Binding Profiles and Transcriptomes and Therapeutic Resistance, Oh My!
The Regulation of ERα Action in Breast Cancer

By

Kim Stauffer

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Molecular Pathology & Immunology

February 28, 2022

Nashville, Tennessee

Approved:

Jay Jerome, Ph.D.

Andries Zijlstra, Ph.D.

William Tansey, Ph.D.

Vivian Weiss, M.D., Ph.D.

Thomas Stricker, M.D., Ph.D.

DEDICATION

This dissertation is dedicated to the memory of:

My grandma Bonita Stauffer, who passed of breast cancer before I was able to get to know her.

My sister-in-law Cyndie Stauffer, who lost her battle with lung cancer while I was in my studies.

Her son, my nephew Jacob Stauffer, a bright young man whose curiosity and resilience will forever inspire me.

ACKNOWLEDGMENTS

The work contained within this dissertation was made possible with the financial support of the Cellular, Biochemical and Molecular Sciences Training Program (CBMS), the Vanderbilt Institute for Clinical and Translational Research (VICTR), and the National Cancer Institute (NCI). The program managers Carolyn Berry (Interdisciplinary Graduate Program, IGP), Whit Adams (retired, Pathology Microbiology and Immunology, PMI), Kristi Hargrove (formerly PMI), and current PMI program manager Liz Roelofsz have provided me with invaluable support along the way. Assistant director of the Graduate Programs in Biomedical Sciences Beth Bowman provided me with encouragement and compassion when I needed it most. I am grateful for the guidance and feedback given by PMI DGS Jay Jerome, as well as from Jim Patton as director of the CBMS training grant. I am so appreciative of the staff at VANTAGE sequencing core for handling my sequencing samples with care and speed, as well as for teaching me how to use the chromatin sonicator. My graduate work has been enriched through the work with all my collaborators, Eric Wright, Kasia Ludwik, Luigi Formisano, Sarah Croessman, and Valerie Jansen. I am very grateful to Deb Lannigan for running a fulfilling and fun journal club for the program. Each member of my dissertation committee has provided me with important advice, scientific and professional, that has contributed to my success in graduate school. I would also like to thank the past members of the Stricker lab, Ben Bulen, Erin Fey, Brian Cholewa, and Vivian Weiss, for being excellent coworkers. The Arteaga lab group, led by Carlos Arteaga, was instrumental to me finding my footing in the early and middle years of my graduate career. Dr. Arteaga was an excellent co-mentor to me, and his weekly journal clubs were very often a highlight of my week. I am indebted to my mentor, Tom Stricker, who never gave up on me despite all the obstacles I faced in lab and in life. He taught me not only how to code, run bench experiments, and give compelling presentations, but how to truly find the most exciting possibilities in science. I am nothing without my graduate school friends and their various forms of support, including group chats, nights dancing on Broadway, training for half marathons, and impromptu ice cream dates to cry over failed experiments. These friends include Caleigh Azumaya, Stephanie Moore-Lotridge, Diana Healey, Diane Saunders, and Meredith Frazier. I want to thank my partner in life, Spencer Alexander, for always being on my team, and for growing with me through these years. I owe the ultimate thanks to my parents, Greg and Vickie Stauffer, who have always supported me and the dogged pursuit of my dreams.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

Abbreviation	Meaning	Page
AACR	American Association for Cancer Research	6
AI	Aromatase Inhibitor	4
AI	Artificial Intelligence	102
ASCOM	Activating Signal Cointegrator-2 -Containing Complex	8
bp	Base-pair	14
ChIP-seq	Chromatin Immunoprecipitation Sequencing	8
DEG	Differentially Expressed Genes	13
ER+	Estrogen Receptor Positive	2
E2	Estradiol	2
ERα	Estrogen Receptor Alpha	1
ERE	Estrogen Response Element	3
ETR	Endocrine Therapy Resistance	6
ESR1	Estrogen Receptor 1	2
FulvPalb	Fulvestrant Palbociclib Combination	39
GSEA	Gene Set Enrichment Analysis	15
GSVA	Gene Set Variation Analysis	40
HR+	Hormone Receptor Positive	1
IDR	Irreproducible Discovery Rate	21
KD	Knockdown	10
KMT2C	Lysine Methyltransferase 2C	8
LBD	Ligand Binding Domain	6
LTED	Long Term Estradiol Deprivation	46
ML	Machine Learning	102
MLL3	Mixed Lineage Leukemia 3	8
MSigDB	Molecular Signature Database	15
NCL	Non-clonogenic Luminal	9
ORA	Overrepresentation Analysis	15
PFS	Progression Free Survival	4
ppg	Peaks Per Gene	30
PR	Progesterone Receptor	1
PTM	Post Translational Modification	8
RNA-seq	RNA-sequencing	8
ROS	Reactive Oxygen Species	72
RTK	Receptor Tyrosine Kinase	45
SERD	Selective Estrogen Receptor Degrader/Downregulator	4
SERM	Selective Estrogen Receptor Modulator	4

SVA	Surrogate Variable Analysis	13
TCGA	The Cancer Genome Atlas	5
TF	Transcription Factor	16
TMD	Transmembrane Domain	45
TNBC	Triple Negative Breast Cancer	1
WT	Wildtype	10

CHAPTER I

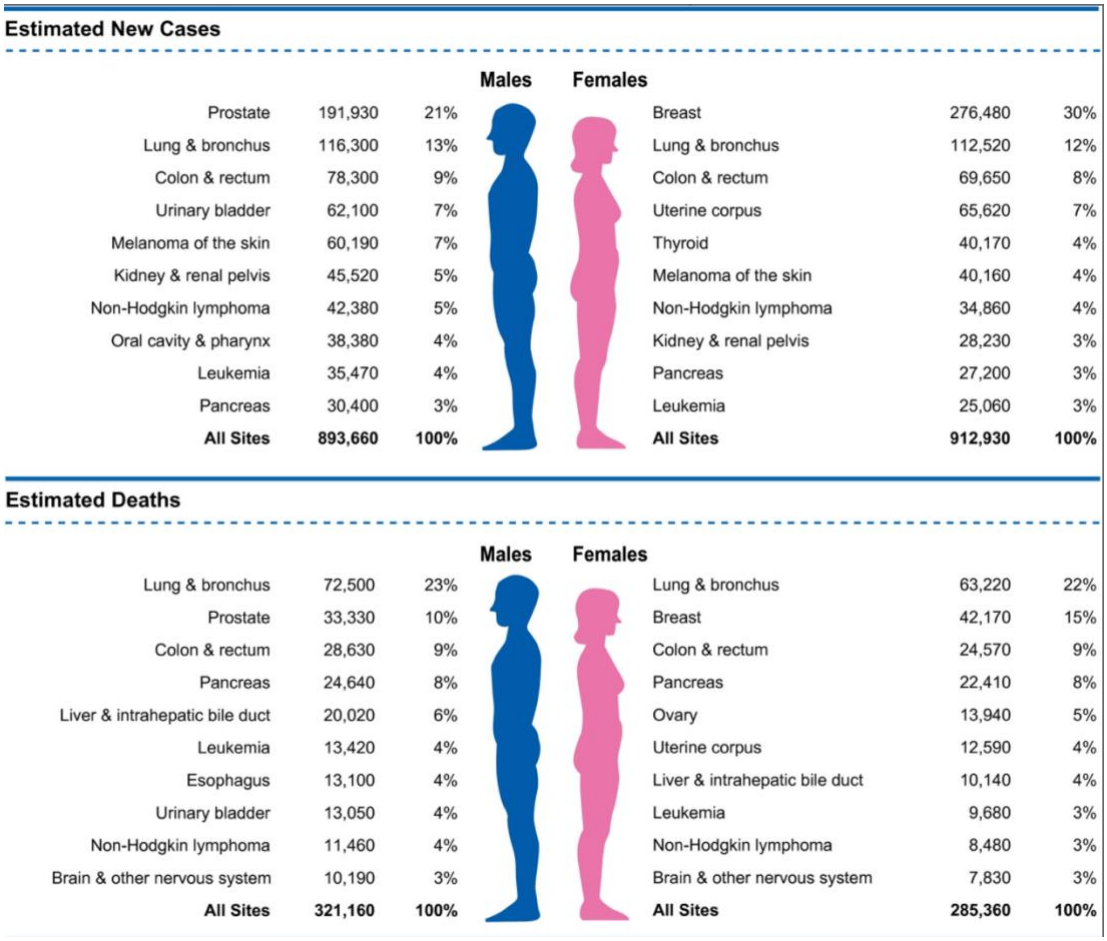
INTRODUCTION

Overview of Breast Cancer in the United States

Breast cancer accounts for 30% of all the cancers diagnosed in women in the United States, and is the second leading cause of cancer-related death (**Fig. 1-1**).[1] One in eight American women will be diagnosed with breast cancer during their lifetime. Approximately 70-80% of these breast cancer cases are considered hormone receptor positive (HR+), meaning they express one or both estrogen receptor alpha (ERα) and one of its transcriptional targets progesterone receptor (PR). The other major groups of breast cancers are comprised of those that either overexpress HER2 (HER2+) or are negative for ER expression, PR expression, and HER2 overexpression. The latter are called triple negative breast cancers (TNBC) and make up 10-15% of all cases of breast cancer. HER2+ breast cancers make up 15% of cases and are eligible for treatment targeting HER2 overexpression such as trastuzumab.

To better understand how breast cancers form, normal development and maintenance of the mammary gland must be understood. Much of the development of the mammary gland occurs after birth, including the stages of ductal morphogenesis, alveologenesis, lactation, and involution. Many different cell types form this dynamic organ; a fat pad formed by adipocytes, vascular endothelial cells for blood vessels, fibroblasts and immune cells in stroma, and two epithelial compartments forming the ductal network. Basal epithelium, made of myoepithelial cells, makes up the outer layer of the gland while luminal epithelium forms ducts and secretory alveoli. The hormones and growth factors that signal for morphological changes at different stages of life are implicated in breast cancer formation.[2–4] While some risk factors for developing breast cancer are straight forward, such as being a woman, advanced age, and family history of breast cancer, others are less intuitive. Breast cancer risk is indeed increased with early menarche, late menopause, late age at first pregnancy and low parity.[5–7] These risk factors are heavily influenced by levels of estrogen, both endogenous and exogenous.

HR+ breast cancer cases have seen a 0.3% increase in incidence per year between 2012-2016. This increase is thought to be driven by a decline in parity rates and physical activity, and increases in age at first birth, obesity, and alcohol consumption.[8] However despite this continued increase in incidence, the standard of care for HR+ cancers has remained anti-estrogen, or endocrine, therapy. While many HR+ breast cancers initially respond to anti-estrogen therapy, up to 40% of them have intrinsic potential to become resistant to endocrine therapies.[9] Furthermore, almost all advanced disease patients will eventually relapse despite being given antiestrogen therapy, due to either de novo or acquired endocrine resistance.[10,11] In order to better treat these patients, causes of endocrine therapy resistance must be identified. Discovery of the molecular events that allow cancer cells to circumvent treatment will not only provide further targets for additional or combinatorial therapy but will also lead to treatment-response biomarkers that inform the best treatment plan for each individual cancer patient. My thesis work focuses on the molecular underpinnings of this clinical dilemma.



Estimated New Cases

Males			Females		
Prostate	191,930	21%	Breast	276,480	30%
Lung & bronchus	116,300	13%	Lung & bronchus	112,520	12%
Colon & rectum	78,300	9%	Colon & rectum	69,650	8%
Urinary bladder	62,100	7%	Uterine corpus	65,620	7%
Melanoma of the skin	60,190	7%	Thyroid	40,170	4%
Kidney & renal pelvis	45,520	5%	Melanoma of the skin	40,160	4%
Non-Hodgkin lymphoma	42,380	5%	Non-Hodgkin lymphoma	34,860	4%
Oral cavity & pharynx	38,380	4%	Kidney & renal pelvis	28,230	3%
Leukemia	35,470	4%	Pancreas	27,200	3%
Pancreas	30,400	3%	Leukemia	25,060	3%
All Sites	**893,660**	**100%**	**All Sites**	**912,930**	**100%**

Estimated Deaths

Males			Females		
Lung & bronchus	72,500	23%	Lung & bronchus	63,220	22%
Prostate	33,330	10%	Breast	42,170	15%
Colon & rectum	28,630	9%	Colon & rectum	24,570	9%
Pancreas	24,640	8%	Pancreas	22,410	8%
Liver & intrahepatic bile duct	20,020	6%	Ovary	13,940	5%
Leukemia	13,420	4%	Uterine corpus	12,590	4%
Esophagus	13,100	4%	Liver & intrahepatic bile duct	10,140	4%
Urinary bladder	13,050	4%	Leukemia	9,680	3%
Non-Hodgkin lymphoma	11,460	4%	Non-Hodgkin lymphoma	8,480	3%
Brain & other nervous system	10,190	3%	Brain & other nervous system	7,830	3%
All Sites	**321,160**	**100%**	**All Sites**	**285,360**	**100%**

Figure 1-1. Leading Types of Cancer for Estimated New Cancer Cases and Deaths. These estimates, which are based on modeled projections, are for the United States from 2020 and are rounded to the nearest 10. Basal cell and squamous cell skin cancers, and in situ carcinoma other than bladder are excluded. This figure and its legend are taken from Siegel et. al.[12]

The Action of Estrogen Receptor Alpha in Breast Cancer

To better understand how HR+ breast cancer cells survive and proliferate despite anti-estrogen therapy, first the importance of estrogen receptor signaling to breast cancer cell behavior must be grasped. Because of the central role the estrogen receptor has in breast cancer behavior, I will refer to HR+ cancer as ER+ cancer from now on. Estrogen receptors are nuclear hormone receptors that act as transcription factors. The main estrogen receptor that functions in breast cells is ERα, the protein product of gene ESR1 (Estrogen Receptor 1).[13] Its behavior is regulated by estrogens, steroid hormones derived from cholesterol that diffuse across the plasma membrane. The main circulating estrogen, 17-β estradiol or E2, is responsible for many physiological maintenance processes. These include, but are not limited to, maintenance of bone mass and cognitive function, regulation of insulin responsiveness, and development of secondary sex characteristics.[14] Deregulation of estrogen signaling is involved in pathophysiological processes, such as the initiation and development of ovarian, endometrial, and breast cancers.[15]

Estradiol can signal through several interconnected pathways; one of the most important to ER+ breast cancer is the ERα canonical signaling pathway which I will focus on here (**Figure 1-2**). The Cys447 residue of ERα is palmitoylated with help from heat shock protein 27; this allows ERα to interact with caveolin-1 for the transport of the receptor to the cell membrane.[16–19] ERα monomers bind to E2 which induces their dimerization.[20] Depalmitoylation of dimerized ERα leads to decreased association with caveolin-1, and the receptor dimer travels to the nucleus. Here E2-bound ER dimers either bind directly to the DNA to regulate transcription at estrogen response elements (EREs), or indirectly via interactions with other transcription factors that act as co-regulators of transcription of its gene targets. EREs are palindromic DNA sequences usually

found in the distant enhancers of target genes.[21–23] They consist either mostly or exactly of the consensus sequence 5′-GGTCAnnnTGACC-3′, where 'n' is a nonspecific nucleotide.[24] Coregulators are recruited to the ERα binding sites to modulate the target gene transcription.[25] Some of these coregulators recruit chromatin-modifying proteins to aid in activation or repression of the transcriptional target.[26] Importantly, pioneer factors such as FOXA1, GATA3, and PBX1 assist in creating an open chromatin conformation for ER-chromatin interactions.[27–29] Once bound to the chromatin either directly or indirectly through coregulators such as AP1 or SP1, ERα and the enhancer region is looped over to interact with the transcriptional machinery at the proximal promoter region of the target gene.

ERα is the major effector of estrogen signaling that leads to breast cancer growth in a hormone-dependent setting.[30] The nuclear receptor regulates the transcription of many target genes important to the survival and proliferation of breast cancer cells. In fact, two of the earliest defined targets of ERα in the history of estrogen receptor research include c-MYC and cyclin D1. Examples of the consequential genes estradiol can stimulate transcription of through ERα are both the oncogene c-MYC, involved in mitogen-stimulated cell growth[31], and cyclin D1, which initiates progression past phase G1 of the cell cycle.[32] Furthermore, ERα propels oncogenic properties in breast cancer through controlling the expression of GREB1, which contributes to cell growth, PR, which utilizes paracrine signaling to induce proliferation of neighboring cells, and Fos, which transcriptionally regulates proliferation and survival genes in combination with Jun.[33,34] ERα also regulates the transcription of Wnt11, an anti-apoptotic factor that increases breast cancer cell survival.[35] These genes are only a few of the well-known targets of ERα transcriptional control; they exist in networks of tens of hundreds of genes that the estrogen receptor regulates transcription of to drive ER+ breast cancer. These networks form the basis for molecular subtypes, summarized later in this chapter, and are the reason that antiestrogen therapy targeting their expression has been a mainstay of ER+ treatment for several decades.

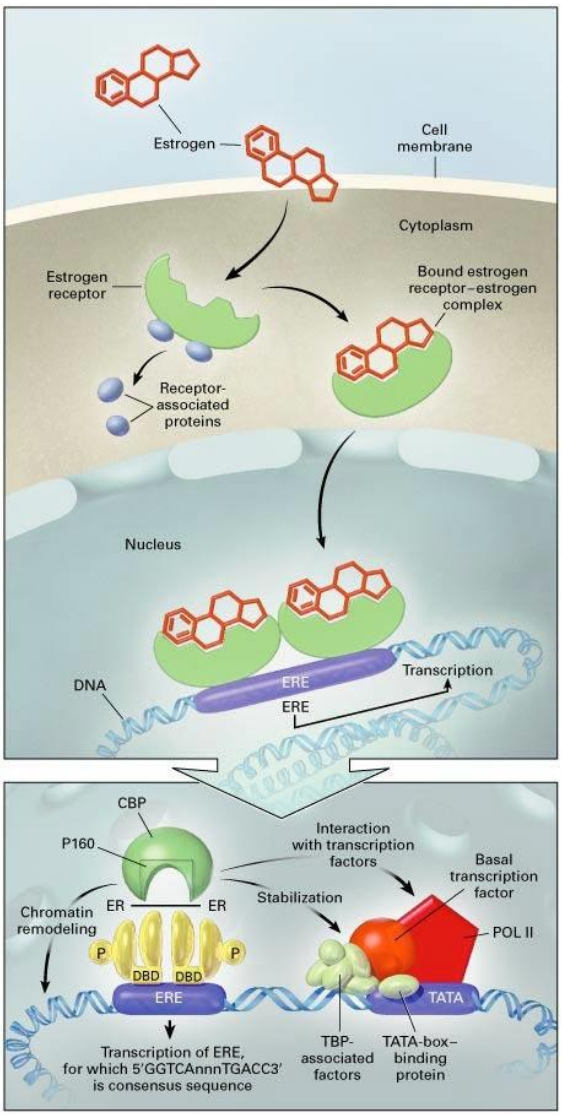


Figure 1-2. The Classical Estrogen Signaling Pathway. First estrogen (E2) binds to ERα, causing a release from its chaperones, or receptor-associated proteins, and inducing dimerization of the E2-ERα complexes. The E2-ERα dimer translocates to the nucleus and binds estrogen response elements in the DNA. The hormone–receptor complex thus initiates transcription of its targets. This figure is from Gruber et al.[36]

The Use of Antiestrogen Therapy in Treating ER+ Breast Cancer

If nuclei-localized ERα is detected with diagnostic testing, breast cancer patients qualify to be treated with anti-estrogen therapy (**Figure 1-3**). For most patients this looks like a combination of the available classes of antiestrogens, as combination therapies are more successful at achieving tumor regression than monotherapy.[37] Part of the combination therapy includes selective estrogen receptor modulators (SERMs), of which tamoxifen is the most widely used.[38,39] This category of drugs works in competition with E2 by binding to ERα, such that the resulting ERα structure can no longer recruit cofactors at the same capacity.[38] The other two categories of antiestrogen therapies are selective estrogen receptor downregulators (SERDs), and aromatase inhibitors (AIs). SERDs, of which the most widely used is fulvestrant, bind to ERα to induce a structural change disallowing cofactors interactions and marking the receptor for proteasomal degradation.[40] AIs come in two flavors: Type I AIs like exemestane are steroidal and irreversibly bind to aromatase, the enzyme that converts androgens to estrogens, causing permanent inactivation and eventual degradation. Type II AIs like anastrozole and letrozole are non-steroidal; they compete with androgens by reversibly binding to aromatase.[41,42]

Despite the initial tumor regression achieved by tamoxifen treatment in many cases of ER+ breast cancer, almost half of patients with advanced cases of ER+ breast cancer present with *de novo* endocrine therapy resistance. This situation is defined as primary resistance, or disease progression within 6 months of antiestrogen therapy in advanced breast cancer. Furthermore, metastatic cases eventually stop responding to tamoxifen.[10,11] This situation is defined as acquired resistance, or disease progression after at least 6 months of antiestrogen therapy in advanced breast cancer.[43] However, most patients who have clinical relapse after initial success with tamoxifen present with retention of ERα expression. Furthermore, clinical trials show that SERD fulvestrant extends progression-free survival (PFS) in patients who became resistant to tamoxifen.[44–47] Thus, ERα continues to play an important role in breast cancers that never respond or stop responding to tamoxifen.

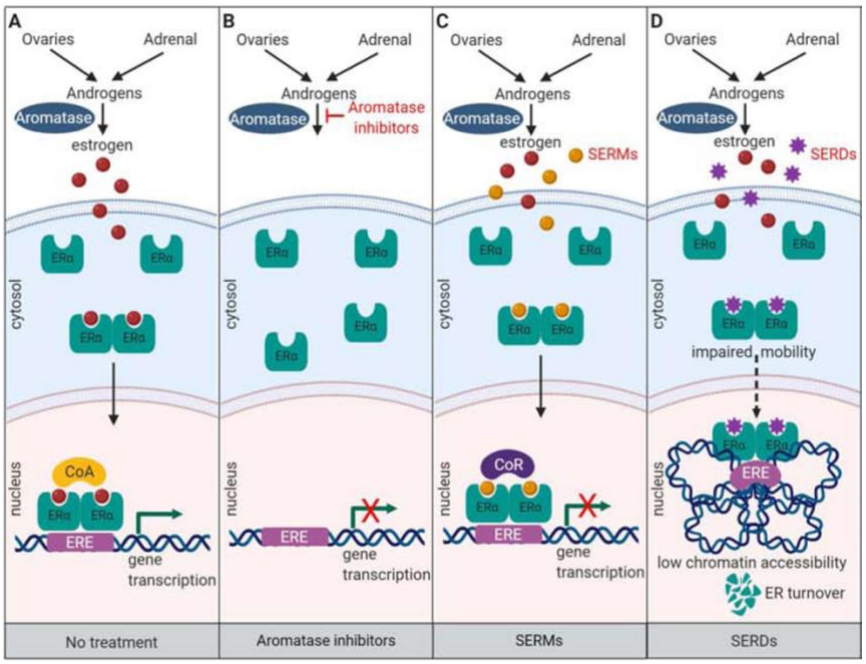


Figure 1-3. Class and Mechanism of Endocrine Therapies. In Panel **A**, several tissues such as the ovaries and adrenal glands produce androgens that aromatase converts to estrogens. The classical estrogen signaling pathway follows. In Panel **B**, aromatase inhibitors block the aromatization of androgens to estrogen. Panel **C** shows selective estrogen receptor modulators that compete with estrogen binding to ERα. While ER dimers bound by SERM may still interact with chromatin, their association with co-repressors (CoR) inhibits transcriptional activation of ERα targets in the breast. Lastly Panel **D** exhibits selective estrogen receptor downregulators, which impair the translocation of ER dimers to the nucleus, lower the accessibility of the chromatin for the ER dimer, and lead to faster ER turnover by degradation. This figure is from Hanker et al.[48]

History and Utility of Genomics in the Fight Against Cancer

The advent of genomics has allowed for deeper classification of breast cancers that, in combination with histological categorization, inform not only the course of treatments available but also the projected behavior and response of the cancer to treatment. Histological categorization of breast cancer utilizes growth patterns and cytological features to distinguish between subtypes, the majority of which are either ductal or lobular. Molecular markers are routinely included in the categorization to better determine treatment strategies and prognosis, including ER, PR, HER2, and p53.[49] Molecular categorization arose when Perou et al. published a seminal molecular portraits paper in 2000, based on gene expression patterns from 65 surgical human breast tumours.[50] This work used hierarchical clustering of expression data using arrays, and established four subtypes of breast cancer: ER+/luminal-like, basal-like, HER2-enriched, and normal-breast like (**Figure 1-4**). These groupings mostly, but not completely, overlapped with ER+/PR+, HER2+, and TNBC, with luminal A and luminal B as two different subtypes of ER+ breast cancer.

Since this original paper has been published, molecular subtypes have been expanded on the basis of further human and murine breast tumor gene expression data to six types overall: claudin-low, basal-like, HER2-enriched, normal breast-like, luminal A, and luminal B.[51–53] The clinical heterogeneity of ER+ BRCA is captured in these expression profiles as well as mutation profiles. Information about risk of recurrence and response to therapy gleaned from the combination of histological and molecular subtypes has become integral to breast cancer management plans.[54,55] Gene expression assays like Oncotype DX takes advantage of this fact; it uses a 21 gene expression panel to stratify ER+ HER2- breast cancer patients into low, intermediate, or high-risk of recurrence after surgical resection. Schaafsma et al. showed that over its first decade of clinical use, Oncotype DX was associated with decreased adjuvant chemotherapy usage and increased survival.[56]

The possibilities provided by utilizing gene expression signatures with the aim to improve disease outcome have only expanded with the establishment of several cancer-associated mutation databases such as the seminal effort The Cancer Genome Atlas (TCGA), Database of Curated Mutations (DoCM), Clinical Interpretation of Variants in Cancer (CIViC), the National Cancer Institute (NCI) Genomic Data Commons (GDC), and the American Association for Cancer Research's (AACR) Genomics Evidence Neoplasia Information Exchange (GENIE). Commonly recurring mutations identified using these databases

give clues to the driving force behind expression profiles and therefore tumor behavior, as well as providing possible biomarkers for PFS or response to therapy, and potential targets for therapy. Some databases only contain sequencing information from pre-treatment primary tumors, such as the TCGA breast cancer dataset, which can be utilized to look for de novo causes of ETR. Some databases contain sequencing information from tumors that have been exposed to general chemotherapy or targeted therapy, which can lend information about both de novo and acquired ETR. With the abundance of tumor sequencing data that has resulted from the decreased cost of sequencing technologies, it is now well established that a patient's breast cancer can have different genomic profiles between tumor cells at primary site and metastatic site. To this end, sequencing data from longitudinal studies that include initial biopsy, biopsy after treatment, and/or biopsy of metastatic cancer provide clues about the genomic changes contributing to the behavior of breast tumors in response to therapy.[57,58]

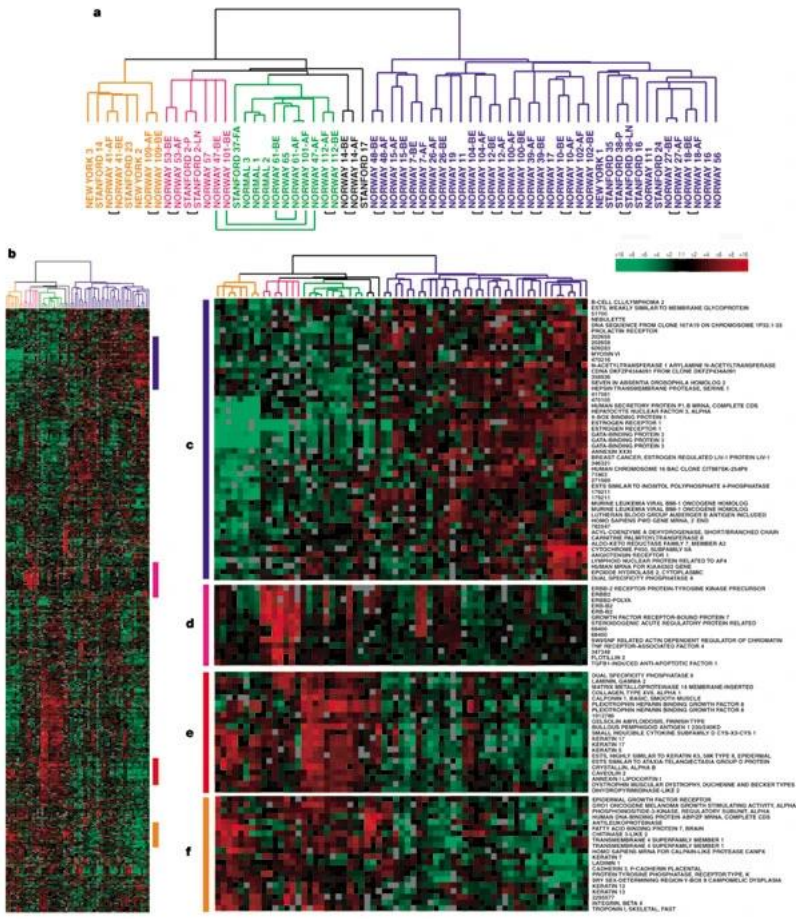


Figure 1-4. Four Molecular Subtypes of Breast Cancer. Dendrogram branch colors indicate subtype: basal-like, orange; HER2+, pink; normal-breast-like, light green; and luminal epithelial/ER+, dark blue. Panel **A** is a close-up of the breast tumor sample cluster dendrogram. Panel **B** shows the intrinsic cluster diagram, which is enlarged for the four subtype clusters in Panels **C-E**, which show the luminal epithelial/ER gene cluster, HER2+ gene cluster, a basal epithelial cell-associated cluster containing keratins 5 and 17 and a second basal epithelial-cell-enriched gene cluster, respectively. This figure is from Perou et al.[50]

<u>Known Causes of Endocrine Therapy Resistance</u>

Several different avenues for resistance to antiestrogen treatment have been identified in ER+ breast cancer. While loss of ERα is perhaps the most obvious explanation for endocrine therapy resistance, only about 10% of breast cancers with ETR exhibit loss of ER.[59] The targets of antiestrogen therapy, E2, ERα, and aromatase, have evaded the effects of the drugs in other ways as well. ESR1 acquires mutations in its ligand-binding domain (LBD), usually at Y357 and D538, in approximately 20% of recurring ER+ breast cancer cases.[60,61] Additionally the gene encoding aromatase, CYP19A1, is amplified in 21.5% of patients that relapse after being treated with AIs.[62]

Antiestrogen treatment can fail when ERα becomes reactivated independently of estrogen, as well. This reactivation can stem from altered interactions with its coactivators or corepressors. For example, MYC, CTCF, TBX3, and FOXA1 are mutated or amplified in many cases of ETR breast cancers.[63] FOXA1, which is amplified or overexpressed in tumors that have decreased PFS under tamoxifen treatment, reprograms its own cistrome to evade the effects antiestrogen therapy.[64,65] ERα can also engage in compensatory crosstalk with other oncogenic signaling pathways to become reactivated. These signaling pathways include, but are not limited to, EGFR, HER2, PI3K/mTOR, and RAS/RAF/MEK/ERK. EGFR is amplified in about 2% of

metastatic breast cancers with ETR.[63] FGFR is amplified in approximately 15% of metastatic breast cancers; its amplification promotes ETR by enabling estrogen-independent transcription of ER target genes.[66] Other routes to ETR involved with oncogenic signaling pathways include activating PIK3CA hotspot mutations acquired after fulvestrant treatment[67], HER2 amplification[68], or activating mutations of HER2 which are found in approximately 5% of ETR non-HER2 amplified metastatic breast cancers.[69,70] NF1, a negative regulator of RAS, can also promote ETR through loss-of-function alterations.[71,72(p1)]

The Search for Unknown Causes of Endocrine Therapy Resistance

ER+ breast cancer is heterogenous in behavior, and this heterogeneity is reflected in the transcriptional profile both pre- and post-therapy. ERα is a tentpole driver in breast cancer, and even in tumors resistant to antiestrogen therapy it remains important to tumor behavior. In fact, ERα expression is maintained in ~80% of tumors resistant to antiestrogen therapy.[73] ETR tumors still recruit ERα to chromatin, but the DNA regions it binds to are associated with poor clinical outcome. Furthermore, these new binding locations of ERα correlate with gene signatures predicting poor clinical outcome.[74] Enhancer-specific chromatin marks and chromatin openness shows differential patterns between breast cancer cells susceptible and resistant to antiestrogen therapy as well.[73] Sequencing data has also shown that certain recurrent mutations are correlated with clinical features of ER+ breast cancer that can be mapped back to molecular pathways involved in tumorigenic behavior. For instance, MAP3K1 is recurrently mutated in ER+ breast cancer that has luminal A molecular classification and low proliferation, but TP53 is recurrently mutated in ER+ breast cancer with luminal B molecular classification and higher proliferation.[75]

With the correlation between ERα binding profile, chromatin landscape, mutational profile, and their shared ties to response to endocrine therapy, the regulation of ERα transcription factor activity is undoubtedly consequential to breast cancer behavior. Even without this knowledge of the above correlations, the fact that typically only a few thousand of the hundreds of thousands of EREs in the breast cancer genome are bound by ERα demonstrates that its regulation is more dynamic and complex that relying on simple binding motifs. Understanding the mechanisms that shape the ERα transcriptome in different settings gets us closer to fully understanding the causes of endocrine therapy resistance.

Despite the abundance of explanations for endocrine therapy resistance listed above, only ~40% of cases of ETR are explained by known mechanisms.[63] To search for currently unidentified causes of antiestrogen therapy resistance in ER+ breast cancer, we can utilize the mutational profile of publicly available breast cancer cases that have been sequenced, as the underlying molecular contributors to cancers can often be found in mutational information.[76] Several ongoing sequencing projects already exist for exactly this purpose; among the most prolific of these are The Cancer Genome Atlas (TCGA), Genomics Evidence Neoplasia Information Exchange (GENIE), and Catalogue of Somatic Mutations In Cancer (COSMIC). Across these databases of sequencing information, there are many categories of genes that are recurrently mutated. The wealth of mutations in chromatin modifying-enzymes implicates alterations in the chromatin landscape in cancer. In fact, the 2013 TCGA Pan-Cancer Nature paper surveyed mutation profiles from 3,281 tumors across 12 tumor types and found that chromatin modifying enzymes made up 13 out of the 127 recurrently mutated genes.[77]

Chromatin modifiers demarcate enhancers and promoters to prime DNA for activation or repression of transcription. They come in three main flavors: writers that mark chromatin with post translational modifications (PTMs) such as methylation or acetylation, readers that identify and interpret the PTMs, and erasers that remove those PTMs. In addition, chromatin marks are associated with either transcriptional poising, activation, or repression depending on what histone residue they are

located on (H3K4, H3K27, H3K9, H4K20, H3K36, etc.), the other histone marks nearby, and whether the DNA element it's marking is a promoter or enhancer. The histone mark H3K4me1, for instance, marks enhancers as poised for transcription if no H3K27ac is present and active if it is.[78,79] Interestingly, ChIP-seq, RNA-seq, and whole exome sequencing have connected changes in epigenetic marks with mutations in chromatin-modifying proteins.[80] Some alterations in the epigenome, such as in H3K4 methylation levels, are associated with poor prognosis in breast cancer.[81] Changes in epigenetic marks such as histone methylation disrupt the function of enhancers, which are vital to the full activity of gene expression[82,83]. Accordingly, this disruption is involved in cancer development.[79,84]

Given the regulatory control that histone modifiers have in determining enhancer function, the frequency of mutations in histone modifiers in breast cancer, and the dysregulation of enhancers that is often seen in cancer, we decided to investigate what histone modifiers are recurrently mutated in ER+ breast cancer. Mutated in approximately 7-10% of breast cancer[75,77,85–88], MLL3 (mixed lineage leukemia 3), also known at KMT2C (lysine methyltransferase 2C), is one of the most frequently mutated histone modifiers in breast cancer. It serves as a major histone methyltransferase for H3K4 monomethylation along with MLL4/KMT2D.[89] The two histone modifier paralogues belong to a histone modifying complex called ASCOM or Activating Signal Cointegrator-2 -Containing Complex, which features either of the histone methyltransferases but not both. H3K4me1 is a chromatin mark that can help poise or activate enhancers for transcription. In addition, MLL3 aids in recruiting p300/CBP,[90,91] and leading KDM6A, another member of ASCOM, to remove H3K27me3 so p300/CBP can create H3K27ac for complete enhancer activation.[92] To successfully methylate histone H3K4, MLL3 must be bound to two of the other ASCOM members, ASH2L and RBBP5.[93]

Considering that ERα functions mainly at enhancers, MLL3 is an important effector of enhancer function, and MLL3 is recurrently mutated in ER+ breast cancer, we decided to focus our investigation on the role MLL3 may play in endocrine therapy resistance. While MLL3 is involved in the epigenetic activation of ERα transcription[94], it also binds to pioneer factor FOXA1 to cooperate in opening up chromatin conformation for ERα transcriptional control.[95(p3)] In addition, mutation of MLL3 leads to a shorter PFS in patients with ER+ breast cancer on antiestrogen therapy,[96] and has been identified as a driver of metastatic cancer.[97] These pieces of evidence show the importance of MLL3 function to the action of ERα.

While it is clear within the literature that MLL3 is involved in the regulation of ER+ breast cancer transcriptomics, it is not entirely known how MLL3 loss or mutation affects the binding profile and transcriptional output of ERα. Moreover, other sources of regulation of ERα transcriptional activity and thus mechanisms of endocrine therapy resistance are important to the discovery of biomarkers for treatment response and potential therapeutic targets. Looking again at sequencing information for breast cancers, FGFR1 amplification is seen in 10% of ER+/HER2- breast cancers. This molecular alteration is associated with early relapse following adjuvant tamoxifen therapy and with poor survival.[98] Treatment settings can also be a driving force behind resistance to combinatorial targeted therapies commonly used to treat ER+ breast cancer. Acquired resistance to fulvestrant and palbociclib presents a molecular setting which, once unraveled, will lead to better biomarkers and treatment plans for ER+/HER2- breast cancers. Even the mechanisms behind regulation of ERα degradation are connected to expression of estrogen target genes and therefore risk for developing breast cancer. Thus the aims of my thesis work are as follows.

Thesis Aims

The overall goal of my dissertation research was to further understand how ERα transcriptional activity is regulated in ER+ breast cancer, and to find biomarkers that could inform cancer prognosis, response to endocrine therapy, and potential targets for further therapy. This was accomplished by utilizing RNA-seq data from ER+ breast cancer cell lines and ER+ breast cancer patients, as well as ChIP-seq data from those cell lines, followed by bioinformatic analyses.

The majority of my work was focused on how MLL3 affects the transcriptional activity of ER+ breast cancer cells when lost or mutated. This part of my research had the specific aims to (a) identify the changes in regulation and output of ERα transcriptional activity upon loss of functional MLL3 and (b) identify the effect of loss of functional MLL3 on endocrine therapy resistance in breast cancer cells. For this research I used ChIP-seq and RNA-seq data generated from knock-down and control cell lines to identify shifts in ERα binding and gene expression caused by loss of MLL3 function. I also analyzed TCGA breast cancer RNA-seq expression data to identify genes that are differentially expressed under MLL3 mutation. Secondly, I utilized cell-based assays with knockdown and control cell lines to identify oncogenic properties and endocrine therapy resistance caused by loss of MLL3 function. I hypothesized that mutation or loss of MLL3 will shift both the enhancer and ERα genomic landscapes, and that this shift will affect ERα transcriptional response and biological behavior such as endocrine therapy resistance. The results of this study can inform future studies of ER biology, and of MLL3 biology in the context of breast cancer. Predictive information can also be gained through discovering the effect of MLL3 mutation on endocrine therapy response, and specific vulnerabilities of MLL3 mutant ER+ tumors identified during the project will lead to targeted therapies.

The other chapters included in this dissertation were completed as collaborations with peers also investigating the regulation of ERα in ER+ breast cancer. The aim of the third chapter was to investigate the transcriptional patterns associated with acquired resistance to combinatorial treatment of fulvestrant and palbociclib in ER+ breast cancer cells, a common therapeutic plan for many ER+/HER2- breast cancers. The aim of the fourth chapter was to determine the relationship between transcriptional regulation activity of FGFR1 and ERα in ER^{+}/*FGFR1*–amplified breast cancer cells in the context of endocrine therapy, given that FGFR1 amplification is present in 10% of ER+/HER2- breast cancers and is associated with poor clinical outcome. Lastly, the aim of the fifth chapter investigates the regulation of ERα homeostasis with respect to the balance between its degradation and transcriptional activity. The transcriptional divergences between non-clonogenic luminal (NCL) cells of mice to delineated the relationship between estrogen-responsive gene expression, estrogen abundance, and RSK2 status, an effector of ER homeostasis. The results of these studies provide further information about the regulation of ERα activity as well as mechanisms of resistance to endocrine therapies. Overall my thesis work contributes to the understanding of ERα genomic regulation in the context of ER+ breast cancer and endocrine therapy.

CHAPTER II

MLL3 IS A DE NOVO CAUSE OF ENDOCRINE THERAPY RESISTANCE

This section is a paper published in *Cancer Medicine* as "MLL3 is a de novo Cause of Endocrine Therapy Resistance" Kim Stauffer*, David Elion, Rebecca Cook, and Thomas Stricker.

Summary

I initially identified MLL3 as a recurrently mutated gene of interest in breast cancer upon reading the 2013 *Nature* paper "Mutational landscape and significance across 12 major cancer types". Upon further research of literature on MLL3 and its complex ASCOM, I hypothesized that the mutation of MLL3 would alter the H3K4me1 landscape of the breast cancer genome. I considered the significance of estrogen receptor alpha (ERα) binding profile on the behavior of ER+ breast cancer and hypothesized that the altered H3K4me1 landscape could affect ERα binding. I believed this would in turn alter the tumorigenic tendencies of the cancer. This hypothesis formed the basis of my thesis project, and eventually became my first author paper.

For this project I chose to utilize an ER+, MLL3 wildtype breast cancer cell line, ZR751. After lentiviral knockdown (KD) of MLL3 in these cells, I worked on proliferation assays to compare the response of MLL3-KD and WT cells to two common endocrine therapies, fulvestrant and tamoxifen. With the help of David Elion, I was able to show endocrine therapy resistance (ETR) in the MLL3-KD cells. From this point, I performed RNA-seq and ChIP-seq to interrogate the accuracy of my hypothesis. Coupled with RNA-seq data from TCGA ER+ breast cancer cases, downstream analyses focused on differential gene expression, changes in the enhancer landscape as defined by H3K4me1, changes in the ERα binding profile, and the intersection of the three.

This study identified MLL3 mutation as a cause of *de novo* endocrine therapy resistance in ER+ breast cancer. Although MLL3 is only mutated in ~10% of ER+ breast cancers, it is likely that this accounts for a large portion of breast cancers with unexplained causes of endocrine therapy resistance. The full manuscript, of which I am first author, is reproduced below.

Introduction

Breast cancer is the second most commonly diagnosed cancer in American women and 75% of cases are estrogen-receptor positive (ER+). Anti-estrogens are the first line of therapy, however 80% of women present with (*de novo*) or develop (acquired) endocrine therapy resistance.[99] Disease recurrence and drug resistance are major drivers of mortality in ER+ breast cancer. While some causes of endocrine therapy resistance, such as *ESR1* mutation, *HER2* amplification, and *FGFR1/CCND1* amplifications are known,[100-101] ~60% of cases do not have an identified mechanism.[63] Furthermore, only 50-70% of ER+ patients respond to initial endocrine therapy, highlighting a need for *de novo* resistance biomarkers. Improved understanding of the mechanisms of endocrine resistance will guide therapeutic development.

ChIP-Seq studies show tumors that respond poorly to endocrine therapy have a unique set of ERα genomic binding locations.[74] Furthermore it has been shown that ER+ breast cancer can adapt to estrogen deprivation through epigenetic reprogramming at enhancers.[73] These patterns therefore suggest that genes regulating ERα binding may affect/alter endocrine therapy

responsiveness. One such gene that has been shown to regulate nuclear receptor activity[102] is *MLL3*, the 6th most frequently mutated gene in ER+ breast cancer.[103] MLL3 primarily monomethylates H3K4 to mark enhancers. Interestingly, ERα binding sites regulate gene transcription largely from enhancers. In MCF7 cells the pioneer factor FOXA1 has been shown to recruit MLL3 to demarcate enhancers for ERα.[95] Further implicating the monomethyltransferase as an important regulator of ERα binding, MLL3 possesses LXXLL domains known to interact with nuclear hormone receptors such as ERα.[104]

Recurrent *MLL3* mutation was first identified in acute myeloid leukemia (AML), where it was determined to be a haploinsufficient tumor suppressor.[105] Similarly, *MLL3* is recurrently mutated in ER+ breast cancer.[106,103] These mutations are predicted to be functional and therefore drivers.[107,108] Not only is *MLL3* recurrently mutated, its mutation is also associated with more aggressive disease characteristics both *in vitro*[109,110] and *in vivo.*[111-96]

Given the above observations, we predicted that mutation of *MLL3* will shift both the enhancer and ERα genomic landscape, and that this shift will affect transcriptional control by ERα and biological behavior such as endocrine resistance.

Methods and Materials

GERP analysis

Hg 19 base-wise GERP scores were downloaded from http://mendel.stanford.edu/SidowLab/downloads/gerp/.[112] To find average GERP scores for the missense mutations in each gene we used 595 TCGA ER+ luminal breast cancer cases and found the GERP score for the location of each missense mutation for the following genes: *MLL2*, *PIK3CA*, *PTEN*, and *TTN*. We calculated GERP averages for each set of missense mutations. We then selected a corresponding number of GERP scores from the entire coding sequence that would potentially lead to missense variants of each gene at random and calculated the average of those GERP scores. We repeated the random selections and average calculation 10,000 times. To get a value of significance, we divided the number of times a random selection GERP average was greater the actual mean GERP score of our gene of interest by 10,000. Values less than 0.05 were considered significant.

Cell culture and antibodies

ZR751 cells (RRID CVCL_0588) were obtained from the Lannigan laboratory[113] and grown in RPMI (Sigma Aldrich #R8758500ml) supplemented with 10% heat-inactivated FBS (Corning™ #35016CV), 0.002% insulin (Sigma Aldrich #11376497001) and 50 IU penicillin, 50 mg/mL streptomycin (Corning™ #MT30001CI). HEK 293T cells (RRID CVCL_0063) were obtained from the Lannigan laboratory[113] lab and grown in DMEM with high glucose, L-glutamine, phenol red, but not sodium pyruvate (Sigma Aldrich D0819-500ML), 5% FBS, 1% Pen/Strep, and 1% Sodium pyruvate (Sigma Aldrich S8636-100ML). The cell culture incubator parameters were as follows: 37˚C, 95% relative humidity, and 5% CO_2 concentration. The antibodies used for ChIP-seq were anti-Erα (Santa Cruz Biotechnology sc-543X), anti-H3K4me1 (Abcam ab8895), anti-SP1 (Abcam ab13370), and sheep anti-rabbit IgG Dynabeads M-280 (Invitrogen™ 11203D).

Lentivirus-mediated RNA-interference (RNAi)

Oligos to use for shRNA were designed and ordered from Sigma/Genosys at the Molecular Cell Biology Core at Vanderbilt. The oligos were annealed, phosphorylated, and ligated into pSuper for transformation into DH5α cells. QIAprep Spin Miniprep Kit (Qiagen 27104) was used to isolate the vector, which was transfected into ZR751 cells and assessed by qPCR for KD. KDs that

worked were then isolated with QIAprep Spin Miniprep Kit (Qiagen 27104), digested, and ligated into pLVTH[114] (Addgene 12262) for transformation into STBL3 cells. A QIAGEN Plasmid Plus Maxi Kit (Qiagen 12963) isolated the pLVTH for transfection into HEK 293T cells, from which lentivirus was collected. The oligo sequence used to silence MLL3 was 5° - CCGGCGCACCTTATAGTAAACAGTTCTCGAGAACTGTTTACTATAAGGTGCGTTTTT -3°, taken from The RNAi Consortium.[115] Negative control Luciferase shRNA Control was donated by the Lannigan laboratory.[116] Cells were stably transduced at 100,000 cells per well in a 6-well plate (Corning 3516) with 4 µl lentivirus, and subsequently flow sorted for GFP expression and propidium iodide (Sigma Aldrich P4864) staining after 3 days. qPCR was performed in biological triplicate to check shRNA KD 3 days after transduction. Experiments were performed in multiple, but early (<=10) passages of the stably transduced cell lines.

RNA-Seq

Cells were harvested at steady-state using the RNAEasy Kit (Qiagen 74104). RNA samples of 600 ng were subjected to Turbo DNAse (Thermo Scientific #AM2238) and Superscript III RT (ThermoFisher 18080093) with Random Hexamers (ThermoFisher N8080127) and dNTPs (ThermoFisher 18427088). qPCR was performed with 2 µl cDNA, 0.5 µl of 10 mM forward and reverse primers each, 10 µl SYBR Green (ThermoFisher 4364346), and 7 µl water in the Molecular Cell Biology Resource Core at Vanderbilt (BioRad CFX96 Touch Real-Time PCR Detection System). An initial denaturation and enzyme activation step of 95°C for 3 minutes was performed, followed by 40 cycles of 95°C for 10 seconds to denature and 55°C for 30 seconds to anneal, and finally a melt curve. Reactions were performed in biological triplicate using SYBER green PCR Master Mix (Thermo Scientific #4344463), and results were analyzed using the delta-delta Ct method. The average of the three biological replicate Ct values for the reference GAPDH gene was subtracted from the 3 individual biological replicate Ct values for the target MLL3 gene. A t-test was performed on the resultant two groups of delta Ct values to give a p-value of 0.0193. The Ct values ranged from 11.77 to 25.06. The qPCR was performed three times to obtain a working assay. The primers were ordered from the DNA Core at Vanderbilt from Sigma Genosys as follows: MLL3 forward, AACTCACGACCACCATCTCC, MLL3 reverse, TCTGGAGGTTTTGCATAGGG, GAPDH (control) forward, GTGAAGGTCGGAGTCAACGAPDH (control) reverse, CCCATACGACTGCAAAGACC. RNA quality was assessed in VANTAGE via Invitrogen Qubit and Agilent BioAnalyzer and samples with RIN >7 were used. RNA libraries were generated with two biological replicates of 2 µg RNA using Illumina's TruSeq Stranded Total RNA Sample Prep Kit (20020597). Libraries were sequenced at VANTAGE with PE75 to a depth of approximately 30 million reads per sample on an Illumina HiSeq3000 (**Table 2-10**)**.** Quality of NGS data was assessed using FastQC, and adapters/low quality bases were trimmed from reads using fastq-mcf from ea-utils, with minimum quality of 7 and minimum length of 25. Fastq files from 595 breast invasive carcinoma samples in TCGA were downloaded from the Cancer Genomics Hub (https://browser.cghub.ucsc.edu/). Tumor classification data was obtained from the TCGA Data Portal (https://tcga-data.nci.nih.gov/tcga/). RNA-seq reads, both in-house and from the TCGA, were aligned to the human genome (hg19) with Tophat (v2.0.13), quantified using cufflinks (v2.2.1) and normalized using cuffnorm (v2.2.1).[117]

Differential Expression Analysis

For ZR751 RNA-seq, differential expression analysis was performed in Rstudio v3.6.1 using a gene-by-gene linear regression model with ANOVA taking MLL3 knockdown status into account. Genes with a mean expression level of log2(fpkm+0.5) greater than 1 were kept for the analysis. A log2(fpkm+0.5) transformation was used on the gene expression table. The sva (surrogate variable analysis) package in Bioconductor was utilized to remove batch effects.[118] DEG were identified as those with an ANOVA FDR q-value less than 0.05; q-values were calculated using the qvalue package in R.

