EFFECTS OF PHOSPHORYLATION ON THE DNA BINDING PROPERTIES OF ESTROGEN RECEPTOR-ALPHA

by

Kyle T. Helzer

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Cancer Biology)

At the

UNIVERSITY OF WISCONSIN-MADISON

2019

Date of final oral examination: 08/21/2019

The dissertation is approved by the following members of the Final Oral Committee:

Shigeki Miyamoto, Professor, Department of Oncology
J. Wesley Pike, Professor, Department of Biochemistry
William Sugden, Professor, Department of Oncology
Beth Weaver, Professor, Department of Cell and Regenerative Biology

EFFECTS OF PHOSPHORYLATION ON THE DNA BINDING PROPERTIES OF ESTROGEN RECEPTOR-ALPHA

Kyle T. Helzer
Under the supervision of Professor Elaine T. Alarid
At the University of Wisconsin-Madison

ABSTRACT

Estrogen receptor-α (ER) is a major driver of breast cancer growth and development making it a common target for therapies. Understanding the activation and control of the ER signaling pathway is critical for developing new breast cancer treatments as well as combating resistance to current treatments. Activation of ER can occur in a ligand-dependent manner by 17ß-estradiol (E2) or in a ligand-independent manner through various growth factors or other stimuli. A common feature to both signaling pathways is the induction of ER phosphorylation at serine 118 (pS118-ER). Previous work has found that pS118-ER is an important modulator of ER-dependent gene transcription. A crucial step in the ER signaling pathway which allows for the altering of target gene expression is the binding of ER to DNA. While much work has been done to define the genome-wide binding profile (cistrome) of ER, the effect of phosphorylation on the distribution of ER across the genome is not well understood. In the work presented here, I aim to define the cistrome of pS118-ER and identify defining characteristics of sites occupied by the phosphorylated receptor compared to all ER occupancy sites. Using a comprehensive approach with multiple pS118-ER specific antibodies, I find that pS118-ER occupies a subset of ER sites in the human genome, with these sites being enriched for enhancer marks as well as the DNA-binding motif for the transcription factor

grainyhead-like 2 (GRHL2). Additionally, analysis of these sites on *in vitro* DNA-binding arrays revealed an association between pS118-ER and direct DNA-binding. Finally, I developed a mutant MCF-7 cell line expressing ER with an S118A mutation to interrogate the biology of pS118-ER in a cell line which is dependent on ER. Consistent with previous reports, I found that E2-dependent ER downregulation was impaired in the S118A ER mutant. Interestingly, the downregulation of the *ESR1* gene in response to E2 was also prevented indicating the control of pS118-ER over ER levels is not limited to post-translational mechanisms. These studies further the knowledge of how phosphorylation affects ER biology and may serve as a target for future breast cancer therapies.

TABLE OF CONTENTS

LIST OF TABLES

Chapter One

Chapter Six

Appendix A

Appendix D

LIST OF FIGURES

Chapter One

Chapter Two

Chapter Three

Chapter Four

Chapter Five

Appendix A

Appendix B

Appendix C

Appendix D

ACKNOWLEDGEMENTS

The thesis work presented here and my journey through my PhD could not have been done without the support and encouragement of so many friends, family members, colleagues, and collaborators. First and foremost, I would like to thank my advisor, Dr. Elaine Alarid for taking me on as a student, guiding me though graduate school, and assisting my growth as a scientist. Her endless enthusiasm for science helped me through some hard times and encouraged me to be a better researcher. She allowed me to follow my own curiosities and supported me in those endeavors and for that I am eternally grateful.

A huge thanks goes out to the members of the Alarid lab, especially Natasha Solodin, who was the bedrock of the group and without whom we would all be perpetually lost. I am also thankful for her encouraging me and others to enjoy life by getting out of the lab from time to time (you were always right). To the two other graduate students in the lab: David Lung, my partner in crime, has been through the whole graduate school process with me since the beginning from when we interviewed to when we join the lab together all the way to today. I am very glad I had someone supporting me from the beginning to talk to and discuss ideas and problems because I don't think I could have done this alone. To Becky Reese, who, despite seeing these two other strange graduate students flounder around in lab, decided to join us in our shenanigans and has since been so incredibly encouraging with everything, both work and life related. Her seemingly endless optimism kept us together and truly made it a joy to come in to lab every day. I can honestly say that both of you feel like my family now and I cannot imagine graduate school without you two.

I would also like to thank past members of the lab, especially Dr. Jessi Lang, for helping me out when I was a lowly first-year student trying to find my way in the lab. Additional thanks goes out to all the undergrads who worked in the lab in my time here: Kelsi Bjorklund, Brant Bickford, Allie Kissel, Nina Anbouba, and Simona Khomutov—your help in the lab was vital for keeping the lab up and running.

I would like to thank all of my collaborators, whose expertise was crucial for competition of this work. A special thanks goes out to the members of the Pike lab, Nancy Benkusky for helping out with the ChIP-seq experiments and to Dr. Mark Meyer whose guidance in all things ChIP-seq related was indispensable for the success of this work. Additional thanks go to Dr. Chris Hooper of the Miyamoto lab for helping me write my first paper and being patient with me throughout the process. I would also like to extend my gratitude to my thesis committee: Dr. Beth Weaver, Dr. Wes Pike, Dr. Shigeki Miyamoto, and Dr. Bill Sugden for the helpful discussion, feedback, and guidance on my project.

There were so many friends that I made along the way whose company has made graduate school so much more bearable. Thanks to the members of the surrounding labs, Dr. Mailee Huynh, Dr. Chris Hooper, Dr. Dominique Lisiero, Dr. Amy Fowler, Dr. Kelley Salem, Manoj Kumar, Tina Mark, Debayan De Bakshi, Hae Yeun Chang, and many others for keeping our floor a fun and interesting place to work. Thanks to the other students in the Cancer Biology program here at UW, especially Dr. Mailee Huynh, Dr. Jordan Becker, Josh Martin, Ed Evans, Dr. Kathleen Makielski, Dalton McLean, Jordan Vellky, and Tamara Rodems for being a fun group of people to hang out and de-stress with. I would especially like to thank the other students in my cohort: David Lung, Ed Evans, and Kathleen Makielski. I couldn’t ask for a better group of people to go through

this journey together, and I am confident they all will be successful in whatever they put their minds to, because that's just what they do. I would also like to thank my childhood friends who have been with me since the beginning: Brian Akers, Andrew Darling, Binh Tran, and Jon Manby. You guys are the best group of people I could even imagine and were always encouraging and enthusiastic about whatever I was doing in life. Although we are all living in different parts of the country now, I still feel a very strong connection to all of you and that is something for which I am forever grateful.

So much love and thanks go out to my family, who has never doubted me any second of the way, even when I doubted myself. My parents, Tom and Connie Helzer, who have unconditionally supported my growth in all aspects of life and provided a rock of encouragement from which I build all my aspirations and goals. To my amazing brothers, Brett and Alan Helzer, who are so incredibly supportive and caring and have been there for me since the beginning. So much thanks go to my grandparents, who are always very curious and supportive of my work and made the journey all the way out to Madison to watch me graduate. I love you all so much.

And last, the person who deserves the most thanks for this whole process is my wonderful girlfriend Anna Cacciaglia. Your love and support lifted me through the hard times in graduate school and I am so ever grateful for your patience and understanding. You were willing to join me on the journey to Madison even though it was terrifying to leave the comfort of our home state and start a new life in a different city. You helped me throughout so many bleak times which I know I could not have weathered alone and encouraged me when I doubted myself. I cannot imagine a world where I make it through

graduate school without you and I cannot wait to start a new chapter of life together with you.

ABBREVIATIONS

AF-1	activation function 1
AF-2	activation function 2
AIB1	Amplified in breast cancer 1 / NCoA3
AP-1	Activator Protein 1
AP-2γ	Activating Enhancer-Binding Protein 2 γ
AR	androgen receptor
bp	base pair
BSA	bovine serum albumin
CAK	cyclin activating kinase
Cas9	CRISPR associated protein 9
CAT	chloramphenicol acetyltransferase
CDK7	cyclin dependent kinase 7
ChIP	Chromatin Immunoprecipitation
ChIP-chip	Chromatin Immunoprecipitation on a chip
ChIP-qPCR	Chromatin Immunoprecipitation quantitative PCR
ChIP-seq	Chromatin Immunoprecipitation followed by sequencing
COUP-TF1	Chicken Ovalbumin Upstream Promoter-Transcription Factor 1
CRISPR	clustered regularly interspaced short palindromic repeats
CTD	C-terminal domain
DBD	DNA-binding domain
DNA-PK	DNA-dependent protein kinase
DMEM	Dulbecco’s Modified Eagle Medium
E2	17ß-estradiol
E6-AP	E6-associated protein
EDTA	ethylenediaminetetraacetic acid
EGF	epidermal growth factor

ENCODE	Encyclopedia of DNA Elements
ER	estrogen receptor α (protein)
ERß	estrogen receptor ß
ERBS	estrogen receptor binding site
ERE	estrogen response element
ESR1	estrogen receptor α (gene)
EtOH	ethanol
FM	full media
FOXA1	forkhead box A1
FOXM1	forkhead box M1
GATA3	GATA binding protein 3
GO	gene ontology
GR	glucocorticoid receptor
GREB1	Growth Regulating Estrogen Receptor Binding 1
GRHL2	grainyhead-like protein 2
GRIP1	glucocorticoid receptor-interacting protein-1
GRO-seq	global run-on sequencing
GSK-3	Glycogen Synthase Kinase 3 ß
h	hours
HDR	homology directed repair
H3K4me1	histone 3 lysine 4 monomethylation
H3K27ac	histone 3 lysine 27 acetylation
HCl	hydrochloric acid
HOMER	Hypergeometric Optimization of Motif EnRichment
HRP	horseradish peroxidase
ICI	ICI-182,780
IgG	immunoglobulin G
IHC	immunohistochemistry

IKKα	Inhibitor of Nuclear Factor Kappa-B Kinase Subunit α
kb	kilobases
kDa	kilodalton
LBD	ligand-binding domain
LTED	long term estrogen depravation
luc	luciferase
MAPK	mitogen-activated protein kinase
MAT1	ménage à trios 1
MCF-7	Michigan Cancer Foundation 7
MegaTrans	mega transcription factor-bound in trans
MR	mineralocorticoid receptor
MW	molecular weight
NaCl	sodium chloride
NCoA1	nuclear receptor coactivator 1
NR	nuclear receptor
p300	E1A binding protein p300
PAM	protospacer adjacent motif
PBS	phosphate buffered saline
PCR	polymerase chain reaction
Pin1	Peptidylprolyl Cis/Trans Isomerase, NIMA-Interacting 1
PR	progesterone receptor
pS118-ER	estrogen receptor α phosphorylated at serine 118
PTM	post-translational modification
qPCR	quantitative real-time PCR
RIME	Rapid Immunoprecipitation Mass spectrometry of Endogenous proteins
RNP	ribonucleoprotein
Runx1	RUNX Family Transcription Factor 1

SBPB	stromelysin-1 platelet-derived growth factor-responsive element-binding protein
SDS-PAGE	sodium dodecyl sulfate–polyacrylamide gel electrophoresis
siRNA	small interfering RNA
sgRNA	short guide RNA
SNAP	Specificity and Affinity for Proteins
Sp1	specificity protein 1
SRC	steroid receptor coactivator
ssODN	single-stranded donor template
Tam	Tamoxifen / 4-hydroxytamoxifen
TBS	tris buffered saline
TFF1	trefoil factor 1
TFIIH	General Transcription Factor IIH

CHAPTER ONE:

Introduction

Overview of Breast Cancer and its Treatment History

Breast cancer is a major health issue worldwide. In 2016, breast cancer was the leading cancer in terms of new cases diagnosed (124.2 per 100,000 people) out of all cancer cases, and had the second highest rate of death (20.0 per 100,000 people) in the United States in the same time period (1). In 2018 it was estimated that 2,088,849 new cases of breast cancer were diagnosed worldwide, accounting for 11.6% of cancers at all sites (2). Of these cases of breast cancer, approximately 70% express estrogen receptor-alpha (ER) (3,4). Due to its control over growth and development of estrogen-responsive tissue as well as breast cancer, ER quickly became a target of interest for breast cancer therapy. This focus led to the development of the antiestrogen tamoxifen (originally named ICI 46,474 and now sold under the brand name Nolvadex), which is still in use today for the treatment and prevention of ER-positive breast cancers (5). However, resistance is still very common and further research on the ER signaling pathway is needed to identify alternative modes of attack for the treatment of breast cancer.

Early Breast Cancer Research

The connection between a secreted factor, now known to be estrogen, and the growth of breast cancer has been known for some time. In the late 1800s, British physician George Beatson, after observing that cows produced milk longer when their ovaries were removed after calving, hypothesized a connection between the ovaries and the growth of breast cancers stating, “it pointed to one organ holding the control over … another and separate organ” (6). Beatson also observed a similarity between the histology in the development of mammary glands during lactation and the histology of

breast carcinomas noting, “lactation is at one point perilously near becoming a cancerous process if it is at all arrested.” (6). Beatson hypothesized that the ovaries may exert some control over the growth of breast cancers, and to test this theory performed an oophorectomy on two women with inoperable breast tumors. He observed a decrease in the size of the tumors after the surgery in both patients, a result which quickly led to other physicians adopting his strategy. However, it was soon discovered that this surgery was not effective in all cases, indicating that the physiology of breast cancers was more complex than originally thought (7). Evidence for a link between secreted hormones and cancer was further strengthened when mice castrated at less than six months of age were shown to have a decreased incidence of tumors compared to control mice (8). The isolation of the first hormone, estrone, was performed by its extraction from urine in the late 1920s, and it wouldn’t be until 1940 when estradiol (originally termed α-dihydrotheelin) was purified from human placenta (9–12). In the 1960s, work by Elwood V. Jensen showed that tritium-labeled estradiol would stay bound to the uterus and vagina, but not to other non-estrogen responsive tissues such as the muscle or liver—the first evidence that there existed a “receptor” for estradiol and that it was also tissue-specific (13). Subsequent studies showed that the bound estradiol in the uterus could be released by proteases, indicating that the receptor for estradiol was a protein (14). The estrogen receptor (now known to be estrogen receptor-alpha) was soon isolated by Toft and Gorski in 1966, and its gene, *ESR1,* was cloned by Green and Chambon in 1986 (15,16). These advancements, along with the isolation of the strongly ER-positive breast cancer cell line MCF-7, led to a substantial expansion of research on the role of ER in breast cancer biology (17–19).

General ER background

ER is a member of the steroid nuclear receptor (NR) family of proteins which also comprises, estrogen receptor-ß (ERß), androgen receptor (AR), progesterone receptor (PR), glucocorticoid receptor (GR), and mineralocorticoid receptor (MR). While the amino acid sequences vary drastically between NRs, they all contain six distinct domains (labeled A-F) which all serve different functions but are conserved between NRs (20,21). For ER (Figure 1-1), the N-terminal A/B domain contains the activation function 1 (AF-1) which controls activation of the receptor in the absence of ligand. This domain is unstructured compared to the other parts of the protein and can function independently to activate the receptor (22,23). The C domain is located roughly in the center and contains the DNA-binding domain (DBD) (24,25). The D domain, also called the hinge region, serves as a flexible linker between the N-terminal AF-1/DBD domains and the C-terminal E domain, which contains the activation function 2 (AF-2) or ligand-binding domain (LBD) (26). Finally, at the C-terminus is the short F domain, which varies in sequence homology between NRs and whose function is still not well understood (Figure 1-1). Generally, NRs are activated through binding of their respective ligand to the LBD, followed by binding to DNA at specific response elements and activation of gene transcription. Unlike other NRs, ER is mainly localized to the nucleus in an unliganded state, through its constitutive nuclear localization signal (27,28). In the case of ER, binding of its major ligand, 17ß-estradiol (E2), leads to a large rearrangement of the tertiary structure of the receptor, which in turn leads to its homodimerization (29,30) and subsequent binding to estrogen response elements (EREs) on DNA (Figure 1-2). This

structural alteration includes the repositioning of a major α-helix at the C-terminal end (commonly referred to as “helix 12”) and allows for the interaction of E2-bound ER with various coactivators (31–33). Binding of tamoxifen to the ligand-binding pocket of ER not only outcompetes E2 for binding, but also causes helix 12 to adopt a position that blocks the interaction of these cofactors, which contributes to its efficacy as an ER antagonist (33). Ligand-dependent activation of ER through the AF-2 domain has been a major focus of drugs for ER, as this is the canonical pathway through which ER is activated. However, ER can also be activated independently through the N-terminal AF-1 domain, mainly through phosphorylation of serine residues (22).

Protein Phosphorylation

Once translated, proteins serve many functions and adopt numerous confirmations to perform their specific tasks. These conformational changes can be caused by interactions with other proteins or by direct covalent modification of the protein through a variety of different molecules. These post-translational modifications (PTMs) are often reversible and serve as switches to allow control over a protein’s function at specific times. ER has been shown to be extensively modified by PTMs, including phosphorylation (34), acetylation (35), methylation (36,37), ubiquitylation (38), palmitoylation (39), and sumoylation (40). Of these modifications, phosphorylation has been the most widely studied in ER (41). Protein phosphorylation—the addition of a phosphoryl group (PO_3^{2-}) to a protein—most commonly occurs on the amino acids serine, threonine, and tyrosine due to the accessible hydroxyl group in their respective side chains. Phosphorylation of these residues adds a negative charge to that region of the protein which can lead to

shifts in its tertiary structure. Phosphorylation of proteins can lead to new protein-protein interactions (42), alterations in sub-cellular localization (43,44), activation of enzymatic activity (45), or alterations in protein stability (46,47). Importantly, phosphorylation is a reversible process, allowing the phosphorylated function of a protein to be switched on or off as needed. A specific group of proteins called kinases performs phosphorylation, and dephosphorylation is performed by phosphatases. Recent mass spectrometry studies have estimated that 75% of cellular proteins can be phosphorylated (48,49), a finding which emphasizes the major role of phosphorylation in cellular biology.

ER Phosphorylation

ER was first found to be phosphorylated by incubating mouse uterus cells with ^{32}P-labeled orthophosphate followed by ER immunoprecipitation (50). Similar results were found in the human breast cancer cell line MCF-7 (34). When the cells were incubated with E2, a drastic increase in ER phosphorylation was observed, indicating that phosphorylation of ER was an E2-dependent process (34,50). Kinetic studies in MCF-7 cells revealed that E2-dependent phosphorylation occurs very rapidly, with increases in phosphorylation seen as early as two minutes post E2 treatment (34). Mutational analysis later found that serine 118 (S118) was a major site of phosphorylation in response to E2, with three other serine sites (S104, S106, and S167) also identified as significant phosphorylation substrates (51–53). More recently, mapping of ER PTMs by mass spectrometry and other methods has identified many additional phosphorylation sites (Y52, S102, S154, S212, Y219, S236, S282, S294, S305, T311, Y537, S554, S559; Figure 1-1; Table 1-1), however many of these phosphorylation events are in low

abundance and their biological significances are not well understood (54–56). Phosphorylation of ER at serine 118 (pS118-ER) has been a major focus of research due to its strong induction and its functional role in ER regulation through the AF-1 region.

Stimulation of pS118-ER and related kinases

Since the identification of S118 as a site of phosphorylation on ER, many studies have been performed to determine its biological function. As many of the phosphorylation sites in ER fall in the AF-1 region, it was hypothesized that phosphorylation exerted control over AF-1-driven transcriptional activity. It was found very early on that pS118-ER could be induced not only by E2, but also by ER antagonists such as tamoxifen and ICI-164384 (a precursor to fulvestrant, also known as ICI-182780, which is more commonly used today) (51). The induction of pS118-ER by tamoxifen and ICI, but lack of gene activation suggested that pS118-ER alone is not sufficient to control AF-1 driven gene transcription, indicating that other factors contributed. Additionally, certain factors that activate the mitogen-activated protein kinase (MAPK) pathway, such as epidermal growth factor (EGF) and insulin-like growth factor (IGF-1) can induce pS118-ER in a ligand-independent manner (57–59). Both of these treatments lead to activation of ER as measured by ERE-reporter constructs (57,58). Interestingly, activation of the MAPK pathway through overexpression of the MAPK pathway protein K-Ras can activate a truncated form of ER containing only the AF-1 and DBD domains in the absence of E2, but cannot activate an ER containing only the DBD and AF-2 domain indicating that these activation pathways work primarily through phosphorylation of the ER AF-1 domain (57). Numerous stimuli have been reported to induce pS118-ER in a ligand-independent

manner, including prolactin (60), heregulin (61), leptin (62), reactive oxygen species (63), and hypoxia (64), however these activation pathways have been less extensively studied.

Many kinases have been observed to phosphorylate ER at S118 as well as at other sites depending on the stimulus and the cell type used (Table 1-1). While activation of the MAPK pathway was originally found to be responsible for activation of ER in a ligand-independent manner through pS118-ER, studies also found that MAPK was not activated upon treatment with E2, suggesting that other kinases were involved in the phosphorylation of ER in an E2-dependent manner (65). Casein kinase II was found to phosphorylate the AF-1 domain of ER in response to E2 *in vitro*, however, the specific serine residue was not determined and it is unknown whether this kinase acts on ER *in vivo* (66). The first kinase discovered to phosphorylate ER at S118 in an E2-dependent manner was cyclin-dependent kinase-7 (Cdk7) (67). Cdk7 is a component of the cyclin activating kinase (CAK) complex, which comprises Cdk7, cyclin H, and ménage à trios 1 (MAT1). CAK itself is a subcomplex of the larger general transcription factor TFIIH and is responsible for phosphorylating the C-terminal domain (CTD) of RNA polymerase II (68). Overexpression of Cdk7 in COS-1 cells was found to enhance ER transactivation in the presence of E2 and the S118A ER mutation abolished the Cdk7-dependent enhancement. ER was also shown to be phosphorylated at S118 by Cdk7 *in vitro* and that this phosphorylation required interaction of ER with TFIIH (67). Further investigations found that MAPK was the primary kinase involved in ligand-independent activation of ER, and Cdk7 was responsible for ligand-dependent activation (69). The ligand-independent phosphorylation of ER at S118 was also found to depend on the interaction of ER with the Chicken Ovalbumin Upstream Promoter-Transcription Factor 1 (COUP-TF1; NR2F1),

which increases the affinity of ER for MAPK (70). Additionally, glycogen synthase kinase 3 (GSK-3) was shown to phosphorylate S118 as well as two other sites, S104 and S106, in response to E2, and an ER-GSK-3 interaction was demonstrated through co-immunoprecipitation, indicating that Cdk7 may not be the sole kinase responsible for pS118-ER (71). However, these studies were performed in a different cell line (MELN cells) than the Cdk7 studies and the results may represent variation in kinase activity between cell lines. ER can also be phosphorylated at S118 in an E2-dependent manner by the IκB kinase (IKK) complex subunit IKKα, which also associates with ER target gene promoters (72). Knockdown of IKKα using siRNA abolished E2-induced pS118-ER and reduced E2-induced ER target gene expression (72). More recently, there has been evidence for the phosphorylation of ER at S118 by the DNA-dependent protein kinase (DNA-PK) holoenzyme, which requires DNA binding for its kinase activity (73,74). The interaction between ER and DNA-PK forms upon E2 stimulation, and pS118-ER led to stabilization of ER protein (73). Mass spectrometry of an ER complex bound to DNA also found DNA-PK as one of the cofactors associated with ER on DNA (74). A full list of the stimuli and kinases known to phosphorylate ER at S118 is shown in Table 1-1.

Effect of pS118-ER on Gene Transcription

Activation of ER through E2 leads to a rapid and profound change in cellular gene transcription. Large sequencing studies of ER-dependent transcription have found that approximately 26% of all annotated genes (6003/22893) are altered by E2 treatment (75). Considering the vast gene alterations occurring, control of ER activation is critical for cellular responses to ligand and other stimuli. As noted previously, ER activity is not solely

modulated by ligand, but can further be fine-tuned by PTMs such as phosphorylation. Many early studies that investigated ER gene activation used an ERE-reporter system, with the reporter gene typically being chloramphenicol acetyltransferase (CAT) or luciferase (luc). When S118 was first discovered to be the major phosphorylation site on ER, its mutation to an alanine led to a decrease in ERE-CAT activity to 60% of wild-type ER in both COS-1 and HeLa cells (52). Using a reporter system constructed from the promoter of the ER-target gene *TFF1* connected to a CAT reporter, the S118A mutation decreased CAT activity to 60% and 25% of wild-type in COS-1 and HeLa cells, respectively, indicating that the effect and magnitude of pS118-ER on ER gene transactivation may depend on the cell type (52). Interestingly, an ER truncation containing the AF-1 and DBD domains alone was sufficient to induce transcription of an ERE-CAT reporter in COS-1 cells, but not in HeLa cells, further highlighting the importance of cell-type context in AF-1-driven ER activity (52). Additionally, activation of the MAPK pathway does not rescue the decrease in ERE-CAT activity seen in the S118A mutation, indicating that MAPK alters ER function through S118 (57,76). Numerous other groups have confirmed the reduction in maximal ER activation in the S118A mutant in both ERE-reporter constructs and endogenous ER-target genes (59,77–81). Intriguingly, mutation of S118 to glutamic acid, which mimics the charge of a phosphate group, causes greater ER activation in response to ligand compared to the wild-type receptor, suggesting that the charge introduced by pS118-ER is key for its action (77,78).

Coactivator Interactions Dependent on pS118-ER

Significant work has investigated the mechanism for pS118-ER control over gene regulation. Given the alteration in electrostatic charge by the addition of a phosphate group, it was hypothesized that phosphorylation of ER at S118 would lead to altered protein-protein interactions which affect ER transactivation. Indeed, ER was shown to interact with the p160 coactivator family of proteins and this interaction could be modified by mutations in S118 and S104/S106 (77). The p160 coactivator family consists of three members: nuclear receptor coactivator-1 (NCoA1), glucocorticoid receptor-interacting protein-1 (GRIP1), and amplified in breast cancer-1 (AIB1, also known as NCoA3). Each has been shown to activate and enhance the transactivation of various NRs (82–87). Additionally, the transcriptional coactivator p300 interacts with ER and forms a complex along with the members of the p160 coactivator family to enhance ER transactivation (88). While these coactivators mainly interact with the AF-2 domain of ER in a ligand-dependent manner (33,89,90), they have also been shown to interact with the ER AF-1 domain to enhance its activity (91–95). The ER S118A mutation reduces the physical interaction of the p160 proteins with full-length receptor and diminishes their ability to enhance ER activity. Interestingly, inserting the S118A mutation into a truncated form of ER containing only the AF-1 and DBD domains did not affect the interaction between the truncated ER and the p160 proteins, but still had an effect on ER activation suggesting that the S118A mutation modulates p160 function rather than physical interaction and that both the AF-1 and AF-2 domains exert control over ER-p160 interactions (77). Studies have shown that the AF-1 and AF-2 domains of ER can interact with each other to modulate their functions and that this interaction can be mediated by one of the p160

family members, NCoA1 (95,96). The S118A mutation reduces the ability of NCoA1 to assist interaction between the AF-1 and AF-2 domains, suggesting a role for pS118-ER in allosteric interactions (77). Additionally, introducing serine to alanine mutations in S104 and S106 in ER further reduced the p160 enhancement activity on ER, indicating that multiple phosphorylation sites on ER could be working in tandem to modulate these interactions (77).

Protein-protein interactions mediated by pS118-ER are not limited solely to coactivators. A phage display screen using S118E ER as the bait protein discovered stromelysin-1 platelet-derived growth factor-responsive element-binding protein (SPBP) as an enriched interactor with S118E ER (97). Interaction of SBPB with ER was found to be E2-dependent and was abolished in the S118A mutant. Interestingly, SBPB was found to inhibit ER activity in response to E2 or EGF, but the interaction between ER and SBPB alone was not sufficient to perform its repressive activity (97). Overexpression of SBPB inhibited the proliferation of MCF-7 cells in response to E2 whereas the ER-negative cell line SK-BR-3 was not affected (97).

pS118-ER Control of ER Protein Levels

Another mode of modulation for ER transcriptional activity is the regulation of ER protein levels. Studies from our group found that increasing ER concentration in the absence of ligand is sufficient to induce ER transcriptional activity (98), demonstrating that control of ER protein levels is crucial for modulating ER transactivation. To shut off ER activation, ER protein is ubiquitylated and degraded by the 26S proteasome upon E2 treatment (99–102). Additionally, ER downregulates its own gene, *ESR1*, upon E2

stimulation (103–106). Studies from our group found that pS118-ER is crucial for the proteasome-dependent degradation of ER by E2, as the S118A ER mutation prevents E2-dependent ER downregulation (78). A computational model of ER kinetics analyzing RNA and protein dynamics found that pS118-ER was critical in regulating *ESR1* expression, and consequently, cellular ER levels (107). ER was found to interact with Peptidylprolyl *cis/trans* Isomerase NIMA-interacting 1 (Pin1) in a pS118-ER-dependent manner, which led to isomerization of the S118-P119 bond in ER from a *cis* to *trans* confirmation (108). This isomerization activity increased AF-1 transcriptional activity in both the presence and absence of ligand and was independent of AF-2-dependent coactivators (108). Analysis of the ER AF-1 region via nuclear magnetic resonance (NMR) imaging with and without Pin1 showed an extensive conformational change in the presence of Pin1, suggesting the alteration in the S118-P119 bond had an effect on the tertiary structure of ER as a whole (108). Importantly, the interaction between Pin1 and pS118-ER prevented the interaction with and subsequent ubiquitylation of ER by the E3-ubiquitin ligase E6-associated protein (E6-AP), which lead to the stabilization of ER (109). Further investigations into the function of Pin1 on ER found that the isomerization of pS118-ER led to an increased affinity for binding to ERE, suggesting that Pin1 and pS118-ER may have an effect on ER-DNA binding (110).

ER-DNA binding studies

A critical component for the transcriptional functionality of ER and all nuclear receptors is their ability to bind DNA. Around the time ER was discovered and isolated, E2 was also found to associate with chromatin in the nucleus (111,112). Isolation of the

nuclear-bound E2 resulted in the discovery of an E2-protein complex bound to DNA (113–115), and it was later confirmed that binding of E2 to ER resulted in conformational changes which increased the affinity of ER for DNA (116). The DBD of each NR preferentially binds to specific DNA sequences, which allows for NRs to preferentially activate different sets of genes. Early studies on NR DBDs found that removing the endogenous ER DBD and replacing it with the GR DBD caused the chimeric receptor to activate GR target genes in response to E2 (16). Therefore, while binding of ligand to NRs activates their transcriptional function, DNA binding specificity and specific target gene activity are primarily controlled through the DBD. While less common, ER can also interact with DNA through a tethering mechanism, whereby ER interacts with a secondary factor such as AP-1 or Sp1 which binds to DNA (117–121). The sequence to which ER binds is called the estrogen response element (ERE), a 15-base pair (bp) DNA palindrome with the consensus sequence 5'-AGGTCAnnnTGACCT-3' where 'n' represents any nucleotide (122,123). ER can additionally bind to non-consensus sites, which can still have a strong effect on gene transcription (124). A more accurate depiction of the ERE can be demonstrated by a position weight matrix, which allows for the importance of each specific bp in the sequence to be determined and visualized, as very few ER genomic binding sites contain a perfect consensus ERE (125–127). Each monomer of the ER dimer binds to an individual "half-site" (5'-AGGTCA-3') with the ER dimer binding to the entire ERE (128–130). The 3-bp spacer between these half-sites in the ERE was found to be critical for ER binding, with modifications of the spacer to 2-bp or 0-bp causing a drastic decrease in affinity of ER for the ERE (129,131). There has additionally been some evidence that the sequences flanking the ERE may play a role in

altering the affinity of the ER DBD for DNA (123,132,133), however the ERE remains the major driver for directing ER-DNA interaction. Interestingly, the ER DBD alone cannot differentiate between the consensus ERE and more degenerate EREs, indicating that ER domains outside of the DBD can alter binding to sequences that deviate slightly from the consensus (134). Indeed, numerous reports have demonstrated that phosphorylation in the AF-1 region of ER can affect DNA binding in the adjacent domain (34,66,80,108,110,135,136).

When ER was first discovered to be a phosphoprotein, it was concurrently found that the phosphorylation affected DNA binding as measured by an electrophoretic mobility shift assay (EMSA) using a consensus ERE as the probe (34). However, these experiments were performed using a general phosphatase and therefore the specific phosphorylation site responsible for modulating DNA binding activity was not determined (34). Similarly, general phosphorylation of the AF-1 domain by casein kinase II (CKII) resulted in a two-fold increase in ER affinity for ERE binding and had no effect on the binding of ligand to the AF-2 domain (66). Early *in vitro* experiments analyzing the effects of specific phosphorylation sites on DNA binding suggested that pS118 had minimal effects on the ability of ER to bind to an ERE and that pS167 was primarily responsible for the increases in ER-ERE affinity (137). However, these experiments had a number of caveats. First, they were performed *in vitro* with purified ER and kinases, and no mutational analyses of ER were performed to confirm that either pS118 or pS167 was solely responsible for the observed results. Phosphorylation at S118 and S167 was confirmed by Edman degradation and was performed only in the regions immediately surrounding S118 and S167. It is unknown in these systems whether the serine residues

purported to be phosphorylated were the only sites phosphorylated by these kinases *in vitro*, as multiple phosphorylation sites on ER could act in tandem to affect ER function. For example, phosphorylation of Y537 in the AF-2 domain affects the binding of ER to EREs as well as ER dimerization, and Y537 is phosphorylated along with serine residues in the AF-1 domain in response to E2 (138). Second, the effect of pS118-ER on DNA binding was investigated only in the ligand-independent context via MAPK phosphorylation. Although it is well accepted that E2 induces phosphorylation at S118 (51,52,65,67,69,71–73,80,81,136,139), the authors reported no phosphorylation of S118 by CKII in the presence of E2 *in vitro*, only phosphorylation at S167. This result is most likely due to the use of CKII as the sole kinase in the *in vitro* assay, as ER can be phosphorylated at S118 by multiple kinases in response to E2 (67,69,71–73). While these experiments suggest that pS118-ER alone is not sufficient to induce DNA binding, other reports performed in cells have found that pS118-ER is important for maximal ER-DNA binding in response to E2 (80,110). Additionally, genome-wide analysis of ER binding sites later demonstrated that ligand-independent activation of ER through EGF and MAPK has profound effects on ER-DNA binding (140).

With the development and optimization of the chromatin immunoprecipitation (ChIP) assay, which allows for the investigation of protein occupancy on endogenous genomic locations (141,142), the interaction between ER and chromatin in its native environment became a major focus in the field. Upon E2 treatment, ER occupancy at the promoters of various ER target genes was found to increase which was accompanied by an increase in H3 and H4 acetylation (143,144). Along with ER, a number of coactivators were found to be recruited to these same promoter sites in an E2-dependent manner

(144). The recruitment of these coactivators is mediated by phosphorylation of ER at S118, as the S118A ER mutation alters binding of the coactivators at certain sites (80). Interestingly, E2 time course experiments revealed that ER and its associated cofactors cycle on and off these promoters in approximately 90 min cycles and that a sequential order of coactivator loading is followed to activation target gene transcription (144,145). However, these experiments were limited to two promoter sites (*TFF1* and *CATD*), and recent genome-wide analyses of ER occupancy has not been able to replicate the previously observed cycling (146). Nevertheless, it is still widely accepted that ER and its coactivators are recruited to ER genomic binding sites upon E2 treatment.

Genome-Wide ER Binding

With the sequencing of the human genome (147,148) and the advances made in sequencing technology, ChIP has been adapted into various high-throughput assays to detect protein-DNA interactions at a genome-wide level. The initial application for analyzing ER occupancy on a large scale was the ChIP-on-chip assay, which consisted of tiling select genomic regions on an array and hybridizing immunoprecipitated DNA after purification and florescent tagging (149–152). Due to size restrictions of the early arrays, the first ER ChIP-on-chip was limited to analyzing approximately 19,000 promoter regions rather than the entire human genome (152). This analysis found 153 promoters bound by ER in the presence of E2 and when these binding regions were analyzed by motif-finding algorithms, the DNA-binding motif for forkhead box A1 (FOXA1) was found to be enriched in the sites occupied by ER (152). Many of the promoters identified were associated with E2-dependent ER-target genes, but other genes were discovered to have no response to

E2 despite recruitment of ER to their respective promoters (152). Additional analysis of these promoter regions by performing ChIP-on-chip for nuclear receptor coactivators (NCoA1, GRIP1, AIB1, and p300) found a strong correlation between binding of ER and binding of the coactivators to promoters in response to E2 (153,154). However, given that ER binding sites have been shown to have effects on gene transcription over long genomic distances (155), additional ChIP-on-chip studies were performed to analyze large portions of the genome in an unbiased manner (156). Rather than selecting promoters, the sequences for chromosomes 21 and 22 were tiled on arrays and analyzed for ER occupancy in response to E2 (156). Across these two chromosomes, 57 ER occupancy sites were identified with many falling outside of promoter regions, indicating that ER binding is not limited to promoters (156). Similar to the promoter-restricted ChIP-on-chip, the ER occupancy sites found on chromosomes 21 and 22 were enriched for the DNA binding motif for FOXA1 (156). FOXA1 is a pioneer factor which alters chromatin from a closed to an open state, allowing other factors to then access and bind DNA (157). Subsequent ChIP-on-chip experiments examining FOXA1 binding sites found that FOXA1 was preferentially recruited to distal enhancers occupied by ER rather than proximal promoters (158). Additionally, knockdown of FOXA1 blocked the association of ER with DNA, suggesting that the pioneering function of FOXA1 is necessary for ER-DNA interaction (152). The first ER ChIP-on-chip spanning all chromosomes and encompassing 1,500 Mb of non-repetitive genomic sequences was performed soon after and confirmed many of the findings from the first two ER ChIP-on-chip reports (159). Interestingly, only 4% of ER occupancy sites identified fell within promoter regions (defined as 1-kb upstream of a transcription start site), indicating that ER can affect

transcriptional activity from long genomic distances and the focus on promoter regions is insufficient to characterize the full genomic action of ER (159). This result contrasted with the genomic distribution of RNA polymerase II which was found to be present mainly at gene promoter regions (153,156). Comparisons of ER binding sites between the osteosarcoma cell line U2OS expressing ER (U2OS-ER) and MCF-7 cells revealed a small amount of overlap (15%) between the two binding profiles, indicating that ER was not the sole factor responsible for directing ER binding (160). Interestingly, U2OS-ER cells do not express FOXA1 and the FOXA1 motif was not enriched at ER binding sites as it was in MCF-7 cells supporting a role for FOXA1 in ER-chromatin interactions (160). Re-expression of FOXA1 in U2OS-ER cells was found to recover ER functionality (161). Given the impact of phosphorylation on ER activity and DNA binding (see previous sections), the effect of phosphorylation on ER was investigated using the ChIP-on-chip assay (162). The genomic binding profiles of parental MCF-7 cells and MCF-7 cells overexpressing the serine/threonine kinase Akt were compared in the presence and absence of E2. The overexpression of Akt lead to phosphorylation of ER and a wide redistribution of ER on the genome, which led to an altered gene activation profile (162). Between the two cell lines, 2651 ER binding sites were conserved and approximately 1700 sites were unique to each individual cell line. With the known effect of ER phosphorylation on ER-coactivator interaction, it was suggested that the phosphorylation of ER by Akt altered coactivator binding which then affected genome occupancy (162). While this report highlights the importance of phosphorylation on ER-chromatin interactions, the study did not identify specific phosphorylation sites in ER responsible for the shift in genome occupancy observed. Akt is involved in the phosphorylation of ER at

S167 and S305 (63,163–165), and the ChIP-on-chip was additionally performed in the presence of E2, which induces phosphorylation of ER at S118 along with other serine residues (see previous sections). Given the numerous substrates for Akt (166) overexpression could lead to phosphorylation of other targets which could alter chromatin structure or cause other secondary effects which could modulate the interaction of ER with DNA.

In order to analyze the entire human genome, including repeated elements, the ChIP-on-chip technique was expanded to the now widely used ChIP followed by sequencing (ChIP-seq) assay (167). ChIP-seq follows the typical ChIP protocol with the addition of sequencing the isolated DNA using high-throughput sequencing techniques. These sequences are then aligned to their respective locations on the genome using bioinformatics tools. The enrichment of aligned sequences in a region is then interpreted as an occupancy site for the analyzed factor. This technique allows for the entire genome to be assayed, but in turn requires a known genome sequence for the assay to be performed.

Bioinformatics Analysis of ChIP-seq Data

With the development of these high-throughput sequencing techniques, computational tools were first required to analyze the large amounts of data collected for ChIP-seq experiments. Currently, ChIP-seq experiments can generate upwards of 10^8 individual sequencing reads, and the analyses of these reads must be both accurate and efficient (168). Although the purified DNA fragments from the ChIP are generally around 500 bp, only 50 bp of sequencing from one or both ends of the fragment is performed for

the purpose of cost and time savings. Additionally, 50 bp is generally sufficient to accurately identify a single sequence on the genome to which it aligns (169). Programs that align short reads to a reference genome need to consider errors that may occur in the sequencing process, and therefore should permit few mismatches in the alignment process. While many alignment programs have been developed (170,171), one of the more commonly used sequence aligners is the Bowtie2 program (172,173) due to its speed, accuracy, and efficiency of memory usage. However, not all sequence aligners are identical, and specific programs should be chosen to match the question being asked in the experiment. After alignment to a reference genome, regions that are enriched for the factor of interest must then be identified. This process has been termed “peak-calling” due to the shape of the mapped sequences in a region of enrichment (Figure 1-3). Because the purified DNA fragments are sequenced from either end, peaks from alignments will appear as doublets, with one set of sequences aligned to the positive DNA strand and the other set aligned to the negative strand (Figure 1-3). These profiles are combined through either shifting of the matched peaks or estimation of the opposite fragment end locations and combining the fragments into one peak (Figure 1-3). For peaks to be called, the treatment conditions need to be compared to a negative control which consists of either sequencing the input chromatin or isolated DNA from a ChIP-seq performed with a non-DNA binding factor such as IgG. These controls reveal background levels of pulldown and control for regions of the genome that are enriched through the DNA isolation process, IP protocol, or sequencing strategy. Generally, peak-calling takes into account both the fold enrichment over control as well as the absolute number of reads mapped to the specific location, although this determination depends on the peak-calling

program used. Additionally, many peak-calling algorithms consider the type of factor being analyzed by ChIP-seq, as transcription factors will generate different binding patterns compared to histone marks or RNA polymerase. Popular peak-calling programs include Model-based Analysis of ChIP-seq (MACS2) (174) and Hypergeometric Optimization of Motif EnRichment (HOMER) (175) although many others have been developed (176–179). The specific parameters used for peak-calling can have drastic effects on downstream analysis, therefore careful consideration should be taken for the selection of these parameters to generate biologically relevant conclusions from the data.

Downstream analysis of peak sets varies widely, and a biological question should ideally be identified before choosing which programs to use. Most ChIP-seq analyses will perform visualization of peak sites, motif analysis, and gene annotation, although this varies depending on the goal of the experiment. Visualization of ChIP-seq peaks is an important aspect of ChIP-seq analysis and is often performed using the University of California at Santa Cruz (UCSC) Genome Browser (genome.ucsc.edu) (180) which contains a large collection of genomes including multiple builds of the human genome and allows viewing of sequencing data aligned to genomic sequences. The UCSC genome browser is not limited to viewing sequencing data and also includes gene expression data, comparative genomics, single nucleotide polymorphism data, and phenotypic data (180). Motif analysis, which searches for enriched DNA motifs in the genomic locations identified by the peak callers, can identify candidate factors that could interact with the protein of interest through the presence of its DNA binding sequence. These analyses can be performed using a database of known consensus sequences or *de novo* sequences in which motif discovery can be performed if the consensus sequence

is unknown. Common motif analysis programs include HOMER (175) and the Multiple EM for Motif Elicitation (MEME) suite of tools (181), which contains programs for motif scanning (182), discovery (183), and positional analysis (184), among others. Another form of downstream analysis commonly performed in ChIP-seq analysis is gene and location annotation. Location analysis is performed by annotating peaks in relation to genomic features (i.e. introns, exons, intragenic regions) and can infer if a factor of interest is enriched in, for example, promoter regions over background. Additionally, peaks can be annotated to their nearest gene to predict transcriptional control. However certain factors, such as ER, are known to exert control over gene transcription from long genomic distances (185) and the gene controlled by a specific enhancer may not be the nearest gene. Some algorithms such as Genomic Regions Enrichment of Annotations Tool (GREAT) perform more complex annotation by compiling peaks into genomic regions and then correlating these gene regions with certain gene ontologies based on the genes in the region (186). ChIP-seq datasets can also be combined with RNA-seq data to predict control of gene targets by specific binding sites, however definitively demonstrating the control of a gene by a specific peak or set of peaks remains a challenge in the field.

ER ChIP-seq Studies

The first ER ChIP-seq was performed in MCF-7 cells in the presence of E2 and utilized paired-end diTag cloning and sequencing (ChIP-PET) (187). This study identified 1,234 ER occupancy sites of which 71% contained a full ERE. In concurrence with the previous ChIP-on-chip reports, the FOXA1 DNA binding motif, as well as motifs for the

tethering factors AP-1 and Sp1, were found to be enriched at ER occupancy sites (187). Additionally, of the ER occupancy sites identified, only 5% were located in promoter regions and 20% were found in “gene desert” regions, where the nearest gene is greater than 100 kb away (187). Deeper sequencing led to a more complete ER binding profile (cistrome), identifying 10,205 ER occupancy sites (188). Similar to previous reports (156,159,187), a low proportion (7%) of ER occupancy sites were found in promoter regions with the majority of sites (41%) located within introns (188). These observations have since been confirmed by many other reports (121,158,185,189,190) and are not features limited to ER, as other NRs such as AR (191) and GR (192) display similar genomic distributions. These distal ER enhancers have been shown to be the sites of enhancer RNA (eRNA) production, which are non-coding RNAs that have important roles in gene transcription and mark active enhancer sites, however, the exact mechanism and purpose of eRNAs are still not well understood (193–195). The observation that many ER occupancy sites fall far from gene promoters led to the study of chromatin looping on ER function. Chromatin interaction analysis by paired-end tag sequencing (ChIA-PET) was developed to identify occupancy sites which interact physically with each other but are genomically distant (185). This analysis found that E2 induced many interactions between distal ER enhancers and proximal promoter sites of ER target genes, particularly at the *TFF1* and *GREB1* loci, and development of these chromatin loops was dependent on ER.

Comparisons of the ER cistrome with the binding profiles of other transcription factors and histone modifications allowed for broad correlations to be made that could previously be investigated only on select genomic sites by conventional ChIP. A summary of these and other large-scale ER ChIP studies is listed in Table 1-2. These comparisons

found that the best predictors of ER occupancy on a site in the genome were a strong ERE, FOXA1 occupancy, the presence of H3K4 methylation, and open chromatin as assessed by Formaldehyde-Assisted Isolation of Regulatory Elements (FAIRE-seq) (190). Additionally, ChIP-seq of the p160 family of NR coactivators was found to overlap with ER sites and only associated with these sites upon E2 treatment (189). The interaction of these coactivators with ER is known to be dependent on ER phosphorylation *in vitro* (77), however the distribution and effect of phosphorylation on the ER genome-wide binding profile are poorly understood and will be investigated in Chapters 2 and 3 of this thesis. Numerous other factors have been found to be associated with ER on chromatin as identified through ChIP-seq including GATA3 (196,197), Runx1 (121), ZNF217 (198), GR (199), and FOXM1 (200) (Table 1-2). To assess factors associated with ER on chromatin in an unbiased manner, techniques involving mass spectrometry of protein complexes on DNA have been performed (74,201). One study utilizing a biotin-DNA pull-down assay with an ERE as bait found that a large complex of proteins was recruited to EREs in response to E2 with the strongest associated factors being ER, the p160 family of coactivators, CREB binding protein (CBP), p300, and the Mediator complex (74). Interestingly, the addition of ATP to the complexes converted them to an active state through phosphorylation by the DNA-dependent protein kinase (DNA-PK). Phosphorylation was shown to occur on multiple components of the identified complex, including pS118-ER. Reduction of DNA-PK activity through either siRNA or small molecule inhibitors led to a reduction in E2-dependent ER transcriptional activity, highlighting the importance of phosphorylation in the ER transcriptional complex (74).

Stimuli other than E2 were also found to have a drastic effect on the genomic binding profile of ER. Both of the ER antagonists tamoxifen and ICI induce binding of ER to DNA, although to a lesser extent compared to E2 (188). Unlike E2, they prevent recruitment of RNA polymerase II and activation of target genes which contributes to their anti-estrogenic effects (188). Additionally, EGF activation of ER, which functions through phosphorylation of ER at S118 (57,59,65), was found to induce a distinct ER binding pattern compared to E2 activation (140). A common aspect of these treatments is that phosphorylation of ER is induced, although through different kinases (see previous section). The effects of ER phosphorylation on the ER cistrome in both ligand-dependent and ligand-independent contexts is not well understood and will be addressed further in this thesis.

Clinical Role of pS118-ER

Clinically, the role of pS118-ER is not well defined. Expression of pS118-ER in primary breast cancer tumor biopsies as measured by immunohistochemistry (IHC) is associated with better patient outcomes and is generally thought of as a marker for an active ER signaling pathway (202,203). Additionally, a clinical trial of breast cancer patients receiving tamoxifen found that patients with high levels of pS118-ER saw better responses to tamoxifen treatment compared to those with low pS118-ER (204). Levels of pS118-ER have been observed to decrease following endocrine therapies, with the magnitude of decrease correlating with therapy response (205). In contrast, other studies have found poorer disease-free survival in patients expressing high pS118-ER (206). A recent large-scale IHC study of 1,036 breast cancer patient samples found a correlation

between phosphorylated MAPK, pS118-ER, and pS167-ER, indicating an activated MAPK pathway, but no correlation was found between any of these phosphorylation marks and either disease-free survival or overall survival (207). While a wealth of knowledge on pS118-ER has come from *in vitro* and cell line studies that demonstrate the importance of pS118-ER in ER biology, ascertaining the clinical significance of pS118-ER remains difficult.

Although mutations in ER leading to constitutive activity have been known for some time (208,209), these ER mutants have recently been identified in metastatic breast cancer tumors (210–213). The mutations occur primarily in the ligand-binding domain near the ligand-binding pocket and are most frequently Y537S and D538G (214). These mutations cause ER transcriptional activity in the absence of E2 (215), and their cistromes are distinct from that of wild-type (wt) ER (216,217). Additionally, when MCF-7 cells expressing these mutations were introduced into mice through xenografts, the mutant ER tumors displayed increased growth and metastasis in the absence of ligand and mimicked the growth phenotype of wt ER exposed to E2 (217). Interestingly, these ER mutants are constitutively phosphorylated at S118 and interact with various coactivators in the absence of ligand (216,217). Inhibition of the kinase CDK7 with the small molecule inhibitor THZ1 led to a decrease in pS118-ER in the Y537S and D538G ER mutants, and a combination treatment of THZ1 with ICI led to a greater decrease in the growth of mutant ER xenografts compared to ICI treatment alone (216,217). Additionally, a translocation between *ESR1* and *YAP1* which generated a fusion protein consisting of the N-terminal AF-1 and DBD of ER and C-terminus of YAP1 was identified in patient samples (218). This ER-YAP1 fusion protein induced growth of MCF-7 and T47-D cells in the absence

of E2 and activated the ER target gene *TFF1*, suggesting that the ER AF-1 domain was active in the ESR1/YAP1 fusion (218). However, it is not yet known whether this fusion protein is phosphorylated at S118 or any of the other phosphorylation sites located in the AF-1 domain.

Summary

Given the critical role played by ER in breast cancer biology, control of the receptor's activity is crucial for the treatment of ER-positive breast cancers. As discussed here, pS118-ER has been implicated in both ligand-dependent and ligand-independent activation of ER and its target genes. Additionally, pS118-ER is known to affect the DNA binding properties of ER, but these analyses have been limited to small-scale experiments. While the genomic binding profile of ER has been investigated in a large range of conditions (Table 1-2), the effects of phosphorylation on its cistrome are not well defined. Elucidating the role of pS118-ER in this manner is critical not only for the understanding of general ER biology, but also for the treatment of metastatic breast cancers given the constitutive nature of pS118-ER in metastatic-specific ER mutants. The goal of my work presented here was to define the pS118-ER cistrome in MCF-7 cells and to identify features associated with this phosphorylation. Chapters 2 and 3 of this thesis will analyze the cistrome of pS118-ER in MCF-7 cells, with Chapter 2 focusing on technical challenges faced in the optimization and development of the assay, and Chapter 3 focusing on the analysis of the ChIP-seq data. With the majority of S118A ER mutational studies performed in ER-negative cells, the functional significance of pS118-ER is still poorly understood in an ER-native cellular environment. Chapter 4 describes the

development of an S118A ER model in the ER-positive cell line MCF-7 through CRISPR. The studies presented in this thesis advance the ER field by performing the first in-depth analysis of a post-translationally modified steroid receptor, and by identifying novel factors associated with pS118-ER. Additionally, the development of the S118A MCF-7 cell line will aid researchers with a robust tool to further ascertain the function of pS118-ER in ER and breast cancer biology.

Figure 1-1: Domain structure of ER protein and phosphorylation sites.

The protein translated from the *ESR1* gene is 595 amino acids and consists of distinct domains with separate functions. Known phosphorylation sites and their approximate locations are displayed above the diagram. At the N-terminus is the Activation Function-1 (AF-1) domain, which is generally activated in ligand-independent activation pathways and is widely phosphorylated. The DNA binding domain (DBD) and hinge (H) region are located in the central part of the protein. The majority of the C-terminus contains the Activation Function-2 domain (AF-2) which itself contains the ligand-binding domain (LBD) and is the portion of the protein that responds to ligand-dependent activation. The AF-2 domain has fewer phosphorylation sites compared to the AF-1 domain.

Figure 1-1

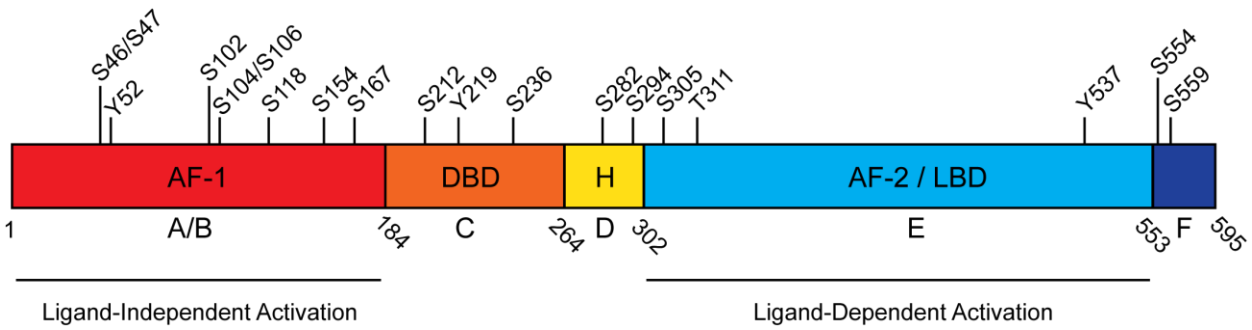

S46/S47
Y52
S102
S104/S106
S118
S154
S167
S212
Y219
S236
S282
S294
S305
T311
Y537
S554
S559
AF-1
DBD
H
AF-2 / LBD
1
A/B
184
C
264
D
302
E
553
F
595
Ligand-Independent Activation
Ligand-Dependent Activation

Figure 1-2: ER activation mechanisms and DNA binding states.

ER can be activated through two major pathways—a ligand-dependent pathway through E2, and a ligand-independent pathway through various signaling factors (see Table 1-1). Both activation pathways lead to phosphorylation of ER and binding of ER to DNA. Ligand-dependent activation by E2 causes dimerization of ER, phosphorylation at S118, binding to DNA, and degradation by the 26S proteasome. Ligand-independent activation generally activates a phosphorylation cascade, which subsequently phosphorylates ER and leads to DNA binding. Additionally, ER can interact with DNA through a tethering mechanism (bottom), where ER is tethered to DNA by a linker cofactor (CoF) or other DNA-associating protein. The role of phosphorylation in tethering mechanisms has not been thoroughly investigated.

Figure 1-2

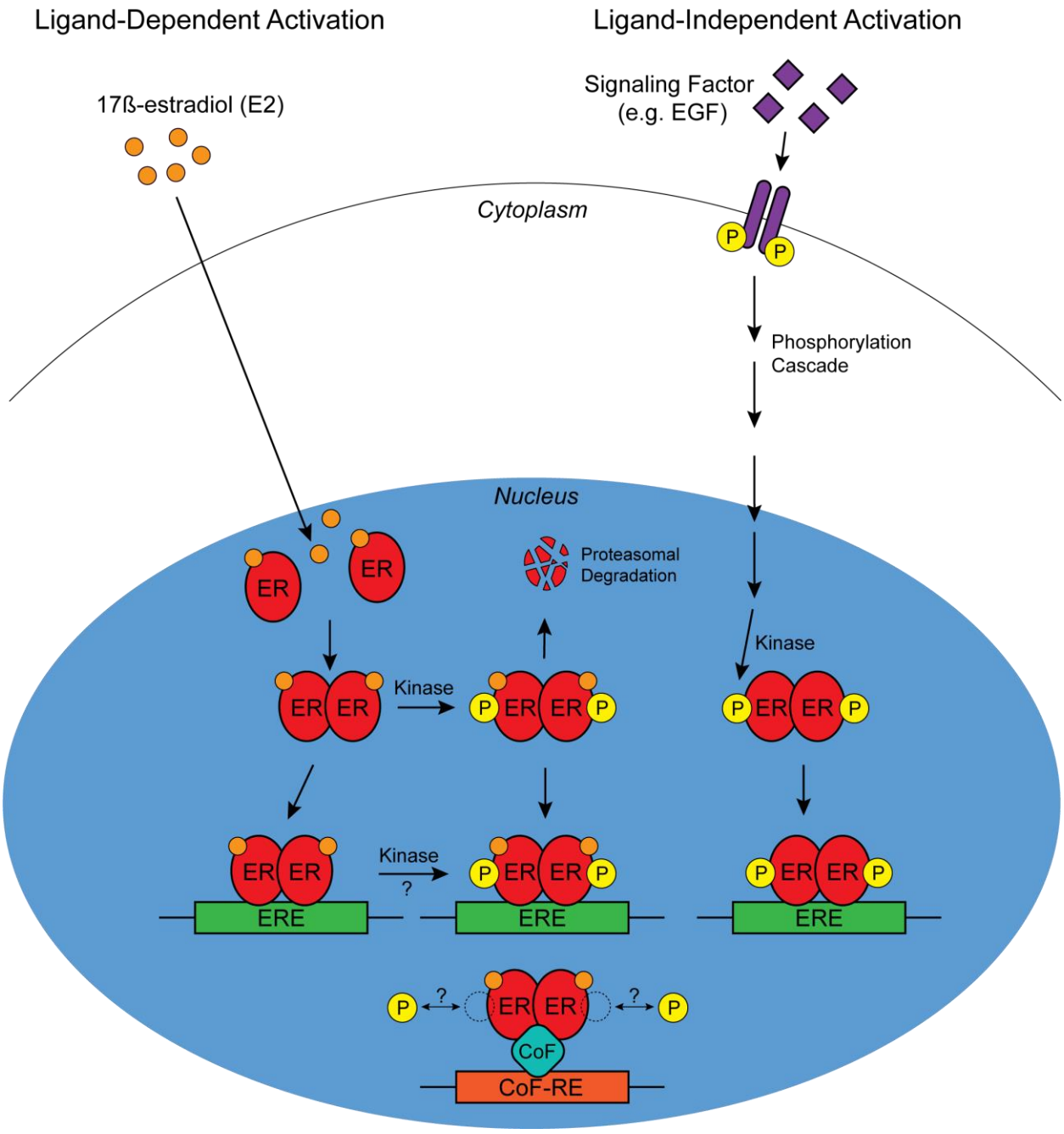

Ligand-Dependent Activation
Ligand-Independent Activation
17ß-estradiol (E2)
Signaling Factor
(e.g. EGF)
Cytoplasm
P
P
Phosphorylation
Cascade
Nucleus
ER
ER
Proteasomal
Degradation
Kinase
ER ER
Kinase
P ER ER P
P ER ER P
ER ER
Kinase
?
P ER ER P
P ER ER P
ERE
ERE
ERE
P
?
ER ER
?
P
CoF
CoF-RE

Figure 1-3: Schematic of ChIP-seq analysis.

After experimental treatment conditions, cells are treated with a fixing agent such as formaldehyde, the DNA is sheared, and the factor of interest along with the attached DNA is immunoprecipitated with an antibody (Step 1). Next, DNA fragments approximately 500 bp in length are isolated and sequenced using high-throughput sequencing technology (Step 2). Generally, only 50 bp are sequenced on one end of each DNA fragment. These reads are then aligned to a reference genome (Step 3). Since only one of the 5' ends from each fragment has been sequenced, alignment can be to the + strand (red) or the – strand (blue). Stacking of the aligned sequences creates a peak doublet (dashed lines) for an individual binding event. These doublets must be corrected by shifting the peaks toward each other (Step 4) to generate one single peak. Peaks are then called throughout the genome (Step 5) and downstream analysis is performed on these regions (Step 6). This list of downstream analysis is not exhaustive, and other methods of analysis should be selected based on the hypothesis being tested.

