate my *in vitro* findings in the CPTAC-BRCA cohort. In accordance with previous reports of treatment-naïve patients with ER+ disease [95-100], only one *ESR1* mutation (p.R352K) in a single patient out of 89 patients with ER+ disease was found. Especially the stratification on per patient basis highlighted the strong inter-patient heterogeneity, as exemplarily illustrated on the pathway activation level (Figure 44). Individual patients with estrogen independent disease showed activation of *in vitro* identified drivers of endocrine therapy resistance, adding strong clinical relevance to my *in vitro* findings.

### 5.9 Limitations of the study

Despite the in-depth analysis and novel findings that are presented in my PhD thesis, the utilized barcoding approach had some limitations. In the bulk barcoding approach by Dr. Simone Borgoni [236], a selection of mainly three different clonal populations as indicated in Figure 10 A by *orange*, *green* and *yellow* barcodes was observed in the T47D +E2 control replicates. A similar neutral drift has previously been described in *Escherichia coli* over the time of 60,000 generations [394] and may have been due to selective advantages of individual clones even in the absence of stress. T47D *orange* +E2_5 clones appeared to have an additional selective advantage such as increased proliferation [395] over *green* clones in the +E2_5 replicate since the *orange* clonal population made up roughly 50% of the +E2_5 replicate while the *green* clonal population made up less than 5%. When analyzing *orange* and *green* +E2_5 clones individually however, T47D *green* +E2_5 clones proliferated significantly more than T47D *orange* +E2_5 clones (Figure 19 A). Accordingly, these findings contrast each other, and such proliferative dynamics would have more closely resembled the T47D +E2_2 replicate, which was dominated by the *green* +E2 clonal population (Figure 10 A). However, by isolating single clones and analyzing them separately, clonal cooperativity and competition which influence the cells' behavior in the tumor bulk [396] were abrogated and could not be taken into account. To address the questions of clonal cooperativity and competition, *orange* and *green* +E2_5 clones could be mixed and cultivated in different ratios. Afterwards, the proportion of *orange* and *green* +E2_5 clones could be determined by a barcode PCR and the results could exemplarily be compared to the +E2_2 and +E2_5 pools.

T47D *orange* clones were of particular interest as short-term estrogen deprivation of +E2_5 clones induced a 'cycling persister' phenotype that was stable for months (*orange* LTED clones). Lifting the applied pressure restored a strongly proliferative phenotype in these cells. AKT signaling appeared to be important for the short-term induction of the cycling persister phenotype but the AKT inhibitor MK-2206 showed limited efficiency, in our *in vitro* model and in patients. *Orange* clones of selected treatment conditions were subjected to RNA-Seq and (phospho-)proteomic analysis by Mass spectrometry. Subjecting the remaining samples (mainly LTED clones supplemented with estrogen) to RNA-Seq and Mass spectrometry may further reveal specific mechanisms responsible for this highly reversible phenotype that can be therapeutically exploited. Further RNA-Seq experiments may also be advisable to perform for Bortezomib treated MCF7 TAMR and LTED clones to determine the treatment's effect on UPR activation and p65 transcriptional activity by GSEA and DoRothEA, respectively.

To add clinical relevance to my *in vitro* findings, I utilized the recently published CPTAC-BRCA cohort [234]. This comparison had two major limitations. First, I had to compare endocrine therapy resistant *in vitro* models with ER- patients since no cohort with endocrine therapy treated patients containing phospho-proteomics data is currently publicly available. Secondly, activation scores of individual patients unavoidably represented assemble average measurements of different cells and potentially of different clonal populations. The utilized MCF7 cell lines I used however represented clonal populations. Recent technological innovations may improve the analysis of clones originating from the same tumor. Exemplarily, scRNA-Seq allows to analyze the expression profiles of single cells and would provide such insight on gene expression level. This technique could be complemented with single cell proteomics [397]. Ultra-high-sensitivity proteomics could further be coupled with high resolution microscopy. A study by Mund *et al.* could show that by utilizing such combination they were able to distinguish normal, pre-cancerous and cancerous lesions from melanoma tissue [398]. Combining multiple techniques, such as WGS, histopathology and spatial transcriptomics would further help to improve our understanding of breast cancer clones' evolutionary trajectories *in situ* as recently shown by Lomakin *et al.* [399].

My study further uncovered the need for drug development targeting TFs. I inhibited STAT1, STAT2 and p65 by gene knockdown. However, this knockdown reduced protein levels by > 90% and thereby the inhibitory effect could not be titrated. Gene knockdown results in a gradual reduction of protein levels, which cells might sense and counteract by activating salvage mechanisms. Drug effects mostly set in instantly thereby blocking protein activities without much delay. Accordingly, further improvements in pharmacologically targeting the ‘undruggable’ TFs [400] may help improve the therapeutic outcome of (breast) cancer patients. An example of targeting STAT TFs is Stattic, an inhibitor for STAT3 [401].

Finally, my study was also limited by aspects I did not look into. First and except for *ESR1* hotspot mutations covering codons 536 to 538, I only analyzed for gene expression and phospho-proteomic changes but did not inspect potential genetic alterations. Accordingly, my study systematically missed out on genetic drivers of endocrine therapy resistance. Additionally, I only utilized cancer cell line models in monolayer and -culture. Further studies should analyze cancer cells in a 3D environment or with cells from the tumor microenvironment, such as cancer associated fibroblasts. *In vivo* experiments could be performed to further validate my *in vitro* findings. The ‘initial’ cells could be injected into immunocompromised mice and these be treated with estrogen, Tamoxifen or an aromatase inhibitor such as Letrozole after successful tumor engraftment. Here, it would be of interest to see if the same barcodes would be enriched *in vitro* and *in vivo*. In the next step, single clones could be isolated from the mouse tumors and analyzed as performed for this thesis. Individual resistance drivers identified *in vitro* should also be challenged *in vivo*. After successful engraftment of different clonal tumor cells, mice could be treated with endocrine therapy alone or in combination, for example, with Bortezomib or Sotrastaurin.

# 6 Conclusion

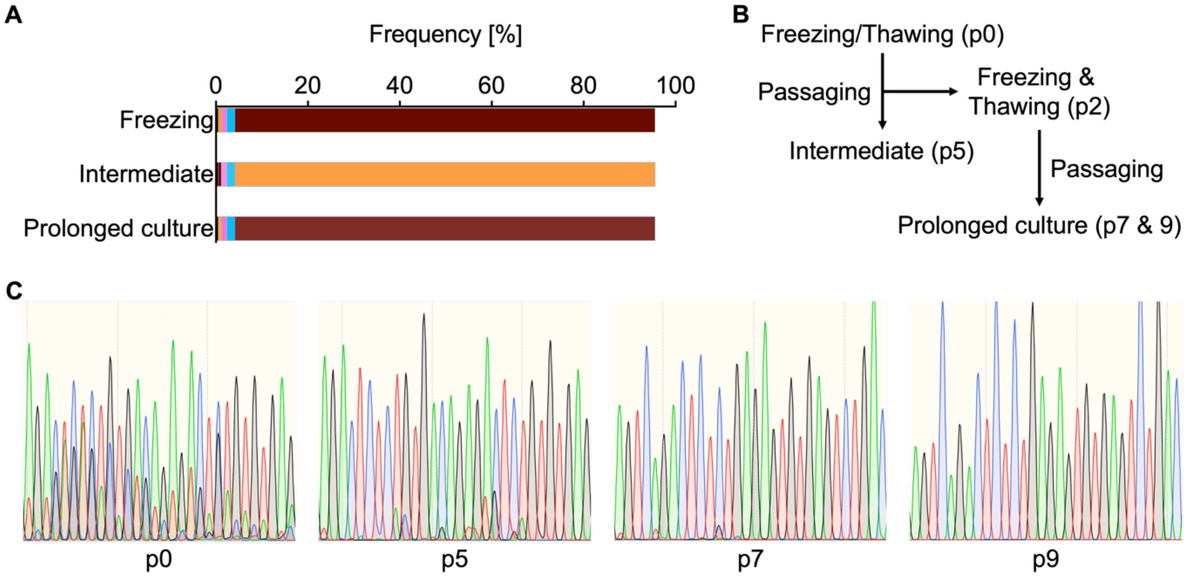
ER+ breast tumors can be effectively targeted by impinging on estrogen signaling using endocrine therapy. Nevertheless, treatment resistance may arise in up to 41% of patients who originally present with late stage disease [84]. Therapeutic failure can be driven by different mechanisms and tumor heterogeneity is increasingly appreciated to also contribute to therapy resistance [173]. Using an *in vitro* barcoding approach, Dr. Simone Borgoni found that Tamoxifen resistance can arise either by the selection of pre-existing clones or by the rewiring of initially treatment persisting clones [236]. In my study, I isolated and deeply characterized clonal endocrine therapy sensitive and resistant populations from these previously established *in vitro* models. Phenotypically, endocrine therapy resistant populations resembled either 'cycling persister' cells or fast proliferating cells. Further, endocrine therapy resistance was accompanied by private (specific for individual clones) and shared (between different clones) molecular alterations on gene expression and phospho-proteomic levels.

T47D *orange* clones were of strong interest since they were enriched in the treatment-naïve (+E2) and treated (-E2) conditions. Estrogen deprivation (-E2) induced a 'cycling persister' phenotype characterized by low proliferation and a strong reduction in cycling cells that was stable for months. This phenotype was reversed when the treatment was lifted, indicating that cells are capable to withstand therapeutic stress for a long period of time and quickly regain proliferative capacity once therapy is lifted. The induction of the 'cycling persister' phenotype was accompanied by significant activation of AKT. Treatment with the AKT inhibitor MK-2206 was effective at reducing cellular viability in 'cycling persister' and strongly proliferating cells but not specifically for short-term persisting cells, which had shown the strongest activation of AKT.

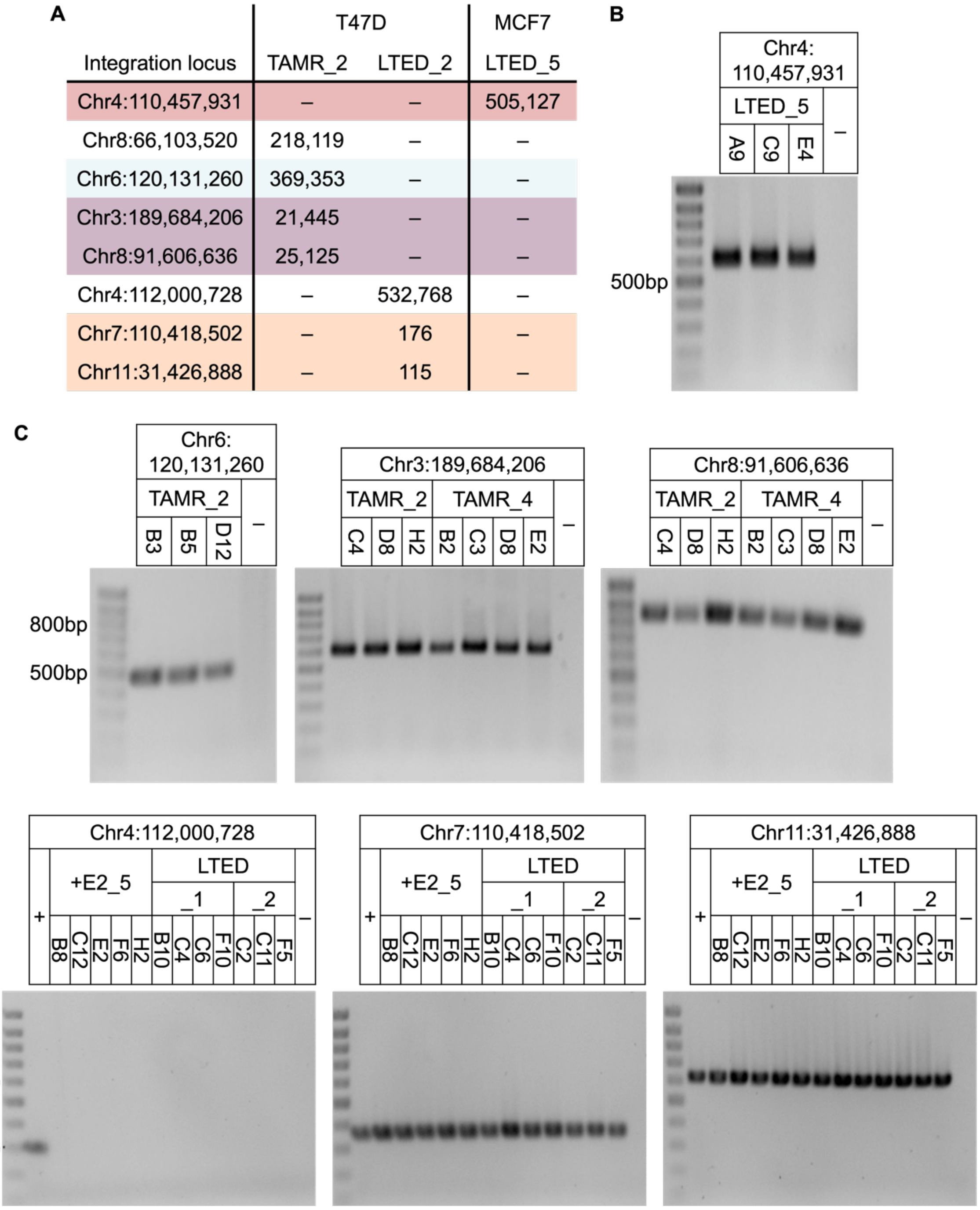
MCF7 TAMR and LTED clones displayed stronger proliferative capacities than their respective T47D counterparts. Here, I could pinpoint UPR and PKC activation as private and shared endocrine therapy resistance drivers, respectively. On the one hand, MCF7 *blue* TAMR_1 clones displayed a mild UPR which was likely potentiated to a detrimental level by low nanomolar concentrations of Bortezomib. On the other hand, strong reduction in cellular viability by PKC inhibition through Sotrastaurin was evident in multiple MCF7 clones. Sensitivity to Sotrastaurin roughly correlated with the stronger activation of different PKC isoforms.

Finally, I could highlight that the activation of individual resistance drivers can be masked by inter-patient heterogeneity in a clinical cohort, as was particularly evident in the case of PKC activation. Deconvoluting the cohort to the individual patient level revealed individual patients closely resembling my *in vitro* models adding clinical relevance to my clonal *in vitro* findings.

# 7 Supplementary Figures



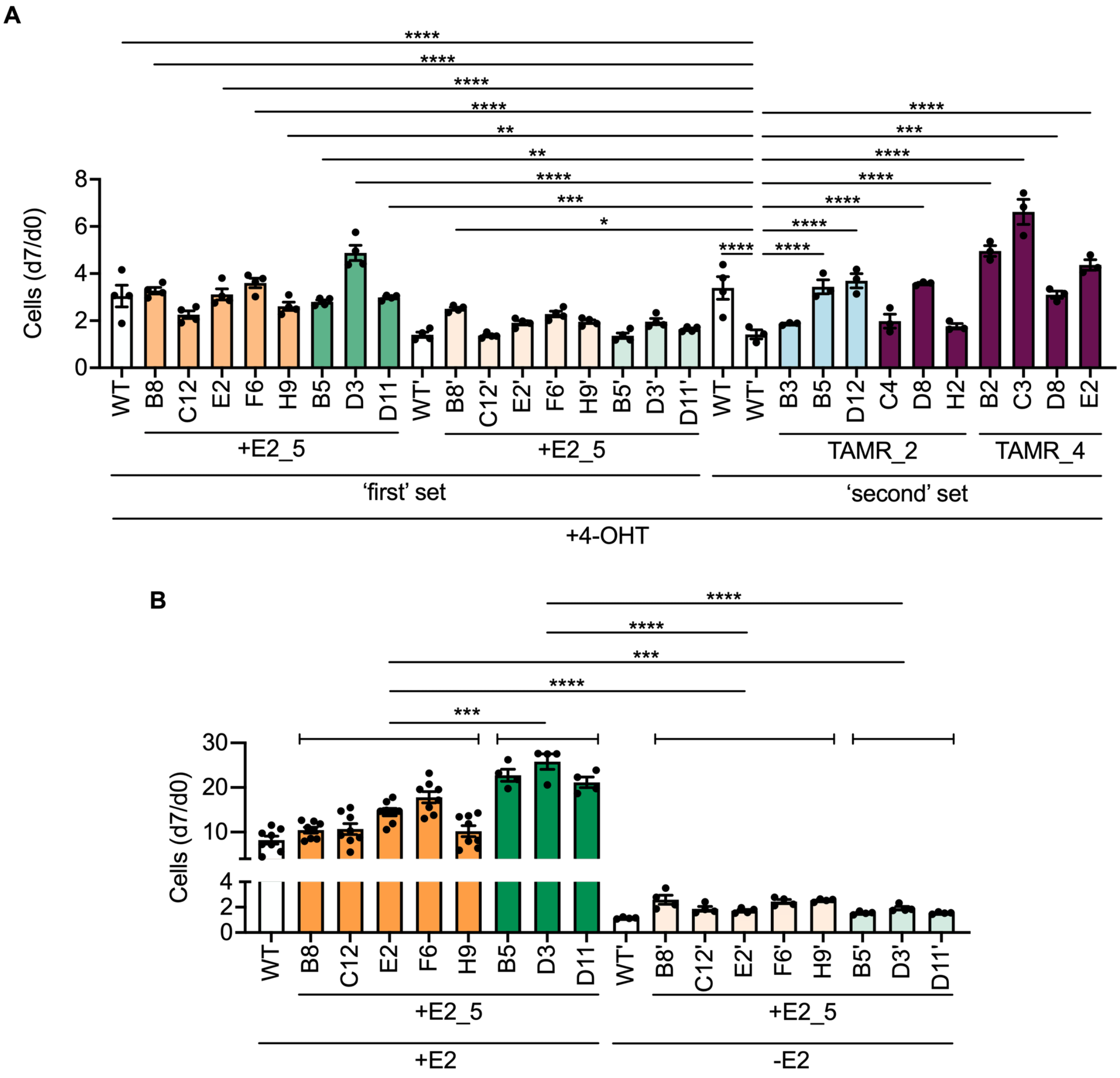
**Supplementary Figure 1: Clonality dynamics in T47D LTED_2 cell pool**. **A** Initial barcode analysis of LTED_2 after five months under estrogen deprivation revealed the *brown* barcode as dominating barcode in the cell pool (‘Freezing’). After further cultivation under estrogen deprivation two shifts in clonal dynamics were observed. In the initial five passages after thawing the *orange* clonal population outcompeted the *brown* clonal population (‘Intermediate’). After further cultivation, the *brown* clonal population reclaimed its dominance again (‘Prolonged culture’). Of note, cells were not cultivated consecutively but were frozen down, thawed and expanded again between the passages as shown in **B**. **C** Sanger sequencing chromatograms depicting the barcode sequences of LTED_2 cell pools under prolonged culture. The dominating sequence in the first two chromatograms represents the *orange* barcode whereas the dominating sequence in the third and fourth chromatograms represents the *brown* barcode.



