Regulation of estrogen receptor alpha by promoter methylation, alternative splicing and microRNA

Albrecht, Juliane

2024

Document Version:
Publisher's PDF, also known as Version of record

Regulation of estrogen receptor alpha by promoter methylation, alternative splicing and microRNA

JULIANE ALBRECHT
DEPARTMENT CLINICAL SCIENCES, LUND | FACULTY OF MEDICINE | LUND UNIVERSITY

Regulation of estrogen receptor alpha by promoter methylation, alternative splicing and microRNA

Juliane Albrecht

DOCTORAL DISSERTATION

Doctoral dissertation for the degree of Doctor of Philosophy (PhD) at the Faculty of Medicine at Lund University to be publicly defended on the 23rd of October at 09.00 in Belfragesalen, BMC, Lund, Sweden

Faculty opponent
Claudia Kutter, PhD

Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet

Stockholm, Sweden

Organization: LUND UNIVERSITY
Document name: Doctoral dissertation **Date of issue** 2024-10-23
Author(s): Juliane Albrecht **Sponsoring organization:**

Title and subtitle: Regulation of estrogen receptor alpha by promoter methylation, alternative splicing and microRNA

Abstract:
Breast cancer, a common and often aggressive malignancy, presents significant challenges due to its ability to develop resistance to therapies. This resistance is intricately linked to the function and regulation of the estrogen receptor, a key driver of breast cancer cell proliferation. Our research focuses on unraveling the complex mechanisms underlying this resistance by investigating the roles of estrogen receptor isoforms, epigenetic modifications, and microRNAs (miRNAs). Through comprehensive analyses, we demonstrate that the estrogen receptor undergoes alternative splicing, producing multiple isoforms with distinct functional properties. These isoforms vary in their sensitivity to breast cancer therapies, potentially leading to differential treatment outcomes. Additionally, we explored how DNA methylation in regulatory regions of the estrogen receptor gene influences the expression and function of these isoforms, revealing a crucial layer of epigenetic control that could contribute to therapy resistance. Furthermore, our study identifies the miRNA miR-4728-3p, encoded by the ERBB2 oncogene, as a significant regulator of estrogen synthesis in breast cancer cells. By modulating the levels of aromatase and other estrogen-related enzymes, miR-4728-3p plays a pivotal role in the intricate network of factors driving breast cancer progression and resistance to treatment. These findings enhance our understanding of the multifaceted mechanisms of breast cancer resistance, providing valuable insights that could inform the development of more effective therapeutic strategies.

Key words: estrogen receptor, methylation, alternative splicing, microRNA

Language: English **ISSN and key title**: 1652-8220

ISBN: 978-91-8021-620-3

Number of pages: 134

Signature Date 2024-10-23

Regulation of estrogen receptor alpha by promoter methylation, alternative splicing and microRNA

Juliane Albrecht

LUND
UNIVERSITY

Faculty of Medicine
Department Clinical Sciences Lund

ISBN 978-91-8021-620-3
ISSN 1652-8220

Printed in Sweden by Media-Tryck, Lund University
Lund 2024

“Knowledge is knowing that a tomato is a fruit;
wisdom is not putting it in a fruit salad.”

— *Miles Kington*

Table of Contents

Abstract

Breast cancer, a common and often aggressive malignancy, presents significant challenges due to its ability to develop resistance to therapies. This resistance is intricately linked to the function and regulation of the estrogen receptor, a key driver of breast cancer cell proliferation. Our research focuses on unraveling the complex mechanisms underlying this resistance by investigating the roles of estrogen receptor isoforms, epigenetic modifications, and microRNAs (miRNAs). Through comprehensive analyses, we demonstrate that the estrogen receptor undergoes alternative splicing, producing multiple isoforms with distinct functional properties. These isoforms vary in their sensitivity to breast cancer therapies, potentially leading to differential treatment outcomes. Additionally, we explored how DNA methylation in regulatory regions of the estrogen receptor gene influences the expression and function of these isoforms, revealing a crucial layer of epigenetic control that could contribute to therapy resistance. Furthermore, our study identifies the miRNA miR-4728-3p, encoded by the *ERBB2* oncogene, as a significant regulator of estrogen synthesis in breast cancer cells. By modulating the levels of aromatase and other estrogen-related enzymes, miR-4728-3p plays a pivotal role in the intricate network of factors driving breast cancer progression and resistance to treatment. These findings enhance our understanding of the multifaceted mechanisms of breast cancer resistance, providing valuable insights that could inform the development of more effective therapeutic strategies.

Populärvetenskaplig sammanfattning

Bröstcancer är en av de vanligaste cancerformerna som drabbar kvinnor världen över. Trots betydande framsteg inom behandling står många patienter fortfarande inför utmaningen att deras cancer blir resistent mot behandlingarna. Att förstå orsakerna till denna resistens är avgörande för att förbättra behandlingsresultaten. Vår forskning är inriktad på tre nyckelområden: olika former av östrogenreceptorn, förändringar i DNA-regleringen och den roll som små molekyler, så kallade mikroRNA, spelar.

Östrogenreceptorn är ett protein som normalt bidrar till att upprätthålla den kvinnliga reproduktiva vävnaden genom att reagera på hormonet östrogen. Vid bröstcancer kan denna receptor dock främja tumörtillväxt om den blir överaktiv eller förlorar sina normala regleringsmekanismer.

Den gen som ansvarar för att producera östrogenreceptorn kan generera flera olika versioner av proteinet, så kallade isoformer, genom processer som alternativ splicing och användning av olika promotorer. Dessa isoformer kan ha olika funktioner och reagera olika på bröstcancerbehandlingar som tamoxifen och fulvestrant. I arbete I identifierade vi flera nya isoformer av östrogenreceptorn och fann att vissa av dem kan bidra till terapiresistens. En särskild variant uppvisade ett dominant-negativt beteende, dvs. den kan störa aktiviteten hos den ”korrekta” östrogenreceptorn och därmed eventuellt minska effekten av behandlingar. Vi upptäckte också att vissa isoformer ackumuleras mer i kärnan, vilket kan påverka deras interaktioner med andra cellulära komponenter och påverka behandlingsresultaten.

I arbete II undersökte vi hur förändringar i DNA-regleringen, framför allt genom en process som kallas DNA-metylering, kan påverka östrogenreceptorns aktivitet och dess protein. DNA-metylering innebär att kemiska markörer läggs till DNA som kan slå på eller av gener utan att ändra den underliggande genetiska koden. Vi fann att högre nivåer av metylering i en specifik region av östrogenreceptorgenen var förknippade med lägre receptoruttryck och förlängd resistens mot läkemedlet fulvestrant i bröstcancerceller. Detta tyder på att metyleringsmönster kan spela en avgörande roll för att upprätthålla resistens.

I arbete III analyserade vi mikroRNA:s roll, små molekyler som reglerar genuttrycket genom att binda till budbärar-RNA. Vid bröstcancer kan vissa mikroRNA påverka koncentrationen av östrogenreceptorisoformer, vilket i sin tur

kan påverka effekten av behandlingar. Vi fokuserade på ett specifikt mikroRNA kallat miR-4728-3p, som kodas av ERBB2-genen, en gen som ofta förknippas med aggressiv bröstcancer. Vår forskning visade att miR-4728-3p spelar en viktig roll i regleringen av östrogenproduktionen genom att rikta in sig på enzymer som är involverade i biosyntesen. Blockering av miR-4728-3p ledde till ökad produktion av aromatas, ett nyckelenzym i östrogensyntesen, vilket kan främja tillväxten av cancerceller och bidra till behandlingsresistens.

Genom att undersöka samspelet mellan östrogenreceptorisoformer, epigenetiska förändringar och mikroRNA ger vår forskning djupare insikter i utvecklingen av terapiresistens vid östrogenreceptorpositiv bröstcancer. Dessa resultat är ett viktigt steg mot att förstå de biologiska mekanismer som ligger bakom resistens. Även om det fortfarande finns mycket kvar att upptäcka bidrar våra studier till utvecklingen av nya behandlingar som kan förbättra utfallet för dem som drabbats av bröstcancer.

Popular scientific summary

Breast cancer is one of the most common types of cancer affecting women worldwide. Despite significant advances in treatment, many patients still face the challenge of their cancer becoming resistant to therapies. Understanding the underlying mechanisms of this resistance is essential to improve treatment outcomes. Our research focuses on three important areas: different forms of the estrogen receptor, changes in DNA regulation and the role of small molecules called microRNAs.

The estrogen receptor is a protein that normally contributes to the maintenance of female reproductive tissue by responding to the hormone estrogen. In breast cancer, however, this receptor can promote tumor growth if it becomes overactive or loses its normal regulatory mechanisms.

The gene that is responsible for producing the estrogen receptor can generate multiple versions of the protein, known as isoforms, through processes such as alternative splicing and the use of different promoters. These isoforms may have different functions and respond differently to breast cancer treatments such as tamoxifen and fulvestrant. In Paper I, we identified several new isoforms of the estrogen receptor and found that some of them may contribute to therapy resistance. For example, one variant showed a dominant-negative behavior, which could interfere with the activity of the "proper" estrogen receptor and thus potentially reduce the efficacy of therapies. We also discovered that certain isoforms accumulate more in the nucleus, which could influence their interactions with other cellular components and affect treatment outcomes.

In Paper II, we investigated how changes in DNA regulation, particularly through a process called DNA methylation, can affect estrogen receptor activity and its protein. DNA methylation involves adding chemical markers to DNA that can turn genes on or off without altering the underlying genetic code. We found that higher levels of methylation in a specific region of the estrogen receptor gene were associated with lower receptor expression and prolonged resistance to the drug fulvestrant in breast cancer cells. This suggests that methylation patterns may play a critical role in maintaining resistance.

In Paper III, we investigated the role of microRNAs, small molecules that regulate gene expression by binding to messenger RNA. In breast cancer, certain microRNAs can affect the concentration of estrogen receptor isoforms, which in

turn can influence the efficacy of treatments. We focused on a specific microRNA called miR-4728-3p, which is encoded by the ERBB2 gene, a gene commonly associated with aggressive breast cancer. Our research revealed that miR-4728-3p plays an important role in regulating estrogen production by targeting enzymes involved in the biosynthetic pathway. Blocking miR-4728-3p led to increased production of aromatase, a key enzyme in estrogen synthesis, which could promote cancer cell growth and contribute to therapy resistance.

Our research provides new insights into the development of therapy resistance in estrogen receptor-positive breast cancer based on the interplay between estrogen receptor isoforms, epigenetic changes and microRNAs. These findings are an important step towards understanding the biological mechanisms underlying resistance. While there is still much to discover, our studies are contributing to the development of new therapies that could improve outcomes for those affected by breast cancer.

Populärwissenschaftliche Zusammenfassung

Brustkrebs ist eine der häufigsten Krebsarten, von denen Frauen weltweit betroffen sind. Trotz erheblicher Fortschritte in der Behandlung stehen viele Patientinnen immer noch vor der Herausforderung, dass ihr Krebs gegen die Therapien resistent wird. Um die Behandlungsergebnisse zu verbessern, ist es von entscheidender Bedeutung, die Ursachen für diese Resistenz zu verstehen. Unsere Forschung konzentriert sich auf drei Schlüsselbereiche: verschiedene Formen des Östrogenrezeptors, Veränderungen in der DNA-Regulation und die Rolle kleiner Moleküle, so genannter microRNAs.

Der Östrogenrezeptor ist ein Protein, das normalerweise zum Erhalt des weiblichen Fortpflanzungsgewebes beiträgt, indem es auf das Hormon Östrogen reagiert. Bei Brustkrebs kann dieser Rezeptor jedoch das Tumorwachstum fördern, wenn er überaktiv wird oder seine normalen Regulationsmechanismen verliert.

Das Gen, das für die Produktion des Östrogenrezeptors verantwortlich ist, kann durch Prozesse wie alternatives Spleißen und die Nutzung von unterschiedlichen Promotoren mehrere Versionen des Proteins erzeugen, die als Isoformen bezeichnet werden. Diese Isoformen können unterschiedliche Funktionen haben und unterschiedlich auf Brustkrebsbehandlungen wie Tamoxifen und Fulvestrant reagieren. Im Paper I haben wir mehrere neue Isoformen des Östrogenrezeptors identifiziert und festgestellt, dass einige von ihnen zur Therapieresistenz beitragen können. So zeigte eine bestimmte Variante ein dominant-negatives Verhalten, d. h. sie könnte die Aktivität des „richtigen“ Östrogenrezeptors beeinträchtigen und damit möglicherweise die Wirksamkeit von Therapien verringern. Wir entdeckten auch, dass sich bestimmte Isoformen stärker im Zellkern anreichern, was ihre Interaktionen mit anderen zellulären Komponenten beeinflussen und sich auf die Behandlungsergebnisse auswirken könnte.

Im Paper II untersuchten wir, wie Veränderungen in der DNA-Regulierung, insbesondere durch einen Prozess namens DNA-Methylierung, die Östrogenrezeptoraktivität und dessen Protein beeinflussen können. Bei der DNA-Methylierung werden der DNA chemische Marker hinzugefügt, die Gene an- oder abschalten können, ohne den zugrunde liegenden genetischen Code zu verändern. Wir fanden heraus, dass höhere Methylierungswerte in einer bestimmten Region des

Östrogenrezeptor-Gens mit einer geringeren Rezeptorexpression und einer verlängerten Resistenz gegenüber dem Medikament Fulvestrant in Brustkrebszellen verbunden waren. Dies deutet darauf hin, dass Methylierungsmuster eine entscheidende Rolle bei der Aufrechterhaltung der Resistenz spielen könnten.

Im Paper III untersuchten wir die Rolle der microRNAs, kleiner Moleküle, die die Genexpression durch Bindung an Boten-RNA regulieren. Bei Brustkrebs können sich bestimmte microRNAs auf die Konzentration von Östrogenrezeptor-Isoformen auswirken, was wiederum die Wirksamkeit von Behandlungen beeinflussen kann. Wir haben uns auf eine bestimmte microRNA namens miR-4728-3p konzentriert, die vom ERBB2-Gen kodiert wird, einem Gen, das häufig mit aggressivem Brustkrebs in Verbindung gebracht wird. Unsere Forschung ergab, dass miR-4728-3p eine wichtige Rolle bei der Regulierung der Östrogenproduktion spielt, indem sie auf Enzyme abzielt, die am Biosyntheseweg beteiligt sind. Die Blockierung von miR-4728-3p führte zu einer erhöhten Produktion von Aromatase, einem Schlüsselenzym der Östrogensynthese, was das Wachstum von Krebszellen fördern und zur Therapieresistenz beitragen könnte.

Durch die Untersuchung des Zusammenspiels zwischen Östrogenrezeptor-Isoformen, epigenetischen Veränderungen und microRNAs liefert unsere Forschung tiefere Einblicke in die Entwicklung von Therapieresistenz bei Östrogenrezeptor-positivem Brustkrebs. Diese Erkenntnisse sind ein wichtiger Schritt zum Verständnis der biologischen Mechanismen, die der Resistenz zugrunde liegen. Auch wenn es noch viel zu entdecken gibt, tragen unsere Studien zur Entwicklung neuer Therapien bei, die die Ergebnisse für die von Brustkrebs Betroffenen verbessern könnten.

List of Papers

Paper I

Alternative promoters and splicing create multiple functionally distinct isoforms of oestrogen receptor alpha in breast cancer and healthy tissues.

Carlos Enrique Balcazar Lopez, Juliane Albrecht, Völundur Hafstað, Cornelia Börjesson Freitag, Johan Vallon-Christersson, Cristian Bellodi, Helena Persson

Cancer Medicine, 2023. (18):18931-18945

Paper II

Dynamic methylation and expression of alternative promoters for oestrogen receptor alpha in cell line models of fulvestrant resistance

Juliane Albrecht, Mirjam Müller, Völundur Hafstað, Kamila Kaminska, Gabriella Honeth, Johan Vallon-Christersson and Helena Persson

Mol Oncology, 2024. doi:10.1002/1878-0261.13713

Paper III

The ERBB2-encoded miRNA miR-4728-3p regulates estrogen signaling in SK-BR-3 cells

Völundur Hafstad, Juliane Albrecht, Euisuk Han, Helena Persson

Manuscript

Author's contribution to the papers

Following the CRediT (Contributor Roles Taxonomy) system:

Paper I

Formal analysis, investigation, software, validation, visualization, writing – original draft, writing – review and editing.

Paper II

Methodology, software, validation, formal analysis, investigation, data curation, writing – original draft, writing – review & editing, visualization.

Paper III

Validation, formal analysis, investigation, writing – review and editing.

Abbreviations

4-OHT	4-Hydroxytamoxifen
AF1	Activation function 1
AF2	Activation function 2
AI	Aromatase inhibitor
ASO	Antisense oligonucleotide
bp	Base pair
cDNA	Complementary DNA
cds	Coding sequence
ChIP	Chromatin immunoprecipitation
CYP	Cytochrome P450
DCIS	Ductal carcinoma in situ
DBD	DNA-binding domain
DMR	Differentially methylated region
DNMT	DNA methyltransferase
dNTPs	Deoxyribonucleotide triphosphate
DNA-Seq	DNA sequencing
DMSO	Dimethyl sulfoxide
EGFR	Epidermal growth factor receptor
ERE	Estrogen response element
E1	Estrone
E2	Estradiol
E3	Estriol
E4	Estetrol
ER	Estrogen receptor alpha
ERBB2	Erythroblastic oncogene B
ERfl	Estrogen receptor full length
ERtr	Estrogen receptor truncated isoform
ESE	Exonic splicing enhancer
ESS	Exonic splicing silencer

FR	Fulvestrant resistant
FR-F	Fulvestrant resistant cultured without fulvestrant
H&E	Hematoxylin and Eosin
HER2	Human epidermal growth factor receptor 2
hnRNPs	Heterogeneous nuclear ribonucleoproteins
HR+	Hormone receptor-positive
HRP	Horseradish peroxidase
IHC	Immunohistochemistry
ISE	Intronic splicing enhancer
ISS	Intronic splicing silencer
kDa	Kilo Dalton
LBD	Ligand-binding domain
miRNA	microRNA
MRI	Magnetic resonance imaging
mRNA	Messenger RNA
NTD	N-terminal domain
ONT	Oxford Nanopore Technology
P	Parental
PacBio	Pacific Biosciences
PGV	Pathogenic germline variants
PPi	Pyrophosphate
PR	Progesterone receptor
PR-A and PR-B	Progesterone receptor isoforms A and B
PROTACs	Proteolysis targeting chimeras
RNAi	RNA interference
RNA-Seq	RNA sequencing
RFI	Recurrence-free interval
RS	Recurrence Score
RT-PCR	Real-time reverse transcription polymerase chain reaction
SCAN-B	Sweden Cancerome Analysis Network – Breast
SERM	Selective estrogen receptor modulator
SERD	Selective estrogen receptor downregulator
SMRT	Single-molecule real-time sequencing
SR	Serine/arginine-rich proteins
snRNPs	Small nuclear ribonucleoproteins
T-DM1	Trastuzumab-emtansine

T-DXd	Trastuzumab-deruxtecan
TCGA	The Cancer Genome Atlas
TDLU	Terminal duct lobular unit
TGFBI	Transforming growth factor beta induced
TNBC	Triple negative breast cancer
TNM	Tumor, Node, Metastasis
TRBP	TAR RNA binding protein
UTR	Untranslated region
XPO5	Exportin 5

Background

Introduction to cancer

Cancer represents a group of genetic diseases that affect different cell types and organs and is a major cause of death worldwide which can originate anywhere in the human body. Usually, human cells grow and multiply through a process known as cell division to form new cells whenever the body needs them. But when cells become old or damaged, they die and new cells take their place. However, occasionally this regulated process collapses and abnormal or damaged cells may grow and multiply even when they should not, leading to uncontrolled proliferation and changes in the surrounding microenvironment through paracrine signaling. In other words, the tumor itself becomes a 'micro-ecosystem' that is selecting for beneficial traits.

Cancer develops through the stepwise creation of genetic and epigenetic changes, allowing normal cells to divide and grow beyond normal tissue boundaries and become malignant. These malignant abnormal cells can form into tumors, which are lumps of tissue. The tumors can be cancerous (malignant), a condition where the cells become more aggressive, invade and metastasis, or non-cancerous (benign) (1). This transformation occurs when various growth control and differentiation programs are disrupted, leading to tumor development as cancer cells keep gaining modifications to adapt to their altering environment. This change from a normal cell to a malignant tumor and finally to a lethal metastatic tumor is well characterized. Once cells lose these control mechanisms, they rapidly acquire changes. Such adaptability is the reason why cancer therapies often fail and why cancer relapses frequently grow back more severe and more difficult to treat by spreading to other organs.

Such cells inactivate apoptotic pathways as normal cells do, but they develop independence from growth signals by interacting with surrounding tissues and the immune system in a different way than normal cells. Different molecular and pathological characteristics are used to classify various tumors into subtypes, which helps to determine prognosis and treatment. However, on an individual level, the challenge lies in the genetically and phenotypically distinct nature of each tumor, requiring precision medicine. It is challenging within an individual tumor due to intratumorally clonal diversity, which makes it difficult to accurately assess the

tumor in the clinic due to sampling bias. Indeed, a single biopsy might not be representative of the whole tumor, and more aggressive clones potentially exist elsewhere in the tumor or in separate foci (2,3).

Over the years, increasing understanding of the complexity of cancer biology has incorporated the study of the Hallmarks of Cancer (Figure 1) which are a collection of functional capabilities that human cells acquire during the transition from normal to neoplastic states initially reported in 2001 and then updated in 2011 and 2022 (4–6). These abilities are essential for the formation of malignant tumors and underlie stringent control mechanisms that prevent normal cells from switching towards selfish survival behavior.

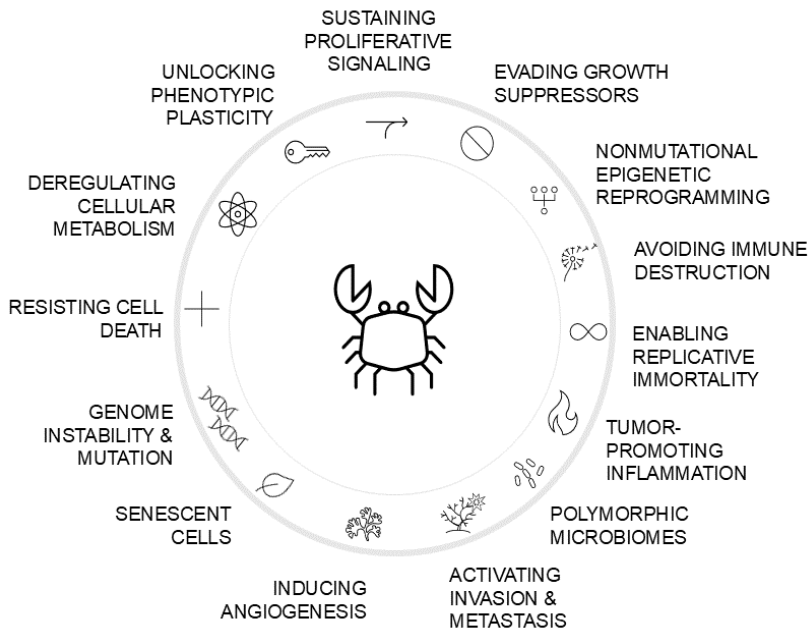

Figure 1 The hallmarks of cancer highlight key biological capabilities that enable cells to grow uncontrollably, evade death, and metastasize, driving tumor development and progression (modified image from (4–6)).

Their characteristics are the following: The sustaining of proliferative signaling where cancer cells undergo continuous self-signaling to divide, the evasion of growth suppressors enabling cells to avoid mechanisms that normally suppress cell growth, the resistance to cell death, avoiding apoptosis, which is the programmed process of cell removal that normally removes damaged or excess cells, the facilitation of replicative immortality, escaping the normal limitations of cell division, the induction of angiogenesis enabling cancer cells to stimulate the

development of new blood vessels to supply nutrients, and the activation of invasion and metastasis, promoting the spread (4).

Other features since 2011 include the deregulation of cellular metabolism to inhibit growth and survival, the prevention of immune destruction, genomic instability and mutation to promote growth, and the tumor-promoting inflammation (5). Finally, since 2022, the proposed hallmarks additionally include the unlocking of phenotypic plasticity allowing the cell to change in phenotype and adapt, non-mutational epigenetic reprogramming, the polymorphic microbiome involving the diversity of microbial communities linked to tumors affecting oncogenesis and treatment responses, and finally, the senescent cells resulting in an accumulation of cells that stop dividing but do not die, leading to tumor progression and therapy resistance (6).

The hallmarks of cancer provide a comprehensive framework for understanding the key features that enable tumor growth, evasion of immune responses and metastasis. Improving this understanding is essential for improving cancer diagnosis, developing targeted therapies, predicting patient outcomes and further refining research towards more effective treatments.

The female breast and female sex hormones

The female breast is a glandular organ on the woman chest with the important function of producing milk, formed by a system of ducts and lobules regulated by female sex hormones. The female breast consists of three types of tissue: glandular, adipose and connective tissue (Figure 2). Among these tissues, communication is critical to breast function, and it is well-established that hormone signaling actively influences breast development and function. The mammary glands have 15-20 milk-productive glands that are connected to ducts transporting milk to the nipple. The functional unit of the breast is the terminal duct lobular unit (TDLU), which consists of a collection of alveoli with surrounding ducts. Breast tissue is repeatedly remodeled during the fertile years, triggered by puberty, menstrual cycles, pregnancy and breastfeeding. In connection with menopause, the mammary gland normally undergoes involution, during a process whereby the TDLUs age and shrink. Incomplete involution is one of the biological processes associated with the development of breast cancer (7).

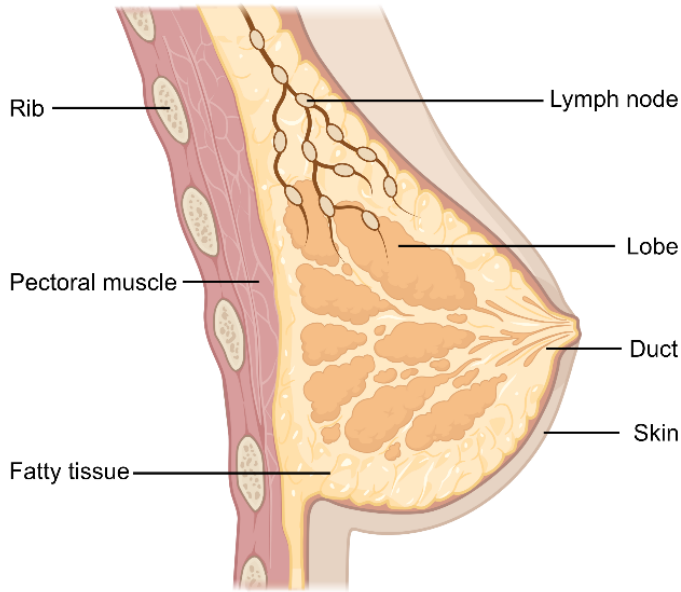

Figure 2 Anatomy of the female breast, showing the internal structures including the skin, fatty tissue, lobes, ducts, lymph nodes, pectoral muscle, and rib (Image created with BioRender).

As mentioned, the female breast is stimulated by the most important female sex hormones, the steroid hormones estrogen and progesterone. In addition to this, these hormones control the processes of female secondary sexual features and reproduction.

Estrogen is responsible for the development of the mammary glands and has a major role in triggering prolactin secretion by the pituitary gland. There are four estrogen hormones: estrone (E1), estradiol (E2), estriol (E3) and estetrol (E4), whereby E2 is considered the dominant hormone and is a powerful mitogen during the fertile years. The key enzyme for the synthesis of estrogen is aromatase (*CYP19A1*), which has been identified to be expressed and active in many human tissues, including endometrium, brain, bone, skin and adipose tissue (8). The biosynthesis of estrogens in women of reproductive age takes place primarily in the ovaries, but is also synthesized in extraglandular tissues and in the breast itself (9,10).

However, other tissues such as adipose tissue, osteoblasts, chondrocytes and the vascular endothelium also provide a significant contribution to estrogen production and estrogen metabolism (11). Through direct ovarian glandular secretion into the bloodstream, estradiol is released into the breast via an endocrine mechanism in premenopausal women. Estrogens stimulate the proliferation of normal and malignant cells via the induction of proteins involved in nucleic acid synthesis and the activation of growth-regulating genes. Increased cell proliferation may increase

the likelihood of errors in DNA repair, causing an accumulation of mutations over time. These mutations can contribute to the transition from normal cell growth to hyperplasia to neoplasia (12–14).

E2 is involved in controlling the growth of many breast tumors and is present in significant amounts in postmenopausal breast tumors, whereby their estrogens are produced by aromatase-mediated conversion of androgens from the adrenal glands and ovaries into estrogens in normal tissues (adipose tissue, muscle, liver or brain) as well as in breast tumors (8,15,16).

The second major sex hormone is progesterone, a small lipophilic steroid hormone with 21 carbon atoms, which is involved in the female menstrual cycle, pregnancy and embryogenesis by binding to progesterone receptors. Furthermore, it plays a fundamental role in normal female biology (17). The breast is a primary target of progesterone, where it regulates the development of the branching epithelial duct and the expansion of the milk-producing alveoli during lactation.

Progesterone acts via the progesterone receptor (PR), and estrogen acts via the estrogen receptor (ER), both belonging to the superfamily of nuclear steroid hormone receptors which will be described in more detail below.

Historical information
The first documented association between estrogens (produced in the ovaries) and breast cancer was observed by George Thomas Beatson in 1896. In a case report, he described a premenopausal breast cancer patient with metastasizing disease. Although not understanding the exact mechanisms of hormonal action in human physiology, he was aware of a method in bovine animals in which prolonged lactation was achieved by removing the ovaries. Inspired by this, Beatson removed the ovaries on both sides of his patient (oophorectomy), leading to a complete remission of the disease, due to the decreased level of estrogen (18).

Breast cancer

Epidemiology and risk factors

Worldwide, female breast cancer has overtaken lung cancer as the most frequently diagnosed type of cancer. This trend is expected to accelerate. By 2040, an estimated 28.4 million new cases of cancer are predicted, representing a significant increase of 47% in relation to estimates for 2020. In 2022, 10,952 women were diagnosed with breast cancer in Sweden, according to cancer statistics (19). However, as a result of early detection and treatment, the number of people who survive after a cancer diagnosis has increased over time (20).

Breast cancer is a heterogeneous disease characterized by several classification systems, ranging from invasive or *in situ* tumors, intrinsic subtypes, the detection of receptors including the ER, the PR and the human epidermal growth factor receptor 2 (HER2), the grade and the staging. Generally, breast cancer develops either in the ducts (80-85% of cases) or in the lobules (10-15% of cases) of the breast. The supportive tissue between the ducts and lobules is the stroma, which can also be the site of cancer development, resulting in sarcomas. Precancerous lesions are limited to the ducts or lobules and are unable to spread beyond the basement membrane; they are termed *in situ* or intraductal cancer (DCIS) (21,22). In contrast, true cancer cells can spread through the blood and lymph vessels as soon as they have invaded the basement membrane and are referred to as invasive or infiltrating cancer due to their infiltration through the ducts or lobules (1).

Approximately 90% of breast cancers occur as sporadic cases, defined as cases without a clear family history and are affected by specific risk factors, which can be classified into non-modifiable factors such as gender, age, genetics, breast density, benign breast diseases and endogenous hormones, and modifiable factors such as breastfeeding, alcohol consumption, smoking, childbirth, obesity and exogenous hormones (15,23–27).

In order to illustrate how risk factors influence breast cancer risk, I will explain an example from each group of factors to help to understand the association between risk factors and the development of breast cancer. Among the non-modifiable factors, one known risk factor for breast cancer is increased breast density. While the causes of increased breast density are not yet fully understood, they have been linked partly to genetic factors. Research with twins has shown that breast density has a hereditary component. Furthermore, estrogen levels are also involved, as external estrogens have the potential to increase breast density, while anti-estrogens can decrease it. High breast density not only makes mammography more challenging, but also increases the risk of developing breast cancer (28).

One other relevant risk factor among the modifiable factors is obesity, especially in post-menopausal women (26). Obesity is linked to a higher risk of breast cancer, as adipose tissue is an important source of estrogen production after menopause. Increased estrogen production can promote the growth of estrogen-dependent tumor (29–31). On top of this, obesity is often associated with chronic inflammation and insulin resistance, which can both further contribute to cancer risk (32,33).

Approximately 10% of breast cancer cases that are not sporadic are related to genetic mutations and/or hereditary breast cancer with a genetic predisposition (34–36). In this context, pathogenic mutations of the *BRCA1* and *BRCA2* genes are commonly referred to as high-risk genetic factors, significantly increasing the likelihood of developing breast and ovarian cancers. Approximately 5-10% of all breast cancer cases are linked to hereditary factors, specifically due to pathogenic

mutations in genes such as *BRCA1* and *BRCA2* (37,38). These mutations are referred to as pathogenic germline variants (PGV). Pathogenic mutations found in the *BRCA1* and *BRCA2* genes contribute significantly to hereditary breast cancer cases, representing around 20% of hereditary cases. There are significant risks associated with PGVs as women carrying these mutations in these genes have a 45-65% lifetime risk of developing breast cancer (39,40).

Diagnosis and symptoms

The most prominent symptom of breast cancer is the development of a new lump in the breast, which is usually detectable from about 1 to 2 centimeters in size, although this also depends on the location of the lump and the density of the breast tissue. Besides lumps, changes in breast dimensions, shape and texture, such as skin rashes, fever or redness, can also be a sign of breast cancer. But the incidence of large tumors (more than 5 cm) has decreased significantly compared to the past (41).

Awareness of breast cancer has increased significantly, and self-examination has often been promoted by the media over the last ten years. Additionally, many countries have introduced screening programs in which mammography is used for early detection. In Sweden, for example, it is recommended for women to have a mammogram every two years between the ages of 40 and 74. In comparison to self-examination, mammography is a more accurate method for the early detection of tumors as it lowers the rate of advanced breast cancer and consequently reduces mortality. A clinically detectable tumor size of about 1 cm^3, which corresponds to about one gram, already contains around one billion tumor cells, indicating that a tumor may contain millions to billions of cells at diagnosis (42). Within the mammogram, small calcifications, so-called microcalcifications, which often reflect the first signs of cancer or precancerous changes, can be visualized. The mammography results are categorized in is normal, benign, unspecific/investigation case, suspected malignancy and malignant (43).

However, breast characteristics such as tissue density can affect the sensitivity of mammograms because the density influences the clarity of underlying abnormalities, making it harder to detect potential cancerous lesions. Other imaging techniques, such as ultrasound, are particularly helpful in detecting tumors in dense breast tissue, as they provide better differentiation between tissue types and tumors. Magnetic resonance imaging (MRI) is mainly used for women at very high risk, for undiagnosed primary tumors or for tumors in underlying areas of the breast but is associated with a high rate of false positives (43).

Currently, histopathological methods are used in combination with molecular tests to classify the different molecular subtypes of breast cancer more appropriately. Molecular diagnosis has an important impact on the management and personalized treatment of breast cancer.

Sampling – histopathology

Once a suspicious lump is found that could indicate breast cancer, a biopsy is performed using a needle guided by imaging. This procedure helps determine whether the lump is cancerous, if the cancer has spread or is localized, its type (such as ductal, lobular, or mucinous), its grade, and which receptors it expresses. When a suspicious nodule is detected through imaging that may indicate breast cancer, a biopsy is performed to confirm the diagnosis. Core needle biopsy, the most commonly used technique, involves using a hollow needle to extract a tissue sample from the nodule for further analysis.