Figure 1-3

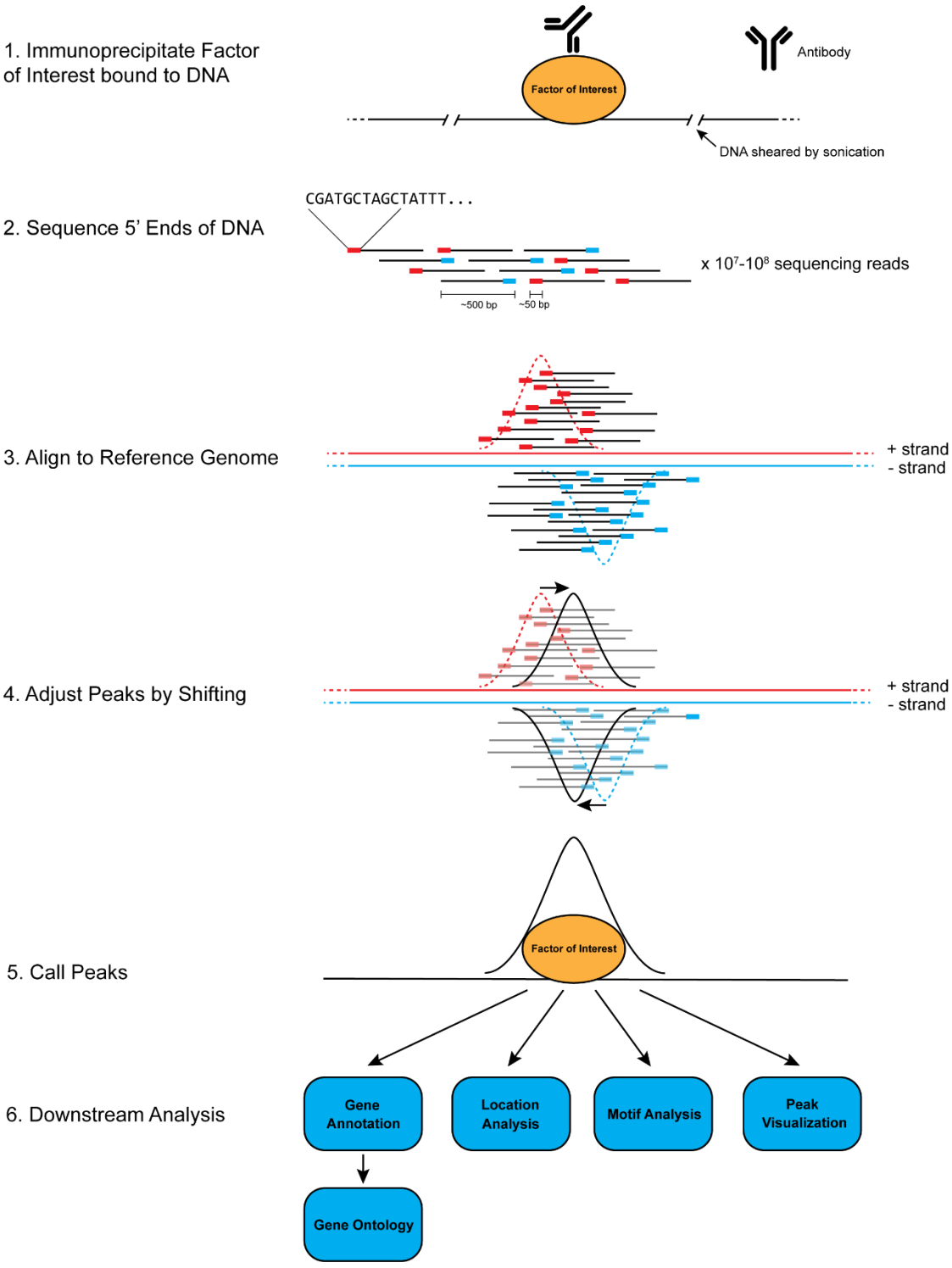

1. Immunoprecipitate Factor of Interest bound to DNA
Factor of Interest
Antibody
DNA sheared by sonication
2. Sequence 5' Ends of DNA
CGATGCTAGCTATTT...
x 10^7-10^8 sequencing reads
~500 bp
~50 bp
3. Align to Reference Genome
+ strand
- strand
4. Adjust Peaks by Shifting
+ strand
- strand
5. Call Peaks
Factor of Interest
6. Downstream Analysis
Gene Annotation
Location Analysis
Motif Analysis
Peak Visualization
Gene Ontology

Table 1-1:

ER Phosphorylation Sites and Associated Kinases

Phosphorylation Site	Activated by	Kinase	Ref.
S46/47	unk.	unk.	(56)
Y52	unk.	c-Abl	(219)
S102	E2	GSK3	(71)
S104/S106	E2	GSK3	(71)
	E2	MAPK	(139)
	E2	cyclin A-CKD2	(220)
	Tam	MAPK	(221)
	ICI	MAPK	(221)
	PMA	MAPK	(221)
S118	E2	unk.	(52)
	E2	CDK7	(67)
	E2	GSK3	(71)
	E2	IKKα	(72)
	E2	MAPK	(139)
	E2	DNA-PK	(73,74)
	Tam	unk.	(52)
	ICI	unk.	(52)
	EGF	MAPK	(57)
	IGF-1	MAPK	(57)
	PMA	unk.	(222)
	Prolactin	unk.	(60)
	ROS	MAPK	(63)
	Heregulin	MAPK	(61)
	Hypoxia	MAPK	(64)
	ER Y537S	CDK7	(216,217)
S154	E2	unk.	(54)
	EGF	unk.	(54)
S167	E2	E2	(53)
	EGF	p90-RSK1	(223)
	PI3K	Akt	(164)
	ROS	Akt	(63)
	PMA	S6K1/MAPK	(224,225)
	Insulin	S6K1/MAPK	(224,225)
S212	E2	unk.	(55)
Y219	unk.	c-Abl	(219)
S236	unk.	PKA	(143)
S282	E2	CK2	(56)

Table 1-1 (cont.)

ER Phosphorylation Sites and Associated Kinases

Phosphorylation Site	Activated by	Kinase	Ref.
S294	E2	unk.	(55,56)
S305	unk.	PKA	(226)
	unk.	Pak1	(227)
	unk.	Akt	(163)
T311	E2	MAPK (p38)	(228)
Y537	E2	calf uterus TK	(229)
	E2	SRC	(138)
S554	E2	unk.	(55)
S559	E2	CK2	(55,56)

Abbreviations: Akt, RAC-alpha serine/threonine-protein kinase; c-Abl, cellular Abelson tyrosine kinase; CDK2, cyclin dependent kinase 2; CK2, casein kinase 2 alpha; DNA-PK, DNA-dependent protein kinase; E2, 17ß-estradiol; EGF, epidermal growth factor; GSK3, glycogen synthase kinase-3; ICI, ICI-182780/fulvestrant; IGF-1, insulin-like growth factor 1; IKKα, IκB kinase alpha; MAPK, mitogen-activated protein kinase; p90-RSK1, ribosomal S6 kinase alpha-1; Pak1, p21 activated kinase 1; PI3K, phosphoinositide 3-kinase; PMA, Phorbol 12-myristate 13-acetate; ROS, reactive oxygen species; S6K1, ribosomal protein S6 kinase beta-1; SRC, src proto-oncogene non-receptor tyrosine kinase; Tam, 4-hydroxytamoxifen; unk., unknown

Table 1-2:

List of Large-Scale ER-DNA Binding Studies

Procedure	Cell Lines	Condition	Genomic Scale	Other Factors ChIPed	Ref.
ChIP-chip	MCF-7	E2	promoters	-	(152)
	MCF-7	E2	chr 21, 22	-	(156)
	MCF-7	E2	1.5 Gbp	RNA PolII	(159)
	MCF-7	E2	promoters	c-MYC, AcH3	(154)
	MCF-7	E2	promoters	RNA PolII, SRC, acH3, acH4	(153)
	MCF-7	E2, Akt	1.5 Gbp	-	(162)
	U2OS+ER	E2	chr 1, 6	-	(160)
	MCF-7	E2	chr 1, 3, 6, 21, 22, X, Y	ERß	(230)
	MCF-7	E2	1.5 Gbp	FOXA1	(158)
	MCF-7	E2	1.5 Gbp	PAX2	(231)
	MCF-7	EGF	tiling array	-	(140)
	MCF-7	E2	tiling array	ERK2	(232)
ChIP-PET	MCF-7	E2	genome-wide	-	(187)
ChIP-seq	MCF-7	E2, Tam, ICI	genome-wide	RNA PolII	(188)
	MCF-7, T-47D	E2	genome-wide	RNA PolII, FOXA1, meH3, acH3	(190)
	MCF-7	E2	genome-wide	FOXA1, GATA3	(196)
	MCF-7, T-47D, ZR-75-1	E2	genome-wide	FOXA1, CTCF	(161)
	mouse	E2	genome-wide	RNA PolII	(233)
	BC patient tumors	-	genome-wide	FOXA1	(234)
	T-47D, Ishikawa	E2	genome-wide	-	(235)
	MCF-7	E2	genome-wide	-	(193)
	mouse mammary	E2, dex	genome-wide	GR	(199)
	MCF-7	E2, siGATA3	genome-wide	FOXA1, GATA3	(197)
	MCF-7	siER, E2, FM	genome-wide	-	(236)
	T-47D	progesterone	genome-wide	PR, p300	(237)
	MCF-7	E2, ER Y537S	genome-wide	-	(216)
	MCF-7	ER Y537S, D538G	genome-wide	FOXA1, acH3	(217)

Table 1-2 (cont.)

List of Large-Scale ER-DNA Binding Studies

Procedure	Cell Lines	Condition	Genomic Scale	Other factors ChIPed	Ref.
ChIP-seq	MCF-7	E2 (tc)	genome-wide	-	(146)
	MCF-7	E2	genome-wide	pS118-ER	(238)
	MCF-7	E2	genome-wide	GRHL2	(239)
ChIA-PET	MCF-7	E2	genome-wide	-	(185)
RIME	MCF-7	E2, Tam	genome-wide	E2F4, RXRa, COT2, KLF4, GREB1, TLE3	(201)
	MCF-7, T-47D	progesterone	genome-wide	FOXA1	(237)

Abbreviations: acH3, acetylated histone H3; acH4, acetylated histone H4; ChIA-PET, chromatin interaction analysis by paired-end tag sequencing; ChIP-chip, chromatin immunoprecipitation on a chip; ChIP-PET, chromatin immunoprecipitation paired-end diTag; ChIP-seq, chromatin immunoprecipitation followed by sequencing; chr, chromosome; c-MYC, cellular-MYC; COT2, Chicken Ovalbumin Upstream Promoter Transcription Factor 2; dex, dexamethasone; E2, 17ß-estradiol; E2F4, E2F Transcription Factor 4; ERK2, mitogen-activated protein kinase 1; ERß, estrogen receptor-ß; FM, full media; FOXA1, forkhead box A1; GATA3, GATA binding protein 3; Gbp, gigabase pairs; GR, glucocorticoid receptor; GREB1, growth regulation by estrogen in breast cancer 1; GRHL2, grainyhead-like 2; ICI, ICI182,780 / fulvestrant; KLF4, Kruppel-like factor 4; meH3, methylated histone H3; p300, histone acetyltransferase p300; PAX2, paired box 2; PR, progesterone receptor; pS118-ER, estrogen receptor-α phosphorylated at serine 118; RIME, rapid immunoprecipitation mass spectrometry of endogenous proteins; RNA PolII, RNA polymerase II; RXRα, retinoid X receptor α; SRC, src proto-oncogene non-receptor tyrosine kinase; Tam, 4-hydroxytamoxifen; tiling array, Affymetrix GeneChip Human Tiling 2.0R Array; tc, time course; TLE3, TLE family member 3, transcriptional corepressor

CHAPTER TWO:

Analysis of pS118-ER Dynamics, Antibodies, and Initial ChIP-seq

Portions of this chapter are published in the following:

Helzer KT, Szatkowski Ozers M, Meyer MB, Benkusky NA, Solodin N, Reese RM, Warren CL, Pike JW, Alarid ET. The Phosphorylated Estrogen Receptor α (ER) Cistrome Identifies a Subset of Active Enhancers Enriched for Direct ER-DNA Binding and the Transcription Factor GRHL2. *Mol. Cell. Biol.* 2019;39(3). doi:10.1128/MCB.00417-18.

Abstract

Post-translational modifications are key regulators of protein function, providing cues that can alter protein interactions and cellular location. Phosphorylation of the estrogen receptor-α (ER) at serine 118 (pS118-ER) occurs in response to multiple stimuli and is involved in modulating ER-dependent gene transcription. Both ligand-dependent and ligand-independent activation mechanisms are known to induce pS118-ER, suggesting an important role in ER signaling. Numerous *in vitro* studies have been performed to analyze the effect of pS118-ER on ER-DNA binding, however the full genome-wide binding profile (cistrome) of pS118-ER is uncharacterized. In this chapter, I first analyze the dynamics of pS118-ER in MCF-7 cells in response to multiple known stimuli as well as its effect on ER-DNA binding at certain genomic locations. I demonstrate that while pS118 is not required for DNA binding, lack of phosphorylation reduces the occupancy of ER at various binding sites. Additionally, I show that DNA binding is not required for phosphorylation of ER, suggesting that phosphorylation occurs prior to DNA binding. To accurately characterize the cistrome of pS118-ER, the reliability of antibodies is tested for their ability to identify only the phosphorylated form of ER. Here I analyze the utility of three different pS118-ER specific antibodies according to rigorous guidelines and conclude that two are suitable for chromatin immunoprecipitation followed by sequencing (ChIP-seq) studies. These results identify key aspects regarding the temporal dynamics of pS118-ER and lay the groundwork for defining the pS118-ER cistrome.

Introduction

The phosphorylation of proteins has proved to be a highly important regulator of various protein functions including enzymatic activity, cellular localization, protein-protein interactions, and degradation (42,44–46). The addition of a phosphate group onto a protein via a kinase can produce a quick and often reversible alteration to the function of the target protein. It has been estimated that 30% of the proteome is phosphorylated (240), but more recent studies utilizing deep phosphoproteomics has demonstrated that the proportion of phosphorylated proteins can be as high as 75% (48) highlighting the ubiquitous role phosphorylation plays in the cell.

Members of the steroid hormone nuclear receptor family are no exception to this vast phosphorylation activity. Estrogen receptor-α (ER) is a nuclear hormone receptor responsible for the growth and maintenance of estrogen-dependent tissue such as the breast. Activation of ER is controlled by binding of its ligand, 17ß-estradiol (E2), to the C-terminal ligand-binding domain which leads to dimerization and binding of the ER dimer to DNA at estrogen response elements (EREs) (128). Transcriptional machinery is then recruited to these sites to activate gene transcription of ER target genes. ER activity can also be altered by post-translational modifications, including phosphorylation. Multiple phosphorylation sites have been identified on ER, most of which occur in the N-terminal AF-1 domain (41,241). Of particular interest is the phosphorylation of ER at serine 118 (pS118-ER). Serine 118 phosphorylation is induced by E2 as well as the ER antagonists tamoxifen and ICI 182,780 (52). Activation through pS118-ER can also occur in a ligand-independent manner and many stimuli have been shown to produce this effect including epidermal growth factor (EGF), reactive oxygen species, prolactin, and hypoxia

(57,59,60,63,64,69). Numerous reports have demonstrated that pS118-ER is required for maximal E2-stimulated ER transactivation through mutational analysis of S118 to alanine (52,59,78,80,81). However, this response is not universal to all ER target genes, as others have shown that activation of some E2-induced genes is not dependent on pS118-ER (79).

Phosphorylation of ER has also been shown to affect its interaction with DNA. Many *in vitro* studies have demonstrated that ER phosphorylation leads to an increased affinity of ER for the ERE (34,66,110,135). Upon E2 stimulation, pS118-ER occupies various ER target gene promoters (136) and recruitment of the p160 family of coactivators at these sites is impaired in the S118A ER mutant (80). In addition to promoters, pS118-ER has also been identified as a component of the large “MegaTrans” transcriptional complex located at enhancer regions (242). Overexpression of the kinase Akt, which is known to phosphorylate ER (164), led to a unique set of ER binding sites as assessed by ChIP-on-chip, indicating that phosphorylation can cause a redistribution of ER on chromatin (162). However, the specific ER phosphorylation sites induced by Akt were not determined in this study.

While numerous studies have been performed to analyze the cistrome of ER in response to different stimuli (Table 1-2), the genome-wide binding profile of pS118-ER and how it compares to the general E2-induced ER cistrome is unknown. Part of the reason for the lack of knowledge on pS118-ER genome occupancy is the lack of well-characterized ChIP-grade antibodies directed toward phosphorylated substrates. Unlike full-length proteins where multiple short epitopes can be selected throughout the protein and tested for viable antibody production, phospho-specific antibodies must be generated

against an epitope that contains the phosphorylation site, providing a constraint on the number of epitopes that can be tested (243,244). Additionally, these antibodies need to be specific for the phosphorylated substrate while avoiding interaction with the same substrate in the unphosphorylated state. Numerous studies have noted that phospho-specific antibodies can produce unreliable results through cross-reactivity (245,246), emphasizing the need for phospho-antibody validation.

In this chapter, I determine the optimal conditions for performing a ChIP-seq experiment on pS118-ER in MCF-7 cells by optimizing E2 treatment and validating three separate pS118-ER antibodies using the standards set up by the ENCODE consortia for ChIP-seq antibody validation (247). All three antibodies passed the criteria, and ChIP-seq was performed using each of the three antibodies in MCF-7 cells treated with E2. Interestingly, although all three antibodies were validated, the ChIP-seq signal from one of the antibodies produced a non-specific binding pattern which resembled an RNA polymerase II-like signal. This along with other inconsistencies from the non-specific pS118-ER antibody lead to it being removed from further analysis. These results provide a high-quality ChIP-seq data set and allow for the analysis of the pS118-ER cistrome.

Results

Dynamics of pS118-ER in Response to Stimuli

Prior to performing ChIP-seq, we sought to identify the optimal E2-treatment conditions to induce maximal pS118-ER. MCF-7 cells were grown in phenol-red free charcoal-stripped media for three days prior to estrogen treatment. Cells were treated with either vehicle (0.1% EtOH) or 10 nM E2 for varying amounts of time between 15 min and 24 h. We found that maximal pS118-ER induction occurs 30 minutes post treatment with 10 nM E2 and decreases after the 30-minute time point (Figure 2-1A). We also observed a decrease in total ER levels with E2 treatment as reported previously (99,100). To further investigate pS118-ER dynamics in response to other stimuli, similar time course treatments were performed using the ER antagonist tamoxifen and the ligand-independent ER activator EGF. Interestingly, treatment of MCF-7 cells with tamoxifen led to sustained phosphorylation at S118 which did not decrease through the duration of the time course (Figure 2-1B). Treatment with EGF, which is known to phosphorylate ER on S118 (57), led to rapid induction of pS118-ER at 15 min post-treatment which then declined for the remainder of the time course (Figure 2-1C).

pS118-ER is Required for Maximal ER Occupancy on DNA

To investigate the necessity of phosphorylation of S118 for ER-DNA binding, we utilized MDA-MB-231 cells containing doxycycline-inducible wild-type (wt) or S118A ER constructs. Upon expression of ER with doxycycline, the wt ER construct was phosphorylated in response to E2 while the S118A mutant was not (Figure 2-2A). ChIP-qPCR was performed for ER on the wt and S118A expressing cells in the presence and absence of E2, and we found that the S118A mutant significantly decreased the E2-

induced ER occupancy at known ER binding sites (Figure 2-2B). Interestingly, the S118A mutant did not entirely abolish ER occupancy at these sites, indicating that phosphorylation at S118 is not required for ER-DNA binding. These results indicate that phosphorylation at S118 drives maximal ER occupancy but is not necessary for ER-DNA binding to occur in response to E2.

A Functional DNA Binding Domain is Not Necessary for Phosphorylation of ER at S118 in Response to E2

The phosphorylation of ER at S118 is mediated by a variety of kinases including MAPK, CDK7, GSK3, and IKKα depending on the stimulus (57,69,71,72,136). Additionally, various studies have shown that ER is capable of being phosphorylated by the DNA-dependent protein kinase (DNA-PK) holoenzyme (73,74,137). To test if DNA binding is required for ER to be phosphorylated at S118 in response to E2, the C202/205H ER mutant was utilized which ablates the zinc finger of the ER DNA binding domain (16,248). HEK293 cells were transiently transfected with either wt ER or C202/205H ER and treated with 10 nM E2 for 30 min. Both wt ER and C202/205H ER were phosphorylated at S118 in response to E2 indicating that a functional DNA binding domain is not required for this phosphorylation event (Figure 2-2C).

Validation of pS118-specific Antibodies for ChIP-seq

The accuracy of a ChIP-seq study is highly dependent on the efficacy of the antibody used. Non-specific antibody interactions with other chromatin-bound factors can cause false positive occupancy sites to be detected, which leads to inaccurate downstream analyses and misleading conclusions. A set of criteria has been defined by the Encyclopedia of DNA Elements (ENCODE) consortium for validation of antibodies for

ChIP-seq (247) which consists of a primary test and a secondary test. For an antibody to be validated it must pass one primary characterization and one secondary characterization. The primary test consists of validating the antibody by either immunoblot or immunofluorescence. To pass primary characterization by immunoblot, 50% of the total signal on the blot must come from the primary band (247). Three pS118-ER specific antibodies were tested (pS118-ER #1, Cell Signaling; pS118-ER #2, Abcam; pS118-ER #3, Santa Cruz) for their ability to detect pS118-ER in MCF-7 cells in response to E2. MCF-7 cells were grown in charcoal-stripped medium for three days, treated for 30 min with E2, and analyzed by immunoblot. All three pS118-ER antibodies detected a strong band at the correct size in the E2 treated sample with little signal in the untreated sample (Figure 2-3A). Additionally, a majority of the signal from all three of the antibodies came from the primary band detected. Because all three antibodies passed primary characterization by immunoblot, immunofluorescence validation was not required.

Next, each antibody must pass one out of five secondary characterization assays. One of these five assays is to mutate or knockdown the protein of interest to demonstrate specificity of the antibody to the target of interest. To test the specificity of the pS118-ER antibodies for pS118 and not any other phosphorylation site on ER, an S118A mutant ER was utilized. MDA-MB-231 cells with a doxycycline-inducible wt or S118A ER construct were treated with doxycycline to express ER, and subsequently treated with E2 for 30 min. Cell lysates were collected and analyzed by immunoblot with all three pS118-ER antibodies. Induction of pS118-ER was observed by all three antibodies in the wt construct and the S118A mutation caused a complete loss of signal with all three antibodies tested (Figure 2-3B) indicating that the signal observed from the pS118-ER

antibodies is specific for S118 phosphorylation and not any other phosphorylation site on ER induced by E2. These results indicate that all three pS118-ER antibodies pass the ENCODE guidelines for ChIP-seq analysis. To provide a comprehensive analysis of the pS118-ER cistrome, all three antibodies were used in separate ChIP-seq experiments and their data compared. As an added measure of validation with a secondary assay (247), DNA motif analysis was performed on the peaks identified within each pS118-ER ChIP-seq dataset and identified the ERE as the top binding motif in all antibodies analyzed, further confirming their effectiveness for ChIP-seq (Figure 2-3C).

ChIP-seq of pS118-ER Using Multiple Antibodies

To investigate the inducible genome-wide occupancy of pS118-ER, MCF-7 cells were treated with either vehicle or 10 nM E2 for 30 min and ChIP-seq was performed using three antibodies directed toward pS118-ER and one for ER. ChIP-seq was performed in triplicate for each antibody and peaks were called using HOMER (175). Replicates were overlapped and sites present in each of the three replicates were used for further analysis (Figure 2-4A). Interestingly, the variability in the pS118-ER replicates was greater than the total ER replicates, with a large proportion of the total pS118-ER peaks identified unique to each replicate indicating that pS118-ER occupancy may be highly dynamic (Figure 2-4A). E2-treated ER ChIP-seq identified 44,050 ER occupancy sites which is consistent with recent ER ChIP-seq studies (197,201). The total number of peaks varied for pS118-ER, depending on the antibody. The E2-treated pS118-ER ChIP-seq identified 2,054 sites with pS118-ER #1, 2,277 sites with pS118-ER #2, and 14,057 with pS118-ER #3 (Figure 2-4A, center overlaps). Multiple sites present in all three pS118-ER ChIP-seq experiments were validated by ChIP-qPCR (Figure 2-5A).

Inconsistencies Observed in pS118-ER Antibodies

Both ER and all three pS118-ER ChIP-seq experiments showed an increase in occupancy sites upon E2 treatment (Figure 2-4B). A large number of ER sites were identified (15,905) in the absence of E2, which is consistent with a previous study demonstrating that unliganded ER is capable of binding chromatin (236). Two of the three pS118-ER antibodies (#2, and #3) displayed a low number of occupancy sites in the vehicle-treated cells compared to their respective E2-treated sites as expected (Figure 2-5B). Surprisingly, the remaining pS118-ER antibody (#1) identified a similar number of sites in vehicle (1,653) and E2 (2,054) conditions (Figure 2-5B). The two concordant pS118-ER antibodies showed a high amount of overlap with the ER antibody (99.7% and 94.3%, respectively), whereas pS118-ER #1 had a relatively low amount of overlap with ER (52.2%) suggesting a non-specific interaction (Figure 2-4D). Additionally, overlapping the peaks identified from the three pS118-ER antibodies revealed considerable overlap between pS118-ER #2 and #3, with pS118-ER #1 identifying many unique sites (Figure 2-4C). Visualization of the ChIP-seq tracks using the UCSC Genome Browser (180) revealed a strong intragenic signal present in the pS118-ER #1 track that was not present in either the ER or the other two pS118-ER tracks (Figure 2-5B). Two hypotheses were considered for this observation: [1] the signal was detecting pS118-ER which was tracking along actively transcribed genes and [2] the pS118-ER #1 antibody was non-specifically binding to a factor associated with these regions. The lack of the intragenic signal in the total ER ChIP-seq and the low amount of overlap between the ER and pS118-ER #1 peak sets (Figure 2-4D) suggested a non-specific interaction was more likely, and the following additional observations support this claim. First, a very small proportion of ER is

phosphorylated at S118 in an unstimulated state (Figure 2-1) and therefore the number of pS118-ER peaks in the vehicle-treated conditions was expected to be low due to the lack of a target for the antibody. However, the pS118-ER #1 antibody detected a high number of peaks in the absence of E2 compared to the E2 treated condition, whereas the other two pS118-ER antibodies did not (Figure 2-4C). Second, comparisons of this intragenic signal to a previously published ChIP-seq analysis of RNA polymerase II (RNA-Pol II) in MCF-7 treated with E2 (188) showed a strong similarity between the signals along the *TFF1* and *GREB1* genes, indicating that the intragenic signal was identifying RNA Pol II or a factor tracking with RNA Pol II. Additionally, the intragenic signal increased on genes known to be induced by E2 (Figure 2-6A) and decreased on genes known to be repressed by E2 (Figure 2-6B). All genes in the genome were analyzed for this difference in intragenic signal and gene ontology (GO) analysis was performed on genes identified to have a significantly increased or decreased signal upon E2 treatment. The top GO term for genes with an increased intragenic signal was mammary gland development suggesting a correlation with E2-induced genes. The top GO term for genes with a decreased intragenic signal was nucleosome assembly. Interestingly, for large genes, the difference between the vehicle and E2-treated intragenic signal was only present at the beginning of the gene (see Figure 2-6 *ESR1* and *GREB1*). The speed of RNA polymerase II has been calculated to be between 1.2 kb/min to 4.3 kb/min depending on the study (249–251). Assuming transcriptional modulation occurs at the beginning of the E2 treatment, with the speeds listed above the alteration to RNA polymerase II occupancy in a 30 min treatment would be observed between 36 to 129 kb from the transcription start site. While this range is large, the point at which the E2 intragenic signal

returns to vehicle levels (~50 kb from the transcription start site) falls within this range suggesting that the intragenic signal is a mark of active transcription. Third, the intragenic signal was observed on actively transcribed genes not known to be affected or controlled by ER and E2 such as *ACTB* and the signal was not altered by E2 treatment (data not shown). Fourth, taking the genomic locations of all peaks identified (both ER and pS118-ER) and calculating the total tag density present at these sites from each antibody revealed a large number of peaks identified by the pS118-ER #1 antibody which had little or no ER occupancy, suggesting interaction with a non-specific factor (Figure 2-7A). In contrast, the tag density of the other two pS118-ER antibodies correlated with the tag density of ER, indicating that these antibodies were detecting the phosphorylated form of ER (Figures 2-7B, 2-7C). Due to the non-specific signal in the pS118-ER #1 antibody and its potential to mask pS118-ER binding events as well as create false positives, data from the pS118-ER #1 ChIP-seq was excluded from further downstream analysis. All three pS118-ER ChIP-seq datasets and the ER ChIP-seq dataset are deposited in the GEO database (GSE117569).

Discussion

ER phosphorylation is known to play an important role in both ligand-dependent and ligand-independent ER gene transcription. Considering the many reported stimuli that can lead to phosphorylation of ER at S118, analyzing its function will be important for understanding both ligand-dependent and ligand-independent activation mechanisms. While much of the current research on pS118-ER has focused on its transcriptional functional and cofactor interactions, the extent to which pS118-ER affects the genome-wide binding profile of ER remains largely unknown. The work presented in this chapter first analyzes pS118-ER dynamics in response to various stimuli and demonstrates that phosphorylation on S118 and DNA binding are independent events. Multiple pS118-ER antibodies were assessed for their use in ChIP-seq and initial ChIP-seq results are presented, with evidence suggesting a non-specific interaction from one of the pS118-ER antibodies. These results are critical for providing the most accurate cistrome analysis for pS118-ER in MCF-7 cells.

ER is known to be downregulated by the proteasome in an E2-dependent manner (99,100), and pS118-ER was also found to be necessary for this downregulation through its regulation of ER ubiquitination (78,109). Consistent with these results, treatment of MCF-7 cells with E2 caused a rapid increase in pS118-ER, followed by a decrease in total ER levels. Treatment with tamoxifen led to sustained levels of pS118-ER and no degradation of ER up to 4 hours post treatment, indicating that pS118-ER alone is not sufficient to cause ER downregulation. Additionally, treatment with EGF led to a similar phosphorylation dynamic as E2, however the ER levels were not affected by the treatment, suggesting that the addition and removal of the phosphorylation is not sufficient

for ER downregulation. While phosphorylation at S118 occurs in response to E2, tamoxifen, and EGF, the set of cofactors recruited to ER in each treatment condition varies (31,33,77,97,201) which contributes to the different responses observed to each ligand. The E3 ligase E6-AP interacts with and ubiquitylates ER in a pS118-dependent manner leading to its degradation by the proteasome (109,252). Mutation of serine 118 to alanine prevents E2-induced ER downregulation by preventing the interaction with E6-AP (78,109). The results presented here suggest that E6-AP is not recruited to ER in response to tamoxifen or EGF, given that ER levels are not altered by either treatment.

This report and others have found that phosphorylation is an early event in the ER activation mechanism, wherein binding of ligand results in phosphorylation of ER followed by DNA binding and activation of target genes. The data presented in this chapter indicate that phosphorylation of ER and DNA binding occur independently, as neither is required for the other (Figure 2-2). Past reports have indicated that ER can be phosphorylated at S118 by the DNA-dependent protein kinase (DNA-PK), however these studies were performed in the MELN breast cancer cell line and it is unknown if this same ER phosphorylation mechanism occurs in MCF-7 cells (73). In a proteomics analysis of elements bound to EREs upon E2 treatment, DNA-PK was found to interact with ER on DNA, but the direct phosphorylation of ER by DNA-PK was not demonstrated (74). Given the numerous kinases which have been reported to phosphorylate ER in response to E2 (Table 1-1), it is likely that a single kinase is not solely responsible for inducing phosphorylation at S118 and that many phosphorylation pathways are activated simultaneously by E2 which all result in ER phosphorylation at S118. While the results in this chapter demonstrate that DNA binding is not required for pS118-ER (Figure 2-2C),

these results do not contradict the evidence from other groups which conclude that DNA-PK is a kinase capable of phosphorylating ER at S118.

Understanding the dynamics of pS118-ER first was critical for selecting suitable conditions for analysis of the pS118-ER cistrome. A 30 minute treatment with E2 was chosen due to its high levels of pS118-ER and because ER downregulation does not substantially occur at this point. Given past inconsistencies with phospho-antibodies (245,246,253,254) and the degree to which antibody specificity can affect ChIP-seq analysis (255), multiple pS118-ER antibodies were assessed for their use in ChIP-seq. Although the ENCODE guidelines for ChIP-seq antibody validation were followed (247) and all three passed the criteria, one of the antibodies still displayed a non-specific RNA-Pol II-like signal. These results demonstrate the importance of scrutinizing ChIP-seq data after the ChIP-seq experiment has been performed and suggests additional methods of validation should be considered in addition to the assays put forth by the ENCODE consortium. For example, in addition to an immunoblot assay, antibodies should be validated with a conventional ChIP assay prior to ChIP-seq since detection of the protein of interest on an immunoblot is in a denatured environment whereas a ChIP assay captures the protein in its native folded conformation. Ideally, in addition to the cell line of interest, ChIP antibody validation would be performed in a cell line not expressing the target of interest or in conditions where the protein of interest was knocked down as a negative control. These additional experiments would make antibody validation more robust and add to the confidence in the data acquired by ChIP-seq. While one of the pS118-ER antibodies did not produce accurate data, the two other pS118-ER antibodies generated datasets for which analysis was possible. The data collected here provide the

groundwork for analyzing the pS118-ER cistrome, and the results of this analysis are provided in the next chapter.

Clinically, the role of pS118-ER is not well understood. Given that pS118-ER is an indicator of an activated ER signaling pathway, the majority of studies analyzing primary breast cancer tumors have tested pS118-ER for prognostic value in terms of overall survival, disease-free survival, or treatment efficacy. Immunohistochemistry (IHC) staining of these tumors has returned conflicting results, with some studies finding pS118-ER associated with better disease outcome (202,203) or better response to tamoxifen (204,205), whereas other reports have found a correlation between pS118-ER and resistance to tamoxifen (206). Additionally, a recent large-scale study of primary breast tumors analyzing multiple phosphorylation marks in the MAPK and Akt pathways including pS118-ER and pS167-ER found no prognostic correlations between any of the phosphorylation marks and clinical outcome (207). Interestingly, most of these studies used the same pS118-ER antibody that produced the non-specific signal in our pS118-ER ChIP-seq. Additionally, one study reported that some ER-negative tumors stained positive for pS118-ER using this antibody (256). We have additionally observed non-specific interactions with this antibody in immunofluorescence staining (data not shown) and its non-specific interaction may underlie the discrepancies observed in clinical data.

Figure 2-1: Temporal dynamics of pS118-ER in response to stimuli.

MCF-7 cells were hormone deprived for three days in phenol-red free DMEM supplemented with charcoal-stripped fetal bovine serum followed by treatment with 10 nM E2 (A), 100 nM 4-hydroxytamoxifen (B), or 100 nM EGF (C) for the times specified. Protein was isolated and analyzed by immunoblot for pS118-ER and total ER. ß-actin served as a loading control.

Figure 2-1

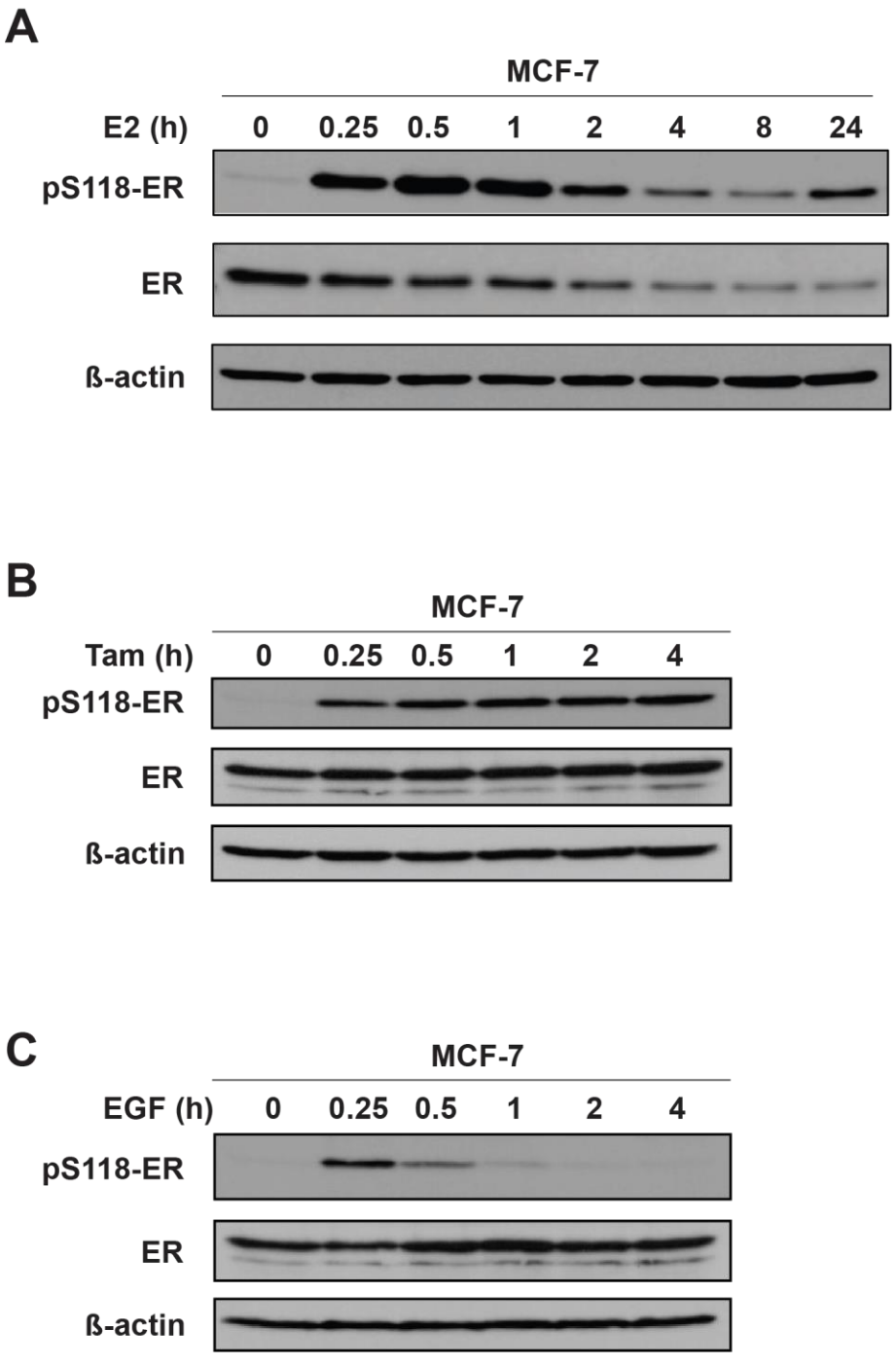

A
MCF-7
E2 (h)
0
0.25
0.5
1
2
4
8
24
pS118-ER
ER
ß-actin
B
MCF-7
Tam (h)
0
0.25
0.5
1
2
4
pS118-ER
ER
ß-actin
C
MCF-7
EGF (h)
0
0.25
0.5
1
2
4
pS118-ER
ER
ß-actin

Figure 2-2: Relationship between pS118-ER and DNA binding.

(A) Western blot analysis of MDA-MB-231 cells containing doxycycline-inducible wt ER or S118A ER. MDA-MB-231 cells were treated with 5.0 µg/mL (wt ER) or 0.5 µg/mL (S118A ER) dox for 24 h followed by a 30 min treatment of vehicle (0.1% EtOH) or 10 nM E2. ß-actin serves as a loading control. (B) ER ChIP-qPCR in MDA-MB-231 wt ER or S118A ER-expressing cells. Data are displayed as a percentage of input. n = 3, mean ± SD is shown. *p < 0.05. (D) HEK293 cells transfected with either vector, wt ER, or the DNA binding ER mutant C202/205H were treated for 30 min with either vehicle or 10 nM E2 and analyzed via immunoblot.

Figure 2-2

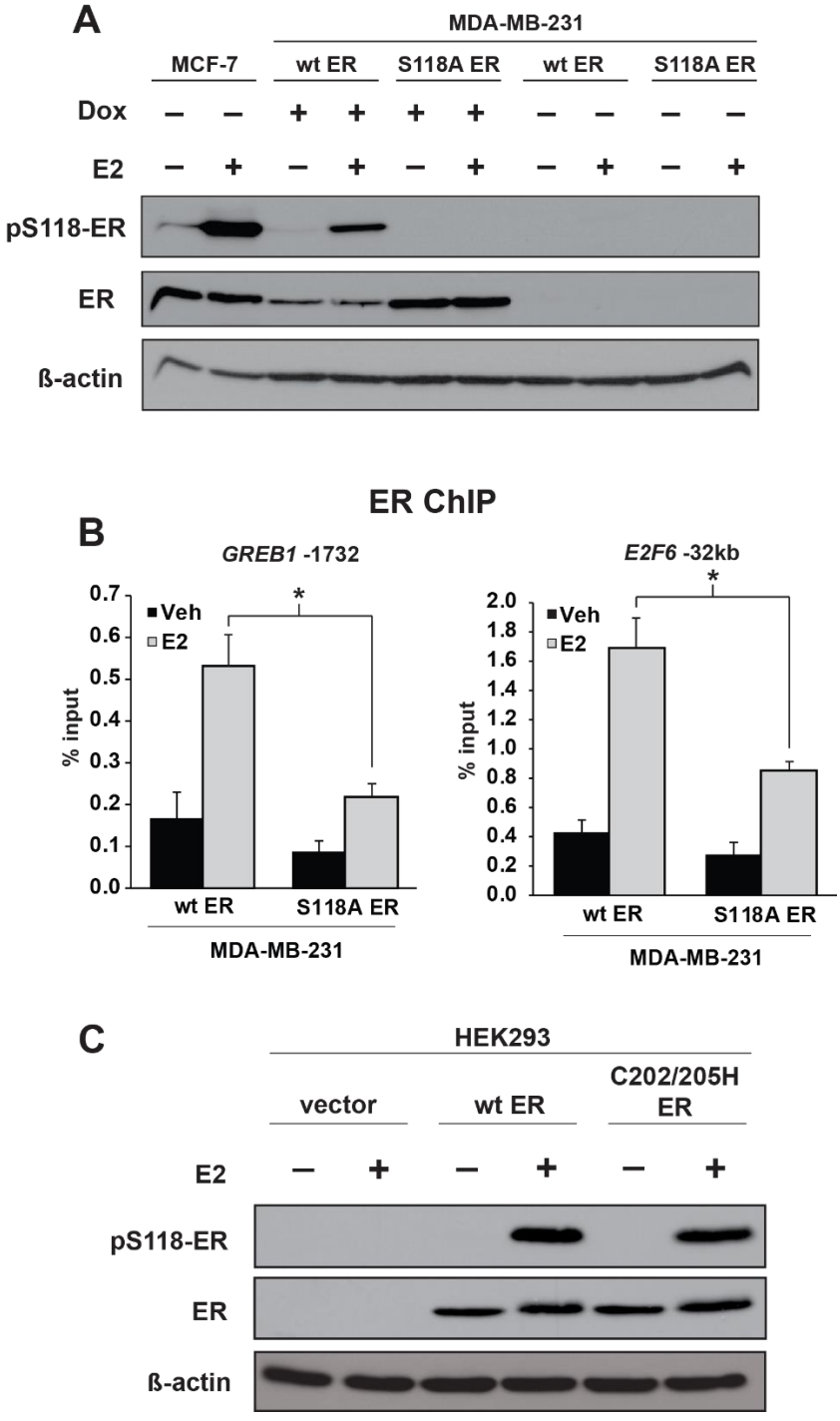

A
MDA-MB-231
MCF-7
wt ER
S118A ER
wt ER
S118A ER
Dox
E2
pS118-ER
ER
ß-actin
ER ChIP
B
GREB1 -1732
E2F6 -32kb
% input
Veh
E2
wt ER
S118A ER
MDA-MB-231
C
HEK293
vector
wt ER
C202/205H ER
E2
pS118-ER
ER
ß-actin

Figure 2-3: Validation of pS118-ER antibodies following the ENCODE guidelines.

(A) Primary validation of antibodies. Immunoblot analysis of MCF-7 cells treated with either vehicle (0.1% EtOH) or 10 nM E2 for 30 minutes. Three pS118-ER antibodies (pS118-ER #1, Cell Signaling; pS118-ER #2, Abcam; pS118-ER #3, Santa Cruz) were tested for detection of pS118-ER in the E2-treated sample and minimal detection in the vehicle sample. Three separate blots were run, one for each antibody to avoid any stripping and re-probing procedures. Full blots for each pS118-ER antibody are shown to display non-specific bands detected. Expected MW for pS118-ER is ~66kDa. ß-actin serves as a loading control. (B) Secondary validation of antibodies with the S118A ER mutant. Portions of Figure 2-2A are reprinted here for clarity. MDA-MB-231 cells expressing either wt ER or S118A were treated with vehicle (0.1% EtOH) or 10 nM E2 for 30 minutes and analyzed by immunoblot. All three pS118-ER antibodies tested displayed a loss of signal in the S118A mutation indicating specificity for pS118 and not other phosphorylation sites on ER. (C) DNA Motif analysis of ChIP-seq peaks detected for each pS118-ER antibody. HOMER *de novo* motif analysis (175) was performed on the peaks identified in the E2 treated samples and the top motif identified is displayed. The presence of the ERE indicates the specificity of the pS118-ER antibodies for ER.

Figure 2-3

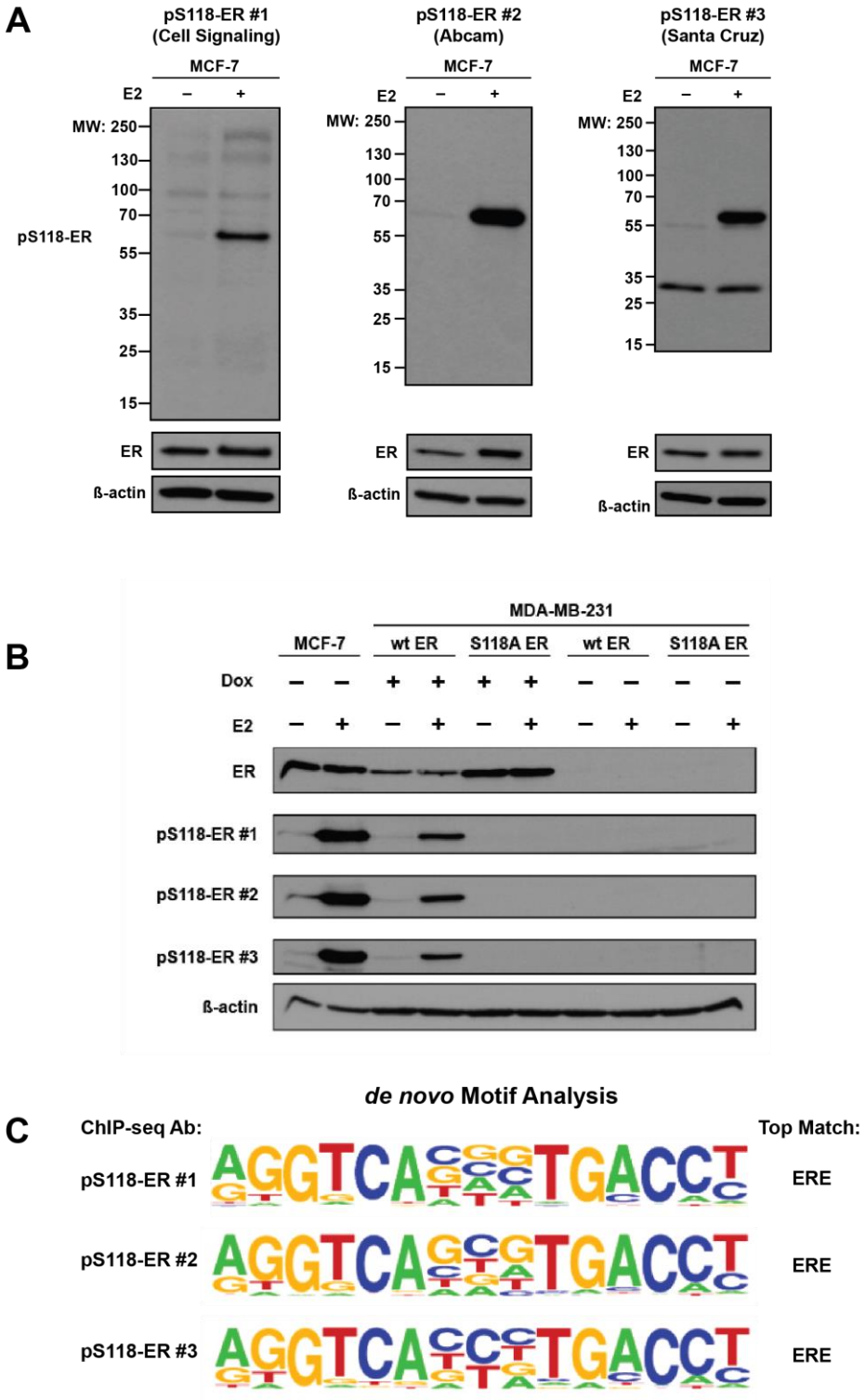

A
pS118-ER #1 (Cell Signaling)
pS118-ER #2 (Abcam)
pS118-ER #3 (Santa Cruz)
MCF-7
E2
MW: 250
130
100
70
55
35
25
15
pS118-ER
ER
ß-actin
B
MDA-MB-231
MCF-7
wt ER
S118A ER
Dox
E2
ER
pS118-ER #1
pS118-ER #2
pS118-ER #3
ß-actin
C
de novo Motif Analysis
ChIP-seq Ab:
Top Match:
pS118-ER #1
pS118-ER #2
pS118-ER #3
ERE

Figure 2-4: ChIP-seq peak overlaps.

ChIP-seq was performed in MCF-7 cells treated with either vehicle (0.1% EtOH) or 10 nM E2 for 30 minutes. One antibody for total ER and three antibodies for pS118-ER were used. Three replicates were performed for each antibody and peaks were called using HOMER (175). Overlaps between the replicates within each antibody were performed and only peaks present in all three replicates were used for further analysis. (A) Overlaps between replicates in the E2-treated samples for each antibody. (B) Overlap between vehicle and E2 peaks within each antibody. (C) Overlap between peaks detected by the three pS118-ER antibodies in the E2-treated samples. (D) Overlap between ER and each separate pS118-ER antibody in the E2-treated samples.

Figure 2-4

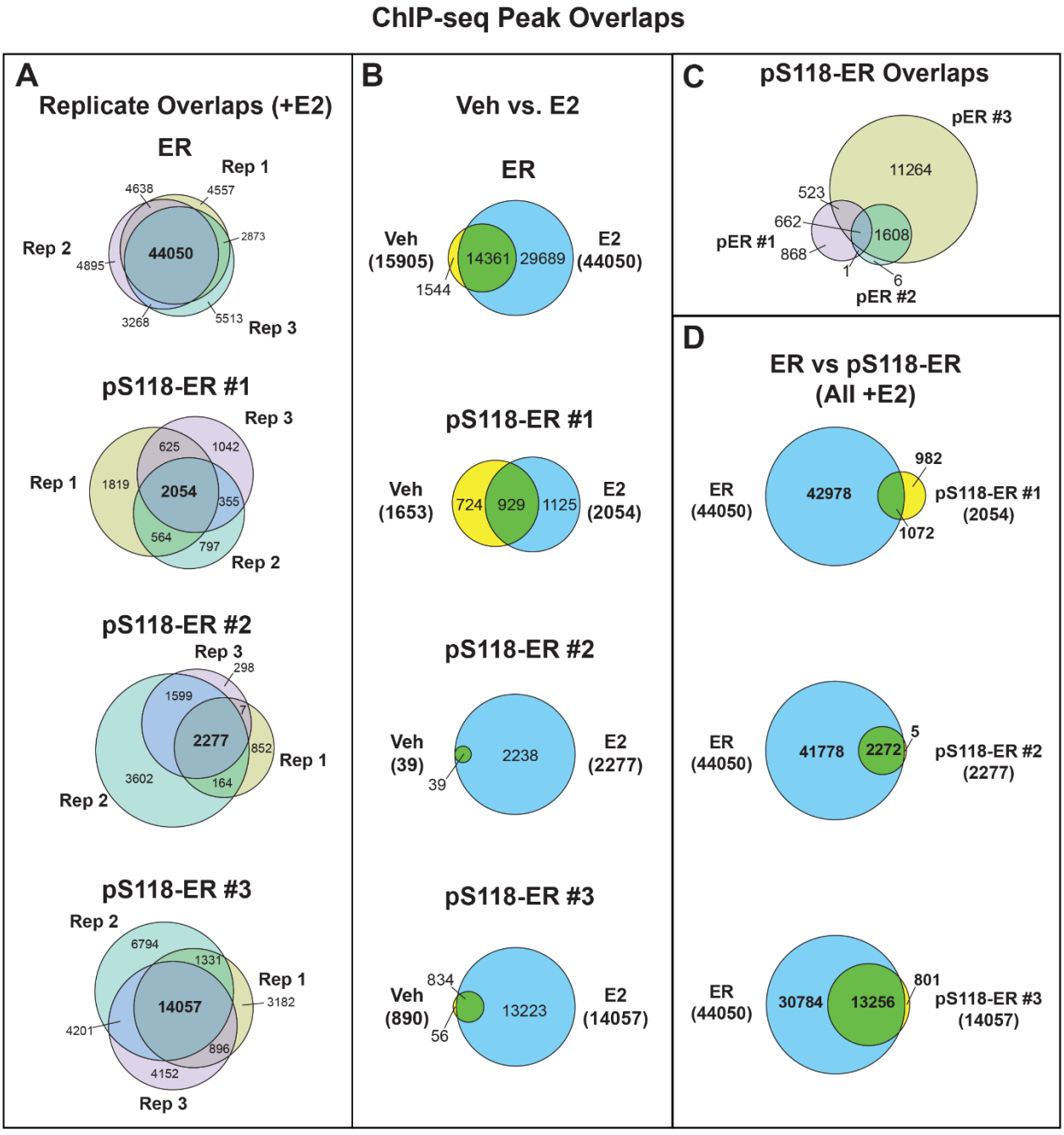

ChIP-seq Peak Overlaps
A
Replicate Overlaps (+E2)
ER
Rep 1
Rep 2
Rep 3
4638
4557
2873
4895
44050
3268
5513
pS118-ER #1
Rep 3
Rep 1
Rep 2
625
1042
1819
2054
355
564
797
pS118-ER #2
Rep 3
298
1599
7
2277
852
Rep 1
Rep 2
3602
164
pS118-ER #3
Rep 2
6794
1331
Rep 1
14057
3182
4201
896
4152
Rep 3
B
Veh vs. E2
ER
Veh
(15905)
14361
29689
E2
(44050)
1544
pS118-ER #1
Veh
(1653)
724
929
1125
E2
(2054)
pS118-ER #2
Veh
(39)
2238
E2
(2277)
39
pS118-ER #3
Veh
(890)
834
13223
E2
(14057)
56
C
pS118-ER Overlaps
pER #3
11264
523
662
pER #1
1608
868
1
6
pER #2
D
ER vs pS118-ER
(All +E2)
ER
(44050)
42978
982
pS118-ER #1
(2054)
1072
ER
(44050)
41778
2272
5
pS118-ER #2
(2277)
ER
(44050)
30784
13256
801
pS118-ER #3
(14057)

Figure 2-5: ChIP-seq validation and visualization.

(A) ChIP-qPCR validation of select pS118-ER occupancy sites. The number after the gene name denotes the approximate distance in bp between the region tested and the transcription start site of the gene with a negative number being upstream and a positive number being downstream. Data are displayed as a percentage of input. n = 3, mean ± SD is shown. *p < 0.05 vs vehicle (veh). (B) ChIP-seq track display from the UCSC Genome Browser showing ER and pS118-ER occupancy sites at the *TFF1* (top), *BRIP1* (middle), and *GREB1* (bottom) loci. For each track, vehicle (yellow) and E2-treated (blue) are overlaid. Overlapping tracks are displayed as green. The y-axis displays tag density normalized to 10^7 tags. Black bars below the tracks denote the approximate amplicon location of primer sets used in (A).

Figure 2-5

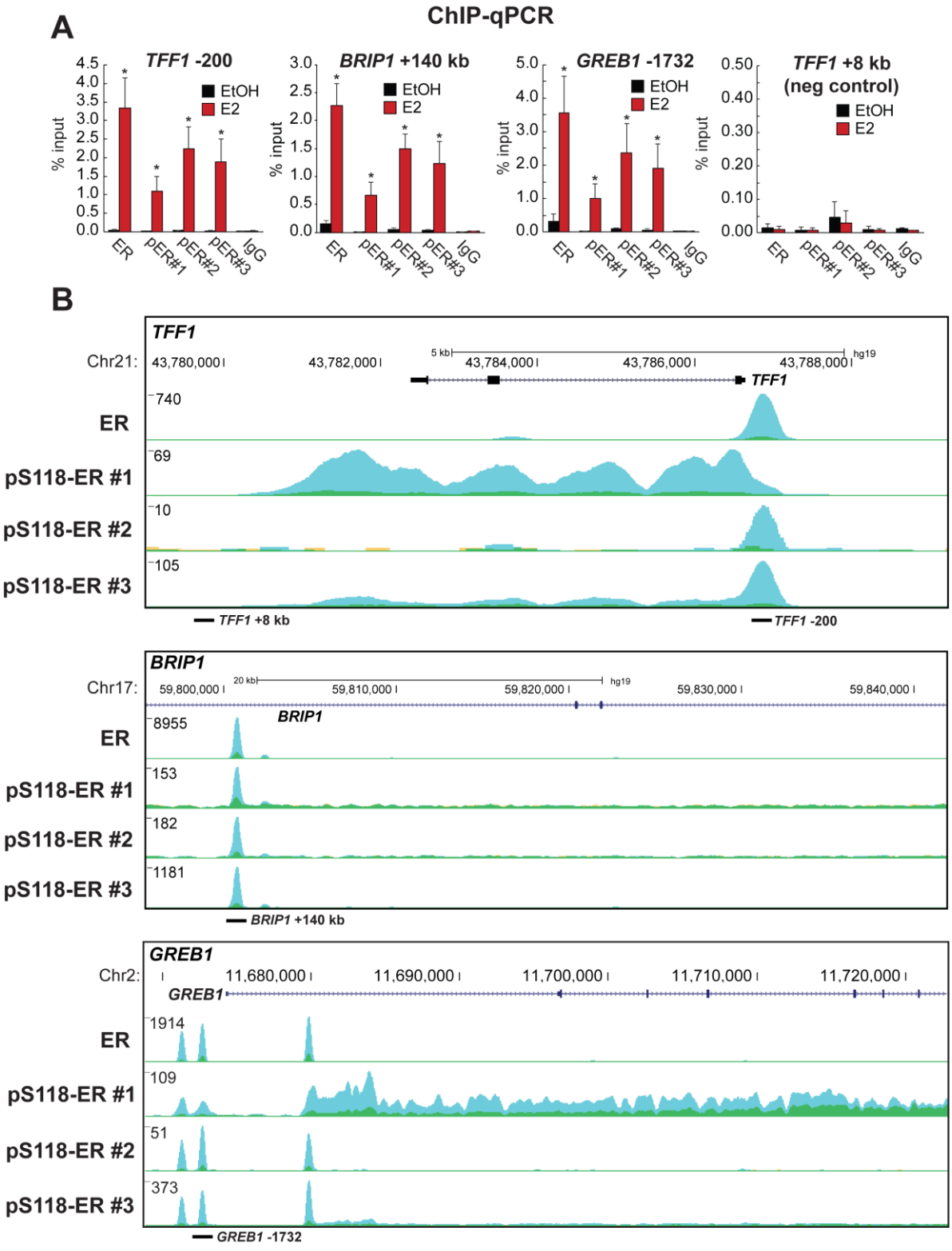

A
ChIP-qPCR
TFF1 -200
BRIP1 +140 kb
GREB1 -1732
TFF1 +8 kb
(neg control)
EtOH
E2
% input
ER
pER#1
pER#2
pER#3
IgG
B
TFF1
Chr21:
43,780,000
43,782,000
43,784,000
43,786,000
43,788,000
hg19
740
69
10
105
ER
pS118-ER #1
pS118-ER #2
pS118-ER #3
TFF1 +8 kb
TFF1 -200
BRIP1
Chr17:
59,800,000
59,810,000
59,820,000
59,830,000
59,840,000
8955
153
182
1181
BRIP1 +140 kb
GREB1
Chr2:
11,680,000
11,690,000
11,700,000
11,710,000
11,720,000
1914
109
51
373
GREB1 -1732

Figure 2-6: Intragenic signal observed in pS118-ER #1 antibody is altered by E2 on E2-responsive genes.

Visualization of ER and pS118-ER #1 (p-ER) ChIP-seq tracks on the UCSC genome browser at selected sites. Vehicle tracks (yellow) and E2-treated tracks (blue) are overlaid. Overlapping tracks are displayed as green. The y-axis displays tag density normalized to 10^7 tags. (A) Tracks for genes known to be upregulated by E2 (*CCND1*, *FOS*, *GREB1*, *XBP1*) display increased intragenic signal in the pS118-ER #1 track in response to E2. (B) Tracks for genes known to be downregulated by E2 (*ESR1*, *GATA3*, *NFKBIZ*, *PCDH10*) display decreased intragenic signal in the pS118-ER #1 track in response to E2. The lack of a similar signal in the ER tracks indicates a non-specific interaction in the pS118-ER #1 antibody.