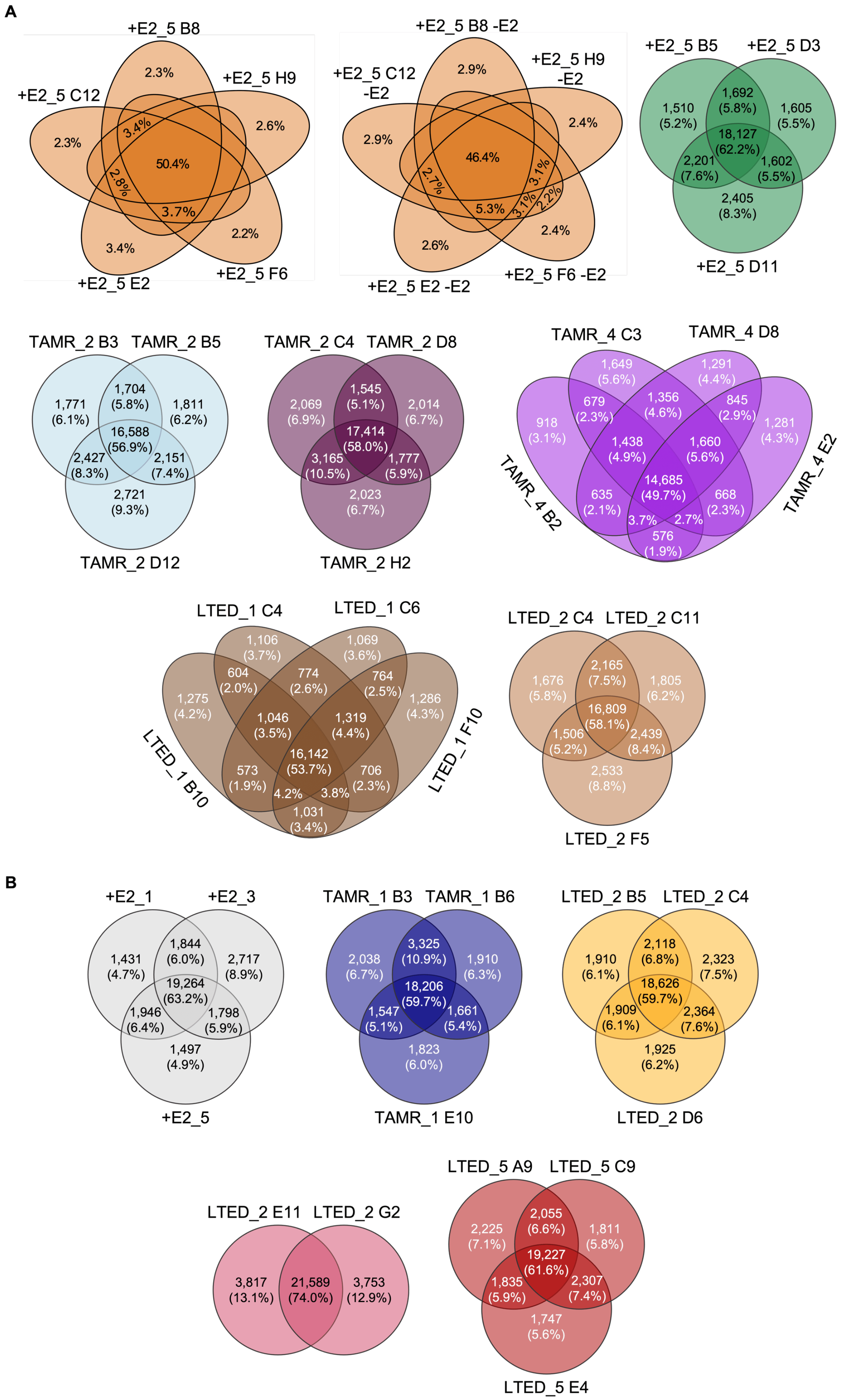
|  | T47D |  | MCF7 |
|---|---|---|---|
| Integration locus | TAMR_2 | LTED_2 | LTED_5 |
| Chr4:110,457,931 | – | – | 505,127 |
| Chr8:66,103,520 | 218,119 | – | – |
| Chr6:120,131,260 | 369,353 | – | – |
| Chr3:189,684,206 | 21,445 | – | – |
| Chr8:91,606,636 | 25,125 | – | – |
| Chr4:112,000,728 | – | 532,768 | – |
| Chr7:110,418,502 | – | 176 | – |
| Chr11:31,426,888 | – | 115 | – |

A
B
Chr4: 110,457,931
LTED_5
A9
C9
E4
–
500bp
C
Chr6: 120,131,260
TAMR_2
B3
B5
D12
–
800bp
500bp
Chr3:189,684,206
TAMR_2
TAMR_4
C4
D8
H2
B2
C3
D8
E2
–
Chr8:91,606,636
TAMR_2
TAMR_4
C4
D8
H2
B2
C3
D8
E2
–
D
Chr4:112,000,728
+
+E2_5
LTED
_1
_2
B8
C12
E2
F6
H2
B10
C4
C6
F10
C2
C11
F5
–
600bp
300bp
Chr7:110,418,502
+
+E2_5
LTED
_1
_2
B8
C12
E2
F6
H2
B10
C4
C6
F10
C2
C11
F5
–
Chr11:31,426,888
+
+E2_5
LTED
_1
_2
B8
C12
E2
F6
H2
B10
C4
C6
F10
C2
C11
F5
–

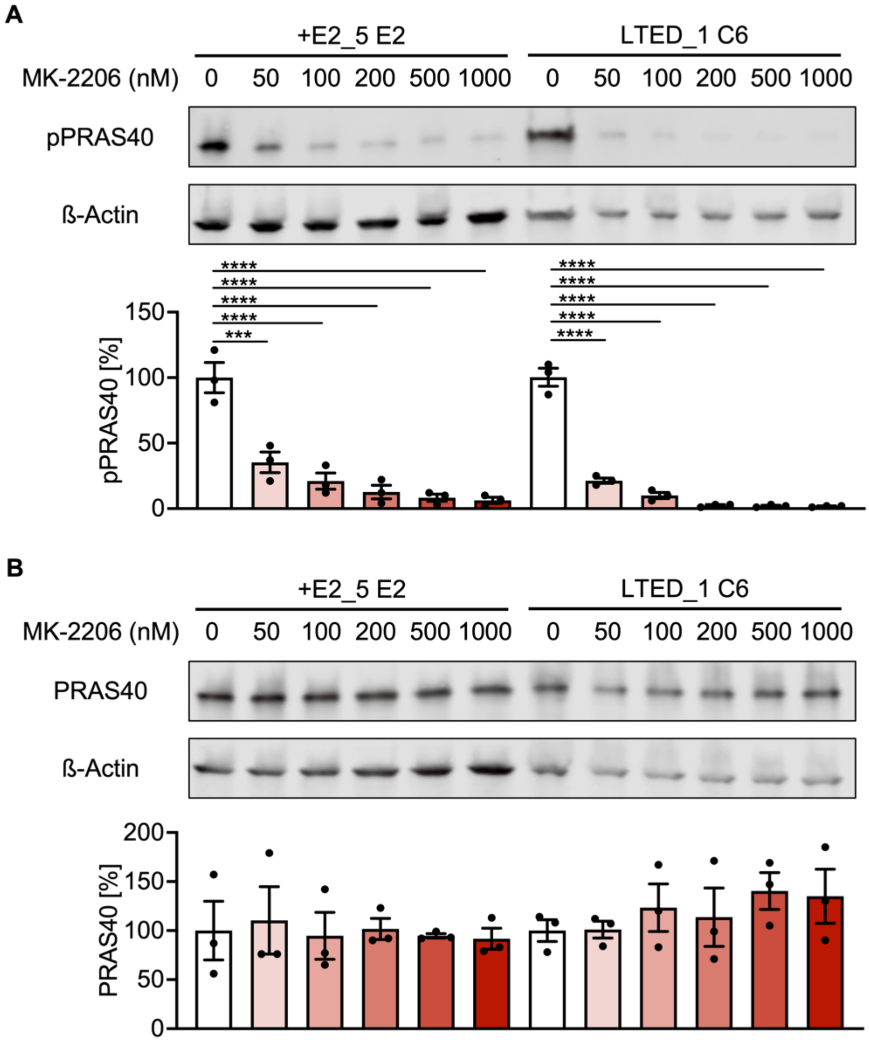
**Supplementary Figure 2: Integration site verification for T47D and MCF7 clones**. **A** List of viral integration sites and sequence counts from the analysis of barcoded cell pools. Color coding indicates the barcode estimated to be represented by the viral integration site according to Figure 14. Integration sites verified in the corresponding barcoded cell pools that could not be identified for isolated clones are lacking a color coding. Exemplarily, the viral integration site in Chr4:112,000,728 could be validated for the T47D LTED_2 cell pool ('+') but not for the *orange* clones isolated from the T47D +E2_5, LTED_1 and LTED_2 replicates (**D**). Primers were designed for a unique part of the viral 3'LTR and regions of the host's genome adjacent to the viral integration site. **B**-**D** Validated integration sites for MCF7 *red* LTED_5 clones (**B**), T47D *turquoise* TAMR_2 (left), T47D *purple* TAMR_2 and TAMR_4 clones (middle and right) (**C**), and T47D *orange* +E2_5, LTED_1 and LTED_2 clones (**D**).



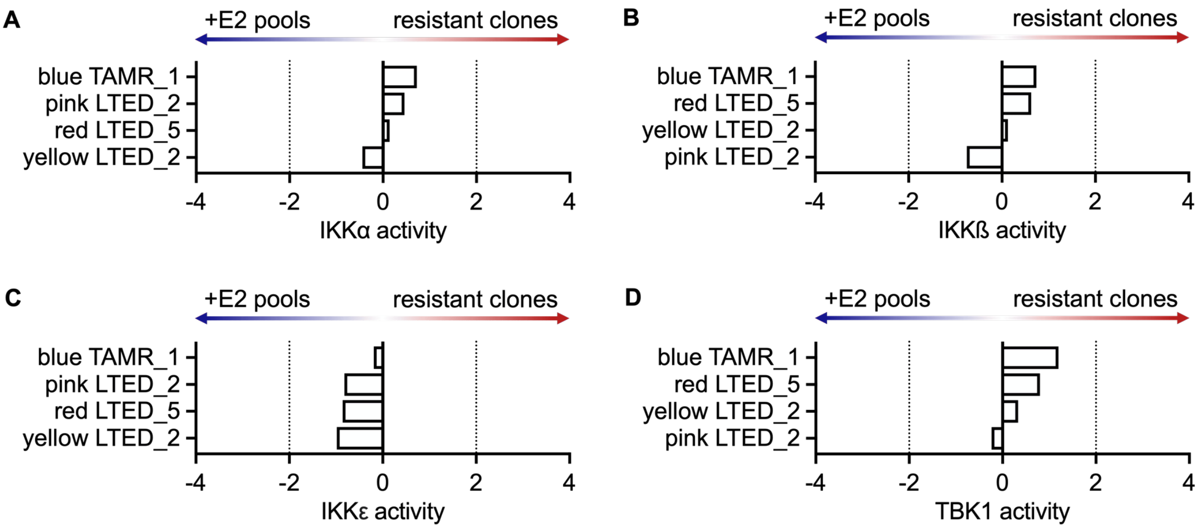
**Supplementary Figure 3: Proliferation of T47D *orange* and *green* +E2_5 clones under endocrine therapy.** **A** Proliferation of *orange* and *green* +E2_5 clones under 4-OHT treatment compared to TAMR_2 and TAMR_4 clones. 'Second' set refers to data shown in Figure 19 B. **B** Proliferation of *orange* and *green* +E2_5 clones under normal growth conditions (+E2) and estrogen deprivation (-E2). For WT and *orange* +E2_5 clones +E2, data from Figure 19 A and C was combined. **A** and **B** Cells were kept in or treated with the indicated media (+E2, +4-OHT, -E2). WT indicates non barcoded cells. ' indicates 14 day pre-treatment with 4-OHT or by estrogen deprivation (-E2) and then continued treatment for the duration (7 days) of the proliferation assay. Cell numbers were determined by microscopy-based nuclei counting at day 0 (d0) and day 7 (d7). **A** Given the different phenotypes depicted by the TAMR clones (non-proliferative: TAMR_2 B3, C4 and H2 vs proliferative: TAMR_2 B5, D12, D8, TAMR_4 B2, C3, D8 and E2), clones with the same barcode were not grouped together but rather compared to pre-treated non-barcoded T47D WT cells as also for Figure 19 B. **B** Clones with the same barcode and treatment (±E2) were grouped and the grouped clones compared. **A** and **B** * represents $p<0.05$, ** represents $p<0.01$, *** represents $p<0.001$ and **** represents $p<0.0001$ as determined by one-way ANOVA with Tukey or Dunnett multiple comparisons test.


A
+E2_5 B8
+E2_5 H9
+E2_5 C12
+E2_5 E2
+E2_5 F6
2.3%
2.6%
2.3%
3.4%
50.4%
2.8%
3.7%
2.2%
3.4%
+E2_5 B8 -E2
+E2_5 H9 -E2
+E2_5 C12 -E2
+E2_5 E2 -E2
+E2_5 F6 -E2
2.9%
2.4%
2.9%
46.4%
2.7%
3.1%
5.3%
3.1%
2.2%
2.4%
2.6%
+E2_5 B5
+E2_5 D3
+E2_5 D11
1,510 (5.2%)
1,692 (5.8%)
1,605 (5.5%)
18,127 (62.2%)
2,201 (7.6%)
1,602 (5.5%)
2,405 (8.3%)
TAMR_2 B3
TAMR_2 B5
TAMR_2 D12
1,771 (6.1%)
1,704 (5.8%)
1,811 (6.2%)
16,588 (56.9%)
2,427 (8.3%)
2,151 (7.4%)
2,721 (9.3%)
TAMR_2 C4
TAMR_2 D8
TAMR_2 H2
2,069 (6.9%)
1,545 (5.1%)
2,014 (6.7%)
17,414 (58.0%)
3,165 (10.5%)
1,777 (5.9%)
2,023 (6.7%)
TAMR_4 C3
TAMR_4 D8
TAMR_4 B2
TAMR_4 E2
1,649 (5.6%)
1,291 (4.4%)
679 (2.3%)
1,356 (4.6%)
845 (2.9%)
918 (3.1%)
1,281 (4.3%)
1,438 (4.9%)
1,660 (5.6%)
14,685 (49.7%)
635 (2.1%)
668 (2.3%)
3.7%
2.7%
576 (1.9%)
LTED_1 C4
LTED_1 C6
LTED_1 B10
LTED_1 F10
1,106 (3.7%)
1,069 (3.6%)
604 (2.0%)
774 (2.6%)
764 (2.5%)
1,275 (4.2%)
1,286 (4.3%)
1,046 (3.5%)
1,319 (4.4%)
16,142 (53.7%)
573 (1.9%)
706 (2.3%)
4.2%
3.8%
1,031 (3.4%)
LTED_2 C4
LTED_2 C11
LTED_2 F5
1,676 (5.8%)
2,165 (7.5%)
1,805 (6.2%)
16,809 (58.1%)
1,506 (5.2%)
2,439 (8.4%)
2,533 (8.8%)
B
+E2_1
+E2_3
+E2_5
1,431 (4.7%)
1,844 (6.0%)
2,717 (8.9%)
19,264 (63.2%)
1,946 (6.4%)
1,798 (5.9%)
1,497 (4.9%)
TAMR_1 B3
TAMR_1 B6
TAMR_1 E10
2,038 (6.7%)
3,325 (10.9%)
1,910 (6.3%)
18,206 (59.7%)
1,547 (5.1%)
1,661 (5.4%)
1,823 (6.0%)
LTED_2 B5
LTED_2 C4
LTED_2 D6
1,910 (6.1%)
2,118 (6.8%)
2,323 (7.5%)
18,626 (59.7%)
1,909 (6.1%)
2,364 (7.6%)
1,925 (6.2%)
LTED_2 E11
LTED_2 G2
3,817 (13.1%)
21,589 (74.0%)
3,753 (12.9%)
LTED_5 A9
LTED_5 C9
LTED_5 E4
2,225 (7.1%)
2,055 (6.6%)
1,811 (5.8%)
19,227 (61.6%)
1,835 (5.9%)
2,307 (7.4%)
1,747 (5.6%)

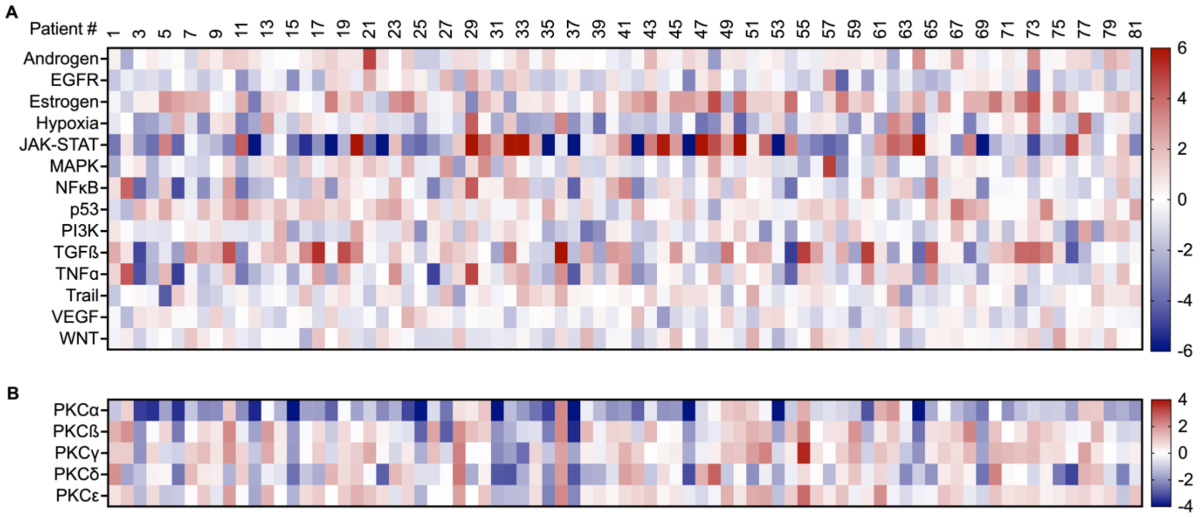
**Supplementary Figure 4: Overlap of detected phosphosites in biological replicates.** Shown are the overlaps in detected phosphosites for biological replicates of T47D (**A**) and MCF7 (**B**) control cell lines and isolated clones. Particularly for T47D *orange* +E2_5 clones (**A**) only percentages are shown. Additionally, only values > 2.0% are highlighted for better visualization.



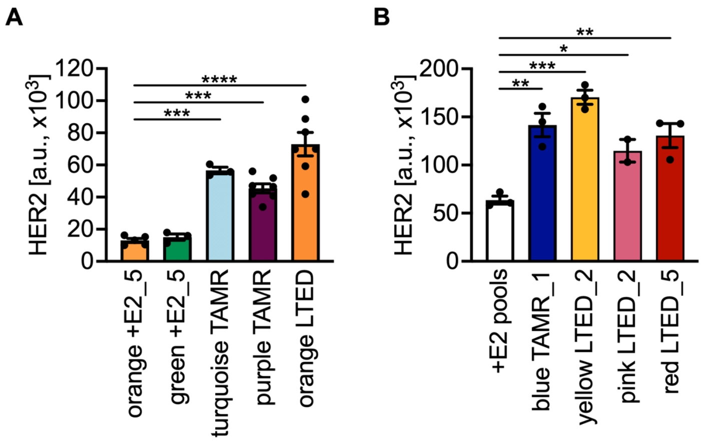
**Supplementary Figure 5: Quantification of AKT inhibition by MK-2206.** T47D +E2_5 E2 and LTED_1 C6 clones were treated with DMSO (‘0’) or increasing MK-2206 concentrations for 1 h and the inhibition of downstream signaling was estimated from PRAS40 Thr246 phosphorylation levels. A representative Western Blot image for pPRAS40 (**A**) and PRAS40 (**B**) and their loading controls are shown. n = 3. (p)PRAS40 signals were normalized to the respective ß-Actin loading controls and then normalized to DMSO-treated (‘0’) +E2_5 E2 and LTED_1 C6 cells, respectively. Highlighted are mean ± SEM. *** represents $p<0.001$ and **** represents $p<0.0001$ as determined by one-way ANOVA with Dunnett multiple comparisons test.



