Dissertation

submitted to the

Combined Faculties for the Natural Sciences and for Mathematics

of the Ruperto Carola University of Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

Presented by

M.Sc. Emre Sofyali

born in: Edirne, Turkey

Oral examination: 09.09.2019

GLYATL1 PROMOTES ENDOCRINE THERAPY RESISTANCE IN BREAST CANCER

Referees: Prof. Dr. Stefan Wiemann

PD Dr. rer. nat. Karin Müller-Decker

Declaration of Authorship

I hereby declare that the work presented in my dissertation was carried out between August 2015 and July 2019 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, _____

Emre Sofyali

Table of Contents

List of Figures

List of Tables

Summary

Breast cancer is the most frequently diagnosed malignancy among women. Due to its molecular heterogeneity, a generalized therapy is not possible. Nuclear estrogen receptor-α (ER-α) is overexpressed in the majority of breast tumors. ER-positive patients benefit from endocrine therapy that abrogates estrogen-induced tumor growth. However, approximately half of the patients do not respond or relapse due to *de novo* or acquired resistance against therapy. Occurrence of resistance is a drawback for long-term efficacy of endocrine therapy and leads to poor prognosis. The mechanisms underlying acquisition of endocrine therapy resistance, however, remain elusive. Recently, epigenetic reprogramming has been proposed as a means to render the tumor cells refractory to treatment. Therefore, the aims of this project were to uncover novel targets that confer resistance and to elucidate the involvement of epigenetics in this process. To this end, two ER-positive cell lines (MCF7 and T47D) were utilized to recapitulate endocrine therapy resistance *in vitro* by treating them either with tamoxifen (TAMR) or depriving them of estrogen (LTED). I identified *GLYATL1* (glycine-N-acyltransferase like 1) as a highly de-regulated gene as revealed by RNA-seq and ATAC-seq integrative analysis comparing resistant cell lines to the sensitive parental. *GLYATL1* encodes for an enzyme that catalyzes the transfer of an acyl group to glutamine. I showed that knockdown of *GLYATL1* sensitizes resistant cell lines while GLYATL1 overexpression renders sensitive luminal cells resistant to endocrine therapy. Furthermore, I found GLYATL1 is involved in acetylation of histone residues H3K9 and H3K14 since the knockdown of *GLYATL1* led to a decrease in these two histone marks in resistant MCF7 cells. Moreover, I showed the expression of *GLYATL1* to be regulated in these cells by methylation, growth factor receptor HER2, and transcription factors ERα, GATA3 and p300. CRISPR/dCas9-mediated epigenetic editing method was adopted to validate the involvement of methylation in *GLYATL1* regulation. This method is a repurposed version of CRISPR/Cas9 system where Cas9 is catalytically inactive and fused to catalytic domain of epigenetic enzymes. Combined with sgRNAs, these effectors can be recruited to target regions to modulate the epigenetic landscape, thereby altering gene expression. Furthermore, I utilized this method of epigenetic editing to investigate endocrine therapy resistance involvement of other genes such as *CD44* and *BAMBI*, expression of which were also found to be elevated in resistant cells compared to parental. I showed that targeting promoter regions of *CD44* and *BAMBI* with dCas9-p300 yielded upregulation of both genes whereas dCas9-G9a combination led to a downregulation which resulted in retarded proliferation in LTED cells. Moreover, altering expression of *BAMBI* elicited similar

changes in *CD44* expression further proving *CD44* as a direct target gene of the Wnt signaling pathway, for which *BAMBI* acts as an activator. In conclusion, my results demonstrate the importance of GLYATL1 in initiation and maintenance of endocrine therapy resistance and identify its involvement in H3K9 and H3K14 acetylation. This study demonstrates the potential of epigenetic reprogramming mediated regulation of target gene expression as a novel method of therapeutic intervention.

Zusammenfassung

Brustkrebs ist die am häufigsten diagnostizierte Krebskrankheit bei Frauen. Aufgrund der molekularen Heterogenität ist eine generalisierte Therapie nicht möglich. Der nukleäre Östrogenrezeptor α (ERα) ist bei der Mehrheit aller Brusttumoren überexprimiert. ER-positive Patienten profitieren von einer endokrinen Therapie die Östrogen-induziertes Tumorwachstum unterbindet. Allerdings spricht ungefähr die Hälfte der Patienten nicht an oder rezidiviert aufgrund von *de novo* oder erworbener Therapieresistenz. Das Auftreten von Resistenz ist ein Rückschritt für die Langzeiteffizienz von endokriner Therapie und führt zu einer schlechten Prognose. Die Mechanismen die dem Erwerb der endokrinen Therapieresistenz unterliegen sind jedoch unbekannt. Kürzlich wurde epigenetische Reprogrammierung als Mittel vorgeschlagen, welches Tumorzellen unempfänglich für die Therapie macht. Daher waren die Ziele dieses Projekts neue Zielmoleküle zu entdecken die Resistenz vermitteln und die Rolle der Epigenetik in diesem Prozess aufzuklären. Zu diesem Zweck wurden zwei ER-positive Zelllinien (MCF7 und T47D) verwendet um die endokrine Therapieresistenz *in vitro* zu rekapitulieren indem die Zellen entweder mit Tamoxifen (TAMR) behandelt wurden oder indem ihnen das Östrogen entzogen wurde (LTED). Ich identifizierte *GLYATL1* (glycine-N-acyltransferase like 1) als ein stark dereguliertes Gen wie die integrative Analyse von RNA-Seq und ATAC-Seq, die resistente Zelllinien mit sensitiven parentalen Zelllinien verglich, zeigte. *GLYATL1* kodiert ein Enzym das den Transfer einer Acylgruppe auf Glutamin katalysiert. Ich konnte zeigen, dass Knockdown von *GLYATL1* resistente Zelllinien sensibilisiert, während die Überexpression von *GLYATL1* sensitiven luminalen Zellen Resistenz gegen endokrine Therapie verleiht. Des Weiteren fand ich heraus, dass GLYATL1 in die Acetylierung der Histonreste H3K9 und K3K14 involviert ist, da Knockdown von *GLYATL1* zur Abnahme der beiden Histonmodifikationen in resistenten MCF7 Zellen führte. Außerdem zeigte ich, dass die Expression von *GLYATL1* in diesen Zellen durch Methylierung, den Wachstumsfaktor-Rezeptor HER2 und die Transkriptionsfaktoren ERα, GATA3 und p300 reguliert wird. CRISPR/dCas9-vermittelte epigenetische Editierung wurde eingesetzt um die Beteiligung von Methylierung an der *GLYATL1* Regulation zu validieren. Diese Methode ist eine umfunktionierte Version des CRISPR/Cas9 Systems bei dem Cas9 katalytisch inaktiv ist und an die katalytische Domäne epigenetischer Enzyme fusioniert wird. In Kombination mit sgRNAs können diese Effektoren zu Zielregionen rekrutiert werden um die epigenetische Landschaft zu modulieren und dadurch die Genexpression zu verändern. Darüber hinaus verwendete ich diese Methode der epigenetischen Editierung um die

Beteiligung anderer Gene an der endokrinen Therapieresistenz zu untersuchen. Zum Beispiel *CD44* und *BAMBI*, deren Expression in resistenten Zellen im Vergleich zu parentalen Zellen ebenfalls erhöht war. Ich zeigte, dass epigenetische Modulierung der Promotorregionen von *CD44* und *BAMBI* mit dCas9-p300 zur Hochregulierung beider Gene führte, wohingegen die Kombination dCas9-G9a zu einer Herunterregulierung führte, die in verlangsamter Proliferation von LTED Zellen resultierte. Außerdem rief eine Änderung der *BAMBI* Expression ähnliche Änderungen der *CD44* Expression hervor, was *CD44* als direktes Ziel des Wnt Signalweges, für den *BAMBI* als Aktivator agiert, bestätigte. Abschließend zeigen meine Ergebnisse die Bedeutung von GLYATL1 in der Initiierung und Erhaltung von endokriner Therapieresistenz und identifizieren seine Beteiligung an der Acetylierung von H3K9 und H3K14. Diese Studie demonstriert das Potential der durch epigenetische Reprogrammierung vermittelte Regulation der Expression von Zielgenen als neue Methode für die therapeutische Intervention.

1. Introduction

1.1 Breast Cancer

Breast cancer is the most frequently diagnosed malignancy among women worldwide (Siegel et al., 2019). According to Global Cancer Observatory (GLOBOCAN) statistics, in 2018, there were more than two million new cases of breast cancer (24.2% of all cancer entities affecting females) and more than 600,000 deaths which accounts for 15% of cancer related decease among females (Bray et al., 2018). Breast cancer is a highly heterogeneous disease rendering a generalization of therapy impossible. Depending on the status of hormone receptors, estrogen (ER) and progesterone (PR) receptors, and the receptor tyrosine kinase (RTK) HER2/neu breast cancer can be subdivided into clinical subtypes including Luminal A, Luminal B, HER2-enriched, Triple-Negative (X. Dai et al., 2015). Luminal subtypes are positive for expression of ER and/or PR, the HER2-enriched subtype is characterized by overexpression of the HER2 RTK, whereas the triple-negative subtype is negative for these three receptors (Neve et al., 2006). Receptor status is mostly assessed by immunohistochemistry (IHC) and *ERBB2* (the gene encoding for HER2) amplification is verified by fluorescence in situ hybridization (FISH). Status for ER, PR and HER2 stratifies the patient according to the subtype of the disease and indicates the treatment strategies combined with clinical indicators such as tumor size, grade and lymph node status. A number of studies have investigated large patient cohorts with the use of state-of-the-art large-scale screening technologies to identify perturbations in genomic, epigenomic and proteomic levels that drive breast cancer initiation and metastasis. These studies have revealed intrinsic subtypes within the classified molecular subtypes that require more precise treatment approaches (Lehmann et al., 2011, 2016; Rueda et al., 2019)

1.1.1 Molecular Subtypes of Breast Cancer

In 2001, Sørlie *et al* published their study which paved the way for classifying breast carcinomas by identifying five molecular breast cancer subtypes (Luminal A, Luminal B, HER2-enriched, Basal, Normal-like) with different clinical prognosis depending on gene expression patterns (Sørlie et al., 2001). This subtype classification has been re-identified with some moderations by

different gene signatures by different groups such as PAM50, MammaPrint and OncoTypeDX (Bernard et al., 2009; Cristofanilli, 2005; Mook et al., 2007).

Luminal

Luminal breast cancer is the most common subtype. Nuclear estrogen receptor-α (ERα) is overexpressed in the majority (up to 70%) of breast tumors. Luminal tumors are sub-classified into Luminal A and Luminal B.

Luminal A tumors are characterized as ER-positive (ER+) and/or PR-positive (PR+), HER2-negative with low Ki67 (fraction of Ki67-positive tumor cells<14%), a marker for proliferation index. Luminal B tumors also express ER (ER+) and/or PR (PR+), however can be either HER2-negative with high Ki67 (≥14%) or HER2-positive with any Ki67 (Fulawka & Halon, 2017; Gnant et al., 2011). Luminal A breast tumors are generally lower-grade with slow growth tendencies compared to luminal B and other subtypes. Luminal B subtype has typically higher proliferative index compared to luminal A.

Luminal A subtype is characterized by the elevated expression of ER-responsive genes such as *ESR1, FOXA1, GATA3, XBP1*. *GATA3, FOXA1* code for transcription factors and are termed as 'pioneer factors' facilitating the transcription of ER-responsive genes by binding to condensed chromatin and changing the chromatin architecture allowing other TFs to bind (Manavathi et al., 2014). Genomic mutations of common pathway related genes are often observed in Luminal A subtype such as *MAP3K1, MAP2K4* of RAS/RAF/MEK/ERK and JNK stress-response signaling pathways, respectively (Koboldt et al., 2012). Mutations in tumor suppressor *TP53* may also contribute to aberrant growth of these tumors. *TP53* mutation rate is higher in luminal B subtype compared to luminal A, providing further explanation for their more aggressive phenotype (Bertheau et al., 2013; Koboldt et al., 2012).

Given their ER+ status and their dependency on estrogen for growth, the standard therapy for the luminal subtypes is endocrine therapy. The luminal A subtype has the best prognosis among all molecular subtypes in terms of overall survival. Patients affected by the luminal B subtype also benefit from endocrine therapy, however, require supplemental chemotherapy. Still, they exhibit less favorable overall prognosis compared to patients with luminal A BCa.

HER2-enriched

Her2-enriched subtype is hormone receptor negative and typically characterized by *ERBB2* gene amplification and HER2 protein overexpression. These tumors tend to have higher grade and patients have worse clinical prognosis compared to those presenting with the luminal A subtype. HER2 is a member of the epidermal growth factor receptor (EGFR) family. Overexpression of HER2 renders these tumors susceptible to targeted therapies against this receptor tyrosine kinase. Discovery of several drugs targeting HER2 increased the survival rate of the patients suffering from Her2-enriched subtype of breast cancer (Slamon et al., 2001). These drugs are mostly comprised of monoclonal antibodies targeting HER2 homodimers and HER2 heterodimers with other EGFR family members (i.e. EGFR, HER3, HER4) such as trastuzumab and pertuzumab, respectively (Ataseven, 2012). However, resistance to these approaches resulted in modification of existing drugs by conjugating them with cytotoxic components, as seen in trastuzumab emtansine (T-DM1) case, which is basically an antibody-drug conjugate between trastuzumab and DM1 (Verma et al., 2012).

Triple Negative

Triple negative subtype (TNBC) is characterized as ER-negative, PR-negative and HER2-negative, based on IHC stainings. Basal-like and triple negative breast cancers are often used to define the same molecular subtype as TNBC is largely comprised of basal-like (BL) tumors. However, Prat *et al* nomenclature subdivides TNBCs into BL and non-basal-like (NBL) (Prat et al., 2013). In short, not all TNBCs are basal-like (although the majority are) and not all basal-like tumors are TNBC. A recent sub-classification of the TNBC subtype by revealed four sub-subtypes: mesenchymal (M), basal-like 1 and basal-like 2 (BL1 and BL2), and luminal androgen receptor (LAR) (Lehmann et al., 2016). Regardless of various different nomenclatures, TNBC tumors are highly aggressive and metastatic. High frequency for *PI3KCA, BRCA1/2* and/or *TP53* mutations are observed for this subtype (L. Yang et al., 2019).

The lack of three targetable receptors renders this subtype ineligible for conventional targeted therapy approaches such as endocrine therapy or therapeutic antibodies, like trastuzumab. Patients mostly receive neo-adjuvant chemotherapy with taxanes and platinum salts. These

tumors are highly aggressive and due to the unavailability of a generalized targeted therapy, they have the worst prognosis compared to other molecular subtypes.

Sørlie *et al* nomenclature also defines normal-like breast cancer as an intrinsic molecular subtype which is characterized by elevated expression of genes related to non-epithelial tissues with reduced expression of luminal signature genes (Sørlie et al., 2001).

Treatment strategy for breast cancer can either be in neo-adjuvant or adjuvant setting in which therapy administration is either before or after surgery, respectively. Treatment approach is largely determined by the molecular profile of the tumor. Endocrine therapy, targeted therapy or chemotherapy constitute standard systemic treatment regimen.

1.1.2 Endocrine Therapy

Endocrine therapies are commonly applied to treat ER+ patients and come in mostly three ways: Tamoxifen is a selective ER modulator (SERM) and blocks estrogen-induced ERα activation by competitive inhibition of estrogen binding. Fulvestrant is a selective ER downregulator (SERD) that leads to degradation of the receptor. Aromatase inhibitors (AIs) block synthesis of estrogen, depriving ER from its ligand estrogen (Ali et al., 2011).

Since its approval by the US Food and Drug Administration (FDA) in the 1970s, tamoxifen has been the backbone of endocrine therapy lowering the breast cancer mortality rates considerably (Smith, 2014). It has been widely used to treat early breast cancer in pre-menopausal females in a 5 year long adjuvant setting (Yeo et al., 2014). Aromatase inhibitors are the choice of endocrine therapy for post-menopausal women. Aromatase is a monooxygenase that converts the steroid testosterone into estrogen. Determining factor for the choice of endocrine therapy agent is the status of ovarian function. Ovaries are the main source of estrogen in pre-menopausal women whereas estrogen production takes place primarily in adipose tissue, bones and muscle after menopause catalyzed by aromatase enzyme in both states. AIs are not effective in pre-menopausal women. This can be explained by the high level production of aromatase in pre-menopausal ovaries. In fact, the exposure to AI brings about a decrease in estrogen produced by peripheral tissues which leads to an increase in gonadotrophin secretion. Secreted gonadotropin, in turn leads to stimulation of ovaries to induce production of more estrogen. In short, AI

administration leads to even increased estrogen levels in pre-menopausal women (Scharl & Salterberg, 2016). AIs are given to pre-menopausal women with the addition of gonadotropin releasing hormone (GnRH) agonists/antagonists (Zelnak & O'Regan, 2015). Tamoxifen treatment can also be accompanied by ablation of ovarian function with gonadotropin releasing hormone (GnRH) agonists/antagonists in pre-menopausal patients (Robertson & Blamey, 2003).

With the introduction of aromatase inhibitors (AIs) such as anastrozole, examestane, letrozole in 1990s, they constitute the first line of therapy for post-menopausal patients or second line therapy for relapsing/non-responding tumors in pre-menopausal women (Rugo et al., 2016). Aromatase inhibitors can be subclassified into two classes: non-steroidal ones which bind reversibly to aromatase and steroidal ones which binds irreversibly (Schneider et al., 2011). Examestane is the sole representative of the steroidal AI among third generation AIs. Switch between non-steroidal to steroidal AIs occur in case of recurrence depending on prior therapy. Fulvestrant is commonly used as second line therapy for resistant tumors (Chia et al., 2008).

De-regulation of cell cycle and activation of alternative growth pathways such as PI3K/AKT/mTOR are major hallmarks of endocrine therapy resistance and disease relapse (Araki & Miyoshi, 2018; A. Gao et al., 2018). Therefore, metastatic or relapsed luminal breast cancers can also be treated with cell cycle inhibitors such as CDK4/6 inhibitor palbociclib as an additional therapy (Cristofanilli et al., 2016; X. Fu et al., 2013; Lu, 2015; Pang et al., 2014) Everolimus, an mTOR inhibitor, provides another alternative to overcome endocrine therapy resistance. It is used in combination with examestane in cases of non-responsive non-steroidal AI treatments (Gnant et al., 2013)(Liedtke & Kolberg, 2016) (Figure 1).

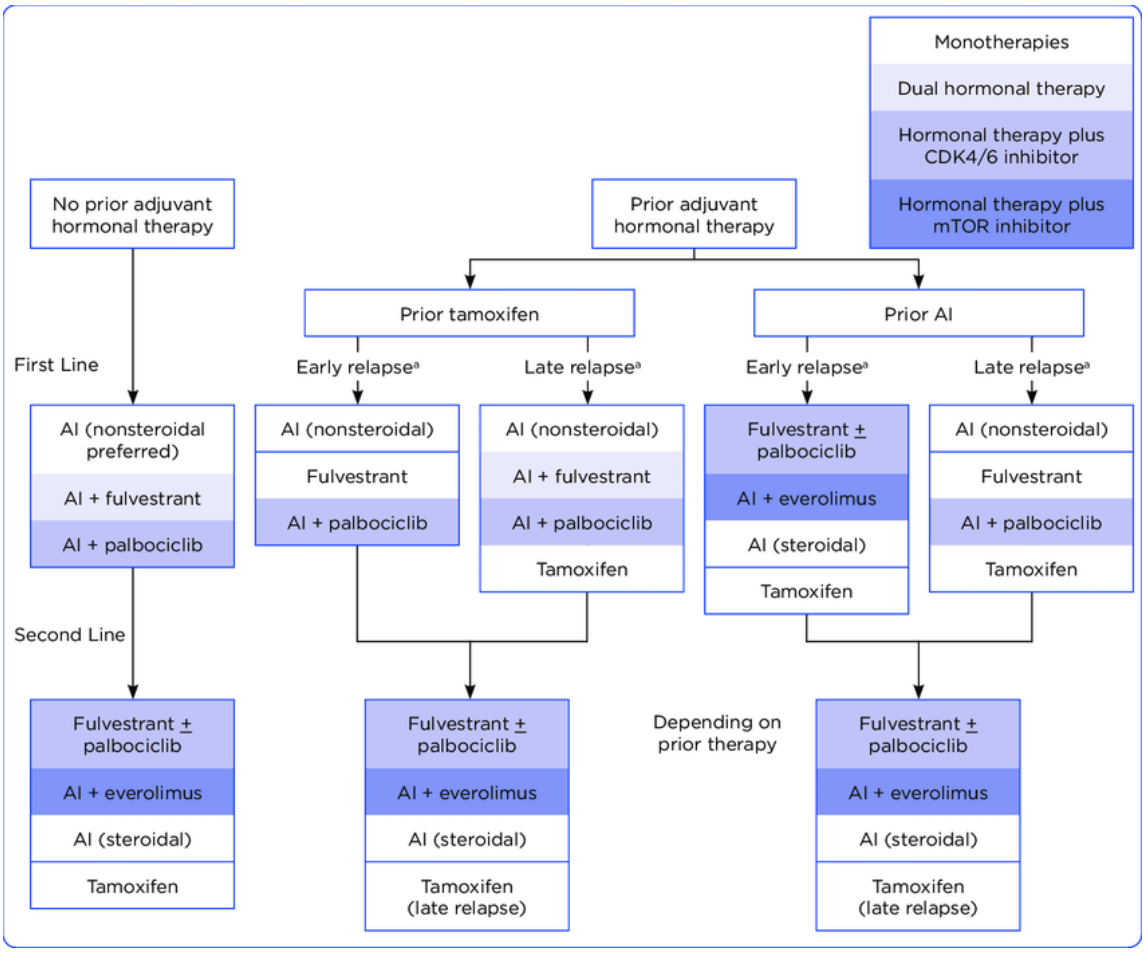


Figure 1: Treatment guidelines for endocrine therapy of ER-positive breast cancer in post-menopausal women. Modified from American Society of Clinical Oncology Guidelines (Rugo et al., 2016). AI:aromatase inhibitor, (a) Early relapse: ≤12 months after adjuvant endocrine therapy; late relapse: > 12 months after endocrine therapy AI:aromatase inhibitor.

1.1.3. Endocrine Therapy Resistance

Although luminal A subtype has better prognosis when compared to triple-negative breast cancer, approximately half of the patients do not respond or relapse due to *de novo* or acquired resistance against endocrine therapy, respectively (Early Breast Cancer Trialists' Collaborative Group (EBCTCG) et al., 2011; Osborne & Schiff, 2011) Occurrence of resistance is a limiting drawback for long-term efficacy of endocrine therapy and leads to poor prognosis 10-20 years after first diagnosis and treatment. Development of resistance against endocrine therapy is a long-term process where genetic and epigenetic changes play a predominant role (Badia et al., 2007; García-Becerra et al., 2012) (Figure 2).

Resistance to hormone therapy may occur due to many factors:

- Aberrant regulation of cell cycle or cell survival molecules
- Mutations of *ESR1* (Hyperactive estrogen signaling)
- Aberrant regulation of ER coactivators/corepressors
- Loss of ERα expression due to (epi)genetic modifications
- Alternative growth pathways (Increased estradiol-independent estrogen signaling)

De novo endocrine therapy resistance occurs primarily due to loss or modification of ERα expression. This can occur via epigenetic repression (methylation of CpG island or histone deacetylation) on *ESR1* expression. It has been shown that use of HDAC and DNMT1 inhibitor resulted in re-expression of *ESR1* in ER-negative breast cancer cell lines and rendered them susceptible to endocrine therapy (J. Fan et al., 2008; Milani, 2014).

Deficiency of cytochrome P450 2D6 (CYP2D6) in the liver has been proposed as an alternative mechanism for intrinsic resistance. *CYP2D6* encodes the enzyme that is responsible for metabolizing tamoxifen to 4-hydroxy-tamoxifen (4-OHT) which is by about 100 fold more potent to inhibit ER. Testing of the CYP2D6 status has become a companion diagnostic for luminal BCa patients as this status determines the dose of tamoxifen that needs to be administered. CYP2D6 also catalyzes formation of 4-hydroxy-*N*-desmethyltamoxifen (endoxifen), active metabolite of tamoxifen with high affinity to ER (Desta, 2004). Lacking this enzyme renders these patients refractory to tamoxifen treatment (Hoskins et al., 2009).

Mutations in the *ESR1* locus can also be accounted for the refractory state of the tumor to endocrine therapy. However loss of function mutations are found in less than 1% of primary breast tumors (Riggins et al., 2007). Activating *ESR1* mutations, which mostly occur in the ligand binding domain may result in hyperactive ER signaling even in the absence of the ligand and have been found mostly in relapsed tumors (W. Fan et al., 2015). Atypical expression levels of ER-coregulators have also been shown to confer therapy resistance as under-expression of co-repressors and/or over-expression of co-activators would contribute to augmented ERα signaling (Zhou et al., 2007).

De-regulation of the cell cycle caused by overexpression of positive regulators (e.g. Myc or Cyclin D1) or by a reduction of cyclin-dependent kinase (Cdk) inhibitors (e.g. p21, p27) might

enable tumor cells to evade cytostatic effects exerted by endocrine therapy drugs (Chu et al., 2008; Jansen et al., 2005; Span et al., 2003).

Activation of alternative growth factor pathways may bypass the effect of estrogen-dependent growth of tumor cells, thereby leading to endocrine therapy resistance. ERα is shown to be activated by growth factor receptors EGFR, HER2 and IGF1R in a ligand-independent fashion (Schiff et al., 2004). This leads to activation of PI3K/AKT/mTOR and MAPK downstream signaling (DeGraffenried et al., 2004; Faridi et al., 2003; Musgrove & Sutherland, 2009). *ERBB2* amplification, loss of *PTEN*, activating mutations of PI3K can account for aberrant activity of these signaling pathways (Arpino et al., 2008; Riggins et al., 2007).

According to the literature, alteration of metabolism (e.g. cholesterol biosynthesis) and activation of alternative signaling pathways (e.g. Notch and Wnt/β-catenin signaling pathway) can also contribute to development of endocrine resistance and have intricate cross-talk mechanisms in this process (Magnani et al., 2013; Nguyen et al., 2015; Riggins et al., 2007)

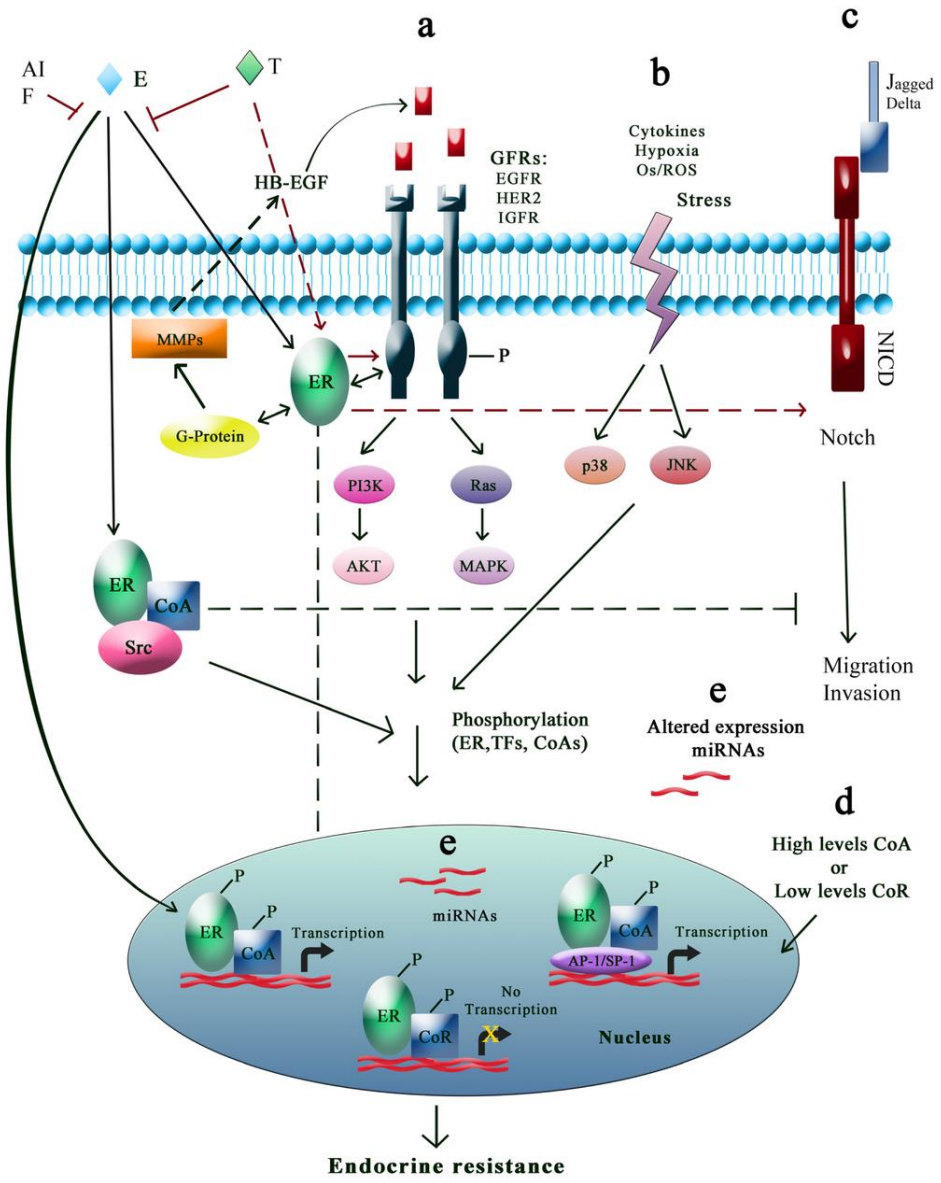


Figure 2: Mechanisms of endocrine therapy resistance. Schematic representation of factors playing a role in acquired endocrine therapy resistance (García-Becerra et al., 2012).

It has been recently shown that tumor cells having acquired a selective advantage against endocrine therapy do not cease proliferating due to genome-wide epigenetic reprogramming (Magnani et al., 2013) (S. V. Sharma et al., 2010) (Nguyen et al., 2015). Epigenetic reprogramming refers to any alteration in the epigenome such as DNA methylation, histone modification and chromatin compaction. Although there are epigenetic drugs available for endocrine resistance-reversal, they lack specificity and act on the whole genome (Cherblanc et al., 2012) (Steele et al., 2009). Identification and targeting of specific (epi)genetic elements that

confer resistance in tumors and discovery of epigenome-based biomarkers to predict prognosis and treatment outcome are fundamental to combat endocrine-resistance.

1.2 Epigenetics

Epigenetics can be described as the study of the heritable, yet reversible, modifications on chromatin that alter gene expression, with no change in genomic content (Marx, 2012). Three fundamental epigenetic mechanisms include non-coding RNAs, histone modifications and DNA methylation (Allis & Jenuwein, 2016).

1.2.1 DNA methylation

DNA methylation is defined as addition of a methyl group to the fifth carbon of cytosine residue which results in 5-methyl-cytosine (5mC). The methyl group donor is typically *S*-adenosyl methionine (SAM). This reaction is catalyzed by DNA methyltransferases (DNMTs). In eukaryotes, DNMT3a and DNMT3b are *de novo* DNMTs that can introduce methylation to unmethylated DNA whereas DNMT1 is responsible from methylation maintenance during DNA (Moore et al., 2013; Okano et al., 1999; Suetake et al., 2004; S. Xie et al., 1999) DNMT3L, although devoid of a catalytic domain, can interact with *de novo* DNMTs and enhance their methyltransferase activity (Aapola et al., 2000; Hata et al., 2006; Jia et al., 2007; Suetake et al., 2004).

Genomic imprinting, X chromosome inactivation and tissue-specific gene expression regulation is established by DNA methylation. DNA methylation is also important for silencing of retroviral and transposable elements that could jeopardize genomic integrity unless inactivated (Michaud et al., 1994; W. Xie et al., 2012)

DNA methylation exerts its effect on the genome in a region-dependent manner. CpG islands are DNA stretches that are 500-1500 bp long with a higher CG ratio compared to the rest of the genome. CpG islands typically contain most of the gene promoters and are often not methylated (Illingworth et al., 2010; Saxonov et al., 2006). A recent publication stated that when combined with particular TF motifs the presence of high CpG abundance within CpG islands is enough to enhance transcriptional activity independent of methyltransferase activity (Hartl et al., 2019)

Methylation of CpG islands leads to gene repression (Mohn et al., 2008). Methylation of CpG islands located in gene promoters mostly exhibit tissue specificity unlike intragenic or gene body CpG islands (Eckhardt et al., 2006; Ghosh et al., 2010; Maunakea et al., 2010; Meissner et al., 2008; Rakyan et al., 2004). Tissue specific methylation is also observed in CpGs located 2 kb upstream or downstream of CpG islands, collectively termed as CpG island shores, where methylation is inversely correlated with gene expression (Irizarry et al., 2009). On the other hand, gene body methylation correlation with gene expression is more complicated. In methylation nomenclature, gene body is defined as the region following the first exon (Brenet et al., 2011). Elevated gene expression is observed for dividing cells after gene body methylation (Aran et al., 2011; Ball et al., 2009; Hellman & Chess, 2007). No correlation was found between expression of a gene and methylation of its gene body in dormant slowly dividing cells (Guo et al., 2011). There has been recent studies suggesting gene body methylation is closely associated to alternative splicing and loss of gene body methylation leads to aberrant splicing (Flores et al., 2012; Maunakea et al., 2013, 2010)

The interplay between DNA methylation and histone modifications dictate the chromatin accessibility for DNA binding proteins such as transcription factors thereby permitting or hindering transcription of a gene. For example, H3K4 tri-methylation (H3K4me3), a permissive histone mark, impede binding of DNMTs. Histone marks associated with closed chromatin such as H3K9 or H3K27 methylation attract or favor DNMT binding, thereby establishing a more inaccessible chromatin landscape (Vaissière et al., 2008).

DNA demethylation can arise either in an active or passive manner. Passive demethylation is closely related with the lack of or a decrease in DNMT activity. It is associated with DNA replication where methylation marks are lost with increasing number of cell cycles in the absence of maintenance methylation. Active demethylation can be orchestrated by three enzyme families. Ten-eleven translocation (TET) enzymes catalyze hydroxylation of 5mC resulting in 5-hydroxymethylcytosine (5hmC) which can be further oxidized to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC). AID/APOBEC family of enzymes catalyzes deamination of 5mC or 5hmC to 5-methyluracil (5mU) or 5-hydroxymethyluracil (5hmU), respectively. 5mU, 5hmU or 5caC are further catalyzed by TDG/SMUG1 (base excision repair glycosylases) to form cytosine,

therefore concluding active DNA demethylation (Bhutani et al., 2011; Kohli & Zhang, 2013) (Figure 3).

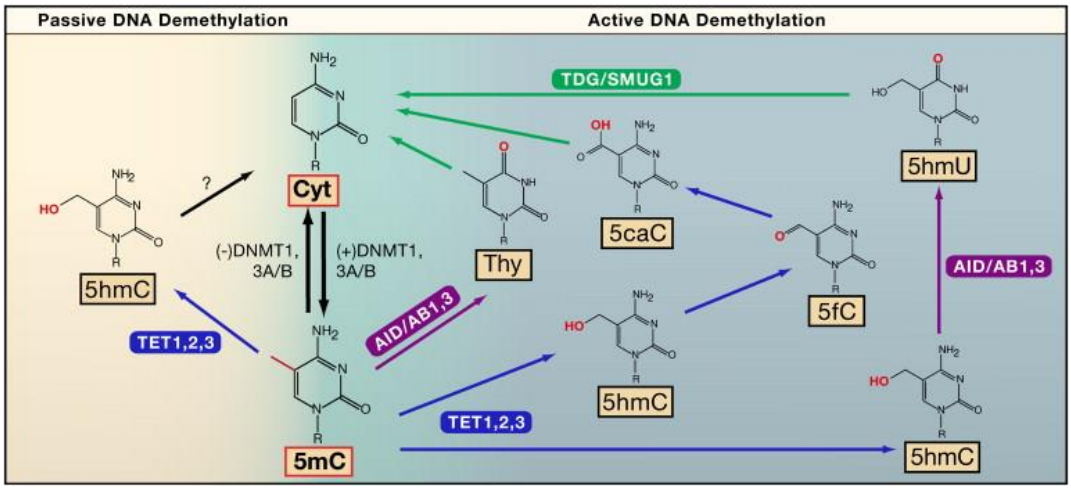


Figure 3: Mechanisms of DNA demethylation. Schematic representation of active and passive DNA demethylation (Bhutani et al., 2011).

1.2.2 Histone modifications

Post-translational modifications of histones are diverse in many ways. These modifications can be present on all five histones: H2A, H2B, H3, H4 and the linker histone H1. They are commonly introduced at the N-terminal tail of histone proteins. Several types of post-translational modifications of histones have been reported, including acetylation, methylation, phosphorylation, ubiquitination, succinylation, sumoylation, citrunnilation, propionylation, butyrylation, formylation and ADP ribosylation (Lennartsson & Ekwall, 2009; Ruthenburg et al., 2007). Moreover, the amino acid residues that are modified may vary. Some residues may only allow one type of modification while others are prone to several. For instance, while phosphorylation is the only mark that is found on serine, threonine or tyrosine residues, lysine residues can be modified by acetylation, biotinylation, methylation, ribosylation and ubiquitination. Additionally, a methylation mark may be present in a mono-, di- or tri-methylated state (H. Dai & Wang, 2014). A recent study revealed 67 novel histone modifications such as lysine crotonylation by mass spectrometry which expands the histone code considerably (Castillo et al., 2017; Tan et al., 2011)

While DNA methylation is a rather stable modification, histone marks undergo constant change by epigenetic modifiers. The histone modifying enzymes that introduce the modifications are termed “writers”, while “erasers” are responsible for removing the marks. This epigenetic code is then read by the “readers”, which specifically bind to certain epigenetic marks and depending on the composition of the epigenetic landscape, either recruit repressive or activating complexes to the gene, thereby modifying gene expression (X. J. Yang & Seto, 2007). While the effect of some epigenetic marks is context dependent, meaning depending on the presence of other epigenetic modifications, histone acetylation is generally associated with transcriptional activation. Positive charge of lysine is neutralized by acetylation thereby the interaction between histone tail and negatively charged DNA is weakened which leads to opening up the chromatin structure and making it accessible for transcription factors or the transcription machinery. The composition of the histone modifications make up the so-called ‘histone code’ which pertains to regulation of gene expression via changes in the chromatin architecture (Carlberg et al., 2018; Turner, 2000, 2007)

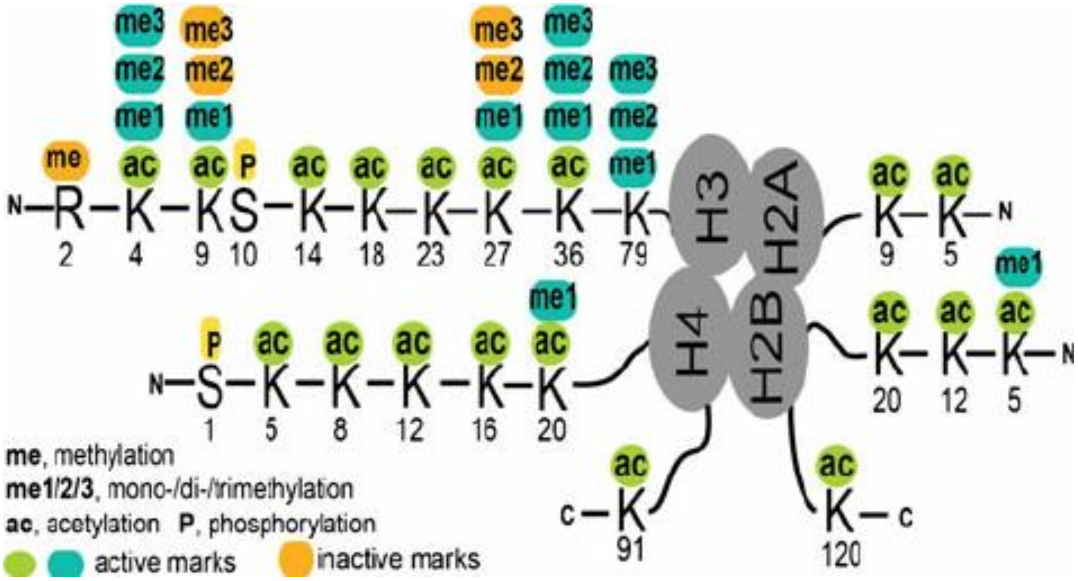


Figure 4: Histone modifications of core histones and histone code. Schematic depiction of modifications that may occur on core histone tails and their implications on activating or repressing gene expression (H. Dai & Wang, 2014).

1.2.2.1. Histone Acetyl Transferases (HATs) and Histone Deacetylases (HDACS)

Histone acetylation is mediated by Histone Acetyl Transferases (HATs). HATs can be either nuclear (Type A) or cytoplasmic (Type B) (Richman et al., 1988). Nuclear HATs are subdivided into five families. MYST HATs (MOF, MOZ, MORF, HBO1 and TIP60) typically contain a chromodomain and/or a PHD finger domain in addition to MYST domain, which is composed of zinc finger and acetyl CoA-binding domains (Avvakumov & Côté, 2007). Gcn5-related N-acetyltransferases (GNATs) family is composed of GCN5 (KAT2A) and PCAF (KAT2B) in humans. They are characterized by the presence of bromodomain in addition to HAT domain (Z. Nagy & Tora, 2007). p300/CBP (CREB-binding protein) HATs constitute another HAT family which are composed of two transcriptional co-activators. Harboring a chromodomain, bromodomain or PHD finger domain enables these enzymes to recognize and bind modified histone residues thereby rendering them as epigenetic "readers" adding to their designated epigenetic "writer" status due to HAT activity (Yun et al., 2011; T. Zhang et al., 2015). Transcription factor-related HATs (TFIIIC, TAF1) and nuclear receptor co-activators (SRC, ACTR, P160, CLOCK) are other HAT families (Doi et al., 2006; Torchia et al., 1998).

HATs (mostly as part of multisubunit HAT complexes) exert their catalytic activity on preferential histone lysine residues. As an example, GCN5/PCAF can acetylate H3K9 and H3K14 in SAGA or ADA complexes and MYST family member MOF acetylates H4K16 (Jin et al., 2011; Krebs et al., 2011; Lee & Workman, 2007; G. G. Sharma et al., 2010). Histone modifications are dynamic and transient processes. Abnormal histone acetylation patterns are closely related to malignancies such as cancer (Di Cerbo & Schneider, 2013).

Post-translational acetylation of a histone lysine residue can be deacetylated by a histone deacetylase (HDAC) (Seto & Yoshida, 2014). Reversible nature of this modification results in transcriptional regulation. Being devoid of DNA-binding domain, HDACs can interact with their targets through binding to coactivators or repressors of transcription (Wang et al., 2009) Hence, the repertoire and range of the available interaction partners for HDACs determine their effect on gene expression regulation (Turner, 2000). Histone deacetylation is commonly associated with closed chromatin and reduced gene expression. However, there is growing evidence suggesting HDACs might also play a role in transcriptional initiation and progression via 'erasing'

acetylation marks which can obstruct elongation (C. B. Greer et al., 2015; Nusinzon & Horvath, 2005)

There are 11 classical HDACs which are further subdivided into four families: Class I, IIa, IIb and IV. They are differing from each other in function, structure and expression. Class III HDACs (sirtuins) deviate from classical HDACs since they depend on nicotine adenine dinucleotide (NAD^+) for their enzymatic activity (Haberland et al., 2009; RUIJTER et al., 2003).

HDACs are notoriously de-regulated in cancer. HDACs are generally found to be elevated in tumorigenesis. Class I HDACs are associated with increased proliferative and invasive phenotype. Breast, prostate and gastrointestinal cancers have reportedly elevated levels of HDAC1 (Choi et al., 2001; Halkidou et al., 2004; Nakagawa et al., 2007). HDAC inhibitors have been used to combat cancer entities of differing origin in pre-clinical and clinical trials in an attempt to attenuate proliferation, angiogenesis and invasion of tumor cells.

Inhibition of HDACs results in cell cycle arrest and apoptosis *in vitro* due to increased levels of p21 and induction of proapoptotic genes (*Bak, Bax*), respectively (Marchion & Münster, 2007). Efficacy of HDAC treatment in humans is limited. They are effective in abrogating proliferation, angiogenesis, invasion and inflammation. However the drawback is their non-specificity. Anti-angiogenesis effect of HDACs diminishes their delivery into solid tumors. Anti-inflammatory effect of HDACs depletes also the immune cells, therefore anti-tumor immune response (Ropero & Esteller, 2007; Somech et al., 2004).

1.2.2.2. Histone methyltransferases and Histone demethylases

Histone methyltransferases mediates methylation of lysine or arginine residues on histone tails. Depending on the methylated amino acid they can be divided into two groups: protein lysine methyltransferases (PKMT) and protein arginine methyltransferases (PRMT) (Dillon et al., 2005; Feng et al., 2002; Lachner & Jenuwein, 2002) PKMTs can be subdivided into SET (Su(var)3–9, Enhancer of Zeste, and Trithorax) domain containing and non-SET domain containing enzymes (Dillon et al., 2005). Much like in DNA methylation, SAM provides the methyl group for HMTs. There are two methylation states (mono-, di-methylation) for arginine while lysine residues can

be either mono-, di- or tri-methylated (Bannister & Kouzarides, 2011; Tiedeken & Chang, 2015) Unlike lysine acetylation, lysine methylation does not alter the side chain's positive charge. For this reason, the effect of lysine methylation is generally context-dependent, meaning it depends on the position and state of methylation and the molecules that 'read' the histone code 'written' by HMTs (Martin & Zhang, 2005). Methylation of H3K4, H3K36 and H3K79 are histone marks associated with active transcription. H3K9, H3K27 and H4K20 methylation, on the other hand, comprise a repressive histone code (Hyun et al., 2017). Bivalent chromatins are characterized by their occupancy with opposing histone marks such as H3K4 methylation (activating) and H3K27 methylation (repressing). Bivalency is crucial for maintaining poised state of transcription ready for activation or inactivation in stem cells (Vastenhouw & Schier, 2012).

Methyl group removal from histone tail residues is catalyzed by histone demethylases. The discovery of jumanji protein domain (JmjC) responsible for demethylase activity led to an expansion in histone demethylases (Agger et al., 2008; Tsukada et al., 2006)

Aberration in histone methylation patterns have been closely related to many maladies. That led to the consideration of histone methyltransferases, histone demethylases or methyl-lysine binding proteins as therapeutic targets (E. L. Greer & Shi, 2012; Song et al., 2016).