Figure 2-6

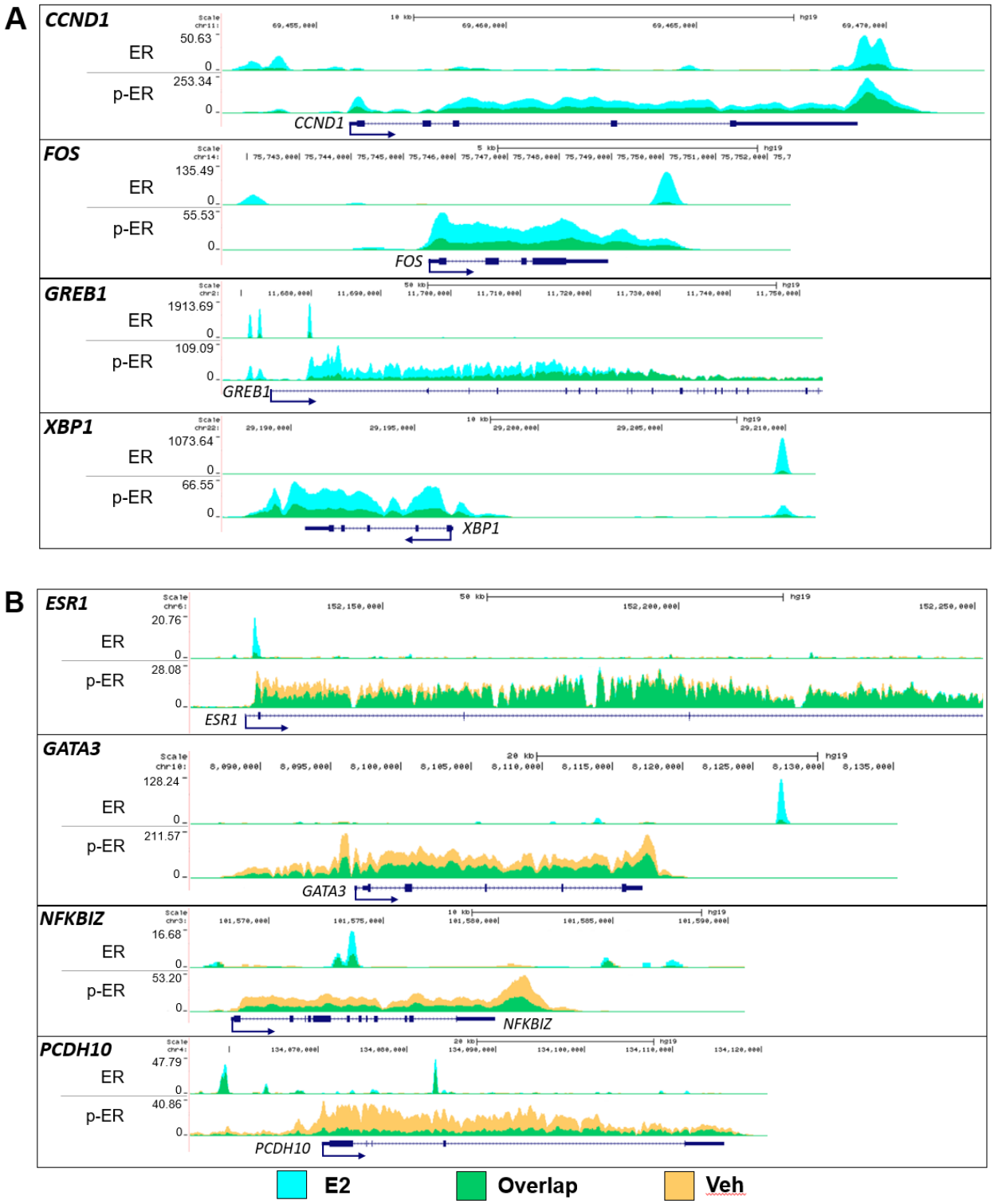
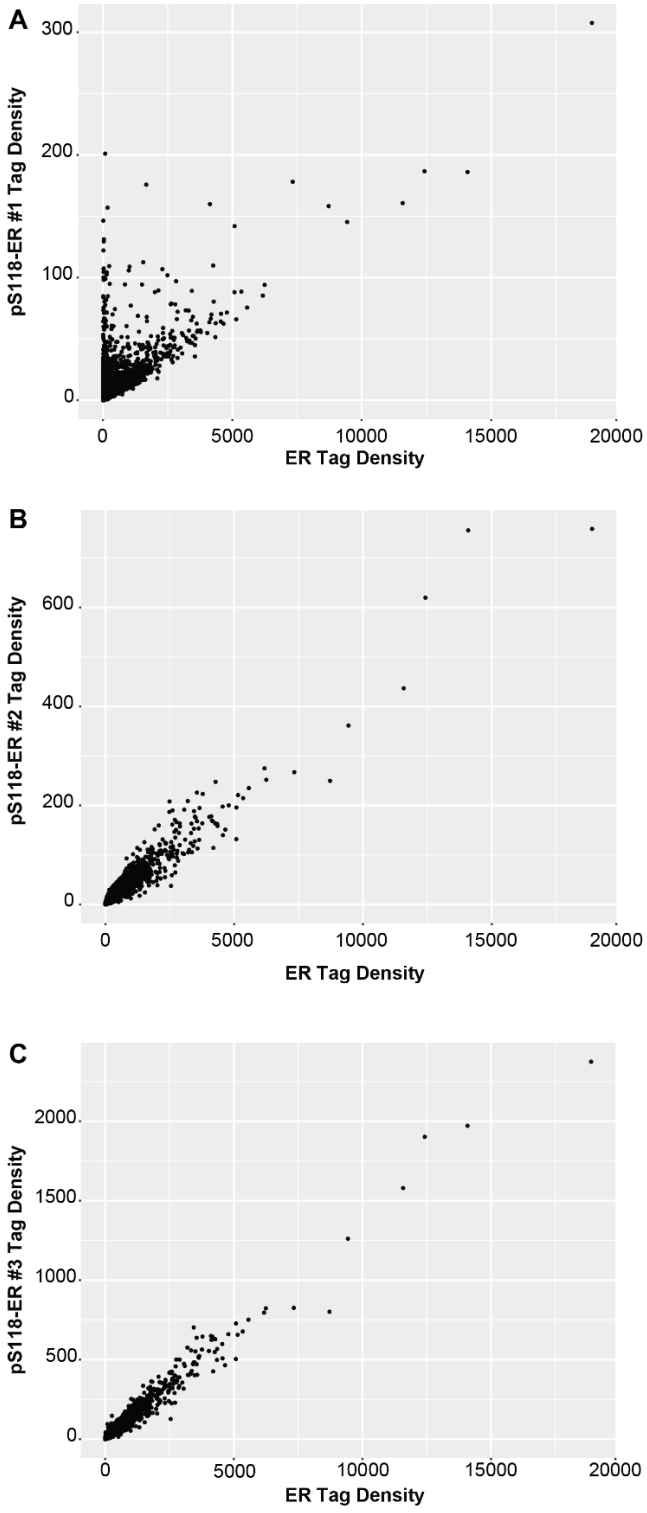

A
CCND1
ER
50.63
0
p-ER
253.34
0
CCND1
FOS
ER
135.49
0
p-ER
55.53
0
FOS
GREB1
ER
1913.69
0
p-ER
109.09
0
GREB1
XBP1
ER
1073.64
0
p-ER
66.55
0
XBP1
B
ESR1
ER
20.76
0
p-ER
28.08
0
ESR1
GATA3
ER
128.24
0
p-ER
211.57
0
GATA3
NFKBIZ
ER
16.68
0
p-ER
53.20
0
NFKBIZ
PCDH10
ER
47.79
0
p-ER
40.86
0
PCDH10
E2
Overlap
Veh

Figure 2-7: Tag density correlation between ER and pS118-ER antibodies.

All genomic locations found to contain either ER or pS118-ER occupancy were analyzed for total tag density within each antibody and plotted. Tag density in each region was calculated using the annotatePeaks.pl program available in HOMER (175). Each dot on the graph represents a single genomic location. The tag density of ER was compared to (A) pS118-ER #1, (B) pS118-ER #2, and (C) pS118-ER #3. Units for tag density are normalized to 10^7 tags. The presence of data points along the y-axis in (A) suggests a non-specific interaction.

Figure 2-7


A
pS118-ER #1 Tag Density
300
200
100
0
0
5000
10000
15000
20000
ER Tag Density
B
pS118-ER #2 Tag Density
600
400
200
0
0
5000
10000
15000
20000
ER Tag Density
C
pS118-ER #3 Tag Density
2000
1500
1000
500
0
0
5000
10000
15000
20000
ER Tag Density

CHAPTER THREE:

Analysis of the pS118-ER Cistrome

Portions of this chapter are published in the following:

Helzer KT, Szatkowski Ozers M, Meyer MB, Benkusky NA, Solodin N, Reese RM, Warren CL, Pike JW, Alarid ET. The Phosphorylated Estrogen Receptor α (ER) Cistrome Identifies a Subset of Active Enhancers Enriched for Direct ER-DNA Binding and the Transcription Factor GRHL2. *Mol. Cell. Biol.* 2019;39(3). doi:10.1128/MCB.00417-18.

Abstract

The phosphorylation of proteins is an integral aspect of cellular function, controlling various aspects of protein activity including sub-cellular localization, substrate interactions, and enzymatic function amongst others. Many nuclear receptors are phosphorylated which contributes to their activity in numerous ways. Estrogen receptor-α (ER) is phosphorylated at S118 (pS118-ER) in response to multiple stimuli and is involved in modulating ER-dependent gene transcription. While the genome-wide DNA binding profile of ER is well established, surprisingly little is understood about how phosphorylation impacts ER-DNA binding activity. To define the pS118-ER cistrome, ChIP-seq was performed on pS118-ER and ER in MCF-7 cells treated with estrogen. pS118-ER occupied a subset of ER binding sites which were associated with the active enhancer mark, acetylated H3K27. Unlike ER, pS118-ER sites were enriched in GRHL2 DNA binding motifs, and estrogen treatment increased GRHL2 recruitment to sites occupied by pS118-ER. Additionally, pS118-ER occupancy sites showed greater enrichment of full-length estrogen response elements relative to ER sites. In an *in vitro* DNA binding array of genomic binding sites, pS118-ER was more commonly associated with direct DNA binding events compared to indirect binding events. These results indicate that phosphorylation of ER at serine 118 promotes direct DNA binding at active enhancers and is a distinguishing mark for associated transcription factor complexes on chromatin.

Introduction

Post-translation modifications play a major role in the control of protein function. Phosphorylation in particular provides a quick and often reversible signal which can alter protein function in a variety of ways. The majority of nuclear receptors are phosphorylated (257) which can affect transcriptional activity, dimerization, nuclear localization, or protein stability (52,57,78,143,258,259). Estrogen receptor-α (ER) is a member of the steroid nuclear receptor family and is involved in the development and maintenance of estrogen-responsive tissue such as the breast. Binding of 17ß-estradiol (E2) to the ligand-binding domain of ER leads to ER dimerization followed by DNA binding and activation of target genes. In addition to ligand-dependent activation mechanisms, ER can be activated in a ligand-independent manner through a variety of stimuli including epidermal growth factor, reactive oxygen species, and prolactin. A commonality between these activation mechanisms is the phosphorylation of ER at S118 (pS118-ER). In addition to the ligand-independent stimuli listed above, pS118-ER can be induced by E2 as well as the ER antagonists tamoxifen and ICI 182,780 (52). Mutational analysis of S118 to alanine (S118A), found that phosphorylation at S118 was required for maximal ER target gene activation (59,77–81), but less is known on how pS118-ER affects ER-DNA binding.

An integral property required for the function of ER is its ability to bind DNA. Early studies analyzing the DNA binding domains (DBDs) of the steroid receptors demonstrated a high level of specificity between nuclear receptor DBDs and their respective response elements (16). The ERE is classically defined as the 15 base pair (bp) palindromic sequence 5’-AGGTCA-nnn-TGACCT-3’ where “n” is any nucleotide (122,123). Studies investigating the sequence specificity of the ERE found that even small changes in either

the ERE sequence or the spacing between the two half-sites can drastically decrease the affinity of ER for the modified ERE (124,129,133). ER can also associate with chromatin indirectly through tethering mechanisms with other transcription factors including AP-1, Sp-1, and RUNX1 (117,118,121,159,260–262). While DNA sequence appears to be the main driver of ER affinity for the ERE, there is some *in vitro* evidence that pS118-ER can increase the affinity of ER for the canonical ERE (135). Structural studies suggest that phosphorylation events in the N-terminus of ER can be propogated through to affect the neighboring DBD (108,110,263), thereby affecting DNA binding affinity and specificity. Chromatin immunoprecipitation studies have shown that pS118-ER is present at the promoters of certain ER target genes as well as at enhancers as part of the “MegaTrans” complex as defined by Liu *et al.*, however the genome-wide binding landscape (cistrome) of pS118-ER has not been thoroughly investigated (136,242).

In this report, we define the pS118-ER cistrome, the first in depth analysis of a cistrome for a post-translationally modified nuclear receptor. We found that while relatively few pS118-ER occupancy sites were present in promoter regions, pS118-ER did associate with both acetylated H3K27, a mark of active enhancers, as well as upregulation of nearby genes. Further analysis revealed pS118-ER preferentially associated with EREs as well as the binding motif for the transcription factor GRHL2 relative to all ER sites. GRHL2 occupancy was found to increase upon E2 treatment at sites co-occupied by GRHL2 and pS118-ER, but not at sites with GRHL2 alone. We also utilized the Specificity and Affinity for Proteins (SNAP) DNA array to assess direct versus indirect binding events and found that direct binding events were more likely to be occupied by pS118-ER. SNAP array analysis also revealed the presence of multiple

potential ER binding events within single ER ChIP-seq peaks. Additionally, the number of potential ER binding events in a binding region, identified by the SNAP array, positively correlated with pS118-ER occupancy by ChIP-seq. These results highlight the role of pS118-ER at the ER-chromatin interface and as a regulator of the ER-ERE interaction.

Results

Analysis of pS118-ER Occupancy Sites

As stated in Chapter 2, three pS118-ER antibodies were used to assess the pS118-ER cistrome in E2-treated MCF-7 cells with one of the antibodies displaying a non-specific interaction. The data from this antibody was removed and the remaining two pS118-ER ChIP-seq data sets were combined into one pS118-ER data set (Figure 3-1A). This resulted in a pS118-ER data set containing 14,063 sites, of which 13,261 (94.3%) overlapped with ER sites. pS118-ER peaks that did not overlap with ER peaks were considered false positives, and therefore only those pS118-ER peaks that overlapped with ER occupancy sites were used for further analysis. The ER peaks were divided into those sites occupied by pS118-ER (pS118-ER; 13,261 peaks) and those sites not occupied by pS118-ER (ER-only; 30,794 peaks) for comparative analysis (Figure 3-1A). To test if the genomic locations of pS118-ER sites differed from ER-only sites, peak locations were analyzed and compared. Consistent with previous studies, we found that the majority of ER-only sites were present in introns or intergenic sites with few sites located in promoter regions (121,156,188,189). Notably, the locations of pS118-ER sites followed this same pattern, with no enrichment seen in promoter sites compared to ER-only sites (Figure 3-1B). Additionally, when peaks were annotated to the nearest transcription start site, we found a similar distribution in distance to transcription start site between ER sites and pS118-ER sites indicating that pS118-ER is not enriched in locations closer to gene promoters relative to ER-only sites (Figure 3-1C). To test if the phosphorylation event is associated with active sites, we analyzed and compared data from a previous report on the active enhancer mark H3K27ac in MCF-7 cells treated with

E2 (66, 67; GSE45822). Interestingly, we found that pS118-ER sites displayed an increase in the H3K27ac signal compared to ER-only sites (Figure 3-1D). To test if pS118-ER is associated with activation of genes, the pS118-ER ChIP-seq data was integrated with previously published RNA-seq data from MCF-7 cells treated with 1 nM E2 for 24 h (68, GSE89888). Genes were first categorized into significantly upregulated (3,692 genes) or significantly downregulated (3,006 genes) by E2. Each gene was then analyzed for the presence of a pS118-ER peak within 100 kb of its transcription start site (TSS). We found that 29.9% of upregulated genes had a pS118-ER site within 100 kb of its TSS compared to only 8.3% of downregulated genes suggesting that pS118-ER is associated with gene activation rather than gene repression (Figures 3-1E, 3-1F). These results indicate that while pS118-ER is not enriched in promoter regions, the phosphorylation event does appear to be a mark of active enhancers and is associated with activation of genes.

Motif Analysis of pS118-ER Sites Reveals Enrichment of EREs and GRHL2 Motifs Relative to ER

Previous work has shown that other accessory factors are important for ER-DNA interaction including the pioneer factor FOXA1, as well as tethering factors such as AP-1, Sp1, and RUNX1 (118,121,156,159,261). These factors are often identified by the presence of their respective DNA binding motif near ER occupancy sites. To determine enriched DNA motifs at ER and pS118-ER occupancy sites, *de novo* motif analysis was performed on ER-only (30,794 sites), and pS118-ER (13,261 sites). As expected, the most significantly enriched motif at both ER-only and pS118-ER sites was the consensus

ERE (Figure 3-2A). Interestingly, we also found enrichment of the grainyhead-like 2 (GRHL2) motif in the pS118-ER data set (p = 1e-493) (Figure 3-2A). The GRHL2 motif was the fourth most enriched motif in the pS118-ER data set after the ERE (p = 1e-4685), Forkhead motifs (p = 1e-652), and NR2F2 (COUP-TFII or an ERE half-site, p = 1e-617) (Figure 3-2A). To determine which motifs were enriched at pS118-ER sites relative to total ER sites (44,050 peaks), differential motif analysis was performed with the MEME suite CentriMo tool using total ER sites as background (184). This analysis revealed that the pS118-ER sites were significantly enriched for both the ERE motif and the GRHL2 motif relative to ER sites (Figure 3-2B). Additionally, we observed an increase in GRHL2 motifs in the center of pS118-ER sites compared to ER-only sites (Figure 3-2D, GRHL2). When we analyzed the proportion of sites that contained certain motifs in each data set, we found that pS118-ER sites were more likely to contain an ERE compared to ER-only sites (Figure 3-2C). On average, these EREs were positioned in the center of peaks in both ER-only and pS118-ER occupancy regions (Figure 3-2D, ERE). We also found a significant decrease in the proportion of pS118-ER sites containing GATA3 motifs compared to ER-only sites (Figures 3-2C, 3-2D). Additionally, we observed a slight but significant increase in the proportion of pS118-ER sites containing AP-2γ motifs (Figures 3-2C, 3-2D). Motifs for Forkhead factors (e.g. FOXA1) were found to be enriched in both ER-only and pS118-ER sites, but they were not found to be enriched compared to each other (Figures 3-2C, 3-2D). These data indicate that pS118-ER sites are more likely to contain ERE and GRHL2 DNA motifs compared to ER-only sites. Additionally, FOXA1 motifs are equally distributed between pS118-ER and ER-only sites suggesting that FOXA1 is not a differentiating factor between sites occupied by pS118-ER and ER.

GRHL2 Occupancy Increases in Response to E2 at pS118-ER Sites

In order to determine the occupancy status of GRHL2 at pS118-ER and ER sites, we analyzed and compared GRHL2 ChIP-seq data from a previously published report which performed ChIP-seq on GRHL2 in MCF-7 cells in full media and identified 30,134 GRHL2 occupancy sites (265). When the GRHL2 sites were overlapped with ER-only and pS118-ER sites, we found that a greater proportion of pS118-ER sites (32.1%) were co-occupied by GRHL2 compared to ER-only sites (13.6%, Figure 3-3A). Additionally, pS118-ER sites were found to have an increased GRHL2 signal on average compared to ER-only sites (Figure 3-3B) indicating that GRHL2 more commonly occupies sites that are occupied by pS118-ER. Given that pS118-ER is induced by E2 (Figure 2-1A) and that GRHL2 associates with pS118-ER occupancy sites, we hypothesized that GRHL2 occupancy would be altered by E2 treatment. MCF-7 cells were stripped and treated with either vehicle or E2 for 30 min and ChIP-qPCR was performed for ER, pS118-ER, and GRHL2. Using the available ChIP-seq data for GRHL2 (265), sites were identified which were either co-bound by pS118-ER and GRHL2, or bound by GRHL2 only. GRHL2 occupancy increased upon E2 treatment at sites co-occupied by pS118-ER and GRHL2 (Figure 3-3C), but not at sites solely occupied by GRHL2 (Figure 3-3D). These results indicate that estrogen induces an increase in GRHL2 binding, preferentially at sites to which pS118-ER is recruited.

Direct ER Binding Events are More Likely to be Occupied by pS118-ER

Because of the enrichment of EREs at pS118-ER sites over ER sites (Figs. 4B, 4C), we hypothesized that pS118-ER was associated with direct ER binding events rather than tethering events. To address this question, we utilized purified phosphorylated ER to probe a genomic Specificity and Affinity for Proteins (SNAP) DNA binding array (110,266). Immunoblot analysis of the purified ER used in our experiments revealed that S118 is phosphorylated while phosphorylation at S104/106 and S294 was not detected. S167 also appeared to be modified. Importantly, the purified receptor co-migrated with the estrogen-induced, post-translationally modified receptor in MCF7 cell extract (Figure 3-4B) indicating that the majority of the purified receptor is modified. A custom-designed SNAP array composed of 359,156 25-bp double-stranded DNA hairpin probes, capable of forming B-form DNA was created to tile across 5,020 genomic regions that had been confirmed by multiple reports to be occupied by ER (140,159,187,188,196,201,234,236). We incubated purified, phosphorylated ER with E2, exposed the ER-E2 complex to the array, and visualized binding of ER to the DNA features by fluorescence (Figure 3-4A). Of the 5,020 regions tiled, Figure 3-4C depicts a representative ER occupancy site, the *GREB1* promoter, which has been identified in previous ChIP-seq studies (140), and the individual probes which tile across its associated DNA sequence that were synthesized on the SNAP array. The resulting SNAP intensity shows binding of pS118-ER within the region that is dependent on the ERE site (Figure 3-4C). SNAP intensity has been shown to correlate with the affinity of ER for its DNA binding sites (110). Due to the lack of any accessory factors in the *in vitro* SNAP array assay, regions that displayed both binding on the SNAP array and ER occupancy in our ChIP-seq data were defined as direct binding

(Figures 3-5A, 3-5B, *TFF1* promoter), whereas regions that displayed no ER binding events on the SNAP array, but contained ER occupancy by ChIP-seq were defined as indirect ER interactions (Figures 3-5A, 3-5B, *IGF1R* +130kb). Of the 5,020 genomic regions represented on the array, 4,431 regions contained at least one probe with a significant level of ER binding whereas 589 regions displayed no significant ER binding (Figure 3-5C). Of the 4,431 direct ER binding regions on the SNAP array, 71.2% (3,575/4,431) were present in the pS118-ER ChIP-seq regions (Figure 3-5C). The remaining 589 regions represented potential indirect ER binding events and of these, 46.9% (276/589) were present in the pS118-ER ChIP-seq regions indicating that pS118-ER tends to associate with direct binding events rather than indirect binding events (Figure 3-5C). Interestingly, we found that many ER occupancy sites analyzed on the array contained multiple ER binding events throughout the region, rather than one located at the center (Figures 3-5B, 3-5D). For example, the *TFF1* promoter region which was detected as a single ChIP-seq peak, was found to have four potential ER binding events throughout the region (Figure 3-5B). Analysis of the number of ER binding events in each region analyzed on the genomic SNAP array showed that 56.4% of these regions contained two or more potential ER binding events (Figure 3-5D). When the regions were binned by number of ER binding events, a positive correlation was found between the number of ER binding events and the likelihood of occupancy by pS118-ER in our ChIP-seq data further supporting the notion that pS118-ER is more likely to associate with direct ER binding events. (Figure 3-5E).

Using the full genomic regions from the binned ER binding events of the SNAP array, we next queried the positional and spatial relationships between the direct ER

binding sites and inferred co-binders. First, we identified the ER binding motif as the top motif from the binned ER binding events (Figure 3-6) and found that a more complete ERE occurred in the regions with one peak as well as the regions with 6 or greater peaks while a 5'-AGGTCA-3' half-site predominated in the 2-, 3-, 4-, and 5-peak bins. These ER binding sites were then input as the primary motif in the Spaced Motif (SpaMo) analysis tool (267) and used to identify non-redundant neighboring secondary motifs. Based on the number of ER binding sites, different transcription factor binding motifs occurred near the ERE. For example, in the one-peak bin, POU5F1, VDR, and GRHL2 were the predominant neighboring motifs occurring in 455, 423, and 337 of the ER peaks, respectively. We next examined the gap in bps between the ERE and adjacent motifs. Interestingly, we found that the most frequent gap between the ERE and the GRHL2 motif was 50 bp (Figure 3-6). Previous work has demonstrated that each ER monomer of the ER dimer binds DNA in the major groove and that the ER dimer binds predominantly on one face of the DNA (268,269). The ER-DNA interaction can also cause bending in the associated DNA (270,271). B-form DNA completes a helical turn approximately once every 10 bp (272) suggesting that the 50 bp gap between the ERE and the GRHL2 motif would allow for potential contact with the ER dimer on the same face of DNA.

Discussion

Phosphorylation of nuclear receptors has been shown to be a critical regulator of their various functions (257). Considering the many reported stimuli that can lead to phosphorylation of ER at S118, understanding its function will be important for both ligand-dependent and ligand-independent activation mechanisms. A critical step in the ER activation pathway is its interaction with chromatin, and current studies on the effect of pS118-ER on this interaction are limited. Here we report the analysis of the pS118-ER cistrome in E2-stimulated MCF-7 cells and provide evidence that pS118-ER is associated with the transcription factor GRHL2 along with direct binding to EREs. By taking a comprehensive approach with multiple pS118-ER specific antibodies, we identified a subset of ER occupancy sites which contain pS118-ER. While these pS118-ER sites were not found to be enriched in promoter regions compared to ER-only occupancy sites, the pS118-ER sites were found to be enriched in the active enhancer mark H3K27ac as well as associated with activation of nearby genes. To our knowledge, this report is the first comprehensive analysis of a cistrome for a post-translationally modified nuclear receptor. Together with data that places pS118-ER as a component of the MegaTrans complex (242), these studies support a role for pS118-ER as a regulator of ER-DNA binding and as a mark of active enhancers.

Phosphorylation is described as an early event in the ER activation mechanism, wherein activation by ligand results in post-translational modification followed by DNA binding. Our data indicate that phosphorylation and DNA binding occur independently, as neither is required for the other. Rather, phosphorylation increases ER-DNA binding at specific enhancer sites. Studies that have performed ER ChIP-seq have reported a low

proportion of occupancy sites within promoter regions and that most ER occupancy sites fall greater than 20 kb away from a transcription start site (121,156,188,189). We additionally found that pS118-ER follows this same pattern indicating that pS118-ER is not enriched in promoter regions compared to ER-only sites (Figure 3-1B). The mechanism by which ER can have effects on gene transcription from long, linear genomic distances has been examined via chromatin interaction analysis by paired-end tag sequencing, and it was found that distal ER occupancy sites could be anchored to gene promoters even over long genomic distances (185). We found that pS118-ER sites correlate with higher levels of H3K27ac at distal enhancers, consistent with the notion that active phosphorylated ER regulates transcription via enhancer control.

Interestingly, we found an enrichment for the GRHL2 binding motif at pS118-ER sites relative to ER-only sites. Overlapping pS118-ER and ER-only sites with a previously published GRHL2 ChIP-seq data set (265) revealed that a higher proportion of pS118-ER sites overlap with GRHL2 sites. GRHL2 occupancy was also found to increase upon E2 treatment at these overlapping sites, but not at GRHL2 sites devoid of pS118-ER. This finding is supported by a recent preprint by Holding *et al.* which reported a genome-wide increase of GRHL2 binding in the presence of E2 and many of the E2-induced GRHL2 occupancy sites overlapped with ER occupancy sites (273). GRHL2 is a transcription factor involved in embryonic development and it has been implicated as both a tumor suppressor and an oncogene depending on the cellular context (274,275). There are limited data on its role in breast cancer, but immunohistochemistry analysis of primary breast tumors revealed a significant positive correlation between ER positivity and GRHL2 expression (275). GRHL2 was also found to be an interacting partner with ER

and the pioneer factor FOXA1 using rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) (201,265,276). The latter also showed association with the methyltransferase MLL3, indicating a transcriptional complex with ER, FOXA1, GRHL2, and MLL3 (265). Interestingly, when motif analysis was performed on MLL3 occupancy sites, the GRHL2 motif was found to be significantly enriched. Unlike FOXA1 motifs which are enriched at both ER-only and pS118-ER sites, GRHL2 motifs are enriched only at pS118-ER binding sites raising the possibility that pS118-ER may distinguish ER complexes containing FOXA1, GRHL2, and MLL3 from those including only FOXA1.

Increasing evidence suggests that the GRHL family of proteins (GRHL1, GRHL2, and GRHL3) can act as pioneer factors as can FOXA1 and GATA3 (156,157,197,277). The GRHL DNA binding motif was found to be enriched in regions of open chromatin as assessed by ATAC-seq, and knockdown of all three GRHL transcription factors in MCF-7 cells led to decreases in chromatin accessibility at GRHL target regions (277). A similar study performed in epiblast-like cells found that knockout of GRHL2 led to a decrease in active enhancer marks at GRHL2 occupancy sites as well as an overall decrease in chromatin accessibility at these locations (278). GRHL2 overexpression in embryonic stem cells resulted in an increase of these same marks (278). These findings have also been corroborated in *Drosophila* grainyhead (Grh), where approximately 50% of Grh binding sites were associated with regions of open chromatin in embryos (279). Enrichment of GRHL2 DNA binding motifs was observed in regions nearby upregulated genes in differentiated urothelial cells, suggesting the opening of chromatin by GRHL2 can lead to gene activation (280). We found that pS118-ER was similarly associated with upregulated genes in E2-treated MCF-7 cells. It is intriguing to speculate that GRHL2

could cause an opening of chromatin that could potentially allow the binding of transcription factors, such as pS118-ER, at specific sets of target genes.

The genomic layout of transcription factor binding sites can yield important information regarding potential complexes that may form at estrogen responsive genes. The positional distribution of transcription factor binding sites within nucleosome-depleted regions of enhancers has been recently shown to reflect the functional role of that factor (281). Multiple reports have noted that ER binding sites are enriched for the co-binding of the FOX family of transcription factors as pioneer factors, as well as the motifs of numerous other transcription factors (118,156,159,187,200,261). Notably, we found that the GRHL2 motif occurred most frequently 50 bp from the ERE. Our positional motif analysis suggests that GRHL2 is close enough to be in complex with ER, and future studies will be directed at the temporal and spatial relationship of GRHL2 with pS118-ER.

Previous studies examining the effect of phosphorylation on the ER-ERE interaction in an *in vitro* context using fluorescence anisotropy reported kinase-specific effects on ER-DNA binding affinity (135). Our group further found that a cis-trans prolyl isomerase, Pin1, interacts with ER in a pS118-specific manner, altering the conformation of ER and increasing its affinity for canonical EREs (110). Here we show that estrogen-bound pS118-ER is more likely to be associated with EREs in a native chromatin environment. These findings are in line with a recent report which showed that the most transcriptionally necessary ER binding sites contained strong EREs (282). We would predict that pS118-ER sites, which tend to contain EREs at a greater frequency and have an increased H3K27ac signal are also predictive of ER site transcriptional necessity. Additionally, our finding that numerous ER and pS118-ER binding sites as defined by

ChIP-seq have two or greater ER binding events possible raises the question of how multiple ER binding events effect the regulation of their respective gene targets.

Clinically, primary ER-positive breast cancer tumor samples displaying high levels of pS118-ER positively correlate with a more differentiated phenotype as well as better outcomes in patients receiving adjuvant tamoxifen therapy (202,204,256,283). More recently, the constitutively active ER mutation Y537S, which is predominantly found in metastatic breast cancers, was found to be constitutively phosphorylated at S118 (211,213,215,216). Similarly, the ESR1/YAP1 translocation, which consists of the N-terminal AF-1 and DBD domains of ER fused to the C-terminus of YAP1, can induce growth in the absence of E2, suggesting a potential role for the AF-1 domain in the translocations (218). The AF-1 domain of ER depends on pS118 for activity, though it is not yet known if ER is phosphorylated at S118 in the ESR1/YAP1 translocation.

Because phosphorylation at S118 is associated with constitutively active receptors and is known to control AF-1 activity, it had been proposed that inhibition of the phosphorylation at S118 would be a potential method for reducing the activity of ER mutants. Indeed, recent studies using the CDK7 inhibitor THZ1 demonstrated a decrease in pS118-ER in MCF-7 cells expressing either wt ER or Y537S ER (216,217). Additionally, THZ1 in combination with Fulvestrant was found to inhibit tumor growth in orthotopic xenografts of MCF-7 cells expressing Y573S ER compared to either drug alone (216,217). Our data suggest that inhibition of pS118-ER by THZ1 could lead to decreased occupancy of both wt and mutant ER on ERE containing sites.

In summary, these findings demonstrate that post-translational modification of ER by phosphorylation provides genome-wide specification of ER binding to DNA by

preferentially occupying direct binding sites and associating with specific transcription factors. Given that the majority of nuclear receptors can be phosphorylated (257), understanding how these phosphorylation events contribute to their respective cistromes will be critical for fully understanding their functions and will aid in refining targeted therapies.

Figure 3-1: Location Analysis of pS118-ER Sites

(A) Combination of pS118-ER sites into one dataset. E2 peaks from pS118-ER #2 and pS118-ER #3 ChIP-seq were combined into one peak set and overlapped with ER sites. (B) Location annotation of pS118-ER and ER occupancy sites. Sites were annotated to either a promoter region, intron, intergenic, or other. A promoter region was defined as -1000 bp to +100 bp from a transcription start site. The outer ring displays the distribution of ER sites and the inner ring displays the distribution of pS118-ER sites. (C) ER and pS118-ER sites were annotated to their nearest transcription start site and binned into various distances. (D) Average H3K27ac signal at ER and pS118-ER sites. H3K27ac data in MCF-7 cells treated with E2 was acquired from a previous published dataset (67; GSE45822). (E) Volcano plot of RNA-seq data from MCF-7 cells treated with 1 nM E2 for 24 h acquired from a previous published dataset (68; GSE89888). Fold change represents E2-treated vs vehicle treated. Each dot represents a gene and red dots are genes with a pS118-ER occupancy site within 100 kb of its transcription start site. The horizontal dashed line denotes a cutoff of significance at p=0.001. (F) Proportion of genes upregulated or downregulated by E2 with a pS118-ER site within 100 kb of their respective transcription start sites.

Figure 3-1

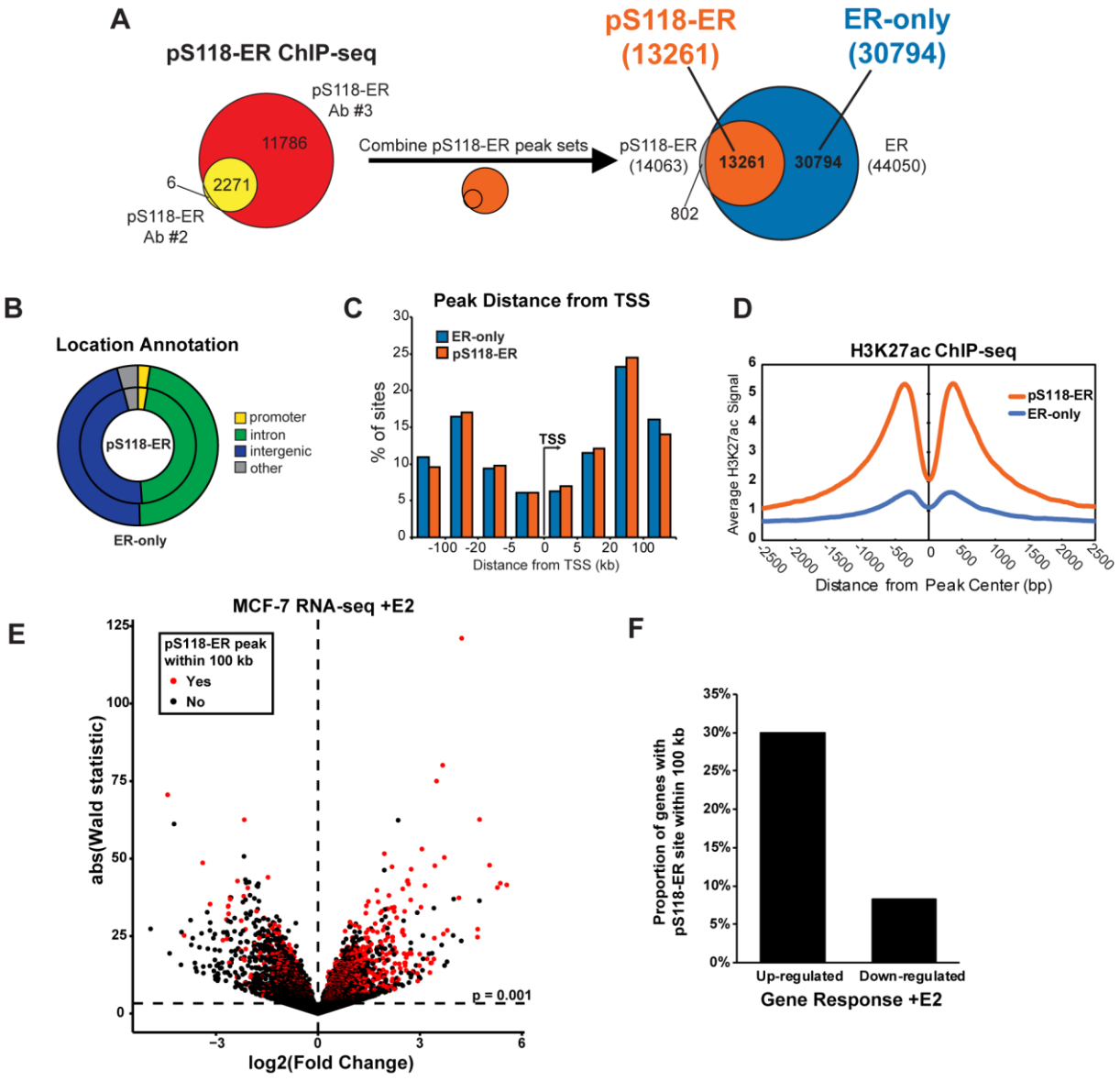

A
pS118-ER ChIP-seq
pS118-ER Ab #3
11786
6
2271
pS118-ER Ab #2
Combine pS118-ER peak sets
pS118-ER (13261)
ER-only (30794)
pS118-ER (14063)
13261
30794
ER (44050)
802
B
Location Annotation
pS118-ER
promoter
intron
intergenic
other
ER-only
C
Peak Distance from TSS
ER-only
pS118-ER
% of sites
TSS
Distance from TSS (kb)
D
H3K27ac ChIP-seq
pS118-ER
ER-only
Average H3K27ac Signal
Distance from Peak Center (bp)
E
MCF-7 RNA-seq +E2
pS118-ER peak within 100 kb
Yes
No
abs(Wald statistic)
p = 0.001
log2(Fold Change)
F
Proportion of genes with pS118-ER site within 100 kb
Up-regulated
Down-regulated
Gene Response +E2

Figure 3-2: Motif analysis of ER and pS118-ER sites.

(A) *de novo* motif analysis of ER sites and pS118-ER sites using HOMER. TF, associated transcription factor for each motif. (B) Differential motif analysis of pS118-ER sites. Enriched motifs were searched for using the CentriMo tool in the MEME suite (184). All ER sites (44,050) were used as the control (background) sequences and pS118-ER sites were used as the primary input sequences. (C) Proportion of sites containing each specific motif from ‘known’ motif analysis using HOMER. p-values from chi-squared test are displayed when significant. (D) Distribution of specific motifs around ER and pS118-ER sites.

Figure 3-2

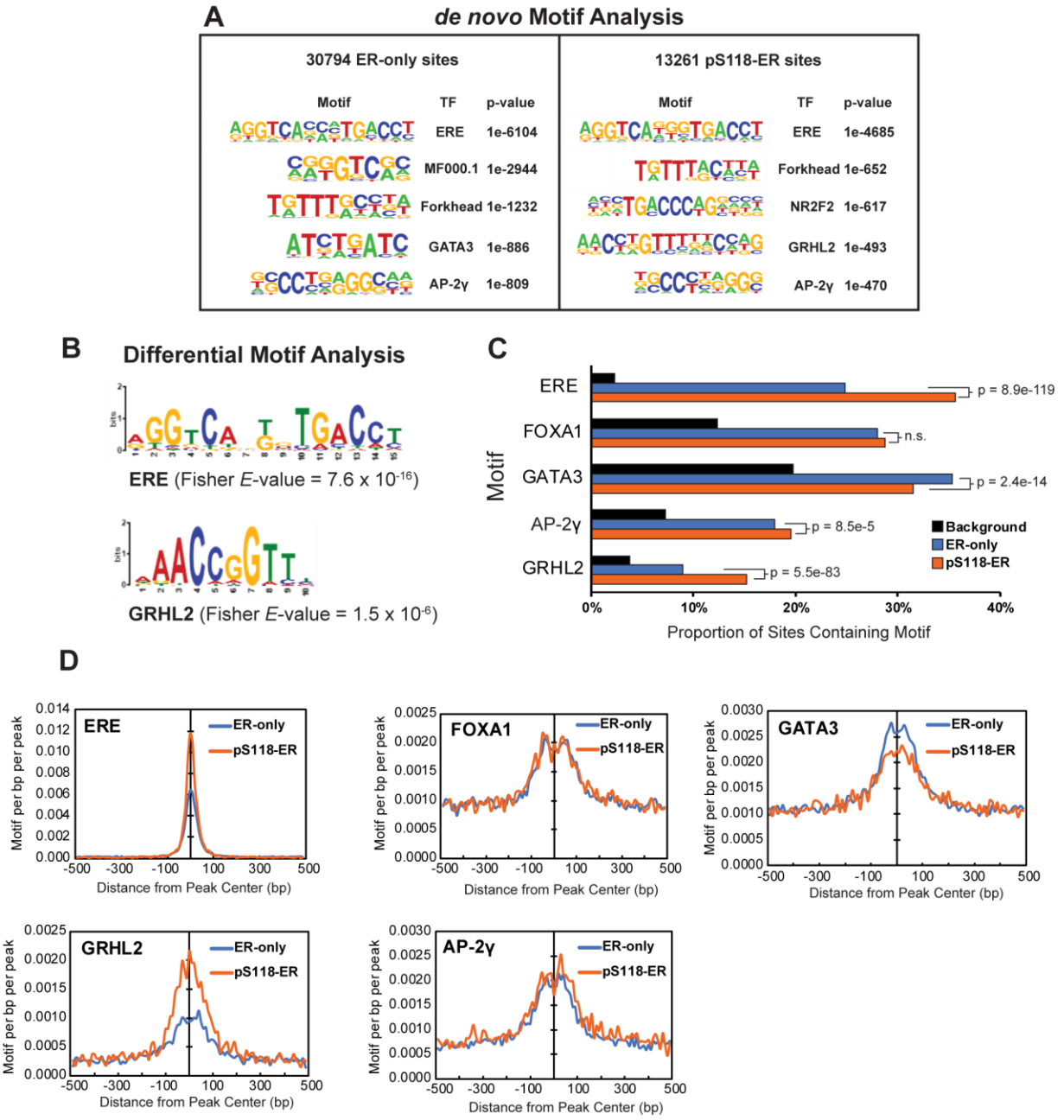

A
de novo Motif Analysis
30794 ER-only sites
Motif
TF
p-value
ERE 1e-6104
MF000.1 1e-2944
Forkhead 1e-1232
GATA3 1e-886
AP-2γ 1e-809
13261 pS118-ER sites
Motif
TF
p-value
ERE 1e-4685
Forkhead 1e-652
NR2F2 1e-617
GRHL2 1e-493
AP-2γ 1e-470
B
Differential Motif Analysis
ERE (Fisher E-value = 7.6 x 10-16)
GRHL2 (Fisher E-value = 1.5 x 10-6)
C
Motif
ERE
FOXA1
GATA3
AP-2γ
GRHL2
p = 8.9e-119
n.s.
p = 2.4e-14
p = 8.5e-5
p = 5.5e-83
Background
ER-only
pS118-ER
0%
10%
20%
30%
40%
Proportion of Sites Containing Motif
D
ERE
FOXA1
GATA3
GRHL2
AP-2γ
ER-only
pS118-ER
Motif per bp per peak
Distance from Peak Center (bp)

Figure 3-3: Analysis of GRHL2 Occupancy at pS118-ER Sites.

(A) Overlap of GRHL2 sites with pS118-ER sites and ER sites identified by ChIP-seq. (B) Average GRHL2 signal at ER and pS118-ER sites. GRHL2 ChIP-seq data was acquired from a previously published dataset (265). (C) ChIP-qPCR analysis of ER, pS118-ER and GRHL2 after 30 min treatment with E2 at sites co-occupied by pS118-ER and GRHL2 (n = 4) or (D) occupied by GRHL2 only. IgG is used as a control. n = 3, mean ± S.E. is shown. *$p<0.05$

Figure 3-3

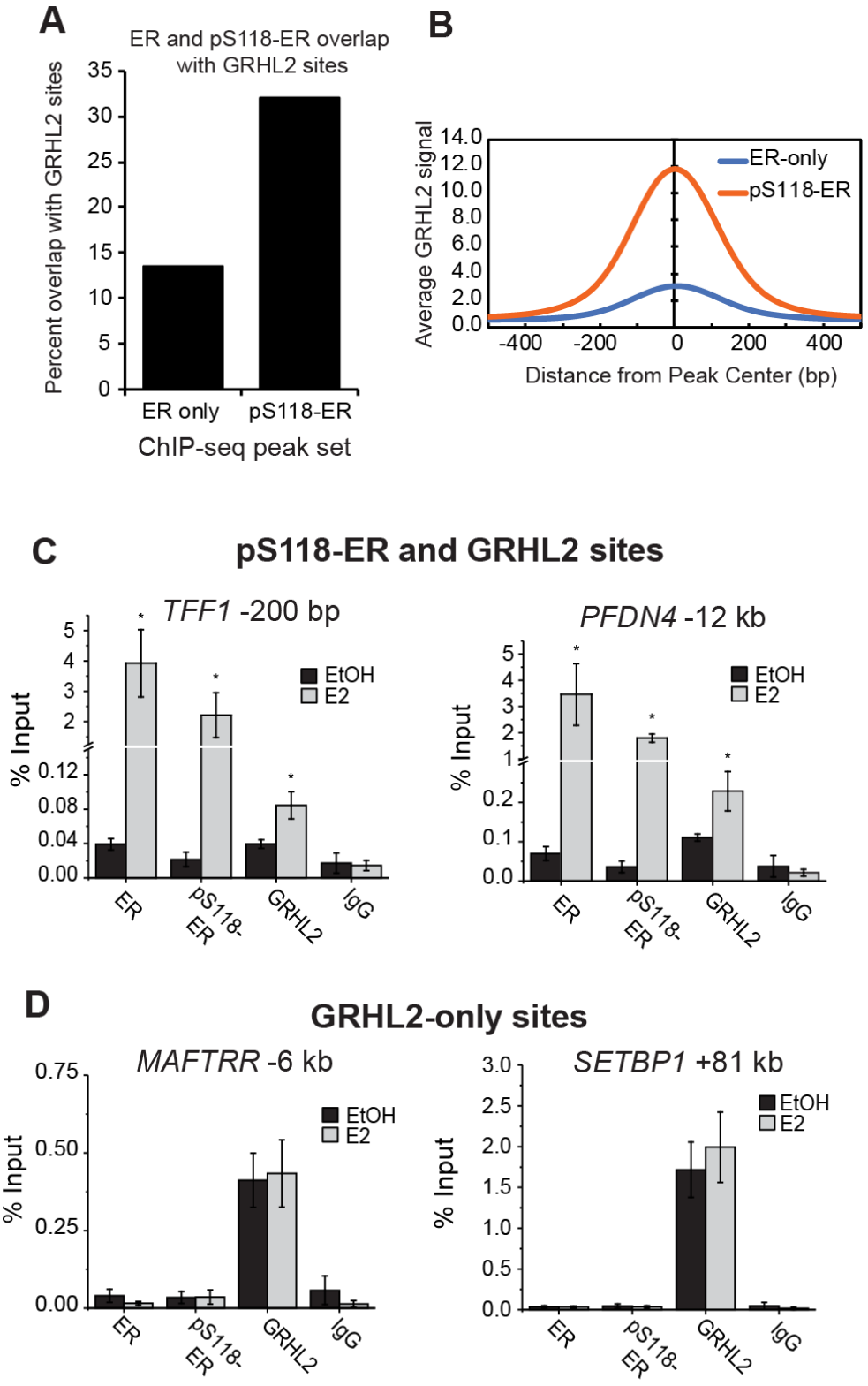

A
ER and pS118-ER overlap with GRHL2 sites
Percent overlap with GRHL2 sites
35
30
25
20
15
10
5
0
ER only
pS118-ER
ChIP-seq peak set
B
Average GRHL2 signal
14.0
12.0
10.0
8.0
6.0
4.0
2.0
0.0
-400
-200
0
200
400
Distance from Peak Center (bp)
ER-only
pS118-ER
C
pS118-ER and GRHL2 sites
TFF1 -200 bp
PFDN4 -12 kb
% Input
EtOH
E2
ER
pS118-ER
GRHL2
IgG
D
GRHL2-only sites
MAFTRR -6 kb
SETBP1 +81 kb

Figure 3-4: Design of SNAP array.

(A) Purified, estrogen-occupied ER was incubated with the SNAP microarray which displays 359,156 different genomic 25 bp DNA sequences configured as double stranded 53–mer hairpins with a 3 base loop. Specific ER binding was detected by fluorescence. (B) Phosphorylation status of purified ER. (Left) 1 pmol of purified ER was analyzed by immunoblot and probed for phosphorylation at S104/S106, S118, S167, and S294 using phospho-specific antibodies. Total ER is shown as a loading control. (Right) Immunoblot demonstrating the proportion of recombinant purified ER (rER) that is phosphorylated. The higher molecular weight shift of the majority of the recombinant purified ER is similar to the shift seen in ER in E2-treated MCF-7 cells. (C) The SNAP DNA probes tiled 5020 genomic regions (~2 Mb of the genome) compiled from 58 published ChIP-seq data sets from both cell lines and tumors, at 6 bp resolution. Schematic shows that direct ER binding events at specific DNA sites across a tiled genomic region are indicated by differing intensity values. Representative probes tiling across a section of the ChIP-seq peak at the *GREB1* promoter are shown as an example. The 6 bp tiling resolution of the DNA probes allows for identification of the likely direct ER binding site (blue underline) indicated by the SNAP intensity values (in red).

Figure 3-4

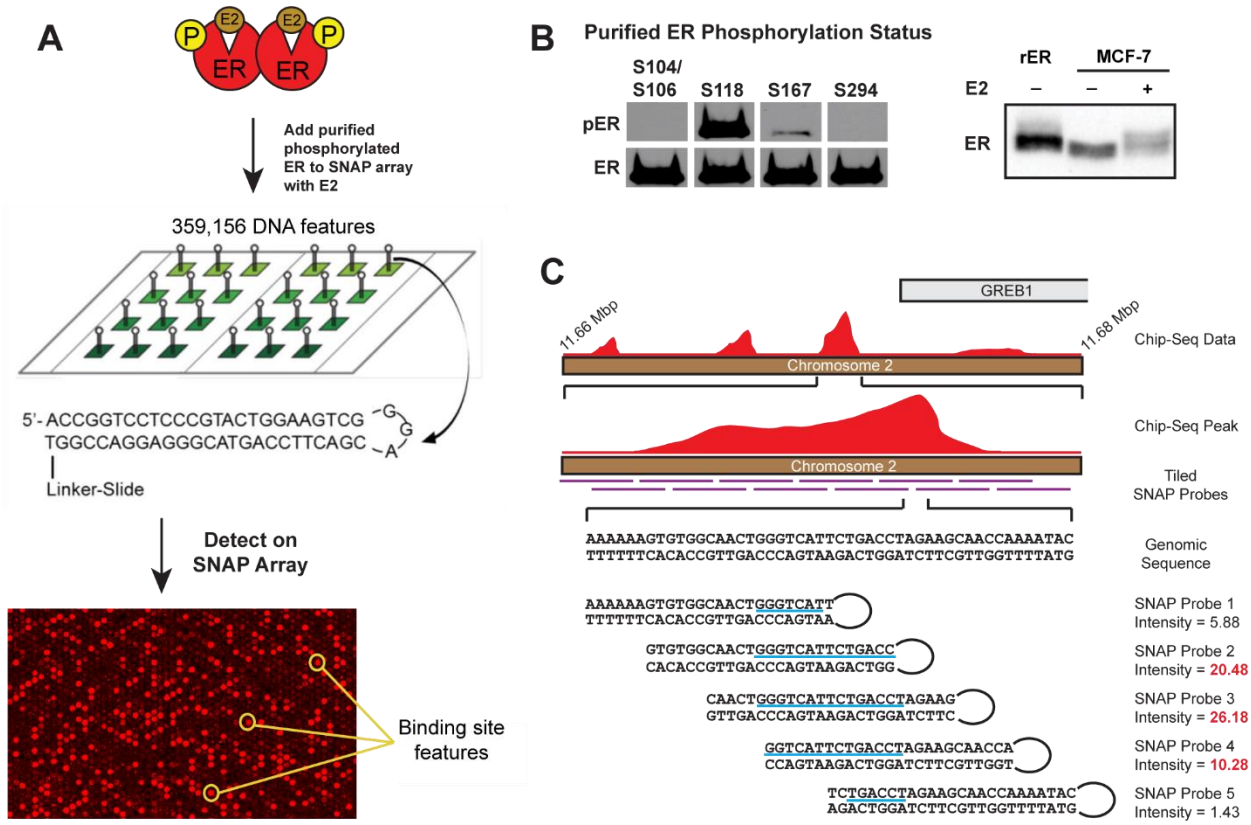

A
P
E2
ER
ER
E2
P
Add purified phosphorylated ER to SNAP array with E2
359,156 DNA features
5'- ACCGGTCCTCCCGTACTGGAAGTCG
TGGCCAGGAGGGCATGACCTTCAGC
G
G
A
Linker-Slide
Detect on SNAP Array
Binding site features
B
Purified ER Phosphorylation Status
S104/ S106
S118
S167
S294
pER
ER
rER
MCF-7
E2
–
–
+
ER
C
GREB1
11.66 Mbp
11.68 Mbp
Chromosome 2
Chip-Seq Data
Chip-Seq Peak
Chromosome 2
Tiled SNAP Probes
AAAAAAGTGTGGCAACTGGGTCATTCTGACCTAGAAGCAACCAAAATAC
TTTTTTCACACCGTTGACCCAGTAAGACTGGATCTTCGTTGGTTTTATG
Genomic Sequence
AAAAAAGTGTGGCAACTGGGTCATT
TTTTTTCACACCGTTGACCCAGTAA
SNAP Probe 1 Intensity = 5.88
GTGTGGCAACTGGGTCATTCTGACC
CACACCGTTGACCCAGTAAGACTGG
SNAP Probe 2 Intensity = 20.48
CAACTGGGTCATTCTGACCTAGAAG
GTTGACCCAGTAAGACTGGATCTTC
SNAP Probe 3 Intensity = 26.18
GGTCATTCTGACCTAGAAGCAACCA
CCAGTAAGACTGGATCTTCGTTGGT
SNAP Probe 4 Intensity = 10.28
TCTGACCTAGAAGCAACCAAAATAC
AGACTGGATCTTCGTTGGTTTTATG
SNAP Probe 5 Intensity = 1.43

Figure 3-5: SNAP array analysis reveals direct and indirect ER binding events

(A) Schematic for identifying direct and indirect binding events. If a site shows binding in both ChIP-seq and the SNAP array, it is categorized as a direct binding event. If a peak identified in ChIP-seq is not detected on the SNAP array, it is categorized as an indirect binding event. (B) Representative examples of a direct and indirect binding event. (Top) ChIP-seq signal from ER and pS118-ER at the *TFF1* promoter and a site within the *IGF1R* gene. Color key and y-axis is the same as in Fig 2C. (Bottom) Relative ER binding intensity at these regions on the SNAP array. Each region contains multiple DNA probes tiled across the region. (C) Breakdown of regions assayed on the SNAP array. Pie charts display the proportion of regions assayed on the SNAP array and present in either the pS118-ER or ER-only ChIP-seq sites. Analysis revealed that directly bound sites were more likely to be occupied by pS118-ER rather than ER. (D) Distribution of the number of ER binding events present in the SNAP array genomic regions assayed. The number of ER binding events was calculated for each genomic region and is displayed as a histogram. A relative ER binding intensity of 5 was used as a cutoff for a binding event. (E) Proportion of SNAP array regions present in ER or pS118-ER ChIP-seq sites. SNAP array regions were binned by the number of ER binding events and these regions were overlapped with ER and pS118-ER sites.

Figure 3-5

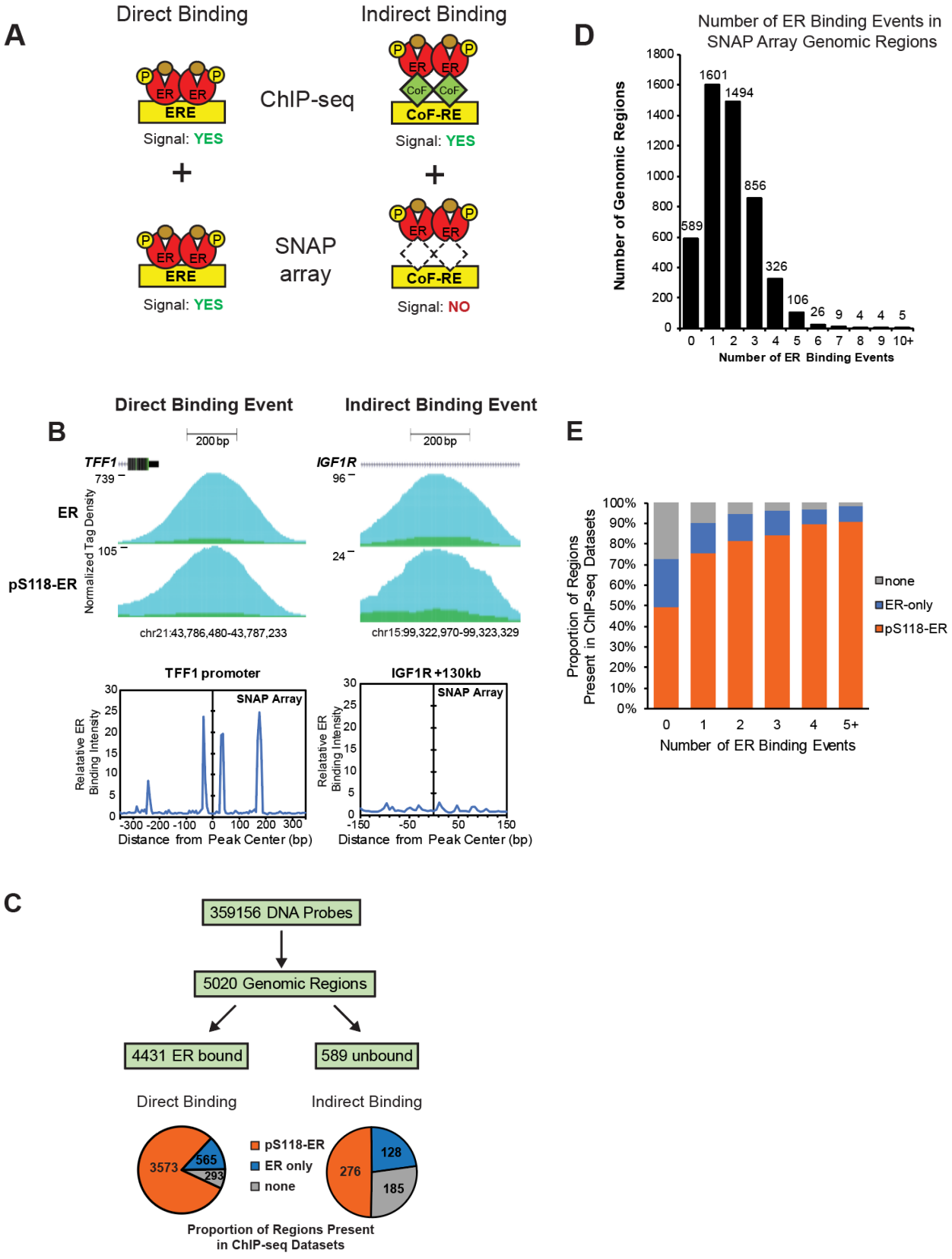
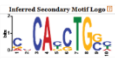

A
Direct Binding
Indirect Binding
ChIP-seq
SNAP array
ERE
CoF-RE
Signal: YES
Signal: NO
B
Direct Binding Event
Indirect Binding Event
200 bp
TFF1
IGF1R
ER
pS118-ER
Normalized Tag Density
chr21:43,786,480-43,787,233
chr15:99,322,970-99,323,329
TFF1 promoter
IGF1R +130kb
SNAP Array
Relatative ER Binding Intensity
Distance from Peak Center (bp)
C
359156 DNA Probes
5020 Genomic Regions
4431 ER bound
589 unbound
Direct Binding
Indirect Binding
3573
565
293
276
128
185
pS118-ER
ER only
none
Proportion of Regions Present in ChIP-seq Datasets
D
Number of ER Binding Events in SNAP Array Genomic Regions
Number of Genomic Regions
Number of ER Binding Events
589
1601
1494
856
326
106
26
9
4
4
5
E
Proportion of Regions Present in ChIP-seq Datasets
Number of ER Binding Events
none
ER-only
pS118-ER

Figure 3-6: Neighboring motif analysis on SNAP array regions.

Spaced motif analysis using the MEME suite SpaMo tool (267) identifies the neighboring motifs (“inferred secondary motif logo”) around ER binding sites binned by number of potential ER binding events on the SNAP array. Displayed are the top motif found followed by motifs neighboring the top motif. Listed under each adjacent motif is the number of peaks with the respective motif nearby and the most frequent gap between the ERE and the motif in bps.

Figure 3-6

ER Peak Count	Top Motif	Neighboring Secondary Motifs
0 peaks (589 regions)	 586 regions - ZNF263, SP2	
1 peak (1601 regions)	 1601 regions - ER	POU5F1  peaks: 455 gap: 70 MAX  peaks: 128 gap: 7 NR2F1  peaks: 126 gap: 6 VDR  peaks: 423 gap: 8 MECOM  peaks: 57 gap: 27 GRHL2  peaks: 337 gap: 50
2 peaks (1494 regions)	 1494 regions - ER	P53  peaks: 76 gap: 38 TFAP2A  peaks: 1269 gap: 30 HNF1A  peaks: 95 gap: 79
3 peaks (856 regions)	 856 regions - ER	FosL1:JUND  peaks: 456 gap: 29 CREB3  peaks: 17 gap: 10 TFAP2A(var.2)  peaks: 586 gap: 35 NFIC  peaks: 462 gap: 1 EVX1  peaks: 203 gap: 55 TFDP1  peaks: 132 gap: 45
4 peaks (326 regions)	 326 regions - ER	BHLHE22  peaks: 19 gap: 28 POU6F2  peaks: 44 gap: 78 HEY1  peaks: 14 gap: 19 HIF1A  peaks: 28 gap: 20 TCF3  peaks: 95 gap: 46 MECOM  peaks: 10 gap: 63
5 peaks (106 regions)	 106 regions - ER	FOXL1:JUND  peaks: 66 gap: 60 MAX  peaks: 12 gap: 50
6-35 peaks (48 regions)	 48 regions - ER	

CHAPTER FOUR:

Development of an S118A ER knock-in MCF-7 model and its effects on ER stability

Abstract

Estrogen receptor-α (ER) is a key driver of breast cancer growth and development and has been a major target for treatment of the disease. ER function is heavily controlled through post-translation modifications such as phosphorylation and ubiquitylation. Phosphorylation of ER at S118 (pS118-ER) is known to regulate ER transcriptional function as well as ER protein stability. Most of the studies analyzing the function of pS118-ER through mutational analysis of serine 118 to alanine utilized expression of this mutant in ER-negative cells to avoid interactions with the endogenously expressed wt ER as would occur in ER-positive cell lines. While these studies have been vital for investigating the function of pS118-ER, the cellular contents are variable between cell lines and it is not known if the phenotypes observed are specific to the cell line or if they translate to a native ER-positive cell environment. Recent advances in CRISPR/Cas9 technology have made single base pair knock-in mutations more efficient in cell lines and have allowed for the interrogation of mutations in their endogenous environments. Here, we describe the generation of an S118A ER knock-in MCF-7 model and examine its effect on E2-dependent gene regulation and ER stability. Consistent with previous reports, we find that E2-dependent ER downregulation is impaired in the S118A model, indicating that ER degradation is regulated by pS118-ER in MCF-7 cells. Interestingly, we found that downregulation of the *ESR1* gene by E2 was also impaired in the S118A model, suggesting that control of ER levels by pS118-ER is not limited to post-translational mechanisms. These results confirm the role of pS118-ER in ER regulation and provide a tool for investigating the function of pS118-ER in an ER-positive cellular environment.

Introduction

Estrogen Receptor-alpha (ER) is a steroid nuclear receptor responsible for the growth and maintenance of estrogen-dependent tissues such as the breast. ER is the major driver of many breast cancers, with approximately 70% of primary breast cancers expressing ER (4). Upon stimulation with 17ß-estradiol (E2), ER homodimerizes, binds to chromatin at estrogen response elements, and recruits coactivators or corepressors to effect target gene expression (89). One of the target genes repressed by ER is its own gene, *ESR1* (104,105). The level of ER in the cell is highly regulated, with an E2-dependent feedback loop causing a decrease in *ESR1* synthesis, mRNA levels, ER protein levels, and an increase in ER ubiquitylation (78,99,103,107,108,110). This negative feedback loop relies heavily on the phosphorylation of ER at serine 118 (pS118-ER). Phosphorylation at this site occurs in response to E2, as well as many ER agonists and antagonists such as tamoxifen, ICI-182780, and epidermal growth factor (EGF) (52,57). Recent studies found that the interaction of ER with the E3 ubiquitin ligase E6-interacting protein (E6-AP) is dependent on pS118-ER (109,284). Additionally, the peptidyl-prolyl cis-trans isomerase NIMA-interacting 1 (Pin1) interacts with ER in a pS118-ER-dependent manner. When phosphorylated, the peptide bond between pS118 and P119 on ER can be isomerized by Pin1, which in turn disrupts the interaction with E6-AP, leading to stabilization of ER (109). These data are supported by computational modeling of ER dynamics which found that pS118-ER is the key axis around which *ESR1* and subsequent ER protein levels are regulated (107). At the transcriptional level, pS118-ER is required for maximal expression of various ER target genes in response to E2, whereas other target genes were found to be independent of pS118-ER (79). Similarly,

mutation of serine 118 to alanine results in a reduction of E2-stimulated ER occupancy on DNA as measured by chromatin immunoprecipitation (ChIP). Notably, the S118A mutant is still able to bind DNA upon E2 treatment indicating that pS118-ER is not required for DNA binding (80,238). However, a major caveat of these S118A mutational analyses is the performance of the experiments in ER-negative cell lines, usually HEK293, MDA-MB-231, or HeLa. These cell lines do not require ER for growth and survival, whereas the ER-positive MCF-7 cells are highly dependent on ER for survival (285). The introduction of ER into a cellular environment which does not require ER could affect the interpretation of observations seen in the S118A mutants. Development of an MCF-7 cell line expressing the S118A ER mutation while lacking the wt ER is therefore critical for determining the role of pS118-ER in breast cancer cells that rely on ER for survival and will allow for a more direct comparison to be made for the effects of phosphorylation on ER biology.