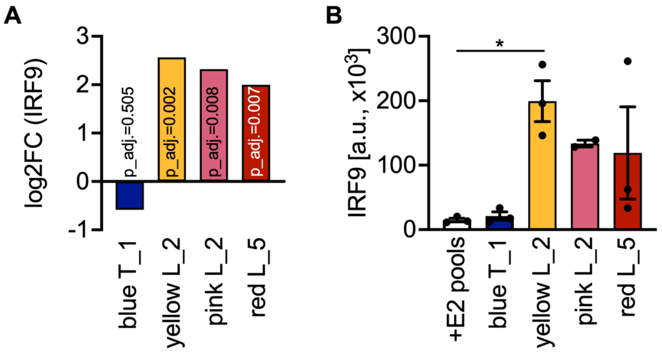
**Supplementary Figure 6: IKK activities in MCF7 TAMR and LTED clones.** IKKα (**A**), IKKß (**B**), IKKε (**C**) and TBK1 (**D**) are capable of phosphorylating IκB. Absolute activity scores ≥ 2 were considered significant as indicated by dashed lines. No significant deregulation of either IKK were evident from the phospho-proteomics data. Analyses were performed by Dr. Efstathios-Iason Vlachavas.



**Supplementary Figure 7: Pathway and PKC activity profiles of individual patients with ER+ disease from the CPTAC cohort.** Supplementary Figure to Figure 44 and Figure 45 C. Pathway (**A**) and PKC activities (**B**) were determined as described above. Activation and repression are indicated by red and blue color coding, respectively. For better visualization extreme values were adjusted to the indicated scales. Significance of deregulations are not indicated. Analyses were performed by Dr. Efstathios-Iason Vlachavas.



**Supplementary Figure 8: HER2 protein overexpression in T47D and MCF7 TAMR and LTED clones.** Protein levels were determined by Mass spectrometry for T47D (**A**) and MCF7 (**B**). * represents p<0.05, ** represents p<0.01, *** represents p<0.001 and **** represents p<0.0001 as determined by one-way ANOVA with Dunnett multiple comparisons test. a.u.: arbitrary units.



**Supplementary Figure 9: *IRF9* gene and IRF9 protein overexpression in MCF7 TAMR and LTED clones. A** Log2FC and adjusted p-values of *IRF9* compared to +E2 control cell lines. **B** Protein levels were determined by Mass spectrometry. * represents p<0.05 as determined by one-way ANOVA with Dunnett multiple comparisons test. T_1: TAMR_1. L_2: LTED_2. L_5: LTED_5. a.u.: arbitrary units.

# 8 Supplementary Tables

**Supplementary Table 1: Sequences and color coding of barcodes representing isolated clones.**

| Barcode sequence | Cell line(s) | Color coding | Individual clones | Designation in Figure 16 |
|---|---|---|---|---|
| AGACTCTCTGTGACAGAGACTCTGTGTGAG | T47D +E2_5 and LTED | *orange* | +E2_5: B8, C12, E2, F6, H9<br>LTED_1: B10, C4, C6, F10<br>LTED_2: C2, C11, F5 | a |
| AGACTCTGTGTCACACTGAGTCTGAGTGTG | T47D +E2_5 | *green* | B5, D3, D11 | b |
| TGTGACTGTGAGACACAGAGAGTGAGTGTG | T47D TAMR_2 | *turquoise* | B3, B5, D12 | c |
| AGAGACAGTGTGACTGTCACTGAGAGAGAG | T47D TAMR_2 and TAMR_4 | *dark red/purple* | TAMR_2: C4, D8, H2<br>TAMR_4: B2, C3, D8, E2 | d |
| ACTGAGACTGAGTCACACTGTCACAGTGAC | T47D TAMR_2 and TAMR_4 | *dark blue/purple* | TAMR_2: C4, D8, H2<br>TAMR_4: B2, C3, D8, E2 | e |
| ACAGAGTGAGAGTGTGAGTCTGAGTCACAG | MCF7 TAMR_1 | *blue* | B3, B6, E10 | a |
| ACTGTGTGTGACAGAGACAGAACTCTGACA | | | | b |
| AGAGACTCAGTCTGTGAGACTGACACAGTC | | | | c |
| AGTCTGTCTGTGTCTCACAGAGTGTGAGAC | | | | d |
| AGTCTGTGTGTCTCTGTCTGTCTCTGAGTG | | | | e |
| AGTGAGAGTGTCAGTCTGACTGACACACAG | | | | f |
| TCAGTGAGAGTGAGTCTCAGTCTGAGTGTG | | | | g |
| TCTGTGTGTGAGTGACTGAGTGTCAGTGTG | | | | h |
| TGACACTGAGAGAGAGAGTGACACAGTGAC | | | | i |
| TGACAGAGTGTGTCTGAGAGTGAGTGAGTG | | | | j |
| TGACTGTGTGTGCAGAGAGTGAGTGTGTCA | | | | k |
| TGTCAGAGTGAGTGTGACTCACTCAGAGTC | | | | l |
| TGTCTGTCAGACTGTGAGAGAGAGTCAGAC | | | | m |
| TGTCTGTGTGACAGTCACACTGTCTCAGTC | | | | n |
| TGTGAGTGTCACTGACTGTCTGAGAGTGAG | | | | o |
| TGTGTGAGACTGTCTCACAGTCAGAGAGAC | | | | p |
| TGTGTGTCAGAGTGTCTGACAGAGAGTGAG | | | | q |
| AGACACAGAGTGAGAGAGAGACACAGAGTG | MCF7 LTED_2 | *yellow* | B5, C4, D6 | r |
| AGAGTGTGTCTGTGTGTGACTGACAGTCTG | | | | s |
| AGTCTGTGAGTGTGTGTCTCACTGACTGTG | | | | t |

| | | | | |
|---|---|---|---|---|
| TCTGTCACACTGCGACACTCTCAGTCAGTC | | | | u |
| TGTGTCACACACTGACTCTGTCAGTGAGTG | | | | v |
| ACAGTGTGTCTGAGACACAGTGACACAGTG | MCF7 LTED_2 | *pink* | E11, G2 | w |
| AGTCTCAGAGTGAGTCTGAGTGTCTGTGTC | | | | x |
| AGTGTCTCAGTGTCTGTCAGTGACACTCTG | | | | y |
| TGACTGAGAGTGAGACACACTGTCACACAC | | | | z |
| TGACTGAGTCTGAGTCTGAGAGAGTGTCTC | | | | ä |
| TGTCTCTCAGAGTGAGTGAGAGTGACTCAC | | | | ö |
| AGACTCTGAGACTGTCTGTGAGTGACTCAC | MCF7 LTED_5 | *red* | A9, C9, E4 | ü |

**Supplementary Table 2: Genomic integration sites of individual barcodes.** Verified clonal integration sites were investigated with the UCSC genome browser for potential disruption of regulatory or coding regions using the GRCh38/hg38 assembly (https://genome.ucsc.edu/cgi-bin/hgGateway).

| Integration locus | Cell line and barcode | Nearest gene (within 30 kb) | Integration |
|---|---|---|---|
| Chr3:189,684,206 | T47D *purple* TAMR | *TP63* | *TP63*<br>Intron 1 |
| Chr4:110,457,931 | MCF7 *red* LTED_5 | – | – |
| Chr6:120,131,260 | T47D *turquoise* TAMR_2 | – | – |
| Chr7:110,418,502 | T47D *orange* +E2_5 and LTED | *ENSG00000226965* | *ENSG00000226965*<br>Intron 4 |
| Chr8:91,606,636 | T47D *purple* TAMR | *ENSG00000253901* | *ENSG00000253901*<br>Intron 2 |
| Chr11:31,426,888 | T47D *orange* +E2_5 and LTED | *DNAJC24* | *DNAJC24*<br>Intron 4 |

**Supplementary Table 3: Gene set enrichment analysis of top 200 most variable genes for T47D.** The GSEA online tool was used (http://www.gsea-msigdb.org/gsea/msigdb/human/annotate.jsp, accessed: 26.01.23).

| Hallmark Gene Set | Genes in Gene Set | Genes in Overlap | Overlap [%] | p-value | FDR q-value |
|---|---|---|---|---|---|
| EPITHELIAL_MESENCHYMAL_TRANSITION | 200 | 11 | 5.50 | 4.89E-09 | 2.45E-07 |
| ESTROGEN_RESPONSE_EARLY | 200 | 8 | 4.00 | 6.89E-06 | 1.15E-04 |
| ESTROGEN_RESPONSE_LATE | 200 | 8 | 4.00 | 6.89E-06 | 1.15E-04 |
| COAGULATION | 138 | 6 | 4.35 | 6.28E-05 | 7.85E-04 |
| INTERFERON_GAMMA_RESPONSE | 200 | 6 | 3.00 | 4.7E-04 | 3.35E-03 |
| KRAS_SIGNALING_UP | 200 | 6 | 3.00 | 4.7E-04 | 3.35E-03 |
| MYOGENESIS | 200 | 6 | 3.00 | 4.7E-04 | 3.35E-03 |
| INTERFERON_ALPHA_RESPONSE | 97 | 4 | 4.12 | 1.33E-03 | 8.34E-03 |
| KRAS_SIGNALING_DN | 200 | 5 | 2.50 | 3.09E-03 | 1.72E-02 |

**Supplementary Table 4: Gene set enrichment analysis of top 200 most variable genes for MCF7.** The GSEA online tool was used (http://www.gsea-msigdb.org/gsea/msigdb/human/annotate.jsp, accessed: 26.01.23).

| Hallmark Gene Set | Genes in Gene Set | Genes in Overlap | Overlap [%] | p-value | FDR q-value |
|---|---|---|---|---|---|
| ESTROGEN_RESPONSE_EARLY | 200 | 15 | 7.50 | 7.61E-14 | 3.8E-12 |
| ESTROGEN_RESPONSE_LATE | 200 | 11 | 5.50 | 4.64E-09 | 1.16E-07 |
| EPITHELIAL_MESENCHYMAL_TRANSITION | 200 | 9 | 4.50 | 6.58E-07 | 1.1E-05 |
| INTERFERON_GAMMA_RESPONSE | 200 | 8 | 4.00 | 6.64E-06 | 6.64E-05 |
| KRAS_SIGNALING_UP | 200 | 8 | 4.00 | 6.64E-06 | 6.64E-05 |
| INTERFERON_ALPHA_RESPONSE | 97 | 6 | 6.19 | 8.28E-06 | 6.9E-05 |
| MYOGENESIS | 200 | 7 | 3.50 | 5.91E-05 | 4.22E-04 |
| APOPTOSIS | 161 | 5 | 3.11 | 1.17E-03 | 7.34E-03 |

**Supplementary Table 5: Transcription factor activities for T47D TAMR and LTED clones.** The top 10 activated and repressed TFs are listed with their NES. Absolute activity scores ≥ 2 were considered significant. TF activities in the TAMR and LTED clones were compared to the activities in *orange* (and *green*) +E2_5 clones. Positive values indicate higher activities in TAMR and LTED clones while negative values indicate higher activities in *orange* (and *green*) +E2_5 clones. Analyses were performed by Dr. Efstathios-Iason Vlachavas. For simplicity and unambiguity, proteins are referred to by their nonitalicized gene names as previously recommended [402].

| ***turquoise* TAMR vs *orange* and *green* +E2_5** | | ***purple* TAMR vs *orange* and *green* +E2_5** | | ***orange* LTED vs *orange* and *green* +E2_5** | | ***orange* LTED vs *orange* +E2_5** | |
|---|---|---|---|---|---|---|---|
| **TF** | **NES** | **TF** | **NES** | **TF** | **NES** | **TF** | **NES** |
| GRHL2 | 4.86 | KLF5 | 4.46 | ZBT11 | 4.34 | ZNF263 | 5.06 |
| KLF5 | 4.86 | SIX5 | 3.81 | MAFF | 4.13 | ZBTB11 | 3.43 |
| ESRRA | 4.13 | RFX1 | 3.80 | GABPA | 3.67 | NCOA1 | 2.82 |
| RFX1 | 4.09 | GRHL2 | 3.58 | NCOA1 | 3.45 | ELF5 | 2.59 |
| PRDM14 | 3.42 | THAP1 | 3.13 | ZNF263 | 3.39 | BHLHE40 | 2.52 |
| SIX5 | 3.13 | PRDM14 | 3.04 | MEF2A | 3.24 | MAFF | 2.43 |
| SREBF1 | 3.00 | RFX2 | 2.90 | GRHL2 | 3.16 | GABPA | 2.38 |
| FOXK2 | 2.71 | ZNF639 | 2.76 | ELF3 | 3.07 | TBP | 2.23 |
| CREB3 | 2.59 | FOXP1 | 2.76 | SREBF1 | 3.00 | SP2 | 2.06 |
| FOXO3 | 2.48 | HIF1A | 2.52 | TBP | 2.66 | KLF13 | 2.03 |
| ETS2 | -2.54 | ZNF263 | -3.79 | NR2F1 | -3.39 | EBF1 | -4.13 |
| E2F7 | -2.55 | ARID2 | -3.90 | MYC | -3.60 | NR2F1 | -4.16 |
| IRF2 | -2.95 | E2F1 | -4.03 | IRF2 | -3.64 | SP1 | -4.67 |
| STAT1 | -3.13 | FOXM1 | -4.21 | SPI1 | -3.71 | IRF2 | -4.89 |
| FOXM1 | -3.14 | IRF2 | -4.23 | RARA | -3.74 | SPI1 | -4.94 |
| ZBTB7A | -3.64 | STAT1 | -4.25 | SP1 | -3.78 | STAT1 | -5.09 |
| E2F4 | -3.64 | ZBTB7A | -4.43 | NFKB1 | -4.04 | NFKB1 | -5.20 |
| ARID2 | -3.66 | MYC | -4.64 | LYL1 | -4.46 | LYL1 | -5.71 |
| TCF7 | -3.73 | STAT2 | -5.36 | IRF1 | -5.90 | IRF1 | -8.23 |
| STAT2 | -7.18 | E2F4 | -6.58 | STAT2 | -7.40 | STAT2 | -10.53 |

**Supplementary Table 6: Transcription factor activities for MCF7 TAMR and LTED clones.** The top 10 activated and repressed TFs are listed with their NES. Absolute activity scores ≥ 2 were considered significant. TF activities in the TAMR and LTED clones were compared to the activities in +E2_1, +E2_3 and +E2_5 control cell lines. Positive values indicate higher activities in TAMR and LTED clones while negative values indicate higher activities in +E2 control cell lines. Analyses were performed by Dr. Efstathios-Iason Vlachavas. For simplicity and unambiguity, proteins are referred to by their nonitalicized gene names as previously recommended [402].

| ***blue* TAMR_1 vs +E2 pools** | | ***yellow* LTED_2 vs +E2 pools** | | ***pink* LTED_2 vs +E2 pools** | | ***red* LTED_5 vs +E2 pools** | |
|---|---|---|---|---|---|---|---|
| **TF** | **NES** | **TF** | **NES** | **TF** | **NES** | **TF** | **NES** |
| TEAD1 | 8.47 | STAT2 | 6.94 | ZNF384 | 5.19 | TEAD1 | 6.00 |
| MEF2A | 7.09 | KLF5 | 6.81 | NCOA3 | 5.17 | STAT2 | 5.07 |
| ARNT | 5.28 | KLF6 | 4.31 | ARID2 | 5.03 | MEF2A | 4.69 |
| KLF6 | 5.04 | GRHL2 | 4.05 | MEF2A | 4.70 | NCOA3 | 3.96 |
| GRHL2 | 4.94 | IRF1 | 3.77 | FOXO3 | 4.66 | GRHL2 | 3.81 |
| TEAD4 | 4.73 | THAP1 | 3.65 | STAT2 | 4.56 | TEAD4 | 3.79 |
| NCOA3 | 4.70 | TBP | 3.57 | MEF2C | 4.24 | ELF1 | 3.65 |
| AR | 4.48 | PRDM1 | 3.38 | MAFG | 4.07 | SMAD4 | 3.61 |
| TCF12 | 4.41 | TEAD1 | 3.31 | ARNT | 3.91 | TBP | 3.55 |
| RBPJ | 4.37 | SOX6 | 3.10 | SIX2 | 3.79 | ZNF263 | 3.38 |
| HBP1 | -2.84 | PAX5 | -2.53 | SP1 | -2.71 | LYL1 | -2.61 |
| FOXA1 | -2.96 | DUX4 | -2.55 | MYC | -3.05 | MBD2 | -2.90 |
| ZNF639 | -3.28 | ESR1 | -3.13 | MITF | -3.10 | E2F1 | -2.90 |
| ATF3 | -3.30 | ATF3 | -3.45 | E2F6 | -3.15 | E2F4 | -3.20 |
| IRF3 | -3.40 | E2F4 | -3.49 | HMBOX1 | -3.16 | MYC | -3.35 |
| BCL6 | -3.45 | MITF | -3.53 | HBP1 | -3.51 | FOXM1 | -3.45 |
| MBD2 | -3.61 | FOXP1 | -3.71 | ESR1 | -4.18 | FOXA1 | -4.13 |
| CTCFL | -4.11 | FOXA1 | -3.78 | FOXA1 | -4.58 | ESR1 | -4.15 |
| ZEB2 | -4.19 | FOXM1 | -3.87 | BCL6 | -4.80 | ZEB2 | -4.51 |
| FOXP1 | -7.22 | ZBTB7A | -5.65 | FOXP1 | -6.98 | FOXP1 | -6.00 |

**Supplementary Table 7: Kinase and phosphatase activities for T47D TAMR and LTED clones.** The top 10 activated and repressed kinases and phosphatases are listed with their NES. Absolute activity scores ≥ 2 were considered significant. Kinase and phosphatase activities in the TAMR and LTED clones were compared to the activities in *orange* (and *green*) +E2_5 clones. Positive values indicate higher activities in TAMR and LTED clones while negative values indicate higher activities in *orange* (and *green*) +E2_5 clones. Analyses were performed by Dr. Efstathios-Iason Vlachavas. For simplicity and unambiguity, proteins are referred to by their nonitalicized gene names as previously recommended [402].

| ***turquoise* TAMR vs *orange* and *green* +E2_5** | | ***purple* TAMR vs *orange* and *green* +E2_5** | | ***orange* LTED vs *orange* and *green* +E2_5** | | ***orange* LTED vs *orange* +E2_5** | |
|---|---|---|---|---|---|---|---|
| **Kinase/ Phosphatase** | **NES** | **Kinase/ Phosphatase** | **NES** | **Kinase/ Phosphatase** | **NES** | **Kinase/ Phosphatase** | **NES** |
| HIPK2 | 3.22 | RPS6KA2 | 3.27 | CSNK2A2 | 4.30 | CSNK2A2 | 4.04 |
| SRC | 2.96 | RPS6KA3 | 3.18 | CSNK2A1 | 2.98 | CSNK2A1 | 2.67 |
| PPP3CB | 2.94 | HIPK2 | 2.99 | PTPN11 | 2.72 | CSNK1E | 2.64 |
| PPP3CC | 2.94 | PTK2 | 2.97 | CSNK1E | 2.55 | PTPN11 | 2.36 |
| CSNK1E | 2.82 | CSNK1E | 2.66 | VRK1 | 2.08 | PTPRG | 1.96 |
| CSNK2B | 2.17 | AKT2 | 2.41 | RPS6KA3 | 2.01 | VRK1 | 1.94 |
| CSNK2A2 | 2.10 | PRKD1 | 2.33 | PRKD1 | 1.98 | PRKD1 | 1.77 |
| PDK1 | 1.93 | PPP2CA | 2.31 | STK40 | 1.92 | PTPN1 | 1.76 |
| PPP3CA | 1.81 | RPS6KA5 | 2.30 | PTPRG | 1.89 | STK40 | 1.63 |
| CLK3 | 1.63 | PAK2 | 2.28 | CSNK2B | 1.82 | CLK1 | 1.49 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| AURKA | -2.68 | PRKCZ | -2.01 | CDK20 | -2.90 | CDK18 | -2.97 |
| PTPRG | -2.70 | MELK | -2.10 | SRC | -2.97 | STK16 | -3.04 |
| RPS6KA1 | -2.73 | WEE1 | -2.25 | MAPK1 | -2.97 | MAPK3 | -3.11 |
| LATS2 | -2.79 | CDC7 | -2.46 | STK16 | -3.03 | AURKB | -3.15 |
| MAPK3 | -3.27 | PTPRG | -2.59 | MAPK3 | -3.14 | SRC | -3.28 |
| CDK2 | -3.30 | AURKB | -2.97 | CDK16 | -3.65 | CDK5 | -3.51 |
| CDK5 | -3.77 | AURKA | -3.73 | AURKB | -4.39 | CDK16 | -3.56 |
| PRKACG | -3.96 | CDK4 | -4.30 | CDK4 | -5.75 | CDK4 | -4.26 |
| CDK1 | -4.41 | CDK2 | -4.58 | CDK1 | -6.89 | CDK2 | -7.27 |
| PRKACA | -4.63 | CDK1 | -4.58 | CDK2 | -7.34 | CDK1 | -7.38 |