Figure 5 provides a general overview of the ‚writers' and ‚erasers' of the histone methylation marks in different organisms, histone methyl transferases and histone demethylases, respectively (Hyun et al., 2017).

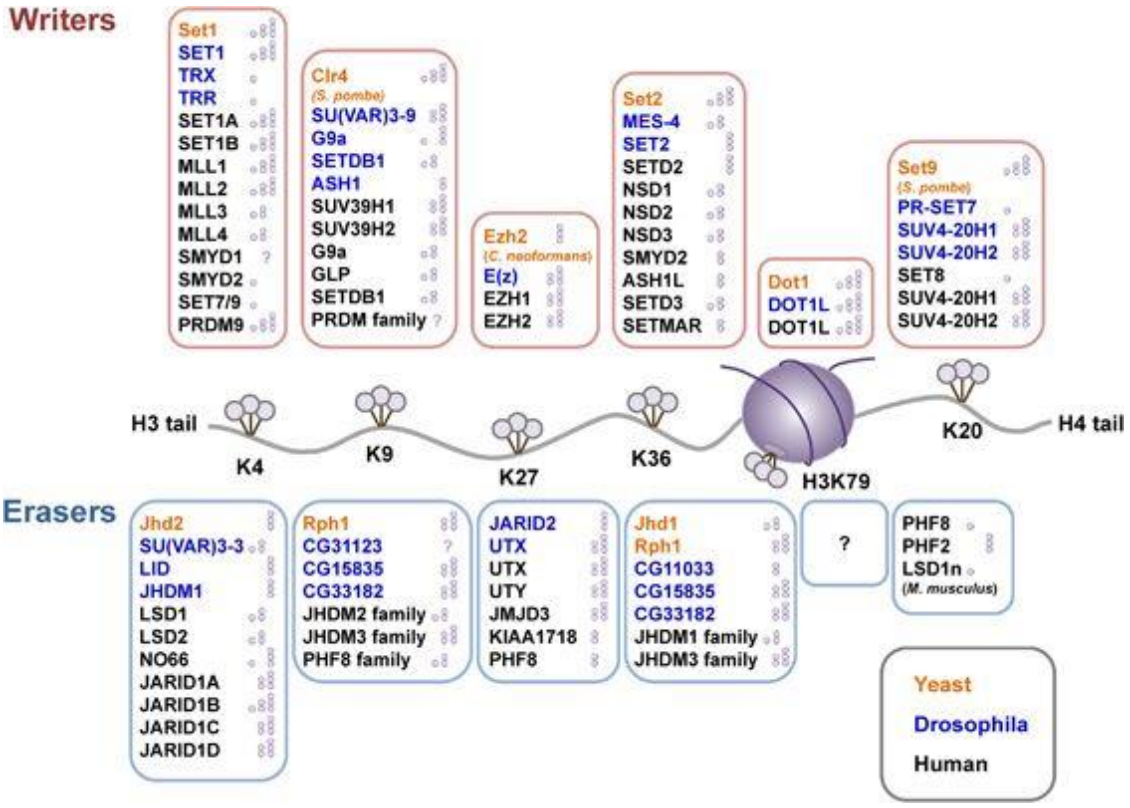


Figure 5: Histone methyltransferases and demethylases. Schematic representation of ´writers´ and ´erasers´ of histone methylation in different organisms with their respective choice of histone tail residues (Hyun et al., 2017).

1.3 CRISPR/dCas9-mediated Epigenetic Editing

The Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/CRISPR associated (Cas) system was introduced by Jennifer Doudna and Emmanuelle Charpentier in 2012 as one of the most ground-breaking scientific methods in recent years, and has since then revolutionized research in genome engineering (Doudna & Charpentier, 2014). It was discovered much earlier, as part of the research on the prokaryotic immune system, where this method is used as an adaptive defense mechanism against viruses or other foreign DNA (Horvath & Barrangou, 2010).

In bacteria, foreign DNA is at times fragmented and integrated as a spacer between short palindromic repeats in the CRISPR locus of the bacterial chromosome. The locus is then

transcribed into crRNA and associates with a so-called tracrRNA to form the guide RNA (gRNA). The gRNA will then in turn form a complex with a Cas protein. In case of invasion of the same foreign DNA, this complex is then guided by the crRNA spacer sequence being complementary to the foreign DNA. The complex binds and cleaves the target DNA through the endonuclease activity of the Cas protein in order to diminish foreign invasion (Jinek et al., 2012).

Doudna and Charpentier have engineered a two component version of this system, that uses Cas9 endonuclease and a fusion of the two RNA components, termed single guide RNA (sgRNA) (Sternberg & Doudna, 2015). In genome engineering, this system is utilized to introduce targeted double strand breaks at a genetic region of interest, guided by the choice of sgRNA. It makes use of the cellular repair mechanisms, which will repair the double strand break in one of two ways. Using non-homologous end joining (NHEJ), the break is ligated without a template, resulting in the occurrence of insertions or deletions (indel) (Figure 6a). Often, premature stop codons or frame shifts appear due to these indels, thereby causing a functional knock out of the gene. The other repair mechanism, homology directed repair (HDR) uses a template, such as the second chromosome or a template DNA, to repair the break (Gilbert et al., 2013; Knott & Doudna, 2018; Sternberg et al., 2014). By specifically adding a repair template to the cell, one is able to modify a genetic region of interest in a targeted fashion

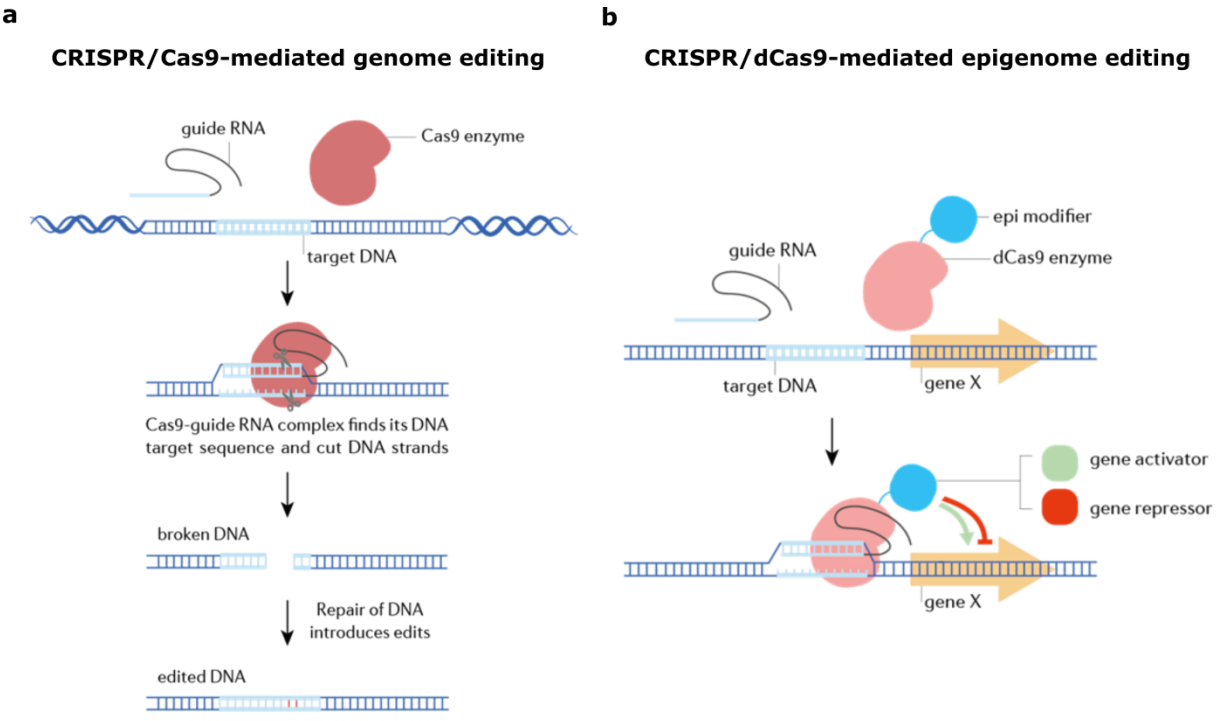


Figure 6: Two distinct applications of CRISPR/Cas9 system. (a) Genome editing by CRISPR/Cas9 system where the targeted region of the genome is cut and repaired via NHEJ, introducing indels, thereby resulting in edited genome. (b) Epigenetic editing via CRISPR/dCas9, where catalytically dead Cas9 (dCas9) is fused to catalytic domains of epigenetic effectors and recruited to targeted regions of the genome to modulate the epigenetic landscape, thereby altering transcriptional status of target genes.

The specificity and activity of the CRISPR/Cas9 system are conveyed by two elements, the target sequence and the Protospacer adjacent motif (PAM) sequence. In the sgRNA, the target sequence is a 20 nucleotide sequence that is complementary to the region of interest and guides the sgRNA/Cas9 complex to the targeted genomic locus. It is followed by the two to six nucleotide PAM sequence, which is specific for every Cas9 protein and essential in facilitating the binding of Cas9 to the DNA and enabling its endonuclease activity. The PAM sequence for Cas9 derived from *Streptococcus pyogenes* is NGG, with N being any base followed by two guanines (Kleinstiver et al., 2015; Knott & Doudna, 2018).

Since the ground-breaking innovation of tailored genetic editing systems, like TALENs, Zinc Finger Nucleases (ZFNs) or the CRISPR/Cas9 system, research has concentrated on also utilizing these systems for other applications. Recently, CRISPR/dCas9 system has been repurposed for epigenetic editing of targeted region of the genome (Brocken et al., 2018; Vojta et al., 2016). It makes use of the CRISPR/Cas9 system's advantage of specifically targeting a region of interest in the genome. The Cas9 nuclease here is catalytically dead (dCas9), and instead fused to effector domains of epigenetic enzymes (De Groote et al., 2012; X. Gao et al., 2014) (Figure 6b). This approach is used to modify the epigenome at a specific locus, eg. a gene regulatory element, and thereby altering the expression of a gene of interest.

2. Aims

The main focus of this project has been the identification of fundamental epigenetic changes that bring about endocrine therapy resistance and interfering with the epigenome to mimic resistance or achieve resistance-reversal and re-sensitization to endocrine therapy. This project has been part of EU funded EpiPredict Consortium (Horizon 2020). The research questions have been tackled in two parts in this thesis:

PART I: Elucidating the Role of GLYATL1 in Endocrine Therapy Resistant Breast Cancer

The aim of this part is to identify novel target gene(s) which confer resistance to endocrine therapy via:

- profiling the changes in transcriptomics (RNA-Seq), methylome (EPIC array) and chromatin accessibility (ATAC-Seq) in resistant cell lines (MCF7 and T47D) compared to their sensitive counterparts
- combining large scale screening techniques such as RNA-Seq and ATAC-Seq to reveal commonalities between distinct therapy resistance models has identified novel target *GLYATL1*
- validating the contribution of *GLYATL1* to endocrine therapy resistance development
- elucidating (epigenetic) regulation of *GLYATL1*
- utilizing established CRISPR/dCas9- mediated epigenetic editing approach in order to validate epigenetic regulation of *GLYATL1*.

PART II: Establishing CRISPR/dCas9-mediated Epigenetic Editing Methodology to Modulate the Epigenetic Landscape of Endocrine Therapy Resistant Breast Cancer

The aims of this part are:

- to establish innovative CRISPR/dCas9-mediated epigenetic editing tools and to modulate expression of potentially endocrine therapy resistance related genes.
- to elucidate the involvement of the epigenetic landscape and expression of *SLC9A3R1*, *BAMBI* and *CD44* in the development and maintenance of endocrine therapy resistance in the parental and resistant cell lines

3. Materials and Methods

3.1 Materials

3.1.1 Instruments

Bacterial Incubator (37°)	Memmert (Schwabach, Germany)
Bacterial shaking Incubator (37°)	INFORS HAT (Surrey, UK)
Biohit Proline multichannel pipette	Sartorius (Göttingen, Germany)
Cell culture hood HERA Safe	Thermo Fisher Scientific (Waltham, USA)
Cell culture incubator	Heraeus (Hanau, Germany)
Centrifuges	Eppendorf AG (Hamburg, Germany)
Extracellular flow bioanalyzer (Seahorse XF96)	Agilent Technologies (Santa Clara USA)
Flow Cytometer FACS Canto II	Becton Dickinson (New Jersey, USA)
Flow Cytometer LSRFortessa™	Becton Dickinson (New Jersey, USA)
Fluorescent microscope Axiovert 40 CFL	Carl Zeiss (Oberkochen, Germany)
Freezer (-20°)	Liebherr (Bulle, Switzerland)
Freezer (-80°)	Sanyo (Osaka, Japan)
Fridge (+4°)	Liebherr (Bulle, Switzerland)
Gel documentation system	Herolab GmbH (Wiesloch, Germany)
Glomax explorer plate reader	Promega (Wisconsin, USA)
Molecular Devices Microscope IXM XLS	Molecular Devices (Sunnyvale, USA)
Nanodrop ND-1000 spectrophotometer	Thermo Fisher Scientific (Waltham, USA)
Neubauer cell counting chamber	BRAND
Odyssey Infrared Imaging System	Li-Cor Biosciences GmbH (Bad Homburg, Germany)
Pipetboy acu pipette	INTEGRA Biosciences (Fernwald, Germany)
Pipetman® pipette	Gilson (Limburg, Germany)
Protein Gel Apparatus MiniProtean II	Bio-Rad (Hercules, USA)
SW41 Ti Rotor and Tubes	Beckman Coulter (Brea, USA)
Thermocycler	Applied Biosystems (Foster City, USA)
Thermomixer	Eppendorf AG (Hamburg, Germany)
Titramax 100 rocking platform	Heidolph (Schwabach, Germany)

Trans-Blot Turbo Transfer	Bio-Rad (Hercules, USA)
Tube Rotator	VWR (Darmstadt, Germany)
Ultracentrifuge, Beckman L8-70M	Beckman Coulter (Brea, USA)
Vacuboy aspiration device	INTEGRA Biosciences (Fernwald, Germany)
Vortex mixer	neoLab (Heidelberg, Germany)

3.1.2 Chemicals and Reagents

(Z)-4-Hydroxytamoxifen ≥98%	Sigma-Aldrich (Saint-Louis, USA)
0.25% Trypsin EDTA Solution	Gibco (New York, USA)
5-Aza-dC (5-Aza-2 deoxycytidine)	Abcam (Cambridge, USA)
Acetic Acid	Sigma-Aldrich (Saint-Louis, USA)
Acrylamide/bisacrylamide 37.5:1	Carl Roth (Karlsruhe, Germany)
Agar	Sigma-Aldrich (Saint-Louis, USA)
Ammoniumperoxodisulfate (APS)	Sigma-Aldrich (Saint-Louis, USA)
Ampicillin	Sigma-Aldrich (Saint-Louis, USA)
Blasticidin	Thermo Fischer Scientific (Waltham, USA)
BSA	Sigma-Aldrich (Saint-Louis, USA)
BsmBI restriction enzyme	New England Biolabs (Massachusetts, USA)
Charcoal Stripped Fetal Bovine Serum	Sigma-Aldrich (Saint-Louis, USA)
cOmplete Mini Protease Inhibitor Cocktail	Roche Diagnostics (Mannheim, Germany)
Competent *E.coli* (DH5α)	New England Biolabs (Massachusetts, USA)
Crystal violet	Sigma-Aldrich (Saint-Louis, USA)
DMEM medium	Gibco (New York, USA)
DMSO	Sigma-Aldrich (Saint-Louis, USA)
Doxycycline	Sigma-Aldrich (Saint-Louis, USA)
DPBS	Gibco (New York, USA)
EDTA	Sigma-Aldrich (Saint-Louis, USA)
Ethanol	Sigma-Aldrich (Saint-Louis, USA)
Fetal Bovine Serum	Gibco (New York, USA)
Geneticin	Gibco (New York, USA)

Hoechst 33258	Sigma (Krefeld, Germany)
Isopropanol	Greiner Bio-One International GmbH (Kremsmünster, Austria)
LB Broth	Sigma-Aldrich (Saint-Louis, USA)
L-glutamine, 200mM	Gibco (New York, USA)
Lipofectamine LTX	Thermo Fischer Scientific (Waltham, USA)
Lipofectamine RNAiMAX	Thermo Fischer Scientific (Waltham, USA)
Matrigel	Corning (Corning, USA)
Methanol	Greiner Bio-One International GmbH (Kremsmünster, Austria)
NaCl	VWR International (Darmstadt, Germany)
NaF	Bernd Kraft (Duisburg, Germany)
NaOH	Sigma-Aldrich (Saint-Louis, USA)
NEB Buffer #3, #3.1	New England Biolabs (Massachusetts, USA)
non-DEPC treated nuclease-free water	Thermo Fisher Scientific (Waltham, USA)
Opti-MEM medium	Gibco (New York, USA)
Paraformaldehyde 16%	Thermo Fisher Scientific (Waltham, USA)
PBS	Gibco (New York, USA)
Penicillin/Streptomycin (Pen/Strep)	Gibco (New York, USA)
PhosSTOP Phosphatase Inhibitor Cocktail	Roche Diagnostics (Mannheim, Germany)
primaQUANT 2x qPCR Probe Master Mix	Steinbrenner Laborsysteme (Weinbach, Germany)
Protein Marker Precision Plus Protein Dual Color	Bio-Rad (Hercules, USA)
Proteinase K	Sigma-Aldrich (Saint-Louis, USA)
Puromycin	Thermo Fischer Scientific (Waltham, USA)
RIPA Lysis and Extraction buffer	Thermo Fischer Scientific (Rockford,

	USA)
Rockland Blocking Buffer	Rockland Immunochemicals Inc. (Limerick, USA)
Roti®-Load 1, 4x sample loading buffer	Carl Roth (Karlsruhe, Germany)
SDS	Carl Roth (Karlsruhe, Germany)
siRNAs	Dharmacon, Thermo Fisher Scientific (Waltham, USA)
SOC Outgrowth Medium	New England Biolabs (Massachusetts, USA)
Sodium pyruvate, 100mM	Gibco (New York, USA)
T4 DNA Ligase	Thermo Fischer Scientific (Rockford, USA)
TEMED	Carl Roth (Karlsruhe, Germany)
Tris-base	Sigma-Aldrich (Saint-Louis, USA)
Triton X-100	Sigma-Aldrich (Saint-Louis, USA)
TrypLE Express Enzyme (1X, no phenol red)	Thermo Fischer Scientific (Rockford, USA)
Tween 20	Sigma-Aldrich (Saint-Louis, USA)
β-estradiol	Sigma-Aldrich (Saint-Louis, USA)

3.1.3 Assay Kits

BCA Protein Assay Kit PierceTM	Thermo Fisher Scientific (Waltham, USA)
CellTiter-Glo Luminescent assay	Promega (Wisconsin, USA)
Click-iT™ EdU Cell Proliferation Kit for Imaging, Alexa Fluor™ 647 dye	Thermo Fisher Scientific (Waltham, USA)
DNeasy Blood & Tissue Kit	Qiagen (Hilden, Germany)
Histone Extraction Kit	Abcam (Cambridge, USA)
Histone H3 acetyl 14 ELISA Kit (H3K14)	Active Motif (California, USA)
Histone H3 acetyl 9 ELISA Kit (H3K9)	Active Motif (California, USA)
Infimium MethylationEPIC BeadChip	Illumina (San Diego, USA)

NE-PER Nuclear Cytoplasmic Extraction Kit	Thermo Fisher Scientific (Waltham, USA)
primaQUANT 2x qPCR Probe Master Mix	Steinbrenner Laborsysteme (Weinbach, Germany)
Qiagen Plasmid Plus Midi Kit	Qiagen (Hilden, Germany)
RevertAid™ H Minus First Strand cDNA synthesis kit	Fermentas, Thermo Fisher Scientific (Waltham, USA)
RNeasy Mini Kit	Qiagen (Hilden, Germany)
Total Histone H3 ELISA Kit	Active Motif (California, USA)
Trans-Blot Turbo mini PVDF Transfer Kit	Bio-Rad (Hercules, USA)
Universal Probe Library (UPL)	Roche Diagnostics (Mannheim, Germany)
Wizard SV Gel and PCR Clean-up System	Promega (Wisconsin, USA)

3.1.4 Consumables

Micro centrifuge tubes (1,5,2 and 5 ml)	Eppendorf AG (Hamburg, Germany)
Petri dishes (100 mm and 150 mm)	Techno Plastic Products (TPP) AG (Trasadingen, Switzerland)
Falcon tubes (15 ml and 50 ml)	Becton Dickinson (New Jersey, USA)
Multi-well plates (6, 12, 24, 48-well), F-bottom, transparent	Thermo Fisher Scientific (Waltham, USA)
96-well plate (F-bottom, transparent)	Becton Dickinson (New Jersey, USA)
96-well plate (F-bottom, white)	PerkinElmer (Waltham, USA)
96-well plate (F-bottom, black)	Greiner Bio-One International GmbH (Kremsmünster, Austria)
96-well Clear V-Bottom Deep Well Plate	Corning (Corning, USA)
384-well plates for TaqMan	Applied Biosystems (Foster City, USA)
Adhesive Optically Clear Plate Seal	Thermo Fisher Scientific (Waltham, USA)

Cell Culture Flasks, T-25, T-75, T-175	Greiner Bio-One International GmbH (Kremsmünster, Austria)
Cell Scraper	Corning (Corning, USA)
Cryo vials 1.8mL	Nunc, Thermo Fisher Scientific (Waltham, USA)
FACS tubes	Becton Dickinson (New Jersey, USA)
Filter tips, 10µL, 20µL, 200µL, 1000µL	Neptune Scientific (San Diego, USA)
4-15% Mini-PROTEAN®TGX™ Precast Protein Gels (10,12 and 15-well)	Bio-Rad (Hercules, USA)
Oncyte® Avid Nitrocellulose Film-Slide	Grace Bio-Labs (Bend, USA)
PCR strips	Steinbrenner Laborsysteme GmbH (Wiesenbach, Germany)
PVDF membrane Immobilon-P	Merck Millipore (Darmstadt, Germany)
Reservoirs 50ml	Corning (Corning, USA)
Round-Bottom polypropylene tubes 5ml	Thermo Fisher Scientific (Waltham, USA)
Serological pipettes 5mL, 10mL, 25mL, 50ml	Becton Dickinson (New Jersey, USA)
Trans-well system (8.0µm pore size)	Corning (Corning, USA)
Whatman 3 MM filter paper	GE Healthcare (Little Chalfont, United Kingdom)

3.1.5 Software

BD FACSDiva™ Software	Becton Dickinson (New Jersey, USA)
GraphPad Prism 5	GraphPad Software, Inc. (La Jolla, USA)
FlowJo 10	Becton Dickinson
Image Studio	LI-COR Biosciences (Lincoln USA)
Inkscape	Software Freedom Conservancy, Inc. (NY USA)
Molecular Devices Analysis Software	Molecular Devices (Sunnyvale, USA)
Odyssey 2.1	LI-COR (Lincoln, USA)

QuantStudio	Applied Biosystems (Foster City, USA)
Roche UPL Design Center	Roche Diagnostics (Mannheim, Germany)
SDS 2.2	Applied Biosystems (Foster City, USA)

3.1.6 Database

METABRIC mRNA data	https://www.synapse.org/#!Synapse:syn1688369/wiki/27311
TCGA mRNA data	TCGA_BRCA_exp_HiSeqV2-2015-02-24
GSE1456 (Pawitan dataset)	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE1456

3.1.7 Buffers and solutions

EdU assay	
Fixation Buffer	4% PFA in ddH_2O
Permeabilization Buffer	0.5% Triton X-100 in PBS
Washing Buffer	3% BSA in PBS
Western Blotting	
Lysis Buffer	RIPA Buffer 10ml 1x tablet PhosSTOP Phosphatase Inhibitor 1x tablet Complete Mini Protease Inhibitor Cocktail
SDS-PAGE Running Buffer (1L)	0.25 M Tris Base (30.3 g) 1.92 M Glycine (144.1 g) 0.1% SDS (w/v)
Transfer Buffer	20% Trans-BlotR Turbo™ 5x Transfer Buffer 20% EtOH 60% ddH_2O
Blocking Buffer	Rockland blocking Buffer:TBS (1:1) 5mM NaF 1mM Na_3VO_4
Washing Buffer (1X TBST)	0.1% Tween®20 in TBS (TBST)
10X TBS	1.37M NaCl 200mM Tris pH 7.6
FACS Buffer	2% FBS in PBS

3.1.8 Antibodies

Primary antibodies

Protein name	Host	Product ID (Company)	Dilution
β-Actin	Mouse	Actin (clone C4) (MP Biomedicals USA)	1:10000
β-Actin	Rabbit	Actin 21-33 (Sigma Aldrich USA)	1:10000
GLYATL1	Rabbit	HPA039501 (Human Protein Atlas)	1:1000
Lamin B1	Rabbit	CST12586 (Cell Signaling Technology, USA)	1:1000
CD44-FITC (monoclonal,IM7)	Rat / IgG2b, kappa	11-0441-82 eBioscience, Thermo Fisher Scientific (Waltham, USA)	1:100
Rat IgG2b kappa Isotype Control-FITC (eB149/10H5)	Rat / IgG2b, kappa	11-4031-82 eBioscience, Thermo Fisher Scientific (Waltham, USA)	1:100
CD44-PE (monoclonal,IM7)	Rat / IgG2b, kappa	12-0441-82 eBioscience, Thermo Fisher Scientific (Waltham, USA)	1:150
Rat IgG2b kappa Isotype Control-PE (eB149/10H5)	Rat / IgG2b, kappa	12-4031-82 eBioscience, Thermo Fisher Scientific (Waltham, USA)	1:150
H3K-pan-ac [Histone 3 (acetyl K9+K14+K18+K23+K27)]	Rabbit	ab47915 Abcam (Cambridge, USA)	1:1000
Histone 3 (total)	Mouse	05-449 (Clone 6.6.2) (Merck, Germany)	1:1000

3.1.9 siRNAs

siRNA	Annotation	Catalogue Number	Target Sequence
ON-TARGETplus non-targeting Pool	siCTRL	D-001810-10	UGGUUUACAUGUCGACUAA
			UGGUUUACAUGUUGUGUGA
			UGGUUUACAUGUUUUCUGA
			UGGUUUACAUGUUUUCCUA
siON-TARGETplus Set of 4 Upgrade siRNA GLYATL1	siGLYATL1	LU-010292	AACCUAAUCUGUAUGGAUA
			GCCAUUGGAACUUGGAUUA
			ACCACUAAGGUGAGGAUUA
			GUGAAACUCCCAACUUUAA

siON-TARGETplus Set of 4 Upgrade siRNA ERBB2	siERBB2	LU-003126	UGGAAGAGAUCACAGGUUA GAGACCCGCUGAACAAUAC GGAGGAAUGCCGAGUACUG GCUCAUCGCUCACAACCAA
siON-TARGETplus Set of 4 Upgrade siRNA ESR1	siESR1	LU-003401	GAUCAAACGUCUAAGAAG GAAUGUGCCUGGCUAGAGA GAUGAAGGUGGGAUACGA GCCAGCAGGUGCCCUACUA
siON-TARGETplus Set of 4 Upgrade siRNA EP300	siEP300	LU-003486	
siON-TARGETplus Set of 4 Upgrade siRNA GATA3	siGATA3	LU-003781	

3.1.10 Primers

Gene	Primer Left	Primer Right	Probe #
ACTB	attggcaatgagcggttc	ggatgccacaggactcca	11
BAMBI	cgccactccagctacatctt	cacagtagcatcgaatttcacc	71
CD44	gacaccatggacaagttttgg	cggcaggttatattcaaatcg	13
EP300	gcagcctgcaactccact	gaggatttgatacctgtccttca	20
ERBB2	gggaaacctggaactcacct	agcgatgagcacgtagcc	4
ESR1	gatgggcttactgaccaacc	aaagcctggcaccctctt	24
GATA3	ctacgtgcccgagtacagc	acacactccctgccttctgt	3
GLYATL1	cacatcaatcacgggaacc	ccatgtcatcagtcatctcctg	72
KAT2A	gttgtgagcacccccttgg	tctccacatccaccacca	3
KAT2B	tcccaatgatgatatttctggat	aactgtggcacgttgcagt	76
PUM1	tcacatggatcctcttcaagc	cctggagcagcagagatgtat	86
SLC9A3R1	cagttcatccggtcagtgg	cttcccctccatgcagac	30
ZMYND8	aaagaaacctggcttactgaaca	agtaaaacggaccatgtcttagttc	68

3.2 Methods

3.2.1 Cell Culture and Growth Conditions

MCF7 and T47D cells were obtained from ATCC (LGC Standards GmbH, Wesel, Germany). Cell lines were regularly authenticated by Multiplexion GmbH (Friedrichshafen, Germany) and tested for mycoplasma contamination. The cell lines were cultured in DMEM media and incubated at 37^{o}C with 5% CO_2 in a humidified atmosphere. The growth media for each cell line is described in Table 1. Cells were passaged when they reached 80% confluency under aseptic conditions in a laminar air-flow hood. Briefly, medium was aspirated from the flask and the cells were washed with PBS following which 0.25% trypsin-EDTA was added and the cells were returned to 37^{o}C. Once the cells detached, growth medium was added to neutralize the trypsin. The cells were counted using the Neubauer cell counting chamber. Depending on their growth and cell size, 1-2 x 10^6 cells were seeded into a 150 mm petri dish with a final volume of 15ml. Cell numbers to seed in other canisters were calculated according to the surface area. Cells were used up to passage number 20.

MCF7 resistant cell lines were kindly provided by Magnani Lab at ICL (London, United Kingdom). MCF7 and T47D cell lines, model human breast cancer cell lines of the Luminal A subtype, were chosen to recapitulate the resistance development *in vitro*. In short, MCF7 cells were treated with 100nM 4-Hydroxytamoxifen (4-OHT), the active metabolite of tamoxifen, for more than one year to make them gain resistance, resulting in the MCF7 TAMR cell line. Depriving MCF7 cells from estrogen for over a year, on the other hand, resulted in MCF7 LTED (long term estrogen deprived) cells, which mimic aromatase inhibitor treatment. Double resistant cell line MCF7 LTEDTAMR was generated by treating LTED cell line with 100 nM 4-OHT for one more year (Nguyen et al., 2015). T47D resistant cells were generated in house in collaboration with PhD student Simone Borgoni. Same approach was adopted for T47D resistance acquisition where parental cells were treated with either 100 nM 4-OHT (T47D TAMR) or deprived from estrogen (T47D LTED) for a year. DNA, RNA and protein samples were collected at different time points to monitor changes during resistance acquisition.

Cell line	Description	Growth Media
MCF7	Breast cancer cell line of the Luminal A subtype	DMEM media + 10%FCS +10^{-8} M 17-β-estradiol + 1%Penicillin/ Streptomycin (P/S)
MCF7 TAMR	MCF7 cell line, resistant to 100nM 4-OHT	DMEM media + 10%FCS + 1%P/S + 100nM 4-OHT
MCF7 LTED (Long term estrogen deprivation)	MCF7 cell line, resistant to the deprivation of estrogen. Mimicking the treatment with Aromatase inhibitors	DMEM media without phenol red + 10% charcoal stripped (CS) FCS + 1%P/S
MCF7 LTED TAMR	MCF7 cell line, resistant to the deprivation of estrogen and 100 nM 4-OHT.	DMEM media without phenol red + 10% charcoal stripped (CS) FCS + 1%P/S + 100 nM 4-OHT
MCF7 empty	MCF7 cell line lentivirally transduced with pLX-304 vector without any ORF	DMEM media + 10%FCS +10^{-8} M 17-β-estradiol + 1%Penicillin/ Streptomycin (P/S) + Blasticidin (1μg/ml)
MCF7 GLYATL1 ox	MCF7 cell line lentivirally transduced with pLX-304 vector with *GLYATL1* ORF (ORFeome clone: HQ448152)	DMEM media + 10%FCS +10^{-8} M 17-β-estradiol + 1%Penicillin/ Streptomycin (P/S) + Blasticidin (1μg/ml)
T47D	Breast cancer cell line of the Luminal A subtype	DMEM media + 10%FCS +10^{-8} M 17-β-estradiol + 1%Penicillin/ Streptomycin (P/S)
T47D TAMR	T47D cell line, resistant to 100nM 4-OHT	DMEM media + 10%FCS + 1%P/S + 100nM 4-OHT
T47D LTED	T47D cell line, resistant to the deprivation of estrogen. Mimicking the treatment with Aromatase inhibitors	DMEM media without phenol red + 10% charcoal stripped (CS) FCS + 1%P/S
T47D empty	T47D cell line lentivirally transduced with pLX-304 vector without any ORF	DMEM media + 10%FCS +10^{-8} M 17-β-estradiol + 1%Penicillin/ Streptomycin (P/S) + Blasticidin (1μg/ml)
T47D GLYATL1 ox	T47D cell line lentivirally transduced with pLX-304 vector with *GLYATL1* ORF (ORFeome clone: HQ448152)	DMEM media + 10%FCS +10^{-8} M 17-β-estradiol + 1%Penicillin/ Streptomycin (P/S) + Blasticidin (1μg/ml)

Table 1: Cell lines and culture conditions

To freeze cell stocks, the trypsinized cells were centrifuged at 1200 rpm for 5min. Cell pellets were re-suspended in 1ml freezing media composed of 70% growth media, 20%FBS, 10%DMSO and stored in 1.5 ml cryovials which were slowly cooled down in an isopropanol bath at -80^{o}C for a minimum of 24h before transferring to a liquid nitrogen container for long term storage.

Frozen vials of cells were recovered by thawing quickly in a 37^{o}C water bath. The cell suspension was centrifuged at 1200 rpm for 5min in order to remove DMSO. Following

centrifugation, cells were re-suspended in fresh growth media and transferred into an appropriate culture flask or dish. Cells were allowed to attach overnight before changing the growth media.

3.2.1.1 Generation of stable overexpressing cell lines

For generation of lentiviral particles, HEK293FT cells (Thermo Fischer Scientific, Germany) were co-transfected with the lentiviral expression constructs (pLX-304 vector with or without GLYATL1 ORF) and 2nd generation viral packaging plasmids VSV.G (Addgene #14888) and psPAX2 (Addgene #12260). 48h after transfection, virus containing supernatant was removed and cleared by centrifugation (5min/500g). The supernatant was passed through a 0.45µm filter to remove remaining cellular debris. Target cells were transduced with lentiviral particles at 70% confluency in the presence of 10 µg/ml polybrene (Merck, Germany). 24h after transduction virus containing medium was replaced with selection medium for the respective expression constructs. Transduced cells were selected with the respective antibiotic concentration (geneticin, puromycin or blasticidin).

3.2.2 siRNA Transfections

Cells were seeded to a confluency of 80%. After overnight incubation, transfections were performed with RNAiMax® according to manufacturer's instructions. Unless otherwise stated, siRNAs were used at a final concentration of 30nM.

Plate Format	Volume of OptiMEM added (µl)	Volume of RNAiMax added (µl)	Final concentration of siRNA(nM)	Total volume of transfection mix (µl)
96-well plate	9.6	0.4	30	20
24-well plate	28.5	1.5	30	60
12-well plate	57.5	2.5	30	120
6-well plate	115	5	30	240

Table 2: Volumes of reagents used for siRNA transfections

A pre-mix of RNAiMax and Opti-MEM was prepared (Table 2), and in parallel siRNAs were diluted in Opti-MEM. The siRNA and RNAiMax pre-mix were mixed and incubated for 5 min. During this incubation, media was aspirated from the cells and replaced with fresh growth media

without antibiotics (Pen/Strep). Post incubation, the transfection mix was added to the cells. Cells were then incubated in 37^{o}C, 5% CO_2 humidified atmosphere for different time points depending on the assay being performed.

3.2.3 Analysis of RNA expression

3.2.3.1 mRNA isolation

mRNA was isolated using the "RNeasy Mini" Kit from Qiagen according to manufacturer's recommendations. Extracted RNA was eluted in 30-50µl of nuclease free water. RNA concentrations were determined by NanoDrop ND-100.

3.2.3.2 cDNA synthesis

Up to 1 µg of total RNA was reverse transcribed to cDNA using the RevertAid™ H minus First strand Kit. In short, RNA was mixed with 1µl OligodT primer to a total volume of 12µl. Following a 5min-long incubation at 70^{o}C, 8µl of master mix composed of following ingredients were added per sample.

Reaction Buffer (5X)	4µl
10mM dNTP mix	2µl
Ribolock (RNase inhibitor)	1µl
Reverse Transcriptase	1µl

After the addition of the master mix, samples were incubated in following conditions allowing reverse transcription.

37^{o}C	5 min
42^{o}C	60 min
70^{o}C	10 min

3.2.3.3 Quantitative RT PCR

Primers and probes were designed for Taqman® qRT-PCR using the Roche UPL Design Center (Refer to section 3.1.11 for primer sequences and corresponding probe numbers).

For the Taqman assay the cDNA was first diluted to 2ng/µl. 5µl of cDNA was pipetted into 384 well plates in triplicate. For each gene, a master mix was prepared. 6µl of the master mix was pipetted into each well.

primaQUANT 2x qPCR Probe Master Mix	5.5µl
Left Primer	0.11µl
Right Primer	0.11µl
Taqman probe	0.11µl
H_2O	0.17µl

A plate layout document was prepared using the SDS or QuantStudio software and the PCR conditions were set as follows:

50°C	2 min	
95°C	15 min	
95°C	15 sec	45 cycles
60° C	1 min	

Raw data was analysed using the SDS or QuantStudio software with the ΔΔCt method (Yuan et al., 2006). The Ct values were normalized to housekeeping genes *PUM1* and *ACTB.*

3.2.3.4 RNA Sequencing

Total RNA was extracted using Qiazol (Qiagen, Germany). Quality control, library preparation and sequencing was done at DKFZ Genomics and Proteomics Core Facility using HiSeq 4000 platform (paired-end 100 bp). Analysis of the sequencing data was performed by Maryam Soleimani-Dodaran and Perry Moerland (UvA, The Netherlands). Raw sequencing data were subjected to quality control using FastQC and trimmed using Trimmomatic (v0.32). Reads were aligned to the human reference genome (hg38) using HISAT2 (v2.0.4). Gene level counts were obtained using HTSeq (v0.6.1) and the human GTF from Ensembl (release 85). Statistical analyses were performed using the edgeR and limma R/Bioconductor packages. Genes with more than 1 count in 3 or more samples were retained. Count data were transformed to log2-counts per million (logCPM), normalized by applying the trimmed mean of M-values method and precision weighted using voom. Differential expression between the conditions of interest

was assessed using an empirical Bayes moderated t-test within limma's linear model framework including the precision weights estimated by voom. Resulting p-values were corrected for multiple testing using the Benjamini-Hochberg false discovery rate (FDR). Additional gene annotation was retrieved from Ensembl (release 90) using the biomaRt R/Bioconductor package.

3.2.4 Analysis of protein expression

3.2.4.1 Protein isolation

Cells were washed with ice cold PBS after indicated experimental durations (e.g. transfection, treatment). Cells were lysed with appropriate amount (100µl per 10 cm petri dishes or 40µl per 1 well of a 6-well plate) of RIPA lysis buffer supplemented with 1x Complete Mini Protease Inhibitor Cocktail and 1x Phospho-Stop phosphatase Inhibitor and detached by cell scraper. Lysate was transferred into an eppendorf tube, vortexed and incubated on ice for 30min with agitation. The lysate was tehncentrifuged at 13,000rpm for 10min and the supernatant was transferred to a new pre-chilled eppendorf tube. Lysates were stored at -80^{o}C.

3.2.4.2 Protein quantification

Protein concentrations of the lysates were determined using the BCA™ protein assay kit according to manufacturer's instructions. Briefly, BSA standards of different concentrations (2mg/ml, 1.5mg/ml, 1mg/ml, 0.75mg/ml, 0.5mg/ml, 0.25mg/ml, 0.125mg/ml, 0.025mg/ml and 0mg/ml) were prepared by serial dilution in PBS. 25µl of each standard and 5µl of each sample were pipetted into a 96-well microplate in duplicates. BCA™ working reagent is prepared freshly for each assay by mixing Reagent A with Reagent be in a 50:1 ratio. 200µl of the working reagent was added to each well of the microplate. Following an incubation at 37^{o}C for 30 min in the dark, the absorbance at 562nm was measured on the GloMax explorer plate reader.

A standard curve was prepared using BSA standards and protein concentrations of the samples were determined based on this curve. The obtained concentrations were multiplied by 5 to account for the dilution factor between samples and the standards.

3.2.4.3 Western Blotting

Samples for gel electrophoresis were then prepared (20-30 µg per well) for gel electrophoresis by mixing with protein loading buffer (4X RotiLoad) in a 1:4 dilution and incubated at 95^{o}C for 5min for denaturation. Mini-PROTEAN TGX Precast Gels were loaded into the MiniProtean gel apparatus filled with 1x running buffer. After the removal of the comb, the gels were loaded with 5 µl molecular weight marker and appropriate amount of samples depending on the well number of the gel. Electrophoresis was performed at 120V for 70min.

After protein separation via SDS-PAGE, the proteins were transferred to a PVDF membrane using the Trans-BlotR Turbo Transfer System according to manufacturer's instruction (mixed weight setting, 7min transfer). The membrane was then blocked for 1h at RT with Rockland blocking buffer and subsequently incubated with a target specific primary antibody diluted in blocking buffer overnight at 4°C on a rocking platform. The membrane was washed 3x 15 min in TBST followed by a 1h incubation with IRDye®680 or IRDye®800 conjugated secondary antibodies (1:15,000 diluted in TBST) depending on the host organism of the primary antibody. Following washing for 3x 20 min in TBST, the membrane was scanned at an excitation wavelength of 685nm or 800nm and a resolution of 84 µm using the Odyssey Infrared Imaging System. Western blot bands were normalized using loading controls β-Actin, LaminB1 or total Histone 3.

3.2.4.4 Nuclear and Cytoplasmic Extraction

Nuclear and cytoplasmic fractions were extracted with NE-PER Nuclear AND cytoplasmic extraction kit (Thermo Fisher Scientific) according to manufacturer's instructions. Briefly, cell pellet of 2x10^{6} cells was resuspended in 200µl ice-cold CERI buffer and incubated on ice for 10 min. 11µl CERII buffer was added to the mix and incubated for 1 min on ice. Following vortexing, the tube was centrifuged at 13,000 rpm for 5 min at 4 ^{o}C. Supernatant (cytopliasmic fraction was transferred into a pre-chilled eppendorf tube. Pellet was resuspended in 100µl NER buffer, vortexed and incubated on ice for 40 min with vortexing briefly every 10 min. Following the incubation, the tube was centrifuged at 13,000 rpm for 10 min at 4^{o}C. Supernatant (nuclear fraction) was transferred into a pre-chilled eppendorf tube. Protein concentrations were

determined by BCA™ protein assay kit. Samples were stored at -80 ^{o}C. For western blotting, Lamin B1 and β-actin were used as loading controls for nuclear and cytoplasmic fractions, respectively.

3.2.4.5 Histone Extraction

Histones were extracted with Histone extraction kit (Abcam) according to manufacturer's instructions. In short, cell pellet of approximately 1x10^{7} cells were resuspended in 1ml of 1X Pre-Lysis Buffer, incubated on ice for 10 min and centrifuged at 10,000 rpm for 1 min. Pellet was re-suspended in 200µl of Lysis Buffer and vortexed briefly. Following a 30 min-long incubation on ice, samples were centrifuged at 13,000 rpm for 10 min and supernatant was transferred to a new eppendorf tube. Approximately 60µl of Balance Buffer (with the addition of DTT) was added to the samples prior to protein concentration determination by BCA™ protein assay kit and long term storage at -80°C.

3.2.4.6 Histone ELISA

Histone ELISA was carried out according to manufacturer's instructions. Histone ELISA for Total H3, H3K9ac and H3K14ac were performed in parallel with the same samples. Purified core histones were diluted (100 ng/well for total H3, 500 ng/well for H3K9ac and H3K14ac) in Assay diluent buffer 1X and pipetted into the wells of 96-well plates pre-coated with anti-H3 (Total Histone 3) in duplicates. Recombinant histones were prepared with Assay diluent buffer 1X as instructed and pipetted into the wells in duplicates and diluted as instructed serving as standards. Following pipetting of protein standard curve and samples, plates were incubated at room temperature for 1 h with agitation. Following 1 h incubation, the wells were washed 3 times with 200µl Wash buffer 1X (diluted from 20X with distilled H_2O). 50µl diluted primary antibody (1:1000 dilution for anti-total H3; 1:2000 dilution for anti-H3K9ac; 1:500 dilution for anti-H3K14ac were done in assay diluent buffer 1X) was added into each corresponding well. Plates were incubated at room temperature for 1 h with agitation. After the incubation, wells were washed 3 times with 200µl Wash Buffer 1X. 50µl diluted secondary antibody (1:2000 HRP conjugated anti-rabbit IgG) was added into each well. Plates were incubated at room temperature

for 1h without agitation. After the incubation, wells were washed 3 times with 200µl Wash Buffer 1X. After complete removal of the final wash, 100µl developing solution was added into each well. Plates were incubated for the optimal developing time (2 min for total H3, 3 min for H3K9ac, 9 min for H3K14ac) in the dark without agitation. After optimal incubation time, 100µl stop solution was added to each well and absorbances were measured on GloMax® microplate reader at 450 nm with a reference wavelength at 600 nm. Optical densities of standards were plotted in order to get the best fit for standard curve and quantify the modified histone abundance in samples. Each modified histone sample was normalized against its total Histone 3.

3.2.4.7 Flow Cytometry

Cells were cultured to 70 to 80% confluence and detached from the cell culture flasks using TrypLE express. Cell pellets were obtained and washed with FACS Buffer (2% FBS in PBS). All further steps were performed on ice and all centrifugation steps at 4°C. Monoclonal antibodies against human CD44 (FITC, eBioscience; CD44-PE, eBioscience) and their isotype controls were added to the cell suspension (in final volume of 100µl FACS Buffer) at concentrations recommended by the manufacturer and incubated on ice in the dark for 30 min. Unstained controls and stained cells were analyzed on FACS Canto II (BD Biosciences). Gating was set to relevant isotype control-labeled cells or unstained cells for each cell line. For transfection efficiency check, cells transfected with only transfection reagent and transfected with mCherry-tagged dCas9 (±effector domain) plasmid or GFP plasmid (pmax GFP) were analyzed on LSRFortessa™ (BD Biosciences).