Within the past five years, genome editing using Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) along with the endonuclease CRISPR associated protein 9 (Cas9) has allowed for highly specific mutations to be made in an organism's genome. While CRISPR/Cas9 technology is useful for creating knockout cell lines, generation of a knock-in cell line, especially a homozygous mutation which requires the introduction of a specific mutation into two alleles, has been less effective. Recent advances in CRISPR/Cas9 technology including the optimization of sgRNA design and the utilization of ribonucleoprotein (RNP) delivery via electroporation have made it more feasible to develop homozygous mutations in cell lines (286). Utilizing these recent advances, we developed two clones of the MCF-7 cell line expressing the S118A ER

mutation. Using these clones, we show that the S118A mutation partially prevents the downregulation of ER protein upon E2 treatment, consistent with previous reports. Interestingly, we additionally found that the S118A mutation is not capable of downregulating the *ESR1* gene upon E2 treatment, indicating an essential role for pS118-ER in the coordinated E2-induced ER downregulation response at multiple levels.

Results

Development of an S118A ER knock-in MCF-7 cell line

In order to investigate the role of pS118-ER in MCF-7 cells, we designed a CRISPR/Cas9 strategy using homology directed repair (HDR) to knock-in the S118A ER mutation (Figure 4-1). The codon for S118 lies within exon 2 of *ESR1* and a short guide RNA (sgRNA) was designed to direct Cas9 cleavage as close as possible to the desired mutation. Recent advances in CRISPR/Cas9 technology have found a strong inverse correlation between (a) the distance in base pairs between the Cas9 cleavage site to the desired mutation and (b) the successful incorporation of the mutation through HDR (286). Additionally, the introduction of silent mutations which ablate either the PAM sequence or the sequence to which the sgRNA anneals increases HDR efficiency by preventing additional editing after the initial edit (286). The successful introduction of these silent mutations along with the desired mutation creates a Cas9-resistant mutant for the specific sgRNA designed. Using this strategy, we chose a sgRNA that directs Cas9 to cut 4 bp away from the desired mutation site (Figure 4-1A). To mutate the S118 to alanine, a single bp change of T to G was introduced to change the serine codon (TCG) to alanine (GCG). Mutation of the PAM sequence (NGG) at the 3' end of the sgRNA was not a feasible strategy because mutation of either guanidine would cause a missense mutation in proline 115. Therefore, we designed the 120 bp HDR donor sequence to contain three silent mutations in addition to the desired S118A mutation which would decrease the affinity of the sgRNA for its analogous genomic annealing site (Figure 4-1A, 4-1B). The introduction of the S118A mutation also introduces a novel SfoI restriction digestion site such that the wt *ESR1* sequence (GTCGCC) cannot be digested by SfoI, whereas the

mutated S118A *ESR1* can (GGCGCC) (Figure 4-1C). Using the designed sgRNA and the HDR donor, we electroporated $5.0x10^5$ MCF-7 cells with a ribonucleoprotein complex consisting of purified Cas9 with the sgRNA and the HDR donor sequence containing the desired mutations. As a control, MCF-7 cells were separately electroporated with purified Cas9 and the HDR donor sequence, but without a sgRNA to direct Cas9.

Prior to single cell selection, the transfected populations were screened for the presence of the S118A mutation. Primers flanking the mutation site were used to amplify the region and the amplicon was subjected to SfoI digestion. Upon SfoI digestion, the cell population transfected with the sgRNA showed the presence of the inserted mutation whereas the control cells did not (Figure 4-2A). Additionally, the presence of the mutation was preserved through cell passages, indicating that negative selection was not occurring to the S118A ER mutants (Figure 4-2B). The cells transfected with sgRNA were then sparsely plated in a 10 cm plate at 100 to 1000 cells per plate to isolate single cell clones. Upon screening, two clones were identified which displayed complete digestion by SfoI, indicating a homozygous clone (Figure 4-2C). Sequencing confirmed the presence of the S118A mutation in both clones, with one clone incorporating all four of the mutations from the HDR donor, and the other incorporating three of the four mutations (Figure 4-2D). These two clones will be referred to as S118A cl.1 and S118A cl.2, respectively.

To confirm that the S118A clones expressed an ER lacking the ability to be phosphorylated at S118, we subjected both clones to an E2 treatment and probed for pS118-ER via immunoblot. Cells were estrogen-deprived for three days prior to a 30-min E2 treatment, which has been shown to induce the highest levels of pS118-ER in MCF-7 cells (136,238). As expected, both S118A clones were not phosphorylated at S118 in

response to E2 whereas the parental MCF-7 cells did display pS118-ER (Figure 4-2E). Importantly, both S118A clones expressed ER at the correct size suggesting that no other major alterations were made to ER during the editing process. These results show that two MCF-7 S118A knock-in clones were generated, each with no detectable wt ER present by sequencing or by immunoblot for pS118-ER.

S118A Mutation Prevents E2-Dependent ER Downregulation

Previous investigations have shown that ER protein is downregulated upon E2 treatment through the ubiquitin-proteasome pathway and that this response is dependent on pS118-ER (78,99–101). The studies identifying pS118-ER as a mediator of E2-dependent ER downregulation were performed by expressing S118A ER in HEK293 cells, an ER-negative cell line, which could affect the interpretation of the results in ER-positive cell lines. To test if pS118-ER similarly plays a role in downregulation of ER by E2 in MCF-7 cells, we treated the two MCF-7 S118A ER clones with either vehicle (0.1% EtOH) or 10 nM E2 for 24 h and measured protein expression by immunoblot. We also treated the S118A ER clones with 100 nM Tam and 100 nM ICI-182780 for 24 h to assess the effectiveness of ER antagonists on the stability of S118A ER. Similar to previous studies, the downregulation of the S118A ER mutants by E2 was attenuated indicating that the mechanisms governing pS118-ER-dependent downregulation of ER by E2 are consistent between MCF-7 cells and the studies performed in ER-negative HEK293 cells (Figure 4-3). Stabilization of ER was observed in the S118A ER clones upon treatment with 4'-hydroxytamoxifen, consistent with previous reports (287–289). Treatment of both mutant S118A ER clones with ICI-182780 lead to degradation of the receptor at similar levels as the wt parental MCF-7 cells indicating that the degradation of ER by ICI-182780 is

independent of the ER phosphorylation status at S118 (Figure 4-3). These results demonstrate that pS118-ER is indeed an important modification for E2-dependent ER downregulation in MCF-7 cells and corroborate the data observed in HEK293 cells expressing S118A ER (78).

S118A Mutation Prevents E2-Depedent ESR1 Downregulation

In addition to the post-translational downregulation of ER protein by E2, the gene encoding for ER, *ESR1*, is also downregulated upon E2 treatment (104–106). To determine if *ESR1* transcriptional regulation by E2 is affected by pS118-ER, we treated the two MCF-7 S118A ER clones with 10 nM E2, 100 nM tamoxifen, or both for 24 h and analyzed mRNA levels by qRT-PCR. Interestingly, while we observed a downregulation of *ESR1* mRNA in response to E2 in the wt parental MCF-7 cells as expected, both MCF-7 S118A ER clones failed to downregulate *ESR1* at a significant level upon E2 treatment (Figure 4-4, *ESR1*). Previous studies have also shown that pS118-ER plays a role in the expression of various ER target genes, however these studies were performed in the ER-negative cervical cancer cell line HeLa rather than an ER-positive breast cancer cell line (79). We analyzed the expression of a number of ER target genes in the MCF-7 S118A mutant clones in response to E2 and found that while the S118A mutation did not abolish the ability of E2 to induce *TFF1* and *GREB1*, the fold change over vehicle control observed in both S118A clones was significantly decreased compared to the parental wt MCF-7 cells indicating dependence on pS118-ER for maximal induction by E2 (Figure 4-4). Another ER target gene, *PGR*, showed partial dependence on pS118-ER for maximal induction, as one clone (MCF-7 S118A cl.1) showed no effect on E2 induction compared to the wt parental MCF-7 cells and the other clone (MCF-7 S118A cl.2) showed a

reduction in its E2 activation (Figure 4-4). Analysis of *GATA3* found that the E2-induced downregulation observed in the wt parental cells was abrogated in one clone (S118A cl.2) but still present in the other (S118A cl.1) (Figure 4-4, *GATA3*). Treatment with the ER antagonist tamoxifen in combination with E2 reduced the E2 gene induction of *TFF1* and *GREB1* at similar levels in the wt parental and S118A clones (Figure 4-4). E2 induction of *PGR* was not affected by tamoxifen in the wt parental cells whereas expression was attenuated in the S118A clones (Figure 4-4). These studies demonstrate that E2-dependent ER downregulation in MCF-7 cells is affected by the S118A ER mutation at both the protein and mRNA levels.

Discussion

Control of cellular ER level is known to be a highly regulated process and alterations from this regulation can lead to aberrant gene expression and growth (98,290). Past studies show that overexpression of ER alone is sufficient to induce ER target gene activation as well as growth independent of ligand (98). ER levels are controlled at multiple stages including mRNA production (104), protein stability through phosphorylation (78,109) and ubiquitylation (99,100). Previous studies demonstrate that pS118-ER is required for E2-induced ER downregulation (78) along with maximal ER-target gene activation (52,69,79). More recently, computational analysis of ER dynamics found that pS118-ER was a key component for regulating *ESR1* and subsequently ER protein levels (107). While these and other past studies have greatly contributed to the current understanding of pS118-ER function, these studies frequently analyzed the effect of pS118-ER through introduction of the S118A ER mutant in ER-negative cells lines which contain a different cellular environment compared to the ER-positive MCF-7 cell line. Given the dependence of MCF-7 cells on ER, analyzing the impact of the S118A ER mutation in the context of an MCF-7 background is important for determining its function. Here, we report the development of an S118A ER MCF-7 cell line and show that the mutation abrogates E2-dependent ER downregulation at both the protein and mRNA levels.

Upon treatment with E2, ER is ubiquitylated and degraded by the 26S proteasome (99–101). E2-dependent ubiquitylation occurs near the beginning of the AF-2 region on residues K302/K303 (291) and additional deletion analysis of ER found that removal of the AF-1 region prevents E2-induced ER downregulation even though the ubiquitylation

sites are still present (78). Subsequent mutational analysis of the phosphorylation sites in the AF-1 region identified S118 as the key residue responsible for ER protein degradation in response to E2 (78). The S118A mutation prevented E2-induced ER downregulation whereas mutations of the other AF-1 phosphorylation sites S104/106A or S167A did not. These studies were all performed in HEK293 cells, an ER-negative human embryonic kidney cell line commonly used in cellular biology. Even within breast cancer cell lines, choosing the correct cell line to use for each experiment is critical and can alter interpretations and conclusions (292), therefore testing these results in an ER-positive cell line is important to understanding the function of pS118-ER in a native ER environment. Using the S118A ER MCF-7 cell line generated in this study, we confirm the observation from Valley *et al.* which demonstrates that the S118A mutation prevents E2-induced ER downregulation (78). We similarly found that degradation of ER by ICI 182,780 was not affected by the S118A mutation, consistent with previous reports (78). The corroborating results observed in both HEK293 cells expressing ER constructs and the S118A ER MCF-7 cell line suggest that the machinery necessary to degrade ER in response to E2 is present in both cell lines. The E3 ligase E6-AP is known to serve both as a coactivator for ER as well as a ubiquitin-ligase responsible for ubiquitylating ER in response to E2 (252,293). Additionally, E6-AP associates with ER in a pS118-ER dependent manner (109) to promote ubiquitylation and degradation of ER. E6-AP is expressed at detectable levels in both HEK293 and MCF-7 cells (294–296) indicating that this mechanism may be responsible for the pS118-ER dependent degradation seen in both cell lines. Further analysis will need to be performed to confirm these hypotheses.

We also observed an increase in ER levels in response to tamoxifen in both wt MCF-7 and the S118A ER mutant MCF-7. Previous studies have reported stabilization of ER by tamoxifen (287–289) and here we report that this stabilization is not affected by phosphorylation of ER at S118. Interestingly, treatment of MCF-7 cells with tamoxifen has been shown to inhibit expression of E6-AP potentially through autoubiquitylation (295). Given that E6-AP ubiquitylates and degrades ER (252,293), reduction of E6-AP levels and therefore decreased ER ubiquitylation and degradation could be a potential mechanism for the stabilization of ER by tamoxifen.

In addition to abrogated ER protein downregulation in the S118A ER mutant, we observed the loss of E2-dependent *ESR1* transcriptional downregulation normally seen in wt *ESR1* (Figure 4-4). In the HEK293 model, Valley *et al.* showed no change in *ESR1* levels upon E2 treatment when wt ER was transfected into HEK293 (78). This discrepancy is likely due to the control of *ESR1* gene by a plasmid construct. While it is unclear if HEK293 cells express endogenous ER when exogenous ER is introduced, the mRNA analysis of this system is confounding since the primers used to detect *ESR1* cannot differentiate between the mRNA produced endogenously and that produced by the introduced plasmid construct. Control of *ESR1* expression via a plasmid does not contain the endogenous machinery in the *ESR1* promoter region or the additional intronic regions, therefore E2 cannot control expression of *ESR1* from a plasmid if these additional components are required. In the system generated here, elements of the surrounding *ESR1* locus are still intact and provide a more accurate depiction of *ESR1* control through stimuli such as E2. In the S118A MCF-7 model, *ESR1* downregulation is impaired indicating that pS118-ER plays an important role in ER autoregulation at the

transcriptional level. ER is known to bind to the *ESR1* promoter as well as enhancer sites around the *ESR1* loci (106). Additionally, the ER AF-1 region is a major hub of interaction for cofactors such as the p160 family of coactivators as well as p300 (77,88,92–95). Phosphorylation of ER at S118 is also known to regulate the interaction with these cofactors (77), suggesting that the control of the *ESR1* gene may also be dependent on the interaction of ER with these cofactors on the *ESR1* promoter.

Many studies have been performed to elucidate the role of pS118-ER by introducing an S118A ER mutational construct into ER-negative cell lines. Through these studies, pS118-ER has been shown to regulate ER transactivation (52,69,79), ER protein levels (78), and coactivator interactions (77). Other strategies have been adopted to investigate the effects of the S118A ER mutation in MCF-7 such as combining shRNA knockdown of ER with transfection of ER mutants with limited success (81). The S118A ER MCF-7 model developed here will serve as an important tool for investigating the effect of pS118-ER on breast cancer biology by providing a more accurate and relevant ER environment. Additionally, given the numerous sites which are post-translationally modified on ER (41), this strategy can be expanded to analyze the functions of these sites in the context of an MCF-7 cellular background.

Figure 4-1: Schematic of S118A Mutational Strategy.

(A) Specific mutations were selected for insertion of the S118A mutation in the *ESR1* gene. The wild-type codon coding for a serine (TCG) was mutated by a single bp change to code for an alanine (GCG) shown in red. Additional mutations shown in blue were silent and prevented the sgRNA from binding once the mutations were inserted. A novel SfoI restriction site (underlined) was introduced in the mutation for screening by PCR. (B) Schematic showing the inability of the Cas9/sgRNA ribonucleoprotein complex to anneal to the mutated sequence. (C) The wt amplicon could not be digested by SfoI and produced a 615 bp fragment. Insertion of the mutation introduced a novel SfoI site which allowed the amplicon to be cut into two fragments of size 152 bp and 463 bp.

Figure 4-1

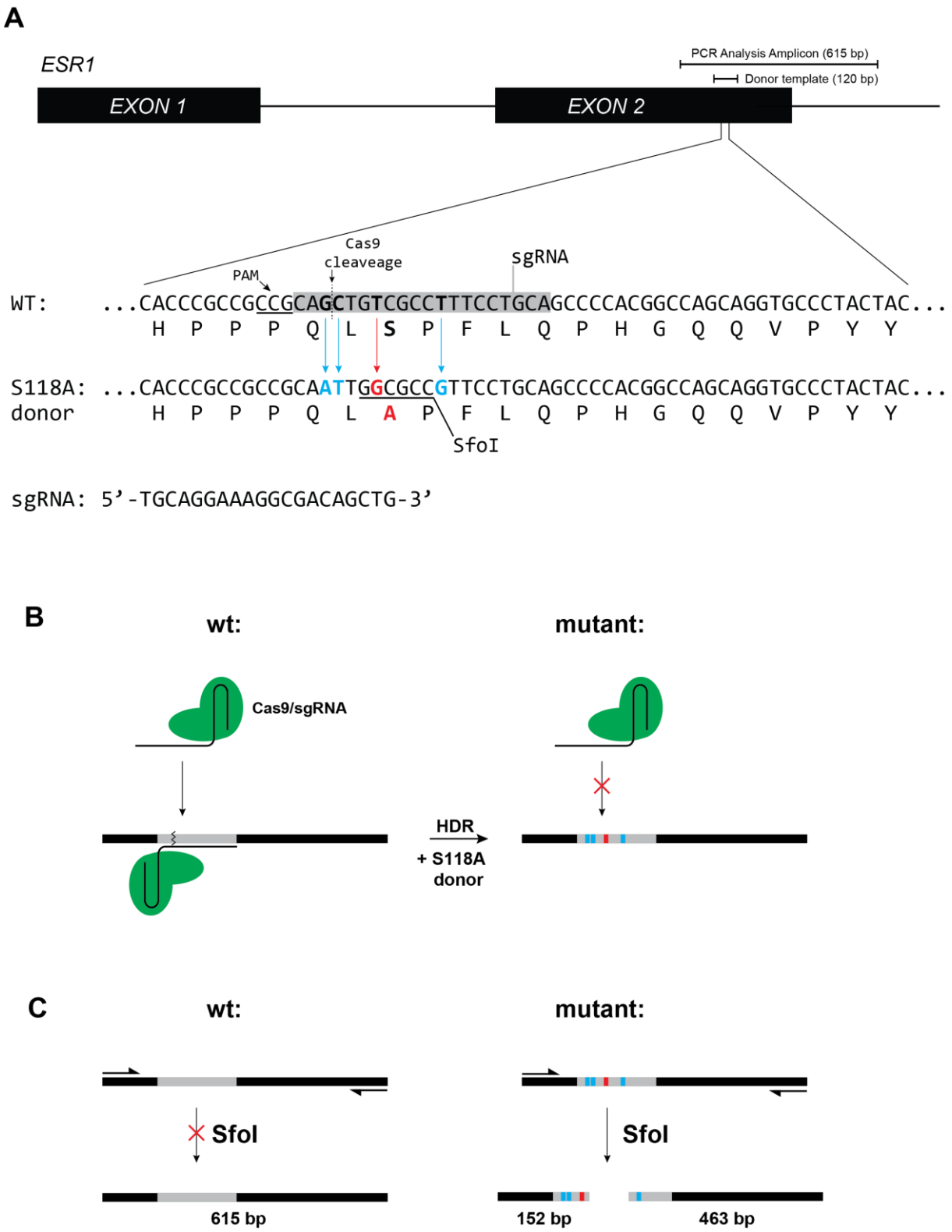

A
ESR1
PCR Analysis Amplicon (615 bp)
Donor template (120 bp)
EXON 1
EXON 2
PAM
Cas9 cleavage
sgRNA
WT: ...CACCCGCCGCCGCAGCTGTCGCCTTTCCTGCAGCCCCACGGCCAGCAGGTGCCCTACTAC...
H P P P Q L S P F L Q P H G Q Q V P Y Y
S118A: ...CACCCGCCGCCGCAATTGGCGCCGTTCCTGCAGCCCCACGGCCAGCAGGTGCCCTACTAC...
donor H P P P Q L A P F L Q P H G Q Q V P Y Y
SfoI
sgRNA: 5'-TGCAGGAAAGGCGACAGCTG-3'
B
wt:
mutant:
Cas9/sgRNA
HDR
+ S118A donor
C
wt:
mutant:
SfoI
SfoI
615 bp
152 bp
463 bp

Figure 4-2: Isolation of S118A ER Clones.

(A) PCR analysis of transfected cell populations. Parental MCF-7 and MCF-7 cells with no gRNA were used as controls. Genomic DNA was isolated and SfoI digestion performed. Arrow denotes processed band at the expected 463 bp indicting presence of the mutation in the S118A gRNA cell population. (B) MCF-7 S118A cell populations were passaged and multiple passages were analyzed for presence of the S118A mutation. The expected band at 463 bp is denoted by the arrow. (C) After single cell isolation and screening, two clones were identified with homozygous insertions of the S118A mutation indicated by full digestion by SfoI. As a positive control, both wt and S118A PCR amplicons were tested for digested by KpnI. Expected fragment sizes for KpnI digest are 255 bp and 360 bp. (D) Sequencing chromatogram of S118A clones confirming insertion of the S118A mutation. (E) Immunoblot analysis of S118A clones treated with E2 for 30 min. Both S118A do not induce phosphorylation at S118 whereas the wt does display pS118-ER.

Figure 4-2

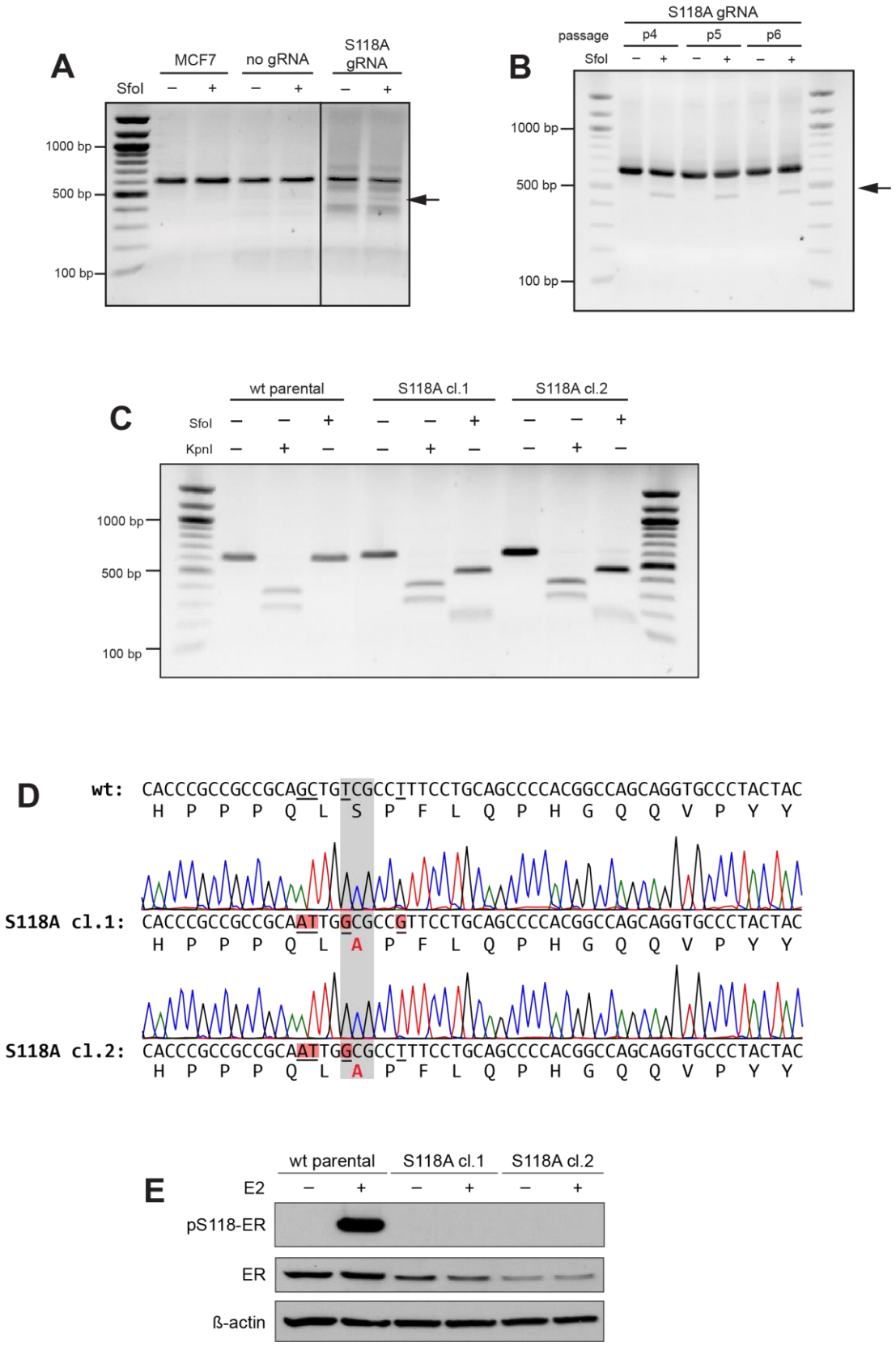

A
MCF7
no gRNA
S118A gRNA
SfoI
1000 bp
500 bp
100 bp
B
S118A gRNA
passage
p4
p5
p6
SfoI
1000 bp
500 bp
100 bp
C
wt parental
S118A cl.1
S118A cl.2
SfoI
KpnI
1000 bp
500 bp
100 bp
D
wt: CACCCGCCGCCGCAGCTGTCGCCTTTCCTGCAGCCCCACGGCCAGCAGGTGCCCTACTAC
H P P P Q L S P F L Q P H G Q Q V P Y Y
S118A cl.1: CACCCGCCGCCGCAATTGGCGCCGTTCCTGCAGCCCCACGGCCAGCAGGTGCCCTACTAC
H P P P Q L A P F L Q P H G Q Q V P Y Y
S118A cl.2: CACCCGCCGCCGCAATTGGCGCCTTTCCTGCAGCCCCACGGCCAGCAGGTGCCCTACTAC
H P P P Q L A P F L Q P H G Q Q V P Y Y
E
wt parental
S118A cl.1
S118A cl.2
E2
pS118-ER
ER
ß-actin

Figure 4-3: S118A mutation prevents E2-dependent ER protein downregulation.

Parental MCF-7 and S118A clones were estrogen-deprived for three days in phenol-red free DMEM supplemented with 10% serum-stripped FBS followed by treatment with vehicle (0.1% EtOH), 10 nM E2, 100 nM 4-hydroxytamoxifen, or 100 nM ICI 182,780 for 24 hr. Protein was isolated and analyzed by immunoblot. ß-actin served as a loading control. Band marked with an asterisks (*) on the ER blot is non-specific.

Figure 4-3

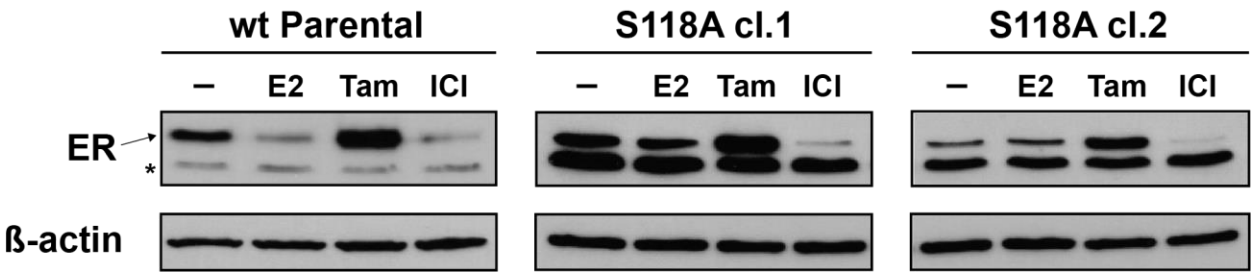
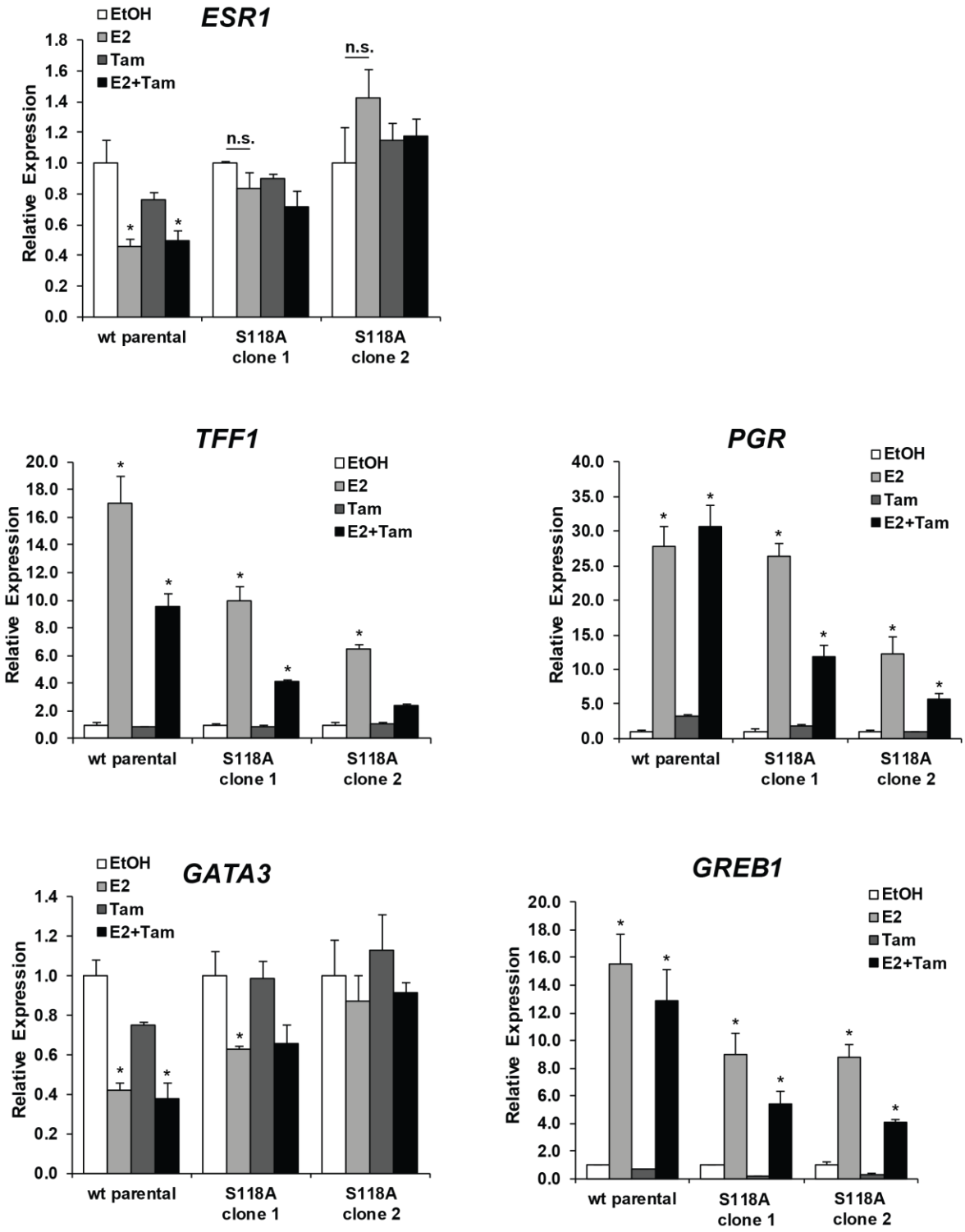


Figure 4-4: S118A ER mutation prevents downregulation of the *ESR1* gene by E2.

Parental MCF-7 and S118A ER clones were estrogen-deprived for three days followed by treatment with vehicle (0.1% EtOH), 10 nM E2, 100 nM tamoxifen (Tam), or E2 and Tam at the same concentrations for 24 hours. RNA was isolated and analyzed by qRT-PCR. Relative expression was calculated using the ΔΔCT method using the *RPLP0* gene as a housekeeping gene. Data are normalized to EtOH control. *$p<0.05$ compared to EtOH.

Figure 4-4


ESR1
EtOH
E2
Tam
E2+Tam
Relative Expression
n.s.
n.s.
wt parental
S118A clone 1
S118A clone 2
TFF1
PGR
GATA3
GREB1

CHAPTER FIVE:

Conclusions and Future Directions

Conclusions

The phosphorylation of proteins has proven to be a critical regulator of protein function. Since the discovery of ER as a phosphoprotein (50), and the subsequent identification of S118 as a phosphorylation site (52) much has been learned regarding the role of pS118-ER in ER transactivation and ER stability. While evidence points to pS118 affecting the DNA binding of ER to EREs and select sites in the genome, prior to this study the genome-wide analysis of pS118-ER occupancy had not been thoroughly investigated. The goal of this work was to fill in this gap in knowledge by defining the pS118-ER cistrome and identifying defining characteristics of pS118-ER sites compared to the general ER cistrome.

While many stimuli are known to induce phosphorylation of ER at S118 (Table 1-1), E2 was used in this work to describe how canonical ligand-dependent signaling is involved in directing pS118-ER specific DNA binding. Initial observations made investigating the necessity of pS118 for ER binding found that the S118A mutation does not abolish DNA binding but does reduce the induction of ER occupancy by E2 compared to the wt receptor indicating that pS118-ER is required for maximal ER occupancy. Additionally, through the use of an ER mutant which does not bind DNA, phosphorylation of ER at S118 was found to be independent of DNA binding, suggesting that phosphorylation occurs prior to the association of ER with chromatin.

After defining the pS118-ER cistrome in MCF-7 cells using a comprehensive approach with multiple antibodies, analysis of the pS118-ER sites through their comparison to ER sites revealed many similarities and differences. Numerous reports analyzing the ER cistrome found that a small proportion of ER occupancy sites fall within

promoter regions (121,156,188,189). Consistent with these observations, few pS118-ER occupancy sites were found to be located in promoter regions with the majority of sites located in introns or intragenic regions, indicating that pS118-ER does not localize to promoters relative to total ER. Although most of the pS118-ER sites fall far in linear genomic distance from transcription start sites, pS118-ER occupancy sites were enriched for the active enhancer mark H3K27ac (264) indicating that pS118-ER associates with active enhancers. Additionally, integration of the pS118-ER cistrome data with RNA-seq data from MCF-7 cells treated with E2 found that a higher proportion of upregulated genes had a pS118-ER occupancy site within 100 kb of their respective transcription start site compared to downregulated genes. These results suggest that pS118-ER is a more transcriptionally active form of ER compared to unphosphorylated ER and is associated with gene activation.

Motif analysis of pS118-ER occupancy sites revealed an enrichment of the ERE as well as the DNA binding motif for the transcription factor GRHL2 compared to ER sites. Further investigation into GRHL2 using previously published GRHL2 ChIP-seq data (265) found a greater proportion of GRHL2 occupancy sites overlapped with pS118-ER occupancy sites compared to ER sites. Additionally, upon E2 treatment GRHL2 occupancy increased at sites occupied by both pS118-ER and GRHL2, but not at GRHL2 occupancy sites absent of pS118-ER. While an association between ER and GRHL2 has been demonstrated in past studies (201,265,275), this report is the first to identify the association between pS118-ER and GRHL2. These results demonstrate that GRHL2 and pS118-ER localize to similar sites on the genome, and that E2 can increase GRHL2 binding as sites where pS118-ER is recruited.

Identification of the ERE as an enriched motif at pS118-ER sites compared to ER sites prompted the investigation of the role of pS118-ER in direct and indirect DNA binding. Using a customized *in vitro* DNA-binding array composed of small DNA segments tiling across selected ER binding sites from the human genome, purified pS118-ER was used to assess binding to these locations. Comparing the data between the genomic DNA array and the pS118-ER ChIP-seq data allowed for the identification of potential direct and indirect pS118-ER binding events. This analysis revealed that pS118-ER preferentially associates with direct binding rather than indirect binding, suggesting that pS118-ER acts as a key anchoring factor for the subsequent formation of protein complexes on chromatin.

In the process of performing this work, it was found that the selection of antibody used for ChIP-seq is critical and can impact downstream analysis. To create a comprehensive analysis of the pS118-ER cistrome, three different antibodies were used for ChIP-seq, and although all three passed the validation guidelines proposed by the ENCODE consortium (247), one of the antibodies produced a non-specific signal which is described in Chapter 2. This antibody produced false-positive pS118-ER peaks and demonstrates that adherence to the ENCODE guidelines does not guarantee meaningful positive results from an antibody. The data here suggest that future phospho-specific ChIP-seq experiments should critically evaluate the utility of the antibody before performing ChIP-seq, and even after ChIP-seq is performed to validate results using additional analysis.

An additional goal of this work was to expand the tool set available to researchers for analysis of pS118-ER, specifically in the ER-positive MCF-7 breast cancer cell line.

Much of the foundational work on pS118-ER was performed in ER-negative cell lines expressing the S118A ER mutant. A common criticism of these experiments is that the environment of ER-negative cell lines may not replicate what is occurring in a cell line dependent on ER expression, such as MCF-7. To address this issue, CRISPR/Cas9 technology was utilized to create two homozygous clones of MCF-7 cells expressing the S118A ER mutant. Using these engineered cells it was confirmed that the S118A ER mutation prevents downregulation of ER protein by E2, consistent with previous reports performed in HEK293 cells (78). Additionally, downregulation of the *ESR1* gene by E2 was also abrogated in the S118A ER mutant cells, indicating that E2-dependent ER downregulation is not limited to post-translational mechanisms. Given that these cells express S118A ER from the endogenous *ESR1* gene locus, these results provide a more relevant environment for the study of pS118-ER in ER biology.

An updated model of pS118-ER signaling and its role in DNA binding is presented in Figure 5-1. In the ligand-dependent signaling pathway, E2 binds to ER in the nucleus causing dimerization of ER. ER can then bind to DNA, interact with cofactors, and become phosphorylated at S118. The unphosphorylated form of ER is more associated with indirect DNA binding, whereas ER phosphorylated at S118 binds directly to DNA at sites enriched for the GRHL2 motif and the active enhancer mark H3K27ac. The phosphorylation at S118 is necessary for E2-dependent proteasome-mediated degradation of ER as well as repression of its gene, *ESR1*. While the work presented in this thesis has advanced the scope of knowledge on pS118-ER, there is much more to be learned and many additional questions have arisen from this study which will be discussed in the following section.

Overall Discussion and Future Directions

Expansion of ER phospho-cistrome to other phosphorylation sites and analysis of the ER "phospho-code"

The work in this thesis presents the cistrome of pS118-ER and observes many characteristics associated with pS118-ER. However, this is only one of many phosphorylation sites identified on ER. With 17 phosphorylation sites identified on ER (Figure 1-1; Table 1-1), the maximum combinations of phosphorylation states is theoretically 2^{17} = 131,072, assuming all sites are independent of each other. Although the actual number of phosphorylation states is likely fewer due to dependence between the phosphorylation sites and the low abundance of some phosphorylation events, this variety of combinations could allow for many functional outcomes which can fine tune the role of ER in a reversible manner. How do these multiple phosphorylation sites on ER coordinate to affect ER function? Understanding this ER "phospho-code" could help to refine our understanding of phospho-ER biology, as confounding data, especially at the clinical level, demonstrates that identifying pS118-ER status alone is not sufficient for understanding ER function or predicting patient outcomes (202,204,207,283). While select phosphorylation sites have been investigated individually in detail, very few studies have analyzed effects of multiple phosphorylation events in tandem on functional outcome. Analysis of these combinations is mainly limited to antibody-based techniques and mutational analysis. For ER, most phospho-specific research has focused on pS118, pS167, and pS104/106 due to the robustness to which they are induced, and the availability of antibodies targeting these phosphorylation sites. Many of these antibody-based methods analyze total ER protein population, and it is highly probable that many

different phospho-ER combinations are present in a cell at one time. Analysis of pS118-ER in MCF-7 has estimated that 67% of ER is phosphorylated at S118 in response to E2 (65). This analysis was performed by western blot which is a population analysis, so it is not clear whether 100% of the cells in the population have 67% of their ER phosphorylated, or 67% of the cells in the population have 100% of their ER phosphorylated, or any value between these ranges. Either way this suggests that there are different pools of ER existing simultaneously within the cell which are unphosphorylated or phosphorylated at different sites. These varying pools of ER could have differing effects on the function of ER.

In terms of how the ER phospho-code affects ER-DNA binding, analysis of the pS167-ER cistrome in response to E2 could provide additional insights into how phosphorylation governs the interaction of ER with chromatin. Past studies have found that phosphorylation of ER at S167 also affects ER-DNA binding (53,135) and comparing the pS167-ER cistrome to the pS118-ER cistrome can identify sites occupied by both pS118 and pS167 as well as sites specific to each of these ER phospho-forms. While the work presented here found an association between GRHL2 and pS118-ER, sites specific to pS167-ER could identify associations with other important factors involved in ER transcriptional function on DNA which interact specifically with pS167-ER. Other analysis which were performed in this work on the pS118-ER cistrome, if repeated on the pS167-ER cistrome could lead to additional insights on how different phosphorylation events further regulate the association of ER with chromatin.

In addition to analyzing the cistromes of other phosphorylation sites on ER, performing ChIP-seq on MCF-7 cells in response to other pS118-ER-inducing stimuli will

reveal insights into ligand-independent activation of ER. Many stimuli are known to activate ER through phosphorylation in a ligand-independent manner (Table 1-1), but how the ER or pS118-ER cistrome is altered in these conditions is largely unknown. Previous work showed that the ER cistromes between EGF-induced ER and E2-induced ER are distinct, with a moderate amount of overlap in between (140). Given that both of these stimuli lead to phosphorylation of ER at S118 (52,57), I would predict that the ER occupancy sites overlapping between the E2-induced sites and EGF-induces sites would be enriched for pS118-ER. Performing ChIP-seq for pS118-ER in cells stimulated with EGF or other stimuli which induce pS118-ER and comparing sites with pS118-ER sites identified in response to E2 would identify sites which are more dependent on pS118-ER compared to all ER sites.

Not limiting to ER, many of the other nuclear receptors are phosphorylated including AR (288,297), PR (298,299), GR (300,301), ERß (302), TR (303,304), all PPARs (305), RARα (306), RARγ (307), RXR (308), and VDR (309). Like ER, many of these phosphorylation sites lie within the N-terminal AF-1 domain of each respective receptor (257). Extensive cistromic analysis of AR, GR, and PR has been performed, but to date no phospho-specific cistrome has been analyzed for any of these nuclear receptors. Given the extensive role phosphorylation plays in the function of many nuclear receptors, phospho-specific cistromic analysis of the nuclear receptors listed above could reveal additional roles in their chromatin interactions.

Analysis of GRHL2 and its role in ER function

In this work, analysis of motifs enriched in the pS118-ER cistrome compared to the ER cistrome discovered the DNA binding motif for the transcription factor grainyhead-like

2 (GRHL2). A major question remaining from this thesis is what is the significance of GRHL2 in ER and pS118-ER signaling? GRHL2 is one of the three proteins in the grainyhead-like transcription factor family in mammals with the others being GRHL1 and GRHL3. The GRHL family of transcription factors is involved in epithelial development and GRHL2 has recently been associated with many cancers, particularly hormone dependent cancers (310–313). For example, an analysis of primary breast cancer tumor samples found a correlation between ER status and GRHL2 expression with ER-positive tumors expressing more GRHL2 on average than ER-negative tumors (275). Additionally, the GRHL2 motif was found to be associated with tumor-specific ER binding sites compared to ER binding sites in normal breast cells (314). Upon discovery of the GRHL2 motif enrichment at pS118-ER sites, it was hypothesized that pS118-ER or GRHL2 may be mediating the occupancy or transcriptional activity of the other. Work performed in Chapter 3 of this thesis found that treatment of MCF-7 cells with E2 leads to an increase in GRHL2 occupancy at sites occupied by pS118-ER and GRHL2, but not at GRHL2 sites absent of pS118-ER. A more recent study corroborated these observations by performing ChIP-seq on GRHL2 in MCF-7 treated with E2 and found that GRHL2 occupancy increased at many sites across the genome (239). These results suggest that E2 treatment and subsequent pS118-ER induction and DNA binding can lead to increased accessibility for GRHL2 at pS118-ER sites which contain the GRHL2 motif. Recent studies have classified the role of GRHL2 as a pioneer factor which is defined as a factor with the ability to promote open chromatin structures thereby allowing additional transcription factors to bind (315,316). Knockdown of all three GRHL family members led to a loss of DNA accessibility at GRHL binding sites as assessed by the Assay for

Transposase-Accessible Chromatin using sequencing (ATAC-seq) method supporting the role of the GRHL family of proteins as pioneer factors (277). In this sense it is also possible that GRHL2 may serve as a pioneer factor for pS118-ER. Conventional ChIP-qPCR for ER and pS118-ER has been attempted in MCF-7 cells with GRHL2 knocked down, and occupancy of both ER and pS118-ER is not significantly affected between the control and the GRHL2 knockdown (Appendix C; Figure C-5). However, it is possible that the sites selected do not respond to a GRHL2 knockdown and more sensitive techniques like ChIP-seq may need to be performed to fully assess the effect of GRHL2 knockdown on ER and pS118-ER occupancy.

In addition to ER, the androgen receptor (AR) has also been associated with GRHL2. Using a technique called Rapid Immunoprecipitation Mass spectrometry of Endogenous proteins (RIME) which allows for the unbiased analysis of proteins associated with a specific factor on chromatin (201,276), GRHL2 was identified as an AR interacting protein (317). GRHL2 was subsequently identified as an AR target gene with GRHL2 mRNA and protein expression increasing upon dihydrotestosterone (DHT) treatment in multiple AR-positive cell lines. Additionally, GRHL2 and AR were found to regulate each other, with knockdown of AR or GRHL2 in LNCaP cells leading to loss of expression of the other (317). Given that ER and AR are both in the hormone receptor family, it was hypothesized that ER may be regulated by GRHL2 in a similar manner. GRHL2 knockdown was performed in MCF-7 cells treated with and without E2 for 24 hours and expression of ER was analyzed by immunoblot. Loss of GRHL2 did not lead to changes in ER expression at the protein level and its loss did not affect the downregulation of ER by E2 (Appendix C; Figure C-3). Expression of AR and E-cadherin

was also assessed as expression of these two proteins was downregulated upon GRHL2 knockdown in the report by Paltoglou *et al.* (317). In MCF-7 cells, loss of GRHL2 did not have any effect on either AR or E-cadherin expression (Appendix C; Figure C-3), indicating that GRHL2 may be functioning differently in MCF-7 and LNCaP cells.

Additionally, the E2 response of ER-target genes was analyzed in the GRHL2 knockdown system. Genes known to be upregulated (*CCND1*, *GREB1*, *PGR*, *TFF1*) or downregulated (*ERBB3*, *ESR1*, *GATA3*) by E2 were analyzed in MCF-7 cells treated with E2 or vehicle along with GRHL2 knockdown. All genes tested displayed the expected response to E2, but no significant difference in E2 activation or repression was observed in the GRHL2 knockdown (Appendix C; Figure C-4). Interestingly, although the GRHL2 gene locus contains many pS118-ER and ER occupancy sites as identified from the ChIP-seq studies in this thesis, expression of *GRHL2* was not affected by E2 treatment (Appendix C; Figure C-1, C-4D). While these ER-target genes were chosen due to their known response to E2, this was done regardless of the presence of a nearby GRHL2 binding site and it is possible that other ER-target genes not assessed here are affected by GRHL2 knockdown. A more effective strategy would be to select genes with strong ER and GRHL2 occupancy sites within a certain proximity and to analyze expression of those genes in the GRHL2 knockdown. Alternatively, RNA-seq could be performed on GRHL2 knockdown cells with and without E2 to analyze the effect of GRHL2 loss on E2-dependent gene expression in an unbiased manner.

Due to the overlap between GRHL2 and pS118-ER sites, it was hypothesized that GRHL2 and pS118-ER interacted with each other in a complex. Co-immunoprecipitation (Co-IP) between GRHL2 and either pS118-ER or ER was attempted in our lab on MCF-

7 cells treated with E2 with limited success due to antibody issues. While our group has not been able to detect an interaction between GRHL2 and ER by Co-IP, other studies have been successful. Like AR, a RIME analysis of ER in MCF-7 cells identified GRHL2 as an interacting protein (201). GRHL2 was identified in all three ER RIME replicates and none of the five negative control replicates. This interaction was not investigated further, as it was one of many proteins identified in this study, but it provides strong evidence that ER and GRHL2 interact in a complex in MCF-7 cells. More recently, computational analysis of ER transcriptional networks combined with ER ChIP-seq data across multiple E2 time points identified GRHL2 as an ER cofactor involved in eRNA production (239). This study was able to Co-IP GRHL2 with ER in both MCF-7 and T47D cells. While these experiments demonstrate an association between ER and GRHL2, it is still unknown in this interaction is dependent on pS118-ER. Given the increased overlap observed between pS118-ER and GRHL2 compared to ER and GRHL2 found in this work, I would hypothesize that the interaction between GRHL2 and ER is dependent on pS118-ER. The S118A ER MCF-7 cells developed in Chapter 4 of this thesis would be an ideal model to test this hypothesis. Co-IP of GRHL2 and ER can be performed in the wt MCF-7 cells and the S118A ER cells to test if the interaction between ER and GRHL2 is dependent on phosphorylation of ER at S118.

In addition to ER, other ER-associated factors involved with histone modification have been associated with GRHL2. A RIME experiment performed on the pioneer factor FOXA1 identified the histone lysine methyltransferase MLL3 as the top FOXA1-interacting factor (265). ChIP-seq of MLL3 and subsequent motif analysis of MLL3 occupancy sites found an enrichment of the GRHL2 DNA binding motifs, similar to the

results found in this thesis on pS118-ER. The FOXA1 RIME also identified GRHL2 as a FOXA1 interacting protein. Comparisons of ChIP-seq data for MLL3, FOXA1, and GRHL2 found that nearly all MLL3 sites we co-occupied by either FOXA1 or GRHL2, indicating that either of these factors may mediate MLL3 occupancy on chromatin. These data suggest that GRHL2 may be involved in the recruitment of MLL3 causing the deposition of the histone marks H3K4me1 and H3K4me2 which are associated with enhancers (318). To test this, GRHL2 knockdown can be performed in MCF-7 cells and H3K4me1 or H3K4me2 status can be analyzed by ChIP-qPCR at specific ER enhancers or by ChIP-seq in an unbiased manner. A loss of H3K4 methylation at specific sites would indicate a loss of enhancer activity. Conversely, GRHL2 was found to interact with and inhibit the histone acetyltransferase activity of the ER coactivator p300 (319). The transcriptional activity of ER is enhanced by p300 (88,320) and it was found that 43.5% of p300 target genes were repressed by knockdown of GRHL2 (319). While this study identified inhibition of p300 by GRHL2, no experiments were performed involving ER despite its well-known association with p300. Although many associations have been made between ER and GRHL2, few mechanistic analyses have been performed to determine the role of GRHL2 in ER biology. Investigating the role of p300 inhibition by GRHL2 in ER signaling may provide insight this understudied field.

Analysis of pS118-ER function using the S118A MCF-7 model

The lack of a reliable model to study the S118A ER mutation in an ER-positive cell line prompted the generation of the S118A ER MCF-7 cell line using CRISPR technology (see Chapter 4). Many biological questions can be addressed on the various functions of pS118-ER using this new model, a few of which are selected and discussed here.

While the work described in Chapters 2 and 3 identified sites occupied by pS118-ER in MCF-7 cells, it did not determine which sites were dependent on pS118-ER for ER occupancy. A few pS118-ER occupancy sites were tested in MDA-MB-231 cells expressing S118A ER (Figure 2-2B) and were found to have partial dependence on pS118-ER. To address this question, ChIP-seq for ER could be performed in the S118A ER MCF-7 cells and the cistromes between S118A and wt ER can be compared. This experiment would determine which ER binding sites are dependent on pS118-ER and would expand on this thesis work which mainly defines where pS118-ER is present and does not assess pS118-ER dependence on a genome-wide level. In this work, only a few individual sites were analyzed for their dependence on pS118-ER in an ER-negative cell line expressing the S118A mutation and these sites were found to be dependent on pS118-ER for maximal occupancy (Figure 2-2B). These experiments would not only expand the analysis of pS118-dependece on ER occupancy to a genome-wide level, it would also be performed in the ER-positive MCF-7 breast cancer model allowing for the effect of S118A ER on ER occupancy to be assessed in an ER-native environment. Additionally, using the S118A ER MCF-7 cells would alleviate the need for phospho-antibodies, as the general ER antibody can detect both wt and mutant S118A ER. While comparisons within the same antibody are preferable to comparisons between antibodies, one drawback is be that similar levels of receptor need to be present in both the wt and the S118A ER cell lines to accurately determine differences between the two cistromes. One of the clones developed here (S118A cl.2) displays lower levels of ER which would affect the number of peaks detected in a ChIP-seq experiment so using S118A cl.1 which

has ER levels more closely resembling the wt control would be more ideal for these experiments.

Many ER-associating cofactors are known to interact with ER in a pS118-specific manner (see Chapter 1). These studies have been limited by their analysis on a protein-by-protein basis, and the full spectrum of pS118-depednent interacting proteins in MCF-7 cells is not known. With the development of the RIME technique to analyze protein complexes on DNA through mass spectrometry (201), analysis of wt ER and S118A ER MCF-7 cells using this technique could lead to the identification of additional factors dependent on pS118-ER for interaction with ER in an unbiased manner. This could be used to not only verify previously reported pS118-dependented interactions, but also to identify novel interactors which are difficult to detect by co-immunoprecipitation. Identification of novel pS118-specific interactors could lead to a deeper understanding of pS118-ER function.

Phosphorylation of ER at S118 has been implicated in the ubiquitylation and degradation of ER in response to E2 (78,109). Studies from this thesis work confirmed that the S118A ER mutation prevents ER from downregulation by E2 (Figure 4-3). The pS118-ER cistrome presented here has additionally demonstrated that pS118-ER is highly involved in ER occupancy, particularly at active enhancers. ER activity on certain E2-responsive promoters is known to be cyclical (145) and proteasome-mediated turnover is required for this cycling to occur (321). Given these links between ER degradation, phosphorylation, and DNA binding, I hypothesize that the S118A mutation prevents release of ER from DNA after it is bound. The time points for pS118-ER occupancy analyzed by ChIP-qPCR and ChIP-seq in this thesis work were all at 30

minutes E2 treatment and only concerned initiation of ER-DNA binding. Using the S118A ER MCF-7 model developed here, a time course ChIP-qPCR comparing wt ER and S118A ER could determine if the S118A mutation affects the cycling of ER on E2-responsive gene promoters as well as the release of ER from DNA by analyzing later time points. The observation that pS118-ER is required for maximal occupancy on DNA and maximal gene expression in response to E2 could be explained by the prevention of DNA release in the S118A mutation, as a single round of ER binding and transcriptional activation would be detected, but further rounds of ER cycling could not be performed to achieve maximal ER binding and transcription of target genes.

Numerous other biological questions regarding pS118-ER could be addressed here to either confirm previous analyses of S118A ER in ER-negative cell lines or to investigate novel functions of pS118-ER. Importantly, the development of this S118A ER MCF-7 cell line provides a tool to analyze the function of pS118-ER in the widely used MCF-7 breast cancer model.

Connecting pS118-ER ChIP-seq peaks to gene regulation

ChIP-seq has revolutionized the study of chromatin-protein interactions by expanding the scope of analysis from single genomic locations to a genome-wide level. It has allowed for unbiased identification of protein-DNA interactions and has contributed to the overall understanding of chromatin structure and protein complex formation on DNA. However, ChIP-seq does have limits which have still not been fully addressed. While ChIP-seq experiments can provide the binding locations of transcription factors across the genome, it cannot assign the functionality of a site to a specific gene. Generally, attempts to tie binding sites to gene regulation have consisted of annotating

peaks to the nearest gene transcription start site. This works well for binding sites which fall within promoter regions or are in close proximity to a gene, but with many transcription factors such as ER, most of the binding sites occur at distal regions which cannot easily be assigned to a gene by a simple distance measurement. In this thesis work, it was demonstrated that genes upregulated by E2 were more likely to have a pS118-ER occupancy site within 100 kb of their respective transcription start sites, however direct connections between specific pS118-ER sites and genes were not determined. It is also unknown if pS118-ER sites are more necessary than ER sites for gene regulation. This could be addressed in various ways.

Recently, a method called Enhancer-interference (Enhancer-i) was developed to investigate the relationship between enhancers and gene regulation (282,322). It consists of a catalytically dead Cas9 (dCas9) combined with the repressive domains of KRAB and the SIN3A interacting domain of MAD1 (SID). The dCas9/repressor complex is directed to specific enhancer sites using sgRNAs to inactivate them and gene expression can be assessed. This method could be used to determine which pS118-ER or ER sites are important for expression of a gene within a certain range. The *TFF1* locus is a good model region to test this as there are limited ER and pS118-ER occupancy sites to assess, and *TFF1* is the only ER target gene in the vicinity. Selecting pS118-ER and ER sites around the *TFF1* locus and designing guides to direct selective repression at each site could determine which binding sites are most important for *TFF1* expression. These experiments could also ascertain if pS118-ER sites are more necessary for gene activation compared to ER sites. Given the association of pS118-ER with the active enhancer mark H3K27ac, I would hypothesize that inactivating pS118-ER occupancy

sites around an E2-activated gene would decrease its activation more than inactivation of an ER site.

Another method of gene regulation that was not addressed in this work was the formation of chromatin loops to drive gene transactivation. ChIP-seq data is displayed on a linear genome, whereas the three-dimensional architecture of chromatin is more complex, involving interactions between sections of the genome which are far in linear distance. These looping structures are known to link enhancers to promoter regions and the presence of a chromatin loop correlates with gene activation (323). A method called Chromatin Interaction Analysis by Paired-End Tag sequencing (ChIA-PET) was developed to identify chromatin interactions mediated by ER at a genome-wide level (185). This study found that ER was involved in extensive chromatin looping which was induced by E2. Loss of ER led to the loss of these chromatin loops indicating that ER is an important mediator for loop formation. A remaining question from this work is how does pS118-ER impact the formation and activity of these ER-dependent chromatin loops? Given the presence of pS118-ER at active enhancer sites and the correlation between chromatin loops and gene activity, I would hypothesize that pS118-ER is more likely to be found at looping sites compared to unphosphorylated ER. If ER-driven chromatin loops are dependent on pS118-ER, it could explain how pS118-ER is required for maximal gene transcription of ER target genes (59,77–81), as these genes could be partially activated from ER binding to the promoter, but transcription could not be further enhanced through looping of enhancer sites to the promoter. This could be tested using the S118A ER MCF-7 model developed in Chapter 4. Individual looping sites could be analyzed by chromatin conformation capture (3C) or global looping could be assessed by Hi-C (324) or ChIA-

PET between the wt and S118A ER to determine if pS118-ER plays a role in ER-dependent chromatin looping.

Of the methods described above, ChIA-PET or Hi-C would allow for the entire pS118-ER interaction network to be determined at a genome-wide level without a readout for gene transcription, whereas Enhancer-i would be utilized on a gene-by-gene basis to determine specific interactions between enhancers and promoters and the gene in question. Identification of these binding site-gene pairings would be useful and could aid in analysis of ER or pS118-ER cistromes performed in patient tumors by identifying enhancer sites in control of aberrantly expressed genes. ER occupancy sites specific to breast cancer patient tumors have been identified (234,314), but specific genes associated with these unique binding sites have not been determined.

Control of gene expression by multiple ER binding events within a single ChIP-seq peak

Interestingly, it was observed that many pS118-ER occupancy sites which were defined as a single ChIP-seq peak contained multiple potential pS118-ER binding sites within the peak when analyzed by the *in vitro* DNA binding array. It was assumed that a single ChIP-seq peak would be generated from a single binding site located in the center of the peak, however this was not found to be the case in all instances. All genomic locations assayed on the array were analyzed for the presence of pS118-ER peaks and it was found that while many sites (31.9%) contained one pS118-ER binding site located roughly in the middle of the peak, the majority of sites (56.4%) contained two or greater potential pS118-ER binding events. Analysis of these sites found a positive correlation between the number of ER binding events present and the percent of sites occupied by

pS118-ER (Figure 3-5E). At the extreme end of the spectrum, one pS118-ER occupancy site, which in ChIP-seq data displayed as a single peak, contained 35 possible pS118-ER binding locations throughout the 1,030 bp region analyzed on the DNA array. This site is located in an intron in the *CRAMP1* gene on chromosome 16. Analysis of the DNA sequence in this binding region found that it contained 53 instances of an ERE half site (AGGTCA) with only two of these being part of a full ERE (Figure 5-2). The number of half sites in this region is greater than the number of binding sites detected from the array analysis because the 25 bp probe length can accommodate multiple half sites which are ultimately detected as one binding event. It is unknown what the biological function of a large number of half sites in a small region is, but one explanation could be that this site requires a low but constant level of ER occupancy for its specific function, as ER binds to a half ERE with lower affinity than a full ERE. Although this example describes an outlier, there are 480 ER ChIP-seq peaks which contain 4 or more potential ER binding sites within their respective ChIP-seq peaks as assessed by the array (Figure 3-5D). Additionally, because of space limitations on the array, all ER and pS118-ER occupancy sites were not analyzed so this number is likely higher. These observations bring up the interesting question of how multiple ER binding events in a single region affect ER activity. One hypothesis is that an increased number of ER binding events allows for increased transcriptional control, as sites with a single ER binding event would function in a binary manner (Figure 5-3B). Previous reports have found that the addition of multiple EREs upstream of a reporter gene increased expression of the reporter in response to E2 (155), however this was performed with perfect EREs and in a reporter construct void of many endogenous chromatin components. Similarly, while the data collected from the *in vitro*

DNA binding array detects binding events at an increased resolution compared to ChIP-seq, analysis is limited to simple protein-DNA interactions in the absence of histones or any accessory factors that would aid in DNA binding. With the data generated here, a correlation between the number of ER binding events and nearby gene induction by E2 could be assessed. I would hypothesize that a greater number of ER binding events would correlate with an increase in gene expression on its associated gene. The number of ER binding events in a peak could also be correlated with other marks known to be associated with active enhancers such as H3K27ac (264), p300 (325), and eRNA production (194) to determine if these multi-ER binding sites are also correlated with increased transcriptional activity. Mechanistically, it is unknown if the function of multiple ER binding events within a single binding site is to allow binding of multiple ER dimers simultaneously, or simply to increase the probability of at least one ER dimer binding to the region, and thus activating transcriptional activity (Figure 5-3A). Further research should be performed in this area to determine how these multi-ER binding regions function to control gene regulation.