**Supplementary Table 8: Kinase and phosphatase activities for MCF7 TAMR and LTED clones.** The top 10 activated and repressed kinases and phosphatases are listed with their NES. Absolute activity scores ≥ 2 were considered significant. Kinase and phosphatase activities in the TAMR and LTED clones were compared to the activities in +E2_1, +E2_3 and +E2_5 control cell lines. Positive values indicate higher activities in TAMR and LTED clones while negative values indicate higher activities in +E2 control cell lines. Analyses were performed by Dr. Efstathios-Iason Vlachavas. For simplicity and unambiguity, proteins are referred to by their nonitalicized gene names as previously recommended [402].

| ***blue* TAMR_1 vs +E2 pools** | | ***yellow* LTED_2 vs +E2 pools** | | ***pink* LTED_2 vs +E2 pools** | | ***red* LTED_5 vs +E2 pools** | |
|---|---|---|---|---|---|---|---|
| **Kinase/ Phosphatase** | **NES** | **Kinase/ Phosphatase** | **NES** | **Kinase/ Phosphatase** | **NES** | **Kinase/ Phosphatase** | **NES** |
| PRKCA | 7.71 | PRKACA | 4.88 | AURKB | 3.65 | PRKCA | 4.80 |
| PRKCB | 5.63 | PRKCB | 4.75 | PRKCB | 3.09 | PRKACA | 3.56 |
| PRKACA | 5.51 | PRKCA | 4.74 | CAMK2D | 3.06 | PRKCB | 3.55 |
| PRKCE | 5.00 | MAP2K2 | 3.78 | MAP2K2 | 2.94 | AURKB | 3.43 |
| PRKCG | 4.33 | PRKCE | 3.72 | DYRK2 | 2.86 | PRKACG | 3.34 |
| CAMK2D | 4.18 | CAMK2D | 3.59 | PRKCD | 2.76 | CAMK2D | 3.22 |
| PRKACG | 4.02 | PDK1 | 3.54 | CDK20 | 2.40 | PRKCE | 3.07 |
| AURKB | 3.81 | PRKACG | 3.23 | NUAK1 | 2.25 | PRKCD | 3.07 |
| PRKCD | 3.51 | AKT2 | 3.14 | ATM | 2.23 | PRKACB | 2.92 |
| AKT1 | 3.27 | AKT3 | 3.08 | MAP3K8 | 2.14 | MAP2K2 | 2.85 |
| PPP2CA | -2.51 | CDK5 | -3.70 | TTK | -2.21 | CDK5 | -2.93 |
| MTOR | -2.77 | CDK7 | -4.47 | DUSP1 | -2.35 | DYRK1A | -3.15 |
| MAPK11 | -2.79 | CDK4 | -4.59 | EPHA2 | -2.48 | CDK4 | -3.30 |
| PPP1CB | -2.82 | CSNK2A1 | -4.61 | RPS6KA2 | -2.56 | MTOR | -3.30 |
| PPP1CC | -2.82 | MAPK10 | -4.83 | RIPK2 | -2.79 | MAPK10 | -3.60 |
| MAPK8 | -2.84 | MAPK8 | -4.96 | CDK1 | -3.10 | MAPK8 | -4.46 |
| CDK1 | -2.97 | MAPK14 | -7.34 | ULK3 | -3.98 | CDK1 | -5.92 |
| MAPK10 | -3.06 | GSK3B | -8.05 | GSK3B | -4.07 | MAPK14 | -5.93 |
| CDK5 | -3.16 | CDK1 | -8.96 | MAPK10 | -4.30 | CDK2 | -6.88 |
| GSK3B | -3.75 | CDK2 | -10.21 | MTOR | -4.99 | GSK3B | -7.04 |

# 9 Abbreviations

| | |
|---|---|
| **4-OHT** | 4-hydroxytamoxifen |
| **ADC** | Antibody-drug conjugates |
| **AI** | Aromatase inhibitor |
| **AKT** | RAC-alpha serine/threonine-protein kinase |
| **ALK** | Anaplastic lymphoma kinase |
| **AmpR** | Ampicillin resistance gene, encodes β-lactamase |
| **APC** | Adenomatous polyposis coli |
| **APOBEC** | Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like |
| **ATP** | Adenosine triphosphate |
| **AURKA** | Aurora kinase A |
| **AURKB** | Aurora kinase B |
| **BCA** | Bicinchoninic acid |
| **BRCA** | Breast cancer |
| **BRCA1** | Breast cancer type 1 susceptibility protein |
| **BRCA2** | Breast cancer type 2 susceptibility protein |
| **BSA** | Bovine serum albumin |
| **$Ca^{2+}$** | Calcium |
| **C/EBPγ** | CCAAT enhancer binding protein gamma |
| **CAF** | Cancer associated fibroblast |
| **CaMK2δ** | Calcium/Calmodulin dependent protein kinase II delta |
| **CDK** | Cycling dependent kinase |
| **cDNA** | Complementary deoxyribonucleic acid |
| **CIN** | Chromosomal instability |
| **CML** | Chronic myeloid leukemia |
| **CNA** | Copy number alteration |
| **CNV** | Copy number variation |
| **$CO_2$** | Carbon dioxide |
| **CPTAC** | Clinical Proteomic Tumor Analysis Consortium |
| **CSFBS** | Charcoal stripped fetal bovine serum |
| **CSNK2A1** | Casein kinase 2 alpha 1 |
| **CSNK2A2** | Casein kinase 1 alpha 2 |
| **CSNK1E** | Casein kinase 1 epsilon |
| **CTC** | Circulating tumor cell |
| **CTRL** | Control |
| **$CuSO_4$** | Copper(II) sulfate |
| **CK1ε** | Casein kinase 1 epsilon |
| **CYP19A1** | Cytochrome P450 family 19 subfamily A member 1 |
| **CYP2D6** | Cytochrome P450 family 2 subfamily D member 6 |
| **CYP3A4/5** | Cytochrome P450 family 3 subfamily A member 4/5 |
| **DAG** | Diacylglycerol |
| **DAPI** | 4′,6-diamidino-2-phenylindole |
| **DCIS** | Ductal carcinoma *in situ* |
| **DMBA** | 7,12-Dimethylbenz[a]anthracene |
| **DMEM** | Dulbecco´s modified eagle´s medium |
| **DMSO** | Dimethyl sulfoxide |
| **DNA** | Deoxyribonucleic acid |
| **E2** | Estrogen |
| **E2F4** | E2F transcription factor 4 |
| **EDTA** | Ethylenediaminetetraacetic acid |
| **EGFR** | Epidermal growth factor receptor |
| **eIF2α** | Eukaryotic translation initiation factor 2 subunit α |
| **eIF4B** | Eukaryotic translation initiation factor 4B |
| **ELF1** | E74 Like ETS transcription factor 1 |
| **ELF3** | E74 Like ETS transcription factor 3 |
| **ER-** | Estrogen receptor negative |
| **ER+** | Estrogen receptor positive |
| **ERAD** | Endoplasmic-reticulum-associated protein degradation |
| **ERK** | Extracellular signal-regulated kinase |
| **ERα** | Estrogen receptor alpha |
| **ESR1** | Estrogen receptor 1 |

| | |
|---|---|
| **EtBr** | Ethidium bromide |
| **EtOH** | Ethanol |
| **FACS** | Fluorescence-activated cell sorting |
| **FBS** | Fetal bovine serum |
| **FOXA1** | Forkhead box A1 |
| **FOXC1** | Forkhead box C1 |
| **FOXM1** | Forkhead box M1 |
| **FOXP1** | Forkhead box P1 |
| **GABPα** | GA binding protein transcription factor subunit alpha |
| **GATA3** | GATA-binding protein 3 |
| **GRB2** | Growth factor receptor-bound protein 2 |
| **GRHL2** | Grainyhead like transcription factor 2 |
| **GSK-3α** | Glycogen synthase kinase 3 alpha |
| **GSK-3ß** | Glycogen synthase kinase 3 beta |
| **H&E staining** | Hematoxylin and eosin staining |
| **HER2** | Human epidermal growth factor receptor 2 |
| **HIFα** | Hypoxia-inducible factor 1 alpha |
| **IDC** | Invasive ductal carcinoma |
| **IFN** | Interferon |
| **IRF9** | IFN-regulatory factor 9 |
| **IκB** | Nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor |
| **IKKα** | IκB kinase alpha |
| **IKKß** | IκB kinase beta |
| **IKKε** | IκB kinase epsilon |
| **ISGF3** | IFN-stimulated gene (ISG) factor 3 |
| **JAK** | Janus kinase |
| **JNK3** | c-Jun N-terminal kinase 3 |
| **KLF5** | KLF transcription factor 5 |
| **KLF6** | KLF transcription factor 6 |
| **KMT2A** | Histone-lysine N methyltransferase 2A complex |
| **LBD** | Ligand-binding domain |
| **LCM** | Laser capture microdissection |
| **LTED** | Long-term estrogen deprivation/deprived |
| **LTR** | Long-terminal repeat |
| **MAPK/ERK** | Mitogen-activated protein kinase |
| **MAP2K/MEK** | Mitogen-activated protein kinase kinase |
| **MAP3K8** | Mitogen-activated protein kinase kinase kinase 8 |
| **MEF2A** | Myocyte enhancer factor 2A |
| **MEF2C** | Myocyte enhancer factor 2C |
| **MEK2** | Mitogen-activated protein kinase kinase 2 |
| **MEIS1** | Meis homeobox 1 |
| **mRNA** | Messenger ribonucleic acid |
| **MSI** | Microsatellite instability |
| **mTOR** | Mammalian target of rapamycin kinase |
| **MYC** | MYC proto-oncogene, BHLH transcription factor |
| **NaF** | Sodium fluoride |
| **NCOA3** | Nuclear receptor coactivator 3 |
| **NES** | Normalized enrichment score |
| **NFκB** | Nuclear factor kappa-light-chain-enhancer of activated B cells |
| **NGS** | Next-generation sequencing |
| **NSAI** | Non-steroidal aromatase inhibitors |
| **NSCLC** | Non-small-cell lung carcinoma |
| **ORF** | Open reading frame |
| **Ori** | Origin of replication |
| **OS** | Overall survival |
| **p38α** | p38 mitogen-activated protein kinase alpha |
| **$p70^{S6K}$** | p70 S6 kinase |
| **pCR** | Pathological complete response |
| **PBS** | Phosphate buffered saline |
| **PDPK1** | Phosphoinositide-dependent kinase 1 |
| **PERK** | Protein kinase R (PKR)-like endoplasmic reticulum kinase |
| **PFA** | Paraformaldehyde |
| **PFS** | Progression-free survival |

| | |
|---|---|
| **PI** | Propidium iodide |
| **PI3K** | Phosphatidylinositol-4,5-bisphosphate 3-kinase |
| **PIK3CA** | Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha |
| **PIP2** | Phosphatidylinositol (4,5)-bisphosphate |
| **PIP3** | Phosphatidylinositol (3,4,5)-trisphosphate |
| **PKCα** | Protein kinase C alpha |
| **PKCß** | Protein kinase C beta |
| **PKCγ** | Protein kinase C gamma |
| **PKCδ** | Protein kinase C delta |
| **PKCε** | Protein kinase C epsilon |
| **PKCη** | Protein kinase C eta |
| **PKCθ** | Protein kinase C theta |
| **PKCι** | Protein kinase C iota |
| **PKCζ** | Protein kinase C zeta |
| **PLC** | Phospholipase C |
| **PR** | Progesterone receptor |
| **PRDM1** | PR/SET domain 1 |
| **PTEN** | Phosphatase and tensin homolog |
| **PuroR** | Puromycin resistance gene, encodes puromycin N-acetyltransferase |
| **PVDF** | Polyvinylidene fluoride |
| **RFP** | Red fluorescent protein |
| **RSK** | Ribosomal S6 kinase |
| **RT-qPCR** | Quantitative reverse transcriptase polymerase chain reaction |
| **RIPA** | Radioimmunoprecipitation |
| **RNA** | Ribonucleic acid |
| **RNAi** | Ribonucleic acid interference |
| **ROS1** | C-ros oncogene 1 |
| **RTK** | Receptor tyrosine kinase |
| **S6K1** | Ribosomal protein S6 kinase beta 1 |
| **scRNA-Seq** | Single-cell RNA sequencing |
| **SD** | Standard deviation |
| **SDS-PAGE** | Sodium dodecyl sulfate polyacrylamide gel electrophoresis |
| **SEM** | Standard error of the mean |
| **SERD** | Selective estrogen receptor degrader |
| **SERM** | Selective estrogen receptor modulator |
| **siRNA** | Small interfering ribonucleic acid |
| **SOX6** | SRY-box transcription factor 6 |
| **SRBP1** | Sterol regulatory element-binding transcription factor 1 |
| **STAT** | Signal transducer and activator of transcription |
| **SWI/SNF** | SWItch/Sucrose non-fermentable |
| **TAMR** | Tamoxifen resistant |
| **TBK1** | TANK-binding kinase 1 |
| **TBS-T** | Tris buffered saline with 0.1% Tween 20 |
| **TEAD4** | TEA Domain transcription factor 4 |
| **TF** | Transcription factor |
| **TGFß** | Transforming growth factor ß |
| **TKI** | Tyrosine kinase inhibitor |
| **TNBC** | Triple negative breast cancer |
| **TNFα** | Tumor necrosis factor α |
| **TP53** | Tumor protein 53 |
| **TPA** | 2-O-Tetra-decanoylphorbol 13-acetate |
| **Trail** | TNF-related apoptosis-inducing ligand |
| **VEGF** | Vascular endothelial growth factor |
| **WES** | Whole exome sequencing |
| **WGD** | Whole genome doubling |
| **WGS** | Whole genome sequencing |
| **Wnt** | Wingless-type MMTV integration site family |
| **WT** | Wild type |

# 10 References

[1] Sung, H., *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA: A Cancer Journal for Clinicians. 2021. 71(3): p. 209-249.

[2] Siegel, R.L., *et al.* Cancer Statistics, 2021. CA Cancer J Clin. 2021. 71(1): p. 7-33.

[3] Vargo-Gogola, T. and Rosen, J.M. Modelling breast cancer: one size does not fit all. Nature Reviews Cancer. 2007. 7(9): p. 659-672.

[4] Russnes, H.G., *et al.* Breast Cancer Molecular Stratification: From Intrinsic Subtypes to Integrative Clusters. The American Journal of Pathology. 2017. 187(10): p. 2152-2162.

[5] Scholzen, T. and Gerdes, J. The Ki-67 protein: From the known and the unknown. Journal of Cellular Physiology. 2000. 182(3): p. 311-322.

[6] Wong, E., *et al.* McMaster Pathophysiology Review on Breast Cancer. 2012. Accessed: 23.01.23. Available at: http://www.pathophys.org/wp-content/uploads/2012/12/breastcancer-copy.png.

[7] Thorpe, S.M. Estrogen and Progesterone Receptor Determinations in Breast Cancer—Technology, Biology and Clinical Significance. Acta Oncologica. 1988. 27(1): p. 1-19.

[8] Dunnwald, L.K., *et al.* Hormone receptor status, tumor characteristics, and prognosis: a prospective cohort of breast cancer patients. Breast Cancer Research. 2007. 9(1): p. R6.

[9] Kohler, B.A., *et al.* Annual Report to the Nation on the Status of Cancer, 1975-2011, Featuring Incidence of Breast Cancer Subtypes by Race/Ethnicity, Poverty, and State. JNCI: Journal of the National Cancer Institute. 2015. 107(6): p. djv048-djv048.

[10] The Cancer Genome Atlas, N., *et al.* Comprehensive molecular portraits of human breast tumours. Nature. 2012. 490: p. 61.

[11] Curtis, C., *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature. 2012. 486: p. 346.

[12] Pereira, B., *et al.* The somatic mutation profiles of 2,433 breast cancers refine their genomic and transcriptomic landscapes. Nature Communications. 2016. 7(1): p. 11479.

[13] Dai, X., *et al.* Breast cancer intrinsic subtype classification, clinical use and future trends. Am J Cancer Res. 2015. 5(10): p. 2929-43.

[14] Rivenbark, A.G., *et al.* Molecular and Cellular Heterogeneity in Breast Cancer: Challenges for Personalized Medicine. The American Journal of Pathology. 2013. 183(4): p. 1113-1124.

[15] Bertheau, P., *et al.* p53 in breast cancer subtypes and new insights into response to chemotherapy. The Breast. 2013. 22: p. S27-S29.

[16] Hennigs, A., *et al.* Prognosis of breast cancer molecular subtypes in routine clinical care: A large prospective cohort study. BMC cancer. 2016. 16(1): p. 734-734.

[17] Lawrenson, R., *et al.* The impact of different tumour subtypes on management and survival of New Zealand women with Stage I-III breast cancer. N Z Med J. 2018. 131(1475): p. 51-60.

[18] Goldhirsch, A., *et al.* Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. Annals of oncology : official journal of the European Society for Medical Oncology. 2013. 24(9): p. 2206-2223.

[19] Gnant, M., *et al.* St. Gallen 2011: Summary of the Consensus Discussion. Breast care (Basel, Switzerland). 2011. 6(2): p. 136-141.

[20] Coates, A.S., *et al.* Tailoring therapies--improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015. Annals of oncology : official journal of the European Society for Medical Oncology. 2015. 26(8): p. 1533-1546.

[21] Hicks, D.G. and Tubbs, R.R. Assessment of the HER2 status in breast cancer by fluorescence in situ hybridization: a technical review with interpretive guidelines. Human Pathology. 2005. 36(3): p. 250-261.

[22] Hudziak, R.M., *et al.* p185HER2 monoclonal antibody has antiproliferative effects in vitro and sensitizes human breast tumor cells to tumor necrosis factor. Molecular and Cellular Biology. 1989. 9(3): p. 1165-1172.

[23] Romond, E.H., *et al.* Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. N Engl J Med. 2005. 353(16): p. 1673-84.

[24] Piccart-Gebhart, M.J., *et al.* Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. N Engl J Med. 2005. 353(16): p. 1659-72.

[25] Jhaveri, K. and Esteva, F.J. Pertuzumab in the Treatment of HER2+ Breast Cancer. Journal of the National Comprehensive Cancer Network. 2014. 12(4): p. 591-598.

[26] Swain, S.M., *et al.* Pertuzumab, trastuzumab, and docetaxel for HER2-positive metastatic breast cancer (CLEOPATRA): end-of-study results from a double-blind, randomised, placebo-controlled, phase 3 study. The Lancet Oncology. 2020. 21(4): p. 519-530.

[27] Verma, S., *et al.* Trastuzumab Emtansine for HER2-Positive Advanced Breast Cancer. New England Journal of Medicine. 2012. 367(19): p. 1783-1791.

[28] Krop, I.E., *et al.* Trastuzumab emtansine versus treatment of physician's choice for pretreated HER2-positive advanced breast cancer (TH3RESA): a randomised, open-label, phase 3 trial. The Lancet Oncology. 2014. 15(7): p. 689-699.