After a biopsy is taken, the sample is processed in the pathology department through a series of steps, including dehydration, fixation, embedding, and staining. The most common staining technique used is Hematoxylin and Eosin (short H&E), which shows the structure and cellular details of the tissue. Beside the H&E, additionally incubation with a panel of specific biomarkers to assess the presence of receptors and other molecular features is done (44). Pathologists then evaluate the biopsy for key characteristics such as cellular morphology, presence of dysplasia, and receptor expression, all of which contribute to determining whether the nodule is benign or malignant and to guiding appropriate therapy recommendations.

Anatomic staging

The TNM staging system (45) is a globally standardized classification system for the staging of cancer, which includes breast cancer. It was developed by the Union for International Cancer Control (UICC) and the American Joint Committee on Cancer (AJCC) and is used to assess the spread and progression of cancer (46). The ranking system stands for Tumor, indicating the size and spread of the primary tumor, Node, describing the occurrence and extent of spread to the local lymph nodes and Metastasis, describing whether distant metastases are present.

Table 1 TNM staging classification system.

TNM Staging	
Tumor size	*Lymph Nodes*
T0: in situ = pre-invasive	N0: negative
T1: <2 cm	N1: 1-3 positive
T2: 2-5 cm	N2: 4-9 positive
T3: > 5 cm	N3: 10 or more positive
T4: inflammatory, invasion through muscle or skin = aggressive tumor type	*Metastases*
	M0: none
	M1: distant (outside of breast and lymph nodes)

T4 includes the invasion of the skin and chest wall, subdivided based on whether the cancer has spread to the chest wall, skin, both, or in the case of inflammatory carcinoma.

The nodes category includes the lymph nodes in the armpit, behind the breastbone (sternum) and around the collarbone. Within this category, the focus is also on the sentinel lymph node, i.e. the lymph node closest to the primary tumor site most likely to harbor metastases due to fluid drainage. For this reason, it is also of prognostic relevance for breast cancer patients, and the results of lymph node biopsies can have a significant influence on surgical and treatment recommendations. Lymph node biopsies can be classified with positive referring to tumor cells in the node (i.e. macro metastases over 2 mm or micro metastases of 0.2 to 2 mm) or negative (where even isolated tumor cells are under 0.2 mm) (43). The metastases category retains unchanged its core definitions and indicates whether distant metastases are present or not. The most common metastases in primary breast cancer include brain, bone, lungs and liver (47). Patients are then classified based on TNM in four different prognostic groups.

Prognostic staging

Besides the commonly used TNM classification (anatomical staging), additional indicators are used to classify the tumor more precisely. Such additional factors include the grade, the receptor status (biomarkers) and genomic tests such as Oncotype DX (48), MammaPrint (49), PAM50 (50), and the Breast Cancer Index (51) and others. All of these have been implemented into the staging system to improve treatment.

Grading

Besides the TNM system, the grading is an important additional indicator based on the semi-quantitative Nottingham grading system (52). It indicates the extent of tubular formation along with the status of nuclear pleomorphism and mitoses of the cancer cells. By comparing breast cancer cells with normal breast epithelial cells using three morphological features, the microscopic analysis assigns a score from 1 to 3 to each feature and then combines these scores to produce three levels: Grade I (well differentiated), Grade II (moderately differentiated) and Grade III (poorly differentiated). The tumors of grade I, commonly referred to as low-grade tumors, appear similar to normal breast tissue and tend to grow and spread more slowly. As opposed to this, grade III tumors, referred to as high-grade tumors, typically show more rapid growth and spread (53).

Biomarker

Biomarkers are another factor in prognostic staging. The term refers to a measurable parameter of a biological process and can range from small molecules such as glucose in diabetes to proteins like the estrogen receptor in ER positive breast cancer. The purpose of biomarkers enables the diagnosis of diseases and/or provides information on disease progression, enabling treatment to be adapted, for example. They are prognostically or diagnostically significant and therefore act as important indicators in the clinic.

Biomarkers are a prognostic tool based on quantitative assessment of the hormone receptors (HR) including ER and PR, human epidermal growth factor receptor (HER2, originating from the *ERBB2* gene), and Ki-67 in addition to the clinicopathological information (54,55). It is of important relevance in the diagnosis of breast cancer based on biopsy samples.

Using an antibody detection method, immunohistochemistry, the tumor can be assessed according to whether it expresses the specific protein or not (in other words, whether it is positive or negative). Briefly, a positive ER and PR status in Sweden means that more than 10% of the cells express the receptor (43), indicates that the cancer cells are likely to respond to hormone therapies (56). A positive HER2 status is represented by different levels, which indicate different levels of aggressive cancer growth and often require targeted therapy with antibodies. Approximately 15-20% of breast cancer cases do not express any of these three proteins and are classified as triple-negative breast cancer (TNBC; negative for ER, PR, and HER2) (57).

The Ki-67 marker is used to assess cell proliferation (58), with high values indicating strong cell proliferation, which correlates with more aggressive tumors. All together give a strong indication of protein expression and tumor behavior which is needed for treatment prediction.

Genomic tests and molecular subtypes

The final element of prognostic staging includes the application of genomic testing, typically gene expression profiling using RNA sequencing or microarrays to detect gene amplification.

In Sweden, these tests are used in postmenopausal women with node-negative, ER-positive and HER2-negative breast cancer with unclear risk assessment, regardless of tumor size, before the determination of the chemotherapy schedule. While several validated tests are available like MammaPrint (49), Oncotype DX (48), or Prosigna PAM50 (50) (Table 2), where the last two multigene signature tests are recommend to use in Sweden (43).

Table 2 Gene expression profiling tests, their results and use.

Test	Number of genes analyzed	Result	Use/Prediction
Prosigna (PAM50)	50	Risk of Recurrence (ROR) Score and Intrinsic Subtype	• Predicts the risk of distant recurrence and classifies tumors into intrinsic subtypes: luminal A, luminal B, HER2-enriched, and basal-like. • Primarily used for early-stage, hormone receptor-positive (HR+), HER2-negative breast cancer.
Oncotype DX	21	Recurrence Score	• Predicts the risk of breast cancer recurrence and the potential benefit from adjuvant chemotherapy in early-stage, HR+, HER2-negative breast cancer. • Guides decisions on whether chemotherapy is needed in addition to hormone therapy.
MammaPrint	70	Risk of Recurrence Score	• Predicts the risk of distant recurrence in early-stage breast cancer for both hormone receptor-positive and hormone receptor-negative cases. • Helps determine the need for adjuvant chemotherapy.

Along with the availability of affordable microarray and high-throughput sequencing technologies, there are increasing efforts to define subtypes based on mutational signatures or gene expression profiles (50,59).The PAM50 classification system is the most important as it divides breast tumors into five molecular subtypes: luminal A, luminal B, HER2-enriched, basal and normal-like (Figure 3). The correlation between molecular subtypes and immunohistochemistry (IHC) phenotypes by using biomarkers is relatively high at approximately 75-90% (60).

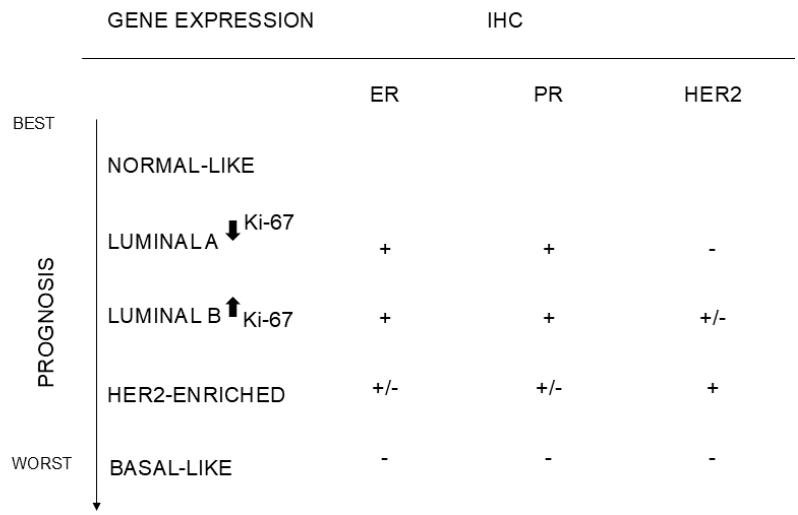

Figure 3 Intrinsic breast cancer subtypes (ranging from normal-like, luminal A and B, HER2-enriched, and basal-like), as well as prognosis and corresponding biomarker status by IHC.

Gene expression profiling has enabled the assignment of a distinct biological unit to each subtype, characterized by differences in cancer incidence, prognosis and treatment outcome (50,59).

Treatment of breast cancer

Multidisciplinary tumor conference

At the beginning of breast cancer treatment, the Swedish national guidelines recommend a preoperative multidisciplinary tumor conference to discuss the treatment strategies for all breast cancer cases. Typically, these conferences are attended by radiologists, pathologists, surgeons, oncologists and specialized nurses. The purpose of this collaborative approach is to optimize the individual treatment strategy both before and after surgery (43). After this conference, various local and systemic therapies are administered, including surgery, radiotherapy, chemotherapy, immunotherapy and targeted therapies. The therapeutic approaches differ clearly between early and advanced breast cancer as well as depending on hormone receptor and HER2 status (43).

Surgery

Surgery is the first main treatment for stage I–III primary breast cancer to control the tumor in the local and surrounding areas. Historically, a more aggressive approach to removing tumor tissue was believed to result in better outcomes. This led to radical mastectomy procedures, where the entire breast and surrounding muscles were removed. Unfortunately, these procedures often caused significant deformities and were frequently associated with painful arm swelling after the complete removal of lymph nodes in the armpit. More recent research has shown that more conservative techniques like lumpectomy along with radiotherapy are equivalent in effectiveness to mastectomy, leading to a transition in favor of breast-conserving surgery preferred approach. (61–63). In the past few years, an average of 70% of patients with invasive breast tumors underwent breast-conserving surgery, resulting in more favorable aesthetic results (64,65). In patients with a family history of breast cancer who have PGVs in genes that are susceptible to breast cancer, a prophylactic mastectomy can be performed to reduce the risk of developing breast cancer in the future (66–68).

Besides tumor removal, another focus is on the lymph nodes. The axillary lymph node removal, which was once routinely performed, is replaced by sentinel node biopsy, which is a more patient beneficial and advanced technique (69–72). The sentinel lymph node is the primary lymph node into which cancer cells are most likely to spread from the first primary tumor until they can be detected in the axilla. The sentinel node biopsy identifies the first lymph nodes draining the breast tumor and these are removed for pathological analysis (71). If a negative biopsy is obtained, it means that cancer cells have not yet spread to nearby lymph nodes or other organs. This technique has replaced axillary dissection as the standard procedure and offers comparable results with significantly fewer side effects.

Radiotherapy

Radiotherapy uses high-precision radiation to destroy cancer cells, shrink tumors, and reduce the risk of recurrence. The type and timing of radiotherapy depend on factors like cancer type, tumor size, location, and the patient's health. For breast cancer, post-operative radiotherapy is standard, particularly after breast-conserving surgery, to minimize recurrence (43,73). This tailored approach maximizes treatment effectiveness while minimizing side effects. Radiotherapy can be used before surgery to shrink tumors or after surgery to prevent cancer from returning. In some cases, radiotherapy is applied during surgery (intraoperative radiotherapy) to target the tumor immediately.

Chemotherapy

An additional option to treat breast cancer is chemotherapy, which aim is to stop the growth of cancer cells that are actively dividing by either killing them or preventing them from dividing, in order to prevent invasion, metastasis and recurrence.

Since cancer cells multiply more frequently than most normal cells, chemotherapy aims to kill them while sparing intact cells that are not actively dividing. Different chemotherapeutic agents achieve this in different ways, e.g. by damaging the cell's control center that regulates division, by disrupting the chemical processes involved in cell division, by damaging the cells during replication of their genes or by damaging the cells at the site of division (74). Conventional chemotherapeutic agents disrupt the synthesis and function of macromolecules in cancer cells by attacking DNA, RNA or protein synthesis or by interfering with the function of already formed molecules. This disruption can either lead directly to cell death or trigger apoptosis. The cell death may be prolonged, requiring repeated treatments to achieve a response. The most cytotoxic drugs target either the S phase (DNA synthesis phase) of the cell cycle, or the M phase and block mitotic spindle formation (74,75). Often, a combination of chemotherapeutic agents is given to patients to target cells at multiple stages of cell division, improving the likelihood of killing more cancer cells. However, this can also lead to toxic effects on normal cells.

Nowadays, the standard protocol for chemotherapy includes both anthracyclines (type of antibiotic with DNA intercalation function) and taxanes (cytostatic disrupting of microtubule function) (43,76), applied in both the neoadjuvant and adjuvant phases, giving equivalent results (77). For example, according to the Swedish guidelines, adjuvant chemotherapy is recommended for patients under 35 years of age (43). The standard treatment for patients with inoperable primary tumors and for patients with certain subtypes such as HER2-positive breast cancer and TNBC is neoadjuvant chemotherapy (43,78,79).

Immunotherapy

Immunotherapy is a treatment option for breast cancer that aims to activate the immune system by recognizing and fighting cancer cells. Using the natural ability of the immune system, this therapy helps it to fight tumors more effectively. There are several different approaches to immunotherapy, which include:

Immune checkpoint inhibitors which are designed to block immune checkpoints that normally prevent the immune response from becoming excessive. By blocking checkpoint receptors such as CTLA-4 and PD-1, the T cell activity against cancer cells is enhanced. Recent developments in the treatment of TNBC include the use of immune checkpoint inhibitors that target PD-1 or its ligand PD-L1, since this specific protein is in some cases express (80,81). They are monoclonal antibodies

that block the PD-L1/PD-1 complex to promote the T cell-mediated destruction of cancer cells (82). The use of such inhibitors is established internationally and also in Sweden (43).

Endocrine therapy

Endocrine therapy, also called hormonal therapy, is one of the most important treatment possibilities for hormone receptor-positive breast cancer. The estrogen receptor is the main target of endocrine therapy since it is expressed in between 75% in all breast cancer patients worldwide (83,84). Around 50% or more of patients with ER-positive breast cancer benefit from adjuvant endocrine treatment (85,86), meaning that this treatment reduces recurrence and increases survival (87,88).

In Sweden, nearly all patients with ER-positive tumors receive endocrine therapy, excluding those with the smallest tumors (< 10 mm) and without lymph node involvement (43), since good results are already achieved by surgical removal.

Estrogen is one of the main regulators of breast tissue growth through activation of the ER and the resulting growth programs. Targeting these signaling pathways as a form of cancer therapy has revolutionized the treatment of ER-positive breast cancer (89). Understanding the estrogen receptor mechanism is essential for the successful treatment of ER-positive breast cancer. The natural ligand estradiol binds to the estrogen receptor in the cytoplasm, dimerizes and translocate into the cell nucleus. Once there, it binds to a specific DNA sequence known as the estrogen response element (ERE) and activates the transcription of the target gene (90,91).

For the treatment of ER-positive breast cancer, there are three classes of anti-hormonal endocrine agents (Figure 4). These include the selective estrogen receptor modulators (SERMs), competing with estrogen for ER binding and exhibiting tissue-specific mixed agonist/antagonist capabilities, acting directly at the receptor and blocking ER activity; a good example is tamoxifen. Secondly, the aromatase inhibitors (AIs), which inhibit estrogen synthesis; one example is letrozole. Finally, the selective estrogen receptor downregulators (SERDs), which lead to destabilization and degradation of the ER; one example is ICI 182,780 (fulvestrant) (92,93). All three treatments were applied in paper I-III.

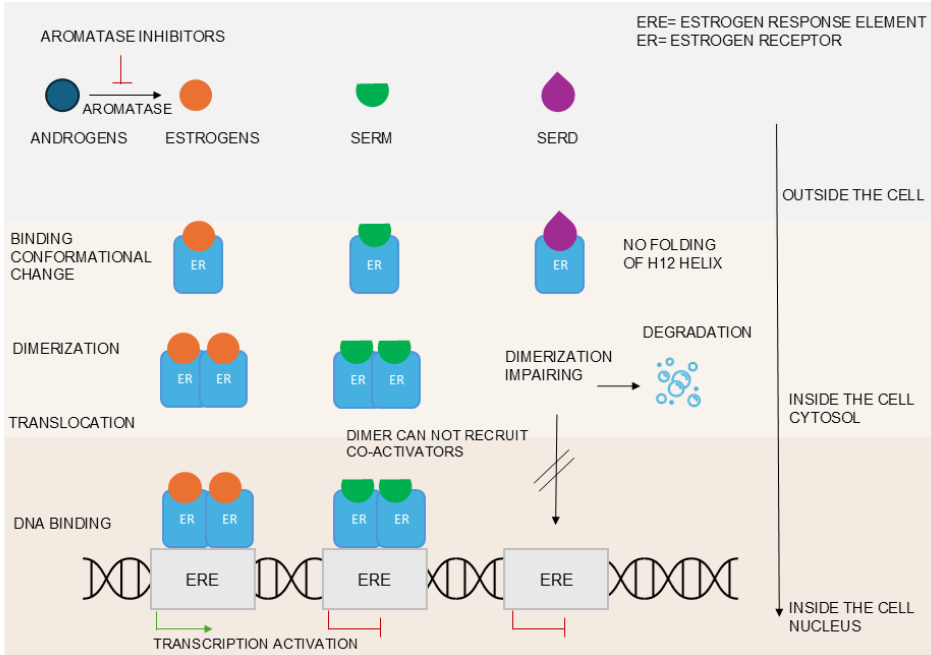

Figure 4 The classical mechanisms of ER action (ligand binding, dimerization, translocation, and DNA binding) along with the mode of action of endocrine therapies: AIs block estrogen synthesis, SERMs inhibit transcription by competing with estrogen for ER binding, and SERDs promote the degradation of the ER.

These approaches have been shown to be clinically effective in women with ER-positive breast cancer. The standard duration of endocrine treatment with tamoxifen for pre- and perimenopausal women with low-risk tumors is currently five years and 10 years for high-risk tumors (T3-4 or lymph node-positive). For postmenopausal women with low-risk tumors it is five years of AI or two years AI and three years tamoxifen and for high-risk tumors previously treated with tamoxifen it is five years with AI (43).

SERMs - Tamoxifen

Tamoxifen is a selective estrogen receptor modulator (SERM) that has both antagonistic and agonistic effects on the estrogen receptor. First discovered in the late 1960s, it was progressively introduced for the treatment of breast cancer from the 1970s onwards (94). It has proven to be effective in reducing the risk of recurrence by 39% and continues to be the first-line treatment for primarily premenopausal women at low risk of recurrence in cases where ovarian suppression is either not required or contraindicated (62). For premenopausal women, tamoxifen induces anti-estrogenic effects on the breast similar to those resulting from surgical removal of the ovaries followed by estrogen withdrawal. The estrogen agonistic

characteristics of tamoxifen in premenopausal women are minimal. In postmenopausal women, tamoxifen acts as an anti-estrogen in the breast tissue, but has estrogen agonistic effects on the uterus, breast, vagina, bone, pituitary and liver (56,95,96). Clinical trials on adjuvant treatment have demonstrated that five years of therapy with tamoxifen reduces the recurrence of breast cancer and the incidence of contralateral second primary breast tumors by 50 % (97). The long-term use of tamoxifen has also been associated with an increased incidence of endometrial cancer in breast cancer patients (98–100).

The SERMs are able to bind to the intracellular ER and compete with estrogen for binding. At the beginning it is important to highlight that tamoxifen is extensively metabolized by the cytochrome P450 enzymes (CYPs) present in the liver and breast tissue (101). Two major CYPs, CYP3A4/5 and CYP2D6, can convert tamoxifen to N-desmethyl-tamoxifen, 4-hydroxy-tamoxifen (4-OHT) and endoxifen through demethylation and hydroxylation (102). However, endoxifen and 4-hydroxy-tamoxifen are reported to be responsible for the antitumor effect of the drug *in vivo* (103,104), but I will just focus on 4-OHT in the following explanation.

Mechanism: The ER-4OHT binding is triggered by the dissociation of heat shock protein 90 (HSP90), resulting in the formation of ER-4OHT complexes that can translocate into the cell nuclei (105). The active metabolite 4-OHT matches the hydrophobic pocket of the ligand-binding domain of the receptor to which estradiol usually binds. The anti-estrogenic side chain of 4-OHT blocks the reorientation of helix 12, required to lock the ligand in the receptor before coactivators can bind and form a transcriptional complex, interfering with the receptor complex (106).

Similar to the classical mechanism, the ER-4-OHT dimers in the nucleus bind to the ERE in the promoter region of estrogen-dependent genes. However, in contrast, the ER-4OHT complexes block the binding of ER coactivators and reduce or inhibit the transcriptional activation of estrogen-dependent genes in breast cancer (103,106). As a result of these nuclear effects, cell cycle arrest, inhibition of proliferation and apoptosis occur in tamoxifen-responsive breast cancer cells.

The permanent presence of corepressors in the complex explains the anti-estrogenic properties of tamoxifen. Indeed, the relative amounts of corepressor and coactivator in specific tissues and the presence of other factors control whether tamoxifen acts as an agonist or antagonist.

Aromatase inhibitors

The aromatase inhibitors are one of the main types of endocrine treatment and are prescribed especially to postmenopausal women. Although ER inhibitors have been used in the clinic since the early 1980s, aromatase inhibitors act differently by lowering plasma estrogen levels by inhibiting or inactivating aromatase (*CYP19A1*), which is the enzyme that synthesizes estrogens from androgenic precursors (107,108).

In other words, these drugs inhibit the enzyme aromatase, which synthesizes estrogen in the liver, muscles and adipose tissue. Since aromatase inhibitors can only sufficiently inhibit extra glandular estradiol production, they are ineffective in premenopausal patients without simultaneously ovarian suppression. Currently used aromatase inhibitors can be divided into two classes: irreversible steroidal inactivators such as exemestane and reversible non-steroidal inhibitors such as anastrozole and letrozole (109–111).

Mechanism: The aromatase enzyme is responsible for catalyzing the key step in the conversion of androgens to estrogens. The steroidal aromatase inhibitor exemestane and the two non-steroidal inhibitors anastrozole and letrozole are extremely potent and specific. The two classes of inhibitors reduce aromatase activity to 1% to 2.5% of baseline in postmenopausal women (112), significantly lower plasma estradiol levels, suppress tissue concentrations of this steroid in breast tumors and provide no estrogen agonistic properties. The benefit of aromatase inhibitors is limited to postmenopausal patients, since the interruption of the negative feedback of estradiol leads to a response-induced increase in luteinizing hormone (important for the processes in the reproductive system) and follicle-stimulating hormone (important for sexual development and fertility), resulting in the bypassing of aromatase blockade in premenopausal women (107,113).

SERDs – Fulvestrant

Fulvestrant, a steroidal 7α-alkylsulfinyl analogue of estradiol, is structurally distinct from the non-steroidal anti-estrogen tamoxifen and other selective SERMs. Fulvestrant has a significantly higher affinity for the ER compared to tamoxifen, with 89% versus 2.5% of the binding affinity of estradiol (114,115). This higher affinity leads to a more complete blockade of the ER (116). Clinical studies confirm that fulvestrant can reduce ER levels after both short-term (21 days) and long-term (6 months) treatments, although ER expression is not completely eliminated (117). Fulvestrant is considered a significant therapeutic approach to treat ER-positive breast cancer in both early stage and more advanced drug-resistant cases (118).

Mechanism: The mechanism of action of fulvestrant involves several steps. The binding of fulvestrant to the ER impairs receptor dimerization and energy-dependent nucleocytoplasmic transport, resulting in the blocking of nuclear localization of the receptor (119,120). At the same time, any fulvestrant-ER complex that enters the nucleus is transcriptionally inactive, as both the activation functions (AF1 and AF2) are deactivated. Finally, the fulvestrant-ER complex is unstable (121), leading to accelerated degradation of the ER protein by the ubiquitin-proteasome system in comparison to estradiol- or tamoxifen-bound ER (121–124).

Such downregulation of the cellular ER protein occurs without a reduction in ER mRNA. Therefore, fulvestrant binds to, blocks, and accelerates the degradation of

the ER protein, leading to a complete inhibition of estrogen signaling through the ER (122). Disruption of both AF1 and AF2 sites means that, unlike SERMs, which do not inhibit AF1 activity and therefore have partial estrogenic agonist activity, fulvestrant has no estrogenic agonist activity in animals or humans. Additionally, it is important to mention the difference between SERMs and SERDs in expression, while fulvestrant eliminated completely the estrogen regulated expression, in the presence of tamoxifen, some genes remained partially transcriptionally responsive to estrogen which highlight the different properties and mode of actions (125).

Targeted therapy

Targeted therapy is a way of treating cancer using specific molecules designed to identify and attack cancer cells while minimizing damage to normal cells. They target specific biomolecules involved in the growth, progression and proliferation of cancer cells, such as proteins, genes or the surrounding environment contributing to cancer growth and survival. In approximately 20-30% of cases of breast cancer, endocrine therapies are not applicable since these patients do not express the ER (126). In addition to the mentioned therapy options, a targeted therapy is available, which primarily benefits patients with HER2-positive breast cancer.

For HER2-positive tumors, anti-HER2 therapies such as monoclonal antibodies or antibody-drug conjugates are the treatment of choice. The first commercially available HER2-targeted drug was the humanized monoclonal antibody trastuzumab, an effective treatment for human breast cancer with antiproliferative effects on cells transformed by HER2 overexpression (127,128). The binding of trastuzumab to HER2 inhibits ligand-independent HER2 signaling and prevents proteolytic cleavage of the extracellular domain, an activation mechanism of HER2.

Along its remarkable clinical success, additional HER2-specific drugs have been developed and approved, for example pertuzumab which is an anti-HER monoclonal antibody that prevents dimerization of the HER2 receptor with other members of the HER family (129). In addition, more advanced molecule like antibody-drug conjugates such as trastuzumab-emtansine (T-DM1), which uses an active agent as ‘cargo’ and binds the tubulin inhibitor emtansine to trastuzumab or trastuzumab-deruxtecan (T-DXd) covalently linked to the topoisomerase I inhibitor (130). Other receptors of tyrosine kinase inhibitors are lapatinib, neratinib and tucatinib. Randomized clinical trial data has demonstrated that following adjuvant treatment with trastuzumab, the recurrence rate in HER2-positive patients decreased by approximately 50% (131,132).

Recently, it has been developed a three-tiered framework for classifying the HER2 status, covering the “HER2-low” status, defined by a positive IHC result with no associated gene amplification (130,133). More attention is being paid to this

classification since new evidence suggests that anti-HER2 therapies may offer potential benefits in advanced HER2-low breast tumors.

Treatment resistance in breast cancer

Even the best therapy carries the risk that the patient will develop resistance. Drug resistance describes the condition in which the treatment is initially or gradually becomes ineffective over time, shows no improvement or reaches a stagnation point. Several factors can influence resistance, including drug inactivation, changes in the drug target, DNA repair mechanisms, inhibition of cell death, cell heterogeneity, epigenetic effects, or a combination of these mechanisms.

Many studies are exploring the mechanisms of endocrine resistance, and several underlying mechanisms are known today. These include loss of ER expression, altered expression of microRNAs (microRNAs), ER interactions with signaling pathways, epigenetic changes and genomic aberrations.

While ER expression is the main indicator of the response to endocrine therapy, it does not guarantee treatment success. The first-line treatment for ER-positive breast cancer is tamoxifen. In more than 30% of cases of hormone therapy with tamoxifen, de novo (initial resistance) or acquired resistance may occur (134–136). Despite resistance, ER often remains expressed, leading to a second therapeutic option of fulvestrant (137,138). Fulvestrant is often used when resistance to tamoxifen develops, as it has a different mechanism of action (described in the chapter on endocrine therapies). Approximately 20% of patients respond to second-line therapy with aromatase inhibitors or fulvestrant, suggesting that patients who have acquired resistance to tamoxifen continue to express ER.

However, it is frequent that resistance to fulvestrant is also developed, often characterized by a complete loss of ER expression (139). This loss is often associated with hypermethylation of the promoter of the ER gene. Epigenetically modified ER signals can trigger new, often ligand-independent transcriptional programs without the activation of established ER-induced genes. Consequently, epigenetic regulation is considered an important regulator of the ER activation network and is clinically relevant for resistance in breast cancer.

An additional aspect in the development of resistance to endocrine therapy is the correlation between the loss of ER and the expression of the epidermal growth factor receptor (EGFR). Increased EGFR expression occurs in parallel with the loss of ER expression in breast cancer cells (140,141). This upregulation of EGFR can activate alternative signaling pathways promoting cell growth and survival, bypassing the need for ER signaling and promoting resistance to ER-targeted endocrine therapies.

Moreover, mutations promote drug resistance, although it is rare in the ER in primary tumors, occurring in about 1% of cases. In advanced ER-positive cases, however, these mutations occur more frequently, especially in patients treated with aromatase inhibitors. Up to 20% of these patients exhibit ER mutations (142,143). Mutations can further impact the efficacy of endocrine therapies by altering the functionality or expression of the receptor, leading to a reduced response to treatment. In such cases, the increased expression of EGFR is often a mechanism by which the cancer cells adapt and continue to proliferate despite therapeutic pressure from endocrine treatments. Understanding and targeting both ER and EGFR signaling pathways is therefore critical to overcoming resistance and improving treatment outcomes in ER-positive breast cancer.

In the case of HER2-positive breast cancer, however, the treatment with trastuzumab can also potentially lead to primary or acquired resistance. The resistance is significantly associated with hypermethylation of the promoter region of the tumor suppressor gene *Transforming Growth Factor Beta Induced (TGFBI)*. When *TGFBI* is hypermethylated, it can inhibit its expression, contributing to the development of resistance by enabling cancer cells to evade the effects of trastuzumab. Additional resistance mechanisms include mutations or amplifications in the *HER2* gene itself, activation of alternative signaling pathways, or changes in the tumor microenvironment that reduce the efficacy of trastuzumab. Gaining an insight into these resistance mechanisms is important for the development of new strategies to overcome therapy resistance and improve treatment outcomes.

Future perspective of breast cancer treatment

Since up to 30% of ER-positive breast cancer have an intrinsic potential to develop hormone resistance at the time of diagnosis, it is extremely important to understand this potentially non-responsive phenotype, which is complex and often unclear.

Recently, the addition of cyclin-dependent kinase 4/6 inhibitors (CDK4/6i) to endocrine treatment has emerged as treatment option (43). The combination has improved the prognosis of patients with advanced luminal breast cancer compared to endocrine therapy alone (144). But also other combinations with inhibitors of other signaling pathways (e.g. PI3K/AKT/mTOR pathway) have shown promising results in preclinical studies and early clinical trials (43).

In development are new-generation (ng) SERDs in the form of oral pills instead of injections (121). The ng-SERDs can significantly reduce the expression of ER protein and block both estrogen-dependent and -independent ER signaling compared to fulvestrant which is less potent in degrading and blocking in signaling. Therefore, ng-SERDs are considered a promising approach therapeutic approach for the treatment of ER-positive breast cancer in both early-stage and advanced, drug-

resistant cases since they have better efficacy, with the potential to overcome resistance (93,121). An additional innovative approach involves PROTACs (PROteolysis TArgeting Chimeras). Such molecules can stimulate ER degradation by targeting the cell's protein degradation mechanism, which represents a novel approach for the treatment of ER-positive breast cancer (145).

The future of cancer treatment is personalized medicine, focusing also on the development of new biomarkers to identify indicators that predict response to treatment or treatment combinations. At the same time, the monitoring of resistance is important, for example by monitoring circulating tumor DNA (ctDNA), which could lead to the early detection of processes like disease progression.

What is a receptor and why are they important?

A receptor is a protein that is commonly found on the surface of cells (cell membrane) or inside cells (intracellular) receiving and responding to chemical signals from both outside and inside the cell. If these signals, frequently molecules known as ligands, such as hormones, neurotransmitters or growth factors, bind to the receptor, it triggers a series of cellular reactions (1).

Receptors can be broadly categorized into cell surface receptors, such as the receptor tyrosine kinases (e.g. HER2), and intracellular receptors, such as the nuclear receptors (e.g. ER and PR). They fulfil various essential functions. For example, they enable signal transduction, involving communication between the cell and its environment by transforming extracellular signals into intracellular signals (1,146). They also regulate cellular activities, including gene expression through acting as transcription factors switching genes on or off (147,148), and regulate various metabolic pathways by regulating enzyme activity or cell metabolism (149). They also play a role in cell growth and differentiation and in the immune response (1,147,150,151).

There are several key explanations why cells express receptors. Most importantly, receptors enable cells to adapt to their environment by recognizing and responding to conditions like the presence of nutrients, toxins or other cells (149). They are important for maintaining homeostasis by reacting to hormonal and other signals to maintain internal cellular and systemic balance by provide coordinated functions in tissues and organs.

Beyond this, they help to identify stress signals and initiate protective responses, including the activation of repair mechanisms or inducing apoptosis in cases where damage is unrepairable (152). Finally, specific receptors play an important role in the specific functions of different cell types (153,154). Receptors are important for cells since they enable them to recognize and respond to their environment, adjust

internal processes, interact with other cells and maintain overall function and survival. Previously, in the chapter describing biomarkers in breast cancer, I mentioned the estrogen receptor, the progesterone receptor and HER2. In the following section I explore in more detail these receptors' structure and function.

Estrogen receptor alpha

Between 70% and 80% of breast cancers express estrogen receptors, primarily ER alpha (ER), which is encoded by the *ESR1* gene. Additionally, there are other estrogen receptors such as ER beta (155), encoded by the *ESR2* gene, and the G-protein coupled estrogen receptor 1 GPER1 (156), but these are not the focus of this discussion.

ER beta, while less studied than ER alpha, plays a significant role in modulating the effects of ER alpha, often exerting opposing actions on gene expression and cell proliferation. ER beta is involved in inhibiting cellular proliferation and has been associated with tumor suppressor functions, making it a potential therapeutic target for certain subtypes of breast cancer (155). The GPER1, also known as GPR30, is distinct from the classical nuclear estrogen receptors and mediates rapid, non-genomic signaling pathways in response to estrogen. GPER1 is implicated in various cellular processes, including the regulation of cell proliferation, apoptosis, and migration, and has been linked to the progression of hormone-dependent cancers. The roles of ER beta and GPER1 are increasingly recognized as critical in understanding the diverse and complex effects of estrogen signaling in both normal physiology and cancer.

ER alpha is considered a critical driver of breast cancer progression, particularly in luminal-type breast cancers, and serves as a primary target for endocrine therapies aimed at blocking estrogen signaling, which is essential for tumor growth. This receptor type is the focus of the present work and will be referred to as the ER.

Mechanism of ER activation and function

The ER is a member of the nuclear hormone receptor family that regulates transcription in a hormone-dependent manner through sequence-specific DNA binding. The ER is stimulated by the binding of its natural ligand, estradiol, to the ligand-binding domain of the receptor. This binding leads to dissociation from the chaperone protein Hsp90 and facilitates ER-chromatin interactions, enabling transcription.

ER functions through both genomic and non-genomic mechanisms. The genomic pathway involves direct binding to DNA at EREs, while the non-genomic pathway involves rapid signal transduction through cytoplasmic ERs interacting with

membrane-bound receptors and kinases, contributing to cell proliferation and survival.