3.2.5 Functional assays

3.2.5.1 Cell Counting Assay

Cell growth under different treatment (siRNA, sgRNA and media) conditions were analyzed with a microscopy based nuclei counting method. Cells were seeded in clear-bottomed 96 well black plates and after overnight incubation they were transfected with siRNA or plasmids encoding dCas9+effector domain and sgRNAs. At different time points DNA was stained with

intercalating dye Hoechst-33258 (1:5000 dilution in growth media) for 30-45min. Subsequently, the plates were imaged with a molecular devices microscope IXM XLS. All nuclei were defined by Hoechst signals within a certain size and intensity and were detected and counted by the Molecular Devices Software. The number obtained was considered as cell number.

3.2.5.2 Cell Titer Glo Assay

Cell titer Glo assay measures the ATP in cells as a proxy of the metabolic activity. The incubation with the assay reagent results in lysis of the cells which can then release its ATP. For this assay cells were seeded into black 96-well plates and treated with siRNA and/or the media containing either 4-OHT or estrogen deprivation media after overnight incubation. The assay was performed 72h or 7 days after treatment according to the manufacturer's recommendation. Briefly, the components of the assay kit were mixed and added into the wells. The incubation with the reagents resulted in the lysis of the cells which released ATP. The plate was placed on a shaker for 2 min to help the lysis. Luciferin, catalyzed in the assay by UtraGlo Luciferase, and the ATP generated oxyluciferin, which was detected via luminescence measurement using the Glomax Explorer Plate Reader.

3.2.5.3 Quantification of S-phase with EdU incorporation assay

For quantification of cells in S-phase, Click-iT™ EdU Cell Proliferation Kit was used according to manufacturer´s instructions. Cells were seeded in a black 96-well plate. Cells were allowed to attach overnight before changing the growth media to endocrine therapy conditions. After 72h of incubation at 37°C with 5% CO_2 EdU (5-ethynyl-2′-deoxyuridine) was added into the cells by replacing half of the media with freash media including 20µM EdU to reach final concentration of 10µM EdU in each well. Following 24h of EdU incorporation, media was aspirated and cells were washed twice with PBS. Then, 50µl of 4% PFA was added into each well and cells were fixed for 15 min at RT. After removing the fixative, cells were washed twice with 100µl 3% BSA in PBS. Following the removal of the washing solution, 100µl 0.5% Triton X-100 in PBS was added into each well and cells were incubated for 20 min at RT. Following the removal of the permeabilization buffer, cells were washed twice with 100µl 3% BSA in PBS. After washing,

50µl of freshly prepared Click-iT® reaction cocktail (ingredients shown below) was added into each well and the plate was incubated for 30 min at RT protected from light.

1X Click-iT® reaction buffer	43µl per sample
$CuSO_4$	2µl per sample
Alexa Fluor® azide (Alexa Fluor 647)	0.12µl per sample
Reaction buffer additive	5µl per sample
Total Volume	50µl per sample

Following the incubation, cells were washed once with 100µl 3% BSA in PBS and 100µl nuclear staining solution (Hoechst-33258 diluted 1:5000 in PBS) was added into each well. Plate was incubated for 30 min at RT protected from light before imaging with molecular devices microscope IXM XLS. All nuclei were defined by Hoechst signals within a certain size and intensity and were detected and counted by the Molecular Devices Software. Cells going through S-phase were defined by Alexa Fluor 647 staining. The ratio of Alexa Fluor 647-positive cells were obtained as percentage of total cell number determined by Hoechst staining.

3.2.5.4 Determination of late apoptotic cells via Propidium Iodide (PI)

Late apoptotic cells were determined via PI staining combined with Hoechst staining (Refer to Part 3.2.16.1). PI (1:5000 dilution in growth media) was added into the wells of 96-well plate 5 min before the end of Hoechst staining incubation time (30-45 min). Plates were imaged with molecular devices microscope IXM XLS. All nuclei were defined by Hoechst signals within a certain size and intensity and were detected and counted by the Molecular Devices Software. Late apoptotic cells were defined by PI staining. The ratio of late apoptotic cells were obtained as percentage of total cell number determined by Hoechst staining.

3.2.5.5 Cell migration and invasion assay

A cell suspension of $2x10^5$ cells in 200µl of DMEM media without FBS were seeded into each upper chamber of Transwell chamber (8-µm pore size, Corning Costar Corp, US), which was

pre-coated with or without 5 mg/ml matrigel. 500µl DMEM supplemented with 20% FBS was used in the lower chamber as chemoattractant. After incubating for 24 h for migration and 72h for invasion at 37°C in a humidified incubator with 5% CO_2, chambers were disassembled. Cells were swiped off the top off the upper chambers with a cotton swab and the lower chambers were by fixed incubating with 4% PFA for 10 min. Transwell inserts were then moved to a new 24 wells plate and stained with crystal violet for 30 min. After staining the inserts were washed with PBS to remove the excess staining and dried overnight. 10% acetic acid was used to elute crystal violet and the absorbance of the dissolved crystal violet was measured at 490 nm using the Glomax Explorer Plate Reader. Absorbance values were normalized to values of parental (WT) cells.

3.2.6 Determination of methylation changes via Illumina EPIC 850k array

Total DNA was extracted using Qiagen DNeasy Blood and Tissue Kit (Qiagen, Germany). Methylation profiling was performed using the Illumina MethylationEPIC BeadChip platform. Analysis of the raw data was performed by Maryam Soleimani-Dodaran (UvA, The Netherlands). Pre-processing was performed using the R package minfi. Data were normalized using functional normalization. Detection p-values were calculated for each methylation probe. 14947 probes showed an unreliable signal ($p>0.01$) in one or more samples and were removed. Probes corresponding to loci that contain a SNP in the CpG site or in the single-base extension site were removed. We also removed probes that have been shown to cross-hybridize to multiple genomic positions (Pidsley et al., 2016). The final data set comprised 781062 CpG loci. Using the resulting M-values CpG-wise linear models were fitted with coefficients for each condition (TAMR, LTED, LTEDTAMR, WT). Differential methylation between the conditions of interest was assessed using empirical Bayes moderated t-statistics (R package limma) and p-values were corrected using the Benjamini-Hochberg FDR.

3.2.7 Assay for Transposase Accessible Chromatin with High Throughput Sequencing (ATAC-Seq)

Libraries for ATAC-sequencing were prepared as previously published with modifications (Buenrostro et al., 2015). Briefly, cells were lysed by 1% NP40 and tagmented at 55°C for 8 minutes in a reaction mix with 2.5µl of TDE1 (Nextera Illumina DNAKit), 25µl Tagmentation buffer (Nextera Illumina DNA Kit) and 25µl of lysed cells. Reaction was stopped by adding 10 µl Guanidium (5 M) and samples were purified using AMpure Beads. Libraries were generated using NEBNext High Fidelity PCR Mix and sequenced on the Illumina HiSeq 2500 platform at DKFZ Genomics and Proteomics Core Facility. Sequencing reads were adaptor-trimmed using cutadapt (v. 1.10). Genomic alignments were performed against the human reference genome (hg19, NCBI build 37.1) using Bowtie2 (v. 2.3.0). The non- default parameters "-q 20 -s" were used. PCR duplicates were removed by Picard MarkDuplicates (v. 1.125; http://broadinstitute.github.io/picard). Signal tracks were generated using deepTools (v. 2.3.3). Peaks were called using Macs2 (v. 2.1.1.) with the parameters "--nomodel --shift -50 --extsize 100 --qvalue 0.01". All peaks were merged to create a common bed file with read counts before differential analysis using edgeR (v. 0.3.16). Gene annotations were made using ChIPpeakAnno. Sample and library preparation were carried out by me. Data analysis was carried out by Dr. Simin Oz (DKFZ).

3.2.8 CRISPR/dCas9-mediated targeted epigenetic editing experiments

For epigenetic editing editing experiments, a hit-and-run approach was adopted in which the cells were transiently transfected with both a sgRNA plasmid and another plasmid encoding dCas9 fused with an effector domain (epi-effector). The latter construct also has an mCherry tag to enable sorting to ensure enrichment of the transfected cells. dCas9-effector domain constructs were either provided by Mihaly Koncz (MTA-SZBK) and Marianne Rots (UMCG) or obtained from Addgene with appropriate MTAs. The general procedure of experimental setup involves transfection for 48h, collection and validation of epigenetic interference in mRNA level (TaqMan) and re-seeding of cells to assess long term effects.

3.2.8.1 Plasmids

Name	Description
pMdCas9-P2A-mCherry	Plasmid expressing the catalytically dead (D10A and H840A mutations) dCas9 without effector domain (dCas9-NED) and mCherry with a self-cleavage linker in between
pMdCas9-p300-P2A-mCherry	Plasmid expressing dCas9 fused to the histone acetyltransferase core of p300 and mCherry with a self-cleavage linker in between
pMdCas9-G9a-P2A-mCherry	Plasmid expressing dCas9 fused to catalytic domains of the methyltransferase G9a. Included are the SET domain, the PreSET domain and one ANK repeat domain. The fluorophore mCherry is also expressed from the plasmid
pMdCas9-PRDM9-P2A-mCherry	Plasmid expressing dCas9 fused to the histone methyltransferase core of PRDM9 and mCherry with a self-cleavage linker in between
pMdCas9-DOT1L-P2A-mCherry	Plasmid expressing dCas9 fused to the histone methyltransferase core of DOT1L and mCherry with a self-cleavage linker in between
pMdCas9-M.SssI(Q147L)-P2A-mCherry	Plasmid expressing dCas9 fused to the reduced activity mutant (Q147L point mutation) of prokaryotic DNMT and mCherry with a self-cleavage linker in between
pMdCas9-TET1-P2A-mCherry	Plasmid expressing dCas9 fused to the DNA demethylase core of TET1 and mCherry with a self-cleavage linker in between
MLM3636	sgRNA backbone, for the expression of *CD44* sgRNAs. MLM3636 was a gift from Keith Joung (Addgene plasmid #43860;http://n2t.net/addgene:43860;RRID:Addgene_43860)
MLM 2.0	sgRNA backbone, for the expression of *BAMBI*, *GLYATL1*, *SLC9A3R1* sgRNAs; with two MS2 stem loops; derived from MLM3636
pmax GFP	Plasmid used to assess transfection efficiency of sgRNA plasmid (MLM2.0) due to similar size.
pHAGE-TRE dCas9	Lentiviral expression vector expressing Tet-inducible catalytically dead Cas9 (dCas9). Used as a backbone vector to produce inducible cell lines stably expressing dCas9+effector domain.

	pHAGE TRE dCas9 was a gift from Rene Maehr & Scot Wolfe (Addgene plasmid # 50915 ; http://n2t.net/addgene:50915 ; RRID:Addgene_50915)

3.2.8.2 Generation of cell lines expressing inducible stable dCas9+effector domain

Stable cell line generation was done using pHAGE-TRE dCas9 (fused to indicated effector domains) lentiviral contructs as explained in section 3.2.1.1. The lentiviral constructs were generated by Mihaly Koncz (MTA-SZBK) using pHAGE-TRE dCas9 as a backbone. Generated cell lines were seeded in the presence of 2µg/ml doxycycline to induce expression of dCas9+effector domain and transfected with the sgRNA plasmid(s) next day.

Name	Description	Growth Media
MCF7 (dCas9-p300)	MCF7 cell line stably expressing Tet-inducible dCas9 fused to the histone acetyltransferase core of p300	DMEM media + 10%FCS +10^{-8}M 17-β-estradiol + 1%Penicillin/ Streptomycin (P/S) + Neomycin (1µg/ml)
MCF7 (dCas9-TET1)	MCF7 cell line stably expressing Tet-inducible dCas9 fused to the catalytic domain of DNA demethylase TET1	DMEM media + 10%FCS +10^{-8}M 17-β-estradiol + 1%Penicillin/ Streptomycin (P/S) + Neomycin (1µg/ml)
MCF7 TAMR (dCas9-Q147L)	MCF7 TAMR cell line stably expressing Tet-inducible dCas9 fused to the reduced activity mutant (Q147L point mutation) of prokaryotic DNMT	DMEM media + 10%FCS + 1%P/S + 100nM 4-OHT+ Neomycin (1µg/ml)
MCF7 TAMR (dCas9-G9a)	MCF7 TAMR cell line stably expressing Tet-inducible dCas9 fused to catalytic domains of the methyltransferase G9a. Included are the SET domain, the PreSET domain and one ANK repeat domain.	DMEM media + 10%FCS + 1%P/S + 100nM 4-OHT+ Neomycin (1µg/ml)
MCF7 LTED (dCas9-	MCF7 LTED cell line stably expressing Tet-inducible dCas9 fused to the	DMEM media without phenol red + 10% charcoal stripped

Q147L)	reduced activity mutant (Q147L point mutation) of prokaryotic DNMT	(CS) FCS + 1%P/S+ Neomycin (1µg/ml)
MCF7 LTED (dCas9-G9a)	MCF7 LTED cell line stably expressing Tet-inducible dCas9 fused to catalytic domains of the methyltransferase G9a. Included are the SET domain, the PreSET domain and one ANK repeat domain.	DMEM media without phenol red + 10% charcoal stripped (CS) FCS + 1%P/S+ Neomycin (1µg/ml)
MCF7 LTED (dCas9-G9a)	MCF7 LTED cell line stably expressing Tet-inducible dCas9 fused to catalytic domains of the methyltransferase G9a. Included are the catalytically inactive SET domain, the PreSET domain and one ANK repeat domain.	DMEM media without phenol red + 10% charcoal stripped (CS) FCS + 1%P/S+ Neomycin (1µg/ml)

3.2.8.3 Design and Cloning of sgRNAs

3.2.8.3.1 Design of sgRNAs

The sgRNAs were designed using the CRISPOR online tool (www.crispor.tefor.net) and cloned into the MLM2.0 vector (except for *CD44* sgRNAs that were cloned in to MLM3636 backbone vector). Genomic locations with histone marker peaks and CpG islands (if existent) were taken into account while designing the sgRNAs. Promoter region of target genes are covered with sgRNAs targeting both upstream and downstream of transcription start site (TSS). Additionally, enhancer region of *SLC9A3R1* gene was also covered with sgRNAs.

sgRNA	Sequence	Genomic Orientation
BAMBI #1	TCTCAAGGGGCGTGCTGACAAGG	antisense
BAMBI #2	AGTGTCGTCTCGTTGGCGCCGGG	sense
BAMBI #3	AGGGTCTCCACAGCACGCAAGGG	antisense
BAMBI #4	CGATCCATTGACGCCCCGCACGG	antisense
BAMBI #5	TGAGGGCACGCGGCAGCTACAGG	antisense
BAMBI #6	TGTAGCACATCATCCTCGCTGGG	antisense
CD44 #1	TATGATATTTCAATCTCAAAAGG	sense
CD44 #2	CGGATGGAAGGATATTTAGGAGG	sense
CD44 #3	GAGTTGGTGAATCTTCCAAGTGG	antisense
CD44 #4	ACGGAGGCACTGCGCCACCCAGG	sense
CD44 #5	TCACAGGATGTTGGATATCCTGG	antisense

CD44 #6	TGGTGCAAGGTTTTACGGTTCGG	sense
CD44 #7	CACTGGCTTTCTTCTCCCGCTGG	antisense
CD44 #8	GGGGAACCTGGAGTGTCGCGGGG	antisense
CD44 #9	GCTGAGGCTGTAAATAATCGGGG	antisense
CD44 #10	TCGCTCCCTCCCTCCGTCTTAGG	sense
SLC9A3R1 #1	AGGCTGTTAAAACGTTGGCCAGG	sense
SLC9A3R1 #2	AGTGATCCGAGATCGCGCCACGG	sense
SLC9A3R1 #3	AATGGCCGGCGCCGTTCACCCGG	sense
SLC9A3R1 #4	TGGGACGCTCAGACGCCGCGCGG	sense
SLC9A3R1 #5	GCGCGACTTGGGGTTCCGATGGG	antisense
SLC9A3R1 #6	AGAAAAACGACGGCGGGCGCGGG	antisense
SLC9A3R1 #7	CCACACCGCCCTGAGATGGTGGG	antisense
SLC9A3R1 #8	AGAAGAGGTTACTAGCCAGAAGG	sense
SLC9A3R1 #9	GATAAATGGGAGTGACGTCCAGG	antisense
SLC9A3R1 #10	CCCAATGGGTTGCCCCATTCTGG	sense
SLC9A3R1 #11	ACACGCAGATTTCGAGTGCTTGG	antisense
GLYATL1 #1	AGCCTCAATTAAGTTTACTCTGG	antisense
GLYATL1 #2	CACCATTGGAGCTAGCCTGCAGG	antisense
GLYATL1 #3	GTTAATTCCGCCTGGTGATTAGG	sense
GLYATL1 #4	GTGTAAGGACCGCTTTCATGAGG	antisense
GLYATL1 #5	CCTTCATAGCTTGCCTTACAAGG	sense

3.2.8.3.2 Oligoduplex formation

In short sgRNA top and bottom strand oligos, containing a BsmBI digestion site, were resuspended to 100 μM in H_2O. 5μl of top strand oligo and 5μl of bottom strand oligo were mixed with 2 μl 10x NEB Buffer #3 and 8 μl of H_2O and heated at 95°C for 5 minutes with subsequent cool down to 10°C at -5°C/minute, to ensure generations of oligo duplexes.

Oligo	**Sequence**
BAMBI #1 forward	ACACCGTCTCAAGGGGCGTGCTGACAG
BAMBI #1 reverse	AAAACTGTCAGCACGCCCCTTGAGACG
BAMBI #2 forward	ACACCGAGTGTCGTCTCGTTGGCGCCG
BAMBI #2 reverse	AAAACGGCGCCAACGAGACGACACTCG
BAMBI #3 forward	ACACCGAGGGTCTCCACAGCACGCAAG
BAMBI #3 reverse	AAAACTTGCGTGCTGTGGAGACCCTCG
BAMBI #4 forward	ACACCGCGATCCATTGACGCCCCGCAG
BAMBI #4 reverse	AAAACTGCGGGGCGTCAATGGATCGCG
BAMBI #5 forward	ACACCGTGAGGGCACGCGGCAGCTACG
BAMBI #5 reverse	AAAACGTAGCTGCCGCGTGCCCTCACG
BAMBI #6 forward	ACACCGTGTAGCACATCATCCTCGCTG
BAMBI #6 reverse	AAAACAGCGAGGATGATGTGCTACACG
SLC9A3R1 #1 forward	ACACCAGGCTGTTAAAACGTTGGCCG
SLC9A3R1 #1 reverse	AAAACGGCCAACGTTTTAACAGCCTG
SLC9A3R1 #2 forward	ACACCAGTGATCCGAGATCGCGCCAG

SLC9A3R1 #2 reverse	AAAACTGGCGCGATCTCGGATCACTG
SLC9A3R1 #3 forward	ACACCAATGGCCGGCGCCGTTCACCG
SLC9A3R1 #3 reverse	AAAACGGTGAACGGCGCCGGCCATTG
SLC9A3R1 #4 forward	ACACCTGGGACGCTCAGACGCCGCGG
SLC9A3R1 #4 reverse	AAAACCGCGGCGTCTGAGCGTCCCAG
SLC9A3R1 #5 forward	ACACCGCGCGACTTGGGGTTCCGATG
SLC9A3R1 #5 reverse	AAAACATCGGAACCCCAAGTCGCGCG
SLC9A3R1 #6 forward	ACACCAGAAAAACGACGGCGGGCGCG
SLC9A3R1 #6 reverse	AAAACGCGCCCGCCGTCGTTTTTCTG
SLC9A3R1 #7 forward	ACACCCCACACCGCCCTGAGATGGTG
SLC9A3R1 #7 reverse	AAAACACCATCTCAGGGCGGTGTGGG
SLC9A3R1 #8 forward	ACACCAGAAGAGGTTACTAGCCAGAG
SLC9A3R1 #8 reverse	AAAACTCTGGCTAGTAACCTCTTCTG
SLC9A3R1 #9 forward	ACACCGATAAATGGGAGTGACGTCCG
SLC9A3R1 #9 reverse	AAAACGGACGTCACTCCCATTTATCG
SLC9A3R1 #10 forward	ACACCCCCAATGGGTTGCCCCATTCG
SLC9A3R1 #10 reverse	AAAACGAATGGGGCAACCCATTGGGG
SLC9A3R1 #11 forward	ACACCACACGCAGATTTCGAGTGCTG
SLC9A3R1 #11 reverse	AAAACAGCACTCGAAATCTGCGTGTG
GLYATL1 #1 forward	ACACCGAGCCTCAATTAAGTTTACTCG
GLYATL1 #1 reverse	AAAACGAGTAAACTTAATTGAGGCTCG
GLYATL1 #2 forward	ACACCGCACCATTGGAGCTAGCCTGCG
GLYATL1 #2 reverse	AAAACGCAGGCTAGCTCCAATGGTGCG
GLYATL1 #3 forward	ACACCGTTAATTCCGCCTGGTGATTG
GLYATL1 #3 reverse	AAAACAATCACCAGGCGGAATTAACG
GLYATL1 #4 forward	ACACCGTGTAAGGACCGCTTTCATGG
GLYATL1 #4 reverse	AAAACCATGAAAGCGGTCCTTACACG
GLYATL1 #5 forward	ACACCGCCTTCATAGCTTGCCTTACAG
GLYATL1 #5 reverse	AAAACTGTAAGGCAAGCTATGAAGGCG

3.2.8.3.3 Digestion of plasmid backbone

2 μg of MLM2.0 vector backbone were mixed with 10μl 10x NEB 3.1 buffer, 1μl of BsmBI and H_2O up to 100μl and digested at 55°C for 4 hours. The digested plasmid was run on a 2% agarose gel and isolated using the Promega Wizard SV gel and PCR clean up system according to manufacturer's instructions.

3.2.8.3.4 Ligation

For the ligation, 30ng of digested plasmid, 1 μl of oligo duplex, 1 μl 10x T4 DNA ligase bugger, 0.5 μl T4 DNA ligase (NEB) and H_2O ad 10μl were incubated at 24°C for 10 minutes and chilled on ice prior to transformation.

3.2.8.3.5 Transformation

80µl of competent *E.coli* were mixed with 1µl of the ligated plasmid mixture and incubated on ice for 30 minutes. Following a heat shock at 42°C for 90 seconds, 800µl of SOC (Super Optimal Broth with Catabolite repression) media were added and incubated shaking at 37°C for 1 hour. 100 µl of the transformed bacteria were spread on LB-Agar plates with Ampicillin (100µg/ml) and incubated over night. Colonies were picked the next day for colony PCR.

3.2.8.3.6 Colony PCR and Sanger sequencing

PCR mix was as shown in the table

Component	**Amount per sample**
10X Buffer	2.5µl
$MgCl_2$	1.5µl
10mM dNTP	0.5µl
10µM forward primer	0.5µl
10µM reverse primer	0.5µl
Taq Polymerase	0.2µl
ddH_20	19.3µl

Forward primer was the sequencing primer for MLM3636 plasmid. Reverse primer was the reverse oligo for the respective sgRNA.

MLM3636 sequencing primer (OS280)	5'-CAGGGTTATTGTCTCATGAGCGG-3'

PCR conditions were as follows:

94°C	3 min	
94°C	20 sec	30 cycles
Tm (depending on the primers)	20 sec	
72°C	20 sec	
72°C	7 min	

Following the agarose gel electrophoresis and validation of the fragment inserted, samples were sent for Sanger sequencing to ensure complete and correct insertion.

3.2.8.3.7 Plasmid Purification

Bacteria were grown in LB-media with Ampicillin (100μg/ml) and pelleted at 5,000 x g for 15 minutes at 4°C. The plasmids were isolated using the Qiagen Plasmid Plus Midi Kit according to manufacturer's instructions. In short, pellets were resuspended in 4ml of buffer P1, mixed with 4ml of buffer P2 and incubated at room temperature for 3 minutes. 4ml of buffer S3 were added and the mixture was inverted 4-6 times. The lysate was transferred to the QIAfilter Cartridge and incubated at room temperature for 10 minutes. The lysate was filtered into a tube, using a plunger, and mixed with 2ml Buffer BB. The mixture was transferred to a QIAGEN Plasmid Plus spin column and drawn through the column with a vacuum pump. The column was washed with 0.7ml Buffer ETR and 0.7 ml Buffer PE and centrifuged at 10,000 x g for 1 minute to dry the membrane. The plasmid was eluted in 150 μl H_2O and stored at -20°C.

3.2.8.4 Plasmid Transfection

The transfection conditions are shown in table. Cells were seeded in P/S-free media and transfected 24 hours later. The corresponding amount of plasmid DNA (1:1 ratio sgRNA plasmid and dCas9 plasmid) was added to Opti-MEM (media) containing the appropriate amount of LTX and plus reagent (ThermoFisher Scientific) and incubated for 30 minutes at room temperature (Table 3). The transfection mix was then added dropwise on the cells.

Culture vessel	P/S-free media	Cells per well	Transfection mix	DNA amount	LTX reagent	Plus reagent
96 well plate	80 μl	1250 (MCF7 WT and TAMR) 2000 (MCF7 LTED)	20 μl	100 ng	0.4 μl	0.2 μl
48 well plate	160 μl	20.000	40 μl	200 ng	0.8 μl	0.4 μl
24 well plate	400 μl	40.000	100 μl	500 ng	1.5 μl	0.75 μl
12 well plate	800 μl	80.000	200 μl	1000 ng	2.5 μl	1.25 μl
6 well plate	1500 μl	150.000	500 μl	2000 ng	5 μl	2.5 μl

Table 3: Volumes of reagents used for plasmid transfections

3.2.9 Dataset Analysis

Publicly available METABRIC and TCGA datasets were downloaded from the website given in Section 3.1.7. METABRIC data comprised of mRNA expression microarray and clinical data of nearly 2000 patients with a follow up of 20 years. Box plots are represented as $\log_2$ transformed

gene expression data. Kaplan-Meier method was used to generate survival curves (Kaplan & Meier, 1958). 33.3% of the patients with the highest and lowest expression of *GLYATL1* and 25% of the patients with the highest and lowest expression of *BAMBI* were used to generate survival curves. Km plotter was used to generate survival curves (Á. Nagy et al., 2018). Survival curve for Pawitan dataset was obtained from Cancertool (Cortazar et al., 2018). TCGA data comprises of RNA-sequencing and clinical data of nearly 1200 patients. Box plots were generated using gene expression data represented as $\log_2$ (x+1) rsem.

3.2.10 Statistical Analysis

Unless otherwise mentioned, data are presented as mean ± SD and statistical analyses were performed by unpaired two-tailed Student's t-test (with Welch's correction) and p-values smaller than 0.05 were considered statistically significant. P values <0.05, <0.01, <0.001 and <0.0001 are indicated with one, two, three and four asterisks respectively.

3.2.11 Graphical Illustrations

All graphs were generated using the GraphPad Prism Software and illustrated via Inkscape v 0.91

4. Results

4.1 Part I: Elucidating the Role of *GLYATL1* in Endocrine Therapy Resistant Breast Cancer

4.1.1 Establishment and characterization of endocrine therapy resistant breast cancer cell lines

In this project, we are focusing on Luminal A subtype due to their ER^+ status. I chose MCF7 as luminal model cell line of interest. Dr. Luca Magnani (Imperial College London) has kindly provided MCF7 cells resistant to 100 nM 4-OH-tamoxifen (TAMR), long term estrogen-deprivation (LTED) which mimics aromatase inhibitor resistance, and LTED MCF7 cells which are also resistant to tamoxifen (LTEDTAMR) as well as a sensitive counterpart (WT) (Nguyen et al., 2015). MCF7 TAMR and LTED cells have been generated either via treatment with 100 nM 4-OH-Tamoxifen (active metabolite of the drug) or deprivation from estrogen for one year, respectively. Treatment of LTED cells with 4-OH-tamoxifen for one more year resulted in LTEDTAMR cells, which is a model for double resistance (Figure 7a) (Nguyen et al., 2015). We have generated also a resistance acquisition model in T47D, another luminal A cell line, in collaboration with Simone Borgoni established via treatment with either 100 nM 4-OH-tamoxifen (T47D TAMR) or long term estrogen deprivation (T47D LTED) for one year (Figure 8a). For T47D resistance acquisition model, we collected DNA, RNA, supernatant and protein samples from the cells in different time-points for further analysis. We also monitored and validated resistance development over time with phenotypic assays such as proliferation, viability, drug response, migration and invasion.

Tamoxifen resistance was validated by viability assays under treatment with increasing doses of 4-OH-Tamoxifen (4-OHT). Tamoxifen resistant cells were found to be more viable compared to their sensitive counterparts under treatment with up to 500 nM 4-OHT, which is five times more than the amount of the drug dose that they have been used for development of resistance (Figure 7b). LTEDTAMR cells showed a similar trend in tamoxifen resistance. Resistant cells also

displayed more invasive phenotype compared to parental MCF7 cells that they are derived from (Figure 7c).

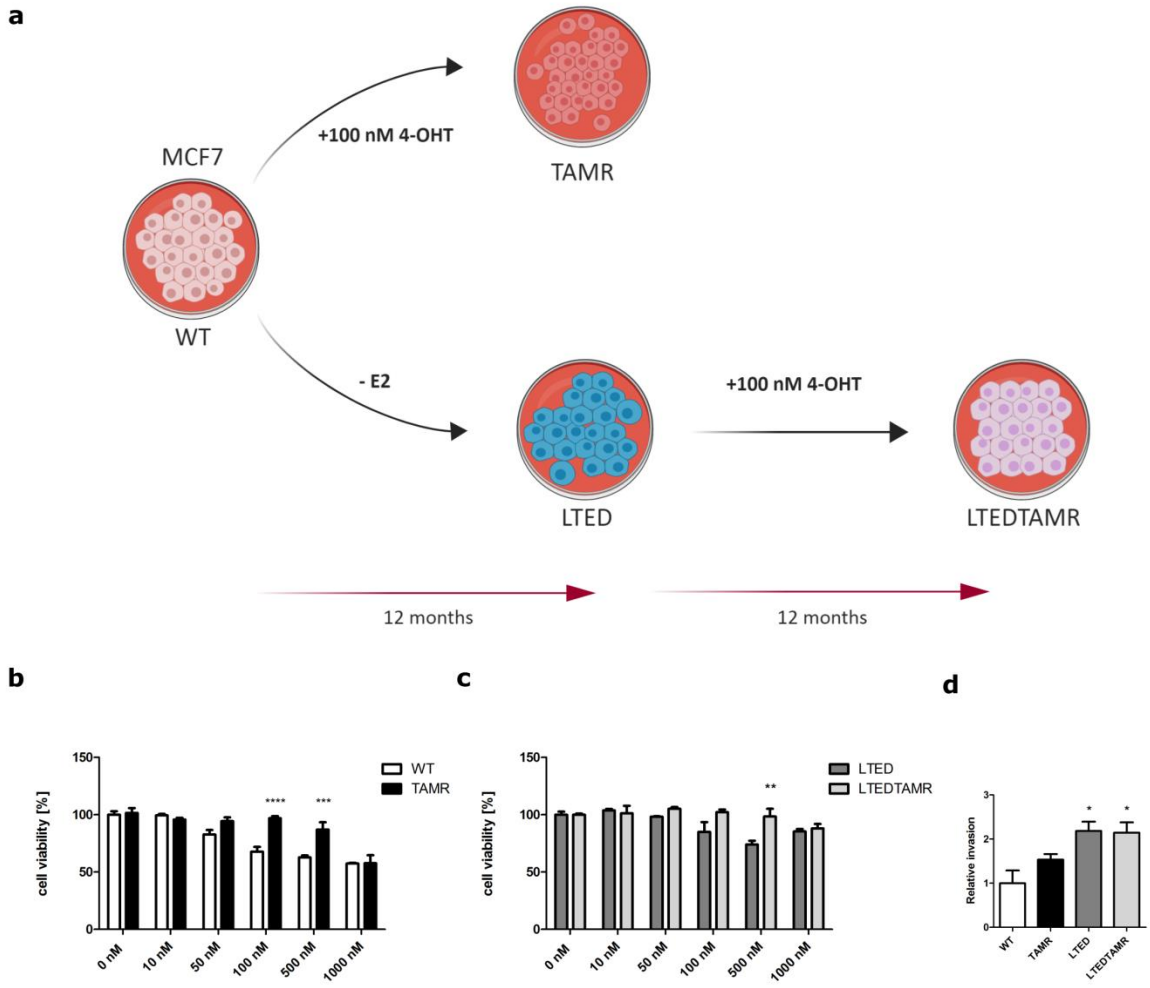


Figure 7: Establishment and characterization of endocrine therapy resistant MCF7 breast cancer cell lines. (a) Scheme for establishment of endocrine therapy resistance acquisition experiment for MCF7 cells. (b-d) Invasion assay was performed at different time-points of resistance acquisition. (b) WT and TAMR cells were grown in the presence of either vehicle (EtOH) or 4-OHT and cell viability was measured after 72h with Cell Titer Glo Assay. (c) LTED and LTEDTAMR cells were grown in the presence of either vehicle (EtOH) or 4-OHT and cell viability was measured after 72h with Cell Titer Glo Assay. (d) Invasion assay of MCF7 cell repertoire. All values are represented as relative values normalized to WT. Data are presented as mean ± SD, for invasion assays n=3 (each with 3 technical replicates), for cell viability assays n=2 (with 6 technical replicates). *** represents $p < 0.001$, ** represents $p<0.01$, * represents $p<0.05$.

The same approach was adopted in characterization of resistant phenotype and validation of resistance acquisition over time for T47D cell line. Large scale screening techniques such as RNA-sequencing and EPIC array was performed on samples acquired at months 1, 2, 5 and 7 of resistance acquisition to fully comprehend the changes occurring during immediate early, early. intermediate and late time-points of exposure to endocrine therapy conditions. T47D TAMR cells displayed resistance to 4-OHT starting from 5 months of drug exposure (Figure 8e). After 7 months under 4-OHT treatment, the cells were even resistant to higher doses of 4-OHT than 100 nM (Figure 8f,g). And similar to their MCF7 counterparts, T47D cells got more invasive phenotype as they acquired resistance to endocrine therapy conditions (Figure 8b-d).

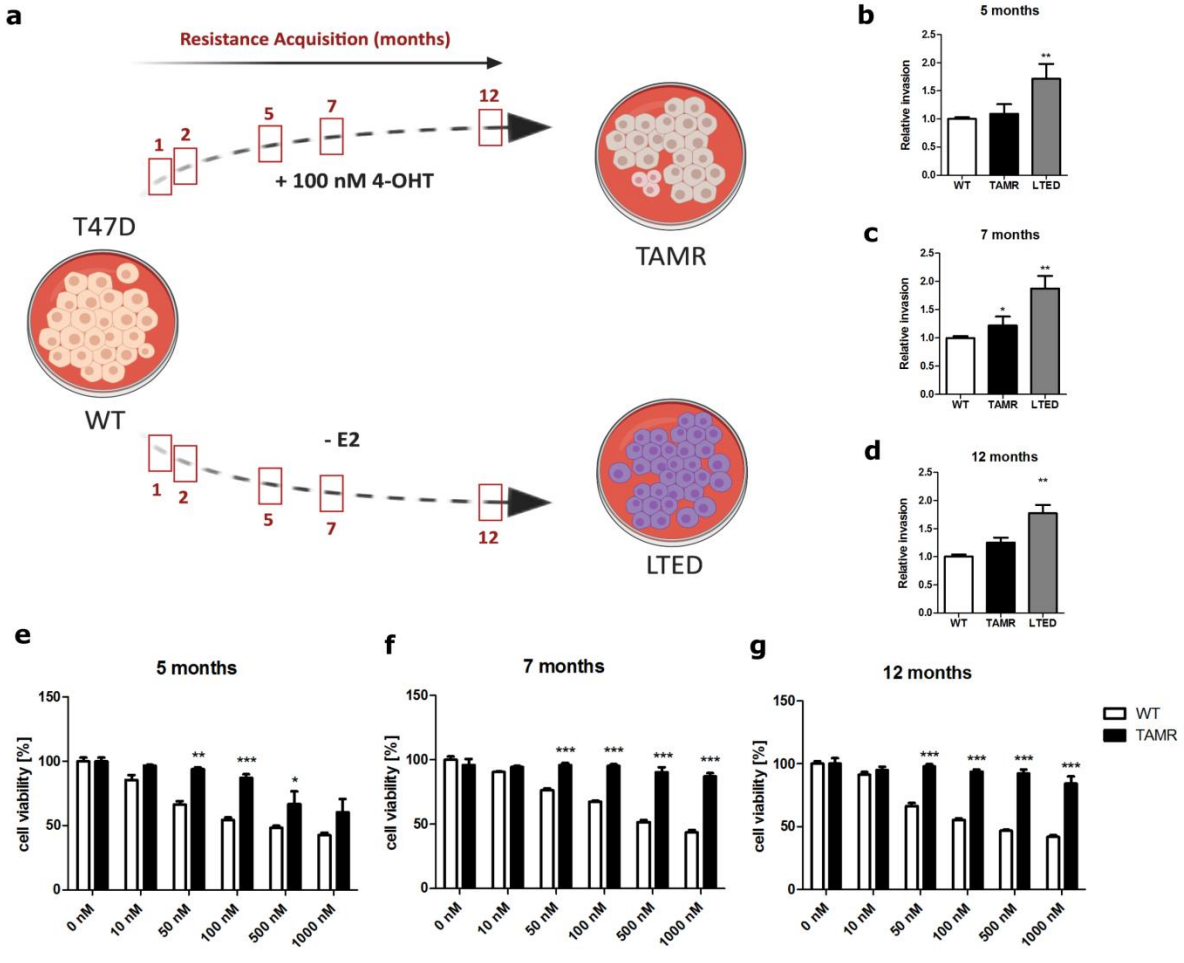


Figure 8: Establishment and characterization of endocrine therapy resistant T47D breast cancer cell lines. (a) Scheme for establishment of endocrine therapy resistance acquisition experiment for T47D cells. The indicated time-points are marking the samples used for RNA-seq, EPIC array (except for month 12) and functional assays. (b-d) Invasion assays were performed at different time-points of resistance acquisition. (e-g) WT and TAMR cells were grown in the presence of either vehicle (EtOH) or 4-OHT and cell viability was measured over time with Cell Titer Glo Assay. All values are represented as relative values normalized to WT. Data are presented as mean ± SD, for invasion assays n=1 for each time point (each with 3 technical replicates), for cell viability assays n=1 for each time point (with 6 technical replicates). *** represents $p < 0.001$, ** represents $p<0.01$, * represents $p<0.05$.

4.1.2 Profiling of resistant cell line repertoire

Resistant MCF7 cell lines and their sensitive counterpart were profiled with next-generation sequencing techniques in order to elucidate changes in transcriptome (RNA-Sequencing), chromatin accesibility (ATAC-Seq) and methylome (EPIC array) that could be accounted for the development of endocrine therapy resistance in these cells. The reason for inclusion of ATAC-Seq and EPIC array was to assess and understand the level of epigenetic regulation of endocrine therapy resistance to a certain extent. RNA-Seq and EPIC array data were analyzed by Maryam Soleimani-Dodaran and Perry Moerland (UvA, The Netherlands). ATAC-Seq data initial analysis was performed by Dr. Simin Oz (DKFZ) and integration of ATAC-Seq and RNA-Seq data was done by Dr. Perry Moerland.

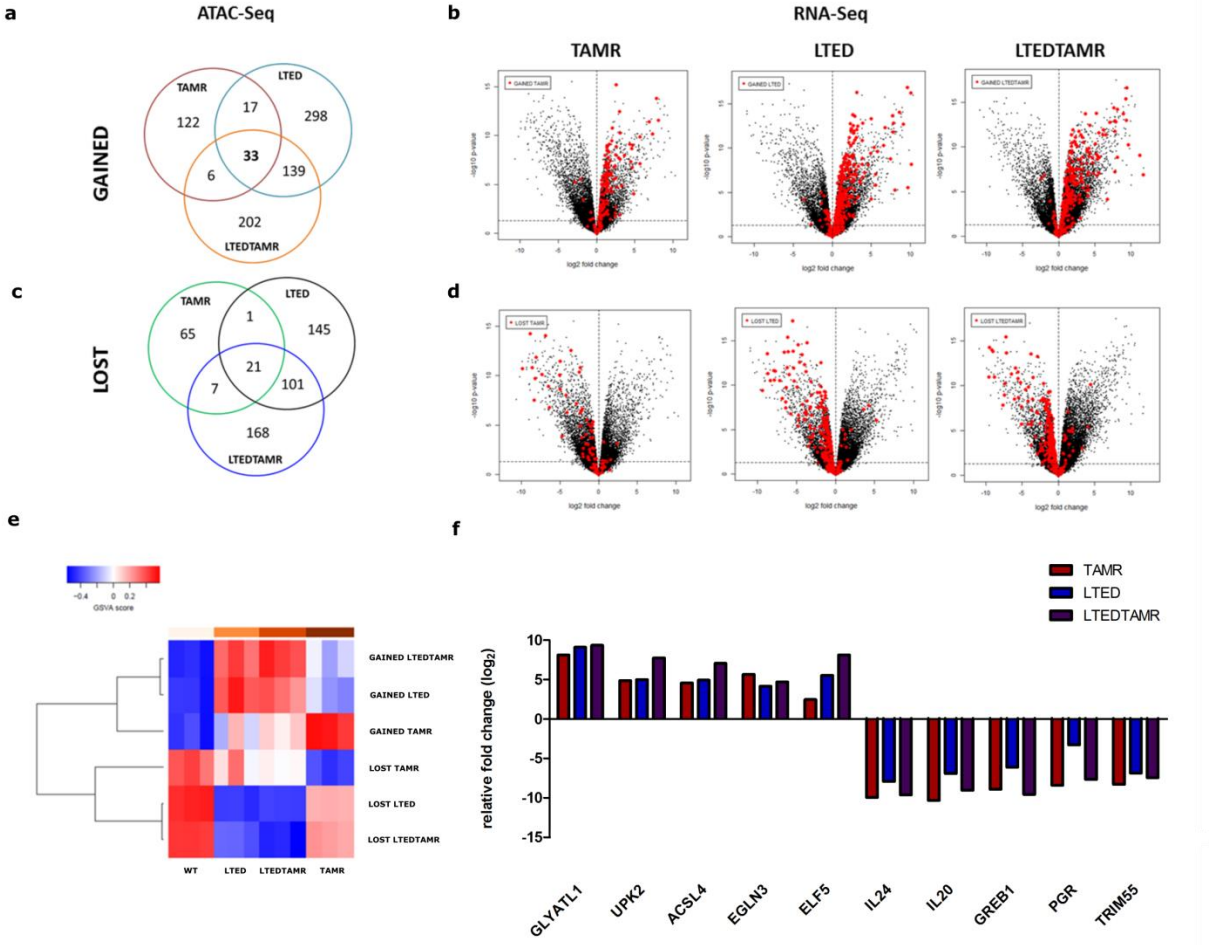


Figure 9: Correlation of RNA-Seq and ATAC-Seq for resistant MCF7 cell line repertoire. The number of genes in associated with gained (a) and lost (c) peaks around TSS (±2 kb) based on ATAC-Seq data and RNA-Seq volcano plots upregulated (b) and downregulated (d) genes in MCF7 resistant cell lines compared to parental. The genes from ATAC-Seq analysis are highlighted in red. (e) Gene set variation analysis (GSVA) comparing the subset of genelists based on gained and lost peaks of ATAC-Seq data with RNA-Seq results. (f) Top 5 up- and down-regulated genes in all three resistant cell lines derived from parental MCF7.

Integration of ATAC-Seq and RNA-Seq data revealed that gained ATAC-Seq peaks around the transcription start site (± 2 kbp TSS) in resistant cell lines were correlating with upregulated genes whereas lost peaks of ATAC-Seq data were belonging to the TSS of the genes that were downregulated in resistant cells compared to parental. This finding was verified via geneset variation analysis comparing the genesets based on gained and lost peaks of ATAC-Seq data with RNA-Seq (Figure 9a-e).

Among the genes common to all MCF7-derived resistant cell lines, top 5 upregulated genes with gained peaks from ATAC-Seq were *GLYATL1*, *EGLN3*, *ACSL4*, *UPK2*, *ELF5* and top 5 downregulated with lost peaks were *IL24*, *IL20*, *GREB1*, *PGR* and *TRIM55* (Figure 9f). I chose *GLYATL1* as target gene potentially implicated in therapy resistance, since this gene was found to be the most strongly upregulated one in MCF7 TAMR, LTED and LTEDTAMR.

GLYATL1 (Glycine-N-acyltransferase Like 1/Glutamine-N-acyltransferase) encodes for an acyltransferase using L-glutamine as its substrate. This gene has been characterized in 2007 and suggested to be implicated in activation of heat shock pathway (H. Zhang et al., 2007). There are three known members of GLYATL family of enzymes: *GLYATL1*, *GLYATL2* and *GLYATL3*. All encode for acyltransferases with differing substrates. *GLYATL1* is highly expressed in liver and kidney, whereas *GLYATL2* is enriched in salivary glands (Matsuo et al., 2012). *GLYATL1* is considered to be a biomarker for prostate cancer since it has been shown to be significantly upregulated in two independent datasets comparing primary prostate cancer samples to healthy control (Barfeld et al., 2014). TCGA dataset comparison between normal and matched primary tumors also revealed that prostate cancer, colon cancer and breast cancer are the only cancer entities with significantly increased expression difference in tumor tissue.

Interestingly, *GLYATL1* has also been implicated in androgen independent prostate cancer refractory to androgen deprivation therapy as a candidate gene (Nalla et al., 2016). Moreover, *GLYATL1* was the only gene common for two independent gene signatures for vascular invasion prediction in hepatocellular carcinoma (W. Liu et al., 2012; Mínguez et al., 2011)

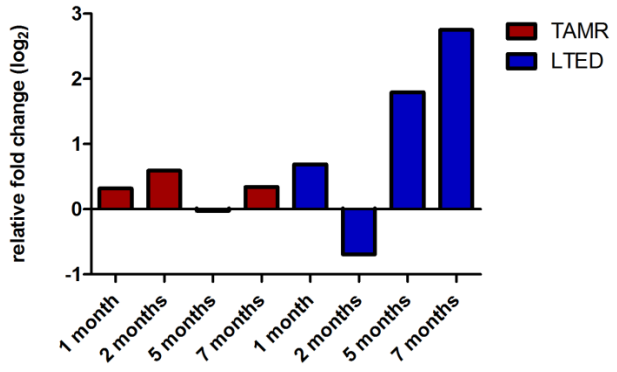


Figure 10: Changes in *GLYATL1* expression during T47D resistance acquisition. RNA was extracted from T47D cells at different time-points of tamoxifen-exposure (TAMR) or estrogen deprivation (LTED). GLYATL1

levels were determined by RNA-Sequencing. All values are represented as relative values normalized to WT. Data are presented as mean of two biological replicates.