[29] Krop, I.E., *et al.* Trastuzumab emtansine versus treatment of physician's choice in patients with previously treated HER2-positive metastatic breast cancer (TH3RESA): final overall survival results from a randomised open-label phase 3 trial. The Lancet Oncology. 2017. 18(6): p. 743-754.

[30] Modi, S., *et al.* Trastuzumab Deruxtecan in Previously Treated HER2-Positive Breast Cancer. New England Journal of Medicine. 2019. 382(7): p. 610-621.

[31] Baselga, J., *et al.* Lapatinib with trastuzumab for HER2-positive early breast cancer (NeoALTTO): a randomised, open-label, multicentre, phase 3 trial. The Lancet. 2012. 379(9816): p. 633-640.

[32] Geyer, C.E., *et al.* Lapatinib plus Capecitabine for HER2-Positive Advanced Breast Cancer. New England Journal of Medicine. 2006. 355(26): p. 2733-2743.

[33] Chan, A., *et al.* Neratinib after trastuzumab-based adjuvant therapy in patients with HER2-positive breast cancer (ExteNET): a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial. The Lancet Oncology. 2016. 17(3): p. 367-377.

[34] Murthy, R.K., *et al.* Tucatinib, Trastuzumab, and Capecitabine for HER2-Positive Metastatic Breast Cancer. New England Journal of Medicine. 2019. 382(7): p. 597-609.

[35] SEER_Program. SEER 5-Year Relative Survival Rates, 2011-2017. 2022. Accessed: 02.03.22. Available at: https://seer.cancer.gov/statfacts/html/breast-subtypes.html.

[36] Bauer, K.R., *et al.* Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple-negative phenotype. Cancer. 2007. 109(9): p. 1721-1728.

[37] Rakha, E.A., *et al.* Prognostic markers in triple-negative breast cancer. Cancer. 2007. 109(1): p. 25-32.

[38] Lehmann, B.D., *et al.* Refinement of Triple-Negative Breast Cancer Molecular Subtypes: Implications for Neoadjuvant Chemotherapy Selection. PLOS ONE. 2016. 11(6): p. e0157368.

[39] Wahba, H.A. and El-Hadaad, H.A. Current approaches in treatment of triple-negative breast cancer. Cancer Biology & Medicine. 2015. 12(2): p. 106-116.

[40] Thomas, R., *et al.* Immune Checkpoint Inhibitors in Triple Negative Breast Cancer Treatment: Promising Future Prospects. Frontiers in Oncology. 2021. 10.

[41] Zeng, P., *et al.* Impact of TROP2 expression on prognosis in solid tumors: A Systematic Review and Meta-analysis. Scientific Reports. 2016. 6(1): p. 33658.

[42] Bardia, A., *et al.* Sacituzumab Govitecan in Metastatic Triple-Negative Breast Cancer. New England Journal of Medicine. 2021. 384(16): p. 1529-1541.

[43] Nilsson, S., *et al.* Mechanisms of Estrogen Action. Physiological Reviews. 2001. 81(4): p. 1535-1565.

[44] Fuentes, N. and Silveyra, P. Estrogen receptor signaling mechanisms. Adv Protein Chem Struct Biol. 2019. 116: p. 135-170.

[45] Davies, C., *et al.* Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial. The Lancet. 2013. 381(9869): p. 805-816.

[46] Gray, R.G., *et al.* aTTom: Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years in 6,953 women with early breast cancer. Journal of Clinical Oncology. 2013. 31(18_suppl): p. 5-5.

[47] Li, L., *et al.* Clinical outcomes comparison of 10 years versus 5 years of adjuvant endocrine therapy in patients with early breast cancer. BMC Cancer. 2018. 18(1): p. 977.

[48] Hanker, A.B., *et al.* Overcoming Endocrine Resistance in Breast Cancer. Cancer Cell. 2020. 37(4): p. 496-513.

[49] Simpson, E.R. and Davis, S.R. Minireview: Aromatase and the Regulation of Estrogen Biosynthesis—Some New Perspectives. Endocrinology. 2001. 142(11): p. 4589-4594.

[50] Smith, I.E. and Dowsett, M. Aromatase Inhibitors in Breast Cancer. New England Journal of Medicine. 2003. 348(24): p. 2431-2442.

[51] Fabian, C.J. The what, why and how of aromatase inhibitors: hormonal agents for treatment and prevention of breast cancer. International Journal of Clinical Practice. 2007. 61(12): p. 2051-2063.

[52] Hong, Y. and Chen, S. Aromatase inhibitors: structural features and biochemical characterization. Ann N Y Acad Sci. 2006. 1089: p. 237-51.

[53] Geisler, J., *et al.* Influence of anastrozole (Arimidex), a selective, non-steroidal aromatase inhibitor, on in vivo aromatisation and plasma oestrogen levels in postmenopausal women with breast cancer. British Journal of Cancer. 1996. 74(8): p. 1286-1291.

[54] Dowsett, M., *et al.* In vivo measurement of aromatase inhibition by letrozole (CGS 20267) in postmenopausal patients with breast cancer. Clinical Cancer Research. 1995. 1(12): p. 1511-1515.

[55] Geisler, J., *et al.* In vivo inhibition of aromatization by exemestane, a novel irreversible aromatase inhibitor, in postmenopausal breast cancer patients. Clinical Cancer Research. 1998. 4(9): p. 2089-2093.

[56] Love, R.R., *et al.* Effects of Tamoxifen on Bone Mineral Density in Postmenopausal Women with Breast Cancer. New England Journal of Medicine. 1992. 326(13): p. 852-856.

[57] Zidan, J., *et al.* Effects of tamoxifen on bone mineral density and metabolism in postmenopausal women with early-stage breast cancer. Medical Oncology. 2004. 21(2): p. 117-121.

[58] Jordan, V.C. Tamoxifen: a most unlikely pioneering medicine. Nature Reviews Drug Discovery. 2003. 2(3): p. 205-213.

[59] Martinkovich, S., *et al.* Selective estrogen receptor modulators: tissue specificity and clinical utility. Clin Interv Aging. 2014. 9: p. 1437-52.

[60] Kuhl, H. Pharmacology of estrogens and progestogens: influence of different routes of administration. Climacteric. 2005. 8(sup1): p. 3-63.

[61] Dehal, S.S. and Kupfer, D. CYP2D6 catalyzes tamoxifen 4-hydroxylation in human liver. Cancer Res. 1997. 57(16): p. 3402-6.

[62] Desta, Z., *et al.* Comprehensive Evaluation of Tamoxifen Sequential Biotransformation by the Human Cytochrome P450 System in Vitro: Prominent Roles for CYP3A and CYP2D6. Journal of Pharmacology and Experimental Therapeutics. 2004. 310(3): p. 1062-1075.

[63] Jacolot, F., *et al.* Identification of the cytochrome P450 IIIA family as the enzymes involved in the N-demethylation of tamoxifen in human liver microsomes. Biochem Pharmacol. 1991. 41(12): p. 1911-9.

[64] Shang, Y., *et al.* Cofactor Dynamics and Sufficiency in Estrogen Receptor–Regulated Transcription. Cell. 2000. 103(6): p. 843-852.

[65] Wu, L.*, et al.* Smad4 as a Transcription Corepressor for Estrogen Receptor α. Journal of Biological Chemistry. 2003. 278(17): p. 15192-15200.

[66] Wakeling, A.E. and Bowler, J. Steroidal pure antioestrogens. J Endocrinol. 1987. 112(3): p. R7-10.

[67] Wakeling, A.E.*, et al.* A Potent Specific Pure Antiestrogen with Clinical Potential. Cancer Research. 1991. 51(15): p. 3867-3873.

[68] Huynh, H.T. and Pollak, M. Insulin-like Growth Factor I Gene Expression in the Uterus Is Stimulated by Tamoxifen and Inhibited by the Pure Antiestrogen ICI 182780. Cancer Research. 1993. 53(23): p. 5585-5588.

[69] Osborne, C.K.*, et al.* Comparison of the Effects of a Pure Steroidal antiestrogen With Those of Tamoxifen in a Model of Human Breast Cancer. JNCI: Journal of the National Cancer Institute. 1995. 87(10): p. 746-750.

[70] Blin, C.*, et al.* Contrasting effects of tamoxifen and ICI 182 780 on estrogen-induced calbindin-D 9k gene expression in the uterus and in primary culture of myometrial cells. The Journal of Steroid Biochemistry and Molecular Biology. 1995. 55(1): p. 1-7.

[71] McClelland, R.A.*, et al.* Effects of short-term antiestrogen treatment of primary breast cancer on estrogen receptor mRNA and protein expression and on estrogen-regulated genes. Breast Cancer Res Treat. 1996. 41(1): p. 31-41.

[72] Hyder, S.M.*, et al.* Selective inhibition of estrogen-regulated gene expression in vivo by the pure antiestrogen ICI 182,780. Cancer Research. 1997. 57(13): p. 2547-2549.

[73] Wardell, S.E.*, et al.* The turnover of estrogen receptor α by the selective estrogen receptor degrader (SERD) fulvestrant is a saturable process that is not required for antagonist efficacy. Biochemical Pharmacology. 2011. 82(2): p. 122-130.

[74] Pike, A.C.W.*, et al.* Structural Insights into the Mode of Action of a Pure Antiestrogen. Structure. 2001. 9(2): p. 145-153.

[75] Guan, J.*, et al.* Therapeutic Ligands Antagonize Estrogen Receptor Function by Impairing Its Mobility. Cell. 2019. 178(4): p. 949-963.e18.

[76] Wijayaratne, A.L. and McDonnell, D.P. The Human Estrogen Receptor-α Is a Ubiquitinated Protein Whose Stability Is Affected Differentially by Agonists, Antagonists, and Selective Estrogen Receptor Modulators. Journal of Biological Chemistry. 2001. 276(38): p. 35684-35692.

[77] Wittmann, B.M.*, et al.* Definition of Functionally Important Mechanistic Differences among Selective Estrogen Receptor Down-regulators. Cancer Research. 2007. 67(19): p. 9549-9560.

[78] Calligé, M.*, et al.* CSN5/Jab1 is involved in ligand-dependent degradation of estrogen receptor {alpha} by the proteasome. Molecular and cellular biology. 2005. 25(11): p. 4349-4358.

[79] Ingle, J.N.*, et al.* Fulvestrant in women with advanced breast cancer after progression on prior aromatase inhibitor therapy: North Central Cancer Treatment Group Trial N0032. J Clin Oncol. 2006. 24(7): p. 1052-6.

[80] Wang, J.*, et al.* Fulvestrant in advanced breast cancer following tamoxifen and aromatase inhibition: a single center experience. Breast J. 2009. 15(3): p. 247-53.

[81] Osborne, C.K.*, et al.* Double-Blind, Randomized Trial Comparing the Efficacy and Tolerability of Fulvestrant Versus Anastrozole in Postmenopausal Women With Advanced Breast Cancer Progressing on Prior Endocrine Therapy: Results of a North American Trial. Journal of Clinical Oncology. 2002. 20(16): p. 3386-3395.

[82] Robertson, J.F.*, et al.* Fulvestrant versus anastrozole for the treatment of advanced breast carcinoma in postmenopausal women: a prospective combined analysis of two multicenter trials. Cancer. 2003. 98(2): p. 229-38.

[83] Hernando, C.*, et al.* Oral Selective Estrogen Receptor Degraders (SERDs) as a Novel Breast Cancer Therapy: Present and Future from a Clinical Perspective. International Journal of Molecular Sciences. 2021. 22(15): p. 7812.

[84] Pan, H.*, et al.* 20-Year Risks of Breast-Cancer Recurrence after Stopping Endocrine Therapy at 5 Years. The New England journal of medicine. 2017. 377(19): p. 1836-1846.

[85] Clarke, R.*, et al.* Cellular and molecular pharmacology of antiestrogen action and resistance. Pharmacol Rev. 2001. 53(1): p. 25-71.

[86] Clarke, R.*, et al.* Antiestrogen resistance in breast cancer and the role of estrogen receptor signaling. Oncogene. 2002. 22: p. 7316.

[87] Ring, A. and Dowsett, M. Mechanisms of tamoxifen resistance. 2004. 11(4): p. 643.

[88] Riggins, R.B.*, et al.* Pathways to tamoxifen resistance. Cancer Letters. 2007. 256(1): p. 1-24.

[89] Wiechmann, L.*, et al.* Crosstalk between the Estrogen Receptor and the HER Tyrosine Kinase Receptor Family: Molecular Mechanism and Clinical Implications for Endocrine Therapy Resistance. Endocrine Reviews. 2008. 29(2): p. 217-233.

[90] Musgrove, E.A. and Sutherland, R.L. Biological determinants of endocrine resistance in breast cancer. Nature Reviews Cancer. 2009. 9: p. 631.

[91] Tryfonidis, K.*, et al.* Endocrine treatment in breast cancer: Cure, resistance and beyond. Cancer Treatment Reviews. 2016. 50: p. 68-81.

[92] Dixon, J.M. Endocrine Resistance in Breast Cancer. New Journal of Science. 2014. 2014: p. 27.

[93] Katzenellenbogen, J.A.*, et al.* Structural underpinnings of oestrogen receptor mutations in endocrine therapy resistance. Nature Reviews Cancer. 2018. 18(6): p. 377-388.

[94] Asghar, U.*, et al.* The history and future of targeting cyclin-dependent kinases in cancer therapy. Nature Reviews Drug Discovery. 2015. 14(2): p. 130-146.

[95] Merenbakh-Lamin, K.*, et al.* D538G Mutation in Estrogen Receptor-α: A Novel Mechanism for Acquired Endocrine Resistance in Breast Cancer. Cancer Research. 2013. 73(23): p. 6856-6864.

[96] Jeselsohn, R.*, et al.* Emergence of Constitutively Active Estrogen Receptor-α Mutations in Pretreated Advanced Estrogen Receptor–Positive Breast Cancer. Clinical Cancer Research. 2014. 20(7): p. 1757-1767.

[97] Toy, W.*, et al.* Activating ESR1 Mutations Differentially Affect the Efficacy of ER Antagonists. Cancer Discovery. 2017. 7(3): p. 277.

[98] Takeshita, T.*, et al.* Droplet digital polymerase chain reaction assay for screening of

ESR1 mutations in 325 breast cancer specimens. Translational Research. 2015. 166(6): p. 540-553.e2.

[99] Chandarlapaty, S., *et al.* Prevalence of ESR1 Mutations in Cell-Free DNA and Outcomes in Metastatic Breast Cancer: A Secondary Analysis of the BOLERO-2 Clinical Trial. JAMA Oncology. 2016. 2(10): p. 1310-1315.

[100] Wang, P., *et al.* Sensitive Detection of Mono- and Polyclonal ESR1 Mutations in Primary Tumors, Metastatic Lesions, and Cell-Free DNA of Breast Cancer Patients. Clin Cancer Res. 2016. 22(5): p. 1130-7.

[101] Najim, O., *et al.* The association between type of endocrine therapy and development of estrogen receptor-1 mutation(s) in patients with hormone-sensitive advanced breast cancer: A systematic review and meta-analysis of randomized and non-randomized trials. Biochimica et Biophysica Acta (BBA) - Reviews on Cancer. 2019. 1872(2): p. 188315.

[102] Jeselsohn, R., *et al.* ESR1 mutations—a mechanism for acquired endocrine resistance in breast cancer. Nature Reviews Clinical Oncology. 2015. 12(10): p. 573-583.

[103] Brett, J.O., *et al.* ESR1 mutation as an emerging clinical biomarker in metastatic hormone receptor-positive breast cancer. Breast Cancer Research. 2021. 23(1): p. 85.

[104] Razavi, P., *et al.* The Genomic Landscape of Endocrine-Resistant Advanced Breast Cancers. Cancer Cell. 2018. 34(3): p. 427-438.e6.

[105] Borgoni, S., *et al.* Time-Resolved Profiling Reveals ATF3 as a Novel Mediator of Endocrine Resistance in Breast Cancer. Cancers. 2020. 12(10): p. 2918.

[106] Hanahan, D. and Weinberg, R.A. The Hallmarks of Cancer. Cell. 2000. 100(1): p. 57-70.

[107] Hanahan, D. and Weinberg, Robert A. Hallmarks of Cancer: The Next Generation. Cell. 2011. 144(5): p. 646-674.

[108] Hanahan, D. Hallmarks of Cancer: New Dimensions. Cancer Discovery. 2022. 12(1): p. 31-46.

[109] Kenny, F.S., *et al.* Overexpression of Cyclin D1 Messenger RNA Predicts for Poor Prognosis in Estrogen Receptor-positive Breast Cancer. Clinical Cancer Research. 1999. 5(8): p. 2069-2076.

[110] Elsheikh, S., *et al.* CCND1 amplification and cyclin D1 expression in breast cancer and their relation with proteomic subgroups and patient outcome. Breast Cancer Research and Treatment. 2008. 109(2): p. 325-335.

[111] Lefebvre, C., *et al.* Mutational Profile of Metastatic Breast Cancers: A Retrospective Analysis. PLOS Medicine. 2016. 13(12): p. e1002201.

[112] Ertel, A., *et al.* RB-pathway disruption in breast cancer. Cell Cycle. 2010. 9(20): p. 4153-4163.

[113] Alves, C.L., *et al.* High CDK6 Protects Cells from Fulvestrant-Mediated Apoptosis and is a Predictor of Resistance to Fulvestrant in Estrogen Receptor–Positive Metastatic Breast Cancer. Clinical Cancer Research. 2016. 22(22): p. 5514-5526.

[114] Kaminska, K., *et al.* Distinct mechanisms of resistance to fulvestrant treatment dictate level of ER independence and selective response to CDK inhibitors in metastatic breast cancer. Breast Cancer Research. 2021. 23(1): p. 26.

[115] Miller, T.W., *et al.* ERα-Dependent E2F Transcription Can Mediate Resistance to Estrogen Deprivation in Human Breast Cancer. Cancer Discovery. 2011. 1(4): p. 338-351.

[116] Hortobagyi, G.N., *et al.* Ribociclib as First-Line Therapy for HR-Positive, Advanced Breast Cancer. New England Journal of Medicine. 2016. 375(18): p. 1738-1748.

[117] Slamon, D.J., *et al.* Phase III Randomized Study of Ribociclib and Fulvestrant in Hormone Receptor–Positive, Human Epidermal Growth Factor Receptor 2–Negative Advanced Breast Cancer: MONALEESA-3. Journal of Clinical Oncology. 2018. 36(24): p. 2465-2472.

[118] Slamon, D.J., *et al.* Overall Survival with Ribociclib plus Fulvestrant in Advanced Breast Cancer. New England Journal of Medicine. 2019. 382(6): p. 514-524.

[119] Tripathy, D., *et al.* Ribociclib plus endocrine therapy for premenopausal women with hormone-receptor-positive, advanced breast cancer (MONALEESA-7): a randomised phase 3 trial. The Lancet Oncology. 2018. 19(7): p. 904-915.

[120] Im, S.-A., *et al.* Overall Survival with Ribociclib plus Endocrine Therapy in Breast Cancer. New England Journal of Medicine. 2019. 381(4): p. 307-316.

[121] George W. Sledge, J., *et al.* MONARCH 2: Abemaciclib in Combination With Fulvestrant in Women With HR+/HER2− Advanced Breast Cancer Who Had Progressed While Receiving Endocrine Therapy. Journal of Clinical Oncology. 2017. 35(25): p. 2875-2884.

[122] Goetz, M.P., *et al.* MONARCH 3: Abemaciclib As Initial Therapy for Advanced Breast Cancer. Journal of Clinical Oncology. 2017. 35(32): p. 3638-3646.

[123] Sledge, G.W., Jr, *et al.* The Effect of Abemaciclib Plus Fulvestrant on Overall Survival in Hormone Receptor–Positive, ERBB2-Negative Breast Cancer That Progressed on Endocrine Therapy—MONARCH 2: A Randomized Clinical Trial. JAMA Oncology. 2020. 6(1): p. 116-124.