Concluding Remarks

Prior to this thesis work, the effects of pS118-ER on ER-DNA binding at a genome-wide scale were unknown. Here, the cistrome of ER phosphorylated at S118 is described, providing insight into the role of pS118-ER in ER biology. These analyses found that pS118-ER is located at a subset of ER occupancy sites and that pS118-ER associates with active enhancers, the transcription factor GRHL2, and direct DNA binding. Additionally, development of the S118A ER MCF-7 knock-in cell line confirmed the role

of pS118-ER in E2-dependent ER degradation in MCF-7 cells and further found that downregulation of the *ESR1* gene by E2 is similarly affected by the S118A mutation, indicating that E2-dependent ER downregulation is not limited to post-translation mechanisms. Taken together, these experiments greatly expand the current knowledge on pS118-ER and provide a basis to build further research on ER phosphorylation and its role in breast cancer.

Figure 5-1: A revised model for pS118-ER function.

In the ligand-dependent signaling pathway, E2 binds to ER in the nucleus causing dimerization of ER. ER can then bind to DNA, interact with cofactors, and become phosphorylated at S118. ER phosphorylated at S118 binds directly to DNA at sites enriched for the GRHL2 motif and the active enhancer mark H3K27ac. Indirect ER binding is more associated with ER not phosphorylated at S118. The phosphorylation at S118 is necessary for proteasome-mediated degradation of ER as well as repression of its gene, *ESR1*.

Figure 5-1

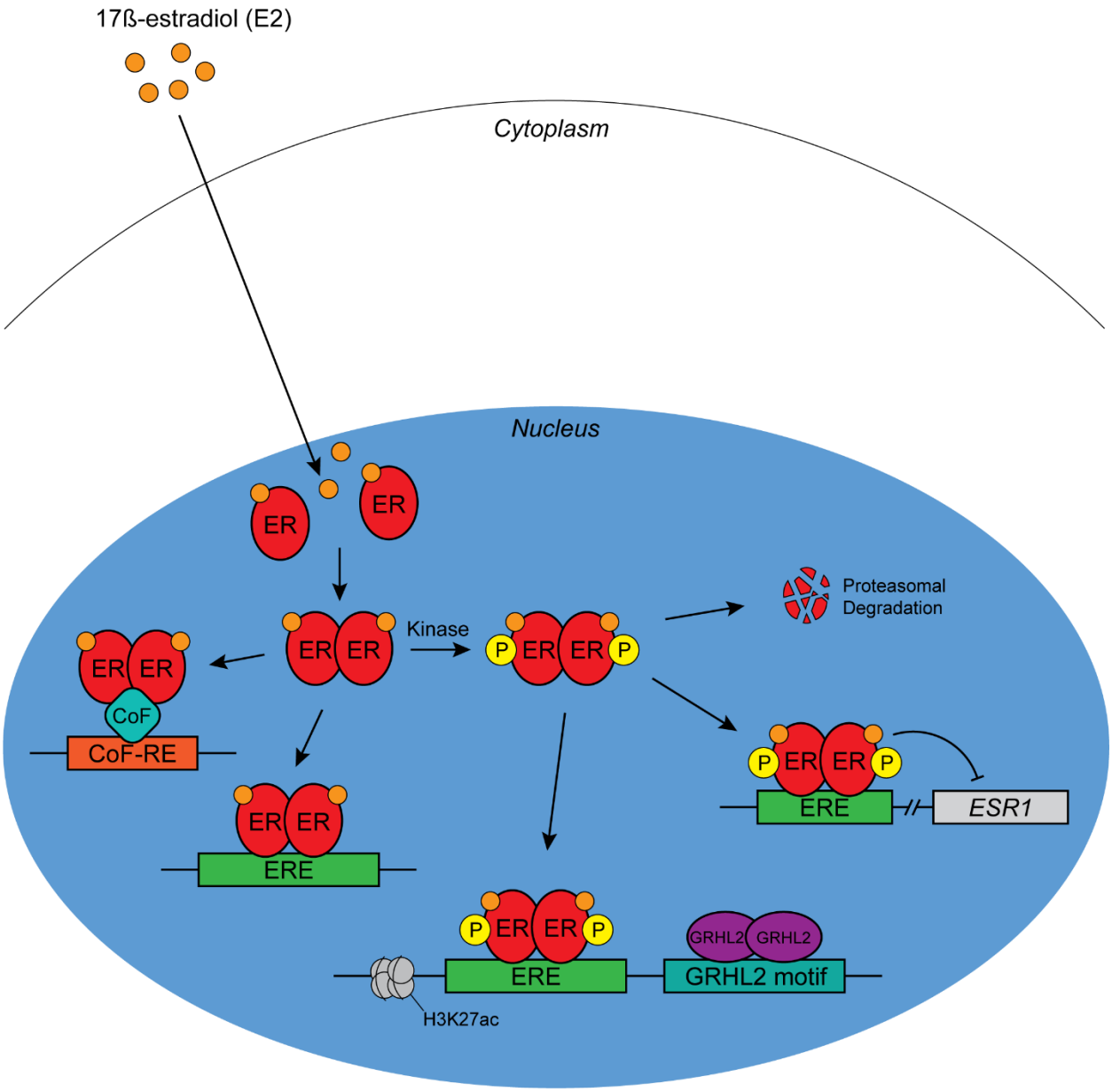
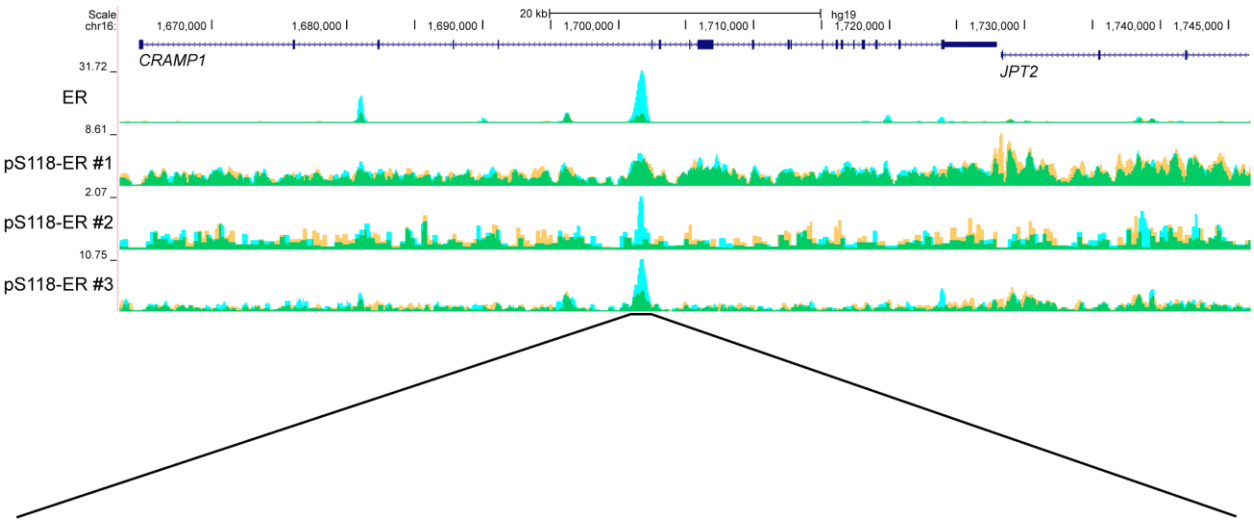

17ß-estradiol (E2)
Cytoplasm
Nucleus
ER
ER
ER ER
Kinase
P ER ER P
Proteasomal Degradation
ER ER
CoF
CoF-RE
ER ER
ERE
P ER ER P
ERE
ESR1
P ER ER P
ERE
GRHL2 GRHL2
GRHL2 motif
H3K27ac

Figure 5-2: A multi-ER binding site in the *CRAMP1* gene.

SNAP array analysis found 35 potential ER binding events located in a single pS118-ER ChIP-seq peak located in an intron of the *CRAMP1* gene on chromosome 16. Sequence analysis of the binding region found 53 instances of a half ERE (5'-AGGTCA-3', underlined and bolded).

Figure 5-2

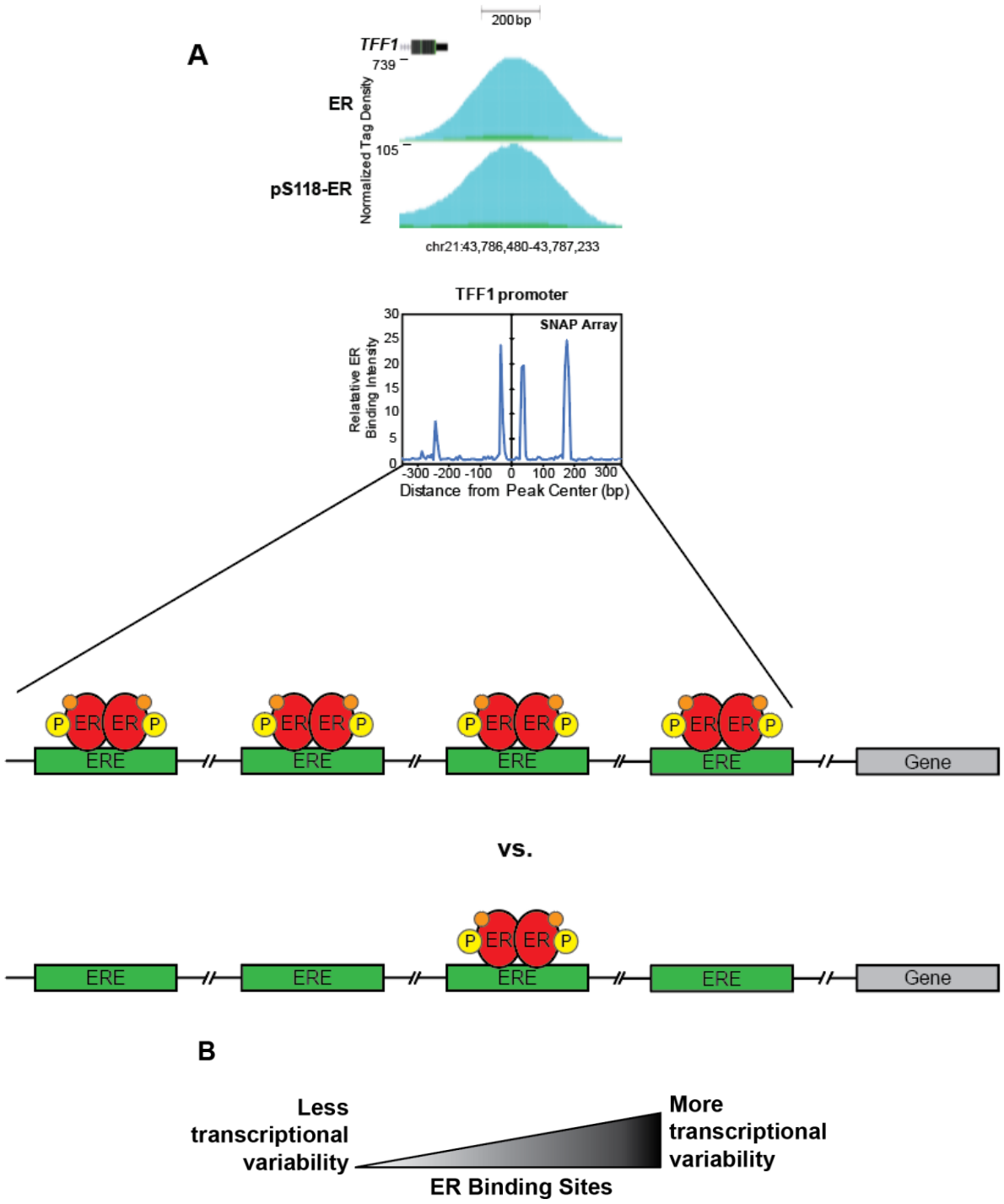


GAGAAGTGGACTGAG**AGGTCA**CAGAGAGGAGGACTCAGAGGCCAC**AGGTCA**TGGAAAGG
TAGATTG**AGGTCA**CAG**AGGTCA**CGAAAAGGTGGACTG**AGGTCA**CGGAGAGGTGGACTG**A**
GGTCACAC**AGGTCA**CAAAAAGGTGGACTCAG**AGGTCA**CAGAGTTGGACAGGTTACACAG**A**
GGTCACACGGTGGACTGAGATCACGGAGAGCTGGACTG**AGGTCA**CACACAGGTTATGGA
GAGGTCGACTGAG**AGGTCA**CACAG**AGGTCA**TCTAGGGGTGGATTT**AGGTCA**CGCAC**AGG**
TCACGGAAAGGTGGACTGAG**AGGTCA**CGGAGAGGTGGACTG**AGGTCA**CAG**AGGTCA**CCT
AGAGGTGAATTGTCACACAG**AGGTCA**TGTTGAGGTGGACTG**AGGTCA**CACAC**AGGTCA**TC
TAGAGGTGGATTG**AGGTCA**CACAC**AGGTCA**TCTAGAGGTGGACTG**AGGTCA**CAC**AGGTC**
ATTTAGAGGTGGATTG**AGGTCA**CACAGTTCACGGAGAGGTGGGCTG**AGGTCA**CACAG**AG**
GTCATGGGGAGGTGGACTGAG**AGGTCA**CCTAGAGGTGGATTGAAGTCACACAGAGGTAA
CCTAGAGGTGGACTGAG**AGGTCA**CGGAGAGGTGGACTG**AGGTCA**CACAG**AGGTCA**CCTA
GAGGTGGATTG**AGGTCA**CACAG**AGGTCA**TAGAGAGGTGGACTG**AGGTCA**CAC**AGGTCA**C
AGAGAGGTGGACTGAG**AGGTCA**CACAG**AGGTCA**CAGAAAGGTGGATTG**AGGTCA**CACAG
AGGTCACAGAAAGGTGGACTG**AGGTCA**CAC**AGGTCA**TCTAGAGATGGATTG**AGGTCA**CAT
GG**AGGTCA**CGGGGAGGTGAACTG**AGGTCA**CAGAG**AGGTCA**CGTAGAAGTGGATTG**AGGT**
CACACAG**AGGTCA**TGGGGAGGTGGACTG**AGGTCA**C**AGGTCA**TGTAGAGCTGGATTT**AGG**
TCACACAG**AGGTCA**CAGGGAGGTGGACTGAG**AGGTCA**CGGAGAGGTAGACTGAG**AGGTC**
ATGGGGAGGAGGACTGAGAGCTTC

Figure 5-3: Potential Model of ER function in multi-ER binding sites.

Many ER occupancy sites as defined by ChIP-seq were found to have multiple potential ER binding events within a single ChIP-seq peak (see Figure 3-5D). (A) The *TFF1* promoter is occupied by pS118-ER and has four pS118-ER binding sites as defined by the *in vitro* SNAP array (See Chapter 3, Figure 3-5). It is not known if all these sites can be bound by ER simultaneously or if only one can be bound at a time. (B) The presence of more ER binding sites within a promoter or enhancer may give the site more variability in gene transcription, whereas a site with only one ER binding event would act in a binary manner.

Figure 5-3


A
200 bp
TFF1
739
ER
Normalized Tag Density
105
pS118-ER
chr21:43,786,480-43,787,233
TFF1 promoter
SNAP Array
Relative ER Binding Intensity
30
25
20
15
10
5
0
-300 -200 -100 0 100 200 300
Distance from Peak Center (bp)
P ER ER P
ERE
Gene
vs.
B
Less transcriptional variability
More transcriptional variability
ER Binding Sites

CHAPTER SIX:

Materials and Methods

Cell Culture

MCF-7, MDA-MB-231, HEK293, and the MCF-7 S118A ER cell lines were maintained in high-glucose (4.5 g/L) Dulbecco's modified Eagles Medium (DMEM; Mediatech, Manassas, VA) containing 10% fetal bovine serum (FBS), 100 U/mL penicillin and 100 µg/mL streptomycin incubated at 10% CO_2 and 37°C. Prior to estrogen treatments, cells were grown for three days in stripping media consisting of phenol-red free DMEM supplemented with 10% charcoal-stripped FBS, 100 U/mL penicillin, 100 µg/mL streptomycin, 4 mM L-glutamine, and 1 mM sodium pyruvate.

Drug Treatments

Prior to all drug treatments, cells were estrogen deprived by culturing for three days in phenol-red free DMEM supplemented with 10% charcoal-stripped FBS, 1% penicillin/streptomycin, 1% sodium pyruvate, and 2% L-glutamine. Cell were then treated with vehicle (0.1% EtOH), 10 nM 17β-estradiol, 100 nM 4'-hydroxytamoxifen, or 100 nM ICI-182780 for various amounts of time depending on the experiment. For dual 17β-estradiol and 4'-hydroxy tamoxifen treatments, the same concentrations as above were used.

Western Blot Analysis

Post-treatment, cells were washed with PBS and lysed in-well with 2x SDS sample buffer (120 mM Tris-HCl, 20% glycerol, 2% SDS, 2% ß-mercaptoethanol, 0.01% bromophenol blue, pH 6.8) followed by boiling at 95°C for 10 min. Protein concentrations were quantified using an RC DC kit (Bio-Rad, Hercules, CA) and equal levels of protein were

run on a 10% SDS-PAGE gel. Blots were transferred to a PVDF membrane followed by blocking with 5% milk in TBST (50 mM Tris-HCl, 150 mM NaCl, 0.1% Tween 20). Antibodies and their primary incubation conditions can be found in Table 6-1. All antibodies were diluted in 5% milk in TBST and incubated with blots overnight at 4°C. Blots were washed 3 x 20 min in TBST followed by 1 h incubation with either sheep α-mouse IgG HRP (NA931V, GE Healthcare, 1:3000) or donkey α-rabbit IgG HRP (NA934V, GE Healthcare, 1:3000) diluted in 5% milk in TBST. Blots were washed 3 x 20 min in TBST followed by 1 x 20 min wash in TBS. Blots were then exposed to Clarity Western ECL solution (BioRad) followed by exposure to autoradiography film (HyBlot CL, Denville Scientific).

RNA analysis

For all cell types analyzed, cells were seeded in a 12-well plate at a concentration of $2.0x10^5$ cells per well and estrogen deprived for three days as described above. After estrogen deprivation, cells were treated with 10 nM E2, 100 nM Tam, or a combination of E2 and Tam for 24h. RNA was then isolated following the specifications of the QIAGEN RNeasy Mini Kit (QIAGEN, cat. No. 74106) and cDNA was generated using the BioRad iScript cDNA Synthesis Kit (BioRad, cat# 1708891) .Equal volumes of samples were analyzed on a 96-well PCR plate (BioRad, cat# HSP9601) with BioRad SYBR Green Master Mix (BioRad, cat# 1708882) using the BioRad CFX Connect Real-Time System qPCR machine. Primers sequences used can be found in Table 6-2. Genes were normalized to the housekeeping gene *RPLP0* using the ΔCT method followed by

normalization to the vehicle control. Statistical analysis was performed using Student's t-test with p-values below 0.05 considered as significant.

MDA-MB-231 ER Constructs

MDA-MB-231 cells stably expressing a doxycycline inducible wild-type (wt) or S118A ER construct were generously provided by Dr. Amy Fowler and were maintained in DMEM as described above in cell culture. Cells were stripped prior to estrogen treatment as described in cell culture methods. On the second day of stripping, cells were treated with doxycycline at either 5.0 µg/mL (wt ER construct) or 0.5 µg/mL (S118A ER construct) and incubated for 24 hours (h) at 37°C and 5% CO_2. Doxycycline concentrations were optimized to express similar levels of protein between constructs. After 24-hour incubation with doxycycline, cells were treated for their respective experiments.

Transfection Procedure for HEK293 Cells

HEK293 cells were stripped for three days as described in cell culture methods. After stripping, 5.0 x 10^5 cells were seeded in 6-well plates in penicillin-free, streptomycin-free stripping media and incubated overnight at 37°C and 5% CO_2. Plasmids containing either wt ER-HA, C202/205H ER-HA, or LHL-CA vector were used for transfections. The LHL-CA vector controls expression of its constructs through the cytomegalovirus-ß-actin fusion promoter and is described elsewhere (326). Cells were transfected for 5 h with 5 µg of DNA using Lipofectamine 2000 (ThermoFisher Scientific). Post transfection, media was replaced with stripping media containing penicillin and streptomycin and incubated overnight at 37°C and 5% CO_2. Cells were then treated with either ethanol vehicle (0.1%

EtOH) or 10 nM E2 for 30 min, harvested, and analyzed by immunoblot as described in Western blot analysis.

Chromatin Immunoprecipitation

MCF-7 cells were stripped as described in cell culture methods. Cells were then treated with either vehicle (0.1% EtOH) or 10 nM E2 for 30 min. After treatment, cells were washed with PBS and then crosslinked with 1.5% formaldehyde for 15 min at room temperature followed by quenching with 0.125 mM glycine for 5 min. All buffers used for preparation of chromatin contained PMSF (1 mM), aprotinin (10 µg/mL), leupeptin (10 µg/mL), and $NaVO_4$ (200 nM). Cell pellets were subjected to nuclear isolation by resuspending in 1 mL Nuclear/Chromatin Preparation (NCP) Buffer 1 (10 mM EDTA, 0.5 mM EGTA, 10 mM Hepes pH 6.5, 0.25% Triton X-100) followed by centrifugation at 1650 x g for 5 min at 4°C. The pellets were resuspended in 1 mL NCP Buffer 2 (1 mM EDTA, 0.5 mM EGTA, 10 mM Hepes pH 6.5, 200 mM NaCl) followed by centrifugation at 1650 x g for 5 min at 4°C. The pellets were then lysed in 1 mL lysis buffer (1% SDS, 5 mM EDTA, 50 mM Tris-HCl pH 8.1) and sonicated three times for 15 seconds (s) at 15% maximum power (Fisher Scientific 550 Sonic Dismembrator). Samples were allowed to cool on ice between each sonication. Sonicated samples were centrifuged at 10000 x g at 4°C, and the supernatant was collected. Aliquots were taken prior to immunoprecipitation (IP) for use as inputs. For each IP, 100 µL of sonicated chromatin solution was taken and diluted 10-fold with IP dilution buffer (1% Triton X-100, 2 mM EDTA, 150 mM NaCl, 20 mM Tris-HCl pH 8.1). The diluted chromatin solution was precleared for 2 h at 4°C with 20 µL 50% Protein A Sepharose Beads (CL-4B, GE

Healthcare), 2 µg/mL herring sperm DNA, and 5 µg/mL BSA. After the preclearing incubation, chromatin was centrifuged at 1000 x g to pellet beads and the supernatant was collected. Precleared chromatin was subjected to IP with 1 µg antibody to ER (HC-20, sc-543; Santa Cruz), pS118-ER #1 (16J4; Cell Signaling), pS118-ER #2 (ab32396; Abcam), pS118-ER #3 (sc-12915; Santa Cruz), GRHL2 (HPA004820, Sigma, St. Louis, MO), or normal rabbit IgG (sc-2027; Santa Cruz) and incubated overnight at 4°C while rotating. To each sample, 45 µL of 50% Protein Sepharose A beads was added and incubated for 1 h at 4°C. The precipitated complexes were washed once with 1 mL TSE I buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.1, 150 mM NaCl), once with 1 mL TSE II buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.1, 500 mM NaCl), once with 1 mL TSE III Buffer (0.25 M LiCl, 1% NP-40, 1% sodium deoxycholate, 10 mM Tris-HCl pH 8.1), and twice with 1 mL TE (10 mM Tris-HCl pH 8.1, 1 mM EDTA). Chromatin complexes were extracted from beads by incubating in 75 µL Extraction Buffer (1% SDS, 0.1 M $NaHCO_3$) for 30 min with vortexing every 5 min. Extraction was repeated 3 times and the extractions were pooled. Samples were reverse crosslinked by incubating at 65°C overnight. DNA was purified using the QIAquick PCR Purification Kit (Qiagen; Hilden, Germany) and 1 µL of the purified DNA was analyzed via qPCR. Data are calculated as a percent of input. Primers used for ChIP-qPCR are included in Table 1.

For ChIP-seq, 10 x 10 cm plates per condition of MCF-7 cells were stripped as described in cell culture methods and then treated with either vehicle or 10 nM E2 for 30 min. Samples were prepared as above and the remaining ChIP-seq procedure was performed as previously described (327). Antibodies used for ChIP-seq were as follows:

ER (HC-20, sc-543, Santa Cruz), pS118-ER #1 (16J4, Cell Signaling), pS118-ER #2 (ab32396, Abcam), pS118-ER #3 (sc-12915, Santa Cruz). ChIP-seq was performed in triplicate for each antibody. Peaks were called as previously described (327) and aligned to the human genome (hg19). Only those peaks which were present in all three replicates were utilized for further analysis. All data files are deposited in the GEO repository (GSE117569)

Motif Analysis and Peak Annotation

Motif analysis was performed using Hypergeometric Optimization of Motif EnRichment (HOMER; (175). Motifs of length 8, 10, 12, and 15 were searched for in the *de novo* analysis and the size parameter was set to 200 (-size 200). Peak annotation was performed using the annotatePeaks.pl program from HOMER and the basic annotations were used. Differential motif analysis was performed using the CentriMo tool in the MEME suite (184). Neighboring secondary motif analysis was performed using the SpaMo analysis tool (267).

Specificity and Affinity for Proteins (SNAP) Array

The Genomic SNAP array was bioinformatically designed (Proteovista LLC, Madison, WI) to tile across ER binding sites compiled from 58 ChIP-seq datasets across conditions, cell lines and tumors. To identify high-confidence ER binding sites, genomic regions were culled from these datasets to include ER binding regions from at least five manuscripts, reported from two independent authors, and represented in both cell lines and in tumors (140,159,187,188,196,201,234,236). These criteria identified a set of ER genomic

sequences that were used to design SNAP probes, tiling across 5020 genomic regions at 6 bp spacing. The array also contains positive and negative controls for ER binding. Purified, estrogen-occupied ER (ThermoFisher Scientific) was incubated with the SNAP microarrays (synthesized by Agilent Technologies) and binding was detected by fluorescence and analyzed according to established protocols (110).

Development of MCF-7 S118A knock-in CRISPR cell lines

Generally, the protocol for “Alt-R CRISPR-Cas9: Delivery of ribonucleoprotein complexes in Jurkat T cells using the Neon Transfections System” was followed, with modifications for MCF-7 cells (IDT). Parental MCF-7 cells were brought up at a low passage (p2) and were cultured for 3 days prior to transfection. The sgRNA duplex was formed by combining 2.2 µL of 200 µM crRNA (5’-TGCAGGAAAGGCGACAGCTG-3’, IDT) with 2.2 µL of 200 µM ATTO-550 tracrRNA (IDT, cat# 1075927) and 5.6 µL nuclease-free IDTE Buffer pH 7.5 (IDT, cat# 11-01-02-02). The mixture was heated to 95°C for 5 min and then cooled to room temp. The ribonucleoprotein complex was formed by combining 0.5 µL sgRNA duplex with 0.5 µL of 36 µM Cas9 nuclease (IDT, cat# 1081058) per reaction and incubating at room temp for 20 min. MCF-7 cells were counted, washed with 1xPBS, and resuspended in Resuspension Buffer R (ThermoFisher, provided in Neon Transfection System Kit, cat# MPK1096) to $7.5x10^5$ cells/µL. For each electroporation, 1 µL of the Cas9/crRNA/tracrRNA ribonucleoprotein complex, 1 uL of single-stranded donor template (ssODN) at 100 µM (5’-CCCCCACTCAACAGCGTGTCTCCGAGCCCGCTGATGCTACTGCACCCGCCGCCG CAATTGGCGCCGTTCCTGCAGCCCCACGGCCAGCAGGTGCCCTACTACCTGGAG

AACGAGCCCAGC-3'), 2 µL of 10.8 µM Alt-R Cas9 Electroporation Enhancer (IDT, cat# 1075915), and 8 µL of MCF-7 cell suspension at $7.5x10^5$ cells were combined. Cells were electroporated with 10 µL of the transfection mixture using the Neon Transfection System (ThermoFisher, cat# MPK5000) with a 10 µL tip using a pulse voltage of 1250V, pulse width of 20 ms, and pulse number of 2. As a control, MCF-7 cells were transfected with Cas9 and the ssODN but no sgRNA. Cells were allowed to recover for 7 days in a 12-well plate containing DMEM without penicillin/streptomycin (pen/strep) before continuing culturing in DMEM with 1% pen/strep. Both S118A and control cells were expanded, and genomic DNA was isolated using the QIAGEN DNeasy Blood & Tissue Kit (QIAGEN, cat# 69506). Isolated DNA was tested for insertion of the S118A mutation by PCR with primers flanking the mutation site (fwd: 5'-GCCAACGCGCAGGTCTA-3'; rev: 5'-AAGCGCCCCGTGTTTATTTG-3') and a subsequent diagnostic digest with SfoI (NEB, cat# R0606S). Insertion of the ER S118A mutation introduces a novel SfoI digestion site. To isolate clones, cells were sparsely seeded in a 10 cm plate (100-1000 cells/plate) containing a 1:1 ratio of fresh DMEM and 0.22 µm filtered MCF-7 conditioned media. Clones were picked, expanded, and screened for the ER S118A mutation as described above.

Sequencing MCF-7 S118A knock-in cell lines

Sequencing of the mutation site within clones of interest was performed by first amplifying a 2,539 bp region around the mutation site (Fwd: 5'-CGTCCTCCAGCACCTTTGTA-3'; Rev: 5'-TCCCCTCTGTATTTGACCAGG-3'). Samples were run on a 1% agarose gel and the major band at ~2.5 kb was excised and purified using the QIAGEN DNeasy Blood &

Tissue Kit (QIAGEN, cat# 69506). The isolated DNA was then sent to the University of Wisconsin Biotechnology Center for Sanger sequencing using the forward primer from the SfoI digestion (5'-GCCAACGCGCAGGTCTA-3')

GRHL2 siRNA Transfection

MCF-7 cells were stripped for two or three days as described in cell culture procedure prior to siRNA transfection. For each transfection, 10 pmol GRHL2 siRNA (Santa Cruz, sc-77606) or control siRNA (QIAGEN AllStars negative control, cat# SI03650318) was added to 100 µL Opti-MEM (ThermoFisher, cat# 11058021). Additionally for each transfection, 3 µL Lipofectamine 2000 (ThermoFisher, cat# 11668027) was added to 100 µL OPTI-MEM. The siRNA and Lipofectamine solutions were added in equal volumes and allowed to sit at room temperature for 20 minutes. Cells were trypsanized and seeded into a 12-well plate at a concentration $2.0x10^5$ cells per well using phenol-red free DMEM supplemented with 10% charcoal dextran stripped FBS and no penicillin/streptomycin. 200 µL of the siRNA/Lipofectamine solution was added directly to the suspended cells in the 12-well plate and mixed well by pipetting up and down. Transfected cells were incubated at 37°C and 10% CO_2 for 24 hours. Cells were then treated according to their respective experiments.

Statistical Methods

For qPCR, p-values were calculated using a two-sided Student's t-test. For comparisons of distributions, a chi-squared test was performed. P-values less than 0.05 were considered statistically significant.

Table 6-1: Primary Antibodies Used for Western Blots

Target	Vendor	Catalog	Concentration
AR	Cell Signaling	#5153	1:1000
E-Cadherin	BD Biosciences	610181	1:1000
ER	Santa Cruz	sc-543	1:1000 – 1:2500
GRHL2	Sigma	HPA004820	1:1000
pS104/106-ER	Cell Signaling	#2517	1:1000
pS118-ER #1	Cell Signaling	#2511	1:1000
pS118-ER #2	Abcam	ab32396	1:1000
pS118-ER #3	Santa Cruz	sc-12915	1:1000
pS167-ER	Cell Signaling	#5587	1:1000
pS294-ER	Abcam	Ab207602	1:1000
ß-actin	Sigma	AC-15	1:5000

Table 6-2: Primer Sequences used for qRT-PCR

Gene	Name	Sequence (5' - 3')	T_a (°C)
CCND1	714f	ATCTACACCGACAACTCCATC	60
	901r	TGTTCTCCTCCGCCTCTG	
ERBB3	702f	GAGATGCTGAGATAGTGGTGAAG	57
	820r	GATGGTCTTGGTCAATGTCTGG	
ESR1	1518f	CCTGATGATTGGTCTCGTCTG	57
	1702r	GGCACACAAACTCCTCTCC	
GATA3	1429f	ACAAAATGAACGGACAGA	60
	1259r	GTGGTGGTCTGACAGTTC	
GREB1	fwd	ATTTGTTTCCAGCCCTCCTT	60
	rev	GTGGTAGCCGAGTGGACAAT	
GRHL2	1319f	GGGTCCTTGACATTGCCGAT	57
	1405r	GCCTCTTCATTCACGTCCCA	
PGR	1720f	GCTGTCATTATGGTGTCCTTAC	55
	1796r	GTAGTTGTGCTGCCCTTCC	
RPLP0	657f	GACAATGGCAGCATCTACAAC	60
	759r	GCAGACAGACACTGGCAAC	
TFF1	17f	CGCCTTTGGAGCAGAGAG	60
	166r	ACCACAATTCTGTCTTTCACG	
VIM	1015f	ACACCCTGCAATCTTTCAGACA	60
	1090r	GATTCCACTTTGCGTTCAAGGT	
ZEB1	fwd	GTGGCGGTAGATGGTAAT	60
	rev	CTGTTTGTAGCGACTGGA	

Table 6-3: Primer Sequences used for ChIP-qPCR

Location	Fwd/Rev	Sequence	T_a (°C)
BRIP1 +140 kb	Fwd	CAATGTGCACCACCAGCAAA	60
	Rev	GAATGTGGGCATGACCTCCC	
E2F6 -32 kb	Fwd	CTGATCTGGGGCACACTGAC	60
	Rev	CAGAGGAGACAAAACGGGCA	
GREB1 -1732	Fwd	AGGGCAGAGCTGATAACGTC	60
	Rev	AGAATGACCCAGTTGCCACA	
TFF1 +8 kb	Fwd	GTTTACTGCGCATTTCCGGG	60
	Rev	AGGCTGACAACTCCTTCAGC	
TFF1 -200	Fwd	GCTGATAGACAGAGACGACATG	60
	Rev	CTTCTGCAGTGAGTACAGTATTTACC	
MAFTRR -6 kb	Fwd	GTCAGCTGAGACAAGAGTGGAA	60
	Rev	GTTACCACGAATGGCAAAGC	
PFDN4 -12 kb	Fwd	GGACGTCTTTGATGCCTTGTC	60
	Rev	AAGCCCCAGGACATGAGGTA	
SETBP1 +81 kb	Fwd	GACCTGGCCCCATTGTTGTA	60
	Rev	CCAGCTGGCCCAGTTCATTA	

Appendix A:

Ubiquitlyation of Nuclear Receptors: New Linkages and Therapeutic Implications

This work has been published in:

Helzer KT, Hooper C, Miyamoto S, Alarid ET. Ubiquitylation of nuclear receptors: new linkages and therapeutic implications. *J. Mol. Endocrinol.* 2015;54(3):R151-167.

Abstract

The nuclear receptor superfamily is a group of transcriptional regulators that controls multiple aspects of both physiology and pathology and are broadly recognized as attractive therapeutic targets. While receptor-modulating drugs have been successful in many cases, the discovery of new drug targets is still an active area of research, because resistance to nuclear receptor-targeting therapies remains a significant clinical challenge. Many successful targeted therapies have harnessed the control of receptor activity by targeting events within the nuclear receptor signaling pathway. In this review, we explore the role of nuclear receptor ubiquitylation and discuss how the expanding roles of ubiquitin might be leveraged to identify additional entry points to control receptor function for future therapeutic development.

Exploiting NR signaling to identify new avenues of drug discovery

Nuclear receptors (NRs) comprise a family of transcriptional regulators that control multiple physiological processes including growth, development, reproduction and metabolism through the control of gene expression (328,329). The founding member of the family, estrogen receptor-α (ERα), was identified via its high affinity binding to radiolabeled ligand, estradiol (15,330). Following the cloning of the glucocorticoid receptor (GR, NR3C1)(331,332), numerous other nuclear receptors were identified and combined into a superfamily composed of a total of 48 receptors in mammals including estrogen receptors α (ERα; NR3A1) and β (ERβ, NR3A2), thyroid hormone receptor (TR, NR1A1), progesterone receptor (PR, NR3C3), androgen receptor (AR, NR3C4), retinoic acid receptor (RAR, NR1B1-3), retinoic X receptor (RXR, NR2B1-3), Vitamin D receptor (VDR, NR1I1), Peroxisome proliferator-activated receptors (PPAR, NR1C1-3), and a number of orphan receptors with no known ligands (333). The receptors share a similar architecture consisting of an intrinsically disordered N-terminus, which in some receptors encodes a ligand-independent transactivation domain, a central DNA binding domain containing two zinc finger motifs, and a C-terminal ligand-binding domain (LBD). The LBD mediates multiple receptor functions including ligand-binding, dimerization, co-regulator interactions, and ligand-dependent transcriptional activation function. It is no surprise then that research has focused largely on the LBD and the modulation of receptor actions through both endogenous and synthetic ligands (334,335).

Originally, NRs were considered a relatively simple signal transduction pathway in which receptors sensed the activating signal and directly mediated the response in the nucleus through direct DNA binding. Though fundamentally correct, the broadening

knowledge of components in the nuclear receptor activation mechanism have greatly expanded the model and simultaneously expanded the opportunity to control receptor function. In the contemporary model for ligand-activated NRs, ligand binding releases inactive receptors from heat shock protein complexes and alters receptor conformation to facilitate the recruitment of multi-protein transcriptional complexes. The activated NR transcriptional complex can include co-regulators (activators and repressors), chromatin modifying and remodeling complexes and components of the basal transcriptional machinery. To date, over 300 co-regulators have been identified (336, www.nursa.org). The addition of a temporal component to receptor transcriptional complexes further expands the potential combinatorial regulation of receptor function (145,337). Advances in the NR field made through the dissection of the molecular events regulating receptor function have contributed significantly to the drug discovery tool box which now includes agents targeting receptors at multiple levels including co-activator interactions (338–340), dimerization, subcellular localization (341) and DNA binding (342–345).

Post-translational modification (PTM) is another regulatory mechanism governing NR function. PTMs represent an important cross-talk mechanism for other signaling pathways and also contributes directly to receptor transcriptional function. In the case of ERα, all domains of the receptor can be phosphorylated in response to ligand and/or growth factor cascades (51,52,59,143,208,226,346–348). Studies in breast cancer cell models have demonstrated that phosphorylation can impact multiple aspects of receptor function including protein stability, dimerization, DNA binding, and co-activator preferences (66,78,135,138,139,143,349–351). ERα is also subject to other modifications including acetylation (35), methylation (36,37), SUMOylation (40,352), and

palmitoylation (353). Readers are referred to a recent comprehensive review of ERα post-translational modifications (41). Importantly, ERα and other NRs are targets of ubiquitylation, a post-translational modification that closely interfaces receptor protein status and transcriptional function at multiple levels of the receptor signaling pathway.

Nuclear Receptor Ubiquitylation

Ubiquitin is a small 76-residue protein that modifies target substrates by covalent attachment of its C-terminal carboxyl group to a lysine residue on the target protein substrate (i.e., "ubiquitylation"). The specificity of ubiquitylation is controlled by three classes of enzymes (354–356). E1 activating enzymes bind two ubiquitin moieties through a cysteine residue in an ATP-dependent mechanism, creating a high-energy thioester bond. The E1 enzyme then transfers the ubiquitin to an E2 conjugating enzyme, also through a thioester linkage. Lastly, the E3 ubiquitin ligases transfer ubiquitin from the E2 conjugating enzyme, directly or indirectly, to the substrate. There are two classes of ubiquitin ligases which determine the transfer mechanism of ubiquitin onto the target substrate. Homologous to E6-AP C-terminus (HECT) ligases accept ubiquitin from E2 conjugating enzymes and transfer ubiquitin to the substrate, whereas Really Interesting New Gene (RING) finger ligases facilitate E2-substrate interaction but do not physically transfer ubiquitin (357). To date, two E1 activating enzymes (UBA1, UBA6), 35 E2 conjugating enzymes, and over 600 E3 ligases have been reported in humans (357–362). The most well-characterized ubiquitin modification is polyubiquitylation of substrates via K48 linkages. This polyubiquitin chain is formed when the lysine 48 residue of one ubiquitin is attached to the C-terminal glycine residue of another. In general, K48-linked polyubiquitin chains target substrate for degradation by the 26S proteasome. Moreover,

it was the discovery of proteasome-dependent degradation of NRs that identified ubiquitylation of receptors as a potential regulatory module.

Early studies investigating the degradation of NRs via the proteasome showed clear evidence that the ubiquitin proteasome pathway mediated the down-regulation of many NRs upon binding of their respective ligands. The first studies investigating the ubiquitin proteasome pathway in NR biology showed that proteasome inhibitors disrupted ligand-mediated decreases in ERα protein (99–101). Pulse-chase analysis showed that the half-life of endogenous ERα protein in pituitary cells was reduced from 3 hours to 1 hour upon treatment with 17β-estradiol and this estradiol-induced degradation could be prevented with proteasome inhibitors (99). *In vitro* studies revealed that ubiquitin-activating E1 enzymes (UBA1) as well as ubiquitin-conjugating E2 enzymes (UbcH5 and UbcH7) could result in accumulation of high molecular weight conjugates (100). Ligand-dependent ubiquitylation of endogenous ERα was demonstrated by Wijayaratne and McDonnell (287). Subsequently, it was observed that retinoic acid receptor γ2 (RARγ2) and retinoic acid receptor α (RARα) were down-regulated in response to their ligand, all-trans-retinoic acid, and the down-regulation was blocked by proteasome inhibitors, MG132 and lactacystin (308,363). Thyroid hormone receptor (TR), glucocorticoid receptor (GR), and mineralocorticoid receptor (MR) were also found to be downregulated in response to ligand binding via a similar pathway (364–366). These studies, along with many others, pointed to the ubiquitin-proteasome pathway as one of the key regulators of NR protein turnover. Interestingly, the studies with proteasome inhibitors also implicated the ubiquitin-proteasome pathway in the regulation of transcription (321,367). It should be noted, however, that the 26S proteasome degradation pathway is not unique

to NRs, and inhibition of such a vital cellular function can lead to both inhibition and activation of other signaling pathways, production of reactive oxygen species, and induction of apoptosis (368–370). The role of ubiquitylation and proteasome activity in transcription is further complicated as NR turnover can be disassociated from transcriptional activity in certain contexts (78,308,371), and proteasome inhibitors can also affect NR expression at the transcriptional level and indirectly lead to downstream changes in NR target gene expression (372,373). Due to the confounding effects of proteasome inhibitors, the study of NR degradation shifted in the past decade to the identification of E3 ligases responsible for targeting NRs for ubiquitin-mediated proteasomal degradation. Considering the relative specificity between ubiquitylation machinery and their substrates, there is potential for the development of new drugs targeting these interactions which could alter the transactivation or degradation of NRs. Table 1 provides a summary of known E3 ligases implicated in the control of NR proteins. The following section will focus on three ligase types that are implicated in the control of multiple NR's as representatives of conserved activities within the NR signaling cascade (Figure 1).

CHIP ligase

One of the most common E3 ligases involved in the ubiquitylation and degradation of NRs is the Carboxyl-terminus of Hsc70-interacting protein (CHIP; Table 1). CHIP is a tetratricopeptide repeat (TPR) containing protein that associates with heat shock proteins to cause a decrease in ATPase activity. This association ultimately decreases the efficiency of the chaperone and impairs its function (374). Many NRs are held stable in their unliganded state by chaperone complexes which include Hsp70 and Hsp90 (375).

CHIP interacts with the TPR site of Hsp90 and incorporates itself into the NR-Hsp90 heterocomplex causing remodeling that favors degradation of the Hsp90 substrate. In the context of specific NRs, CHIP can ubiquitylate GR and directly target it for degradation by interacting with the S5a subunit of the proteasome (376). Disruption of Hsp90 using molecular inhibitors has also been shown to cause down-regulation of ERα and PR in a proteasome-dependent manner (377–379). The key observation that CHIP–null cells retain the ability to down-regulate ERα in the presence of estrogen and that CHIP depletion by siRNA caused ERα to accumulate in the absence of ligand identified CHIP as an important ligase in basal ligand-independent NR protein stability (380). These studies identified CHIP as an essential ligase in basal ligand-independent NR protein stability. Other NRs also appear to be poly-ubiquitylated by CHIP, including ERβ, AR, GR, and MR suggesting that ubiquitylation of NRs by CHIP is a conserved mechanism (381–384). Therefore, the control of NRs by CHIP represents specific ligase activity at one of the earliest steps in NR signaling, controlling basal NR expression and receptor availability prior to ligand binding, nuclear localization, and transcriptional activation.

Mdm2 and E6-AP ligases

The RING finger-containing E3 ligase mouse double minute 2 (Mdm2) has been implicated in the control of transactivation and turnover for many NRs, including AR, ERα, ERβ, and GR (Table 1, references therein). Mdm2 is most commonly known for its role in the control of the tumor suppressor p53 via ubiquitylation and subsequent degradation (385). Early studies analyzing the interaction between Mdm2 and p53 revealed that phosphorylation of Mdm2 by the serine/threonine protein kinase Akt was required for both translocation of Mdm2 to the nucleus and its ability to ubiquitylate p53 (43,386). Akt was

also found to phosphorylate AR which resulted in the suppression of AR target genes as well as the inhibition of AR-mediated apoptosis (288). These results prompted the hypothesis that phosphorylation of AR by Akt leads to Mdm2-dependent, ubiquitin-proteasome degradation of AR. Indeed, AR was shown to be ubiquitylated by Mdm2 and subsequently degraded by the proteasome and this action was dependent on the E3 ligase activity of Mdm2 as well as the phosphorylation of Mdm2 by Akt. Phosphorylation of AR at S515 by cdk7 kinase also leads to the recruitment of Mdm2 (387). Impairment of AR phosphorylation, using an S515A mutant, blocks AR transactivation through defects in the recruitment of transcriptional machinery. Interestingly, this mutant caused the preferred recruitment of CHIP over that of Mdm2 (387). These results suggest that phosphorylation is a common step necessary for AR ubiquitylation and degradation.

Mdm2 has also been implicated in other NR pathways. Like AR, Mdm2 is recruited to ERα and β complexes associated with phosphorylation (78,284,388). Overexpression of Mdm2 in MCF-7 breast cancer cells stimulates growth while knockdown of Mdm2 inhibits growth with varying stimuli (371,389,390). In the case of ERβ, Mdm2 works in concert with the coregulator, CREB-Binding Protein (CBP) to form a complex that results in ubiquitylation and ultimate degradation of ERβ. Degradation of ERα is also increased with Mdm2 over-expression, but in the case of ERα, degradation mediated through Mdm2 requires p53 (371). The requirement of p53 in ubiquitylation have also been shown for GR and Hdm2 (human homologue of Mdm2) (391). Interestingly, unlike ERβ, the Mdm2-CBP complex was unable to target ERα for degradation. These observations suggest that receptor ubiquitylation by Mdm2 is mediated through complex formations with

coactivators, and that the specificity of the coactivator complex dictates the receptor target.

The HECT E3 ligase E6-AP is itself a NR coactivator and can direct the turnover of PR and ER (392). Like Mdm2, E6-AP ubiquitylation of ERα and β is dependent on phosphorylation of receptors (109,284). Thus, ubiquitin ligases can direct receptors for degradation directly via recognition of post-translational modifications on the receptor, as in the study of Mdm2 and AR, or indirectly via co-regulator complexes. This presents the possibility of utilizing the ubiquitin system to target subpopulations of receptor based on post-translational modifications or coactivator complexes (Figure 1).

BRCA1/BARD1 and RNF31 (HOIP) ligases

While most E3 ligases associated with the ubiquitylation of NRs involve the addition of polyubiquitin chains, BRCA1/BARD1 and RNF31 have both been shown to monoubiquitylate ERα (393,394). Monoubiquitylation results in the addition of one ubiquitin moiety on a given lysine residue on the receptor, which increases the molecular weight of the protein by ~8kD. Multiple lysine residues on the receptor can also be simultaneously monoubiquitylated (i.e. multiubiquitylation) giving the appearance of a “ladder” or “smear” which are often difficult to distinguish from a polyubiquitin ladder or smear in Western blot analysis. BRCA1 contains a RING domain at its N-terminus and its heterodimerization with BARD1 is required for its E3 ligase activity. Mutations in the BRCA1 RING domain that abolish the BRCA1/BARD1 E3 activity have been identified as cancer-predisposing (395), indicating that the E3 ligase activity is required for its function as a tumor suppressor. BRCA1 has been shown to regulate ERα and PR transcriptional activity and both of these NRs were shown to be ubiquitylated by BRCA1/BARD1

(393,396). Interestingly, *in vitro* studies revealed that ERα was monoubiquitylated by BRCA1/BARD1. This monoubiquitylation was dependent on the E3 ligase activity of BRCA1 and cancer pre-disposing BRCA1 mutations (C61G and C64G) were shown to abolish the ability of BRCA1 to monoubiquitylate ERα. The site of monoubiquitylation on ERα was identified through mass spectrometry to be K302, however, the K302A ERα mutant was still monoubiquitylated *in vitro*. The adjacent lysine residue, K303, can be targeted for monoubiquitylation in lieu of K302 (393). The function of ERα monoubiquitylation by BRCA1/BARD1 is still unknown, although is it hypothesized to play a role in inhibition of ERα transcriptional activity (397).

More recently, PR has been shown to be a substrate for BRCA1/BARD1-mediated ubiquitylation. Unlike ERα, PR was observed to be polyubiquitylated rather than monoubiquitylated (396). Overexpression of wild-type BRCA1 lead to increased polyubiquitylation of PR followed by degradation by the 26S proteasome whereas overexpression of a mutant BRCA1 deficient in its E3 ligase activity could not induce polyubiquitylation. However, the distinction between polyubiquitylation and multiubiquitylation was not resolved leaving the possibility that PR is multiubiquitylated by BRCA1/BARD1. BRCA1 was also shown to effect PR transcriptional activity when overexpressed (396).

RNF31—also known as HOIP or ZIBRA—was originally identified as an upregulated gene in breast cancer (398). RNF31 has also emerged as a critical component of LUBAC (linear ubiquitin assembly complex) that is implicated in NF-κB signaling (see below). Knockdown of RNF31 decreased estrogen-dependent growth as well as cell cycle progression of MCF-7 breast cancer cells (394). These studies also

revealed a positive correlation between RNF31 levels and ERα levels, indicating that RNF31 could be stabilizing ERα. Overexpression of RNF31 caused an increase in ERα-mediated transcription as measured by luciferase assay while siRNA knockdown of RNF31 decreased ERα-mediated transcription. Importantly, these effects were shown to be dependent on the E3 ligase activity of RNF31. ERα was found to be monoubiquitylated by RNF31 suggesting that the monoubiquitylation of ERα is regulating its transcriptional activation (394). The control of ERα by RNF31 and BRCA1/BARD1 suggests that there may be other NRs regulated in the same manner via monoubiquitylation. Moreover, given that RNF31 can act in the context of the LUBAC E3 ligase complex, it is plausible that linear (or M1-linked) polyubiquitylation also plays a regulatory role in NR function (see below).

Polyubiquitin chains with non-classical ubiquitin linkages

With the exceptions above, most E3 ligases thus far revealed add polyubiquitin chains to their respective NR substrate and direct them to the proteasome for degradation; however, there are both degradative and non-degradative functions of ubiquitin, the latter of which is only beginning to be explored in the NR field. In addition to K48-linked polyubiquitin chains described above, an emerging “ubiquitin code” (399) exists at the forefront of the ubiquitin field wherein polyubiquitin chains can form through ubiquitin-ubiquitin linkages at six other lysine residue—K6, K11, K27, K29, K33, or K63—in addition to K48 (Figure 2). Each type of polyubiquitin chain has been shown to be present *in vivo* through analysis by mass spectrometry (400). Moreover, the N-terminal methionine residue of ubiquitin can serve as a target of ubiquitin to assemble polyubiquitin chains with "M1-linkage". Many substrates can also be multiubiquitylated at multiple

lysine sites. Finally, different types of ubiquitin configuration could occur in a single substrate (i.e., "mixed" ubiquitin linkages). The different linkages can regulate protein activities in different contexts, including cell signaling, cell cycle regulation, DNA damage repair, and lysosomal degradation, among others (Figure 2). The role for these less commonly studied forms of polyubiquitin linkages in NR regulation is a virtually untapped area of research.

A New Frontier in Ubiquitylation

One example of atypical ubiquitin regulation of NR function was reported for the AR. Xu et al. discovered that the androgen receptor (AR) was polyubiquitylated by the E3 ubiquitin ligase, RNF6, in prostate cancer cells (401). Identified as an interacting partner of AR in GST-pull down assays, RNF6 overexpression leads to an increase in polyubiquitylated AR without changes in AR protein levels. *In vitro* ubiquitylation assays revealed that RNF6 added K6- or K27-linked polyubiquitin chains to the AR. The authors went on to show that overexpression of RNF6 increased the recruitment of co-factors, specifically ARA54, to androgen response elements (AREs), which suggests the possibility that specific ubiquitin linkages may contribute to the specificity of coregulator complexes on DNA. This may have implications in hormone-refractory prostate cancer where RNF6 has been shown to be overexpressed (401).

The study of RNF6 in AR biology provides a proof of principle that atypical ubiquitin chains can contribute to NR transcriptional function, but the array of potential linkages that remain unexplored provides a great opportunity for the NR field to better characterize the functional role of different ubiquitin moieties in the cell. The ability to tightly regulate NR activation with ligand provides the potential to study the formation of different ubiquitin

linkages in a controlled biological setting. Lessons from the NF-κB field provide a valuable guide into the potential of the “ubiquitin code”. It is well documented that ubiquitin regulates multiple steps in the NF-κB signaling pathway using both degradative and non-degradative polyubiquitin chains (402). NF-κB signaling is initiated by activation of cell surface receptors, such as the tumor necrosis factor receptor (TNFR) family or the toll-like receptor (TLR) family, which leads to the activation of the IκB kinase (IKK) complex and subsequent phosphorylation and degradation of IκB (Inhibitor of κB). Degradation of IκB releases cytoplasmically sequestered NF-κB to migrate to the nucleus to initiate transcription. Coupling of the activated cell surface receptors to the IKK complex involves K63- and M1-linked polyubiquitin chains on adaptor proteins that are assembled on activated cell surface receptors. An E2, Ubc13/Uev1A, can specifically assemble K63-linked polyubiquitin chains in conjunction with multiple E3s belonging to the TRAF (TNFR associated factor) family, such as TRAF3 and TRAF6. As briefly mentioned above, LUBAC, composed of the catalytic HOIP (HOIL-1-interacting protein), regulatory HOIL-1L (heme-oxidized IRP2 Ub ligase-1L) and SHARPIN (SHANK-associated RH domain interacting protein in postsynaptic density) assembles M1-linked polyubiquitin chains. In a simplified model, these K63 and M1-linked polyubiquitin chains are recognized by TAB2/3 (TAK1 binding proteins 2 and 3), regulatory subunits of TAK1 (TGFβ-activated kinase 1), and NEMO (NF-κB essential modulator), a regulatory subunit of the IKK complex (403) to induce TAK1-dependent IKK activation. The active IKKβ then phosphorylates IκB which allows it to be recognized by β-TrCP E3 ligase, and polyubiquitylated with K48-linked polyubiquitin chains for targeting to the 26S

proteasome. In this example, the non-degradative polyubiquitin chains couple the cell surface receptor to key components of the activation mechanism.

The RNF6 study in AR suggests that non-degradative polyubiquitin chains may interface with NR activation mechanism at the level of co-activators, but other components of the NR signaling pathway (post-translational modifications, dimerization, DNA binding) are all potential interactions that have not yet been tested. Reciprocally, despite the great strides in the ubiquitin field toward uncovering and characterizing nondegradative ubiquitin roles, the currently limited functional assignments for the different ubiquitin linkages are likely an underestimation of its involvement in signaling. The limited studies of ubiquitin linkages in other biological contexts, including the NR field, is due in large part to the difficulty to detect and quantify proteins modified by specific ubiquitin linkages. Three major techniques for detecting different ubiquitin linkages include linkage specific antibodies, mutagenic analysis/genetic complementation, and mass-spectrometry that could be applied to the analyses of NRs.

Linkage-specific antibodies. Antibodies are available for K48, K63 or M1 linkages from commercial sources. The antibodies can be used in immunoprecipitation (IP)-Western analysis wherein a specific receptor of interest is immunoprecipitated under denaturing conditions followed by Western blot analysis using one of these linkage-specific ubiquitin antibodies to detect ubiquitin modified proteins via specific linkages. These approaches added to the understanding of linkage-specific roles in NF-κB signaling by providing a function for non-degradative ubiquitin chains (404,405) and DNA repair (406). Depending on the quality of the antibodies, a large amount of NR may need to be immunoprecipitated to be detected by these chain-specific antibodies.

Mutagenesis: Atypical ubiquitin linkages can also be implicated by overexpression or knockdown/knockout approaches to attenuate specific E2 (e.g. Ubc13 for K63 chains) and E3 components (e.g. LUBAC subunits for M1 chains). In addition, mutating specific lysines in ubiquitin to arginine and either over-expressing or genetically complementing these mutant species in cells has helped to define a cellular role for different ubiquitin linkages. For example, Xu et al. showed that the involvement of non-degradative K63-linked polyubiquitin chains in NF-κB signaling by generating a knock-in system to replace endogenous ubiquitin with a K63R mutant form (407). Interestingly, the use of this system allowed them to determine which stimuli (interleukn-1β, but not TNFα) required K63-linked ubiquitin chains in the activation of NF-κB. While transient expression analysis with ubiquitin mutants may be readily applicable, generation of stable cell lines harboring mutant ubiquitin replacing endogenous ubiquitin is exceedingly challenging due to the presence of multiple genes that encode ubiquitin and the impacts on cell viability when certain ubiquitin mutants (e.g. K48R) are employed (407).

Mass spectrometry: In the process of ubiquitylation, the C-terminal tail of ubiquitin is typically attached to a lysine residue on its target substrate. Trypsin digestion of the ubiquitylated protein will result in a peptide where by a Gly-Gly motif is attached to a Lys residue on the substrate. In a similar fashion, polyubiquitin chains formed through specific Lys residues on ubiquitin also result in unique peptide sequences surround the Lys linkage, which allow for detection of different ubiquitin linkages (408). For example, Nishikawa et al. have used this method to identify that BRCA-1/BARD-1 is able to generate K6-linked polyubiquitin chains *in vitro* (409). One of the major caveats to using mass-spectrometry is that it requires a large amount of ubiquitylated product for

identification. Thus, the majority of ubiquitylated substrates used for this technique have been generated from *in vitro* ubiquitylation of recombinant substrates. Moving into cells requires a large amount of starting material (e.g., 50-100 $x10^6$ cells or 20-50 150mm plates) (410,411), which is feasible for cell line studies but may be limiting in the analysis of primary cells isolated from animal sources or patient samples. The use of anti-di-Gly antibodies is useful in the enrichment of ubiquitylated peptide species following trypsin digestion to reduce the complexity of MS analysis (412).

The role of ubiquitin in NR function is just in its infancy and despite some of the current technical challenges, understanding how this protein modification regulates NR function may open new avenues of research and therapeutic design. There are many critical reagents being generated (e.g., antibodies that specifically detect different ubiquitin linkages (413,414)), new techniques being developed (e.g., advanced, sensitive and quantitative MS analyses (415–417)) and specific ubiquitin E2s, E3s and deubiquitinases (DUBs) that act on specific ubiquitin linkages (361,418,419). These advances have had tremendous impact on the NF-κB signaling field in particular and may in the future have a similarly profound impact on NR family of proteins.

Current and future ubiquitin-targeting therapeutics

The examples provided above indicate the potential for targeting multiple steps (protein folding, coactivator interactions, transcriptional function) in the NR signaling pathway via control of ubiquitylation. Clinical approaches in cancer therapy have thus far focused on inhibiting the 26S proteasome (420). Bortezomib (Velcade, PS-341) is a

general proteasome inhibitor that is FDA approved for the treatment of multiple myeloma. Second generation proteasome inhibitors have also been developed including Carfilzomib, which was approved in 2012 for multiple myeloma patients that are refractory to bortezomib therapy (421). While the preclinical data supports the efficacy of proteasome inhibitors in other cancer types, the results outside of hematological malignancies have been disappointing (422). Hence, efforts to more specifically target the ubiquitylation machinery and their substrates are underway. A glimpse into this potential with relevance to the NR field is provided by studies of SCF-Skp2 and $p27^{kip1}$. Skp2 is an F-box protein and component of the Skp1-Cullin1-F-Box (SCF) ubiquitin E3 ligase complex. It is overexpressed in human cancers and deregulation of Skp2 is implicated in progression through loss of control of cell cycle control and transcription (423–425). Skp2 ligase activity was shown to be dependent on phosphorylation which can be induced by 17β-estradiol, leading to ubiquitylation of $p27^{kip1}$ (426,427). The loss of nuclear $p27^{kip1}$ has been shown to occur in 17β-estradiol-induced type I endometrial carcinogenesis (426). Using a small molecule screen, specific agents have been identified that block Skp2-dependent ubiquitylation of $p27^{kip1}$, thus preventing its degradation. Treatment of 17β-estradiol-induced endometrial carcinoma cell lines with these small molecules resulted in increased levels of $p27^{kip1}$ along with decreased proliferation (428). These experiments demonstrate that alterations of E3 ligase activity using small molecule inhibitors could be a viable strategy for future therapeutic development. Parenthetically, ERα is also a target for Skp2 ligase activity and high Skp2 expression in human tumors correlates with loss of ERα-negative status (351). By extension, the regulation of Skp2 may have broader implications in ER-dependent

diseases. To date, three E3-targeting drugs have been approved by the FDA, and all three target the same enzyme, Cereblon (CRBN). CRBN is a part of the Cul4-Rbx1-DDB1-CRBN E3 ubiquitin ligase complex and the three drugs that target CRBN—thalidomide, lenalidomide, and pomalidomide; commonly referred to as IMiDs (immunomodulatory drugs)—bind to CRBN and promote the recruitment of substrates, including Ikaros (IKZF1) and Aiolos (IKZF3), which are subsequently ubiquitylated and degraded (429–432). Currently these drugs are approved for multiple myeloma therapy (433,434).

Therapeutic targeting of NRs have been successfully developed by capitalizing on the in-depth knowledge of NR signaling and transcription. Likewise, development of ubiquitin-targeting drugs requires knowledge of substrates and their role in the pathogenesis of disease. The marriage of NR and ubiquitin fields presents the opportunity to move two potential therapeutic approaches forward by dually capitalizing on this rapidly advancing field.

Figure A-1: A model of conserved roles of the ubiquitin proteasome pathway in NR signaling.

The primary function ascribed to ubiquitin in NR signaling is targeting receptors to the proteasome. Shown is a simplified NR signaling pathway in which receptor (R) is activated either by ligand or by kinases from growth factor or membrane bound NRs. Receptors held bound by heat shock proteins can be targeted for degradation by the 26S proteasome following ubiquitylation by E3 ligases, such as CHIP. Upon binding ligand, receptors undergo dimerization (homo- and heterodimerization are not distinguished here) and can be decorated by multiple post-translational modifications including phosphorylation, shown as a star. Phosphorylated receptors can be directly recognized by E3 ligases, ubiquitylated, and directed to the proteasome for degradation. Alternatively, post-translationally modified receptors can be incorporated into active transcriptional complexes of variable co-regulator components, represented by a grey multiprotein-complex. The make-up of the co-regulator/receptor complexes can recruit E3 ligases that direct ubiquitylation and degradation of the transcriptional complex. In the degradative pathways, post-translational modifications and co-regulator complexes guide E3 targeting that allows subpopulations of receptors to be selectively degraded in cells.