GLYATL1 was also found to be upregulated in T47D resistant cells in mRNA level. T47D LTED cells had higher *GLYATL1* expression compared to T47D TAMR. Detailed analysis of different time-points of T47D resistance acquisition revealed that *GLYATL1* mRNA expression was elevated already at 1 month of both tamoxifen administration and estrogen-deprivation (Figure 10). Even though it fluctuated in between, the expression got even higher for T47D LTED cells at 12 months (Figure 11c). I then continued with checking the protein levels of GLYATL1 in both MCF7 and T47D resistant cell line repertoire that we have. Western blotting results confirmed the gene expression data obtained via RNA-Seq and TaqMan qRT-PCR as resistant cells derived from both luminal cell lines showed higher levels of GLYATL1 in protein level (Figure 11).

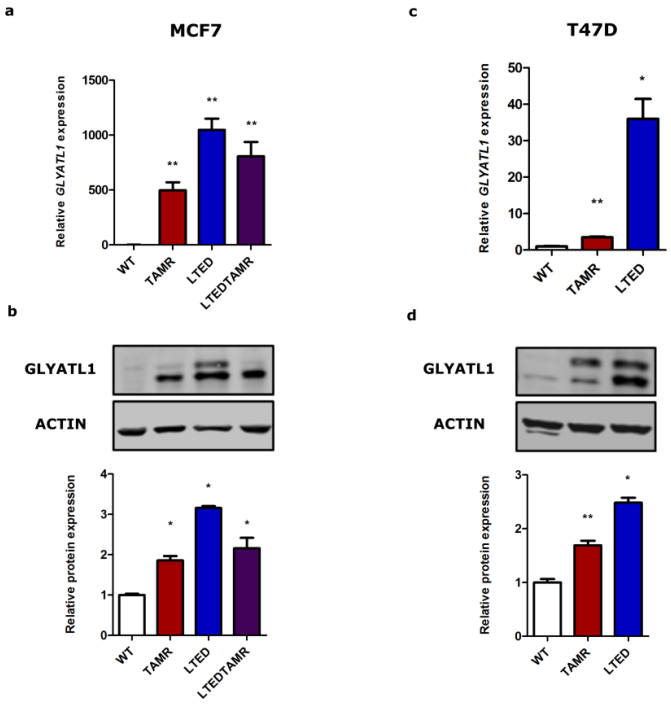


Figure 11: GLYATL1 increases both at mRNA and protein level in endocrine therapy resistant cells. (a) RNA and protein were extracted from resistant and sensitive MCF7 and T47D cells. GLYATL1 levels were determined by qPCR (a,c) and western blot (b,d) respectively. All values are represented as relative values normalized to WT. Data are presented as mean ± SD. Values for mRNA expression were first normalized to *PUM1* and *ACTB* levels, n=3 (each with 3 technical replicates), Protein expression values were first normalized to actin levels. Data shown are representative of 3 independent experiments (n=3) ** represents $p<0.01$, * represents $p<0.05$.

4.1.3 Clinical Significance of *GLYATL1*

Assessment of *GLYATL1* expression among 42 different breast cancer cell lines revealed that breast cancer cell lines belonging to luminal or HER2-enriched subtype had higher *GLYATL1* expression compared to the ones of basal subtype (Figure 12a). Cell lines provide the opportunity to recapitulate endocrine therapy resistance acquisition *in vitro*. Established cell line models enable the researchers to shed more light into the intricate molecular mechanisms that bring about resistance to therapies and relapse in patients. However, it is of utmost importance to understand and evaluate the clinical relevance of *in vitro* findings. For this purpose, the expression patterns of *GLYATL1* among breast cancer molecular subtypes were analyzed using two publicly available datasets TCGA (1200 patients) and METABRIC (2000 patients) (Collins, 2007; Curtis et al., 2012). *GLYATL1* expression was found to be significantly higher in the luminal and HER2-enriched subtype of breast cancer compared to basal in both datasets corroborating *in vitro* cell line data (Figure 12b). However, the MCF7 and T47D cell lines did not show elevated levels of basal *GLYATL1* expression compared to cell lines of other subtypes. Interestingly, HER2-enriched subtype had the highest *GLYATL1* expression among all subtypes in METABRIC dataset (Figure 12c).

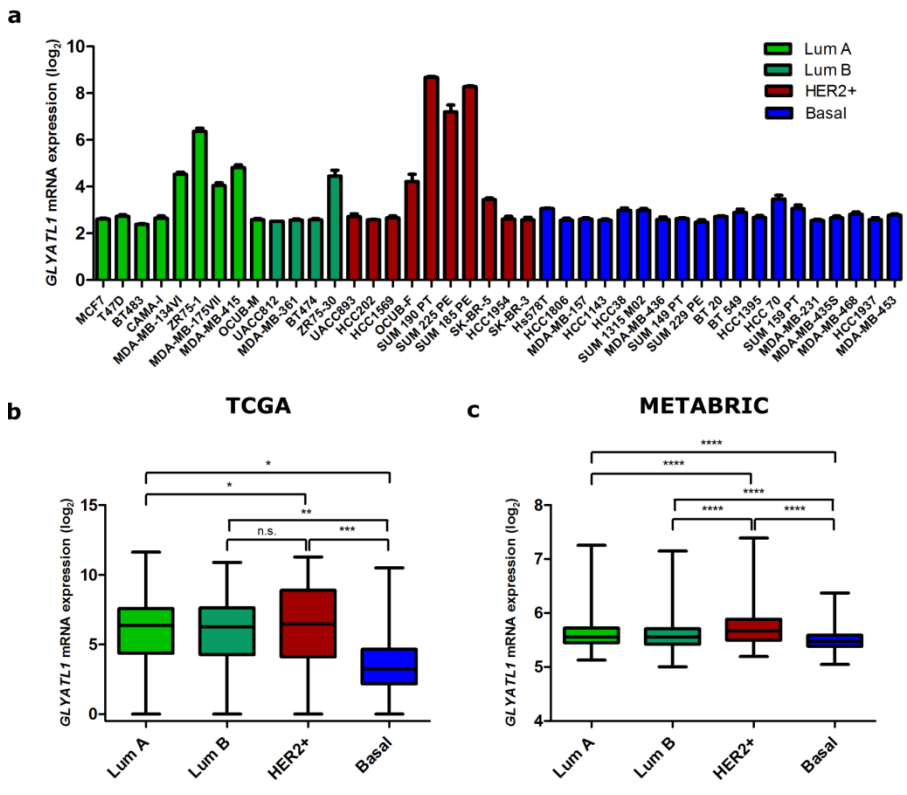


Figure 12: *GLYATL1* expression in breast cancer cell lines and patients. (a) Comparison of GLYATL1 mRNA expression levels among 42 different breast cancer cell lines belonging to different subtypes Lum A (n=9), Lum B (n=4), HER2+ (n=10), Basal (n=19). (b) Comparison of *GLYATL1* mRNA expression analysis [$\log_2$ (x+1) rsem] in the TCGA dataset among the intrinsic subtypes Lum A (n=421), Lum B (n=192), HER2+ (n=67) and Basal (n=141). (c) Comparison of *GLYATL1* mRNA expression analysis ($\log_2$ a.u.) in the METABRIC dataset among the intrinsic subtypes Lum A (n=721), Lum B (n=492), HER2+ (n=240) and Basal (n=331). **** represents p<0.0001, *** represents p < 0.001, ** represents p<0.01, * represents p<0.05.

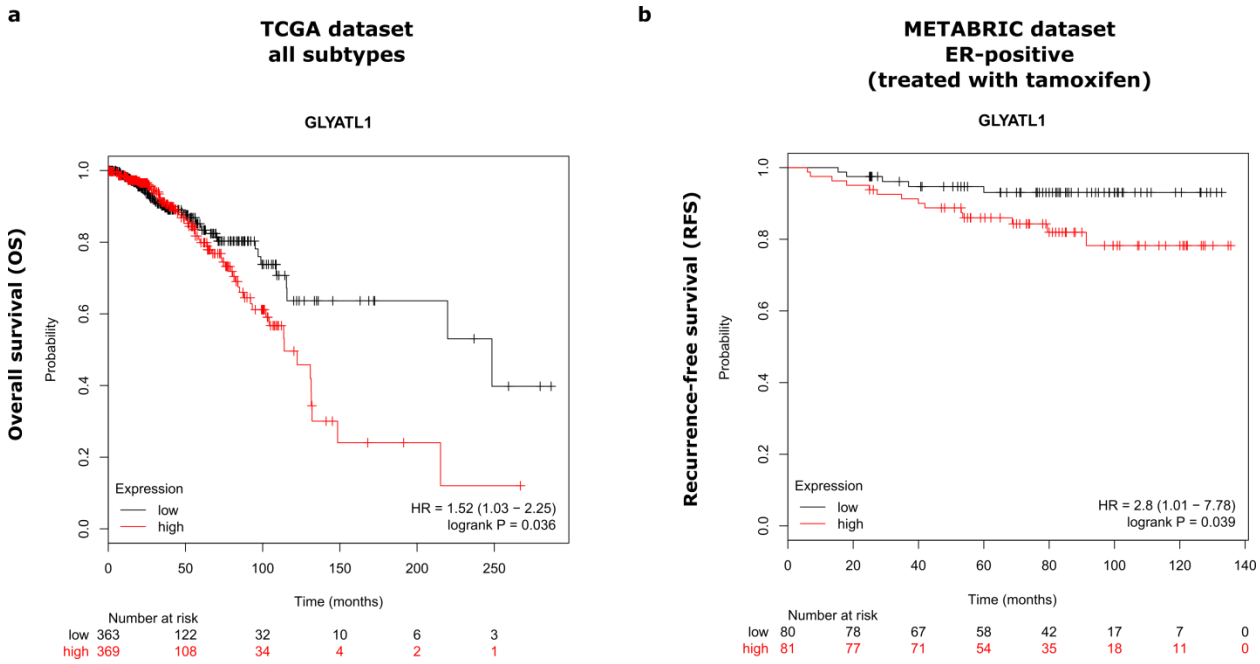


Figure 13: Correlation of *GLYATL1* expression with patient survival. (a) Tertile –based survival analysis of TCGA dataset of *GLYATL1* low versus high gene expression (for each tertile n=363). (b) **.** Recurrence-free survival analysis of METABRIC dataset of *GLYATL1* low versus high gene expression for ER^+ patients undergoing tamoxifen treatment (n=161). Plotted with Kaplan-Meier Plotter (Á. Nagy et al., 2018)

As for the role of *GLYATL1* on overall patient survival, TCGA dataset showed that high *GLYATL1* expression was significantly correlated with poor patient prognosis (Á. Nagy et al., 2018) (Figure 13a). Survival analysis of METABRIC dataset demonstrated a similar trend in recurrence-free survival (RFS) for patients with ER-positive tumors who are under tamoxifen treatment (n=161) indicating a proxy for *GLYATL1* in the context of endocrine therapy resistance (Figure 13b).

This trend was further confirmed in an independent cohort (n=159) where patients with higher *GLYATL1* expression levels and disease-free survival were significantly inversely correlated (Cortazar et al., 2018; Pawitan et al., 2005). In this dataset, recurrence within 5 years was also monitored and *GLYATL1* expression was found to be significantly higher in patients with recurred breast cancer compared to disease-free individuals (Pawitan et al., 2005)(Figure 14).

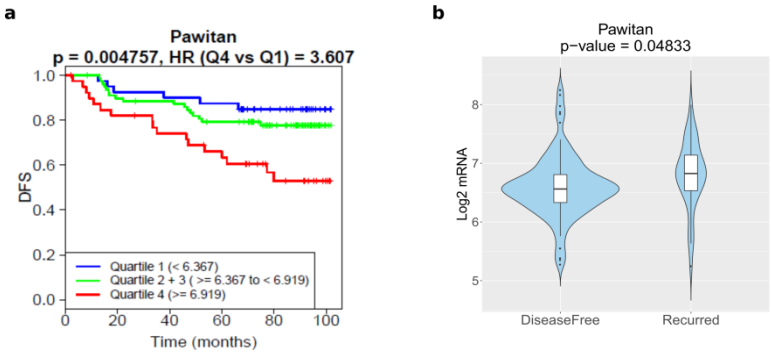


Figure 14: Correlation of *GLYATL1* expression with disease-free patient survival. (a) Quartile–based survival analysis of Pawitan dataset of *GLYATL1* low, moderate and high gene expression. (b) *GLYATL1* expression difference between disease-free individuals and recurrence patients (Cortazar et al., 2018; Pawitan et al., 2005).

4.1.4 *GLYATL1* as a novel gene in endocrine therapy resistance context

4.1.4.1 Effect of *GLYATL1* knockdown on cell viability

Integrative analysis of RNA-Seq and ATAC-Seq revealed *GLYATL1* as one of the top differentially regulated genes in all resistant cell lines compared to their sensitive counterparts. In order to investigate the effect of *GLYATL1* loss on both parental and resistant cell lines, RNAi approach was used. I used a pool of four siRNAs targeting *GLYATL1* to transfect MCF7 cells and checked the viability of the cells after 7 days in their respective treatment conditions with Cell Titer Glo assay from Promega which measures the ATP production of the cells as a proxy for metabolic activity and therefore viability.

Viability of parental MCF7 cells did not change after transfection. TAMR cells were found to be less viable under tamoxifen treatment compared to vehicle control when *GLYATL1* was knocked down. LTED cells also showed less metabolic activity following siRNA transfection. LTEDTAMR cells showed sensitivity to both estrogen deprivation and tamoxifen treatment after *GLYATL1* knockdown (Figure 15).

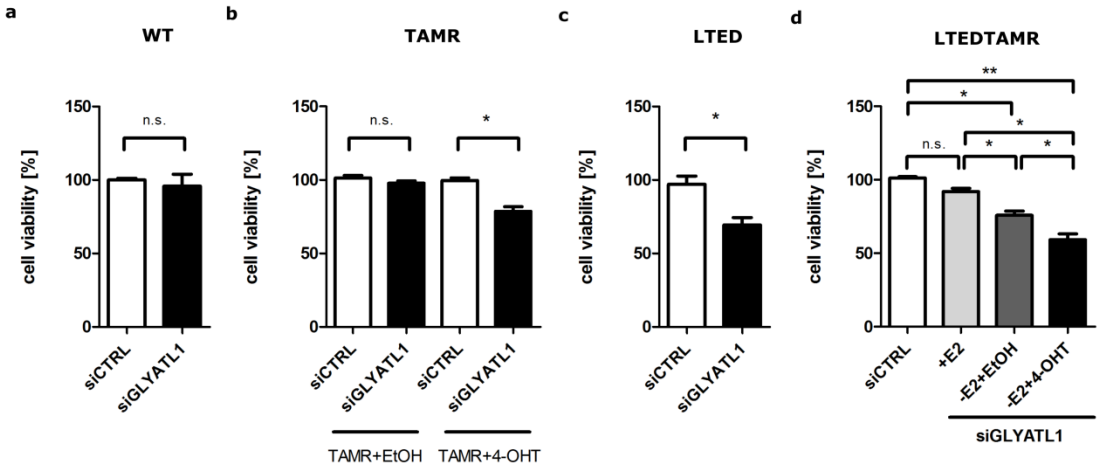


Figure 15: *GLYATL1* knockdown affects resistant MCF7 cell viability. a-e, Cells were transfected with siGLYATL1 and cell viability was assessed after 7 days with the Cell Titer Glo assay. All values are represented as relative values normalized to non-targeting control siRNA (siCTRL). Data are presented as mean ± SD, n=3 (each with 5 technical replicates) ** represents $p < 0.01$,* represents $p<0.05$ (n.s: not significant).

T47D resistant cell lines displayed similar tendency in re-sensitization to endocrine therapy conditions upon *GLYATL1* knockdown. T47D TAMR cells were significantly less resistant to 4-OHT exposure and T47D LTED cells were less viable following siGLYATL1 transfection (Figure 16).

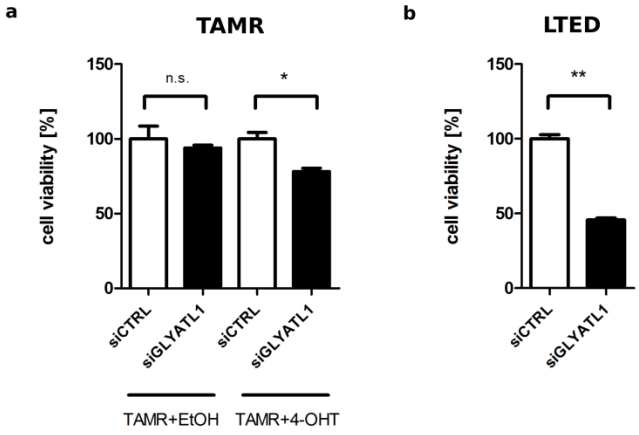


Figure 16: *GLYATL1* knockdown affects resistant T47D cell viability. T47D TAMR (a) and LTED (b) cells were transfected with siGLYATL1 and cell viability was assessed after 7 days with the Cell Titer Glo assay. All values are represented as relative values normalized to non-targeting control siRNA (siCTRL). Data are presented as mean ± SD, n=3 (each with 5 technical replicates) ** represents $p < 0.01$,* represents $p<0.05$ (n.s: not significant).

Deconvolution of the 4 siRNAs of the *GLYATL1* pool revealed that all siRNAs show a significant decrease in *GLYATL1* mRNA levels (Figure 17). Nuclei count of the cells transfected with individual and pool GLYATL1 siRNA showed all sgRNAs were able to significantly, albeit variably, inhibit cell growth.

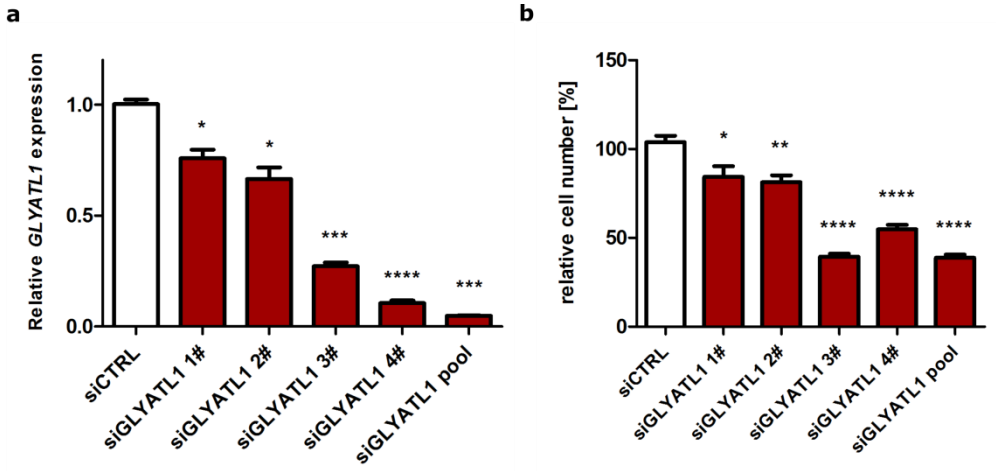


Figure 17: Knockdown efficiency and deconvolution of siGLYATL1 pool. MCF7 LTED cells were transfected with siGLYATL1 (pool and individual). (a) RNA was harvested 72 hours post-transfection and knockdown efficiency was determined by qRT-PCR. (b) Microscopy-based nuclei counting of the cells were performed 72h post transfection. All values were normalized to siCTRL. Values for mRNA expression were first normalized to *PUM1* and *ACTB* levels and are presented as mean ± SD, n=3 (each with 3 technical replicates). Proliferation data are represented as mean ± SD, n=3 (each with 5 technical replicates) ****represents $p<0.0001$, *** represents $p < 0.001$, ** represents $p<0.01$, * represents $p<0.05$.

In conclusion, knockdown of *GLYATL1* resulted in significant decrease in cell viability under endocrine therapy conditions in resistant cell lines derived from both MCF7 and T47D.

4.1.4.2 Effect of GLYATL1 overexpression on cell proliferation

Following the observation that knockdown of *GLYATL1* partially, albeit significantly, re-sensitizes the resistant cells to endocrine therapy, I wanted to investigate whether overexpression of this gene would confer resistance to parental cells or not. In order to answer this research question, I used lentivirally transduced stable cell lines overexpressing *GLYATL1* (MCF7 GLYATL1 ox and T47D GLYATL1 ox) generated by Dr. Rainer Will (DKFZ). Overexpression of target gene and protein was validated by TaqMan qRT-PCR and Western Blotting,

respectively (Figure 18) for both cell lines comparing overexpressing cells (GLYATL1 ox) to their empty vector controls (MCF7 empty and T47D empty). The two bands observed in western blotting in GLYATL1 overexpressing cells suggest a post-translational modification for this protein.

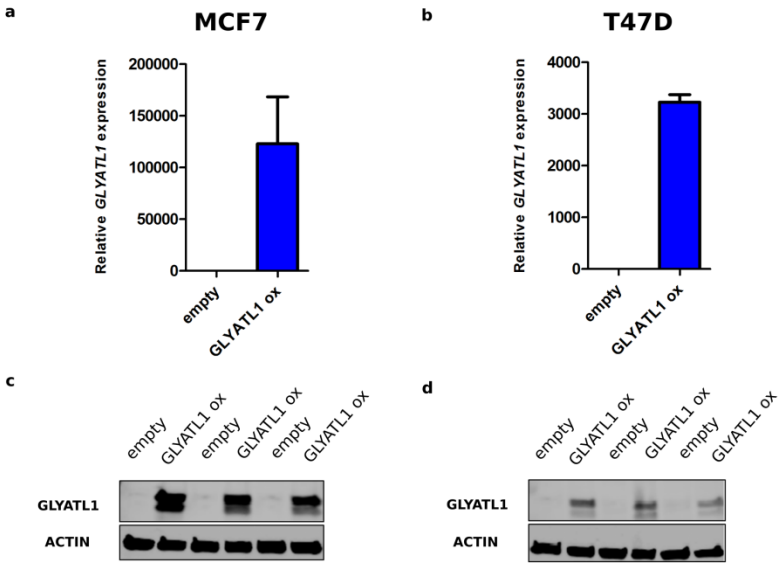


Figure 18: GLYATL1 overexpression in MCF7 and T47D cells. RNA and protein were harvested from GLYATL1 overexpresing cells (GLYATL1 ox) and empty vector controls (empty). GLYATL1 levels were determined by qRT-PCR (a,b) and western blotting (c,d). Values for mRNA expression were first normalized to *PUM1* and *ACTB* levels. Data are presented as mean ± SD, n=2 for qRT-PCR (each with 3 technical replicates)., n=3 for western blotting.

Cell proliferation assays were performed over a course of one week via microscopy-based nuclei count following Hoechst staining. Cell growth media was changed to endocrine therapy conditions 24 hours post-seeding to assess the effects of therapy conditions. The conditions were normal growth media supplemented with E2 (+E2), growth media with vehicle EtOH, growth media with 100 nM 4-OHT, estrogen deprivation media (-E2) or estrogen deprivation media with 100 nM 4-OHT. For MCF7 cells, overexpression of GLYATL1 gave the cells a slight proliferation advantage in normal growth conditions (+E2). Moreover, MCF7 GLYATL1 ox cells grew significantly better in estrogen deprivation conditions compared to empty vector

control. Impressively, GLYATL1 overexpressing MCF7 cells suffered significantly less from tamoxifen treatment. They also coped better with both estrogen deprivation and tamoxifen treatment significantly than their empty vector counterparts (Figure 19).

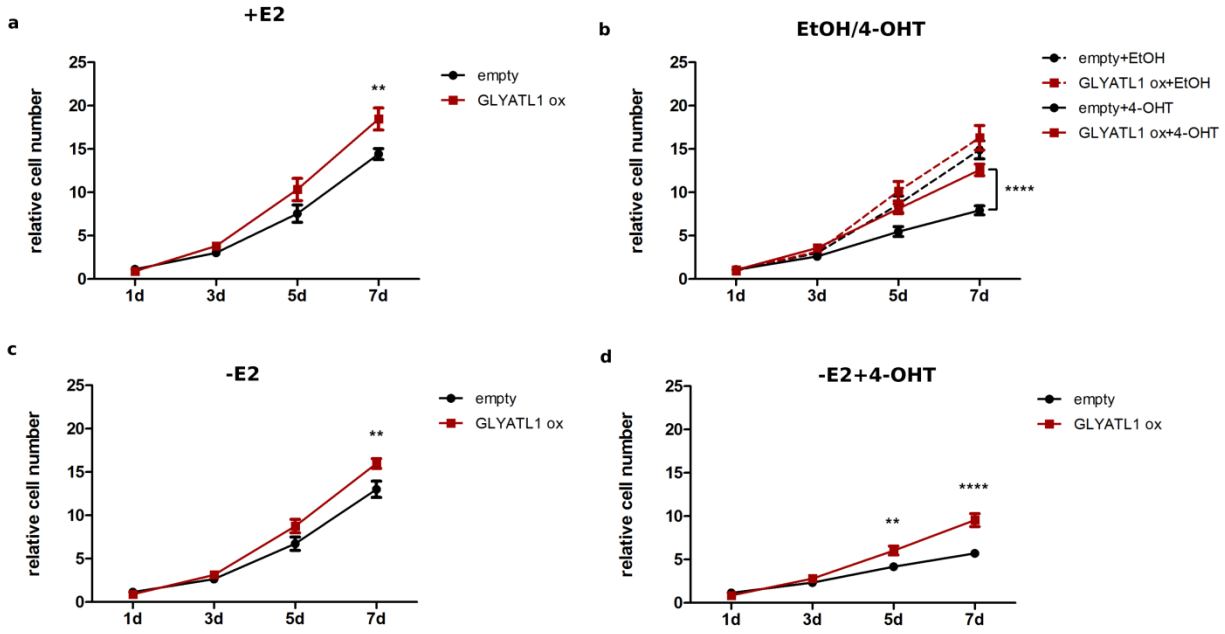


Figure 19: Effect of GLYATL1 overexpression on MCF7 cell proliferation under endocrine therapy conditions. (a-d) MCF7 GLYATL1 overexpressing (GLYATL1 ox) and empty vector control (empty) cells were seeded and treated with different growth media to simulate endocrine therapy conditions. Microscopy based nuclei counting was performed at different time points during 7 days. All values for proliferation assays were normalized to seeding controls. Data are presented as mean ± SEM, n=4 (each with 5 technical replicates). **** represents $p<0.0001$, *** represents $p<0.001$, ** represents $p<0.01$.

The effect of GLYATL1 overexpression followed a similar trend for T47D cells. GLYATL1 overexpressing T47D cells grew significantly better under tamoxifen treatment conditions and in estrogen deprivation conditions, especially at day 7 of treatment. T47D GLYATL1 overexpressing cells did not show any significant change in proliferation dynamics when exposed to both tamoxifen and estrogen deprivation at the same time (Figure 20).

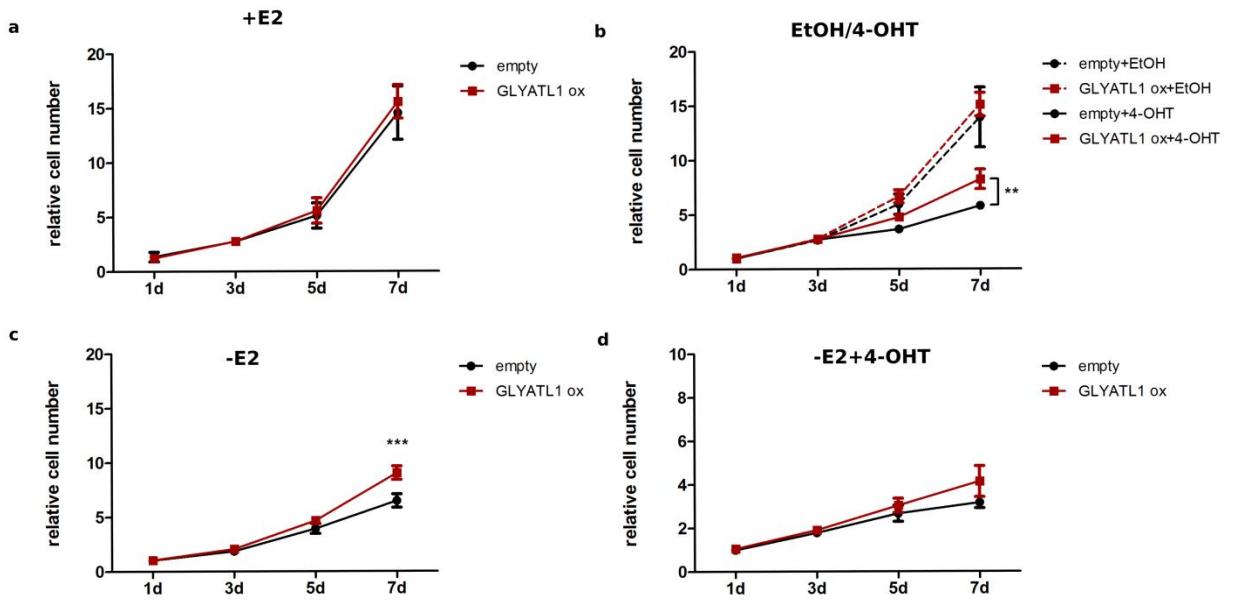


Figure 20: Effect of GLYATL1 overexpression on T47D cell proliferation under endocrine therapy conditions. (a-d) T47D GLYATL1 overexpressing (GLYATL1 ox) and empty vector control (empty) cells were seeded and treated with different growth media to simulate endocrine therapy conditions. Microscopy based nuclei counting was performed at different time points during 7 days. All values for proliferation assays were normalized to seeding controls. Data are presented as mean ± SEM, n=3 (each with 5 technical replicates). *** represents $p<0.001$, ** represents $p<0.01$.

4.1.4.3 Effect of GLYATL1 overexpression on cell cycle

In order to elucidate the changes in proliferation dynamics of GLYATL1 overexpressing cells compared to their empty vector controls, I decided to check the cell cycle of the these cells under normal growth conditions and growth conditions mimicking endocrine therapy. For this purpose, EdU incorporation assay (Click-iT EdU Imaging Kit) was utilized. This assay simply provides detection of cells undergoing S-phase of the cell cycle via fluorescent tagged antibodies targeting incorporated EdU. I could determine the percentage of the cells in S-phase via combining this assay with Hoechst staining of the cells. EdU incorporation was performed in order to distinguish any changes in cell cycle kinetics following a 24h pulse. In short, the percentages of the GLYATL1 overxpressing MCF7 cells undergoing S-phase was found to be slightly more in

normal growth conditions. While a decrease in the percentages of cycling cells was observed under tamoxifen treatment, estrogen deprivation or simultaneous administration of both, GLYATL1 overexpressing cells had significantly more cells cycling through S-phase (Figure 21a).

T47D cells demonstrated a similar trend, as the percentages of cells going through S-phase were found to be higher in GLYATL1 overexpressing cells under endocrine therapy conditions. However, the difference here was not significant (Figure 21b).

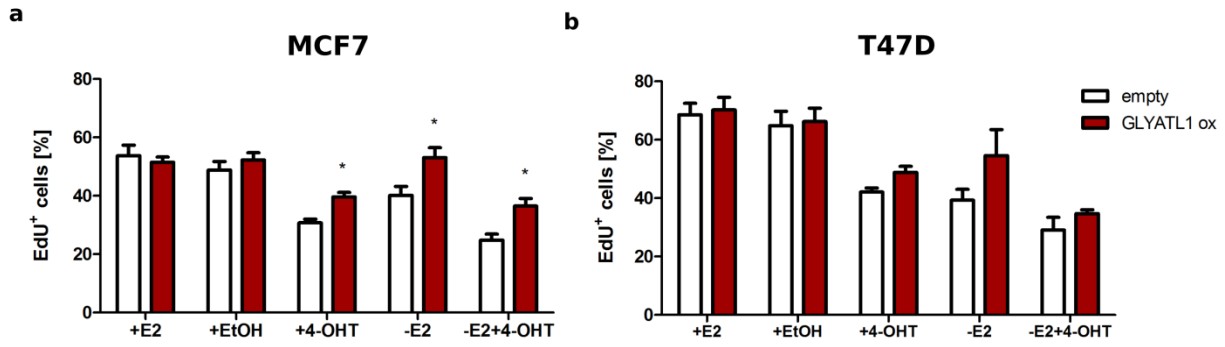


Figure 21: Effect of GLYATL1 overexpression on cell cycle. MCF7 (a) and T47D (b) cells were pulsed with EdU for 24h. Percentages of EdU^+ cells were determined by microscopy-based counting. Data are presented as mean ± SD, n=3. * represents $p<0.05$.

4.1.4.4 Effect of GLYATL1 overexpression on apoptosis

Next, I wanted to check the changes in apoptosis (namely late apoptosis), which could be accounted for the differential growth dynamics of overexpressing cells. For this purpose, percentage of cells undergoing late apoptosis was determined by PI addition to the cells before microscopic nuclei counting with Hoechst staining. Here, the percentage of late apoptotic cells for MCF7 empty vector were found to be increasing as the cells are introduced to tamoxifen, estrogen deprivation or both at the same time whereas GLYATL1 overexpressing cells had significantly less apoptotic cells under endocrine therapy conditions (Figure 22a).

For T47D GLYATL1 overexpressing cells, a similar trend was observed regarding apoptosis. Percentage of late apoptotic cells was significantly less under tamoxifen treatment and estrogen

deprivation whereas no significant change was observed when GLYATL1 overexpressing cells were challenged with both endocrine therapy conditions at the same time (Figure 22b).

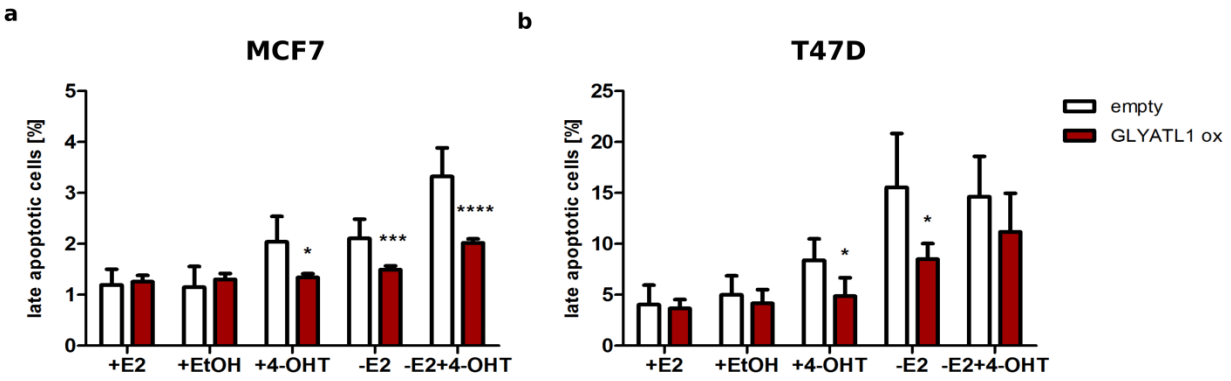


Figure 22: Effect of GLYATL1 overexpression on apoptosis. Percentage of late apoptotic cells were determined via PI staining combined with Hoechst-based microscopic nuclei count. Data are presented as mean ± SD, n=3 for MCF7 and n=2 for T47D (each with 4 technical replicates). **** represents $p<0.0001$, *** represents $p<0.001$, * represents $p<0.05$.

Taken together, these data suggest overexpression of GLYATL1 confers resistance to estrogen deprivation and especially tamoxifen treatment as revealed with differing proliferation dynamics for both MCF7 and T47D cells. Assays assessing cell cycle kinetics and apoptosis further confirmed the changes in dynamics in overexpressing cells indicating they cope better with endocrine therapy conditions compared to their empty vector control.

4.1.5 Investigating the role of GLYATL1 as a potential HAT

The STRING database suggests KAT2A and KAT2B as potential interaction partners of GLYATL1 (Szklarczyk et al., 2019). KAT2A and KAT2B are histone acetyltransferases belonging to GCN5-related N-acetyltransferases (GNATs) (Figure 23).

These HATs predominantly catalyze acetylation of H3 or H4 tail lysine residues with a preference for H3K9 and H3K14 (Howe et al., 2001; Jin et al., 2011). For this reason, first I re-assessed the RNA-Seq data with a focus on expression levels of HATs. Among the probed HATs, *KAT2B* was found to be the one that was the most upregulated in all MCF7 resistant cells compared to parental. Interestingly, *KAT2A* expression was upregulated only in MCF7 TAMR and LTEDTAMR. In contrast, in T47D cells, although both HATs were downregulated, *KAT6B*,

a HAT from MYST family, was found to be upregulated for both resistance models (TAMR and LTED). *KAT6A* was found to be upregulated only at the latest time-point (7 months) probed for T47D resistance acquisition with RNA-sequencing. *CREBBP* and *EP300* (p300/CBP HATs) were also found to be upregulated in later time-points for T47D and for MCF7 LTED and LTEDTAMR cells (Figure 24).

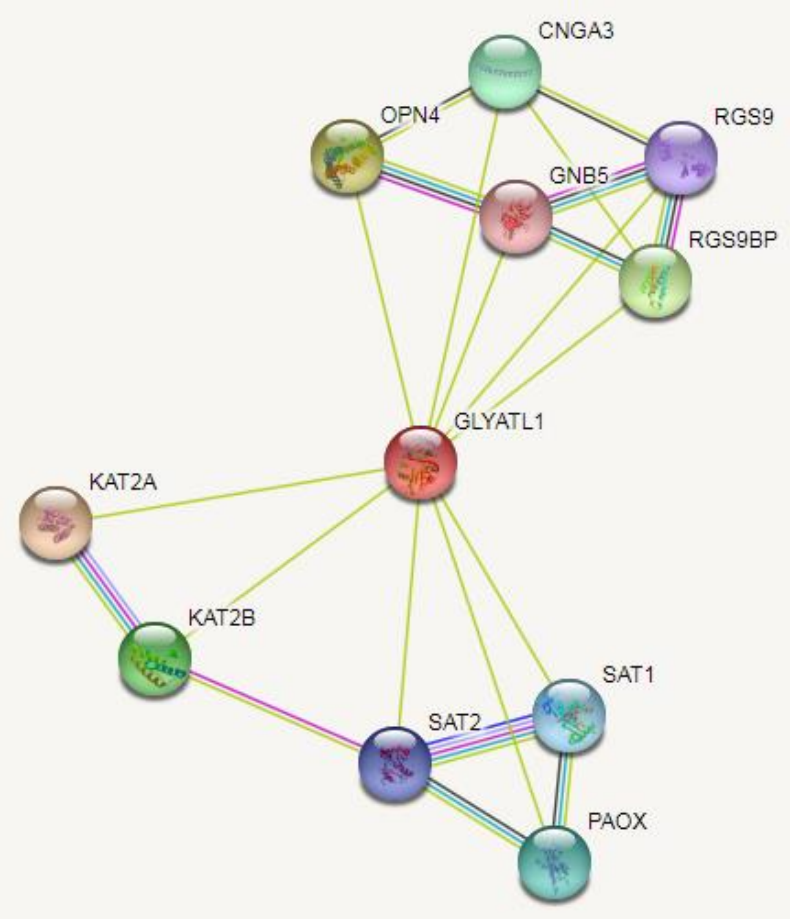


Figure 23: Putative interaction partners of GLYATL1. Possible interaction partners of GLYATL1 were determined using STRING database (version 11.0, accessed on 21.06.2019) (Szklarczyk et al., 2019).

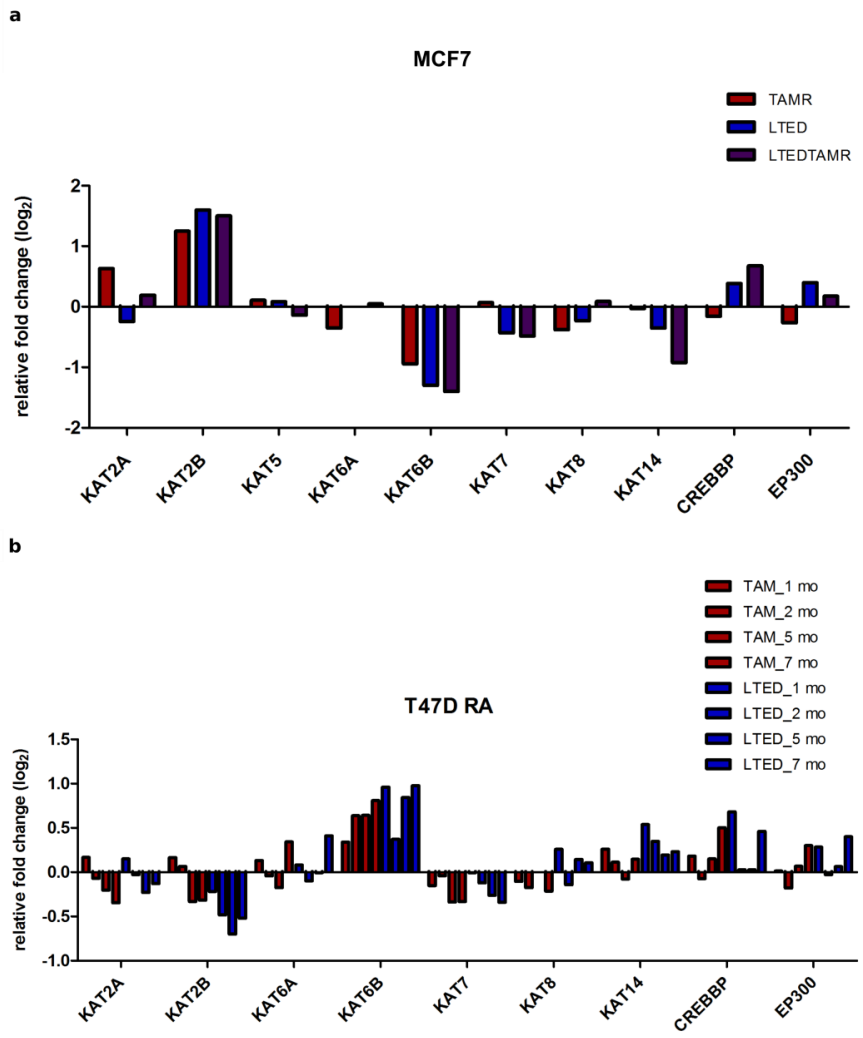


Figure 24: Expression patterns of histone acetyltransferases in resistant MCF7 and T47D cell lines. Expression levels for family members of HATs in MCF7 (a) and T47D (b) cells were determined via RNA-Sequencing. All values are represented as relative values normalized to their respective WT.

Given the putative function of GLYATL1 as a histone acetyltransferase (or taking part in a multi-subunit HAT complexes), I wanted to check whether there was a global increase in histone 3 (H3) acetylation. For this purpose, histones were extracted and changes in pan-histone 3 acetylation patterns were assessed by western blotting. Pan-acetylation of H3 was found to be increased especially in LTED cells compared to parental MCF7 (Figure 25).

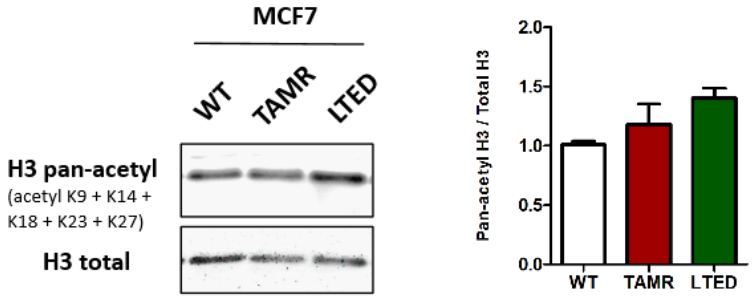


Figure 25: Pan-acetylation of H3 in MCF7 cell line repertoire. Pan-acetylation of histone 3 was determined via western blotting of purified core histones. Protein expression levels were normalized to total H3 levels (n=2).

ZMYND8, a gene encoding the epigenetic reader of H3K14ac mark (Savitsky et al., 2016), also came out of the MCF7 ATAC-Seq and RNA-Seq correlation analysis as one of the top significantly gained peaks around TSS and upregulated genes. This gene was also found to be upregulated for T47D resistance acquisition samples (Figure 26). Intriguingly, KAT2A, KAT2B, KAT6A and KAT6B are also epigenetic readers of histone marks H3K9ac and H3K14ac in addition to their role as histone acetyltransferases introducing these marks (Gong et al., 2016). That is the reason why I chose to focus specifically on acetylation of these two histone tail residues.

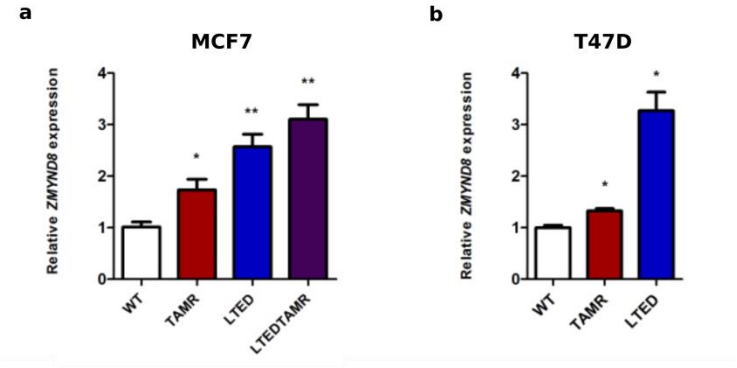


Figure 26: *ZMYND8* expression levels in MCF7 and T47D cell line repertoire. *ZMYND8* mRNA levels in MCF7 (a) and T47D (b) were determined via qRT-PCR. Values for mRNA expression were first normalized to PUM1 and ACTB levels and are represented as mean ± SD, n=3 (each with 3 technical replicates). ** represents $p<0.01$, * represents $p<0.05$.

GLYATL1 has been reported to to localize to mitochondria (Matsuo et al., 2012) while the Human Protein Atlas suggests localization in the cytosol and Golgi (Uhlen et al., 2017) GeneCards and the compartments database annotate also the nucleus as subcellular location, based on a prediction by PSORT (Binder et al., 2014; Database GeneCards, 2017). In order to further investigate the potential role of GLYATL1 in histone acetylation, I next checked whether it localizes to the nucleus or not. Western blotting results of cytoplasmic and nuclear fractions from GLYATL1 overexpressing cells and their empty vector controls showed GLYATL1 both in cytoplasmic and nuclear fractions of MCF7 and T47D (Figure 27). This finding strengthened the hypothesis that GLYATL1 might indeed have a potential role in the nucleus, such as histone acetylation.