For TCGA RNA-seq, we limited our search to breast cancer cases that were marked as ER+ in the clinical file. To decrease the variance in the control ER transcriptional activity profile, we also limited the breast cancer cases that were marked as molecular subtypes luminal A and luminal B in the clinical file. Samples that did not have information in the clinical file were discarded. Samples with an internal size factor of less than 0.35 were discarded from the analysis. Samples from men were excluded. Genes with a mean expression level of log2(fpkm+0.5) greater than 1.5 were kept for the analysis. A transformation of log2(fpkm+0.5) was performed on the gene expression set. The sva (surrogate variable analysis) package in Bioconductor was utilized to remove batch effects.[118] A gene-by-gene linear regression model with multivariate ANOVA accounting for histological subtype, molecular subtype, and MLL3 mutation status was utilized to find differential gene expression. DEG were identified as those with an ANOVA FDR q-value for the MLL3-mutation status variable less than 0.05; q-values were calculated using the qvalue package in R.

ChIP-seq

ChIPs were performed for two biological replicates, for one experimental repetition. Cells were grown to 80% confluency, washed 3 times in ice-cold PBS (8 g NaCl, 0.2 g KCl, 1.44 g Na_2HP0_4, 0.24 g KH_2PO_4, H_2O up to 1 L, adjusted to pH 7.4 with HCl) and then fixed for 10 minutes at room temperature using 1.85% formaldehyde (50 ml cold PBS, 2.5 ml 37% formaldehyde solution Sigma Aldrich 252549), followed by quenching with 2.5 ml of 2.5 M glycine (93.8 g glycine Sigma Aldrich G7126 in 500 ml H_2O) for two minutes at room temperature. After aspirating and washing with 50 ml cold PBS, we lysed the cells using 20 ml Farnham lysis buffer (5 mM HEPES pH 8, 85 mM KCl, 0.5% NP-40) and 400 µl protease inhibitor cocktail (PIC, Roche 11873580001) to scrape the cells off (Corning™ 3008) into a 50 ml conical tube (Corning 352098). These tubes were spun down at 425 g for five minutes at 4° Celsius.

Nuclei lysis buffer (50 mM Tris-HCl pH 8, 10 mM EDTA pH 8, 1% SDS), 1X PIC, and 10 mM sodium butyrate (Sigma Aldrich B5887) were added to a concentration of 20,000,000 cells per 400 µl and resuspended until homogenous. Chromatin was sonicated using a Covaris LE220 for 35 minutes, then centrifuged at max speed for 10 minutes at 4°C to obtain supernatant. Per 0.1 ml of supernatant, we diluted with 0.9 ml ChIP dilution buffer (50 mM Tris-HCl pH 8, 0.167 M NaCl, 1.1% Triton X-100, 0.11% sodium deoxycholate), 0.5 ml RIPA-150 (50 mM Tris-HCl pH 8, 0.15 M NaCl, 1 mM EDTA pH 8, 0.1% SDS, 1% Triton X-100, 0.1% sodium deoxycholate), 28 µl 50X PIC, and 14 µl 1 M sodium butyrate.

Anti-ERα (3 µl/IP), anti-H3K4me1 (1 µl/IP), and anti-SP1 (3 µl/IP) were linked to 100 µl/IP, 60 µl/IP, and 100 µl/IP magnetic anti-rabbit Dynabeads respectively with RIPA-150 to a final volume of 500 µl for 6 hours at 4°C in low-bind tubes (Eppendorf Z666505), and then incubated with 150 µg of chromatin overnight at 4°C. Immunoprecipitants were washed with RIPA-150 once, followed by RIPA-500 (50 mM Tris-HCl pH 8, 0.5 M NaCl, 1 mM EDTA pH 8, 0.1% SDS, 1% Triton X-100, 0.1% sodium deoxycholate) twice, then RIPA-LiCl (50 mM Tris-HCl pH 8, 1 mM EDTA pH8, 1% Nonidet P-40, 0.7% sodium deoxycholate, 0.5 M $LiCl_2$) twice, and finally 1X TE Buffer pH 8 (10 mM Tris-HCl pH 8, 1 mM EDTA pH 8) twice for 5 minutes each. Chromatin-IPs were eluted from the beads in 200 µl freshly made Direct Elution Buffer (10mM Tris-HCl pH 8, 0.3 M NaCl, 5 mM EDTA pH 8, 0.5% SDS), and then treated with 1 µl of 1 mg/ml RNase A (Fisher Scientific FEREN0531) at 65°C with shaking for 4 hours. This was followed by 3 µl proteinase-K (Sigma-Aldrich 3115879001) overnight at 55°C to reverse crosslinks. DNA was purified using phenol–chloroform extraction. Samples were transferred to a spun-down 2 ml phase lock gel tube (Qiagen 129056) and an equal volume of phenol/chloroform/isoamyl alcohol (Sigma Aldrich P3803100ML) was added and vortexed. This was spun at room temperature for 5 minutes at 14,000 g, and the sample was moved to a new 1.5 ml tube. One tenth

volume sodium acetate (Invitrogen AM9740), 1 µl glycogen (Roche 10901393001), twice volume 100% ethanol (Sigma Aldrich E7023500ML) was added, and the samples were incubated at -80°C for 30 minutes. The sample was spun at 20000 g for 30 minutes at 4°C, and the supernatant was carefully aspirated. The pellet was washed with 1 ml cold 70% ethanol, and spun at 20000 g for 30 minutes at 4°C. The supernatant was aspirated, and the spin was repeated a final time. The supernatant was removed, and pellet was allowed to dry. The pellet was then resuspended in 25 µl elution buffer (Qiagen 19086) and subsequently quantified by Qubit 2.0 Fluorometer.

Standard Illumina ChIP-seq Library Kits (IP-202-1012, IP-202-1024) were used to build sequencing libraries for two biological replicates per condition for one experimental repetition, with inputs used as control. Libraries were sequenced at VANTAGE using an SR50 flow cell on the Illumina HiSeq3000 to a depth of approximately 20 million reads (**Table 2-10**)**.** Quality of NGS data was assessed using FastQC v0.11.5, and adapters/low quality bases were trimmed from reads using fastq-mcf from ea-utils, with minimum quality of 7 and minimum length of 25. The fastq files were aligned to human genome version 19 by BWA (Burrows–Wheeler aligner Version 0.7.5a-r405).[119] Post-alignment filtering was performed with Samtools 1.7[120] and Picard 1.126 MarkDuplicates. PhantomPeakQualTools v1.2.1[121] was used to assess ChIP-seq enrichment quality prior to inclusion in the study, and all replicates used in this study passed. Self-pseudoreplicates, pooled data, and pooled-pseudoreplicates were generated and used to call peaks for creation of peak thresholds. Peaks were called against matching input using SPP v1.15.5 according to best practices ENCODE 3 Pipeline v1. SPP uses a normalization factor is implicitly used to linearly scale the control sample for comparison with the ChIP sample; it does this by identifying a subset of background bins with a tag count exceeding Poisson density ($p < 0.0001$). Those background regions can then be normalized to the input channel. The Irreproducible Discovery Rate (IDR) framework version 2.0.3 was used to measure the reproducibility of ChIP-seq peaks identified from replicate experiments and find thresholds based on reproducibility.[123] All call sets used for this study met IDR benchmarks for reproducibility (**2-5a, 2-11a, Table 2-11**). Final peak thresholds were chosen from this structured comparison of number of peaks called from original replicates, self-pseudoreplicates, and pooled-pseudoreplicates; these peak thresholds were applied to a pooled reads file composed of the two biological ChIP replicate libraries. The DiffBind package in R was utilized to find differential binding of ZR751shLucif vs ZR751shMLL3 H3K4me1, ERα, and SP1 ChIP-seq peaks (**2b**, **S6c**).

Peak Assignment

Using Bedtools v2.26.0 we assigned each ChIP-seq peak to the two closest DEGs rather than the closest gene in the human genome.[124] We then removed all assignments that had a peak-to-gene distance greater than 1 million base pairs (bp), ranging from 16% to 26% of assignments, because most chromatin-chromatin interactions span 1 million bp or less.[125]

To determine whether our ChIP-seq peaks are closer to our DEG than we would expect by chance, we randomly selected a matched number (6,677 to equal the number of differentially expressed genes) of genes from the reference genome file to assign to our peaks, calculated distances, and then repeated this process 1000 times. A one-sided Kolmogorov–Smirnov test between our DEG-peak assignments and randomly chosen gene set-peak assignments was performed for each of the 1000 repetitions, and then created a final measure of robustness by subtracting the number of p-values less than 0.05 divided by 1000 from 1. Peak categories with a final measure of robustness less than 0.05 were kept.

Bioinformatic Tools

Mutation information, survival plots, and TCGA for breast cancer samples were acquired from the National Cancer Institute Genomic Data Commons Data Portal. GRMetrics R package usage included GRfit by cell line and time point to calculate GR values. For IDR plots, peak files and an hg19 genome file was loaded into R. Parameters included half.width = NULL, overlap.ratio = 0, is.broadpeak = F, sig.value = "signal.value". Data was processed and IDR output generated with process.narrowpeak, compute.pair.uri, and fit.em with fix.rho2=T as a parameter. NGS Plot heatmaps and histograms were created at command line using ngs.plot.r with hg19 genome, with final bed files as region to plot, configuration files to plot both control and KD bam files, length from gene body of 3000 bp, ensemble as the gene database, and chipseq and protein_coding as the annotations to use. Diffbind in R utilized the DBA_EDGER analysis method with a reporting threshold of 0.1 and bUsePval = TRUE. The DBA__BLACKLIST_HG19 blacklist was applied, and a greylist.pval of 0.9 was applied afterwards. A consensus peakset with a minOverlap of 0.66 and consensus of DBA_CONDITION was created and used to count reads in dba.count. These reads were normalized with dba.normalize and method = DBA_ALL_METHODS, and then contrasted with dba.contrast by condition and minMembers = 2. Analysis of differential peak enrichment was carried out using dba.analyze using DBA_ALL_METHODS. GREAT webtool version 3.0.0 was used to identify gene set enrichment analysis with ChIP-seq data[126] with human genome UCSC hg19 for species assembly, whole genome as background, and basal plus extension with 5.0 kb upstream, 1.0 kb downstream, and distal up to 1000 kb for associating genomic regions with genes. Curated regulatory domains were included. WebGestalt 2019 version was utilized for gene set enrichment analysis with RNA-seq and ChIP-seq data.[127] RNA-seq data was submitted to WebGestalt Gene Set Enrichment Analysis (GSEA) as rank (rnk) files sorted by -log10(p-value) from the differential expression analysis in R, and the Molecular Signatures Database (MSigDB) curated gene sets of chemical and genetic perturbations (C2 CGP) database as the functional database to survey. All genes expressed in the specific dataset (ZR751 or TCGA) were used as the reference set. The minimum number of genes for a category was set at 3, and the maximum was set at 2000. P-values from this analysis were adjusted for multiple hypothesis testing using Benjamin-Hochberg method, and the top 50 most significant terms by FDR were retrieved. Gene groups from the integration of RNA-seq and ChIP-seq data were submitted to WebGestalt using an Over-Representation Analysis (ORA) using all the same parameters except for use of protein-coding portion of the human genome as the background. The iRegulon tool v1.3 (build 2015-02-12) in Cytoscape software version 3.7.1 was utilized to identify enriched transcription factor motifs in DEG from RNA-seq data[128] with the "Predict regulators and targets" option. The species and gene nomenclature chosen was Homo sapiens, HGNC symbols, the type of search space was gene-based, the motif collection was 10k (9713 PWMs), the track collection was ENCODE raw signals, the putative regulatory region was 20kb centered around TSS, and the motif rankings database was 7 species. The Enrichment score threshold was 3.0, the ROC threshold for AUC calculation was 0.03, and the rank threshold was 5000. The minimum identity between orthologous genes for TF prediction was 0, and the maximum FDR on motif similarity was 0.001. MEME-suite command-line tools version 4.11.2 was used to identify enriched transcription factor motifs in ChIP-seq data.[129] Fasta files were used with MEME command and max dataset size of 5000000 letters, using the DNA alphabet, and a max number of motifs at three. Tomtom was utilized with the HOCOMOCOv11_full_HUMAN_mono_meme_format.meme database to identify known motifs within the MEME results. Dependence scores for ER+ breast cancer cell lines were acquired from the DEMETER dependence tool online at the Dependency Map (DepMap) Portal, https://depmap.org/portal/.[41,130] IGV version 2.9.4 was utilized to visualize RNA-seq and ChIP-seq data in the form of bigwig files, hosted at data.cyverse.org.[131] Bigwig files were generated using command line bamCoverage program from deepTools version 3.3.1-Python-3.7.2 on merged bam files with the parameters bin size of 100, smoothing length of 250, normalizing using RPKM, and effective genome size using hg19.

Proliferation Assays

Cells were plated in 96-well plates (Fisher Scientific 07-200-95) with 10,000 cells per well and three biological replicates per experiment in phenol-red free RPMI (Sigma-Aldrich R8758500ml) with 10% heat-inactivated charcoal-stripped FBS (Corning™ 35016CV), 10 nM β-estradiol (Sigma-Aldrich E8875-5G), 0.002% insulin (Sigma-Aldrich 11376497001), and 50 U/mL penicillin, 50 mg/mL streptomycin (Corning™ MT30001CI), and either DMSO (Sigma-Aldrich D8418-100ML), Tamoxifen (Sigma-Aldrich 579002-5MG), or Fulvestrant (Sigma-Aldrich I4409-25MG). Media was switched out every four days and plates were fixed on days 4 and 8. All plates were stained with crystal violet (Sigma-Aldrich C0775-25G) and quantification by spectrophotometric detection at 490 nm using plate reader Molecular Devices Spectramax M3. Ten experimental replicates were performed to obtain parameters (cells per well, estradiol amount, time points) that gave consistent results. Effects were analyzed using GRmetrics version 1.10.0, one-sided Wilcoxon Rank Sum Test, n=3.

Statistical Analyses

All significance level thresholds are $p<0.05$ unless otherwise noted. For all bar-and-whisker plots, the center line signifies the median, box limits signify upper and lower quartiles, and whiskers signify the 1.5x interquartile range. All data points are shown as dots. For histograms and line plots, error bars represent standard deviation. Significance of survival curves (1H, S1E) were evaluated by Log-Rank test. Quantification of gene expression (qPCR, S2A) was evaluated by a one-tailed unpaired t-test of the calculated delta CT values. For differential expression analyses RNA-seq FPKM files were log2 transformed. The R SVA package[132] was utilized to estimate artifacts in the form of surrogate variables from the RNA-seq data, which were then removed from the data. The cleaned data was then analyzed with a gene-by-gene multivariate linear regression model accounting for KD status for ZR751 data and histological subtype, intrinsic molecular subtype, and binary MLL3 mutation status for TCGA data. An ANOVA was used to evaluate the model. Estimated log expression change and Pr(>|t|) for MLL3 mutation or KD status from the linear regression and Pr(>F) for MLL3 mutation or KD status from the ANOVA were recorded for each expressed gene. Multiple hypotheses correction was achieved through use of the qvalue R package on the ANOVA p-values[133]. Overlap between groups of genes was tested with the GeneOverlap R package[134] which employs the Fisher's exact test. For proliferation assays the R package GRMetrics was utilized to find GR values, which are the growth-rate inhibition value of a given treatment at a given concentration. The GR values were then assessed by Wilcoxon Rank Sum Exact test, for each concentration and time point. The SP1 Dependency scores were assessed for effect by MLL3 mutation using a Wilcoxon Rank Sum exact test. The number of peaks assigned to DEG was assessed for patterns of loss or gain using both a proportions test where gain of peaks assigned to DEG in the KD condition =1 and a loss of peaks = 0, as well as a two-sided Wilcoxon paired signed rank test with continuity correction.

Data Availability Statement

The TCGA data that support the findings of this study are openly available in the Genomic Data Commons at https://portal.gdc.cancer.gov/. The ZR751 RNA-seq and ChIP-seq data that support the findings of this study are available at https://www.ncbi.nlm.nih.gov/geo under series GSE163264 . For codes, see online at https://github.com/staufferalexander/MLL3.

Results

MLL3 mutation pattern in ER+ breast cancer suggests that MLL3 is a haploinsufficient tumor suppressor.

MLL3 has been reported to be a haploinsufficient tumor suppressor in AML[105], and thus we hypothesized that most *MLL3* mutations in breast cancer would be heterozygous (**2-1a**).[135,136] We expect a 1:1 mutant-to-wildtype allele ratio in the TCGA ER+ breast cancer sample set to present as a 35:65 mutant-to-wildtype allele ratio for a few reasons: TCGA ER+ breast cancer samples have approximately 75% tumor purity,[137] and copy number data from the TCGA demonstrate that no amplifications or deletions coincide with MLL3 mutations for these samples (**2-2d**). Analysis of TCGA data demonstrates that the average *MLL3* mutant allele frequency, corresponding to the percent of sequencing reads containing a mutation, is approximately 30% across the different categories of mutation (**2-1b**). This suggests that only one of two alleles is mutated, and that heterozygosity is not lost upon mutation of that one allele. This trend persists across multiple breast cancer datasets (**2-2b**), and in some of the other most frequently mutated genes in ER+ breast cancer (**2-2a**). Indeed, *MAP2K4* and *TP53*, tumor suppressors associated with loss of heterozygosity, [138,139] have a higher mutant allele fraction of approximately 50-60%. These ratios are more consistent with mutation of one allele, followed by loss of heterozygosity of the other allele in the tumor cells, given the aforementioned tumor purity.

With evidence to support that *MLL3* mutations in ER+ luminal breast cancer are heterozygous, we next considered whether the effect of the mutations would be deleterious to the function of the methyltransferase. Mutations were a mix of nonsense (16/49), frameshift (18/49,), missense (14/49), and splice (1/49) mutations spread across the length of the gene with no mutational hotspots (**2-1c, 2-2c**). **Table 2-1** shows that while there are no mutations within the catalytic SET domain of MLL3, there are 34 truncating mutations that occur 5' to the SET domain. In addition, missense mutations within the PHD domains of *MLL3* have been shown to be oncogenic.[109] Considering this information, we speculated that the 10 missense mutations outside defined regions of the protein would still lead to deleterious effects on MLL3 function.

Domain Name	Function	Amino Acids	# Truncating Mutations In/Prior To	TCGA Mutations
PHD1	Putative H3/Zn binding	247-330	2	NA
PHD2	Putative H3/Zn binding	390-435	3	1ns, 2ms
PHD3	Putative H3/Zn binding	466-517	4	1ns
PHD4	Binds to H4R3me0, H4R3me2a	952-1008	10	-
PHD5	Binds to H4R3me0, H4R3me2a	1009-1055	10	-
PHD6	Binds to H4R3me0, H4R3me2a	1086-1136	10	-
LXXLL Motif	Nuclear Receptor Interacting	1408-1412	12	-
HMG-1	DNA Binding	1655-1703	15	1fs
LXXLL Motif	Nuclear Receptor Interacting	2745-2749	23	-
LXXLL Motif	Nuclear Receptor Interacting	2918-2922	23	-
LXXLL Motif	Nuclear Receptor Interacting	3055-3059	23	-
LXXLL Motif	Nuclear Receptor Interacting	3777-3781	26	-
PHD7	Putative H3/Zn binding	4402-4506	32	1ns, 1ms
FYRN	Unknown	4550-4604	33	1fs
FYRC	Unknown	4606-4691	33	-
SET	Catalytic Domain, Methylates H3K4	4772-4893	34	-

Table 2-1: Domains of MLL3 and TCGA ER+ Luminal Breast Cancer Mutations. The number of truncating mutations occurring within or prior to each domain is listed in the 4th column.

To interrogate the effect of missense mutations in the ER+ luminal TCGA cases we performed an analysis using GERP scores, an evolutionary calculation of nucleotide constraint. Genomic positions with higher scores are thought to be more deleterious if altered.[112,140] We hypothesized that the GERP scores for mutations observed in MLL3 in breast cancer would be higher, i.e. more deleterious, than randomly selected missense variants, indicating that the residues mutated in TCGA samples are more conserved, and thus mutation of these conserved residues will likely be detrimental to protein function. For positive controls, we chose *PIK3CA* as an oncogene with hotspot mutations, and *PTEN* as a tumor suppressor with mutations throughout the gene.[107] For a negative control we chose *TTN*, a known false-positive in cancer resequencing studies. In *PIK3CA* and *PTEN* the average GERP score of missense mutations for each gene were significantly higher, and therefore more deleterious, than the simulated GERP score averages (*PIK3CA* $p < 0.0001$, *PTEN* $p=0.007$) (**2-1d, e**). In *TTN* the average GERP score was within the middle of the distribution of simulated GERP averages ($p = 0.7522$) (**2-1f**). The average GERP score of missense *MLL3* mutations was on the tail of the distribution of simulated GERP averages, very similar to that of *PTEN* ($p=0.0004$) (**2-1g**). This analysis suggests that missense mutations in *MLL3* in ER+ luminal breast cancers are deleterious to the function of the protein. Of note, a similar analysis, using the ratio of nonsynonymous to synonymous mutations in cancer also found that *MLL3* is enriched for missense mutations with evidence of selection.[107]

A Kaplan-Meier plot of TCGA breast cancer patients demonstrated that untreated ER+ breast cancer patients with *MLL3*-mutant breast tumors have a significantly poorer overall survival than those with *MLL3*-wildtype tumors (**2-1h**), suggesting that loss of MLL3 function contributes to poor outcome in breast cancer patients. This trend remains true when comparing patients with *MLL3* missense mutations to patients with *MLL3*-wildtype tumors (**2-2e**). The analyses above, along with the lack of hotspots and the number of loss-of-function mutations, illustrates that MLL3 is a haploinsufficient tumor suppressor in ER+ breast cancer. Thus, we decided to model *MLL3* mutation with lentiviral shRNA knockdown (KD) in the ER+ breast cancer cell line ZR751 in order to maintain some residual expression of wildtype *MLL3* (**2-3a**).

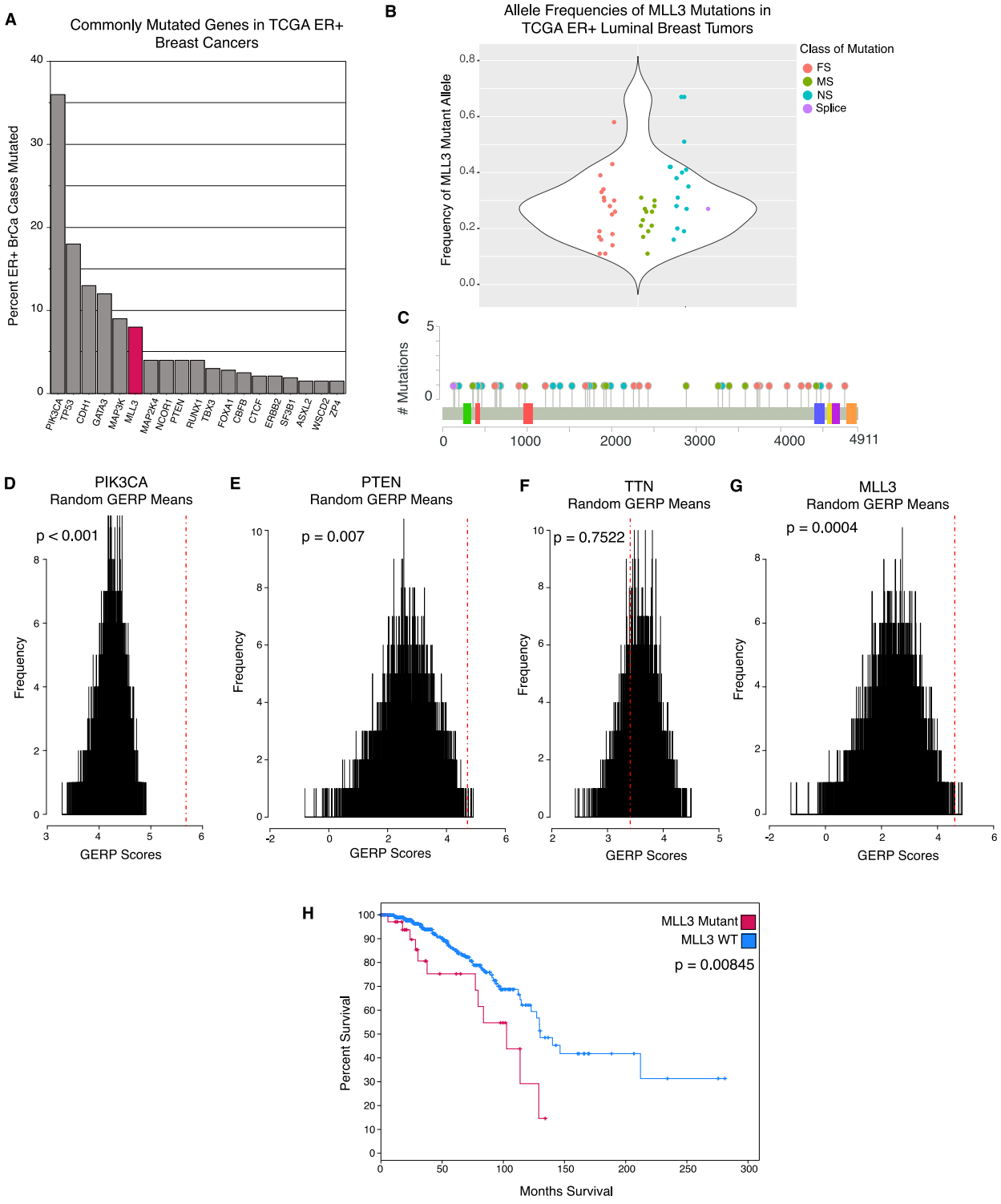


Figure 2-1. MLL3 is significantly mutated in ER+ breast cancer; its mutation confers poor outcome. (**a**) The most commonly mutated genes in the provisional TCGA ER+ breast cancer RNA-seq dataset (n = 581) BrCa = breast cancer. ER+ = estrogen receptor positive. (**b**) Frequency of mutant MLL3 allele in TCGA ER+ luminal breast cancer cases (n=581). FS = frameshift. MS = missense. NS = nonsense. (**c**) MLL3 mutation lollipop plot of luminal TCGA breast cancer cases with RNA-seq data (n= 46 mutations). Red lollipops indicate frameshift mutations, green indicate missense mutations, blue indicate nonsense mutations, and purple indicate splice mutations. Colored boxes indicate specialty domains as follows: PHD-like zinc-binding (green), PHD finger (red), F/Y-rich N-terminus (blue), F/Y-rich C-terminus (yellow), catalytic SET domain (purple). (**d**) Histograms of (#) simulations of averages of randomly-chosen GERP scores in PIK3CA (**e**) PTEN (**f**) TTN and (**g**) MLL3. The number of randomly-chosen GERP scores matches the number of mutations in each respective gene in the TCGA luminal breast cancer cases (n= 581). Simulated averages are shown by black lines, the actual average GERP score is shown by the red dotted line. P-values are calculated by dividing number of simulated averages higher than the actual average GERP score by the total number of simulated averages. (**h**) Survival curve showing luminal cases from TCGA breast cancer cohort (n= 581) that are either mutant (red) or wildtype (blue) for MLL3. Log-rank Test p-value = 0.00845. WT = wildtype.

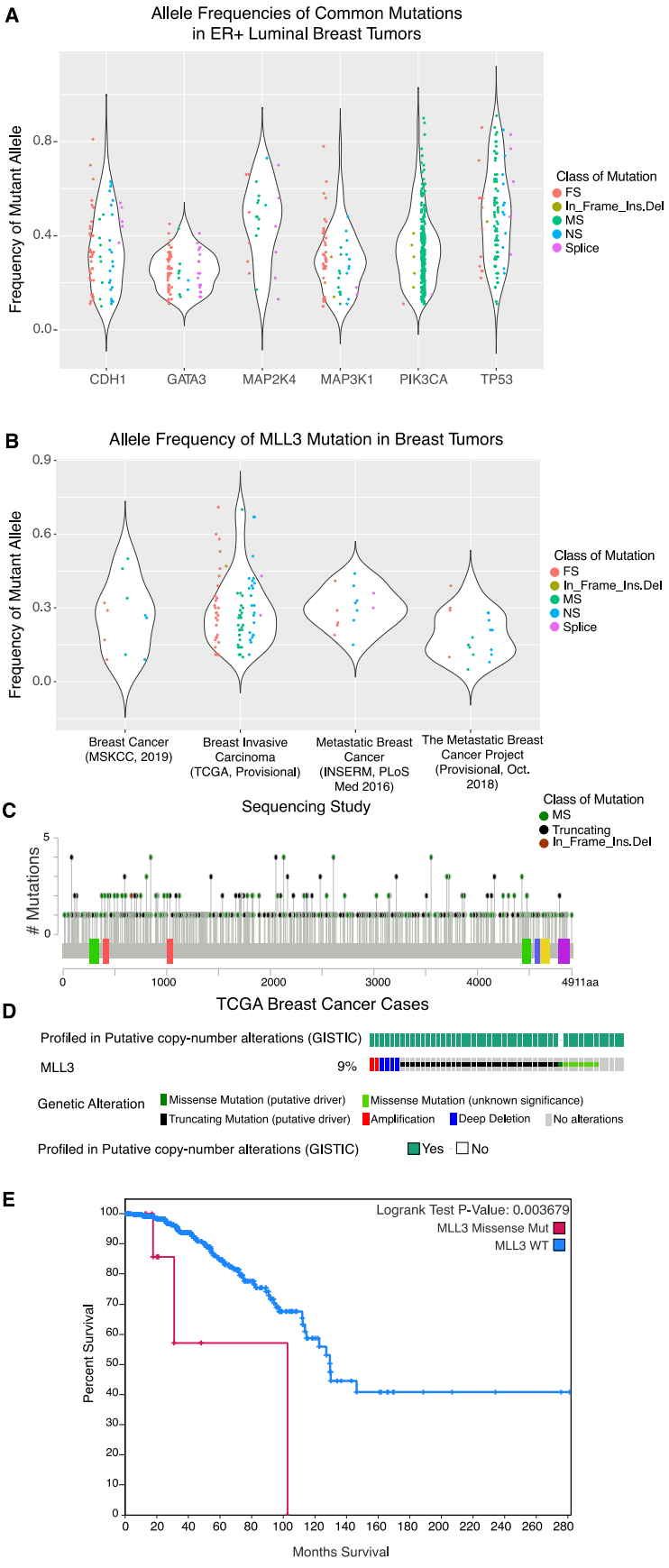


Figure 2-2. Supplement to Figure 2-1. (**a**) Frequency of mutant MLL3 alleles in four different breast cancer sequencing studies (n =125 mutations). ER+ = estrogen receptor positive. FS = frameshift. In_Frame_Ins.Del = in frame insertion or deletion. MS = missense. NS = nonsense. (**b**) Frequency of mutant alleles in the top recurrently mutated genes in TCGA ER+ breast cancer cases (n = 554 mutations) (**c**) MLL3 mutation lollipop plot of all TCGA breast cancer cases with RNA-seq data (n= 982 breast cancer cases). Green lollipops indicate missense mutations, black indicate truncating mutations, red indicate inframe mutations. Colored boxes indicate specialty domains as follows: PHD-like zinc-binding (green), PHD finger (red), F/Y-rich N-terminus (blue), F/Y-rich C-terminus (yellow), catalytic SET domain (purple). (**d**) Snapshot of the copy number alterations and mutations in MLL3 in ER+ luminal breast cancer samples from TCGA. (n= 581 ER+ luminal breast cancer samples, only 45 samples which have copy number alterations and mutations shown) (**e**) Overall survival curve from cBioPortal comparing TCGA ER+ luminal breast cancers that has a missense mutation in MLL3 to those that are WT for MLL3. Logrank Test, p = 0.003679. (n = 581 samples) WT = wildtype.

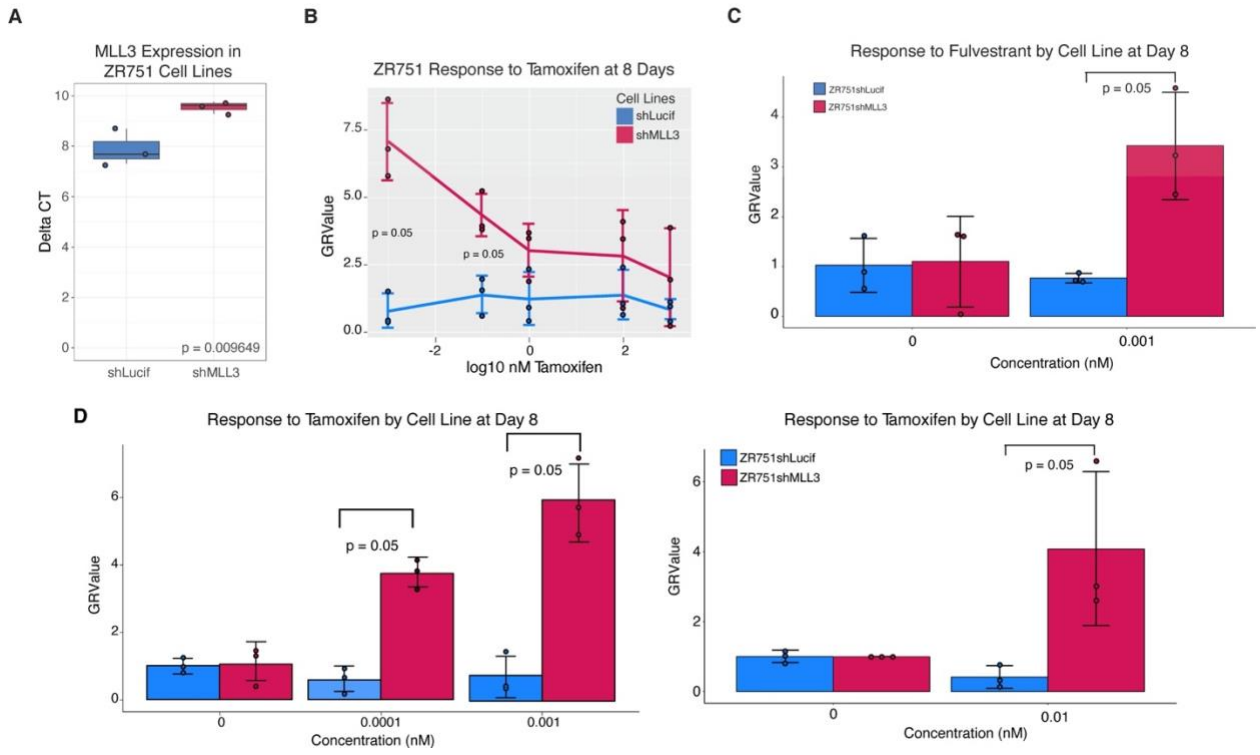


Figure 2-3. MLL3 KD Confers Endocrine Therapy Resistance. (**a**) Delta CT values from qPCR. The center line signifies the median, box limits signify upper and lower quartiles, and whiskers signify the 1.5x interquartile range. All data points are shown as dots. (n = 3 biological replicates) (p = 0.009649, one-tailed unpaired t-test) (**b**) Crystal violet assay for ZR751shLucif (blue) and ZR751shMLL3 (red) treated with Tamoxifen for 8 days. GRValues reflect the effect of a treatment such as Tamoxifen on the growth rate of a cell population on a per-division basis rather than on the percent viability. $GR(c)=(2^{*}(log2(x(c)/x_0))/(log2(x(o)/x_0)))-1$, where *x(c)* is the number of cells in a treated well at concentration *c*, x_0 is the number of cells in a well at beginning of treatment, and *x(o)* is the number of cells in an untreated well. Error bars represent standard deviation. (n = 3 biological replicates) (p = 0.05, p = 0.05, one-sided Wilcoxon rank sum test) (**c**) Crystal violet assay for ZR751shLucif (blue) and ZR751shMLL3 (red) treated with Fulvestrant for 8 days. Error bars represent standard deviation. (n = 3 biological replicates) (p = 0.05, one-sided Wilcoxon rank sum test) (**d**) Crystal violet assays for ZR751shLucif (blue) and ZR751shMLL3 (red) treated with Tamoxifen for 8 days. Error bars represent standard deviation. (n=3 biological replicates each experiment) (left, p = 0.05, p= 0.05; right, p = 0.05; one-sided Wilcoxon Rank Sum test)

Knockdown of *MLL3* changes the genomic enhancer landscape.

MLL3, as part of the coregulator complex ASCOM, monomethylates histone H3K4.[141] Loss of MLL3 leads to a loss of H3K4me1 across the genome in MEF cells.[142] We posited that loss of MLL3 function would result in a similar loss of global H3K4me1 in ER+ breast cancer. We chose to test this hypothesis in ZR751, an ER+ breast cancer cell line wildtype for *MLL3*. ChIP-seq for H3K4me1 was performed with two biological replicates for ZR751shMLL3 and ZR751shLucif each, with inputs used as background controls. Samples were processed according to the ENCODE (phase-3) transcription factor and histone ChIP-seq best practices. Peak calling was accomplished with SPP[143] and reproducibility between replicate experiments was examined to provide thresholds for optimal peak selection with the Irreproducible Discovery Rate (IDR) framework.[123] The resulting set of peaks demonstrated a massive decrease in the number of H3K4me1 sites upon *MLL3* KD (**2-4a**). This loss is global, and comparison of H3K4me1 peaks directly shows that, on average, there is more H3K4me1 deposited at ZR751shMLL3 H3K4me1 genomic locations in control cells than in KD cells, suggesting that H3K4me1 genomic locations common to both cell lines have lower amounts of H3K4me1 in ZR751shMLL3 compared to control (**2-5b**-**c**). Comparison of the H3K4me1 ChIP-seq samples with DiffBind[74,144] proved this to be true, with 97.3% (19,619/20,166, FDR<0.05) of common H3K4me1 genomic locations having a positive fold change and therefore more H3K4me1 deposited in the control than in the KD (**2-4b**).

We reasoned that changes in the H3K4me1 enhancer landscape due to *MLL3* KD would be accompanied by genomic shifts in ERα binding. Indeed, ERα ChIP-seq revealed a substantial shift in ERα binding upon KD of *MLL3* (**2-4a**). At genomic locations bound by ERα in ZR751shLucif, there was a greater intensity of ERα binding in ZR751shLucif cells than in ZR751shMLL3 cells, and vice versa (**2-5b-c**). Upon analysis with DiffBind we saw that indeed the differentially bound genomic locations with an FDR less than 0.05 were enriched in the ZR751shMLL3 condition if they overlapped a peak called for ZR751shMLL3, and vice versa (**2-4b**). We predicted that the altered enhancer landscape created by loss of MLL3, comprised of major H3K4me1 loss and an altered ERα binding profile, would affect genes in pathways associated with cancer phenotypes. Assessment with GREAT, which assigns peaks to genes using both proximity and gene annotation categories, was used to evaluate pathway and gene signature enrichment for our ChIP-seq data (**2-5d**). This analysis showed that, as a whole, H3K4me1 peaks in the *MLL3* KD, but not in the control, are enriched for the Creighton 'group 4 set' of genes associated with acquired endocrine therapy resistance in breast tumors (**2-4c**).[126] In the *MLL3* KD, ERα peaks are enriched for genes downregulated in breast cancers formed by MCF-7 xenografts resistant to Tamoxifen (**2-4c**). Enrichment in these gene terms suggest that *MLL3* KD confers endocrine therapy resistance to breast cancer cells via a global loss of H3K4me1 and a shift in ERα binding profile. Given these results, we assessed the response of *MLL3* KD cells to endocrine therapies Tamoxifen and Fulvestrant and found that *MLL3* KD results in increased resistance to endocrine therapies (**2-4d, 2-3b-d**).

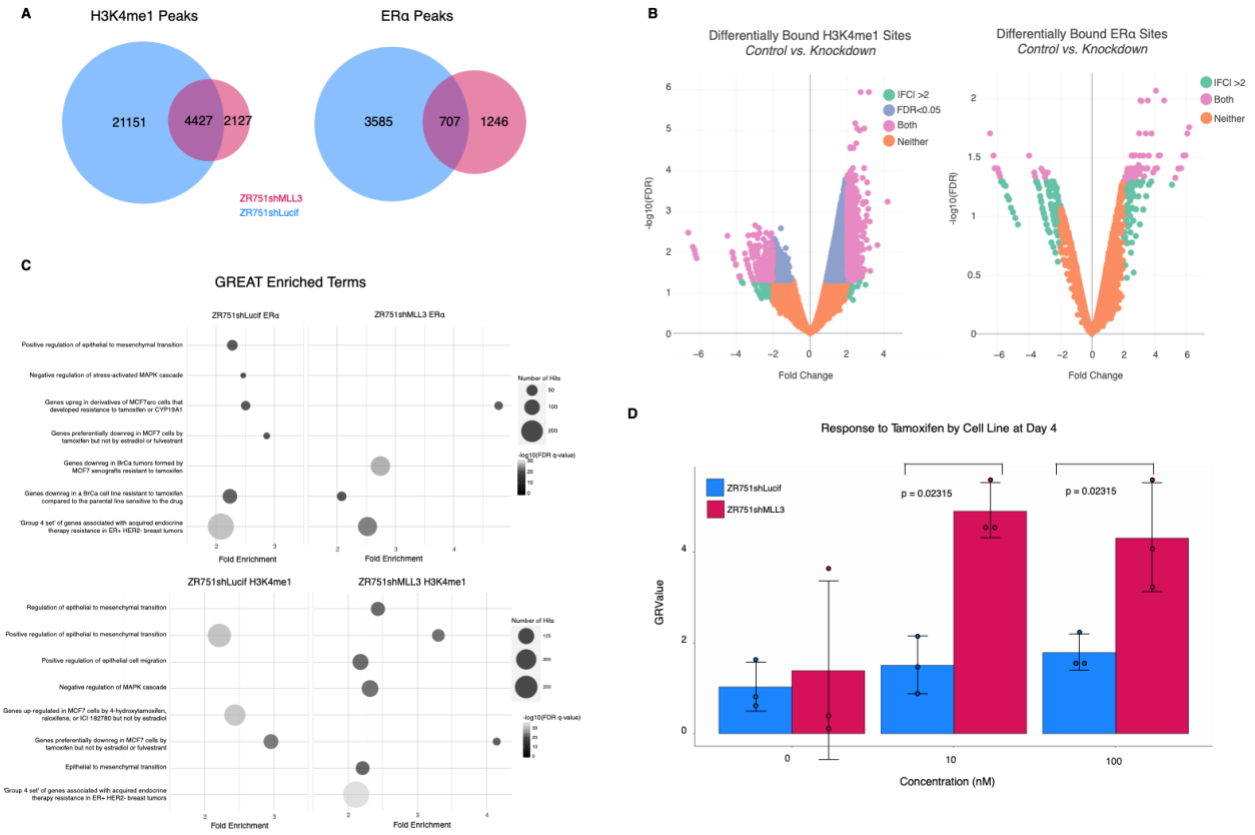
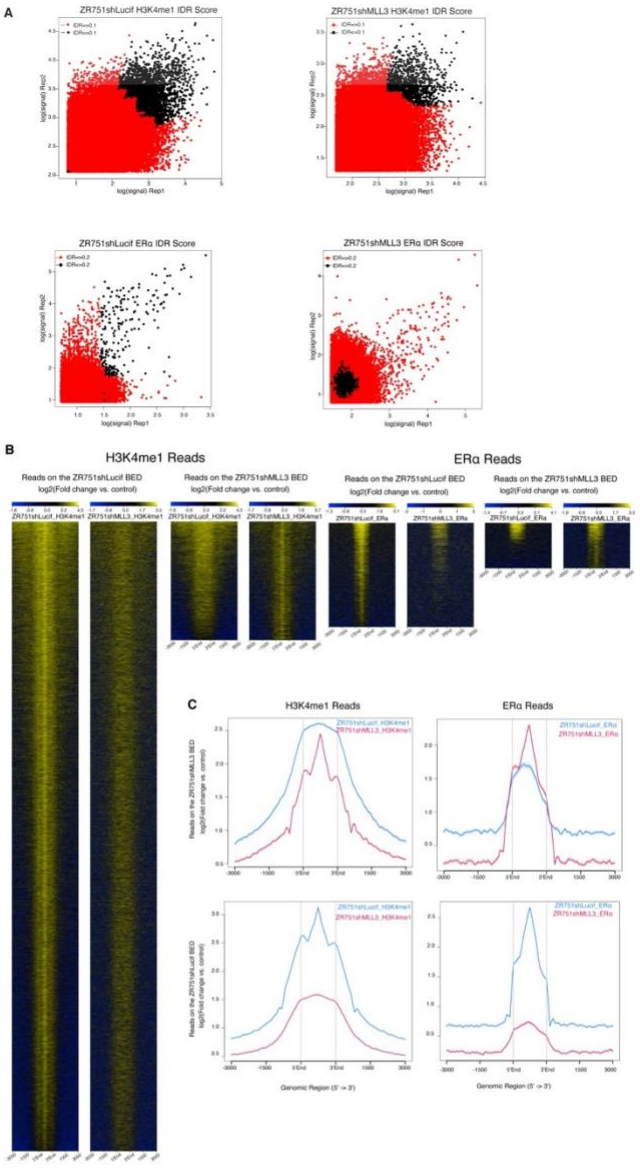


Figure 2-4. Knockdown of MLL3 leads to a reduction in H3K4me1 that correlates with a shift in ERα-binding. (**a**) Venn diagrams showing either ERα or H3K4me1 peaks between merged ZR751shLucif (blue) and merge ZR751shMLL3 (red) (2 biological replicates per experiment, pooled samples with peaks chosen through IDR protocol) (**b**) Differentially bound H3K4me1 (left) and ERα (right) sites upon MLL3 knockdown in ZR751. Fold change and -log10(FDR) are plotted for the sites found by DiffBind to be differentially bound between ZR751shLucif and ZR751shMLL3. Genomic sites that have an absolute value fold change of 2 or greater are green if they do not have an FDR of less than 0.05, and pink if they do. Sites that have an FDR of less than 0.05 but do not have an absolute fold change greater than 2 are blue. Sites with an FDR of more than 0.05 and an absolute fold change of less than 2 are orange. Positive fold enrichment indicates higher amounts of binding in ZR751shLucif compared to ZR751shMLL3. FC = fold change. (**c**) Gene enrichment terms from GREAT for peaks that were from either ZR751shLucif or ZR751shMLL3 cells for ERα-binding or H3K4me1 deposition. The results are displayed in matching

graphs where each line on the y-axis is a gene-term, the x-axis shows increasing fold enrichment, the color of the circle denotes the significance, and the size of the circle denotes the number of genes from the dataset belonging to the respective gene-term. GREAT tool's binomial test was employed. (2 biological replicates per experiment, pooled samples with peaks chosen through IDR protocol) (**d**) Crystal violet assay for ZR751shLucif (blue) and ZR751shMLL3 (red) treated with Tamoxifen for 4 days. Error bars represent standard deviation. (n=3 biological replicates) (p = 0.02315, p = 0.02315, one-sided Wilcoxon Rank Sum test of GRValues). GRValues reflect the effect of a treatment such as Tamoxifen on the growth rate of a cell population on a per-division basis rather than on the percent viability. $GR(c)$=(2*($log2(x(c)/x_0)$)/($log2(x(o)/x_0)$))−1, where *x(c)* is the number of cells in a treated well at concentration *c*, x_0 is the number of cells in a well at beginning of treatment, and *x(o)* is the number of cells in an untreated well.



D

GREAT Terms Enriched for ZR751sLucif ERα	q-value
Negative regulation of immune response	1.76E-31
Cis-eQTL in the liver that contribute to metabolic quantitative traits	4.57E-30
The 'group 4 set' of genes associated with acquired endocrine therapy resistance in breast tumors expressing ESR1 but not ERBB2	1.54E-27
Up-regulated genes from the optimal set of 550 markers discriminating breast cancer samples by ESR1 expression	1.89E-24
Genes up-regulated in MCF7 cells at 6 h of estradiol treatment	8.94E-17
Genes up-regulated in the luminal B subtype of breast cancer	2.00E-16
Genes up-regulated by estradiol in U2OS cells expressing ESR1 or ESRG	2.96E-15
Genes down-regulated in skin upon knockout of AOX4	3.81E-15
Negative regulation of stress-activated MAPK cascade	6.63E-19
Genes up-regulated in breast cancer samples positive for ESR1 compared to the ESR1 negative tumors	7.51E-14
Genes within amplicon 12p13-q21 identified in a copy number alterations study of 191 breast tumor samples	1.14E-13
Genes up-regulated in MCF7 cells after stimulation with NRG1	1.21E-13
Genes down-regulated in U2-OS Tet-On cells after induction of ESX1 expression	1.74E-12
Genes from the 12p region that were down-regulated in choriocarcinoma cells compared to normal testis	5.74E-13
Genes up-regulated in ductal carcinoma vs normal lobular breast cells	9.90E-13
Termination of signal transduction	6.20E-15
Genes up-regulated in HUVEC cells after stimulation with leukotriene or thrombin for 1 h	1.74E-12
Regulation of transforming growth factor beta production	7.02E-15
Genes down-regulated in ST1 cells after knockdown of FRA2 by RNAi	2.16E-12
Termination of G-protein coupled receptor signaling pathway	8.00E-15
Genes up-regulated in ER(-) / PR(-) breast tumors with molecular similarity to ER(+) relative to the rest of the ER(-) / PR(-) samples	4.03E-12
Genes from the 12p region that were down-regulated in seminoma tumors compared to normal testis	1.13E-11
Genes from the 12p region that were down-regulated in embryonic carcinoma tumors compared to normal tissue	1.52E-11
Genes from common regions of gains observed in more than 15% of 148 primary breast cancer tumors	3.13E-11
Genes important for ovulation, based on mouse models with female fertility defects	5.31E-11

GREAT Terms Enriched for ZR751shMLL3 ERα	q-value
Genes down-regulated in breast cancer tumors formed by MCF-7 xenografts resistant to tamoxifen	5.3469E-26
The 'group 4 set' of genes associated with acquired endocrine therapy resistance in breast tumors expressing ESR1 but not ERBB2	2.1545E-21
Genes up-regulated in MCF7 cells at 6 h of estradiol treatment	3.2890E-18
Genes rapidly up-regulated in breast cancer cell cultures by estradiol	7.60E-16
Genes bound by ESR1 and up-regulated by estradiol in MCF-7 cells	1.19E-15
Genes bound by ESR1 and up-regulated by estradiol in MCF-7 cells expressing constitutively active form of AKT1	5.25E-15
Genes up-regulated in MCF7 cells at 24 h of estradiol treatment	9.69E-15
Up-regulated genes from the optimal set of 550 markers discriminating breast cancer samples by ESR1 expression	1.17E-11
Genes up-regulated in breast cancer samples positive for ESR1 compared to the ESR1 negative tumors	9.16E-11
Genes up-regulated by estradiol in U2OS cells expressing ESR1 or ESRG	2.07E-10
Genes important for ovulation, based on mouse models with female fertility defects	2.26E-09
Genes up-regulated in derivatives of MCF-7aro cells that developed resistance to tamoxifen or inhibitors of aromatase	3.40E-09
Genes up-regulated in lobular carcinoma vs normal lobular breast cells	3.01E-07
Genes down-regulated in H1975 cells resistant to gefitinib after treatment with EGFR inhibitor CL-387785 for 6h	3.31E-06
The 'ER-alpha profile' genes up-regulated in T47D cells upon activation of ESR1 by estradiol	5.10E-06
Genes up-regulated by estradiol and not modulated by ESRRA in MCF-7 cells	9.64E-06
Cellular response to steroid hormone stimulus	2.09E-04
Cis-eQTL in the liver that contribute to metabolic quantitative traits	9.42E-06
Genes from the red module which are down-regulated in HAEC cells after exposure to cefHPC	9.36E-06
The 'CLASS-1' set of pro- and anti-inflammatory cytokines whose expression identifies stage I lung adenocarcinoma patients with poor prognosis	2.58E-05
Cellular response to alcohol	2.38E-04
Genes most down-regulated in hypoxia tolerant NCI H460 cells	2.59E-05
Genes down-regulated in bladder cancer cells resistant to cisplatin compared to the parental cells sensitive to the drug	2.53E-05
Response to activity	2.50E-04
Up-regulated genes in HepG2 cells overexpressing an oncogenic variant of tumor suppressor TP50 compared to its wild type form	2.95E-05

Figure 2-5. Supplement to Figure 2-4. (**a**) IDR scatterplot of log(signal) of ZR751 H3K4me1 ChIP replicates (top) and ERα replicates (bottom). Red dots signify peaks that have an IDR score greater than the chosen threshold, 0.1 for H3K4me1 and 0.2 for ERα. Black dots signify peaks that have an IDR score of less than or equal to the chosen threshold. IDR = Irreproducibile Discovery Rate. (**b**) Heatmaps of ERα and H3K4me1 ChIP-seq reads plotted on ZR751shLucif and ZR751shMLL3 bed files of the respective ChIP-seq experiment. (n = 2 biological replicates per experiment, shown is one pooled bed file per experiment with peaks chosen through IDR protocol) (**c**) Histograms of either H3K4me1 or ERα ChIP read enrichment over control, plotted over mapped peaks from either ZR751shLucif or ZR751shMLL3. (2 biological replicates per experiment, pooled samples with peaks chosen through IDR protocol) (**d**) Chosen terms enriched in the GREAT analysis of ERα ChIP-seq experiments for ZR751shLucif and ZR751shMLL3. Q-value shown in right column of the tables (binomial test from GREAT). (n = 2 biological replicates per experiment, one pooled bed file per experiment with peaks chosen through IDR protocol).