When estradiol is not present, a fraction of ER resides in the nucleus, loosely bound as dimers to specific DNA sequences (consensus sequence GGTCAnnnTGACC), called EREs, which regulate estrogen-responsive genes (genomic pathway). The binding of estradiol to the ER promotes the translocation of the estradiol-ER complex to the nucleus and increases its binding to the EREs, regulating gene expression.

This regulation is not limited to direct DNA interactions (non-genomic pathway). ER also modulates gene expression through interaction with other transcription factors like AP-1 and impacting a broader range of cellular functions (157,158). The transcriptional control by ER involves two activating functions, AF1 and AF2, recruiting co-activator and co-repressor proteins to the transcription complex and resulting in the expression of target genes. Furthermore, the ER activity is also regulated by phosphorylation through several pathways, among them the epidermal growth factor receptor family, the insulin-like growth factor receptor and the chemokine receptor pathways (159,160). Both canonical and non-canonical ER signaling mechanisms are complementary and synergistic contributing to the regulation of estrogen signaling and the cellular reaction to hormonal therapies.

Structure and functionality of ER

The ER is a modular protein with six functional domains, A to F (Figure 5). The A/B domain is the amino-terminal domain encoding a hormone-independent transcriptional activation function 1 (AF1), responsible for protein-protein interactions and transcriptional activation of target gene expression. The domain C is a highly conserved DNA-binding domain (DBD) consisting of two functionally distinct zinc finger motifs responsible for the specific binding of receptors to the ERE or ERE-like sequences in the promoter of target genes (161). This is followed by region D, the hinge region separating the DBD and the ligand-binding domain. One hypothesis is that the flexibility of the secondary structure of this region allows conformational changes of the receptor molecule during activation and plays a role in dimerization. The region E/F encodes the ligand-binding domain (LBD), located in the carboxy-terminal part of the receptors (162).

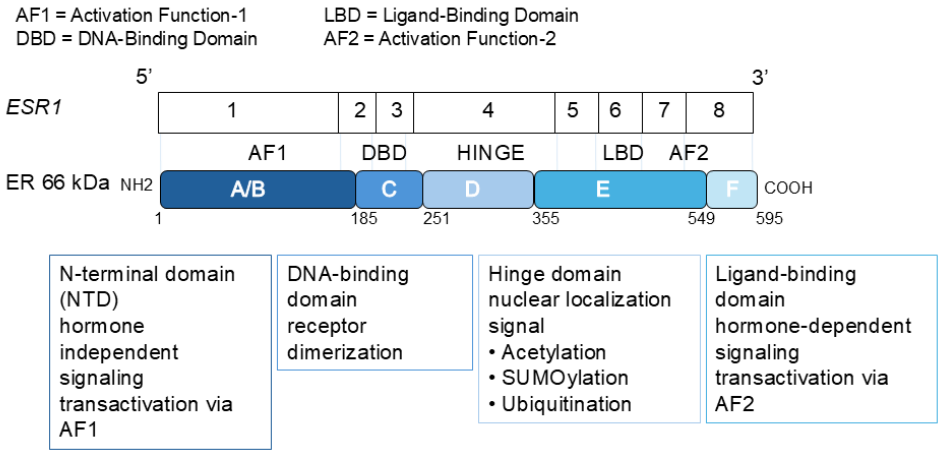

Figure 5 Estrogen receptor protein domains (full-length) encoded by the *ESR1* gene with a short explanation of each domain (modified from (163)).

An essential feature of ER structure is its ligand-binding capability, where the ligand is locked into a hydrophobic pocket that is closed by helix 12. The position of helix 12 above this pocket is necessary for the recruitment of co-activators to the activation function 2 (AF2) site and the initiation of RNA polymerase activity. Repositioning of helix 12 after ligand binding is an important mechanism for achieving the full estrogenic effect of ER (164,165). The structural integrity and functionality of these domains are vital for the precise regulation of gene expression and the subsequent physiological effects mediated by ER.

Role of ER in breast cancer

ER drives the proliferation and growth of luminal-type breast cancer and is the target of endocrine therapies for this disease. Although the ER is primarily expressed in the mammary glands and female reproductive tract (including ovaries, fallopian tubes, uterus, and cervix), it is also present in various other tissues such as the liver, muscle, adipose tissue, and the pituitary gland. The functions mediated by ER include fertility regulation, cancer progression, osteoporosis, and endometriosis. In breast cancer, ER's role is multifaceted, involving the activation of pathways that promote cell proliferation and survival. ER is involved in the up- and downregulation of a wide range of genes. It stimulates cell cycle progression through the transcriptional upregulation of key genes such as *CCND1*, *CCNA1* or *CCNA2*, and *MYC*. Additionally, ER targets transcription factors and cofactors such as *PR* or *GREB1*, as well as genes related to growth hormone signaling, including *VEGFA*, *IGF1*, and *EGFR*. The dysregulation of these pathways by ER contributes

to the oncogenic processes in breast cancer, making it a pivotal target for therapeutic intervention.

Progesterone receptor

The progesterone receptor (PR) is a multidomain protein which is a highly regulated target gene of the ER, exhibiting estrogen-dependent expression and can modulate the action of the ER (166). The PR represents one of the best characterized ER target genes. Because it is often expressed together with ER in breast cancers, PR also serves as a useful prognostic biomarker. Therefore, PR testing is often combined with ER testing to evaluate the hormone receptor status of a breast tumor. However, PR expression status is not a robust predictor of response to endocrine therapy, indicating that PR expression is not exclusively controlled by ER activity (167). Research studies show that 70% of ER-positive/PR-positive tumors respond effectively to tamoxifen, while only 34% of ER-positive/PR-negative tumors respond to tamoxifen therapy (168). One potential explanation for this result could be that PR-positive tumors may retain a functional ER signaling pathway. Especially since PR expression is an estrogen-regulated event, ER might be present in these tumors at a level beneath the detection limit for ligand binding or immunohistochemical assays.

The PR shares conserved functional domains with other members of the nuclear receptor family of transcription factors. Such domains include an N-terminal region modified extensively post-translationally and containing transcriptional activation functions, a central DNA-binding domain with two cysteine-linked zinc fingers and a C-terminal ligand-binding domain (169–171).

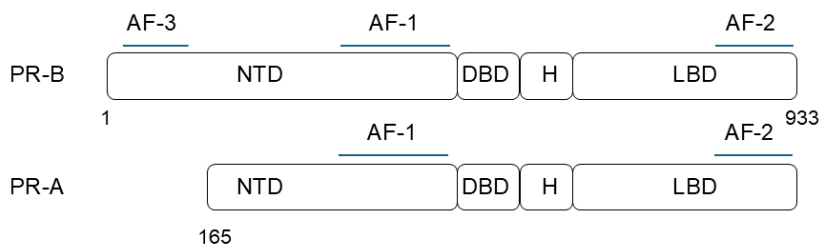

Figure 6 Progesterone receptor PRA and PRB protein domains, PRB is a full-length isoform, PRA lacks 164 amino acids at the N-terminal domain (NTD), ligand-binding domain (LBD), hinge region (H), DNA-binding domain (DBD), and activation function domains (AFs) (Modified from (172)).

In humans, there are two major isoforms of PR expressed from a single gene on chromosome 11q22.1: the full-length, 933 amino acid PR-B and a truncated, 769 amino acid PR-A which is transcribed from an internal start site (Figure 6) (173).

The PR gene is stimulated by estrogens, and both PR-A and PR-B are expressed in approximately one third of the luminal epithelial cells of the normal breast. There is also evidence of PR expression in basal-like epithelial cells (174).

Human Epidermal Growth Factor Receptor 2

The receptor HER2 (Human Epidermal Growth Factor Receptor 2) encoded by the gene *ERBB2* (Erythroblastic Oncogene B), is overexpressed in about 15-20% of all breast cancers and is associated with more aggressive tumors and poor prognosis (175,176).

The HER2 belongs to the human epidermal growth factor receptor (HER or ErbB) family, which consists of four members including HER1 (EGFR, ERBB1), HER2 (Neu, ERBB2), HER3 (ERBB3) and HER4 (ERBB4) (177,178). The ligand-induced homo- or heterodimerization of HER proteins induces a downstream phosphorylation signaling cascade that stimulates cell growth, proliferation and differentiation. The four HER proteins are strongly associated with tumorigenesis, with EGFR and HER2 considered to be the most potent oncoproteins. Indeed, their overexpression is associated with many types of cancer, including breast and lung cancers. Within the HER family, HER2 is a unique protein since it has no known ligands and is unable to form ligand-dependent homodimers. Moreover, HER2 is unusual among ErbB receptors in that it is able to transform cells in a ligand-independent manner when overexpressed (177,178).

In order to promote downstream signaling, HER2 must either form heterodimers with other HER proteins once they have bound specific ligands or combine into homodimers in a ligand-independent way under conditions of overexpression. The most relevant combinations of HER2-containing heterodimers are EGFR/HER2 and HER2/HER3 based on their impact on cellular function and disease (178,179). While the mechanisms of HER family homodimerization in terms of extracellular ligand binding and intracellular kinase domains are well characterized, the molecular mechanisms of heterodimerization of HER2 with other family members remain relatively unclear.

Regulation mechanisms

The regulation of cellular mechanisms is essential for normal cell growth, differentiation and adaptation to environmental influences. This includes all processes that affect the frequency, relative proportion or level of cellular activity. Although these processes can occur at the cellular level, they are not restricted to a single cell. Accurate regulation is fundamental to the survival and viability of the cell, regardless of whether it is normal or malignant. In the field of proteins, several

regulatory mechanisms have been researched over the years. Among these are splicing such as alternative splicing, epigenetics such as methylation and microRNAs. Those mechanisms help to contribute to maintaining the correct function and balance in the cell and play a key role in cell physiology.

For cancer cells, these mechanisms are mostly disruptive structures using abnormal splicing patterns to promote cancer growth and create a supportive environment. This also includes changes in epigenetic regulation to activate genes for cell proliferation or to deactivate tumor suppressors. Moreover, cancer cells can have a specific microRNA profile that further promotes gene expression and tumor progression.

Splicing and alternative splicing

Splicing describes a process that removes the non-coding sequences (introns) from the pre-mRNA and connects the coding sequences (exons). It is a fundamental part of the eukaryotic genome and is required to generate a functional mature mRNA that contains only exons which is suitable to translate it into a protein (1).

Alternative splicing is a specialized mechanism of splicing that combines the exons and introns of a pre-mRNA in distinct ways to produce various mRNA transcripts (180). It enables a single gene to produce several unique proteins by combining exons, introns and alternative splice sites in different ways. This leads to different mRNA transcripts, also called isoforms, that originate from the same gene locus but differ then in the transcription start sites, protein coding sequence, 3' untranslated region (UTR) and/or 5' UTR (181,182).

The different mRNA transcripts produce different protein variants with potentially different functions that enhance the genetic diversity and the ability of an organism to respond to different needs and environmental influences. Essentially, it allows cells to produce specific proteins in different tissues or developmental stages to regulate the amount and type of proteins produced. Different protein isoforms may contain or lack different functional domains, which also affects protein function (181).

Mechanism

Splicing is a complex, multi-step process carried out by a large complex of proteins and RNA molecules, the spliceosome. First, the spliceosome recognizes specific sequences at the boundaries of exons and introns. These include the 5' splice site (donor site), the 3' splice site (acceptor site) and the branching site within the intron. This recognition is essential for the correct splicing process. The next step is the assembly of the spliceosome. It consists of several small nuclear ribonucleoproteins (snRNPs) and other associated proteins, whereby the snRNPs U1, U2, U4, U5 and U6 play a central role (1,183,184).

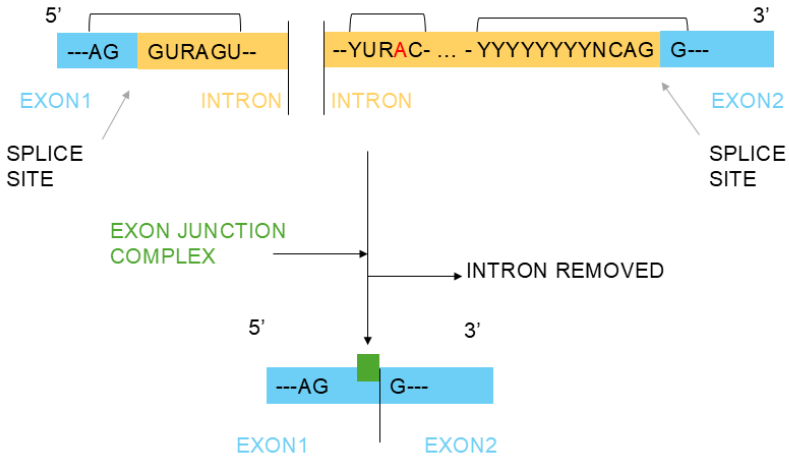

Figure 7 Simplified representation of the splicing process and its key sequences for intron removal, recognized by snRNPs. The image highlights the 5' splice site, containing the GU sequence, and the 3' splice site, containing the AG sequence, as well as the branch point sequence (YURAC). The exon junction complex (EJC) is formed after the intron is removed, joining the exons together (modified from (1)).

Then, the spliceosome catalyzes the cutting of the RNA at the 5'- and 3'-splice site. Two simultaneous interactions are involved in this process: First, the 2'-OH of the branch site attacks the 5'-splice site and forms a loop from the intron, the so-called lariat. Afterwards, the free 3'-OH of the exon at the 5'-end attacks the 3'-splice site, whereby the two exons are joined together and the lariat is separated. Once the splicing reaction is complete, the lariat intron is degraded. Now the mature mRNA, which consists only of exons, is ready for export from the cell nucleus and subsequent translation (Figure 7) (1,183,184).

The decision to use alternative splicing is influenced by several factors and mechanisms. Splicing regulators such as serine/arginine-rich (SR) proteins and heterogeneous nuclear ribonucleoproteins (hnRNPs) play a central role (183). These regulatory proteins bind to specific RNA sequences and influence whether and how the spliceosome is recruited to specific splice sites. Furthermore, cis-regulatory elements on the pre-mRNA play an important role. While exonic splicing enhancers (ESEs) and intronic splicing enhancers (ISEs) promote spliceosome recruitment to splice sites, exonic splicing silencers (ESSs) and intronic splicing silencers (ISSs) can inhibit this process (185). Also the chromatin structure and the speed of transcription influence the splicing pattern. Slow transcription gives the spliceosome more time to recognize and process weaker splice sites. Changes in the environment and external signals such as stress, hormones or growth factors can also activate signaling pathways that influence the activity of splicing factors and thus modulate alternative splicing. Alternative splicing occurs in different patterns

(Figure 8). The most frequent type is exon skipping, in which certain exons in the pre-mRNA are skipped to generate different mRNA transcripts and thus different protein variants.

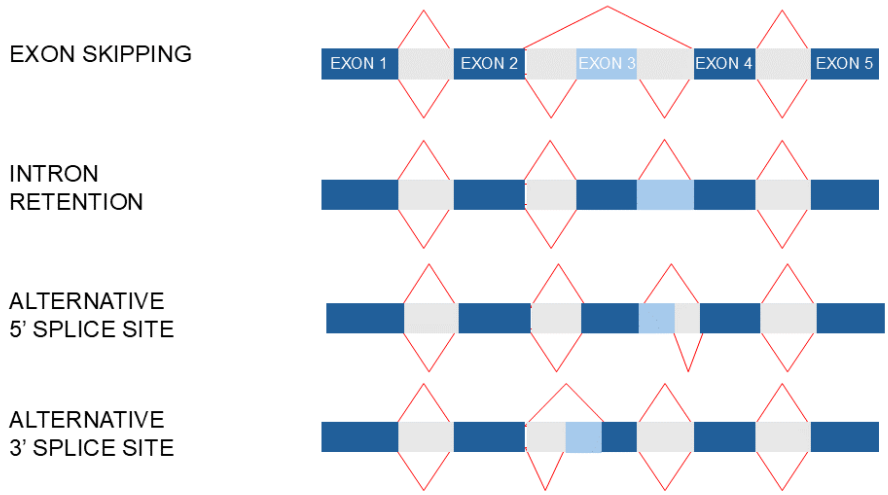

Figure 8 Major types of alternative splicing include exon skipping, intron retention, and alternative 5′ or 3′ splice sites. Dark blue boxes represent exon sequences, light blue boxes indicate possible exons included in only one mRNA variant, and red lines show where introns (grey) are removed (modified from (1)).

Alternative splicing of the estrogen receptor alpha - ESR1

Since the beginning of the 1990s, several splice variants of the estrogen receptor have been identified and multiple alternative protein isoforms have been characterized in structure and function (186). Among these are the common isoforms of 46 kDa and 36 kDa (187–190) in size as a well as several isoforms comprising cryptic alternative exons (Figure 9). The 46 kDa isoform is an example of exon skipping which lacks the N-terminal (173 amino acids) because it arises from a downstream promoter and results from splicing exon 1 directly to exon 2, skipping the region that codes for the N-terminal domain (190). Functional analyses showed that the 46 kDa isoform is an effective ligand-inducible transcription factor in a cell context responsive to the transactivation function AF-2. On the other hand, in a cell context in which the transactivation function of AF-1 dominates, the 46 kDa isoform is a strong inhibitor of the full-length, 66 kDa isoform (190). The identification of a 36 kDa isoform was first reported in 2006 (191). For the 36 kDa isoform, an alternative start codon is used that leads to translation of a truncated protein and additionally several exons are skipped, including those coding for the transcriptional activation domains AF-1 and AF-2 but retains the DNA binding domain. It translocates primarily between the cytoplasm and the cell membrane

(192) and it has been shown that the 36 kDa isoform can interact with the 66 kDa isoform and inhibit gene expression, making it a negative modulator of the full-length protein (187,193,194). In this context, the regulation of the 36 and 66 kDa isoforms differs because they have different promoters.

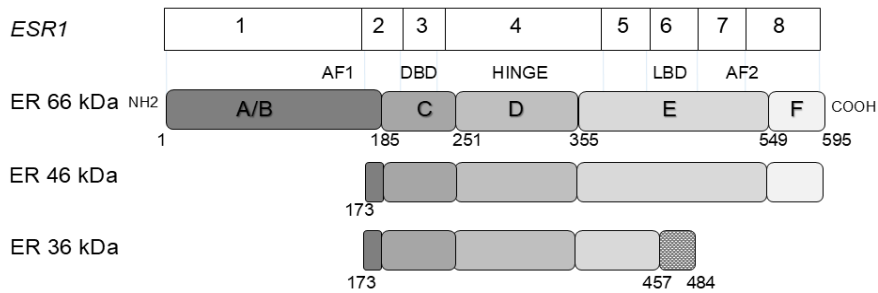

Figure 9 ER protein encoded by the *ESR1* gene, showing the full-length ER (66 kDa) and two protein isoforms (46 kDa and 36 kDa). They by the absence of certain domains, indicating variations in their structure and potentially their function (modified from (195)).

While these isoforms originate from the same *ESR1* gene, each of them has its own amino acid sequence and distinct biological function as a result of different promoters and alternative splice sites. The majority of these naturally occurring variants are mRNA splice variants and are characterized by the deletion of one or more exons in the ER mRNA. Apart from variants without exon 3 or 4, in most ER splice variants, the translation downstream of the splice site happens out of frame and leads to a truncated protein (196,197).

The ERΔ7 mRNA has been reported as the main alternative splice form in most human breast tumors and cancer cell lines (198–200). ERΔ7 is of specific interest as the LBD, AF-2 and DBD domain are found partially located in exon 7 and works as dominant negative isoform which inhibit the 66 kDa the full-length protein (198,200,201).

However, splice variants represent a challenge when it comes to ER diagnosis in the clinic in terms of IHC, since they are a potential source of false-positive cases. Since the antibodies used for ER in IHC often recognize epitopes encoded by the first exon of the ER gene (202), these splice variants may be recognized as ER positive, despite potentially differing in function from the normal ER protein. The alternative ER proteins can exhibit functional activities that are either negative, dominant negative, or dominant active on ER target genes. The dominant negative variants are not only inactive themselves, but they also inhibit the full-length ER through heterodimerization (198). The expression of the dominant-negative ER variant ERΔ7 may help to clarify why tumors show the basal molecule type while staining immunohistochemically as ER-positive. Potentially, these tumors may lack a

functional response to estrogen and consequently may not respond to hormonal therapies.

Epigenetics - DNA methylation

Epigenetics describes a heritable molecular mechanism with modifications in gene expression or function not caused by changes in the DNA sequence itself (1,203). It is an important regulatory mechanism as it controls gene expression by adjusting the accessibility of DNA to the transcription machinery. By adding or removing chemical groups to the DNA or histones, it is possible for the cell to determine which genes should be activated or deactivated. Such epigenetic modifications are necessary for adapting gene activity to different requirements and are heritable during cell division. However, epigenetic mechanisms are also a consequence of environmental factors, since they can be affected by factors such as nutrition, toxins and stress. It also enables more differentiated cells to express their specific genes and repress the genes of other cell types, which is important for the maintenance of tissue and organ function.

Key epigenetic regulatory mechanisms involve DNA methylation, histone modifications and RNA-mediated gene targeting (203,204). DNA methylation describes the addition of a methyl group to cytosine bases of DNA, resulting in the long-term silencing of genes (Figure 10) (1,205,206). By contrast, chemical modifications to histone proteins affect the tightness of the DNA wrapped around the histones, which determines the accessibility of the DNA to the transcription machinery and enables a more dynamic adjustment of gene activity, allowing it to be modified if necessary (207,208).

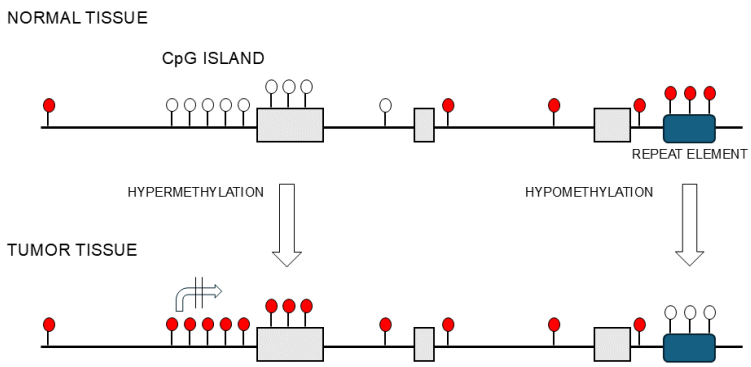

Figure 10 DNA methylation in normal tissue vs. tumor tissue (red = methylated, white = unmethylated). In normal tissue, CpG islands are usually unmethylated, allowing gene expression, whereas in tumor tissue, hypermethylation can silence genes (modified from (209)).

DNA methylation

The DNA methylation represents the process of adding methyl groups to the 5' position of the cytosine ring in a CpG dinucleotide by enzymes called DNA methyltransferases (DNMTs) (205,210). There are three catalytically active DNA methyltransferases: DNMT1, DNMT3A and DNMT3B (205,206,210), that do not alter the DNA sequence. While DNMT1 maintains previous DNA methylation patterns after DNA replication, DNMT3A and DNMT3B are involved in de novo methylation, which is stimulated by DNMT3L (205,211–214).

The CpG dinucleotides, consisting of a cytosine and a guanine, are commonly located at higher frequencies in the promoter regions (215,216), leading to CpG islands or large repetitive sequence regions such as centromeres and retrotransposon elements (217–219). They are defined as regions with a GC content of more than 55% and a ratio of observed to expected sequences of greater than 0.65 with a minimum size of 500 base pairs (220). Almost 70% of CpG islands are known to be linked to human genes, most of them are found in the promoter regions (221).

DNA methylation has been reported to be dysregulated in many malignant diseases (222,223) and inhibits transcription by either directly blocking transcription factor binding sites or by recruiting methyl-CpG binding proteins (206,224). The mentioned molecular subtypes of breast cancer are linked to particular DNA methylation patterns. It was found that basal-like, luminal A and luminal B tumors had distinct methylation profiles, whereas tumors of the normal-like and HER2-enriched molecular subtypes lacked distinct methylation profiles (225,226). If a CpG island is methylated its DNA becomes inaccessible to the transcription machinery, resulting in induced gene silencing. Within normal cells, CpG islands remain mostly hypomethylated, while the whole genome is methylated including repetitive sequences to prevent chromosome instability (227). In contrast, cancer cells often show hypermethylation of CpG islands, which is associated with global genome hypomethylation. This inhibits tumor suppressor genes and contributes to genomic instability (227).

Comparing methylation analysis of tissue samples from healthy controls, DCIS and invasive breast cancer, a striking change in methylation profiles was observed from one cellular disease progression to the other (228,229). Most of the methylation changes were observed during transition from healthy breast tissue to DCIS, supporting the hypothesis that methylation changes play an early role in the development of breast cancer and represent a reasonable target to improve early detection. In fact, abnormal methylation occurs in a variety of cancers including thyroid cancer, lung cancer, liver cancer, stomach cancer, prostate cancer, bladder cancer, ovarian cancer, colorectal cancer and breast cancer (230). This highlights the importance of epigenetics in cancer research, since nearly all cancers are linked to abnormal DNA methylation.

Regulation by microRNAs

MicroRNAs (miRNAs) are small, non-coding single-stranded RNAs with a length of between 18 and 25 nucleotides (231). They regulate gene expression by blocking mRNA translation or by inducing mRNA transcripts to degrade by binding to the 3'UTR within the target mRNAs in the cytoplasm. Their function is to regulate the expression of a broad range of genes directly involved in important cellular processes including growth, homeostasis, differentiation and apoptosis (232–234). Through perfect or incomplete base-pairing within the 6 to 8 nucleotide seed sequence each miRNA is capable of targeting many mRNAs (235,236). Indeed, each miRNA can modulate the expression of hundreds of genes at the same time, with around 30% of human proteins believed to be regulated by miRNAs (237). MicroRNAs can function as oncogenes (oncomiRs) or tumor suppressors (238), could provide cancer biomarkers and are also being tested as cancer therapeutics in clinical trials (235).

Biogenesis of miRNAs

There are three main steps in the biogenesis of miRNAs. Illustrated in Figure 11 is the canonical pathway of miRNA biogenesis. There are also additional alternative pathways leading to miRNA generation, but these will not be discussed further.

The initiation of the process involves transcription by RNA polymerase II (239) or III (240), which generates primary miRNA transcripts (pri-miRNAs) of approximately 3000 to 5000 bases in length . Many miRNAs are encoded within host genes, often located in introns or exons. The pri-miRNA is processed by the microprocessor complex, including DROSHA and the DiGeorge syndrome critical region gene 8 (DGCR8), to form short hairpin precursor miRNAs (pre-miRNAs) of approximately 70 bases in length. Together with Ran-GTP, the resulting pre-miRNAs are exported from the nucleus to the cytoplasm via exportin 5 (XPO5). Once exported, they are processed by Dicer (DICER1) assisted by the TAR RNA binding protein (TRBP) and transformed into mature miRNAs of 21-25 nucleotides in length. After separation into single-stranded RNAs and binding to Argonaute proteins (AGO1-4), miRNAs form the RNA-induced silencing complexes (RISC). When incorporated into the RISC complex, the miRNA hybridizes with the 3’UTR of a target mRNA. As a result, it blocks protein synthesis post-transcriptionally by inhibiting translation and/or degrading the mRNA (241).

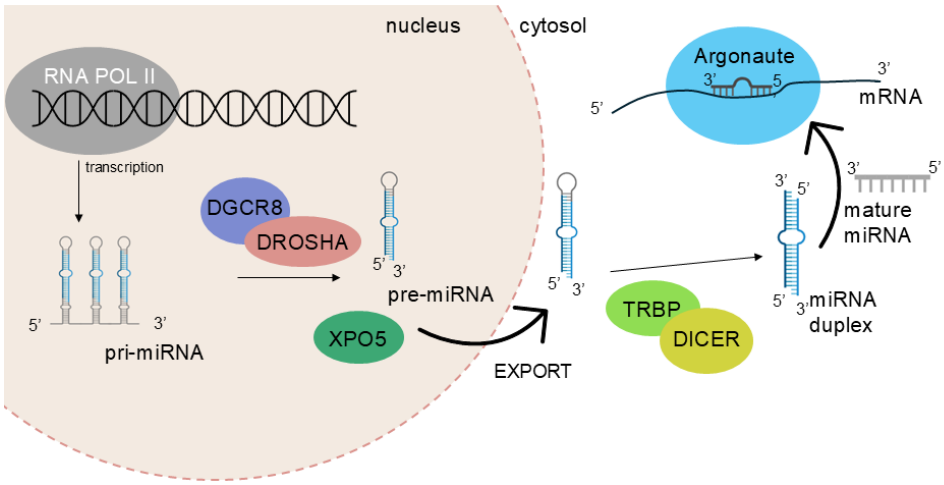

Figure 11 The canonical miRNA biogenesis pathway involves pri-miRNA being processed by the microprocessor complex into pre-miRNA, which is exported to the cytosol and further processed by Dicer into a mature miRNA strand, then incorporated into Argonaute (AGO) protein complexes (modified from (242)).

MicroRNA functions and their contributions to cancer development

The interaction of miRNAs with their target genes is dynamic depending on many factors, including the subcellular localization of miRNAs, frequency of miRNAs and target mRNAs, as well as the affinity of miRNA-mRNA binding (243).

A critical component of this interaction is the seed region, a sequence typically spanning nucleotides 2 to 7 at the 5' end of the mature miRNA, which plays a pivotal role in target recognition. This region determines which mRNAs a miRNA will target, but perfect complementarity between the seed and the target mRNA is not necessary for effective gene silencing. A canonical miRNA target site can be established with as few as 6 nucleotides, and the efficiency of silencing generally improves with increased base pairing beyond the seed region (244). Given the compactness of the seed sequence and the vast diversity of miRNAs in humans, most protein-coding genes contain conserved miRNA binding sites within their 3' UTRs. This allows for a single mRNA transcript to be regulated by multiple different miRNAs, while a single miRNA can target the 3' UTRs of various mRNAs (245,246). This interaction results in a complex, interconnected network of post-transcriptional gene regulation.

The mature miRNA that is incorporated into an AGO-protein complex is able to bind to mRNAs and can activate a series of regulatory activities. In the event of perfect or near-perfect complementarity occurring between the miRNA and the mRNA, AGO2 is able to catalyze the slicing of the mRNA, resulting in the induction of its degradation (247). However, it is rare in most animals, compared to plants

(245). An imperfect base pairing between the miRNA and the mRNA occurs much more frequently.

Abnormal regulation of miRNAs during tumorigenesis could lead to a role for miRNAs in cancer. In fact, miRNAs influence several stages of breast cancer, including tumor development, metastasis and treatment escape. In the last 15 years, research has highlighted the role of miRNA regulation in contributing to the risk or prevention of ER-positive breast cancer. More recently, studies have highlighted that miRNAs function not only as targets of ER/hormonal signaling, but the ER is also a regulatory target of multiple miRNAs (248). As an example, miRNA microarray analyses revealed that estradiol can upregulate the expression of a large set of individual miRNAs and miRNA families (249). Importantly, a unique miRNA expression profile could be potentially associated with certain subtypes.

The most prominent oncomiR is miR-21 (250,251), which is commonly overexpressed in various human tumors and cancer cell lines, including glioblastoma, ovarian carcinoma, cervical carcinoma and lung cancer (250,252). It is overexpressed in TNBC breast cancer and correlates with poorer disease-free survival and overall patient survival (253). Also functional studies with cancer cell lines have demonstrated that miR-21 plays a role in the oncogenesis process (254). Underlining the importance of miR-21 is the strong association with increased cell proliferation, low apoptosis, enhanced invasion and increased metastatic potential (255–257).

MicroRNAs significantly influence the development, progression, and therapy resistance of cancers, including breast cancer. Aberrant expression of specific miRNAs is linked to oncogenesis, often resulting from methylated CpG sequences in promoter and enhancer regions that regulate miRNA expression. This epigenetic modification indirectly affects the expression of cancer-related genes and proteins (258,259).

In breast cancer, certain miRNAs modulate the estrogen signaling pathway and ER-dependent gene expression, contributing to resistance against anti-estrogen therapies (249,260). Notably, hypomethylation of ER-targeted miRNAs correlates with dysfunctional ER activity (261). Altered miRNA expression levels are pivotal regulators of oncogenic signaling pathways, impacting therapeutic outcomes. For instance, variations in specific miRNAs are associated with tamoxifen resistance and can predict patient prognosis and response to treatment (262,263). Given the roles of estrogen and miRNAs in tumor dynamics, exploring how ERs influence non-coding miRNAs offers valuable insights into mechanisms underpinning endocrine treatment resistance (260).

mir-4728

The discovery and understanding of the biological functions of miRNAs has wide-ranging relevance for cancer research, which includes the relatively unexplored miRNA *mir-4728*. This miRNA is encoded in an intron of the *ERBB2* oncogene and was discovered in 2011(264). It is encoded directly upstream of exon 24 and is an example of a "mirtron", a type of miRNA created by a special splicing process.

The main mature miRNA product is miR-4728-3p, it is found at much higher levels than its 5p counterpart, which emphasizes its primary role in regulation. Moreover, the observation that *mir-4728* is encoded in an *ERBB2* intron implies that this locus not only produces the receptor protein, but also generates a regulating miRNA that may contribute to the development of cancer. Research reveals that miR-4728-3p is associated with downregulation of the poly(A) polymerase TENT4B, which affects the degradation of miR-21 (265). The tumor suppressor PTEN is a target of miR-21-5p potentially influencing the effectiveness of anti-ERBB2 therapies (266).

At the same time, miR-4728-3p also regulates ER, which highlights its role in the regulation of hormone signaling (267–269). This suggests the potential role of miR-4728-3p in the mechanism of ERBB2- and ER-mediated signaling. The characterization of miR-4728-3p could provide an important insight into the development of resistance to anti-ERBB2 therapies. The interaction of miR-4728-3p with the regulation of miR-21 and ER suggests that it may be involved in developing resistance to treatments like trastuzumab (266).

Targeting miR-4728-3p could potentially enhance treatment effectiveness and help address existing resistance to therapy. However, to be noted is that miR-4728-3p may also have tumor-suppressive effects, as demonstrated in studies of colorectal cancer and other tumor types (270,271). Together, these two conflicting findings highlight the complexity and context-dependence of microRNA functions.

Future perspective from regulatory mechanisms to the clinic

Advancements in the understanding of regulatory mechanisms such as DNA methylation, alternative splicing, and miRNAs hold significant promise for the future of breast cancer treatment, especially in overcoming therapeutic resistance. These mechanisms play important roles in regulating gene expression and cellular behavior, which are central to cancer development and treatment response.

Recent research has underscored the potential of Differentially Methylated Regions (DMRs) as biomarkers for breast cancer. These regions, which exhibit distinct methylation patterns in cancerous versus healthy tissues, could lead to more precise diagnostic tools. However, the variability across studies, with minimal overlap in

identified DMRs (272,273), highlights the need for more standardized approaches to identify clinically relevant methylation biomarkers. Moreover, while epigenetic drugs like 5-aza-2′-deoxycytidine and histone deacetylase inhibitors like trichostatin A have shown potential in reactivating silenced genes such as the ER in ER-negative breast cancer (274), their clinical success is highly dependent on the specific context and levels of gene expression (275). This suggests that further research is essential to optimize these therapies and to understand better how epigenetic modifications can be manipulated for therapeutic benefit.

Alternative splicing affects the production of different protein isoforms, which can significantly impact therapy effectiveness. In breast cancer, splicing variants of key receptors like ER, PR, and HER2 may contribute to resistance to targeted therapies. Identifying and characterizing these variants are critical areas for future research, as they could lead to the development of more personalized treatment strategies and improve therapeutic outcomes.