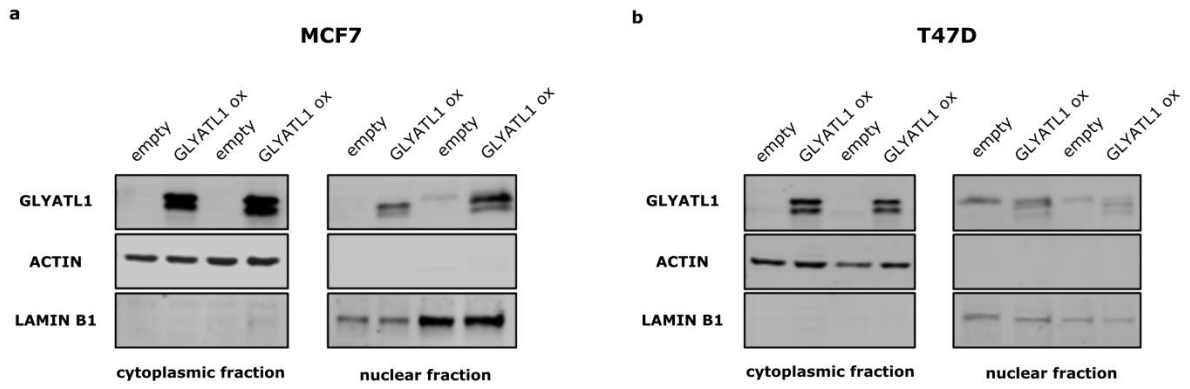


Figure 27: Nuclear and cytoplasmic localization of GLYATL1 in MCF7 and T47D. GLYATL1 protein was detected in cytoplasmic and nuclear fractions of MCF7 (a) and T47D (b) GLYATL1 overexpressing (GLYATL1 ox) cell lines. Actin was used as a loading control for cytoplasmic fraction whereas Lamin B1 was a loading control for nuclear fraction (n=2).

Next step was to assess the changes in acetylation of H3K9 and H3K14 residues in MCF7 cell repertoire. For this, core histones were purified and alterations in H3K9ac and H3K14ac marks were quantified by histone ELISA. Each histone modification mark was normalized to total H3 levels to negate any fluctuation in the probed histone mark due to abundance of total histone proteins. Both histone acetylation marks were found to be increased in TAMR and LTED cells compared to parental MCF7 (Figure 28a,c). H3K14ac increase was significant for both TAMR and LTED, whereas H3K9 was significantly more acetylated only in LTED cells. Furthermore, GLYATL1 overexpressing cells showed increased acetylation for both histone marks (Figure 28b,d).

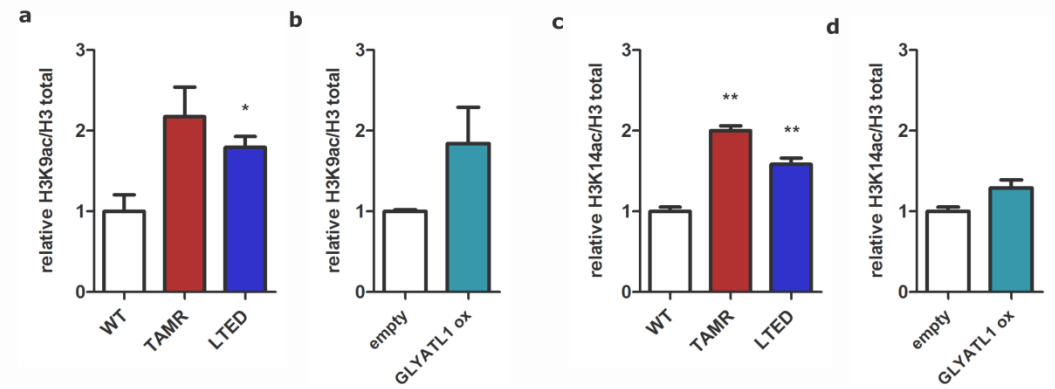


Figure 28: H3K9ac and H3K14ac levels in MCF7 and T47D cells. H3K9ac (a,b) and H3K14ac (c,d) levels were detected by Histone ELISA in purified core histones of MCF7 parental and resistant cell lines and MCF7 GLYATL1 overexpressing cells, respectively. Values for each histone mark was normalized to total H3 levels detected. Data are represented as mean ± SD, n=3 (each with 2 technical replicates). ** represents $p<0.01$, * represents $p<0.05$.

In order to link the increase in H3K9ac and H3K14ac to an increase in GLYATL1 levels, the changes in both marks were investigated via histone ELISA upon RNA interference. *GLYATL1* knockdown on both MCF7 TAMR and LTED cells resulted in significantly less H3K9 and H3K14 acetylation, confirming the role of GLYATL1 as part of the machinery introducing these marks (Figure 29).

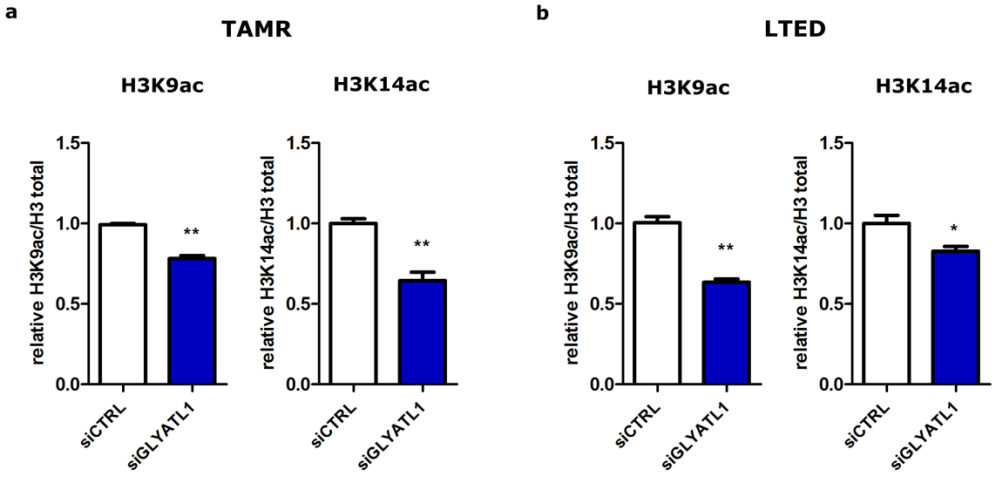


Figure 29: Effect of *GLYATL1* knockdown on H3K9ac and H3K14ac levels in MCF7 TAMR and LTED cells. MCF7 TAMR (a) and LTED (b) cells were transfected with siGLYATL1 and non-targeting siRNA pool (siCTRL). Core histones were purified 96h post-transfection. H3K9ac and H3K14ac levels were determined by Histone ELISA. Values for each histone mark was normalized to total H3 levels detected. Data are represented as mean ± SD, n=3 (each with 2 technical replicates). ** represents $p<0.01$, * represents $p<0.05$.

Moreover, I wanted to check whether loss of GLYATL1 has an effect on expression of two family members of GCN5-related N-acetyltransferases (GNATs). Interestingly, *KAT2B* expression was significantly downregulated whereas a slight (albeit not significant) increase in *KAT2A* levels was observed upon *GLYATL1* knockdown in MCF7 LTED cells. On the other hand, overexpression of GLYATL1 resulted in an opposite trend for these two HATs where *KAT2A* was significantly downregulated and *KAT2B* was significantly upregulated (Figure 30).

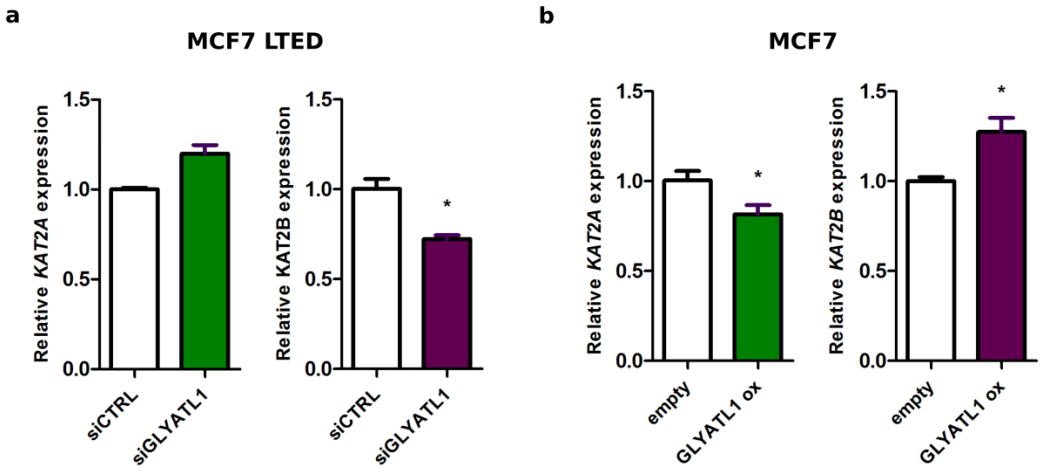


Figure 30: Effect of *GLYATL1* knockdown and GLYATL1 overexpression on *KAT2A* and *KAT2B* expression levels. (a) MCF7 LTED cells were transfected with siGLYATL1 and non-targeting siRNA pool (siCTRL). Total RNA were purified 72h post-transfection. *KAT2A* and *KAT2B* expression levels were determined from *GLYATL1* knockdown and overexpression samples (b) via qRT-PCR. Values for mRNA expression were normalized to *PUM1* and *ACTB* levels. Data are represented as mean ± SD, n=3 (each with 3 technical replicates). * represents $p<0.05$.

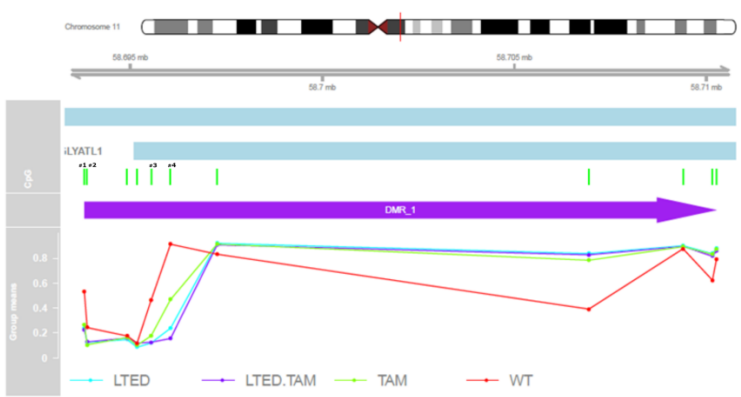
Altogether, these findings suggest a possible role for GLYATL1 involving in introduction of histone marks H3K9ac and H3K14ac in MCF7 endocrine therapy resistant cell lines, possibly through interaction with GCN5-related HATs, especially KAT2B.

4.1.6 Regulation of *GLYATL1*

4.1.6.1 Epigenetic Regulation

The above results demonstrate the importance of *GLYATL1* in the context of endocrine therapy resistance since knockdown of this gene results in partial re-sensitization of the MCF7 resistant cells and overexpression renders the sensitive cells more resistant to endocrine therapy conditions while hinting at the involvement of GLYATL1 in histone acetylation. To understand whether this gene itself might be under epigenetic regulation or not I next investigated the EPIC array data generated for resistant cell line repertoire. The analysis results for MCF7 cells (analysis done by Maryam Soleimani-Dodaran) indeed revealed significantly differentially methylated CpGs spanning the promoter region and the 1st exon of *GLYATL1*. Majority of the differentially methylated CpGs were found to be hypomethylated (between 11.5% - 75.6% hypomethylation) in all MCF7 resistant cells compared to their sensitive counterpart (Figure 31). This hypomethylation could explain the observed increase in expression of this gene in resistant MCF7 cells.

a



b

CpG	Position (hg19)	TAMR vs. WT (%)	LTED vs. WT (%)	LTEDTAMR vs. WT (%)
cg24440196 **(#1)**	58693803	-28.7630784	-29.758	-30.7530967
cg19111544 **(#2)**	58693878	-12.6250418	-12.45854	-11.5203101
cg06080398 **(#3)**	58695542	-31.0534408	-34.59911	-33.6117423
cg04226898 **(#4)**	58696048	-60.049961	-73.03169	-75.6361017

Figure 31: Methylation status of *GLYATL1* promoter region. (a) Schematic depiction of methylation changes in MCF7 cell line repertoire in GLYATL1 differentially methylated region (DMR). Methylation of CpGs covering GLYATL1 promoter region and 1st exon were determined via Illumina Methylation EPIC 850k array. (b) CpGs that are significantly differentially methylated are depicted with their genomic location (GRCh37/hg19 build) and changes in methylation levels (%) compared to parental MCF7.

In order to validate regulation of *GLYATL1* via methylation, I first treated MCF7 and T47D parental cells with 5-Aza-dC (5-aza-2'-deoxycytidine), a global DNMT1 inhibitor, for 72 hours and checked the changes in expression levels for *GLYATL1*. Interestingly, *GLYATL1* expression was significantly upregulated (approximately 600-fold increase) upon induction of DNA hypomethylation (Figure 32a). This increase was reminiscent of the increase in MCF7 TAMR cells compared to their sensitive counterpart. A moderate upregulation of *GLYATL1* expression was also observed in T47D cells upon 5 μM 5-Aza-dC treatment (Figure 32b). However, EPIC array results did not indicate any significant methylation changes in probed CpGs between months 1-7 of resistance acquisition in T47D (data not shown). These results suggested that methylation might be responsible for regulation of *GLYATL1* in MCF7 cells, however, not in T47D.

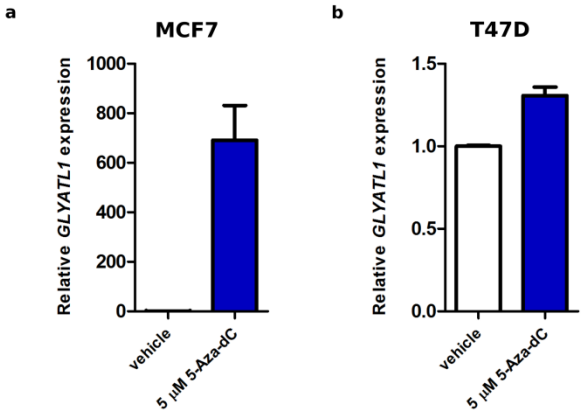


Figure 32: Effect of 5-Aza-dC treatment on MCF7 and T47D parental cells. MCF7 (a) and T47D (b) cells were treated either with 5 μM 5-Aza-dC or vehicle control (water) for 72h. Total RNA was extracted and *GLYATL1* expresssion levels were determined via qRT-PCR. All values are normalized to vehicle control. mRNA expression values were first normalized to *PUM1* and *ACTB* levels. Data are represented as mean ± SD, n=2 (each with 3 technical replicates).

In order to explore and elucidate the epigenetic layer of regulation on *GLYATL1*, CRISPR/dCas9-mediated targeted epigenetic editing method was adopted (establishment of this method is explained more in detail in Part II of results). For this purpose, I designed five different sgRNAs covering CpGs that were probed by EPIC array and that were found to be significantly hypomethylated (Figure 33). I then utilized dCas9 fused to the catalytic domain of DNA demethylase TET1 or reduced activity mutant of prokaryotic *M.SssI* DNA methyltransferase (*M.SssI*-Q147L). *M.SssI* is a bacterial methyltransferase with a high enzymatic

activity. However it has been recently shown that the enzymatic activity can be reduced (~10%) by a point mutation (Q147L) which renders the enzyme more specific for introducing methylation on target loci. This specificity thereby overcomes off-target methylation spreading effects of eukaryotic DNMTs (Lei et al., 2017).

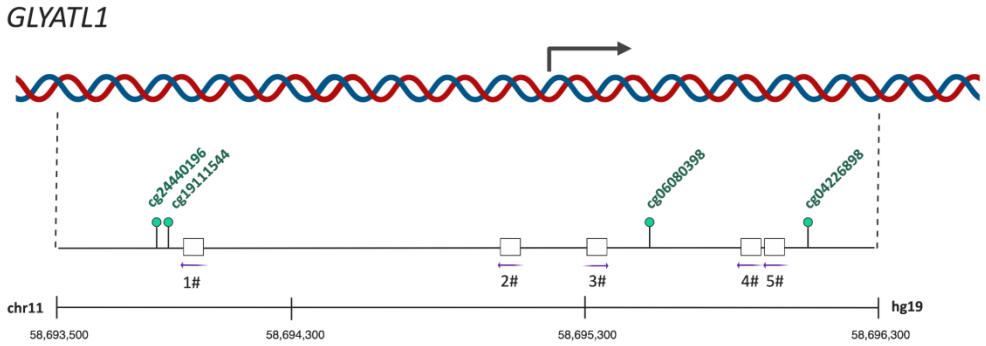


Figure 33: Designed sgRNAs for *GLYATL1*. Schematic depiction of location of sgRNAs designed to target *GLYATL1* promoter region with CRISPR/dCas9-mediated epigenetic editing with their genomic locations and proximity to annotated differentially methylated CpGs in MCF7 resistant cell lines. Arrows point the PAM sequence and indicate the directionality of sgRNA binding. Arrow on the genome depicts TSS.

Utilizing doxycycline-inducible MCF7 cells stably expressing dCas9 fused to the catalytic domain of TET1, I could verify an increase in *GLYATL1* expression after 10 days of transfection with sgRNAs targeting differentially methylated CpGs. This increase was found to be significant for two individual sgRNAs used (Figure 34). Moreover, the upregulation was found to be higher, albeit not significant, when CpGs were targeted with the combination of 5 sgRNAs.

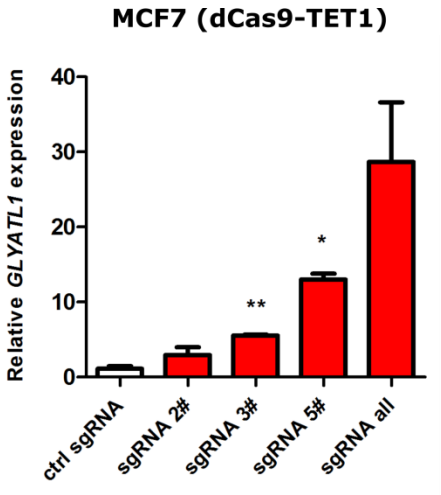


Figure 34: Effect of epigenetic editing on *GLYATL1* expression in MCF7 cells. MCF7 cells stably expressing inducible dCas9 fused to TET1 were transfected with three individual sgRNAs as well as combination of all targeting CpGs of *GLYATL1* and control sgRNA. (a) *GLYATL1* mRNA levels were determined by qRT-PCR 10 days after transfection. All values are normalized to control sgRNA. mRNA expression values were first normalized to *PUM1* and *ACTB* levels Data are represented as mean ± SD, n=2 (each with 3 technical replicates). ** represents p<0.01, * represent p<0.05.

Next, I wanted to check whether introduction of methylation would suffice to suppress *GLYATL1* expression in resistant cell lines. For this purpose, doxycycline-inducible MCF7 LTED cells stably overexpressing dCas9-Q147L were transfected with sgRNAs. Three individual sgRNAs significantly downregulated *GLYATL1* expression.(Figure 35a). Downregulation of GLYATL1 was found to confer LTED cells a proliferative disadvantage as seen in retarded growth of LTED cells tranfected either with sgRNA 1# or sgRNA 3# compared to control sgRNA (Figure 35b,c).

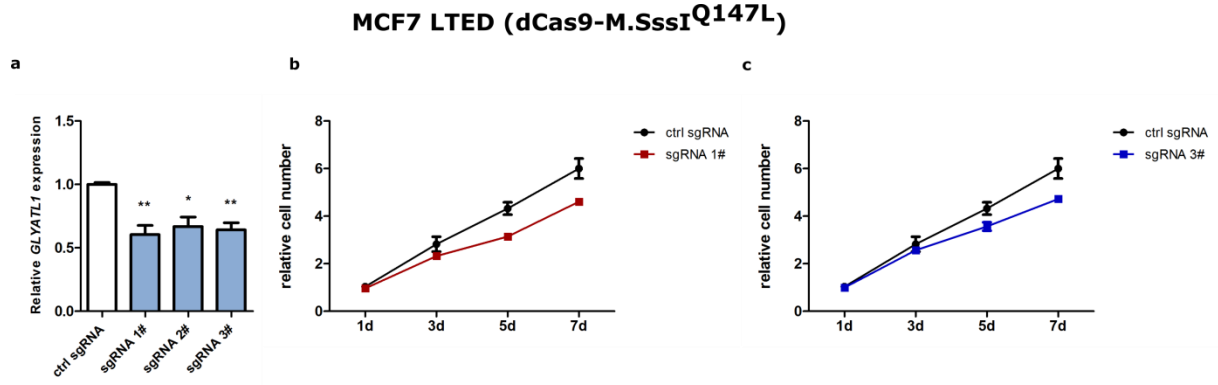


Figure 35: Effect of epigenetic editing on *GLYATL1* expression in MCF7 LTED cells. (a) MCF7 LTED cells stably expressing inducible dCas9 fused to M.SssI (Q147L) were transfected with sgRNAs targeting CpGs of *GLYATL1* and control sgRNA. *GLYATL1* mRNA levels were determined by qRT-PCR 7 days after transfection. (b,c) Cell numbers were determined via microscopy-based nuclei count at different time-points over a course of 1 week following transfection with two individual sgRNAs and control sgRNA. All values of mRNA expression are normalized to control sgRNA. mRNA expression values were first normalized to *PUM1* and *ACTB* levels. Data are represented as mean ± SD, n=2 (each with 3 technical replicates) All values for proliferation assays were normalized to seeding controls. n=1 (with 5 technical replicates). ** represents p<0.01, * represents p<0.05.

I also wanted to investigate whether methylation would have a similar effect also in MCF7 TAMR cells. For this purpose, doxycycline-inducible MCF7 TAMR cells stably overexpressing dCas9-Q147L were transfected with two individual sgRNAs (sgRNA 1# and 2#), which had been found to be effective in inducible LTED cells. Results indicate a significant downregulation

in *GLYATL1* expresssion after 7 days of transfection (Figure 36a). Interestingly, proliferation assay of TAMR cells over a week after transfection revealed that the decrease in *GLYATL1* expression levels rendered these cells more susceptible to tamoxifen treatment (4-OHT) as they proliferated less in the presence of tamoxifen compared to cells transfected with the same sgRNA but treated with vehicle control (EtOH) or the cells transfected with control sgRNA (Figure 36b,c). This finding further confirmed that downregulation of *GLYATL1* results in re-sensitization of resistant cells to endocrine therapy, in particular to tamoxifen exposure.

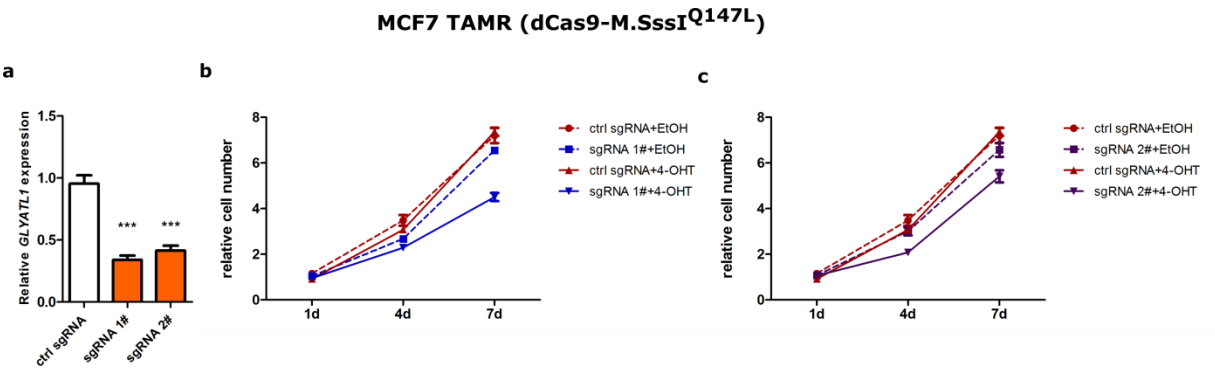


Figure 36: Effect of epigenetic editing on *GLYATL1* expression in MCF7 TAMR cells. MCF7 TAMR cells stably expressing inducible dCas9 fused to M.SssI (Q147L) were transfected with two individual sgRNAs targeting CpGs of *GLYATL1* and control sgRNA. (a) *GLYATL1* mRNA levels were determined by qRT-PCR 7 days after transfection. (b,c) Cell numbers were determined via microscopy-based nuclei count at different time-points over a course of 1 week following transfection with two individual sgRNAs and treatment with either 4-OHT or vehicle control (EtOH). All values of mRNA expression are normalized to control sgRNA. mRNA expression values were first normalized to *PUM1* and *ACTB* levels. Data are represented as mean ± SD, n=2 (each with 3 technical replicates). All values for proliferation assays were normalized to seeding controls. n=1 for proliferation assay (with 5 technical replicates). *** represents $p<0.001$.

4.1.6.2 Regulation by HER2

Investigation of public breast cancer patient datasets (Collins, 2007; Curtis et al., 2012) and of the breast cancer cell line panel of 42 individual cell lines representing different molecular subtypes showed that *GLYATL1* expression is higher in luminal, however, also in HER2-enriched subtypes (Figure 12). This finding led to the hypothesis that HER2 might play a role in regulation of *GLYATL1*. In order to test this hypothesis, first I analyzed the mRNA expression levels of *ERBB2* in the MCF7 cell line repertoire and compared these with expression levels of *GLYATL1*. Both genes showed a similar trend in expression as they both were upregulated in

TAMR, LTED and LTEDTAMR cells, LTED having the highest expression levels for both genes (Figure 37a,b). Next, RNA interference targeting *ERBB2* was applied to assess a putative role of *ERBB2* in the regulation of *GLYATL1*. Knockdown of *ERBB2* in both TAMR and LTED cells of MCF7 origin resulted in significant downregulation of *GLYATL1*, suggesting an involvement of *ERBB2* in the regulation of *GLYATL1* (Figure 37c,d).

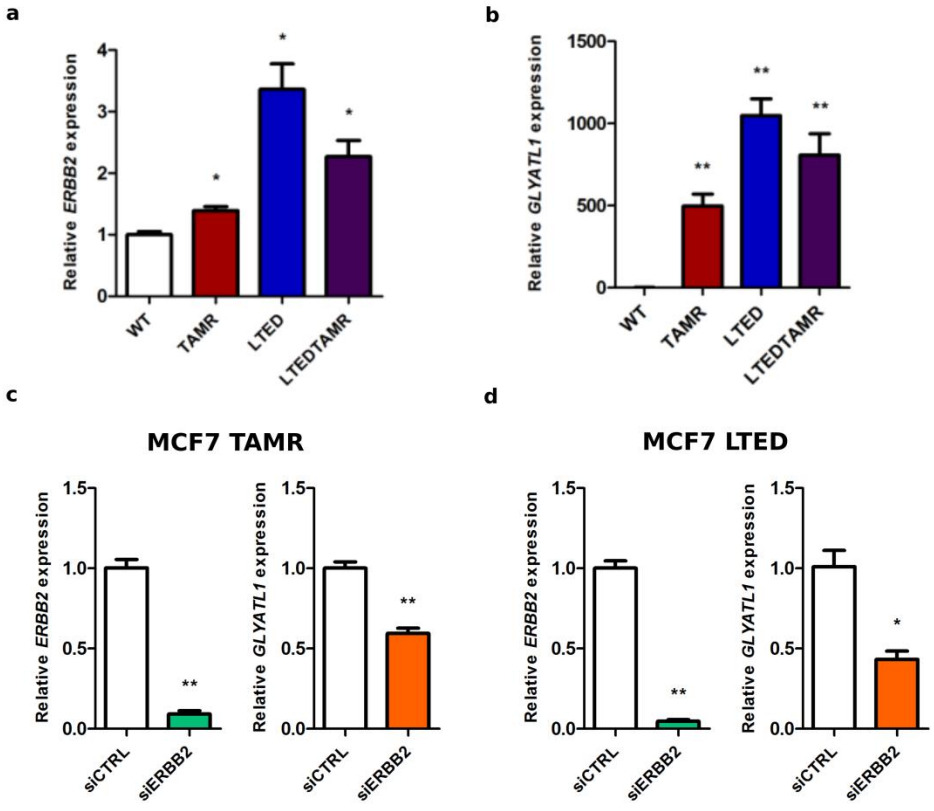


Figure 37: Effect of *ERBB2* knockdown on *GLYATL1* expression. Levels of *ERBB2* (a) and *GLYATL1* (b) mRNA expression in MCF7 cell line repertoire were determined by qRT-PCR. MCF7 TAMR (c) and LTED cells (d) were transfected with siERBB2 and non-targeting control siRNA (siCTRL). *GLYATL1* mRNA levels were determined by qRT-PCR 72h after transfection. All values of mRNA expression are normalized to parental MCF7 (a,b) or siCTRL (c,d). mRNA expression values were first normalized to *PUM1* and *ACTB* levels. Data are represented as mean ± SD , n=3 (each with 3 technical replicates). ** represents $p<0.01$, * represents $p<0.05$.

4.1.6.3 Regulation by transcription factors

Next step was to investigate the involvement of transcription factors associated with luminal subtype in regulation of *GLYATL1* expression. For this, I chose to focus on *ESR1* and *GATA3*. I also included *EP300* as it was shown to code for a transcription factor with a predicted binding

site in the promoter region of *GLYATL1* based on UCSC Genome Browser (Kent et al., 2001) (Figure 38a). I went to check the expression levels of these genes in resistant MCF7 cells compared to parental. RNA-sequencing data revealed that both *ESR1* and *EP300* were downregulated in TAMR cells whereas LTED cells had elevated expression of these genes. GATA3, on the other hand, was found to be downregulated in both *TAMR* and *LTED* cells.

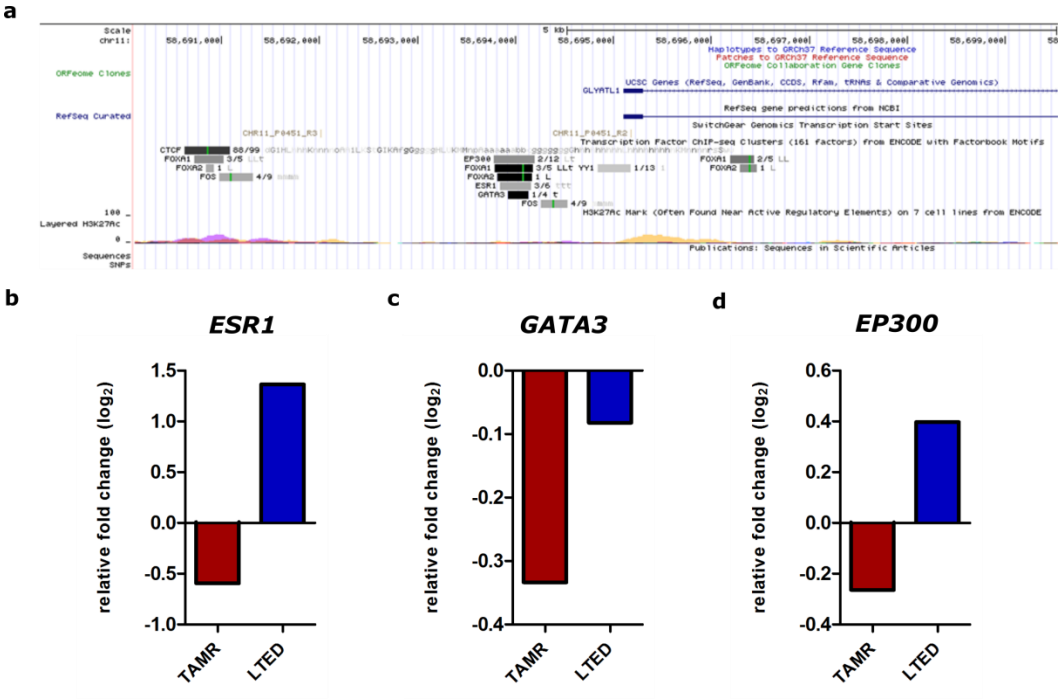


Figure 38: Transcription factor occupancy of *GLYATL1* promoter. (a) Transcription factors with potential binding sites for GLYATL1 promoter were assessed based TF-ChIP data of UCSC Genome Browser (Kent et al., 2001)**.** Expression levels for (b) *ESR1*, (c) *GATA3*, (d) *EP300* in resistant MCF7 cells were determined via RNA-Sequencing. All values are represented as relative values normalized to their respective WT.

Knockdown of each transcription factor resulted in a significant decrease in *GLYATL1* expression levels in both MCF7 TAMR and LTED cells, apart from siGATA3 in MCF7 LTED cells (Figure 39).

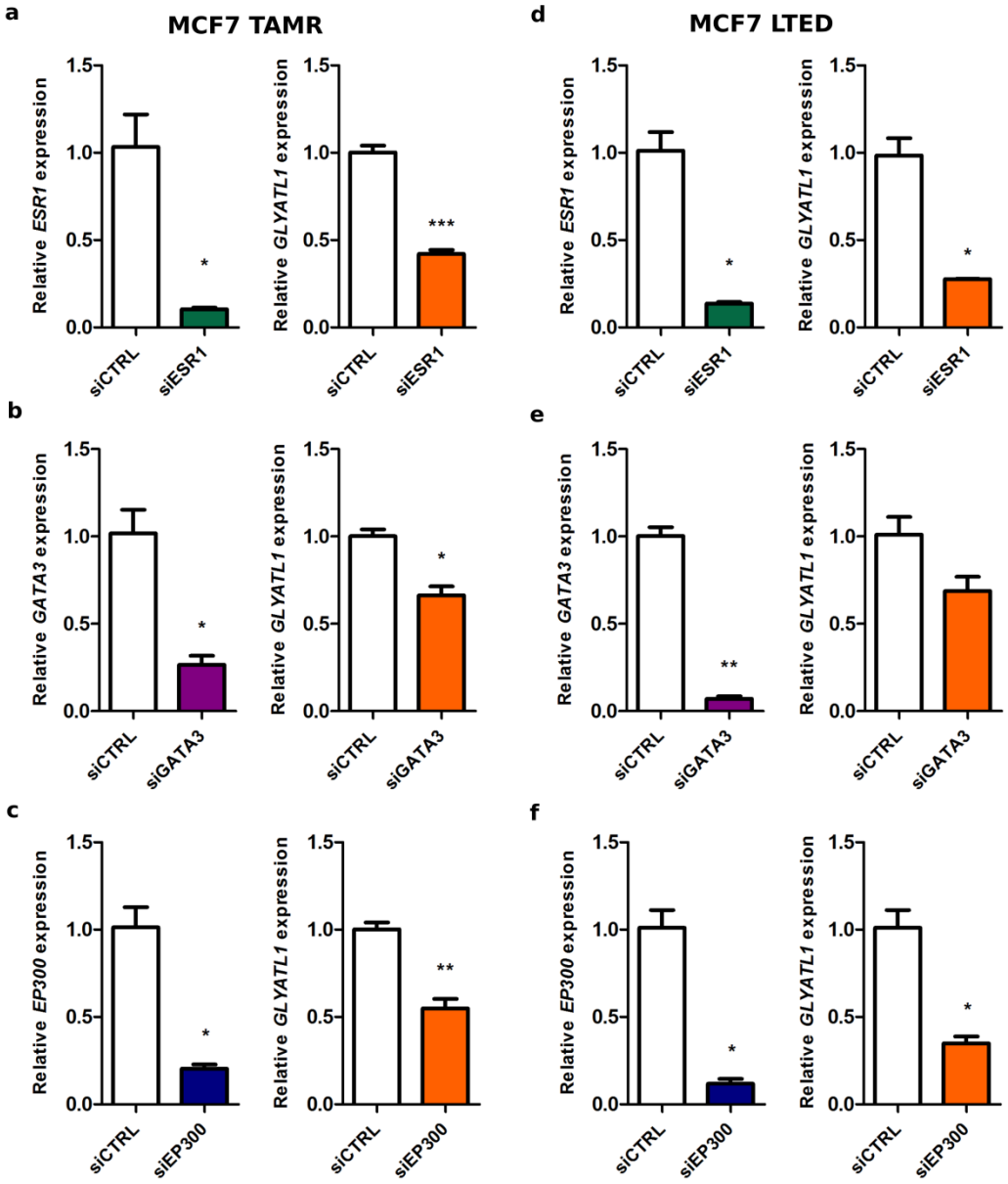


Figure 39: Effect of transcription factor knockdown on *GLYATL1* expression. MCF7 TAMR and LTED cells were transfected with siESR1 (a,d), siGATA3 (b,e), siEP300 (c,f) and non-targeting control siRNA (siCTRL). *GLYATL1* mRNA levels were determined by qRT-PCR 72h after transfection. All values of mRNA expression are normalized to siCTRL. mRNA expression values were first normalized to *PUM1* and *ACTB* levels. Data are represented as mean ± SD, n=3 (each with 3 technical replicates). *** represents $p<0.001$, ** represents $p<0.01$, * represents $p<0.05$.

Taken altogether, these results suggest that *GLYATL1* is an ER-responsive gene with implications of methylation, HER2, luminal pioneer factor GATA3 and histone acetyltransferase p300 playing a role in regulating its transcription.

4.2 Part II: Establishing CRISPR/dCas9-mediated Epigenetic Editing Methodology to Modulate the Epigenetic Landscape of Endocrine Therapy Resistant Breast Cancer

4.2.1 CRISPR/dCas9-mediated targeted epigenetic editing

Epigenetic editing can be defined as any alteration on epigenetic level which in turn can be exploited to regulate expression of specific genes. Genome can be epigenetically interfered by complexes comprised of specific DNA recognition domains such as zinc finger, transcription activator-like (TAL) effector or CRISPR/dCas9 and chromatin modifying enzyme with catalytic activity (Kungulovski & Jeltsch, 2016). In this study I aimed to utilize epigenetic interference tools based on the CRISPR/dCas9 system with deactivated Cas9 (dCas9) coupled to catalytic domains of DNA or histone modifying enzymes. to induce epigenetic changes at specific genomic loci implicated in endocrine therapy resistance.. For this approach, I adopted a hit-and-run approach in which the cells were transiently transfected with both a sgRNA plasmid and another plasmid encoding dCas9 fused with an effector domain (epi-effector). All dCas9-effector domain constructs I used contain a self-cleavable mCherry reporter tag and were provided by Mihaly Koncz (MTA-SZBK) or Marianne Rots (UMCG). The general experimental procedure involved transfection for 48h, collection and validation of epigenetic interference at the mRNA level (TaqMan) and re-seeding of cells to assess potential long term effects. Transient transfection efficiency of dCas9 constructs were modest for MCF7 cells (approximately 30% for parental MCF7 and 40% for MCF7 LTED) (Figure 40a,b). I also utilized cells stably transduced with doxycycline-inducible dCas9 fused to catalytic domain of an epigenetic modifier where upon doxycycline induction the cells were only transfected with sgRNA plasmids (thereby increasing the transfection efficiency) (Figure 40c).

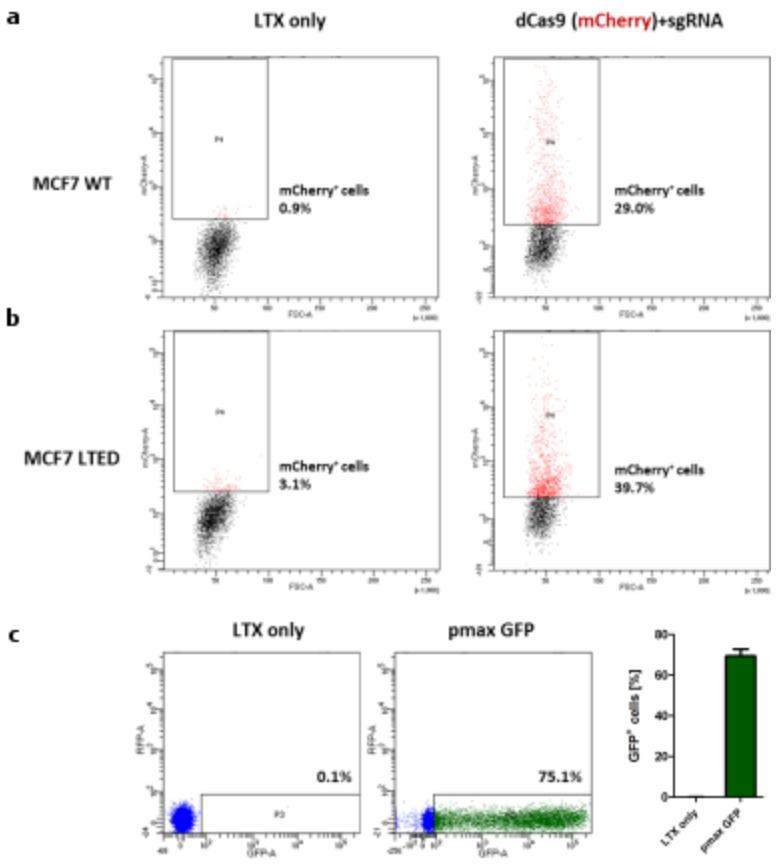


Figure 40: Transfection efficiency of dCas9-effector domain plasmids and sgRNA plasmid. Transfection efficiency of dCas9 plasmids were determined by mCherry^{+} cells in flow cytometry in (a) MCF7 parental and (b) LTED cells. (c) sgRNA plasmid transfection efficiency was determined via GFP^{+} cells in flow cytometry upon transfection with a similar-sized plasmid (pmaxGFP). Flow cytometry dot blot gating was done according to the transfection reagent control. LTX only:Transfection reagent control.

In order to develop the epigenetic interference method in the context of endocrine therapy resistance, I chose target genes *SLC9A3R1* and *CD44* based on recent findings from project partner Dr. Luca Magnani's lab (ICL) and prior knowledge, respectively, and *BAMBI* based on RNA-seq data of resistant cell line repertoire (Refer to Part I).

4.2.1.1 *SLC9A3R1*

The first target gene I investigated is *SLC9A3R1* (Solute Carrier Family 9 (Sodium/Hydrogen Exchanger), Isoform 3 Regulatory Factor 1). This gene was identified in a published study, where enhancer ranking of metastatic versus primary breast tumors had been performed based on H3K27ac ChIP-Seq (Patten et al., 2018). Among the ranked enhancers, that of *SLC9A3R1* was

found to be clonal in metastatic tumors compared to primary ones. *SLC9A3R1* encodes for NHERF-1, a PDZ-scaffold protein that is involved in the regulation of major cancer signaling pathways such as PDGFR/EGFR, PI3K/PTEN/AKT and Wnt/β-catenin (Vaquero et al., 2017). The PDZ2 domain of SLC9A3R1 was found to be related with metastasis and invadopodia-dependent invasion (Cardone et al., 2012). The expression of *SLC9A3R1* was significantly upregulated in MCF7 LTED cells compared to their sensitive counterparts (Figure 42a). Knowing the exact location of the enhancer for *SLC9A3R1* -which is located in the first intron- enabled us to tackle with the epigenetic regulation of this gene. For this purpose, I designed sgRNAs targeting promoter and enhancer regions (Figure 41). I used PRDM9 (H3K4 methyltransferase) or DOT1L (H3K79 methyltransferase) fused with dCas9 in combination with two sgRNAs (sgRNA 2# and 3#) targeting upstream of the CGI to increase gene expression since H3K4me3 and H3K79me are activating histone markers. Both constructs indeed resulted in increased gene expression in MCF7 cells compared to dCas9 (with no effector domain) transfected cells. The increase in gene expression was found to be significantly upregulated after 12 days of transfection with both constructs (Figure 42b,c).

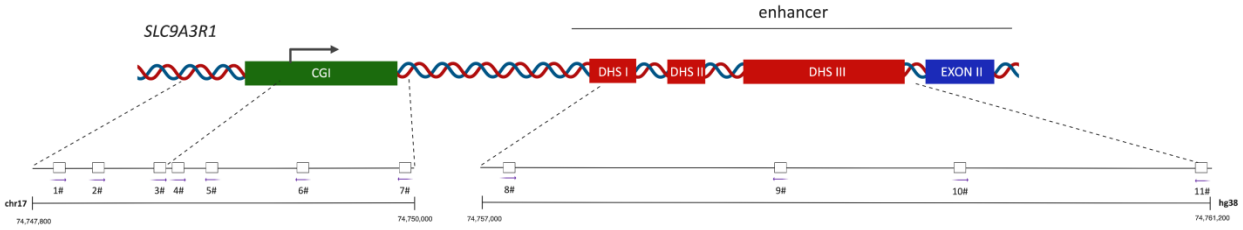


Figure 41: Designed sgRNAs for *SLC9A3R1*. Schematic depiction of location of sgRNAs designed to target *SLC9A3R1* promoter and enhancer regions (Patten et al., 2018) with CRISPR/dCas9-mediated epigenetic editing with their genomic locations. Arrows point the PAM sequence and indicate the directionality of sgRNA binding. Arrow on the genome depicts TSS: transcription start site, CGI:CpG island, DHS:DNase I-hypersensitive site.

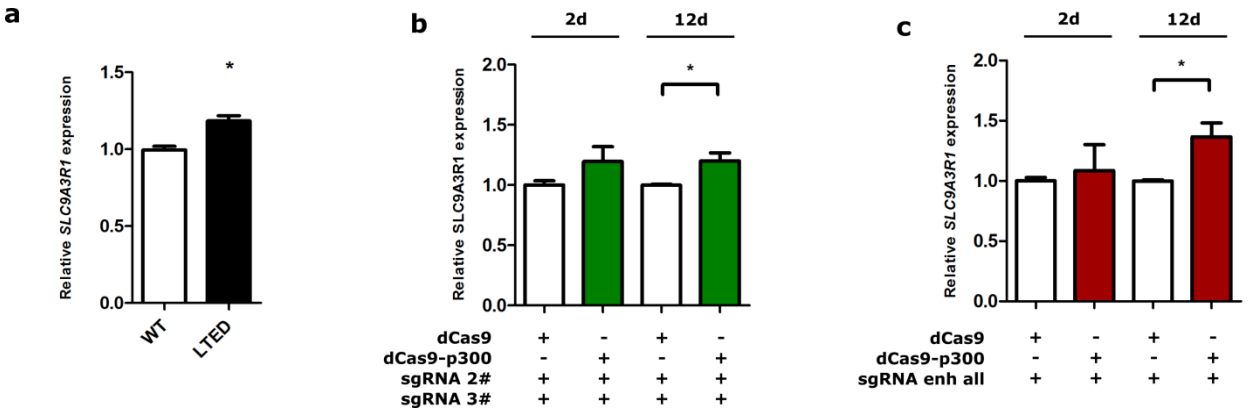


Figure 42: Effect of dCas9-p300 on *SLC9A3R1* expression. (a) Levels of *SLC9A3R1* expression were determined by qRT-PCR. (b) MCF7 cells were transfected with either dCas9 (with no effector domain) or dCas9-p300 in combination with two sgRNAs targeting promoter region of *SLC9A3R1*. (c) MCF7 cells were transfected with either dCas9 (with no effector domain) or dCas9-p300 in combination with four sgRNAs targeting enhancer region of *SLC9A3R1*. *SLC9A3R1* mRNA levels were determined by qRT-PCR 2 days and 12 days after transfection. All values of mRNA expression are normalized to dCas9+sgRNA. mRNA expression values were first normalized to *PUM1* and *ACTB* levels. Data are represented as mean ± SD, n=3 (each with 3 technical replicates). * represents $p<0.05$.