<u>Loss of functional MLL3 leads to enhanced transcription of genes associated with aggressive tumor behavior.</u>

Differential expression of RNA-seq in ZR751shLucif and ZR751shMLL3 identified 3,037 upregulated and 3,518 downregulated genes upon KD of *<u>MLL3</u>*, q<0.05. To determine if the same gene expression changes were occurring in clinical breast tumors with *<u>MLL3</u>* mutations, we utilized RNA-seq data from TCGA ER+ luminal breast cancer patients; this analysis revealed 688 upregulated and 693 downregulated genes based on *<u>MLL3</u>* mutation status, q<0.05. Comparison of the two sets of DEG from the ZR751 (q<0.05) and TCGA (p<0.05) analyses revealed a significant overlap between both upregulated (3,036 ZR751, 1,185 TCGA) and downregulated (3,643 ZR751, 3,638 TCGA) gene sets (**2-6a**, **2-10b, 2-7a-d**). This MLL3-deficiency signature consisted of 208 upregulated genes (p = 0.0000072, Fisher's exact test) and 750 downregulated genes (750 genes, $p = 4 \times 10^{-17}$, Fisher's exact test).

Given the enhanced endocrine therapy resistance displayed in proliferation assays and poorer overall survival curves, we reasoned that the transcriptional program of *<u>MLL3</u>* KD cells would be enriched for cancer progression pathways. Webgestalt over-representation analysis (ORA) of ZR751 DEG identified terms associated with aggressive tumor behavior due to AKT1 activation, including "genes bound by ERα and up-regulated by estradiol in MCF7 cells expressing constitutively active AKT1" (**Table 2-3**).[127] Webgestalt ORA of the TCGA DEG illuminated positive enrichment in *<u>MLL3</u>* mutants for "genes upregulated in ER+ breast cancer samples" and "KRAS-dependency signature genes", and negative enrichment for "genes downregulated in ER+ breast cancer samples" (**Table 2-4**). Interestingly, Gene Set Enrichment Analysis (GSEA) for both TCGA and ZR751 DEG revealed a

significant positive enrichment score for "genes induced by Akt and sensitive to everolimus" (**2-6b, c**). This gene signature is correlated with an increased incidence of metastases and a shorter disease-free survival time in several breast tumor datasets.[145] It is worth noting that mutations in genes in the ASCOM complex, which includes MLL3, and PIK3CA pathway mutations co-occur in breast cancer more than we would expect by chance.[146] The mTOR pathway activation gene signature is also enriched in *MLL3* KD and mutant breast cancer samples compared to WT (**Table 2-5**). This signature is associated with poorer outcome in breast cancer compared to the pAKT pathway activation signature.[147] These results demonstrate that canonical ERα target genes important to aggressive cancer behavior are upregulated upon loss of *MLL3*.

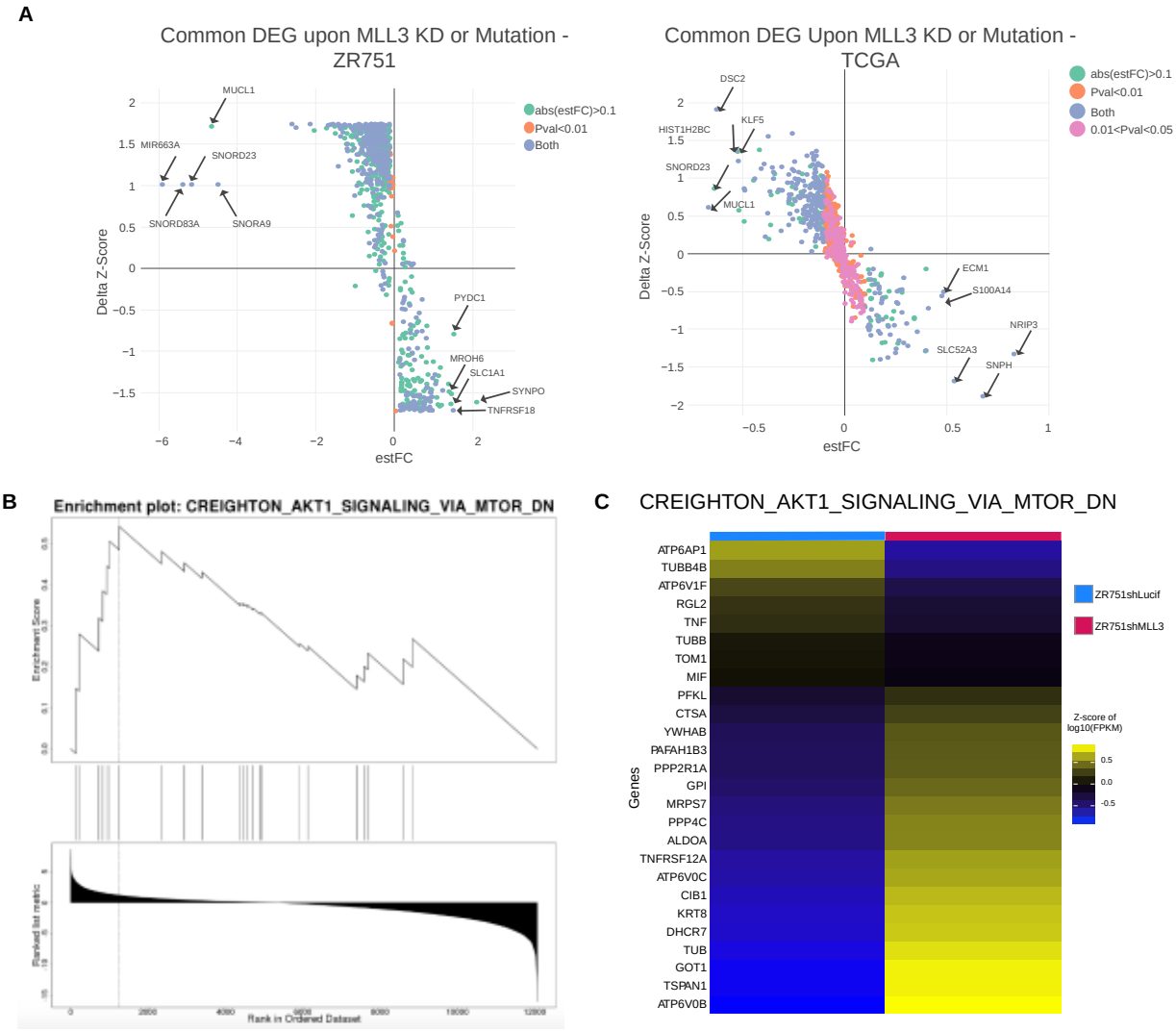


Figure 2-6. Knockdown of MLL3 and mutation of MLL3 share an MLL3-deficiency transcriptional signature. (**a**) Scatterplot of the differentially expressed genes in common between ZR751 breast cancer cells upon MLL3 knockdown (left) and TCGA ER+ luminal breast cancer samples with MLL3 mutations (right). Estimated log fold change from the gene-by-gene linear regression model with ANOVA is plotted against the change in Z-score between the control (ZR751shLucif on left, MLL3 wildtype samples on right) and the experimental (ZR751shMLL3 on left, MLL3 mutant samples on right). Genes with an absolute estimated log fold change greater than 0.1 are colored green if the p-value is larger than 0.01, and blue is the p-value is less than 0.01. Genes with a p-value less than 0.01 and absolute estimated log fold change less than 0.1 are orange. DEG = differentially expressed genes. estFC = estimated log fold change. (**b**) TCGA enrichment plot for selected MSigDB term CREIGHTON_AKT1_SIGNALING_BY_MTOR_DN by WebGestalt GSEA. Normalized enrichment score 1.8757, FDR q-value 0.026442. (**c**) ZR751 cell lines heatmap of Z-scores for merged-sample log10 normalized FPKM for genes in the CREIGHTON_AKT1_SIGNALING_BY_MTOR_DN term by WebGestalt GSEA, normalized enrichment score 2.0273, FDR q-value 0.0026447. n = 2 biological replicates per experiments.

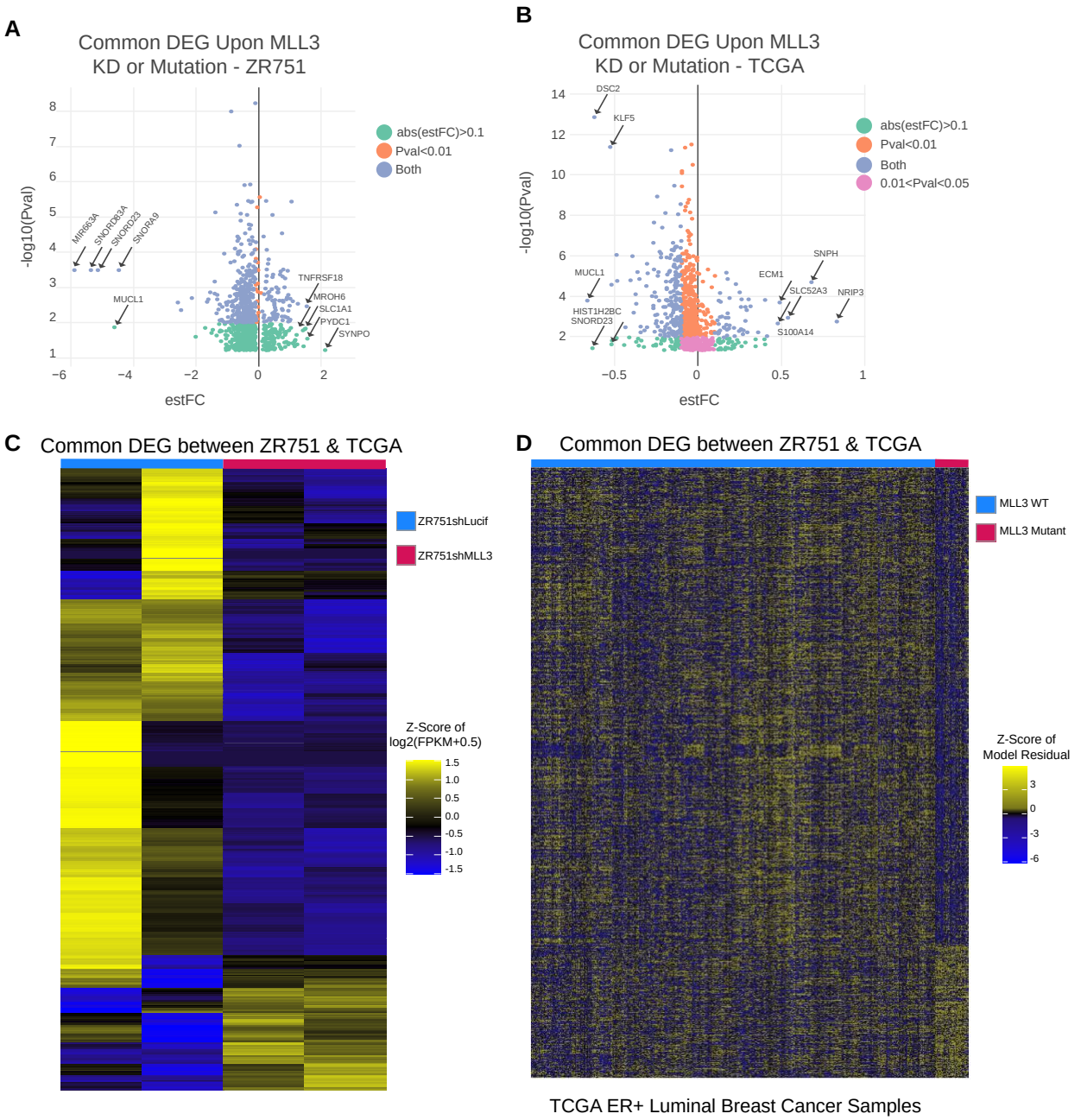


Figure 2-7. Supplement to Figure 2-6. (**a**) Scatterplot of the differentially expressed genes in common between ZR751 breast cancer cells upon MLL3 knockdown and TCGA ER+ luminal breast cancer samples with MLL3 mutations. Estimated log fold change between the control (ZR751shLucif) and the experimental (ZR751shMLL3) from the gene-by-gene linear regression model with ANOVA is plotted against the -log10(p-value). Genes with an absolute estimated log fold change greater than 0.1 are colored green if the p-value is larger than 0.01, and blue is the p-value is less than 0.01. Genes with a p-value less than 0.01 and absolute estimated log fold change less than 0.1 are orange. DEG = differentially expressed genes. estFC = estimated log fold change. (**b**) Scatterplot of the differentially expressed genes in common between ZR751 breast cancer cells upon MLL3 knockdown and TCGA ER+ luminal breast cancer samples with MLL3 mutations. Estimated log fold change between the control (TCGA ER+ luminal MLL3 wildtype breast cancer samples) and the experimental (TCGA ER+ luminal MLL3 mutant breast cancer samples) from the gene-by-gene linear regression model with ANOVA is plotted against the -log10(p-value). Genes with an absolute estimated log fold change greater than 0.1 are colored green if the p-value is larger than 0.01, and blue is the p-value is less than 0.01. Genes with a p-value less than 0.01 and absolute estimated log fold change less than 0.1 are orange. Genes with a p-value between 0.01 and 0.05 and an absolute estimated log fold change less than 0.1 are pink. (**c**) Heatmap of the differentially expressed genes in common between the ZR751shLucif vs ZR751shMLL3 analysis and the TCGA MLL3 WT vs mutant analysis, consisting of 750 downregulated genes and 208 upregulated genes. Z-scores of ZR751 expression counts of replicates are shown. ANOVA FDR q<0.05. n = 2 biological replicates per experiment. (**d**) Heatmap of the differentially expressed genes in common between the ZR751shLucif vs ZR751shMLL3 analysis and the TCGA MLL3 WT vs mutant analysis, consisting of 750 downregulated genes and 208 upregulated genes. Residuals from linear regression model *not* accounting for MLL3 mutation status used for expression values to calculate z-scores. Mutant samples are denoted by red and WT by blue. ANOVA FDR q<0.05. WT = wildtype.

MLL3 KD-driven H3K4me1 loss and ERα binding shifts contribute to differential gene expression programs in breast cancer.
To investigate the relationship between the changes in the genomic enhancer landscape and ERα binding profiles with the transcriptional changes upon KD of *MLL3*, we assigned H3K4me1 and ERα ChIP-seq peaks to ZR751 DEG by proximity. To check the robustness of these assignments, we used a permutation-based analysis that demonstrated our experimentally determined binding sites were closer to DEG than expected by chance (**2-9a).** H3K4me1 peaks and ERα peaks in both cell lines gave us a robustness measure of p = 0, and >80% of peaks were assigned for all conditions**.** 4,179 genes out of the 6,677 DEG were assigned to at least one peak (**2-8a**). We hypothesized that there would be an association between losing ERα peaks, losing H3K4me1 peaks, and decreased gene expression, and vice versa. To test this hypothesis, we next assigned each DEG to a category based on whether the number of peaks assigned to it was larger in the control or *MLL3* KD. This categorization showed a pattern in which DEG with a higher number of H3K4me1 peaks assigned to ZR751shMLL3 than ZR751shLucif tend to be downregulated rather than upregulated in ZR751shMLL3. The converse is also true (**2-8b**).

To quantify this trend we used a two-sided Wilcoxon paired signed-rank test with continuity correction, which confirmed that while *MLL3* KD has a sizeable effect on the number of ERα peaks assigned to DEG in both the top 100 up- and downregulated gene sets ($p = 5.106 \times 10^{-9}$, r = 0.59; $p = 2.924 \times 10^{-11}$, r = 0.67 respectively), a more robust effect on the number of H3K4me1 peaks assigned to DEG in the top 100 up- and downregulated gene sets is evident ($p = 3.198 \times 10^{-15}$, r = 0.853; $p = 3.28 \times 10^{-12}$, r = 0.871). To investigate this relationship further, the top 100 upregulated DEGs and the top 100 downregulated DEGs were dichotomized to ERα peak gain or loss and H3K4me1 peak gain or loss (**2-8c, 2-9b**). Interestingly, the proportions of the top 100 upregulated and downregulated genes that gained ERα peaks were similar (20% of upregulated genes gained ERα peaks, and 15% percent of downregulated genes gained ERα peaks). This difference was not significant (p-value = 0.4566, 2-sample test for equality of proportions with continuity correction), suggesting that ERα peak number, per se, is not a dominating factor in determining the direction of gene expression change. However, a similar analysis for H3K4me1 peaks showed that the proportion of gained H3K4me1 peaks were vastly different (4% of the top 100 upregulated peaks, 64% of the top downregulated peaks). This difference was significant (p-value < 2.2×10^{-16}, 2-sample test for equality of proportions with continuity correction), suggesting that, interestingly, gain of H3K4me1 peaks upon loss of MLL3 is strongly associated with downregulation of gene expression.

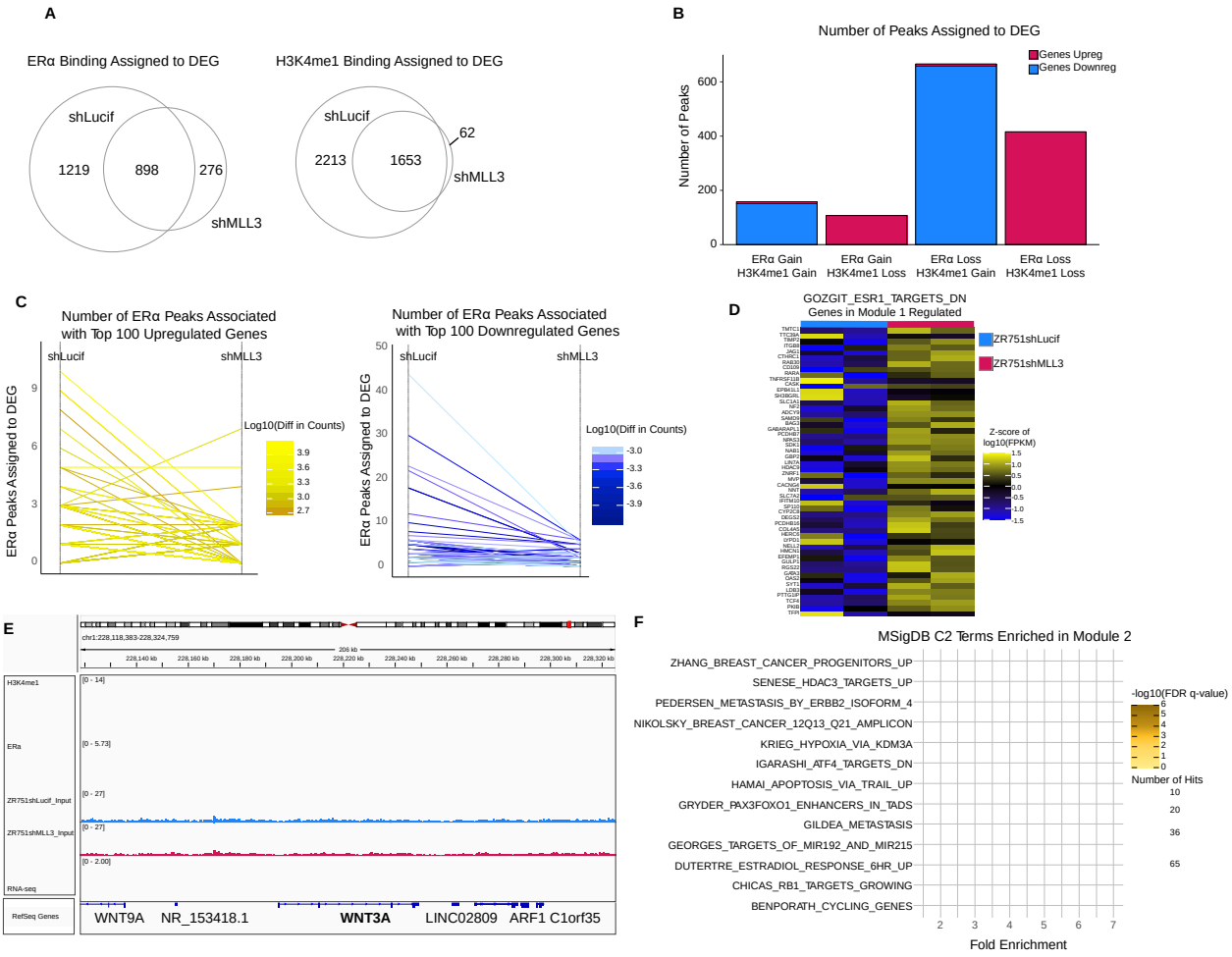


Figure 2-8. Knockdown of MLL3 in leads to a new transcriptional regulation program of ERα targets in conjunction with changes in H3K4me1 deposition. (**a**) Venn diagram of ERα and H3K4me1 ChIP-seq peak assignments to differentially expressed genes (DEG) in ZR751 MLL3 KD cells. (n = 2 biological ChIP-seq replicates per experiment) DEG = differentially expressed genes. (**b**) DEG upon MLL3 KD in ZR751 cells grouped into four categories based on the number of ERα and H3K4me1 ChIP-seq peaks assigned to each gene in the control and MLL3 KD conditions. (n = 2 biological ChIP-seq replicates per experiment) Upreg = upregulated expression. Downreg = downregulated expression. (**c**) Slope graph showing difference in number of ERα ChIP-seq peaks assigned to each DEG in ZR751s upon MLL3 KD, between the control and MLL3 KD conditions. The left graph shows the top 100 upregulated genes, and the right shows the top 100 downregulated genes. The color of each individual line represents the difference in log10-normalized counts. (n = 2 biological ChIP-seq replicates per experiment) (**d**) Heatmap of Z-score of the log10 normalized FPKM of genes in the GOZGIT_ESR1_TARGETS_DN MSigDB term, which was significantly enriched in the Group 1 genes using WebGestalt Over Representation Analysis (ORA) (number of hits = 38, enrichment ratio = 2.1328, FDR q-value = 0.0129) (n = 2 biological RNA-seq replicates per experiment) (**e**) IGV Genome Browser snapshot of WNT3A, which belongs to Group 1 where gene expression is increased, but number of H3K4me1 and ERα ChIP-seq peaks assigned to the gene are decreased upon MLL3 KD. (n = 2 biological ChIP-seq replicates per experiment (**f**) Bubble plot showing significant MSigDB C2 terms for Group 2 genes by WebGestalt ORA. (n = 2 biological ChIP-seq replicates per experiment)

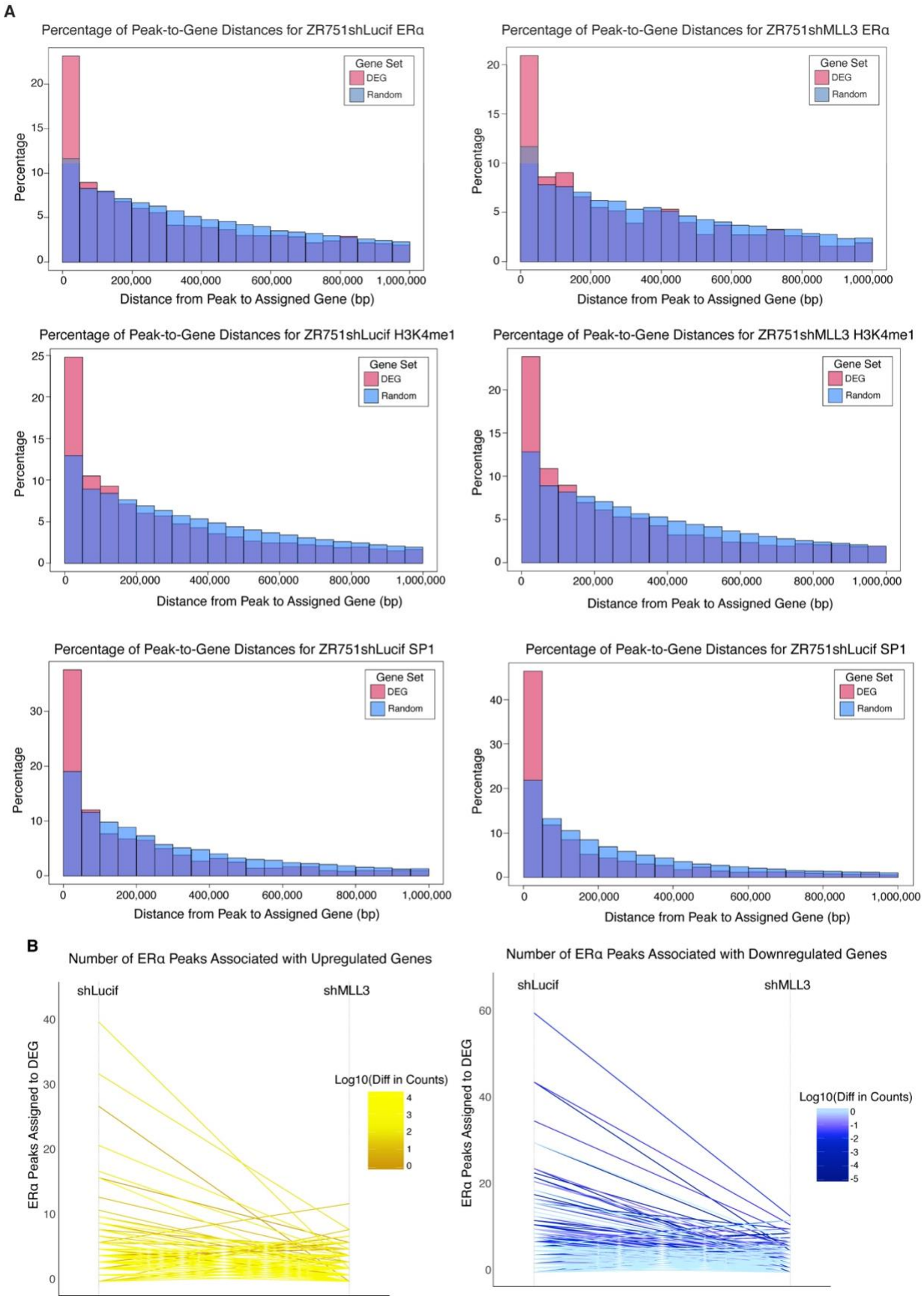


Figure 2-9. Supplement to Figure 2-8. (**a**) Peak-gene assignment distance for the ZR751 DEG in red, and a matched number of randomly chosen genes from hg19 for 1000 repetitions in teal. Bp = basepair. DEG = differentially expressed gene. (**b**) Slope graph showing difference in number of ERα ChIP-seq peaks assigned to each DEG in ZR751s upon MLL3 KD, between the control and MLL3 KD conditions. The left graph shows the upregulated genes and the right shows the downregulated genes. The color of each individual line represents the difference in log10-normalized counts. (n = 2 biological replicates per experiment for both RNA-seq and ChIP-seq)

Two new regulatory programs on the H3K4me1-ERα axis drive transcriptional enrichment for *ESR1* target genes and genes associated with aggressive tumor behavior upon *MLL3* KD.

To further refine our model of how loss of MLL3 enhances endocrine therapy resistance through histone mark changes and shifts in the ERα binding profile, we proposed that genes with similar changes in enhancer landscape, ERα binding, and direction of expression upon loss of *MLL3* would share similar biological functions. To identify genes with similar regulatory profiles and expression levels, we took an unbiased approach, grouping genes with at least one assigned H3K4me1 or ERα peak from either cell line into the four possible H3K4me1 categories: 1) H3K4me1 gain, expression upregulated 2) H3K4me1 loss, expression upregulated 3) H3K4me1 gain, expression downregulated 4) H3K4me1 loss, expression downregulated, as well as the four possible ERα categories: 1) ERα gain, expression upregulated 2) ERα loss, expression upregulated 3) ERα gain, expression downregulated 4) ERα loss, expression downregulated. Then, all 16 possible pairwise overlaps were assessed using Fisher's Exact Test with Bonferroni correction; overlaps with a significant p-value indicates that there is a module of co-regulated genes with that ERα and H3K4me1 status (**Table 2-2**). We tested for significant overlap between groups of DEG with either a gain or loss of assigned ERα ChIP-seq peaks upon *MLL3* KD, and either a gain or loss of assigned H3K4me1 peaks upon *MLL3* KD. Four out of eight comparisons showed a significant overlap by one-sided Fisher's Exact Test ($p < 0.05$) with a non-zero Jaccard index: 1) upregulated genes with a loss in assigned H3K4me1 peaks per gene (ppg) and a loss in assigned ERα ppg upon *MLL3* KD (416 genes, $p = 2.90 \times 10^{-227}$), 2) upregulated genes with a loss in H3K4me1 ppg and a gain in ERα ppg upon *MLL3* KD (107 genes, $p = 1.5 \times 10^{-38}$), 3) downregulated genes with a gain in H3K4me1 ppg and a loss in ERα ppg upon *MLL3* KD (658 genes, $p = 3.4 \times 10^{-266}$), and 4) downregulated genes with a gain in H3K4me1 ppg and gain in ERα ppg upon *MLL3* KD (151 genes, 8.9×10^{-35}). We collapsed the four groups into two modules based on direction of effect in conjunction with H3K4me1 loss/gain (**Table 2-2**), as ERα can both drive and repress transcription of its targets. Taken together, these patterns in differential gene expression and number of associated peaks suggest that H3K4me1 peaks in WT cells that are lost after KD of *MLL3* are associated with gene upregulation, while the H3K4me1 peaks gained after KD are primarily associated with gene repression.

Module	Category	H3K4me1	ERα	Direction of Effect	Overlap	P-value	Bonferroni (α < 0.003125)	Jaccard Index	Odds Ratio
1	**1**	**Loss - 1356**	**Loss - 494**	**Upreg**	**416**	**2.90E-227**	**Yes**	**0.3**	**29.7**
	2	**Loss - 1356**	**Gain - 159**	**Upreg**	**107**	**1.50E-38**	**Yes**	**0.1**	**8.7**
2	**3**	**Gain - 1150**	**Gain - 333**	**Downreg**	**151**	**8.90E-35**	**Yes**	**0.1**	**4.4**
	4	**Gain - 1150**	**Loss - 1170**	**Downreg**	**658**	**3.40E-266**	**Yes**	**0.4**	**13.1**
3	5	Loss - 0	Loss - 1170	Downreg	0	1	No	0	0
4	6	Loss - 0	Gain - 333	Downreg	0	1	No	0	0

5	7	Gain - 53	Loss - 494	Upreg	8	0.039	No	0	2.2
6	8	Gain - 53	Gain - 159	Upreg	7	2.30E-04	Yes	0	6.5

Table 2-2: Categories of Regulons Affected by Knockdown of *MLL3.* The table displays the organization of ZR751 differentially expressed genes (DEG) based on whether a gain or loss of associated ERα and H3K4me1 peaks were observed in the *MLL3* KD compared to the control. The background size used for the one-sided Fisher's exact test was 6,677 genes, as this was the number of DEG to which the peaks were matched. The GeneOverlap R package, by Li Shen was utilized. Categories in bold had significant overlaps. Significant categories sharing two characteristic changes were collapsed into modules.

Pathway analysis of genes belonging to the group 1 module showed significant enrichment for *ESR1* targets (52 genes, enrichment ratio 2.7435, FDR q = 6.915×10^{-8}) (**2-8d, Table 2-7**). This implies that despite a decrease in regulatory H3K4me1 and ERα peaks per upregulated gene upon *MLL3* KD, *ESR1* targets are being transcribed at a higher level in *MLL3* KD cells, for example, *WNT3A* (**2-8e, 2-12a-f**). Module 2 is enriched for several carcinogenic signatures, including "top genes down-regulated in metastatic vs non-metastatic bladder cancer cell lines" and "genes up-regulated in primary melanoma, sensitive to TRAIL compared to metastatic melanoma, resistant to TRAIL" (**2-8f, Table 2-8**). These results suggest that apoptosis via TRAIL is being evaded in *MLL3* KD cells, and pathways involved in metastasis are being expressed at higher levels than in the control.

SP1 binding increases upon *MLL3* KD.

It is probable that the change in the enhancer landscape and ERα binding profile upon *MLL3* KD would be accompanied by a new milieu of transcriptional regulators responsible for aggressive behavior. To find these regulators, we interrogated motifs found in our ChIP-seq and RNA-seq data. We first analyzed ERα peaks that were gained upon the loss of MLL3 using MEME, which looks at the DNA sequences of the peaks to identify enrichment of binding motifs, which were then classified as belonging to transcription factors using TOMTOM. This analysis identified GATA3, FOS, and SP1 motifs enriched in ERα peaks gained after *MLL3* knockdown (**2-10a**).[129] The iRegulon plug-in in Cytoscape leverages both precomputed motifs and ChIP-seq data to identify enriched transcription factor binding sites when presented with a gene list.[128] Thus, we used genes that were differentially expressed in the same direction in both our ZR751 *MLL3* KD and in the TCGA mutant tumors to define a set of 958 differentially expressed genes as a MLL3-deficient signature. iRegulon analysis of this gene list identified *SP1* as a candidate transcription factor for one of the top ten most-enriched motifs for upregulated genes in the MLL3-deficient signature (**2-10c**). Intriguingly, *SP1* was significantly upregulated in MLL3 mutants in our TCGA dataset (p = 2.32e-6), although there was no statistically significant differential expression in the ZR751 cell line.

Changes in gene expression are not the only mechanism of regulation, and we hypothesized that the change in enhancer landscape might change the transcription factor milieu regardless of expression. As SP1 motifs demonstrated enrichment in both our ChIP-seq and RNA-seq datasets upon *MLL3* KD, we hypothesized that loss of MLL3 leads to increased activity of SP1. While the DEMETER tool for cancer-cell line dependencies illuminated no trend toward increased or decreased dependence on SP1 for *MLL3*-mutant ER+ breast cancer cells lines compared to those that are *MLL3*-WT (**2-10d**),[130] we investigated the SP1 binding patterns in ZR751shLucif and ZR751shMLL3. ChIP-seq for SP1 demonstrated a massive gain of 2,182 binding sites in *MLL3* KD cells (**2-10e, 2-11b-d**). This suggests that while SP1 was not transcribed at a significantly higher rate in the *MLL3* KD, it is differentially bound to the genome depending on *MLL3* status in ZR751s.

To identify which genes SP1 regulates in control and *MLL3* KD cells, SP1 peaks were assigned to DEG in ZR751 cells using the method described for H3K4me1 and ERα. Strikingly, the largest group of DEG with both ERα and SP1 assignments are those that have an ERα peak loss and an SP1 peak gain upon *MLL3* KD**.** Figure **5f** illustrates that upon *MLL3* KD, there is a switch from ERα to SP1 regulation of genes. Furthermore, when gene assignments between ERα and SP1 categories are compared, there is a significant overlap by one-sided Fisher's exact test between DEG with a change in number of ERα peak assignments in *MLL3* KD cells and those with a gain in the number of assigned SP1 peaks ($p = 1.4 \times 10^{-34}$, **Table 2-9**, **2-11e-f**). Thus, SP1 may play a role in creating a transcriptome resistant to endocrine therapy by regulating the transcription of ERα targets that have altered ERα binding upon *MLL3* KD. In fact, 381 (nearly half of the 809 genes in Module 2) gain SP1 peaks upon MLL3 KD, while 22 of the Module 1 genes lose SP1 peaks upon MLL3 KD.

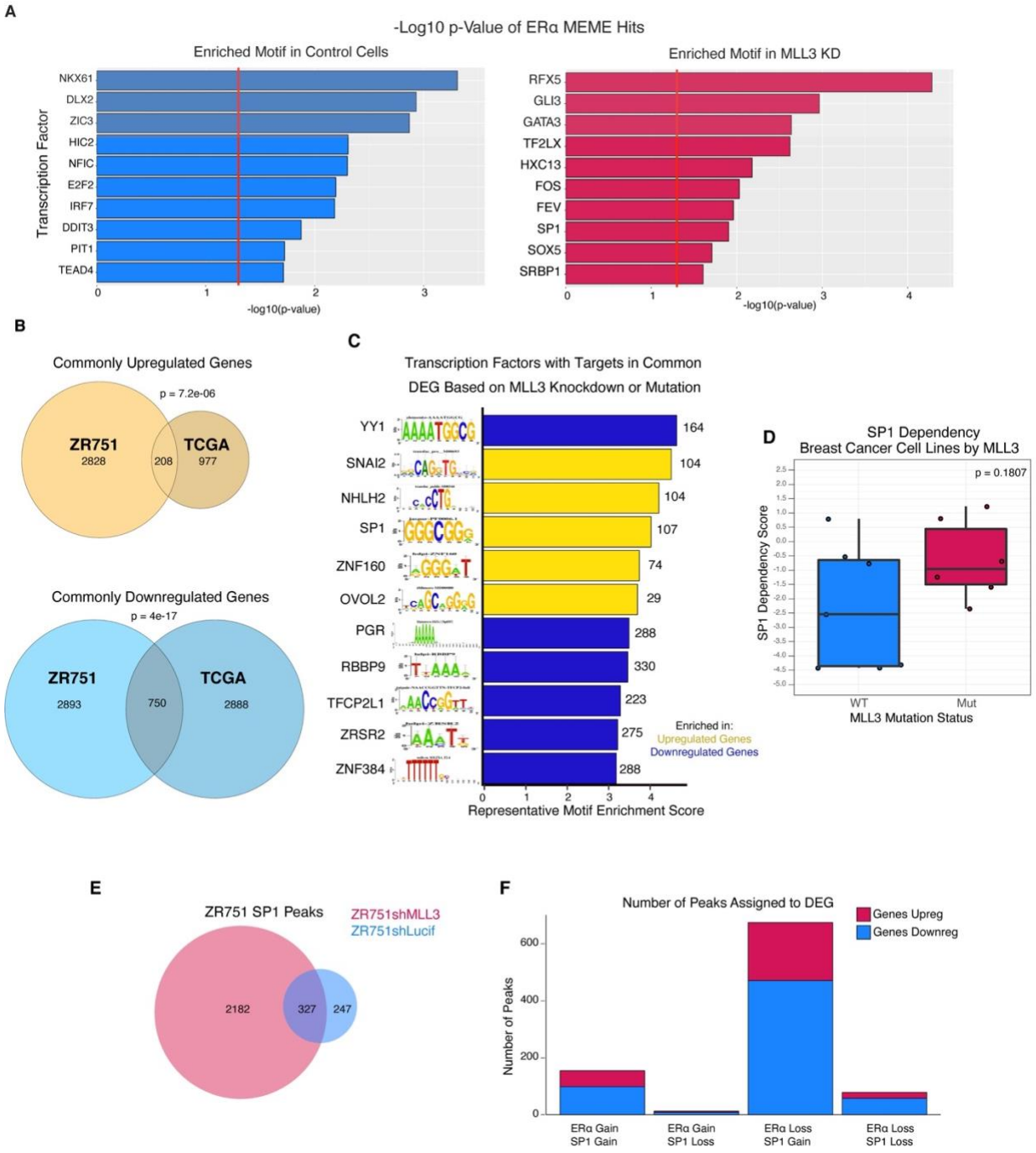


Figure 2-10. SP1 binding increases upon MLL3 KD in ER+ breast cancer cell line. (**a**) Representative enriched transcription factor motifs in ERα ChIP-seq samples by MEME analysis. (n = 2 biological ChIP-seq replicates per experiment) KD = knockdown. (**b**) Venn diagram of upregulated genes in the ZR751shLucif vs ZR751shMLL3 analysis as well as in the TCGA ER+ luminal breast cancer MLL3 WT vs MLL3 mutant analysis. Fisher's test, $p = 7.2 \times 10^{-06}$. Venn diagram of downregulated genes in the ZR751shLucif vs ZR751shMLL3 analysis as well as in the TCGA ER+ luminal breast cancer MLL3 WT vs MLL3 mutant analysis. Fisher's test, $p = 4 \times 10^{-17}$. (n = 2 biological RNA-seq replicates per experiment) (**c**) Representative enriched transcription factor motifs in the common differentially expressed genes between TCGA MLL3 WT vs. mutants and ZR751 control and MLL3 KD cells, by iRegulon analysis in Cytoscape. (n = 2 biological RNA-seq replicates per experiment) DEG = differentially expressed genes. (**d**) SP1 dependency scores of ER+ luminal breast cancer cell lines from the DEMETER tool where a lower score denotes a higher dependency. The center line signifies the median, box limits signify upper and lower quartiles, and whiskers signify the 1.5x interquartile range. All data points are shown as dots. Wilcoxon Rank Sum test, p = 0.1807 (n = 13 ER+ luminal breast cancer cell lines) (**e**) Venn diagram showing number of SP1 ChIP-seq peaks in ZR751 control and MLL3 KD cell lines. (n = 2 biological ChIP-seq replicates per experiment) (**f**) DEG upon MLL3 KD in ZR751 cells grouped into four categories based on the number of ERα and SP1 ChIP-seq peaks assigned to each gene in the control and MLL3 KD conditions. (n = 2 biological ChIP-seq replicates per experiment)

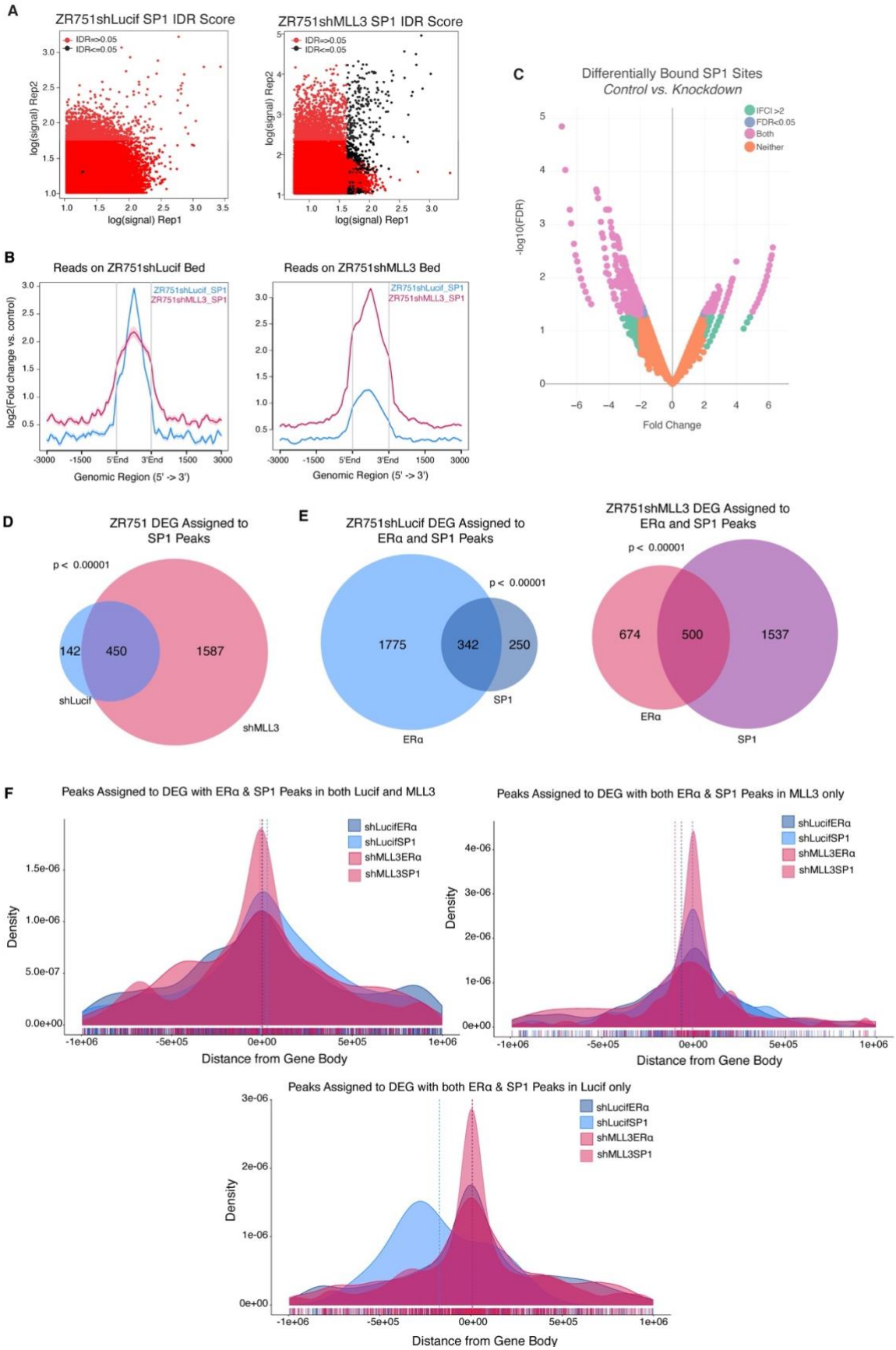


Figure 2-11. Supplement to Figure 2-10. (**a**) IDR scatterplot of log(signal) of ZR751 SP1 ChIP replicates. Red dots signify peaks that have an IDR score greater than the chosen threshold of 0.05 for SP1. Black dots signify peaks that have an IDR score of less than or equal to the chosen threshold. IDR = Irreproducible Discovery Rate. (**b**) Histograms of merged SP1 ChIP-seq peaks in ZR751 cells plotted on the control and MLL3 KD genomic locations. (n = 2 biological ChIP-seq replicates per experiment) (**c**) Differentially bound SP1 sites upon MLL3 knockdown in ZR751. Fold change and -log10(FDR) are plotted for the sites found by DiffBind to be differentially bound between ZR751shLucif and ZR751shMLL3. Differentially bound H3K4me1 (left) and ERα

(right) sites upon MLL3 knockdown in ZR751. Fold change and -log10(FDR) are plotted for the sites found by DiffBind to be differentially bound between ZR751shLucif and ZR751shMLL3. Genomic sites that have an absolute value fold change of 2 or greater are green if they do not have an FDR of less than 0.05, and pink if they do. Sites that have an FDR of less than 0.05 but do not have an absolute fold change greater than 2 are blue. Sites with an FDR of more than 0.05 and an absolute fold change of less than 2 are orange. Positive fold enrichment indicates higher amounts of binding in ZR751shLucif compared to ZR751shMLL3. FC = fold change. (**d**) Venn diagram showing the number of ZR751 differentially expressed genes assigned to ChIP-seq SP1 peaks in ZR751shLucif and ZR751shMLL3 cells. Chi-square test of independence X^2 (1, N = 6263) = 563.4442, p = < 0.00001 (n = 2 biological replicates per experiment for both RNA-seq and ChIP-seq) DEG = differentially expressed genes. (**e**) Venn diagram showing the overlap of ERα and SP1 peak-to-DEG assignments in ZR751shLucif and ZR751shMLL3 cells. Chi-square test of independence ZR751shLucif X^2 (1, N = 6263) = 167.8586, p = < 0.00001, ZR751shMLL3 X^2 (1, N = 6263) = 66.6957, p = < 0.00001 (n = 2 biological replicates per experiment for both RNA-seq and ChIP-seq) (**f**) Density histograms showing the distance in base pairs of the SP1 and ERα peaks in both cell lines from the gene body of the respective assigned DEGs that they regulate. (n = 2 biological replicates per experiment for both RNA-seq and ChIP-seq)

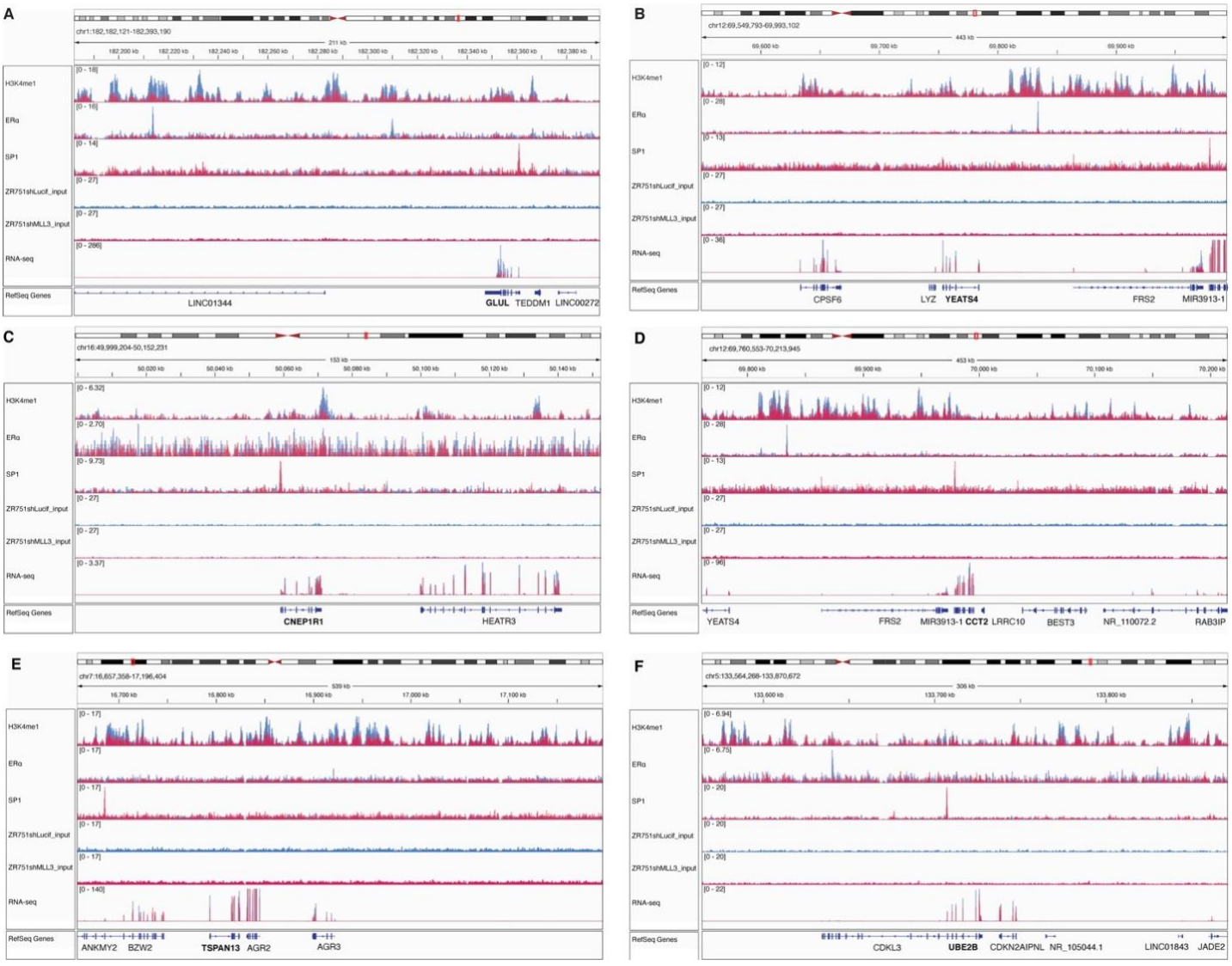


Figure 2-12. IGV Genome Browser Snapshots of H3K4me1, ERα, and SP1 binding. (**a**) IGV Genome Browser snapshot of GLUL, which has decreased gene expression, a higher number of H3K4me1 peaks, less ERα peaks, and more SP1 peaks assigned upon MLL3 KD. (**b**) IGV Genome Browser snapshot of YEATS4, which has decreased gene expression, a higher number of H3K4me1 peaks, less ERα peaks, and the same number of SP1 peaks assigned upon MLL3 KD. (**c**) IGV Genome Browser snapshot of CENP1R1, which has decreased gene expression, more H3K4me1 peaks, less ERα peaks, and more SP1 peaks assigned upon MLL3 KD. (**d**) IGV Genome Browser snapshot of CCT2, which has decreased gene expression, more H3K4me1 peaks, less ERα peaks, and same number of SP1 peaks assigned upon MLL3 KD. (**e**) IGV Genome Browser snapshot of TSPAN13, which has decreased gene expression, less ERα peaks, and more SP1 peaks assigned upon MLL3 KD. (**f**) IGV Genome Browser snapshot of UBE2B, which has decreased gene expression, less ERα peaks, and more SP1 peaks assigned upon MLL3 KD.