MicroRNAs are key players in the regulation of gene expression and can contribute to cancer progression and therapy resistance. For instance, the overexpression of miR-221/222 also associates with fulvestrant-resistance confirming that they are essential for cell growth and cell cycle progression and in conclusion also resulted in deregulation of multiple oncogenic signaling pathways previously associated with drug resistance (276). Changes in miRNA expression correlate with diagnostic and prognostic markers in breast cancer therapy (277).

By targeting specific miRNAs to modulate their expression, it may be possible to restore sensitivity to existing therapies or develop new therapeutic strategies. However, the complexity of miRNA interactions and their broad regulatory effects necessitate deeper investigation to translate these findings into clinical practice.

Overall, while significant progress has been made, further research is still needed to standardize the identification of methylation biomarkers, understand the effects of alternative splice variants, and explore the therapeutic potential of miRNA modulation. Addressing these research needs will be critical to integrating these regulatory mechanisms into the clinic and ultimately lead to more effective and personalized treatments for breast cancer.

Aims of this thesis

Overall aims

The studies in this thesis aimed to investigate how different regulatory mechanisms including alternative splicing, promoter methylation and differential promoter usage, as well as miRNAs affect the ER and ER signaling in breast cancer.

Specific aims

Paper I

The ER has multiple transcripts where the splicing pattern may influence breast cancer biology and clinical outcomes. The aim of this paper is to comprehensively understand the diversity of ER isoforms in breast cancer, reveal novel mRNA isoforms generated through alternative promoter usage and splicing, and functionally characterize several protein isoforms. Our goal is to provide insights into the role of alternative mRNA splicing in breast cancer, enable the identification of mechanisms of endocrine therapy resistance, and uncover novel therapeutic targets.

Paper II

Resistance to endocrine therapy in ER-positive breast cancer patients is a major challenge, with mechanisms that are complex and still under extensive investigation. We use a panel of six different ER-positive breast cancer cell lines, including matched sensitive and fulvestrant-resistant cells, as well as resistant cells cultured without the drug, which differ in characteristics. We study the methylation of several different promoter regions of the ER to identify the mechanisms behind the stability changes of resistance and determine if the methylation status correlates with the stability of resistance. We measure how the expression of the ER changes between these cell lines and compare this to the methylation of CpG dinucleotides in promoter regions. Our aim was to determine whether the promoter methylation of different alternative first exons of the ER correlates with the stability of

resistance, thereby enhancing our understanding of the underlying mechanisms driving resistance in ER-positive breast cancer.

Paper III

The miRNA *mir-4728*, located in the *ERBB2* (*HER2*) oncogene, was identified by the research group in 2011. In HER2-positive breast cancer, *mir-4728* is co-amplified with its host gene, making it a potential biomarker for predicting HER2 status. In our manuscript, we aim to examine the overall impact of this miRNA on gene expression and translation. We use antisense oligonucleotides to block the function of miR-4728-3p in a HER2-negative breast cancer cell line and use polysome fractionation and RNA sequencing (RNA-Seq) to study the genes and pathways affected by miR-4728-3p. Our research indicates that miR-4728-3p is associated with steroid hormone biosynthesis, indicating a potential role in regulating estrogen synthesis.

Materials and methods

Listed below are the main methods used in papers I-III of this thesis.

Cancer cell lines

Cancer cell lines are used to study the biological and molecular mechanisms of cancer in a simplified and reproducible environment. Those valuable models are useful to understand the behavior of cancer cells such as proliferation, migration and response to different treatments or to study resistance without the complexity of a whole organism.

Cancer cell lines are derived from tumors and have the ability to multiply indefinitely without entering cellular senescence. These cell lines represent populations of cells that can be maintained in culture for extended periods while retaining stability in certain phenotypes and functions. This stability allows researchers to manipulate the cells for studying various aspects of cancer biology, including gene expression, drug response, and cell signaling pathways. In this thesis, we primarily used common breast cancer cell lines obtained from the American Type Culture Collection (ATCC), expressing different receptors, such as MCF7 (ER+, HER2-), T-47D (ER+, HER2-), BT-474 (ER+, HER2+), and SK-BR-3 (ER-, HER2+), as well as one human liver cancer cell line, HepG2 (ER-, HER2-). These cell lines were used for both functional *in vitro* assays and *in silico* analysis to investigate the role of different receptors in breast cancer, explore mechanisms of drug resistance, and validate findings from other experimental approaches. We also analyzed fulvestrant-resistant cell lines developed by Kaminska *et al.* (278), which will be explained in the results section. We chose these specific cell lines because they reflect the diversity of breast cancer subtypes and provided a solid foundation for exploring mechanisms of drug resistance and other aspects of cancer biology.

One limitation of using cancer cell lines is that they may not fully reflect the complexity of human tumors, as they lack the tumor microenvironment and may undergo mutations or other changes during long-term culture. While human cell lines are suitable for laboratory work and are widely used in cancer research, the results obtained with these cell lines do not necessarily translate directly to clinical outcomes.

Cloning

Bacterial cloning is an important molecular biology technique for producing multiple copies of a specific DNA sequence. This technique is used to generate assembled plasmids containing specific promoters, coding sequences and resistance markers for use in cell line experiments.

In bacterial cloning, the desired DNA fragment is inserted into a bacterial plasmid using restriction enzymes and DNA ligase. The plasmid, which now contains the desired DNA sequence, is then introduced into a competent bacterial cell, e.g. by heat shock, a method that allows the bacteria to take up foreign DNA through temperature changes. Once the bacterium containing the plasmid multiplies and recovers, the plasmid can be isolated. This isolated plasmid can then be transfected into a cancer cell line, where "transfected" means that the cell takes up the plasmid and subsequently produces the desired protein.

Using bacterial cloning, we have produced composite plasmids with specific promoters, coding sequences and resistance markers for use in cell line experiments. In this way, we were able to introduce these plasmids into cancer cell lines so that the cells could express the desired proteins and we could better study the functions of these proteins and their role in cancer biology. This method allows precise control over the DNA sequences introduced into the cancer cell lines and ensures that the desired genes are accurately expressed in our experiments.

A limitation of bacterial cloning is that the process can be time-consuming and requires multiple steps to ensure that the correct DNA sequence is inserted and expressed. In addition, the method depends on the efficiency of the bacterial cells in taking up the plasmid, which can sometimes be low.

Validating ER gene activity

Luciferase assay

The luciferase assay is a reporter gene assay that can be used to detect and measure the activity of transcription factors, like the ER. It determines whether a protein can activate or repress the transcription of a gene of interest by quantifying the resulting expression of luciferase reporter genes.

In this method we use two plasmids where one is expressing the ER and the other one which is a dual luciferase plasmid. It contains two genes which is the *Renilla* luciferase gene, always expressing at a constant level and serves as an internal control for normalization and the luciferase gene, regulated by an ERE promoter.

When a ligand, such as estrogen or tamoxifen, binds to the ER, the dimerized receptor binds to the ERE, leading to the expression of the firefly luciferase gene. If the ER does not bind to the ERE, the firefly luciferase gene is not expressed (Figure 12).

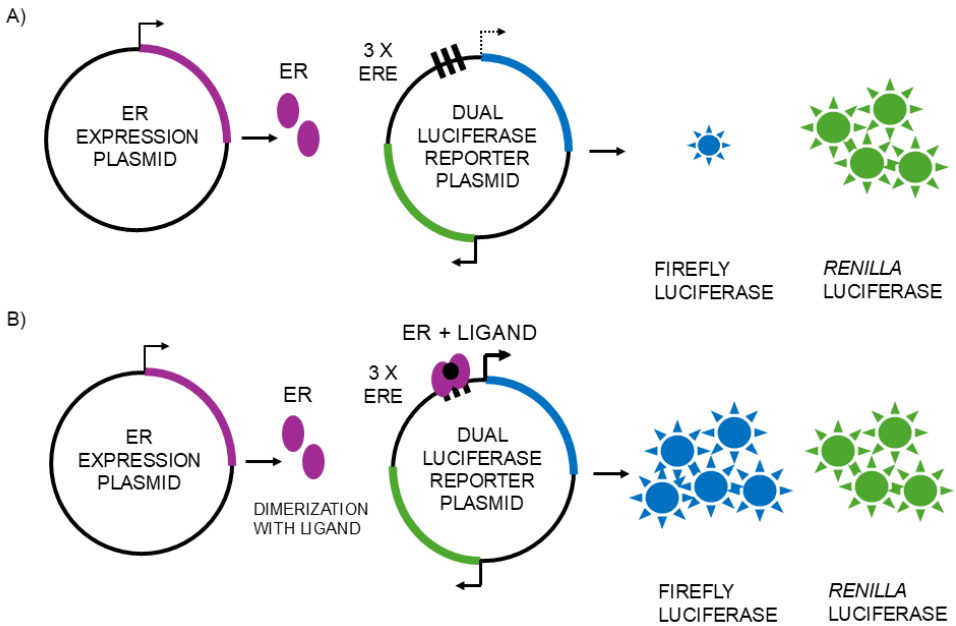

Figure 12 The mechanism of the luciferase assay involves a plasmid that expresses the ER protein, and a second plasmid containing a *Renilla* gene that is continuously expressed and used for normalization. Additionally, there is a firefly gene regulated by 3xERE sequences. Without a ligand firefly luciferase is lowly expressed (A). When the ER binds with the ligand to the ERE, the transcription of the firefly gene is activated (B).

Luminescence signals from both *Renilla* and firefly luciferases are measured using a luminometer. The ratio of these signals reflects the activity of the ER as a transcription factor and its response to various ligands.

In Paper I, we used the luciferase assay to understand how ER isoforms behave in cells, expressing endogenous ER (ER-positive cell line) or transfected with exogenous ER (ER-negative cell line). This technique allowed us to analyze and measure the transcriptional activity of ER isoforms in response to different treatments, such as estrogen or tamoxifen. The use of reporter genes allows for accurate normalization and differentiation of specific transcriptional activity, making it ideal for studying the functional impact of ER isoforms and their interactions with ligands. One limitation of the luciferase assay is that it relies on the efficiency of the transfection process and the expression levels of the reporter

genes, which can vary between experiments. Additionally, the assay measures only the transcriptional activity of the ER and does not provide information on other aspects of ER function or its interactions with co-regulators.

Alternatives to the luciferase assay is a chromatin immunoprecipitation (ChIP), which can provide more detailed information on protein-DNA interactions and ER binding to specific genomic regions.

Western blotting

Western blotting is a technique widely used in molecular biology for the detection and identification of specific proteins in complex protein mixtures. It is suitable for determining protein expression and for detecting specific target proteins in biological samples.

The method starts with the extraction of proteins from cells that can be expressed endogenously or by transfection with a plasmid. The protein concentration is measured using a protein assay kit. The quantified protein samples are then loaded onto a gel and separated according to size by electrophoresis. After separation, the proteins are transferred from the gel to a membrane. The membrane is blocked and washed to prevent non-specific binding. Specific primary antibodies are used to detect the target protein. After incubation with the primary antibodies, a horseradish peroxidase (HRP)-conjugated secondary antibody is applied to facilitate visualization. The detection reagent is then added to trigger a luminescence reaction catalyzed by the HRP-conjugated antibody. The resulting luminescence is captured and quantified using an imaging system (Figure 13). Additional confirmation proteins, such as anti-α-tubulin or anti-Lamin B2, are used to verify sample quality and equal loading.

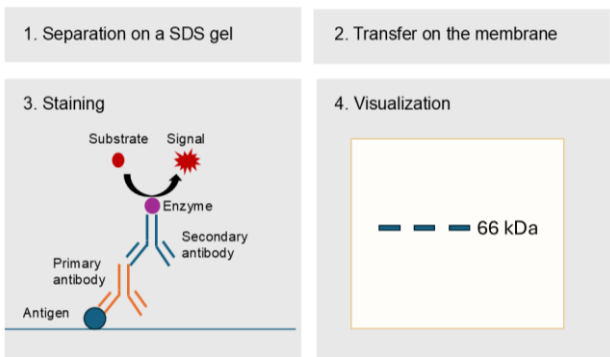

Figure 13 Western blot workflow includes protein separation by gel electrophoresis, transfer to a membrane, incubation with primary antibodies to detect the target protein, followed by secondary antibodies for signal amplification, and finally, visualization of the detected proteins.

In Papers I and II, western blotting was used to detect full-length ER and ER isoforms using C-terminal and N-terminal anti-ER antibodies. For Paper III, it was used to detect the CYP19A1 protein with a mouse anti-CYP19A1 clone H4 antibody.

Western blotting allowed us to quantify and identify these specific proteins and confirm their expression levels and localization in the samples. It provides a reliable method for the detection and quantification of specific proteins in complex mixtures. For example, we were able to distinguish between different ER protein isoforms, protein expression and localization in different cell compartments, making it an ideal experiment for our research questions.

However, Western blotting has some limitations, including the potential for non-specific binding and the need to optimize antibodies and detection conditions. It is also relatively time-consuming and requires careful handling to avoid errors. Alternatives to Western blotting include techniques such as mass spectrometry, which allow for more comprehensive identification and characterization of proteins. We tested mass spectrometry in paper I, but as mentioned in the paper, each isoform generated only a single peptide, and in general the sensitivity in our initial tests was not sufficient to detect them. However, in theory, these methods may offer advantages such as higher throughput or greater detail in protein analysis, but they may also involve more complex protocols or higher costs.

Real time RT-PCR

Real-time reverse transcription polymerase chain reaction (RT-PCR) is used to amplify and simultaneously quantify a specific RNA molecule, providing a measure of gene expression. It indicates the amount of mRNA copies of a particular gene transcript in a given sample (279).

The RT-PCR process begins with the extraction of RNA from cell lines and its conversion into complementary DNA (cDNA) by reverse transcription. For this purpose, oligo(dT) primers, nucleotides and reverse transcriptase are used, which are incubated in a thermal cycle to produce cDNA from the RNA template. The primers are added in the quantitative PCR step. These primers are complementary to the sequences at the beginning and end of the specific cDNA segment of interest. The cDNA sample is mixed with nucleotides, specific primers and a reagent containing a polymerase and a fluorescent dye, such as SYBR Green (Figure 14). The mixture is incubated to repeated heat cycles in a PCR thermocycler, during which the polymerase amplifies the specific DNA segment flanked by the primers. The fluorescent dye binds to the double-stranded DNA, and the emitted fluorescence signal is detected. The number of cycles required to produce a detectable fluorescence signal is inversely proportional to the initial concentration of the target RNA in the sample.

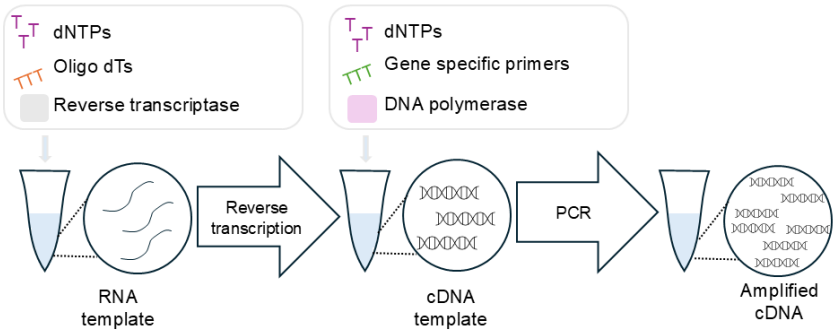

Figure 14 RT-PCR workflow includes the extraction of RNA, reverse transcription to convert RNA into cDNA, followed by the amplification of the cDNA using specific primers, allowing for the detection and quantification of target sequences.

Additionally, mRNA expression levels can be compared across different samples by normalizing the results to the total mRNA content using housekeeping genes – genes that are expressed consistently across all cells (e.g., *MRLP19* from Paper I or *ACTB* from Paper II). This normalization helps account for variations in RNA input and ensures accurate comparison of gene expression levels.

In our studies, we used RT-PCR to measure the gene expression of specific ER mRNA transcripts. By quantifying the amount of cDNA corresponding to the target genes, we were able to assess the expression levels of these genes under different experimental conditions. By choosing this method, we were able to achieve a precise and quantitative measurement of gene expression. The sensitivity and specificity of RT-PCR make it an effective tool for assessing changes in mRNA levels in response to different treatments or conditions.

However, a major limitation of RT-PCR is the quality and integrity of the RNA sample (e.g. whether it is degraded or not), which can be affected by variations in RNA extraction or reverse transcription efficiency. In addition, the method requires careful optimization of primers and reaction conditions to avoid non-specific amplification. Alternatives to RT-PCR include techniques such as RNA-Seq, which provide a more comprehensive overview of gene expression and can detect a wider range of RNA species, but are more complex and expensive compared to RT-PCR.

Polysome fractionation

Polysome fractionation is used to study the translation of mRNA into proteins in cells. It specifically detects which mRNAs are actively being translated into proteins by separating ribosomes and associated mRNAs based on their size and density in gradient (280).

This method works by isolating ribosome-mRNA complexes from cells. The process begins with cell lysis performed under conditions that preserve the ribosome-mRNA interactions. The lysate is then applied to a sucrose gradient, which ranges from low to high concentration (Figure 15). During an ultracentrifugation step, ribosome-mRNA complexes are separated based on their size and density. The gradient is collected in small volumes, allowing for the distinction between monosomes (single ribosomes), light polysomes (few ribosomes), and heavy polysomes (many ribosomes). These fractions can be analyzed using techniques such as real-time RT-PCR (Paper I) or RNA-Seq (Paper III).

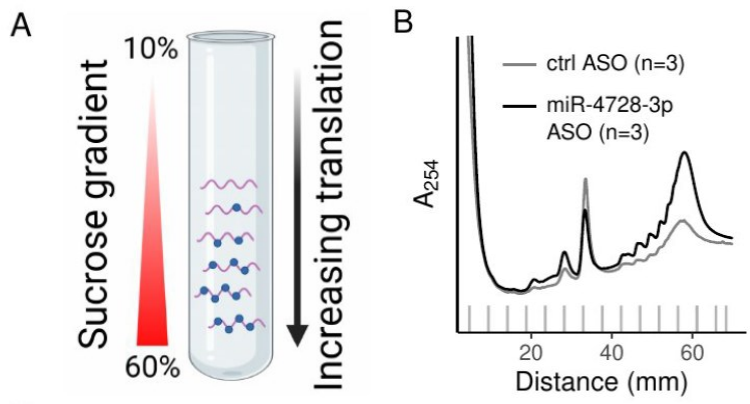

Figure 15 Polysome fractionation workflow includes the separation of ribosome-mRNA complexes using a sucrose gradient (A), followed by the analysis of mRNA distribution across the gradient (B) to assess the effect of treatments on translation (Images from Paper III).

We used polysome fractionation to gain insights into the distribution of mRNAs across different fractions of the gradient, where we were able to determine the translational activity of specific mRNAs. It provides detailed information about the efficiency of translation and helps to understand the regulation of gene expression at the translational level.

Polysome fractionation has its limitations including the risk of sample loss, the need for specialized equipment such as an ultracentrifuge and the careful handling to preserve the integrity of the ribosomal mRNA complexes. Additionally, the method may not fully capture the dynamics of translation *in vivo* due to the extraction and separation process.

Sequencing

Sequencing has redefined molecular biology by moving from Sanger sequencing to massively parallel sequencing, which generates huge amounts of biological and medical data. It enables the precise identification of nucleotide sequences in DNA molecules. These base sequences (labeled with the letters A, T, C and G) contain the essential biological information required for the development and function of cells.

Illumina - sequencing by synthesis

Illumina sequencing is a method designed to sequence large quantities of DNA or RNA rapidly and accurately. It detects the nucleotide sequence of DNA or cDNA by incorporating fluorescently labeled deoxyribonucleotide triphosphates (dNTPs) into DNA strands and measuring the emitted fluorescence signals (Figure 16).

The process begins with library preparation, where DNA, RNA or cDNA samples are processed depending on the specific application. This may involve random fragmentation of the samples or the use of pre-amplified PCR products, followed by ligation of adapter oligonucleotides to the ends of each fragment or amplicon. These adapters serve several purposes, such as providing priming sites for amplification and sequencing, including sequencing primer sequences that contain unique barcodes for sample identification and capture sequences for binding to the flow cell. The fragmented DNA sequences are then immobilized on a flow cell, a solid surface coated with oligonucleotide probes complementary to the adapter sequences, where each fragment binds to a probe and is hybridized. This is followed by bridge amplification, in which the DNA fragments are amplified by bridge amplification and form clusters of identical molecules on the surface of the flow cell. Each cluster contains thousands of copies of the same DNA fragment (Figure 16).

The sequencing itself, known as sequencing by synthesis, involves cycles of nucleotide incorporation and detection. Illumina sequencing typically generates short reads ranging from 50 to 300 base pairs (bp), although advancements have allowed for longer read lengths in some applications. Paired-end sequencing can be employed to read both ends of each DNA fragment, providing paired-end reads, which improve the accuracy of alignment and variant detection by capturing additional context from both ends of the fragment. Each cycle of sequencing includes the introduction of fluorescently labeled dNTPs and DNA polymerase. DNA polymerase incorporates a complementary dNTP into the new strand, and if the insertion is correct, a fluorescent signal is emitted, which is captured by a camera, indicating the nucleotide added to each cluster. After each cycle, the fluorophore and terminator are removed in a deblocking step, and the template is

prepared for the next cycle with a new set of dNTPs. This process is repeated for multiple cycles to sequence the entire fragment (Figure 16) (281).

In Papers I and II, Illumina RNA-Seq data was used to analyze gene expression and investigate transcriptome-wide changes. In Paper II, DNA sequencing (DNA-Seq) was applied to explore methylation status of ER promoter target regions.

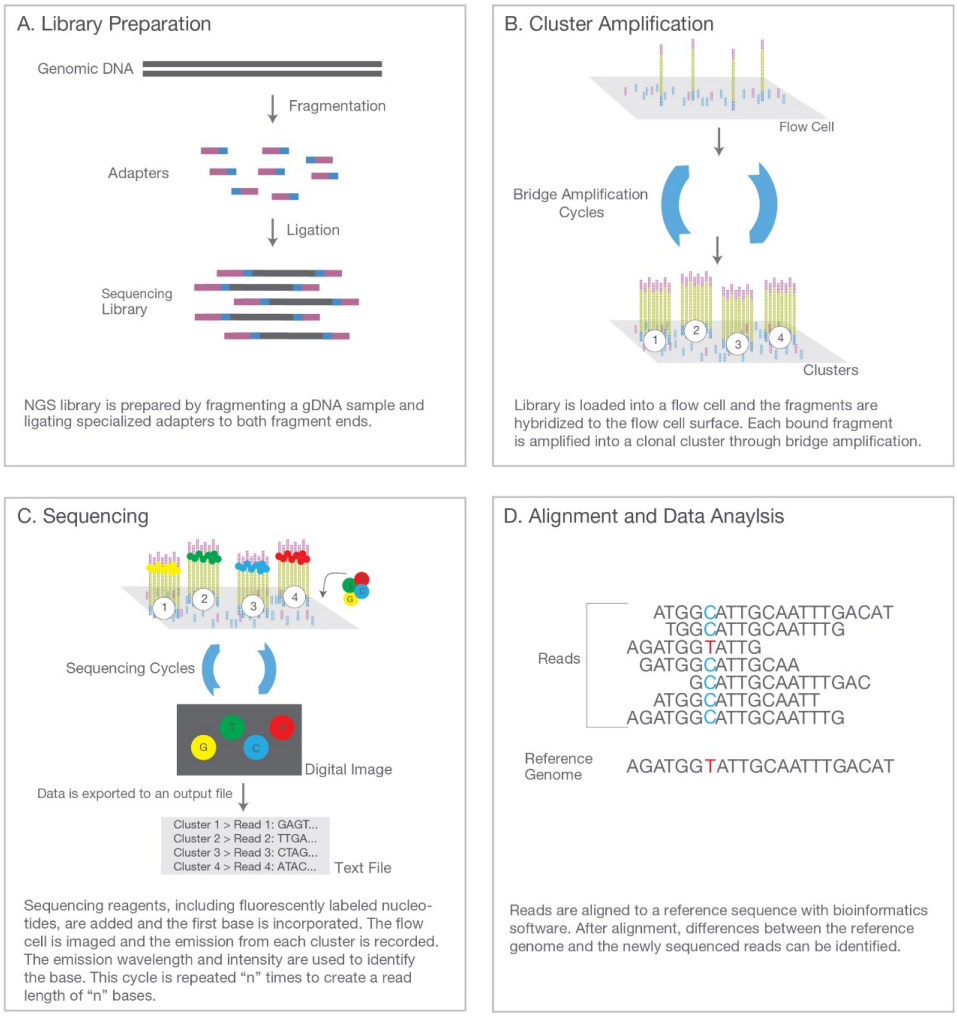

Figure 16 Illumina sequencing workflow for Next-Generation Sequencing includes four key steps, including library preparation, cluster generation, sequencing, and alignment with data analysis (From (281)).

We chose Illumina sequencing because it offers high throughput, accuracy and cost-effectiveness for high-volume sequencing. Its ability to provide high-resolution data and support paired-end sequencing enhances our ability to obtain comprehensive and reliable genomic and transcriptomic information. However, the technique also has its limitations, such as the need for high-quality samples and the requirement for expertise in sequencing protocols to achieve optimal results. The method generally produces shorter reads compared to some other technologies, which can affect the assembly of complex genomes.

Pyrosequencing

Another sequencing method we used is pyrosequencing which is a DNA sequencing technology designed to rapidly and accurately determine DNA sequences by detecting the release of pyrophosphate (PPi) during nucleotide incorporation (282,283). It provides real-time data on the DNA sequence as each nucleotide is added to the growing DNA strand.

Pyrosequencing begins with sample preparation, where the DNA is first amplified by PCR to produce double-stranded DNA. These double-strands are then separated to obtain single-stranded DNA templates. The single-stranded DNA is prepared with specific biotinylated primers at the 5' end, ensuring that only one strand of DNA is bound to the surface during sequencing as the complementary strand is removed.

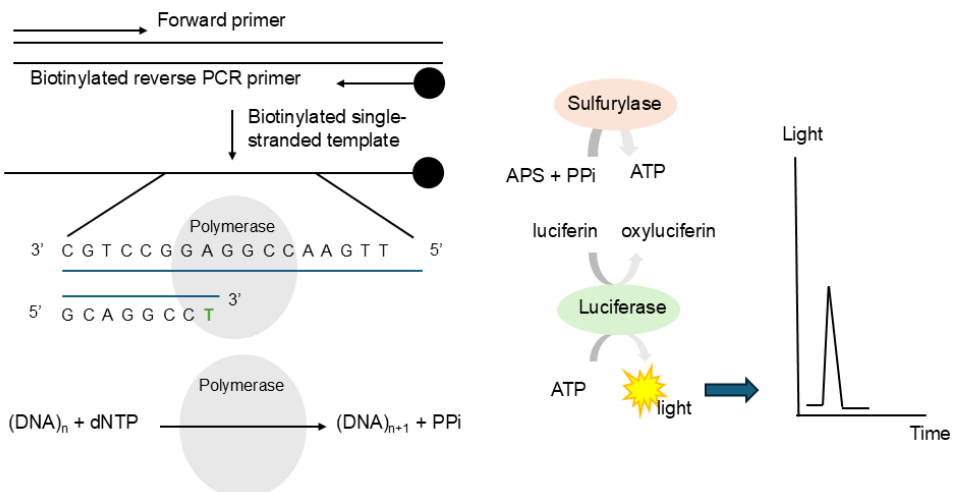

Figure 17 The Pyrosequencing workflow involves annealing primers to a single-stranded DNA template, followed by nucleotide incorporation by DNA polymerase. The release of PPi triggers a reaction that produces light, which is detected and measured to determine the DNA sequence (modified from (284)).

Following preparation, the sequencing reaction is initiated by mixing the single-stranded DNA with various components, including DNA polymerase, beads, luciferase, and substrates necessary for nucleotide incorporation. The sequencing progresses through nucleotide incorporation, where matching nucleotides are added to the growing DNA strand, releasing pyrophosphate (PPi) as a byproduct. This release of PPi triggers a series of enzymatic reactions, beginning with ATP sulfurylase converting PPi to ATP (285). In the presence of luciferase, ATP is then used to produce light, with the intensity of the emitted light corresponding to the nucleotide added. This light is detected and recorded by a camera, providing the sequence data with a read length of around 150 bp (Figure 17) (282).

Pyrosequencing was used as a validation method to Illumina sequencing to determine DNA sequences and analyzing methylation status in the target ER promoter region. It is particularly useful for applications requiring high-throughput sequencing with accurate real-time data, such as mutation detection and sequencing of small genomic regions.

We chose pyrosequencing because it is simple, cost-effective and able to accurately quantify methylation at a limited number of sites, making it an ideal method for validating our previous Illumina sequencing results. While pyrosequencing provides detailed sequence information quickly, it is limited by low throughput and shorter read lengths, making it less suitable for capturing complex or larger sequences. Developing effective assays for pyrosequencing can also be a challenge. Illumina sequencing would have required more work, higher costs and provided far more data than was required for this particular aspect of our study. Therefore, pyrosequencing was the more practical and efficient choice.

Long-read sequencing

Long-read sequencing is a genomic technology designed to generate extended sequences of DNA fragments, often few kilobases (kb), compared to short-read sequencing methods that typically produce shorter reads of up to 300 bp (286). This method is used to obtain comprehensive information about complex genomes, including accurate assembly of repetitive regions, detection of structural variants, and full-length transcript visualization.

Oxford Nanopore Technology (ONT) and Pacific Biosciences (PacBio) are the two main long-read sequencing methods. In ONT, DNA molecules are passed through a nanopore, and as the DNA travels through the pore, changes in electrical conductivity are measured to determine the DNA sequence. This real-time measurement of electrical changes enables long-read sequencing.

PacBio, on the other hand, uses single-molecule real-time sequencing (SMRT), in which the DNA polymerase incorporates nucleotides into a growing DNA strand in real time. The technology detects the incorporation of nucleotides by measuring the

fluorescence emitted as each nucleotide is added, which also enables long-read sequencing.

In Paper I, we used ONT to generate long reads for mRNA isoform annotation. This approach allowed us to visualize full-length transcripts, which cannot be achieved with short-read sequencing methods. We chose long-read sequencing to generate extended sequence information that improves the assembly of full-length ER isoforms.

However, long-read sequencing methods have significant limitations. Especially ONT can have higher error rates compared to short-read technologies. Long-read sequencing has a lower throughput compared to short-read sequencing methods.

In conclusion, Illumina sequencing is recognized for its high accuracy and cost-efficiency in generating short reads, though it may be limited in resolving complex genomic regions and full-length transcripts that are better addressed by long-read sequencing platforms such as ONT and PacBio. Pyrosequencing, offering intermediate read lengths, represents a compromise but has seen decreased utilization with the development of more advanced sequencing technologies. The selection of a sequencing method should be guided by the specific objectives of the study, with careful consideration of factors such as read length, accuracy, and cost.

Study Cohorts

SCAN-B

The Sweden Cancerome Analysis Network – Breast (SCAN-B) initiative is designed to enhance the understanding of breast cancer biology through comprehensive genomic characterization and to establish a detailed, population-based archive of breast cancer cases in southern Sweden by analyzing patient samples (287). This population-based study, conducted across nine hospital centers, invites patients with newly diagnosed breast cancer to participate by providing blood and tumor tissue samples. The SCAN-B study adheres to the ethical principles outlined in the Declaration of Helsinki and has received approval from the Regional Ethical Review Board of Lund (2007/155, 2009/658, 2009/659, 2014/8), the county government's biobanking center, and the Swedish Data Inspection Group (364-2010) (287).

The genomic characterization component of SCAN-B involves RNA-Seq of tumor samples, enabling a thorough analysis of gene expression and molecular characteristics specific to each patient's cancer. In Papers I and II, RNA-Seq data and clinical information from the SCAN-B study were utilized to explore breast cancer biology. For Paper I, data from 3,478 breast tumors with varying ER, PR,

and HER2 statuses were analyzed. For Paper II, we selected a subset of 1,429 patients from the original 3,478 for survival analysis, focusing specifically on postmenopausal women with ER-positive breast cancer who received endocrine therapy but no chemotherapy or anti-HER2 treatment. The SCAN-B dataset was chosen for its comprehensive, population-based approach, offering a broad and detailed analysis of cancer-specific genetic and molecular features, enabling robust investigation of breast cancer biology and treatment responses within a real-world clinical context. However, the SCAN-B data's specificity to southern Sweden may limit generalizability to other populations, and the availability of specific patient subgroups can affect the scope of analysis.

TCGA data

The Cancer Genome Atlas (TCGA), a landmark cancer genome program funded by the National Cancer Institute and the National Human Genome Research Institute in the United States, provides molecular characterization of over 20,000 primary cancer and matched normal samples from 33 cancer types including breast cancer (288). It is a key initiative in cancer research, beginning with the collection of samples from patients with various cancers. These samples undergo analysis using multiple high-throughput technology platforms, including DNA sequencing, RNA sequencing, and methylation arrays, to generate detailed molecular profiles. The resulting data are used to characterize the genetic, epigenetic, and transcriptomic features of tumors and normal tissues.

In Papers I and II, we utilized the TCGA-BRCA dataset to examine breast cancer-related molecular features and validate findings from other datasets. The extensive molecular data provided by TCGA enabled us to expand and deepen our research on breast cancer biology, facilitating robust analysis and comparison across different technologies. We selected TCGA data for its comprehensive, multi-platform approach to cancer characterization, including methylation data that were not available for SCAN-B. The large dataset and broad scope offer valuable insights into the molecular mechanisms underlying breast cancer. However, the use of TCGA data comes with limitations, such as its representation of a wide range of cancer types, which may lack specific focus on certain subtypes or patient demographics, and the possibility of incomplete or missing data in some patient samples.

Comparison SCAN-B and TCGA

The SCAN-B initiative and TCGA are both important data platforms for advancing breast cancer research through molecular characterization, but they serve different purposes and offer unique strengths. SCAN-B, with its population-based approach focused on southern Sweden, provides highly relevant, context-specific data

reflecting the real-world incidence and progression of breast cancer within a specific population. This specificity makes SCAN-B particularly valuable for local research, enabling detailed studies of gene expression and molecular profiles that are directly applicable to patient care in this region. However, SCAN-B results may be less generalizable to other populations due to its limited geographic scope.

On the other hand, TCGA provides a large and diverse dataset that supports a broad range of molecular analyses for different cancer types, including breast cancer. This breadth enables comprehensive comparative studies across different cancer types and research into general cancer mechanisms. However, the extensive and diverse nature of TCGA data can result in a diluted focus on a single cancer type or population, making it less suitable for highly targeted breast cancer research. In addition, some samples in TCGA may have incomplete data, limiting the depth of analysis possible for certain breast cancer subtypes.

The difference in RNA sequencing between the two initiatives illustrates this contrast: SCAN-B uses RNA sequencing to generate detailed gene expression profiles specific to breast cancer in a given population, while TCGA applies RNA sequencing more broadly to multiple cancer types, resulting in less detailed data specific to breast cancer. While SCAN-B is characterized by providing detailed, population-specific insights that are important for understanding regional clinical scenarios, TCGA is better suited for broad, comparative studies aimed at identifying novel biomarkers and exploring cancer mechanisms in different cancer types and populations. SCAN-B is thus characterized by providing detailed, population-specific data, while TCGA offers comprehensive, cross-cancer insights, each with their own advantages and limitations.