[124] Zhang, Q.Y., *et al.* MONARCH plus: abemaciclib plus endocrine therapy in women with HR+/HER2- advanced breast cancer: the multinational randomized phase III study. Therapeutic advances in medical oncology. 2020. 12: p. 1758835920963925-1758835920963925.

[125] Finn, R.S., *et al.* The cyclin-dependent kinase 4/6 inhibitor palbociclib in combination with letrozole versus letrozole alone as first-line treatment of oestrogen receptor-positive, HER2-negative, advanced breast cancer (PALOMA-1/TRIO-18): a randomised phase 2 study. The Lancet Oncology. 2015. 16(1): p. 25-35.

[126] Finn, R.S., *et al.* Overall survival results from the randomized phase 2 study of palbociclib in combination with letrozole versus letrozole alone for first-line treatment of ER+/HER2− advanced breast cancer (PALOMA-1, TRIO-18). Breast Cancer Research and Treatment. 2020. 183(2): p. 419-428.

[127] Finn, R.S., *et al.* Palbociclib and Letrozole in Advanced Breast Cancer. New England Journal of Medicine. 2016. 375(20): p. 1925-1936.

[128] Rugo, H.S., *et al.* Palbociclib plus letrozole as first-line therapy in estrogen receptor-positive/human epidermal growth factor receptor 2-negative advanced breast cancer with extended

follow-up. Breast Cancer Research and Treatment. 2019. 174(3): p. 719-729.

[129] Cristofanilli, M.*, et al.* Fulvestrant plus palbociclib versus fulvestrant plus placebo for treatment of hormone-receptor-positive, HER2-negative metastatic breast cancer that progressed on previous endocrine therapy (PALOMA-3): final analysis of the multicentre, double-blind, phase 3 randomised controlled trial. The Lancet Oncology. 2016. 17(4): p. 425-439.

[130] Turner, N.C.*, et al.* Overall Survival with Palbociclib and Fulvestrant in Advanced Breast Cancer. New England Journal of Medicine. 2018. 379(20): p. 1926-1936.

[131] Rugo, H.S.*, et al.* 234P Effect of palbociclib (PAL) + endocrine therapy (ET) on time to chemotherapy (TTC) across subgroups of patients (pts) with hormone receptor-positive/human epidermal growth factor receptor 2-negative (HR+/HER2-) advanced breast cancer (ABC): Post hoc analyses from PALOMA-2 (P2) and PALOMA-3 (P3). Annals of Oncology. 2021. 32: p. S461.

[132] Tian, Q.*, et al.* Overall survival and progression-free survival with cyclin-dependent kinase 4/6 inhibitors plus endocrine therapy in breast cancer: an updated meta-analysis of randomized controlled trials. Eur Rev Med Pharmacol Sci. 2021. 25(23): p. 7252-7267.

[133] Du, Z. and Lovly, C.M. Mechanisms of receptor tyrosine kinase activation in cancer. Molecular Cancer. 2018. 17(1): p. 58.

[134] Metibemu, D.S.*, et al.* Exploring receptor tyrosine kinases-inhibitors in Cancer treatments. Egyptian Journal of Medical Human Genetics. 2019. 20(1): p. 35.

[135] Elbauomy Elsheikh, S.*, et al.* FGFR1 amplification in breast carcinomas: a chromogenic in situhybridisation analysis. Breast Cancer Research. 2007. 9(2): p. R23.

[136] Turner, N.*, et al.* FGFR1 Amplification Drives Endocrine Therapy Resistance and Is a Therapeutic Target in Breast Cancer. Cancer Research. 2010. 70(5): p. 2085-2094.

[137] Giltnane, J.M.*, et al.* Genomic profiling of ER<sup>+</sup> breast cancers after short-term estrogen suppression reveals alterations associated with endocrine resistance. Science Translational Medicine. 2017. 9(402): p. eaai7993.

[138] Formisano, L.*, et al.* Aberrant FGFR signaling mediates resistance to CDK4/6 inhibitors in ER+ breast cancer. Nature Communications. 2019. 10(1): p. 1373.

[139] Levine, K.M.*, et al.* FGFR4 overexpression and hotspot mutations in metastatic ER+ breast cancer are enriched in the lobular subtype. npj Breast Cancer. 2019. 5(1): p. 19.

[140] Coombes, R.C.*, et al.* Results of the phase IIa RADICAL trial of the FGFR inhibitor AZD4547 in endocrine resistant breast cancer. Nature Communications. 2022. 13(1): p. 3246.

[141] Nicholson, R.I.*, et al.* Modulation of epidermal growth factor receptor in endocrine-resistant, estrogen-receptor-positive breast cancer. Ann N Y Acad Sci. 2002. 963: p. 104-15.

[142] Knowlden, J.M.*, et al.* Elevated Levels of Epidermal Growth Factor Receptor/c-erbB2 Heterodimers Mediate an Autocrine Growth Regulatory Pathway in Tamoxifen-Resistant MCF-7 Cells. Endocrinology. 2003. 144(3): p. 1032-1044.

[143] Gutierrez, M.C.*, et al.* Molecular changes in tamoxifen-resistant breast cancer: relationship between estrogen receptor, HER-2, and p38 mitogen-activated protein kinase. J Clin Oncol. 2005. 23(11): p. 2469-76.

[144] Croessmann, S.*, et al.* Combined Blockade of Activating ERBB2 Mutations and ER Results in Synthetic Lethality of ER+/HER2 Mutant Breast Cancer. Clinical Cancer Research. 2019. 25(1): p. 277-289.

[145] Nayar, U.*, et al.* Acquired HER2 mutations in ER+ metastatic breast cancer confer resistance to estrogen receptor–directed therapies. Nature Genetics. 2019. 51(2): p. 207-216.

[146] Ma, C.X.*, et al.* The phase II MutHER study of neratinib alone and in combination with fulvestrant in HER2 mutated, non-amplified metastatic breast cancer. Clinical Cancer Research. 2022: p. clincanres.CCR-21-3418-E.2021.

[147] Miller, T.W.*, et al.* Hyperactivation of phosphatidylinositol-3 kinase promotes escape from hormone dependence in estrogen receptor-positive human breast cancer. J Clin Invest. 2010. 120(7): p. 2406-13.

[148] Sanchez, C.G.*, et al.* Preclinical modeling of combined phosphatidylinositol-3-kinase inhibition with endocrine therapy for estrogen receptor-positive breast cancer. Breast Cancer Research. 2011. 13(2): p. R21.

[149] Huang, D.*, et al.* PIK3CA mutations contribute to fulvestrant resistance in ER-positive breast cancer. American journal of translational research. 2019. 11(9): p. 6055-6065.

[150] Faridi, J.*, et al.* Expression of Constitutively Active Akt-3 in MCF-7 Breast Cancer Cells Reverses the Estrogen and Tamoxifen Responsivity of these Cells in Vivo. Clinical Cancer Research. 2003. 9(8): p. 2933-2939.

[151] deGraffenried, L.A.*, et al.* Inhibition of mTOR Activity Restores Tamoxifen Response in Breast Cancer Cells with Aberrant Akt Activity. Clinical Cancer Research. 2004. 10(23): p. 8059-8067.

[152] Campbell, R.A.*, et al.* Phosphatidylinositol 3-Kinase/AKT-mediated Activation of Estrogen Receptor α: A NEW MODEL FOR ANTI-ESTROGEN RESISTANCE. Journal of Biological Chemistry. 2001. 276(13): p. 9817-9824.

[153] Fritsch, C.*, et al.* Characterization of the Novel and Specific PI3Kα Inhibitor NVP-BYL719 and Development of the Patient Stratification Strategy for Clinical Trials. Molecular Cancer Therapeutics. 2014. 13(5): p. 1117-1129.

[154] André, F.*, et al.* Alpelisib for PIK3CA-Mutated, Hormone Receptor–Positive Advanced Breast Cancer. New England Journal of Medicine. 2019. 380(20): p. 1929-1940.

[155] Rugo, H.S.*, et al.* Alpelisib plus fulvestrant in *PIK3CA*-mutated, hormone receptor-positive advanced breast cancer after a CDK4/6 inhibitor (BYLieve): one cohort of a phase 2, multicentre, open-label, non-comparative study. The Lancet Oncology. 2021. 22(4): p. 489-498.

[156] Baselga, J.*, et al.* Everolimus in Postmenopausal Hormone-Receptor–Positive Advanced Breast Cancer. New England Journal of Medicine. 2011. 366(6): p. 520-529.

[157] Yardley, D.A.*, et al.* Everolimus Plus Exemestane in Postmenopausal Patients with HR+ Breast Cancer: BOLERO-2 Final Progression-Free

Survival Analysis. Advances in Therapy. 2013. 30(10): p. 870-884.

[158] Bachelot, T., *et al.* Randomized Phase II Trial of Everolimus in Combination With Tamoxifen in Patients With Hormone Receptor–Positive, Human Epidermal Growth Factor Receptor 2–Negative Metastatic Breast Cancer With Prior Exposure to Aromatase Inhibitors: A GINECO Study. Journal of Clinical Oncology. 2012. 30(22): p. 2718-2724.

[159] Schmid, P., *et al.* Fulvestrant Plus Vistusertib vs Fulvestrant Plus Everolimus vs Fulvestrant Alone for Women With Hormone Receptor–Positive Metastatic Breast Cancer: The MANTA Phase 2 Randomized Clinical Trial. JAMA Oncology. 2019. 5(11): p. 1556-1564.

[160] Dhillon, A.S., *et al.* MAP kinase signalling pathways in cancer. Oncogene. 2007. 26: p. 3279.

[161] Bertucci, F., *et al.* Genomic characterization of metastatic breast cancers. Nature. 2019. 569(7757): p. 560-564.

[162] Griffith, O.L., *et al.* The prognostic effects of somatic mutations in ER-positive breast cancer. Nature Communications. 2018. 9(1): p. 3476.

[163] Pearson, A., *et al.* Inactivating NF1 Mutations Are Enriched in Advanced Breast Cancer and Contribute to Endocrine Therapy Resistance. Clinical Cancer Research. 2020. 26(3): p. 608-622.

[164] Fribbens, C., *et al.* Tracking evolution of aromatase inhibitor resistance with circulating tumour DNA analysis in metastatic breast cancer. Annals of Oncology. 2018. 29(1): p. 145-153.

[165] Kato, S., *et al.* Activation of the Estrogen Receptor Through Phosphorylation by Mitogen-Activated Protein Kinase. Science. 1995. 270(5241): p. 1491-1494.

[166] Thomas, R.S., *et al.* Phosphorylation at serines 104 and 106 by Erk1/2 MAPK is important for estrogen receptor-α activity. Journal of Molecular Endocrinology. 2008. 40(4): p. 173.

[167] Bunone, G., *et al.* Activation of the unliganded estrogen receptor by EGF involves the MAP kinase pathway and direct phosphorylation. The EMBO Journal. 1996. 15(9): p. 2174-2183.

[168] Flaherty, K.T., *et al.* Improved Survival with MEK Inhibition in BRAF-Mutated Melanoma. New England Journal of Medicine. 2012. 367(2): p. 107-114.

[169] Sullivan, R.J., *et al.* First-in-Class ERK1/2 Inhibitor Ulixertinib (BVD-523) in Patients with MAPK Mutant Advanced Solid Tumors: Results of a Phase I Dose-Escalation and Expansion Study. Cancer Discovery. 2018. 8(2): p. 184-195.

[170] Lian, T., *et al.* Trametinib in the treatment of multiple malignancies harboring MEK1 mutations. Cancer Treatment Reviews. 2019. 81: p. 101907.

[171] McGranahan, N. and Swanton, C. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. Cell. 2017. 168(4): p. 613-628.

[172] Dagogo-Jack, I. and Shaw, A.T. Tumour heterogeneity and resistance to cancer therapies. Nature Reviews Clinical Oncology. 2018. 15(2): p. 81-94.

[173] Marusyk, A., *et al.* Intratumor Heterogeneity: The Rosetta Stone of Therapy Resistance. Cancer Cell. 2020. 37(4): p. 471-484.

[174] Burrell, R.A., *et al.* The causes and consequences of genetic heterogeneity in cancer evolution. Nature. 2013. 501(7467): p. 338-345.

[175] Koren, S. and Bentires-Alj, M. Breast Tumor Heterogeneity: Source of Fitness, Hurdle for Therapy. Molecular Cell. 2015. 60(4): p. 537-546.

[176] Nowell, P. The clonal evolution of tumor cell populations. Science. 1976. 194(4260): p. 23-28.

[177] Malhotra, G., *et al.* Shared signaling pathways in normal and breast cancer stem cells. Journal of Carcinogenesis. 2011. 10(1): p. 38-38.

[178] Lessard, J. and Sauvageau, G. Bmi-1 determines the proliferative capacity of normal and leukaemic stem cells. Nature. 2003. 423(6937): p. 255-60.

[179] Shackleton, M., *et al.* Heterogeneity in Cancer: Cancer Stem Cells versus Clonal Evolution. Cell. 2009. 138(5): p. 822-829.

[180] de Bruin, E.C., *et al.* Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. Science. 2014. 346(6206): p. 251-6.

[181] Jamal-Hanjani, M., *et al.* Tracking the Evolution of Non–Small-Cell Lung Cancer. New England Journal of Medicine. 2017. 376(22): p. 2109-2121.

[182] Reya, T., *et al.* Stem cells, cancer, and cancer stem cells. Nature. 2001. 414(6859): p. 105-111.

[183] Wang, J., *et al.* Cancer stem cells in glioma: challenges and opportunities. Translational cancer research. 2013. 2(5): p. 429-441. This is the reference Simone used for cancer stem cells.

[184] Batlle, E. and Clevers, H. Cancer stem cells revisited. Nature Medicine. 2017. 23(10): p. 1124-1134.

[185] Al-Hajj, M., *et al.* Prospective identification of tumorigenic breast cancer cells. Proceedings of the National Academy of Sciences. 2003. 100(7): p. 3983-3988.

[186] Mani, S.A., *et al.* The Epithelial-Mesenchymal Transition Generates Cells with Properties of Stem Cells. Cell. 2008. 133(4): p. 704-715.

[187] Morel, A.P., *et al.* Generation of breast cancer stem cells through epithelial-mesenchymal transition. PLoS One. 2008. 3(8): p. e2888.

[188] Chaffer, C.L., *et al.* Normal and neoplastic nonstem cells can spontaneously convert to a stem-like state. Proceedings of the National Academy of Sciences. 2011. 108(19): p. 7950-7955.

[189] Gupta, Piyush B., *et al.* Stochastic State Transitions Give Rise to Phenotypic Equilibrium in Populations of Cancer Cells. Cell. 2011. 146(4): p. 633-644.

[190] Wang, Y.-H. and Scadden, D.T. Harnessing the apoptotic programs in cancer stem-like cells. EMBO reports. 2015. 16(9): p. 1084-1098.

[191] Hata, A.N., *et al.* Tumor cells can follow distinct evolutionary paths to become resistant to epidermal growth factor receptor inhibition. Nature Medicine. 2016. 22(3): p. 262-269.

[192] Shaw, A.T., *et al.* Resensitization to Crizotinib by the Lorlatinib ALK Resistance Mutation L1198F. New England Journal of Medicine. 2015. 374(1): p. 54-61.

[193] Choi, Y.L., *et al.* EML4-ALK Mutations in Lung Cancer That Confer Resistance to ALK

Inhibitors. New England Journal of Medicine. 2010. 363(18): p. 1734-1739.

[194] Li, J., *et al.* BRAF V600E Mediates Crizotinib Resistance and Responds to Dabrafenib and Trametinib in a ROS1-Rearranged Non-Small Cell Lung Cancer: A Case Report. The Oncologist. 2021. 26(12): p. e2115-e2119.

[195] Bhang, H.-e.C., *et al.* Studying clonal dynamics in response to cancer therapy using high-complexity barcoding. Nature Medicine. 2015. 21: p. 440.

[196] Dhanyamraju, P.K., *et al.* Drug-Tolerant Persister Cells in Cancer Therapy Resistance. Cancer Res. 2022. 82(14): p. 2503-2514.

[197] Raha, D., *et al.* The Cancer Stem Cell Marker Aldehyde Dehydrogenase Is Required to Maintain a Drug-Tolerant Tumor Cell Subpopulation. Cancer Research. 2014. 74(13): p. 3579-3590.

[198] Carcereri de Prati, A., *et al.* Metastatic Breast Cancer Cells Enter Into Dormant State and Express Cancer Stem Cells Phenotype Under Chronic Hypoxia. Journal of Cellular Biochemistry. 2017. 118(10): p. 3237-3248.

[199] Alowaidi, F., *et al.* Assessing stemness and proliferation properties of the newly established colon cancer ‘stem’ cell line, CSC480 and novel approaches to identify dormant cancer cells. Oncol Rep. 2018. 39(6): p. 2881-2891.

[200] Zhou, N., *et al.* Stem cell characteristics of dormant cells and cisplatin-induced effects on the stemness of epithelial ovarian cancer cells. Mol Med Rep. 2014. 10(5): p. 2495-2504.

[201] Murakami, A., *et al.* Hypoxia Increases Gefitinib-Resistant Lung Cancer Stem Cells through the Activation of Insulin-Like Growth Factor 1 Receptor. PLOS ONE. 2014. 9(1): p. e86459.

[202] Sharma, S.V., *et al.* A Chromatin-Mediated Reversible Drug-Tolerant State in Cancer Cell Subpopulations. Cell. 2010. 141(1): p. 69-80.

[203] Ramirez, M., *et al.* Diverse drug-resistance mechanisms can emerge from drug-tolerant cancer persister cells. Nature Communications. 2016. 7(1): p. 10690.

[204] Hangauer, M.J., *et al.* Drug-tolerant persister cancer cells are vulnerable to GPX4 inhibition. Nature. 2017. 551(7679): p. 247-250.

[205] Ravindran Menon, D., *et al.* A stress-induced early innate response causes multidrug tolerance in melanoma. Oncogene. 2015. 34(34): p. 4448-4459.

[206] Vinogradova, M., *et al.* An inhibitor of KDM5 demethylases reduces survival of drug-tolerant cancer cells. Nature Chemical Biology. 2016. 12(7): p. 531-538.

[207] Navin, N., *et al.* Inferring tumor progression from genomic heterogeneity. Genome Res. 2010. 20(1): p. 68-80.

[208] Shah, S.P., *et al.* The clonal and mutational evolution spectrum of primary triple-negative breast cancers. Nature. 2012. 486(7403): p. 395-399.

[209] Yates, L.R., *et al.* Subclonal diversification of primary breast cancer revealed by multiregion sequencing. Nature medicine. 2015. 21(7): p. 751-759. Metasteses can derive from subclones in the primary tumor.

[210] Casasent, A.K., *et al.* Multiclonal Invasion in Breast Tumors Identified by Topographic Single Cell Sequencing. Cell. 2018. 172(1): p. 205-217.e12.

[211] Luca, F.D., *et al.* Mutational analysis of single circulating tumor cells by next generation sequencing in metastatic breast cancer. Oncotarget. 2016. 7(18).

[212] Møller, E., *et al.* Next-Generation Sequencing of Disseminated Tumor Cells. Frontiers in Oncology. 2013. 3(320).

[213] Shah, S.P., *et al.* Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. Nature. 2009. 461(7265): p. 809-813.

[214] Ding, L., *et al.* Genome remodelling in a basal-like breast cancer metastasis and xenograft. Nature. 2010. 464(7291): p. 999-1005.

[215] Maley, C.C., *et al.* Genetic clonal diversity predicts progression to esophageal adenocarcinoma. Nature Genetics. 2006. 38(4): p. 468-473.

[216] Dillekås, H., *et al.* Are 90% of deaths from cancer caused by metastases? Cancer medicine. 2019. 8(12): p. 5574-5576.

[217] Brown, D., *et al.* Phylogenetic analysis of metastatic progression in breast cancer using somatic mutations and copy number aberrations. Nature Communications. 2017. 8(1): p. 14944.