Discussion

Over 40,000 women will die from breast cancer this year[148], and over 50% of those deaths will be due to ER+ breast cancer.[149] ERα drives the growth of ER+ breast cancers and is the target of endocrine therapy. In randomized clinical trials, endocrine therapies have effectively prevented cancer recurrence.[150] However, approximately 20% of ER+ breast cancers will present with de novo resistance[151,152], and many patients with early stage disease will recur after endocrine therapy.[153] The majority of patients with metastatic ER+ breast cancer have or develop endocrine resistance, and thus both *de novo* and acquired resistance to endocrine therapy present significant hurdles to the effective treatment of breast cancer. The mechanisms underlying both *de novo* and acquired endocrine resistance remain incompletely understood, however. Somatic mutations such as *ERBB2* amplification[154,155], ligand binding domain ERα mutations[60], and co-amplification of FGFR1 and CCND1 have been associated with endocrine resistance, but these mechanisms do not explain even the majority of endocrine resistance.

Interestingly, both preclinical and clinical observations suggest that the majority of endocrine-resistant tumors remain dependent on ERα. Most ER+ breast tumors retain protein expression of ERα after developing resistance[156,157] . Furthermore, about 30% of patients that develop resistance to aromatase inhibition (AI) respond to fulvestrant[158,159], and in first line therapy for metastatic disease, the combination of fulvestrant and AI is superior to AI alone.[160,161] Importantly, ERα binds to different genomic locations in tumors will good vs. poor outcomes, and studies show that ERα binds to different locations in endocrine-sensitive and endocrine-resistant cell lines[74], or in cell lines expressing ERα with mutations in the ligand binding domain (LBD). These results suggest that dysfunction of the regulatory mechanisms governing ERα genomic binding contribute to the development of endocrine resistant ER+ breast cancer, and we hypothesized that chromatin remodeling enzymes that can regulate the ERα genomic landscape may contribute to endocrine resistance.

MLL3, a histone monomethylase that is known to interact with nuclear hormone receptors such as ERα, is recurrently mutated in many cancers. *MLL3* is the 6th most mutated gene in ER+ breast cancer. Indeed, *MLL3* is altered in 9% of ER+ breast cancer patients in the TCGA dataset and 8.5% in the AACR GENIE dataset[95,103]. In the work above, we identify mutation of *MLL3* as a potential common cause of endocrine resistance in ER+ breast cancer. We demonstrate that the mutation pattern of *MLL3* in breast cancer is most consistent with a haploinsufficient tumor suppressor.

Modeling loss of MLL3 function using shRNA knockdown in the ER+ PIK3CA-wildtype breast cancer cell line ZR751, we found that knockdown of *MLL3* led to a major loss of H3K4me1 marked peaks across the genome. This loss was associated with a major shift in ERα binding, including to genes in signatures associated with endocrine resistance. Indeed, loss of MLL3 expression increased resistance to endocrine therapy. The loss of MLL3 function was not only associated with massive changes to the H3K4me1-marked enhancer landscape and to ERα genomic binding sites, but also significant changes in gene expression. Assigning peaks to DEG, we were able to identify two groups of genes that were altered upon loss of MLL3. Module 1 genes demonstrate that when functional MLL3 is lost, a substantial amount of H3K4me1 marks is also lost, accompanied by a

loss of ERα at those genomic locations. However, the canonical ERα target genes controlled by those lost peaks are upregulated. Module 2 genes demonstrate that a loss of functioning MLL3 results in a compensatory H3K4 methyltransferase activity that is accompanied by a change in number of regulatory ERα peaks and decreased gene expression. These two ERα-H3K4me1-gene modules allow breast cancer cells with a loss in functional MLL3 to increase the expression of canonical ERα targets, while also deploying transcriptional programs shown to mediate aggressive tumor behaviors.

The changes in gene expression attributed to changes in ERα regulation could be due to changes in the milieu of regulatory factors coordinating the binding of ERα to the genome. Motif analysis of both our ChIP-seq and RNA-seq data suggested that an SP1 transcriptional program might be activated upon inactivation of MLL3, global reduction of H3K4me1, and re-organization of ERα genomic binding sites. Indeed, loss of MLL3 was associated with a massive increase in SP1 peaks. Strikingly, the largest group of differentially expressed genes with both ERα and SP1 peaks are those that have an ERα peak loss and a SP1 peak gain upon MLL3 KD. This suggests that the reorganization of the ERα-driven transcriptome caused by loss of MLL3 results in a substantial fraction of genes being driven by SP1. Future studies will seek to identify the mechanism that unleashes SP1 in MLL3 mutant cells and its contribution to aggressive tumor behavior.

MLL3 is a member of multi-protein epigenetic complexes, ASCOM and COMPASS[141,142]. Both ASCOM and COMPASS complexes interact with nuclear hormone receptors, including ERα. Importantly, MLL3 is not the only histone methyltransferase that can be a component of these complexes. MLL4 can also serve as the histone methyltransferase in ASCOM and COMPASS. However, each individual complex contains either MLL3 or MLL4, and the difference in their function is not well understood. Both MLL3 and MLL4 have been shown to help regulate ERα transcriptional activity, for targets such as *EZH2*, *HOX* genes, and *HOTAIR*.[142,162–164] Interestingly, *MLL4* is also recurrently mutated in many cancers, such as lung adenocarcinoma and bladder cancer[108], but is NOT recurrently mutated in breast cancer. It is thus possible that loss of MLL3, and its replacement with MLL4 in ASCOM complexes leads to unique histone monomethylation locations and changes in regulatory partners, like SP1, altering the transcriptional program and driving endocrine resistance. Interestingly, MLL4 has been shown to be regulated by AKT1, leading to ERα-driven therapeutic resistance to PIK3CA inhibition. Targeted treatment of *PIK3CA*-mutant breast cancers with anti-PIK3CA therapy is known to lead to a compensatory increase in ER-dependent transcription and shift in ERα genomic binding that limits therapeutic efficacy.[165,166] These changes are dependent on MLL4, suggesting that increased MLL4 function can lead to a shift in the genomic location of ERα binding that may contribute to therapeutic resistance. It has been shown that in the TCGA breast cancer dataset, increased MLL4 mRNA expression leads to shorter overall survival ($p = 0.0398$).[167] Unsurprisingly, MLL3 KD in ZR751 cells upregulated expression of MLL4, albeit not to statistical significance, and in TCGA ER+ luminal breast cancer studies *MLL3*-mutant cases had significantly higher expression of *MLL4* ($p = 0.01176938$, ANOVA of multivariate linear regression).

Curiously, one recent paper found that in MCF7 cells loss of MLL3 leads to decreased proliferation, decreased ERα transcriptional activity, and increased growth in estrogen-absent media[96]. However, MCF7 cells have a *PIK3CA* mutation, while ZR751 are wild-type for *PIK3CA*. Thus, in MCF7 cells PIK3CA may be restraining MLL4 and ERα through activated AKT1, while MLL4 is free to activate transcription in ZR751. Future studies will focus on the interplay of MLL3, MLL4, and the PIK3CA signaling pathway. Synergies between loss of MLL3, inhibition of PIK3CA, and anti-estrogen therapies may provide new avenues for therapy of endocrine resistant tumors. Furthermore, we have established that MLL3 is a haploinsufficient tumor suppressor, which suggests the possibility that loss of the remaining allele of *MLL3* could be detrimental to cancer cell survival. MLL3 is an enzyme and is thus a potential target for small molecule inhibitors. As such, subsequent studies will focus on the

possibility that MLL3 and/or MLL4 may represent a therapeutic target in *MLL3*-mutant breast cancers, as well as present a mechanism for reversal of endocrine resistance.

CHAPTER III

FULVESTRANT/PALBOCICLIB RESISTANCE IN ER+ BREAST CANCER

Summary

In collaboration with Valerie Jansen of the Carlos Arteaga laboratory, I investigated the transcriptional patterns associated with resistance to a common combinatorial treatment aimed at targeting the abnormal activation of the CDK4/6/cyclin D complex that occurs in many cases of ETR. This combinatorial treatment includes Fulvestrant, a common SERD, and Palbociclib, a CDK4/6 inhibitor. We hypothesized that there is a Fulvestrant/Palbociclib resistance transcriptional signature that would allow for the identification of patients unlikely to respond to treatment. I analyzed RNA-seq data from MCF7 parental and MCF7 Fulvestrant/Palbociclib – resistant cells treated with and without Fulvestrant/Palbociclib. We found that while pieces of the RB pathway show differential expression in resistant cells, treatment with FulvPalb altered expression of even more genes within the RB pathway. This work reveals a possible diagnostic indicator of likelihood to show resistance to FulvPalb treatment in ER+ breast cancer. This work is currently unpublished.

Introduction

While endocrine therapy is the standard of care for patients with metastatic ER+ HER2- breast cancer, many tumors present with de novo resistance. Those that do respond initially will eventually become resistant to endocrine therapy. To better manage this subset of breast cancers, endocrine therapy is often combined with other drugs in the clinic. Combination therapies have enabled longer progression-free times and overall survival times in many clinical settings. One target of therapy used in combination with endocrine therapies is CDK4/6.

Cyclin-dependent kinases 4 and 6 (CDK4/6) are activated in complex with their cyclins because of signaling through growth factor and hormone receptors. One of these participatory hormone receptors is ERα, which can signal through the MAPK/Ras/Raf, Wnt/β-catenin, and PI3K/AKT/mTOR pathways[168] to reach CDK4/6. After the complex of CDK4/6/cyclin D is activated, it phosphorylates and inactivates tumor suppressor retinoblastoma protein (pRb), and E2F transcription factors are free to start the cell cycle transition from G1-phase to S-phase. This makes CDK4/6 major contributors to tumor growth in ER+ breast cancer.

Importantly, several pieces of the above pathway are recurrently mutated in breast cancer. This abnormal activation of the CDK4/6/cyclin D complex is implicated in endocrine therapy resistance in breast cancer cell lines,[169,170] and accordingly CDK4/6 inhibition has shown to be effective in killing breast cancer cells with acquired endocrine therapy resistance[170,171]. With this plentiful evidence, CDK4/6 inhibitors such as palbociclib have had several clinical trials. A series of randomized Phase II and III clinical trials showed improved progression-free survival (PFS) in advanced ER+ breast cancer with combination CDK4/6 inhibitor and anti-estrogen therapy compared to anti-estrogen therapy alone.[172,173,174] The selective CDK4/6 inhibitors are approved by the FDA for combinatorial use with anti-estrogen therapy in ER+ HER2- breast cancer as both frontline and ETR treatment. However as in most cancer therapeutics, resistance to this combinatorial treatment does inevitably arise. Identification of biomarkers to predict resistance are needed for better treatment planning, as well as identification of resistance mechanisms so that further therapies can be engineered.

This project was undertaken to investigate the role of combinatorial treatment with Fulvestrant and Palbociclib on ERα-mediated transcriptional activity. The goals of this experiment are to identify differentially expressed ERα target genes causally associated with drug resistance. To accomplish this, we compared the expression of known ERα targets between MCF7 parental and MCF7-FulvPalbResistant (FPR) cells with and without the combinatorial treatment of Fulvestrant + Palbociclib. We hypothesized that maintenance of ERα transcriptional programs is associated with resistance.

Methods and Materials

RNA-Seq

MCF7 parental and MCF7-Fulv+palbo-resistant cells, treated with and without drugs (fulv+palbo, DMSO) were utilized for RNA collection. Total RNA was isolated from cells using the Maxwell® 16 Total RNA Purification Kit, in biological triplicate for each condition. Stranded mRNA libraries were built in the VANTAGE core using the Illumina Tru-seq RNA sample prep kit, following manufacturer's protocols. Libraries were sequenced using paired-end 75bp reads to a depth of 45e6 pairs on a HiSeq 3000 in the VANTAGE core. Sequencing quality was assessed with FastQC and reads were trimmed to remove adapters and low quality sequencing. Reads were aligned to human genome version 19 with TopHat2, a splice-aware aligner, using UCSC gene models (version 19) as a guide. Expression levels were quantified as counts using FeatureCounts. Unsupervised hierarchical clustering and principal components analysis was performed in R. Supervised analyses, such as differential ER target gene expression, were determined using Bioconductor software package DESeq2, which utilizes a model based on negative binomial distribution and correction for multiple hypothesis testing using Benjamini-Hochberg method. Pathway analysis using GSVA in R and MSigDB terms were performed to identify enrichment of pathways and biological processes.

Results

Upon creating at PCA plot for the RNA-seq samples (**3-1**), we noticed that the resistance phenotype grouped together closely regardless of whether they were treated with DMSO or the Fulv/Palb combination. The parental MCF7 cells however showed great divergence depending upon which treatment they received. The sample distance heatmap (**3-2**), based on the regularized log transformation of the count data in DEseq2, recapitulated this pattern.

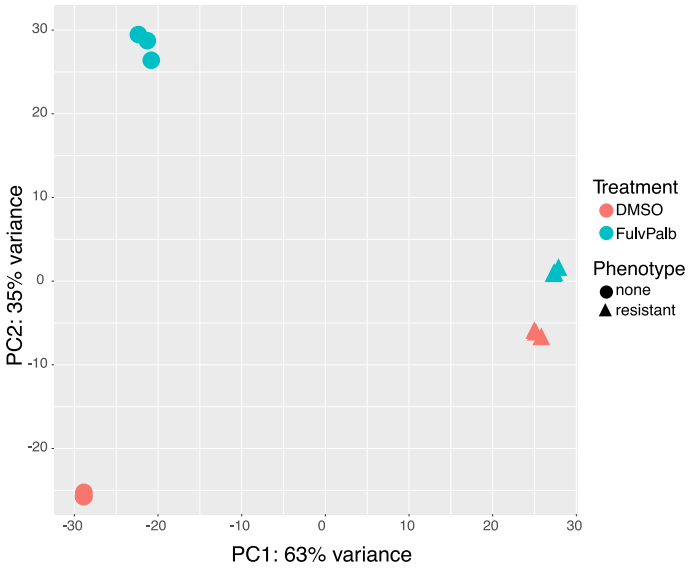


Figure 3-1. PCA Plot of RNA-seq Samples.

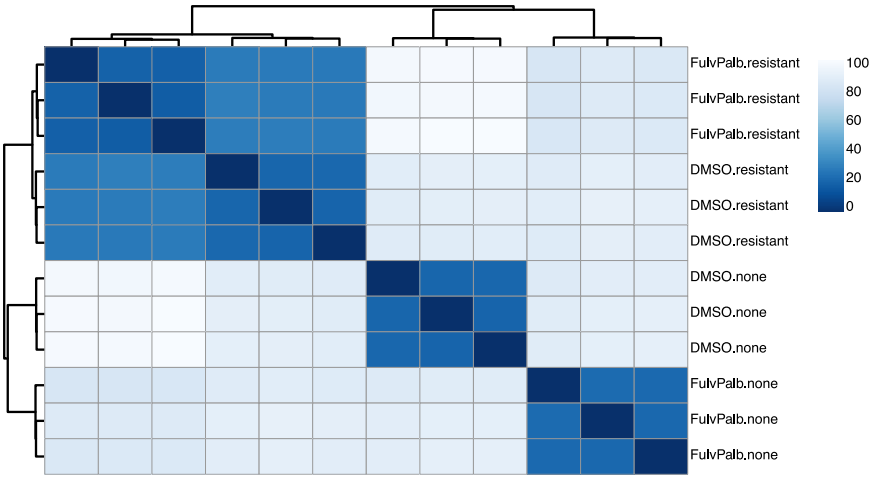


Figure 3-2. RNA-seq Sample Distance Heatmap. This is an overview over similarities and dissimilarities between samples with respect to the expression counts, including hierarchical clustering based on the sample distances.

Because of the divergence of the sample based on Fulv/Palb resistance, we decided to focus our comparison on treatment response phenotype – in untreated and treated samples. First we found differentially expressed genes in the DMSO treated group based on whether the cell line used was sensitive or resistant to Fulv/Palb treatment. Utilizing a cutoff of BH-adjusted $p < 0.01$, there were 10,336 DEG. Repeating the analysis for the Fulv/Palb treated group between parental and resistant phenotypes with the same cutoff, 10,319 genes were differentially expressed. When comparing the two DEG lists, there was an overlap of 7,526 genes. This portion accounted for nearly 73% of each full DEG set, lending credibility to the hypothesis that a transcriptional program for Fulv/Palb resistance persists in untreated and treated conditions alike.

To probe what the molecular underpinnings of Fulv/Palb resistance consist of, we utilized the GSVA R package with the DMSO-treated and Fulv/Palb-treated groups separately. The entirety of the MSigDB catalog was employed so that we could gain insights about any big-picture pathways that may be at play in creating Fulv/Palb resistance in breast cancer cells. The top 200 differentially enriched terms between parental and resistant cells (**3-3** top 40 terms of each group; DMSO-treated on top, Fulv/Palb-treated on bottom) revealed 40 common differentially enriched terms. Among these were "GO_REGULATION_OF_PHOSPHATIDYLINOSITOL_3_KINASE_ACTIVITY", and "MASRI_RESISTANCE_TO_TAMOXIFEN_AND_AROMATASE_INHIBITORS_UP". In addition, there were several terms centered around TP53 activity interspersed throughout the two groups, separately. Upon inspection of the overlapping top 40 MSigDB terms, we noticed SOX9 as a recurring gene member. Given the role of SOX9 and SOX2 in regulating stem and progenitor cells, we examined the expression of these two transcription factors (**Table 3-1**). While Sox2 was not differentially expressed between parental and FulvPalb-resistant, Sox9 was significantly upregulated in the FulvPalb-resistant cells, in both treated and untreated cells (**3-4**).

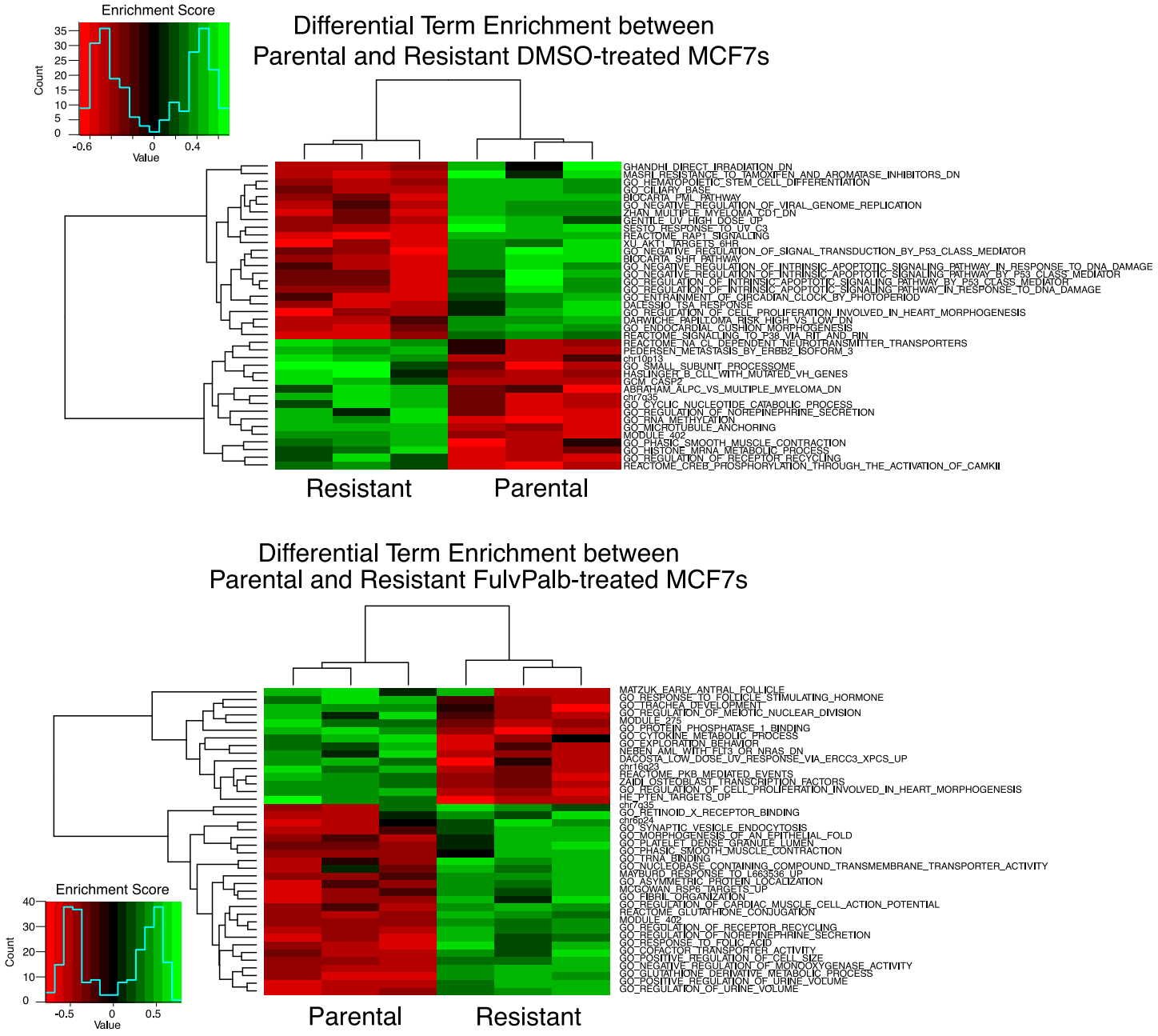


Figure 3-3. Heatmap of Top MSigDB Terms from GSVA.

	Mean Expression (normalized counts)	Log2(Fold Change)	Raw p-value	Adjusted p-value (Benjamini-Hochberg)	Conclusion
Sox2					
DMSO	2.020693	0.5614652	0.1882725	0.2658239	No
FulvPalb	2.097788	0.1916897	0.6663436	0.7521814	No
Sox9					
DMSO	50.41421262	1.558218637	1.10E-07	3.17E-07	Yes, Resis Higher
FulvPalb	114.5154674	0.915046181	2.55E-05	6.46E-05	Yes, Resis Higher

Table 3-1. Sox2 and Sox9 expression between Parental and Resistant MCF7 Cells.

Figure 3-4. SOX9 counts.

In addition to increased SOX9 expression in cells resistant to Fulv/Palb treatment, RB1 expression was downregulated. This is interesting as SOX2 and/or SOX9 upregulation is sometimes accompanied by RB1 downregulation in clinically advanced tumors.[175,176] In addition, RB1 loss presents a possibly-targetable vulnerability for tumors resistant to first-line therapy.[177] We postulated that expression of other genes downstream in the RB1 pathway would be affected as well. In particular we were interested in E2Fs, which are regulated by pRB and control the expression of cellular proliferation genes, and one of its targets ARF, a modulator of MDM2-mediated degradation of p53.[178] Indeed with the resistance phenotype, regardless of treatment-type, MDM2 is upregulated. Interestingly under FulvPalb treatment, but not DMSO treatment, resistant cells have upregulated TP53, MDM4, and E2Fs (**3-5**).

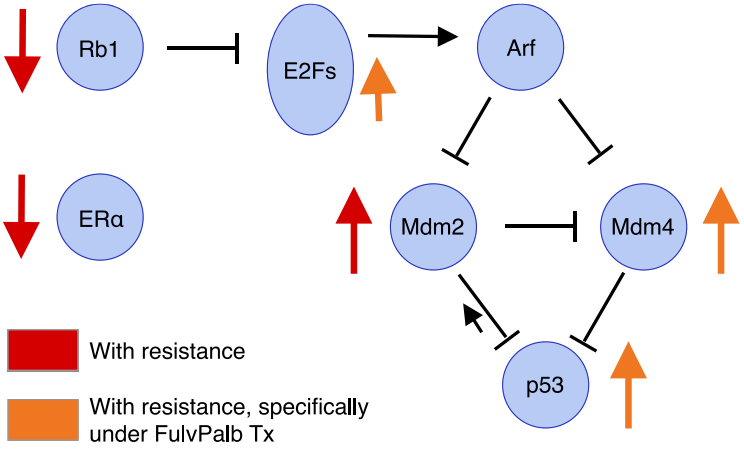


Figure 3-5. RB Signature Schematic for FulvPalb Resistance.

To further interrogate how the downstream RB pathway is involved in FulvPalb resistance in ER+ breast cancer, we examined expression of a 20-gene RB-loss signature from the Perou group.[179] We observed that the RB-loss signature is generally downregulated in parental cells when treated with FulvPalb, but upregulated in the resistant cells treated when treated with FulvPalb (**3-6**). Furthermore, parental cells have higher expression of the RB20 signature genes than resistant cells when treated with DMSO; the reverse is true when cells are treated with FulvPalb.

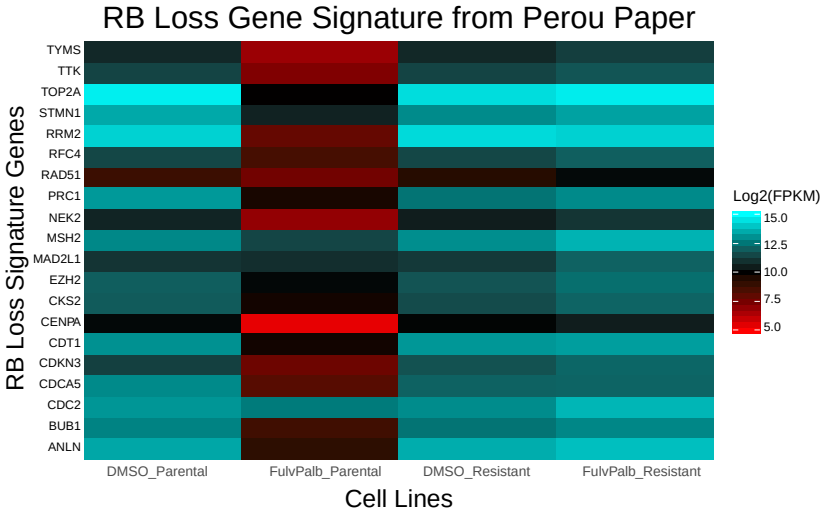


Figure 3-6. RB 20 Gene Signature from the Perou et al.

Discussion

The findings presented here suggest that breast cancer cells that have become resistant to Fulv/Palb combination treatment employ a transcriptional signature to survive and proliferate that persists under both control (DMSO) and Fulv/Palb treatment. This transcriptional signature featured TP53 activity, as well as SOX9 upregulation in resistance and upon FulvPalb treatment. Because of the SOX9 upregulation we investigated whether RB loss might be occurring in the resistant setting. In fact 14 out of the 20 genes in the RB 20 Gene Signature from the Perou Group were included in our FulvPalb resistance signature, regardless of

whether cells were treated with DMSO or FulvPalb. The remaining six genes in the RB 20 Gene Signature (TYMS, CDT1, RRM2, CDKN3, ANLN, and MSH2) are contained in the DEG found between parental and resistant cells treated with FulvPalb, but not in the same comparison for DMSO-treated cells. Upon examination of the direction of effect we saw that the RB 20 signature genes are downregulated upon FulvPalb treatment in parental cells, but upregulated in resistant cells upon FulvPalb treatment. High expression of this 20 gene signature was found to correlate with a good response to neoadjuvant paclitaxel and fluorouracil-doxorubicin-cyclophosphamide in breast cancer patients[179,180], thus it is possible prior FulvPalb treatment may make RB1-low breast cancers increasingly susceptible to standard chemotherapeutics. This may be more effective in RB1-low breast cancer cells that acquire resistance to FulvPalb, as they appear to have elevated levels of the downstream RB1 pathway and are able to cycle through G1 to S phase. This notion is bolstered by the significantly upregulated expression of E2F1, E2F2 and E2F3 in the resistant cells compared to the parental cells with Fulvestrant treatment, which occurs before entry to S-phase of the cell cycle.[181] Of note PI3K inhibitors may also be of use in FulvPalb resistant breast cancers, as we saw enrichment for GSEA term "GO_REGULATION_OF_PHOSPHATIDYLINOSITOL_3_KINASE_ACTIVITY" in resistant cells. In support of this hypothesis, new PIK3CA driver mutations were found in patients from the PALOMA-3 randomized phase III trial that progressed[67], and another study showed that combination therapy simultaneously targeting PI3K, CDK4/6, and ER prevented and/or delayed the onset of resistance in breast cancer models.[182]

CHAPTER IV

ASSOCIATION OF FGFR1 WITH ERA MAINTAINS LIGAND-INDEPENDENT ER TRANSCRIPTION AND MEDIATES RESISTANCE TO ESTROGEN DEPRIVATION IN ER + BREAST CANCER

This section is a paper published in *Clinical Cancer Research* as "Association of FGFR1 with ERα Maintains Ligand-Independent ER Transcription and Mediates Resistance to Estrogen Deprivation in ER + Breast Cancer" Luigi Formisano, Kimberly M Stauffer, ... Thomas Stricker, Carlos L Arteaga.

Summary

In collaboration with Luigi Formisano of the Carlos Arteaga laboratory, I investigated the FGFR1 and ERα binding patterns and transcriptional activity in ER$^+$/*FGFR1*–amplified breast cancer cells treated with fulvestrant and/or lucitanib. We hypothesized that FGFR1 and ERα physically interact to take advantage of FGFR1 amplification to allow cancer cell survival and proliferation despite endocrine therapy. We found that FGFR1 amplification sustains estrogen-independent breast tumor growth through cooperation between FGFR1 and ERα in binding to ERα target genes. This binding was enriched upon treatment with FGF3 ligand and reduced upon treatment with lucitanib and/or fulvestrant. This work supports use of combination ERα and FGFR antagonists for patients with ER+, FGFR1-amplified breast cancer.

Introduction

Amplification of the chromosomal region 8p11-12, the genomic location of *FGFR1*, has been reported in breast, ovarian, bladder, lung and oral squamous cancers, and in rhabdomyosarcoma.[183–189] *FGFR1* amplification occurs in approximately 10% of patients with estrogen receptor–positive (ER$^+$)/HER2$^-$ breast cancer, where it is associated with early relapse following adjuvant tamoxifen therapy and with poor survival.[98] Blockade of FGFR1 signaling by pharmacologic or genetic approaches in human breast cancer cells harboring *FGFR1* amplification leads to decreased cell growth and survival, suggesting *FGFR1* gene amplification is a surrogate of cancer cell dependence on aberrant FGFR activity.[98]

FGFRs belong to the family of receptor tyrosine kinases (RTK) that consist of an extracellular ligand-binding domain linked to an intracellular catalytic protein kinase core via a single-pass transmembrane domain (TMD).[190] Binding of FGF ligands induces receptor dimerization, activation of the kinase domain, and phosphorylation of C-terminal tyrosines to which adaptor proteins dock, followed by activation of signal transduction pathways, including PI3K/AKT, RAS/RAF/MEK/ERK, phospholipase Cγ (PLCγ) and STATs.[191] In addition, there is strong evidence that FGFRs traffic to the nucleus, where they may function in a different manner to classic transmembrane RTKs.[192] For example, nuclear FGFR3 has been shown in the nucleus of malignant and nonmalignant breast epithelial cells.[193] A nuclear interaction of FGFR2, STAT5, and progesterone receptor (PR), associated with PR/STAT5–regulated gene expression and breast cancer progression was also reported.[194] Other studies have reported nuclear localization and a nucleus-specific function of FGFR1 in nonmammary cells.[195–197,198(p1)] Medulloblastoma cells transfected with FGFR1-eGFP and evaluated by immunofluorescence (IF) have shown FGFR1 is associated with cell membranes, cytosol, and nuclear compartments.[198(p1)] Substitution of the atypical TMD of FGFR1 (β-sheet containing polar amino acids) with the typical TMD of FGFR4 (α-helical, hydrophobic) prevents the nuclear localization of FGFR1.[199] Inability of both the full-length and cleaved forms of FGFR1 to localize to the nucleus results in reduced migration and invasiveness of cancer cells.[196,197] Finally, Chromatin immunoprecipitation sequencing (ChIP-seq) studies revealed that FGFR1 binds nuclear transcription factors involved in neural and muscle development.[200(p1)]

Amplification of the *FGF3/4/19* ligand genes on chromosome 11q12-14 occurs in approximately 15% of human breast cancers[201,202]. Notably, one third of *FGFR1*-amplified tumors also harbor amplification of *CCND1, FGF3, FGF4*, and *FGF19*.[203] This coamplification has also been associated with resistance to estrogen deprivation in ER$^+$ breast cancer and poor patient outcome[203], suggesting the possibility of ligand–receptor cooperativity.

Herein, we investigated mechanisms by which FGFR1 amplification confers resistance to antiestrogens in ER+ breast cancer. In a cohort of patients with ER+ breast cancer treated with the aromatase inhibitor letrozole, we observed that tumors with FGFR1 amplification maintained their proliferation despite drug-induced estrogen deprivation and exhibited nuclear localization of FGFR1. Estrogen deprivation also resulted in an increase of total and nuclear FGFR1 as well as FGF3/4/19 ligand expression in ER+/FGFR1–amplified breast cancer cells. FGFR1 coupled with ERα to drive estrogen-independent transcription of ERα-responsive genes. The association of FGFR1 with ERα was inhibited upon transfection with a kinase-dead FGFR1 mutant and by pharmacologic inhibition of FGFR1. Finally, combined inhibition of FGFR1 and ERα with fulvestrant and lucitanib reduced the association of FGFR1 and ERα and growth of ER+/FGFR1–amplified patient-derived xenografts (PDX). We propose a physical interaction between FGFR1 and ERα provides a mechanistic explanation for how FGFR1 amplification contributes to resistance to endocrine therapy in ER+ breast cancer.

Methods and Materials

Clinical trial and tumor biopsies

Tumor samples were obtained from patients with stage I–III operable ER+/HER2− breast cancer enrolled in a clinical trial of the aromatase inhibitor letrozole administered for 2 weeks prior to surgery (NCT00651976).[204] Patients provided written informed consent according to a protocol approved by the Vanderbilt-Ingram Cancer Center Institutional Review Board. Intraoperative biopsies or surgical specimens, snap-frozen in liquid nitrogen and formalin-fixed paraffin-embedded (FFPE), were obtained from each patient's tumor. IHC was conducted in both the pretreatment biopsy and in the posttreatment surgical biopsy of both tumors for Ki67 (Dako #M7240), ER (Santa Cruz Biotechnology #sc542) and PR (Dako #M3569). IHC for ER and PR was conducted according to methods reported elsewhere.[205] FFPE tumor sections were scanned at ×100 magnification, and the area containing the highest number of positive cells was selected. Positive and negative tumor cells were manually counted at 400×; the percentage of positive cells was calculated with at least 1,000 viable cells. Ki67 IHC was scored by two independent pathologists (M.M. Estrada and J.M. Giltnane).

Cell lines

Cell lines were obtained from ATCC between 2014 and 2016 and maintained in DMEM/10% FBS (Gibco). Long-term estrogen-deprived (LTED) cells were generated upon long-term culture in phenol red-free IMEM/10% dextran charcoal–treated FBS [DCC-FBS; Hyclone, contains <0.0367 pmol/L 17β-estradiol (E2)] for 3 to 8 months until exponentially growing, hormone-independent cells emerged as described previously.[206] Cell lines were authenticated by ATCC prior to purchase by the STR method. Cell lines were not authenticated after purchase. Mycoplasma testing was conducted for each cell line before use. All experiments were performed less than 2 months after thawing early passage cells.

FGFR1 and CCND1 FISH

FGFR1 and CCND1 copy number was measured by FISH analysis in FFPE tumor sections (see Supplementary Methods).

Viral transduction

FGFR1 wild-type and GFP-expressing lentiviral constructs were generated in the pLX302 Gateway vector (Open Biosystems); FGFR1/TK– (K514M) pLX302 was created by site-directed mutagenesis by Genewiz. To generate stably transduced lines, 4 μg of the FGFR1, FGFR1/TK– (K514M), and GFP-pLX302 constructs was cotransfected with 3 μg psPAX2 (plasmid encoding gag, pol, rev, and Tat genes), and 1 μg pMD2G envelope plasmid (Sigma Aldrich) into 293FT cells using Lipofectamine 2000 (Thermo Fisher Scientific). 293FT growth media were changed 24 hours posttransfection; virus-containing supernatants were harvested 48 and 72 hours posttransfection, passed through a 0.45-μm filter, diluted 1:4, and applied to target cells with 8 μg/mL polybrene (Sigma Aldrich). Virus-producing cells were selected in 1 μg/mL puromycin.

Proximity ligation assay

Proximity ligation assay (PLA) was performed in cultured cells and in FFPE primary tumor sections to detect FGFR1/ERα localization using the Duolink Detection Kit (#DUO92101, Sigma; see Supplementary Methods).

Gene expression analyses

CAMA1 cells were plated in estrogen-free media and treated ± 100 ng/mL FGF3/19 (Sigma) for 6 hours. Cells were harvested and RNA was purified using the RNeasy Kit (Qiagen). cDNA was generated using High-Capacity cDNA Reverse Transcription Kits (Applied Biosystems), followed by analysis of ERα pathway genes using the Estrogen Receptor PCR Array (Qiagen, PAHS-005Z). RNA sequencing (RNA-seq) data (see "RNA-seq and cDNA library construction" below) were aligned to human genome version 19 using the splice-aware aligner TopHat (v2.0.9), and isoform level expression was quantified using cufflinks. Expression levels were normalized across the data set using cuffnorm. We compared genes upregulated in *FGFR1*-amplified versus *FGFR1* nonamplified cancers (≥2.0-fold, FDR-adjusted $P \leq 0.05$). These genes were entered into gene set enrichment analysis (GSEA) as a ranked list. Gene sets with an FDR of <0.01 were considered to be enriched in *FGFR1*-amplified versus nonamplified tumors (see Supplementary Methods).

ChIP/DNA sequencing

ChIP was done using CAMA1 cells plated in estrogen-free media ± 100 ng/mL FGF3 and treated with 2 μmol/L lucitanib, 1 μmol/L fulvestrant, or the combination. Cells were grown to 80% confluency, washed 3 times in ice-cold PBS, and then fixed for 10 minutes at room temperature using 7% formaldehyde, followed by quenching with 2.5 mol/L glycine. Cells were first lysed using Farnham lysis buffer and then with nuclei lysis buffer (50 mmol/L Tris-HCl pH 8.0, 10 mmol/L EDTA pH 8.0, 1% SDS). Chromatin was sonicated using a Covaris LE220 with the following conditions: 35 minutes at peak power 350, duty factor 15, 200 cycles/burst, and average power 52.5; 200 μL of the chromatin was saved for input. Sonicated chromatin was diluted using ChIP Dilution Buffer (50 mmol/L Tris-HCl pH 8.0, 0.167 mol/L NaCl, 1.1% Triton X-100, 0.11% sodium deoxycholate), RIPA-150, protease inhibitors, and sodium butyrate. ERα (sc-8002) and FGFR1 (ab10646) antibodies were linked to magnetic anti-mouse and anti-rabbit Dynabeads, respectively, and then incubated with chromatin for >12 hours at 4°C. Immunoprecipitates (IP) were washed with the following buffers (RIPA-150, RIPA-500, RIPA-LiCl, and TE Buffer) for 5 minutes each. Chromatin-IPs were eluted from the beads, treated with RNase A at 65°C with shaking for 4 hours to reverse crosslinks, followed by proteinase-K treatment at 55°C for 1 hour. Next, DNA was purified using phenol–chloroform extraction, followed by ethanol precipitation and subsequent quantification by Qubit. Standard Illumina ChIPseq Library Kits were used to build sequencing libraries. Libraries were sequenced at Vanderbilt Technologies for Advanced Genomics (VANTAGE) Core Resource as SR50 on a HiSeq3000. Each antibody pulldown and the corresponding matching input was sequenced. The resulting sequencing files were aligned to human genome version 19 by BWA (Burrows–Wheeler aligner). For each replicate, peaks were called comparing with matched input, using MACS14 and default settings. The intersection of peak calls from each

replicate was used to define the peak call set for each condition. Peaks were assigned to closest genes using annotatePeaks.pl in the HOMER analysis suite, and heatmaps were generated using ngs.plot.

ChIP/quantitative PCR

ChIP was performed in CAMA1 cells as described above. DNA was analyzed by real-time qPCR in triplicate with Sso Advanced SYBR Green Supermix (Bio-Rad) in a CFX qPCR machine (Bio-Rad). The fold enrichment of ChIP samples was calculated using the $2^{\Delta}C_t$ (threshold cycle) method. C_t values for ERα-ChIP and FGFR1-ChIP samples were normalized to input DNA C_t values, and then independently to respective negative control C_t values to account for antibody background. Primer sequences are listed in **Table 4-1**.

Genomic Region	Fwd primer	Rev primer
38477209:38479862	TGGGTGTCTCTTGCTTCGTC	CATGATGTGTGCTGGAGGGT
38477209:38479862	AACCTTCAGCCCAGGAATCG	ATCTGCACAGTGGGTCACAG
17271502:17274328	GCCCCGCATAAAGAAAGCAG	AGCAAAAGCCGCAGTAGAGT

Table 4-1. Primer sequences used for ChIP-qPCR.

RNA-seq and cDNA library construction

Core biopsies were flash frozen in liquid N2 and stored at −80°C until RNA extraction was performed as described elsewhere.[207] Total RNA was quantified using the Quant-iT RiboGreen RNA Assay Kit (Invitrogen) and normalized to 4 ng/μL; 200 ng of each sample was used for library preparation in an automated variant of the Illumina Tru Seq RNA Sample Preparation protocol (Revision A, 2010). This method uses oligo(dT) beads to select mRNA from the total RNA sample and is followed by heat fragmentation and cDNA synthesis from the RNA template. The resultant cDNA went through library preparation (end repair, base "A" addition, adapter ligation, and enrichment) using Broad Institute–designed indexed adapters for multiplexing. After enrichment, libraries were quantitated with qPCR using the KAPA Library Quantification Kit for Illumina Sequencing Platforms and pooled equimolarly. The entire process was performed in a 96-well format with all pipetting done by either the Agilent Bravo or PerkinElmer JANUS Mini liquid handlers.

Nonstranded Illumina RNA-seq

Pooled libraries were normalized to 2 nmol/L and denatured using 0.2 N NaOH prior to sequencing. Flowcell cluster amplification and sequencing were performed according to the manufacturer's protocol using either the HiSeq 2000 v3 or HiSeq 2500. Each run was a 76-bp, paired-end run with an eight-base index barcode. Data were analyzed using the Broad Picard Pipeline, which includes demultiplexing and data aggregation. TopHat spliced aligner software was used to map sequencing reads and to generate a BAM file for each tumor.[208] RNA-seq GCT files were generated from BAM files using RNA-SeQC.[209]

Xenograft studies

These studies were approved and performed in accordance with the Vanderbilt Institutional Animal Care and Use Committee. We used two ER+/HER2−/FGFR1–amplified PDXs. PDX T272 (XenTech) required estrogen supplementation in the drinking water (8.5 mg/L estrogen) to grow as tumors in female athymic nude mice (Envigo). The second PDX, TM00368 (The Jackson Laboratory), was implanted in female ovariectomized SCID/beige mice (The Jackson Laboratory) supplemented with a subcutaneous 21-day release, 0.25-mg 17β-estradiol pellet (Innovative Research of America). Tumors were serially transplanted in athymic or SCID/beige mice under general anesthesia. When xenografts reached a volume ≥200 mm3, mice were randomized to treatment with vehicle, lucitanib (10 mg/kg/day orally for T272 or 7 mg/kg/day orally for TM00368), fulvestrant (5 mg/week s.c.) or both drugs (n = 10 per group for T272 and n = 8 per group for TM00368). Tumor diameters were measured using calipers twice a week, and volume in mm3 calculated with the formula: volume = width2 × length/2. When tumor volume exceeded 2 cm3 or at the end of treatment, mice were sacrificed and tumors harvested 1 hours after the last dose of lucitanib. Portions of tumors were snap frozen or fixed in 10% neutral-buffered formalin and embedded in paraffin for subsequent analyses. Five-micron paraffinized sections were used for IHC using Y653/54 phosphorylated FGFR1 (Abcam #111124) and ERα (Santa Cruz Biotechnology #8002). Sections were scored by an expert pathologist (M.M. Estrada) blinded to treatment arm.

Statistical analysis

Results are representative of three independent experiments and are expressed as the mean ± SEM. A *P* value of less than 0.05, determined by Student *t* test, was considered statistically significant.

FGFR1 and CCND1 fluorescence in situ hybridization (FISH) analysis

Four-µm tissue sections were mounted on charged slides and hybridized overnight with the SPEC FGFR1/CEN8 Dual Color Probe (ZytoVision, catalog# Z-2072-200) and CCND1/CEN11 Dual Color Probe (ZytoVision, catalog# Z-2071-200). Briefly, deparaffinization, protease treatment and washes were performed as per standard protocols. After this pretreatment, the slides were denatured in the presence of 10 µL of the probe for 6 min at 72°C, and hybridized at 37°C overnight in StatSpin (Thermobrite, Abbott Molecular, Inc.). Post-hybridization saline-sodium citrate washes were performed at 72°C and the slides were then stained with DAPI before analysis. Normal vessels, fibroblasts and/or non-tumor tissues served as internal positive controls. Cases were further evaluated only if diploid nuclei in normal tissues displayed one or two clearly distinct signals of each color. Tumor tissue was scanned for amplification hot spots under 40× magnification (Olympus BX60 Fluosescent microscope). If the FGFR1 or CCND1 signals were homogeneously distributed, then random areas were used for counting the signals. Twenty to sixty tumor cell nuclei from random areas were individually evaluated with the 100× oil immersion objective by counting green FGFR1 and orange centromere 8 (for FGFR1), or orange CCND1 and green centromere 11 (for CCND1) signals. The FGFR1/CEN8 or CCND1/CEN11 ratio and the average FGFR1 or CCND1 copy number per cell were calculated next. Cases were considered to be FGFR1 or CCND1 amplified under one of the following conditions:

a) FGFR1/CEN8 or CCND1/CEN11 ratio ≥2.0;

b) average number of FGFR1 or CCND1 signals per tumor cell nucleus ≥6

Cell proliferation assays

Clonogenic assays. Cells (5x104 /well) were seeded in triplicate in medium with 10% DCC-FBS in 6-well plates and then treated with ± 100 ng/mL FGF3 ± 2 µM lucitanib or 1 µM ICNB054828. Media, FGF ligands, and drugs were replenished every 3 days until 50-70% confluency was observed in control wells. Monolayers were then fixed and stained with 20% methanol/80% water/0.5% crystal violet for 20 min, washed with water, and dried. After photographic images of the plates

were obtained, the crystal violet stain was solubilized with 20% acid acetic and the image intensity of the monolayers was quantified by spectrophotometric detection at 490 nm using a plate reader (GloMax®-Multi Detection System, Promega).
siRNA transfection experiments. Cells were reverse transfected into 100-mm dishes using Lipofectamine RNAiMAX® (Invitrogen) and 25 nM siRNA [siControl- Ambion cat. #4390843; siFGFR1- Ambion cat. #AM16708; siFGFR1- Ambion cat. #AM51331]. The next day, 5x104 cells/well were reseeded in IMEM/10% DCC-FBS in 6-well plates for proliferation assays or in 60-mm plates for immunoblot analysis. For proliferation assays, media was changed 72 h after transfection to IMEM/10% DCC-FBS + 100 ng/mL FGF3 and every 3 days thereafter. Cells were trypsinized 7 days post-transfection and counted using a Coulter Counter (Beckman Coulter). For immunoblot analyses, cells were harvested and protein lysates prepared on day 3 post-transfection.

Three-dimensional Matrigel culture

Cells (~1x104 /well) were seeded in 48-well plates in triplicate. Before seeding, cells were suspended in their respective medium on growth factor–reduced Matrigel (BD Biosciences) as described previously.[210] Ligands and/or inhibitors were added at the time of cell seeding and replenished with fresh medium every 3 days. After 6 or 12 days, images were captured from at least 3 different fields using a CK 40 microscope. Cell viability was measured by MTT assay and the number of colonies per well was quantified by Gelcount® scanning.