Results and discussion

Paper I: Alternative promoters and splicing create multiple functionally distinct isoforms of oestrogen receptor alpha in breast cancer and healthy tissues

Background

The ER is an oncogene in breast cancer and acts as an important prognostic factor, therapeutic target and predictor of response to treatment. The gene encoding the ER, *ESR1*, generates multiple transcripts through alternative splicing and promoter usage. While these transcript variants may not alter breast cancer biology or clinical outcomes dramatically, their differentiated effects on ER function and signaling pathways remain important and worth careful consideration.

We identified significant gaps in the GENCODE and RefSeq databases commonly relied upon for transcript annotation in genomic research. Notably, many splicing events detected in the SCAN-B breast cancer cohort were not included in these databases, making it difficult to comprehensively analyze the ER transcriptome. For example, we identified an isoform known as ERΔ7 in RefSeq but missing in GENCODE, and another isoform reported as ER∂E3 missing in both databases but here were also many novel exons and splice junctions that had not been previously described.

Our aim was to expand the understanding of ER isoform diversity in breast cancer by discovering novel mRNA isoforms resulting from alternative promoter usage and splicing, followed by functional characterization of several identified protein isoforms. Through this comprehensive analysis, we intended to gain insights into the role of alternative mRNA splicing in breast cancer and provide the basis for investigating mechanisms of endocrine therapy resistance and discovering novel therapeutic targets.

Results

In this study, we focused on creating a comprehensive annotation of ER transcripts by integrating short-read and long-read RNA sequencing data from breast tumors, cell lines, and normal tissues and investigating biological effects of a subset of protein coding isoforms by experimental approaches. Short-read data is highly

effective for annotating individual features due to its depth and the large number of samples, but it cannot be used for complete isoform annotation. Long-read sequencing, on the other hand, allows for the mapping of whole transcripts, though it is comparatively shallow.

Through this combined approach, which integrated known annotations from GENCODE and RefSeq, we identified numerous novel first, internal, and last exons, as well as splice junctions and splice sites, significantly expanding the known transcript map of ER. This resulted in the identification of 146 exons, 74 exon starts, 77 exon ends, and 326 splice junctions.

We then assessed the expression patterns of these extended exons and splice junctions across ER-positive and ER-negative tumors, discovering varied expression levels. Some novel exons and splice junctions showed weak expression, while others exhibited patterns inconsistent with previously annotated events. To further explore these features, we performed long-read sequencing of RT-PCR products to assemble them into full-length transcripts. This approach unveiled numerous alternatively spliced isoforms, including some with novel protein-coding sequences that were differentially expressed.

Among these full-length transcripts, we searched for open reading frames to compare with the known coding sequences and alternative protein isoforms. Among these, six were selected for experimental validation studies (Figure 18, Table 3). We called the protein-coding isoforms 'cds', which stands for 'coding sequence' together with an arbitrary numbering.

For instance, cds1 results from an exon-skipping event leading to a frameshift that produces a truncated protein; cds11 uses an alternative splice donor that inserts two amino acids into the DBD; cds13 has a 39-amino acid deletion in the DBD; cds16 has a single-amino acid deletion in the DBD, while cds38 and cds60 have larger internal deletions affecting multiple domains. Notably, cds13 and cds60 were not present in existing databases, although cds13 has been previously described, while the others could be translated from GENCODE or RefSeq transcripts.

To investigate whether these selected isoforms have transcription factor activity, we cloned six isoforms and tested them in luciferase assays. The results showed that two isoforms, cds11 and cds16, exhibited transcription factor activity, although this activity was lower than that of the full-length ER. We further examined the impact of co-expressing these alternative isoforms with the full-length in ER-negative HepG2 cells. One isoform, cds13, displayed dominant-negative behavior, while others, like cds1, cds38, and cds60, either enhanced or reduced transcriptional activity when co-expressed with the full-length ER.

Table 3 Characteristics of protein isoforms selected for functional validation, including differences in coding sequence according to HGVS nomenclature, number of amino acids, correspondence to annotated transcripts and underlying alternative splicing events.

Isoform	Alteration	Amino acids	Annotated transcripts	Transcriptional event
cds1	Gly457Valfs*10	466	NM_001385570, NM_001385571, NM_001385572	Skipping exon 7
cds11	Gly215_His216ins AsnArg	597	NM_001291230	Alternative donor exon 2
cds13	His216_Gly254del	556	Novel	Skipping exon 3
cds16	Gly254del	594	NM_001291241	Alternative donor exon 3
cds38	Pro152_Arg412del	334	ENST00000406599.5	Skipping exons 2-5
cds60	Gly254_Pro365del	483	Novel	Skipping exon 4

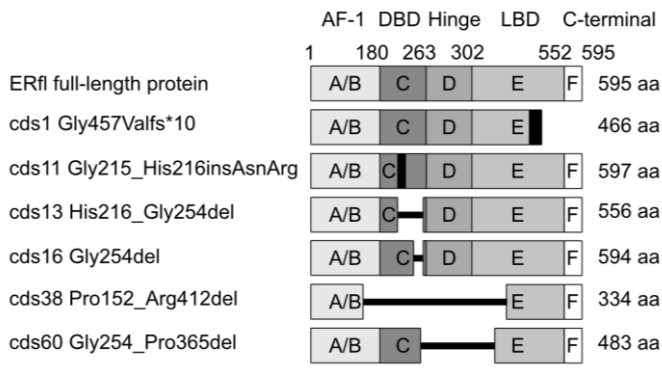

Figure 18 Schematic illustration of the domain structure of the selected CDS isoforms, showing the unique alterations of each isoform (Image from paper I).

In addition to their functional properties, these isoforms also exhibited differences in subcellular localization. For instance, cds11 showed greater accumulation in the nucleus, as determined by subcellular fractionation followed by western blotting, which assessed the relative distribution of these isoforms between the cytoplasm and nucleus using both N-terminal and C-terminal anti-ESR1 antibodies (Figure 19).

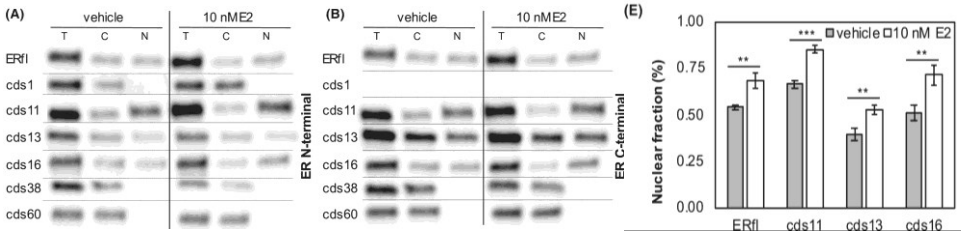

Figure 19 Subcellular fractionation followed by western blotting to determine the relative distribution between cytoplasm and nucleus for each isoform using (A) an N-terminal or (B) a C-terminal anti-ESR1 antibody. Quantitation of nuclear fraction in vehicle control and after 30min treatment with 10nM E2. Bars indicate the mean of n=3 replicates ± standard deviation. T=total lysate, C=cytoplasmic fraction and N=nuclear fraction, vehicle control=0.1% ethanol. **p<0.001 (Images from paper I).

Lastly, we tested the sensitivity of these isoforms to fulvestrant. Isoforms lacking parts of the ligand-binding domain were resistant to fulvestrant, which correlated with their lack of transcriptional activity. In contrast, the isoforms that retained transcriptional activity were sensitive to fulvestrant. These findings suggest that isoforms lacking transcriptional activity due to alterations in the ligand-binding domain could contribute to resistance to endocrine therapy in breast cancer, underscoring the importance of understanding the functional diversity of ER isoforms in the context of breast cancer treatment.

Our study highlights the complexity of ER transcript variants in breast cancer and demonstrates that combining short-read sequencing for depth with long-read sequencing for isoform completeness is important for accurate annotation. This work lays a strong foundation for advancing our understanding of ER's role in breast cancer and its implications for treatment strategies.

Limitations

One limitation of this study is our reliance primarily on tumor data, with only some normal tissue samples included. Tumors are known to have a more chaotic transcriptome, which can complicate the analysis of isoforms and their biological relevance. However, this challenge reflects a broader issue in the field of gene expression analysis which is the incompleteness and inconsistency of databases like GENCODE and RefSeq. These databases often lack comprehensive annotation, as demonstrated by our isoform cds13 (identified as ER∂E3 in other studies) being unannotated. Additionally, databases sometimes include predicted transcripts without biological evidence, making functional interpretations more complex and less reliable.

In terms of experimental design, we initially planned to use mass spectrometry for protein identification to validate that the alternative isoforms were expressed at the protein level as an alternative method to western blotting. However, this approach

proved ineffective due to the production of only one unique peptide per isoform. Mass spectrometry typically requires multiple unique peptides for reliable isoform identification, as a single peptide peak may not provide sufficient evidence to distinguish between closely related isoforms, particularly when peptides are short or shared among multiple proteins. Our preliminary tests showed that the sensitivity was insufficient to detect these peptides, even in cells transfected to overexpress the isoforms at high levels. This limitation highlights the technical challenges of quantifying protein isoforms, especially when using less efficient digestion enzymes like chymotrypsin for mass spectrometry.

The selection of antibodies also presented a challenge. Our experiments revealed that the ability and sensitivity to detect different ER isoforms using N- or C-terminal anti-ER antibodies varied significantly depending on the specific protein domains included in the isoforms. For example, cds1 was detected with an N-terminal antibody but not with a C-terminal antibody, as it lacked the C-terminal region.

Additionally, our study focused on only six selected CDS isoforms, offering a limited view of the diverse behaviors of the full spectrum of ER isoforms. The co-transfection experiments, which involved only the full-length ER, may not fully capture the complex interactions between different isoforms. Investigating these interactions, especially when multiple isoforms are present simultaneously, is a significant technical challenge. Transfection efficiency can decrease, and cells may respond unfavorably to the introduction of multiple plasmids, complicating the analysis.

Discussion

The ER holds significant prominence in breast cancer research, having been extensively studied since the early 1960s. However, ongoing challenges in transcript annotation continue to hinder the performance of high-quality research. A major problem lies in the lack of comprehensive and consistent data on transcript isoforms, as well as discrepancies across databases like GENCODE and RefSeq. These issues were evident in our study, where different database versions yielded varying transcript annotations, leading to potential misunderstandings and misinformation that could compromise the reproducibility and validity of research findings. This problem is increased by the presence of predicted transcripts that lack biological evidence, which further complicates functional interpretation.

We used long-read sequencing to achieve an in-depth analysis of full-length ER transcripts. This method, unlike short-read sequencing, enabled us to accurately map all exons simultaneously, offering comprehensive insights into transcript isoforms, alternative splicing events, and potential promoter usage. While long-read sequencing is proficient at detecting large structural variants, it is not without its limitations, including higher error rates and lower read depth compared to short-read sequencing. Despite these weak spots, our approach provided critical insights

into the complexity of ER transcriptomes, particularly the identification of novel exons, splice junctions, and protein isoforms.

Our study also highlighted the differential expression of ER isoforms across breast cancer subtypes, which may play a role in therapy response and the development of resistance to endocrine treatments. However, these results should be viewed carefully as the full-length ER protein shows the highest expression level and the exact significance of the alternative isoforms is not yet fully understood. It is challenging to determine their impact on the biology and treatment of breast cancer without a comprehensive understanding of the isoforms produced, their expression levels and their functional role. Continued research is therefore needed to clarify the significance of these alternative isoforms.

One important conclusion from our study is the variability in antibody detection, which is a potential challenge in the clinical treatment of breast cancer. IHC is routinely used in pathology laboratories to determine ER expression and serves as a basis for treatment decisions (83). However, the use of different antibodies can lead to substantial differences in the detection of alternative isoforms. For instance, two commonly used N-terminal antibodies (EP1 and 6F11) are capable of detecting all six alternative isoforms included in our functional studies, while the C-terminal antibody SP1 would miss isoforms like cds1 (289). Indeed, such variability could have significant clinical implications, indicating that some tumors classified as ER-negative may actually express alternative isoforms that could influence treatment response. Using a more refined approach to ER detection, potentially including isoform-specific expression analysis, could improve diagnostic accuracy and help to tailor treatment strategies to individual patients.

Our results also emphasize the potential of alternative ER isoforms to contribute to resistance to endocrine therapies. Isoforms lacking parts of the LBD, showed resistance to fulvestrant. These findings suggest that these isoforms may provide alternative survival pathways for cancer cells when conventional ER-targeted treatments are used. Together, these findings emphasize the importance of isoform-specific research for the development of more effective therapeutic strategies.

By identifying a previously unexpected complexity in the expression of alternative ER mRNA isoforms, our research provides a comprehensive functional characterization of six alternative protein isoforms. Among these, two isoforms exhibit transcription factor activity, one functions as a dominant-negative receptor when co-expressed with the full-length ER, and three are transcriptionally inactive but resistant to fulvestrant-induced degradation. These insights may aid in the interpretation of clinical and experimental data and support future studies aimed at improving our understanding of the ER in health and disease.

Paper II: Dynamic methylation and expression of alternative promoters for oestrogen receptor alpha in cell line models of fulvestrant resistance

1. Background

Resistance to endocrine therapies in ER-positive breast cancer patients remains a serious challenge in treatment and often leads to more aggressive tumor behavior and a poor prognosis (290). The mechanisms of resistance are complex and multifaceted and include mutations in the ER, kinase activation, phosphorylation of the ER and epigenetic changes such as DNA methylation. Such alterations can lead to altered ER signaling and activation of alternative growth pathways that contribute to the development of resistance. One important epigenetic mechanism is the methylation of cytosines in CpG dinucleotides, which occurs characteristically in promoter regions and can repress gene expression and contribute to resistance.

In our study, we focused on the investigation of the role of DNA methylation in the regulation of ER expression in six different ER-positive breast cancer cell lines that have developed resistance to fulvestrant. Among these were fulvestrant-sensitive cells (P), fulvestrant-resistant cells (FR) and fulvestrant-resistant cells cultured without fulvestrant (FR-F), in which was resistance either maintained or lost (Table 4). Those cell lines, previously studied by Kaminska *et al.,* were associated with overexpression of cyclin E2 as a biomarker for prolonged resistance and shorter progression-free survival in patients (278).

Table 4 Characteristics of cell line panel receptor expression and fulvestrant resistance stability

Cell line	Receptor expressed	Stability of resistance
CAMA-1	ER+ HER2-	long term
ZR-75-1	ER+ HER2-	long term
EFM-19	ER+ HER2-	moderate
HCC1428	ER+ HER2-	moderate
MCF7	ER+ HER2-	low
T-47D	ER+ HER2-	low

These cell lines showed striking differences in resistance stability under various culture conditions. For example, the HCC1428 cell line exhibited moderate resistance stability and returned to sensitivity after a few weeks, while the T-47D cell line became sensitive again within only one week. These results emphasize the dynamic and potentially reversible nature of resistance mechanisms.

The aim of our study was to investigate the methylation of different ER promoter regions to uncover the mechanisms underlying variations in resistance stability,

specifically focusing on how methylation of alternative promoter regions within the ER gene correlates with changes in ER expression and resistance stability in breast cancer cell lines. Indeed, our results showed that DNA methylation plays a distinct role in regulating ER expression, with high methylation levels often leading to downregulated ER expression and sustained drug resistance. Vice versa, reduced methylation has been associated with ER re-expression and loss of resistance.

Besides expanding our understanding of the molecular basis of endocrine resistance, these findings suggest that targeting DNA methylation and alternative promoter usage may be effective strategies to potentially improve cancer prognosis and therapy. Through the integration of these findings, our study contributes to a more comprehensive understanding of epigenetic regulation in endocrine resistance and lays the basis for future research to develop more effective treatment strategies for ER-positive breast cancer.

Results

We explored the dynamic methylation and expression of alternative ER promoters in breast cancer cell lines resistant to the selective ER degrader fulvestrant. We analyzed six ER-positive breast cancer cell lines as well as their fulvestrant-resistant (FR) sublines and resistant cells cultured without fulvestrant (FR-F). Interestingly, we found that the HCC1428 cell line showed increased ER expression under resistance conditions, while the other five cell lines showed decreased ER expression, matching the results of Kaminska *et al*.

We used bisulfite conversion together with Illumina sequencing to analyze DNA methylation in the ER promoter regions. In this method, DNA is treated with sodium bisulfite, which converts unmethylated cytosines to uracil, while methylated cytosines remain unchanged. We designed 12 primer pairs targeting multiple first exons of ER, covering a total of 108 CpG sites in the ER promoter regions. For each cell line, 12 amplicons of approximately 400-450 bp each were generated, pooled, purified and sequenced after PCR amplification to allow accurate identification of methylated cytosine positions.

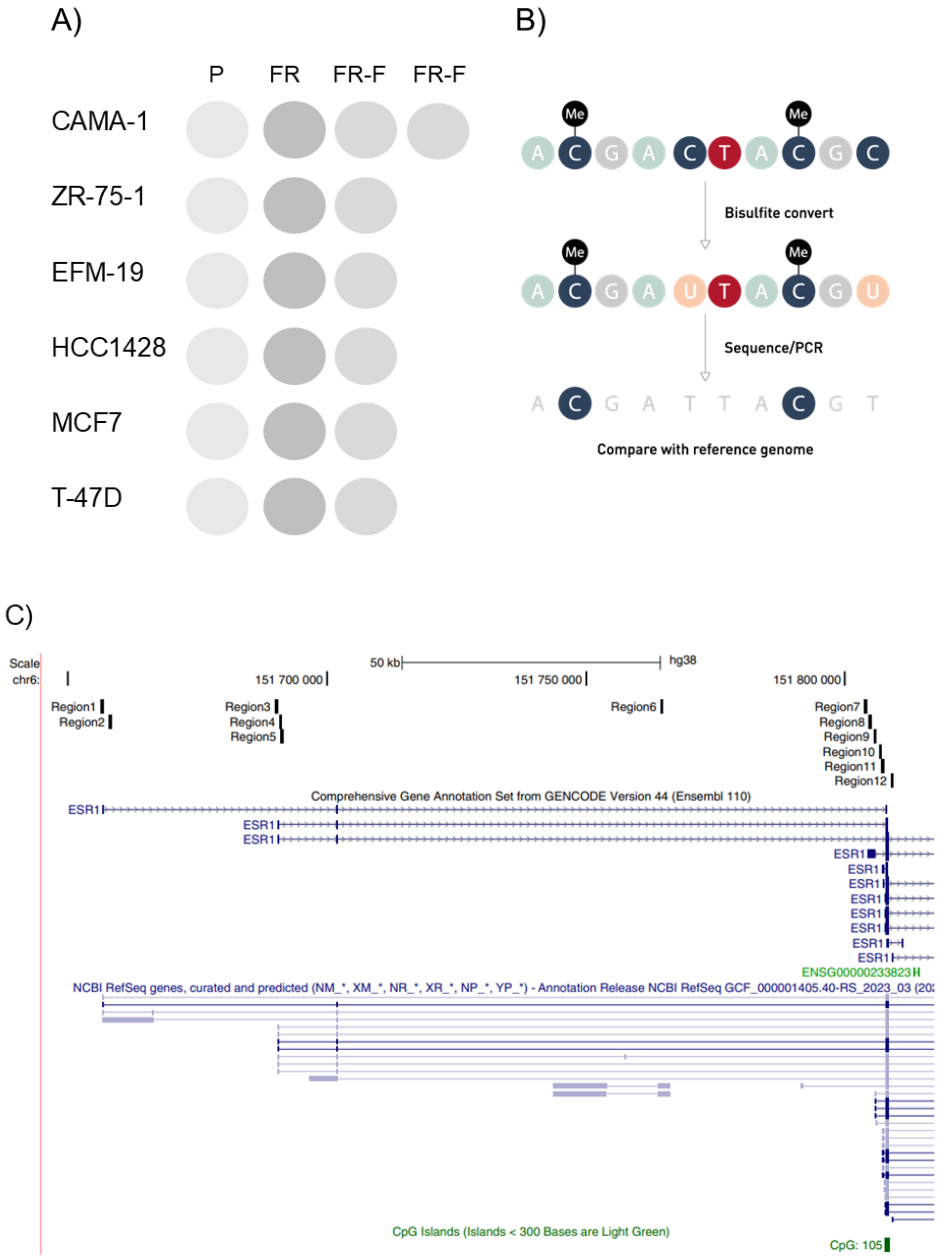

Figure 20 Used breast cancer cell lines (P, FR,FR-F) n=19 (A), principle of bisulfite conversion (B) and the 12 selected regions for PCR amplification followed by sequencing (C) (Images from paper II and (291)).

Analysis revealed significant differences in promoter methylation between cell lines and between different subline groups (P, FR, FR-F). Prolonged resistance and ER downregulation were linked in particular to high methylation levels in certain promoter regions (regions 11 and 12) close to the main promoter. Conversely, loss of resistance was associated with re-expression of ER and reduced methylation levels, indicating that methylation status is closely related to the stability of resistance.

In order to validate these results, we performed pyrosequencing on two selected cell lines, ZR-75-1 and HCC1428, focusing on regions 10-12. The results were consistent with the Illumina sequencing data and confirmed the observed methylation patterns. Analysis in comparison with published data on tamoxifen-resistant cell lines showed similar methylation changes, underlining further the pattern of consistency of these epigenetic modifications across different endocrine therapies.

Besides validating these epigenetic changes, we performed functional annotation of CpG sites and identified that methylation at sites overlapping with binding sites of transcription factors such as ETS2 and GATA3 correlates with changes in ER expression. This indicates that methylation could directly interfere with transcription factor binding, thereby influencing gene expression and contributing to the resistance phenotype. Moreover, evolutionary analysis revealed that CpG sites within ER promoter regions are evolutionarily less conserved among other genes, highlighting their potential role in species-specific regulation of ER expression.

We analyzed methylation data from 1095 breast cancer patients from TCGA cohort to assess the clinical relevance of our findings. According to this analysis, CpG sites within ER promoter regions, particularly those with higher methylation levels, were associated with lower ER expression, further linking methylation status to clinical outcomes. We also extended our findings to a subset of the SCAN-B cohort, which included 1429 post-menopausal women with ER-positive breast cancer who received endocrine therapy. Survival analysis performed in this cohort showed that higher expression of certain alternative ER exons, particularly promE and promC (Figure 21), was associated with poorer overall survival (OS), indicating that these exons may be markers of aggressive tumor behavior. On the other hand, increased expression of promA correlated with better patient outcomes, including longer OS and recurrence-free interval (RFI). These results suggest that specific methylation patterns and alternative uses of the promoter have a significant impact on patient prognosis.

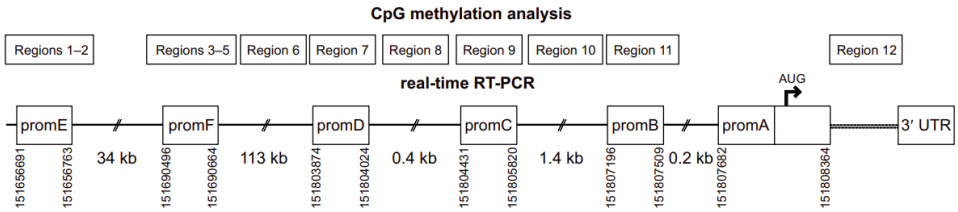

Figure 21 Representation of the 12 regions with CpG sites together with the different first exons of the ESR1 locus analyzed by real-time RT-PCR, labeled with promA-F (Image from paper II)

This study provides a comprehensive analysis of how DNA methylation of ER promoter regions influences ER expression and the stability of resistance in breast cancer. By linking these epigenetic changes to transcription factor binding and clinical outcomes, we may provide critical insights into the molecular basis of endocrine resistance. These findings not only strengthen our understanding of resistance mechanisms, but also suggest potential biomarkers and therapeutic targets to improve patient prognosis in ER-positive breast cancer.

Limitations

Our study focused on fulvestrant, which is not the first line treatment for ER-positive breast cancer. Usually, tamoxifen is the initial treatment, followed by fulvestrant if resistance or disease progression occurs. Having directly compared the effects of tamoxifen and fulvestrant in a similar experimental context may have been insightful. Although we included a number of published studies that provided insights into methylation patterns in tamoxifen-resistant ER-positive cell lines, they often lacked the comprehensive extent of our work, which covered a broader panel of cell lines and conditions.

Our study uses an extensive cell line panel extending beyond the usual two to three cell lines found in most other studies and incorporates fulvestrant-resistant lines both with and without continuous drug exposure over time. Nevertheless, the study could have benefited from an even broader inclusion of ER-positive cell lines and from sampling at additional time points during the development of resistance, particularly in the FR-F state. Fulvestrant is administered over longer durations in the clinical setting, and studying temporal differences could have provided deeper insights into the dynamics of resistance.

Another important starting point for future research would be to examine whether demethylation of ER promoter regions can restore ER function in resistant cell lines or even in TNBC cell lines, where the ER is often epigenetically silenced. Furthermore, while we have focused on methylation, resistance to endocrine therapy relies on a variety of other mechanisms, including mutations. Although Kaminska *et al*. examined some *ESR1* mutations, our study did not address these other

potential resistance factors, which may be important for a more integrated understanding of resistance mechanisms.

Experimentally, one limitation of our study was the size of the amplicons, which were kept to a maximum of 450 bp. This cut-off was necessary to ensure complete coverage of the amplicon during sequencing and to maintain the reliability of the bisulfite-converted PCR, given the maximum read length of the MiSeq kits used. However, a known challenge with PCR amplification of bisulfite-converted DNA is its low efficiency, which can be influenced by both DNA fragmentation and secondary structure formation. This limitation meant that some CpG sites were excluded from our analysis despite their potential significance. Although a design with more, shorter amplicons could have been used, this would have required running more PCRs with fewer CpG sites covered in each region. Additionally, primer design would still have been challenging since we aimed to avoid placing primers in CpG positions, not necessarily making the process easier. Despite this, the chosen approach was validated as it helped avoid the increased risk of errors and unreliable results associated with longer amplicons.

In addition, while the inclusion of the SCAN-B methylation data for primary tumors would have provided valuable insights, the methylation data from the SCAN-B project, particularly regarding the association between DNA methylation and clinical outcomes, is still in production and has not yet been released for analysis. This limitation currently prevents us from fully investigating how methylation correlates with clinical outcomes in ER-positive breast cancer. Once these data are available, they may provide a more complete understanding of the role of methylation in resistance and its impact on patient prognosis.

Overall, while our study makes an important contribution to the understanding of epigenetic regulation of ER in breast cancer, these limitations and challenges highlight areas for further exploration and refinement. Future studies that include a broader range of cell lines, additional time points, and incorporation of clinical outcome data will be critical to better understanding mechanisms of resistance and improving therapeutic strategies for ER-positive breast cancer.

Discussion

Breast cancer cell line models are essential resources for studying the mechanisms of resistance to endocrine therapies and provide valuable insights into how resistance develops and progresses. Several models, including those involving cells cultured in increasing concentrations of drugs such as tamoxifen or fulvestrant, and those exposed to long-term estrogen deprivation, have been developed to mimic clinical settings. In particular, the MCF7 cell line is a well-established model for tamoxifen resistance, in which tamoxifen can stimulate paradoxical growth in resistant cells, highlighting the complexity of endocrine resistance. Altered growth factor receptor signaling, such as upregulation of *ERBB2*, is often associated with these resistant phenotypes, particularly in cells that remain hormone independent.

Fulvestrant-resistant MCF7 and T-47D cells, for example, typically exhibit downregulated ER expression, further complicating the therapeutic landscape.

Addressing this, our study focused on fulvestrant-resistant ER-positive breast cancer cell lines to investigate the epigenetic mechanisms underlying resistance. We focused in particular on the methylation patterns in the promoter regions of the ER using NGS on the Illumina MiSeq platform. This approach offers several key advantages over conventional methylation arrays. By targeting more CpG sites in the regions of interest, we achieved higher sensitivity and single-base resolution, which allowed us to detect subtle but potentially significant methylation changes that might have been missed by other methods. Moreover, by identifying methylation patterns within DNA fragments, we were able to investigate the clonality and heterogeneity of cell populations, which gave us insights into the dynamics of resistance. Importantly, the cost per sample was also significantly lower compared to methylation arrays, as our focus on specific regions of interest allowed for more efficient and targeted analysis.

Our findings showed significant differences in the methylation patterns of the different cell lines, with methylation proving to be a potential biomarker for monitoring disease progression. In contrast to mutations, which are permanent changes in the DNA sequence, methylation is an epigenetic modification that can be reversible, making it an attractive target for therapeutic intervention. The correlation between specific CpG site methylation and ER expression in our study suggests that methylation may play a role in regulating gene expression during resistance development.

The validation of our sequencing results was performed by pyrosequencing, focusing on key regions of interest. While expanding the number of samples and regions analyzed would have provided a more comprehensive picture, our approach was limited by time and cost considerations.

Furthermore, our data analysis suggests that methylation at CpG sites within ER promoter regions may affect transcription factor binding, particularly at sites that overlap with ETS2 and GATA3 binding sites, thereby affecting ER expression and contributing to the resistance phenotype.

Future research should aim to incorporate in vivo samples to validate these findings in a clinical context. While population cohorts such as SCAN-B and TCGA were used in our study, direct examination of *in vivo* tumor samples, particularly comparisons between primary and recurrent tumors, would provide deeper insights into the role of methylation in the clinical setting and could potentially lead to more effective therapeutic strategies if such material were available. In addition, the inclusion of a broader range of ER-positive cell lines and the investigation of more time points during resistance development could further elucidate the dynamics of methylation and resistance.

Overall, our study makes an insightful contribution to the understanding of epigenetic regulation of ER in breast cancer, particularly in the context of fulvestrant resistance. By integrating detailed methylation analyses with protein expression studies and considering potential therapeutic interventions, our research provides valuable insights that could ultimately improve treatment outcomes in ER-positive breast cancer. The methodological advantages of our approach have set the stage for future studies that could lead to further clarification of resistance mechanisms and the development of new therapeutic strategies.

Paper III The *ERBB2*-encoded miRNA miR-4728-3p regulates estrogen signaling in SK-BR-3 cells

Background

The miRNA *miR-4728*, specifically its mature form miR-4728-3p, was discovered in 2011 within an intron of the *ERBB2* gene. In HER2-positive breast cancer, miR-4728-3p is co-amplified with its host gene, potentially playing a significant role in cancer biology. In our study, we found that this miRNA regulates crucial cancer-related pathways. MicroRNAs, such as miR-4728-3p, are known to regulate gene expression post-transcriptionally, often leading to the suppression of their target genes. Given its co-amplification in HER2-positive breast cancer, miR-4728-3p is hypothesized to influence critical pathways involved in cancer progression, including those related to steroid hormone biosynthesis. The regulation of enzymes like aromatase (*CYP19A1*), which is essential for estrogen synthesis, is particularly significant in this context, as it may impact breast cancer progression and contribute to potential resistance to therapy.

The study aimed to investigate the broad effects of miR-4728-3p on gene expression and translation, with a focus on its potential as a therapeutic target in breast cancer. By understanding the specific pathways and targets affected by miR-4728-3p, the research seeks to provide insights that could lead to the development of novel treatment strategies for breast cancer, particularly in cases with ERBB2 amplification.

Results

Our study aimed to understand the role of miR-4728-3p in ERBB2-positive breast cancer cells, focusing on its effects on gene expression and translation. Therefore, we used antisense oligonucleotides (ASO), short synthetic nucleic acid strands that specifically bind to their target miRNA and block its function. By blocking miR-4728-3p in the HER2-positive, ER-negative breast cancer cell line SK-BR-3 we were interested in exploring the broader biological functions of this miRNA.

After blocking miR-4728-3p, we used polysome fractionation to separate mRNAs based on ribosomal content, which allowed us to distinguish between actively translated mRNAs (polysomally bound) and less actively translated or untranslated mRNAs (monosomally bound). Afterwards, we performed RNA-Seq analysis on these fractions as well as on total RNA to obtain a comprehensive overview of transcriptional and translational changes.

Our analysis revealed a significant upregulation of genes involved in the biosynthesis of steroid hormones, especially those related to estrogen synthesis. Among these, aromatase (*CYP19A1*), an important enzyme responsible for the conversion of testosterone to E2, showed the most distinct upregulation. This indicates that miR-4728-3p normally represses these genes and thus affects the biosynthesis of estrogen.

We performed an experiment with conditioned medium to further investigate whether the increased expression of aromatase leads to functional biological outcomes. We transferred medium from miR-4728-3p-inhibited SK-BR-3 cells to ER-positive MCF7 cells. The results indicated that the conditioned medium from miR-4728-3p-blocked cells stimulated the proliferation of MCF7 cells, suggesting increased estrogen production. This proliferative effect was removed when we added the aromatase inhibitor letrozole to the SK-BR-3 cells, supporting that the observed effect could indeed be caused by estrogen. Moreover, the proliferative effects were recovered when the conditioned medium was enriched with additional estrogen after letrozole treatment, which provides further evidence for the role of miR-4728-3p in the regulation of estrogen synthesis.

Blocking miR-4728-3p resulted in positive enrichment of targets for both miR-4728-3p and miR-21-5p. The co-enrichment of miR-21-5p targets suggests a functional interplay between the two miRNAs, whereby miR-4728-3p may increase the activity of miR-21-5p. These findings suggest a complex regulatory network in which miR-4728-3p influences not only its direct targets but also broader gene expression pathways, thereby contributing to the regulation of processes in breast cancer cells.

Beyond the role of miR-4728-3p in steroid hormone biosynthesis, our differential gene expression analysis revealed further effects of miR-4728-3p on cellular processes. We found that miR-4728-3p affects pathways involved in development, cytoskeletal organization, and mRNA metabolism. The enrichment of miR-4728-3p targets in these processes suggests that this miRNA may have far-reaching effects on cell behavior, potentially influencing cancer cell growth and invasion.

Our study highlights the regulatory role of miR-4728-3p in HER2-positive breast cancer, not only in estrogen synthesis but also in other important cellular processes. These results form the basis for further research on the potential of miR-4728-3p as a therapeutic target in breast cancer.

Limitations

The main limitation of our study is the exclusive use of the SK-BR-3 cell line for polysome fractionation. Even though this is relevant for HER2-positive, ER-negative breast cancer, it limits the extent to which our results can be generally applied. The time-consuming approach of polysome fractionation, where only six samples can be processed at once, and the high cost of RNA-Seq analysis also limited the possibility of including additional cell lines. However, extending the analysis to other breast cancer cell types would give a more comprehensive understanding of the role of miR-4728-3p.

One additional limitation is the lack of *in vivo* validation. While our experiments with conditioned media provide insights into the functional outcomes of miR-4728-3p inhibition, these results may not fully reflect the role of miR-4728-3p *in vivo*. Using *in vivo* models, such as mouse xenografts, would provide stronger validation and deeper insights into the clinical relevance of miR-4728-3p-mediated regulation of estrogen synthesis and cell proliferation.

Finally, we focused on transcriptional and translational changes without investigating possible post-translational modifications or protein interactions affected by miR-4728-3p. For example, miR-4728-3p can bind to the mRNA of a kinase, an enzyme responsible for adding phosphate groups to proteins (phosphorylation). By binding to this mRNA, miR-4728-3p can either inhibit its translation or lead to its degradation, reducing the kinase's levels in the cell. This reduction could decrease the phosphorylation of target proteins, potentially altering their activity, stability, or interactions with other proteins, which could impact key cellular pathways involved in breast cancer progression. Future studies should target these areas to gain a more global understanding of the role of miR-4728-3p in breast cancer biology.