[218] Hoadley, K.A., *et al.* Tumor Evolution in Two Patients with Basal-like Breast Cancer: A Retrospective Genomics Study of Multiple Metastases. PLOS Medicine. 2016. 13(12): p. e1002174.

[219] Savas, P., *et al.* The Subclonal Architecture of Metastatic Breast Cancer: Results from a Prospective Community-Based Rapid Autopsy Program “CASCADE”. PLOS Medicine. 2016. 13(12): p. e1002204.

[220] De Mattos-Arruda, L., *et al.* The Genomic and Immune Landscapes of Lethal Metastatic Breast Cancer. Cell Reports. 2019. 27(9): p. 2690-2708.e10.

[221] Zhu, Z., *et al.* Genome profiles of pathologist-defined cell clusters by multiregional LCM and G&T-seq in one triple-negative breast cancer patient. Cell Reports Medicine. 2021. 2(10): p. 100404.

[222] Miller, W.R., *et al.* Changes in breast cancer transcriptional profiles after treatment with the aromatase inhibitor, letrozole. Pharmacogenetics and Genomics. 2007. 17(10).

[223] Iwamoto, T., *et al.* Immunohistochemical Ki67 after short-term hormone therapy identifies low-risk breast cancers as reliably as genomic markers. Oncotarget. 2017. 8(16).

[224] Selli, C., *et al.* Molecular changes during extended neoadjuvant letrozole treatment of breast cancer: distinguishing acquired resistance from dormant tumours. Breast Cancer Research. 2019. 21(1): p. 2.

[225] Miller, C.A., *et al.* Aromatase inhibition remodels the clonal architecture of estrogen-receptor-positive breast cancers. Nature Communications. 2016. 7(1): p. 12498.

[226] Almendro, V., *et al.* Inference of Tumor Evolution during Chemotherapy by Computational Modeling and In Situ Analysis of Genetic and Phenotypic Cellular Diversity. Cell Reports. 2014. 6(3): p. 514-527.

[227] Chung, W., *et al.* Single-cell RNA-seq enables comprehensive tumour and immune cell

profiling in primary breast cancer. Nature Communications. 2017. 8(1): p. 15081.

[228] Gaudet, S.*, et al.* Exploring the Contextual Sensitivity of Factors that Determine Cell-to-Cell Variability in Receptor-Mediated Apoptosis. PLOS Computational Biology. 2012. 8(4): p. e1002482.

[229] Scott, J. and Marusyk, A. Somatic clonal evolution: A selection-centric perspective. Biochimica et Biophysica Acta (BBA) - Reviews on Cancer. 2017. 1867(2): p. 139-150.

[230] Rodrigo, G. and Stocks, N.G. Suprathreshold Stochastic Resonance behind Cancer. Trends in Biochemical Sciences. 2018. 43(7): p. 483-485.

[231] Griffiths, J.I.*, et al.* Serial single-cell genomics reveals convergent subclonal evolution of resistance as patients with early-stage breast cancer progress on endocrine plus CDK4/6 therapy. Nature Cancer. 2021. 2(6): p. 658-671.

[232] Hinohara, K.*, et al.* KDM5 Histone Demethylase Activity Links Cellular Transcriptomic Heterogeneity to Therapeutic Resistance. Cancer Cell. 2018. 34(6): p. 939-953.e9.

[233] Hong, S.P.*, et al.* Single-cell transcriptomics reveals multi-step adaptations to endocrine therapy. Nature communications. 2019. 10(1): p. 3840-3840.

[234] Krug, K.*, et al.* Proteogenomic Landscape of Breast Cancer Tumorigenesis and Targeted Therapy. Cell. 2020. 183(5): p. 1436-1456.e31.

[235] Wahjudi, L.W.*, et al.* Integrating proteomics into precision oncology. International Journal of Cancer. 2021. 148(6): p. 1438-1451.

[236] Borgoni, S. Time-resolved profiling reveals ATF3 as a novel mediator of endocrine resistance in breast cancer. in Molecular Genome Analysis. 2020. Heidelberg.

[237] Berthois, Y.*, et al.* Phenol red in tissue culture media is a weak estrogen: implications concerning the study of estrogen-responsive cells in culture. Proceedings of the National Academy of Sciences of the United States of America. 1986. 83(8): p. 2496-2500.

[238] Beumers, L.*, et al.* Deconvoluting clonal complexity of barcoded cell populations. 2022. Cytena.

[239] Ahmann, F.R.*, et al.* Intracellular adenosine triphosphate as a measure of human tumor cell viability and drug modulated growth. In Vitro Cellular & Developmental Biology. 1987. 23(7): p. 474-480.

[240] Sevin, B.U.*, et al.* Application of an ATP-bioluminescence assay in human tumor chemosensitivity testing. Gynecol Oncol. 1988. 31(1): p. 191-204.

[241] Crouch, S.P.M.*, et al.* The use of ATP bioluminescence as a measure of cell proliferation and cytotoxicity. Journal of Immunological Methods. 1993. 160(1): p. 81-88.

[242] Kepp, O.*, et al.* Cell death assays for drug discovery. Nature Reviews Drug Discovery. 2011. 10(3): p. 221-237.

[243] Yan, L. Abstract #DDT01-1: MK-2206: A potent oral allosteric AKT inhibitor. Cancer Research. 2009. 69(9_Supplement): p. DDT01-1-DDT01-1.

[244] Stottrup, C.*, et al.* Upregulation of AKT3 Confers Resistance to the AKT Inhibitor MK2206 in Breast Cancer. Molecular cancer therapeutics. 2016. 15(8): p. 1964-1974.

[245] Manasanch, E.E. and Orlowski, R.Z. Proteasome inhibitors in cancer therapy. Nature Reviews Clinical Oncology. 2017. 14(7): p. 417-433.

[246] Quintás-Cardama, A.*, et al.* Preclinical characterization of the selective JAK1/2 inhibitor INCB018424: therapeutic implications for the treatment of myeloproliferative neoplasms. Blood. 2010. 115(15): p. 3109-3117.

[247] Wagner, J.*, et al.* Discovery of 3-(1H-Indol-3-yl)-4-[2-(4-methylpiperazin-1-yl)quinazolin-4-yl]pyrrole-2,5-dione (AEB071), a Potent and Selective Inhibitor of Protein Kinase C Isotypes. Journal of Medicinal Chemistry. 2009. 52(20): p. 6193-6196.

[248] Geiges, D.*, et al.* Activation of protein kinase C subtypes α, γ, δ, ϵ, ζ, and η by tumor-promoting and nontumor-promoting agents. Biochemical Pharmacology. 1997. 53(6): p. 865-875.

[249] Goel, G.*, et al.* Phorbol Esters: Structure, Biological Activity, and Toxicity in Animals. International Journal of Toxicology. 2007. 26(4): p. 279-288.

[250] Sanger, F.*, et al.* DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A. 1977. 74(12): p. 5463-7.

[251] Morgan, M.*, et al.* ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data. Bioinformatics. 2009. 25(19): p. 2607-2608.

[252] R_Core_Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Url.

[253] RStudio_Team. RStudio: Integrated Development Environment for R}. Url.

[254] Huber, W.*, et al.* Orchestrating high-throughput genomic analysis with Bioconductor. Nature Methods. 2015. 12(2): p. 115-121.

[255] Liao, Y.*, et al.* The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. Nucleic Acids Research. 2019. 47(8): p. e47-e47.

[256] Liao, Y.*, et al.* featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2013. 30(7): p. 923-930.

[257] Reisinger, E.*, et al.* OTP: An automatized system for managing and processing NGS data. Journal of Biotechnology. 2017. 261: p. 53-62.

[258] Robinson, M.D.*, et al.* edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2009. 26(1): p. 139-140.

[259] Ritchie, M.E.*, et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Research. 2015. 43(7): p. e47-e47.

[260] Law, C.W.*, et al.* voom: precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biology. 2014. 15(2): p. R29.

[261] Schubert, M.*, et al.* Perturbation-response genes reveal signaling footprints in cancer gene expression. Nature Communications. 2018. 9(1): p. 20.

[262] Holland, C.H.*, et al.* Transfer of regulatory knowledge from human to mouse for functional genomics analysis. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms. 2020. 1863(6): p. 194431.

[263] Garcia-Alonso, L.*, et al.* Benchmark and integration of resources for the estimation of human transcription factor activities. Genome Research. 2019. 29(8): p. 1363-1375.

[264] Badia-i-Mompel, P.*, et al.* decoupleR: ensemble of computational methods to infer biological activities from omics data. Bioinformatics Advances. 2022. 2(1): p. vbac016.

[265] Wu, T.*, et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. The Innovation. 2021. 2(3): p. 100141.

[266] Mori, R.*, et al.* Both β-actin and GAPDH are useful reference genes for normalization of quantitative RT-PCR in human FFPE tissue samples of prostate cancer. The Prostate. 2008. 68(14): p. 1555-1560.

[267] Livak, K.J. and Schmittgen, T.D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. Methods. 2001. 25(4): p. 402-408.

[268] Laemmli, U.K. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. Nature. 1970. 227: p. 680.

[269] Wessel, D. and Flügge, U.I. A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. Analytical Biochemistry. 1984. 138(1): p. 141-143.

[270] Bekker-Jensen, D.B.*, et al.* Rapid and site-specific deep phosphoproteome profiling by data-independent acquisition without the need for spectral libraries. Nature Communications. 2020. 11(1): p. 787.

[271] Tyanova, S.*, et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. Nature Methods. 2016. 13(9): p. 731-740.

[272] Kim, H.J.*, et al.* PhosR enables processing and functional analysis of phosphoproteomic data. Cell Reports. 2021. 34(8): p. 108771.

[273] Huber, W.*, et al.* Variance stabilization applied to microarray data calibration and to the quantification of differential expression. Bioinformatics. 2002. 18(suppl_1): p. S96-S104.

[274] Türei, D.*, et al.* Integrated intra- and intercellular signaling knowledge for multicellular omics analysis. Molecular systems biology. 2021. 17(3): p. e9923.

[275] Cerami, E.*, et al.* The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. Cancer Discovery. 2012. 2(5): p. 401-404.

[276] Gao, J.*, et al.* Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. Science Signaling. 2013. 6(269): p. pl1-pl1.

[277] Barbie, D.A.*, et al.* Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature. 2009. 462(7269): p. 108-112.

[278] Wickham, H. *ggplot2*. Use R! 2016: Springer Cham. 260.

[279] Li, S.*, et al.* Endocrine-Therapy-Resistant ESR1 Variants Revealed by Genomic Characterization of Breast-Cancer-Derived Xenografts. Cell Reports. 2013. 4(6): p. 1116-1130.

[280] Robinson, D.R.*, et al.* Activating ESR1 mutations in hormone-resistant metastatic breast cancer. Nature Genetics. 2013. 45: p. 1446.

[281] Toy, W.*, et al.* ESR1 ligand-binding domain mutations in hormone-resistant breast cancer. Nature Genetics. 2013. 45: p. 1439.

[282] Karrison, T.G.*, et al.* Dormancy of mammary carcinoma after mastectomy. Journal of the National Cancer Institute. 1999. 91(1): p. 80-85.

[283] Meng, S.*, et al.* Circulating Tumor Cells in Patients with Breast Cancer Dormancy. Clinical Cancer Research. 2004. 10(24): p. 8152-8162.

[284] Meltzer, A. Dormancy and breast cancer. J Surg Oncol. 1990. 43(3): p. 181-8.

[285] Cabanos, H.F. and Hata, A.N. Emerging Insights into Targeted Therapy-Tolerant Persister Cells in Cancer. Cancers. 2021. 13(11): p. 2666.

[286] Xing, Y.*, et al.* Phase II trial of AKT inhibitor MK-2206 in patients with advanced breast cancer who have tumors with PIK3CA or AKT mutations, and/or PTEN loss/PTEN mutation. Breast Cancer Research. 2019. 21(1): p. 78.

[287] Kovacina, K.S.*, et al.* Identification of a Proline-rich Akt Substrate as a 14-3-3 Binding Partner*. Journal of Biological Chemistry. 2003. 278(12): p. 10189-10194.

[288] Hayden, M.S. and Ghosh, S. Shared Principles in NF-κB Signaling. Cell. 2008. 132(3): p. 344-362.

[289] Chen, J. and Chen, Z.J. Regulation of NF-κB by ubiquitination. Current Opinion in Immunology. 2013. 25(1): p. 4-12.

[290] Taniguchi, K. and Karin, M. NF-κB, inflammation, immunity and cancer: coming of age. Nature Reviews Immunology. 2018. 18(5): p. 309-324.

[291] Wong, A.H.*, et al.* Targeting NF-κB Signaling for Multiple Myeloma. Cancers (Basel). 2020. 12(8).

[292] Obeng, E.A.*, et al.* Proteasome inhibitors induce a terminal unfolded protein response in multiple myeloma cells. Blood. 2006. 107(12): p. 4907-4916.

[293] Shapiro, D.J.*, et al.* Anticipatory UPR Activation: A Protective Pathway and Target in Cancer. Trends in Endocrinology & Metabolism. 2016. 27(10): p. 731-741.

[294] Hetz, C.*, et al.* Mechanisms, regulation and functions of the unfolded protein response. Nature Reviews Molecular Cell Biology. 2020. 21(8): p. 421-438.

[295] Mootha, V.K.*, et al.* PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nature Genetics. 2003. 34(3): p. 267-273.

[296] Subramanian, A.*, et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences. 2005. 102(43): p. 15545.

[297] Subramanian, A.*, et al.* GSEA-P: a desktop application for Gene Set Enrichment Analysis. Bioinformatics. 2007. 23(23): p. 3251-3253.

[298] EBCTCG. Relevance of breast cancer hormone receptors and other factors to the efficacy of

adjuvant tamoxifen: patient-level meta-analysis of randomised trials. Lancet. 2011. 378(9793): p. 771-784.

[299] Regan, M.M., *et al.* Assessment of letrozole and tamoxifen alone and in sequence for postmenopausal women with steroid hormone receptor-positive breast cancer: the BIG 1-98 randomised clinical trial at 8.1 years median follow-up. The Lancet Oncology. 2011. 12(12): p. 1101-1108.

[300] EBCTCG. Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. The Lancet. 2015. 386(10001): p. 1341-1352.

[301] Antonarakis, E.S. Targeting lineage plasticity in prostate cancer. The Lancet Oncology. 2019. 20(10): p. 1338-1340.

[302] Quintanal-Villalonga, Á., *et al.* Lineage plasticity in cancer: a shared pathway of therapeutic resistance. Nature Reviews Clinical Oncology. 2020. 17(6): p. 360-371.

[303] Oren, Y., *et al.* Cycling cancer persister cells arise from lineages with distinct programs. Nature. 2021. 596(7873): p. 576-582.

[304] Güth, U., *et al.* Target and reality of adjuvant endocrine therapy in postmenopausal patients with invasive breast cancer. British Journal of Cancer. 2008. 99(3): p. 428-433.

[305] Kimmick, G., *et al.* Adjuvant Hormonal Therapy Use Among Insured, Low-Income Women With Breast Cancer. Journal of Clinical Oncology. 2009. 27(21): p. 3445-3451.

[306] Ziller, V., *et al.* Adherence to adjuvant endocrine therapy in postmenopausal women with breast cancer. Annals of Oncology. 2009. 20(3): p. 431-436.

[307] Neugut, A.I., *et al.* Association Between Prescription Co-Payment Amount and Compliance With Adjuvant Hormonal Therapy in Women With Early-Stage Breast Cancer. Journal of Clinical Oncology. 2011. 29(18): p. 2534-2542.

[308] Chang, C.-H., *et al.* ESX: a structurally unique Ets overexpressed early during human breast tumorigenesis. Oncogene. 1997. 14(13): p. 1617-1622.

[309] Neve, R.M., *et al.* ErbB2 Activation of ESX gene expression. Oncogene. 2002. 21(24): p. 3934-3938.

[310] Kar, A. and Gutierrez-Hartmann, A. ESE-1/ELF3 mRNA expression associates with poor survival outcomes in HER2(+) breast cancer patients and is critical for tumorigenesis in HER2(+) breast cancer cells. Oncotarget. 2017. 8(41): p. 69622-69640.

[311] Kumegawa, K., *et al.* GRHL2 motif is associated with intratumor heterogeneity of cis-regulatory elements in luminal breast cancer. npj Breast Cancer. 2022. 8(1): p. 70.

[312] Cocce, K.J., *et al.* The Lineage Determining Factor GRHL2 Collaborates with FOXA1 to Establish a Targetable Pathway in Endocrine Therapy-Resistant Breast Cancer. Cell Reports. 2019. 29(4): p. 889-903.e10.

[313] Che, M., *et al.* Opposing transcriptional programs of KLF5 and AR emerge during therapy for advanced prostate cancer. Nature Communications. 2021. 12(1): p. 6377.

[314] Tate, J.G., *et al.* COSMIC: the Catalogue Of Somatic Mutations In Cancer. Nucleic Acids Research. 2018. 47(D1): p. D941-D947.

[315] Nguyen, V.T.M., *et al.* Differential epigenetic reprogramming in response to specific endocrine therapies promotes cholesterol biosynthesis and cellular invasion. Nature Communications. 2015. 6: p. 10044.

[316] Wang, X. and Wang, S. Identification of key genes involved in tamoxifen-resistant breast cancer using bioinformatics analysis. Translational Cancer Research. 2021. 10(12): p. 5246-5257.

[317] Asghari, A., *et al.* Abstract P2-11-20: Endocrine resistance in breast cancer: A dynamic gene expression analysis approach reveals potential new responsible genes. Cancer Research. 2020. 80(4_Supplement): p. P2-11-20-P2-11-20.

[318] Creighton, C.J., *et al.* Development of Resistance to Targeted Therapies Transforms the Clinically Associated Molecular Profile Subtype of Breast Tumor Xenografts. Cancer Research. 2008. 68(18): p. 7493-7501.

[319] Hatami, R., *et al.* KLF6-SV1 Drives Breast Cancer Metastasis and Is Associated with Poor Survival. Science Translational Medicine. 2013. 5(169): p. 169ra12-169ra12.

[320] Xiao, Q., *et al.* MEF2A transcriptionally upregulates the expression of ZEB2 and CTNNB1 in colorectal cancer to promote tumor progression. Oncogene. 2021. 40(19): p. 3364-3377.

[321] He, L., *et al.* Glucocorticoid Receptor Signaling Activates TEAD4 to Promote Breast Cancer Progression. Cancer Research. 2019. 79(17): p. 4399-4411.

[322] Brown, F.C., *et al.* MEF2C Phosphorylation Is Required for Chemotherapy Resistance in Acute Myeloid Leukemia. Cancer Discovery. 2018. 8(4): p. 478-497.

[323] Rosales-Aviña, J.A., *et al.* MEIS1, PREP1, and PBX4 Are Differentially Expressed in Acute Lymphoblastic Leukemia: Association of MEIS1 Expression with Higher Proliferation and Chemotherapy Resistance. Journal of Experimental & Clinical Cancer Research. 2011. 30(1): p. 112.

[324] Gupta, A., *et al.* NCOA3 coactivator is a transcriptional target of XBP1 and regulates PERK–eIF2α–ATF4 signalling in breast cancer. Oncogene. 2016. 35(45): p. 5860-5871.

[325] Rangel, J., *et al.* Prognostic Significance of Nuclear Receptor Coactivator-3 Overexpression in Primary Cutaneous Melanoma. Journal of Clinical Oncology. 2006. 24(28): p. 4565-4569.

[326] Li, R., *et al.* Impact of next-generation sequencing (NGS) for primary endocrine resistance in breast cancer patients. Int J Clin Exp Pathol. 2018. 11(11): p. 5450-5458.

[327] Zhao, W.-L., *et al.* PRDM1 is involved in chemoresistance of T-cell lymphoma and down-regulated by the proteasome inhibitor. Blood. 2008. 111(7): p. 3867-3871.