Furthermore, I wanted to see the effect of targeting both promoter and enhancer region with dCas9 fused to p300 (H3K27 acetyltransferase) since H3K27ac is known to be an activating histone mark for promoters and enhancers (Hilton et al., 2015a). This approach resulted in the significant upregulation of *SLC9A3R1* when promoter and enhancer region were targeted simultaneously compared to targeting either alone (Figure 43).

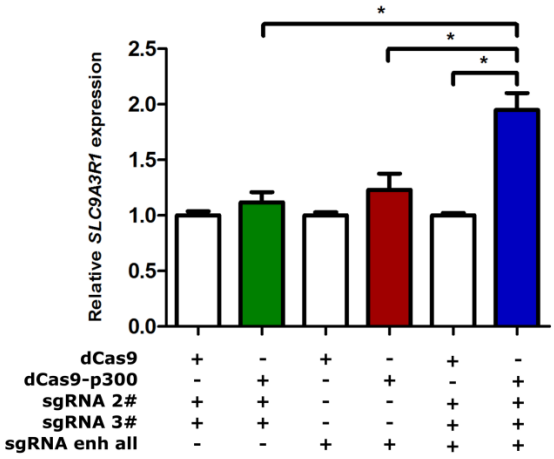


Figure 43: Effect of combinatorial targeting of promoter and enhancer with dCas9-p300 on *SLC9A3R1* expression. MCF7 cells were transfected with either dCas9 (with no effector domain) or dCas9-p300 in combination with either two sgRNAs targeting promoter region of *SLC9A3R1*, four sgRNAs targeting enhancer region of *SLC9A3R1* or both. *SLC9A3R1* mRNA levels were determined by qRT-PCR 48h after transfection. All values of mRNA expression are normalized to their respective dCas9+sgRNA combination. mRNA expression values were first normalized to *PUM1* and *ACTB* levels. Data are represented as mean ± SD, n=3 (each with 3 technical replicates). * represents $p<0.05$.

I also wanted to check whether introducing methylation would alter the expression of this gene. For this purpose, I used dCas9-M.SssI(Q147L) construct. Coupling dCas9-M.SssI(Q147L) with sgRNA 2# and 3# yielded a significant downregulation in MCF7 WT cells which persisted even after 15 days post-transfection (Figure 44).

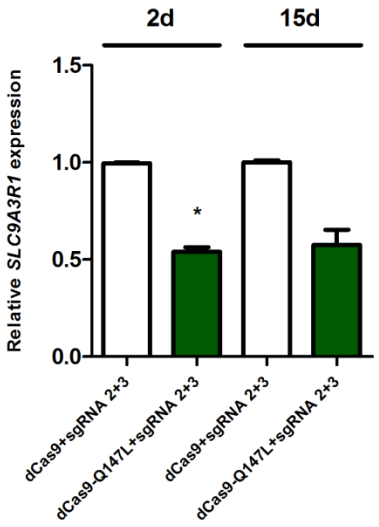


Figure 44: Effect of methylation on *SLC9A3R1* expression in MCF7 cells. MCF7 cells were transfected with either dCas9 (with no effector domain) or dCas9-M.SssI(Q147L) in combination with two sgRNAs targeting promoter region. *SLC9A3R1* mRNA levels were determined by qRT-PCR 48h and 15 days after transfection. All values of mRNA expression were normalized to dCas9+sgRNA. mRNA expression values were first normalized to *PUM1* and *ACTB* levels. Data are represented as mean ± SD, n=2 (each with 3 technical replicates). * represents $p<0.05$.

Taken altogether, these results indicate that it is possible to modulate *SLC9A3R1* expression by CRISPR/dCas9-mediated epigenetic editing in a hit-and-run approach.

4.2.1.2 *CD44*

I selected *CD44* as a second gene to investigate in the context of endocrine therapy resistance. *CD44* is known to be associated with invasiveness and stemness due to the fact that it is a target gene of Wnt signaling pathway(Orian-Rousseau, 2015)**.** Recently it has been shown in Dr. Magnani's lab that long term estrogen deprivation leads to an increase in the $CD44^{hi}$ population in MCF7 cells (Hong et al., 2018) (Figure 45a,b). This finding led to the hypothesis that CD44 might be related to resistance acquisition especially in LTED cells. Given this knowledge, I was interested to see whether it is possible to fine-tune *CD44* expression in MCF7 and MCF7 LTED cells using different epigenetic effectors fused to dCas9 in combination with sgRNAs targeting the promoter region of the *CD44* gene (kindly provided by Marianne Rots´ lab) (Figure 45c).

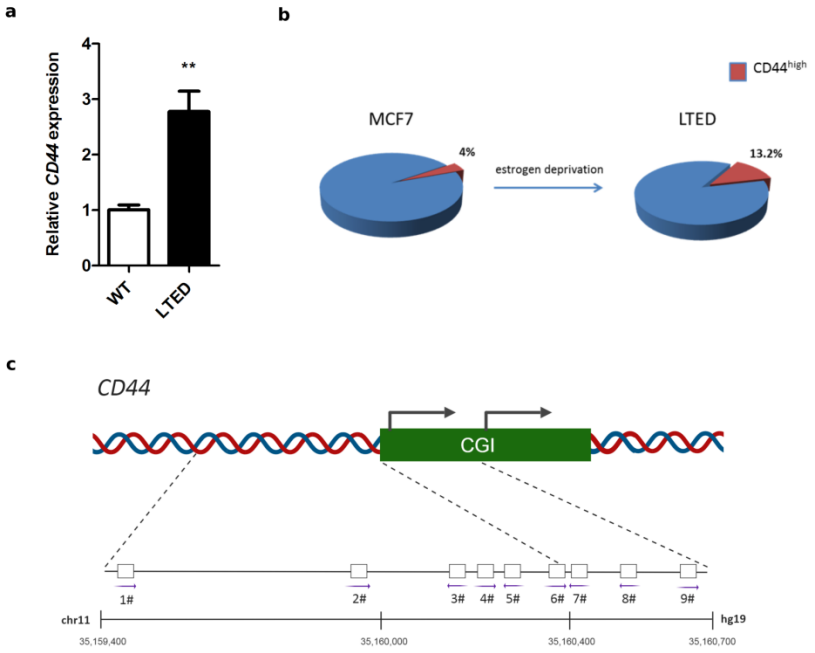


Figure 45: Increased CD44 levels in MCF7 LTED cells and designed sgRNAs for the promoter region of *CD44*. CD44 mRNA and protein levels in LTED cells were determined by qRT-PCR (a) and flow cytometry (b), repectively. (c) Schematic depiction of location of sgRNAs designed to target *CD44* promoter region with CRISPR/dCas9-mediated epigenetic editing with their genomic locations. Arrows point the PAM sequence and indicate the directionality of sgRNA binding. Arrows on the genome depict TSS. CGI:CpG island. All values of mRNA expression are normalized to parental MCF7. mRNA expression values were first normalized to *PUM1* and *ACTB* levels. Data are represented as mean ± SD, n=3 (with 3 technical replicates). ** represents $p<0.01$.

In order to achieve this, first I decided to utilize dCas9-p300 aiming to increase gene expression. Targeting *CD44* promoter region with four individual sgRNAs yielded a significant upregulation in *CD44* expression 48h post-transfection (Figure 46a). Furthermore, one of the sgRNAs (sgRNA 3#) was proven to be also effective in significantly increasing gene expression when used in combination with dCas9-DOT1L (Figure 46b).

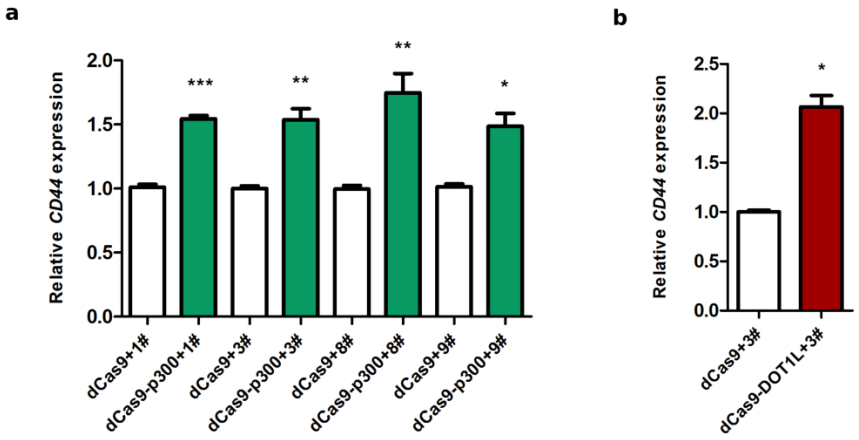


Figure 46: Effect of dCas9-p300 and dCas9-DOT1L on *CD44* expression in MCF7 cells. MCF7 cells were transfected with either dCas9 (with no effector domain) or dCas9-p300 in combination with four individual sgRNAs targeting promoter region. (b) MCF7 cells were transfected with either dCas9 (with no effector domain) or dCas9-DOT1L in combination with a sgRNA targeting promoter region. *CD44* mRNA levels were determined by qRT-PCR 48h after transfection. All values of mRNA expression are normalized to dCas9+sgRNA. mRNA expression values were first normalized to *PUM1* and *ACTB* levels. Data are represented as mean ± SD, n=2 (with 3 technical replicates). *** represents $p<0.001$, ** represents $p<0.01$, * represents $p<0.05$.

For downregulation of *CD44*, I chose histone methyltransferase G9a (fused to dCas9) which is a writer of methylation on H3K9 and H3K27(Mozzetta et al., 2014) both of which are repressive histone marks. Targeting two regions in the promoter resulted in downregulation in *CD44* expression in MCF7 parental cells even after 15 days post-transfection. This decrease in gene expression corroborated with CD44 protein levels as shown in surface staining results of CD44 in dCas9-G9a transfected cells compared to dCas9 (with no effector domain) transfected cells in combination with two individual sgRNAs (Figure 47).

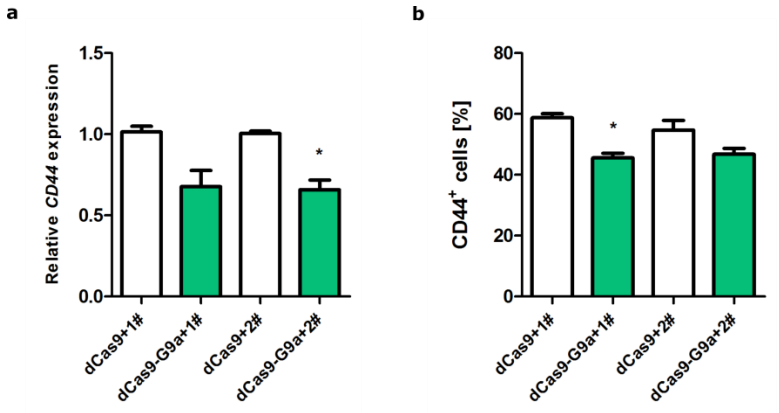


Figure 47: Effect of dCas9-G9a on *CD44* expression in MCF7 cells. MCF7 cells were transfected with either dCas9 (with no effector domain) or dCas9-G9a in combination with two individual sgRNAs targeting promoter region. *CD44* mRNA levels were determined by qRT-PCR (a) and CD44 protein levels were determined by flow cytometry 15 days after transfection. All values of mRNA expression are normalized to dCas9+sgRNA. mRNA expression values were first normalized to *PUM1* and *ACTB* levels. Data are represented as mean ± SD, n=2 (with 3 technical replicates) * represents $p<0.05$.

Alternatively, utilizing dCas9-G9a approach for downregulation of *CD44* was adopted also in MCF7 LTED cells in an attempt to see whether downregulation of *CD44* would re-sensitize these cells to endocrine therapy conditions (estrogen deprivation in particular). Here four individual sgRNAs yielded a significant decrease in *CD44* expression 48h after transfection with the dCas9-G9a construct (Figure 48a). Two of these sgRNAs (sgRNA 4# and sgRNA 8#) were effective in downregulating *CD44* after 7 days as shown in MCF7 LTED cells with inducible stable dCas9-G9a expression (Figure 49b). Furthermore, these two sgRNAs resulted in a significant decrease in proliferation of LTED cells, rendering them more susceptible to estrogen–deprivation conditions compared to controls transfected with a non-targeting control sgRNA (Figure 49c,d).

These results showcase the ability of an engineered CRISPR/dCas9 system to stably fine-tune the expression of *CD44* by editing the epigenetic landscape of the promoter in both sensitive and resistant cell lines. These findings also hint at the possibility of its utilization to alter the gene expression by modifying the epigenetic layer to achieve re-sensitization to endocrine therapy in resistant cell line.

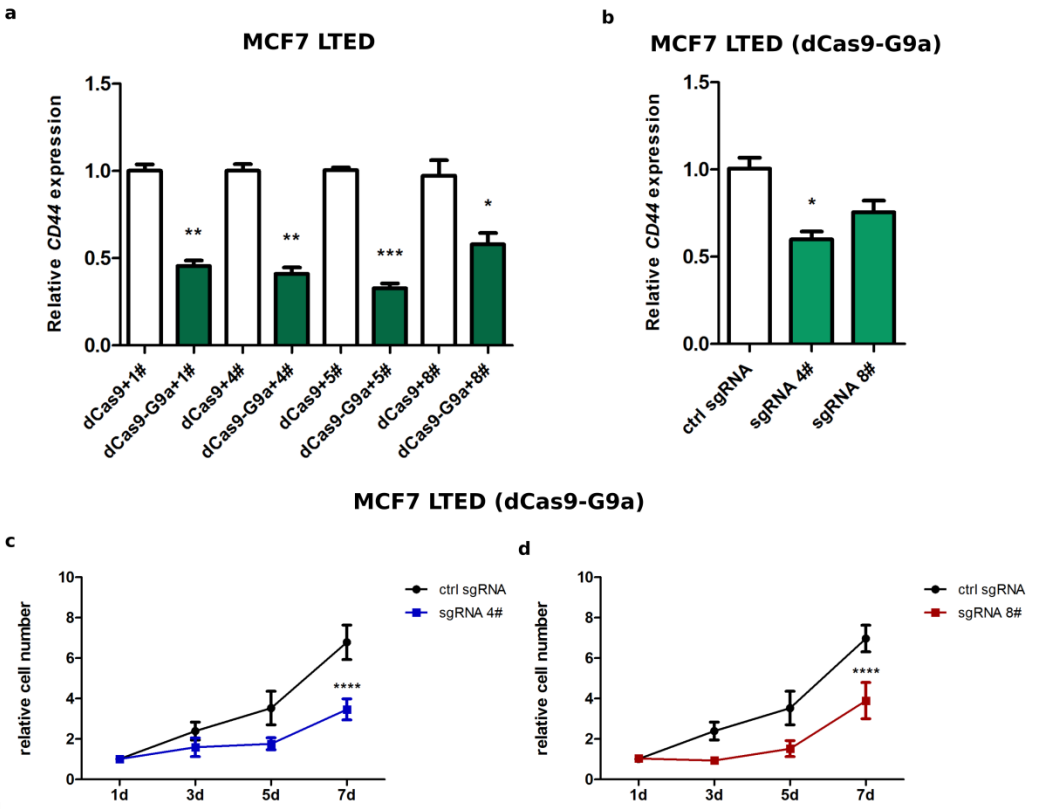


Figure 48: Effect of dCas9-G9a on *CD44* expression in MCF7 LTED cells. (b) MCF7 LTED cells stably expressing dCas9-G9a were induced with doxycycline 24h prior to transfection with two individual sgRNAs targeting promoter region of *CD44* or ctrl sgRNA. *CD44* mRNA levels were determined by qRT-PCR after 48h (a) and 7 days (b) of transfection. Cell numbers (MCF7 LTED cells with inducible stable expression of dCas9-G9a) were determined via microscopy-based nuclei count over 7 days of transfection with either sgRNA 4# (c) or sgRNA 8# (d). All mRNA expression values are normalized to either dCas9+sgRNA (a) or ctrl sgRNA (b). mRNA expression levels were first normalized to *PUM1* and *ACTB* levels. All values for proliferation assays were normalized to seeding controls. Data are represented as mean ± SD, n=2 (each with 3 technical replicates), n=2 for proliferation experiments (each with 4 technical replicates).

4.2.1.3 *BAMBI*

CD44 is considered to be a target of the Wnt signaling pathway. Several publications indicate the involvement of Wnt signaling in endocrine therapy resistance and its contribution to a more aggressive phenotype. Hence, I wondered whether altering the activation status of the Wnt signaling pathway would have an impact in the context of endocrine therapy resistance. I re-assessed the RNA-sequencing data of the wildtype and resistant MCF7 cell lines, searching for differentially regulated genes more upstream in the Wnt signaling cascade.

Among genes involved in Wnt signaling, mRNA expression of BMP and activin bound inhibitor (*BAMBI*) was significantly increased in both MCF7 TAMR and LTED compared to their parental MCF7. This increase was verified via qRT-PCR (Figure 49a). *BAMBI* was also found to be upregulated in T47D resistant cell lines in both treatment conditions (Figure 49b). *BAMBI* is involved in the canonical Wnt pathway, where it can relay Wnt signaling at the beginning of the cascade by binding and enhancing the interaction between the Wnt receptor Frizzled and its co-receptor LRP5/6. Next, I wanted to check the correlation of *BAMBI* expression with overall patient survival in two independent public datasets. The TCGA breast cancer dataset showed a trend for high *BAMBI* expression being correlated with poor patient survival in all subtypes (Figure 49c). When I looked at the overall survival of patients belonging to luminal A subtype specifically, the trend got more pronounced indicating a significant negative correlation between *BAMBI* expression and survival (Figure 49d). The METABRIC dataset showed a similar trend significantly for all subtypes (Figure 49e).

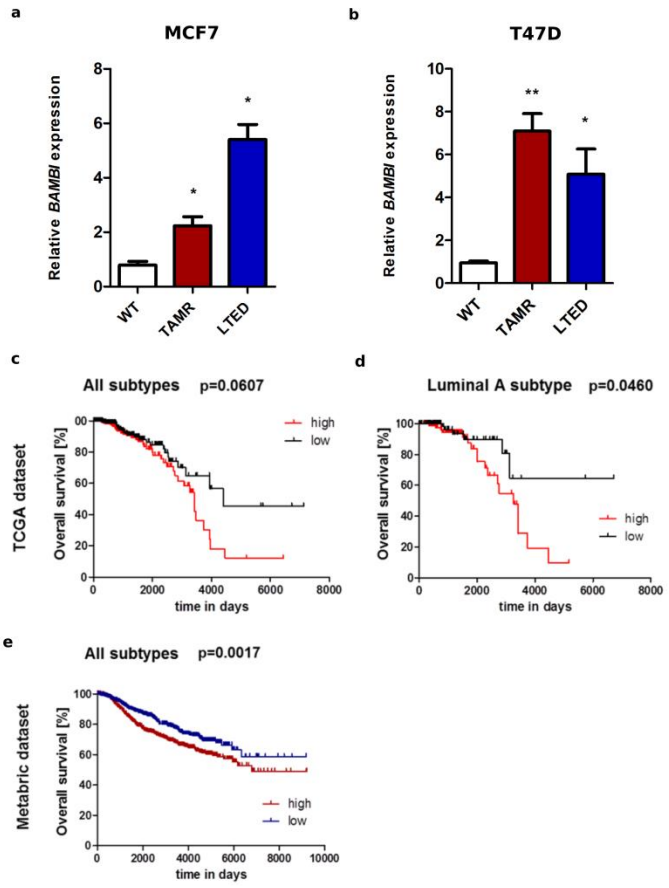


Figure 49: *BAMBI* expression in MCF7 and T47D cell line repertoire and correlation with overall patient survival. BAMBI mRNA levels were determined by qRT-PCR in MCF7 (a) and T47D (b) cells. Quartile-based overall survival analysis of TCGA dataset comparing high versus low *BAMBI* expression for all subtypes (c) and for luminal A subtype (d). (e) Quartile-based overall survival analysis of METABRIC dataset of low versus high *BAMBI* expression (for each quartile n=492). mRNA expression values are normalized to levels of respective WT. mRNA expression levels were first normalized to *PUM1* and *ACTB* levels, n=3 (each with 3 technical replicates). ** represents $p<0.01$, * represents $p<0.05$.

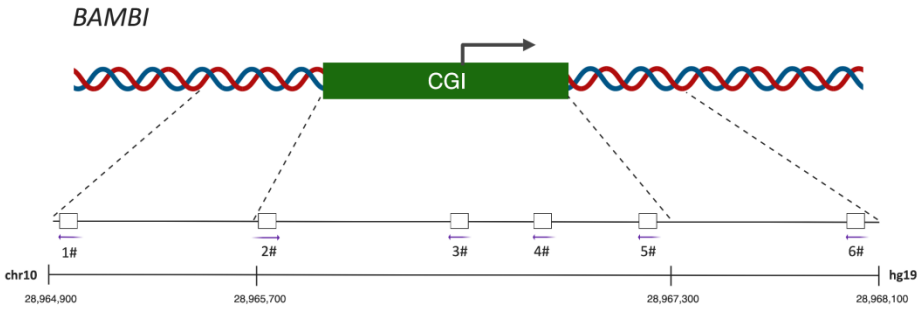


Figure 50: Designed sgRNAs for *BAMBI*. Schematic depiction of location of sgRNAs designed to target *BAMBI* promoter region with CRISPR/dCas9-mediated epigenetic editing with their genomic locations. Arrows point the PAM sequence and indicate the directionality of sgRNA binding. Arrow on the genome depicts TSS. CGI:CpG island.

Given this information, I regarded *BAMBI* to be an interesting target to follow up and designed and tested six sgRNAs targeting the promoter region of *BAMBI* (including the CpG island at the promoter) with the help of Vinona Wagner, a master student (Figure 50). For upregulation of *BAMBI*, dCas9-p300 was chosen. Two individual sgRNAs (sgRNA 1# and sgRNA 6#) were effective in significant upregulation of *BAMBI* 48h post-transfection (Figure 51a). Interestingly, *CD44* expression was found to be increased (approximately 2.5 fold with sgRNA 1#) as well when *BAMBI* was manipulated by CRISPR/dCas9-mediated epigenetic editing (Figure 51b). MCF7 cells stably expressing inducible dCas9 fused to the catalytic domain of p300 showed that these two individual sgRNAs are indeed effective in upregulating *BAMBI* after 10 days of transfection. Here, the effect of sgRNA 1# was found to be significant (Figure 51c).

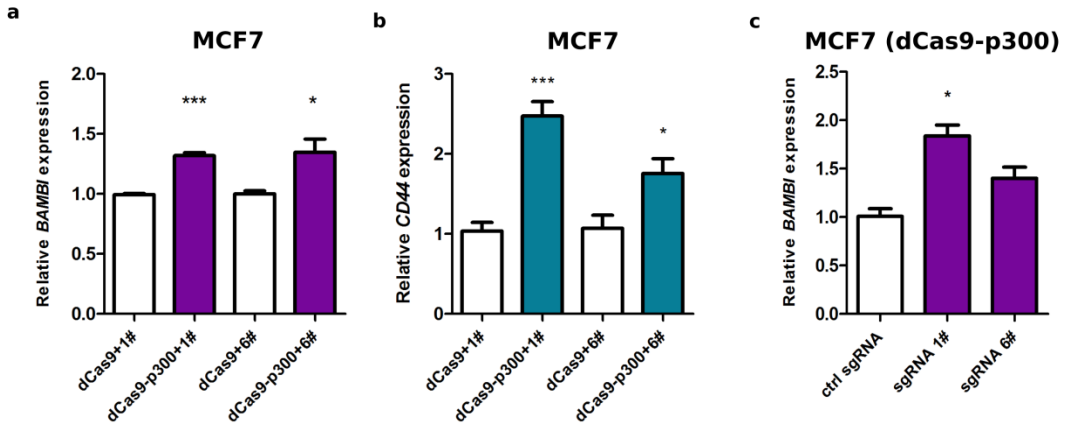


Figure 51: Effect of dCas9-p300 on *BAMBI* expression in MCF7 cells. MCF7 cells were transfected with either dCas9 (with no effector domain) or dCas9-p300 in combination with two individual sgRNAs targeting promoter region of *BAMBI*. Doxycycline- induced MCF7 cells stably expressing dCas9-p300 were transfected with two individual sgRNAs. *BAMBI* (a,c) and *CD44* (b). mRNA levels were determined by qRT-PCR 48h (a,b) and 10 days (c) after transfection. All values of mRNA expression are normalized to either dCas9+sgRNA (a,b) or ctrl sgRNA (c). mRNA expression values were first normalized to *PUM1* and *ACTB* levels. Data are represented as mean ± SD, n=2 (with 3 technical replicates). *** represents $p<0.001$, * represents $p<0.05$.

Alternatively, I also wanted to assess the involvement of methylation on *BAMBI* expression due to the existence of a CpG island at its promoter. For this purpose, I treated parental MCF7 cells with DNMT1 inhibitor 5-Aza-dC and checked *BAMBI* mRNA levels after 72h. *BAMBI* expression was indeed found upregulated upon global DNA demethylation (Figure 52a). Given this information, MCF7 cells stably expressing doxycycline-inducible dCas9-TET1 were transfected with sgRNAs targeting *BAMBI* promoter. *BAMBI* expression was found to be increased with all sgRNAs tested after 10 days (Figure 52b). Remarkably, *CD44* expression was also increased with all the sgRNAs, verifying once more the close correlation between Wnt signaling and *CD44* expression (Figure 52c). Here, the upregulating effect of sgRNA 1# and sgRNA 5# were found to be significant for both genes.

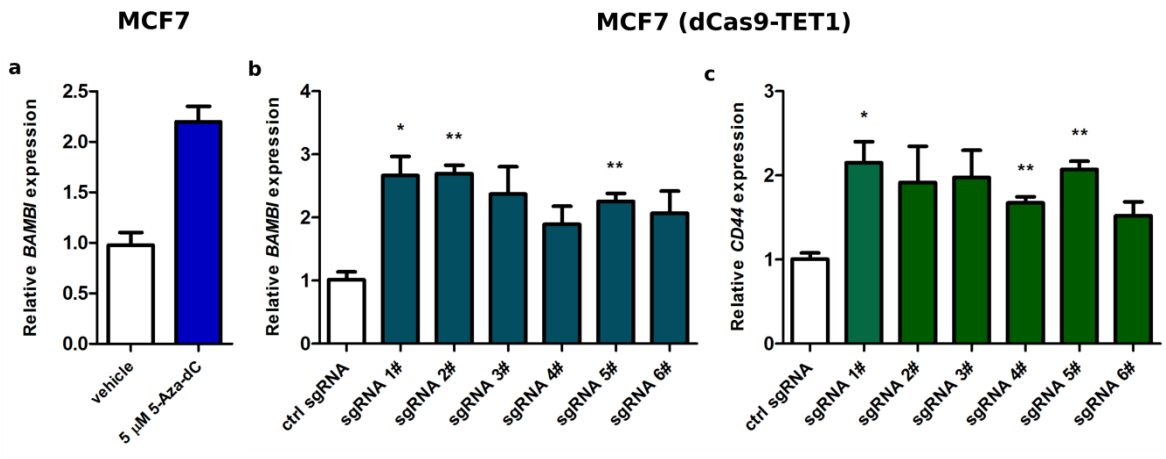


Figure 52: Effect of demethylation on *BAMBI* expression in MCF7 cells. (a) MCF7 cells were treated with either 5 µM 5-Aza-dC or vehicle control (water) for 72h. Total RNA was extracted and *BAMBI* expresssion levels were determined via qRT-PCR. Doxycycline- induced MCF7 cells stably expressing dCas9-TET1 were transfected with six individual sgRNAs targeting the promoter region of *BAMBI*. *BAMBI* (b) and *CD44* (c) mRNA levels were determined by qRT-PCR 48h after transfection. All values of mRNA expression are normalized to either vehicle control (a) or ctrl sgRNA (b,c). mRNA expression values were first normalized to *PUM1* and *ACTB* levels. Data are represented as mean ± SD, n=3 (with 3 technical replicates). ** represents $p<0.01$, * represents $p<0.05$.

For downregulation of *BAMBI*, dCas9-G9a was used since it was found to be effective in downregulating *CD44*. Indeed, *BAMBI* expression in MCF7 LTED cells was found to be significantly decreased after 48h of transfection when the promoter region was targeted with two individual sgRNAs (sgRNA 4# and sgRNA 6#) (Figure 53a). *BAMBI* expression remained significantly downregulated even after 11 days of transfection (Figure 53b). *CD44* levels were concordantly and significantly decreased as well (Figure 53c).

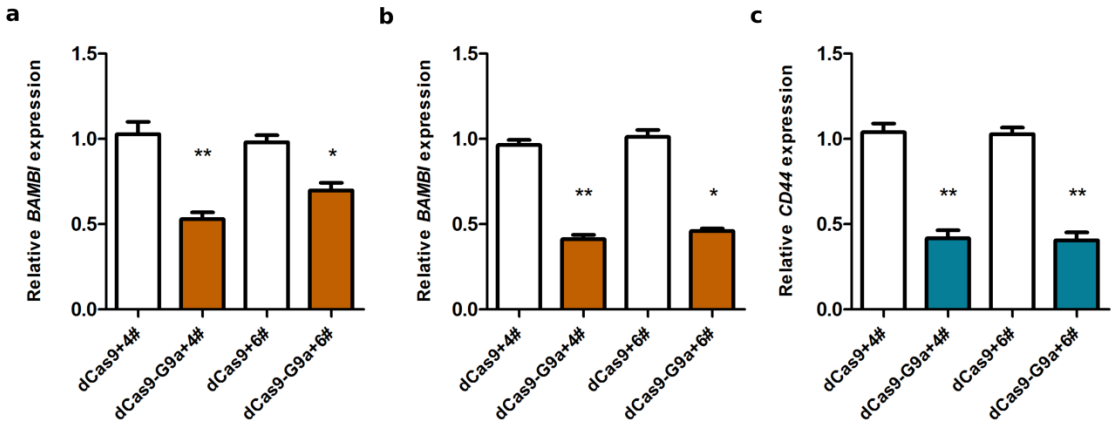


Figure 53: Effect of dCas9-G9a on *BAMBI* expression in MCF7 LTED cells. MCF7 LTED cells transfected with either dCas9 (with no effector domain) or dCas9-p300 in combination with two individual sgRNAs targeting promoter region of *BAMBI*. (a) *BAMBI* mRNA levels were determined by qRT-PCR 48h after transfection. *BAMBI* (b) and *CD44* (c) expression levels were determined by qRT-PCR after 11 days of transfection. All values of mRNA expression are normalized to dCas9+sgRNA. mRNA expression values were first normalized to *PUM1* and *ACTB* levels. Data are represented as mean ± SD, n=2 (with 3 technical replicates). ** represents $p<0.01$, * represents $p<0.05$.

Utilizing MCF7 LTED cells stably expressing inducible dCas9 fused to the catalytic domain of either wild-type G9a or catalytically inactive mutant G9a, I could verify the significant downregulation of *BAMBI*, and concordantly *CD44*, with the same sgRNAs after 7 days of transfection. Here, *BAMBI* and *CD44* expression were found to be significantly decreased in MCF7 LTED cells expressing dCas9-G9a^{WT}, but not in MCF7 LTED cells expressing dCas9-G9a^{mut}, hinting at the fact that the effect of downregulation comes indeed from the methyltransferase activity of G9a (Figure 54a-d). Next step was to assess the effect of *BAMBI* downregulation on the proliferation of LTED cells. LTED cells stably expressing dCas9-G9a^{WT} proliferated significantly less when transfected with sgRNA 4# and sgRNA 6# compared to their non-targeting sgRNA transfected counterparts suggesting a re-sensitization to estrogen deprivation conditions (Figure 53 e,f).

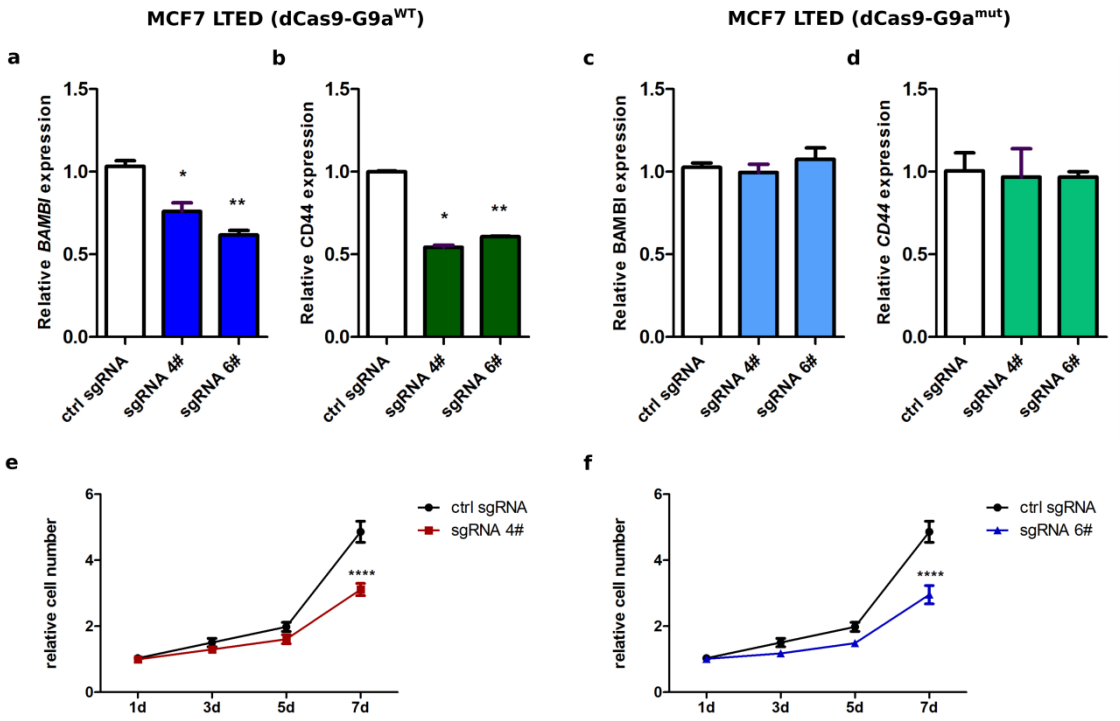


Figure 54: Effect of epigenetic editing on *BAMBI* expression in MCF7 LTED cells with inducible stable expression of dCas9 fused to wild-type or mutant G9a. Cells were induced with doxycycline 24h prior to transfection with two individual sgRNAs targeting promoter region of *BAMBI* or ctrl sgRNA. *BAMBI* (a,c) and *CD44* (b,d) mRNA levels were determined by qRT-PCR after 7 days of transfection in cells expressing either wild-type of mutant G9a fused to dCas9, respectively. Cell numbers (MCF7 LTED cells with inducible stable expression of dCas9-G9a) were determined via microscopy-based nuclei count over 7 days of transfection with either sgRNA 4# (e) or sgRNA 6# (f). All values are normalized to ctrl sgRNA. mRNA expression levels were first normalized to *PUM1* and *ACTB* levels. All values for proliferation assays were normalized to seeding controls. Data are represented as mean ± SD, n=2 (each with 3 technical replicates), n=2 for proliferation experiments (each with 4 technical replicates).

Taken altogether, these results showcase the ability, flexibility and stability of hit-and-run CRISPR/dCas9-mediated epigenetic editing in fine-tuning expression of target genes as depicted here on four independent examples (*GLYATL1*, *SLC9A3R1*, *CD44* and *BAMBI*). Moreover, these findings provide insights regarding the components that might play a role in epigenetic regulation of these genes and additional proof that epigenetic regulation indeed constitutes an integral part of regulation of transcription.

.

5. Discussion

Breast cancer is the second leading cause of cancer-related death in women (Siegel et al., 2019) Due to its heterogeneity, it is really challenging to come up with a generalized therapy against breast cancer. Even though majority of breast cancer cases are ER-positive and therefore can benefit from endocrine therapy, intrinsic or acquired resistance results in relapse for these patients in the long term. Resistance to endocrine therapy still remains a major clinical predicament for luminal breast cancer cases.

Recently, endocrine therapy occurrence is pertained to alterations in epigenetic layer as well as genetic composition (Magnani et al., 2013; Nguyen et al., 2015). Epigenetic reprogramming plays an integral role in adaptation to environmental changes such as drug treatment. This phenomena has been proposed as a mechanism to render the sensitive cells refractory to treatment. In this study, we tried to shed some light on this elusive manner by recapitulating endocrine therapy resistance *in vitro* and profiling the resistant cells of two different origin (MCF7 and T47D) in order to uncover several layers of transcriptional regulation, namely transcriptome, methylome and chromatin accessibility.

5.1 Validation of *in vitro* models

MCF7 cell lines having been made resistant against endocrine therapies (i.e., 4OH-tamoxifen, LTED) and the wildtype MCF7 cells were profiled for changes in transcriptomes and chromatin accessibility using RNA-Seq and ATAC-Seq, respectively. This revealed genes that were affected at RNA and chromatin levels, respectively. *ACSL4* and *ELF5* were among the top upregulated and *GREB1* and *PGR* among the top downregulated genes in the cell line models of endocrine resistance, each mimicking different endocrine therapy conditions (TAMR, LTED and LTEDTAMR). These genes provided a proof of concept as they had previously been implicated in endocrine therapy resistance: *ACSL4* was shown to be responsible for downregulation of ERα in triple-negative MDA-MB-231 cells. Knock-down of *ACSL4* restored ERα expression, rendering these cells susceptible to tamoxifen treatment both *in vitro* and *in vivo* (Orlando et al., 2015). *ACSL4* was also found to induce proliferation and invasion of prostate cancer cells and

was involved in hormonal resistance since it induces proliferation in the absence of androgens which makes it a candidate for castration-resistant prostate cancer (X. Wu et al., 2015). *ELF5* was found to be upregulated in *in vitro*-generated tamoxifen resistant cell lines and methylation status of the *ELF5* promoter was closely linked to its transcription as *ELF5* promoter was less methylated in relapse breast cancer patients compared to primary tumors (Fitzgerald et al., 2016). *GREB1*, an ERα co-factor, silencing by histone methyltransferase EZH2 was suggested as an underlying reason for tamoxifen resistance (Y. Wu et al., 2018). It has been shown in literature that ER+/PR+ breast tumors respond to endocrine therapy better than ER+/PR- tumors (Cui et al., 2005). The loss of progesterone receptor is linked to HER2 overexpression and associated with high relapse risk (Sun et al., 2016). Hence, the cell line models having been generated and/or analyzed in the course of this study seem to be valid *in vitro* models for endocrine resistance.

5.2 *GLYATL1* depletion leads to partial re-sensitization to endocrine therapy conditions

GLYATL1 was one of the top upregulated genes in resistant MCF7 cell lines. This gene was also upregulated in resistance acquisition model derived from another luminal subtype model cell line T47D. This finding led to the hypothesis that *GLYATL1* might indeed be important for endocrine therapy resistance development for the luminal subtype and not be a cell line-dependent effect. In order to understand whether *GLYATL1* is necessary for resistance to endocrine therapy, RNA interference experiments were performed using MCF7 and T47D cell lines that were either parental wildtype cells or had induced therapy-resistance. Loss of *GLYATL1* expression resulted in reduced cell viability of tamoxifen-resistant (TAMR) and estrogen-deprivation resistant (LTED) cells, under tamoxifen exposure and estrogen-deprivation, respectively. MCF7 LTEDTAMR were re-sensitized to both estrogen deprivation and tamoxifen exposure as they proliferated significantly less in both conditions (Figure 15). T47D resistant cells reacted to *GLYATL1* loss in a similar fashion, indicating this gene's importance in a cell-independent manner (Figure 16). These results suggested that loss/downregulation of *GLYATL1* might indeed contribute to re-sensitization of resistant cells to endocrine therapy conditions.

According to literature, *GLYATL1* codes for an N-acyltransferase enzyme which catalyzes the addition of an acyl group to L-Glutamine (H. Zhang et al., 2007). The gene was found to be significantly upregulated in tumor versus normal tissues of colon, prostate and breast cancer among 14 different cancer entities in respective TCGA datasets. *GLYATL1* is suggested as a biomarker for prostate cancer due to its significant upregulation in two independent datasets comparing primary prostate cancer samples to healthy control (Barfeld et al., 2014). All these findings are very well in line with my findings obtained in loss-of-function experiments and helped to prioritize this gene and protein, as the data suggests *GLYATL1* to be involved in the acquisition and/or maintenance of endocrine therapy resistance of breast cancer.

5.3 GLYATL1 overexpression confers endocrine therapy resistance

In order to unravel the necessity of *GLYATL1* for initiation of therapy resistance I next utilized an overexpression model of GLYATL1 in breast cancer cell lines. Exposing wildtype MCF7 cell lines stably overexpressing GLYATL1 to treatment media mimicking different endocrine therapy conditions for a period of one week indicated that GLYATL1 is indeed able to confer resistance to endocrine therapy conditions, especially tamoxifen treatment, since overexpressing cells were able to proliferate better compared to their empty vector control under tamoxifen treatment (Figure 19). However, this effect was observed to be milder for T47D cells under tamoxifen exposure. For T47D cells, overexpression of *GLYATL1* was found to be conferring a proliferative advantage under estrogen deprivation conditions (Figure 20). This can be explained by the fact that for T47D cell line repertoire, *GLYATL1* expression is highest in T47D LTED cells compared to parental control (Figure 11c) and therefore this gene might be more fundamental for overcoming the lack of estrogen in T47D cells.

5.4 GLYATL1 overexpression counteracts induction of cell cycle arrest and apoptosis in endocrine therapy-sensitive cells

Induction of cell cycle arrest and/or apoptosis are two of the hallmark effects of endocrine therapy (Amaral et al., 2013; Moriai et al., 2009; Otto et al., 1996). Hence, I analyzed the effects

of GLYATL1 on the cell cycle and late apoptotic index within MCF7 cells overexpressing the protein. And indeed, cells under endocrine therapy had a higher cell-cycle activity as measured by the fraction of cells in S-phase while the fraction of cells in late apoptosis was reduced. Interestingly, a similar reduction of cells in late apoptosis was observed in T47D cells under the same experimental conditions, however, there was no significant change in cell cycle progression in these cells (Figures 21 & 22). These findings suggest that *GLYATL1* might confer resistance through inducing cell-cycle progression and/or inhibiting apoptosis. These results also hint at the fact that overexpression of the same gene may lead to different mechanisms in two independent biological entities.

5.5 GLYATL1 is involved in histone acetylation

Recently, *GLYATL1* has been proposed to code for a histone acetyltransferase (Vecellio et al., 2014) and nomenclatured as an epigenetic writer (Keating et al., 2018). GLYATL1 has an established status as a mitochondrial enzyme. Therefore it primarily localizes to mitochondria (H. Zhang et al., 2007). To our knowledge, there are no publications indicating a nuclear localization for GLYATL1 to this date. Before further investigation of its putative role in histone acetylation, I wanted to check the cytosolic and nuclear fractions of GLYATL1 overexpressing cells. This led to the finding of GLYATL1 localizing both in cytoplasm and nucleus in MCF7 and T47D cell lines overexpressing GLYATL1 (Figure 27). Such localization has also been suggested in the GeneCards and the compartments database (Binder et al., 2014; Database GeneCards, 2017). Ectopic overexpression of proteins may lead to their artificial localization within cells (Simpson et al., 2000). Hence, nuclear localization of GLYATL1 should also be verified in resistant cell lines expressing elevated levels endogenously. However, GLYATL1 is found to be present in the nucleus also for cells transduced with the empty vector (without GLYATL1 ORF) which suggests nuclear localization even in the absence of overexpression (Figure 27). These results thus suggested that GLYATL1 could indeed play a role also in the nucleus.

mRNA expression levels of HAT family members were elevated in both *in vitro* resistance model cell lines revealing an upregulation of *KAT2A* and *KAT2B* for all resistant MCF7 cell lines and *KAT6A* and *KAT6B* in the resistant T47D cell lines. *KAT2A* was found to be increased

moderately in TAMR and LTEDTAMR cells (Figure 24). KAT2A(GCN5) and KAT2B(PCAF) are responsible for introduction of acetylation on H3K9 and H3K14 residues (Krebs et al., 2011; Z. Nagy & Tora, 2007). These HATs take part in multi-subunit HAT complexes hATAC and SAGA in a mutually exclusive fashion (Spedale et al., 2012).

KAT6A (MOZ) is implicated in acetylation of H3K9 residue as Moz-deficient mice displayed hypoacetylation of H3K9ac in Hox gene loci (Voss et al., 2009). Moreover, both KAT6A and KAT6B have been shown to be involved in the acetylation of H3K9 and H3K14 and to act as writers and readers of these modifications (Dreveny et al., 2014; Voss et al., 2009). KAT6Á and KAT6B exhibit tumorigenesis-inducing properties since their chemical inhibition in lymphoma led to tumor growth arrest (Baell et al., 2018). KAT6B (MORF) is implicated in prostate cancer cell proliferation through the PI3K-AKT signaling axis (He et al., 2013). It has been shown to introduce H3K23ac mark in small cell lung cancer (Simó-Riudalbas et al., 2015). Recent findings also annotate KAT6A as a novel epigenetic activator of ERα promoter (Yu et al., 2017).

Taken altogether, increased *KAT2A* and *KAT2B* in MCF7 cells as well as increased *KAT6A* and *KAT6B* within T47D suggest elevated levels of H3K9ac and H3K14ac marks in the affected cells. Accordingly, levels of both acetylation marks were found to be increased in resistant MCF7 cells. Additionally, GLYATL1 overexpressing MCF7 cells had higher levels of H3K9ac and H3K14ac, suggesting a link between *GLYATL1* expression and acetylation of these marks in the histone tail (Figure 28). I further confirmed the connection between GLYATL1 and histone modifications in RNAi experiments where knockdown of *GLYATL1* resulted in a significant reduction of H3K9ac and H3K14ac in MCF7 TAMR and LTED cells where *GLYATL1* was found to be upregulated (Figure 29). This novel finding draws attention to a previously unknown involvement of GLYATL1 in the acetylation of H3K9ac and H3K14ac via *KAT2B* in the context of endocrine therapy resistance.

In addition, GLYATL1 knockdown led to a significant decrease in *KAT2B* expression in MCF7 LTED cells, while a concomitant increase in *KAT2A* expression was observed. An opposite trend was detected for GLYATL1 overexpressing cells where *KAT2B* expression was downregulated. There, *KAT2A* was found to be increased in mRNA level (Figure 30). These results suggest a correlation between *GLYATL1* and *KAT2B* while corroborating with previous findings indicating

mutual exclusivity of these two GNAT family member HATs (Krebs et al., 2011; Spedale et al., 2012).