Discussion

We have demonstrated that miR-4728-3p plays an important role in HER2-positive breast cancer by regulating estrogen synthesis and promoting cell proliferation. The involvement of this miRNA in the biosynthesis of steroid hormones emphasizes its potential impact on estrogen production. Our results suggest that miR-4728-3p normally represses genes related to estrogen synthesis, such as *CYP19A1*, indicating a potential association between this miRNA and the modulation of estrogen levels.

The use of antisense oligonucleotides to block miR-4728-3p in the SK-BR-3 cell line allowed us to observe the wider biological consequences of its blockage. RNA-Seq analysis following polysome fractionation revealed significant upregulation of genes involved in estrogen synthesis, suggesting that miR-4728-3p exerts its regulatory effects by repressing these targets. The conditioned media experiments also suggested that increased expression of aromatase led to increased estrogen production, which in turn stimulated proliferation of ER-positive MCF7 cells. This

effect was estrogen-dependent, as shown by inhibition of cell proliferation with letrozole, an aromatase inhibitor, and subsequent restoration with added estrogen.

Furthermore, blocking miR-4728-3p resulted in target enrichment for both miR-4728-3p and miR-21-5p, suggesting a functional interplay between these miRNAs. This co-enrichment suggests that miR-4728-3p may enhance the activity of miR-21-5p, contributing to a broader regulatory network that affects not only its direct targets but also other gene expression pathways. These findings point to the complex role of miR-4728-3p in regulating processes critical for breast cancer cell behavior, including development, cytoskeletal organization and mRNA metabolism.

We provide a basis for future research on the potential of miR-4728-3p to act as a therapeutic target in breast cancer. However, to completely understand the range of regulatory roles of miR-4728-3p, it is essential to include additional breast cancer cell lines, especially HER2-negative ones, and to perform *in vivo* studies that could validate these findings and evaluate their clinical relevance.

Overall, miR-4728-3p is an active regulator in HER2-positive breast cancer, affecting both estrogen synthesis and broader cellular processes. Our findings highlight the need for further research to explore its therapeutic potential and to develop targeted strategies that could improve outcomes for patients with breast cancer.

Future perspective

Our research highlights the fundamental roles of alternative splicing, miRNA-mediated regulation, and epigenetic modifications such as methylation in breast cancer progression and resistance to therapy. The study of ER alternative splicing isoforms has revealed their potential impact on tumor behavior, particularly in terms of how they influence the response to endocrine therapies such as tamoxifen and fulvestrant. While we have identified different ER isoforms in our work, the extent of their influence on breast cancer complexity and treatment resistance requires further exploration. This complexity is underscored by our ongoing project, which emphasizes the need to integrate findings from our isoform expression studies, with a particular focus on their role in fulvestrant-resistant breast cancer cells. In this study, we analyzed RNA-Seq data from the ER-positive breast cancer cell line HCC1428, including both the parental and fulvestrant-resistant (FR) lines. The resistant cell line overexpresses the full-length isoform (ERfl) at both RNA and protein levels, along with some alternative transcripts, including three truncated isoforms (Figure 22). These truncated isoforms lack either the entire LBD or additional exons.

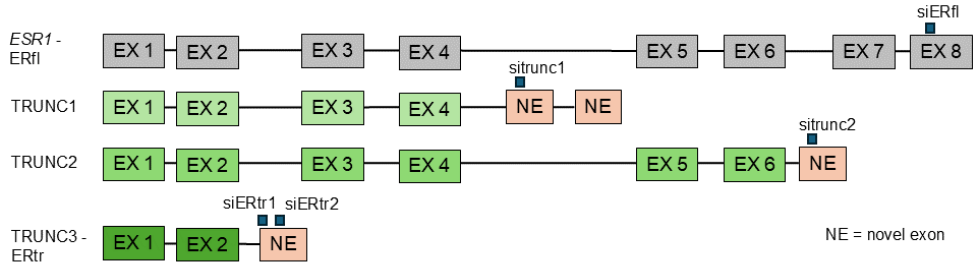

Figure 22 Splicing patterns of the truncated isoforms of the ER identified in HCC1428 breast cancer cell line, highlighted are the locations of selected siRNA binding sites that were used to target these isoforms.

All of the truncated isoforms had been detected in RNA-Seq data for breast tumors from SCAN-B but were not further studied in Balcazar Lopez *et al.* where we focused on functional studies of alternative isoforms with changes in internal exons. While the protein sequence of the full-length ER has 595 amino acids (aa) these transcripts would encode proteins of 220, 413, and 483 aa, respectively. All three isoforms have alternative last exons that begin with a cryptic splice site in intronic transposable elements which leads to the introduction of an early stop codon and a truncated protein.

The literature search indicated that the three isoforms have not been reported directly as such. However, the results of Hattori *et al*. show isoforms that are similar but not identical to ours. For example, while there is no exact match for Trunc1, they report ERαi45bL (GenBank LC120325), which utilizes the same intronic splice acceptor sites as the first novel alternative exon described in our study. The ERαi67 variant (GenBank LC120330) appears to correspond to Trunc2 (197). Similarly, H. Dotzlaw *et al*. report an mRNA transcript corresponding to our Trunc3 that contains the splice site but has a much longer last exon than what we see (292). As mentioned in the first paper, not all transcripts are consistently in the databases, so it is challenging to verify with absolute certainty that the same transcripts have been annotated.

At first, we hypothesized that the two longer isoforms were more likely to be functional due to their extended coding sequences. To test this hypothesis, we designed isoform-specific siRNAs to target the alternative exons and cloned the coding sequences of these isoforms into an expression vector (Figure 22). The presence of both isoforms was confirmed through real-time RT-PCR. However, Western blot analysis performed after siRNA-mediated knockdown revealed that neither isoform was expressed at the protein level in HCC1428 FR cells line (Figure 23). Specifically, the proteins were expected to appear as bands at approximately 46 kDa and 54 kDa, respectively. The absence of these bands suggests that, despite the transcription of these isoforms, they are not being translated into detectable protein products in this cell.

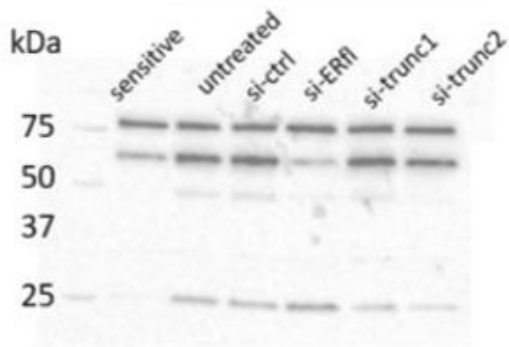

Figure 23 Expression of ER in HCC1428-FR cells, transfected with siRNA-control, siRNA -ERfl, siRNA-trunc1, and siRNA -trunc2, was determined by immunoblotting. Band intensities were normalized to total protein content per lane and expressed relative to the untreated or control.

However, there was a clearly visible band corresponding to an approximately 25 kDa protein that may represent the third, shortest truncated isoform (Figure 23). This isoform is now referred to as ERtr. To further investigate this isoform, we designed two isoform-specific siRNAs for ERtr (Figure 22, siERtr1 and siERtr2). Typically, siRNA binding sites are strategically positioned in unique regions of alternative exons or intronic sequences, allowing for isoform-specific silencing and aiding in understanding the role of these variants. However, developing effective siRNAs proved challenging due to the high copy number of the LINE1 element harboring the alternative last exon, which complicated the targeting process. By using two siRNAs with distinct sequences, we aimed to evaluate their efficacy and compare their effects on ERtr isoform expression.

We then analyzed the effect of knocking down ERfl and ERtr in HCC1428 FR cells, on resistance to fulvestrant using the alamarBlue assay to measure cell proliferation. AlamarBlue is a redox indicator that assesses cell viability by changing color in response to cellular metabolic activity, with the color change from blue to pink reflecting dye reduction by cellular reductases.

Interestingly, knockdown of both ERfl and ERtr led to significantly decreased proliferation in the absence of fulvestrant (p = 0.0002 for siERfl, p = 0.02 for siERtr1, p = 0.008 for siERtr2) (Figure 24). This finding is particularly interesting as, considering the fulvestrant resistance of the HCC1428 cell line, we initially expected that knockdown of these isoforms would primarily affect fulvestrant resistance.

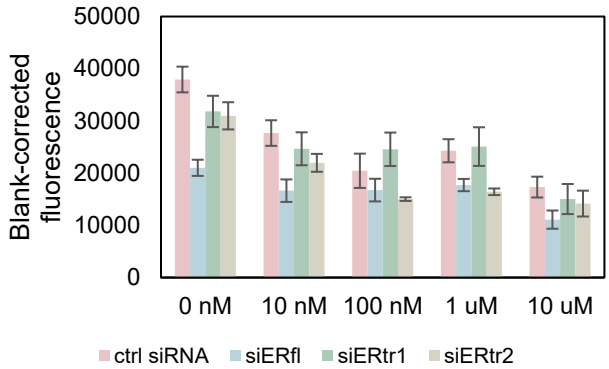

Figure 24 AlamarBlue assay measuring the proliferation of HCC1428-FR cells after siRNA knockdown of ERfl and ERtr isoforms, with absolute emission values recorded at increasing concentrations of fulvestrant (blank corrected).

However, the results were inconsistent between repeated experiments, sometimes showing increased toxicity of fulvestrant upon knockdown and sometimes not, which makes it difficult to draw any definitive conclusions. Since ERfl is the target of fulvestrant, we would expect to see an added effect (i.e., increased sensitivity) when fulvestrant is used in combination with siRNA.

Western blotting of cells transfected with the different siRNAs confirmed the efficient knockdown after 48 hours. In contrast to our expectations, the analysis showed that knockdown of ERtr not only reduced ERtr expression, but also led to a concomitant decrease in ERfl expression. Furthermore, knockdown of ERfl led to an unexpected increase in the expression of the truncated isoform ERtr. We had assumed that each siRNA would specifically affect only the target isoform without affecting the expression of other isoforms (Figure 25).

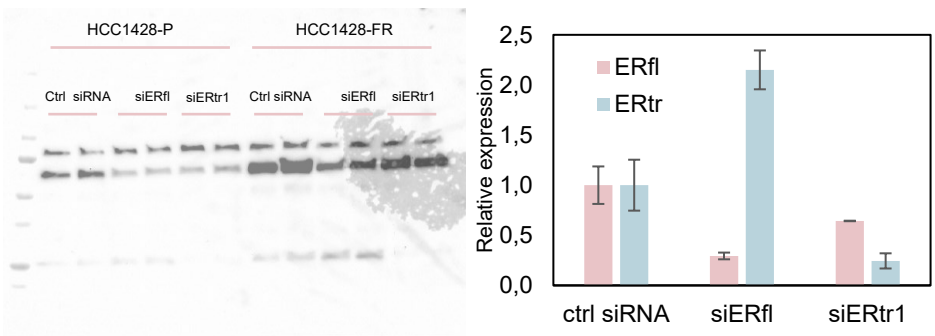

Figure 25 Expression of ER in HCC1428-P and HCC1428-FR cells transfected with siRNA-control, siERfl, and siERtr1 was determined by immunoblotting (n=2). Band intensities were normalized to total protein content per lane and expressed relative to the control.

To validate this result, we also measured the mRNA levels of the full-length and truncated isoforms in cells transfected with siRNAs at 24 and 48 hours using real-time RT-PCR. This analysis confirmed that the observed effects were already present at the RNA level (Figure 26).

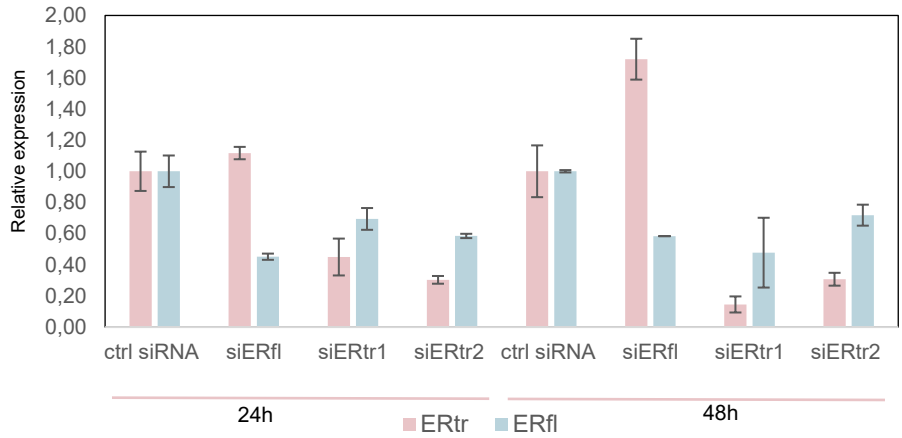

Figure 26 Real-time RT-PCR of ERfl and ERtr in cells transfected with siRNA-control, siERfl, siERtr1, and siERtr2 for 24 and 48 hours, normalized against *MRPL19*.

The effects observed from siRNAs on the non-targeted isoform are difficult to interpret, raising the question of whether these effects are caused by experimental artifacts or reflect true biological processes. Since RNA interference (RNAi) is a cytoplasmic process, it is unclear how siRNA directed against ERtr targeting an intronic region in the canonical transcript could affect the expression of ERfl.

The siRNA against ERfl targets a region downstream of the end of the truncated isoform. Still, the downregulation of the full-length protein observed after transfection with siRNA against ERtr could possibly explain the reduced proliferation observed in the alamarBlue assay.

These unexpected results suggest that there may be more complex interactions between the isoforms than originally thought. It is possible that the silencing of ERtr could affect the stability or translation of ERfl through indirect mechanisms, such as changes in cellular signaling pathways or protein interactions.

Cloning and overexpression of ERtr

Since functional studies of ERtr with siRNAs were problematic, we decided to test for overexpression by cloning the coding sequence of the isoform into an expression vector. We used restriction enzymes to cut both the ERtr coding sequence and the expression vector at compatible sites so that we could efficiently insert the isoform into the vector. The recombinant plasmid was then confirmed by Sanger sequencing.

We transfected the ERtr vector into MCF7 cells, which do not naturally express ERtr, and into HCC1428-FR cells, which express it endogenously. We also included pEGFP-C1, a control vector that does not express ERfl or ERtr, as well as a plasmid expressing ERfl. Western blot analysis confirmed successful protein expression from the plasmid, with greater overexpression observed in MCF7 cells compared to HCC1428-FR cells (Figure 27).

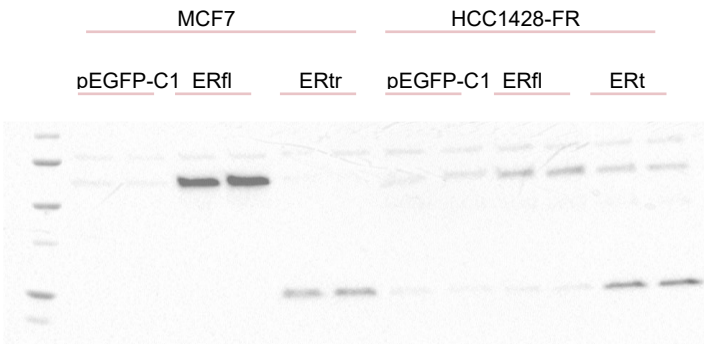

Figure 27 Expression of pEGFP-C1, ERfl and ERtr in MCF7 and HCC1428-FR cells, n=2, after 24h, was determined by immunoblotting. Band intensities were normalized to total protein content per lane and expressed relative to peGFP-C1 per cell line.

Overexpression of ERtr in MCF7 cells resulted in a lower level of ERfl ($p = 0.07$), while in HCC1428-FR cells it resulted in a higher level of ERfl ($p = 0.07$). However, neither result reached statistical significance. In contrast overexpression of ERfl in MCF7 cells show no visible ERtr protein, as expected due to the absence of endogenous ERtr (Figure 27, Figure 28).

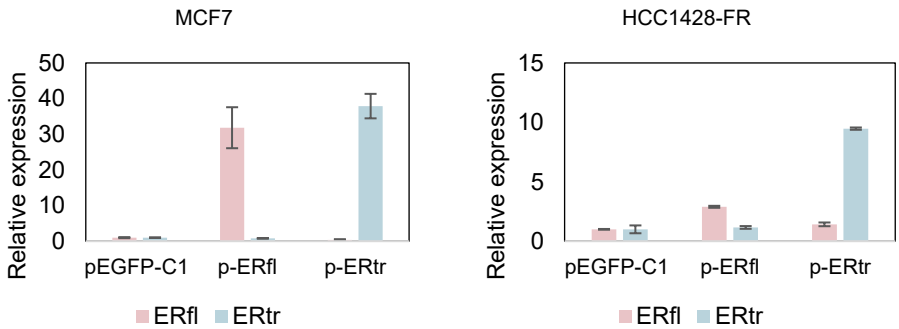

Figure 28 Quantification of the expression of pEGFP-C1, ERfl and ERtr in MCF7 (left) and HCC1428-FR cells (right), n=2, after 24 h, was determined by immunoblotting. Band intensities were normalized to total protein content per lane and expressed relative to pEGFP-C1 per cell line, p=plasmid.

This was followed by the measurement of mRNA levels using real-time RT-PCR, which showed that overexpression of either isoform had no effect on the other (Figure 29). For this experiment, we used primers targeting the 3' UTR to measure the effects on the expression of endogenous ERfl.

The effects observed in the siRNA knockdown experiment were inconclusive, since they mostly looked like experimental artifacts, but could possibly also represent a biological phenomenon. Compared to the siRNA knockdown experiment, we did not observe the same effects in the overexpression experiment.

In the case of HCC1428-FR, we know now that the expression levels are higher after 48 hours than after 24 hours. The western blot only included a 24-hour time point, therefore this experiment should be repeated to allow a more accurate conclusion and comparison (Figure 27).

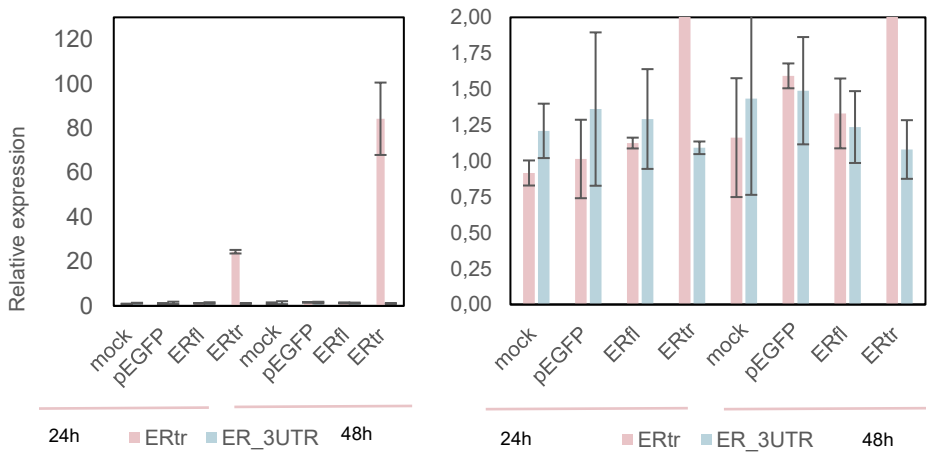

Figure 29 Real-time RT-PCR was performed on HCC1428-FR cells transfected with ERfl, ERtr, pEGFP-C1, and a mock control at 24 and 48 hours, normalized to *MRPL19*. The left panel shows the full range of relative expressions, while the right panel provides a zoomed-in view since the high expression of ERtr at 48 hours.

Next, we repeated the proliferation experiment using alamarBlue with HCC1428 parental cells and FR cells overexpressing ERtr. The two cell lines showed different behavior due to their different sensitivity to fulvestrant. In particular, the plasmid itself was toxic to the parental cells, resulting in only a modest additional effect on cell survival at the highest fulvestrant doses used. Surprisingly, no protective effect of ERfl or ERtr was observed after a 4-day treatment with different fulvestrant concentrations (Figure 30).

However, this experiment was conducted to explore the function of ERtr without the specific expectation that it would increase resistance to fulvestrant, although it would have been quite remarkable if it had. The ERfl plasmid was included as a form of a positive control based on the assumption that its higher expression would lead to increased resistance due to its role as a target of fulvestrant. The absence of the expected effects seems to indicate that there may have been technical problems with the experiment, making it difficult to draw definitive conclusions about the role of ERtr.

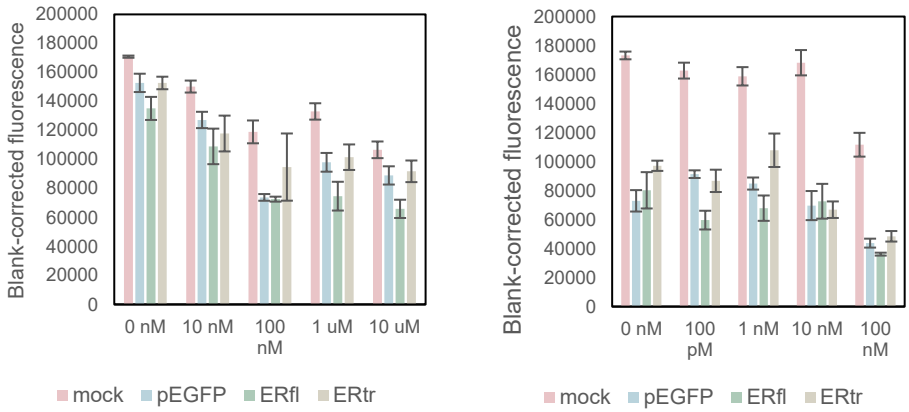

Figure 30 AlamarBlue assay measuring cell proliferation in HCC1428-FR cells (left) and HCC1428 parental cells (right), overexpressed with pEGFP-C1, ERfl, ERtr, or a mock control. Both cell lines were exposed to increasing concentrations of fulvestrant for 4 days.

We wanted to investigate the effects of ER degradation to understand how it is regulated. We performed an experiment to measure the degradation of endogenous ERfl after treatment with fulvestrant in HCC1428-FR cells with and without overexpression of ERtr by using two different time points (6 hours and 24 hours). After 6 hours, cells were treated with DMSO, 100 nM or 10 µM fulvestrant to observe the behavior of ERtr at different concentrations. After 24 hours, the cells were treated with DMSO and 100 nM fulvestrant (Figure 31). This approach allowed us to evaluate the effects of different fulvestrant concentrations and time points on ER degradation in the presence or absence of ERtr overexpression.

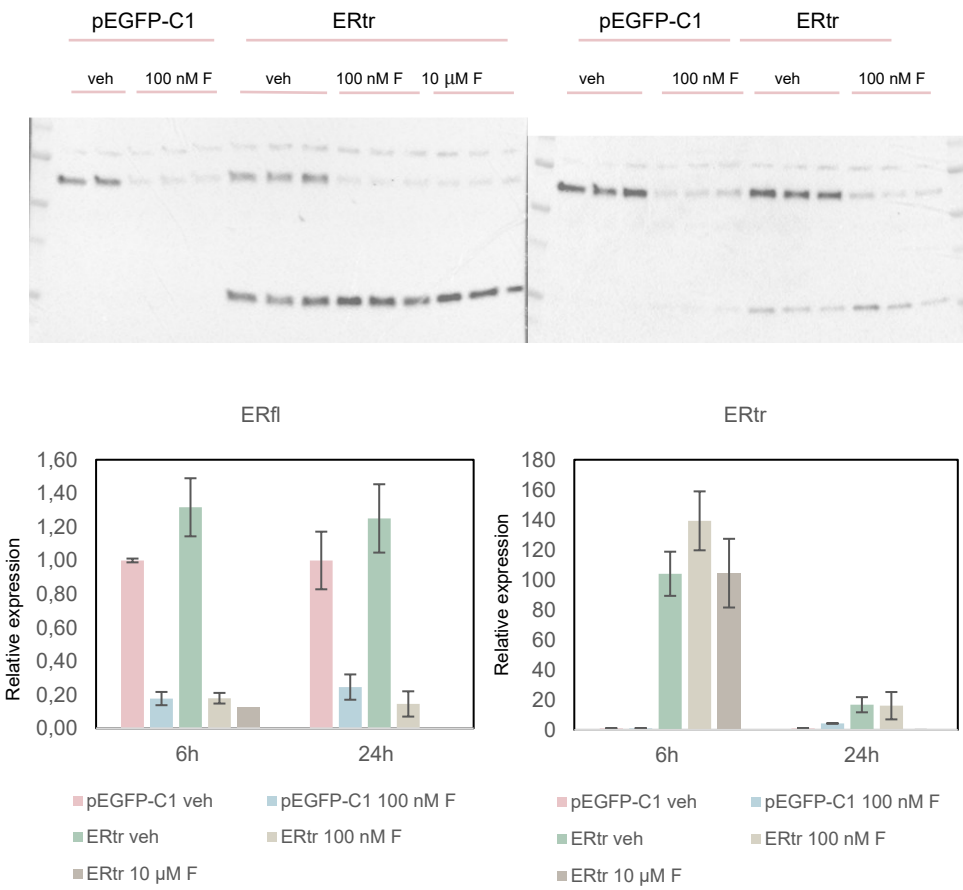

Figure 31 Expression of ERfl and ERtr in HCC1428-FR cells after transfection with pEGFP-C1 and ERtr plasmids and treatment with DMSO, 100 nM fulvestrant, or 10 µM fulvestrant at 6 hours was determined by immunoblotting (left); expression of ERfl and ERtr after transfection and treatment with DMSO or 100 nM fulvestrant at 24 hours was determined by immunoblotting (right). Band intensities were normalized to total protein content per lane and expressed relative to the pEGFP-C1 veh per cell line / blot. F= fulvestrant.

Fulvestrant led to strongly decreased levels of ERfl at both 6 and 24 h and overexpression of ERtr did not mitigate this effect. Different concentrations of fulvestrant at the 6 hour time point showed no differences. We have also attempted to perform this experiment in the parental HCC1428 cells, but the western blot failed for unknown technical reasons.

Finally, we investigated the effect of ERtr on the regulation of a luciferase reporter gene with a promoter containing three EREs. The experiment in parental HCC1428 cells was difficult to interpret, as it unexpectedly did not produce any response to

estradiol (E2). However, overexpression of ERtr in HCC1428-FR cells resulted in higher expression of the reporter gene in response to E2 (Figure 32).

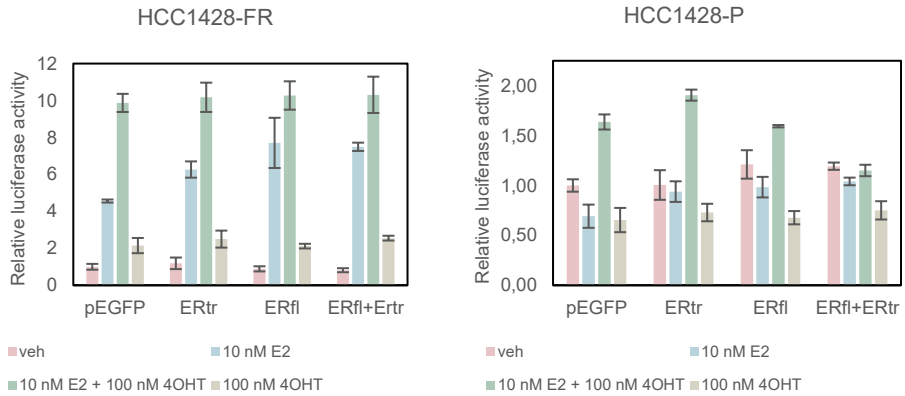

Figure 32 Luciferase assay left in HCC1428-FR and right in HCC1428-P, overexpressing ERfl and ERtr and combination, treated with vehicle (EtOH), 10 nM E2, 10 nM E2 + 100 nM 4OHT, 100 nM 4OHT, n=3, relative to pEGFP-C1 veh.

In addition to HCC1428, we also tested the plasmid in two other cell lines, HepG2 and MCF7, as both have previously provided reliable and reproducible results for this assay. Specifically, we tested the ER-negative hepatoma-derived cell line HepG2 and the ER-positive breast cancer cell line MCF7 (Figure 33).

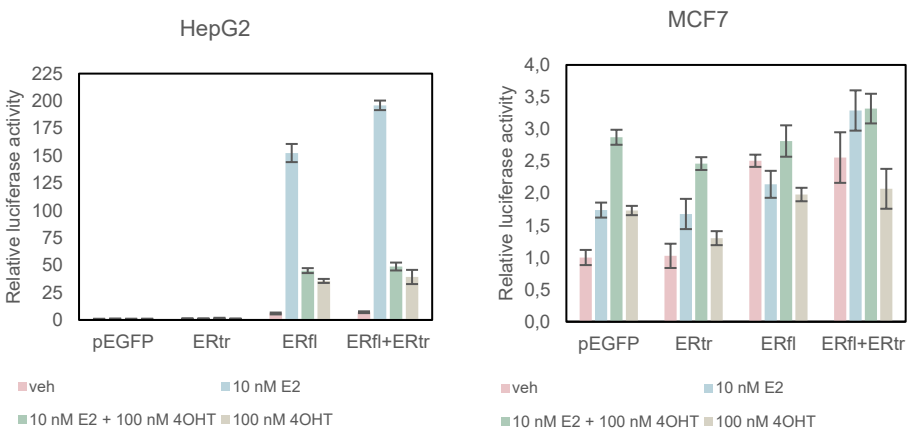

Figure 33 The luciferase assay was conducted in HepG2 cells (left) and MCF7 cells (right), each overexpressing peGFP-C1, ERfl, ERtr, or a combination of ERfl/ERtr. Treated with vehicle (EtOH), 10 nM E2, 100 nM 4OHT+ 10 nM E2, 100 nM 4OHT, n=3, relative to pEGFP-C1 veh.

In both cell lines, co-transfection of ERfl and ERtr led to significantly higher expression of the reporter gene in response to E2 compared to transfection of either isoform alone. Similar to the dominant-negative ERΔexon3 isoform (201,293), which also lacks most the DNA-binding domain, we did not expect ERtr to have any transcriptional activity. More careful examination of the sequence revealed that there is one intact zinc finger that could potentially interact with DNA and affect transcriptional activation. Interestingly, some of our luciferase assay experiments show a very modest increase in reporter gene activity above background levels (~20%) for the ERtr and ERΔexon3 isoforms in the ER-negative HepG2 cell line.

The results presented are preliminary, indicating potential biological effects of ERtr that require further validation and optimization. Insights into the function of this truncated isoform will help us to better understand its potential role in the development of breast cancer. These results need to be confirmed through repeated experiments under optimized conditions. For example, subcellular fractionation could be employed to elucidate the localization and function of these isoforms. Proliferation experiments should be repeated with improved conditions, and degradation studies should be conducted using multiple concentrations. Additionally, RNA-Seq could be utilized to compare the activation of target genes between cells expressing only full-length ER and those expressing ER alongside the truncated isoform.

Parallel to our studies on splice variants, our investigation of miR-4728-3p has highlighted its role in modulating estrogen receptor pathways independently of its well-known protein-coding host gene. This miRNA is implicated in regulating both ER and aromatase, suggesting its potential as a therapeutic target for altering estrogen synthesis in HER2-positive breast cancer. Furthermore, the dynamic methylation patterns of *ESR1* promoter regions across different breast cancer cell lines and resistance conditions introduce another layer of complexity. Our research indicates that methylation patterns are not only cell-specific (considering that tissues are composed of a mix of cell types) but also vary with resistance status. These variations influence ER expression and may impact the efficacy of epigenetic therapies.

Future research should focus on elucidating the detailed mechanisms by which these isoforms and epigenetic factors contribute to treatment resistance. This includes functional studies of alternative ER isoforms, investigation into the role of miRNAs like miR-4728-3p in estrogen receptor signaling, and comprehensive analysis of methylation changes in breast cancer subtypes. Such studies could lead to the identification of novel biomarkers for resistance and the development of more precise therapeutic approaches, including the potential reactivation of ER expression through targeted epigenetic modifications.

Conclusion

The main aim of this work was to investigate the complex regulatory mechanisms that modulate ER signaling in breast cancer, with a particular focus on alternative splicing, differential promoter usage, promoter methylation and miRNAs. Through a combination of *in vitro* functional assays and high-throughput sequencing analyses, our research has provided critical insights into how these regulatory elements contribute to the complexity of ER expression and function and their impact on breast cancer progression and therapy resistance.

Our studies have identified several ER isoforms resulting from alternative splicing and alternative promoter usage that respond differently to endocrine therapies such as tamoxifen and fulvestrant. These isoforms may play a role in mediating resistance to standard treatments, emphasizing the importance of isoform-specific analyses in clinical diagnostics. In addition, we have demonstrated that promoter methylation, by modulating ER expression, can influence ER activity and therapeutic response. This may have important insights for the development of epigenetic therapies aimed at modulating ER expression through targeted methylation sites.

In addition, our study of miR-4728-3p has shown that it has a regulatory effect on ER signaling independent of canonical protein-coding sequences, providing evidence that miRNAs may serve as both biomarkers and therapeutic targets in ER-positive breast cancer.

Altogether, these findings expand our understanding of the molecular mechanisms underlying ER signaling in breast cancer and prepare the long way for the integration of these findings into clinical practice. By considering the specific ER isoforms, methylation status and miRNA profiles of individual tumors, there is an opportunity to refine patient stratification and tailor therapeutic approaches to ultimately improve the clinical outcomes of breast cancer treatment.

Ethical considerations

The studies presented in this paper involve the use of RNA-Seq, DNA-Seq and other raw genomic data obtained from patients' tumors, normal tissue and breast cancer cell lines, which raises important ethical considerations related to patient privacy, informed consent and data security. The SCAN-B initiative, which focuses on comprehensive genomic characterization of breast cancer cases in southern Sweden, and TCGA provided the primary datasets used in this study. Both SCAN-B and TCGA adhere to strict ethical guidelines for the protection of patient data.

In SCAN-B, the ethical considerations revolve around the use of patient samples collected from newly diagnosed breast cancer patients at nine hospitals in southern Sweden. These samples include both tumor and normal tissue. The study strictly follows the ethical principles of the Declaration of Helsinki and has been approved by the regional ethics committee in Lund and other relevant authorities. Informed consent will be obtained from all participants to ensure that they know how their samples and data will be used. Given the sensitive nature of the genomic data, measures will be taken to ensure that the data is stored securely and that only authorized personnel have access to it. The focus of the SCAN-B study on a specific population provides detailed insights into breast cancer biology in this region, but also requires careful consideration of the generalizability of the results.

The handling of TCGA data is also subject to a high degree of ethical control. Although TCGA's processed data is publicly available, access to sensitive raw data such as RNA-Seq, DNA-Seq and Whole Genome Sequencing (WGS) data, including tumor and normal tissue data, is tightly controlled through a project-specific approval process. Researchers must obtain permission to access this data and may only use it for the approved purposes, with an obligation to delete the data upon completion of the project. TCGA uses a “whole consent” model in which patients provide broad consent for the use of their anonymized genetic data in any cancer research approved by TCGA. While this approach facilitates wide-ranging research, it raises questions about the specificity and scope of patient consent.

Ethical considerations also apply to the use of human breast cancer cell lines in our studies that involved DNA sequencing. Although the data from these cell lines are not as sensitive as patient data, they still raise important ethical issues. Many widely used cell lines were produced before current standards for informed consent were

introduced, requiring reflection on the ethical implications of using biological material for which proper informed consent may not have been obtained.

The research conducted in this thesis follows strict ethical guidelines to ensure the protection of patient data and respect for patient autonomy. The use of SCAN-B and TCGA data as well as DNA sequencing of both normal tissue and breast cancer cell lines emphasizes the need for constant ethical vigilance in the handling and use of sensitive genetic information.