Proximity ligation assay (PLA)

FGFR1 expression and localization. PLA was performed using FGFR1 (Abcam, cat. #10646, rabbit) antibody. Cells (5x104 /well) were seeded in 16-well chamber slides (Lab-Tek) in triplicate in their respective growth medium and then serum-starved for 24 h. PLA was performed as per the Duolink in situ PLATM protocol (Olink Bioscience, Sweden). To visualize the bound antibody pairs, the Duolink Detection Kit (#DUO92101 –Sigma) with PLA plus and minus probes for rabbit (anti-rabbit plus #DUO92002, anti-rabbit minus #DUO92005 -Sigma for FGFR1) were used according to the manufacturer's protocol. Slides were mounted with the Duolink Mounting Medium and stained with DAPI (82040-0005). Analysis was performed by confocal microscopy (LSM710, ZEISS) and the number of red dots (FGFR1) was quantitated by Duolink Image Tool software; 8-15 random fields per sample were analyzed.
FGFR1-ERα association and localization. PLA was performed using FGFR1 (Abcam, cat. #10646, rabbit) and ERα (Santa-Cruz, cat. #8002, mouse) antibodies. Cells (5x104 /well) were seeded in 16-well chamber slides (Lab-Tek) in triplicate in their respective growth medium and serum starved for 24 h. PLA (Duolink in situ PLATM; Olink Bioscience, Sweden) was performed to detect FGFR1/ERα complexes. To visualize the bound antibody pairs, the Duolink Detection Kit (#DUO92101, Sigma) with PLA plus and minus probes for rabbit (anti-rabbit plus, #DUO92002, Sigma) and PLA plus and minus probes for mouse (anti-mouse minus, #DUO92004, Sigma) were used according to the manufacturer's protocol. Slides were mounted with the Duolink Mounting Medium and stained with DAPI (Sigma 82040-0005). Analysis was performed by confocal microscopy (LSM710, ZEISS) and the number of red dots (indicating FGFR1/ERα complexes) was quantitated by the Duolink Image Tool software; 8-15 random fields per sample were analyzed. In addition to cells on slides, PLA was performed in 5-µm thick sections from paired pre- and post-letrozole FFPE tumor blocks from patients in the clinical trial. Tumor sections were de-paraffinized and subjected to antigen retrieval by microwave cooking in 0.01M citrate buffer (pH 6.0) at 1000 W for 30 min. After incubation in blocking buffer (1X PBS + 10% BSA + 0.3 % Triton X-100), the slides were incubated overnight with FGFR1 (Abcam, cat. #10646, rabbit) and ERα (Santa-Cruz, cat. #8002, mouse) antibodies. PLA (Duolink in situ PLATM; Olink Bioscience, Sweden) was performed to detect FGFR1/ER complexes and their localization. To visualize the bound antibody pairs, the Duolink Detection Kit (#DUO92101, Sigma) with PLA plus and minus probes for rabbit (anti-rabbit plus,

#DUO92002, Sigma) and PLA plus and minus probes for mouse (anti-mouse minus, #DUO92004, Sigma) was used according to the manufacturer's protocol. Slides were mounted with the Duolink Mounting Medium and stained with DAPI (Sigma 82040-0005). Analysis was performed by confocal microscopy (LSM710, ZEISS) and the number of red dots (FGFR1/ERα complexes) was quantified by Duolink Image Tool software; 8-15 random fields per sample were analyzed.

Gene expression analyses

RNA was purified from cells using RNeAsy kit (Qiagen, Valencia, CA) and cDNA was generated using High Capacity cDNA Reverse Transcription Kits (Applied Biosystems, Carlsbad, CA). qPCR was performed with a cDNA equivalent of 50 ng RNA, 1 μM each of the forward and reverse primers, and Sso Advanced SYBR Green Supermix (Bio-Rad) following the manufacturer's protocol using a CFX qPCR machine (Bio-Rad). We used primers against the following targets: GAPDH (QIAGEN-PPH00150F), FGF3 (QIAGEN-PPH00174C), FGF4 (QIAGEN-PPH00356A), FGF19 (QIAGEN-PPH01290B), CCL2 (QIAGEN-PPH00192F), CCND1 (QIAGEN-PPH00128F), EGR3 (QIAGEN-PPH01479C) and THSB1 (QIAGEN-PPH00799F). CT (threshold cycle) values were determined in triplicate samples by subtracting the target gene CT from the GAPDH CT; 2ΔCT was used to determine the expression of each target gene relative to GAPDH.

Immunoprecipitation and immunoblot analyses

Cells were lysed in RIPA buffer (for immunoblot) or in NP-40 buffer containing protease and phosphatase inhibitors (for immunoprecipitation), and sonicated for 10 sec; debris was separated by centrifugation at 18,000 xg for 10 min at 4°C. Protein concentration in the supernatants was measured using the BCA assay (Pierce). FGFR1 was precipitated from cell lysates with a FGFR1 C-terminal antibody (Abcam #76464) or a FGFR1 N-terminal antibody (Cell Signaling #3472) for 16 h at 4ºC. Whole cell lysates and immune complexes were separated by SDS-PAGE, transferred to nitrocellulose, and subjected to immunoblot analyses as described previously[211] using primary antibodies against ERα, FRS2α (Santa Cruz Biotech.), AIF, tubulin, lamin A/C, actin, phosphorylated FRS2α (T436), phosphorylated ERK1/2 (T202/T204), total ERK1/2, phosphorylated AKT (S473), total AKT, phosphorylated ERα (S167) (Cell Signaling) and FGFR1 (Abcam). HRP-conjugated anti-rabbit and anti-mouse were used as secondary antibodies (Santa Cruz Biotechnology). Immunoreactive proteins were visualized by enhanced chemiluminescence (Pierce, Rockford, IL, USA). Membranes were cut horizontally to probe with multiple antibodies. Blots probed with phospho-antibodies were stripped with Restore Western Blot Stripping Buffer (Thermo Fisher Scientific) and re-probed with antibodies to the total protein.

Membrane, cytoplasmic, nuclear soluble and chromatin-bound fractionation

CAMA1 cells were subjected to fractionation using a cell fractionation kit (Thermo Scientific #78840) according to the manufacturer's protocol. Adequacy of fractionation was confirmed by immunoblot of cell fractions with antibodies against apoptosis-inducing factor (AIF; plasma membrane), tubulin (cytoplasm), lamin A/C (nucleus) and histone H1 (chromatin bound).

Inhibition of nuclear export

CAMA1 (1x105) cells were grown in chamber slides and treated with vehicle or 30 ng/mL leptomycin B for 2 h and then fixed with PBS containing 3.7% formaldehyde, washed with PBS, permeabilized with PBS containing 0.25% Triton-X-100, blocked with PBS containing 10% BSA and 0.1% Tween-20, and incubated overnight with a FGFR1 (Abcam, cat. #10646, rabbit) primary antibody diluted in blocking solution. Slides were washed and incubated with goat-derived Alexa Fluor® 594-conjugated antibodies and mounted with ProLong® Gold Antifade mounting media (Life Technologies). IF analysis was

performed by confocal microscopy (LSM710, ZEISS); nuclear cell fluorescence was quantified by ImageJ using 8-15 random fields per sample.

ER transcriptional reporter assay

Cells were reverse transfected into 100-mm dishes using Lipofectamine RNAiMAX® (Invitrogen) and 25 nM siRNA [siControl-Ambion cat. #4390843; siFGFR1- Ambion cat. #AM16708; siFGFR1- Ambion cat. #AM51331; siFRS2 – Ambion cat. #AM4392420; siFRS2 – Ambion cat. #AM16708]. After 48hr, cells were transfected with pGLB-MERE (encodes two consecutive estrogen response elements) and pCMV-Renilla (Promega, encodes CMV-driven Renilla luciferase) plasmids in 100-mm dishes using Lipofectamine 2000 (Invitrogen). Cells were then reseeded in 96-well plates in 10% DCC-FBS. The next day, cells were treated with 1 μM fulvestrant or followed by measurement of ERE-luciferase activity 18 h later as previously described.

Results

FGFR1 amplification and overexpression is associated with endocrine resistance in ER+ breast cancer

We studied 72 tumor biopsies from postmenopausal women with clinical stage I–III operable, ER+/HER2− breast cancer treated with the aromatase inhibitor letrozole for 2 weeks prior to surgery (NCT00651976). Earlier studies have demonstrated that a Ki67 score 2 weeks after antiestrogen therapy can be utilized to predict which tumors are endocrine sensitive or resistant, as measured by their odds of recurrence following adjuvant endocrine therapy.[151] We applied these metrics to our tumor set and categorized 40 tumors as sensitive [natural log (ln) of post-letrozole Ki67 ≤1.0 or ≤2.4% tumor cells], 11 tumors as intermediate responders (ln = 1.1–1.9 or 2.5%–7.3% tumor cells), and 21 tumors as resistant (ln ≥ 2.0 or ≥7.4% tumor cells; **Fig. 4-1a**). FGFR1 copy number was determined in tumor sections by FISH. We observed FGFR1 amplification in 9 of 21 (43%) resistant tumors compared with 3 of 40 (7%) sensitive tumors and 1 of 11 (10%) intermediate tumors (resistant vs. intermediate and sensitive tumors; P = 0.0011; **Fig. 4-1b**). To correlate FGFR1 copy number with protein levels, we performed IHC. FGFR1 protein levels correlated with gene amplification by FISH. In FGFR1-amplified cancers, we observed a significant increase in total and nuclear FGFR1 in posttreatment compared with pretreatment biopsies ($P < 0.05$; **Fig. 4-1c and e**). A letrozole-induced increase in both total and nuclear FGFR1 was not observed in tumors without FGFR1 amplification (**Fig. 4-2a**). There was no statistical correlation between FGFR1 amplification and histologic tumor grade in the letrozole-resistant group (**Table 4-2**).

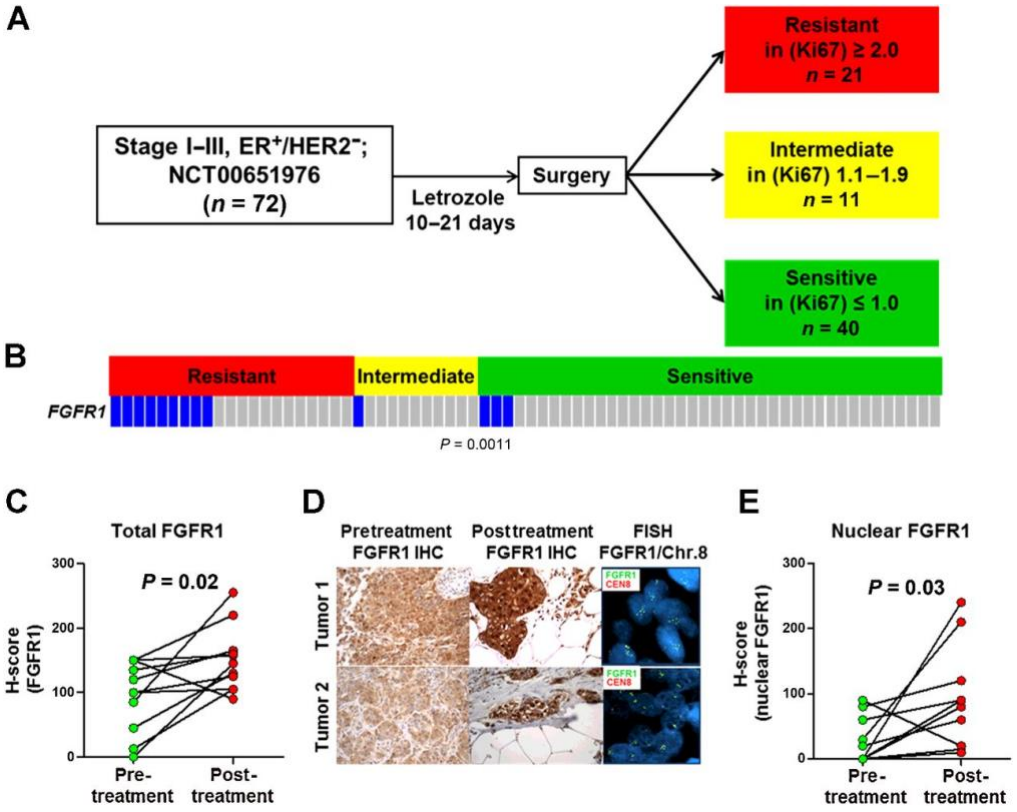


Figure 4-1. FGFR1 amplification and overexpression associate with endocrine resistance in ER+ breast cancer. **A**, Clinical trial schema: patients with stage I–III, ER+/HER2− breast cancer were treated with letrozole for 10 to 21 days. Surgery was performed following treatment, and tumor response was categorized by calculating the natural log (ln) of the post-letrozole Ki67 score as determined by IHC analysis. **B**, FGFR1 amplification, determined by FISH, was significantly associated with resistant versus intermediate or sensitive tumors ($P < 0.05$, Student t test). **C–E**, FFPE sections from FGFR1-amplified tumors were stained for FGFR1; the percent of FGFR1-positive tumor cells and staining intensity were assessed in both the cytoplasmic and nuclear compartments by a blinded expert breast pathologist (M.M. Estrada) to generate an H-score (**D**). The percent of cytoplasmic and nuclear FGFR1+ tumor cells and their staining intensity were assessed by a blinded expert pathologist (M.M. Estrada) to generate an H-score. Total and nuclear FGFR1 H-scores are shown in **C** and **E**, respectively (Student t test). Both total and nuclear FGFR1 staining was higher in posttreatment tumor sections.

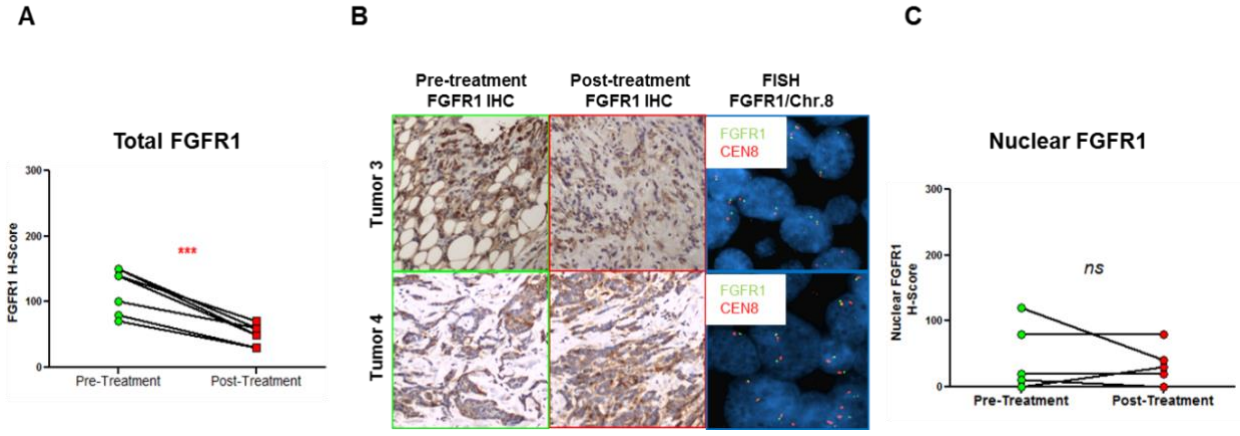


Figure 4-2. Effect of letrozole on expression and localization of FGFR1 in primary breast tumors without FGFR1 amplification. FGFR1 gene copy number and protein expression were determined by FISH (FGFR1:Chr.8 ratio, 100x magnification) and IHC, respectively. Total (**A**) and nuclear (**B**) FGFR1 expression was decreased in post-letrozole compared to paired pre-letrozole patient tumors (***p<0.001 vs. pre-letrozole, Student's t-test). Representative IHC and FISH images are shown in **B**.

Data analyzed	G1	G2/G3	Total
FGFR1 amplification	1	8	9
No amplified tumors	1	11	12
Total	2	19	21

P value and statistical significance	
Test	Fisher's exact test
P value	>0.9999
P value summary	ns
One- or two-sided	Two-sided
Statistically significant (P < 0.05)?	No

Data analyzed	G1/G2	G3	Total
FGFR1 amplification	8	1	9
No amplified tumors	8	4	12
Total	16	5	21

P value and statistical significance	
Test	Fisher's exact test
P value	0.3383
P value summary	ns
One- or two-sided	Two-sided
Statistically significant (P < 0.05)?	No

Table 4-2. FGFR1 amplification does not correlate with a specific histological tumor grade.

Estrogen deprivation increases nuclear and cytosolic FGFR1 expression

To examine whether this same modulation of FGFR1 levels occurred in more controlled experimental conditions, we tested five ER+/HER2− human breast cancer cell lines with and without FGFR1 gene amplification as determined by FISH: CAMA1, MDA-MB-134, and HCC1500 cells are FGFR1 amplified, while MCF-7 and ZR75.1 cells are not (**Fig. 4-3a**). FGFR1 amplification correlated with FGFR1 protein levels; MDA-MB-134 and HCC1500 cells express both full-length and cleaved FGFR1, while only full-length FGFR1 was detected in CAMA1 cells (**Fig 4-3b**). To mirror the acute estrogen deprivation induced by letrozole in primary tumors in the clinical trial[212], we cultured the FGFR1-amplified cell lines in estrogen-free medium for 4 to 6 days. Estrogen withdrawal resulted in an increase in FGFR1 expression in all FGFR1-amplified lines (**Fig. 4-4a**).

To determine whether long-term estradiol deprivation also affected FGFR1 expression, we generated three LTED cell lines as described previously[206]: CAMA1LTED and MDA-MB-134LTED (FGFR1-amplified) and MCF-7LTED (FGFR1 nonamplified). As we had observed with acute estrogen deprivation, CAMA1LTED and MDA-MB-134LTED cells exhibited increased expression of full-length and cleaved FGFR1, respectively, whereas MCF-7LTED cells showed a reduction in FGFR1 expression compared with parental MCF-7 cells. The LTED lines also displayed an increase in ERα levels compared with their parental counterparts (**Fig. 4-4b**). IF by confocal microscopy highlighted the increase in total and nuclear FGFR1 in CAMA1LTED versus parental cells (**Fig. 4-4c**). We next treated CAMA1 cells with nuclear export inhibitor leptomycin B[213]; this resulted in an increase in nuclear FGFR1 as measured by IF (**Fig. 4-4d**). Knockdown of FGFR1 with siRNA confirmed the specificity of the FGFR1 antibody used for both immunoblot and IF analyses (**Fig. 4-3c, d**).

A

Cell line	Ratio FGFR1:Chr.8	Status
MCF7	0.77	*No Amp*
ZR75.1	1.69	*No Amp*
MDA-MB-134	6.8	*Amp*
CAMA1	2.31	*Amp*
HCC1500	2.03	*Amp*

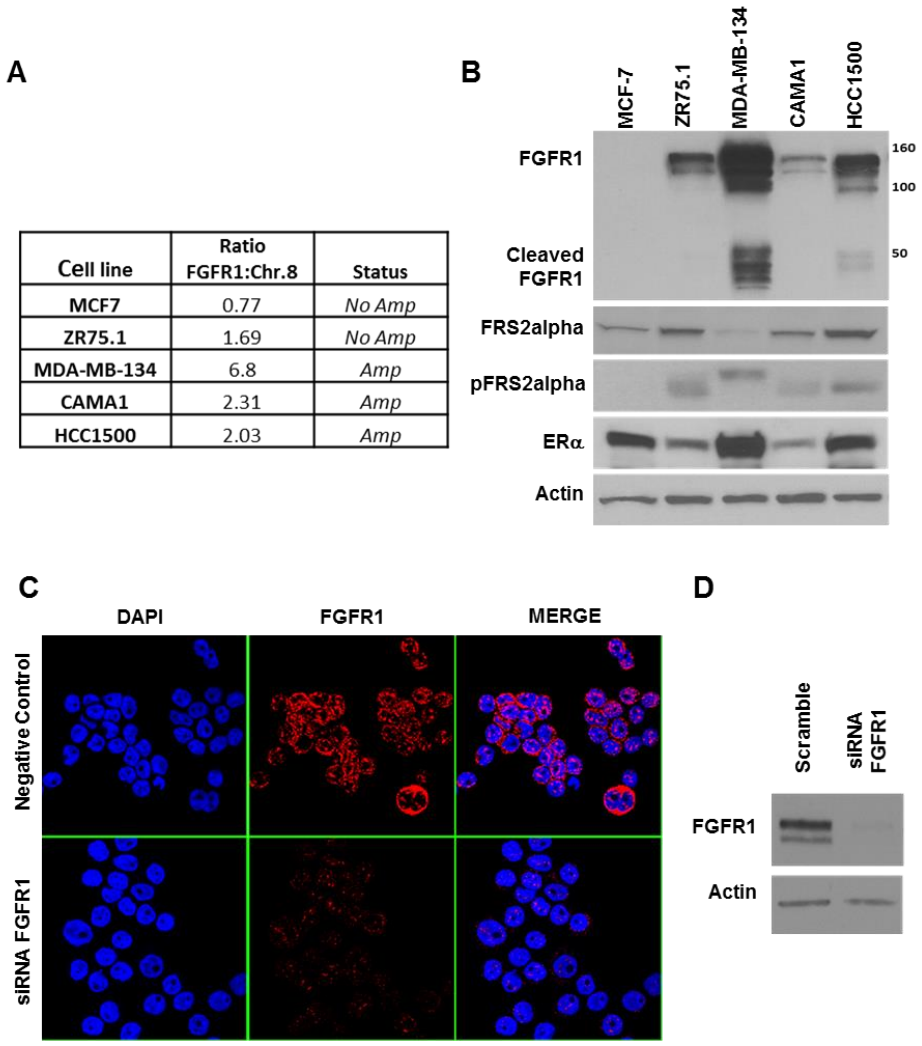


Figure 4-3. FGFR1 amplification and protein expression in ER+ human breast cancer cell lines. **A**, Table depicts the FGFR1:Chr.8 ratio in a panel of ER+ human breast cancer cell lines as determined by FISH. MDA-MB-134, CAMA1, and HCC1500 cells are FGFR1 amplified, whereas MCF7 and ZR75.1 cells are not. **B**, Immunoblot analysis of cell lysates displays the relative content of the full-length and cleaved forms of FGFR1, FRS2, phosphorylated FRS2 and ERα, using actin as a loading control. **C-D**, PLA was used to assess FGFR1 expression in CAMA1 cells transfected with FGFR1 siRNA or a negative (scrambled) control as described in Methods. Cell lysates from identically-treated parallel plates were prepared and subjected to immunoblot analysis with the indicated antibodies to confirm siRNA-mediated FGFR1 knockdown.

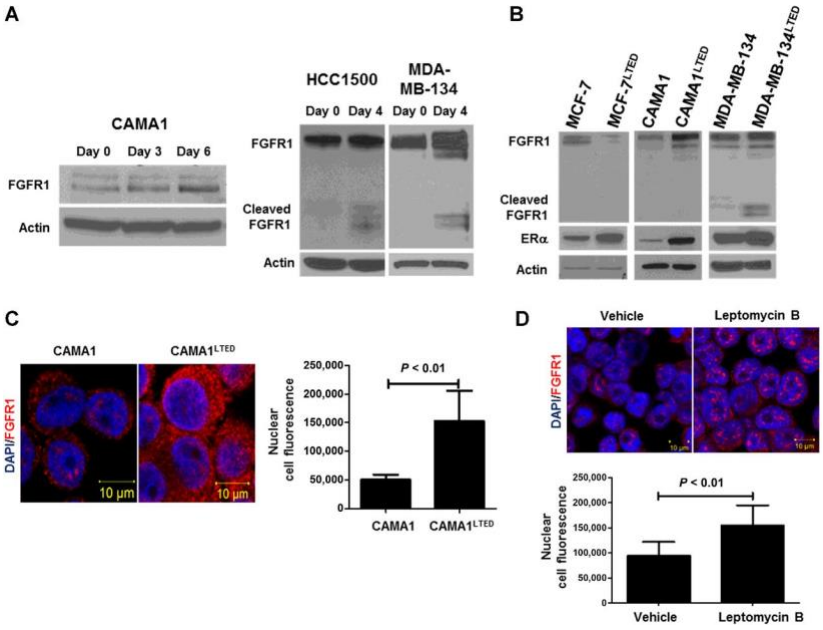


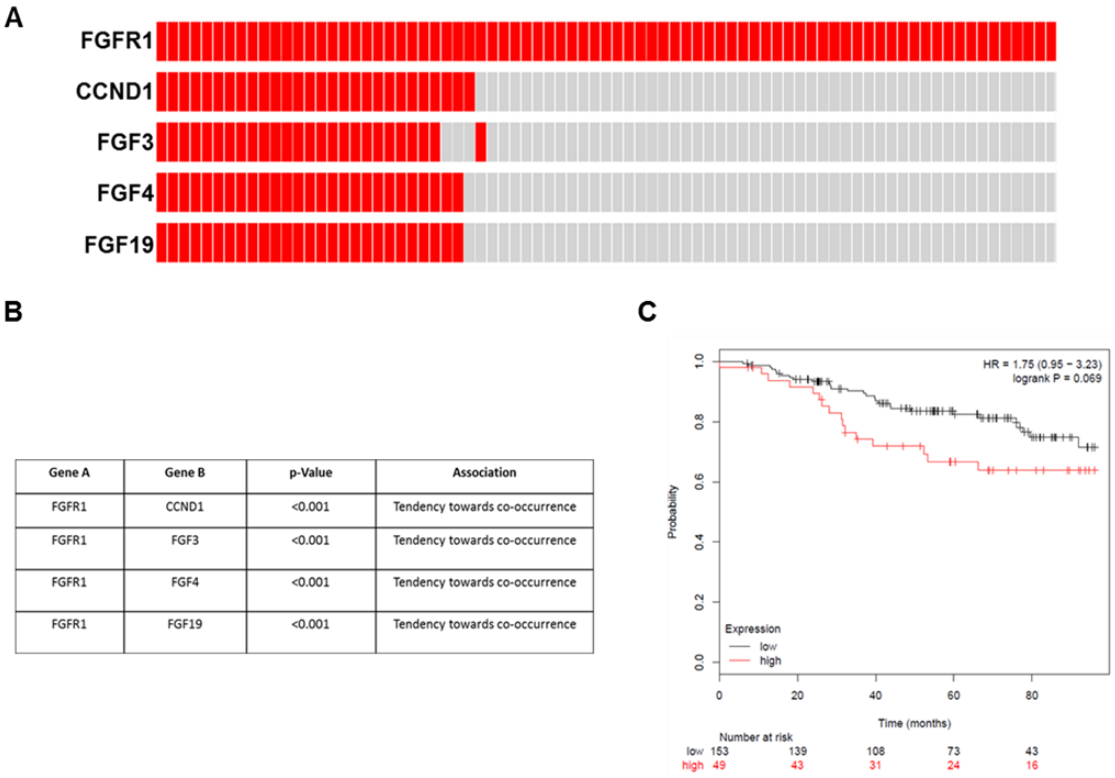
Figure 4-4. Estrogen deprivation increases nuclear and cytosolic FGFR1 expression. **A**, Immunoblot analysis of lysates from CAMA1, HCC1500, and MDA-MB-134 cells exposed to short-term estrogen deprivation up to 6 days revealed an increase in FGFR1 expression over time. HCC1500 cells showed increased expression of the cleaved form of FGFR1. **B**, Immunoblot analysis of parental and LTED ER+ cell lines following 24 hours of estrogen deprivation revealed an increase in FGFR1 and ERα in FGFR1-amplified CAMA1LTED and MDA-MB-134LTED cells but not in FGFR1 nonamplified MCF-7 cells. **C**, Proximity ligation assay (PLA) to detect FGFR1 expression. Analysis of red, amplified loci by confocal microscopy confirmed immunoblot and FISH results in that CAMA1LTED cells harbor more cytosolic and nuclear FGFR1 compared with CAMA1 parental cells. Each bar in the graph to the right of the PLA image represents the mean nuclear fluorescent signals ± SD of 3 wells. **D**, Immunofluorescence (IF) was performed in CAMA1 cells treated with vehicle or 30 ng/mL leptomycin B for 2 hours. Nuclear localization of FGFR1 was detected by confocal microscopy. Each bar represents the mean nuclear fluorescent signals ± SD of 3 wells.

FGF3/4/19 expression is upregulated upon estrogen deprivation

Approximately 30% to 40% of FGFR1-amplified breast cancers exhibit amplification of CCND1, FGF3, FGF4, and FGF19 in chromosome 11q12-14.[214] Coamplification of these genes has been associated with reduced patient survival.[203] By interrogating The Cancer Genome Atlas, we found that among the 13% of breast cancers with FGFR1 amplification, 36% of these tumors also harbor 11q-12-14 amplification (**Fig. 4-5a, b**).[135,136] Outcomes analysis of Kaplan–Meier plotter (breast cancer) showed that patients with coamplification of FGFR1 and CCND1/FGF3/FGF4/FGF19 treated with antiestrogen therapy exhibit a shorter time to relapse compared with patients without coamplified tumors (HR = 1.75; **Fig. 4-5c**).[215] Thus, we next investigated coamplification of FGFR1 and 11q12-14 in our cohort of patients treated with letrozole. In this study, 8 of 9 (90%) FGFR1-amplified tumors exhibited coamplification of FGF3/4/19, and this coamplification strongly correlated with resistance to estrogen deprivation with letrozole (P = 0.0001; **Fig. 4-6a**). These data suggest that coamplification of 11q12-14 and FGFR1 may play a causal role in endocrine resistance.

Notably, all FGFR1-amplified cell lines but not MCF-7 cells exhibited coamplification of 11q12-14 (**Fig. 4-6b, Fig. 4-7a**). All 11q12-14–amplified cell lines expressed markedly higher FGF3/4/19 mRNA levels by qRT-PCR compared with MCF-7 cells (**Fig. 4-6c**). Similar to the effect on FGFR1 protein levels, 24 hours of estrogen deprivation increased FGF3/4/19 mRNA expression 1.5- to 2-fold in all FGFR1-amplified cells (**Fig. 4-7b**). This increase in FGF3/4/19 was even more substantial in LTED FGFR1-amplified cells. In contrast, MCF7LTED cells exhibited little or no increase in FGF ligands mRNA compared with MCF-7 parental cells (**Fig. 4-6d**).

These results also suggested that FGFs can provide a growth advantage to ER+/FGFR1-amplified cells in estrogen-free conditions. To test this, we stimulated estrogen-starved CAMA1 cells with FGF3 in hormone-depleted media. Exogenous FGF3 enhanced estrogen-independent cell growth compared with unstimulated cells. Both treatment with the FGFR1 tyrosine kinase inhibitor (TKI), lucitanib (**Fig. 4-6e**)[216], and transfection with FGFR1 siRNA prevented this outgrowth (**Fig. 4-6f**).



Gene A	Gene B	p-Value	Association
FGFR1	CCND1	<0.001	Tendency towards co-occurrence
FGFR1	FGF3	<0.001	Tendency towards co-occurrence
FGFR1	FGF4	<0.001	Tendency towards co-occurrence
FGFR1	FGF19	<0.001	Tendency towards co-occurrence

Figure 4-5. Breast cancers with co-amplification of FGFR1 and 11q12-14 genes exhibit decreased time to recurrence. **A**, Tile plot of ER+ breast cancers in TCGA (Cell 2015) with co-amplification of FGF3/4/19 and CCND1 on chr.11q12-14 and of FGFR1 on chr. 8p11. **B**, Analysis of TCGA breast whole exome sequencing (WES) data showed significant co-occurrence of FGFR1 and FGF3/4/19 amplification (n=594 samples; p<0.001, Fisher's t-test). **C**, Kaplan Meier plot from the KMPLOT gene expression database showing the probability of relapse for patients with ER+ breast cancer treated with endocrine therapy comparing the high and low tertiles of both FGFR1 and FGF3/4/19 mRNA expression by microarray. Patients in the high tertile tended toward a shorter relapse-free survival compared to patients in the low tertile (HR 1.75, p =0.069).

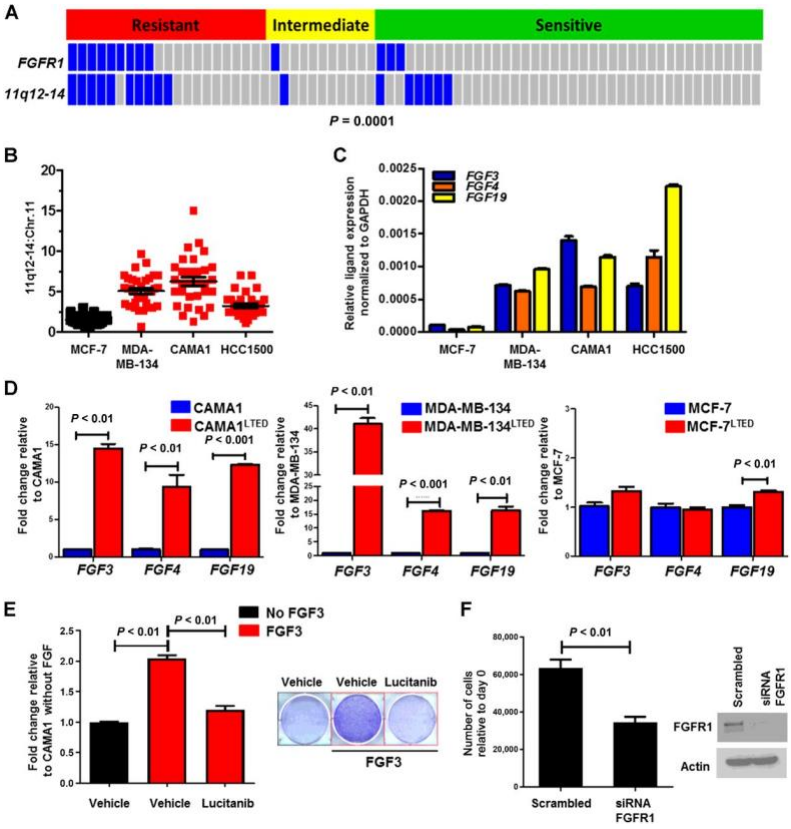


Figure 4-6. FGF3/4/19 expression is upregulated upon estrogen deprivation. **A**, FISH analysis of primary tumor sections showed coamplification of FGFR1 and 11q12-14 mainly in letrozole-resistant versus intermediate and sensitive cancers (P = 0.0001, Student t test). **B**, Coamplification of 11q12-14 was observed in ER+/FGFR1-amplified cell lines MDA-MB-134, CAMA1, and HCC1500; the y-axis shows the 11q12-14:Chr.11 ratio. **C**, Relative transcript expression of FGF3/4/19 in the indicated cell lines was determined by qPCR as described in Materials and Methods. **D**, Transcript levels of FGF3/4/19 were higher in FGFR1-amplified LTED cells (CAMA1 and MDA-MB-134) but not in FGFR1 nonamplified MCF-7LTED cells compared with their parental counterparts (Student t test). **E**, CAMA1 cells were treated with 100 ng/mL FGF3 ± 2 μmol/L lucitanib in estrogen-free medium. After 15 days, plates were washed and stained with crystal violet, and their imaging intensity was quantified by spectrophotometric detection. Representative images and quantification of the integrated intensity values as % of vehicle-treated controls are shown (Student t test). **F**, CAMA1 cells were plated in 100-mm dishes and transfected with FGFR1 or control siRNAs as described in Materials and Methods. Medium

containing 100 ng/mL FGF3 was replenished every 3 days. Seven days later, monolayers were harvested and cell counts determined using a Coulter Counter. Each bar in the left panel represents the mean cell number ± SD of triplicate wells (Student t test). FGFR1 knockdown was confirmed by immunoblot analysis of cell lysates from plates treated identically in parallel (right).

A

Cell line	Ratio 11q12-14:Chr.11	Status
MCF7	1.9	*No Amp*
MDA134	4.9	*Amp*
CAMA1	5.4	*Amp*
HCC1500	3.0	*Amp*

B

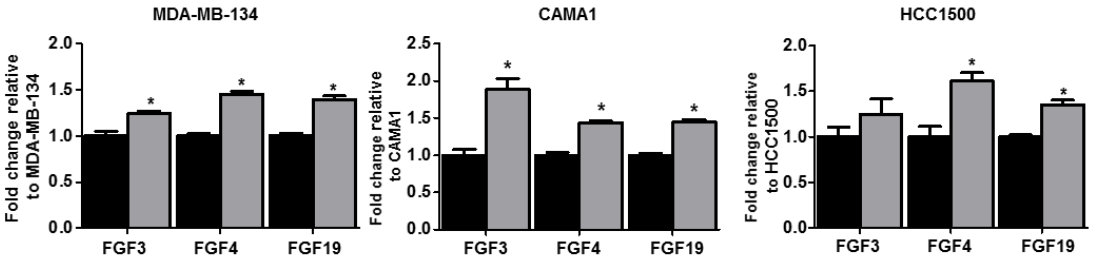


Figure 4-7. Estrogen deprivation upregulates FGF ligand expression in ER+/FGFR1-amplified cells. **A**, 11q12-14 amplification was determined in a panel of ER+ cell lines by FISH using CCND1 and chromosome 11 centromere probes. FGFR1-amplified MDA-MB-134, CAMA1, and HCC1500 cell lines exhibited co-amplification at 11q12-14. **B**, MDA-MB-134, CAMA1 and HCC1500 cells were cultured in full media or estrogen-free medium for 24 h. RNA was collected at that time and subjected to mRNA expression analysis by qPCR as described in Methods. Estrogen-deprivation (grey bars) resulted in an increase in FGF3/4/19 transcript levels compared to non-deprived conditions (black bars) (*p<0.05 vs. control, Student's t-test).

Long-term estradiol deprivation increases the interaction of FGFR1 with ERα

An association of FGFR1 with other nuclear proteins, such as ribosomal S6 kinase (RSK1) and CREB-binding protein (CBP), has been shown to be required for nuclear FGFR1 to induce gene expression in medulloblastoma and neuroblastoma cells.[217] An interaction between FGFR1 and ERα has been reported to mediate lactotroph proliferation in the pituitary gland.[218] Nuclear colocalization of PR, FGFR2, and STAT5 at DNA progesterone-responsive elements with increased transcription of PR/STAT5–regulated genes was also reported in human breast cancer cells.[194] Thus, we next investigated whether ER and FGFR1 interacted in ER+/FGFR1–amplified breast cancer cells. Antibody pulldown of FGFR1 from MDA-MB-134, CAMA1, and CAMA1LTED whole-cell lysates coprecipitated ERα in all three cell lines (**Fig. 4-8a**). This association was stronger in MDA-134 and CAMA1LTED cells compared with parental CAMA1 cells. We next confirmed the FGFR1–ERα association in CAMA1 and CAMA1LTED nuclear extracts after precipitation with both C-terminal and N-terminal FGFR1 antibodies (**Fig. 4-8b**), suggesting with presence of full-length FGFR1 in cell nucleus. To quantitate this interaction, we performed PLAs. An interaction between FGFR1 and ERα was observed in the cytoplasm and nucleus of CAMA1 and CAMA1LTED cells by PLA, particularly in the latter (**Fig. 4-8c and d**), in line with the immunoprecipitation experiments. Treatment with lucitanib reduced FGFR1/ERα complexes (**Fig. 4-8e and f**), suggesting this interaction requires FGFR1 tyrosine kinase (TK) activity.

To explore further whether the TK function of FGFR1 is required for FGFR1–ERα complex formation, CAMA1 cells were transduced with constructs expressing GFP, wild-type FGFR1, or a TK dead K514M FGFR1 mutant (FGFR1/TK–).

Overexpression of wild-type FGFR1 increased detectable FGFR1-ERα complexes, while overexpression of FGFR1/TK– decreased them as measured by PLA (**Fig. 4-8g and h**). Steady-state levels of pFRS2 were upregulated in cells transduced with wild-type FGFR1 but not with the FGFR1/TK– mutant (**Fig. 4-9a**). Importantly, the CAMA1$^{FGFR1/TK-}$ cells were not able to grow in the absence of estradiol (**Fig. 4-9b and c**). These data suggest that FGFR1 TK activity is important for estrogen-independent growth and the association of FGFR with ERα. Finally, we observed an increase of FGFR1-ERα complexes in post-letrozole compared with paired pre-letrozole FFPE tumor sections from 2 breast cancer patients harboring tumor coamplification of FGFR1 and 11q12-14 (**Fig. 4-8i and j**).

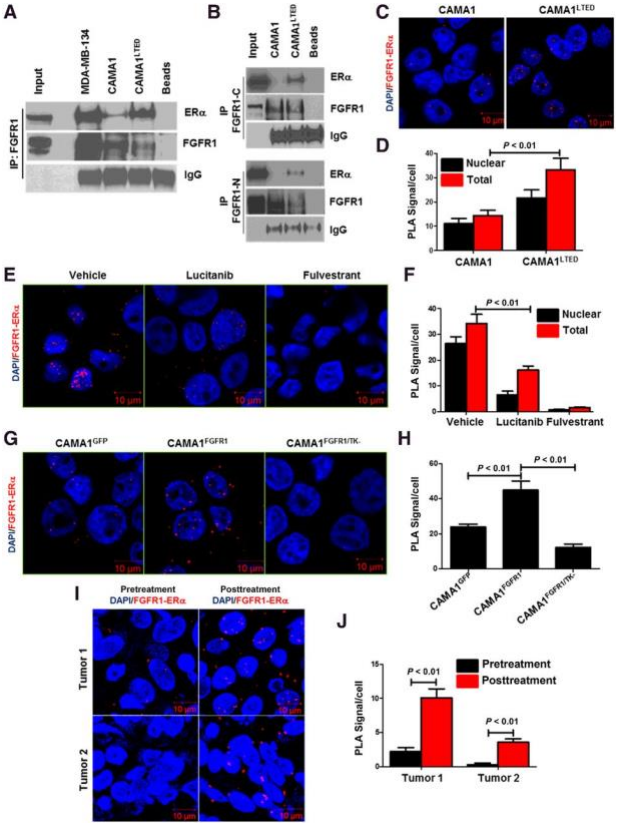


Figure 4-8. Long-term estradiol deprivation increases the interaction of FGFR1 with ERα. **A**, FGFR1 was precipitated from MDA-MB-134, CAMA1 and CAMA1LTED cell lysates; immune complexes were separated by SDS-PAGE and subjected to immunoblot analysis with an ERα antibody. CAMA1LTED cells exhibited greater levels of FGFR1-ERα coimmunoprecipitation compared with CAMA1 cells. **B**, FGFR1 was precipitated from CAMA1 and CAMA1LTED nuclear extracts with C-terminal (Abcam) and N-terminal (Cell Signaling Technology) FGFR1 antibodies; immune complexes were separated by SDS-PAGE and analyzed by ERα immunoblot. **C** and **D**, PLA of CAMA1LTED cells showed greater nuclear colocalization of FGFR1 and ERα compared with parental CAMA1 cells. PLA foci/cell are quantified in **D**. **E** and **F**, CAMA1LTED cells were treated with 2 μmol/L lucitanib or 1 μmol/L fulvestrant for 6 hours. Monolayers were subjected to PLA as described in Materials and Methods. Quantification of FGFR1–ERα complexes as PLA signals/cell is shown in **F**. Each bar represents the mean ± SD of 3 wells. **G** and **H**, CAMA1 cells were stably transfected with expression vectors encoding GFP, FGFR1, and FGFR1/TK– (K514M TK mutant), as described in Materials and Methods, and then plated in chamber slides followed by PLA. Quantification of FGFR1–ERα complexes as PLA signals/cell is shown in **H**. Each bar represents the mean ± SD of 3 wells. **I** and **J**, Paired pre- and post-letrozole primary tumor sections were subjected to PLA as described in Materials and Methods. Post-letrozole tumor cells exhibited more FGFR1–ERα complexes compared with pretreatment tumor cells as quantitated in **J**. Each bar represents the mean PLA signals/cell ± SD of 20 cells counted in each of 4 high-power fields.

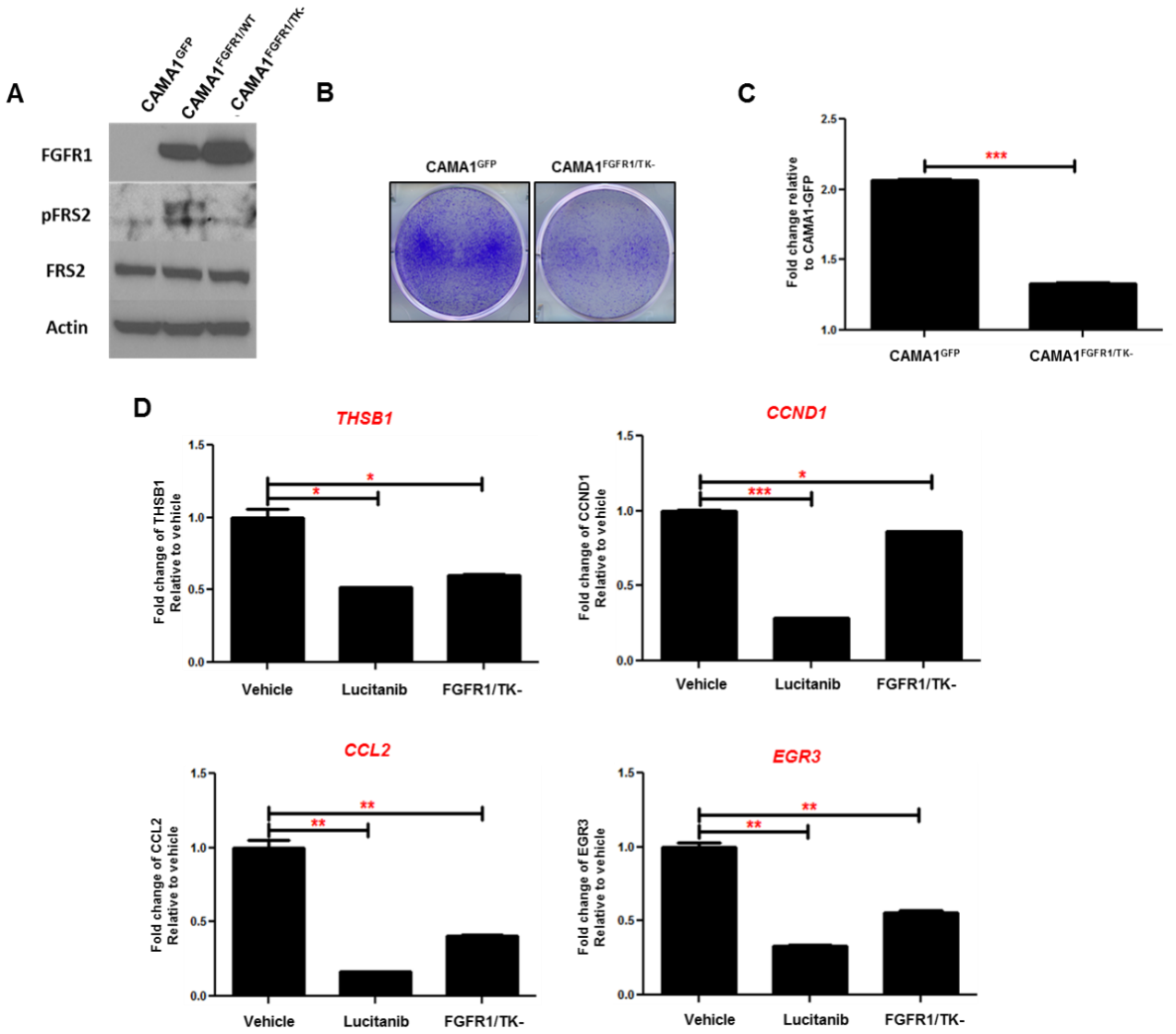


Figure 4-9. FGFR1 TK activity is important for estrogen-independent growth and the association of FGFR with ERα. A, CAMA1^{GFP}, CAMA1$^{FGFR1/WT}$ and CAMA1$^{FGFR1/TK-}$ cells were treated with 100 ng/mL FGF3 for 6 h and then lysed for immunoblot analysis with the indicated antibodies. **B-C**, CAMA1^{GFP} and CAMA1$^{FGFR1/TK-}$ cells were seeded in 6-well plates in estrogen-free media. After 14 days, monolayers were stained with crystal violet. Images of the plates were obtained (**B**) and image intensity was quantitated as described in Methods. Quantitation is shown in (**C**; ***p<0.001 vs. CAMA1GFP, Student's t-test). **D**, CAMA1^{GFP} ± 2 μM lucitanib and CAMA1$^{FGFR1/TK-}$ cells were plated in estrogen-free medium and treated with 100 ng/mL FGF3 for 6 h. At this time, cells were harvested and RNA was prepared and analyzed for THBS1, CCND1, CCL2 and EGR3 mRNA changes by qRT-PCR. Each bar represents the mean transcript level ± SD (*p<0.05, **p<0.01, ***p<0.001 vs. CAMA1GFP, Student's t-test).

FGF/FGFR pathway modulates ERα–DNA binding

To evaluate estrogen-independent genomic functions of ERα in ER+/FGFR1–amplified cells, we performed ChIP followed by next-generation sequencing (ChIP-seq) in estrogen-deprived CAMA1 cells ± FGF3. First, we confirmed by cells fractionation that parental CAMA1 cells exhibited nuclear soluble and chromatin-bound FGFR1 at steady state (**Fig. 4-10a**). Treatment with FGF3 shifted both ERα and FGFR1 to new binding sites that were unoccupied in the absence of the FGF ligand (**Fig. 4-10b** and **c**). We identified 1,120 and 553 regions (peaks) by ERα-ChIP and FGFR1-ChIP, respectively, that were significantly enriched upon FGF3 treatment. Treatment of CAMA1 cells with each fulvestrant or lucitanib alone or in combination reduced ERα or FGFR1 DNA binding to these new sites (**Fig. 4-11a** and **b**). These results were validated by ChIP-PCR (**Fig. 4-10d** and **e**). As

shown in **Fig. 4-9c** and **d**, ERα and FGFR1 bound to different ERα-related genes, but treatment with lucitanib, fulvestrant, or the combination reduced or abrogated this binding.

To interrogate the functional output of estrogen-independent ERα activity, we classified the genes identified by FGFR1 and ERα ChIP-seq using GSEA. The top enriched gene sets included epithelial mesenchymal transition, STAT5 signaling, estrogen response early genes, and p53-pathways (all FDR < 0.009) after FGFR1 ChIP-seq (**Fig. 4-11e**); and estrogen response early genes, estrogen response late genes, K-Ras signaling, and p53-pathways (all FDR < 0.0001) after ERα ChIP-seq (**Fig. 4-11f**). To apply these findings to primary ER+ breast cancers, we performed RNA-seq analysis on 7 FGFR1-amplified and 25 FGFR1 nonamplified tumors treated with letrozole in the clinical trial. The Volcano plot in Fig. 5G shows that of >24,000 genes analyzed, 280 gene transcripts were increased >2-fold in FGFR1-amplified compared with FGFR1 nonamplified cancers (P < 0.01; red dots in **Fig. 4-11g**). The top enriched genes by GSEA in FGFR1-amplified patients included G2–M checkpoint genes, E2F target genes, estrogen response late genes, and estrogen response early genes (all FDR < 0.01; **Fig. 4-11h**; **Fig. 4-12**). These results further suggest that the ERα pathway is still active in estrogen-deprived (upon letrozole treatment) ER+/FGFR1–amplified primary tumors.