Histone modifications are crucial determinants of gene transcription. The epigenetic make-up of histone marks, the so-called ´histone code´, influences the epigenetic landscape cross-talking with DNA methylation status, thereby altering the chromatin architecture to allow or restrict binding of factors of the transcriptional machinery to certain genomic regions (Carlberg et al., 2018). Investigation of changes in histone code is a crucial step in elucidating epigenetic reprogramming.

A recent study on mouse embryonic stem cells further confirmed the co-occurrence of active histone marks H3K9ac & H3K14ac and revealed the occupancy of bivalent promoters by these two marks (Karmodiya et al., 2012). It was also argued in the same publication that these two marks can be used to distinguish active enhancers from poised ones, an attribute associated with the H3K27ac mark. Additionally, it was shown that H3K14ac also occupies some inactive promoters, adding a controversial layer to designated feature of histone acetylation as a transcriptional activation mark (Karmodiya et al., 2012). ChIP-seq of these two histone marks would elucidate the genomic regions that they are enriched in as well as the spectrum of their target genes. Correlation of these genes with the available transcriptomic data would further validate their contribution to epigenetic reprogramming in the context of endocrine therapy resistance.

The increase in *ZMYND8* mRNA levels in resistant MCF7 and T47D cell lines (Figure 26) supported the finding of the increase in H3K14ac, since ZMYND8 is identified as epigenetic reader of dual mark H3K4me1-H3K14ac (Li et al., 2016; Savitsky et al., 2016). In another publication, *ZMYND8* is annotated as a direct target gene of hypoxia inducible factor-1 (HIF-1) and HIF-2 and to encode an epigenetic reader of histone marks H3K36me/me2, H3K14ac and H4K16ac (Y. Chen et al., 2018). Taken altogether, *ZMYND8* proves to be an interesting target linking histone acetylation and breast cancer progression.

This study only focuses on alterations in two histone marks, H3K9ac and H3K14ac in resistant MCF7 cell lines. However, the increase in pan acetylation of H3 hints at the possibility that other H3 residues such as H3K18, H3K23 and H3K27 might be acetylated more as well (Figure 25).

The increase of *KAT6A* and *KAT6B* in T47D resistant cell lines highlights the possibility of elevated H3K9ac and H3K14ac also for this resistance model. Further investigation of the "acetylome" in both resistant cell lines will provide more information regarding the importance of histone acetylation in the context of endocrine therapy resistance.

Moreover, a recent publication highlights the gene fusion between *GLYATL1* and *TAF6L* [TAF6-like RNA polymerase II, p300/CBP-associated factor (PCAF)-associated factor] in breast cancer which can also account for its contribution to histone acetylation (Yoshihara et al., 2015). This fusion could be assessed in endocrine therapy resistant cell lines compared to sensitive ones by fluorescence *in situ* hybridization (FISH) using available probes to investigate the existence or enrichment of such a fusion.

5.6 *GLYATL1* expression correlates with poor prognosis in breast cancer patients

Cell lines provide the opportunity to work in depth on different cancer entity models. However, they may not reflect what is actually happening in patients. In the case of a highly heterogeneous disease as breast cancer, it is extremely important to investigate clinical implications of *in vitro* findings. Overall survival analysis for *GLYATL1* expression on the TCGA breast cancer dataset showed that *GLYATL1* expression was negatively correlated with survival. This finding demonstrates the clinical relevance of *GLYATL1* in the context of patient survival. However, this does not necessarily give any information about its potential role in endocrine therapy resistance due to the lack of matched primary and relapsed tumor material. Investigation of the patient data from METABRIC dataset after filtering for ER+ and tamoxifen-treated patients indicated that higher GLYATL1 expression was found to be correlated with poor recurrence-free patient survival which corroborates previous findings (Á. Nagy et al., 2018) (Figures 13). An independent cohort of breast cancer patients revealed a negative correlation of *GLYATL1* expression with recurrence-free survival as GLYATL1 expression was found to be significantly increased in relapsed patients (Cortazar et al., 2018; Pawitan et al., 2005) (Figure 14). Additionally, overall survival analysis for *ZMYND8*, epigenetic reader of H3K14ac mark, shows a significant negative correlation between *ZMYND8* expression and patient survival in TCGA breast cancer dataset (Á. Nagy et al., 2018) (data not shown). Together, these previous findings

support a tumor promoting role *GLYATL1* has which is associated with therapy resistance, tumor recurrence and patient survival. My data adds functional information to this association suggesting that GLYATL1 is directly involved in the disease process.

A deeper look at the subtype specificity of *GLYATL1* expression revealed the HER2-enriched subtype to be the molecular subtype with the most abundant *GLYATL1* expression in the METABRIC dataset. The basal subtype was the lowest among other subtypes in both datasets investigated suggesting that expression of *GLYATL1* might be related to one of the three receptors (ER, PR, HER2) which the basal subtype lacks (Figure 12b,12c).

5.7 *GLYATL1* expression is regulated by HER2 and luminal transcription factors

It has been shown that endocrine therapy resistance can be a result of activation of alternative growth factor pathways (Araki & Miyoshi, 2018). *ERBB2* upregulation or amplification is observed in several endocrine therapy resistance models (Grabinski et al., 2014; Houston et al., 1999), and also in the two independent *in vitro* resistance cell line models generated and used in this study. In the light of this information, the effect of *ERBB2* on *GLYATL1* expression was assessed by RNA interference. Knockdown of *ERBB2* in MCF7 resistant cell lines resulted in a significant decrease in *GLYATL1* mRNA levels (Figure 37). This suggested that *GLYATL1* expression might be partly regulated by HER2 in resistant MCF7 cell line model.

Investigating the transcription factors that might bind to the promoter region of *GLYATL1* from UCSC Genome Browser revealed TFs encoded by *ESR1*, *GATA3*, *EP300* as potential TFs driving the transcription of this gene (Kent et al., 2001) (Figure 38). Following up on this lead, knockdown experiments on each transcription factor indicated that *GLYATL1* expression was significantly downregulated when *ESR1* and *EP300* was knocked down for both MCF7 TAMR and LTED cells whereas knockdown of *GATA3* led to significant reduction in *GLYATL1* mRNA levels only in MCF7 TAMR cells (Figure 39).

These findings hint at *GLYATL1* being an estrogen receptor-responsive gene. Regulation by ERα, HER2, EP300 and GATA3 might provide an explanation for the increased levels of *GLYATL1* expression in luminal as well as in HER2-enriched subtypes of breast cancer cell lines

and patients. Regulation by the estrogen-receptor was confirmed by the significant increase in growth of MCF7 cells overexpressing GLYATL1 when grown in estrogen-supplemented media as compared to cells that were grown in media supplemented with EtOH, the vehicle for 4-OHT (Figure 19a, 19b). ChIP-seq/PCR experiments for ERα, p300 and GATA3 occupancy on the *GLYATL1* promoter could further validate these findings. Moreover, other transcription factors such as FOXA1 and FOXA2 (as suggested by UCSC Genome Browser) (Figure 38) could also be investigated to assess their role in regulation of *GLYATL1*.

Taken altogether, these discoveries exhibit novel insights into the regulation of this relatively unstudied gene. It might be surprising that ERα can apparently regulate the expression of a tragte gene even in estrogen deprived conditions (LTED). However, it has been shown before that ligand–independent activation of ERα is one of the hallmarks of endocrine therapy resistance (Ding et al., 2003). In fact, the increase in *ERBB2* mRNA levels suggests an activation of HER2 signaling which can induce phosphorylation and binding of ERα to its estrogen-responsive elements (ERE) in the genome. In the case of tamoxifen resistance, this leads to substitution of corepressor- ERα complexes bound to tamoxifen with phosphorylated ERα-activator complexes leading to conversion of tamoxifen into an agonist from its conventional antagonist role on tumor growth (Shou et al., 2004). Moreover, RNA-seq results indicate an increase in *ESR1* expression in MCF7 LTED cells and both in T47D TAMR and LTED cells which can support cross-talk between ERα- and HER2 which could thus constitute an integral part in resistance acquisition. *ESR1* levels in MCF7 TAMR and LTEDTAMR cells, on the other hand, were found to be decreased compared to parental MCF7. This draws attention to the possibility of *GLYATL1* expression being regulated also by other mechanisms such as the epigenetic layer.

5.8 *GLYATL1* is regulated by methylation

Epigenetic regulation gained much recognition recently since it has been associated with adaptation to any changes in the milieu such as constant exposure to drugs. Changes in DNA methylation patterns and the histone code, therefore the whole chromatin architecture, favor expression of certain genes while hindering expression of some others resulting in a phenotype

differing from the original status (i.e. resistant versus sensitive). For this reason, uncovering the changes in the epigenetic layer was essential to discern the role of epigenetics in endocrine therapy resistance. Methylome profiling of MCF7 resistant cell line repertoire showed that there are CpGs significantly differentially hypomethylated in the promoter region of *GLYATL1* in resistant cells compared to parental (Figure 31). This finding led to the hypothesis that this gene might be regulated also by methylation. Treating endocrine therapy sensitive MCF7 cells with a DNA methylation inhibitor (5-Aza-dC) resulted in a significant increase in *GLYATL1* expression (Figure 32a). In fact, the change in gene expression was similar to the change between resistant and sensitive MCF7 cells, hinting at the fundamental role methylation plays in regulating this gene's expression (Figure 11).

Recognition of the epigenetic layer as an integral part in regulation of gene expression resulted in approaches to manipulate the epigenetic states in order to modulate aberrant gene expression. Administration of epigenetic drugs (HDAC inhibitors etc.) can be shown as an example. However, these drugs lack specificity since they act on the whole epigenome which constitutes a major drawback (Marchion & Münster, 2007). Therefore, the need for a more targeted and specific approach has arisen. A modified version of the revolutionary genome editing method CRISPR/Cas9 offered a solution for this predicament. In this CRISPR/dCas9-mediated epigenetic editing approach, a catalytically dead Cas9 (unable to cut) fused to the catalytic domain of desired epigenetic modifier (HAT, DNMT, HMT, HDAC etc.) with a sgRNA targeting a specific region of the genome gives opportunity to modulate the epigenetic composition of a region of choice in a very specific manner (Dominguez et al., 2016; Vojta et al., 2016).

5-Aza-dC treatment was an example for administration of epigenetic drugs. In order to circumvent any non-specific effect of global DNA hypomethylation a CRISPR/dCas9-mediated epigenetic approach has been adopted. With this method, dCas9 fused to the catalytic domains of either TET1 (DNA demethylase) or M.SssI-Q147L (reduced activity mutant of prokaryotic DNMT) was utilized in combination with sgRNAs targeting CpGs that have been identified in Illumina EPIC 850k methylation array analysis. Recently it has been shown that a 10% activity mutant of M.SssI-derived DNMT (Q147L) is effective in introducing targeted methylation in a

more specific manner compared to its eukaryotic counterparts (Lei et al., 2017) and I tus chose to use this protein in my study.

Targeting differentially methylated CpGs in parental MCF7 cells stably expressing dCas9-TET1 led to an upregulation in *GLYATL1* expression as assessed 10 days after transfection with sgRNAs, thereby validating the importance of methylation (Figure 34). However it is important to note that the upregulation achieved was less than the upregulation in resistant cell lines. This hints at the existence of putative other CpGs regulating the expression of *GLYATL1*. Indeed there is a CpG island upstream of *GLYATL1* that might be responsible for regulation of this gene along with others (such as neighbor gene and family member *GLYATL2*). Detailed scrutinization of EPIC array data would reveal whether there are any probed CpGs differentially methylated in this CpG island, providing new target regions to fine-tune *GLYATL1* expression. Hypomethylation of the promoter region in resistant cell lines might also allow binding of transcription factors such as ERα, GATA3 and p300 which I found to have implications on regulation of transcription of *GLYATL1*.

In contrast, deploying dCas9-Q147L construct led to a decreased expression of *GLYATL1* in both TAMR and LTED cells targeting the same CpGs (Figure 35 & 36). In the case of downregulation of *GLYATL1* in TAMR, sgRNAs used were shown to re-sensitize resistant cells to tamoxifen treatment as they proliferated less in the presence of tamoxifen compared to vehicle control and transfection with control sgRNA (Figure 36b, 36c). These findings are in line with the effect of re-sensitization effect of *GLYATL1* knockdown in MCF7 TAMR cells (Figure 15b).

5.9 GLYATL1 as a link between metabolism and epigenetics

GLYATL1 encodes for an enzyme that uses L-Glutamine as substrate (H. Zhang et al., 2007). Elevated levels of this enzyme in both resistant cell line models (MCF7 and T47D) suggest an increase in glutamine levels in resistant cells compared to their parental counterparts. Recently there has been growing number of publications indicating a close interplay between glutamine metabolism and epigenetic reprogramming (Z. Chen et al., 2015; Cluntun et al., 2017). Many metabolites (e.g. Acetyl-CoA, NAD/NADH, SAM/SAH, ATP) that are crucial for histone modification are products of metabolic processes and it has been shown that availability of these

metabolites influence the epigenetic composition of the cell, and therefore its chromatin architecture (Katada et al., 2012; Phang et al., 2013).

Here, I have shown that increased GLYATL1 is contributing to elevated H3K9ac and H3K14ac in MCF7 resistant cells. It can be argued that this increase can be closely related with metabolic rewiring towards glutamine metabolism in resistant cells. Furthermore, a deeper look at the RNA-Seq data reveals an increase in ATP citrate synthase (*ACLY*) mRNA levels in resistant MCF7 cells compared to sensitive counterpart (data not shown). *ACLY* encodes for an enzyme that catalyzes synthesis of acetyl-CoA. This can be seen as further evidence for elevated histone acetylation since acetyl-CoA accumulation in the cell might be sensed by HATs (Katada et al., 2012; Shi & Tu, 2015). Even though the metabolomics were not investigated in this study, it is important to note that GLYATL1 might be an ideal example linking metabolism to epigenetic reprogramming in the context of endocrine therapy resistance. Further investigation to validate the proposed alterations in metabolism and increased intracellular acetyl-CoA availability would provide invaluable insights to better understand the cross-talk between the metabolic and the epigenetic landscapes of endocrine therapy resistant breast cancer.

5.10 CRISPR/dCas9-mediated epigenetic editing is effective in fine tuning target genes

CRISPR/dCas9-mediated epigenetic editing allows alteration of gene expression while providing insights about into the epigenetic layer regulating it. This system offers the advantage and flexibility to fine-tune expression of a gene targeting the same region with different effector domains fused to dCas9 or utilizing the same effector domain to up- or down-regulate the same gene targeting different genomic regulatory regions. This flexibility offers an alternative to the use of two different approaches such as RNA interference and overexpression to investigate the function of a gene. CRISPRa (CRISPR activation) and CRISPRi (CRISPR interference) approaches can be shown as good examples in which dCas9 is fused to either an artificial transcriptional activator or a repressor, in order to achieve transcriptional activation and silencing, respectively (Sanson et al., 2018).

In this study, a hit-and-run CRISPR/dCas9-mediated epigenetic editing approach was established to investigate the involvement of target genes in endocrine therapy resistance in luminal breast cancer. This has been achieved either via transient transfection of the cells with two plasmids, one encoding dCas9 fused to catalytic domain of chosen effector domain and another other encoding sgRNA, or transient transfection of doxycycline-inducible cells stably overexpressing dCas9-effector domain with sgRNA plasmids. The latter approach was adopted in an attempt to increase the transfection efficiency which is modest for a two plasmid transient transfection system in the cell lines of interest and was mostly utilized to assess functional effects of epigenetic editing on target genes. (Figure 40).

Epigenetic editing has gained substantial interest recently. Several studies exhibit examples of this approach and its implications in the stable modulation of gene expression and epigenetic reprogramming (Cano-Rodriguez et al., 2016; Hilton et al., 2015b) Reprogramming mouse primary T-cells by dCas9-p300 induced expression of *FOXP3* is one example to such studies. In that study, dCas9-p300 construct and sgRNA were stably integrated into the genome and led to a substantial increase (approximately 500-fold) in *FOXP3* gene expression which could explain the effect to achieve stable reprogramming (Okada et al., 2017).

The changes that are presented in my study are rather modest compared to the ones exemplified in the literature. This can be explained by the transient transfection approach I adopted (Figure 40). Transfection efficiencies could well account for the moderate expression changes. The effect size might thus be masked by untransfected cells in the bulk population. This was circumvented in part using doxycycline-inducible expression of dCas9-effector domain in target cells. Inducible expression of dCas9 was chosen to negate any non-specific binding of constitutively expressed dCas9 protein to genome even in the absence of sgRNA as previously shown (Galonska et al., 2018).

The endogenous epigenetic landscape of target gene is yet another and very important limiting factor for the efficacy of any CRISPR/dCas9-mediated epigenetic editing. dCas9 is a rather large molecule and sterical hindrance introduced by the target region such as existing CpG methylation or respressive histone marks restrict its binding, therefore lowering its effectivity to introduce epigenetic marks to desired loci. This has been demonstrated in a recent study where demethylation of a target region prior to introduction of activating histone marks resulted in

improved and stable upregulation or target genes (Cano-Rodriguez et al., 2016). Combination of more than one sgRNA might also help improving targeting efficiency as indicated in a recent publication where a pool of sgRNAs in combination with dCas9-p300 resulted in a significantly higher upregulation of target gene expression compared to singular targeting in a hit-and-run approach (Hilton et al., 2015b; Vojta et al., 2016).

Remarkably, even though the effect sizes in the up- or down-regulation of target genes were modest, the epigenetic alterations were found to be stable and remained significant even long after transfection (up to 15 days) suggesting that these marks were maintained during several cell cycles (Figure 44 & 47). Impressively, the effect of introducing H3K9me/me2 or H3K27me/me2 with dCas9-G9a proved to exert improved downregulation of *BAMBI* (and Wnt target gene *CD44*) as shown after 11 days of transfection compared to the effect size 48h post-transfection (Figure 53). Utilizing inducible stable cell lines improved the effect size in terms of stability as shown in the long-term effects of epigenetic editing for almost all target genes (Figures 34, 35b, 36, 48b, 52, 54). This further proved that increased transfection efficiency leads to improved alteration of gene expression.

Off-target effects remain an infamous drawback of the CRISPR/dCas9 system and this does not change for the re-purposed CRISPR/dCas9-mediated epigenetic editing approach. Even though the Cas9 variant used here is catalytically inactive, sufficient sequence identity of sgRNAs with other than the intended target region in the genome would alow binding of dCas9 and could lead to off-target effects. This needs to be taken into consideration even though off-target binding of the Cas9 protein has been reported to have limited impact on the expression of non-targeted genes, (Y. Fu et al., 2013; Galonska et al., 2018; Pattanayak et al., 2013)In order to minimize off-target effects several sgRNAs were designed for the target genes targeting the region of interest (mainly being a gene regulatory sequence such as promoter, CpG island, enhancer). The effect of the effector domains in this study presented were achieved with at least two individual sgRNAs targeting different parts of the region of interest. Moreover, the off-target effects of a hit-and-run approach is likely to be considerably less than stable expression of a dCas9 variant and sgRNAs since the occupancy of genomic regions by dCas9 occurs only for a limited time. Without a question, further experiments such as ChIP-seq for the dCas9 protein to pinpoint

potential other binding sites in the whole genome would be the ultimate experiment to interrogate off-target binding.

It is also important to note that the baseline expression level of a target gene influences the extent its expression can be manipulated. This can be seen in the example of *SLC9A3R1*. This gene is highly expressed in MCF7 cells. The increase in MCF7 LTED cells is 1.2 fold compared to parental (Figure 41a). Epigenetic editing of this gene using the dCas9-p300 construct resulted in a moderate increase, increasing the expression just to the levels of LTED cells. Nevertheless, the increase in gene expression was found to be statistically significant and remained upregulated even after 12 days of transient transfection (Figure 42).

What was more interesting regarding *SLC9A3R1* was the availability of combinatorial targeting of both promoter and enhancer region using dCas9-p300 since the H3K27ac mark is associated with active promoters and enhancers. The enhancer region for the gene was uncovered as a result of a recent study conducted by a collaborator where clonality of enhancers in primary versus metastatic breast cancer was assessed via H3K27ac ChIP-seq (Patten et al., 2018). Here the combinatorial targeting of the promoter and enhancer led to a significantly stronger increase in gene expression compared to targeting of either regulatory element individually (Figure 43).

Further experiments on functional effect of *SLC9A3R1* manipulation in both sensitive and resistant cell lines will likely reveal its contribution to endocrine therapy resistance while providing insights regarding epigenetic regulation of this gene.

5.11 Wnt signaling pathway: *BAMBI –CD44* axis

Wnt signaling has been implicated to play a role in endocrine therapy resistance. Knockdown of a long non-coding RNA was shown to cause a reduction in Wnt pathway activation, thereby rendering tamoxifen-resistant MCF7 cells sensitive to tamoxifen exposure (H. Liu et al., 2016). Increased Wnt pathway activity in another *in vitro*-generated MCF7 tamoxifen resistance model could be reversed by inhibition of the pathway (Loh et al., 2013).

CD44, a target gene of Wnt signaling pathway, encodes for a cell surface glycoprotein that is involved in cell adhesion and migration. Higher expression of this gene has been linked to a

more aggressive and invasive phenotype (Orian-Rousseau, 2015; Senbanjo & Chellaiah, 2017). Recent findings indicate an elevation in the percentage of cell expressing high CD44 levels ($CD44^{high}$) in MCF7 cells under estrogen deprivation conditions (Hong et al., 2018). CD44 is also shown to be a marker for a subset of breast cancer stem cells ($CD44^{high}$ $CD24^{low}$) (Yan et al., 2015).

Indeed MCF7 LTED cells are shown to express more CD44 both at mRNA and protein levels (Figure 45). In light of this finding, *CD44* was chosen as a target gene for epigenetic editing experiments. The experiments presented in this study demonstrated that the expression of the CD44 gene can be strongly modified by a hit-and-run CRISPR/Cas9-based epigenetic editing approach (Figures 46, 47, 48). This change in expression was found to be stable even 7-15 days post transfection (Figures 47, 48).

Moreover, reduction in *CD44* levels as shown in inducible MCF7 LTED cells stably expressing dCas9-G9a resulted in re-sensitization of these cells to estrogen deprivation conditions manifested as retardation in proliferation (Figure 48). This finding strongly suggests the involvement of *CD44* in the maintenance of endocrine therapy resistance in MCF7 LTED cell line model.

CD44 marks the endpoint of a signaling pathway, namely Wnt signaling as it is one of the target genes that gets transcribed upon Wnt stimulation (Zeilstra et al., 2008). Therefore it was intriguing to check whether activation status of Wnt pathway is linked to elevated *CD44* levels in endocrine therapy resistance. RNA-sequencing data of *in vitro*generated cell linesoriginating from two different luminal breast cancer cell lines revealed that *BAMBI* was differentially regulated in resistant cell lines. Its expression was found to be increased in both MCF7 and T47D TAMR as well as LTED cells (Figure 49). A negative correlation between *BAMBI* expression and overall patient survival in two independent publicly available breast cancer dataset pointed tothe clinical significance of this gene in the context of breast cancer (Figure 49).

Epigenetic experiments targeting the *BAMBI* promoter was effective in modulating its expression using dCas9-p300 for upregulation and dCas9-G9a for downregulation (Figures 51, 52, 53, 54). Alteration of expression changes induced by epigenetic editing proved to be stable over time. Intriguingly, *BAMBI* expression was increased (up to 3-fold) with all individual sgRNAs

targeting its promoter in inducible MCF7 cells stably expressing dCas9-TET1, strongly suggesting methylation as a regulatory mechanism for the expression of this gene. Changes in *BAMBI* expression upon epigenetic editing was closely correlated with alteration of *CD44* expression, confirming the status of *CD44* as a target gene of Wnt pathway as *BAMBI* acts as the activator of canonical Wnt signaling pathway (Lin et al., 2008).

Downregulation of BAMBI by dCas9-G9a led to a stable decrease in the expression of both *BAMBI* and *CD44* expressions. Impressively, reduction in *CD44* levels was even stronger upon regulation of *BAMBI* compared to targeting the *CD44* promoter directly using the same dCas9 variant (Figures 48b, 54b). This hints at the fact that targeting a signaling molecule at the beginning of a signaling cascade might have a bigger impact than targeting a downstream molecule of the same signaling cascade. Moreover, a comparison between doxycycline-inducible MCF7 LTED cells stably expressing dCas9-G9a^{WT} and dCas9-G9a^{mut} revealed that downregulation of *BAMBI* and *CD44* was achieved only in dCas9-G9aWT cells, validating the effect on expression reduction is indeed mediated by the methyltransferase activity of G9a (Figure 54). Here it is important to note a limitation of this study as validation of the introduced epigenetic marks by bisulfite sequencing (for DNA methylation) and ChIP-qPCR (for histone modifications) following epigenetic editing is still missing and should be done in the future.

Proliferation assays ın MCF7 LTED cells stably expressing dCas9-G9aWT demonstrated that the reduction of *BAMBI* expression by two individual sgRNAs resulted in growth retardation, in other words the re-sensitization of LTED cells to estrogen deprivation (Figure 54). The effect observed was reminiscent of the *CD44* reduction effect in the same cell line. This finding once again suggested an involvement of Wnt signaling in the maintenance of endocrine therapy resistance, in particular the resistance to aromatase inhibition. Recently it has been demonstrated that the activation of *CD44* and Wnt signaling might confer a positive feedback loop, as CD44 overexpression was also shown to increase activation of the Wnt pathway (Schmitt et al., 2015). Either way, activation status of Wnt pathway and *CD44* remain to be important players of endocrine therapy resistance.

BAMBI had been initially classified as a pseudoreceptor and inhibitor of the Transforming Growth Factor-β (TGF-β) pathway. Owing to its homology with the TGF-β receptor as well as its lack of an intracellular domain, BAMBI could scavenger TGF-β ligands and attenuate TGF-β

signaling (Onichtchouk et al., 1999). Downregulation of *BAMBI* could potentially lead to activation of the TGF-β signaling pathway while causing a reduction in Wnt signaling at the same time. The TGF- β pathway is known for its dual and context dependent role in being either pro- or anti-tumorigenic (Buck & Knabbe, 2006). However, it has been implicated in a more metastatic, invasive and stem-like phenotype in breast cancer (Buck & Knabbe, 2006; Shipitsin et al., 2007). Therefore, it is important to check the status of TGF-β signaling when epigenetically interfering with *BAMBI*.

5.12 Conclusions and Outlook

In conclusion, this study identifies *GLYATL1* as a novel gene with functional relevance in endocrine therapy resistance in luminal breast cancer along with its involvement in histone acetylation of H3K9 and H3K14 residues. Expression of *GLYATL1* was found to be regulated by several factors including methylation, growth factor signaling pathway and luminal transcription factors.

As an outlook, interaction partners of GLYATL1 could be detected via mass spectrometry and signaling pathways that GLYATL1 is part of could be unraveled via Reverse Phase Protein Array (RPPA). Additionally, elucidating the proposed link between glutamine metabolism and GLYATL1 might directly connect epigenetics and metabolism. These information could provide further insights to better understand how or why this rather overlooked protein is involved in endocrine therapy resistance in breast cancer and reveal its vulnerabilites for future therapeutic approaches.

This study denotes the efficiency and flexibility of established hit-and-run CRISPR/dCas9-mediated epigenetic editing with four different examples as target genes; *GLYATL1*, *SLC9A3R1*, *CD44* and *BAMBI*. It was shown to be possible to fine-tune the expression of all target genes utilizing dCas9 fused to different variants of effector domains depending on the epigenetic landscape of the target gene of interest.

This study has also shown that the epigenetic landscapes and expression of the *BAMBI* and *CD44* genes are involved in the maintenance of resistance to endocrine therapy in ER-positive breast cancer. This hints at a general involvement of the Wnt signaling pathway in therapy resistance.

As an outlook to this part of my study, it would be interesting to show the activation of the Wnt pathway upon epigenetic editing of *BAMBI* either via nuclear translocation of β–catenin or utilizing luciferase activity of Wnt reporter cell lines. Phenotypical assays associated with activated Wnt signaling such as migration and invasion could be also of interest to showcase functional effect of epigenetic editing on Wnt pathway components.

References

Aapola, U., Shibuya, K., Scott, H. S., Ollila, J., Vihinen, M., Heino, M., … Peterson, P. (2000). Isolation and initial characterization of a novel zinc finger gene, DNMT3L, on 21q22.3, related to the cytosine-5-methyltransferase 3 gene family. *Genomics*. https://doi.org/10.1006/geno.2000.6168

Agger, K., Christensen, J., Cloos, P. A., & Helin, K. (2008). The emerging functions of histone demethylases. *Current Opinion in Genetics and Development*. https://doi.org/10.1016/j.gde.2007.12.003

Ali, S., Buluwela, L., & Coombes, R. C. (2011). Antiestrogens and Their Therapeutic Applications in Breast Cancer and Other Diseases. *Annual Review of Medicine*, *62*(1), 217–232. https://doi.org/10.1146/annurev-med-052209-100305

Allis, C. D., & Jenuwein, T. (2016). The molecular hallmarks of epigenetic control. *Nature Reviews Genetics*, *17*(8), 487–500. https://doi.org/10.1038/nrg.2016.59

Amaral, C., Varela, C., Borges, M., Tavares Da Silva, E., Roleira, F. M. F., Correia-Da-Silva, G., & Teixeira, N. (2013). Steroidal aromatase inhibitors inhibit growth of hormone-dependent breast cancer cells by inducing cell cycle arrest and apoptosis. *Apoptosis*. https://doi.org/10.1007/s10495-013-0879-6

Araki, K., & Miyoshi, Y. (2018). Mechanism of resistance to endocrine therapy in breast cancer: the important role of PI3K/Akt/mTOR in estrogen receptor-positive, HER2-negative breast cancer. *Breast Cancer*. https://doi.org/10.1007/s12282-017-0812-x

Aran, D., Toperoff, G., Rosenberg, M., & Hellman, A. (2011). Replication timing-related and gene body-specific methylation of active human genes. *Human Molecular Genetics*. https://doi.org/10.1093/hmg/ddq513

Arpino, G., Wiechmann, L., Osborne, C. K., & Schiff, R. (2008). Crosstalk between the estrogen receptor and the HER tyrosine kinase receptor family: Molecular mechanism and clinical implications for endocrine therapy resistance. *Endocrine Reviews*. https://doi.org/10.1210/er.2006-0045

Ataseven, B. (2012). Pertuzumab plus trastuzumab plus docetaxel for metastatic breast cancer: Commentary. *Breast Care*. https://doi.org/10.1159/000338412

Avvakumov, N., & Côté, J. (2007). The MYST family of histone acetyltransferases and their intimate links to cancer. *Oncogene*. https://doi.org/10.1038/sj.onc.1210608

Badia, E., Oliva, J., Balaguer, P., & Cavaillès, V. (2007). Tamoxifen resistance and epigenetic modifications in breast cancer cell lines. *Current Medicinal Chemistry*, *14*(28), 3035–3045. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18220739

Baell, J. B., Leaver, D. J., Hermans, S. J., Kelly, G. L., Brennan, M. S., Downer, N. L., … Thomas, T. (2018). Inhibitors of histone acetyltransferases KAT6A/B induce senescence

and arrest tumour growth. *Nature*. https://doi.org/10.1038/s41586-018-0387-5

Ball, M. P., Li, J. B., Gao, Y., Lee, J. H., Leproust, E. M., Park, I. H., … Church, G. M. (2009). Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nature Biotechnology*. https://doi.org/10.1038/nbt.1533

Bannister, A. J., & Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Research*. https://doi.org/10.1038/cr.2011.22

Barfeld, S. J., East, P., Zuber, V., & Mills, I. G. (2014). Meta-Analysis of prostate cancer gene expression data identifies a novel discriminatory signature enriched for glycosylating enzymes. *BMC Medical Genomics*. https://doi.org/10.1186/s12920-014-0074-9

Bernard, P. S., Parker, J. S., Mullins, M., Cheung, M. C. U., Leung, S., Voduc, D., … Perou, C. M. (2009). Supervised risk predictor of breast cancer based on intrinsic subtypes. *Journal of Clinical Oncology*. https://doi.org/10.1200/JCO.2008.18.1370

Bertheau, P., Lehmann-Che, J., Varna, M., Dumay, A., Poirot, B., Porcher, R., … de Thé, H. (2013). P53 in breast cancer subtypes and new insights into response to chemotherapy. *Breast*. https://doi.org/10.1016/j.breast.2013.07.005

Bhutani, N., Burns, D. M., & Blau, H. M. (2011). DNA demethylation dynamics. *Cell*. https://doi.org/10.1016/j.cell.2011.08.042

Binder, J. X., Pletscher-Frankild, S., Tsafou, K., Stolte, C., O'Donoghue, S. I., Schneider, R., & Jensen, L. J. (2014). COMPARTMENTS: Unification and visualization of protein subcellular localization evidence. *Database*. https://doi.org/10.1093/database/bau012

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*. https://doi.org/10.3322/caac.21492

Brenet, F., Moh, M., Funk, P., Feierstein, E., Viale, A. J., Socci, N. D., & Scandura, J. M. (2011). DNA methylation of the first exon is tightly linked to transcriptional silencing. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0014524

Brocken, D. J. W., Tark-Dame, M., & Dame, R. T. (2018). dCas9: A Versatile Tool for Epigenome Editing. *Current Issues in Molecular Biology*. https://doi.org/10.21775/cimb.026.015

Buck, M. B., & Knabbe, C. (2006). TGF-beta signaling in breast cancer. *Annals of the New York Academy of Sciences*. https://doi.org/10.1196/annals.1386.024

Buenrostro, J. D., Wu, B., Chang, H. Y., & Greenleaf, W. J. (2015). ATAC-seq: A method for assaying chromatin accessibility genome-wide. *Current Protocols in Molecular Biology*. https://doi.org/10.1002/0471142727.mb2129s109

Cano-Rodriguez, D., Gjaltema, R. A. F., Jilderda, L. J., Jellema, P., Dokter-Fokkens, J., Ruiters, M. H. J., & Rots, M. G. (2016). Writing of H3K4Me3 overcomes epigenetic silencing in a

sustained but context-dependent manner. *Nature Communications*, *7*, 12284. https://doi.org/10.1038/ncomms12284

Cardone, R. A., Greco, M. R., Capulli, M., Weinman, E. J., Busco, G., Bellizzi, A., … Reshkin, S. J. (2012). NHERF1 acts as a molecular switch to program metastatic behavior and organotropism via its PDZ domains. *Molecular Biology of the Cell*, *23*(11), 2028–2040. https://doi.org/10.1091/mbc.E11-11-0911

Carlberg, C., Molnár, F., Carlberg, C., & Molnár, F. (2018). The Histone Code. In *Human Epigenomics*. https://doi.org/10.1007/978-981-10-7614-5_5

Castillo, J., López-Rodas, G., & Franco, L. (2017). Histone post-translational modifications and nucleosome organisation in transcriptional regulation: Some open questions. In *Advances in Experimental Medicine and Biology*. https://doi.org/10.1007/5584_2017_58

Chen, Y., Zhang, B., Bao, L., Jin, L., Yang, M., Peng, Y., … Luo, W. (2018). ZMYND8 acetylation mediates HIF-dependent breast cancer progression and metastasis. *Journal of Clinical Investigation*. https://doi.org/10.1172/JCI95089

Chen, Z., Wang, Y., Warden, C., & Chen, S. (2015). Cross-talk between ER and HER2 regulates c-MYC-mediated glutamine metabolism in aromatase inhibitor resistant breast cancer cells. *Journal of Steroid Biochemistry and Molecular Biology*. https://doi.org/10.1016/j.jsbmb.2015.02.004

Cherblanc, F., Chapman-Rothe, N., Brown, R., & Fuchter, M. (2012). Current limitations and future opportunities for epigenetic therapies. *Future Medicinal Chemistry*, *4*(4), 425–446. https://doi.org/10.4155/fmc.12.7

Chia, S., Gradishar, W., Mauriac, L., Bines, J., Amant, F., Federico, M., … Piccart, M. (2008). Double-blind, randomized placebo controlled trial of fulvestrant compared with exemestane after prior nonsteroidal aromatase inhibitor therapy in postmenopausal women with hormone receptor-positive, advanced breast cancer: Rsults from EFECT. *Journal of Clinical Oncology*. https://doi.org/10.1200/JCO.2007.13.5822

Choi, J. H., Kwon, H. J., Yoon, B. I., Kim, J. H., Han, S. U., Joo, H. J., & Kim, D. Y. (2001). Expression profile of histone deacetylase 1 in gastric cancer tissues. *Japanese Journal of Cancer Research*. https://doi.org/10.1111/j.1349-7006.2001.tb02153.x

Chu, I. M., Hengst, L., & Slingerland, J. M. (2008). The Cdk inhibitor p27 in human cancer: Prognostic potential and relevance to anticancer therapy. *Nature Reviews Cancer*. https://doi.org/10.1038/nrc2347

Cluntun, A. A., Lukey, M. J., Cerione, R. A., & Locasale, J. W. (2017). Glutamine Metabolism in Cancer: Understanding the Heterogeneity. *Trends in Cancer*. https://doi.org/10.1016/j.trecan.2017.01.005

Collins, F. S. (2007). The Cancer Genome Atlas (TCGA). *Online*.

Cortazar, A. R., Torrano, V., Martín-Martín, N., Caro-Maldonado, A., Camacho, L., Hermanova, I., … Carracedo, A. (2018). Cancertool: A visualization and representation interface to

exploit cancer datasets. *Cancer Research*. https://doi.org/10.1158/0008-5472.CAN-18-1669

Cristofanilli, M. (2005). A multigene assay for predicting the recurrence of tamoxifen-treated, node-negative breast cancer. *Breast Diseases*. https://doi.org/10.1016/S1043-321X(05)80168-7

Cristofanilli, M., Turner, N. C., Bondarenko, I., Ro, J., Im, S. A., Masuda, N., … Slamon, D. (2016). 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, phas. *The Lancet Oncology*. https://doi.org/10.1016/S1470-2045(15)00613-0

Cui, X., Schiff, R., Arpino, G., Osborne, C. K., & Lee, A. V. (2005). Biology of progesterone receptor loss in breast cancer and its implications for endocrine therapy. *Journal of Clinical Oncology*. https://doi.org/10.1200/JCO.2005.09.004

Curtis, C., Shah, S. P., Chin, S. F., Turashvili, G., Rueda, O. M., Dunning, M. J., … Caldas, C. (2012). The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature*. https://doi.org/10.1038/nature10983

Dai, H., & Wang, Z. (2014). Histone Modification Patterns and Their Responses to Environment. *Current Environmental Health Reports*. https://doi.org/10.1007/s40572-013-0008-2

Dai, X., Li, T., Bai, Z., Yang, Y., Liu, X., Zhan, J., & Shi, B. (2015). Breast cancer intrinsic subtype classification, clinical use and future trends. *American Journal of Cancer Research*.

Database GeneCards. (2017). GeneCards - Human Genes | Gene Database | Gene Search. *Genecards.Org*.

De Groote, M. L., Verschure, P. J., & Rots, M. G. (2012). Epigenetic Editing: Targeted rewriting of epigenetic marks to modulate expression of selected target genes. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gks863

DeGraffenried, L. A., Friedrichs, W. E., Russell, D. H., Donzis, E. J., Middleton, A. K., Silva, J. M., … Hidalgo, M. (2004). Inhibition of mTOR activity restores tamoxifen response in breast cancer cells with aberrant Akt activity. *Clinical Cancer Research*. https://doi.org/10.1158/1078-0432.CCR-04-0035

Desta, Z. (2004). 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*. https://doi.org/10.1124/jpet.104.065607

Di Cerbo, V., & Schneider, R. (2013). Cancers with wrong HATs: The impact of acetylation. *Briefings in Functional Genomics*. https://doi.org/10.1093/bfgp/els065

Dillon, S. C., Zhang, X., Trievel, R. C., & Cheng, X. (2005). The SET-domain protein superfamily: Protein lysine methyltransferases. *Genome Biology*. https://doi.org/10.1186/gb-2005-6-8-227

Ding, L., Yan, J., Zhu, J., Zhong, H., Lu, Q., Wang, Z., … Ye, Q. (2003). Ligand-independent activation of estrogen receptor α by XBP-1. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkg731

Doi, M., Hirayama, J., & Sassone-Corsi, P. (2006). Circadian Regulator CLOCK Is a Histone Acetyltransferase. *Cell*. https://doi.org/10.1016/j.cell.2006.03.033

Dominguez, A. A., Lim, W. A., & Qi, L. S. (2016). Beyond editing: Repurposing CRISPR-Cas9 for precision genome regulation and interrogation. *Nature Reviews Molecular Cell Biology*. https://doi.org/10.1038/nrm.2015.2

Doudna, J. A., & Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. *Science*. https://doi.org/10.1126/science.1258096

Dreveny, I., Deeves, S. E., Fulton, J., Yue, B., Messmer, M., Bhattacharya, A., … Heery, D. M. (2014). The double PHD finger domain of MOZ/MYST3 induces α-helical structure of the histone H3 tail to facilitate acetylation and methylation sampling and modification. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkt931

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

Eckhardt, F., Lewin, J., Cortese, R., Rakyan, V. K., Attwood, J., Burger, M., … Beck, S. (2006). DNA methylation profiling of human chromosomes 6, 20 and 22. *Nature Genetics*. https://doi.org/10.1038/ng1909

Fan, J., Yin, W. J., Lu, J. S., Wang, L., Wu, J., Wu, F. Y., … Shao, Z. M. (2008). ERα negative breast cancer cells restore response to endocrine therapy by combination treatment with both HDAC inhibitor and DNMT inhibitor. *Journal of Cancer Research and Clinical Oncology*. https://doi.org/10.1007/s00432-008-0354-x

Fan, W., Chang, J., & Fu, P. (2015). Endocrine therapy resistance in breast cancer: Current status, possible mechanisms and overcoming strategies. *Future Medicinal Chemistry*. https://doi.org/10.4155/fmc.15.93

Faridi, J., Wang, L., Endemann, G., & Roth, R. A. (2003). 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*.

Feng, Q., Wang, H., Ng, H. H., Erdjument-Bromage, H., Tempst, P., Struhl, K., & Zhang, Y. (2002). Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. *Current Biology*. https://doi.org/10.1016/S0960-9822(02)00901-6

Fitzgerald, L. M., Browne, E. P., Christie, K. D., Punska, E. C., Simmons, L. O., Williams, K. E., … Arcaro, K. F. (2016). ELF5 and DOK7 regulation in anti-estrogen treated cells and tumors. *Cancer Cell International*. https://doi.org/10.1186/s12935-016-0282-9

Flores, K., Wolschin, F., Corneveaux, J. J., Allen, A. N., Huentelman, M. J., & Amdam, G. V. (2012). Genome-wide association between DNA methylation and alternative splicing in an invertebrate. *BMC Genomics*. https://doi.org/10.1186/1471-2164-13-480

Fu, X., Osborne, C. K., & Schiff, R. (2013). Biology and therapeutic potential of PI3K signaling in ER+/HER2-negative breast cancer. *Breast (Edinburgh, Scotland), 22 Suppl 2*(0 2), S12-8. https://doi.org/10.1016/j.breast.2013.08.001

Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reyon, D., Joung, J. K., & Sander, J. D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nature Biotechnology*. https://doi.org/10.1038/nbt.2623

Fulawka, L., & Halon, A. (2017). Ki-67 evaluation in breast cancer: The daily diagnostic practice. *Indian Journal of Pathology and Microbiology*. https://doi.org/10.4103/ijpm.ijpm_732_15

Galonska, C., Charlton, J., Mattei, A. L., Donaghey, J., Clement, K., Gu, H., … Meissner, A. (2018). Genome-wide tracking of dCas9-methyltransferase footprints. *Nature Communications*. https://doi.org/10.1038/s41467-017-02708-5

Gao, A., Sun, T., Ma, G., Cao, J., Hu, Q., Chen, L., … Zhu, Z. (2018). LEM4 confers tamoxifen resistance to breast cancer cells by activating cyclin D-CDK4/6-Rb and ERα pathway. *Nature Communications*. https://doi.org/10.1038/s41467-018-06309-8

Gao, X., Tsang, J. C. H., Gaba, F., Wu, D., Lu, L., & Liu, P. (2014). Comparison of TALE designer transcription factors and the CRISPR/dCas9 in regulation of gene expression by targeting enhancers. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gku836

García-Becerra, R., Santos, N., Díaz, L., & Camacho, J. (2012). Mechanisms of resistance to endocrine therapy in breast cancer: focus on signaling pathways, miRNAs and genetically based resistance. *International Journal of Molecular Sciences*, *14*(1), 108–145. https://doi.org/10.3390/ijms14010108

Ghosh, S., Yates, A. J., Frühwald, M. C., Miecznikowski, J. C., Plass, C., & Smiraglia, D. J. (2010). Tissue specific DNA methylation of CpG islands in normal human adult somatic tissues distinguishes neural from non-neural tissues. *Epigenetics*. https://doi.org/10.4161/epi.5.6.12228

Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., Torres, S. E., … Qi, L. S. (2013). XCRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell*. https://doi.org/10.1016/j.cell.2013.06.044

Gnant, M., Greil, R., Hubalek, M., & Steger, G. (2013). Everolimus in postmenopausal, hormone receptor-positive advanced breast cancer: Summary and results of an austrian expert panel discussion. *Breast Care*. https://doi.org/10.1159/000354121

Gnant, M., Harbeck, N., & Thomssen, C. (2011). St. Gallen 2011: Summary of the consensus discussion. *Breast Care*. https://doi.org/10.1159/000328054

Gong, F., Chiu, L. Y., & Miller, K. M. (2016). Acetylation Reader Proteins: Linking Acetylation

Signaling to Genome Maintenance and Cancer. *PLoS Genetics*. https://doi.org/10.1371/journal.pgen.1006272

Grabinski, N., Möllmann, K., Milde-Langosch, K., Müller, V., Schumacher, U., Brandt, B., … Jücker, M. (2014). AKT3 regulates ErbB2, ErbB3 and estrogen receptor α expression and contributes to endocrine therapy resistance of ErbB2+ breast tumor cells from Balb-neuT mice. *Cellular Signalling*. https://doi.org/10.1016/j.cellsig.2014.01.018

Greer, C. B., Tanaka, Y., Kim, Y. J., Xie, P., Zhang, M. Q., Park, I. H., & Kim, T. H. (2015). Histone Deacetylases Positively Regulate Transcription through the Elongation Machinery. *Cell Reports*. https://doi.org/10.1016/j.celrep.2015.10.013

Greer, E. L., & Shi, Y. (2012). Histone methylation: a dynamic mark in health, disease and inheritance. *Nature Reviews. Genetics*. https://doi.org/10.1038/nrg3173

Guo, J. U., Ma, D. K., Mo, H., Ball, M. P., Jang, M. H., Bonaguidi, M. A., … Song, H. (2011). Neuronal activity modifies the DNA methylation landscape in the adult brain. *Nature Neuroscience*. https://doi.org/10.1038/nn.2900

Haberland, M., Montgomery, R. L., & Olson, E. N. (2009). The many roles of histone deacetylases in development and physiology: Implications for disease and therapy. *Nature Reviews Genetics*. https://doi.org/10.1038/nrg2485

Halkidou, K., Gaughan, L., Cook, S., Leung, H. Y., Neal, D. E., & Robson, C. N. (2004). Upregulation and Nuclear Recruitment of HDACl in Hormone Refractory Prostate Cancer. *Prostate*. https://doi.org/10.1002/pros.20022

Hartl, D., Krebs, A. R., Grand, R. S., Baubec, T., Isbel, L., Wirbelauer, C., … Schübeler, D. (2019). CG dinucleotides enhance promoter activity independent of DNA methylation. *Genome Research*. https://doi.org/10.1101/gr.241653.118

Hata, K., Kusumi, M., Yokomine, T., Li, E., & Sasaki, H. (2006). Meiotic and epigenetic aberrations in Dmnt3L-deficient male germ cells. *Molecular Reproduction and Development*. https://doi.org/10.1002/mrd.20387

He, W., Zhang, M. G., Wang, X. J., Zhong, S., Shao, Y., Zhu, Y., & Shen, Z. J. (2013). KAT5 and KAT6B are in positive regulation on cell proliferation of prostate cancer through PI3K-AKT signaling. *International Journal of Clinical and Experimental Pathology*.