Acknowledgements

Over the past four years, since my arrival in Sweden, I have been surrounded by an incredible network of support, guidance, and friendship. This PhD journey has been shaped by the invaluable contributions of so many wonderful people, and I would like to take this opportunity to express my deepest gratitude to each and every one of you.

First and foremost, I would like to express my deepest gratitude to **Helena**. Your belief in me—both in accepting me as a PhD student and in introducing me to the challenging yet fascinating world of breast cancer research—has had a profound impact on me. Your extraordinary scientific expertise, optimism, and unwavering support have been foundational pillars throughout my journey. From international conferences in Heidelberg and San Diego to my time in Japan, your encouragement has continually inspired me to push my boundaries. This dissertation, and indeed my entire experience, would not have been possible without your guidance. I am so grateful for all that you have taught me.

To the research group—**Völundur, Mirjam, Izabela, Robin**, and all the students who have come and gone over the years. Your collaboration, friendship, and shared knowledge have been invaluable to my development, both as a researcher and as a person. The times we worked, laughed, and learned together will always hold a special place in my heart.

To my co-supervisors, **Håkan, Anders**, and **Johann**, thank you for your insightful suggestions and advice, which I always appreciated. Even though I may not have sought your help as often as I could have, your contributions were essential, and I am grateful for your continued support throughout my PhD.

A special thanks to **Susanne André** for your excellent administrative support. You were always there to ensure everything ran smoothly and solved every problem with ease. I am truly grateful for your kindness, patience, and efficiency.

To my friends at work: you were the life and soul of this journey. Our laughter, conversations, and the support we shared made even the most difficult moments more manageable. I am so grateful for every smile, every joke, and every word of encouragement.

Kamila, I cannot thank you enough for your constant and unwavering support. You were always there to answer my questions, share your expertise, and give me advice

on everything from lab techniques to life in general. The generosity with which you have shared your time and knowledge has been invaluable, and I am forever grateful.

Bengt, your willingness to help whenever I encountered a problem—whether it was an issue with reagents or a general question—was always greatly appreciated. Thank you for your kindness and for always being there when I needed you.

Pontus, thank you so much for your support and for stepping in when we needed help like the camera work. Your assistance was indispensable, and I am so grateful for your friendship.

Thank you to **Deborah** and **Elsa** for always offering your help and advice whenever I had a question.

To **Katja, Lena, Linnéa, Pei, Sergii, Jacob**, **Anna E**. and **Anna K.** and the TCR members—**Lexy, Lennart, Suze, Inaki**—thank you for the inspiring conversations and the wonderful camaraderie we shared over lunch and fika. You all made these moments something to look forward to, and I cherish the memories we created together.

Wonde, I am so grateful for your constant companionship and the friendship we have built. I have learned so much from you, and I appreciate your support.

Ahmed, our conversations have always been a pleasure, and your advice has meant a lot to me. Thank you for your guidance and wisdom.

Matteo, even though you are no longer at LU, I want to express my gratitude for everything you have taught me. Your mentorship was truly invaluable, and I appreciate the time and effort you dedicated to helping me grow.

Rosalia, thank you for being such a helpful and supportive friend. From our Swedish lessons to our countless conversations, your friendship has been a source of strength for me.

Carina, your constant support and openness have been essential throughout my PhD and in general the past 4 years. Your emotional support and friendship have truly made a difference, and I am so grateful for everything we have experienced together. At the same time, I would also like to thank **Ricardo** for your commitment and the great squash matches!

To the **MV oncology department**—especially **Karin, Eva**, **Ingrid**, **Frida**, **Jeanette** and **Tessa**—thank you for always being there when I had questions or needed help. Your assistance was indispensable, and I really appreciate the support you provided.

A big thank you to **MedBioInfo** for providing excellent courses, organizing annual meetings, and offering financial support. The opportunity to connect with new

people and learn from other specialties has enriched my research experience tremendously.

Thanks also to **MDR** for the opportunity to represent medical doctoral students at the Faculty of Medicine, which was an enriching experience.

To the organizers of the **MentLife** program and my mentor, Helena Pettersson, thank you for the insightful conversations and strategic advice that helped ease my concerns about the future. Your guidance has been invaluable.

I am also very grateful for the financial support I received during my PhD, particularly for conferences and my stay in Japan in 2023. My special thanks to **JAP, FUN,** and **Cancerfonden** for making these incredible opportunities possible.

To my friends in Germany—**Mayline, Sandra, Katja, Resi, Steffi,** and **Melly**—thank you for your constant support, enthusiasm, and encouragement throughout this journey. Your friendship means the world to me.

Vor allem möchte ich **meiner Familie** danken: Eure Unterstützung und Ermutigung haben mir viel bedeutet. Ihr habt mich in jeder Phase begleitet und mir stets den Rücken gestärkt. Vielen Dank, dass ihr immer für mich da wart. Ich hab euch ganz doll lieb.

And of course, to **Max**, the heart of my journey. We have been an unbreakable team since 2017, and I wouldn't trade that for anything. You bring laughter and calm even in the most stressful moments, and your unwavering patience, reliability, and humor are my greatest source of strength. None of this would have been possible without you. Thank you from the bottom of my heart for everything.

References

1. Molecular Biology of the Cell, Sixth edition. 6. Garland Science; 2014.

2. Martelotto LG, Ng CK, Piscuoglio S, Weigelt B, Reis-Filho JS. Breast cancer intra-tumor heterogeneity. Breast Cancer Research. 2014 May 20;16(3):210.

3. Gerlinger M, Rowan AJ, Horswell S, Math M, Larkin J, Endesfelder D, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med. 2012 Mar 8;366(10):883–92.

4. Hanahan D, Weinberg RA. The Hallmarks of Cancer. Cell. 2000 Jan 7;100(1):57–70.

5. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. Cell. 2011 Mar 4;144(5):646–74.

6. Hanahan D. Hallmarks of Cancer: New Dimensions. Cancer Discov. 2022 Jan;12(1):31–46.

7. Figueroa JD, Pfeiffer RM, Patel DA, Linville L, Brinton LA, Gierach GL, et al. Terminal Duct Lobular Unit Involution of the Normal Breast: Implications for Breast Cancer Etiology. J Natl Cancer Inst. 2014 Oct 1;106(10):dju286.

8. Zhao H, Zhou L, Shangguan AJ, Bulun SE. Aromatase expression and regulation in breast and endometrial cancer. J Mol Endocrinol. 2016 Jul;57(1):R19-33.

9. Grodin JM, Siiteri PK, MacDonald PC. Source of estrogen production in postmenopausal women. J Clin Endocrinol Metab. 1973 Feb;36(2):207–14.

10. Brown M, Cato L, Jeselsohn R. Chapter 29 - Hormone-Responsive Cancers. In: Strauss JF, Barbieri RL, editors. Yen and Jaffe's Reproductive Endocrinology (Eighth Edition) [Internet]. Philadelphia: Elsevier; 2019 [cited 2024 Aug 8]. p. 717-741.e8. Available from: https://www.sciencedirect.com/science/article/pii/B9780323479127000299

11. Simpson ER. Sources of estrogen and their importance. J Steroid Biochem Mol Biol. 2003 Sep;86(3–5):225–30.

12. Russo J, Hasan Lareef M, Balogh G, Guo S, Russo IH. Estrogen and its metabolites are carcinogenic agents in human breast epithelial cells. J Steroid Biochem Mol Biol. 2003 Oct;87(1):1–25.

13. Yue W, Santen RJ, Wang JP, Li Y, Verderame MF, Bocchinfuso WP, et al. Genotoxic metabolites of estradiol in breast: potential mechanism of estradiol induced carcinogenesis. J Steroid Biochem Mol Biol. 2003 Sep;86(3–5):477–86.

14. Russo J, Hu YF, Tahin Q, Mihaila D, Slater C, Lareef MH, et al. Carcinogenicity of estrogens in human breast epithelial cells. APMIS. 2001;109(1):39–52.

15. Clemons M, Goss P. Estrogen and the Risk of Breast Cancer. New England Journal of Medicine. 2001 Jan 25;344(4):276–85.

16. Bulun SE, Lin Z, Imir G, Amin S, Demura M, Yilmaz B, et al. Regulation of aromatase expression in estrogen-responsive breast and uterine disease: from bench to treatment. Pharmacol Rev. 2005 Sep;57(3):359–83.

17. Taraborrelli S. Physiology, production and action of progesterone. Acta Obstet Gynecol Scand. 2015 Nov;94 Suppl 161:8–16.

18. Beatson GT. On the Treatment of Inoperable Cases of Carcinoma of the Mamma: Suggestions for a New Method of Treatment, with Illustrative Cases. Trans Med Chir Soc Edinb. 1896;15:153–79.

19. Socialstyrelsen. Statistik om cancer [Internet]. 2024 Aug. Available from: https://www.socialstyrelsen.se/statistik-och data/statistik/alla-statistikamnen/cancer/

20. socialstyrelsen. Statistics on Cancer Incidence 2019 [Internet]. 2020 Dec. Available from: https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/statistik/2020-12-7133.pdf

21. van Seijen M, Lips EH, Thompson AM, Nik-Zainal S, Futreal A, Hwang ES, et al. Ductal carcinoma in situ: to treat or not to treat, that is the question. Br J Cancer. 2019 Aug;121(4):285–92.

22. Tan PH, Ellis I, Allison K, Brogi E, Fox SB, Lakhani S, et al. The 2019 World Health Organization classification of tumours of the breast. Histopathology. 2020 Aug;77(2):181–5.

23. Santen RJ, Boyd NF, Chlebowski RT, Cummings S, Cuzick J, Dowsett M, et al. Critical assessment of new risk factors for breast cancer: considerations for

development of an improved risk prediction model. Endocr Relat Cancer. 2007 Jun;14(2):169–87.

24. Opdahl S, Alsaker MDK, Janszky I, Romundstad PR, Vatten LJ. Joint effects of nulliparity and other breast cancer risk factors. Br J Cancer. 2011 Aug 23;105(5):731–6.

25. van den Brandt PA, Ziegler RG, Wang M, Hou T, Li R, Adami HO, et al. Body size and weight change over adulthood and risk of breast cancer by menopausal and hormone receptor status: a pooled analysis of 20 prospective cohort studies. Eur J Epidemiol. 2021 Jan;36(1):37–55.

26. Bjørge T, Häggström C, Ghaderi S, Nagel G, Manjer J, Tretli S, et al. BMI and weight changes and risk of obesity-related cancers: a pooled European cohort study. Int J Epidemiol. 2019 Dec 1;48(6):1872–85.

27. Jones ME, Schoemaker MJ, Wright LB, Ashworth A, Swerdlow AJ. Smoking and risk of breast cancer in the Generations Study cohort. Breast Cancer Res. 2017 Nov 22;19(1):118.

28. Byng JW, Boyd NF, Fishell E, Jong RA, Yaffe MJ. The quantitative analysis of mammographic densities. Phys Med Biol. 1994 Oct;39(10):1629–38.

29. Bhardwaj P, Au CC, Benito-Martin A, Ladumor H, Oshchepkova S, Moges R, et al. Estrogens and breast cancer: Mechanisms involved in obesity-related development, growth and progression. The Journal of Steroid Biochemistry and Molecular Biology. 2019 May 1;189:161–70.

30. White AJ, Nichols HB, Bradshaw PT, Sandler DP. Overall and central adiposity and breast cancer risk in the Sister Study. Cancer. 2015 Oct 15;121(20):3700–8.

31. Cleary MP, Grossmann ME. Minireview: Obesity and breast cancer: the estrogen connection. Endocrinology. 2009 Jun;150(6):2537–42.

32. Gallagher EJ, LeRoith D. Obesity and Diabetes: The Increased Risk of Cancer and Cancer-Related Mortality. Physiol Rev. 2015 Jul;95(3):727–48.

33. Roberts DL, Dive C, Renehan AG. Biological mechanisms linking obesity and cancer risk: new perspectives. Annu Rev Med. 2010;61:301–16.

34. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians. 2018;68(6):394–424.

35. Liu L, Hao X, Song Z, Zhi X, Zhang S, Zhang J. Correlation between family history and characteristics of breast cancer. Sci Rep. 2021 Mar 18;11(1):6360.

36. Claus EB, Schildkraut JM, Thompson WD, Risch NJ. The genetic attributable risk of breast and ovarian cancer. Cancer. 1996 Jun 1;77(11):2318–24.

37. Hu C, Hart SN, Gnanaolivu R, Huang H, Lee KY, Na J, et al. A Population-Based Study of Genes Previously Implicated in Breast Cancer. N Engl J Med. 2021 Feb 4;384(5):440–51.

38. Breast Cancer Association Consortium, Dorling L, Carvalho S, Allen J, González-Neira A, Luccarini C, et al. Breast Cancer Risk Genes - Association Analysis in More than 113,000 Women. N Engl J Med. 2021 Feb 4;384(5):428–39.

39. Balmaña J, Díez O, Rubio IT, Cardoso F. BRCA in breast cancer: ESMO Clinical Practice Guidelines. Annals of Oncology. 2011 Sep 1;22:vi31–4.

40. Kuchenbaecker KB, Hopper JL, Barnes DR, Phillips KA, Mooij TM, Roos-Blom MJ, et al. Risks of Breast, Ovarian, and Contralateral Breast Cancer for BRCA1 and BRCA2 Mutation Carriers. JAMA. 2017 Jun 20;317(23):2402–16.

41. Güth U, Huang DJ, Huber M, Schötzau A, Wruk D, Holzgreve W, et al. Tumor size and detection in breast cancer: Self-examination and clinical breast examination are at their limit. Cancer Detect Prev. 2008;32(3):224–8.

42. Del Monte U. Does the cell number 10(9) still really fit one gram of tumor tissue? Cell Cycle. 2009 Feb 1;8(3):505–6.

43. Reginoala Cancercentrum i Samverkan. Bröstcancer Nationellt vårdprogram [Internet]. 2022 May. Report No.: 4.0. Available from: https://www.swebcg.se/wp-content/uploads/2022/06/nationellt-vardprogram-brostcancer_220510.pdf

44. Slaoui M, Bauchet AL, Fiette L. Tissue Sampling and Processing for Histopathology Evaluation. In: Gautier JC, editor. Drug Safety Evaluation: Methods and Protocols [Internet]. New York, NY: Springer; 2017 [cited 2024 Aug 9]. p. 101–14. Available from: https://doi.org/10.1007/978-1-4939-7172-5_4

45. O'Sullivan B, Brierley J, Byrd D, Bosman F, Kehoe S, Kossary C, et al. The TNM classification of malignant tumours—towards common understanding and reasonable expectations. Lancet Oncol. 2017 Jul;18(7):849–51.

46. Amin MB, Greene FL, Edge SB, Compton CC, Gershenwald JE, Brookland RK, et al. The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. CA Cancer J Clin. 2017 Mar;67(2):93–9.

47. Liang Y, Zhang H, Song X, Yang Q. Metastatic heterogeneity of breast cancer: Molecular mechanism and potential therapeutic targets. Semin Cancer Biol. 2020 Feb;60:14–27.

48. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med. 2004 Dec 30;351(27):2817–26.

49. Cardoso F, van't Veer LJ, Bogaerts J, Slaets L, Viale G, Delaloge S, et al. 70-Gene Signature as an Aid to Treatment Decisions in Early-Stage Breast Cancer. N Engl J Med. 2016 Aug 25;375(8):717–29.

50. Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature. 2000 Aug;406(6797):747–52.

51. Zhang Y, Schnabel CA, Schroeder BE, Jerevall PL, Jankowitz RC, Fornander T, et al. Breast cancer index identifies early-stage estrogen receptor-positive breast cancer patients at risk for early- and late-distant recurrence. Clin Cancer Res. 2013 Aug 1;19(15):4196–205.

52. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology. 1991 Nov;19(5):403–10.

53. Henson DE, Ries L, Freedman LS, Carriaga M. Relationship among outcome, stage of disease, and histologic grade for 22,616 cases of breast cancer. The basis for a prognostic index. Cancer. 1991 Nov 15;68(10):2142–9.

54. Iwamoto T, Booser D, Valero V, Murray JL, Koenig K, Esteva FJ, et al. Estrogen Receptor (ER) mRNA and ER-Related Gene Expression in Breast Cancers That Are 1% to 10% ER-Positive by Immunohistochemistry. JCO. 2012 Mar;30(7):729–34.

55. Duffy MJ, Harbeck N, Nap M, Molina R, Nicolini A, Senkus E, et al. Clinical use of biomarkers in breast cancer: Updated guidelines from the European Group on Tumor Markers (EGTM). Eur J Cancer. 2017 Apr;75:284–98.

56. Early Breast Cancer Trialists' Collaborative Group (EBCTCG), Davies C, Godwin J, Gray R, Clarke M, Cutter D, et al. Relevance of breast cancer

hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. Lancet. 2011 Aug 27;378(9793):771–84.

57. Foulkes WD, Smith IE, Reis-Filho JS. Triple-negative breast cancer. N Engl J Med. 2010 Nov 11;363(20):1938–48.

58. Gerdes J, Lemke H, Baisch H, Wacker HH, Schwab U, Stein H. Cell cycle analysis of a cell proliferation-associated human nuclear antigen defined by the monoclonal antibody Ki-67. J Immunol. 1984 Oct;133(4):1710–5.

59. Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A. 2001 Sep 11;98(19):10869–74.

60. Kaufmann M, Pusztai L, Biedenkopf Expert Panel Members. Use of standard markers and incorporation of molecular markers into breast cancer therapy: Consensus recommendations from an International Expert Panel. Cancer. 2011 Apr 15;117(8):1575–82.

61. Fisher B, Anderson S, Bryant J, Margolese RG, Deutsch M, Fisher ER, et al. Twenty-year follow-up of a randomized trial comparing total mastectomy, lumpectomy, and lumpectomy plus irradiation for the treatment of invasive breast cancer. N Engl J Med. 2002 Oct 17;347(16):1233–41.

62. EBCTCG (Early Breast Cancer Trialists' Collaborative Group), McGale P, Taylor C, Correa C, Cutter D, Duane F, et al. Effect of radiotherapy after mastectomy and axillary surgery on 10-year recurrence and 20-year breast cancer mortality: meta-analysis of individual patient data for 8135 women in 22 randomised trials. Lancet. 2014 Jun 21;383(9935):2127–35.

63. Early Breast Cancer Trialists' Collaborative Group (EBCTCG), Darby S, McGale P, Correa C, Taylor C, Arriagada R, et al. Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: meta-analysis of individual patient data for 10,801 women in 17 randomised trials. Lancet. 2011 Nov 12;378(9804):1707–16.

64. Rosenkranz KM, Ballman K, McCall L, McCarthy C, Kubicky CD, Cuttino L, et al. Cosmetic Outcomes Following Breast-Conservation Surgery and Radiation for Multiple Ipsilateral Breast Cancer: Data from the Alliance Z11102 Study. Ann Surg Oncol. 2020 Nov 1;27(12):4650–61.

65. Nationellt Kvalitetsregister för Bröstcancer. Arsrapport 2022 från Nationellt Kvalitetsregister för Bröstcancer [Internet]. 2022. Available from:

https://cancercentrum.se/contentassets/c36b580a94ab4c3794aa9d41bb954871/ett-urval-av-data-fran-nkbc-rapporten-for-2022.pdf

66. Lostumbo L, Carbine NE, Wallace J. Prophylactic mastectomy for the prevention of breast cancer. Cochrane Database Syst Rev. 2010 Nov 10;(11):CD002748.

67. Carbine NE, Lostumbo L, Wallace J, Ko H. Risk-reducing mastectomy for the prevention of primary breast cancer. Cochrane Database Syst Rev. 2018 Apr 5;4(4):CD002748.

68. Tesson S, Richards I, Porter D, Phillips KA, Rankin N, Costa D, et al. Women's preferences for contralateral prophylactic mastectomy following unilateral breast cancer: What risk-reduction makes it worthwhile? Breast. 2017 Feb;31:233–40.

69. Krag DN, Weaver DL, Alex JC, Fairbank JT. Surgical resection and radiolocalization of the sentinel lymph node in breast cancer using a gamma probe. Surg Oncol. 1993 Dec;2(6):335–9; discussion 340.

70. Veronesi U, Paganelli G, Galimberti V, Viale G, Zurrida S, Bedoni M, et al. Sentinel-node biopsy to avoid axillary dissection in breast cancer with clinically negative lymph-nodes. Lancet. 1997 Jun 28;349(9069):1864–7.

71. Albertini JJ, Lyman GH, Cox C, Yeatman T, Balducci L, Ku N, et al. Lymphatic mapping and sentinel node biopsy in the patient with breast cancer. JAMA. 1996 Dec 11;276(22):1818–22.

72. Veronesi U, Paganelli G, Viale G, Galimberti V, Luini A, Zurrida S, et al. Sentinel lymph node biopsy and axillary dissection in breast cancer: results in a large series. J Natl Cancer Inst. 1999 Feb 17;91(4):368–73.

73. Vaidya JS, Wenz F, Bulsara M, Tobias JS, Joseph DJ, Keshtgar M, et al. Risk-adapted targeted intraoperative radiotherapy versus whole-breast radiotherapy for breast cancer: 5-year results for local control and overall survival from the TARGIT-A randomised trial. Lancet. 2014 Feb 15;383(9917):603–13.

74. Goldberg G. Cancer Chemotherapy: Basic Science to the Clinic, 2nd Edition | Wiley [Internet]. 2nd ed. Wiley; [cited 2024 Aug 9]. 320 p. Available from: https://www.wiley.com/en-us/Cancer+Chemotherapy%3A+Basic+Science+to+the+Clinic%2C+2nd+Edition-p-9781118963852

75. Skeel RT, Khleif SN. Handbook of Cancer Chemotherapy. Lippincott Williams & Wilkins; 2011. 896 p.

76. Goldhirsch A, Winer EP, Coates AS, Gelber RD, Piccart-Gebhart M, Thürlimann B, et al. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. Ann Oncol. 2013 Sep;24(9):2206–23.

77. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Long-term outcomes for neoadjuvant versus adjuvant chemotherapy in early breast cancer: meta-analysis of individual patient data from ten randomised trials. Lancet Oncol. 2018 Jan;19(1):27–39.

78. Curigliano G, Burstein HJ, Gnant M, Loibl S, Cameron D, Regan MM, et al. Understanding breast cancer complexity to improve patient outcomes: The St Gallen International Consensus Conference for the Primary Therapy of Individuals with Early Breast Cancer 2023. Ann Oncol. 2023 Nov;34(11):970–86.

79. Cortazar P, Zhang L, Untch M, Mehta K, Costantino JP, Wolmark N, et al. Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. Lancet. 2014 Jul 12;384(9938):164–72.

80. Bertucci F, Finetti P, Colpaert C, Mamessier E, Parizel M, Dirix L, et al. PDL1 expression in inflammatory breast cancer is frequent and predicts for the pathological response to chemotherapy. Oncotarget. 2015 May 30;6(15):13506–19.

81. Mittendorf EA, Philips AV, Meric-Bernstam F, Qiao N, Wu Y, Harrington S, et al. PD-L1 expression in triple-negative breast cancer. Cancer Immunol Res. 2014 Apr;2(4):361–70.

82. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. Nat Rev Cancer. 2012 Mar 22;12(4):252–64.

83. Waks AG, Winer EP. Breast Cancer Treatment: A Review. JAMA. 2019 Jan 22;321(3):288–300.

84. Dustin D, Gu G, Fuqua SA. ESR1 Mutations in Breast Cancer. Cancer. 2019 Nov 1;125(21):3714–28.

85. Bonneterre J, Thürlimann B, Robertson JF, Krzakowski M, Mauriac L, Koralewski P, et al. Anastrozole versus tamoxifen as first-line therapy for advanced breast cancer in 668 postmenopausal women: results of the Tamoxifen or Arimidex Randomized Group Efficacy and Tolerability study. J Clin Oncol. 2000 Nov 15;18(22):3748–57.

86. Baum M, Budzar AU, Cuzick J, Forbes J, Houghton JH, Klijn JGM, et al. Anastrozole alone or in combination with tamoxifen versus tamoxifen alone for adjuvant treatment of postmenopausal women with early breast cancer: first results of the ATAC randomised trial. Lancet. 2002 Jun 22;359(9324):2131–9.

87. Burstein HJ, Curigliano G, Loibl S, Dubsky P, Gnant M, Poortmans P, et al. Estimating the benefits of therapy for early-stage breast cancer: the St. Gallen International Consensus Guidelines for the primary therapy of early breast cancer 2019. Ann Oncol. 2019 Oct 1;30(10):1541–57.

88. Cardoso F, Kyriakides S, Ohno S, Penault-Llorca F, Poortmans P, Rubio IT, et al. Early breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up†. Ann Oncol. 2019 Aug 1;30(8):1194–220.

89. Miller WR, Bartlett JMS, Canney P, Verrill M. Hormonal therapy for postmenopausal breast cancer: the science of sequencing. Breast Cancer Res Treat. 2007 Jun;103(2):149–60.

90. Klein-Hitpass L, Ryffel GU, Heitlinger E, Cato AC. A 13 bp palindrome is a functional estrogen responsive element and interacts specifically with estrogen receptor. Nucleic Acids Res. 1988 Jan 25;16(2):647–63.

91. Clusan L, Ferrière F, Flouriot G, Pakdel F. A Basic Review on Estrogen Receptor Signaling Pathways in Breast Cancer. International Journal of Molecular Sciences. 2023 Jan;24(7):6834.

92. Licciulli S. Endocrine Therapy in Breast Cancer: Making Sense of the "Word Salad" [Internet]. American Association for Cancer Research (AACR). 2022 [cited 2024 Aug 9]. Available from: https://www.aacr.org/blog/2022/07/18/endocrine-therapy-in-breast-cancer-making-sense-of-the-word-salad/

93. Lawson M, Cureton N, Ros S, Cheraghchi-Bashi A, Urosevic J, D'Arcy S, et al. The Next-Generation Oral Selective Estrogen Receptor Degrader Camizestrant (AZD9833) Suppresses ER+ Breast Cancer Growth and Overcomes Endocrine and CDK4/6 Inhibitor Resistance. Cancer Research. 2023 Dec 1;83(23):3989–4004.

94. Jordan VC. Tamoxifen: a most unlikely pioneering medicine. Nat Rev Drug Discov. 2003 Mar;2(3):205–13.

95. O'Regan RM, Jordan VC. The evolution of tamoxifen therapy in breast cancer: selective oestrogen-receptor modulators and downregulators. Lancet Oncol. 2002 Apr;3(4):207–14.

96. Fabian CJ, Kimler BF. Selective Estrogen-Receptor Modulators for Primary Prevention of Breast Cancer. JCO. 2005 Mar 10;23(8):1644–55.

97. Tamoxifen for early breast cancer: an overview of the randomised trials. Early Breast Cancer Trialists' Collaborative Group. Lancet. 1998 May 16;351(9114):1451–67.

98. Bergman L, Beelen ML, Gallee MP, Hollema H, Benraadt J, van Leeuwen FE. Risk and prognosis of endometrial cancer after tamoxifen for breast cancer. Comprehensive Cancer Centres' ALERT Group. Assessment of Liver and Endometrial cancer Risk following Tamoxifen. Lancet. 2000 Sep 9;356(9233):881–7.

99. Fisher B, Costantino JP, Redmond CK, Fisher ER, Wickerham DL, Cronin WM. Endometrial cancer in tamoxifen-treated breast cancer patients: findings from the National Surgical Adjuvant Breast and Bowel Project (NSABP) B-14. J Natl Cancer Inst. 1994 Apr 6;86(7):527–37.

100. Jones ME, van Leeuwen FE, Hoogendoorn WE, Mourits MJ, Hollema H, van Boven H, et al. Endometrial cancer survival after breast cancer in relation to tamoxifen treatment: Pooled results from three countries. Breast Cancer Res. 2012;14(3):R91.

101. Singh MS, Francis PA, Michael M. Tamoxifen, cytochrome P450 genes and breast cancer clinical outcomes. The Breast. 2011 Apr 1;20(2):111–8.

102. Iacopetta D, Ceramella J, Catalano A, Scali E, Scumaci D, Pellegrino M, et al. Impact of Cytochrome P450 Enzymes on the Phase I Metabolism of Drugs. Applied Sciences. 2023 Jan;13(10):6045.

103. Rondón-Lagos M, Villegas VE, Rangel N, Sánchez MC, Zaphiropoulos PG. Tamoxifen Resistance: Emerging Molecular Targets. Int J Mol Sci. 2016 Aug 19;17(8):1357.

104. Johnson MD, Zuo H, Lee KH, Trebley JP, Rae JM, Weatherman RV, et al. Pharmacological characterization of 4-hydroxy-N-desmethyl tamoxifen, a novel active metabolite of tamoxifen. Breast Cancer Res Treat. 2004 May;85(2):151–9.

105. Fliss AE, Benzeno S, Rao J, Caplan AJ. Control of estrogen receptor ligand binding by Hsp90. J Steroid Biochem Mol Biol. 2000 Apr;72(5):223–30.

106. Heldring N, Nilsson M, Buehrer B, Treuter E, Gustafsson JÅ. Identification of Tamoxifen-Induced Coregulator Interaction Surfaces within the Ligand-

Binding Domain of Estrogen Receptors. Molecular and Cellular Biology. 2004 Apr;24(8):3445.

107. Santen RJ, Brodie H, Simpson ER, Siiteri PK, Brodie A. History of aromatase: saga of an important biological mediator and therapeutic target. Endocr Rev. 2009 Jun;30(4):343–75.

108. Chumsri S, Howes T, Bao T, Sabnis G, Brodie A. Aromatase, Aromatase Inhibitors, and Breast Cancer. J Steroid Biochem Mol Biol. 2011 May;125(1–2):13–22.

109. Smith IE, Dowsett M. Aromatase inhibitors in breast cancer. N Engl J Med. 2003 Jun 12;348(24):2431–42.

110. Miller WR. Aromatase inhibitors: mechanism of action and role in the treatment of breast cancer. Seminars in Oncology. 2003 Aug 1;30:3–11.

111. Geisler J. Differences between the non-steroidal aromatase inhibitors anastrozole and letrozole--of clinical importance? Br J Cancer. 2011 Mar 29;104(7):1059–66.

112. Geisler J, Lønning PE. Endocrine effects of aromatase inhibitors and inactivators in vivo: review of data and method limitations. J Steroid Biochem Mol Biol. 2005 May;95(1–5):75–81.

113. Casper RF, Mitwally MFM. Aromatase Inhibitors for Ovulation Induction. The Journal of Clinical Endocrinology & Metabolism. 2006 Mar 1;91(3):760–71.

114. Howell A. Pure oestrogen antagonists for the treatment of advanced breast cancer. Endocr Relat Cancer. 2006 Sep;13(3):689–706.

115. Wakeling AE, Dukes M, Bowler J. A potent specific pure antiestrogen with clinical potential. Cancer Res. 1991 Aug 1;51(15):3867–73.

116. Bundred N. Preclinical and clinical experience with fulvestrant (Faslodex) in postmenopausal women with hormone receptor-positive advanced breast cancer. Cancer Invest. 2005;23(2):173–81.

117. Howell A. Fulvestrant ('Faslodex'): current and future role in breast cancer management. Crit Rev Oncol Hematol. 2006 Mar;57(3):265–73.

118. Cardoso F, Paluch-Shimon S, Senkus E, Curigliano G, Aapro MS, André F, et al. 5th ESO-ESMO international consensus guidelines for advanced breast cancer (ABC 5). Ann Oncol. 2020 Dec;31(12):1623–49.

119. Fawell SE, White R, Hoare S, Sydenham M, Page M, Parker MG. Inhibition of estrogen receptor-DNA binding by the “pure” antiestrogen ICI 164,384 appears to be mediated by impaired receptor dimerization. Proc Natl Acad Sci U S A. 1990 Sep;87(17):6883–7.

120. Dauvois S, White R, Parker MG. The antiestrogen ICI 182780 disrupts estrogen receptor nucleocytoplasmic shuttling. J Cell Sci. 1993 Dec;106 (Pt 4):1377–88.

121. Hernando C, Ortega-Morillo B, Tapia M, Moragón S, Martínez MT, Eroles P, et al. Oral Selective Estrogen Receptor Degraders (SERDs) as a Novel Breast Cancer Therapy: Present and Future from a Clinical Perspective. Int J Mol Sci. 2021 Jul 22;22(15):7812.

122. Osborne CK, Wakeling A, Nicholson RI. Fulvestrant: an oestrogen receptor antagonist with a novel mechanism of action. Br J Cancer. 2004 Mar;90 Suppl 1(Suppl 1):S2-6.

123. Long X, Nephew KP. Fulvestrant (ICI 182,780)-dependent interacting proteins mediate immobilization and degradation of estrogen receptor-alpha. J Biol Chem. 2006 Apr 7;281(14):9607–15.

124. Wijayaratne AL, McDonnell DP. The human estrogen receptor-alpha is a ubiquitinated protein whose stability is affected differentially by agonists, antagonists, and selective estrogen receptor modulators. J Biol Chem. 2001 Sep 21;276(38):35684–92.

125. Wang G. Fulvestrant as a reference antiestrogen and estrogen receptor (ER) degrader in preclinical studies: treatment dosage, efficacy, and implications on development of new ER-targeting agents. Transl Cancer Res. 2020 Aug;9(8):4464–8.

126. Davoli A, Hocevar BA, Brown TL. Progression and treatment of HER2-positive breast cancer. Cancer Chemother Pharmacol. 2010 Mar 1;65(4):611–23.

127. Baselga J, Albanell J. Mechanism of action of anti-HER2 monoclonal antibodies. Ann Oncol. 2001;12 Suppl 1:S35-41.

128. Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, et al. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. N Engl J Med. 2001 Mar 15;344(11):783–92.

129. von Minckwitz G, Procter M, de Azambuja E, Zardavas D, Benyunes M, Viale G, et al. Adjuvant Pertuzumab and Trastuzumab in Early HER2-Positive Breast Cancer. N Engl J Med. 2017 Jul 13;377(2):122–31.

130. Cortés J, Kim SB, Chung WP, Im SA, Park YH, Hegg R, et al. Trastuzumab Deruxtecan versus Trastuzumab Emtansine for Breast Cancer. N Engl J Med. 2022 Mar 24;386(12):1143–54.

131. Slamon D, Eiermann W, Robert N, Pienkowski T, Martin M, Press M, et al. Adjuvant trastuzumab in HER2-positive breast cancer. N Engl J Med. 2011 Oct 6;365(14):1273–83.

132. Early Breast Cancer Trialists' Collaborative group (EBCTCG). Trastuzumab for early-stage, HER2-positive breast cancer: a meta-analysis of 13 864 women in seven randomised trials. Lancet Oncol. 2021 Aug;22(8):1139–50.

133. Marchiò C, Annaratone L, Marques A, Casorzo L, Berrino E, Sapino A. Evolving concepts in HER2 evaluation in breast cancer: Heterogeneity, HER2-low carcinomas and beyond. Semin Cancer Biol. 2021 Jul;72:123–35.

134. Ring A, Dowsett M. Mechanisms of tamoxifen resistance. Endocr Relat Cancer. 2004 Dec;11(4):643–58.

135. Schiff R, Massarweh S, Shou J, Osborne CK. Breast cancer endocrine resistance: how growth factor signaling and estrogen receptor coregulators modulate response. Clin Cancer Res. 2003 Jan;9(1 Pt 2):447S-54S.

136. Piggott L, Silva A, Robinson T, Santiago-Gómez A, Simões BM, Becker M, et al. Acquired Resistance of ER-Positive Breast Cancer to Endocrine Treatment Confers an Adaptive Sensitivity to TRAIL through Posttranslational Downregulation of c-FLIP. Clinical Cancer Research. 2018 May 14;24(10):2452–63.