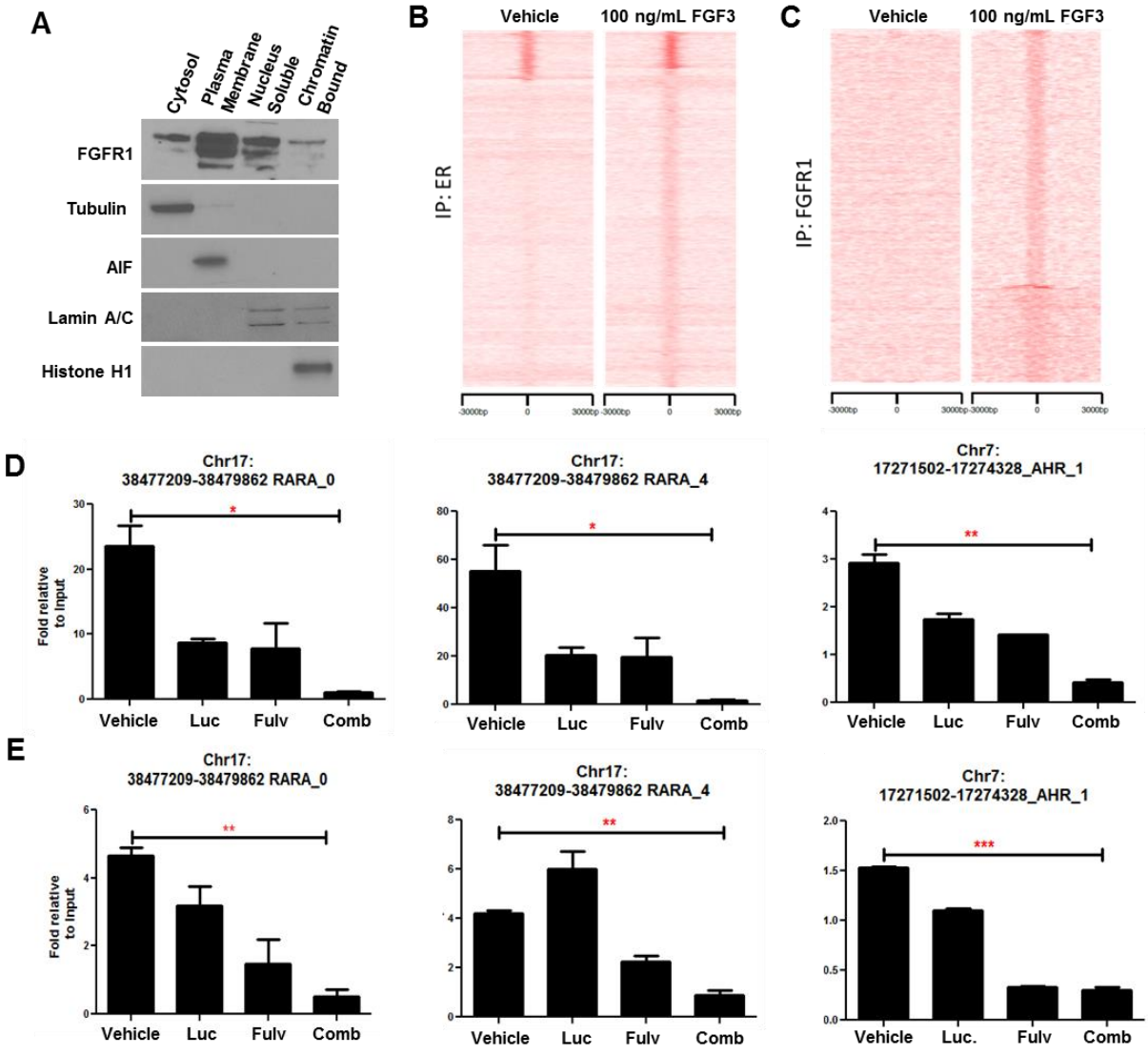


Figure 4-10. FGF3 induces binding of ERα and FGFR1 to DNA. **A**, Membrane, cytoplasmic, nuclear and chromatin-bound fractions of CAMA1 cells revealed full-length FGFR1 in both nuclear soluble and chromatin-bound fractions. Apoptosis-

inducing factor (AIF), tubulin, lamin A/C and histone H1 antibodies were used as controls. **B-C**, Heatmaps from ChIP-seq analysis of estrogen-deprived CAMA1 cells ± 100 ng/mL FGF3 (6 h) showing ERα (**B**) and FGFR1 (**C**) DNA binding peaks. The heatmaps represent the mean of two different experiments. **D-E**, CAMA1 cells were plated in estrogen-free media and treated with vehicle, 2 μM lucitanib, 1 μM fulvestrant or the combination for 6 h. ChIP was performed with ERα (**D**) or FGFR1 (**E**) antibodies. Primers to amplify FGFR1 or ERα binding regions were used in qPCR to determine fold enrichment relative to input. Two-tailed Student's unpaired t test was performed to compare mean signal amplification in cells treated with vehicle vs. lucitanib plus fulvestrant. Each bar represents the mean fold-enrichment in ERα-ChIP (**D**) or FGFR1-ChIP (**E**) ± SD of two independent experiments with three technical replicates each.

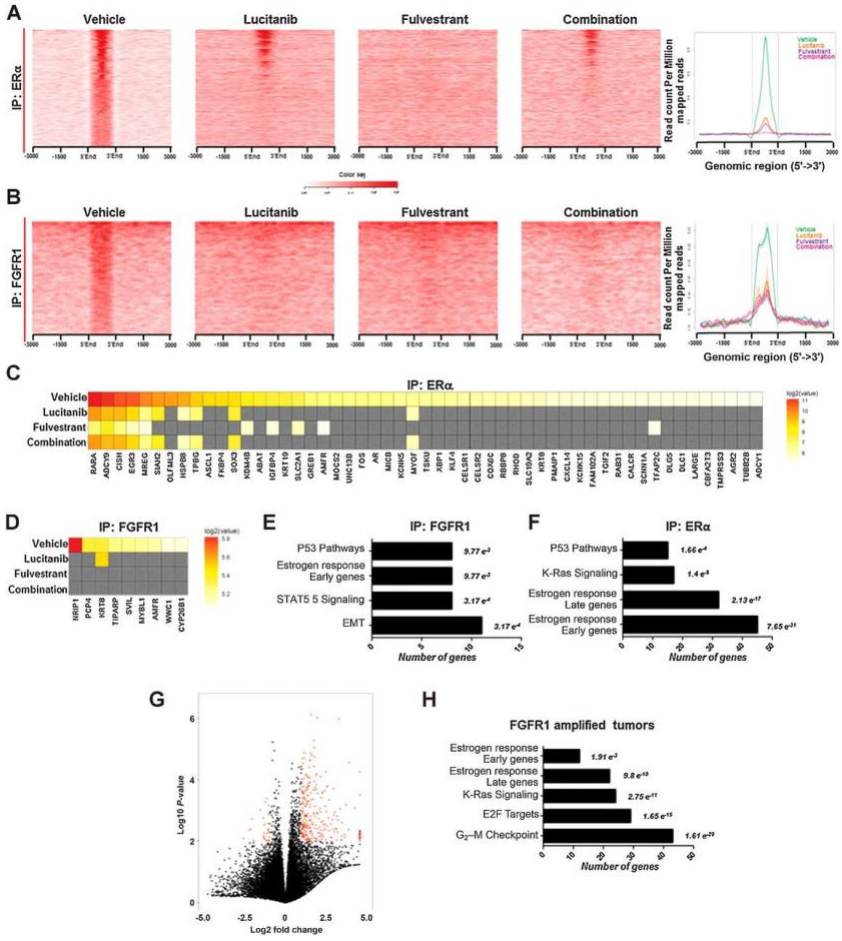


Figure 4-11. Identification of FGF-sensitive ERα and FGFR1 genomic binding sites. **A** and **B**, CAMA1 cells were plated in estrogen-free medium and stimulated with 100 ng/mL FGF3 for 6 hours in the presence of 1 μmol/L fulvestrant, 2 μmol/L lucitanib, or the combination. Cells were harvested and subjected to ChIP-seq as described in Materials and Methods. Shown are heatmaps generated from ChIP-seq analysis of ERα (**A**) and FGFR1 (**B**) DNA binding. Treatment with fulvestrant, lucitanib, or the combination reduced binding of ERα (**A**) or FGFR1 (**B**) binding to DNA. Heatmaps represent the mean of two different experiments. **C** and **D**, Heatmaps of ChIP-seq data showing the effects of fulvestrant, lucitanib, or the combination on DNA/ERα-associated (**C**) and DNA/FGFR1-associated (**D**) genes, respectively, as shown in **A** and **B**. **E** and **F**, Gene set enrichment analysis (GSEA) of FGFR1- and ERα-associated genes. Numbers to the right of each bar represent the FDR q-value. **G**, Volcano plot analysis of differentially expressed genes in tumors from patients treated with letrozole in the clinical trial. Each data point represents the ratio of the average expression for a particular gene in FGFR1-amplified tumors (n = 7) versus FGFR1 nonamplified tumors (n = 25). The red dots in the Volcano plot represent genes that are significantly up- or downregulated >2-fold with $P < 0.01$. **H**, GSEA of significantly enriched genes in FGFR1-amplified relative to FGFR1 nonamplified tumors showed that ERα-related pathways are still active in estrogen-deprived (by letrozole treatment) ER+/FGFR1-amplified primary tumors (**G**). Numbers to the right of each bar represent the FDR q-value.

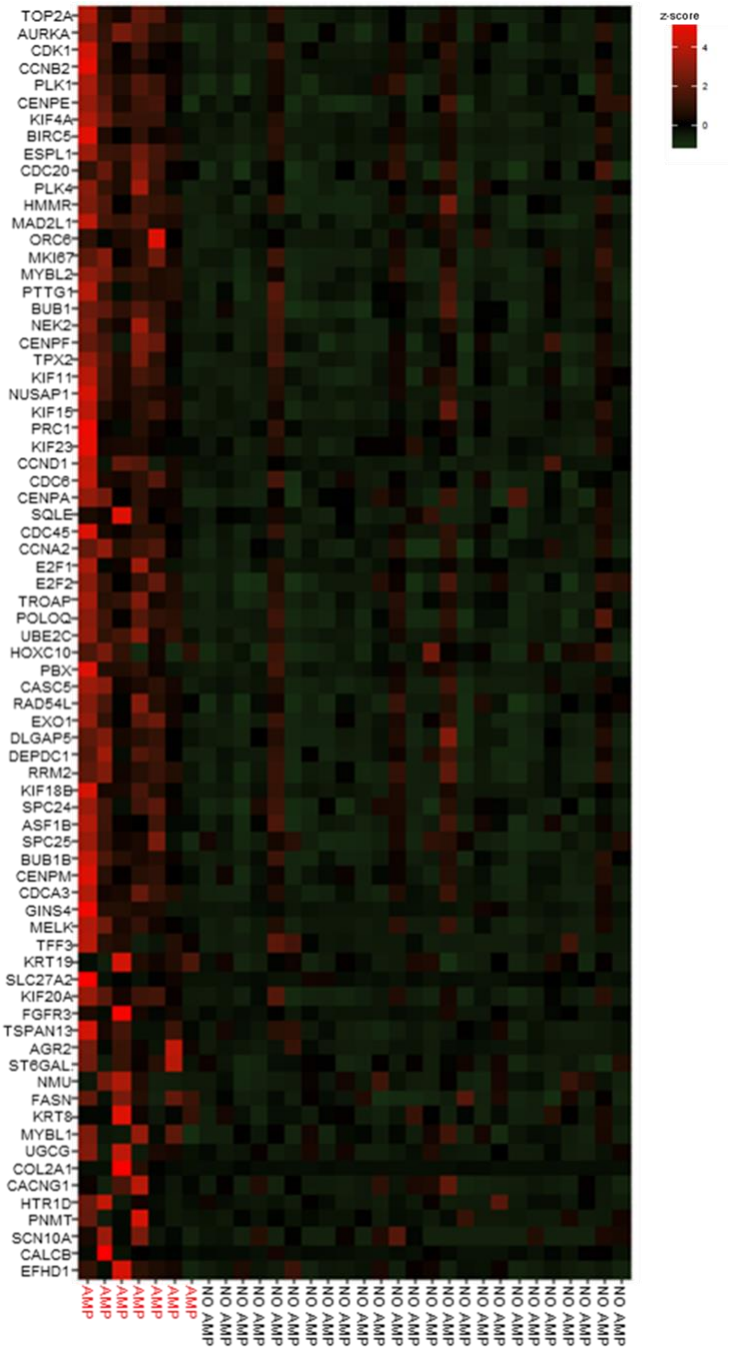


Figure 4-12. ER+/FGFR1-amplified tumors exhibit differential gene expression compared to ER+/FGFR1 non-amplified breast cancers. Heatmap of G2M checkpoint genes, E2F target genes and estrogen-response genes identified by Gene Set Enrichment Analysis (GSEA) in ER+/FGFR1-amplified vs. ER+/FGFR1 non-amplified breast tumors from patients treated with letrozole in the clinical trial.

To further elucidate the role of the FGF/FGFR1 axis on ERα signaling, we performed a qRT-PCR profiling assay including 84 ERα regulated genes. FGF3/19 stimulation of estrogen-deprived CAMA1 cells induced >2-fold expression of a subset of ERα target genes, including TFF1, CCND1, THSB1, CTGF, CCL2, and EGR3 (**Fig. 4-13a**). Both FGF3 and FGF19 induced EGR3, CCND1, and THSB1 mRNA; this induction was inhibited by treatment with lucitanib, fulvestrant, or the combination (**Fig. 4-13b** and **c**), and also by transfection of a TK dead K514M FGFR1 mutant into CAMA1 cells (**Fig. 4-9d**). In line with their higher levels of ERα, FGFR1, and FGF3/4/19 (**Figs. 4-4c** and **4-6d**), CAMA1LTED cells expressed higher levels of ERα-regulated genes than CAMA1 parental cells (**Fig. 4-13d**). Finally, to support our results with lucitanib were not due to off-target effects of the small molecule, we tested the FGFR inhibitor INCB054828.[219(p054828)] Treatment with INCB054828 also blocked FGF3-induced pFRS2, CAMA1 cell growth, and ERα target gene expression (**Fig. 4-14a–c**).

We next examined the effect of the knockdown of FGFR1 on ERα transcriptional activity. Compared with scrambled control siRNA, knockdown of FGFR1, but not of its major signal transducer FRS2, reduced ERE-luciferase reporter activity (**Fig. 4-15a**). FGFR1 and FRS2 downregulation was confirmed by immunoblot or RT-PCR, respectively (**Fig. 4-15b** and **c**). The inability of siFRS2 to reduce ER reporter activity suggested an MEK-independent and PI3K-independent role of the FGFR1 TK on ERα transcriptional function. Supporting this speculation, treatment of CAMA1 cells with lucitanib reduced mRNA levels of the ERα-regulated genes CCND1 and THSB1 more potently than the MEK1/2 inhibitor trametinib[220] and the pan-PI3K inhibitor buparlisib (**Fig. 4-15e**).[221] In parallel experiments, we confirmed by immunoblot analyses drug-mediated inhibition of their molecular targets: ERα for fulvestrant, pFRS2 and pERK for lucitanib, pERK for trametenib, and pAKT for buparlisib (**Fig. 4-15d**).

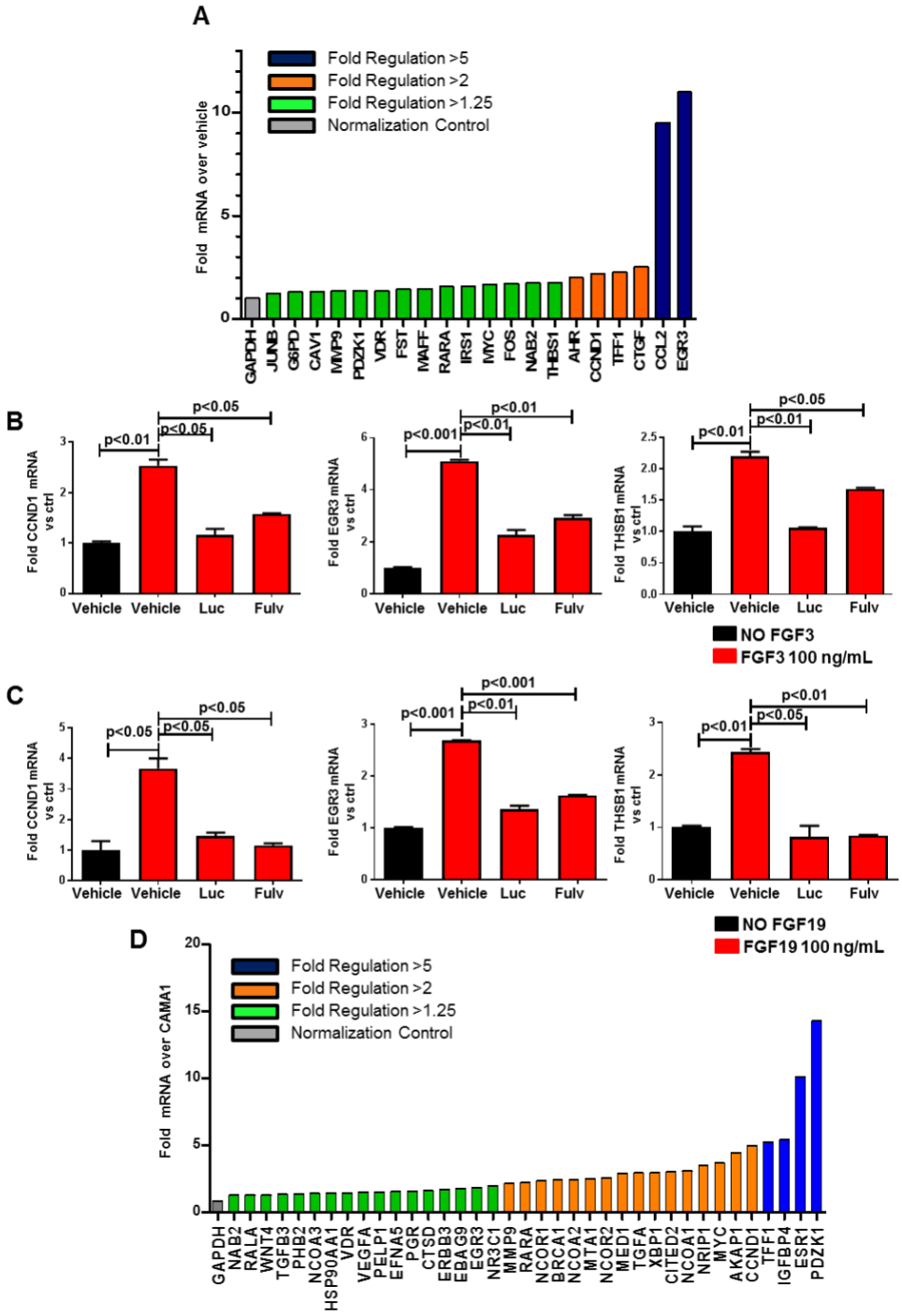


Figure 4-13. Treatment with FGFs induces expression of ERα-dependent genes. **A**, CAMA1 cells were plated in estrogen-free medium and treated with 100 ng/mL FGF3/19 for 6 h. At this time, cells were harvested and RNA prepared and analyzed for mRNA expression changes in ERα pathway genes using the RT2 Profiler Estrogen Receptor Signaling PCR Array (Qiagen). **B-C**, CCND1, EGR3 and THSB1 mRNA expression was confirmed by qRT-PCR in CAMA1 cells treated with FGF3 (**B**) or FGF19 (**C**) for 6 h ± 2 μM lucitanib or 1μM fulvestrant. Each bar represents the mean CCND1, EGR3 and THSB1 transcript levels ± SD (*p<0.05, **p<0.01, ***p<0.001 vs. vehicle without FGF, Student's t-test). **D**, CAMA1 and CAMA1LTED cells were plated in estrogen-free medium and treated with 100 ng/mL FGF3 for 6 h. At this time, cells were harvested and RNA prepared and analyzed for mRNA expression changes in ERα pathway genes using the RT2 Profiler Estrogen Receptor Signaling PCR Array (Qiagen).

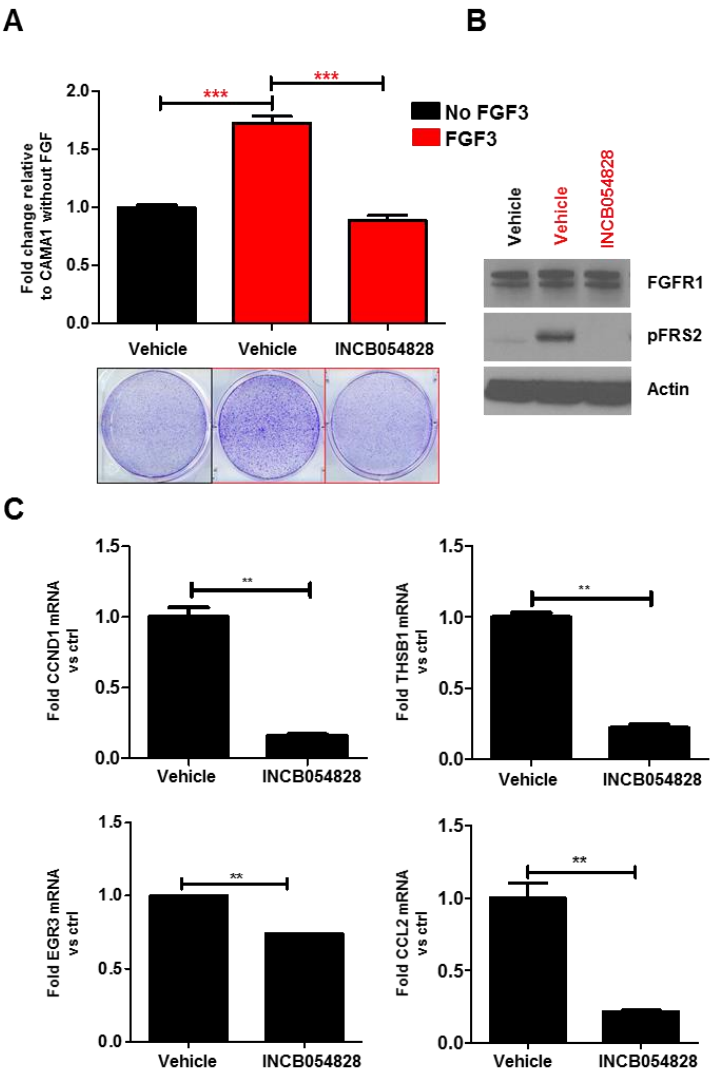


Figure 4-14. Treatment with INCB054828 also blocked FGF3-induced pFRS2, CAMA1 cell growth, and ERα target gene expression. **A**, CAMA1 cells were treated with 100 ng/mL FGF3 ± 1 μM INCB054828 in estrogen-free medium. After 15 days, plates were washed and stained with crystal violet and their imaging intensity was quantified as described in Methods. Representative images and quantification of the imaging intensity values as % of vehicle-treated controls are shown (**p<0.01 vs. controls, Student's t-test). **B**, CAMA1 cells in identically treated parallel plates were treated for 6 h after which lysates were prepared and subjected to immunoblot analyses with the indicated antibodies. **C**, CAMA1 ± 1 μM INCB054828 were plated in estrogen-free medium and treated with 100 ng/mL FGF3 for 6 h. At this time, cells were harvested and RNA was prepared and analyzed for THBS1, CCND1, CCL2 and EGR3 mRNA changes by qRT-PCR. Each bar represents the mean transcript level ± SD (*p<0.05, **p<0.01, ***p<0.001 vs. $CAMA1^{GFP}$, Student's t-test).

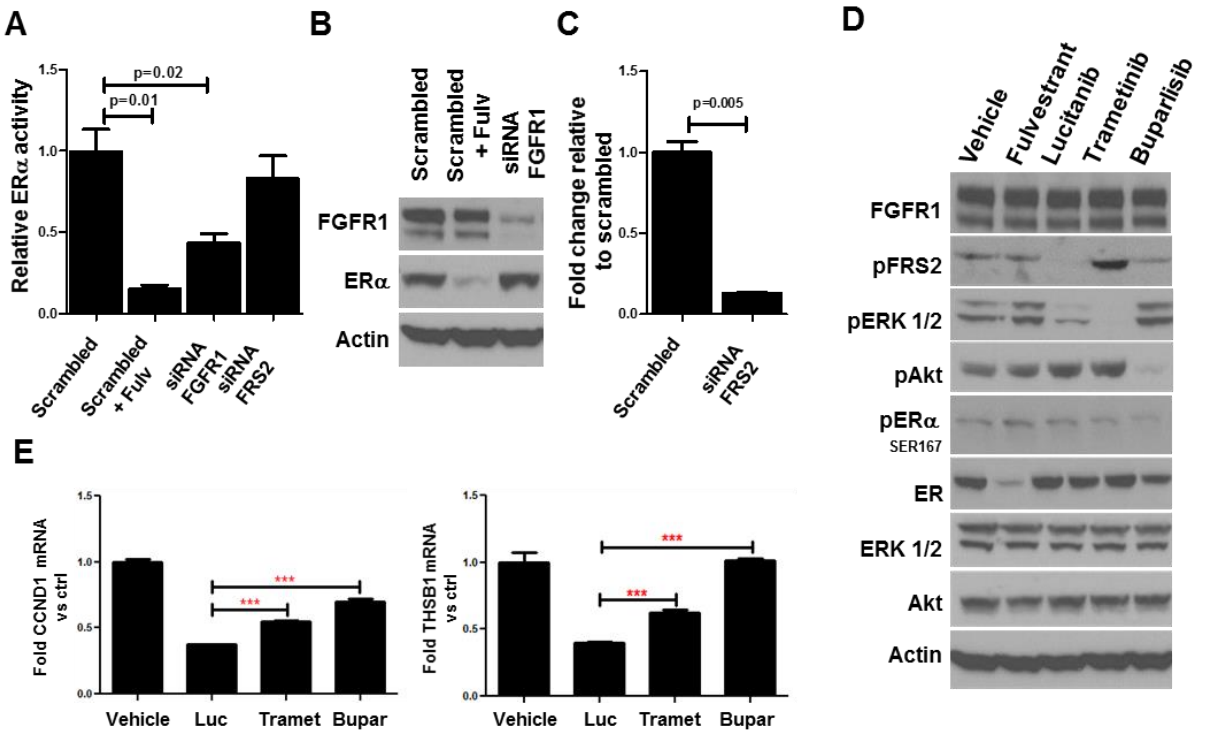


Figure 4-15. FGFR1 signaling is associated with ERα transcriptional activity. **A-C**, CAMA1 cells were transfected with siRNA specific for FGFR1 or FRS2 or a non-targeting (scrambled) control followed by assessment of ERα reporter activity as described in Methods. Treatment with 1 μM fulvestrant reduced ERα reporter activity and was used as a positive control. Each bar represents the mean luciferase activity ± SD of five replicate wells each read twice. Cells from identically-treated parallel plates were lysed for immunoblot analysis (**B**) or qPCR (**C**) to detect FRS2 protein and mRNA, respectively. **D**, CAMA1 cells were plated in estrogen-free medium containing FGF3 (100 ng/mL) and treated with vehicle, 1 μM fulvestrant, 2 μM lucitanib, 1 μM buparlisib or 0.5 μM trametinib for 18 h. Cells were then harvested and lysates prepared followed by SDS-PAGE, transfer to nitrocellulose, and immunoblot analyses with the indicated antibodies as described in Methods. **E**, CCND1 and THSB1 mRNA expression was confirmed by qRT-PCR in CAMA1 cells treated with FGF3 for 6 h ± 2 μM lucitanib, ± 1 μM buparlisib or ± 0.5 μM trametinib. Each bar represents the mean CCND1 and THSB1 transcript levels ± SD (*p<0.05, **p<0.01, ***p<0.001 vs. vehicle without FGF, Student's t-test).

Combined blockade of FGFR1 and ERα potently inhibits growth of ER+/FGFR1–amplified breast cancers

To follow the effect of fulvestrant and lucitanib on ER-dependent gene expression, we next examined whether FGFR1 and/or ERα inhibitors would have an effect on ER+/FGFR1–amplified tumor cell growth. Treatment with the combination of lucitanib and fulvestrant suppressed CAMA1 colony formation in 3D Matrigel significantly more potently than each drug alone (**Fig. 4-16a** and **b**). Western blot analysis of lysates from cells treated for 6 hours showed that only the combination simultaneously reduced levels of pFRS2, pERK1/2, and ERα (**Fig. 4-16c**). We next examined the effect of these drugs against two ER+/HER2−/FGFR1–amplified PDXs, T272 and TM00368 (**Fig. 4-16d**). Ovariectomized mice with established xenografts (≥250 mm3) were treated with vehicle, lucitanib, fulvestrant, or both drugs. PDX T272 but not PDXTM00368 required brief estrogen supplementation to generate tumors. In mice bearing PDX T272, the dose of lucitanib was reduced from 10 to 7 mg/kg/day after 3 weeks of therapy due to toxicity in both lucitanib-containing arms. Mice with TM00368 PDXs were treated with 7 mg/kg/day lucitanib. Treatment with the combination of fulvestrant and lucitanib inhibited growth of both PDXs more potently than either drug alone (**Fig. 4-16e**; **Fig. 4-17a**). All mice bearing TM00368 xenografts exhibited a ≥50% reduction in tumor size from baseline after 3 weeks of treatment with fulvestrant/lucitanib (**Fig. 4-16f**). Biomarkers of response were assessed by IHC in TM00368 tumors harvested at the completion of therapy. Treatment with the combination of lucitanib plus fulvestrant markedly reduced detectable levels of Y653/4 p-FGFR1 and total ERα (**Fig. 4-16g** and **h**). FGFR1 antibody pulldowns of tumor lysates from vehicle- and lucitanib-treated mice coprecipitated ERα. This was not observed in tumors

treated with fulvestrant or the combination (**Fig. 4-16i**). No change in mouse weight was observed in any of the treatment arms (**Fig. 4-17b** and **c**).

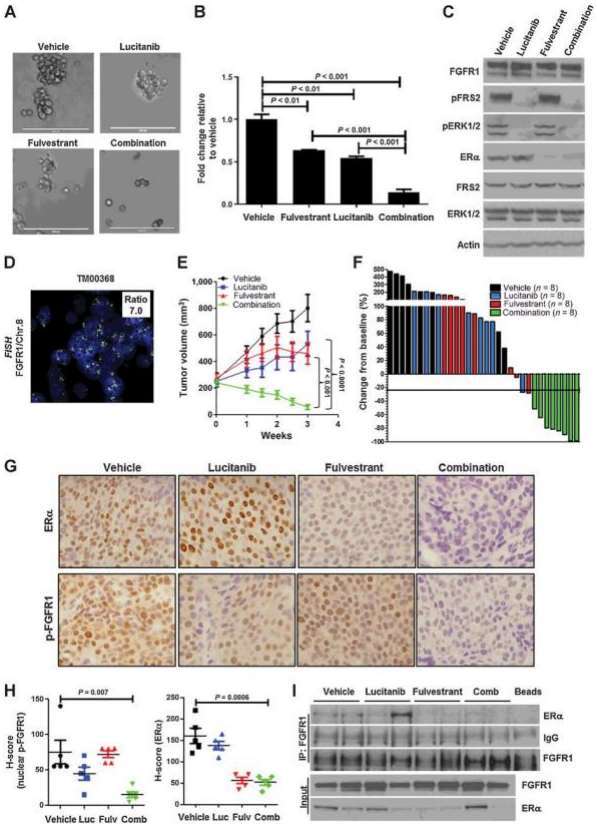


Figure 4-16. Combined blockade of FGFR1 and ERα potently inhibits growth of ER+/FGFR1-amplified breast cancers. **A** and **B**, CAMA1 cells were cultured in 3D Matrigel as described in Materials and Methods and treated with vehicle, 2 μmol/L lucitanib, 1 μmol/L fulvestrant, or the combination. After 15 days, images were captured from three different fields using a CK40 microscope. Quantitation of representative images is shown in **B**. Each bar represents the fold change in colony number relative to vehicle ± SD of three replicate wells repeated twice (Student t test). **C**, CAMA1 cells were treated as in **A** and **B** for 6 hours, after which lysates were prepared and subjected to immunoblot analyses with the indicated antibodies. **E**, ER+/HER2−/FGFR1-amplified TM00368 PDXs were established in ovariectomized SCID/beige mice implanted with a subcutaneous 21-day release, 0.25-mg 17β-estradiol pellet. Once tumors reached ≥200 mm3, mice were randomized to treatment with vehicle, fulvestrant (5 mg/kg/week), lucitanib (7 mg/kg/day), or both drugs for 3 weeks. Each data point represents the mean tumor volume in mm3 ± SD (n = 8 per arm; ANOVA test). **F**, Bar graph showing the percent change in volume in individual TM00368 PDXs after 3 weeks of treatment relative to tumor volumes on day 0 (baseline). **G** and **H**, TM00368 tumors were harvested at the end of treatment. FFPE tumor sections were prepared and subjected to IHC with Y653/4 phosphorylated FGFR1 and ERα antibodies as described in Materials and Methods. The percent of phospho-FGFR1+ and ER+ tumor cells and their staining intensity was assessed by an expert breast pathologist (M.M. Estrada) blinded to the treatment to generate an H-score. Nuclear phospho-FGFR1 and ERα H-scores are shown (Student t test). **I**, FGFR1 was precipitated from lysates of TM00368 tumors harvested at the end of treatment; immune complexes were separated by SDS-PAGE and subjected to immunoblot analysis with the indicated antibodies. Bottom two lanes show FGFR1 and ERα content in lysates before intraperitoneal injection.

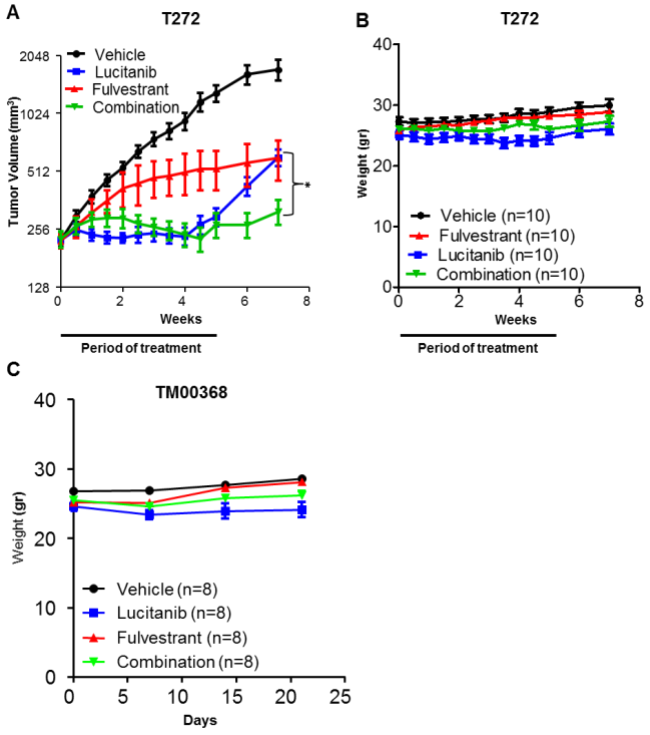


Figure 4-17. Treatment with the combination of fulvestrant and lucitanib inhibited growth of both PDXs more potently than either drug alone. **A**, ER+/HER2−/FGFR1-amplified T272 PDXs were established in female athymic nude mice supplemented with estrogen 8.5 mg/L in the drinking water. Once tumors reached ≥200 mm3, mice were randomized to treatment with vehicle, fulvestrant (5 mg/kg/week), lucitanib (10 mg/kg/day), or both drugs for 5 weeks. Each data point represents the mean tumor volume in mm3 ± SD (n=10 per arm; *p<0.05 vs. lucitanib; Student's t-test). **B-C** Weight of SCID/beige mice bearing T272 or TM00368 PDXs during treatment with vehicle, fulvestrant, lucitanib or the combination for a total of 7 and 3 weeks, respectively. The number of mice in each treatment arm is shown in parentheses. Each data point represents mean weight in grams ± SD.

Discussion

We report herein a novel mechanism by which FGFR1 amplification confers resistance to antiestrogens in ER+ breast cancers. In a cohort of postmenopausal patients treated with the aromatase inhibitor letrozole, cancers with FGFR1 amplification retained tumor cell proliferation, suggesting aberrant FGFR1 signaling is associated with resistance to estrogen deprivation. Short and long-term estrogen deprivation

increased total and nuclear FGFR1 and FGF ligand expression in ER+/FGFR1–amplified breast cancer cells and primary tumors. This was associated with an increase in nuclear FGFR1/ERα complexes and maintenance of estrogen-independent transcription of ER-responsive genes. The interaction between FGFR1 and ERα was blocked by a kinase-dead FGFR1 mutant or by FGFR TKIs. ChIP-seq analysis of FGF-stimulated FGFR1-amplified cells showed binding of FGFR1 and of ERα to DNA, which was inhibited by the FGFR TKI lucitanib and by the ER downregulator fulvestrant, respectively, suggesting a possible interdependence between FGFR1 and ERα at transcription start sites. Of note, RNA-seq data from ER+/FGFR1–amplified tumors from patients treated with letrozole suggested the ERα pathway is still active (**Fig. 4-11g** and **h**), thus providing a plausible explanation for maintenance of proliferation in these estrogen-deprived cancers. Finally, dual pharmacologic inhibition of FGFR1 and ERα potently inhibited growth of ER+/FGFR1–amplified breast cancer cells and PDX models, supporting the clinical development of this combination in patients with this subtype of breast cancer.

FGFR1 in association with nuclear proteins, such as RSK1 and CBP, has been shown to induce gene expression in other cancers.[217] As FGFR1 inhibition reduced the transcription of ERα-related genes (**Fig. 4-13b** and **c**), we speculated the previously reported transcriptional function of FGFR1[197,200,217,220] may play a role in resistance to estrogen deprivation. Of note, we precipitated both FGFR1 and ERα with C-terminal and N-terminal FGFR1 antibodies from FGFR1-amplified CAMA1 cell nuclei (**Fig. 4-8b**). These findings were supported by PLA and confocal microscopy studies (**Fig. 4-8c**). Inhibition of FGFR1 TK activity with lucitanib and expression of a TK dead K514M FGFR1 mutant into CAMA1 cells reduced ERα-dependent gene transcription (**Fig. 4-13b** and **c**; **Fig. 4-9d**) and inhibited the association of FGFR1 with ERα (**Fig. 4-8e–h**). Taken together, these data support a novel TK-dependent role of nuclear FGFR1 on ERα-dependent gene transcription in estrogen-independent ER+/FGFR1-amplified breast cancers.

To the best of our knowledge, this is the first report of a physical association of FGFR1 and ERα associated with antiestrogen resistance. It follows studies supporting both the nuclear localization and function of FGFR1. FGFR1 can enter the nucleus by retrograde transport from the endoplasmic reticulum lumen to the cytosol via Sec61p channels before endoplasmic vesicles deliver the receptor to the plasma membrane.[198,221] This process is possible because of the atypical TMD of FGFR1, which consists of nonpolar amino acid chains interrupted by polar regions in a β-sheet structure, thus allowing mobilization of the receptor out of the membrane.[198,221] Cell surface biotinylation assays show that nuclear FGFR1 can also originate from the cell surface[222], suggesting FGFR1 is internalized and traffics to the nucleus via endosomal pathways. Indeed, FGFR1 and FGFR2 can translocate to the nucleus following ligand stimulation in pancreatic stellate cells; this process requires the interaction of FGFR1 with nuclear import proteins, like importin β.[196,223] Once in the nucleus, FGFR1 has been shown to regulate gene transcription.[197,200(p1),217,220] Nuclear targeting of FGFR1 by substituting its signal peptide for a nuclear localization sequence is sufficient to initiate DNA synthesis and transcription of c-Jun, an activator of cyclin D1. Removal of the kinase region of nuclear-targeted FGFR1 ablates this effect.[224] These data suggest the TK function of FGFR1 is necessary for its transcriptional role, consistent with our data from ER$^+$/*FGFR1*–amplified breast cancer cells shown herein.

In summary, we have identified a mechanism by which amplified FGFR1 can sustain estrogen-independent breast tumor growth. We propose this mechanism explains, in part, the limited effects of estrogen deprivation on ER+/FGFR1–amplified breast cancers in the clinical trial with letrozole. On the basis of these data, we propose combinations of ERα and FGFR antagonists should be tested in patients with ER+/FGFR1–amplified breast cancer.

CHAPTER V

RSK2 MAINTAINS ADULT ESTROGEN HOMEOSTASIS BY INHIBITING ERK1/2-MEDIATED DEGRADATION OF ESTROGEN RECEPTOR ALPHA

This section is a paper published in *Cell Reports* as "RSK2 Maintains Adult Estrogen Homeostasis by Inhibiting ERK1/2-Mediated Degradation of Estrogen Receptor Alpha" Katarzyna A. Ludwik, Zachary M. Sandusky, Kimberly M. Stauffer, Yu Li, Kelli L. Boyd, George A. O'Doherty, Thomas P. Stricker, Deborah A. Lannigan.

Summary

In collaboration with Katarzyna Ludwik of the Deb Lannigan laboratory, I investigated the transcriptional divergences between non-clonogenic luminal (NCL) cells of mice in either estrus or diestrus, with WT-RSK2 or RSK2-KO, to delineate the relationship between estrogen-responsive gene expression, estrogen abundance, and RSK2 status. We hypothesized that a negative regulatory mechanism involving RSK2 must exist to limit ERα degradation and maintain estrogen responsiveness. My RNA-seq analysis highlighted that RSK2 negatively regulates estrogen-responsive gene expression in both mouse and human mammary cells. This regulation occurs through redox homeostasis and therefore prevention of ERK1/2 activation. These results have implications for the increased risk for breast cancer in people consuming exogenous estrogens, as low RSK2 levels are associated with an increase in DNA damage.

Introduction

The importance of estrogen signaling is highlighted by the numerous physiological alterations that occur during menopause, oophorectomy, or anti-estrogen therapy.[225] In the adult, human estrogen levels are highest in the follicular phase, reaching a level of 1 nM and decrease approximately 5-fold in the luteal phase of the menstrual cycle.[226] In the mouse, the estrous cycle is divided into four stages, which are based on vaginal cytology and comprise proestrus, estrus, metestrus, and diestrus. The highest level of estrogen 0.2 nM occurs during proestrus and then decreases by approximately 3-fold in diestrus.[227] All estrogen-receptor-positive (ER+) tissues respond to fluctuations in estrogen levels. The mammary gland undergoes extensive morphological changes as estrogen levels change[228]; therefore, it is an ideal organ in which to investigate the mechanisms that regulate estrogen homeostasis. Estrogen acts primarily through the steroid hormone receptors, estrogen receptor alpha (ERα) and ERβ.[229] In the mammary gland, ERα is of particular importance for its contributions to gland development.[230] Mammary gland development is normal in the absence of ERβ[231]; therefore, to examine estrogen homeostasis in the mammary gland, we focused our studies on ERα.

Estrogen binding to the receptor results in ERα degradation through the 26S proteasome pathway, and both are required for activation of estrogen-responsive gene expression.[232–236] However, it is puzzling how this degradation-coupled transcription is regulated to maintain ERα protein levels because, theoretically, estrogen levels are sufficient throughout the menstrual cycle to continuously drive degradation.[237] Therefore, it might be expected that ERα levels would eventually drop below that required to generate a physiological response. However, both ERα levels and estrogen responsiveness are maintained to allow progression into the next menstrual cycle, but the mechanisms regulating ERα degradation are unknown.[238,239] Maintenance of responsiveness is of particular relevance to individuals taking estrogen-containing oral contraceptives in which estrogen levels do not fluctuate compared with that of the normal menstrual cycle[226] and in transgendered individuals, in which estrogen levels can reach supra-physiological levels.[240]

We hypothesized that a negative regulatory mechanism must exist to limit ERα degradation to preserve ERα levels and, as a result, maintain estrogen responsiveness. To identify that mechanism, we focused on estrogen and its control of the epidermal growth factor receptor (EGFR) signaling pathway because of the importance of EGFR in mammary gland development[241], cell fate specification[242] and breast cancer.[243] Stimulation of EGFR activates the MEK-ERK1/2 signaling cascade. Activated ERK1/2 and its downstream effector, p90 ribosomal S6 kinase (RSK), directly phosphorylate ERα at Ser-118 and Ser-167, respectively.[244,245] These sites increase ERα transcriptional activity in cell-based systems. In a transgenic mouse model RSK2 nuclear accumulation in the mammary gland drives high-grade ER+ ductal carcinoma in situ.[113] Because cancer often exploits mechanisms important in development and homeostasis, we investigated the contributions of ERK1/2-RSK2 signaling to normal ERα biology.

Unexpectedly, we discovered a novel regulatory mechanism in which the ERK1/2-RSK2 pathway acts as a developmentally regulated switch that is required for maintaining ERα protein levels in the mammary gland and uterus in the adult but not in the juvenile. ERK1/2 is activated during the estrus phase of the cycle as a consequence of an estrogen pulse that occurs in proestrus. Activated ERK1/2 phosphorylates ERα, driving the degradation of ERα and estrogen-responsive gene expression. To enable estrogen responsiveness for the subsequent cycle ERK1/2 is inactivated when estrogen levels are low. Active RSK2 limits the response to estrogen by maintaining redox homeostasis, which prevents ERK1/2 activation in response to reactive oxygen species (ROS). In the RSK2 knockout (RSK2- KO) and in individuals taking oral contraceptives decreased RSK2 levels are correlated with an enriched signature for estrogen-responsive gene expression. These observations may explain the mechanism underlying the increase in breast cancer risk that is observed for individuals taking exogenous estrogens because reduced RSK2 is correlated with increased DNA damage.

Methods and Materials

Key Resources Table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rat anti-keratin 8	University of Iowa	TROMA-I; RRID:AB_531826
Chicken anti-keratin14	BioLegend	SIG-3476; RRID:AB_10718041
Rabbit anti-pRSK (Thr359/Ser363) (Tris)	Santa Cruz Biotechnology, Inc.	sc-12898-R; RRID:AB_2181303
Mouse anti-ERα 6F11 (Citrate)	Thermo Fisher Scientific Inc.	MA5-13304; RRID:AB_11002193
Mouse anti-γH2A.X (Ser139) (Tris)	EMD Millipore	JBW301; RRID:AB_568825
Rabbit anit-pERK1/2 (pTEpY) (Tris)	Promega	V803A; RRID:AB_2335893
Rabbit anti-peEF2 (Thr56) (Tris)	Cell Signal Technology	2331; RRID:AB_10015204
Mouse anti-GATA3 (Tris)	Thermo Fisher Scientific Inc.	1A12-1D9; RRID:AB_2536713
Rabbit anti-AR	Thermo Fisher Scientific Inc.	MA5-13426; RRID:AB_11000751

Rabbit anti-E cadherin	Cell Signal Technology	3195; RRID:AB_2291471
Mouse anti-ERK	BD Biosciences	610124; RRID:AB_397530
Donkey anti-rabbit 647	Invitrogen	A31573; RRID:AB_2536183
Donkey anti-mouse 647	Invitrogen	A31571; RRID:AB_162542
Goat anti-rat 546	Invitrogen	A11081; RRID:AB_2534125
Goat anti-chicken 488	Invitrogen	A11039; RRID:AB_2534096
Biotin anti-CD140	Biolegend	APA5; RRID:AB_11211998
Biotin anti-CD31	Biolegend	MEC13.3; RRID:AB_312910
Biotin anti-Ter-119	Biolegend	TER-119; RRID:AB_313704
Biotin anti-CD45	Biolegend	30-F11; RRID:AB_312968
Brilliant Violet 510 Streptavidin	Biolegend	405233
Anti-Sca1-PerCP	Biolegend	108121; RRID:AB_893618
Anti-Sca1-FITC	Biolegend	D7; RRID:AB_313342
Anti-CD49b-APC/Cy7	Biolegend	DX5; RRID:AB_313416
Anti-EpCAM-APC	Biolegend	G8.8; RRID:AB_1134105
Anti-CD49f-PE/Cy7	Biolegend	GoH3; RRID:AB_2561704
Chemicals, Peptides, and Recombinant Proteins		
Cell Trace Violet	Life Technologies Corp.	C34557
Zombie Yellow	Biolegend	423104
EdU (5-Ethynyl-2′-deoxyuridine)	Life Technologies Corp.	NEO87011604
Bortezomib (PS-341)	Calbiochem	50-431-40001
BI-D1870	Enzo Life Sciences	BML-EI407
Trametinib	Selleck Chem	S2673
U0126	Sigma	U120
MG-132	Calbiochem	474790
17-β estradiol (E2)	Sigma	E2758
Phorbol 12-myristate 13-acetate (PMA)	Sigma	P1585
EGF	Calbiochem	324831
FGF7	R&D Systems	251KG010CF
Critical Commercial Assays		
Click-iT Plus EdU Alexa Fluor 488 Flow Cytometry Assay Kit	Thermo Fisher Scientific Inc.	C10632
Click-iT Plus OPP Alexa Fluor 647 Protein Synthesis Assay Kit	Thermo Fisher Scientific Inc.	C10458
CellROX Green Reagent	Thermo Fisher Scientific Inc.	C10444
RNeasy Micro Kit	QIAGEN	74004
Deposited Data		

RNA sequencing data	This paper	GEO: GSE113323
Experimental Models: Cell Lines		
Mouse: TM3 cell line	ATCC	CRL-1714
Experimental Models: Organisms/Strains		
Mouse: RSK2-KO: C57BL/6JRSK2−/−	Andre Hanauer, PhD. Institut de Genetique et Biologie Moleculaire et Cellulaire, C.U. de Strasbourg, France	N/A
Oligonucleotides		
f-GAPDHm	AGAACATCATCCCTGCATCCA	N/A
r-GAPDHm	CAGATCCACGACGGACACATT	N/A
f-GATA3m	GATGTAAGTCGAGGCCCAAG	N/A
r-GATA3m	GCAGGCATTGCAAAGGTAGT	N/A
f-ESR1m	TTACGAAGTGGGCATGATGA	N/A
r-ESR1m	CCTGAAGCACCCATTTCATT	N/A
Recombinant DNA		
pLVTHM	Wiznerowicz and Trono, 2003	Addgene Cat #12247
psPAX2	Provided by Dr. Didier Trono	Addgene Cat #12260
pMD2.G	Provided by Dr. Didier Trono	Addgene Cat #12259
Software and Algorithms		
LSM-FCS/ ZEN	Carl Zeiss, Inc.	N/A
Openlab 5.5.0 / Volocity 6.2.1	PerkinElmer Inc.	N/A
GraphPad Prism 6.0a	GraphPad Spftware Inc.	N/A
Morpheus	Broad Institute	https://software.broadinstitute.org/morpheus/
BioRender	BioRender	https://biorender.com/

Mice

All procedures involving animals were done in accordance with current federal (NIH Guide for Care and Use of Laboratory Animals) and university guidelines and were approved by the University of Virginia and Vanderbilt University Institutional Animal Care and Use Committee.

Female WT or RSK2-KO mice[246] between six and fourteen weeks old were studied. The age of animals in specific experiments are indicated in the figures with adult animals ranging from twelve to fourteen weeks. For whole mount analysis the 4th mammary gland was fixed and stained in Carmine Alum. Ductal distance was measured from the nipple to the tip of the longest duct. The number of secondary branches along the longest primary branch were counted.