[328] Crosby, M.E., *et al.* E2F4 regulates a stable G2 arrest response to genotoxic stress in prostate carcinoma. Oncogene. 2007. 26(13): p. 1897-1909.

[329] Bretones, G., *et al.* Myc and cell cycle control. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms. 2015. 1849(5): p. 506-516.

[330] Millour, J., *et al.* FOXM1 is a transcriptional target of ERα and has a critical role in breast cancer endocrine sensitivity and resistance. Oncogene. 2010. 29(20): p. 2983-2995.

[331] Amaddeo, G.*, et al.* Next-generation sequencing identified new oncogenes and tumor suppressor genes in human hepatic tumors. OncoImmunology. 2012. 1(9): p. 1612-1613.

[332] Guichard, C.*, et al.* Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. Nature Genetics. 2012. 44(6): p. 694-698.

[333] Ijichi, N.*, et al.* Association of Double-Positive FOXA1 and FOXP1 Immunoreactivities with Favorable Prognosis of Tamoxifen-Treated Breast Cancer Patients. Hormones and Cancer. 2012. 3(4): p. 147-159.

[334] Fish, K.J.*, et al.* Isolation and Characterization of Human Casein Kinase I∊ (CKI), a Novel Member of the CKI Gene Family (∗). Journal of Biological Chemistry. 1995. 270(25): p. 14875-14883.

[335] Schittek, B. and Sinnberg, T. Biological functions of casein kinase 1 isoforms and putative roles in tumorigenesis. Molecular Cancer. 2014. 13(1): p. 231.

[336] Hole, S.*, et al.* Aurora kinase A and B as new treatment targets in aromatase inhibitor-resistant breast cancer cells. Breast Cancer Research and Treatment. 2015. 149(3): p. 715-726.

[337] Larsen, S.L.*, et al.* Aurora kinase B is important for antiestrogen resistant cell growth and a potential biomarker for tamoxifen resistant breast cancer. BMC Cancer. 2015. 15(1): p. 239.

[338] Wang, Y.-y.*, et al.* The emerging role of CaMKII in cancer. Oncotarget. 2015. 6(14).

[339] Ma, Y.*, et al.* Oestrogen confers cardioprotection by suppressing Ca2+/calmodulin-dependent protein kinase II. British Journal of Pharmacology. 2009. 157(5): p. 705-715.

[340] Xu, X.*, et al.* Overexpression of SMARCA2 or CAMK2D is associated with cisplatin resistance in human epithelial ovarian cancer. Oncol Lett. 2018. 16(3): p. 3796-3804.

[341] Hagemann, D.*, et al.* Cot protooncoprotein activates the dual specificity kinases MEK-1 and SEK-1 and induces differentiation of PC12 cells. Oncogene. 1999. 18(7): p. 1391-400.

[342] Jeong, J.H.*, et al.* TPL2/COT/MAP3K8 (TPL2) Activation Promotes Androgen Depletion-Independent (ADI) Prostate Cancer Growth. PLOS ONE. 2011. 6(1): p. e16205.

[343] Sabnis, G.J.*, et al.* Effect of selumetinib on the growth of anastrozole-resistant tumors. Breast Cancer Research and Treatment. 2013. 138(3): p. 699-708.

[344] McCubrey, J.A.*, et al.* Diverse roles of GSK-3: Tumor promoter–tumor suppressor, target in cancer therapy. Advances in Biological Regulation. 2014. 54: p. 176-196.

[345] Sokolosky, M.*, et al.* Inhibition of GSK-3β activity can result in drug and hormonal resistance and alter sensitivity to targeted therapy in MCF-7 breast cancer cells. Cell Cycle. 2014. 13(5): p. 820-833.

[346] Tang, L.*, et al.* Epigenetic identification of mitogen-activated protein kinase 10 as a functional tumor suppressor and clinical significance for hepatocellular carcinoma. PeerJ. 2021. 9: p. e10810.

[347] Chen, L.*, et al.* Inhibition of the p38 Kinase Suppresses the Proliferation of Human ER-Negative Breast Cancer Cells. Cancer Research. 2009. 69(23): p. 8853-8861.

[348] Ghayad, S.E.*, et al.* mTOR inhibition reverses acquired endocrine therapy resistance of breast cancer cells at the cell proliferation and gene-expression levels. Cancer Science. 2008. 99(10): p. 1992-2003.

[349] Ma, C.X.*, et al.* A Phase II Trial of Neoadjuvant MK-2206, an AKT Inhibitor, with Anastrozole in Clinical Stage II or III PIK3CA-Mutant ER-Positive and HER2-Negative Breast Cancer. Clinical Cancer Research. 2017. 23(22): p. 6823-6832.

[350] Hu, X.*, et al.* The JAK/STAT signaling pathway: from bench to clinic. Signal Transduction and Targeted Therapy. 2021. 6(1): p. 402.

[351] Platanias, L.C. Mechanisms of type-I- and type-II-interferon-mediated signalling. Nature Reviews Immunology. 2005. 5(5): p. 375-386.

[352] Bluyssen, H.R. and Levy, D.E. Stat2 Is a Transcriptional Activator That Requires Sequence-specific Contacts Provided by Stat1 and p48 for Stable Interaction with DNA. Journal of Biological Chemistry. 1997. 272(7): p. 4600-4605.

[353] Nan, J.*, et al.* IRF9 and unphosphorylated STAT2 cooperate with NF-κB to drive IL6 expression. Proc Natl Acad Sci U S A. 2018. 115(15): p. 3906-3911.

[354] Ulane, C.M. and Horvath, C.M. Paramyxoviruses SV5 and HPIV2 Assemble STAT Protein Ubiquitin Ligase Complexes from Cellular Components. Virology. 2002. 304(2): p. 160-166.

[355] Le, V.T.K.*, et al.* Human cytomegalovirus interferes with signal transducer and activator of transcription (STAT) 2 protein stability and tyrosine phosphorylation. Journal of General Virology. 2008. 89(10): p. 2416-2426.

[356] Ashour, J.*, et al.* NS5 of Dengue Virus Mediates STAT2 Binding and Degradation. Journal of Virology. 2009. 83(11): p. 5408-5418.

[357] Lee, C.J.*, et al.* FBXW7-mediated stability regulation of signal transducer and activator of transcription 2 in melanoma formation. Proc Natl Acad Sci U S A. 2020. 117(1): p. 584-594.

[358] Hu, R.*, et al.* NF-κB Signaling Is Required for XBP1 (Unspliced and Spliced)-Mediated Effects on Antiestrogen Responsiveness and Cell Fate Decisions in Breast Cancer. Molecular and Cellular Biology. 2015. 35(2): p. 379-390.

[359] Chen, X.*, et al.* XBP1 promotes triple-negative breast cancer by controlling the HIF1α pathway. Nature. 2014. 508(7494): p. 103-107.

[360] Nagelkerke, A.*, et al.* Hypoxia stimulates migration of breast cancer cells via the PERK/ATF4/LAMP3-arm of the unfolded protein response. Breast Cancer Research. 2013. 15(1): p. R2.

[361] Cook, K.L.*, et al.* Knockdown of estrogen receptor-α induces autophagy and inhibits antiestrogen-mediated unfolded protein response activation, promoting ROS-induced breast cancer cell death. The FASEB Journal. 2014. 28(9): p. 3891-3905.

[362] Andruska, N.*, et al.* Anticipatory estrogen activation of the unfolded protein response is linked to cell proliferation and poor survival in estrogen receptor α-positive breast cancer. Oncogene. 2015. 34(29): p. 3760-3769.

[363] Ishii, Y.*, et al.* Bortezomib Enhances the Efficacy of Fulvestrant by Amplifying the Aggregation of the Estrogen Receptor, Which Leads

to a Proapoptotic Unfolded Protein Response. Clinical Cancer Research. 2011. 17(8): p. 2292-2300.

[364] Adelson, K*., et al.* Randomized phase II trial of fulvestrant alone or in combination with bortezomib in hormone receptor-positive metastatic breast cancer resistant to aromatase inhibitors: a New York Cancer Consortium trial. npj Breast Cancer. 2016. 2(1): p. 16037.

[365] Newton, A.C. Protein kinase C: perfectly balanced. Crit Rev Biochem Mol Biol. 2018. 53(2): p. 208-230.

[366] Parker, P.J*., et al.* Equivocal, explicit and emergent actions of PKC isoforms in cancer. Nature Reviews Cancer. 2021. 21(1): p. 51-63.

[367] He, S*., et al.* Targeting Protein Kinase C for Cancer Therapy. Cancers. 2022. 14(5): p. 1104.

[368] Castagna, M*., et al.* Direct activation of calcium-activated, phospholipid-dependent protein kinase by tumor-promoting phorbol esters. J Biol Chem. 1982. 257(13): p. 7847-51.

[369] Filler, R.B*., et al.* Cutaneous Two-Stage Chemical Carcinogenesis. Cold Spring Harbor Protocols. 2007. 2007(9): p. pdb.prot4837.

[370] Ono, Y*., et al.* Phorbol ester binding to protein kinase C requires a cysteine-rich zinc-finger-like sequence. Proceedings of the National Academy of Sciences. 1989. 86(13): p. 4868-4871.

[371] Garg, R*., et al.* Protein kinase C and cancer: what we know and what we do not. Oncogene. 2014. 33(45): p. 5225-5237.

[372] Cho, Y*., et al.* Retinoic acid induced growth arrest of human breast carcinoma cells requires protein kinase C alpha expression and activity. J Cell Physiol. 1997. 172(3): p. 306-13.

[373] Nakagawa, S*., et al.* Cell growth inhibition by all-trans retinoic acid in SKBR-3 breast cancer cells: involvement of protein kinase Calpha and extracellular signal-regulated kinase mitogen-activated protein kinase. Mol Carcinog. 2003. 38(3): p. 106-16.

[374] Teegala, L.R*., et al.* Protein Kinase C α and β compensate for each other to promote stem cell factor-mediated KIT phosphorylation, mast cell viability and proliferation. The FASEB Journal. 2022. 36(5): p. e22273.

[375] Meggio, F*., et al.* Different Susceptibility of Protein Kinases to Staurosporine Inhibition. European Journal of Biochemistry. 1995. 234(1): p. 317-322.

[376] Gonzalez-Guerrico, A.M. and Kazanietz, M.G. Phorbol Ester-induced Apoptosis in Prostate Cancer Cells via Autocrine Activation of the Extrinsic Apoptotic Cascade: a key role for protein kinase c delta*. Journal of Biological Chemistry. 2005. 280(47): p. 38982-38991.

[377] Cerda, S.R*., et al.* Protein kinase C delta inhibits Caco-2 cell proliferation by selective changes in cell cycle and cell death regulators. Oncogene. 2006. 25(22): p. 3123-3138.

[378] Bowles, D.K*., et al.* PKCδ mediates anti-proliferative, pro-apoptic effects of testosterone on coronary smooth muscle. American Journal of Physiology-Cell Physiology. 2007. 293(2): p. C805-C813.

[379] Zhao, M*., et al.* Protein Kinase Cδ in Apoptosis: A Brief Overview. Archivum Immunologiae et Therapiae Experimentalis. 2012. 60(5): p. 361-372.

[380] Pal, D*., et al.* Upregulation of PKCη by PKCε and PDK1 involves two distinct mechanisms and promotes breast cancer cell survival. Biochimica et Biophysica Acta (BBA) - General Subjects. 2013. 1830(8): p. 4040-4045.

[381] Osada, S*., et al.* Predominant expression of nPKC eta, a Ca(2+)-independent isoform of protein kinase C in epithelial tissues, in association with epithelial differentiation. Cell Growth Differ. 1993. 4(3): p. 167-75.

[382] Ishino, K*., et al.* Phorbol Ester-induced G1 Arrest in BALB/MK-2 Mouse Keratinocytes Is Mediated by δ and η Isoforms of Protein Kinase C. Japanese Journal of Cancer Research. 1998. 89(11): p. 1126-1133.

[383] Cabodi, S*., et al.* A PKC-η/Fyn-Dependent Pathway Leading to Keratinocyte Growth Arrest and Differentiation. Molecular Cell. 2000. 6(5): p. 1121-1129.

[384] Black, A. and Black, J. Protein kinase C signaling and cell cycle regulation. Frontiers in Immunology. 2013. 3.

[385] Albert, V*., et al.* Protein kinase C targeting of luminal (T-47D), luminal/HER2-positive (BT474), and triple negative (HCC1806) breast cancer cells in-vitro with AEB071 (Sotrastaurin) is efficient but mediated by subtype specific molecular effects. Archives of Gynecology and Obstetrics. 2022. 306(4): p. 1197-1210.

[386] Assender, J.W*., et al.* Protein kinase C isoform expression as a predictor of disease outcome on endocrine therapy in breast cancer. Journal of Clinical Pathology. 2007. 60(11): p. 1216-1221.

[387] Tonetti, D.A*., et al.* PKCa and ERß Are Associated with Triple-Negative Breast Cancers in African American and Caucasian Patients. International Journal of Breast Cancer. 2012. 2012: p. 740353.

[388] Tan, M*., et al.* Upregulation and activation of PKCα by ErbB2 through Src promotes breast cancer cell invasion that can be blocked by combined treatment with PKCα and Src inhibitors. Oncogene. 2006. 25(23): p. 3286-3295.

[389] Ways, D.K*., et al.* MCF-7 breast cancer cells transfected with protein kinase C-alpha exhibit altered expression of other protein kinase C isoforms and display a more aggressive neoplastic phenotype. The Journal of Clinical Investigation. 1995. 95(4): p. 1906-1915.

[390] Tonetti, D.A*., et al.* Stable transfection of protein kinase C alpha cDNA in hormone-dependent breast cancer cell lines. British Journal of Cancer. 2000. 83(6): p. 782-791.

[391] Li, H. and Weinstein, I.B. Protein Kinase C β Enhances Growth and Expression of Cyclin D1 in Human Breast Cancer Cells. Cancer Research. 2006. 66(23): p. 11399-11408.

[392] Pan, Q*., et al.* Protein Kinase Cε Is a Predictive Biomarker of Aggressive Breast Cancer and a Validated Target for RNA Interference Anticancer Therapy. Cancer Research. 2005. 65(18): p. 8366-8371.

[393] Belguise, K. and Sonenshein, G.E. PKCθ promotes c-Rel–driven mammary tumorigenesis in mice and humans by repressing estrogen receptor α synthesis. The Journal of Clinical Investigation. 2007. 117(12): p. 4009-4021.

[394] Good, B.H*., et al.* The dynamics of molecular evolution over 60,000 generations. Nature. 2017. 551(7678): p. 45-50.

[395] Marusyk, A., *et al.* Non-cell-autonomous driving of tumour growth supports sub-clonal heterogeneity. Nature. 2014. 514(7520): p. 54-58.

[396] Caswell, D.R. and Swanton, C. The role of tumour heterogeneity and clonal cooperativity in metastasis, immune evasion and clinical outcome. BMC Medicine. 2017. 15(1): p. 133.

[397] Brunner, A.-D., *et al.* Ultra-high sensitivity mass spectrometry quantifies single-cell proteome changes upon perturbation. Molecular Systems Biology. 2022. 18(3): p. e10798.

[398] Mund, A., *et al.* Deep Visual Proteomics defines single-cell identity and heterogeneity. Nature Biotechnology. 2022.

[399] Lomakin, A., *et al.* Spatial genomics maps the structure, nature and evolution of cancer clones. Nature. 2022.

[400] Henley, M.J. and Koehler, A.N. Advances in targeting ‘undruggable’ transcription factors with small molecules. Nature Reviews Drug Discovery. 2021. 20(9): p. 669-688.

[401] Schust, J., *et al.* Stattic: a small-molecule inhibitor of STAT3 activation and dimerization. Chem Biol. 2006. 13(11): p. 1235-42.

[402] Bruford, E.A., *et al.* Guidelines for human gene nomenclature. Nature Genetics. 2020. 52(8): p. 754-758.

## 11 Contributions

It is stated clearly that I performed most of the *in vitro* experiments but received support mainly from Sara Burmester (Technician, under my supervision). I received further support for experiments from Daniela Heiss (Technician, under my supervision) and Luisa Schwarzmüller (PhD student). The bioinformatic analyses were performed mainly by Dr. Efstathios-Iason Vlachavas. Further bioinformatic analyses were performed by Dr. Birgitta Michels.

## 12 Acknowledgements

My PhD studies have been supported by different people and without their support, this thesis would not have been possible. First of all, I want to express my deepest gratitude to Prof. Dr. Stefan Wiemann for giving me the opportunity to perform my PhD thesis in the division of Molecular Genome Analysis (MGA). I would like to thank Prof. Dr. Stefan Wiemann for his constant support and his encouragement to develop into an independent researcher.

My Thesis Advisory Committee (TAC) members Prof. Dr. Stefan Wiemann, Prof. Dr. Peter Angel and Prof. Dr. Luca Magnani I want to thank for critical discussions, their guidance and support during my PhD studies. I would further want to thank Prof. Dr. Stefan Wiemann, Prof. Dr. Peter Angel, Prof. Dr. Ursula Kummer and Prof. Dr. Dr. Georg Stoecklin for agreeing to be part of my thesis defense committee.

For generous funding of my PhD studies, I would like to thank the DKFZ International PhD Program and the Cancer TRAX program.

My PhD project was a continuation of a previous PhD project established by Dr. Simone Borgoni and I am happy that I could take over this interesting project. In line, I am thankful for the further discussions with and input from Dr. Angela Goncalves and Dr. Luca Penso Dolfin. The initial barcoding and singularization of cells were repeated as a proof-of-concept Application Note in collaboration with the DKFZ Cellular Tools Core Facility. I am very grateful for the technical support and discussions with Dr. Rainer Will, Rita Schatten, Birgit Kaiser, Heike Wilhelm and Katja Bauer. I was further privileged to work in their sacred S2 cell culture. I am particularly thankful to Dr. Rainer Will and Rita Schatten for the discussions and their input. For the deconvolution of complex replicates I am grateful for the assistance I received from the DKFZ scOpenLab, Solentim and Cytena. In line, I would like to acknowledge the DKFZ NGS and Mass Spectrometry Core Facilities for providing excellent services and technical support for the RNA-Seq, NGS of barcode region and (phospho-)proteomics analysis.

For the last couple of years, first as Master thesis student and then as PhD student, I was part of MGA. I met many people along the way that trained, helped and shaped me. Thank you very much! I particularly want to acknowledge Dr. Cindy Körner for her constant help. Further, I am grateful for Dr. Simone Borgoni for initially supervising me in the lab and his constant input. Additionally, I would like to thank Luisa Schwarzmüller for the protein sample preparation and phospho-proteomics measurement by Mass spectrometry as well as the final figure check. I am moreover indebted to Daniela Heiss for the great Western Blots. I would also like to thank Dr. Birgitta Michels for her data analysis, our everyday discussions and for reminding me (as often as needed) that *et kütt wie et kütt*.

I am particularly grateful for Dr. Efstathios-Iason Vlachavas for his constant support, our discussions and the analysis of many datasets. Without you, this thesis would not have been the same. Moreover, I appreciate the friendship we built and that I now know the proper way to drink espresso.

Furthermore, I want to thank Sara Burmester for working (and bearing) with me on the project for three years. You have been a source of great support and I could not have done it without you!

I would like to thank my friends for providing the best antidotes to my PhD studies. Particularly, Cathi, Rebi, Marvin, Haidar, Carlos, Fabian, Ana, Dilay, Aga and Susana: Thank you very much for always making me laugh and being there for me.

Ein großes Dankeschön geht an meine Familie. Vielen Dank, dass Ihr immer für mich da seid und mich bei allem unterstützt, was ich mir so in den Kopf setze!

Finally, I would like to thank Pia for making my life so much easier and more enjoyable. Please stay as you are and never change! I am looking forward to what the future will bring. Next stop: Berlin.