137. Vergote I, Amant F, Leunen K, Van Gorp T, Berteloot P, Neven P. Metastatic breast cancer: sequencing hormonal therapy and positioning of fulvestrant. Int J Gynecol Cancer. 2006;16 Suppl 2:524–6.

138. Wang J, Jain S, Coombes CR, Palmieri C. Fulvestrant in advanced breast cancer following tamoxifen and aromatase inhibition: a single center experience. Breast J. 2009;15(3):247–53.

139. Shiino S, Kinoshita T, Yoshida M, Jimbo K, Asaga S, Takayama S, et al. Prognostic Impact of Discordance in Hormone Receptor Status Between Primary and Recurrent Sites in Patients With Recurrent Breast Cancer. Clin Breast Cancer. 2016 Aug;16(4):e133-140.

140. Liu H, Cheng D, Weichel AK, Osipo C, Wing LK, Chen B, et al. Cooperative effect of gefitinib and fumitremorgin c on cell growth and chemosensitivity in estrogen receptor alpha negative fulvestrant-resistant MCF-7 cells. Int J Oncol. 2006 Nov;29(5):1237–46.

141. Levin ER. Bidirectional signaling between the estrogen receptor and the epidermal growth factor receptor. Mol Endocrinol. 2003 Mar;17(3):309–17.

142. Fribbens C, O'Leary B, Kilburn L, Hrebien S, Garcia-Murillas I, Beaney M, et al. Plasma ESR1 Mutations and the Treatment of Estrogen Receptor-Positive Advanced Breast Cancer. J Clin Oncol. 2016 Sep 1;34(25):2961–8.

143. Chandarlapaty S, Chen D, He W, Sung P, Samoila A, You D, et al. Prevalence of ESR1 Mutations in Cell-Free DNA and Outcomes in Metastatic Breast Cancer: A Secondary Analysis of the BOLERO-2 Clinical Trial. JAMA Oncol. 2016 Oct 1;2(10):1310–5.

144. Engler T, Fasching PA, Lüftner D, Hartkopf AD, Müller V, Kolberg HC, et al. Implementation of CDK4/6 Inhibitors and its Influence on the Treatment Landscape of Advanced Breast Cancer Patients - Data from the Real-World Registry PRAEGNANT. Geburtshilfe Frauenheilkd. 2022 Oct;82(10):1055–67.

145. Dogheim GM, Amralla MT. Proteolysis Targeting Chimera (PROTAC) as a promising novel therapeutic modality for the treatment of triple-negative breast cancer (TNBC). Drug Development Research. 2023;84(4):629–53.

146. Rostron C. Drug Design and Development [Internet]. Oxford University press; 2020 [cited 2024 Aug 11]. 376 p. Available from: https://global.oup.com/academic/product/drug-design-and-development-9780198749318

147. Shupnik MA. Crosstalk between steroid receptors and the c-Src-receptor tyrosine kinase pathways: implications for cell proliferation. Oncogene. 2004 Oct 18;23(48):7979–89.

148. Wärnmark A, Treuter E, Wright APH, Gustafsson JÅ. Activation Functions 1 and 2 of Nuclear Receptors: Molecular Strategies for Transcriptional Activation. Molecular Endocrinology. 2003 Oct 1;17(10):1901–9.

149. Ward PS, Thompson CB. Signaling in Control of Cell Growth and Metabolism. Cold Spring Harb Perspect Biol. 2012 Jul;4(7):a006783.

150. Williams C, Edvardsson K, Lewandowski SA, Ström A, Gustafsson JÅ. A genome-wide study of the repressive effects of estrogen receptor beta on

estrogen receptor alpha signaling in breast cancer cells. Oncogene. 2008 Feb;27(7):1019–32.

151. Klapper LN, Glathe S, Vaisman N, Hynes NE, Andrews GC, Sela M, et al. The ErbB-2/HER2 oncoprotein of human carcinomas may function solely as a shared coreceptor for multiple stroma-derived growth factors. Proc Natl Acad Sci U S A. 1999 Apr 27;96(9):4995–5000.

152. Fulda S, Gorman AM, Hori O, Samali A. Cellular Stress Responses: Cell Survival and Cell Death. Int J Cell Biol. 2010;2010:214074.

153. Couse JF, Lindzey J, Grandien K, Gustafsson JA, Korach KS. Tissue distribution and quantitative analysis of estrogen receptor-alpha (ERalpha) and estrogen receptor-beta (ERbeta) messenger ribonucleic acid in the wild-type and ERalpha-knockout mouse. Endocrinology. 1997 Nov;138(11):4613–21.

154. Scarpin KM, Graham JD, Mote PA, Clarke CL. Progesterone action in human tissues: regulation by progesterone receptor (PR) isoform expression, nuclear positioning and coregulator expression. Nucl Recept Signal. 2009 Dec 31;7:e009.

155. Kuiper GGJM, Enmark E, Pelto-Huikko M, Nilsson S, Gustafsson JÅ. Cloning of a novel estrogen receptor expressed in rat prostate and ovary. Proceedings of the National Academy of Sciences of the United States of America. 1996;93(12):5925–30.

156. Prossnitz ER, Arterburn JB, Sklar LA. GPR30: a G protein-coupled receptor for estrogen. Mol Cell Endocrinol. 2007 Feb;265–266:138–42.

157. Tsai MJ, O'Malley BW. Molecular mechanisms of action of steroid/thyroid receptor superfamily members. Annu Rev Biochem. 1994;63:451–86.

158. Saville B, Wormke M, Wang F, Nguyen T, Enmark E, Kuiper G, et al. Ligand-, cell-, and estrogen receptor subtype (alpha/beta)-dependent activation at GC-rich (Sp1) promoter elements. J Biol Chem. 2000 Feb 25;275(8):5379–87.

159. Miricescu D, Totan A, Stanescu-Spinu II, Badoiu SC, Stefani C, Greabu M. PI3K/AKT/mTOR Signaling Pathway in Breast Cancer: From Molecular Landscape to Clinical Aspects. Int J Mol Sci. 2020 Dec 26;22(1):173.

160. Shou J, Massarweh S, Osborne CK, Wakeling AE, Ali S, Weiss H, et al. Mechanisms of tamoxifen resistance: increased estrogen receptor-HER2/neu cross-talk in ER/HER2-positive breast cancer. J Natl Cancer Inst. 2004 Jun 16;96(12):926–35.

161. Schwabe JW, Chapman L, Finch JT, Rhodes D. The crystal structure of the estrogen receptor DNA-binding domain bound to DNA: how receptors discriminate between their response elements. Cell. 1993 Nov 5;75(3):567–78.

162. Nardulli AM. Estrogen Receptor-α Structure and Function. In: Henry HL, Norman AW, editors. Encyclopedia of Hormones [Internet]. New York: Academic Press; 2003 [cited 2024 Aug 12]. p. 594–9. Available from: https://www.sciencedirect.com/science/article/pii/B0123411033000905

163. Patel JM, Jeselsohn RM. Estrogen Receptor Alpha and ESR1 Mutations in Breast Cancer. In: Campbell MJ, Bevan CL, editors. Nuclear Receptors in Human Health and Disease [Internet]. Cham: Springer International Publishing; 2022 [cited 2024 Aug 21]. p. 171–94. Available from: https://doi.org/10.1007/978-3-031-11836-4_10

164. Brzozowski AM, Pike AC, Dauter Z, Hubbard RE, Bonn T, Engström O, et al. Molecular basis of agonism and antagonism in the oestrogen receptor. Nature. 1997 Oct 16;389(6652):753–8.

165. Moras D, Gronemeyer H. The nuclear receptor ligand-binding domain: structure and function. Curr Opin Cell Biol. 1998 Jun;10(3):384–91.

166. Mohammed H, Russell IA, Stark R, Rueda OM, Hickey TE, Tarulli GA, et al. Corrigendum: Progesterone receptor modulates ERα action in breast cancer. Nature. 2015 Oct 1;526(7571):144.

167. Dowsett M, Cuzick J, Ingle J, Coates A, Forbes J, Bliss J, et al. Meta-analysis of breast cancer outcomes in adjuvant trials of aromatase inhibitors versus tamoxifen. J Clin Oncol. 2010 Jan 20;28(3):509–18.

168. Riggins RB, Schrecengost RS, Guerrero MS, Bouton AH. Pathways to Tamoxifen Resistance. Cancer Lett. 2007 Oct 18;256(1):1–24.

169. Mangelsdorf DJ, Thummel C, Beato M, Herrlich P, Schütz G, Umesono K, et al. The nuclear receptor superfamily: the second decade. Cell. 1995 Dec 15;83(6):835–9.

170. Takimoto GS, Tung L, Abdel-Hafiz H, Abel MG, Sartorius CA, Richer JK, et al. Functional properties of the N-terminal region of progesterone receptors and their mechanistic relationship to structure. J Steroid Biochem Mol Biol. 2003 Jun;85(2–5):209–19.

171. Shahab M, Ziyu P, Waqas M, Zheng G, Bin Jardan YA, Fentahun Wondmie G, et al. Targeting human progesterone receptor (PR), through

pharmacophore-based screening and molecular simulation revealed potent inhibitors against breast cancer. Sci Rep. 2024 Mar 21;14(1):6768.

172. Pateetin P, Hutvagner G, Bajan S, Padula MP, McGowan EM, Boonyaratanakornkit V. Triple SILAC identified progestin-independent and dependent PRA and PRB interacting partners in breast cancer. Sci Data. 2021 Apr 12;8(1):100.

173. Kastner P, Krust A, Turcotte B, Stropp U, Tora L, Gronemeyer H, et al. Two distinct estrogen-regulated promoters generate transcripts encoding the two functionally different human progesterone receptor forms A and B. EMBO J. 1990 May;9(5):1603–14.

174. Hilton HN, Graham JD, Kantimm S, Santucci N, Cloosterman D, Huschtscha LI, et al. Progesterone and estrogen receptors segregate into different cell subpopulations in the normal human breast. Mol Cell Endocrinol. 2012 Sep 25;361(1–2):191–201.

175. Perez EA, Romond EH, Suman VJ, Jeong JH, Sledge G, Geyer CE, et al. Trastuzumab plus adjuvant chemotherapy for human epidermal growth factor receptor 2-positive breast cancer: planned joint analysis of overall survival from NSABP B-31 and NCCTG N9831. J Clin Oncol. 2014 Nov 20;32(33):3744–52.

176. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science. 1987 Jan 9;235(4785):177–82.

177. Yarden Y, Sliwkowski MX. Untangling the ErbB signalling network. Nat Rev Mol Cell Biol. 2001 Feb;2(2):127–37.

178. Wieduwilt MJ, Moasser MM. The epidermal growth factor receptor family: biology driving targeted therapeutics. Cell Mol Life Sci. 2008 May;65(10):1566–84.

179. Bai X, Sun P, Wang X, Long C, Liao S, Dang S, et al. Structure and dynamics of the EGFR/HER2 heterodimer. Cell Discov. 2023 Feb 13;9(1):18.

180. Wright CJ, Smith CWJ, Jiggins CD. Alternative splicing as a source of phenotypic diversity. Nat Rev Genet. 2022 Nov;23(11):697–710.

181. Pal S, Gupta R, Davuluri RV. Alternative transcription and alternative splicing in cancer. Pharmacol Ther. 2012 Dec;136(3):283–94.

182. Donaldson LF, Beazley-Long N. Alternative RNA splicing: contribution to pain and potential therapeutic strategy. Drug Discov Today. 2016 Nov;21(11):1787–98.

183. Ward AJ, Cooper TA. The Pathobiology of Splicing. J Pathol. 2010 Jan;220(2):152–63.

184. Wahl MC, Will CL, Lührmann R. The spliceosome: design principles of a dynamic RNP machine. Cell. 2009 Feb 20;136(4):701–18.

185. Smith CWJ, Valcárcel J. Alternative pre-mRNA splicing: the logic of combinatorial control. Trends in Biochemical Sciences. 2000 Aug 1;25(8):381–8.

186. Flouriot G, Griffin C, Kenealy M, Sonntag-Buck V, Gannon F. Differentially expressed messenger RNA isoforms of the human estrogen receptor-alpha gene are generated by alternative splicing and promoter usage. Mol Endocrinol. 1998 Dec;12(12):1939–54.

187. Pagano MT, Ortona E, Dupuis ML. A Role for Estrogen Receptor alpha36 in Cancer Progression. Front Endocrinol (Lausanne). 2020;11:506.

188. Chantalat E, Boudou F, Laurell H, Palierne G, Houtman R, Melchers D, et al. The AF-1-deficient estrogen receptor ERα46 isoform is frequently expressed in human breast tumors. Breast Cancer Res. 2016;18:123.

189. Klinge CM, Riggs KA, Wickramasinghe NS, Emberts CG, McConda DB, Barry PN, et al. Estrogen receptor alpha 46 is reduced in tamoxifen resistant breast cancer cells and re-expression inhibits cell proliferation and estrogen receptor alpha 66-regulated target gene transcription. Mol Cell Endocrinol. 2010 Jul 29;323(2):268–76.

190. Flouriot G, Brand H, Denger S, Metivier R, Kos M, Reid G, et al. Identification of a new isoform of the human estrogen receptor-alpha (hER-alpha) that is encoded by distinct transcripts and that is able to repress hER-alpha activation function 1. EMBO J. 2000 Sep 1;19(17):4688–700.

191. Wang Z, Zhang X, Shen P, Loggie BW, Chang Y, Deuel TF. A variant of estrogen receptor-{alpha}, hER-{alpha}36: transduction of estrogen- and antiestrogen-dependent membrane-initiated mitogenic signaling. Proc Natl Acad Sci U S A. 2006 Jun 13;103(24):9063–8.

192. Gu Y, Chen T, López E, Wu W, Wang X, Cao J, et al. The therapeutic target of estrogen receptor-alpha36 in estrogen-dependent tumors. J Transl Med. 2014 Jan 21;12:16.

193. Mahboobifard F, Dargahi L, Jorjani M, Ramezani Tehrani F, Pourgholami MH. The role of ERα36 in cell type-specific functions of estrogen and cancer development. Pharmacol Res. 2021 Jan;163:105307.

194. Lee LMJ, Cao J, Deng H, Chen P, Gatalica Z, Wang ZY. ER-alpha36, a novel variant of ER-alpha, is expressed in ER-positive and -negative human breast carcinomas. Anticancer Res. 2008;28(1B):479–83.

195. Estrogen receptor isoforms. The 8 exons of the ESR1 [Internet]. [cited 2024 Aug 22]. Available from: https://www.researchgate.net/figure/Estrogen-receptor-isoforms-The-8-exons-of-the-ESR1-red-and-ESR2-purple-genes-that_fig1_51687937

196. Zhang QX, Borg Å, Fuqua SAW. An Exon 5 Deletion Variant of the Estrogen Receptor Frequently Coexpressed with Wild-Type Estrogen Receptor in Human Breast Cancer1. Cancer Research. 1993 Dec 1;53(24):5882–4.

197. Hattori Y, Ishii H, Munetomo A, Watanabe H, Morita A, Sakuma Y, et al. Human C-terminally truncated ERα variants resulting from the use of alternative exons in the ligand-binding domain. Molecular and Cellular Endocrinology. 2016 Apr 15;425:111–22.

198. García Pedrero JM, Zuazua P, Martínez-Campa C, Lazo PS, Ramos S. The naturally occurring variant of estrogen receptor (ER) ERDeltaE7 suppresses estrogen-dependent transcriptional activation by both wild-type ERalpha and ERbeta. Endocrinology. 2003 Jul;144(7):2967–76.

199. Bollig A, Miksicek RJ. An estrogen receptor-alpha splicing variant mediates both positive and negative effects on gene transcription. Mol Endocrinol. 2000 May;14(5):634–49.

200. Groenendijk FH, Zwart W, Floore A, Akbari S, Bernards R. Estrogen receptor splice variants as a potential source of false-positive estrogen receptor status in breast cancer diagnostics. Breast Cancer Res Treat. 2013;140(3):475–84.

201. Wang Y, Miksicek RJ. Identification of a dominant negative form of the human estrogen receptor. Mol Endocrinol. 1991 Nov;5(11):1707–15.

202. Pathologists' Guideline Recommendations for Immunohistochemical Testing of Estrogen and Progesterone Receptors in Breast Cancer. Breast Care (Basel). 2010 Jun;5(3):185–7.

203. Bird A. Perceptions of epigenetics. Nature. 2007 May 24;447(7143):396–8.

204. Gibney ER, Nolan CM. Epigenetics and gene expression. Heredity (Edinb). 2010 Jul;105(1):4–13.

205. Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell. 1999 Oct 29;99(3):247–57.

206. Baylin SB. DNA methylation and gene silencing in cancer. Nat Clin Pract Oncol. 2005 Dec;2 Suppl 1:S4-11.

207. Kouzarides T. Chromatin Modifications and Their Function. Cell. 2007 Feb 23;128(4):693–705.

208. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. Cell Research. 2011 Mar;21(3):381.

209. Pfeifer GP. Defining Driver DNA Methylation Changes in Human Cancer. International Journal of Molecular Sciences. 2018 Apr;19(4):1166.

210. Feng L, Lou J. DNA Methylation Analysis. Methods Mol Biol. 2019;1894:181–227.

211. Chuang LS, Ian HI, Koh TW, Ng HH, Xu G, Li BF. Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. Science. 1997 Sep 26;277(5334):1996–2000.

212. Chedin F, Lieber MR, Hsieh CL. The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. Proc Natl Acad Sci U S A. 2002 Dec 24;99(26):16916–21.

213. Jia D, Jurkowska RZ, Zhang X, Jeltsch A, Cheng X. Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. Nature. 2007 Sep 13;449(7159):248–51.

214. Hermann A, Goyal R, Jeltsch A. The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. J Biol Chem. 2004 Nov 12;279(46):48350–9.

215. Saxonov S, Berg P, Brutlag DL. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proc Natl Acad Sci U S A. 2006 Jan 31;103(5):1412–7.

216. Shen L, Kondo Y, Guo Y, Zhang J, Zhang L, Ahmed S, et al. Genome-wide profiling of DNA methylation reveals a class of normally methylated CpG island promoters. PLoS Genet. 2007 Oct;3(10):2023–36.

217. Rollins RA, Haghighi F, Edwards JR, Das R, Zhang MQ, Ju J, et al. Large-scale structure of genomic methylation patterns. Genome Res. 2006 Feb;16(2):157–63.

218. Bird A. DNA methylation patterns and epigenetic memory. Genes Dev. 2002 Jan 1;16(1):6–21.

219. Szyf M. DNA methylation signatures for breast cancer classification and prognosis. Genome Med. 2012 Mar 30;4(3):26.

220. Takai D, Jones PA. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. Proc Natl Acad Sci U S A. 2002 Mar 19;99(6):3740–5.

221. Wang Y, Leung FCC. An evaluation of new criteria for CpG islands in the human genome as gene markers. Bioinformatics. 2004 May 1;20(7):1170–7.

222. Quintas-Granados LI, Cortés H, Carmen MGD, Leyva-Gómez G, Bustamante-Montes LP, Rodríguez-Morales M, et al. The high methylation level of a novel 151-bp CpG island in the ESR1 gene promoter is associated with a poor breast cancer prognosis. Cancer Cell International. 2021 Dec 4;21(1):649.

223. Robertson KD. DNA methylation and human disease. Nat Rev Genet. 2005 Aug;6(8):597–610.

224. Bogdanović O, Veenstra GJC. DNA methylation and methyl-CpG binding proteins: developmental requirements and function. Chromosoma. 2009 Oct;118(5):549–65.

225. Holm K, Hegardt C, Staaf J, Vallon-Christersson J, Jönsson G, Olsson H, et al. Molecular subtypes of breast cancer are associated with characteristic DNA methylation patterns. Breast Cancer Res. 2010;12(3):R36.

226. Holm K, Staaf J, Lauss M, Aine M, Lindgren D, Bendahl PO, et al. An integrated genomics analysis of epigenetic subtypes in human breast tumors links DNA methylation patterns to chromatin states in normal mammary cells. Breast Cancer Res. 2016 Feb 29;18(1):27.

227. Esteller M, Herman JG. Cancer as an epigenetic disease: DNA methylation and chromatin alterations in human tumours. J Pathol. 2002 Jan;196(1):1–7.

228. Fleischer T, Frigessi A, Johnson KC, Edvardsen H, Touleimat N, Klajic J, et al. Genome-wide DNA methylation profiles in progression to in situ and

invasive carcinoma of the breast with impact on gene transcription and prognosis. Genome Biol. 2014;15(8):435.

229. Fleischer T, Klajic J, Aure MR, Louhimo R, Pladsen AV, Ottestad L, et al. DNA methylation signature (SAM40) identifies subgroups of the Luminal A breast cancer samples with distinct survival. Oncotarget. 2017 Jan 3;8(1):1074–82.

230. Geissler F, Nesic K, Kondrashova O, Dobrovic A, Swisher EM, Scott CL, et al. The role of aberrant DNA methylation in cancer initiation and clinical impacts. Ther Adv Med Oncol. 2024 Jan 28;16:17588359231220512.

231. Treiber T, Treiber N, Meister G. Regulation of microRNA biogenesis and its crosstalk with other cellular pathways. Nat Rev Mol Cell Biol. 2019 Jan;20(1):5–20.

232. Fu G, Brkić J, Hayder H, Peng C. MicroRNAs in Human Placental Development and Pregnancy Complications. Int J Mol Sci. 2013 Mar 8;14(3):5519–44.

233. Lutter D, Marr C, Krumsiek J, Lang EW, Theis FJ. Intronic microRNAs support their host genes by mediating synergistic and antagonistic regulatory effects. BMC Genomics. 2010 Apr 6;11:224.

234. Bartel DP. MicroRNAs: target recognition and regulatory functions. Cell. 2009 Jan 23;136(2):215–33.

235. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell. 2004 Jan 23;116(2):281–97.

236. Lim LP, Lau NC, Garrett-Engele P, Grimson A, Schelter JM, Castle J, et al. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature. 2005 Feb 17;433(7027):769–73.

237. Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell. 2005 Jan 14;120(1):15–20.

238. Calin GA, Dumitru CD, Shimizu M, Bichi R, Zupo S, Noch E, et al. Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc Natl Acad Sci U S A. 2002 Nov 26;99(24):15524–9.

239. Lee Y, Kim M, Han J, Yeom KH, Lee S, Baek SH, et al. MicroRNA genes are transcribed by RNA polymerase II. EMBO J. 2004 Oct 13;23(20):4051–60.

240. Borchert GM, Lanier W, Davidson BL. RNA polymerase III transcribes human microRNAs. Nat Struct Mol Biol. 2006 Dec;13(12):1097–101.

241. Ha M, Kim VN. Regulation of microRNA biogenesis. Nat Rev Mol Cell Biol. 2014 Aug;15(8):509–24.

242. Lin S, Gregory RI. MicroRNA biogenesis pathways in cancer. Nat Rev Cancer. 2015 Jun;15(6):321–33.

243. O'Brien J, Hayder H, Zayed Y, Peng C. Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. Front Endocrinol (Lausanne). 2018;9:402.

244. Ameres SL, Zamore PD. Diversifying microRNA sequence and function. Nat Rev Mol Cell Biol. 2013 Aug;14(8):475–88.

245. Bartel DP. Metazoan MicroRNAs. Cell. 2018 Mar 22;173(1):20–51.

246. Brennecke J, Stark A, Russell RB, Cohen SM. Principles of microRNA-target recognition. PLoS Biol. 2005 Mar;3(3):e85.

247. Meister G, Landthaler M, Patkaniowska A, Dorsett Y, Teng G, Tuschl T. Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. Mol Cell. 2004 Jul 23;15(2):185–97.

248. Archer A, Kutter C, Williams C. Expression Profiles of Estrogen-Regulated MicroRNAs in Cancer Cells. Methods Mol Biol. 2022;2418:313–43.

249. Bhat-Nakshatri P, Wang G, Collins NR, Thomson MJ, Geistlinger TR, Carroll JS, et al. Estradiol-regulated microRNAs control estradiol response in breast cancer cells. Nucleic Acids Res. 2009 Aug;37(14):4850–61.

250. Krichevsky AM, Gabriely G. miR-21: a small multi-faceted RNA. J Cell Mol Med. 2009 Jan;13(1):39–53.

251. Yan LX, Huang XF, Shao Q, Huang MY, Deng L, Wu QL, et al. MicroRNA miR-21 overexpression in human breast cancer is associated with advanced clinical stage, lymph node metastasis and patient poor prognosis. RNA. 2008 Nov;14(11):2348–60.

252. Hatley ME, Patrick DM, Garcia MR, Richardson JA, Bassel-Duby R, van Rooij E, et al. Modulation of K-Ras-dependent lung tumorigenesis by MicroRNA-21. Cancer Cell. 2010 Sep 14;18(3):282–93.

253. Radojicic J, Zaravinos A, Vrekoussis T, Kafousi M, Spandidos DA, Stathopoulos EN. MicroRNA expression analysis in triple-negative (ER, PR and Her2/neu) breast cancer. Cell Cycle. 2011 Feb 1;10(3):507–17.

254. Wickramasinghe NS, Manavalan TT, Dougherty SM, Riggs KA, Li Y, Klinge CM. Estradiol downregulates miR-21 expression and increases miR-21 target gene expression in MCF-7 breast cancer cells. Nucleic Acids Research. 2009 May 1;37(8):2584–95.

255. Bautista-Sánchez D, Arriaga-Canon C, Pedroza-Torres A, De La Rosa-Velázquez IA, González-Barrios R, Contreras-Espinosa L, et al. The Promising Role of miR-21 as a Cancer Biomarker and Its Importance in RNA-Based Therapeutics. Mol Ther Nucleic Acids. 2020 Mar 13;20:409–20.

256. Aguilar-Martínez SY, Campos-Viguri GE, Medina-García SE, García-Flores RJ, Deas J, Gómez-Cerón C, et al. MiR-21 Regulates Growth and Migration of Cervical Cancer Cells by RECK Signaling Pathway. International Journal of Molecular Sciences. 2024 Jan;25(7):4086.

257. Iorio MV, Ferracin M, Liu CG, Veronese A, Spizzo R, Sabbioni S, et al. MicroRNA gene expression deregulation in human breast cancer. Cancer Res. 2005 Aug 15;65(16):7065–70.

258. Saito Y, Friedman JM, Chihara Y, Egger G, Chuang JC, Liang G. Epigenetic therapy upregulates the tumor suppressor microRNA-126 and its host gene EGFL7 in human cancer cells. Biochem Biophys Res Commun. 2009 Feb 13;379(3):726–31.

259. Aure MR, Fleischer T, Bjørklund S, Ankill J, Castro-Mondragon JA, Børresen-Dale AL, et al. Crosstalk between microRNA expression and DNA methylation drives the hormone-dependent phenotype of breast cancer. Genome Med. 2021 Apr 29;13:72.

260. Klinge CM. miRNAs regulated by estrogens, tamoxifen, and endocrine disruptors and their downstream gene targets. Molecular and Cellular Endocrinology. 2015 Dec 15;418:273–97.

261. Ma L, Li C, Yin H, Huang J, Yu S, Zhao J, et al. The Mechanism of DNA Methylation and miRNA in Breast Cancer. Int J Mol Sci. 2023 May 27;24(11):9360.

262. Cittelly DM, Das PM, Spoelstra NS, Edgerton SM, Richer JK, Thor AD, et al. Downregulation of miR-342 is associated with tamoxifen resistant breast tumors. Mol Cancer. 2010 Dec 20;9:317.

263. Egeland NG, Lunde S, Jonsdottir K, Lende TH, Cronin-Fenton D, Gilje B, et al. The Role of MicroRNAs as Predictors of Response to Tamoxifen Treatment in Breast Cancer Patients. Int J Mol Sci. 2015 Oct 14;16(10):24243–75.

264. Persson H, Kvist A, Rego N, Staaf J, Vallon-Christersson J, Luts L, et al. Identification of new microRNAs in paired normal and tumor breast tissue suggests a dual role for the ERBB2/Her2 gene. Cancer Res. 2011 Jan 1;71(1):78–86.

265. Newie I, Søkilde R, Persson H, Jacomasso T, Gorbatenko A, Borg Å, et al. HER2-encoded mir-4728 forms a receptor-independent circuit with miR-21-5p through the non-canonical poly(A) polymerase PAPD5. Sci Rep. 2016 Oct 18;6:35664.

266. Gong C, Yao Y, Wang Y, Liu B, Wu W, Chen J, et al. Up-regulation of miR-21 mediates resistance to trastuzumab therapy for breast cancer. J Biol Chem. 2011 May 27;286(21):19127–37.

267. Rui T, Xiang A, Guo J, Tang N, Lin X, Jin X, et al. Mir-4728 is a Valuable Biomarker for Diagnostic and Prognostic Assessment of HER2-Positive Breast Cancer. Front Mol Biosci. 2022;9:818493.

268. Floros KV, Lochmann TL, Hu B, Monterrubio C, Hughes MT, Wells JD, et al. Coamplification of miR-4728 protects HER2-amplified breast cancers from targeted therapy. Proc Natl Acad Sci U S A. 2018 Mar 13;115(11):E2594–603.

269. Newie I, Søkilde R, Persson H, Grabau D, Rego N, Kvist A, et al. The HER2-encoded miR-4728-3p regulates ESR1 through a non-canonical internal seed interaction. PLoS One. 2014;9(5):e97200.

270. Pekow J, Hutchison AL, Meckel K, Harrington K, Deng Z, Talasila N, et al. miR-4728-3p functions as a tumor suppressor in ulcerative colitis-associated colorectal neoplasia through regulation of focal adhesion signaling. Inflamm Bowel Dis. 2017 Aug;23(8):1328–37.

271. Schmitt DC, Madeira da Silva L, Zhang W, Liu Z, Arora R, Lim S, et al. ErbB2-intronic microRNA-4728: a novel tumor suppressor and antagonist of oncogenic MAPK signaling. Cell Death Dis. 2015 May 7;6(5):e1742.

272. Schröder R, Illert AL, Erbes T, Flotho C, Lübbert M, Duque-Afonso J. The epigenetics of breast cancer – Opportunities for diagnostics, risk stratification and therapy - PMC. Epigenetics. 17(6):612–24.

273. Bodelon C, Ambatipudi S, Dugué PA, Johansson A, Sampson JN, Hicks B, et al. Blood DNA methylation and breast cancer risk: a meta-analysis of four prospective cohort studies. Breast Cancer Research. 2019 May 17;21(1):62.

274. Sharma D, Saxena NK, Davidson NE, Vertino PM. Restoration of Tamoxifen Sensitivity in Estrogen Receptor–Negative Breast Cancer Cells: Tamoxifen-Bound Reactivated ER Recruits Distinctive Corepressor Complexes - PMC. Cancer Res. 2006 Jun 15;66(12):6370–8.

275. Garcia M, Derocq D, Freiss G, Rochefort H. Activation of estrogen receptor transfected into a receptor-negative breast cancer cell line decreases the metastatic and invasive potential of the cells. Proc Natl Acad Sci U S A. 1992 Dec 1;89(23):11538–42.

276. Rao X, Di Leva G, Li M, Fang F, Devlin C, Hartman-Frey C, et al. MicroRNA-221/222 confers breast cancer fulvestrant resistance by regulating multiple signaling pathways. Oncogene. 2011 Mar 3;30(9):1082–97.

277. Katchy A, Edvardsson K, Aydogdu E, Williams C. Estradiol-activated estrogen receptor α does not regulate mature microRNAs in T47D breast cancer cells. J Steroid Biochem Mol Biol. 2012 Feb;128(3–5):145–53.

278. Kaminska K, Akrap N, Staaf J, Alves CL, Ehinger A, Ebbesson A, et al. Distinct mechanisms of resistance to fulvestrant treatment dictate level of ER independence and selective response to CDK inhibitors in metastatic breast cancer. Breast Cancer Res. 2021 Feb 18;23(1):26.

279. Nygard AB, Jørgensen CB, Cirera S, Fredholm M. Selection of reference genes for gene expression studies in pig tissues using SYBR green qPCR. BMC Mol Biol. 2007 Aug 15;8:67.

280. Panda AC, Martindale JL, Gorospe M. Polysome Fractionation to Analyze mRNA Distribution Profiles. Bio Protoc. 2017 Feb 5;7(3):e2126.

281. Illumina. An Introduction to Next-Generation Sequencing Technology [Internet]. 240818 p. 16. Available from: www.illumina.com/technology/next-generation-sequencing.html

282. Tost J, Gut IG. DNA methylation analysis by pyrosequencing. Nat Protoc. 2007 Sep;2(9):2265–75.

283. Ronaghi M. Pyrosequencing Sheds Light on DNA Sequencing. Genome Res. 2001 Jan 1;11(1):3–11.

284. Pyrosequencing Technology and Platform Overview [Internet]. [cited 2024 Aug 18]. Available from: http://www.qiagen.com/us/knowledge-and-support/knowledge-hub/technology-and-research/pyrosequencing-resource-center/pyrosequencing-technology-and-platform-overview

285. Ronaghi M, Uhlén M, Nyrén P. A sequencing method based on real-time pyrophosphate. Science. 1998 Jul 17;281(5375):363, 365.

286. Ermini L, Driguez P. The Application of Long-Read Sequencing to Cancer. Cancers. 2024 Jan;16(7):1275.

287. Saal LH, Vallon-Christersson J, Häkkinen J, Hegardt C, Grabau D, Winter C, et al. The Sweden Cancerome Analysis Network - Breast (SCAN-B) Initiative: a large-scale multicenter infrastructure towards implementation of breast cancer genomic analyses in the clinical routine. Genome Med. 2015 Feb 2;7(1):20.

288. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature. 2012 Oct 4;490(7418):61–70.

289. Kornaga EN, Klimowicz AC, Guggisberg N, Ogilvie T, Morris DG, Webster M, et al. A systematic comparison of three commercial estrogen receptor assays in a single clinical outcome breast cancer cohort. Mod Pathol. 2016 Aug;29(8):799–809.

290. Nicholson RI, Hutcheson IR, Knowlden JM, Jones HE, Harper ME, Jordan N, et al. Nonendocrine pathways and endocrine resistance: observations with antiestrogens and signal transduction inhibitors in combination. Clin Cancer Res. 2004 Jan 1;10(1 Pt 2):346S-54S.

291. Bisulfite-conversion-acgautac [Internet]. [cited 2024 Aug 18]. Available from: https://www.diagenode.com/img/categories/bisulfite-conversion/bisulfite-conversion-acgautac.png

292. Human estrogen receptor-related protein (variant ER from breast cancer) mRNA, complete cds [Internet]. 1994 [cited 2024 Aug 13]. Available from: http://www.ncbi.nlm.nih.gov/nuccore/M69297.1

293. Balcazar Lopez CE, Albrecht J, Hafstad V, Freitag C, Vallon-Christersson J, Bellodi C, et al. Alternative promoters and splicing create multiple functionally distinct isoforms of oestrogen receptor alpha in breast cancer and healthy tissues - Balcazar Lopez - 2023 - Cancer Medicine - Wiley Online Library. 230912;12(18):18931–45.

Printed by Media-Tryck, Lund 2024 NORDIC SWAN ECOLABEL 3041 0903

LUND UNIVERSITY

FACULTY OF MEDICINE

Department Clinical Sciences, Lund

Lund University, Faculty of Medicine
Doctoral Dissertation Series 2024:123
ISBN 978-91-8021-620-3
ISSN 1652-8220