The stages of the estrous cycle were determined by cytological analysis of vaginal swabs.[247,248] For all experiments requiring matched estrous stages, the cycles were monitored for 2 weeks prior to end point to ensure continuous cycling.

Mammary epithelial cells were isolated with modifications.[242] Briefly, mammary glands were isolated from donor mice, minced, and digested in DMEM/F12 supplemented with 2mg/ml Collagenase A and 100U/ml Pen/Strep for 2.5h in 37°C 5% CO2 incubator. Digested material was pelleted at 180 g for 5 min and the pellet was suspended in DNase I (1000U/ml) for 3-5 min in 37°C in 5% CO2. Fetal bovine serum (FBS) was added and the digested tissue was pelleted at 180 g for 10 min. The pellet was washed with phosphate-buffer saline, pelleted, suspended in Accumax (StemCell Technologies Inc.) and placed in Thermomixer at 37°C for 10 min. Digested material was pelleted at 180 g for 3 min, suspended in 5x trypsin for 5 min at 37°C. Trypsin was quenched with FBS and cells were pelleted and suspended in phosphate buffered saline (PBS) or DMEM/F12. The cell preparation was filtered through 70-μm mesh to obtain single cell suspensions. For mammary gland regenerations, 4x107 cells/ml of single cells in DMEM/F12 were mixed 1:1 with matrigel. 10 μL of cell suspension in matrigel was injected into the cleared 4th mammary fat pad of a recipient 3wk old mouse (Brill et al., 2008).[249]

To inhibit the 26S proteasome pathway or RSK1/2 *in vivo* female mice in estrus (12 wk) were injected intraperitoneally (IP) with vehicle or PS-341 at 5 mg/kg in 2% DMSO, 30% PEG, and 68% saline or *C5″-n*-propyl cyclitol SL0101 at 40mg/kg in one part DMSO and nine parts 25% hydroxypropyl-beta-cyclodextrin. Animals in the PS-341 study were euthanized 4h after injection and animals in the RSK1/2 study were injected twice at 7 h interval before euthanasia.

Cell Line Studies

TM3 cells were purchased and cultured according to ATCC. Cells were maintained in log-phase and screened for Mycoplasma by PCR. Prior to experiments, cells were serum-starved in phenol red-free media for 48 h followed by addition of vehicle, C5”-n-propyl-cyclitol SL0101 (20 μM, 6h), BI-D1870 (10 μM, 6h), trametinib (1 μM, 6h), or U0126 (10 μM, 6h). In experiments with MG-132, cells were pretreated (10 μM, 1h). For analysis of Ser-118 upshift, cells were serum-starved as above and treated with phorbol 12-myristate 13-acetate (PMA) (0.5 μM, 20 min), EGF and FGF7 cocktail (12.5 nM each, 5 min), C5” (20 μM for 2 h). In experiments with trametinib, cells were pretreated (1 μM, 2h). Cells were lysed and analyzed.[244]

Transduction

Constructs to generate lentivirus including psPAX2, pMD2.G, and pLVTHM were provided by D. Trono, Ph.D. (Swiss Institute of Technology, Lausanne, Switzerland). The pLV-Venus lentivirus construct was provided by Ian Macara, Ph.D. (Vanderbilt University, Nashville, TN). Lentiviral production was performed using Lipofectamine 3000 (Invitrogen) according to the manufacturer’s instructions. S118A-ERα and S167A-ERα were generated using Q5 site-directed mutagenesis.

Fluorescence Activated Cell Sorting (FACS)

For FACS, single epithelial cells (106 cells/ml) obtained from mammary glands in PBS were incubated with Cell Trace Violet (1 μM) and Zombie Yellow (1:250) for 20 min at room temperature. Cells were washed and suspended in 5% FBS in PBS. Cells were blocked with 10% normal rat serum in 5% FBS for 10 min at 4°C, followed by incubation with biotin-conjugated primary antibodies against lineage markers for 10 min at 4°C. The cells were incubated with primary antibodies for 20 min at 4°C, washed and suspended in 5% FBS. Cells were analyzed using FACSCantoII or sorted using FACSAriaII. Flow cytometry data were analyzed using Cytobank version 6.2. Further reagents details are provided in the Key Resources Table.

EdU labeling was performed[250] in mice staged in proestrus were injected intraperitoneally with 10 mg/ml EdU in PBS (100 mg/kg) and then administered EdU in the drinking water (1mg/ml). The estrus stage was monitored, and mammary glands were isolated in metestrus (2 days after EdU injection). Mammary cells were isolated and analyzed for EdU incorporation

using the Click-iT Edu Flow Cytometry Assay Kit, followed by the antibody staining as described above carried out in 1xClick-iT saponin based permeabilization buffer. Further reagents details are provided in the Key Resources Table.

Immunostaining

Mouse organs were fixed in buffered 10% formalin for 2 d and then placed in 70% ethanol. The fixed samples were paraffin-embedded, and sectioned. Sections were deparaffinized and antigen retrieval performed in tris-EDTA buffer pH 8.0 or citrate buffer pH 6.0 or pH 7.0 (Key Resources Table). The sections were blocked in 10% bovine serum albumin (BSA) in PBS and incubated with primary antibody in 3% BSA in PBS o/n at 4°C. The sections were washed and incubated with secondary antibody for 1 h in room temperature. For detection of Venus-tagged ERα in TM3 cells, $1x10^4$ cells were seeded on laminin-coated glass coverslips. After treatment, cells were fixed in 4% PFA in PBS (pH 7.4, 15 min). Antibodies are listed in the Key Resources Table. For immunofluorescence staining, cells were fixed in 4% PFA in PBS (pH 7.4, 15 min) and permeabilized with 0.1% Triton X-100 in PBS (15 min), DNA was stained with Hoechst in PBS (10 min) and coverslips mounted using Fluoro-Gel (Electron Microscopy Sciences). Images were collected with a laser-scanning microscope (LSM 510/Meta/FCS, Carl Zeiss Inc.).

RNA Analysis

For RNA isolation, 5x104 EpCAMhiCD49f+Sca1+Cd49b- cells were FACS sorted and total RNA extraction (RNeasy Micro Kit) was performed. The RNA quality was tested using Agilent 100 Bioanalyzer (RIN 8). Libraries were constructed and sequenced by Genewiz LLC. Reads were aligned to the mm10 mouse genome with STAR, the transcripts were assembled using Gencode version 15 as gene models. Genes and transcripts were quantified with HTSeq. Two samples were clear outliers and were discarded. Batch correction was done with SVA, and differential gene expression analysis was performed with DESeq2. Gene set enrichment was done with GSEA using MSigDB and GSVA using GSKB mouse gene sets. RNASeq data is available at Gene Expression Omnibus under accession GSE113323.

For qRT-PCR RNA (1 μg) was reverse transcribed using High Capacity cDNA Reverse Transcription Kit. Analysis was performed using IQ RealTime SyberGreen PCR Supermix (BioRad Laboratories) on the C1000Thermal Cycler CFX96 Real-Time System.[113] The ΔΔCt was calculated using GAPDH as a control. Primers are listed in the Key Resources Table.

Raw reads from the sequencing of normal breast tissue at different stages of menstrual cycle[251] were normalized using DESeq2 according to the estimated size of the libraries. Based on unsupervised hierarchical clustering, 5 samples were rejected as outliers and Z-scores were calculated correcting for sequencing batch.

Quantification and Statistical Analysis

Statistical analyses were performed using GraphPad Prism 6. The statistical test used is reported in the figure legends. Additional ANOVA values for complex comparisons are provided (Table S1).

Results

RSK2 Is Required to Maintain ERα Homeostasis in the Adult Mammary Gland

In the mouse, estrogen levels are highest during proestrus, akin to the follicular stage in humans.[252] Analysis of ERα in the mammary gland of wild-type (WT) mice in situ using quantitative immunofluorescence (IF) revealed that ERα protein levels varied during the estrous cycle (**Fig. 5-1a and 1b**). In the WT mice, the lowest ERα protein levels occurred during estrus,

which is consistent with observations that ERα protein degradation increases in response to the estrogen pulse in proestrus (**Table 5-1**).[233] Staging of the estrous cycle was determined by analysis of vaginal cytology and uterine wet weight (**Fig. 5-2a**).[247] WT and RSK2-KO mice moved through the estrous cycle in a similar manner (**Fig. 5-2b**). In the RSK2-KO glands, ERα levels were consistently lower than in the WT glands across all estrous stages (**Fig. 5-1a, b,** and **5-2c**). These results were unexpected because RSK2 phosphorylation of ERα stimulates transcription[244] and would, presumably, increase ERα degradation. Therefore, based on these observations, we would expect that loss of RSK2 would increase ERα protein levels.

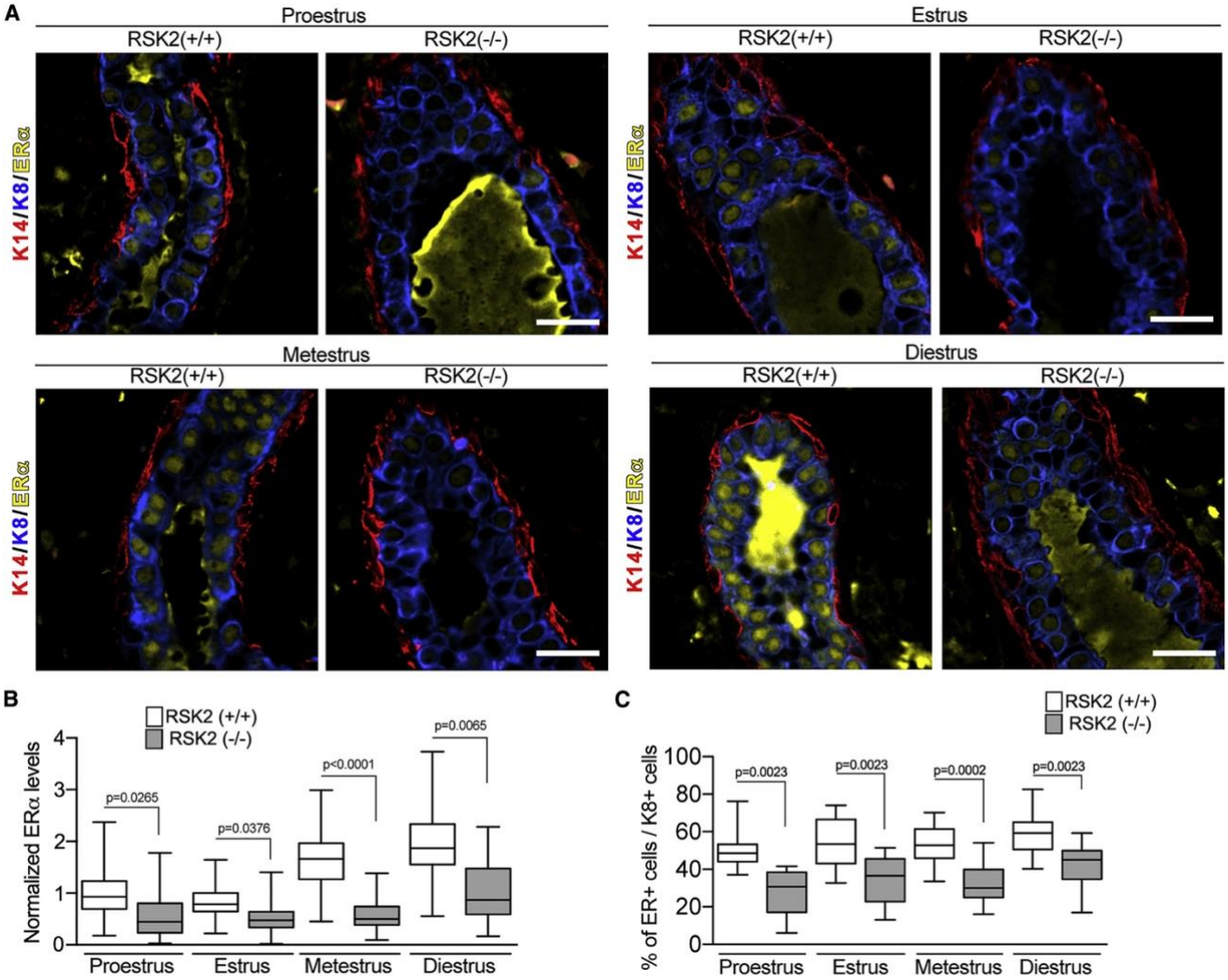


Figure 5-1. RSK2 Regulates ERα Protein Levels in the Adult Mammary Gland throughout the Estrous Cycle. (**A**) ERα protein expression in the adult mammary gland of WT and RSK2-KO mice during the estrous cycle. Scale bar: 20 μm. (**B**) ERα protein levels are lower in the RSK2-KO mice at all stages of the estrous cycle in adult mammary glands as determined by IF. ERα protein levels normalized to the average level observed in the WT mice at proestrus (median ± quartile, n ≥ 3 mice/genotype and stage, one-way ANOVA with Holm-Sidak's correction for multiple comparisons). (**C**) Loss of RSK2 results in a decrease in the number of ERα cells relative to $K8^+$ cells at all stages of the estrous cycle in adult mammary glands (median ± quartile, n ≥ 4 mice/genotype, ≥150 cells/mouse, one-way ANOVA with Holm-Sidak's correction for multiple comparisons). See **Figure 5-2** and **Table 5-1**.

Table 5-1. See Appendix, Supplemental Table 10.

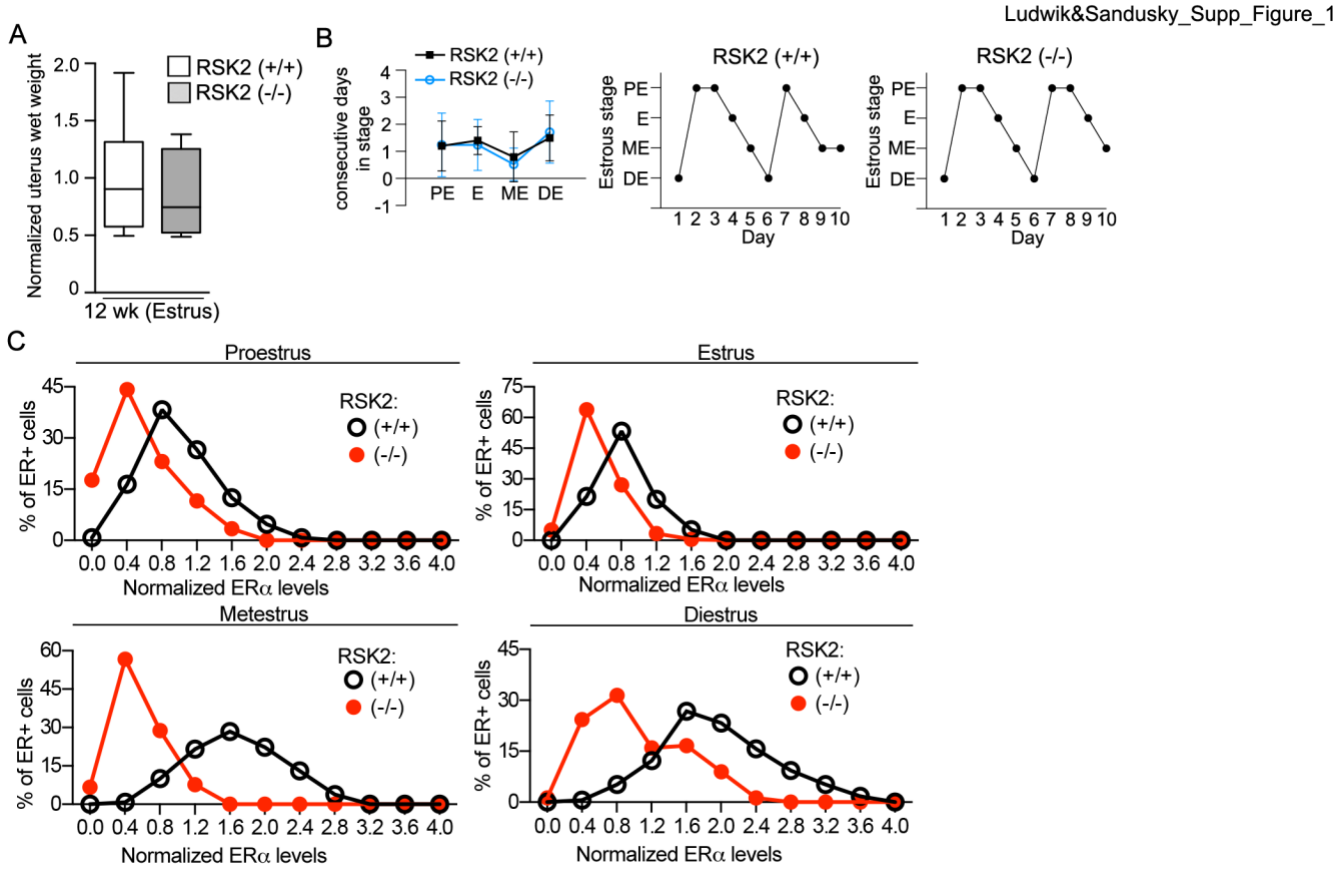


Figure 5-2. Estrogen responsiveness in WT and RSK2-KO mice. (**A**) Uterine wet weight is similar in WT and RSK2-KO. (median ± quartile, n≥8 mice/genotype, Student's t- test). (**B**) Cycling through the estrous cycle is similar in WT and RSK2-KO mice. Left graph: (mean ± S.D., n≥10 mice/genotype); Right graph: Representative cycle. (**C**) ERa protein expression levels are reduced in RSK2-KO at all stages of the estrous cycle in adult mammary glands. The graphs were generated from data shown in Figs. 1B and 1C.

To further investigate the decrease in ERα levels that occur in the RSK2-KO glands, we analyzed cell populations within the adult mammary glands by fluorescence-activated cell sorting (FACS). A novel FACS protocol that allowed the simultaneous analysis of WT and RSK2-KO mammary epithelial cells (MECs) was developed in which one of the genotypes was permanently marked, and equal numbers of cells from the marked and unmarked genotypes were mixed (**Fig. 5-3a**). The marked genotype was varied, and live cells and lineage-negative MECs were determined (**Fig. 5-4a**). The luminal and basal populations were clearly separated using epithelial cell adhesion molecule (EpCAM) and integrin alpha 6 (CD49f) (**Fig. 5-3b**). The distributions were fairly similar in adult WT and RSK2-KO mice at each stage of the estrous cycle (**Fig. 5-4b**). Further fractionation of the luminal cells by stem cells antigen-1 (Sca1) and integrin alpha 2 (CD49b) resulted in four populations with the gates for each experiment established using a fluorescence-minus-one strategy (Fig 5-4c).[253] The EpCAM^{hi}CD49f^{+}Sca1^{+}CD49b^{-} population, which consists primarily of ERα cells[250], is referred to as non-clonogenic luminal (NCL) because of its lack of colony-forming potential *in vitro* and engrafting ability *in vivo*. The EpCAM^{hi}CD49f^{+}Sca^{-}CD49b^{+} and EpCAM^{hi}CD49f^{+}Sca^{+}CD49b^{+} are luminal progenitors, which express low or high levels of luminal differentiation markers, respectively.[250] The EpCAM^{hi}CD49f^{+}Sca^{-}Cd49b^{-} population is currently undefined. In comparison to the WT population at estrus, the NCL population was decreased in the RSK2-KO mice, with a concomitant increase in the undefined population but no change in the luminal progenitor populations (**Fig. 5-3b** and **5-4d**). These observations are consistent with those observed in situ in which fewer ERα cells were observed in the RSK2-KO population (**Fig. 5-1c** and **5-2c**). A decrease in ERα protein levels was also

observed in NCL cells isolated during FACS, consistent with our in situ analysis (**Fig. 5-1a, b,** and **5-3d**). At each stage of the estrous cycle, a reduction in the NCL population was observed (**Fig. 5-3c** and **5-4d**).

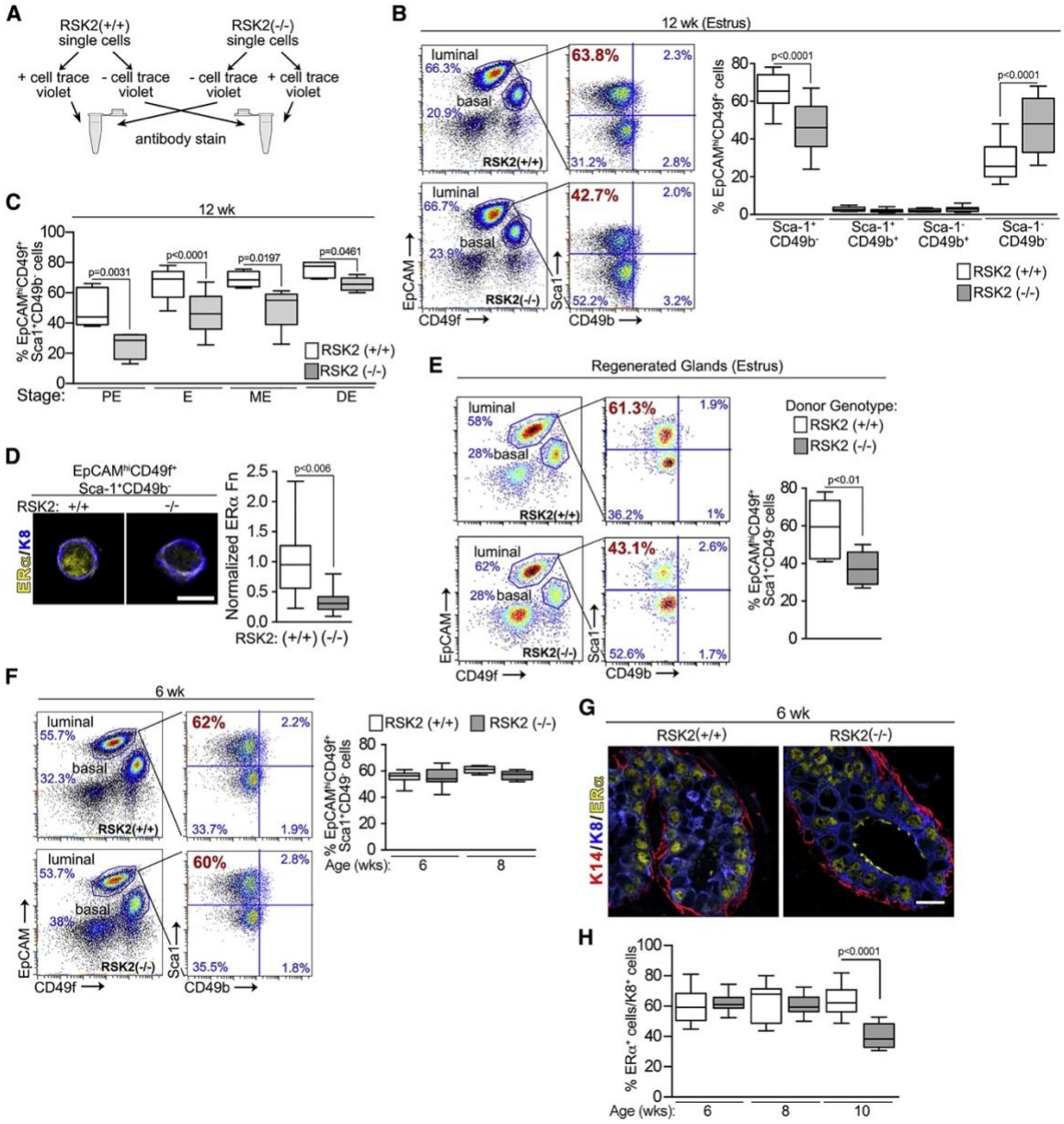


Figure 5-3. RSK2 Maintains the EpCAM^{hi}CD49f^{+}Sca1^{+}CD49b^{-} (NCL) Population within the Adult Mammary Gland throughout the Estrous Cycle. (**A**) Schematic of FACS protocol. (**B**) FACS analysis of adult mammary glands isolated from females during estrus. Gating strategy of luminal cells by further subdivision using Sca-1 and CD49b. The percentage of NCL cells within the luminal population at estrus decreases in adult RSK2-KO mice (median ± quartile, n ≥ 6 mice/genotype, one-way ANOVA with Holm-Sidak's correction for multiple comparisons). (**C**) Loss of RSK2 results in a reduction in the percentage of NCL cells at all stages of the estrous cycle in adult mammary glands (median ± quartile, n ≥ 3 mice/genotype and stage, one-way ANOVA with Holm-Sidak's correction for multiple comparisons). PE, proestrus; E, estrus; ME, metestrus; DE, diestrus. (**D**) ERα protein levels are decreased in cells isolated from the NCL population of RSK2-KO mice (median ± quartile, n = 3 mice/genotype, >20 cells/mouse, Student's t test). Scale bar: 10 μm. Fn, fluorescence. (**E**) RSK2 regulation of the NCL population is intrinsic to the epithelium (median ± quartile, n = 3 mice/genotype, Student's t test). (**F**) The percentage of NCL cells within the luminal population is similar between WT and RSK2-KO juvenile female mice (median ± quartile, n ≥ 3 mice/genotype and age group, one-way ANOVA with Holm-Sidak's correction for multiple comparisons). (**G** and **H**) The levels of ERα protein expression (**G**) and the number of ERα cells (**H**) relative to K8^{+} cells are similar in WT and RSK2-KO juvenile female mice (median ± quartile, n = 3 mice/genotype, ≥5 fields/mouse, one-way ANOVA with Holm-Sidak's correction for multiple comparisons). Scale bar: 20 μm. See **Figures 5-2** and **5-4** and **Table 5-1.**

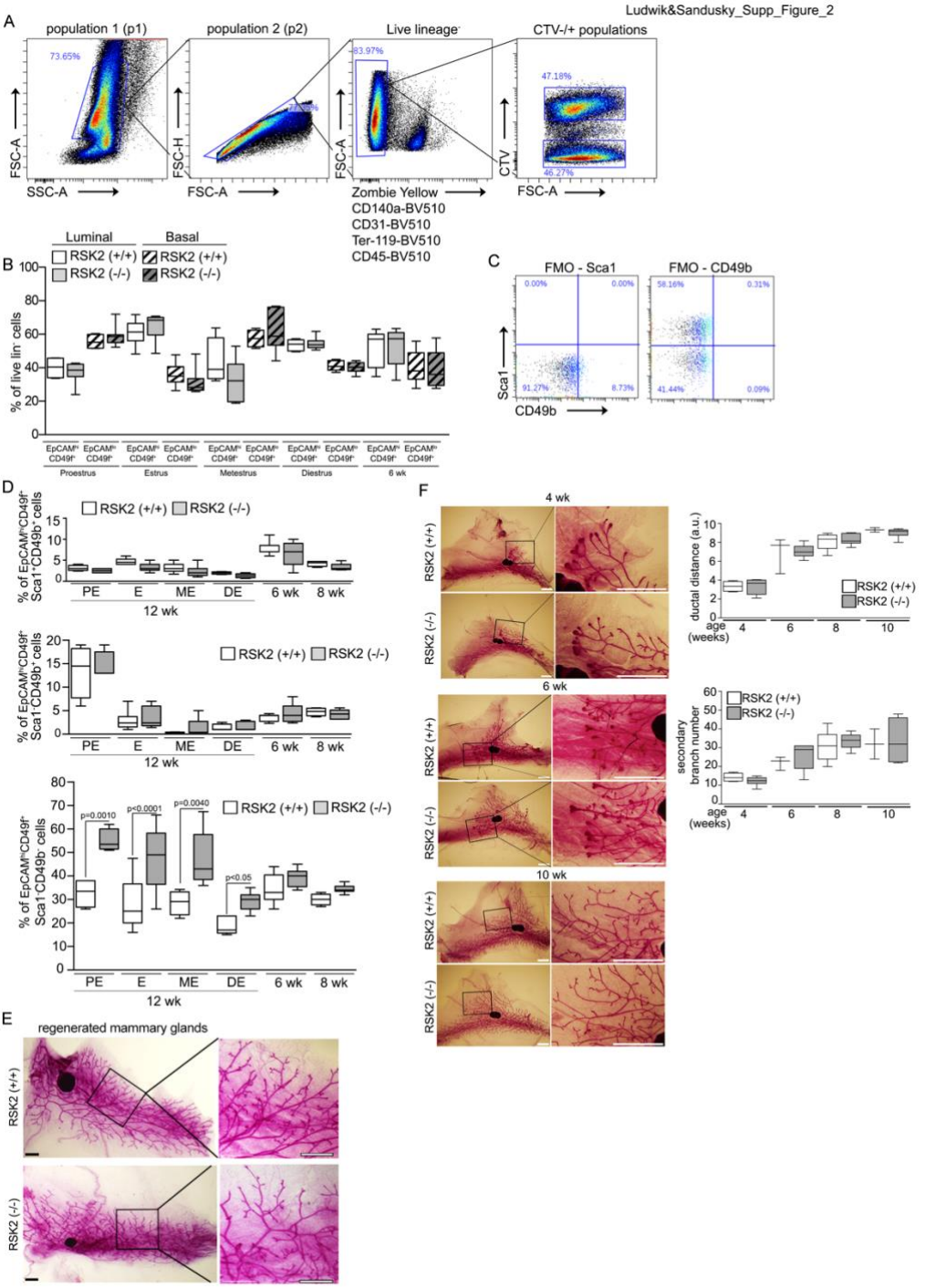


Figure 5-4. Analysis of WT and RSK2-KO mammary glands. (**A**) Gating strategy for flow cytometry analysis and sorting of mouse mammary epithelium. Cells were gated for forward (FCS-A) and side (SSC-A) scatter to remove debris. Single cells (p2) gated by FSC-H/A were then gated for live cells (ZombieYellow negative). Lineage+ (Cd140a+; CD31+; Ter-119+; and CD45+) cells were gated out. CellTraceViolet (CTV) positive and negative populations were separated. (**B**) FACS analysis of luminal and basal epithelial populations in the mammary gland (median ± quartile, n≥4 mice/genotype and stage, one-way ANOVA with Holm-Sidak's correction for multiple comparisons) (**Table 5-1**). solid=luminal, hatched = basal. (**C**) Fluorescence minus one strategy for determining the gates for Sca1 and CD49b. (**D**) FACS analysis of luminal progenitor and undefined epithelial populations (median ± quartile, n ≥ 3 mice/genotype and stage, one-way ANOVA with Holm-Sidak's correction for multiple comparisons) (**Table 5-1**). (**E**) Representative whole mount image of the regenerated 4th mammary gland from WT or RSK2-KO ~ 20 wk after transplantation at 3 wk. Scale bar = 1 mm. (**F**) Mammary gland development is similar in WT and RSK2-KO. (median ± quartile, n≥2 mice/genotype, one- way ANOVA with Holm-Sidak's correction for multiple comparisons) (**Table 5-1**). Scale bar = 2 mm.

RSK2-KO is a constitutive knockout, and therefore, we evaluated the contributions of systemic and intrinsic mechanisms that facilitate RSK2 regulation of the ERα population. To perform these analyses, mammary epithelial cells from WT and RSK2-KO mice were separately introduced into the cleared fourth mammary fat pads of a WT recipient. The glands from the transplanted cells regenerated to similar extents (**Fig. 5-4e**). In regenerated glands, loss of RSK2 also resulted in a decrease in the NCL population (**Fig. 5-3e**), indicating that the effects on the ERα population caused by the loss of RSK2 are intrinsic to the mammary epithelial cells.

Because ERα is absolutely required for mammary gland development[254], we analyzed the mammary gland at different ages starting at puberty. No detectable difference in the expansion of the mammary gland into the fat pad or branching during development was observed (**Fig. 5-4f**). Analysis by FACS showed that all cell populations were similar between RSK2-KO and WT mice in juveniles (**Fig. 5-3f**, **5-4b**, and **5-4d**). Consistent with these data, in situ analysis of the juvenile mammary glands showed similar ERα protein levels (**Fig. 5-3g**) and numbers of ERα cells (**Fig. 5-3h**). We conclude that RSK2 regulates the ERα population only in the adult, which explains the absence of a developmental defect.

ERK1/2-RSK2 Signaling Is Dependent on Estrogens

At the onset of puberty estrogen increases the levels of growth factors[255,256], which, theoretically, would result in RSK activation through its upstream activator, ERK1/2.[257] C57BL/6J mice initiate cycling by ~6 weeks[258], although we observed that cycling was irregular until ~10–12 weeks old. Interestingly, ERK1/2, as shown by Thr202/Tyr204 phosphorylation (pERK1/2), was not active until the animals were ≥10 weeks old, and the levels of active ERK1/2 were similar at estrus between the WT and RSK2-KO mice at the same age (**Fig. 5-5a**). A causal relationship between estrogen and ERK1/2 activity was demonstrated by the observations that ERK1/2 activation was prevented by oophorectomy at 6 weeks (**Fig. 5-5b**). ERK1/2 activation in the WT mouse occurs in estrus after the estrogen burst in proestrus and then decreases during diestrus when estrogen levels are lowest (**Fig. 5-5c**). The inactivation of ERK1/2 appears to be consistent with increased phosphatase activity because the protein levels of ERK1/2 do not change between estrus and diestrus (**Fig. 5-6b**). We conclude that the ability of estrogen to activate ERK1/2 and regulate its cyclic activation appears as the mice sexually mature.

Active ERK1/2 was primarily confined to the luminal compartment and was present in ERα cells (**Fig. 5-6c**). To confirm that RSK was activated in the WT mammary gland, an anti-active RSK antibody (pRSK) was used. RSK is activated in response to coordinated inter- and intra-molecular phosphorylation events[259], which are identical within the RSK family, and therefore, identification of the active state of a particular RSK is not possible. However, active RSK was not detectable in the adult RSK2-KO mammary glands, indicating that RSK2 is the predominant active RSK isoform (**Fig. 5-5d**). These results demonstrate that, in the WT mouse, estrogen activates ERK1/2-RSK2 signaling, and that this activation corresponds with the ability of RSK2 to regulate ERα protein levels (**Fig. 5-1b** and **5-5c**).

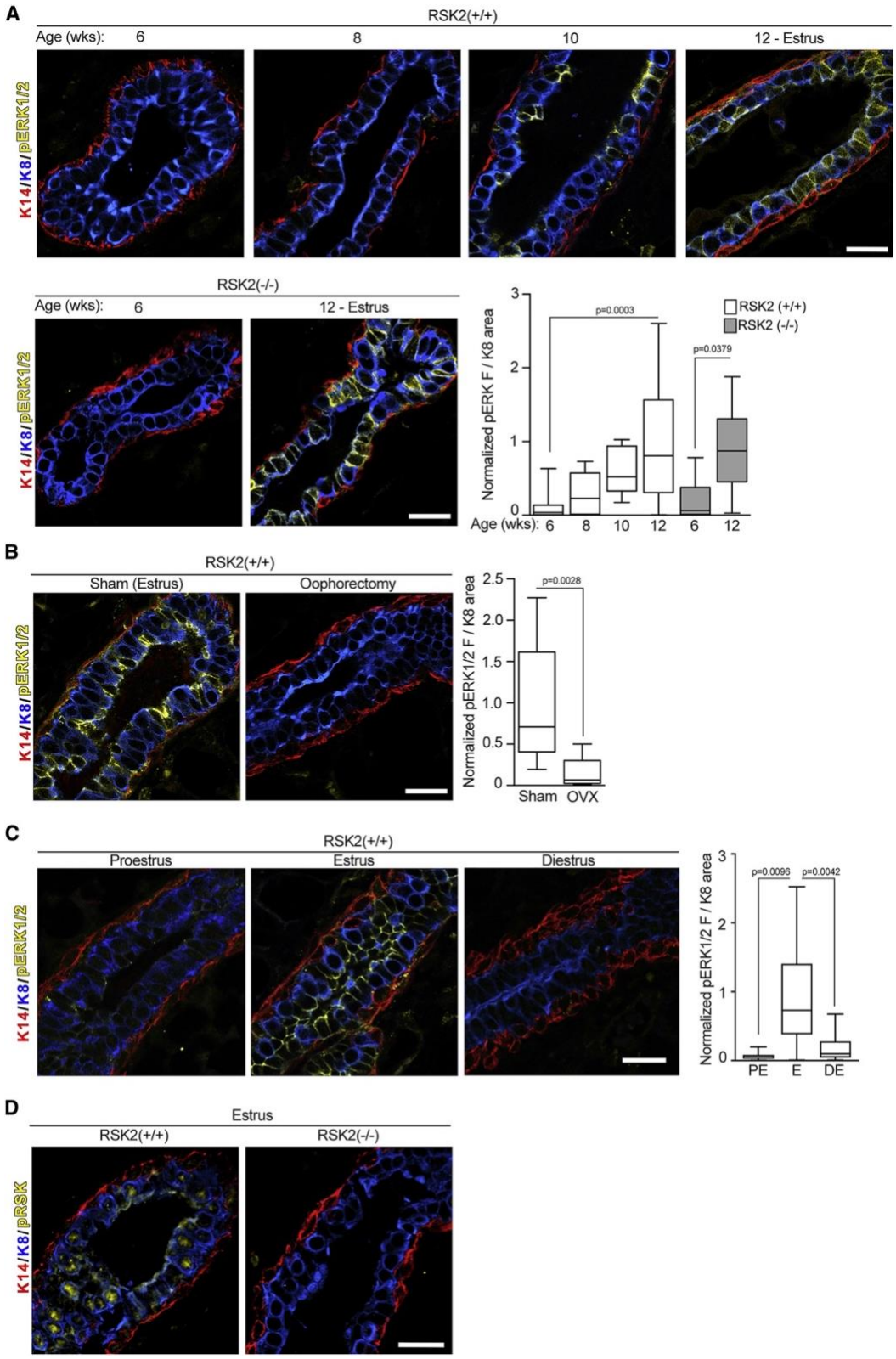


Figure 5-5. ERK1/2-RSK2 Signaling Is Activated Only in the Adult Mammary Gland. (**A**) ERK1/2 activity is increased in the adult compared with juvenile animals (median ± quartile, n ≥ 2 mice/genotype and age, ≥3 fields/mouse, one-way ANOVA with Holm-Sidak's correction for multiple comparisons). Scale bar: 20 μm. (**B**) ERK1/2 activity in the mammary gland depends on estrogen (median ± quartile, n ≥ 2 mice/genotype and procedure, ≥3 fields/mouse, Student's t test). Scale bar: 20 μm. (**C**) ERK1/2 activity varies during the estrous cycle in the WT mice adult mammary gland (median ± quartile, n ≥ 2 mice/genotype, ≥3 fields/mouse, one-way ANOVA with Tukey's correction for multiple comparisons). Scale bar: 20 μm. (**D**) Active nuclear RSK2 is the predominant RSK in adult mammary glands. Scale bar: 20 μm. See **Figure 5-6** and **Table 5-1**.

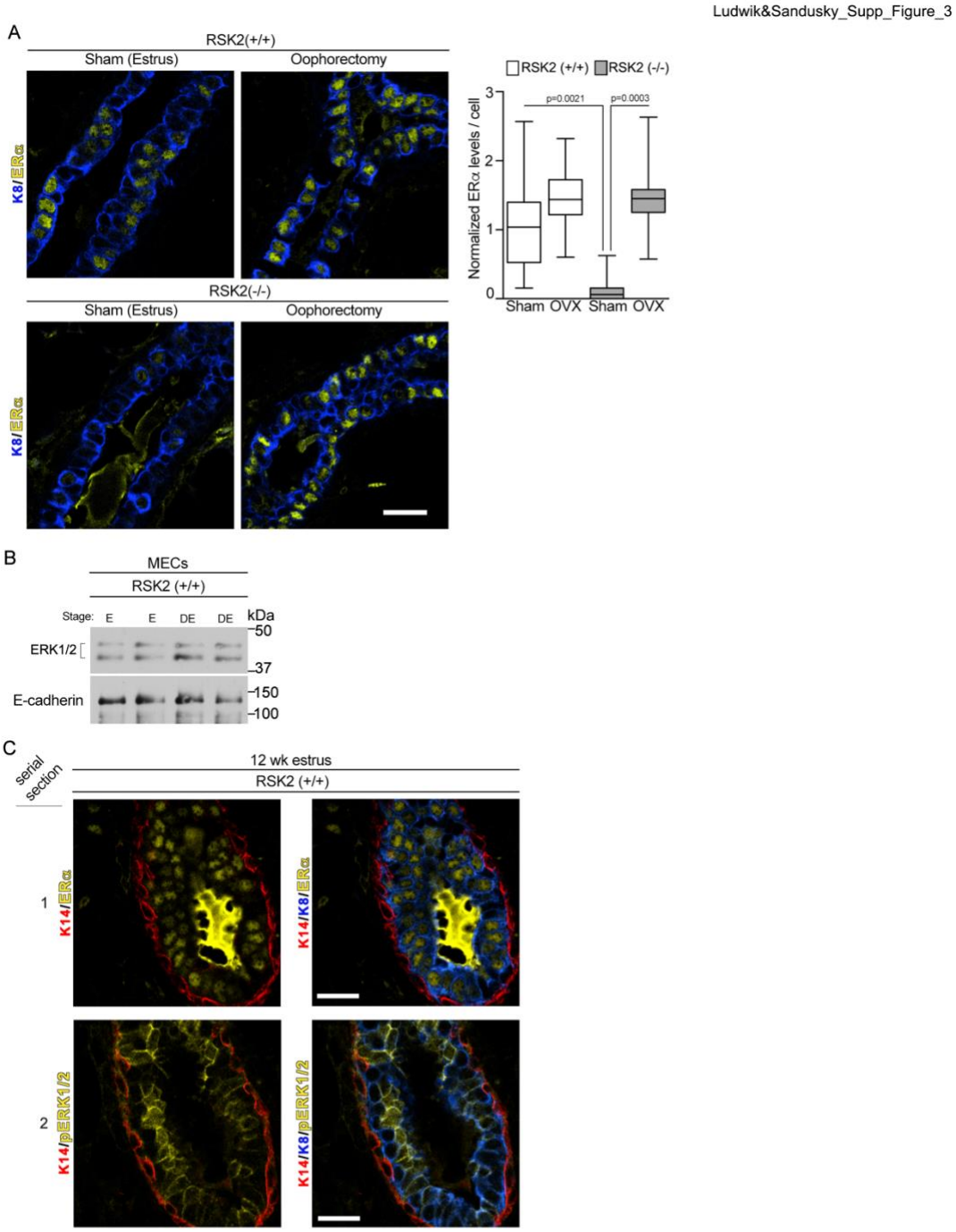


Figure 5-6. ERK1/2 is active in ER+ cells. (**A**) ERa protein levels increase in response to oophorectomy (median ± quartile, n ≥2 mice/genotype and procedure, ≥ 3 fields/mouse, one-way ANOVA with Holm-Sidak's correction for multiple comparisons) (**Table 5-1**). Scale bar= 20 µm. (**B**) ERK1/2 protein levels are similar in estrus and diestrus in mammary epithelial cells isolated from WT adult mammary glands. (**C**) The image on the left is shown without K8 to facilitate the visualization of ERa and pERK1/2. Serial sections were necessary to avoid antibody interference. Scale bar = 20 µm.

RSK2 Negatively Regulates Proteasome-Coupled Transcription in the Adult Mammary Gland

To identify a mechanism that would explain the reduced ERα protein levels with the loss of RSK2, we performed transcriptomic analyses on the NCL population. Estrus was chosen because changes in gene expression would be occurring in

response to the estrogen pulse that happened in proestrus. We contrasted these data with those obtained in diestrus, which has the lowest estrogen levels. The transcriptomic analysis of the RSK2-KO mice showed 2,747 differentially expressed genes (DEGs) between estrus and diestrus as compared with 39 in the WT mice between estrus and diestrus (**Fig. 5-7a** and **5-8a**). The transcriptomic data of RSK2-KO mice at estrus showed a significant correlation with a signature obtained from the ERα breast cancer cell line MCF-7 at 24 h after estrogen treatment (**Fig. 5-7b and c**; **Table 5-2**).[260] This correlation was not driven by cell cycle genes (**Fig. 5-8b; Table 5-2**). No significant correlation with the estrogen-responsive gene signature was obtained for the WT mice at estrus (**Fig. 5-7b and c**). ESR1 (gene encoding ERα) mRNA levels were similar between WT and RSK2-KO (**5-8c**), eliminating the possibility that ERα mRNA expression levels accounted for the transcriptomic differences. Taken together, these data demonstrate that estrogen signaling is higher in the RSK2-KO than in the WT mice, and therefore, we conclude that RSK2 acts to inhibit estrogen-responsive gene expression.

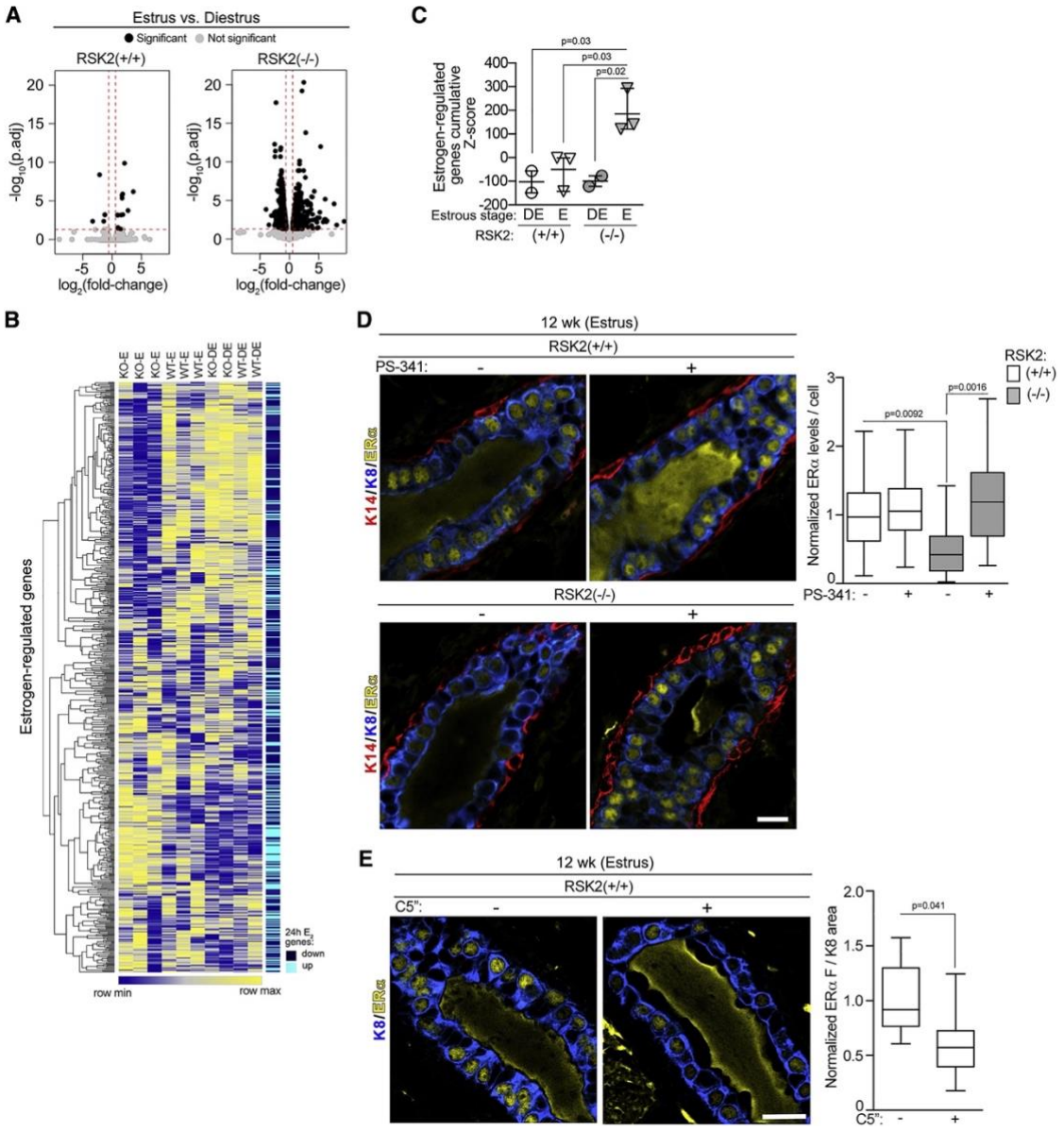


Figure 5-7. RSK2 Is a Negative Regulator of ERα-Mediated Signaling. (**A**) RSK2-KO mice show greater numbers of DEGs between estrus and diestrus (right panel) than do WT mice (left panel). Genes with a fold-change ≥ |1.5| (log_2[fold-change] ≥ |0.5|) and a false discovery rate (FDR)-adjusted $p < 0.05$ are shown as black dots, and genes with a fold-change < |1.5| (log_2[fold-change] < |0.5|) and an FDR-adjusted p value > 0.05 are shown as gray dots. The dashed line indicates the cutoff values. (**B**) Heatmap illustrating that the gene expression of NCL cells isolated from RSK2-KO mice in estrus correlates with a 24-h estrogen-regulated gene signature identified from MCF-7 cells.[260] (**C**) Quantitative assessment of enrichment for

estrogen-regulated genes. Cumulative *Z* scores were generated for each mouse by summing individual *Z* scores of genes upregulated in estrogen-regulated signature and subtracting individual *Z* scores of genes downregulated (mean ± SD, each point represents a mouse; one-way ANOVA with Holm-Sidak's correction for multiple comparisons). (**D**) Loss of RSK2 increases ERα turnover. Adult mice staged at estrus were treated with vehicle or PS-341 (5 mg/kg) intraperitoneally (i.p.) for 4 h before euthanasia and isolation of the mammary gland. ERα protein levels were normalized to those observed in the WT mice at estrus (median ± quartile, n = 3 mice/genotype and condition, ≥200 cells/mouse, one-way ANOVA with Holm-Sidak's correction for multiple comparisons). Scale bar: 20 µm. (**E**) RSK2 kinase activity is necessary to maintain ERα protein levels. Adult mice staged at estrus were treated with vehicle or C5″-*n*-propyl cyclitol SL0101 (C5″) (40 mg/kg) IP twice every 7 h before euthanasia and isolation of the mammary gland (median ± quartile, n ≥ 3 mice/genotype, ≥3 fields/mouse, Student's t test). See **Figures 5-8** and **5-11** and **Table 5-1**, **5-4**, and **5-6**.