Figure A-1

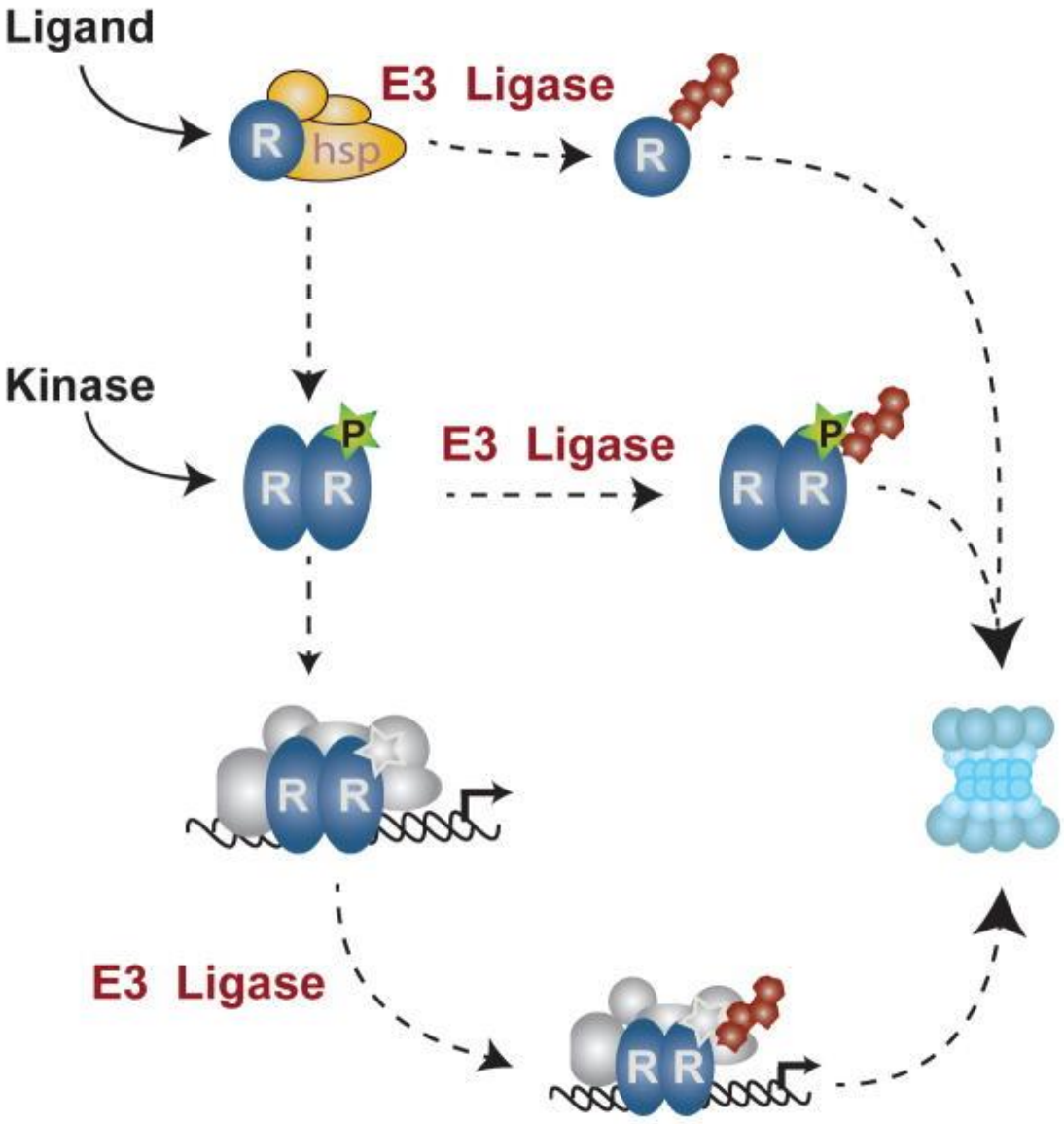

Ligand
E3 Ligase
R
hsp
R
Kinase
P
R R
E3 Ligase
P
R R
R R
E3 Ligase
R R

Figure A-2: A schematic of the cellular processes in which different polyubiquitin chain species have been implicated based on linkage type.

The different types of ubiquitylated species are represented as cartoons in the left-hand column. Polyubiquitin chains can be organized into a “closed” or “open” conformation based solely on the type of linkage that connects them. A (*) symbolizes that structural data is currently unavailable for these linkages; however, modeling of these structures predicts the conformation of each chain type. Cellular roles are determined based on the identification of each chain type in a specific cellular process. Currently the function of many of these chains is still unknown. This list is not meant to be comprehensive but rather to highlight the many diverse roles of ubiquitin.

Figure A-2

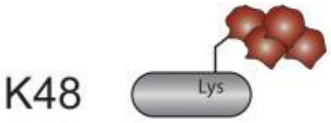
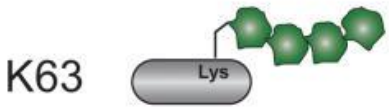
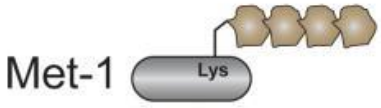
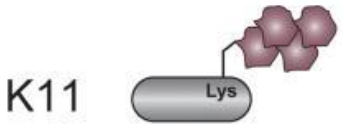
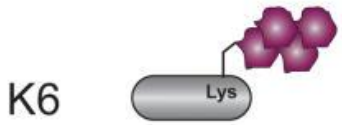
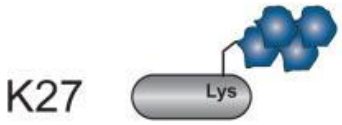
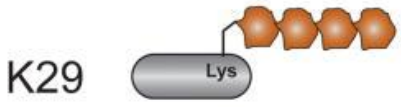
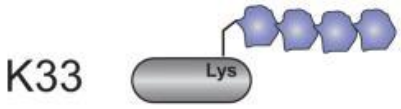
Linkage Type	Cellular Role	References
 K48	Proteosomal Degradation	(415,435)
 K63	Signal Transduction Translation Lysosomal Degradation Endosomal Trafficking DNA Damage Response	(407,436–438)
 Met-1	Signal Transduction	(405,439–441)
 Mono	Lysosomal Degradation DNA Damage Response DNA Repair	(406,442,443)
 K11	Cell Cycle Regulation Proteasomal Degradation Assembled by APC/C ERAD Membrane Trafficking TNFα signaling	(400,444–448)
 K6	DNA Repair DNA Damage Response AR transcription	(400,409,449–452)
 * K27	Mitochondrial maintenance Mitophagy T-cell (Treg) development Assembled by parkin AR transcription	(400,453–456)
 * K29	Lysosomal degradation Ubiquitin-fusion degradation (UFD	(457–460)
 * K33	Reduces T-cell activation	(459,460)

Figure A-3: A hypothetical model of non-degradative ubiquitin in NR signaling.
Non-degradative polyubiquitin chains can serve as protein assembly scaffolds. Examples shown in the middle and right panels are assemblages described in NF-κB signaling and DNA damage repair. Based on these models, we speculate the potential for non-degradative ubiquitin chains providing similar scaffolding in NR signaling, bringing together tertiary complexes with enzymatic activity to post-translationally modify receptor or other coregulator proteins to affect transcription. Such scaffolds could also bring other ubiquitin-binding domain (UBD) proteins into the NR complex without directly binding to NR.

Figure A-3

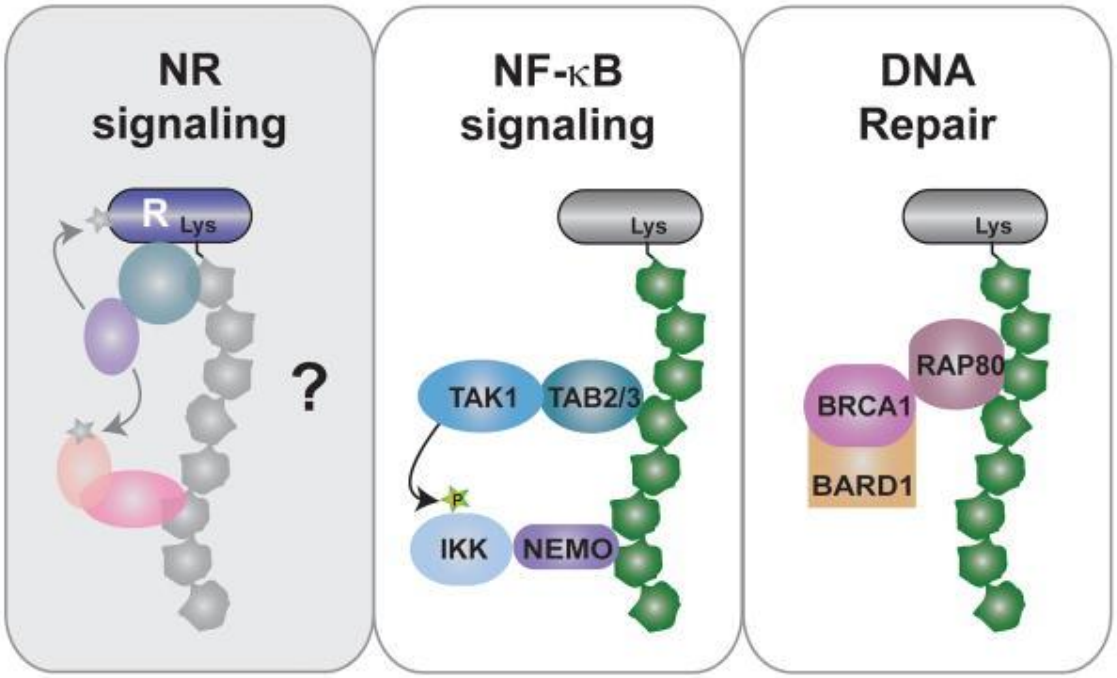

NR
signaling
R
Lys
?
NF-κB
signaling
Lys
TAK1
TAB2/3
P
IKK
NEMO
DNA
Repair
Lys
RAP80
BRCA1
BARD1

Table A-1: E3 Ligases Involved in NR Ubiquitylation

Nuclear Receptor	E3 Ligase	Class of Ligase	Type of Ub	References
Estrogen Receptor α	E6-AP	HECT	poly	(252,293)
	CHIP	RING	poly	(461)
	Mdm2	RING	poly	(371)
	BRCA1/BARD1	RING	mono	(393)
	EFP	RING	poly K48	(462)
	SPOP	RING	poly	(463)
	RBCK1	RING	?	(464)
	CUEDC2	?*	?	(465)
	Skp2	RING	poly	(351)
	RNF31	RING	mono	(394)
Estrogen Receptor β	CHIP	RING	poly K48	(383)
	E6-AP	HECT	poly	(284)
	Mdm2	RING	poly	(388)
Androgen Receptor	Mdm2	RING	poly	(47)
	CHIP	RING	poly	(381)
	RNF6	RING	poly-K6 or -K27	(400)
	Siah2	RING	poly K48	(466)
	UBR1	RING	?	(467)
	Skp2	RING	poly	(468)
Glucocorticoid Receptor	Hdm2	RING	poly	(391)
	CHIP	RING	poly	(382)
	FBXW7	RING	poly	(469)
	UBR1	RING	?	(467)
Progesterone Receptor	CUEDC2	?*	?	(470)
	BRCA1/BARD1	RING	poly	(396)
Retinoic Receptor α	FLRF (Rnf41)	RING	?	(471)
Mineralocorticoid Receptor	CHIP	RING	poly	(384)
PPARγ	Siah2	RING	poly	(472)
Thyroid Hormone Receptor	?			(364)
Estrogen-Related Receptors	Parkin	RING	poly	(473)

Table A-2: NR Coregulators that Contain RING Finger Domains

Coregulator	Ring Finger Designation	Receptor	References
ARNIP	RNF199	AR	(474)
BRCA1	RNF53	ER, PR	(396,475–477)
EFP	RNF147	ER	(462,478)
MAT1	RNF66	ER, PPARγ	(479,480)
RNF8	RNF8	RXR	(481)
RLIM	RNF12	ER	(482)
TIF1	RNF82, RNF96	RXR, RAR, ER, VDR, AR, TR	(483–485)
SNURF	RNF4	AR, GR, PR, ER	(486–488)

Abbreviations: RNF, Ring Finger protein; ARNIP, Androgen receptor N-terminal Interacting Protein; BRCA1, breast cancer type 1 susceptibility protein; EFP, Estrogen-responsive Finger Protein; MAT1, Menage-a trois homologue 1; RLIM, Ring finger protein, LIM Domain Interacting, TIF1, Transcriptional intermediary factor, SNURF, SNRPN Upstream Reading Frame, see references.

Appendix B:

17ß-Estradiol and ICI182,780 Differentially Regulate STAT5 Isoforms in Female Mammary Epithelium, With Distinct Outcomes

This appendix contains select text and data produced for the following report in collaboration with Dr. Fatou Jallow and Dr. Linda Schuler:

Jallow F, Brockman JL, Helzer KT, Rugowski DE, Goffin V, Alarid ET, Schuler LA. 17β-Estradiol and ICI182,780 Differentially Regulate STAT5 Isoforms in Female Mammary Epithelium, With Distinct Outcomes. *J Endocr Soc* 2018;2(3):293–309.

Prolactin (PRL) and estrogen cooperate in lobuloalveolar development of the mammary gland and jointly regulate gene expression in breast cancer cells *in vitro*. Canonical PRL signaling activates STAT5A/B, homologous proteins that have different target genes and functions. Although STAT5A/B are important for physiological mammary function and tumor pathophysiology, little is known about regulation of their expression, particularly of STAT5B, and the consequences for hormone action. In this study, we examined the effect of two estrogenic ligands, 17*β*-estradiol (E2) and the clinical antiestrogen, ICI182,780 (ICI, fulvestrant) on expression of STAT5 isoforms and resulting crosstalk with PRL in normal and tumor murine mammary epithelial cell lines. In all cell lines, E2 and ICI significantly increased protein and corresponding nascent and mature transcripts for STAT5A and STAT5B, respectively. Transcriptional regulation of STAT5A and STAT5B by E2 and ICI, respectively, is associated with recruitment of estrogen receptor alpha and increased H3K27Ac at a common intronic enhancer 10 kb downstream of the *Stat5a* transcription start site. Further, E2 and ICI induced different transcripts associated with differentiation and tumor behavior. In tumor cells, E2 also significantly increased proliferation, invasion, and stem cell-like activity, whereas ICI had no effect. To evaluate the role of STAT5B in these responses, we reduced STAT5B expression using short hairpin (sh) RNA. shSTAT5B blocked ICI-induced transcripts associated with metastasis and the epithelial mesenchymal transition in both cell types. shSTAT5B also blocked E2-induced invasion of tumor epithelium without altering E2-induced transcripts. Together, these studies indicate that STAT5B mediates a subset of protumorigenic responses to both E2 and ICI, underscoring the need to understand regulation of its expression and suggesting exploration as a possible therapeutic target in breast cancer.

In this study, I assisted in analyzing the occupancy of ER at the *Stat5a* and *Stat5b* loci in response to E2 and ICI 182,780. First, to investigate estrogenic regulation of STAT5 isoforms, the effect of E2 and ICI on STAT5A/B transcription in TC11 cells was examined. Initial analysis of *Stat5a* and *Stat5b* mRNA showed that E2 and ICI had differential effects on the two Stat5 isoforms. E2 induced *Stat5a*, whereas ICI induced *Stat5b* mature transcripts (Figure B-1A, Figure B-1B), consistent with the observed changes in levels of these proteins. To assess whether these are direct transcriptional effects, nascent RNA was evaluated. Similar to the mRNA, E2 induced nascent *Stat5a* transcripts, whereas ICI induced nascent *Stat5b* transcripts (Figure B-2A, Figure B-2B), confirming a contribution for transcriptional regulation. Therefore, we investigated the ability of these ligands to recruit ERα to genomic sites near the *Stat5a/Stat5b* locus. ER ChIP-sequencing data from Miranda *et al.* (199) (GEO dataset GSE46123) on a mouse mammary carcinoma cell line treated with E2 for 30 minutes identified five ER occupancy sites in the *Stat5a/Stat5b* locus within 50 kb of either the *Stat5a or Stat5b* coding regions (Figure B-1). We performed ChIP-qPCR on cells treated with Veh, E2, or ICI for 30 minutes and observed a robust increase in ERα occupancy at site *Stat5a* 10k after either E2 or ICI treatment (Figure B-2C). The remaining sites tested (*Stat5b* 51k, *Stat5b* 34k, *Stat5a* 6k, *Stat5a* 27k) showed no substantial change in ER occupancy compared with vehicle upon treatment with either E2 or ICI (Figure B-2C). H3K27 acetylation (H3K27Ac), a marker associated with functional ER*α* binding sites (194,264), was similarly increased by both E2 and ICI at the *Stat5a* 10k site (Figure B-2D). These data show that both E2 and ICI can recruit ER and induce consequent H3K27 acetylation at a common intronic site 10k from the *Stat5a* transcription start site, with distinct effects on STAT5A and 5B expression.

Figure B-1: Location of ER occupancy sites around the *Stat5a/5b* genomic locus.

ER ChIP-seq data from Miranda *et al.* (199) was analyzed for locations of ER occupancy around the *Stat5a* and *Stat5b* locus. The transcriptional start site for *Stat5a* and the upstream start site for *Stat5b* (exon 1b) are shown. ER occupancy sites have been labeled by their proximity to the nearest transcription start site of either *Stat5a* or *Stat5b*.

Figure B-1

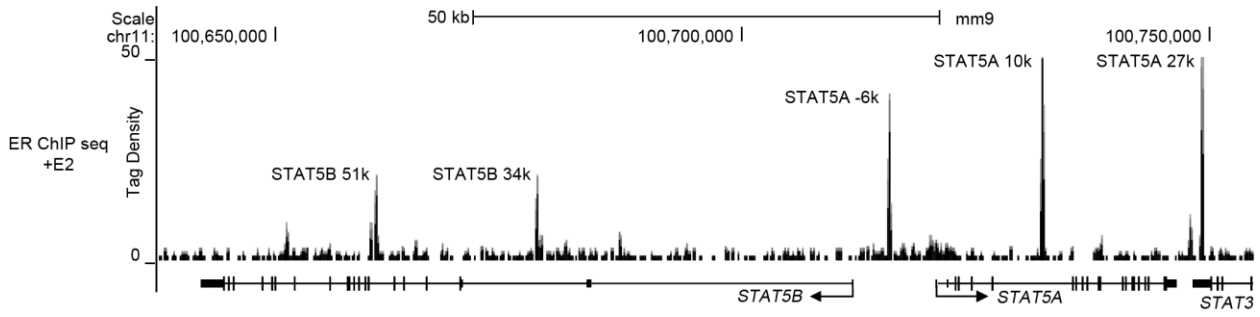
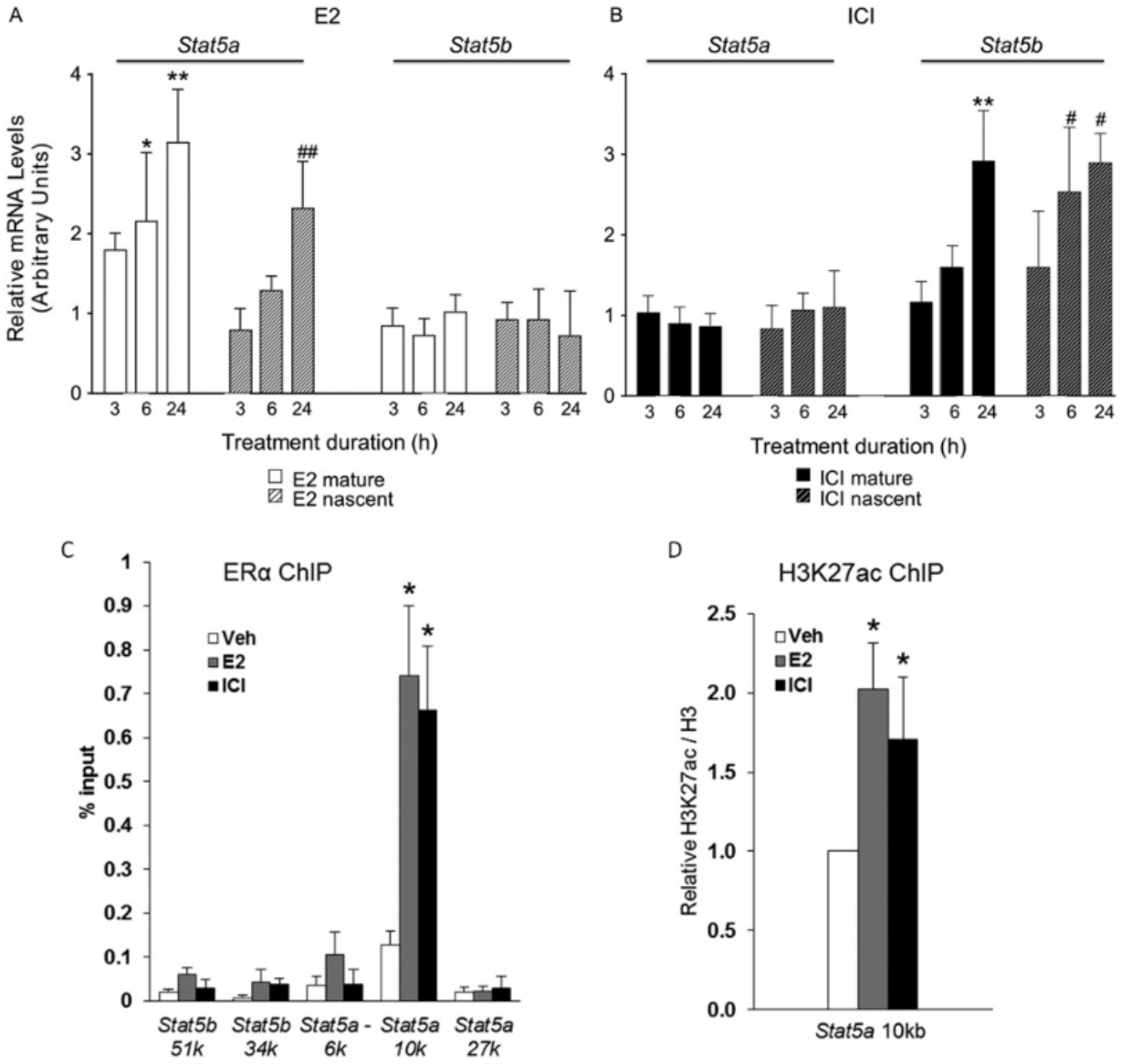

Scale
chr11:
100,650,000
50 kb
mm9
100,700,000
100,750,000
50
0
Tag Density
ER ChIP seq
+E2
STAT5B 51k
STAT5B 34k
STAT5A -6k
STAT5A 10k
STAT5A 27k
STAT5B
STAT5A
STAT3

Figure B-2: Estrogen and ICI recruit ER to the *Stat5* genomic locus in TC11 cells.

(A, B) Serum-starved TC11 cells were treated with Veh, E2, or ICI for 3, 6, or 24 hours, as shown. Levels of mature (solid bars) and nascent (stripped bars) *Stat5a* and *Stat5b* transcripts were determined by qRT-PCR and normalized to 18S RNA. Mean changes relative to levels in Veh-treated cells ± SD are shown. N = 3. Significant differences were determined by two-way ANOVA, followed by the Tukey posttest (differences in levels of mature transcripts from Veh, **P* < 0.01, ***P* < 0.0001; differences in levels of nascent transcripts from Veh, #*P* < 0.001, ##*P* < 0.0001). (C) Serum-starved TC11 cells were treated with Veh, E2, or ICI for 30 minutes; ChIP was performed as described in the Materials and Methods. Mean ± standard error of the mean is shown. N = 3. Significant differences from Veh were determined by one-way ANOVA, followed by the Tukey *post hoc* test (**P* < 0.05) using percent input values. (D) H3K27Ac ChIP at the *Stat5a* 10 kb site. TC11 cells were treated as described in panel C. H3K27ac data were normalized to total H3 followed by normalization to Veh. Statistical differences from Veh were calculated by Wilcoxon signed-rank test, N = 6; mean ± standard error of the mean (**P* < 0.05).

Figure B-2


A
E2
Stat5a
Stat5b
Relative mRNA Levels (Arbitrary Units)
Treatment duration (h)
E2 mature
E2 nascent
B
ICI
Stat5a
Stat5b
Treatment duration (h)
ICI mature
ICI nascent
C
ERα ChIP
Veh
E2
ICI
% input
Stat5b 51k
Stat5b 34k
Stat5a - 6k
Stat5a 10k
Stat5a 27k
D
H3K27ac ChIP
Veh
E2
ICI
Relative H3K27ac / H3
Stat5a 10kb

Appendix C:

Additional GRHL2 Experiments

In Chapter 3, motif analysis was performed on pS118-ER occupancy regions and it was found that the GRHL2 motif was enriched at pS118-ER sites compared to ER-only sites (Figure 3-2). This appendix contains experiments performed to determine if GRHL2 plays a role in ER signaling. Other studies have reported an interaction between ER and GRHL2 through co-immunoprecipitation or RIME experiments (201,239), but no experiments had been performed to test if GRHL2 in involved in E2-dependent ER gene transcription or if GRHL2 or ER control expression of the other. Analysis of the ER and pS118-ER ChIP-seq data revealed that the *GRHL2* locus contain many ER and pS118-ER occupancy sites within 50 kb of its transcription start site (Figure C-1), suggesting that *GRHL2* may be an ER target gene. However, MCF-7 cells treated with E2 for 4 hours showed no change in *GRHL2* mRNA levels, indicating that E2 does not control *GRHL2* (Figure C-4D). Additionally, GRHL2 protein is not affected by E2 treatment for 24 hours (Figure C-3A). To test if loss of GRHL2 had an effect on ER expression, an siRNA protocol for knocking down GRHL2 was optimized (Figure C-2; see Chapter 6 for methods). First, GRHL2 knockdown was performed in MCF-7 cells using 80 pmol siGRHL2 as described in methods section (Chapter 6) and substantial knockdown was achieved (Figure C-2A). Next, to analyze the length of time the GRHL2 knockdown lasts, MCF-7 cells were transfected as before, and after 24 hours the media was replaced and cells were allowed to grow for between one to five days followed by protein isolation and analysis by immunoblot. GRHL2 knockdown was achieved at day 1 and lasted throughout the time course to day 5 (Figure C-2B). Time points longer than 5 days have not been tested. Additionally, a dose response of siGRHL2 was performed to test if lower amounts of siGRHL2 can produce similar levels of GRHL2 knockdown. MCF-7 cells were transfected

with varying amounts of siGRHL2 and analyzed by immunoblot. Knockdown of GRHL2 was achieved at amounts as low as 10 pmol siGRHL2 and the knockdown appeared to be as efficient as 80 pmol siGRHL2 (Figure C-2C). With the GRHL2 knockdown optimized, the effects of GRHL2 knockdown on ER function was analyzed. GRHL2 knockdown was found to have no effect on ER protein levels or the downregulation of ER by E2 (Figure C-3). A previous report found that the loss of GRHL2 decreased levels of androgen receptor (AR) and E-cadherin (E-Cad) in LNCaP cells (317). Given that ER and AR are in the same nuclear receptor family, it was hypothesized that AR and E-Cad would similarly be lost in MCF-7 cells upon loss of GRHL2. However, the loss of GRHL2 in MCF-7 cells led to not change in the levels or AR or E-Cad (Figure C-3) indication that the GRHL2 signaling axis is not the same between MCF-7 and LNCaP cells. Gene analysis of ER target genes was also performed in GRHL2 knockdown MCF-7 cells. While many genes known to be upregulated and downregulated by E2 were tested, no genes were found to be affected by GRHL2 knockdown (Figure C-4A, C-4B). Additional genes known to be regulated by GRHL2 (*VIM* and *ZEB1*) were also tested, but no significant differences were observed (Figure C-4C). *GRHL2* levels were found to be reduced to approximately 40% of the control levels (Figure C-4D) suggesting that low levels of GRHL2 may still be present, and a full knockout of GRHL2 may need to be performed to confirm the results observer here. To test if loss of GRHL2 had an effect on DNA occupancy of ER or pS118-ER, MCF-7 cells were transfected with either control or siGRHL2 followed by treatment with vehicle (0.1% EtOH) or 10 nM E2 and ChIP-qPCR was performed on a set of sites known to be occupied by both GRHL2 and pS118-ER. Similarly, knockdown of GRHL2 had no statistically significant effect on the occupancy of ER or pS118-ER in either control

of E2 conditions (Figure C-5). Although not statistically significant, GRHL2 occupancy tended to increase upon E2 treatment in line with results from Figure 3-3C.

Figure C-1: ER and pS118-ER Occupancy Sites are Present at the GRHL2 Locus.

UCSC Genome Browser view of ER and pS118-ER ChIP-seq tracks at the *GRHL2* locus. Vehicle (yellow) and E2 treated (blue) tracks are overlapped and overlapping regions are displayed as green. Various ER and pS118-ER occupancy sites are labeled with grey lines.

Figure C-1

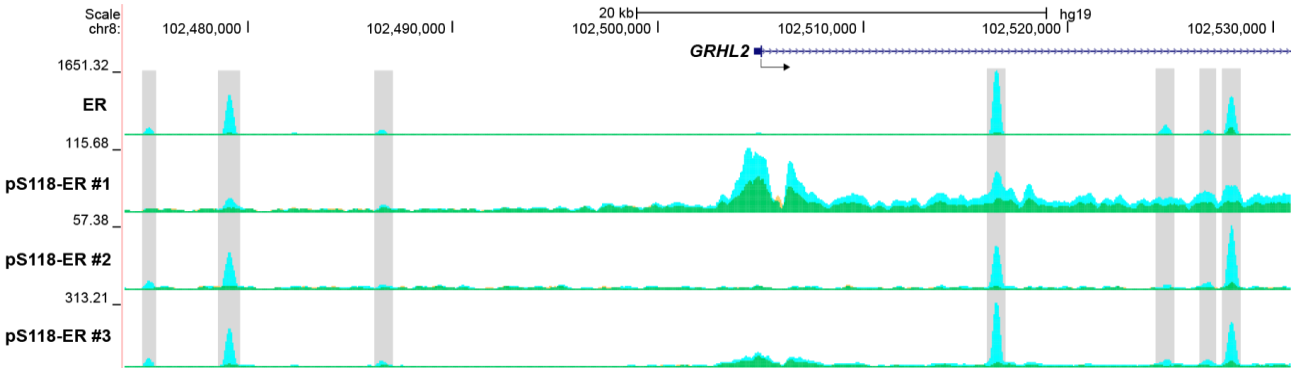

Scale
chr8:
102,480,000
102,490,000
20 kb
102,500,000
102,510,000
102,520,000
hg19
102,530,000
GRHL2
1651.32
ER
115.68
pS118-ER #1
57.38
pS118-ER #2
313.21
pS118-ER #3

Figure C-2: GRHL2 knockdown efficiency in MCF-7 cells.

(A) MCF-7 cells in full media were transfected with 80 pmol siGRHL2 for 24 hours as described in materials and methods (see Chapter 6). Protein was collected and analyzed by immunoblot. (B) Time course analysis of siGRHL2. MCF-7 cells were transfected with 80 pmol siGRHL2 for 24 hours. After transfection media was replaced and cells were allowed to grow for 1 to 5 days followed by protein isolation and analysis. (C) MCF-7 cells were transfected with varying amounts of siGRHL2 for 24 hours and protein was analyzed by immunoblot. No lipofectamine 2000 was used in the negative control and the 0 pmol treatment contains lipofectamine but with no siGRHL2. A long and short exposure of GRHL2 is shown.

Figure C-2

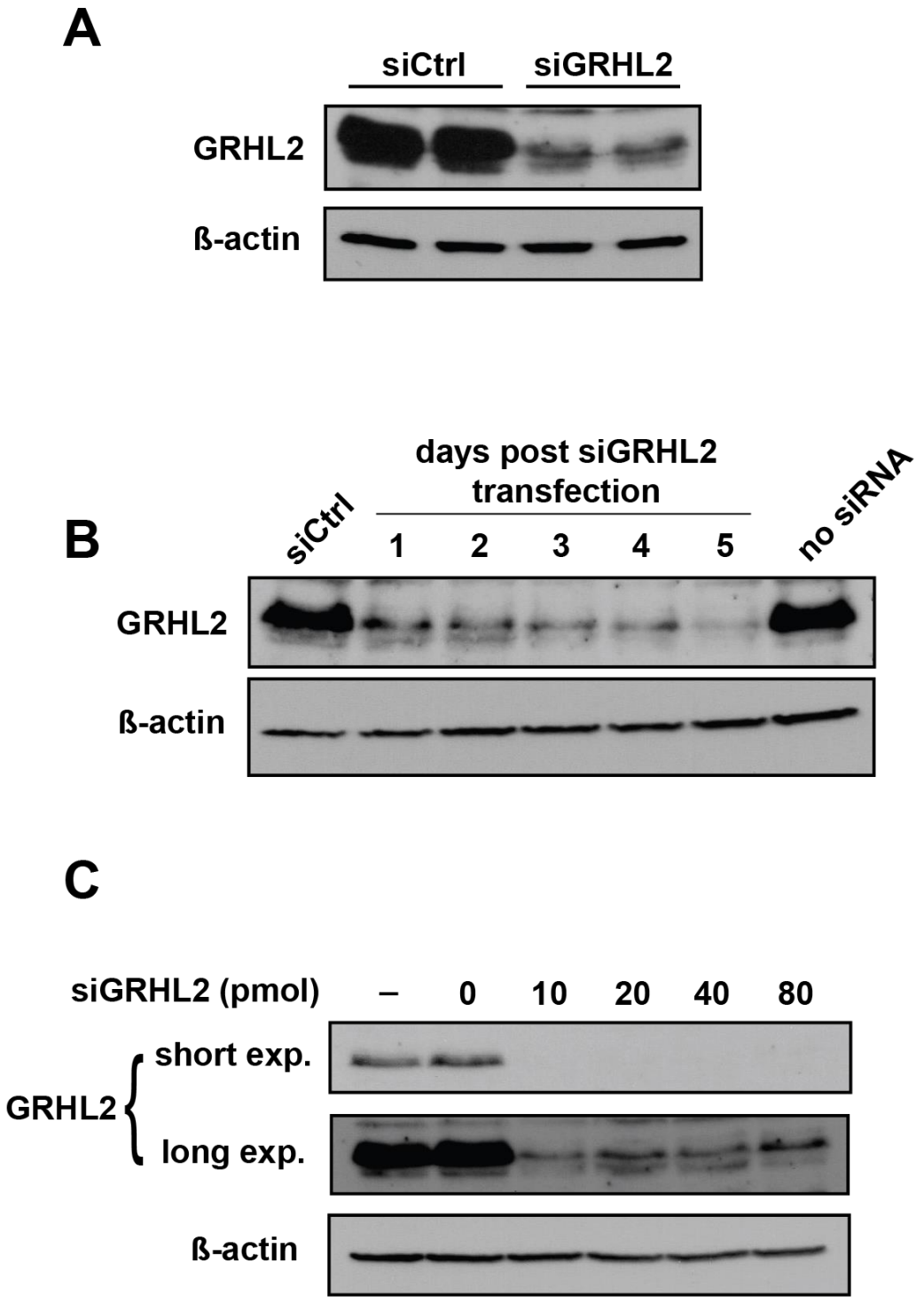

A
siCtrl
siGRHL2
GRHL2
ß-actin
B
days post siGRHL2 transfection
siCtrl
1
2
3
4
5
no siRNA
GRHL2
ß-actin
C
siGRHL2 (pmol)
–
0
10
20
40
80
GRHL2
short exp.
long exp.
ß-actin

Figure C-3: Effect of GRHL2 knockdown on ER and other proteins known to be regulated by GRHL2 in LNCaP cells.

MCF-7 cells were stripped for three days followed by transfected with 80 pmol siGRHL2 for 24 hours. Cells were then treated with either vehicle (0.1% EtOH) or 10 nM E2 for 24 hours and protein was collected and analyzed by immunoblot. Androgen receptor (AR) and E-Cadherin (E-Cad) were analyzed as these were found to be regulated by GRHL2 in LNCaP cells (317).

Figure C-3

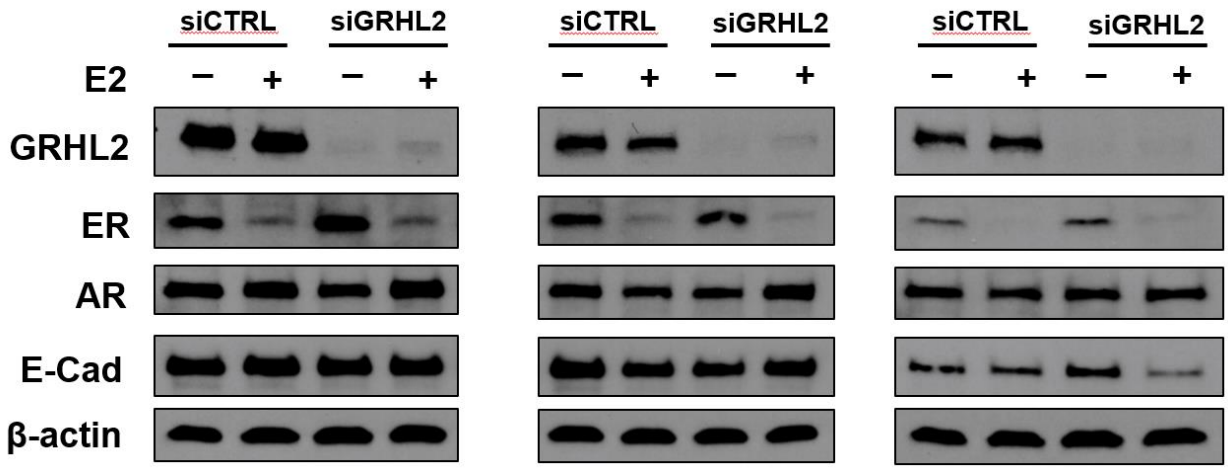

siCTRL
siGRHL2
siCTRL
siGRHL2
siCTRL
siGRHL2
E2
– + – +
– + – +
– + – +
GRHL2
ER
AR
E-Cad
β-actin

Figure C-4: GRHL2 knockdown analysis of ER target genes.

MCF-7 cells were stripped for three days followed by transfected with siGRHL2 for 24 hours. Cells were then treated with either vehicle (0.1% EtOH) or 10 nM E2 for 4 hours and RNA was isolated. Genes known to be (A) upregulated or (B) downregulated by E2 were tested for alteration by siGRHL2. Additionally, genes known to be affected by GRHL2 loss (C) were analyzed. (D) *GRHL2* knockdown efficiency from siGRHL2. Three replicates are shown.

Figure C-4

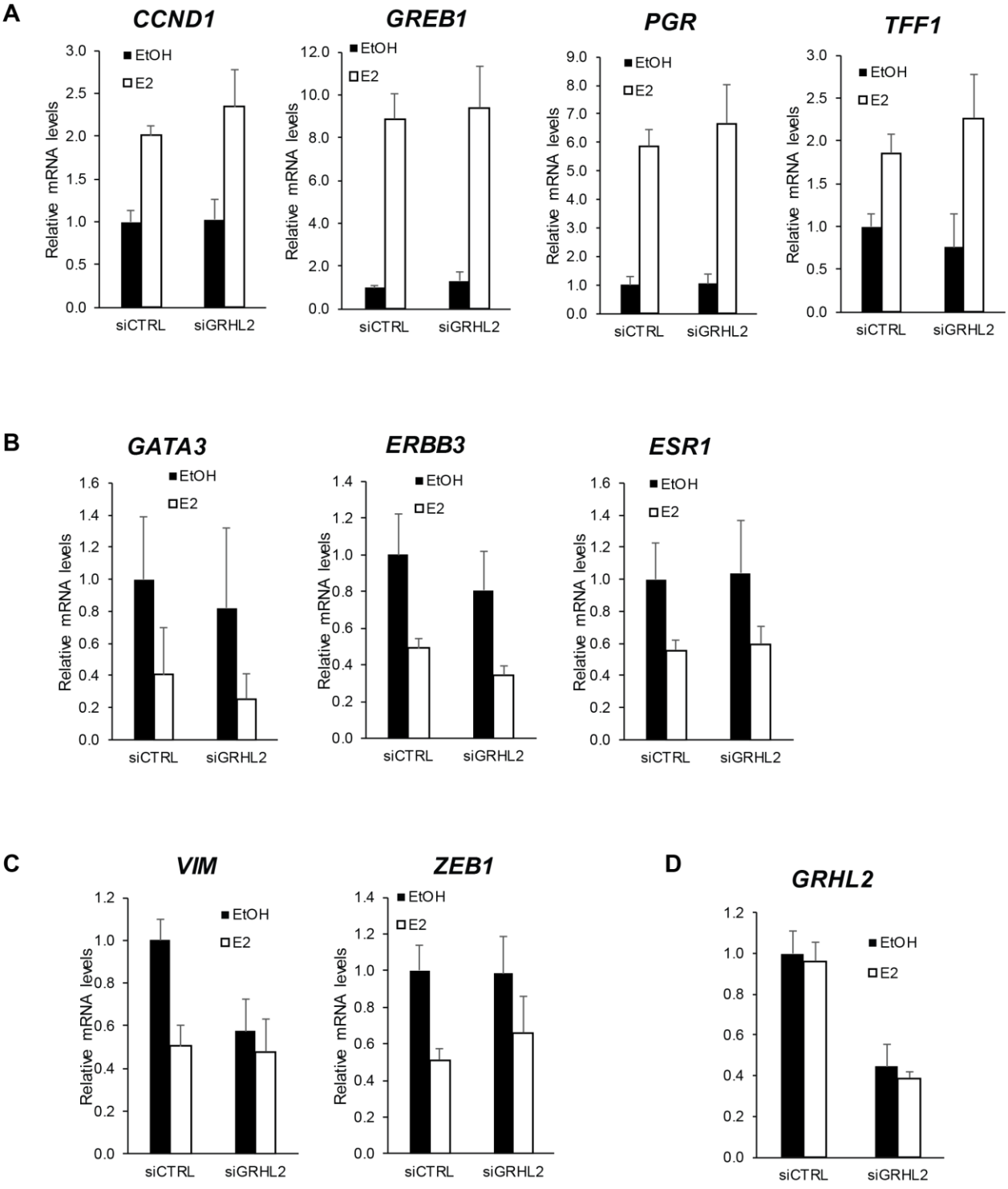
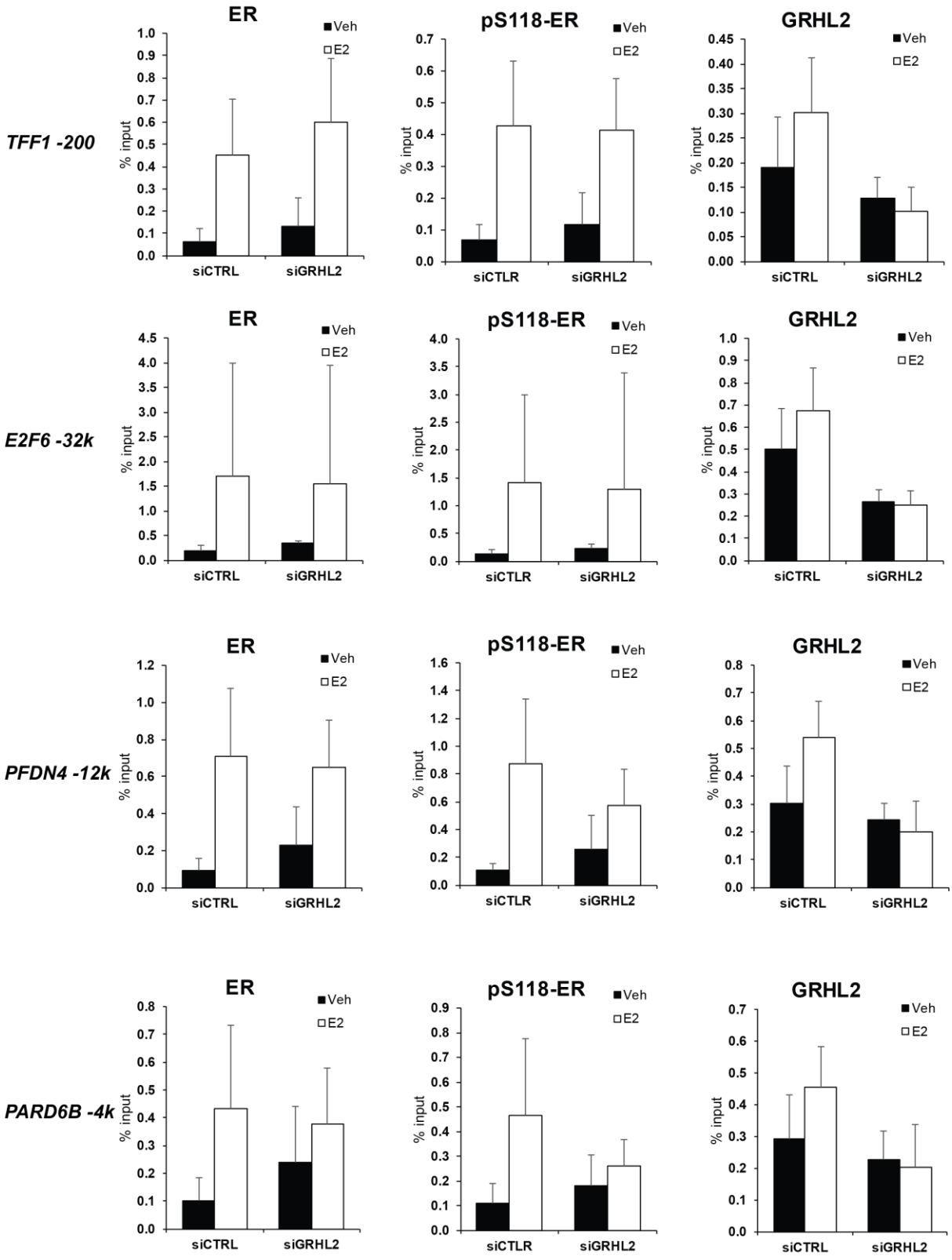

A
CCND1
GREB1
PGR
TFF1
EtOH
E2
Relative mRNA levels
siCTRL
siGRHL2
B
GATA3
ERBB3
ESR1
C
VIM
ZEB1
D
GRHL2

Figure C-5: ER and pS118-ER ChIP in GRHL2 knockdown MCF-7 cells.

MCF-7 cells were stripped for three days and transfected with siGRHL2 as previously described (see Chapter 6 for methods) followed by treatment with 10 nM E2 for 30 min. Chromatin immunoprecipitation was performed for ER, pS118-ER, and GRHL2 in the presence and absence of E2. Four regions were tested for occupancy of each of these factors (displayed on left). No statistical differences were observed between respective siCTRL treatments and siGRHL2.

Figure C-5


TFF1 -200
E2F6 -32k
PFDN4 -12k
PARD6B -4k
ER
pS118-ER
GRHL2
Veh
E2
% input
siCTRL
siCTLR
siGRHL2

Appendix D:

Analysis of Breast Cancer Cells Isolated from Pleural Effusions

This appendix contains data relating to the isolation and analysis of breast cancer cells from pleural effusions or ascites collected from patients with metastatic breast cancer and was done in collaboration with the lab of Dr. Amy Fowler. The MCF-7 cell line was originally isolated from a pleural effusion performed on a patient with ER-positive breast cancer (17,18) and the goal of this project was to isolate breast cancer cells from patients receiving a similar procedure and to analyze the expression and activity of ER and pS118-ER in these cells. Patients were enrolled from the University of Wisconsin Carbone Cancer Center who had ER-positive primary breast cancer which had metastasized to various sites. Each patient's age, metastasis sites, and fluid type are shown in Table D-1. The cells are named WICC cells after the Wisconsin Carbone Cancer Center. Upon receiving the pleural effusion or ascites, cells were pelleted at 500 x g and the supernatant was removed. The cell pellet was resuspended in TAC buffer (150 mM NH_4Cl, 170 mM Tris-HCl, pH 7.4) and incubated at 37°C for 10 min to lyse red blood cells. The remaining cells were pelleted, and the supernatant removed. These cells were then washed three times with 1x PBS, counted, and either frozen down or plated into various medias for growth. Upon initial isolation of the cells, samples were taken to perform protein analysis and to test for ER status. WICC1 though WICC9 were tested for expression of ER and pS118-ER. WICC4 was not analyzed due to insufficient material. WICC1, WICC5, WICC6, and WICC9 were found to be ER-positive and expressed ER at similar levels to MCF-7 (Figure D-1). All ER-positive WICC cells were also found to express pS118-ER, indicating that ER is being activated in some manner. To test if ER was actively bound to DNA, cells from WICC1 and WICC9 cells were collected during isolation and ChIP-qPCR was performed. A number of sites known to be occupied by ER

and pS118-ER in MCF-7 cells were tested and all sites were similarly observed to be significantly occupied by ER and pS118-ER in both WICC1 and WICC9 cells compared to the IgG control (Figure D-2). The presence of pS118-ER and the occupancy observed on DNA of both ER and pS118-ER in the WICC cells indicate that although these patients have been treated with anti-estrogen therapies, ER still appears to be active to some extent.

Table D-1: WICC Patient Information

Patient ID	Age	Met Sites	Fluid Type
WICC1	62	Bone, Brain	Pleural Effusion
WICC2	59	Bone	Ascites
WICC3	59	Bone	Pleural Effusion
WICC4	57	Liver	Ascites
WICC5	58	Bone	Pleural Effusion
WICC6	57	Brain	Ascites
WICC7	56	Liver	Ascites
WICC8	62	Omentum, Bone	Ascites
WICC9	75	Chest Wall	Ascites

Figure D-1: Protein Analysis of WICC Cells.

After isolation of cells from pleural effusions, cells were lysed and protein was quantified. Equal amounts of protein were analyzed by immunoblot and blot was probed for ER and pS118-ER. The ER band is denoted with an arrow. WICC4 cells were not analyzed due to insufficient material.

Figure D-1

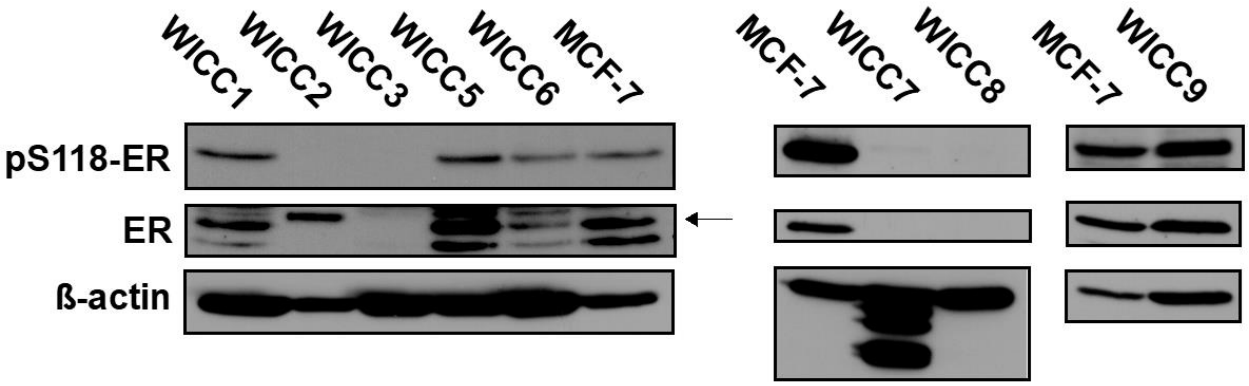
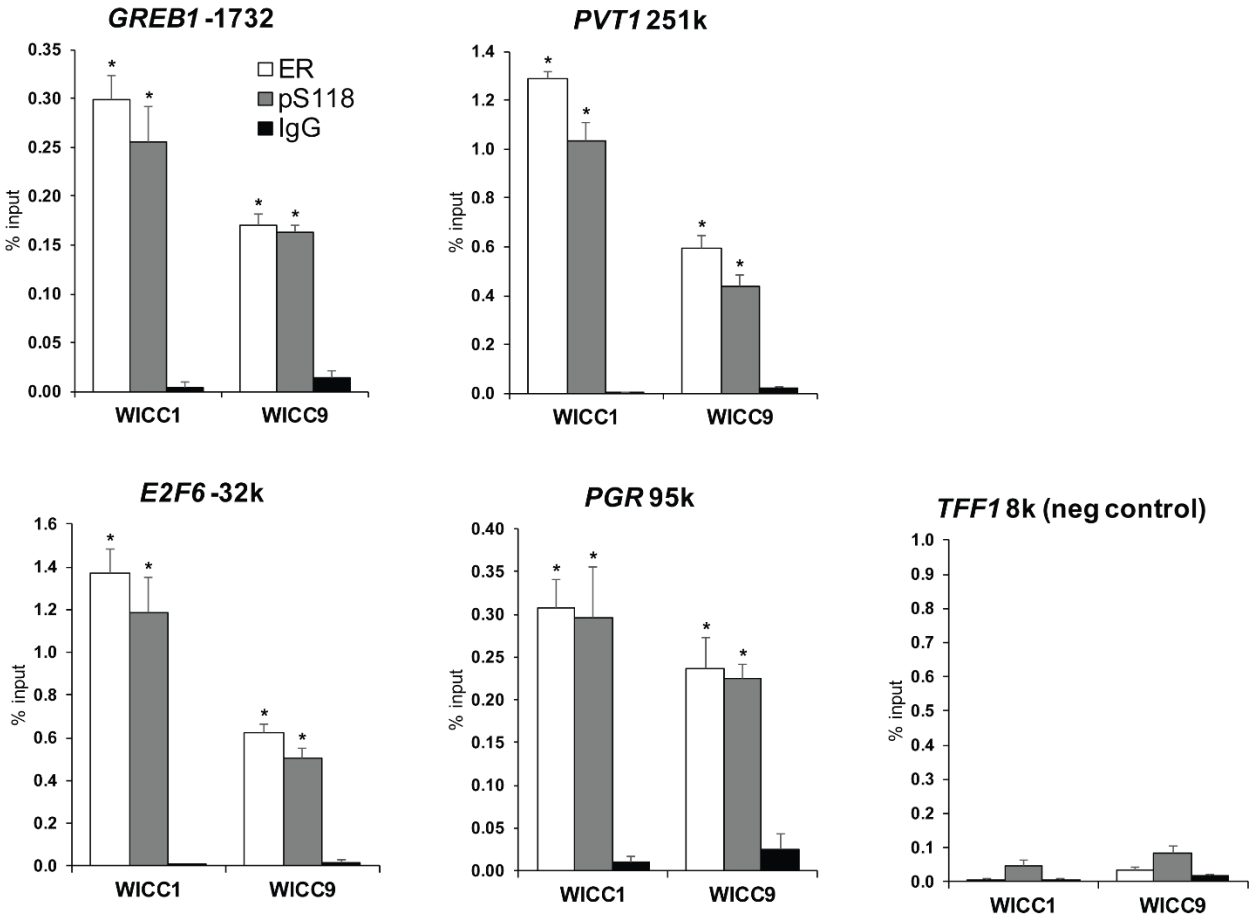

WICC1
WICC2
WICC3
WICC5
WICC6
MCF-7
MCF-7
WICC7
WICC8
MCF-7
WICC9
pS118-ER
ER
ß-actin

Figure D-2: ER and pS118-ER Occupy DNA in WICC1 and WICC9 cells.

Upon isolation of cells from WICC1 and WICC9 pleural effusions, cells were fixed in suspension with 1.5% formaldehyde and chromatin immunoprecipitation was performed for ER and pS118-ER on three technical replicates for each cell line. IgG is used as a control. Known ER and pS118-ER occupancy sites as identified from the ER and pS118-ER ChIP-seq study in this thesis were analyzed for occupancy in the WICC cells. Data are calculated as a percent input. Student's t-test was performed on percent input values and an asterisk (*) denotes $p<0.05$ compared to IgG control.

Figure D-2


GREB1 -1732
ER
pS118
IgG
% input
WICC1
WICC9
PVT1 251k
E2F6 -32k
PGR 95k
TFF1 8k (neg control)

References

1. **U.S. Cancer Statistics Working Group.** U.S. Cancer Statistics Data Visualizations Tool, based on November 2018 submission data (1999-2016): U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute. Available at: https://gis.cdc.gov/Cancer/USCS/DataViz.html. Accessed June 28, 2019.

2. **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 Cancer J Clin* 2018;68(6):394–424.

3. **Stierer M, Rosen H, Weber R, Hanak H, Spona J, Tüchler H.** Immunohistochemical and biochemical measurement of estrogen and progesterone receptors in primary breast cancer. Correlation of histopathology and prognostic factors. *Ann. Surg.* 1993;218(1):13–21.

4. **Harvey JM, Clark GM, Osborne CK, Allred DC.** Estrogen receptor status by immunohistochemistry is superior to the ligand-binding assay for predicting response to adjuvant endocrine therapy in breast cancer. *J Clin Oncol.* 1999;17(5):1474–81.

5. **Cuzick J, Sestak I, Cawthorn S, Hamed H, Holli K, Howell A, Forbes JF, IBIS-I Investigators.** Tamoxifen for prevention of breast cancer: extended long-term follow-up of the IBIS-I breast cancer prevention trial. *Lancet Oncol.* 2015;16(1):67–75.

6. **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;15:153–179.

7. **Lett H.** An Analysis of Ninety-nine Cases of Inoperable Carcinoma of the Breast treated by Oophorectomy. *Med Chir Trans* 1905;88:147–189.

8. **Lathrop AE, Loeb L.** Further investigations on the origin of tumors in mice. III. On the part played by internal secretion in the spontaneous development of tumors. *J Cancer Res* 1916;1(1):1–19.

9. **Doisy E, Veler C, Thayer S.** Folliculin from urine of pregnant women. *Am. J. Physiol.* 1929;90(2):329–330.

10. **Huffman MN, Thayer SA, Doisy EA.** THE ISOLATION OF α-DIHYDROTHEELIN FROM HUMAN PLACENTA. *J. Biol. Chem.* 1940;133(2):567–571.

11. **Simoni RD, Hill RL, Vaughan M.** The Discovery of Estrone, Estriol, and Estradiol and the Biochemical Study of Reproduction. The Work of Edward Adelbert Doisy. *J. Biol. Chem.* 2002;277(28):e17–e17.

12. **Simpson E, Santen RJ.** Celebrating 75 years of oestradiol. *J. Mol. Endocrinol.* 2015;55(3):T1-20.

13. **Jensen EV.** On the mechanism of estrogen action. *Perspect. Biol. Med.* 1962;6:47–59.

14. **Noteboom WD, Gorski J.** Stereospecific binding of estrogens in the rat uterus. *Arch. Biochem. Biophys.* 1965;111(3):559–568.

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

16. **Green S, Chambon P.** Oestradiol induction of a glucocorticoid-responsive gene by a chimaeric receptor. *Nature* 1987;325(6099):75–78.

17. **Brooks SC, Locke ER, Soule HD.** Estrogen receptor in a human cell line (MCF-7) from breast carcinoma. *J. Biol. Chem.* 1973;248(17):6251–6253.

18. **Soule HD, Vazguez J, Long A, Albert S, Brennan M.** A human cell line from a pleural effusion derived from a breast carcinoma. *J. Natl. Cancer Inst.* 1973;51(5):1409–1416.

19. **Comşa Ş, Cîmpean AM, Raica M.** The Story of MCF-7 Breast Cancer Cell Line: 40 years of Experience in Research. *Anticancer Res.* 2015;35(6):3147–3154.

20. **Kumar R, Thompson EB.** The structure of the nuclear hormone receptors. *Steroids* 1999;64(5):310–319.

21. **Bain DL, Heneghan AF, Connaghan-Jones KD, Miura MT.** Nuclear receptor structure: implications for function. *Annu. Rev. Physiol.* 2007;69:201–220.

22. **Tora L, White J, Brou C, Tasset D, Webster N, Scheer E, Chambon P.** The human estrogen receptor has two independent nonacidic transcriptional activation functions. *Cell* 1989;59(3):477–487.

23. **Lavery DN, McEwan IJ.** Structure and function of steroid receptor AF1 transactivation domains: induction of active conformations. *Biochem. J.* 2005;391(Pt 3):449–464.

24. **Kumar V, Green S, Staub A, Chambon P.** Localisation of the oestradiol-binding and putative DNA-binding domains of the human oestrogen receptor. *EMBO J.* 1986;5(9):2231–2236.

25. **Kumar V, Green S, Stack G, Berry M, Jin JR, Chambon P.** Functional domains of the human estrogen receptor. *Cell* 1987;51(6):941–951.

26. **Ekena K, Weis KE, Katzenellenbogen JA, Katzenellenbogen BS.** Identification of amino acids in the hormone binding domain of the human estrogen receptor important in estrogen binding. *J. Biol. Chem.* 1996;271(33):20053–20059.

27. **Welshons WV, Lieberman ME, Gorski J.** Nuclear localization of unoccupied oestrogen receptors. *Nature* 1984;307(5953):747–749.

28. **Ylikomi T, Bocquel MT, Berry M, Gronemeyer H, Chambon P.** Cooperation of proto-signals for nuclear accumulation of estrogen and progesterone receptors. *EMBO J.* 1992;11(10):3681–3694.

29. **Notides AC, Lerner N, Hamilton DE.** Positive cooperativity of the estrogen receptor. *Proc. Natl. Acad. Sci. U.S.A.* 1981;78(8):4926–4930.

30. **Wang H, Peters GA, Zeng X, Tang M, Ip W, Khan SA.** Yeast two-hybrid system demonstrates that estrogen receptor dimerization is ligand-dependent in vivo. *J. Biol. Chem.* 1995;270(40):23322–23329.

31. **Brzozowski AM, Pike ACW, Dauter Z, Hubbard RE, Bonn T, Engström O, Öhman L, Greene GL, Gustafsson J-Å, Carlquist M.** Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature* 1997;389(6652):753–758.

32. **Heery DM, Kalkhoven E, Hoare S, Parker MG.** A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature* 1997;387(6634):733–736.

33. **Shiau AK, Barstad D, Loria PM, Cheng L, Kushner PJ, Agard DA, Greene GL.** The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. *Cell* 1998;95(7):927–937.

34. **Denton RR, Koszewski NJ, Notides AC.** Estrogen receptor phosphorylation. Hormonal dependence and consequence on specific DNA binding. *J. Biol. Chem.* 1992;267(11):7263–7268.

35. **Kim MY, Woo EM, Chong YTE, Homenko DR, Kraus WL.** Acetylation of estrogen receptor alpha by p300 at lysines 266 and 268 enhances the deoxyribonucleic acid binding and transactivation activities of the receptor. *Mol. Endocrinol.* 2006;20(7):1479–1493.

36. **Le Romancer M, Treilleux I, Leconte N, Robin-Lespinasse Y, Sentis S, Bouchekioua-Bouzaghou K, Goddard S, Gobert-Gosse S, Corbo L.** Regulation of estrogen rapid signaling through arginine methylation by PRMT1. *Mol. Cell* 2008;31(2):212–221.