Hellman, A., & Chess, A. (2007). Gene body-specific methylation on the active X chromosome. *Science*. https://doi.org/10.1126/science.1136352

Hilton, I. B., D'Ippolito, A. M., Vockley, C. M., Thakore, P. I., Crawford, G. E., Reddy, T. E., & Gersbach, C. A. (2015a). Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nature Biotechnology*, *33*(5), 510–517. https://doi.org/10.1038/nbt.3199

Hilton, I. B., D'Ippolito, A. M., Vockley, C. M., Thakore, P. I., Crawford, G. E., Reddy, T. E., & Gersbach, C. A. (2015b). Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nature Biotechnology*.

https://doi.org/10.1038/nbt.3199

Hong, S. P., Chan, T. E., Lombardo, Y., Corleone, G., Rotmensz, N., Pruneri, G., … Magnani, L. (2018). Single-cell Transcriptomics reveals multi-step adaptations to endocrine therapy. *BioRxiv*. https://doi.org/10.1101/485136

Horvath, P., & Barrangou, R. (2010). CRISPR/Cas, the immune system of bacteria and archaea. *Science (New York, N.Y.)*. https://doi.org/10.1126/science.1179555

Hoskins, J. M., Carey, L. A., & McLeod, H. L. (2009). CYP2D6 and tamoxifen: DNA matters in breast cancer. *Nature Reviews Cancer*. https://doi.org/10.1038/nrc2683

Houston, S. J., Plunkett, T. A., Barnes, D. M., Smith, P., Rubens, R. D., & Miles, D. W. (1999). Overexpression of c-erbB2 is an independent marker of resistance to endocrine therapy in advanced breast cancer. *British Journal of Cancer*. https://doi.org/10.1038/sj.bjc.6690196

Howe, L., Auston, D., Grant, P., John, S., Cook, R. G., Workman, J. L., & Pillus, L. (2001). Histone H3 specific acetyltransferases are essential for cell cycle progression. *Genes and Development*. https://doi.org/10.1101/gad.931401

Hyun, K., Jeon, J., Park, K., & Kim, J. (2017). Writing, erasing and reading histone lysine methylations. *Experimental and Molecular Medicine*. https://doi.org/10.1038/emm.2017.11

Illingworth, R. S., Gruenewald-Schneider, U., Webb, S., Kerr, A. R. W., James, K. D., Turner, D. J., … Bird, A. P. (2010). Orphan CpG Islands Identify numerous conserved promoters in the mammalian genome. *PLoS Genetics*. https://doi.org/10.1371/journal.pgen.1001134

Irizarry, R. A., Ladd-Acosta, C., Wen, B., Wu, Z., Montano, C., Onyango, P., … Feinberg, A. P. (2009). The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nature Genetics*. https://doi.org/10.1038/ng.298

Jansen, M. P. H. M., Foekens, J. A., Van Staveren, I. L., Dirkzwager-Kiel, M. M., Ritstier, K., Look, M. P., … Berns, E. M. J. J. (2005). Molecular classification of tamoxifen-resistant breast carcinomas by gene expression profiling. *Journal of Clinical Oncology*. https://doi.org/10.1200/JCO.2005.05.145

Jia, D., Jurkowska, R. Z., Zhang, X., Jeltsch, A., & Cheng, X. (2007). Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature*. https://doi.org/10.1038/nature06146

Jin, Q., Yu, L. R., Wang, L., Zhang, Z., Kasper, L. H., Lee, J. E., … Ge, K. (2011). Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. *EMBO Journal*. https://doi.org/10.1038/emboj.2010.318

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. https://doi.org/10.1126/science.1225829

Kaplan, E. L., & Meier, P. (n.d.). *Nonparametric Estimation from Incomplete Observations*.

Karmodiya, K., Krebs, A. R., Oulad-Abdelghani, M., Kimura, H., & Tora, L. (2012). H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. *BMC Genomics*. https://doi.org/10.1186/1471-2164-13-424

Katada, S., Imhof, A., & Sassone-Corsi, P. (2012). Connecting threads: Epigenetics and metabolism. *Cell*. https://doi.org/10.1016/j.cell.2012.01.001

Keating, S. T., van Diepen, J. A., Riksen, N. P., & El-Osta, A. (2018). Epigenetics in diabetic nephropathy, immunity and metabolism. *Diabetologia*. https://doi.org/10.1007/s00125-017-4490-1

Kent, W. J., Sugnet, C. W., & Haussler, D. (2001). The Human Genome Browser at UCSC Using the Browser. *Genome Research*. https://doi.org/10.1101/gr.229102. Article published online before print in May 2002

Kleinstiver, B. P., Prew, M. S., Tsai, S. Q., Topkar, V. V., Nguyen, N. T., Zheng, Z., … Joung, J. K. (2015). Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature*. https://doi.org/10.1038/nature14592

Knott, G. J., & Doudna, J. A. (2018). CRISPR-Cas guides the future of genetic engineering. *Science*. https://doi.org/10.1126/science.aat5011

Koboldt, D. C., Fulton, R. S., McLellan, M. D., Schmidt, H., Kalicki-Veizer, J., McMichael, J. F., … Palchik, J. D. (2012). Comprehensive molecular portraits of human breast tumours. *Nature*. https://doi.org/10.1038/nature11412

Kohli, R. M., & Zhang, Y. (2013). TET enzymes, TDG and the dynamics of DNA demethylation. *Nature*. https://doi.org/10.1038/nature12750

Krebs, A. R., Karmodiya, K., Lindahl-Allen, M., Struhl, K., & Tora, L. (2011). SAGA and ATAC histone acetyl transferase complexes regulate distinct sets of genes and ATAC defines a class of p300-independent enhancers. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2011.08.037

Kungulovski, G., & Jeltsch, A. (2016). Epigenome Editing: State of the Art, Concepts, and Perspectives. *Trends in Genetics*, *32*(2), 101–113. https://doi.org/10.1016/j.tig.2015.12.001

Lachner, M., & Jenuwein, T. (2002). The many faces of histone lysine methylation. *Current Opinion in Cell Biology*.

Lee, K. K., & Workman, J. L. (2007). Histone acetyltransferase complexes: One size doesn't fit all. *Nature Reviews Molecular Cell Biology*. https://doi.org/10.1038/nrm2145

Lehmann, B. D., Bauer, J. A., Chen, X., Sanders, M. E., Chakravarthy, A. B., Shyr, Y., & Pietenpol, J. A. (2011). Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *Journal of Clinical Investigation*. https://doi.org/10.1172/JCI45014

Lehmann, B. D., Jovanović, B., Chen, X., Estrada, M. V., Johnson, K. N., Shyr, Y., … Pietenpol,

J. A. (2016). Refinement of Triple-Negative Breast Cancer Molecular Subtypes: Implications for Neoadjuvant Chemotherapy Selection. *PloS One*. https://doi.org/10.1371/journal.pone.0157368

Lei, Y., Zhang, X., Su, J., Jeong, M., Gundry, M. C., Huang, Y.-H., … Goodell, M. A. (2017). Targeted DNA methylation in vivo using an engineered dCas9-MQ1 fusion protein. *Nature Communications*, *8*, 16026. https://doi.org/10.1038/ncomms16026

Lennartsson, A., & Ekwall, K. (2009). Histone modification patterns and epigenetic codes. *Biochimica et Biophysica Acta - General Subjects*. https://doi.org/10.1016/j.bbagen.2008.12.006

Li, N., Li, Y., Lv, J., Zheng, X., Wen, H., Shen, H., … Lee, M. G. (2016). ZMYND8 Reads the Dual Histone Mark H3K4me1-H3K14ac to Antagonize the Expression of Metastasis-Linked Genes. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2016.06.035

Liedtke, C., & Kolberg, H.-C. (2016). Systemic Therapy of Advanced/Metastatic Breast Cancer - Current Evidence and Future Concepts. *Breast Care (Basel, Switzerland)*, *11*(4), 275–281. https://doi.org/10.1159/000447549

Lin, Z., Gao, C., Ning, Y., He, X., Wu, W., & Chen, Y. G. (2008). The pseudoreceptor BMP and activin membrane-bound inhibitor positively modulates Wnt/β-catenin signaling. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.M804039200

Liu, H., Wang, G., Yang, L., Qu, J., Yang, Z., & Zhou, X. (2016). Knockdown of long non-Coding RNA UCA1 Increases the tamoxifen sensitivity of breast cancer cells through inhibition of wnt/β-catenin pathway. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0168406

Liu, W., He, F., & Jiang, Y. (2012). Network-based discovery of gene signature for vascular invasion prediction in HCC. *Journal of Hepatology*. https://doi.org/10.1016/j.jhep.2011.11.028

Loh, Y. N., Hedditch, E. L., Baker, L. A., Jary, E., Ward, R. L., & Ford, C. E. (2013). The Wnt signalling pathway is upregulated in an in vitro model of acquired tamoxifen resistant breast cancer. *BMC Cancer*. https://doi.org/10.1186/1471-2407-13-174

Lu, J. (2015). Palbociclib: a first-in-class CDK4/CDK6 inhibitor for the treatment of hormone-receptor positive advanced breast cancer. *Journal of Hematology and Oncology*. https://doi.org/10.1186/s13045-015-0194-5

Magnani, L., Stoeck, A., Zhang, X., Lanczky, A., Mirabella, A. C., Wang, T.-L., … Lupien, M. (2013). Genome-wide reprogramming of the chromatin landscape underlies endocrine therapy resistance in breast cancer. *Proceedings of the National Academy of Sciences*, *110*(16), E1490–E1499. https://doi.org/10.1073/pnas.1219992110

Manavathi, B., Samanthapudi, V. S. K., & Gajulapalli, V. N. R. (2014). Estrogen receptor coregulators and pioneer factors: the orchestrators of mammary gland cell fate and development. *Frontiers in Cell and Developmental Biology*.

https://doi.org/10.3389/fcell.2014.00034

Marchion, D., & Münster, P. (2007). Development of histone deacetylase inhibitors for cancer treatment. *Expert Review of Anticancer Therapy*. https://doi.org/10.1586/14737140.7.4.583

Martin, C., & Zhang, Y. (2005). The diverse functions of histone lysine methylation. *Nature Reviews. Molecular Cell Biology*. https://doi.org/10.1038/nrm1761

Marx, V. (2012). Epigenetics: Reading the second genomic code. *Nature, 491*, 143. https://doi.org/10.1038/491143a

Matsuo, M., Terai, K., Kameda, N., Matsumoto, A., Kurokawa, Y., Funase, Y., … Kishimoto, T. (2012). Designation of enzyme activity of glycine-N-acyltransferase family genes and depression of glycine-N-acyltransferase in human hepatocellular carcinoma. *Biochemical and Biophysical Research Communications*. https://doi.org/10.1016/j.bbrc.2012.03.099

Maunakea, A. K., Chepelev, I., Cui, K., & Zhao, K. (2013). Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. *Cell Research*. https://doi.org/10.1038/cr.2013.110

Maunakea, A. K., Nagarajan, R. P., Bilenky, M., Ballinger, T. J., Dsouza, C., Fouse, S. D., … Costello, J. F. (2010). Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature*. https://doi.org/10.1038/nature09165

Meissner, A., Mikkelsen, T. S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., … Lander, E. S. (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature*. https://doi.org/10.1038/nature07107

Michaud, E. J., van Vugt, M. J., Bultman, S. J., Sweet, H. O., Davisson, M. T., & Woychik, R. P. (1994). Differential expression of a new dominant agouti allele (A(iapy)) is correlated with methylation state and is influenced by parental lineage. *Genes and Development*. https://doi.org/10.1101/gad.8.12.1463

Milani, A. (2014). Overcoming endocrine resistance in metastatic breast cancer: Current evidence and future directions. *World Journal of Clinical Oncology*. https://doi.org/10.5306/wjco.v5.i5.990

Mínguez, B., Hoshida, Y., Villanueva, A., Toffanin, S., Cabellos, L., Thung, S., … Llovet, J. M. (2011). Gene-expression signature of vascular invasion in hepatocellular carcinoma. *Journal of Hepatology*. https://doi.org/10.1016/j.jhep.2011.02.034

Mohn, F., Weber, M., Rebhan, M., Roloff, T. C., Richter, J., Stadler, M. B., … Schübeler, D. (2008). Lineage-Specific Polycomb Targets and De Novo DNA Methylation Define Restriction and Potential of Neuronal Progenitors. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2008.05.007

Mook, S., Van't Veer, L. J., Rutgers, E. J. T., Piccart-Gebhart, M. J., & Cardoso, F. (2007). Individualization of therapy using mammaprint®^{TM}: From development to the MINDACT trial. *Cancer Genomics and Proteomics*.

Moore, L. D., Le, T., & Fan, G. (2013). DNA methylation and its basic function. *Neuropsychopharmacology*. https://doi.org/10.1038/npp.2012.112

Moriai, R., Tsuji, N., Moriai, M., Kobayashi, D., & Watanabe, N. (2009). Survivin plays as a resistant factor against tamoxifen-induced apoptosis in human breast cancer cells. *Breast Cancer Research and Treatment*. https://doi.org/10.1007/s10549-008-0164-5

Mozzetta, C., Pontis, J., Fritsch, L., Robin, P., Portoso, M., Proux, C., … Ait-Si-Ali, S. (2014). The Histone H3 Lysine 9 Methyltransferases G9a and GLP Regulate Polycomb Repressive Complex 2-Mediated Gene Silencing. *Molecular Cell*, *53*(2), 277–289. https://doi.org/10.1016/J.MOLCEL.2013.12.005

Musgrove, E. A., & Sutherland, R. L. (2009). Biological determinants of endocrine resistance in breast cancer. *Nature Reviews Cancer*. https://doi.org/10.1038/nrc2713

Nagy, Á., Lánczky, A., Menyhárt, O., & Gyorffy, B. (2018). Validation of miRNA prognostic power in hepatocellular carcinoma using expression data of independent datasets. *Scientific Reports*. https://doi.org/10.1038/s41598-018-27521-y

Nagy, Z., & Tora, L. (2007). Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation. *Oncogene*. https://doi.org/10.1038/sj.onc.1210604

Nakagawa, M., Oda, Y., Eguchi, T., Aishima, S. I., Yao, T., Hosoi, F., … Tsuneyoshi, M. (2007). Expression profile of class I histone deacetylases in human cancer tissues. *Oncology Reports*.

Nalla, A. K., Williams, T. F., Collins, C. P., Rae, D. T., & Trobridge, G. D. (2016). Lentiviral vector-mediated insertional mutagenesis screen identifies genes that influence androgen independent prostate cancer progression and predict clinical outcome. *Molecular Carcinogenesis*, *55*(11), 1761–1771. https://doi.org/10.1002/mc.22425

Neve, R. M., Chin, K., Fridlyand, J., Yeh, J., Baehner, F. L., Fevr, T., … Gray, J. W. (2006). A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell*, *10*(6), 515–527. https://doi.org/10.1016/j.ccr.2006.10.008

Nguyen, V. T. M., Barozzi, I., Faronato, M., Lombardo, Y., Steel, J. H., Patel, N., … Magnani, L. (2015). Differential epigenetic reprogramming in response to specific endocrine therapies promotes cholesterol biosynthesis and cellular invasion. *Nature Communications*, *6*(1), 10044. https://doi.org/10.1038/ncomms10044

Nusinzon, I., & Horvath, C. M. (2005). Histone Deacetylases as Transcriptional Activators? Role Reversal in Inducible Gene Regulation. *Science Signaling*. https://doi.org/10.1126/stke.2962005re11

Okada, M., Kanamori, M., Someya, K., Nakatsukasa, H., & Yoshimura, A. (2017). Stabilization of Foxp3 expression by CRISPR-dCas9-based epigenome editing in mouse primary T cells. *Epigenetics and Chromatin*. https://doi.org/10.1186/s13072-017-0129-1

Okano, M., Bell, D. W., Haber, D. A., & Li, E. (1999). DNA methyltransferases Dnmt3a and

Dnmt3b are essential for de novo methylation and mammalian development. *Cell*. https://doi.org/10.1016/S0092-8674(00)81656-6

Onichtchouk, D., Chen, Y. G., Dosch, R., Gawantka, V., Delius, H., Massagué, J., & Niehrs, C. (1999). Silencing of TGF-β signalling by the pseudoreceptor BAMBI. *Nature*. https://doi.org/10.1038/46794

Orian-Rousseau, V. (2015). CD44 Acts as a Signaling Platform Controlling Tumor Progression and Metastasis. *Frontiers in Immunology*, *6*, 154. https://doi.org/10.3389/fimmu.2015.00154

Orlando, U. D., Castillo, A. F., Dattilo, M. A., Solano, A. R., Maloberti, P. M., & Podesta, E. J. (2015). Acyl-CoA synthetase-4, a new regulator of mTOR and a potential therapeutic target for enhanced estrogen receptor function in receptor-positive and -negative breast cancer. *Oncotarget*. https://doi.org/10.18632/oncotarget.5822

Osborne, C. K., & Schiff, R. (2011). Mechanisms of Endocrine Resistance in Breast Cancer. *Annual Review of Medicine*, *62*(1), 233–247. https://doi.org/10.1146/annurev-med-070909-182917

Otto, A. M., Paddenberg, R., Schubert, S., & Mannherz, H. G. (1996). Cell cycle arrest, micronucleus formation, and cell death in growth inhibition of MCF-7 breast cancer cells by tamoxifen and cisplatin. *Journal of Cancer Research and Clinical Oncology*. https://doi.org/10.1007/BF01221192

Pang, B., Cheng, S., Sun, S.-P., An, C., Liu, Z.-Y., Feng, X., & Liu, G.-J. (2014). Prognostic role of PIK3CA mutations and their association with hormone receptor expression in breast cancer: a meta-analysis. *Scientific Reports*, *4*, 6255. https://doi.org/10.1038/srep06255

Pattanayak, V., Lin, S., Guilinger, J. P., Ma, E., Doudna, J. A., & Liu, D. R. (2013). High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nature Biotechnology*. https://doi.org/10.1038/nbt.2673

Patten, D. K., Corleone, G., Győrffy, B., Perone, Y., Slaven, N., Barozzi, I., … Magnani, L. (2018). Enhancer mapping uncovers phenotypic heterogeneity and evolution in patients with luminal breast cancer. *Nature Medicine*. https://doi.org/10.1038/s41591-018-0091-x

Pawitan, Y., Bjöhle, J., Amler, L., Borg, A. L., Egyhazi, S., Hall, P., … Bergh, J. (2005). Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts. *Breast Cancer Research : BCR*. https://doi.org/10.1186/bcr1325

Phang, J. M., Liu, W., & Hancock, C. (2013). Bridging epigenetics and metabolism. *Epigenetics*. https://doi.org/10.4161/epi.24042

Pidsley, R., Zotenko, E., Peters, T. J., Lawrence, M. G., Risbridger, G. P., Molloy, P., … Clark, S. J. (2016). Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biology*. https://doi.org/10.1186/s13059-016-1066-1

Prat, A., Adamo, B., Cheang, M. C. U., Anders, C. K., Carey, L. A., & Perou, C. M. (2013). Molecular Characterization of Basal-Like and Non-Basal-Like Triple-Negative Breast Cancer. *The Oncologist*. https://doi.org/10.1634/theoncologist.2012-0397

Rakyan, V. K., Hildmann, T., Novik, K. L., Lewin, J., Tost, J., Cox, A. V., … Beck, S. (2004). DNA methylation profiling of the human major histocompatibility complex: A pilot study for the Human Epigenome Project. *PLoS Biology*. https://doi.org/10.1371/journal.pbio.0020405

Richman, R., Chicoine, L. G., Collini, M. P., Cook, R. G., & Allis, C. D. (1988). Micronuclei and the cytoplasm of growing Tetrahymena contain a histone acetylase activity which is highly specific for free histone H4. *Journal of Cell Biology*. https://doi.org/10.1083/jcb.106.4.1017

Riggins, R. B., Schrecengost, R. S., Guerrero, M. S., & Bouton, A. H. (2007). Pathways to tamoxifen resistance. *Cancer Letters*. https://doi.org/10.1016/j.canlet.2007.03.016

Robertson, J. F. R., & Blamey, R. W. (2003). The use of gonadotrophin-releasing hormone (GnRH) agonists in early and advanced breast cancer in pre- and perimenopausal women. *European Journal of Cancer*. https://doi.org/10.1016/S0959-8049(02)00810-9

Ropero, S., & Esteller, M. (2007). The role of histone deacetylases (HDACs) in human cancer. *Molecular Oncology*. https://doi.org/10.1016/j.molonc.2007.01.001

Rueda, O. M., Sammut, S. J., Seoane, J. A., Chin, S. F., Caswell-Jin, J. L., Callari, M., … Curtis, C. (2019). Dynamics of breast-cancer relapse reveal late-recurring ER-positive genomic subgroups. *Nature*. https://doi.org/10.1038/s41586-019-1007-8

Rugo, H. S., Rumble, R. B., Macrae, E., Barton, D. L., Connolly, H. K., Dickler, M. N., … Burstein, H. J. (2016). Endocrine therapy for hormone receptor-positive metastatic breast cancer: American society of clinical oncology guideline. *Journal of Clinical Oncology*. https://doi.org/10.1200/JCO.2016.67.1487

RUIJTER, A. J. M. de, GENNIP, A. H. van, CARON, H. N., KEMP, S., & KUILENBURG, A. B. P. van. (2003). Histone deacetylases (HDACs): characterization of the classical HDAC family. *Biochemical Journal*. https://doi.org/10.1042/bj20021321

Ruthenburg, A. J., Li, H., Patel, D. J., & David Allis, C. (2007). Multivalent engagement of chromatin modifications by linked binding modules. *Nature Reviews Molecular Cell Biology*. https://doi.org/10.1038/nrm2298

Sanson, K. R., Hanna, R. E., Hegde, M., Donovan, K. F., Strand, C., Sullender, M. E., … Doench, J. G. (2018). Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. *Nature Communications*. https://doi.org/10.1038/s41467-018-07901-8

Savitsky, P., Krojer, T., Fujisawa, T., Lambert, J. P., Picaud, S., Wang, C. Y., … Filippakopoulos, P. (2016). Multivalent Histone and DNA Engagement by a PHD/BRD/PWWP Triple Reader Cassette Recruits ZMYND8 to K14ac-Rich Chromatin. *Cell Reports*. https://doi.org/10.1016/j.celrep.2016.11.014

Saxonov, S., Berg, P., & Brutlag, D. L. (2006). A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.0510310103

Scharl, A., & Salterberg, A. (2016). Significance of Ovarian Function Suppression in Endocrine Therapy for Breast Cancer in Pre-Menopausal Women. *Geburtshilfe Und Frauenheilkunde*. https://doi.org/10.1055/s-0042-106389

Schiff, R., Massarweh, S. A., Shou, J., Bharwani, L., Mohsin, S. K., Osborne, C. K., … Brown, M. (2004). Cross-Talk between Estrogen Receptor and Growth Factor Pathways As A Molecular Target for Overcoming Endocrine Resistance. *Clinical Cancer Research*. https://doi.org/10.1158/1078-0432.CCR-031212

Schmitt, M., Metzger, M., Gradl, D., Davidson, G., & Orian-Rousseau, V. (2015). CD44 functions in Wnt signaling by regulating LRP6 localization and activation. *Cell Death and Differentiation*. https://doi.org/10.1038/cdd.2014.156

Schneider, R. E., Barakat, A., Pippen, J., & Osborne, C. (2011). Aromatase inhibitors in the treatment of breast cancer in post-menopausal female patients: An update. *Breast Cancer: Targets and Therapy*. https://doi.org/10.2147/BCTT.S22905

Senbanjo, L. T., & Chellaiah, M. A. (2017). CD44: A Multifunctional Cell Surface Adhesion Receptor Is a Regulator of Progression and Metastasis of Cancer Cells. *Frontiers in Cell and Developmental Biology*. https://doi.org/10.3389/fcell.2017.00018

Seto, E., & Yoshida, M. (2014). Erasers of histone acetylation: The histone deacetylase enzymes. *Cold Spring Harbor Perspectives in Biology*. https://doi.org/10.1101/cshperspect.a018713

Sharma, G. G., So, S., Gupta, A., Kumar, R., Cayrou, C., Avvakumov, N., … Pandita, T. K. (2010). MOF and Histone H4 Acetylation at Lysine 16 Are Critical for DNA Damage Response and Double-Strand Break Repair. *Molecular and Cellular Biology*. https://doi.org/10.1128/MCB.01476-09

Sharma, S. V., Lee, D. Y., Li, B., Quinlan, M. P., Takahashi, F., Maheswaran, S., … Settleman, J. (2010). A Chromatin-Mediated Reversible Drug-Tolerant State in Cancer Cell Subpopulations. *Cell*, *141*(1), 69–80. https://doi.org/10.1016/j.cell.2010.02.027

Shi, L., & Tu, B. P. (2015). Acetyl-CoA and the regulation of metabolism: Mechanisms and consequences. *Current Opinion in Cell Biology*. https://doi.org/10.1016/j.ceb.2015.02.003

Shipitsin, M., Campbell, L. L., Argani, P., Weremowicz, S., Bloushtain-Qimron, N., Yao, J., … Polyak, K. (2007). Molecular Definition of Breast Tumor Heterogeneity. *Cancer Cell*. https://doi.org/10.1016/j.ccr.2007.01.013

Shou, J., Massarweh, S., Osborne, C. K., Wakeling, A. E., Ali, S., Weiss, H., & Schiff, R. (2004). Mechanisms of tamoxifen resistance: increased estrogen receptor-HER2/neu cross-talk in ER/HER2-positive breast cancer. *Journal of the National Cancer Institute*.

Siegel, R. L., Miller, K. D., & Jemal, A. (2019). Cancer statistics, 2019. *CA: A Cancer Journal for Clinicians*. https://doi.org/10.3322/caac.21551

Simó-Riudalbas, L., Pérez-Salvia, M., Setien, F., Villanueva, A., Moutinho, C., Martínez-Cardús, A., … Esteller, M. (2015). KAT6B is a tumor suppressor histone H3 lysine 23 acetyltransferase undergoing genomic loss in small cell lung cancer. *Cancer Research*. https://doi.org/10.1158/0008-5472.CAN-14-3702

Simpson, J. C., Wellenreuther, R., Poustka, A., Pepperkok, R., & Wiemann, S. (2000). Systematic subcellular localization of novel proteins identified by large-scale cDNA sequencing. *EMBO Reports*. https://doi.org/10.1093/embo-reports/kvd058

Slamon, D. J., Leyland-Jones, B., Shak, S., Fuchs, H., Paton, V., Bajamonde, A., … Norton, L. (2001). Use of Chemotherapy plus a Monoclonal Antibody against HER2 for Metastatic Breast Cancer That Overexpresses HER2. *New England Journal of Medicine*, *344*(11), 783–792. https://doi.org/10.1056/NEJM200103153441101

Smith, G. L. (2014). The Long and Short of Tamoxifen Therapy: A Review of the ATLAS Trial. *Journal of the Advanced Practitioner in Oncology*.

Somech, R., Izraeli, S., & Simon, A. J. (2004). Histone deacetylase inhibitors - A new tool to treat cancer. *Cancer Treatment Reviews*. https://doi.org/10.1016/j.ctrv.2004.04.006

Song, Y., Wu, F., & Wu, J. (2016). Targeting histone methylation for cancer therapy: Enzymes, inhibitors, biological activity and perspectives. *Journal of Hematology and Oncology*. https://doi.org/10.1186/s13045-016-0279-9

Sørlie, T., Perou, C. M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., … Børresen-Dale, A. L. (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.191367098

Span, P. N., Tjan-Heijnen, V. C. G., Manders, P., Beex, L. V. A. M., & Sweep, C. G. J. (2003). Cyclin-E is a strong predictor of endocrine therapy failure in human breast cancer. *Oncogene*. https://doi.org/10.1038/sj.onc.1206818

Spedale, G., Timmers, H. T. M., & Pijnappel, W. W. M. P. (2012). ATAC-king the complexity of SAGA during evolution. *Genes and Development*. https://doi.org/10.1101/gad.184705.111

Steele, N., Finn, P., Brown, R., & Plumb, J. A. (2009). Combined inhibition of DNA methylation and histone acetylation enhances gene re-expression and drug sensitivity in vivo. *British Journal of Cancer*, *100*(5), 758–763. https://doi.org/10.1038/sj.bjc.6604932

Sternberg, S. H., & Doudna, J. A. (2015). Expanding the Biologist's Toolkit with CRISPR-Cas9. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2015.02.032

Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C., & Doudna, J. A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*. https://doi.org/10.1038/nature13011

Suetake, I., Shinozaki, F., Miyagawa, J., Takeshima, H., & Tajima, S. (2004). DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct

interaction. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.M400181200

Sun, J. Y., Wu, S. G., Li, F. Y., Lin, H. X., & He, Z. Y. (2016). Progesterone receptor loss identifies hormone receptor-positive and HER2-negative breast cancer subgroups at higher risk of relapse: A retrospective cohort study. *OncoTargets and Therapy*. https://doi.org/10.2147/OTT.S98666

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., … Von Mering, C. (2019). STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gky1131

Tan, M., Luo, H., Lee, S., Jin, F., Yang, J. S., Montellier, E., … Zhao, Y. (2011). Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell*. https://doi.org/10.1016/j.cell.2011.08.008

Tiedeken, S. D., & Chang, C. (2015). Principles of Epigenetic Treatment. In *Epigenetics and Dermatology*. https://doi.org/10.1016/B978-0-12-800957-4.00020-5

Torchia, J., Glass, C., & Rosenfeld, M. G. (1998). Co-activators and co-repressors in the integration of transcriptional responses. *Current Opinion in Cell Biology*. https://doi.org/10.1016/S0955-0674(98)80014-8

Tsukada, Y. I., Fang, J., Erdjument-Bromage, H., Warren, M. E., Borchers, C. H., Tempst, P., & Zhang, Y. (2006). Histone demethylation by a family of JmjC domain-containing proteins. *Nature*. https://doi.org/10.1038/nature04433

Turner, B. M. (2000). Histone acetylation and an epigenetic code. *BioEssays*. https://doi.org/10.1002/1521-1878(200009)22:9<836::AID-BIES9>3.0.CO;2-X

Turner, B. M. (2007). Defining an epigenetic code. *Nature Cell Biology*. https://doi.org/10.1038/ncb0107-2

Uhlen, M., Zhang, C., Lee, S., Sjöstedt, E., Fagerberg, L., Bidkhori, G., … Ponten, F. (2017). A pathology atlas of the human cancer transcriptome. *Science*. https://doi.org/10.1126/science.aan2507

Vaissière, T., Sawan, C., & Herceg, Z. (2008). Epigenetic interplay between histone modifications and DNA methylation in gene silencing. *Mutation Research - Reviews in Mutation Research*. https://doi.org/10.1016/j.mrrev.2008.02.004

Vaquero, J., Nguyen Ho-Bouldoires, T. H., Clapéron, A., & Fouassier, L. (2017). Role of the PDZ-scaffold protein NHERF1/EBP50 in cancer biology: from signaling regulation to clinical relevance. *Oncogene*, *36*(22), 3067–3079. https://doi.org/10.1038/onc.2016.462

Vastenhouw, N. L., & Schier, A. F. (2012). Bivalent histone modifications in early embryogenesis. *Current Opinion in Cell Biology*. https://doi.org/10.1016/j.ceb.2012.03.009

Vecellio, M., Spallotta, F., Nanni, S., Colussi, C., Cencioni, C., Derlet, A., … Gaetano, C. (2014). The histone acetylase activator pentadecylidenemalonate 1b rescues proliferation

and differentiation in the human cardiac mesenchymal cells of type 2 diabetic patients. *Diabetes*. https://doi.org/10.2337/db13-0731

Verma, S., Miles, D., Gianni, L., Krop, I. E., Welslau, M., Baselga, J., … EMILIA Study Group. (2012). Trastuzumab emtansine for HER2-positive advanced breast cancer. *The New England Journal of Medicine*. https://doi.org/10.1056/NEJMoa1209124

Vojta, A., Dobrinic, P., Tadic, V., Bockor, L., Korac, P., Julg, B., … Zoldos, V. (2016). Repurposing the CRISPR-Cas9 system for targeted DNA methylation. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkw159

Voss, A. K., Collin, C., Dixon, M. P., & Thomas, T. (2009). Moz and Retinoic Acid Coordinately Regulate H3K9 Acetylation, Hox Gene Expression, and Segment Identity. *Developmental Cell*. https://doi.org/10.1016/j.devcel.2009.10.006

Wang, Z., Zang, C., Cui, K., Schones, D. E., Barski, A., Peng, W., & Zhao, K. (2009). Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. *Cell*. https://doi.org/10.1016/j.cell.2009.06.049

Wu, X., Deng, F., Li, Y., Daniels, G., Du, X., Ren, Q., … Lee, P. (2015). ACSL4 promotes prostate cancer growth, invasion and hormonal resistance. *Oncotarget*. https://doi.org/10.18632/oncotarget.6438

Wu, Y., Zhang, Z., Cenciarini, M. E., Proietti, C. J., Amasino, M., Hong, T., … Xu, K. (2018). Tamoxifen resistance in breast cancer is regulated by the EZH2–ERa–GREB1 transcriptional axis. *Cancer Research*. https://doi.org/10.1158/0008-5472.CAN-17-1327

Xie, S., Wang, Z., Okano, M., Nogami, M., Li, Y., He, W. W., … Li, E. (1999). Cloning, expression and chromosome locations of the human DNMT3 gene family. *Gene*. https://doi.org/10.1016/S0378-1119(99)00252-8

Xie, W., Barr, C. L., Kim, A., Yue, F., Lee, A. Y., Eubanks, J., … Ren, B. (2012). Base-resolution analyses of sequence and parent-of-origin dependent DNA methylation in the mouse genome. *Cell*. https://doi.org/10.1016/j.cell.2011.12.035

Yan, Y., Zuo, X., & Wei, D. (2015). Concise Review: Emerging Role of CD44 in Cancer Stem Cells: A Promising Biomarker and Therapeutic Target. *STEM CELLS Translational Medicine*. https://doi.org/10.5966/sctm.2015-0048

Yang, L., Ye, F., Bao, L., Zhou, X., Wang, Z., Hu, P., … Bu, H. (2019). Somatic alterations of TP53, ERBB2, PIK3CA and CCND1 are associated with chemosensitivity for breast cancers. *Cancer Science*. https://doi.org/10.1111/cas.13976

Yang, X. J., & Seto, E. (2007). HATs and HDACs: From structure, function and regulation to novel strategies for therapy and prevention. *Oncogene*. https://doi.org/10.1038/sj.onc.1210599

Yeo, B., Turner, N. C., & Jones, A. (2014). An update on the medical management of breast cancer. *BMJ (Online)*. https://doi.org/10.1136/bmj.g3608

Yoshihara, K., Wang, Q., Torres-Garcia, W., Zheng, S., Vegesna, R., Kim, H., & Verhaak, R. G. W. (2015). The landscape and therapeutic relevance of cancer-associated transcript fusions. *Oncogene*. https://doi.org/10.1038/onc.2014.406

Yu, L., Liang, Y., Cao, X., Wang, X., Gao, H., Lin, S. Y., … Li, K. (2017). Identification of MYST3 as a novel epigenetic activator of ERα frequently amplified in breast cancer. *Oncogene*. https://doi.org/10.1038/onc.2016.433

Yuan, J., Reed, A., Chen, F., & Stewart, C. N. (2006). Statistical analysis of real-time PCR data. *BMC Bioinformatics, 7*(1), 85. https://doi.org/10.1186/1471-2105-7-85

Yun, M., Wu, J., Workman, J. L., & Li, B. (2011). Readers of histone modifications. *Cell Research*. https://doi.org/10.1038/cr.2011.42

Zeilstra, J., Joosten, S. P. J., Dokter, M., Verwiel, E., Spaargaren, M., & Pals, S. T. (2008). Deletion of the WNT target and cancer stem cell marker CD44 in Apc(Min/+) mice attenuates intestinal tumorigenesis. *Cancer Research*. https://doi.org/10.1158/0008-5472.CAN-07-2940

Zelnak, A. B., & O'Regan, R. M. (2015). Optimizing endocrine therapy for breast cancer. *JNCCN Journal of the National Comprehensive Cancer Network*. https://doi.org/10.6004/jnccn.2015.0125

Zhang, H., Lang, Q., Li, J., Zhong, Z., Xie, F., Ye, G., … Yu, L. (2007). Molecular cloning and characterization of a novel human glycine-N-acyltransferase gene GLYATL1, which activates transcriptional activity of HSE pathway. *International Journal of Molecular Sciences*. https://doi.org/10.3390/i8050433

Zhang, T., Cooper, S., & Brockdorff, N. (2015). The interplay of histone modifications - writers that read. *EMBO Reports*. https://doi.org/10.15252/embr.201540945

Zhou, Y., Yau, C., Gray, J. W., Chew, K., Dairkee, S. H., Moore, D. H., … Benz, C. C. (2007). Enhanced NFκB and AP-1 transcriptional activity associated with antiestrogen resistant breast cancer. *BMC Cancer*. https://doi.org/10.1186/1471-2407-7-59

Abbreviations

4-OHT	4-Hydroxytamoxifen
5-Aza-dC	5-Aza-2′-deoxycytidine
AI	Aromatase inhibitor
AKT	Serine/Threonine Kinase 1
ANK	Ankyrin
ATAC	Assay for Transposase-Accessible Chromatin
ATP	Adenosine triphosphate
BAMBI	BMP and activin membrane bound inhibitor
BCA	Bicinchoninic acid protein assay
BCa	Breast cancer
BL	Basal-like
BRCA	Breast cancer
BSA	Bovine Serum Albumin
Cas9	CRISPR associated protein 9
CDK	Cycling dependent kinase
cDNA	Complementary deoxyribonucleic acid
CGI	CpG island
ChIP	Chromatin immunoprecipitation
CO_2	Carbon dioxide
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeat
CRISPRa	CRISPR activation
CRISPRi	CRISPR interference
CTRL	Control
CYP2D6	Cytochrome P450 Family 2 Subfamily D Member 6
dCas9	Dead Cas9
DMEM	Dulbecco´S Modified Eagle´s Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid

DNMT	DNA methyltransferase
DOT1L	DOT1 Like Histone Lysine Methyltransferase
E2	Estrogen
EDTA	Ethylenediaminetetraacetic acid
EdU	5-Ethynyl-2'-deoxyuridine
EGFR	Epidermal growth factor receptor
ER-	Estrogen receptor negative
ER+	Estrogen receptor positive
ERBB2	Erb-B2 Receptor Tyrosine Kinase 2
ERα	Estrogen receptor-alpha
ESR1	Estrogen Receptor 1
EtOH	Ethanol
FACS	Fluorescence-activated cell sorting
FBS	Fetal Bovine Serum
FISH	Fluorescence in situ hybridisation
FITC	Fluorescein isothiocyanate
G9a	Histone-lysine N-methyltransferase EHMT2
G9a	Histone-lysine N-methyltransferase EHMT2
GATA3	5-Ethynyl-2'-deoxyuridine
GLYATL1	Glycine-N-acyltransferase-like 1
GNAT	GCN5-Related N-Acetyltransferases
HAT	Histone acetyltransferase
HDAC	Histone deacetylase
HER2	Human Epidermal Growth Factor Receptor 2
HMT	Histone methyltransferase
IHC	Immunohistochemistry
JNK	c-Jun N-terminal kinase
LAR	Luminal androgen receptor
LTED	Long term estrogen deprivation
M	Mesenchymal

MAPK	Mitogen-Activated Protein Kinase
METABRIC	Molecular Taxonomy of Breast Cancer International Consortium
mRNA	Messenger ribonucleic acid
mTOR	Mechanistic Target Of Rapamycin Kinase
NAD	Nicotinamide adenine dinucleotide
NADH	Nicotinamide adenine dinucleotide + hydrogen
NaF	Sodium Fluoride
ORF	Open reading frame
OS	Overall survival
PAM	Protospacer adjacent motif
PBS	Phosphate buffered saline
PE	Phycoerythrin
PFA	Paraformaldehyde
PI	Propidium iodide
PI3K	Phosphatidylinositol-4,5-Bisphosphate 3-Kinase
PI3KCA	Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha
PR	Progesterone receptor
PUM1	Pumilio RNA Binding Family Member 1
PVDF	Polyvinylidene fluoride
qRT-PCR	Quantitative reverse transcriptase polymerase chain reaction
RFS	Recurrence-free survival
RIPA	Radio-immuno precipitation assay
RNA	Ribonucleic acid
RNAi	Ribonucleic acid interference
RPPA	Reverse phase protein array
RT	Room temperature
RTCA	Real time cell analyzer
SAM	S-Adenosyl methionine
SD	Standard deviation
SDS PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis

SEM	Standard error of mean
SERD	Selective estrogen receptor downregulator
SERM	Selective estrogen receptor modulator
SET	Su(var)3-9, Enhancer-of-zeste and Trithorax
sgRNA	Single guide RNA
siRNA	Small interfering ribonucleic acid
SLC9A3R1	Solute Carrier Family 9 (Sodium/Hydrogen Exchanger), Isoform 3 Regulatory Factor 1
TAMR	Tamoxifen resistant
TBST	Tris buffered saline with Tween 20
TCGA	The cancer genome atlas
TEMED	Tetramethylethylenediamine
TET	Ten-eleven Translocation
TGFβ	Transforming growth factor β
TNBC	Triple negative breast cancer
TP53	Tumor protein 53
TSS	Transcription start site
UCSC	University of California, Santa Cruz
Wnt	Wingless-type MMTV integration site family
WT	Wild type

Acknowledgements

"One´s destination is never a place, but a new way of seeing things."
— Henry Miller

I would like to start by thanking my supervisor, Prof. Dr. Stefan Wiemann, for giving me the opportunity to work in the Division of Molecular Genome Analysis and for providing me the freedom to pursue my own research ideas.

I would also like to thank my Thesis Advisory Committee members, Prof. Dr. Christoph Plass, Dr. Karin Müller Decker, Prof. Dr. Robert Brown and Prof. Dr. Marianne Rots for guiding and advising me through my PhD even beyond the TAC meetings. I am grateful to my examiners, Prof. Dr. Stefan Wiemann, Dr. Karin Müller-Decker, Dr. Odilia Popanda and Prof. Dr. Marius Lemberg for agreeing to take part in my defense.

This project was a result of collaborations that extended beyond my lab. I would like to thank Dr. Luca Magnani from Imperial College London for kindly providing the resistant cells. I would also like to extend my gratitude to Prof Dr. Marianne Rots from UMCG for hosting me in Gröningen where I learned the epigenetic editing method. I would like to thank Dr. Perry Moerland and Maryam Soleimani-Dodaran.for the bioinformatics analysis. I would like to thank Dr. Simin Oz for the in-house collaborations with ATAC-seq experiments and for analyzing the data. I would like to thank Mihaly Koncz for all the plasmids he generated and for being a great friend and a lab mate during our time in Gröningen. I would like to thank every member of the EpiPredict Consortium for their guidance, friendship and support.

I would like to thank each member of the MGA division for creating a great research environment. I would especially like to thank Daniela for her awesome western blots, Angelika for her help with TaqMan and Rainer for generating countless numbers of stable cell lines for me. I am grateful to Sara, Sabine, Heike, Corinna, Mireia, Xiaoya, Janine, Zhivka, Cindy, Birgit, Ewald, Rita, Katja and Daniela each of whom has helped me along the way. I would like to thank Vino for choosing our lab to do her master thesis and choosing me as her supervisor (though still I cannot fathom why). "BAMBI-Lost in the forest of epigenetics" would not have been possible without you. I would like to thank Lukas for his undying enthusiasm and his devotion for the formatting of this thesis and his patience with the lack of my Microsoft Word skills. I would like to thank Ana for always being willing to help. I would also like to thank my flatmates, Edoardo and Andreas, who not only helped me find accommodation and but also created a mediterranean atmosphere in the midst of Germany.

Exclusive thanks go to kindred spirits who made this journey an unforgettable one. I would like to thank Alex, for being a constant source of joy and entertainment and proving that a person can turn 50 shades of red in the slightest embarrassment; and Chiara for organizing all the fancy celebrations and for reminding us of the passion for science. I would like to thank my office pals (a.k.a. party people), Devina, for her inspiring quotes that were immortalized in the form of post-it notes and calling my English impeccable once; and Nese, for being my intervention partner and for always helping me to get in touch with reality. I am still curious to see whether indeed nothing good happens after 2 am or not. I could not have asked for a crazier office environment. I would like to thank Simone, with whom I started this journey together and without whom this journey would have been a very different one. Thank you for being my PhD brother and my friend. I consider myself lucky to have known y'all. Thank you for all the memories I will cherish for the rest of my life and for making Heidelberg feel like home for the last 4 years.

I would like to thank my close friends who are spread all around the world for their support and encouragement even from afar. I am especially grateful to Ezgi, Ferhat and Ekin who have always been a phone or skype call away.

I am grateful to my family for providing me with unconditional love and support all through my life.

Thank you all for this invaluable and fascinating experience!

"But we're never gonna survive, unless

We get a little crazy"