37. **Subramanian K, Jia D, Kapoor-Vazirani P, Powell DR, Collins RE, Sharma D, Peng J, Cheng X, Vertino PM.** Regulation of estrogen receptor alpha by the SET7 lysine methyltransferase. *Mol. Cell* 2008;30(3):336–347.

38. **Nirmala PB, Thampan RV.** Ubiquitination of the rat uterine estrogen receptor: dependence on estradiol. *Biochem. Biophys. Res. Commun.* 1995;213(1):24–31.

39. **Acconcia F, Ascenzi P, Fabozzi G, Visca P, Marino M.** S-palmitoylation modulates human estrogen receptor-alpha functions. *Biochem. Biophys. Res. Commun.* 2004;316(3):878–883.

40. **Sentis S, Le Romancer M, Bianchin C, Rostan M-C, Corbo L.** Sumoylation of the estrogen receptor alpha hinge region regulates its transcriptional activity. *Mol. Endocrinol.* 2005;19(11):2671–2684.

41. **Le Romancer M, Poulard C, Cohen P, Sentis S, Renoir J-M, Corbo L.** Cracking the Estrogen Receptor's Posttranslational Code in Breast Tumors. *Endocrine Reviews* 2011;32(5):597–622.

42. **Moran MF, Polakis P, McCormick F, Pawson T, Ellis C.** Protein-tyrosine kinases regulate the phosphorylation, protein interactions, subcellular distribution, and activity of p21ras GTPase-activating protein. *Mol. Cell. Biol.* 1991;11(4):1804–1812.

43. **Mayo LD, Donner DB.** A phosphatidylinositol 3-kinase/Akt pathway promotes translocation of Mdm2 from the cytoplasm to the nucleus. *PNAS* 2001;98(20):11598–11603.

44. **Zhou BP, Liao Y, Xia W, Spohn B, Lee MH, Hung MC.** Cytoplasmic localization of p21Cip1/WAF1 by Akt-induced phosphorylation in HER-2/neu-overexpressing cells. *Nat. Cell Biol.* 2001;3(3):245–252.

45. **Zhou YJ, Hanson EP, Chen YQ, Magnuson K, Chen M, Swann PG, Wange RL, Changelian PS, O'Shea JJ.** Distinct tyrosine phosphorylation sites in JAK3 kinase domain positively and negatively regulate its enzymatic activity. *Proc. Natl. Acad. Sci. U.S.A.* 1997;94(25):13850–13855.

46. **Liu C, Li Y, Semenov M, Han C, Baeg GH, Tan Y, Zhang Z, Lin X, He X.** Control of beta-catenin phosphorylation/degradation by a dual-kinase mechanism. *Cell* 2002;108(6):837–847.

47. **Lin H-K, Wang L, Hu Y-C, Altuwaijri S, Chang C.** Phosphorylation-dependent ubiquitylation and degradation of androgen receptor by Akt require Mdm2 E3 ligase. *The EMBO Journal* 2002;21(15):4037–4048.

48. **Sharma K, D'Souza RCJ, Tyanova S, Schaab C, Wiśniewski JR, Cox J, Mann M.** Ultradeep human phosphoproteome reveals a distinct regulatory nature of Tyr and Ser/Thr-based signaling. *Cell Rep* 2014;8(5):1583–1594.

49. **Vlastaridis P, Kyriakidou P, Chaliotis A, Van de Peer Y, Oliver SG, Amoutzias GD.** Estimating the total number of phosphoproteins and phosphorylation sites in eukaryotic proteomes. *Gigascience* 2017;6(2):1–11.

50. **Washburn T, Hocutt A, Brautigan DL, Korach KS.** Uterine estrogen receptor in vivo: phosphorylation of nuclear specific forms on serine residues. *Mol. Endocrinol.* 1991;5(2):235–242.

51. **Le Goff P, Montano MM, Schodin DJ, Katzenellenbogen BS.** Phosphorylation of the human estrogen receptor. Identification of hormone-regulated sites and examination of their influence on transcriptional activity. *J. Biol. Chem.* 1994;269(6):4458–4466.

52. **Ali S, Metzger D, Bornert JM, Chambon P.** Modulation of transcriptional activation by ligand-dependent phosphorylation of the human oestrogen receptor A/B region. *EMBO J.* 1993;12(3):1153–1160.

53. **Arnold SF, Obourn JD, Jaffe H, Notides AC.** Serine 167 is the major estradiol-induced phosphorylation site on the human estrogen receptor. *Mol. Endocrinol.* 1994;8(9):1208–1214.

54. **Britton DJ, Scott GK, Schilling B, Atsriku C, Held JM, Gibson BW, Benz CC, Baldwin MA.** A novel serine phosphorylation site detected in the N-terminal domain of estrogen receptor isolated from human breast cancer cells. *J. Am. Soc. Mass Spectrom.* 2008;19(5):729–740.

55. **Atsriku C, Britton DJ, Held JM, Schilling B, Scott GK, Gibson BW, Benz CC, Baldwin MA.** Systematic mapping of posttranslational modifications in human estrogen receptor-alpha with emphasis on novel phosphorylation sites. *Mol. Cell Proteomics* 2009;8(3):467–480.

56. **Williams CC, Basu A, El-Gharbawy A, Carrier LM, Smith CL, Rowan BG.** Identification of four novel phosphorylation sites in estrogen receptor alpha: impact on receptor-dependent gene expression and phosphorylation by protein kinase CK2. *BMC Biochem.* 2009;10:36.

57. **Kato S, Endoh H, Masuhiro Y, Kitamoto T, Uchiyama S, Sasaki H, Masushige S, Gotoh Y, Nishida E, Kawashima H, Metzger D, Chambon P.** Activation of the estrogen receptor through phosphorylation by mitogen-activated protein kinase. *Science* 1995;270(5241):1491–1494.

58. **Aronica SM, Katzenellenbogen BS.** Stimulation of estrogen receptor-mediated transcription and alteration in the phosphorylation state of the rat uterine estrogen receptor by estrogen, cyclic adenosine monophosphate, and insulin-like growth factor-I. *Mol. Endocrinol.* 1993;7(6):743–752.

59. **Bunone G, Briand PA, Miksicek RJ, Picard D.** Activation of the unliganded estrogen receptor by EGF involves the MAP kinase pathway and direct phosphorylation. *EMBO J.* 1996;15(9):2174–2183.

60. **González L, Zambrano A, Lazaro-Trueba I, Lopéz E, González JJA, Martín-Pérez J, Aranda A.** Activation of the unliganded estrogen receptor by prolactin in breast cancer cells. *Oncogene* 2009;28(10):1298–1308.

61. **Shou J, Massarweh S, Osborne CK, Wakeling AE, Ali S, Weiss H, Schiff R.** Mechanisms of tamoxifen resistance: increased estrogen receptor-HER2/neu cross-talk in ER/HER2-positive breast cancer. *J. Natl. Cancer Inst.* 2004;96(12):926–935.

62. **Khanal T, Kim HG, Do MT, Choi JH, Won SS, Kang W, Chung YC, Jeong TC, Jeong HG.** Leptin induces CYP1B1 expression in MCF-7 cells through ligand-independent activation of the ERα pathway. *Toxicology and Applied Pharmacology* 2014;277(1):39–48.

63. **Weitsman GE, Weebadda W, Ung K, Murphy LC.** Reactive oxygen species induce phosphorylation of serine 118 and 167 on estrogen receptor alpha. *Breast Cancer Res Treat* 2009;118(2):269–279.

64. **Park J, Lee Y.** Hypoxia induced phosphorylation of estrogen receptor at serine 118 in the absence of ligand. *J. Steroid Biochem. Mol. Biol.* 2017. doi:10.1016/j.jsbmb.2017.08.013.

65. **Joel PB, Traish AM, Lannigan DA.** Estradiol-induced phosphorylation of serine 118 in the estrogen receptor is independent of p42/p44 mitogen-activated protein kinase. *J. Biol. Chem.* 1998;273(21):13317–13323.

66. **Tzeng DZ, Klinge CM.** Phosphorylation of purified estradiol-liganded estrogen receptor by casein kinase II increases estrogen response element binding but does not alter ligand stability. *Biochem. Biophys. Res. Commun.* 1996;223(3):554–560.

67. **Chen D, Riedl T, Washbrook E, Pace PE, Coombes RC, Egly JM, Ali S.** Activation of estrogen receptor alpha by S118 phosphorylation involves a ligand-dependent interaction with TFIIH and participation of CDK7. *Mol. Cell* 2000;6(1):127–137.

68. **Yankulov KY, Bentley DL.** Regulation of CDK7 substrate specificity by MAT1 and TFIIH. *EMBO J.* 1997;16(7):1638–1646.

69. **Chen D, Washbrook E, Sarwar N, Bates GJ, Pace PE, Thirunuvakkarasu V, Taylor J, Epstein RJ, Fuller-Pace FV, Egly J-M, Coombes RC, Ali S.** Phosphorylation of human estrogen receptor alpha at serine 118 by two distinct signal transduction pathways revealed by phosphorylation-specific antisera. *Oncogene* 2002;21(32):4921–4931.

70. **Métivier R, Gay FA, Hübner MR, Flouriot G, Salbert G, Gannon F, Kah O, Pakdel F.** Formation of an hER alpha-COUP-TFI complex enhances hER alpha AF-1 through Ser118 phosphorylation by MAPK. *EMBO J.* 2002;21(13):3443–3453.

71. **Medunjanin S, Hermani A, De Servi B, Grisouard J, Rincke G, Mayer D.** Glycogen synthase kinase-3 interacts with and phosphorylates estrogen receptor alpha and is involved in the regulation of receptor activity. *J. Biol. Chem.* 2005;280(38):33006–33014.

72. **Park K-J, Krishnan V, O'Malley BW, Yamamoto Y, Gaynor RB.** Formation of an IKKalpha-dependent transcription complex is required for estrogen receptor-mediated gene activation. *Mol. Cell* 2005;18(1):71–82.

73. **Medunjanin S, Weinert S, Schmeisser A, Mayer D, Braun-Dullaeus RC.** Interaction of the double-strand break repair kinase DNA-PK and estrogen receptor-alpha. *Mol. Biol. Cell* 2010;21(9):1620–1628.

74. **Foulds CE, Feng Q, Ding C, Bailey S, Hunsaker TL, Malovannaya A, Hamilton RA, Gates LA, Zhang Z, Li C, Chan D, Bajaj A, Callaway CG, Edwards DP, Lonard DM, Tsai SY, Tsai M-J, Qin J, O'Malley BW.** Proteomic analysis of coregulators bound to ERα on DNA and nucleosomes reveals coregulator dynamics. *Mol. Cell* 2013;51(2):185–199.

75. **Hah N, Danko CG, Core L, Waterfall JJ, Siepel A, Lis JT, Kraus WL.** A rapid, extensive, and transient transcriptional response to estrogen signaling in breast cancer cells. *Cell* 2011;145(4):622–634.

76. **El-Tanani MK, Green CD.** Two separate mechanisms for ligand-independent activation of the estrogen receptor. *Mol. Endocrinol.* 1997;11(7):928–937.

77. **Dutertre M, Smith CL.** Ligand-independent interactions of p160/steroid receptor coactivators and CREB-binding protein (CBP) with estrogen receptor-alpha: regulation by phosphorylation sites in the A/B region depends on other receptor domains. *Mol. Endocrinol.* 2003;17(7):1296–1314.

78. **Valley CC, Métivier R, Solodin NM, Fowler AM, Mashek MT, Hill L, Alarid ET.** Differential regulation of estrogen-inducible proteolysis and transcription by the estrogen receptor alpha N terminus. *Mol. Cell. Biol.* 2005;25(13):5417–5428.

79. **Cheng J, Zhang C, Shapiro DJ.** A functional serine 118 phosphorylation site in estrogen receptor-alpha is required for down-regulation of gene expression by 17beta-estradiol and 4-hydroxytamoxifen. *Endocrinology* 2007;148(10):4634–4641.

80. **Duplessis TT, Williams CC, Hill SM, Rowan BG.** Phosphorylation of Estrogen Receptor α at Serine 118 Directs Recruitment of Promoter Complexes and Gene-Specific Transcription. *Endocrinology* 2011;152(6):2517–2526.

81. **Huderson BP, Duplessis TT, Williams CC, Seger HC, Marsden CG, Pouey KJ, Hill SM, Rowan BG.** Stable inhibition of specific estrogen receptor α (ERα) phosphorylation confers increased growth, migration/invasion, and disruption of estradiol signaling in MCF-7 breast cancer cells. *Endocrinology* 2012;153(9):4144–4159.

82. **Oñate SA, Tsai SY, Tsai MJ, O'Malley BW.** Sequence and characterization of a coactivator for the steroid hormone receptor superfamily. *Science* 1995;270(5240):1354–1357.

83. **Voegel JJ, Heine MJ, Zechel C, Chambon P, Gronemeyer H.** TIF2, a 160 kDa transcriptional mediator for the ligand-dependent activation function AF-2 of nuclear receptors. *EMBO J.* 1996;15(14):3667–3675.

84. **Hong H, Kohli K, Trivedi A, Johnson DL, Stallcup MR.** GRIP1, a novel mouse protein that serves as a transcriptional coactivator in yeast for the hormone binding domains of steroid receptors. *Proc. Natl. Acad. Sci. U.S.A.* 1996;93(10):4948–4952.

85. **Anzick SL, Kononen J, Walker RL, Azorsa DO, Tanner MM, Guan XY, Sauter G, Kallioniemi OP, Trent JM, Meltzer PS.** AIB1, a steroid receptor coactivator amplified in breast and ovarian cancer. *Science* 1997;277(5328):965–968.

86. **Chen H, Lin RJ, Schiltz RL, Chakravarti D, Nash A, Nagy L, Privalsky ML, Nakatani Y, Evans RM.** Nuclear receptor coactivator ACTR is a novel histone acetyltransferase and forms a multimeric activation complex with P/CAF and CBP/p300. *Cell* 1997;90(3):569–580.

87. **Li H, Gomes PJ, Chen JD.** RAC3, a steroid/nuclear receptor-associated coactivator that is related to SRC-1 and TIF2. *Proc. Natl. Acad. Sci. U.S.A.* 1997;94(16):8479–8484.

88. **Hanstein B, Eckner R, DiRenzo J, Halachmi S, Liu H, Searcy B, Kurokawa R, Brown M.** p300 is a component of an estrogen receptor coactivator complex. *Proc. Natl. Acad. Sci. U.S.A.* 1996;93(21):11540–11545.

89. **McKenna NJ, Xu J, Nawaz Z, Tsai SY, Tsai MJ, O'Malley BW.** Nuclear receptor coactivators: multiple enzymes, multiple complexes, multiple functions. *J. Steroid Biochem. Mol. Biol.* 1999;69(1–6):3–12.

90. **Stenoien DL, Nye AC, Mancini MG, Patel K, Dutertre M, O'Malley BW, Smith CL, Belmont AS, Mancini MA.** Ligand-mediated assembly and real-time cellular dynamics of estrogen receptor alpha-coactivator complexes in living cells. *Mol. Cell. Biol.* 2001;21(13):4404–4412.

91. **Onate SA, Boonyaratanakornkit V, Spencer TE, Tsai SY, Tsai MJ, Edwards DP, O'Malley BW.** The steroid receptor coactivator-1 contains multiple receptor interacting and activation domains that cooperatively enhance the activation function 1 (AF1) and AF2 domains of steroid receptors. *J. Biol. Chem.* 1998;273(20):12101–12108.

92. **Webb P, Nguyen P, Shinsako J, Anderson C, Feng W, Nguyen MP, Chen D, Huang SM, Subramanian S, McKinerney E, Katzenellenbogen BS, Stallcup MR, Kushner PJ.** Estrogen receptor activation function 1 works by binding p160 coactivator proteins. *Mol. Endocrinol.* 1998;12(10):1605–1618.

93. **Kobayashi Y, Kitamoto T, Masuhiro Y, Watanabe M, Kase T, Metzger D, Yanagisawa J, Kato S.** p300 mediates functional synergism between AF-1 and AF-2 of estrogen receptor alpha and beta by interacting directly with the N-terminal A/B domains. *J. Biol. Chem.* 2000;275(21):15645–15651.

94. **Benecke A, Chambon P, Gronemeyer H.** Synergy between estrogen receptor alpha activation functions AF1 and AF2 mediated by transcription intermediary factor TIF2. *EMBO Rep.* 2000;1(2):151–157.

95. **McInerney EM, Tsai MJ, O'Malley BW, Katzenellenbogen BS.** Analysis of estrogen receptor transcriptional enhancement by a nuclear hormone receptor coactivator. *Proc. Natl. Acad. Sci. U.S.A.* 1996;93(19):10069–10073.

96. **Kraus WL, McInerney EM, Katzenellenbogen BS.** Ligand-dependent, transcriptionally productive association of the amino- and carboxyl-terminal regions of a steroid hormone nuclear receptor. *Proc. Natl. Acad. Sci. U.S.A.* 1995;92(26):12314–12318.

97. **Gburcik V, Bot N, Maggiolini M, Picard D.** SPBP is a phosphoserine-specific repressor of estrogen receptor alpha. *Mol. Cell. Biol.* 2005;25(9):3421–3430.

98. **Fowler AM, Solodin N, Preisler-Mashek MT, Zhang P, Lee AV, Alarid ET.** Increases in estrogen receptor-α concentration in breast cancer cells promote serine 118/104/106-independent AF-1 transactivation and growth in the absence of estrogen. *FASEB J* 2004;18(1):81–93.

99. **Alarid ET, Bakopoulos N, Solodin N.** Proteasome-mediated proteolysis of estrogen receptor: a novel component in autologous down-regulation. *Mol. Endocrinol.* 1999;13(9):1522–1534.

100. **Nawaz Z, Lonard DM, Dennis AP, Smith CL, O'Malley BW.** Proteasome-dependent degradation of the human estrogen receptor. *PNAS* 1999;96(5):1858–1862.

101. **El Khissiin A, Leclercq G.** Implication of proteasome in estrogen receptor degradation. *FEBS Letters* 1999;448(1):160–166.

102. **Preisler-Mashek MT, Solodin N, Stark BL, Tyriver MK, Alarid ET.** Ligand-specific regulation of proteasome-mediated proteolysis of estrogen receptor-alpha. *Am. J. Physiol. Endocrinol. Metab.* 2002;282(4):E891-898.

103. **Read LD, Greene GL, Katzenellenbogen BS.** Regulation of estrogen receptor messenger ribonucleic acid and protein levels in human breast cancer cell lines by sex steroid hormones, their antagonists, and growth factors. *Mol. Endocrinol.* 1989;3(2):295–304.

104. **Saceda M, Lippman ME, Chambon P, Lindsey RL, Ponglikitmongkol M, Puente M, Martin MB.** Regulation of the estrogen receptor in MCF-7 cells by estradiol. *Mol. Endocrinol.* 1988;2(12):1157–1162.

105. **Saceda M, Lindsey RK, Solomon H, Angeloni SV, Martin MB.** Estradiol regulates estrogen receptor mRNA stability. *J. Steroid Biochem. Mol. Biol.* 1998;66(3):113–120.

106. **Ellison-Zelski SJ, Solodin NM, Alarid ET.** Repression of ESR1 through actions of estrogen receptor alpha and Sin3A at the proximal promoter. *Mol. Cell. Biol.* 2009;29(18):4949–4958.

107. **Tian D, Solodin NM, Rajbhandari P, Bjorklund K, Alarid ET, Kreeger PK.** A kinetic model identifies phosphorylated estrogen receptor-α (ERα) as a critical regulator of ERα dynamics in breast cancer. *FASEB J.* 2015;29(5):2022–2031.

108. **Rajbhandari P, Finn G, Solodin NM, Singarapu KK, Sahu SC, Markley JL, Kadunc KJ, Ellison-Zelski SJ, Kariagina A, Haslam SZ, Lu KP, Alarid ET.** Regulation of Estrogen Receptor α N-Terminus Conformation and Function by Peptidyl Prolyl Isomerase Pin1. *Mol. Cell. Biol.* 2012;32(2):445–457.

109. **Rajbhandari P, Schalper KA, Solodin NM, Ellison-Zelski SJ, Ping Lu K, Rimm DL, Alarid ET.** Pin1 modulates ERα levels in breast cancer through inhibition of phosphorylation-dependent ubiquitination and degradation. *Oncogene* 2014;33(11):1438–1447.

110. **Rajbhandari P, Ozers MS, Solodin NM, Warren CL, Alarid ET.** Peptidylprolyl Isomerase Pin1 Directly Enhances the DNA Binding Functions of Estrogen Receptor α. *J. Biol. Chem.* 2015;290(22):13749–13762.

111. **Maurer HR, Chalkley GR.** Some properties of a nuclear binding site of estradiol. *J. Mol. Biol.* 1967;27(3):431–441.

112. **Teng CS, Hamilton TH.** The role of chromatin in estrogen action in the uterus, I. The control of template capacity and chemical composition and the binding of H3-estradiol-17 beta. *Proc. Natl. Acad. Sci. U.S.A.* 1968;60(4):1410–1417.

113. **Giannopoulos G, Gorski J.** Estrogen-binding protein of the rat uterus. Different molecular forms associated with nuclear uptake of estradiol. *J. Biol. Chem.* 1971;246(8):2530–2536.

114. **King RJ, Gordon J.** Involvement of DNA in the acceptor mechanism for uterine oestradiol receptor. *Nature New Biol.* 1972;240(101):185–187.

115. **Toft D.** The interaction of uterine estrogen receptors with DNA. *J. Steroid Biochem.* 1972;3(3):515–522.

116. **Skafar DF, Notides AC.** Modulation of the estrogen receptor's affinity for DNA by estradiol. *J. Biol. Chem.* 1985;260(22):12208–12213.

117. **Kushner PJ, Agard DA, Greene GL, Scanlan TS, Shiau AK, Uht RM, Webb P.** Estrogen receptor pathways to AP-1. *J. Steroid Biochem. Mol. Biol.* 2000;74(5):311–317.

118. **Cheung E, Acevedo ML, Cole PA, Kraus WL.** Altered pharmacology and distinct coactivator usage for estrogen receptor-dependent transcription through activating protein-1. *Proc. Natl. Acad. Sci. U.S.A.* 2005;102(3):559–564.

119. **Heldring N, Isaacs GD, Diehl AG, Sun M, Cheung E, Ranish JA, Kraus WL.** Multiple sequence-specific DNA-binding proteins mediate estrogen receptor signaling through a tethering pathway. *Mol. Endocrinol.* 2011;25(4):564–574.

120. **Porter W, Saville B, Hoivik D, Safe S.** Functional synergy between the transcription factor Sp1 and the estrogen receptor. *Mol. Endocrinol.* 1997;11(11):1569–1580.

121. **Stender JD, Kim K, Charn TH, Komm B, Chang KCN, Kraus WL, Benner C, Glass CK, Katzenellenbogen BS.** Genome-wide analysis of estrogen receptor alpha DNA binding and tethering mechanisms identifies Runx1 as a novel tethering factor in receptor-mediated transcriptional activation. *Mol. Cell. Biol.* 2010;30(16):3943–3955.

122. **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;16(2):647–663.

123. **Peale FV, Ludwig LB, Zain S, Hilf R, Bambara RA.** Properties of a high-affinity DNA binding site for estrogen receptor. *Proc. Natl. Acad. Sci. U.S.A.* 1988;85(4):1038–1042.

124. **Murdoch FE, Byrne LM, Ariazi EA, Furlow JD, Meier DA, Gorski J.** Estrogen receptor binding to DNA: affinity for nonpalindromic elements from the rat prolactin gene. *Biochemistry* 1995;34(28):9144–9150.

125. **D'haeseleer P.** What are DNA sequence motifs? *Nat. Biotechnol.* 2006;24(4):423–425.

126. **Bourdeau V, Deschênes J, Métivier R, Nagai Y, Nguyen D, Bretschneider N, Gannon F, White JH, Mader S.** Genome-wide identification of high-affinity estrogen response elements in human and mouse. *Mol. Endocrinol.* 2004;18(6):1411–1427.

127. **Coons LA, Hewitt SC, Burkholder AB, McDonnell DP, Korach KS.** DNA Sequence Constraints Define Functionally Active Steroid Nuclear Receptor Binding Sites in Chromatin. *Endocrinology* 2017;158(10):3212–3234.

128. **Kumar V, Chambon P.** The estrogen receptor binds tightly to its responsive element as a ligand-induced homodimer. *Cell* 1988;55(1):145–156.

129. **Klinge CM, Peale FV, Hilf R, Bambara RA, Zain S.** Cooperative estrogen receptor interaction with consensus or variant estrogen responsive elements in vitro. *Cancer Res.* 1992;52(5):1073–1081.

130. **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;75(3):567–578.

131. **Näär AM, Boutin JM, Lipkin SM, Yu VC, Holloway JM, Glass CK, Rosenfeld MG.** The orientation and spacing of core DNA-binding motifs dictate selective transcriptional responses to three nuclear receptors. *Cell* 1991;65(7):1267–1279.

132. **Anolik JH, Klinge CM, Hilf R, Bambara RA.** Cooperative binding of estrogen receptor to DNA depends on spacing of binding sites, flanking sequence, and ligand. *Biochemistry* 1995;34(8):2511–2520.

133. **Driscoll MD, Sathya G, Muyan M, Klinge CM, Hilf R, Bambara RA.** Sequence requirements for estrogen receptor binding to estrogen response elements. *J. Biol. Chem.* 1998;273(45):29321–29330.

134. **Ozers MS, Hill JJ, Ervin K, Wood JR, Nardulli AM, Royer CA, Gorski J.** Equilibrium binding of estrogen receptor with DNA using fluorescence anisotropy. *J. Biol. Chem.* 1997;272(48):30405–30411.

135. **Likhite VS, Stossi F, Kim K, Katzenellenbogen BS, Katzenellenbogen JA.** Kinase-Specific Phosphorylation of the Estrogen Receptor Changes Receptor Interactions with Ligand,

Deoxyribonucleic Acid, and Coregulators Associated with Alterations in Estrogen and Tamoxifen Activity. *Molecular Endocrinology* 2006;20(12):3120–3132.

136. **Weitsman GE, Li L, Skliris GP, Davie JR, Ung K, Niu Y, Curtis-Snell L, Tomes L, Watson PH, Murphy LC.** Estrogen receptor-alpha phosphorylated at Ser118 is present at the promoters of estrogen-regulated genes and is not altered due to HER-2 overexpression. *Cancer Res.* 2006;66(20):10162–10170.

137. **Arnold SF, Obourn JD, Jaffe H, Notides AC.** Phosphorylation of the human estrogen receptor by mitogen-activated protein kinase and casein kinase II: consequence on DNA binding. *J. Steroid Biochem. Mol. Biol.* 1995;55(2):163–172.

138. **Arnold SF, Vorojeikina DP, Notides AC.** Phosphorylation of tyrosine 537 on the human estrogen receptor is required for binding to an estrogen response element. *J. Biol. Chem.* 1995;270(50):30205–30212.

139. **Sheeler CQ, Singleton DW, Khan SA.** Mutation of serines 104, 106, and 118 inhibits dimerization of the human estrogen receptor in yeast. *Endocr. Res.* 2003;29(2):237–255.

140. **Lupien M, Meyer CA, Bailey ST, Eeckhoute J, Cook J, Westerling T, Zhang X, Carroll JS, Rhodes DR, Liu XS, Brown M.** Growth factor stimulation induces a distinct ER(alpha) cistrome underlying breast cancer endocrine resistance. *Genes Dev.* 2010;24(19):2219–2227.

141. **Gilmour DS, Lis JT.** Detecting protein-DNA interactions in vivo: distribution of RNA polymerase on specific bacterial genes. *Proc. Natl. Acad. Sci. U.S.A.* 1984;81(14):4275–4279.

142. **Gilmour DS, Lis JT.** In vivo interactions of RNA polymerase II with genes of Drosophila melanogaster. *Mol. Cell. Biol.* 1985;5(8):2009–2018.

143. **Chen D, Pace PE, Coombes RC, Ali S.** Phosphorylation of human estrogen receptor alpha by protein kinase A regulates dimerization. *Mol. Cell. Biol.* 1999;19(2):1002–1015.

144. **Shang Y, Hu X, DiRenzo J, Lazar MA, Brown M.** Cofactor dynamics and sufficiency in estrogen receptor-regulated transcription. *Cell* 2000;103(6):843–852.

145. **Métivier R, Penot G, Hübner MR, Reid G, Brand H, Kos M, Gannon F.** Estrogen receptor-alpha directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter. *Cell* 2003;115(6):751–763.

146. **Holding AN, Cullen AE, Markowetz F.** Genome-wide Estrogen Receptor-α activation is sustained, not cyclical. *Elife* 2018;7. doi:10.7554/eLife.40854.

147. **Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, Halpern A, Hannenhalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-Threideh J, Beasley E, Biddick K, Bonazzi V, Brandon R, Cargill M, Chandramouliswaran I, Charlab R, Chaturvedi K, Deng Z, Di Francesco V, Dunn P, Eilbeck K, Evangelista C, Gabrielian AE, Gan W, Ge W, Gong F, Gu Z, Guan P, Heiman TJ, Higgins ME, Ji RR, Ke Z, Ketchum KA, Lai Z, Lei Y, Li Z, Li J, Liang Y, Lin X, Lu F, Merkulov GV, Milshina N, Moore HM, Naik AK, Narayan VA, Neelam B, Nusskern D, Rusch DB, Salzberg S, Shao W, Shue B, Sun J, Wang Z, Wang A, Wang X, Wang J, Wei M, Wides R, Xiao C, Yan C, Yao A, Ye J, Zhan M, Zhang W, Zhang H, Zhao Q, Zheng L, Zhong F, Zhong W, Zhu S, Zhao**

S, Gilbert D, Baumhueter S, Spier G, Carter C, Cravchik A, Woodage T, Ali F, An H, Awe A, Baldwin D, Baden H, Barnstead M, Barrow I, Beeson K, Busam D, Carver A, Center A, Cheng ML, Curry L, Danaher S, Davenport L, Desilets R, Dietz S, Dodson K, Doup L, Ferriera S, Garg N, Gluecksmann A, Hart B, Haynes J, Haynes C, Heiner C, Hladun S, Hostin D, Houck J, Howland T, Ibegwam C, Johnson J, Kalush F, Kline L, Koduru S, Love A, Mann F, May D, McCawley S, McIntosh T, McMullen I, Moy M, Moy L, Murphy B, Nelson K, Pfannkoch C, Pratts E, Puri V, Qureshi H, Reardon M, Rodriguez R, Rogers YH, Romblad D, Ruhfel B, Scott R, Sitter C, Smallwood M, Stewart E, Strong R, Suh E, Thomas R, Tint NN, Tse S, Vech C, Wang G, Wetter J, Williams S, Williams M, Windsor S, Winn-Deen E, Wolfe K, Zaveri J, Zaveri K, Abril JF, Guigó R, Campbell MJ, Sjolander KV, Karlak B, Kejariwal A, Mi H, Lazareva B, Hatton T, Narechania A, Diemer K, Muruganujan A, Guo N, Sato S, Bafna V, Istrail S, Lippert R, Schwartz R, Walenz B, Yooseph S, Allen D, Basu A, Baxendale J, Blick L, Caminha M, Carnes-Stine J, Caulk P, Chiang YH, Coyne M, Dahlke C, Mays A, Dombroski M, Donnelly M, Ely D, Esparham S, Fosler C, Gire H, Glanowski S, Glasser K, Glodek A, Gorokhov M, Graham K, Gropman B, Harris M, Heil J, Henderson S, Hoover J, Jennings D, Jordan C, Jordan J, Kasha J, Kagan L, Kraft C, Levitsky A, Lewis M, Liu X, Lopez J, Ma D, Majoros W, McDaniel J, Murphy S, Newman M, Nguyen T, Nguyen N, Nodell M, Pan S, Peck J, Peterson M, Rowe W, Sanders R, Scott J, Simpson M, Smith T, Sprague A, Stockwell T, Turner R, Venter E, Wang M, Wen M, Wu D, Wu M, Xia A, Zandieh A, Zhu X. The sequence of the human genome. *Science* 2001;291(5507):1304–1351.

148. **Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann Y, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, de la Bastide M, Dedhia N, Blöcker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kaspryzk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AF, Stupka E, Szustakowki J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Patrinos A, Morgan MJ, de Jong P, Catanese JJ, Osoegawa K, Shizuya H, Choi S, Chen YJ, Szustakowki J, International Human Genome Sequencing Consortium.** Initial sequencing and analysis of the human genome. *Nature* 2001;409(6822):860–921.

149. **Blat Y, Kleckner N.** Cohesins bind to preferential sites along yeast chromosome III, with differential regulation along arms versus the centric region. *Cell* 1999;98(2):249–259.

150. **Ren B, Robert F, Wyrick JJ, Aparicio O, Jennings EG, Simon I, Zeitlinger J, Schreiber J, Hannett N, Kanin E, Volkert TL, Wilson CJ, Bell SP, Young RA.** Genome-wide location and function of DNA binding proteins. *Science* 2000;290(5500):2306–2309.

151. **Weinmann AS, Yan PS, Oberley MJ, Huang TH-M, Farnham PJ.** Isolating human transcription factor targets by coupling chromatin immunoprecipitation and CpG island microarray analysis. *Genes Dev.* 2002;16(2):235–244.

152. **Laganière J, Deblois G, Lefebvre C, Bataille AR, Robert F, Giguère V.** From the Cover: Location analysis of estrogen receptor alpha target promoters reveals that FOXA1 defines a domain of the estrogen response. *Proc. Natl. Acad. Sci. U.S.A.* 2005;102(33):11651–11656.

153. **Kininis M, Chen BS, Diehl AG, Isaacs GD, Zhang T, Siepel AC, Clark AG, Kraus WL.** Genomic analyses of transcription factor binding, histone acetylation, and gene expression reveal mechanistically distinct classes of estrogen-regulated promoters. *Mol. Cell. Biol.* 2007;27(14):5090–5104.

154. **Cheng ASL, Jin VX, Fan M, Smith LT, Liyanarachchi S, Yan PS, Leu Y-W, Chan MWY, Plass C, Nephew KP, Davuluri RV, Huang TH-M.** Combinatorial analysis of transcription factor partners reveals recruitment of c-MYC to estrogen receptor-alpha responsive promoters. *Mol. Cell* 2006;21(3):393–404.

155. **Sathya G, Li W, Klinge CM, Anolik JH, Hilf R, Bambara RA.** Effects of multiple estrogen responsive elements, their spacing, and location on estrogen response of reporter genes. *Mol. Endocrinol.* 1997;11(13):1994–2003.

156. **Carroll JS, Liu XS, Brodsky AS, Li W, Meyer CA, Szary AJ, Eeckhoute J, Shao W, Hestermann EV, Geistlinger TR, Fox EA, Silver PA, Brown M.** Chromosome-Wide Mapping of Estrogen Receptor Binding Reveals Long-Range Regulation Requiring the Forkhead Protein FoxA1. *Cell* 2005;122(1):33–43.

157. **Cirillo LA, Lin FR, Cuesta I, Friedman D, Jarnik M, Zaret KS.** Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol. Cell* 2002;9(2):279–289.

158. **Lupien M, Eeckhoute J, Meyer CA, Wang Q, Zhang Y, Li W, Carroll JS, Liu XS, Brown M.** FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. *Cell* 2008;132(6):958–970.

159. **Carroll JS, Meyer CA, Song J, Li W, Geistlinger TR, Eeckhoute J, Brodsky AS, Keeton EK, Fertuck KC, Hall GF, Wang Q, Bekiranov S, Sementchenko V, Fox EA, Silver PA, Gingeras TR, Liu XS, Brown M.** Genome-wide analysis of estrogen receptor binding sites. *Nat. Genet.* 2006;38(11):1289–1297.

160. **Krum SA, Miranda-Carboni GA, Lupien M, Eeckhoute J, Carroll JS, Brown M.** Unique ERalpha cistromes control cell type-specific gene regulation. *Mol. Endocrinol.* 2008;22(11):2393–2406.

161. **Hurtado A, Holmes KA, Ross-Innes CS, Schmidt D, Carroll JS.** FOXA1 is a key determinant of estrogen receptor function and endocrine response. *Nat. Genet.* 2011;43(1):27–33.

162. **Bhat-Nakshatri P, Wang G, Appaiah H, Luktuke N, Carroll JS, Geistlinger TR, Brown M, Badve S, Liu Y, Nakshatri H.** AKT alters genome-wide estrogen receptor alpha binding and impacts estrogen signaling in breast cancer. *Mol. Cell. Biol.* 2008;28(24):7487–7503.

163. **Barone I, Iacopetta D, Covington KR, Cui Y, Tsimelzon A, Beyer A, Andò S, Fuqua S a. W.** Phosphorylation of the mutant K303R estrogen receptor alpha at serine 305 affects aromatase inhibitor sensitivity. *Oncogene* 2010;29(16):2404–2414.

164. **Campbell RA, Bhat-Nakshatri P, Patel NM, Constantinidou D, Ali S, Nakshatri H.** Phosphatidylinositol 3-kinase/AKT-mediated activation of estrogen receptor alpha: a new model for anti-estrogen resistance. *J. Biol. Chem.* 2001;276(13):9817–9824.

165. **Park S, Song J, Joe CO, Shin I.** Akt stabilizes estrogen receptor alpha with the concomitant reduction in its transcriptional activity. *Cell. Signal.* 2008;20(7):1368–1374.

166. **Manning BD, Toker A.** AKT/PKB Signaling: Navigating the Network. *Cell* 2017;169(3):381–405.

167. **Wei C-L, Wu Q, Vega VB, Chiu KP, Ng P, Zhang T, Shahab A, Yong HC, Fu Y, Weng Z, Liu J, Zhao XD, Chew J-L, Lee YL, Kuznetsov VA, Sung W-K, Miller LD, Lim B, Liu ET, Yu Q, Ng H-H, Ruan Y.** A global map of p53 transcription-factor binding sites in the human genome. *Cell* 2006;124(1):207–219.

168. **Trapnell C, Salzberg SL.** How to map billions of short reads onto genomes. *Nat. Biotechnol.* 2009;27(5):455–457.

169. **Whiteford N, Haslam N, Weber G, Prügel-Bennett A, Essex JW, Roach PL, Bradley M, Neylon C.** An analysis of the feasibility of short read sequencing. *Nucleic Acids Res.* 2005;33(19):e171.

170. **Li R, Li Y, Kristiansen K, Wang J.** SOAP: short oligonucleotide alignment program. *Bioinformatics* 2008;24(5):713–714.

171. **Li H, Ruan J, Durbin R.** Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res.* 2008;18(11):1851–1858.

172. **Langmead B, Trapnell C, Pop M, Salzberg SL.** Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 2009;10(3):R25.

173. **Langmead B, Salzberg SL.** Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 2012;9(4):357–359.

174. **Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS.** Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 2008;9(9):R137.

175. **Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK.** Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* 2010;38(4):576–589.

176. **Fejes AP, Robertson G, Bilenky M, Varhol R, Bainbridge M, Jones SJM.** FindPeaks 3.1: a tool for identifying areas of enrichment from massively parallel short-read sequencing technology. *Bioinformatics* 2008;24(15):1729–1730.

177. **Jothi R, Cuddapah S, Barski A, Cui K, Zhao K.** Genome-wide identification of in vivo protein-DNA binding sites from ChIP-Seq data. *Nucleic Acids Res.* 2008;36(16):5221–5231.

178. **Valouev A, Johnson DS, Sundquist A, Medina C, Anton E, Batzoglou S, Myers RM, Sidow A.** Genome-wide analysis of transcription factor binding sites based on ChIP-Seq data. *Nat. Methods* 2008;5(9):829–834.

179. **Rozowsky J, Euskirchen G, Auerbach RK, Zhang ZD, Gibson T, Bjornson R, Carriero N, Snyder M, Gerstein MB.** PeakSeq enables systematic scoring of ChIP-seq experiments relative to controls. *Nat. Biotechnol.* 2009;27(1):66–75.

180. **Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D.** The human genome browser at UCSC. *Genome Res.* 2002;12(6):996–1006.

181. **Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS.** MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* 2009;37(Web Server issue):W202-208.

182. **Grant CE, Bailey TL, Noble WS.** FIMO: scanning for occurrences of a given motif. *Bioinformatics* 2011;27(7):1017–1018.

183. **Bailey TL.** DREME: motif discovery in transcription factor ChIP-seq data. *Bioinformatics* 2011;27(12):1653–1659.

184. **Bailey TL, Machanick P.** Inferring direct DNA binding from ChIP-seq. *Nucleic Acids Res.* 2012;40(17):e128.

185. **Fullwood MJ, Liu MH, Pan YF, Liu J, Xu H, Mohamed YB, Orlov YL, Velkov S, Ho A, Mei PH, Chew EGY, Huang PYH, Welboren W-J, Han Y, Ooi HS, Ariyaratne PN, Vega VB, Luo Y, Tan PY, Choy PY, Wansa KDSA, Zhao B, Lim KS, Leow SC, Yow JS, Joseph R, Li H, Desai KV, Thomsen JS, Lee YK, Karuturi RKM, Herve T, Bourque G, Stunnenberg HG, Ruan X, Cacheux-Rataboul V, Sung W-K, Liu ET, Wei C-L, Cheung E, Ruan Y.** An oestrogen-receptor-alpha-bound human chromatin interactome. *Nature* 2009;462(7269):58–64.

186. **McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G.** GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* 2010;28(5):495–501.

187. **Lin C-Y, Vega VB, Thomsen JS, Zhang T, Kong SL, Xie M, Chiu KP, Lipovich L, Barnett DH, Stossi F, Yeo A, George J, Kuznetsov VA, Lee YK, Charn TH, Palanisamy N, Miller LD, Cheung E, Katzenellenbogen BS, Ruan Y, Bourque G, Wei C-L, Liu ET.** Whole-genome cartography of estrogen receptor alpha binding sites. *PLoS Genet.* 2007;3(6):e87.

188. **Welboren W-J, van Driel MA, Janssen-Megens EM, van Heeringen SJ, Sweep FC, Span PN, Stunnenberg HG.** ChIP-Seq of ERalpha and RNA polymerase II defines genes differentially responding to ligands. *EMBO J.* 2009;28(10):1418–1428.

189. **Zwart W, Theodorou V, Kok M, Canisius S, Linn S, Carroll JS.** Oestrogen receptor-co-factor-chromatin specificity in the transcriptional regulation of breast cancer. *EMBO J.* 2011;30(23):4764–4776.

190. **Joseph R, Orlov YL, Huss M, Sun W, Kong SL, Ukil L, Pan YF, Li G, Lim M, Thomsen JS, Ruan Y, Clarke ND, Prabhakar S, Cheung E, Liu ET.** Integrative model of genomic factors for determining binding site selection by estrogen receptor-α. *Mol. Syst. Biol.* 2010;6:456.

191. **Yu J, Yu J, Mani R-S, Cao Q, Brenner CJ, Cao X, Wang X, Wu L, Li J, Hu M, Gong Y, Cheng H, Laxman B, Vellaichamy A, Shankar S, Li Y, Dhanasekaran SM, Morey R, Barrette T, Lonigro RJ, Tomlins SA, Varambally S, Qin ZS, Chinnaiyan AM.** An integrated network of

androgen receptor, polycomb, and TMPRSS2-ERG gene fusions in prostate cancer progression. *Cancer Cell* 2010;17(5):443–454.

192. **Reddy TE, Pauli F, Sprouse RO, Neff NF, Newberry KM, Garabedian MJ, Myers RM.** Genomic determination of the glucocorticoid response reveals unexpected mechanisms of gene regulation. *Genome Res.* 2009;19(12):2163–2171.

193. **Li W, Notani D, Ma Q, Tanasa B, Nunez E, Chen AY, Merkurjev D, Zhang J, Ohgi K, Song X, Oh S, Kim H-S, Glass CK, Rosenfeld MG.** Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature* 2013;498(7455):516–520.

194. **Hah N, Murakami S, Nagari A, Danko CG, Kraus WL.** Enhancer transcripts mark active estrogen receptor binding sites. *Genome Res.* 2013;23(8):1210–1223.

195. **Lam MTY, Li W, Rosenfeld MG, Glass CK.** Enhancer RNAs and regulated transcriptional programs. *Trends Biochem. Sci.* 2014;39(4):170–182.

196. **Kong SL, Li G, Loh SL, Sung W-K, Liu ET.** Cellular reprogramming by the conjoint action of ERα, FOXA1, and GATA3 to a ligand-inducible growth state. *Mol. Syst. Biol.* 2011;7:526.

197. **Theodorou V, Stark R, Menon S, Carroll JS.** GATA3 acts upstream of FOXA1 in mediating ESR1 binding by shaping enhancer accessibility. *Genome Res.* 2013;23(1):12–22.

198. **Frietze S, O'Geen H, Littlepage LE, Simion C, Sweeney CA, Farnham PJ, Krig SR.** Global analysis of ZNF217 chromatin occupancy in the breast cancer cell genome reveals an association with ERalpha. *BMC Genomics* 2014;15:520.

199. **Miranda TB, Voss TC, Sung M-H, Baek S, John S, Hawkins M, Grøntved L, Schiltz RL, Hager GL.** Reprogramming the chromatin landscape: interplay of the estrogen and glucocorticoid receptors at the genomic level. *Cancer Res.* 2013;73(16):5130–5139.

200. **Sanders DA, Ross-Innes CS, Beraldi D, Carroll JS, Balasubramanian S.** Genome-wide mapping of FOXM1 binding reveals co-binding with estrogen receptor alpha in breast cancer cells. *Genome Biol.* 2013;14(1):R6.

201. **Mohammed H, D'Santos C, Serandour AA, Ali HR, Brown GD, Atkins A, Rueda OM, Holmes KA, Theodorou V, Robinson JLL, Zwart W, Saadi A, Ross-Innes CS, Chin S-F, Menon S, Stingl J, Palmieri C, Caldas C, Carroll JS.** Endogenous purification reveals GREB1 as a key estrogen receptor regulatory factor. *Cell Rep* 2013;3(2):342–349.

202. **Murphy LC, Niu Y, Snell L, Watson P.** Phospho-serine-118 estrogen receptor-alpha expression is associated with better disease outcome in women treated with tamoxifen. *Clin. Cancer Res.* 2004;10(17):5902–5906.

203. **Bergqvist J, Elmberger G, Ohd J, Linderholm B, Bjohle J, Hellborg H, Nordgren H, Borg A-L, Skoog L, Bergh J.** Activated ERK1/2 and phosphorylated oestrogen receptor alpha are associated with improved breast cancer survival in women treated with tamoxifen. *Eur. J. Cancer* 2006;42(8):1104–1112.

204. **Kok M, Holm-Wigerup C, Hauptmann M, Michalides R, Stål O, Linn S, Landberg G.** Estrogen Receptor-α Phosphorylation at Serine-118 and Tamoxifen Response in Breast Cancer. *JNCI J Natl Cancer Inst* 2009;101(24):1725–1729.

205. **Zoubir M, Mathieu MC, Mazouni C, Liedtke C, Corley L, Geha S, Bouaziz J, Spielmann M, Drusche F, Symmans WF, Delaloge S, Andre F.** Modulation of ER phosphorylation on serine

118 by endocrine therapy: a new surrogate marker for efficacy. *Ann. Oncol.* 2008;19(8):1402–1406.

206. **Chen M, Cui Y-K, Huang W-H, Man K, Zhang G-J.** Phosphorylation of estrogen receptor α at serine 118 is correlated with breast cancer resistance to tamoxifen. *Oncol Lett* 2013;6(1):118–124.

207. **Szijgyarto Z, Flach KD, Opdam M, Palmieri C, Linn SC, Wesseling J, Ali S, Bliss JM, Cheang MCU, Zwart W, Coombes RC.** Dissecting the predictive value of MAPK/AKT/estrogen-receptor phosphorylation axis in primary breast cancer to treatment response for tamoxifen over exemestane: a Translational Report of the Intergroup Exemestane Study (IES)-PathIES. *Breast Cancer Res. Treat.* 2019. doi:10.1007/s10549-018-05110-x.

208. **Weis KE, Ekena K, Thomas JA, Lazennec G, Katzenellenbogen BS.** Constitutively active human estrogen receptors containing amino acid substitutions for tyrosine 537 in the receptor protein. *Mol. Endocrinol.* 1996;10(11):1388–1398.

209. **Zhang Q-X, Borg Å, Wolf DM, Oesterreich S, Fuqua SAW.** An Estrogen Receptor Mutant with Strong Hormone-independent Activity from a Metastatic Breast Cancer. *Cancer Res* 1997;57(7):1244–1249.

210. **Toy W, Shen Y, Won H, Green B, Sakr RA, Will M, Li Z, Gala K, Fanning S, King TA, Hudis C, Chen D, Taran T, Hortobagyi G, Greene G, Berger M, Baselga J, Chandarlapaty S.** ESR1 ligand-binding domain mutations in hormone-resistant breast cancer. *Nat. Genet.* 2013;45(12):1439–1445.

211. **Robinson DR, Wu Y-M, Vats P, Su F, Lonigro RJ, Cao X, Kalyana-Sundaram S, Wang R, Ning Y, Hodges L, Gursky A, Siddiqui J, Tomlins SA, Roychowdhury S, Pienta KJ, Kim SY, Roberts JS, Rae JM, Van Poznak CH, Hayes DF, Chugh R, Kunju LP, Talpaz M, Schott AF, Chinnaiyan AM.** Activating ESR1 mutations in hormone-resistant metastatic breast cancer. *Nat Genet* 2013;45(12):1446–1451.

212. **Merenbakh-Lamin K, Ben-Baruch N, Yeheskel A, Dvir A, Soussan-Gutman L, Jeselsohn R, Yelensky R, Brown M, Miller VA, Sarid D, Rizel S, Klein B, Rubinek T, Wolf I.** D538G Mutation in Estrogen Receptor-α: A Novel Mechanism for Acquired Endocrine Resistance in Breast Cancer. *Cancer Res* 2013;73(23):6856–6864.

213. **Jeselsohn R, Yelensky R, Buchwalter G, Frampton G, Meric-Bernstam F, Gonzalez-Angulo AM, Ferrer-Lozano J, Perez-Fidalgo JA, Cristofanilli M, Gómez H, Arteaga CL, Giltnane J, Balko JM, Cronin MT, Jarosz M, Sun J, Hawryluk M, Lipson D, Otto G, Ross JS, Dvir A, Soussan-Gutman L, Wolf I, Rubinek T, Gilmore L, Schnitt S, Come SE, Pusztai L, Stephens P, Brown M, Miller VA.** Emergence of Constitutively Active Estrogen Receptor-α Mutations in Pretreated Advanced Estrogen Receptor–Positive Breast Cancer. *Clin Cancer Res* 2014;20(7):1757–1767.

214. **Toy W, Weir H, Razavi P, Lawson M, Goeppert AU, Mazzola AM, Smith A, Wilson J, Morrow C, Wong WL, De Stanchina E, Carlson KE, Martin TS, Uddin S, Li Z, Fanning S, Katzenellenbogen JA, Greene G, Baselga J, Chandarlapaty S.** Activating ESR1 Mutations Differentially Affect the Efficacy of ER Antagonists. *Cancer Discov* 2016. doi:10.1158/2159-8290.CD-15-1523.

215. **Bahreini A, Li Z, Wang P, Levine KM, Tasdemir N, Cao L, Weir HM, Puhalla SL, Davidson NE, Stern AM, Chu D, Park BH, Lee AV, Oesterreich S.** Mutation site and context dependent effects of ESR1 mutation in genome-edited breast cancer cell models. *Breast Cancer Res.* 2017;19(1):60.

216. **Harrod A, Fulton J, Nguyen VTM, Periyasamy M, Ramos-Garcia L, Lai C-F, Metodieva G, de Giorgio A, Williams RL, Santos DB, Gomez PJ, Lin M-L, Metodiev MV, Stebbing J, Castellano L, Magnani L, Coombes RC, Buluwela L, Ali S.** Genomic modelling of the ESR1 Y537S mutation for evaluating function and new therapeutic approaches for metastatic breast cancer. *Oncogene* 2017;36(16):2286–2296.

217. **Jeselsohn R, Bergholz JS, Pun M, Cornwell M, Liu W, Nardone A, Xiao T, Li W, Qiu X, Buchwalter G, Feiglin A, Abell-Hart K, Fei T, Rao P, Long H, Kwiatkowski N, Zhang T, Gray N, Melchers D, Houtman R, Liu XS, Cohen O, Wagle N, Winer EP, Zhao J, Brown M.** Allele-Specific Chromatin Recruitment and Therapeutic Vulnerabilities of ESR1 Activating Mutations. *Cancer Cell* 2018;33(2):173-186.e5.

218. **Li S, Shen D, Shao J, Crowder R, Liu W, Prat A, He X, Liu S, Hoog J, Lu C, Ding L, Griffith OL, Miller C, Larson D, Fulton RS, Harrison M, Mooney T, McMichael JF, Luo J, Tao Y, Goncalves R, Schlosberg C, Hiken JF, Saied L, Sanchez C, Giuntoli T, Bumb C, Cooper C, Kitchens RT, Lin A, Phommaly C, Davies SR, Zhang J, Kavuri MS, McEachern D, Dong YY, Ma C, Pluard T, Naughton M, Bose R, Suresh R, McDowell R, Michel L, Aft R, Gillanders W, DeSchryver K, Wilson RK, Wang S, Mills GB, Gonzalez-Angulo A, Edwards JR, Maher C, Perou CM, Mardis ER, Ellis MJ.** Endocrine-therapy-resistant ESR1 variants revealed by genomic characterization of breast-cancer-derived xenografts. *Cell Rep* 2013;4(6):1116–1130.

219. **He X, Zheng Z, Song T, Wei C, Ma H, Ma Q, Zhang Y, Xu Y, Shi W, Ye Q, Zhong H.** c-Abl regulates estrogen receptor alpha transcription activity through its stabilization by phosphorylation. *Oncogene* 2010;29(15):2238–2251.

220. **Rogatsky I, Trowbridge JM, Garabedian MJ.** Potentiation of human estrogen receptor alpha transcriptional activation through phosphorylation of serines 104 and 106 by the cyclin A-CDK2 complex. *J. Biol. Chem.* 1999;274(32):22296–22302.

221. **Thomas RS, Sarwar N, Phoenix F, Coombes RC, Ali S.** Phosphorylation at serines 104 and 106 by Erk1/2 MAPK is important for estrogen receptor-alpha activity. *J. Mol. Endocrinol.* 2008;40(4):173–184.

222. **Joel PB, Traish AM, Lannigan DA.** Estradiol and phorbol ester cause phosphorylation of serine 118 in the human estrogen receptor. *Mol. Endocrinol.* 1995;9(8):1041–1052.

223. **Joel PB, Smith J, Sturgill TW, Fisher TL, Blenis J, Lannigan DA.** pp90rsk1 regulates estrogen receptor-mediated transcription through phosphorylation of Ser-167. *Mol. Cell. Biol.* 1998;18(4):1978–1984.

224. **Yamnik RL, Digilova A, Davis DC, Brodt ZN, Murphy CJ, Holz MK.** S6 kinase 1 regulates estrogen receptor alpha in control of breast cancer cell proliferation. *J. Biol. Chem.* 2009;284(10):6361–6369.

225. **Yamnik RL, Holz MK.** mTOR/S6K1 and MAPK/RSK signaling pathways coordinately regulate estrogen receptor alpha serine 167 phosphorylation. *FEBS Lett.* 2010;584(1):124–128.

226. **Michalides R, Griekspoor A, Balkenende A, Verwoerd D, Janssen L, Jalink K, Floore A, Velds A, van `t Veer L, Neefjes J.** Tamoxifen resistance by a conformational arrest of the estrogen receptor α after PKA activation in breast cancer. *Cancer Cell* 2004;5(6):597–605.

227. **Wang R-A, Mazumdar A, Vadlamudi RK, Kumar R.** P21-activated kinase-1 phosphorylates and transactivates estrogen receptor-alpha and promotes hyperplasia in mammary epithelium. *EMBO J.* 2002;21(20):5437–5447.

228. **Lee H, Bai W.** Regulation of estrogen receptor nuclear export by ligand-induced and p38-mediated receptor phosphorylation. *Mol. Cell. Biol.* 2002;22(16):5835–5845.

229. **Castoria G, Migliaccio A, Green S, Di Domenico M, Chambon P, Auricchio F.** Properties of a purified estradiol-dependent calf uterus tyrosine kinase. *Biochemistry* 1993;32(7):1740–1750.

230. **Liu Y, Gao H, Marstrand TT, Ström A, Valen E, Sandelin A, Gustafsson J-A, Dahlman-Wright K.** The genome landscape of ERalpha- and ERbeta-binding DNA regions. *Proc. Natl. Acad. Sci. U.S.A.* 2008;105(7):2604–2609.

231. **Hurtado A, Holmes KA, Geistlinger TR, Hutcheson IR, Nicholson RI, Brown M, Jiang J, Howat WJ, Ali S, Carroll JS.** Regulation of ERBB2 by oestrogen receptor-PAX2 determines response to tamoxifen. *Nature* 2008;456(7222):663–666.

232. **Madak-Erdogan Z, Lupien M, Stossi F, Brown M, Katzenellenbogen BS.** Genomic collaboration of estrogen receptor alpha and extracellular signal-regulated kinase 2 in regulating gene and proliferation programs. *Mol. Cell. Biol.* 2011;31(1):226–236.

233. **Hewitt SC, Li L, Grimm SA, Chen Y, Liu L, Li Y, Bushel PR, Fargo D, Korach KS.** Research Resource: Whole-Genome Estrogen Receptor α Binding in Mouse Uterine Tissue Revealed by ChIP-Seq. *Molecular Endocrinology* 2012;26(5):887–898.

234. **Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ, Brown GD, Gojis O, Ellis IO, Green AR, Ali S, Chin S-F, Palmieri C, Caldas C, Carroll JS.** Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* 2012;481(7381):389–393.

235. **Gertz J, Savic D, Varley KE, Partridge EC, Safi A, Jain P, Cooper GM, Reddy TE, Crawford GE, Myers RM.** Distinct properties of cell-type-specific and shared transcription factor binding sites. *Mol. Cell* 2013;52(1):25–36.

236. **Caizzi L, Ferrero G, Cutrupi S, Cordero F, Ballaré C, Miano V, Reineri S, Ricci L, Friard O, Testori A, Corà D, Caselle M, Di Croce L, De Bortoli M.** Genome-wide activity of unliganded estrogen receptor-α in breast cancer cells. *Proc Natl Acad Sci U S A* 2014;111(13):4892–4897.

237. **Mohammed H, Russell IA, Stark R, Rueda OM, Hickey TE, Tarulli GA, Serandour AAA, Birrell SN, Bruna A, Saadi A, Menon S, Hadfield J, Pugh M, Raj GV, Brown GD, D'Santos C, Robinson JLL, Silva G, Launchbury R, Perou CM, Stingl J, Caldas C, Tilley WD, Carroll JS.** Progesterone receptor modulates ERα action in breast cancer. *Nature* 2015;523(7560):313–317.

238. **Helzer KT, Szatkowski Ozers M, Meyer MB, Benkusky NA, Solodin N, Reese RM, Warren CL, Pike JW, Alarid ET.** The Phosphorylated Estrogen Receptor α (ER) Cistrome Identifies a Subset of Active Enhancers Enriched for Direct ER-DNA Binding and the Transcription Factor GRHL2. *Mol. Cell. Biol.* 2019;39(3). doi:10.1128/MCB.00417-18.

239. **Holding AN, Giorgi FM, Donnelly A, Cullen AE, Nagarajan S, Selth LA, Markowetz F.** VULCAN integrates ChIP-seq with patient-derived co-expression networks to identify GRHL2 as a key co-regulator of ERa at enhancers in breast cancer. *Genome Biol.* 2019;20(1):91.

240. **Cohen P.** The origins of protein phosphorylation. *Nat. Cell Biol.* 2002;4(5):E127-130.

241. **Ward RD, Weigel NL.** Steroid receptor phosphorylation: Assigning function to site-specific phosphorylation. *Biofactors* 2009;35(6):528–536.

242. **Liu Z, Merkurjev D, Yang F, Li W, Oh S, Friedman MJ, Song X, Zhang F, Ma Q, Ohgi KA, Krones A, Rosenfeld MG.** Enhancer activation requires trans-recruitment of a mega transcription factor complex. *Cell* 2014;159(2):358–373.

243. **Blaydes JP, Vojtesek B, Bloomberg GB, Hupp TR.** The development and use of phospho-specific antibodies to study protein phosphorylation. *Methods Mol. Biol.* 2000;99:177–189.

244. **Czernik AJ, Girault JA, Nairn AC, Chen J, Snyder G, Kebabian J, Greengard P.** Production of phosphorylation state-specific antibodies. *Meth. Enzymol.* 1991;201:264–283.

245. **Mandell JW.** Phosphorylation state-specific antibodies: applications in investigative and diagnostic pathology. *Am. J. Pathol.* 2003;163(5):1687–1698.

246. **Mandell JW.** Immunohistochemical assessment of protein phosphorylation state: the dream and the reality. *Histochem. Cell Biol.* 2008;130(3):465–471.

247. **Landt SG, Marinov GK, Kundaje A, Kheradpour P, Pauli F, Batzoglou S, Bernstein BE, Bickel P, Brown JB, Cayting P, Chen Y, DeSalvo G, Epstein C, Fisher-Aylor KI, Euskirchen G, Gerstein M, Gertz J, Hartemink AJ, Hoffman MM, Iyer VR, Jung YL, Karmakar S, Kellis M, Kharchenko PV, Li Q, Liu T, Liu XS, Ma L, Milosavljevic A, Myers RM, Park PJ, Pazin MJ, Perry MD, Raha D, Reddy TE, Rozowsky J, Shoresh N, Sidow A, Slattery M, Stamatoyannopoulos JA, Tolstorukov MY, White KP, Xi S, Farnham PJ, Lieb JD, Wold BJ, Snyder M.** ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res.* 2012;22(9):1813–1831.

248. **Fowler AM, Solodin NM, Valley CC, Alarid ET.** Altered target gene regulation controlled by estrogen receptor-alpha concentration. *Mol. Endocrinol.* 2006;20(2):291–301.

249. **O'Brien T, Lis JT.** Rapid changes in Drosophila transcription after an instantaneous heat shock. *Mol. Cell. Biol.* 1993;13(6):3456–3463.

250. **Tennyson CN, Klamut HJ, Worton RG.** The human dystrophin gene requires 16 hours to be transcribed and is cotranscriptionally spliced. *Nat. Genet.* 1995;9(2):184–190.

251. **Darzacq X, Shav-Tal Y, de Turris V, Brody Y, Shenoy SM, Phair RD, Singer RH.** In vivo dynamics of RNA polymerase II transcription. *Nat. Struct. Mol. Biol.* 2007;14(9):796–806.

252. **Li L, Li Z, Howley PM, Sacks DB.** E6AP and Calmodulin Reciprocally Regulate Estrogen Receptor Stability. *J. Biol. Chem.* 2006;281(4):1978–1985.

253. **Skliris GP, Rowan BG, Al-Dhaheri M, Williams C, Troup S, Begic S, Parisien M, Watson PH, Murphy LC.** Immunohistochemical validation of multiple phospho-specific epitopes for estrogen receptor alpha (ERalpha) in tissue microarrays of ERalpha positive human breast carcinomas. *Breast Cancer Res. Treat.* 2009;118(3):443–453.

254. **Bordeaux J, Welsh A, Agarwal S, Killiam E, Baquero M, Hanna J, Anagnostou V, Rimm D.** Antibody validation. *BioTechniques* 2010;48(3):197–209.

255. **Egelhofer TA, Minoda A, Klugman S, Lee K, Kolasinska-Zwierz P, Alekseyenko AA, Cheung M-S, Day DS, Gadel S, Gorchakov AA, Gu T, Kharchenko PV, Kuan S, Latorre I, Linder-Basso D, Luu Y, Ngo Q, Perry M, Rechtsteiner A, Riddle NC, Schwartz YB, Shanower GA, Vielle A, Ahringer J, Elgin SCR, Kuroda MI, Pirrotta V, Ren B, Strome S, Park PJ, Karpen GH, Hawkins RD, Lieb JD.** An assessment of histone-modification antibody quality. *Nat. Struct. Mol. Biol.* 2011;18(1):91–93.

256. **Sarwar N, Kim J-S, Jiang J, Peston D, Sinnett HD, Madden P, Gee JM, Nicholson RI, Lykkesfeldt AE, Shousha S, Coombes RC, Ali S.** Phosphorylation of ERalpha at serine 118 in primary breast cancer and in tamoxifen-resistant tumours is indicative of a complex role for ERalpha phosphorylation in breast cancer progression. *Endocr. Relat. Cancer* 2006;13(3):851–861.

257. **Rochette-Egly C.** Nuclear receptors: integration of multiple signalling pathways through phosphorylation. *Cell. Signal.* 2003;15(4):355–366.

258. **Krstic MD, Rogatsky I, Yamamoto KR, Garabedian MJ.** Mitogen-activated and cyclin-dependent protein kinases selectively and differentially modulate transcriptional enhancement by the glucocorticoid receptor. *Mol. Cell. Biol.* 1997;17(7):3947–3954.

259. **Gioeli D, Black BE, Gordon V, Spencer A, Kesler CT, Eblen ST, Paschal BM, Weber MJ.** Stress kinase signaling regulates androgen receptor phosphorylation, transcription, and localization. *Mol. Endocrinol.* 2006;20(3):503–515.

260. **Gaub MP, Bellard M, Scheuer I, Chambon P, Sassone-Corsi P.** Activation of the ovalbumin gene by the estrogen receptor involves the fos-jun complex. *Cell* 1990;63(6):1267–1276.

261. **O'Lone R, Frith MC, Karlsson EK, Hansen U.** Genomic targets of nuclear estrogen receptors. *Mol. Endocrinol.* 2004;18(8):1859–1875.

262. **Scholz A, Truss M, Beato M.** Hormone-induced recruitment of Sp1 mediates estrogen activation of the rabbit uteroglobin gene in endometrial epithelium. *J. Biol. Chem.* 1998;273(8):4360–4366.

263. **Rastinejad F, Ollendorff V, Polikarpov I.** Nuclear receptor full-length architectures: confronting myth and illusion with high resolution. *Trends Biochem. Sci.* 2015;40(1):16–24.

264. **Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, Boyer LA, Young RA, Jaenisch R.** Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. U.S.A.* 2010;107(50):21931–21936.

265. **Jozwik KM, Chernukhin I, Serandour AA, Nagarajan S, Carroll JS.** FOXA1 Directs H3K4 Monomethylation at Enhancers via Recruitment of the Methyltransferase MLL3. *Cell Rep* 2016;17(10):2715–2723.

266. **Warren CL, Kratochvil NCS, Hauschild KE, Foister S, Brezinski ML, Dervan PB, Phillips GN, Ansari AZ.** Defining the sequence-recognition profile of DNA-binding molecules. *Proc. Natl. Acad. Sci. U.S.A.* 2006;103(4):867–872.

267. **Whitington T, Frith MC, Johnson J, Bailey TL.** Inferring transcription factor complexes from ChIP-seq data. *Nucleic Acids Res.* 2011;39(15):e98.

268. **Klinge CM.** Estrogen receptor interaction with estrogen response elements. *Nucleic Acids Res.* 2001;29(14):2905–2919.

269. **Koszewski NJ, Notides AC.** Phosphate-sensitive binding of the estrogen receptor to its response elements. *Mol. Endocrinol.* 1991;5(8):1129–1136.

270. **Nardulli AM, Greene GL, Shapiro DJ.** Human estrogen receptor bound to an estrogen response element bends DNA. *Mol. Endocrinol.* 1993;7(3):331–340.

271. **Sabbah M, Le Ricousse S, Redeuilh G, Baulieu EE.** Estrogen receptor-induced bending of the Xenopus vitellogenin A2 gene hormone response element. *Biochem. Biophys. Res. Commun.* 1992;185(3):944–952.

272. **Wang JC.** Helical repeat of DNA in solution. *Proc. Natl. Acad. Sci. U.S.A.* 1979;76(1):200–203.

273. **Holding AN, Giorgi FM, Donnelly A, Cullen AE, Selth LA, Markowetz F.** VULCAN: Network Analysis of ER-α activation reveals reprogramming of GRHL2. *bioRxiv* 2018:266908.

274. **Cieply B, Farris J, Denvir J, Ford HL, Frisch SM.** Epithelial-mesenchymal transition and tumor suppression are controlled by a reciprocal feedback loop between ZEB1 and Grainyhead-like-2. *Cancer Res.* 2013;73(20):6299–6309.

275. **Werner S, Frey S, Riethdorf S, Schulze C, Alawi M, Kling L, Vafaizadeh V, Sauter G, Terracciano L, Schumacher U, Pantel K, Assmann V.** Dual roles of the transcription factor grainyhead-like 2 (GRHL2) in breast cancer. *J. Biol. Chem.* 2013;288(32):22993–23008.

276. **Mohammed H, Taylor C, Brown GD, Papachristou EK, Carroll JS, D'Santos CS.** Rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) for analysis of chromatin complexes. *Nat Protoc* 2016;11(2):316–326.

277. **Jacobs J, Atkins M, Davie K, Imrichova H, Romanelli L, Christiaens V, Hulselmans G, Potier D, Wouters J, Taskiran II, Paciello G, González-Blas CB, Koldere D, Aibar S, Halder G, Aerts S.** The transcription factor Grainy head primes epithelial enhancers for spatiotemporal activation by displacing nucleosomes. *Nat. Genet.* 2018. doi:10.1038/s41588-018-0140-x.

278. **Chen AF, Liu AJ, Krishnakumar R, Freimer JW, DeVeale B, Blelloch R.** GRHL2-Dependent Enhancer Switching Maintains a Pluripotent Stem Cell Transcriptional Subnetwork after Exit from Naive Pluripotency. *Cell Stem Cell* 2018;23(2):226-238.e4.

279. **Nevil M, Bondra ER, Schulz KN, Kaplan T, Harrison MM.** Stable Binding of the Conserved Transcription Factor Grainy Head to its Target Genes Throughout Drosophila melanogaster Development. *Genetics* 2017;205(2):605–620.

280. **Fishwick C, Higgins J, Percival-Alwyn L, Hustler A, Pearson J, Bastkowski S, Moxon S, Swarbreck D, Greenman CD, Southgate J.** Heterarchy of transcription factors driving basal and luminal cell phenotypes in human urothelium. *Cell Death Differ.* 2017;24(5):809–818.

281. **Grossman SR, Engreitz J, Ray JP, Nguyen TH, Hacohen N, Lander ES.** Positional specificity of different transcription factor classes within enhancers. *Proc. Natl. Acad. Sci. U.S.A.* 2018. doi:10.1073/pnas.1804663115.

282. **Carleton JB, Berrett KC, Gertz J.** Multiplex Enhancer Interference Reveals Collaborative Control of Gene Regulation by Estrogen Receptor α-Bound Enhancers. *Cell Syst* 2017;5(4):333-344.e5.

283. **Murphy LC, Seekallu SV, Watson PH.** Clinical significance of estrogen receptor phosphorylation. *Endocr. Relat. Cancer* 2011;18(1):R1-14.

284. **Picard N, Charbonneau C, Sanchez M, Licznar A, Busson M, Lazennec G, Tremblay A.** Phosphorylation of Activation Function-1 Regulates Proteasome-Dependent Nuclear Mobility and E6-Associated Protein Ubiquitin Ligase Recruitment to the Estrogen Receptor β. *Molecular Endocrinology* 2008;22(2):317–330.

285. **Cook KL, Clarke PAG, Parmar J, Hu R, Schwartz-Roberts JL, Abu-Asab M, Wärri A, Baumann WT, Clarke R.** Knockdown of estrogen receptor-α induces autophagy and inhibits

antiestrogen-mediated unfolded protein response activation, promoting ROS-induced breast cancer cell death. *FASEB J.* 2014;28(9):3891–3905.

286. **Paquet D, Kwart D, Chen A, Sproul A, Jacob S, Teo S, Olsen KM, Gregg A, Noggle S, Tessier-Lavigne M.** Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* 2016;533(7601):125–129.

287. **Wijayaratne AL, McDonnell DP.** The Human Estrogen Receptor-α Is a Ubiquitinated Protein Whose Stability Is Affected Differentially by Agonists, Antagonists, and Selective Estrogen Receptor Modulators. *J. Biol. Chem.* 2001;276(38):35684–35692.

288. **Lin H-K, Yeh S, Kang H-Y, Chang C.** Akt suppresses androgen-induced apoptosis by phosphorylating and inhibiting androgen receptor. *PNAS* 2001;98(13):7200–7205.

289. **Li Z, Joyal JL, Sacks DB.** Calmodulin enhances the stability of the estrogen receptor. *J. Biol. Chem.* 2001;276(20):17354–17360.

290. **Fan M, Nakshatri H, Nephew KP.** Inhibiting proteasomal proteolysis sustains estrogen receptor-alpha activation. *Mol. Endocrinol.* 2004;18(11):2603–2615.

291. **Berry NB, Fan M, Nephew KP.** Estrogen receptor-alpha hinge-region lysines 302 and 303 regulate receptor degradation by the proteasome. *Mol. Endocrinol.* 2008;22(7):1535–1551.

292. **Holliday DL, Speirs V.** Choosing the right cell line for breast cancer research. *Breast Cancer Res* 2011;13(4):215.

293. **Nawaz Z, Lonard DM, Smith CL, Lev-Lehman E, Tsai SY, Tsai M-J, O'Malley BW.** The Angelman Syndrome-Associated Protein, E6-AP, Is a Coactivator for the Nuclear Hormone Receptor Superfamily. *Mol. Cell. Biol.* 1999;19(2):1182–1189.

294. **Singhmar P, Kumar A.** Angelman syndrome protein UBE3A interacts with primary microcephaly protein ASPM, localizes to centrosomes and regulates chromosome segregation. *PLoS ONE* 2011;6(5):e20397.

295. **Lochab S, Pal P, Kanaujiya JK, Tripathi SB, Kapoor I, Bhatt MLB, Sanyal S, Behre G, Trivedi AK.** Proteomic identification of E6AP as a molecular target of tamoxifen in MCF7 cells. *Proteomics* 2012;12(9):1363–1377.

296. **Uhlen M, Zhang C, Lee S, Sjöstedt E, Fagerberg L, Bidkhori G, Benfeitas R, Arif M, Liu Z, Edfors F, Sanli K, von Feilitzen K, Oksvold P, Lundberg E, Hober S, Nilsson P, Mattsson J, Schwenk JM, Brunnström H, Glimelius B, Sjöblom T, Edqvist P-H, Djureinovic D, Micke P, Lindskog C, Mardinoglu A, Ponten F.** A pathology atlas of the human cancer transcriptome. *Science* 2017;357(6352). doi:10.1126/science.aan2507.

297. **Kemppainen JA, Lane MV, Sar M, Wilson EM.** Androgen receptor phosphorylation, turnover, nuclear transport, and transcriptional activation. Specificity for steroids and antihormones. *J. Biol. Chem.* 1992;267(2):968–974.

298. **Weigel NL, Carter TH, Schrader WT, O'Malley BW.** Chicken progesterone receptor is phosphorylated by a DNA-dependent protein kinase during in vitro transcription assays. *Mol. Endocrinol.* 1992;6(1):8–14.

299. **Lange CA, Shen T, Horwitz KB.** Phosphorylation of human progesterone receptors at serine-294 by mitogen-activated protein kinase signals their degradation by the 26S proteasome. *PNAS* 2000;97(3):1032–1037.

300. **Webster JC, Jewell CM, Bodwell JE, Munck A, Sar M, Cidlowski JA.** Mouse glucocorticoid receptor phosphorylation status influences multiple functions of the receptor protein. *J. Biol. Chem.* 1997;272(14):9287–9293.

301. **Bodwell JE, Ortí E, Coull JM, Pappin DJ, Smith LI, Swift F.** Identification of phosphorylated sites in the mouse glucocorticoid receptor. *J. Biol. Chem.* 1991;266(12):7549–7555.

302. **Tremblay A, Tremblay GB, Labrie F, Giguère V.** Ligand-independent recruitment of SRC-1 to estrogen receptor beta through phosphorylation of activation function AF-1. *Mol. Cell* 1999;3(4):513–519.

303. **Goldberg Y, Glineur C, Gesquière JC, Ricouart A, Sap J, Vennström B, Ghysdael J.** Activation of protein kinase C or cAMP-dependent protein kinase increases phosphorylation of the c-erbA-encoded thyroid hormone receptor and of the v-erbA-encoded protein. *EMBO J.* 1988;7(8):2425–2433.

304. **Glineur C, Bailly M, Ghysdael J.** The c-erbA alpha-encoded thyroid hormone receptor is phosphorylated in its amino terminal domain by casein kinase II. *Oncogene* 1989;4(10):1247–1254.

305. **Diradourian C, Girard J, Pégorier J-P.** Phosphorylation of PPARs: from molecular characterization to physiological relevance. *Biochimie* 2005;87(1):33–38.

306. **Rochette-Egly C, Adam S, Rossignol M, Egly JM, Chambon P.** Stimulation of RAR alpha activation function AF-1 through binding to the general transcription factor TFIIH and phosphorylation by CDK7. *Cell* 1997;90(1):97–107.

307. **Bastien J, Adam-Stitah S, Riedl T, Egly JM, Chambon P, Rochette-Egly C.** TFIIH interacts with the retinoic acid receptor gamma and phosphorylates its AF-1-activating domain through cdk7. *J. Biol. Chem.* 2000;275(29):21896–21904.

308. **Kopf E, Plassat J-L, Vivat V, Thé H de, Chambon P, Rochette-Egly C.** Dimerization with Retinoid X Receptors and Phosphorylation Modulate the Retinoic Acid-induced Degradation of Retinoic Acid Receptors α and γ through the Ubiquitin-Proteasome Pathway. *J. Biol. Chem.* 2000;275(43):33280–33288.

309. **Hsieh JC, Jurutka PW, Nakajima S, Galligan MA, Haussler CA, Shimizu Y, Shimizu N, Whitfield GK, Haussler MR.** Phosphorylation of the human vitamin D receptor by protein kinase C. Biochemical and functional evaluation of the serine 51 recognition site. *J. Biol. Chem.* 1993;268(20):15118–15126.

310. **Reese RM, Harrison MM, Alarid ET.** Grainyhead-like Protein 2: The Emerging Role in Hormone-Dependent Cancers and Epigenetics. *Endocrinology* 2019;160(5):1275–1288.

311. **Frisch SM, Farris JC, Pifer PM.** Roles of Grainyhead-like transcription factors in cancer. *Oncogene* 2017;36(44):6067–6073.

312. **Ma L, Yan H, Zhao H, Sun J.** Grainyhead-like 2 in development and cancer. *Tumour Biol.* 2017;39(5):1010428317698375.

313. **Mlacki M, Kikulska A, Krzywinska E, Pawlak M, Wilanowski T.** Recent discoveries concerning the involvement of transcription factors from the Grainyhead-like family in cancer. *Exp. Biol. Med. (Maywood)* 2015. doi:10.1177/1535370215588924.

314. **Chi D, Singhal H, Li L, Xiao T, Liu W, Pun M, Jeselsohn R, He H, Lim E, Vadhi R, Rao P, Long H, Garber J, Brown M.** Estrogen receptor signaling is reprogrammed during breast tumorigenesis. *Proc. Natl. Acad. Sci. U.S.A.* 2019;116(23):11437–11443.

315. **Zaret KS, Mango SE.** Pioneer transcription factors, chromatin dynamics, and cell fate control. *Curr. Opin. Genet. Dev.* 2016;37:76–81.

316. **Mayran A, Drouin J.** Pioneer transcription factors shape the epigenetic landscape. *J. Biol. Chem.* 2018;293(36):13795–13804.

317. **Paltoglou S, Das R, Townley SL, Hickey TE, Tarulli GA, Coutinho I, Fernandes R, Hanson AR, Denis I, Carroll JS, Dehm SM, Raj GV, Plymate SR, Tilley WD, Selth LA.** Novel Androgen Receptor Coregulator GRHL2 Exerts Both Oncogenic and Antimetastatic Functions in Prostate Cancer. *Cancer Res.* 2017;77(13):3417–3430.

318. **Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, Ye Z, Lee LK, Stuart RK, Ching CW, Ching KA, Antosiewicz-Bourget JE, Liu H, Zhang X, Green RD, Lobanenkov VV, Stewart R, Thomson JA, Crawford GE, Kellis M, Ren B.** Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* 2009;459(7243):108–112.

319. **Pifer PM, Farris JC, Thomas AL, Stoilov P, Denvir J, Smith DM, Frisch SM.** Grainyhead-like 2 inhibits the coactivator p300, suppressing tubulogenesis and the epithelial-mesenchymal transition. *Mol. Biol. Cell* 2016;27(15):2479–2492.

320. **Kraus WL, Kadonaga JT.** p300 and estrogen receptor cooperatively activate transcription via differential enhancement of initiation and reinitiation. *Genes Dev.* 1998;12(3):331–342.

321. **Reid G, Hübner MR, Métivier R, Brand H, Denger S, Manu D, Beaudouin J, Ellenberg J, Gannon F.** Cyclic, Proteasome-Mediated Turnover of Unliganded and Liganded ERα on Responsive Promoters Is an Integral Feature of Estrogen Signaling. *Molecular Cell* 2003;11(3):695–707.

322. **Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA, Lim WA, Weissman JS, Qi LS.** CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* 2013;154(2):442–451.

323. **Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, Aiden EL.** A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 2014;159(7):1665–1680.

324. **Belton J-M, McCord RP, Gibcus JH, Naumova N, Zhan Y, Dekker J.** Hi-C: a comprehensive technique to capture the conformation of genomes. *Methods* 2012;58(3):268–276.

325. **Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A, Plajzer-Frick I, Shoukry M, Wright C, Chen F, Afzal V, Ren B, Rubin EM, Pennacchio LA.** ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* 2009;457(7231):854–858.

326. **Van Antwerp DJ, Verma IM.** Signal-induced degradation of I(kappa)B(alpha): association with NF-kappaB and the PEST sequence in I(kappa)B(alpha) are not required. *Mol. Cell. Biol.* 1996;16(11):6037–6045.

327. **Meyer MB, Benkusky NA, Pike JW.** The RUNX2 cistrome in osteoblasts: characterization, down-regulation following differentiation, and relationship to gene expression. *J. Biol. Chem.* 2014;289(23):16016–16031.

328. **Di Croce L, Okret S, Kersten S, Gustafsson JA, Parker M, Wahli W, Beato M.** Steroid and nuclear receptors. *EMBO J* 1999;18(22):6201–6210.

329. **Aranda A, Pascual A.** Nuclear hormone receptors and gene expression. *Physiol. Rev.* 2001;81(3):1269–1304.

330. **Jensen EV, Desombre ER, Kawashima T, Suzuki T, Kyser K, Jungblut PW.** Estrogen-binding substances of target tissues. *Science* 1967;158(3800):529–530.

331. **Miesfeld R, Okret S, Wikström AC, Wrange O, Gustafsson JA, Yamamoto KR.** Characterization of a steroid hormone receptor gene and mRNA in wild-type and mutant cells. *Nature* 1984;312(5996):779–781.

332. **Hollenberg SM, Weinberger C, Ong ES, Cerelli G, Oro A, Lebo R, Thompson EB, Rosenfeld MG, Evans RM.** Primary structure and expression of a functional human glucocorticoid receptor cDNA. *Nature* 1985;318(6047):635–641.

333. **Evans RM, Mangelsdorf DJ.** Nuclear Receptors, RXR, and the Big Bang. *Cell* 2014;157(1):255–266.

334. **Gronemeyer H, Gustafsson J-Å, Laudet V.** Principles for modulation of the nuclear receptor superfamily. *Nat Rev Drug Discov* 2004;3(11):950–964.

335. **McDonnell DP, Wardell SE.** The molecular mechanisms underlying the pharmacological actions of ER modulators: implications for new drug discovery in breast cancer. *Curr Opin Pharmacol* 2010;10(6):620–628.

336. **Malovannaya A, Lanz RB, Jung SY, Bulynko Y, Le NT, Chan DW, Ding C, Shi Y, Yucer N, Krenciute G, Kim B-J, Li C, Chen R, Li W, Wang Y, O'Malley BW, Qin J.** Analysis of the human endogenous coregulator complexome. *Cell* 2011;145(5):787–799.

337. **Nagaich AK, Walker DA, Wolford R, Hager GL.** Rapid periodic binding and displacement of the glucocorticoid receptor during chromatin remodeling. *Mol. Cell* 2004;14(2):163–174.

338. **Norris JD, Paige LA, Christensen DJ, Chang CY, Huacani MR, Fan D, Hamilton PT, Fowlkes DM, McDonnell DP.** Peptide antagonists of the human estrogen receptor. *Science* 1999;285(5428):744–746.

339. **Parent AA, Gunther JR, Katzenellenbogen JA.** Blocking estrogen signaling after the hormone: pyrimidine-core inhibitors of estrogen receptor-coactivator binding. *J. Med. Chem.* 2008;51(20):6512–6530.

340. **Gunther JR, Parent AA, Katzenellenbogen JA.** Alternative inhibition of androgen receptor signaling: peptidomimetic pyrimidines as direct androgen receptor/coactivator disruptors. *ACS Chem. Biol.* 2009;4(6):435–440.

341. **Tran C, Ouk S, Clegg NJ, Chen Y, Watson PA, Arora V, Wongvipat J, Smith-Jones PM, Yoo D, Kwon A, Wasielewska T, Welsbie D, Chen CD, Higano CS, Beer TM, Hung DT, Scher HI, Jung ME, Sawyers CL.** Development of a second-generation antiandrogen for treatment of advanced prostate cancer. *Science* 2009;324(5928):787–790.

342. **Wang LH, Yang XY, Zhang X, An P, Kim H-J, Huang J, Clarke R, Osborne CK, Inman JK, Appella E, Farrar WL.** Disruption of estrogen receptor DNA-binding domain and related intramolecular communication restores tamoxifen sensitivity in resistant breast cancer. *Cancer Cell* 2006;10(6):487–499.

343. **Mao C, Patterson NM, Cherian MT, Aninye IO, Zhang C, Montoya JB, Cheng J, Putt KS, Hergenrother PJ, Wilson EM, Nardulli AM, Nordeen SK, Shapiro DJ.** A new small molecule inhibitor of estrogen receptor alpha binding to estrogen response elements blocks estrogen-dependent growth of cancer cells. *J. Biol. Chem.* 2008;283(19):12819–12830.

344. **Andersen RJ, Mawji NR, Wang J, Wang G, Haile S, Myung J-K, Watt K, Tam T, Yang YC, Bañuelos CA, Williams DE, McEwan IJ, Wang Y, Sadar MD.** Regression of castrate-recurrent prostate cancer by a small-molecule inhibitor of the amino-terminus domain of the androgen receptor. *Cancer Cell* 2010;17(6):535–546.

345. **Caboni L, Lloyd DG.** Beyond the Ligand-Binding Pocket: Targeting Alternate Sites in Nuclear Receptors. *Med. Res. Rev.* 2013;33(5):1081–1118.

346. **Yudt MR, Vorojeikina D, Zhong L, Skafar DF, Sasson S, Gasiewicz TA, Notides AC.** Function of Estrogen Receptor Tyrosine 537 in Hormone Binding, DNA Binding, and Transactivation†. *Biochemistry* 1999;38(43):14146–14156.

347. **Clark DE, Poteet-Smith CE, Smith JA, Lannigan DA.** Rsk2 allosterically activates estrogen receptor alpha by docking to the hormone-binding domain. *EMBO J.* 2001;20(13):3484–3494.

348. **Held JM, Britton DJ, Scott GK, Lee EL, Schilling B, Baldwin MA, Gibson BW, Benz CC.** Ligand binding promotes CDK-dependent phosphorylation of ER-alpha on hinge serine 294 but inhibits ligand-independent phosphorylation of serine 305. *Mol. Cancer Res.* 2012;10(8):1120–1132.

349. **Henrich LM, Smith JA, Kitt D, Errington TM, Nguyen B, Traish AM, Lannigan DA.** Extracellular signal-regulated kinase 7, a regulator of hormone-dependent estrogen receptor destruction. *Mol. Cell. Biol.* 2003;23(17):5979–5988.

350. **Calligé M, Kieffer I, Richard-Foy H.** CSN5/Jab1 is involved in ligand-dependent degradation of estrogen receptor {alpha} by the proteasome. *Mol. Cell. Biol.* 2005;25(11):4349–4358.

351. **Bhatt S, Xiao Z, Meng Z, Katzenellenbogen BS.** Phosphorylation by p38 Mitogen-Activated Protein Kinase Promotes Estrogen Receptor α Turnover and Functional Activity via the SCFSkp2 Proteasomal Complex. *Mol. Cell. Biol.* 2012;32(10):1928–1943.

352. **Hilmi K, Hussein N, Mendoza-Sanchez R, El-Ezzy M, Ismail H, Durette C, Bail M, Rozendaal MJ, Bouvier M, Thibault P, Gleason JL, Mader S.** Role of SUMOylation in full antiestrogenicity. *Mol. Cell. Biol.* 2012;32(19):3823–3837.

353. **Acconcia F, Ascenzi P, Bocedi A, Spisni E, Tomasi V, Trentalance A, Visca P, Marino M.** Palmitoylation-dependent estrogen receptor alpha membrane localization: regulation by 17beta-estradiol. *Mol. Biol. Cell* 2005;16(1):231–237.

354. **Hershko A, Ciechanover A.** The Ubiquitin System. *Annual Review of Biochemistry* 1998;67(1):425–479.

355. **Pickart CM.** Mechanisms Underlying Ubiquitination. *Annual Review of Biochemistry* 2001;70(1):503–533.

356. **Komander D.** The emerging complexity of protein ubiquitination. *Biochem. Soc. Trans.* 2009;37(5):937.

357. **Deshaies RJ, Joazeiro CAP.** RING Domain E3 Ubiquitin Ligases. *Annual Review of Biochemistry* 2009;78(1):399–434.

358. **Bernassola F, Karin M, Ciechanover A, Melino G.** The HECT family of E3 ubiquitin ligases: multiple players in cancer development. *Cancer Cell* 2008;14(1):10–21.

359. **Markson G, Kiel C, Hyde R, Brown S, Charalabous P, Bremm A, Semple J, Woodsmith J, Duley S, Salehi-Ashtiani K, Vidal M, Komander D, Serrano L, Lehner P, Sanderson CM.** Analysis of the human E2 ubiquitin conjugating enzyme protein interaction network. *Genome Res.* 2009;19(10):1905–1911.

360. **Schulman BA, Harper JW.** Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways. *Nat. Rev. Mol. Cell Biol.* 2009;10(5):319–331.

361. **Ye Y, Rape M.** Building ubiquitin chains: E2 enzymes at work. *Nat. Rev. Mol. Cell Biol.* 2009;10(11):755–764.

362. **van Wijk SJL, Timmers HTM.** The family of ubiquitin-conjugating enzymes (E2s): deciding between life and death of proteins. *FASEB J.* 2010;24(4):981–993.

363. **Zhu J, Gianni M, Kopf E, Honoré N, Chelbi-Alix M, Koken M, Quignon F, Rochette-Egly C, Thé H de.** Retinoic acid induces proteasome-dependent degradation of retinoic acid receptor α (RARα) and oncogenic RARα fusion proteins. *PNAS* 1999;96(26):14807–14812.

364. **Dace A, Zhao L, Park KS, Furuno T, Takamura N, Nakanishi M, West BL, Hanover JA, Cheng S.** Hormone binding induces rapid proteasome-mediated degradation of thyroid hormone receptors. *PNAS* 2000;97(16):8985–8990.

365. **Wallace AD, Cidlowski JA.** Proteasome-mediated Glucocorticoid Receptor Degradation Restricts Transcriptional Signaling by Glucocorticoids. *J. Biol. Chem.* 2001;276(46):42714–42721.

366. **Yokota K, Shibata H, Kobayashi S, Suda N, Murai A, Kurihara I, Saito I, Saruta T.** Proteasome-mediated mineralocorticoid receptor degradation attenuates transcriptional response to aldosterone. *Endocr. Res.* 2004;30(4):611–616.

367. **Lonard DM, Nawaz Z, Smith CL, O'Malley BW.** The 26S Proteasome Is Required for Estrogen Receptor-α and Coactivator Turnover and for Efficient Estrogen Receptor-α Transactivation. *Molecular Cell* 2000;5(6):939–948.

368. **Shinohara K, Tomioka M, Nakano H, Toné S, Ito H, Kawashima S.** Apoptosis induction resulting from proteasome inhibition. *Biochem. J.* 1996;317 (Pt 2):385–388.

369. **Emanuele S, Calvaruso G, Lauricella M, Giuliano M, Bellavia G, D'Anneo A, Vento R, Tesoriere G.** Apoptosis induced in hepatoblastoma HepG2 cells by the proteasome inhibitor MG132 is associated with hydrogen peroxide production, expression of Bcl-XS and activation of caspase-3. *International Journal of Oncology* 2002. doi:10.3892/ijo.21.4.857.

370. **Cirit M, Grant KG, Haugh JM.** Systemic Perturbation of the ERK Signaling Pathway by the Proteasome Inhibitor, MG132. *PLoS ONE* 2012;7(11):e50975.

371. **Duong V, Boulle N, Daujat S, Chauvet J, Bonnet S, Neel H, Cavaillès V.** Differential Regulation of Estrogen Receptor α Turnover and Transactivation by Mdm2 and Stress-Inducing Agents. *Cancer Res* 2007;67(11):5513–5521.

372. **Powers GL, Ellison-Zelski SJ, Casa AJ, Lee AV, Alarid ET.** Proteasome inhibition represses ERalpha gene expression in ER+ cells: a new link between proteasome activity and estrogen signaling in breast cancer. *Oncogene* 2010;29(10):1509–1518.

373. **Prenzel T, Begus-Nahrmann Y, Kramer F, Hennion M, Hsu C, Gorsler T, Hintermair C, Eick D, Kremmer E, Simons M, Beissbarth T, Johnsen SA.** Estrogen-dependent gene transcription in human breast cancer cells relies upon proteasome-dependent monoubiquitination of histone H2B. *Cancer Res.* 2011;71(17):5739–5753.

374. **Ballinger CA, Connell P, Wu Y, Hu Z, Thompson LJ, Yin LY, Patterson C.** Identification of CHIP, a novel tetratricopeptide repeat-containing protein that interacts with heat shock proteins and negatively regulates chaperone functions. *Mol. Cell. Biol.* 1999;19(6):4535–4545.

375. **Smith DF, Toft DO.** Steroid receptors and their associated proteins. *Mol. Endocrinol.* 1993;7(1):4–11.

376. **Connell P, Ballinger CA, Jiang J, Wu Y, Thompson LJ, Höhfeld J, Patterson C.** The co-chaperone CHIP regulates protein triage decisions mediated by heat-shock proteins. *Nat Cell Biol* 2001;3(1):93–96.

377. **Pratt WB, Toft DO.** Steroid receptor interactions with heat shock protein and immunophilin chaperones. *Endocr. Rev.* 1997;18(3):306–360.

378. **Bagatell R, Khan O, Paine-Murrieta G, Taylor CW, Akinaga S, Whitesell L.** Destabilization of steroid receptors by heat shock protein 90-binding drugs: a ligand-independent approach to hormonal therapy of breast cancer. *Clin. Cancer Res.* 2001;7(7):2076–2084.

379. **Lee MO, Kim EO, Kwon HJ, Kim YM, Kang HJ, Kang H, Lee JE.** Radicicol represses the transcriptional function of the estrogen receptor by suppressing the stabilization of the receptor by heat shock protein 90. *Mol. Cell. Endocrinol.* 2002;188(1–2):47–54.

380. **Tateishi Y, Kawabe Y, Chiba T, Murata S, Ichikawa K, Murayama A, Tanaka K, Baba T, Kato S, Yanagisawa J.** Ligand-dependent switching of ubiquitin–proteasome pathways for estrogen receptor. *The EMBO Journal* 2004;23(24):4813–4823.

381. **Cardozo CP, Michaud C, Ost MC, Fliss AE, Yang E, Patterson C, Hall SJ, Caplan AJ.** C-terminal Hsp-interacting protein slows androgen receptor synthesis and reduces its rate of degradation. *Archives of Biochemistry and Biophysics* 2003;410(1):134–140.

382. **Wang X, DeFranco DB.** Alternative Effects of the Ubiquitin-Proteasome Pathway on Glucocorticoid Receptor Down-Regulation and Transactivation Are Mediated by CHIP, an E3 Ligase. *Molecular Endocrinology* 2005;19(6):1474–1482.

383. **Tateishi Y, Sonoo R, Sekiya Y, Sunahara N, Kawano M, Wayama M, Hirota R, Kawabe Y, Murayama A, Kato S, Kimura K, Yanagisawa J.** Turning Off Estrogen Receptor β-Mediated Transcription Requires Estrogen-Dependent Receptor Proteolysis. *Mol. Cell. Biol.* 2006;26(21):7966–7976.

384. **Faresse N, Ruffieux-Daidie D, Salamin M, Gomez-Sanchez CE, Staub O.** Mineralocorticoid receptor degradation is promoted by Hsp90 inhibition and the ubiquitin-protein ligase CHIP. *American Journal of Physiology - Renal Physiology* 2010;299(6):F1462–F1472.

385. **Honda R, Tanaka H, Yasuda H.** Oncoprotein MDM2 is a ubiquitin ligase E3 for tumor suppressor p53. *FEBS Letters* 1997;420(1):25–27.

386. **Zhou BP, Liao Y, Xia W, Zou Y, Spohn B, Hung MC.** HER-2/neu induces p53 ubiquitination via Akt-mediated MDM2 phosphorylation. *Nat. Cell Biol.* 2001;3(11):973–982.

387. **Chymkowitch P, May NL, Charneau P, Compe E, Egly J-M.** The phosphorylation of the androgen receptor by TFIIH directs the ubiquitin/proteasome process. *The EMBO Journal* 2011;30(3):468–479.

388. **Sanchez M, Picard N, Sauvé K, Tremblay A.** Coordinate regulation of estrogen receptor β degradation by Mdm2 and CREB-binding protein in response to growth signals. *Oncogene* 2013;32(1):117–126.

389. **Saji S, Okumura N, Eguchi H, Nakashima S, Suzuki A, Toi M, Nozawa Y, Saji S, Hayashi S.** MDM2 Enhances the Function of Estrogen Receptor α in Human Breast Cancer Cells. *Biochemical and Biophysical Research Communications* 2001;281(1):259–265.

390. **Kim K, Burghardt R, Barhoumi R, Lee S, Liu X, Safe S.** MDM2 regulates estrogen receptor α and estrogen responsiveness in breast cancer cells. *J Mol Endocrinol* 2011;46(2):67–79.

391. **Sengupta S, Wasylyk B.** Ligand-dependent interaction of the glucocorticoid receptor with p53 enhances their degradation by Hdm2. *Genes Dev.* 2001;15(18):2367–2380.

392. **Ramamoorthy S, Nawaz Z.** E6-associated protein (E6-AP) is a dual function coactivator of steroid hormone receptors. *Nucl Recept Signal* 2008;6:e006.

393. **Eakin CM, MacCoss MJ, Finney GL, Klevit RE.** Estrogen receptor α is a putative substrate for the BRCA1 ubiquitin ligase. *PNAS* 2007;104(14):5794–5799.

394. **Zhu J, Zhao C, Kharman-Biz A, Zhuang T, Jonsson P, Liang N, Williams C, Lin C-Y, Qiao Y, Zendehdel K, Strömblad S, Treuter E, Dahlman-Wright K.** The atypical ubiquitin ligase RNF31 stabilizes estrogen receptor α and modulates estrogen-stimulated breast cancer cell proliferation. *Oncogene* 2014;33(34):4340–4351.

395. **Hashizume R, Fukuda M, Maeda I, Nishikawa H, Oyake D, Yabuki Y, Ogata H, Ohta T.** The RING Heterodimer BRCA1-BARD1 Is a Ubiquitin Ligase Inactivated by a Breast Cancer-derived Mutation. *J. Biol. Chem.* 2001;276(18):14537–14540.

396. **Calvo V, Beato M.** BRCA1 Counteracts Progesterone Action by Ubiquitination Leading to Progesterone Receptor Degradation and Epigenetic Silencing of Target Promoters. *Cancer Res* 2011;71(9):3422–3431.

397. **Ma Y, Fan S, Hu C, Meng Q, Fuqua SA, Pestell RG, Tomita YA, Rosen EM.** BRCA1 Regulates Acetylation and Ubiquitination of Estrogen Receptor-α. *Molecular Endocrinology* 2010;24(1):76–90.

398. **Thompson HGR, Harris JW, Lin L, Brody JP.** Identification of the protein Zibra, its genomic organization, regulation, and expression in breast cancer cells. *Exp. Cell Res.* 2004;295(2):448–459.

399. **Kulathu Y, Komander D.** Atypical ubiquitylation — the unexplored world of polyubiquitin beyond Lys48 and Lys63 linkages. *Nat Rev Mol Cell Biol* 2012;13(8):508–523.

400. **Xu P, Duong DM, Seyfried NT, Cheng D, Xie Y, Robert J, Rush J, Hochstrasser M, Finley D, Peng J.** Quantitative Proteomics Reveals the Function of Unconventional Ubiquitin Chains in Proteasomal Degradation. *Cell* 2009;137(1):133–145.

401. **Xu K, Shimelis H, Linn DE, Jiang R, Yang X, Sun F, Guo Z, Chen H, Li W, Chen H, Kong X, Melamed J, Fang S, Xiao Z, Veenstra TD, Qiu Y.** Regulation of Androgen Receptor

Transcriptional Activity and Specificity by RNF6-Induced Ubiquitination. *Cancer Cell* 2009;15(4):270–282.

402. **Chen J, Chen ZJ.** Regulation of NF-κB by ubiquitination. *Curr. Opin. Immunol.* 2013;25(1):4–12.

403. **Yamaoka S, Courtois G, Bessia C, Whiteside ST, Weil R, Agou F, Kirk HE, Kay RJ, Israël A.** Complementation cloning of NEMO, a component of the IkappaB kinase complex essential for NF-kappaB activation. *Cell* 1998;93(7):1231–1240.

404. **Haglund K, Dikic I.** Ubiquitylation and cell signaling. *EMBO J.* 2005;24(19):3353–3359.

405. **Emmerich CH, Ordureau A, Strickson S, Arthur JSC, Pedrioli PGA, Komander D, Cohen P.** Activation of the canonical IKK complex by K63/M1-linked hybrid ubiquitin chains. *Proc. Natl. Acad. Sci. U.S.A.* 2013;110(38):15247–15252.

406. **Jackson SP, Durocher D.** Regulation of DNA damage responses by ubiquitin and SUMO. *Mol. Cell* 2013;49(5):795–807.

407. **Xu M, Skaug B, Zeng W, Chen ZJ.** A ubiquitin replacement strategy in human cells reveals distinct mechanisms of IKK activation by TNFalpha and IL-1beta. *Mol. Cell* 2009;36(2):302–314.

408. **Kirkpatrick DS, Denison C, Gygi SP.** Weighing in on ubiquitin: the expanding role of mass-spectrometry-based proteomics. *Nat. Cell Biol.* 2005;7(8):750–757.

409. **Nishikawa H, Ooka S, Sato K, Arima K, Okamoto J, Klevit RE, Fukuda M, Ohta T.** Mass spectrometric and mutational analyses reveal Lys-6-linked polyubiquitin chains catalyzed by BRCA1-BARD1 ubiquitin ligase. *J. Biol. Chem.* 2004;279(6):3916–3924.

410. **Danielsen JMR, Sylvestersen KB, Bekker-Jensen S, Szklarczyk D, Poulsen JW, Horn H, Jensen LJ, Mailand N, Nielsen ML.** Mass spectrometric analysis of lysine ubiquitylation reveals promiscuity at site level. *Mol. Cell Proteomics* 2011;10(3):M110.003590.

411. **Udeshi ND, Mertins P, Svinkina T, Carr SA.** Large-scale identification of ubiquitination sites by mass spectrometry. *Nat Protoc* 2013;8(10):1950–1960.

412. **Kim W, Bennett EJ, Huttlin EL, Guo A, Li J, Possemato A, Sowa ME, Rad R, Rush J, Comb MJ, Harper JW, Gygi SP.** Systematic and quantitative assessment of the ubiquitin-modified proteome. *Mol. Cell* 2011;44(2):325–340.

413. **Newton K, Matsumoto ML, Wertz IE, Kirkpatrick DS, Lill JR, Tan J, Dugger D, Gordon N, Sidhu SS, Fellouse FA, Komuves L, French DM, Ferrando RE, Lam C, Compaan D, Yu C, Bosanac I, Hymowitz SG, Kelley RF, Dixit VM.** Ubiquitin chain editing revealed by polyubiquitin linkage-specific antibodies. *Cell* 2008;134(4):668–678.

414. **Matsumoto ML, Dong KC, Yu C, Phu L, Gao X, Hannoush RN, Hymowitz SG, Kirkpatrick DS, Dixit VM, Kelley RF.** Engineering and structural characterization of a linear polyubiquitin-specific antibody. *J. Mol. Biol.* 2012;418(3–4):134–144.

415. **Peng J, Schwartz D, Elias JE, Thoreen CC, Cheng D, Marsischky G, Roelofs J, Finley D, Gygi SP.** A proteomics approach to understanding protein ubiquitination. *Nat. Biotechnol.* 2003;21(8):921–926.

416. **Xu P, Peng J.** Dissecting the ubiquitin pathway by mass spectrometry. *Biochim. Biophys. Acta* 2006;1764(12):1940–1947.

417. **Phu L, Izrael-Tomasevic A, Matsumoto ML, Bustos D, Dynek JN, Fedorova AV, Bakalarski CE, Arnott D, Deshayes K, Dixit VM, Kelley RF, Vucic D, Kirkpatrick DS.** Improved quantitative mass spectrometry methods for characterizing complex ubiquitin signals. *Mol. Cell Proteomics* 2011;10(5):M110.003756.

418. **Komander D, Clague MJ, Urbé S.** Breaking the chains: structure and function of the deubiquitinases. *Nat. Rev. Mol. Cell Biol.* 2009;10(8):550–563.

419. **Kar G, Keskin O, Nussinov R, Gursoy A.** Human proteome-scale structural modeling of E2-E3 interactions exploiting interface motifs. *J. Proteome Res.* 2012;11(2):1196–1207.

420. **Teicher BA, Ara G, Herbst R, Palombella VJ, Adams J.** The proteasome inhibitor PS-341 in cancer therapy. *Clin. Cancer Res.* 1999;5(9):2638–2645.

421. **Mitsiades CS, Davies FE, Laubach JP, Joshua D, San Miguel J, Anderson KC, Richardson PG.** Future directions of next-generation novel therapies, combination approaches, and the development of personalized medicine in myeloma. *J. Clin. Oncol.* 2011;29(14):1916–1923.

422. **Yang CH, Gonzalez-Angulo AM, Reuben JM, Booser DJ, Pusztai L, Krishnamurthy S, Esseltine D, Stec J, Broglio KR, Islam R, Hortobagyi GN, Cristofanilli M.** Bortezomib (VELCADE) in metastatic breast cancer: pharmacodynamics, biological effects, and prediction of clinical benefits. *Ann. Oncol.* 2006;17(5):813–817.

423. **Bloom J, Pagano M.** Deregulated degradation of the cdk inhibitor p27 and malignant transformation. *Semin. Cancer Biol.* 2003;13(1):41–47.

424. **Kamata Y, Watanabe J, Nishimura Y, Arai T, Kawaguchi M, Hattori M, Obokata A, Kuramoto H.** High expression of skp2 correlates with poor prognosis in endometrial endometrioid adenocarcinoma. *J. Cancer Res. Clin. Oncol.* 2005;131(9):591–596.

425. **Davidovich S, Ben-Izhak O, Shapira M, Futerman B, Hershko DD.** Over-expression of Skp2 is associated with resistance to preoperative doxorubicin-based chemotherapy in primary breast cancer. *Breast Cancer Res.* 2008;10(4):R63.

426. **Lecanda J, Parekh TV, Gama P, Lin K, Liarski V, Uretsky S, Mittal K, Gold LI.** Transforming growth factor-beta, estrogen, and progesterone converge on the regulation of p27Kip1 in the normal and malignant endometrium. *Cancer Res.* 2007;67(3):1007–1018.

427. **Huang K-T, Pavlides SC, Lecanda J, Blank SV, Mittal KR, Gold LI.** Estrogen and progesterone regulate p27kip1 levels via the ubiquitin-proteasome system: pathogenic and therapeutic implications for endometrial cancer. *PLoS ONE* 2012;7(9):e46072.

428. **Pavlides SC, Huang K-T, Reid DA, Wu L, Blank SV, Mittal K, Guo L, Rothenberg E, Rueda B, Cardozo T, Gold LI.** Inhibitors of SCF-Skp2/Cks1 E3 Ligase Block Estrogen-Induced Growth Stimulation and Degradation of Nuclear p27kip1: Therapeutic Potential for Endometrial Cancer. *Endocrinology* 2013;154(11):4030–4045.

429. **Ito T, Ando H, Suzuki T, Ogura T, Hotta K, Imamura Y, Yamaguchi Y, Handa H.** Identification of a primary target of thalidomide teratogenicity. *Science* 2010;327(5971):1345–1350.

430. **Chamberlain PP, Lopez-Girona A, Miller K, Carmel G, Pagarigan B, Chie-Leon B, Rychak E, Corral LG, Ren YJ, Wang M, Riley M, Delker SL, Ito T, Ando H, Mori T, Hirano Y, Handa H, Hakoshima T, Daniel TO, Cathers BE.** Structure of the human Cereblon-DDB1-lenalidomide complex reveals basis for responsiveness to thalidomide analogs. *Nat. Struct. Mol. Biol.* 2014;21(9):803–809.

431. **Fischer ES, Böhm K, Lydeard JR, Yang H, Stadler MB, Cavadini S, Nagel J, Serluca F, Acker V, Lingaraju GM, Tichkule RB, Schebesta M, Forrester WC, Schirle M, Hassiepen U, Ottl J, Hild M, Beckwith REJ, Harper JW, Jenkins JL, Thomä NH.** Structure of the DDB1-CRBN E3 ubiquitin ligase in complex with thalidomide. *Nature* 2014;512(7512):49–53.

432. **Lu G, Middleton RE, Sun H, Naniong M, Ott CJ, Mitsiades CS, Wong K-K, Bradner JE, Kaelin WG.** The myeloma drug lenalidomide promotes the cereblon-dependent destruction of Ikaros proteins. *Science* 2014;343(6168):305–309.

433. **Martiniani R, Di Loreto V, Di Sano C, Lombardo A, Liberati AM.** Biological activity of lenalidomide and its underlying therapeutic effects in multiple myeloma. *Adv Hematol* 2012;2012:842945.

434. **Terpos E, Kanellias N, Christoulas D, Kastritis E, Dimopoulos MA.** Pomalidomide: a novel drug to treat relapsed and refractory multiple myeloma. *Onco Targets Ther* 2013;6:531–538.

435. **Chau V, Tobias JW, Bachmair A, Marriott D, Ecker DJ, Gonda DK, Varshavsky A.** A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. *Science* 1989;243(4898):1576–1583.

436. **Al-Hakim A, Escribano-Diaz C, Landry M-C, O'Donnell L, Panier S, Szilard RK, Durocher D.** The ubiquitous role of ubiquitin in the DNA damage response. *DNA Repair (Amst.)* 2010;9(12):1229–1240.

437. **Mukhopadhyay D, Riezman H.** Proteasome-independent functions of ubiquitin in endocytosis and signaling. *Science* 2007;315(5809):201–205.

438. **Spence J, Gali RR, Dittmar G, Sherman F, Karin M, Finley D.** Cell cycle-regulated modification of the ribosome by a variant multiubiquitin chain. *Cell* 2000;102(1):67–76.

439. **Kirisako T, Kamei K, Murata S, Kato M, Fukumoto H, Kanie M, Sano S, Tokunaga F, Tanaka K, Iwai K.** A ubiquitin ligase complex assembles linear polyubiquitin chains. *EMBO J.* 2006;25(20):4877–4887.

440. **Tokunaga F, Sakata S, Saeki Y, Satomi Y, Kirisako T, Kamei K, Nakagawa T, Kato M, Murata S, Yamaoka S, Yamamoto M, Akira S, Takao T, Tanaka K, Iwai K.** Involvement of linear polyubiquitylation of NEMO in NF-kappaB activation. *Nat. Cell Biol.* 2009;11(2):123–132.

441. **Rahighi S, Ikeda F, Kawasaki M, Akutsu M, Suzuki N, Kato R, Kensche T, Uejima T, Bloor S, Komander D, Randow F, Wakatsuki S, Dikic I.** Specific recognition of linear ubiquitin chains by NEMO is important for NF-kappaB activation. *Cell* 2009;136(6):1098–1109.

442. **Hoege C, Pfander B, Moldovan G-L, Pyrowolakis G, Jentsch S.** RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. *Nature* 2002;419(6903):135–141.

443. **Freudenthal BD, Gakhar L, Ramaswamy S, Washington MT.** Structure of monoubiquitinated PCNA and implications for translesion synthesis and DNA polymerase exchange. *Nat. Struct. Mol. Biol.* 2010;17(4):479–484.

444. **Bremm A, Komander D.** Emerging roles for Lys11-linked polyubiquitin in cellular regulation. *Trends Biochem. Sci.* 2011;36(7):355–363.

445. **Jin L, Williamson A, Banerjee S, Philipp I, Rape M.** Mechanism of ubiquitin-chain formation by the human anaphase-promoting complex. *Cell* 2008;133(4):653–665.

446. **Wickliffe KE, Williamson A, Meyer H-J, Kelly A, Rape M.** K11-linked ubiquitin chains as novel regulators of cell division. *Trends Cell Biol.* 2011;21(11):656–663.

447. **Dynek JN, Goncharov T, Dueber EC, Fedorova AV, Izrael-Tomasevic A, Phu L, Helgason E, Fairbrother WJ, Deshayes K, Kirkpatrick DS, Vucic D.** c-IAP1 and UbcH5 promote K11-linked polyubiquitination of RIP1 in TNF signalling. *EMBO J.* 2010;29(24):4198–4209.

448. **Goto E, Yamanaka Y, Ishikawa A, Aoki-Kawasumi M, Mito-Yoshida M, Ohmura-Hoshino M, Matsuki Y, Kajikawa M, Hirano H, Ishido S.** Contribution of lysine 11-linked ubiquitination to MIR2-mediated major histocompatibility complex class I internalization. *J. Biol. Chem.* 2010;285(46):35311–35319.

449. **Morris JR, Solomon E.** BRCA1 : BARD1 induces the formation of conjugated ubiquitin structures, dependent on K6 of ubiquitin, in cells during DNA replication and repair. *Hum. Mol. Genet.* 2004;13(8):807–817.

450. **Wu-Baer F, Lagrazon K, Yuan W, Baer R.** The BRCA1/BARD1 heterodimer assembles polyubiquitin chains through an unconventional linkage involving lysine residue K6 of ubiquitin. *J. Biol. Chem.* 2003;278(37):34743–34746.

451. **Sobhian B, Shao G, Lilli DR, Culhane AC, Moreau LA, Xia B, Livingston DM, Greenberg RA.** RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites. *Science* 2007;316(5828):1198–1202.

452. **Wu W, Nishikawa H, Hayami R, Sato K, Honda A, Aratani S, Nakajima T, Fukuda M, Ohta T.** BRCA1 ubiquitinates RPB8 in response to DNA damage. *Cancer Res.* 2007;67(3):951–958.

453. **Glauser L, Sonnay S, Stafa K, Moore DJ.** Parkin promotes the ubiquitination and degradation of the mitochondrial fusion factor mitofusin 1. *J. Neurochem.* 2011;118(4):636–645.

454. **Geisler S, Holmström KM, Skujat D, Fiesel FC, Rothfuss OC, Kahle PJ, Springer W.** PINK1/Parkin-mediated mitophagy is dependent on VDAC1 and p62/SQSTM1. *Nat. Cell Biol.* 2010;12(2):119–131.

455. **Peng D-J, Zeng M, Muromoto R, Matsuda T, Shimoda K, Subramaniam M, Spelsberg TC, Wei W-Z, Venuprasad K.** Noncanonical K27-linked polyubiquitination of TIEG1 regulates Foxp3 expression and tumor growth. *J. Immunol.* 2011;186(10):5638–5647.

456. **Durcan TM, Kontogiannea M, Thorarinsdottir T, Fallon L, Williams AJ, Djarmati A, Fantaneanu T, Paulson HL, Fon EA.** The Machado-Joseph disease-associated mutant form of ataxin-3 regulates parkin ubiquitination and stability. *Hum. Mol. Genet.* 2011;20(1):141–154.

457. **Al-Hakim AK, Zagorska A, Chapman L, Deak M, Peggie M, Alessi DR.** Control of AMPK-related kinases by USP9X and atypical Lys(29)/Lys(33)-linked polyubiquitin chains. *Biochem. J.* 2008;411(2):249–260.

458. **Chastagner P, Israël A, Brou C.** Itch/AIP4 mediates Deltex degradation through the formation of K29-linked polyubiquitin chains. *EMBO Rep.* 2006;7(11):1147–1153.

459. **Licchesi JDF, Mieszczanek J, Mevissen TET, Rutherford TJ, Akutsu M, Virdee S, El Oualid F, Chin JW, Ovaa H, Bienz M, Komander D.** An ankyrin-repeat ubiquitin-binding domain determines TRABID's specificity for atypical ubiquitin chains. *Nat. Struct. Mol. Biol.* 2012;19(1):62–71.

460. **Hwang C-S, Shemorry A, Auerbach D, Varshavsky A.** The N-end rule pathway is mediated by a complex of the RING-type Ubr1 and HECT-type Ufd4 ubiquitin ligases. *Nat. Cell Biol.* 2010;12(12):1177–1185.

461. **Fan M, Park A, Nephew KP.** CHIP (Carboxyl Terminus of Hsc70-Interacting Protein) Promotes Basal and Geldanamycin-Induced Degradation of Estrogen Receptor-α. *Molecular Endocrinology* 2005;19(12):2901–2914.

462. **Nakajima A, Maruyama S, Bohgaki M, Miyajima N, Tsukiyama T, Sakuragi N, Hatakeyama S.** Ligand-dependent transcription of estrogen receptor α is mediated by the ubiquitin ligase EFP. *Biochemical and Biophysical Research Communications* 2007;357(1):245–251.

463. **Byun B, Jung Y.** Repression of transcriptional activity of estrogen receptor alpha by a Cullin3/SPOP ubiquitin E3 ligase complex. *Mol. Cells* 2008;25(2):289–293.

464. **Gustafsson N, Zhao C, Gustafsson J-Å, Dahlman-Wright K.** RBCK1 Drives Breast Cancer Cell Proliferation by Promoting Transcription of Estrogen Receptor α and Cyclin B1. *Cancer Res* 2010;70(3):1265–1274.

465. **Pan X, Zhou T, Tai Y-H, Wang C, Zhao J, Cao Y, Chen Y, Zhang P-J, Yu M, Zhen C, Mu R, Bai Z-F, Li H-Y, Li A-L, Liang B, Jian Z, Zhang W-N, Man J-H, Gao Y-F, Gong W-L, Wei L-X, Zhang X-M.** Elevated expression of CUEDC2 protein confers endocrine resistance in breast cancer. *Nat. Med.* 2011;17(6):708–714.

466. **Qi J, Tripathi M, Mishra R, Sahgal N, Fazil L, Ettinger S, Placzek WJ, Claps G, Chung LWK, Bowtell D, Gleave M, Bhowmick N, Ronai ZA.** The E3 Ubiquitin Ligase Siah2 Contributes to Castration-Resistant Prostate Cancer by Regulation of Androgen Receptor Transcriptional Activity. *Cancer Cell* 2013;23(3):332–346.

467. **Sultana R, Theodoraki MA, Caplan AJ.** Specificity in the actions of the UBR1 ubiquitin ligase in the degradation of nuclear receptors. *FEBS Open Bio* 2013;3:394–397.

468. **Li B, Lu W, Yang Q, Yu X, Matusik RJ, Chen Z.** Skp2 regulates androgen receptor through ubiquitin-mediated degradation independent of Akt/mTOR pathways in prostate cancer. *Prostate* 2014;74(4):421–432.

469. **Malyukova A, Brown S, Papa R, O'Brien R, Giles J, Trahair TN, Dalla Pozza L, Sutton R, Liu T, Haber M, Norris MD, Lock RB, Sangfelt O, Marshall GM.** FBXW7 regulates glucocorticoid response in T-cell acute lymphoblastic leukaemia by targeting the glucocorticoid receptor for degradation. *Leukemia* 2013;27(5):1053–1062.

470. **Zhang P-J, Zhao J, Li H-Y, Man J-H, He K, Zhou T, Pan X, Li A-L, Gong W-L, Jin B-F, Xia Q, Yu M, Shen B-F, Zhang X-M.** CUE domain containing 2 regulates degradation of progesterone receptor by ubiquitin–proteasome. *The EMBO Journal* 2007;26(7):1831–1842.

471. **Jing X, Infante J, Nachtman RG, Jurecic R.** E3 ligase FLRF (Rnf41) regulates differentiation of hematopoietic progenitors by governing steady-state levels of cytokine and retinoic acid receptors. *Experimental Hematology* 2008;36(9):1110–1120.

472. **Kilroy G, Kirk-Ballard H, Carter LE, Floyd ZE.** The Ubiquitin Ligase Siah2 Regulates PPARγ Activity in Adipocytes. *Endocrinology* 2012;153(3):1206–1218.

473. **Ren Y, Jiang H, Ma D, Nakaso K, Feng J.** Parkin degrades estrogen-related receptors to limit the expression of monoamine oxidases. *Hum. Mol. Genet.* 2011;20(6):1074–1083.

474. **Beitel LK, Elhaji YA, Lumbroso R, Wing SS, Panet-Raymond V, Gottlieb B, Pinsky L, Trifiro MA.** Cloning and characterization of an androgen receptor N-terminal-interacting protein with ubiquitin-protein ligase activity. *J. Mol. Endocrinol.* 2002;29(1):41–60.

475. **Fan S, Wang J-A, Yuan R, Ma Y, Meng Q, Erdos MR, Pestell RG, Yuan F, Auborn KJ, Goldberg ID, Rosen EM.** BRCA1 Inhibition of Estrogen Receptor Signaling in Transfected Cells. *Science* 1999;284(5418):1354–1356.

476. **Zheng L, Annab LA, Afshari CA, Lee WH, Boyer TG.** BRCA1 mediates ligand-independent transcriptional repression of the estrogen receptor. *Proc. Natl. Acad. Sci. U.S.A.* 2001;98(17):9587–9592.

477. **Kawai H, Li H, Chun P, Avraham S, Avraham HK.** Direct interaction between BRCA1 and the estrogen receptor regulates vascular endothelial growth factor (VEGF) transcription and secretion in breast cancer cells. *Oncogene* 2002;21(50):7730–7739.

478. **Inoue S, Orimo A, Hosoi T, Kondo S, Toyoshima H, Kondo T, Ikegami A, Ouchi Y, Orimo H, Muramatsu M.** Genomic binding-site cloning reveals an estrogen-responsive gene that encodes a RING finger protein. *PNAS* 1993;90(23):11117–11121.

479. **Talukder AH, Mishra SK, Mandal M, Balasenthil S, Mehta S, Sahin AA, Barnes CJ, Kumar R.** MTA1 interacts with MAT1, a cyclin-dependent kinase-activating kinase complex ring finger factor, and regulates estrogen receptor transactivation functions. *J. Biol. Chem.* 2003;278(13):11676–11685.

480. **Helenius K, Yang Y, Alasaari J, Mäkelä TP.** Mat1 inhibits peroxisome proliferator-activated receptor gamma-mediated adipocyte differentiation. *Mol. Cell. Biol.* 2009;29(2):315–323.

481. **Takano Y, Adachi S, Okuno M, Muto Y, Yoshioka T, Matsushima-Nishiwaki R, Tsurumi H, Ito K, Friedman SL, Moriwaki H, Kojima S, Okano Y.** The RING finger protein, RNF8, interacts with retinoid X receptor alpha and enhances its transcription-stimulating activity. *J. Biol. Chem.* 2004;279(18):18926–18934.

482. **Johnsen SA, Güngör C, Prenzel T, Riethdorf S, Riethdorf L, Taniguchi-Ishigaki N, Rau T, Tursun B, Furlow JD, Sauter G, Scheffner M, Pantel K, Gannon F, Bach I.** Regulation of estrogen-dependent transcription by the LIM cofactors CLIM and RLIM in breast cancer. *Cancer Res.* 2009;69(1):128–136.

483. **vom Baur E, Zechel C, Heery D, Heine MJ, Garnier JM, Vivat V, Le Douarin B, Gronemeyer H, Chambon P, Losson R.** Differential ligand-dependent interactions between the AF-2 activating domain of nuclear receptors and the putative transcriptional intermediary factors mSUG1 and TIF1. *EMBO J.* 1996;15(1):110–124.

484. **Thénot S, Henriquet C, Rochefort H, Cavaillès V.** Differential interaction of nuclear receptors with the putative human transcriptional coactivator hTIF1. *J. Biol. Chem.* 1997;272(18):12062–12068.

485. **Teyssier C, Ou C-Y, Khetchoumian K, Losson R, Stallcup MR.** Transcriptional intermediary factor 1alpha mediates physical interaction and functional synergy between the coactivator-associated arginine methyltransferase 1 and glucocorticoid receptor-interacting protein 1 nuclear receptor coactivators. *Mol. Endocrinol.* 2006;20(6):1276–1286.

486. **Moilanen AM, Poukka H, Karvonen U, Häkli M, Jänne OA, Palvimo JJ.** Identification of a novel RING finger protein as a coregulator in steroid receptor-mediated gene transcription. *Mol. Cell. Biol.* 1998;18(9):5128–5139.

487. **Poukka H, Aarnisalo P, Santti H, Jänne OA, Palvimo JJ.** Coregulator small nuclear RING finger protein (SNURF) enhances Sp1- and steroid receptor-mediated transcription by different mechanisms. *J. Biol. Chem.* 2000;275(1):571–579.

488. **Saville B, Poukka H, Wormke M, Janne OA, Palvimo JJ, Stoner M, Samudio I, Safe S.** Cooperative coactivation of estrogen receptor alpha in ZR-75 human breast cancer cells by SNURF and TATA-binding protein. *J. Biol. Chem.* 2002;277(4):2485–2497.