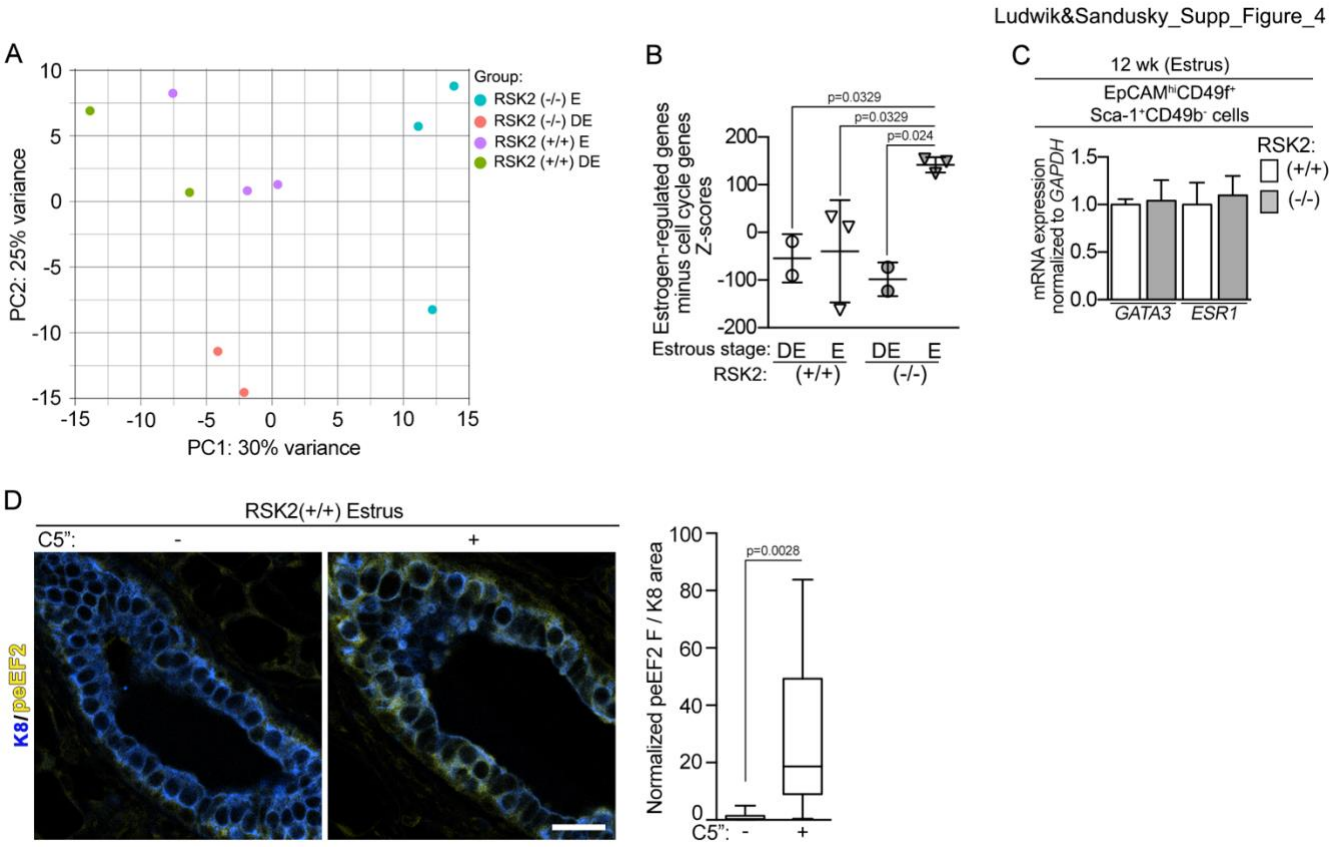


Figure 5-8. Transcriptomic analysis of the NCL population. (**A**) Principal component (PC) analysis of the transcriptomic data. (**B**) Proliferation genes do not drive the enrichment for estrogen -regulated signature in RSK2 KO estrus mice. Cumulative Z-scores were generated for each mouse by summing individual Z-scores of genes up regulated in estrogen-regulated signature in which the cell cycle genes were removed and subtracting individual Z-scores of genes down regulated. (median ± quartile, one-way ANOVA with Holm-Sidak's correction for multiple comparisons) (**Table 5-1**). (**C**) ESR1 and GATA3 mRNA levels are similar in NCL cells isolated from RSK2-KO and WT mice during the estrus stage (mean ± S.D., n=3 mice/genotype in triplicate, Student's t-test). (**D**) On target increase in peEF2 *in vivo* by *C5″-n* Adult mice staged at estrus were treated with vehicle or *C5″* gland (median ± quartile, n ≥ 2 mice/genotype in triplicate, Student's t-test).

Gene sets in overlap		Fisher's exact test for overlap
Genes UP in R2KO-E vs R2KO-DE	E2_24h_UP	0.00001
Genes DOWN in R2KO-E vs R2KO-DE	E2_24h_DOWN	0.00001
Genes UP in R2KO-E vs R2KO-DE	Cell cycle genes	0.0007
Genes UP in WT-E vs WT-DE	E2_24h_UP	No overlap
Genes DOWN in WT-E vs WT-DE	E2_24h_DOWN	No overlap
Genes UP in WT-E vs WT_DE	Cell cycle genes	0.5785

Table 5-2. Statistical analysis of gene set overlaps from the NCL populations.

Estrogen-responsive gene expression is interconnected with ERα destruction through the 26S proteasome pathway.[232–236] Therefore, it would be expected that ERα degradation would be greater in the RSK2-KO, than the WT, mice. To investigate this

possibility, the rate of *in vivo* ERα degradation was determined using the 26S proteasome inhibitor PS-341.[235] In the WT gland, ERα levels did not substantially change in response to proteasome inhibition. However, in the RSK2-KO gland, ERα protein levels increased by ~5-fold in response to PS-341; therefore, ERα degradation is much higher in the absence of RSK2 (**Fig. 5-7d**). This increased degradation explains our in situ observations that ERα levels are lower in the RSK2-KO, compared with those in the WT, mice (**Fig. 5-1a, b**, and **5-2c**).

Reduced ERα protein levels in the RSK2-KO could be the result of decreased RSK2 kinase activity or the loss of the RSK2 protein. To distinguish between these mechanisms, RSK2 activity was inhibited *in vivo* by the specific RSK1/2 inhibitor C5″-*n*-propyl cyclitol SL0101 (C5″).[261,262(p2)] ERα protein levels were reduced by the RSK1/2 inhibitor (**Fig. 5-7e**). To demonstrate that the inhibitor was on target, we used phosphorylation of the elongation translation factor 2 (peEF2) as a biomarker (**Fig. 5-8d**).[263] We conclude that RSK2 kinase activity is important in ERα degradation.

RSK2 Maintains ERα Protein Levels in Adult Reproductive Tissue

We next investigated whether RSK2 preserved ERα protein levels in other estrogen-responsive tissues. We focused on the female reproductive tract because we observed a 40% reduction in the fertility rate in crosses between RSK2-KO female and male mice (**Fig. 5-9a**). RSK2-KO male mice crossed with heterozygote female mice had similar fertility rates to those of the WT mice crosses, indicating that the reduced fertility is associated with the RSK2-KO female mice. Ovaries in the RSK2-KO and WT mice showed all stages of follicular development and the presence of the corpora luteum (**Fig. 5-9b** and **5-10a**), demonstrating that hormonal signaling[264] through the hypothalamic-pituitary-ovarian axis is not impaired in RSK2-KO mice. The uterus expresses high levels of ERα, which is present in stromal cells as well as in glandular and luminal epithelium. In comparison to the WT mice, the ERα protein levels were substantially decreased in the epithelial, but not in the stromal cells, in RSK2-KO mice (**Fig. 5-9c**). Interestingly, ERK1/2 activity was detected in the uterine epithelium but not in the stroma cells, providing further evidence of the connection between ERK1/2-RSK2 signaling and the regulation of ERα protein levels (**Fig. 5-9d**). Uterine wet weight and total uterine width were similar in the WT and RSK2-KO mice, which is consistent with the literature because stromal cells are thought to mediate uterine expansion (**Fig. 5-2a** and **5-10b**).[247,265] These data indicate that RSK2 regulates ERα protein levels in multiple tissues.

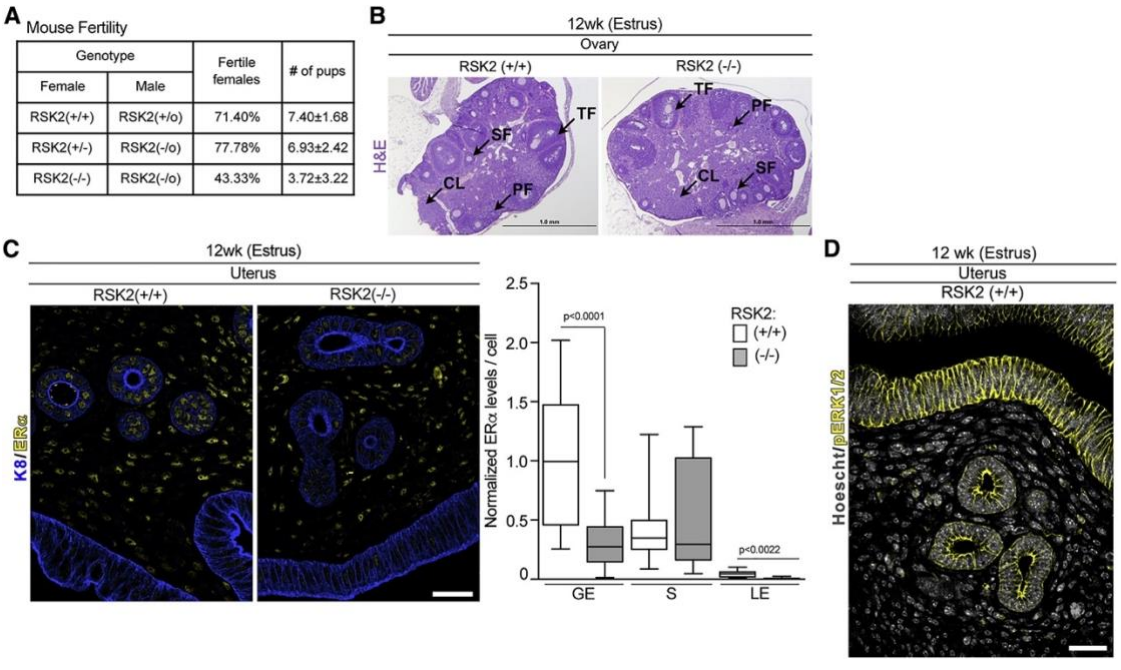


Figure 5-9. RSK2 Maintains ERα Protein Levels in the Uterine Epithelium. (**A**) RSK2-KO mice show a fertility defect (n ≥ 15 dams/genotype, Chi-square test p = 0.0299). (**B**) The hypothalamic-pituitary-ovarian axis is not disrupted in RSK2-KO female mice. H&E sections of ovaries isolated from adult mice in estrus. Scale bar: 1 mm, PF, primary follicle; SF, secondary follicle; TF, tertiary follicle; CL, corpus luteum. (**C**) RSK2-KO mice have reduced ERα protein levels in the glandular and luminal epithelium of the uterus (median ± quartile, n = 3 mice/genotype, >120 cells/mouse, Student's t test). Scale bar: 40 µm. GE, glandular epithelium; S, stroma; LE, luminal epithelium. (**D**) Active ERK1/2 is confined to the epithelium of the uterus. Scale bar: 40 µm. See **Figures 5-2** and **5-10**.

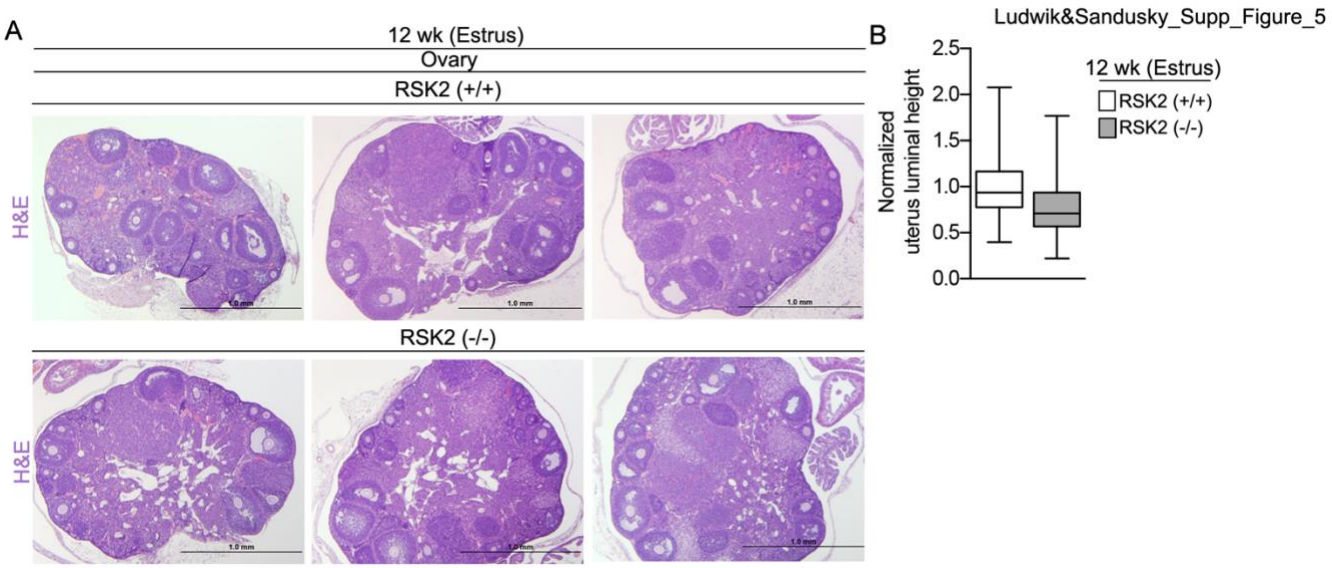


Figure 5-10. The hypothalamic-pituitary-ovarian axis is not impaired in RSK2-KO mice. (**A**) Representative H&E images of ovaries. Scale bar = 1 mm. (**B**) Luminal height in the uterus in the WT and RSK2-KO are similar. Measurements from ≥30 randomly selected regions from each animal (median ± quartile, n≥3 mice/genotype, ≥ 3 fields/mouse, Student's t-test).

ERK1/2 Drives ERα Degradation through Phosphorylation of Ser-118 on ERα

To address the mechanism by which RSK2 regulates ERα protein levels, we initially focused on GATA3 because GATA3 and ERα regulate each other's expression via a positive-feedback mechanism in breast cancer.[266] Therefore, it is conceivable that RSK2 indirectly regulates ERα protein levels through GATA3. However, no difference in GATA3 mRNA levels was observed between WT and RSK2-KO mice (**Fig. 5-8c**). Furthermore, GATA3 protein levels in the uterine glandular epithelium[267] are extremely low, whereas ERα protein levels are very high (**Fig. 5-11a**). We conclude that RSK2 regulation of ERα through GATA3 is unlikely.

Interestingly, in contrast to that of the WT mice, ERK1/2 activity remains elevated in the RSK2-KO mice during diestrus (**Fig. 5-12a**), and coincident with these observations, ERα protein levels remain lower in the RSK2-KO (**Fig. 5-1a, b**, and **5-2c**). Therefore, we investigated whether ERK1/2 activity was a driver of ERα degradation. In support of this hypothesis, when we prevented ERK1/2 activation by oophorectomizing RSK2-KO mice, we observed that the levels of ERα were rescued to WT levels (**Fig. 5-6a**). To perform further mechanistic studies, we used the normal mouse Leydig cell line TM3, which expresses ERα but does not form tumors *in vivo*. Survival of the TM3 line was dependent on RSK2, which prevented knockout approaches. However, short-term treatment with two structurally distinct RSK inhibitors decreased ERα protein levels, which was prevented by the inhibition of the 26S proteasome (**Fig. 5-12b** and **c**). This effect is specific because androgen receptor protein levels do not change in response to RSK1/2 inhibition (**Fig. 5-12d**). ERK1/2 activity increased in response to the RSK inhibitors (**Fig. 5-12b**), which is consistent with our observations at diestrus in the RSK2-KO mice. MEK inhibition by trametinib or U0126 did not decrease ERα levels. Taken together, these results indicate that ERK1/2 activity increases ERα degradation through the 26S proteasome.

It is hypothesized that degradation of phosphorylated ERα occurs at a faster rate than that of the unphosphorylated.[268] Therefore, we investigated whether the ERK1/2 and RSK2 phosphorylation of ERα[244,245] regulated ERα turnover. GFP-tagged ERα mutants were generated, in which the ERK1/2 phosphorylation site, Ser-118 (S118A-ERα), or the RSK2 site, Ser-167 (S167A- ERα), was mutated to Ala. In response to ERK1/2 activation, mutation of Ser-167 did not alter ERα turnover; however, mutation of Ser-118 prevented ERα destruction (**Fig. 5-12e**). An electrophoretic mobility-shift assay was used to confirm phosphorylation of Ser-118 in response to RSK1/2 inhibition because phospho-specific antibodies to human ERα do not recognize the mouse protein (**Fig. 5-11b**).[269] We conclude that ERK1/2 phosphorylation of Ser-118 targets ERα for destruction and that RSK2 negatively regulates ERK1/2 activity to protect ERα from degradation.

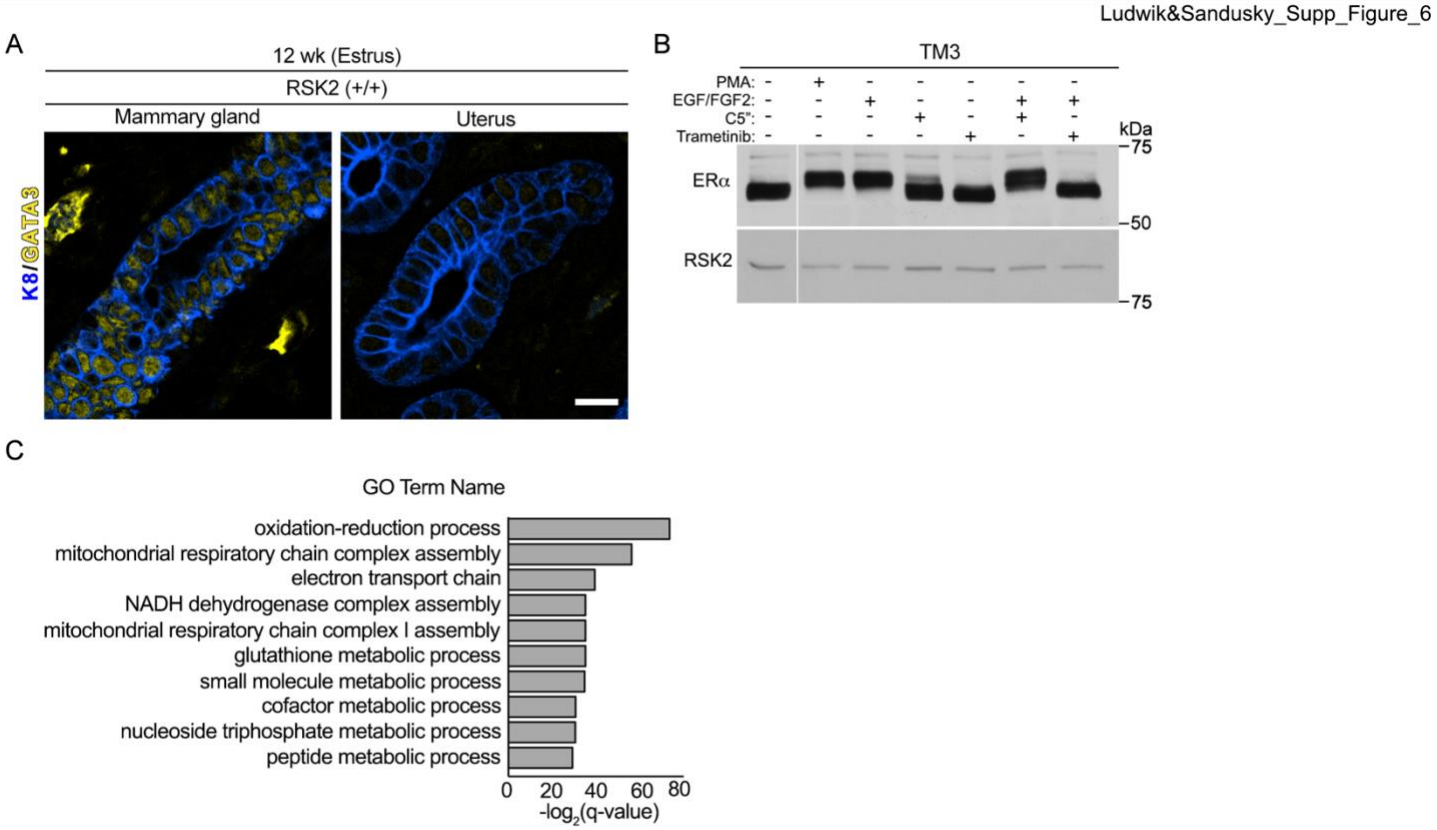


Figure 5-11. Phosphorylation of Ser-118 ERα correlates with degradation of ERα. (**A**) GATA3 is expressed at very low levels in the uterus compared to the mammary gland. Scale bar = 20 μm. (**B**) Ser118- ERa phosphorylation occurs in response to agents that stimulate ERa degradation. Serum starved TM3 were treated with PMA (0.5 μM, 20 min) or an EGF/FGF7 cocktail (12.5 nM each, 5 min) with or without C5" (20 μm, 2h) or trametinib (1 μm, 1 h as a pretreatment). The white vertical line indicates that conditions not relevant to the manuscript were removed. (**C**) GO enrichment analysis for NCL population in RSK2-KO glands at estrus.

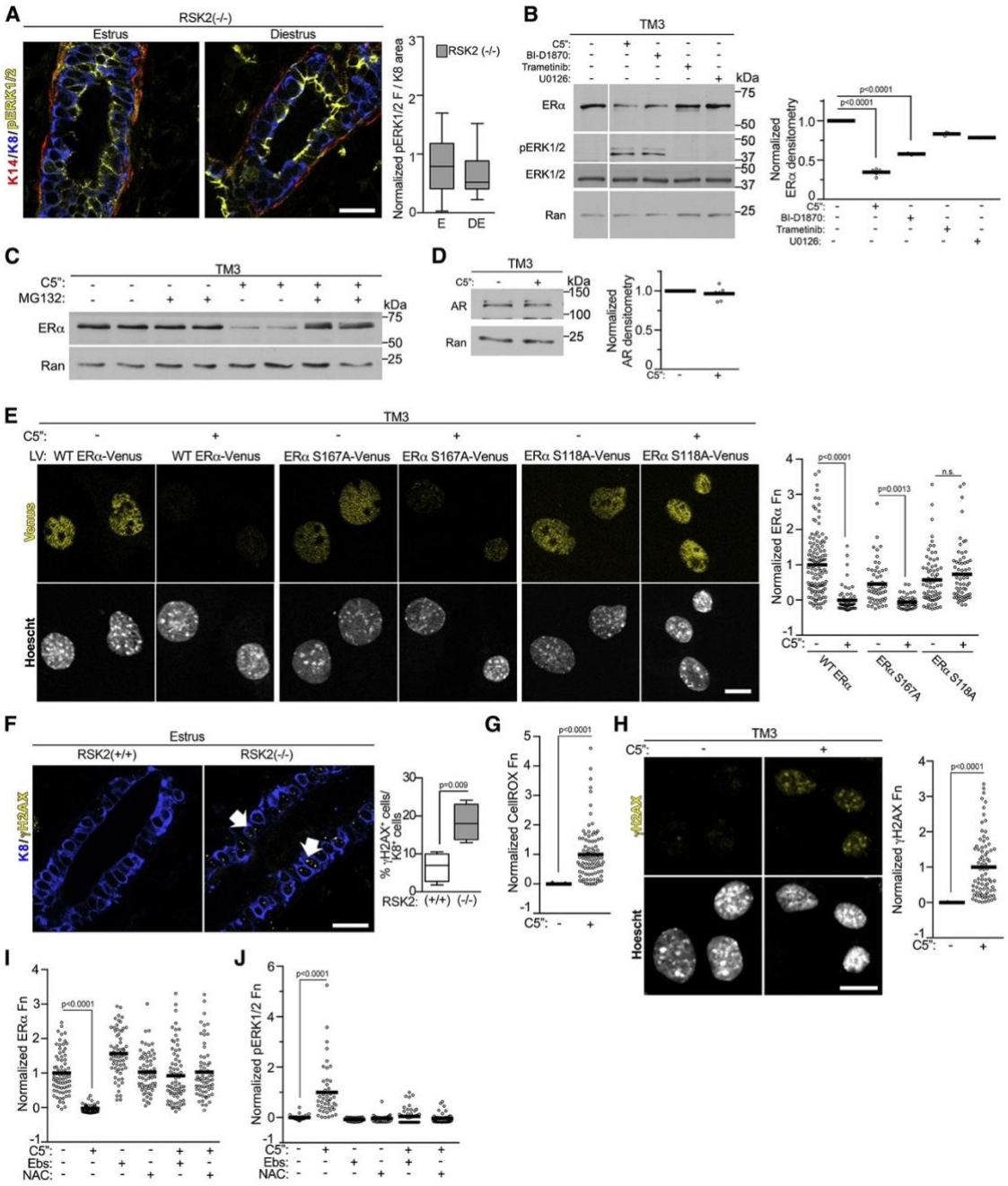


Figure 5-12. ERK1/2 Drives ERα Degradation through Phosphorylation of Ser-118. (**A**) ERK1/2 activity remains elevated during diestrus in the adult mammary gland (median ± quartile, n = 3 mice, ≥3 fields/mice, Student's t test). (**B**) RSK2 is a negative regulator of ERK1/2 activity. Serum-starved TM3 was treated for 6 h with vehicle, C5″-*n*-propyl cyclitol SL0101 (C5″) (20 μM), BI-D1870 (10 μM), trametinib (1 μM), or U0126 (10 μM). The white vertical line indicates that conditions not relevant to the manuscript were removed. ERα levels were normalized to Ran and then to the vehicle (mean, n = 3, one-way ANOVA with Dunnett's correction for multiple comparisons). (**C**) RSK1/2 inhibition stimulates ERα degradation through the 26S proteasome pathway. Serum-starved TM3 was treated for 6 h with vehicle, C5″ (20 μM) with or without a 1 h of pre-treatment with MG132 (10 μM). (**D**) RSK2 does not regulate androgen receptor (AR) degradation. Serum-starved TM3 was treated for 6 h with vehicle or C5″ (20 μM). AR levels were normalized to Ran and then to the vehicle (mean, n = 3 in duplicate, Student's t test). (**E**) Phosphorylation of Ser-118A is required for ERα degradation. Cells transduced with WT or mutant ERα-VENUS were treated with vehicle or C5″ (20 μM) as in (B). The range was normalized to WT ERα (mean, n = 3, >150 cells/condition/experiment, one-way ANOVA with Holm-Sidak's correction for multiple comparisons). Scale bar: 10 μm. (**F**) Loss of RSK2 increases double-stranded DNA breaks (γ-H2AX foci) in the mammary gland (median ± quartile, n ≥ 4 mice/genotype, ≥3 fields/mouse, Student's t test). (**G**) RSK1/2 inhibition increases ROS. Serum-starved TM3 was treated as in (**B**). The data were normalized to the range with and without C5″ (20 μM) (mean, n = 3, >100 cells/condition/experiment, Student's t test). (**H**) RSK1/2 inhibition increases DNA damage *in vitro*. Cells treated for 72 h with vehicle or C5″ (20 μM). The data were normalized to the range with and without C5″ (mean, n = 3, >80 cells/condition/experiment, Student's t test). (**I**) Inhibition of ROS rescues ERα levels. Serum-starved TM3 was treated for 6 h with vehicle or C5″ (20 μM) with or without

ebselen (Ebs) (50 μM) or *N*-acetyl cysteine (NAC) (15 mM) for the final 2 h. The range was normalized to ERα levels in the absence of anti-oxidants (mean, n = 3, >50 cells/condition/experiment, one-way ANOVA with Dunnett's correction for multiple comparisons). (**J**) Inhibition of ROS inhibits ERK1/2 activation. Cells treated and analyzed as in (**I**). See **Figures 5-6** and **5-11** and **Table 5-1**.

RSK2 Negatively Regulates ERK1/2 Activity by Controlling Oxidative Stress Levels

To investigate the mechanism by which RSK2 negatively regulates ERK1/2, we determined whether a loss of RSK2 resulted in increased oxidative stress. In support of this hypothesis, an increase in ROS is associated with estrogen-regulated transcription[270], and ROS activates ERK1/2.[271(p3)] Therefore, in the RSK2-KO mouse, the increased estrogen-regulated transcription could result in elevated ROS levels compared with that of the WT mouse, resulting in ERK1/2 activation. The presence of γ-H2AX provides a readout for the formation of DNA double-stranded breaks, which occur in response to oxidative stress.[272] Consistent with our hypothesis, γ-H2AX was elevated in the RSK2-KO (**Fig. 5-12f**). Analysis of the genes upregulated in the RSK2-KO at estrus compared with diestrus revealed enrichment for genes associated with oxidative stress (**Fig. 5-11c**). Additionally, an over-representation of genes was associated with the glutathione metabolic process, suggesting that the cells were experiencing oxidative stress and attempting to compensate by increasing glutathione production. Consistent with the *in vivo* data, RSK2 inhibition in the TM3 line exhibited elevated ROS (**Fig. 5-12g**) and DNA damage (**Fig. 5-12h**). Importantly, reduction of ROS by two structurally distinct anti-oxidants rescued ERα levels in the presence of RSK2 inhibition (**Fig. 5-12i**) and prevented ERK1/2 activation (**Fig. 5-12j**). Taken together, these data demonstrate that RSK2 maintains estrogen homeostasis by preventing the activation of ERK1/2 by ROS.

RSK2 Integrates Estrogen-Mediated Transcription and Translational Responses to Maintain Homeostasis

There was no evidence of hyperplasia in the RSK2-KO glands, which was surprising because of their increased expression of cell cycle genes. In fact, the rate of proliferation was decreased in the NCL population of RSK2-KO mice (**Fig. 5-13a**, **5-14a**, and **5-14b**), which is consistent with the reduced number of ERα cells observed in these mice (**Fig. 5-1c, 5-3b**, and **5-3c**). Because of this disconnect between the gene expression and proliferation data, we investigated whether RSK2 was important in translational regulation in the ERα population.[273] As a readout for translational activity *in vivo*, we measured eEF2 phosphorylation (peEF2). The levels of peEF2 were higher at diestrus in the RSK2-KO mice (**Fig. 5-13b**), which is consistent with inhibition of protein synthesis.[263] We also observed that RSK1/2 inhibition decreased protein synthesis in the TM3 line (**Fig. 5-13c**). Taken together, these data support a model in which RSK2 regulation of translation contributes to the physiological responses induced by estrogen.

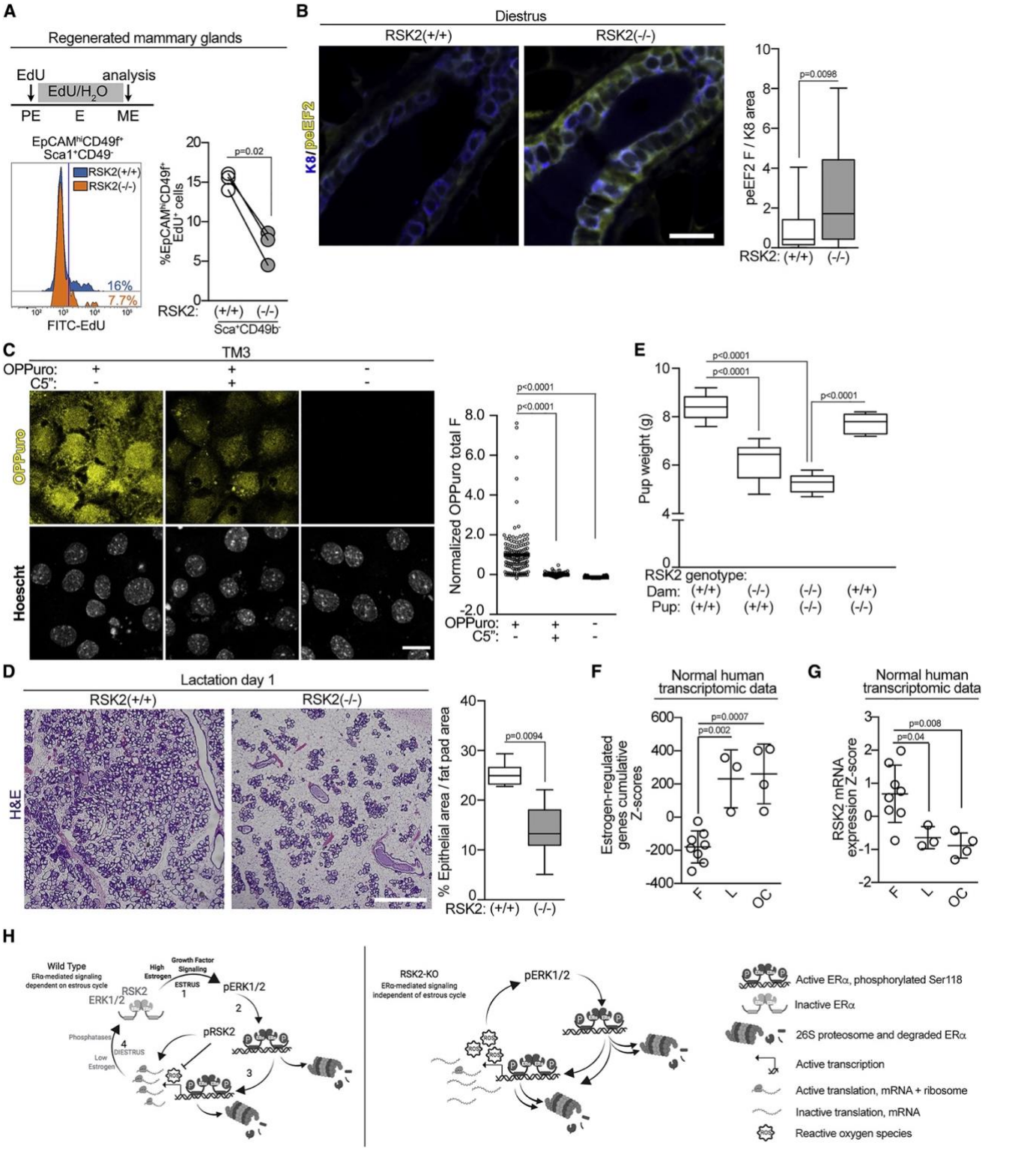


Figure 5-13. RSK2 Is Necessary for Alveolar Expansion. (**A**) RSK2-KO NCL cells show a decrease in proliferation as compared with the WT cells. RSK2-KO or WT MECs were used to regenerate the mammary gland in a WT mouse. These mice were staged in proestrus and administered 5-ethynyl-2′-deoxyuridine (EdU) throughout one estrus cycle. The mammary glands were isolated and analyzed by FACS (n = 3 glands/genotype; paired Student's t test). (**B**) RSK2 regulates eEF2K activity *in vivo* (median ± quartile, n ≥ 3 mice/genotype, ≥5 fields/mouse, Student's t test). (**C**) Inhibition of RSK1/2 decreases translation *in vitro*. Serum-starved TM3 was treated for 6 h with vehicle or C5″ (20 μM). The range was normalized to the *o*-propargyl-puromycin (OPP) in the absence and presence of C5″ (mean, n ≥ 3, >150 cells/condition/experiment, one-way ANOVA with Holm-Sidak's correction for multiple comparisons). (**D**) Alveolar expansion is reduced in RSK2-KO dams as shown by H&E stains of mammary glands isolated from dams 1 d after birth (median ± quartile, n ≥ 3 mice/genotype, ≥3 fields/mouse, Student's t test). Scale bar: 1 mm. (**E**) Pups nursed by RSK2-KO dams are smaller than those nursed by WT dams. Weanling weight at 21 d nursed by a dam with the indicated genotype (median ± quartile, n = 3 litters matched for size/dam

genotype, one-way ANOVA with Holm-Sidak's correction for multiple comparisons). (**F**) The estrogen-regulated signature is enriched in the luteal phase or with oral contraceptive use. Cumulative patient *Z* scores were generated for each individual by summing individual *Z* scores of genes upregulated in estrogen-regulated signature and subtracting individual *Z* scores of genes downregulated (mean ± SD, n = 8; F, follicular, 3 L, luteal; 4 OC, oral contraceptive; one-way ANOVA with Tukey's correction for multiple comparisons). (**G**) RSK2 mRNA levels are decreased in response to the luteal phase or oral contraceptives based on *Z* score analysis as in (**F**). (**H**) Schematic illustrating maintenance of estrogen homeostasis by RSK2. See Discussion for further explanation. See **Figure 5-14** and **Table 5-1**.

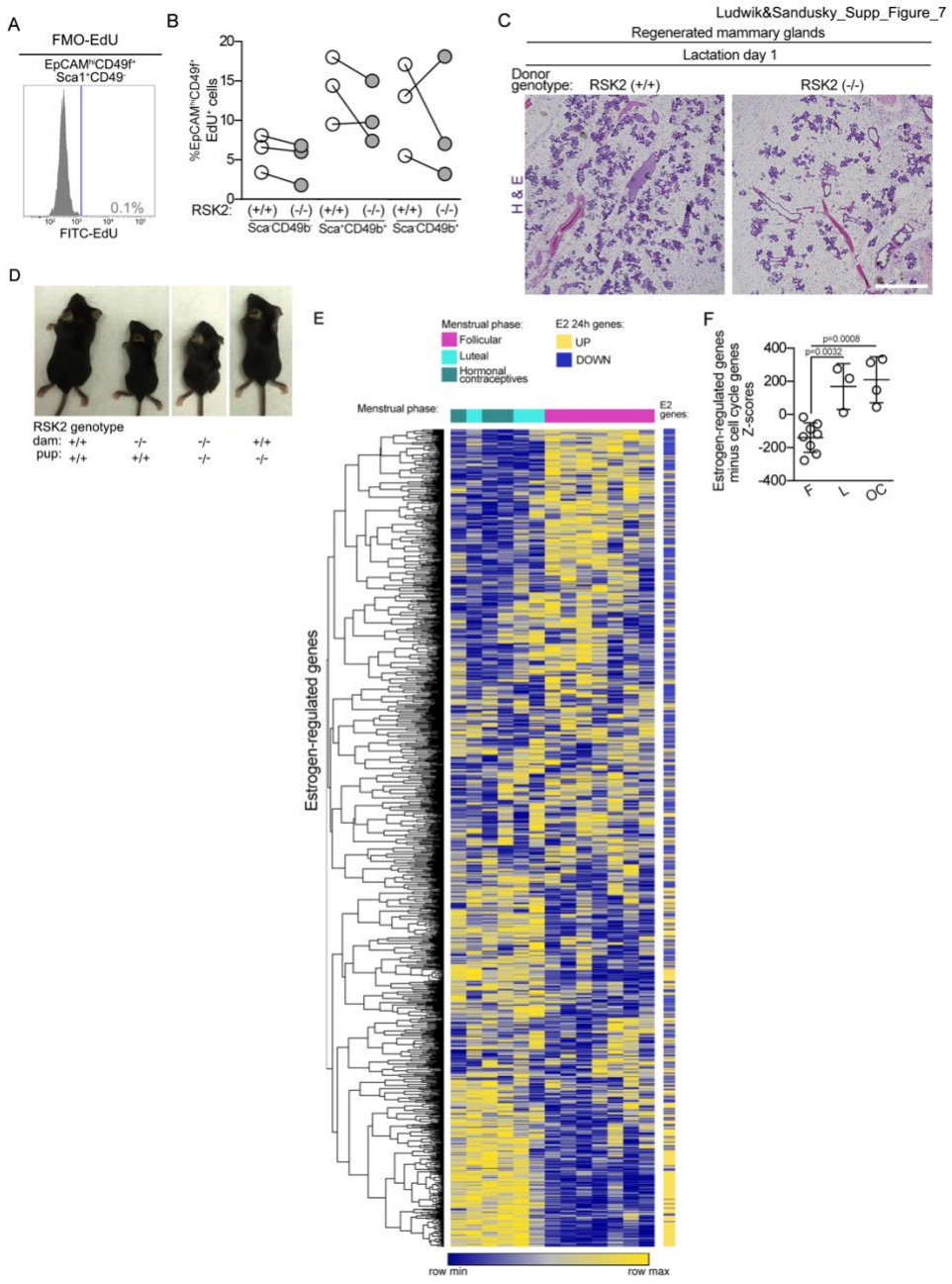


Figure 5-14. RSK2-KO dams fail to provide adequate nutrition for their pups. (**A**) Fluorescence minus one strategy for determining the gate for FITC-EdU. (**B**) FACS analysis of proliferation of mammary glands using RSK2-KO or WT MECs regenerated in a WT mouse. (n=3 glands/genotype; paired Student's t-test). (**C**) Alveolar expansion is reduced in mammary glands regenerated from RSK2-KO mammary epithelial cells as shown by the H&E stains of mammary glands isolated from the same WT dam 1 d after birth. Scale bar = 1 mm. (**D**) Representative images of WT and RSK2-KO pups at 21 d nursed by either WT or RSK2-KO dams. (**E**) Heat map illustrating that estrogen-regulated signature is enriched in the luteal phase and by oral contraceptive use. (**F**) Proliferation genes do not drive the enrichment for the estrogen-regulated signature in individuals in the luteal phase or those taking oral contraceptives. Cumulative Z-scores were generated for each individual by summing

individual Z-scores of genes up regulated in estrogen-regulated signature and subtracting individual Z-scores of genes down regulated. (mean ± S.D., one-way ANOVA with Holm-Sidak's correction for multiple comparisons) (**Table 5-1**).

To further evaluate the physiological importance of RSK2 in estrogen responsiveness, we investigated the remodeling of the mammary gland that occurs during pregnancy. This remodeling is dependent on the ERα cells within the mammary gland, which act as sensors to facilitate alveolar expansion and lactation.[230,255] Alveolar expansion in the whole-animal knockout (**Fig. 5-13d**) and in glands regenerated from RSK2-KO (**Fig. 5-14c**) were reduced, consistent, with the decrease in the ERα population observed in the RSK2-KO glands. Pup weight was reduced in litters arising from RSK2-KO crosses, which could be rescued by fostering RSK2-KO pups to WT dams (**Fig. 5-13e** and **5-14d**). These results argue that the reduced alveolar expansion in the RSK2-KO dams does not provide sufficient nutrition for the pups, rather than a developmental defect in the offspring. These results support our hypothesis that RSK2 is a critical regulator of estrogen responsiveness *in vivo*.

Estrogen Homeostasis in the Human Breast

To evaluate whether RSK2 also functions in regulating estrogen responsiveness in humans, we examined transcriptomic data obtained from normal breast tissue at different stages of the menstrual cycle or from women who were taking oral contraceptives.[251] In women taking oral contraceptives, the levels of synthetic estrogen remain elevated over the time the drugs are administered. In individuals in the luteal phase and in those taking oral contraceptives (**5-13a**), a significant correlation was observed with the estrogen-responsive gene signature obtained from the ERα breast cancer cell line MCF-7.[260] Interestingly, RSK2 mRNA levels were inversely correlated with the estrogen-responsive gene signature (**5-13b** and **5-14a**), which is consistent with the RSK2-KO data. This correlation was not driven by cell cycle genes (**5-14b**). We propose that individuals who take oral contraceptives are subject to prolonged estrogen-responsive gene expression in comparison to individuals who are normally cycling.

Discussion

All ER$^+$ tissues respond to estrogen signaling and, therefore, are subject to the normal fluctuations in the levels of estrogen that occur throughout the estrous cycle. The importance of estrogen signaling is highlighted by the numerous physiological alterations, which occur during menopause, oophorectomy, or anti-estrogen therapy.[225] Here, we provide the first evidence that growth-factor signaling through the ERK1/2-RSK2 pathway is required to maintain cyclic estrogen responsiveness *in vivo*. In the schematic for the WT mice (**5-13c, left**) in step one, we propose that the estrogen pulse in proestrus activates growth factor pathway signaling. This hypothesis is based on observations in neuroendocrine tissues that ERK1/2 is activated after the estrogen surge.[274] Consistent with these data, we found in the mammary gland that ERK1/2 was activated in estrus and that activation was dependent on estrogen. The second step of the schematic shows we identified that ERK1/2 phosphorylates ERα to enhance degradation through the 26S proteasome pathway because mutation of the ERK1/2 phosphorylation site Ser-118 prevents ERα degradation. The most likely mechanism for the increased ERα turnover is through creation of a phosphodegron at Ser-118, which results in E3 ligase recruitment.[275] In step three, we determined that activated ERK1/2 drives ERα degradation to enhance estrogen-responsive gene expression. Additionally, activated RSK2, which regulates protein synthesis[273], was identified to be important in translation of the estrogen-mediated gene program. The physiological importance of RSK2 translational regulation is demonstrated by the reduced pup size and decreased fertility in the RSK2-KO female mice. We propose that the fertility defect is most likely explained by decreased translation in the glandular epithelium because of the loss of RSK2 because estrogen-induced glandular secretions are known to be important for implantation.[276] To reset the cycle, we propose in step four that ERK1/2 is dephosphorylated and inactivated by phosphatases. This hypothesis is based on observations in neutrophils that estrogen upregulates expression of ERK1/2 phosphatases[277] and our data

demonstrating that total ERK1/2 protein levels do not vary with estrogen levels. The cycle is then reinitiated at the next proestrus.

In contrast to temporal activation of ERK1/2 during the estrus cycle, we show in the schematic for the RSK2-KO mice (**5-13c, right**) the disruption of this homeostatic mechanism because of the loss of RSK2. We determined that the loss of RSK2 maintains activation of ERK1/2 in diestrus, which results in increased estrogen-responsive gene expression. We identified that loss of RSK2 resulted in elevated ROS levels, and we hypothesize that this increased ROS inhibits phosphatase activity. This hypothesis is supported by studies showing that oxidation of the reactive-site cysteine in ERK1/2 phosphatases results in their inactivation.[278,279] We speculate that the increased ROS is a result of elevated estrogen-responsive gene expression, which is known to occur[270] and to increased energy requirements. This later hypothesis is supported by gene ontology analysis of the NCL population at estrus, which showed an over-representation of genes associated with the mitochondria. We conclude that RSK2 regulates estrogen-responsive gene expression by controlling redox homeostasis. These findings represent a previously unidentified function for RSK2. Negative regulation of estrogen-responsive gene expression by RSK2 was unexpected because its contributions to ERα breast cancer are well established.[280–282]

We show the importance of ERK1/2 in regulating ERα degradation *in vivo*. Phosphorylation of ERα at Ser-118 has been reported to occur by a number of different kinases and has been associated with increased ERα-mediated transcription in breast cancer cells.[236,245,269,283–286] Furthermore, mutation of Ser-118 to Ala in an ectopic expression system prevented degradation. Numerous ubiquitin ligases have been reported to regulate ERα stability, and components of the 26S proteasome are found in association with ERα on the chromatin in studies using breast cancer cells.[239] It is unclear whether the degradation mechanism differs between breast cancer and normal physiology because ERα protein levels are higher in breast cancer[287], which does suggest that the homeostatic mechanisms have been disrupted.

We also report the analysis of gene expression in the purified NCL population. Relatively few differences in DEGs were detected in the WT mice between estrus and diestrus as compared with the RSK2-KO mice. We propose that the increased gene expression is driven by the continuous ERα transcriptional activation in response to activated ERK1/2 in the RSK2-KO mice. To accurately compare gene expression in the WT and RSK2-KO mice, we developed a FACS protocol that permitted mixing the genotypes and sorting simultaneously. This approach eliminated artifacts from differences in staining among preparations.

RSK2 regulation of estrogen responsiveness occurs in the mature gland but not during puberty. It is possible that unopposed estrogen action is required to facilitate the extensive remodeling of the gland that begins at puberty. However, in the adult, this extensive proliferative response could lead to dysfunction and hyperproliferation within the gland. In human females, we observed an inverse relationship between RSK2 mRNA levels and an estrogen-responsive gene signature in the breast tissue of women in the luteal phase or on oral contraceptives. Consistent with those observations, RSK2 mRNA levels also decreased in endometrial tissue of women in the luteal, compared with the follicular, phase.[288] ERα protein levels are known to decrease in women taking hormone-replacement therapy, suggesting increased ERα-mediated transcription-coupled degradation occurs in those individuals.[287] We speculate that RSK2 levels are decreased in individuals in which normal estrogen levels are disrupted resulting in chronic activation of ERK1/2 and dysregulated estrogen-mediated transcription. This increase in estrogen-mediated signaling could lead to an increase in DNA damage as we observed in the RSK2-KO mice and may account

for the higher risk of breast cancer associated with the use of hormonal contraceptives and hormone-replacement therapy.[289–293]

CHAPTER VI

GENERAL DISCUSSION AND CONCLUSIONS

When I began my graduate studies, the well-known causes of endocrine therapy resistance were limited to TP53 mutations and activation of various growth factor signaling pathways.[294–298] This meant a gap in knowledge about how ETR occurs in the majority of ER+ breast cancers, a large dearth of biomarkers for identification of which patients would respond to certain antiestrogen treatments, and the absolute need for further treatment options for non-responsive tumors. Given that the behavior of breast cancers is mirrored by profound alterations in the transcriptome of their cells[299], next-generation sequencing techniques have been key to the identification of several aspects of tumorigenesis and malignant phenotypes. Furthermore, investigation of the epigenetic landscape and its regulation of the cancer transcriptome has provided key insight to tumor behavior. In line with these observations, the overarching goal of my graduate research was to further our understanding of how ERα transcriptional activity is regulated in ER+ breast cancer by utilizing genomics. This goal included finding biomarkers that could inform cancer prognosis, response to endocrine therapy, and potential targets for further therapy. To achieve this goal, I generated RNA-seq data from ER+ breast cancer cell lines, analyzed RNA-seq data from ER+ breast cancer patients, as well as ChIP-seq data from cell lines, and performed downstream bioinformatic analyses. The data presented in this dissertation furthers our knowledge of regulation of the ERα cistrome, and how that regulation contributes to ETR.

ERα Cistrome in ER+ Breast Cancer

Binding Profile Effectors and Consequences of Altered ERα Binding Profiles

I identified mutation of MLL3 as an effector of the ERα binding profile in my first author paper, chapter 2. MLL3 is one of the main H3K4 methyltransferases at enhancers, along with its paralogue MLL4. Loss of MLL3 through KD in ER+/HER2- breast cancer cells leads to decreased H3K4me1 globally, as well as a major shift in the binding locations for ERα. These new binding locations, along with the new transcriptome in MLL3 KD cells, are enriched for endocrine therapy resistance terms. In addition, the breast cancer cells with MLL3 KD have increased resistance to antiestrogen therapies fulvestrant and tamoxifen.

While the enhancer landscape, defined in my study by H3K4me1 marks as found by ChIP-seq, was drastically altered upon MLL3 KD, recent studies have shown that loss of MLL3 has additional roles in transcriptional regulation outside its catalytic activity. Loss of MLL3 results in loss of Mediator and Pol II binding on active enhancers; this exerts a transcriptional effect by diminishing enhancer activities.[300,301] Moreover, MLL3 is vital to the stabilization and recruitment of KDM6A to the chromatin, which aids in activation of enhancers via its demethylation of H3K27me2/3.[302] Another study showed that MLL3's paralogue MLL4 was only able to promote long-range transcriptional activation through enhancer-promoter looping in the event that its catalytic SET domain was intact.[302] These recent revelations about the transcriptional control MLL3/4 wields are further evidence that mutations in many different domains of MLL3 may affect its function in breast cancer, whether through recruitment of Mediator and Pol II, stabilization or KDM6A, or the ability of the enhancer-promoter loop to form. In the future studies portion of this chapter, I will explore how to refine our mechanistic model for how MLL3 mutation alters the enhancer landscape, ERα binding profile, and transcriptional output of ER+ breast cancer cells.

FGFR1 amplification also alters the ERα cistrome in breast cancer. We found that FGFR1 amplification in post-menopausal ER+ breast cancer patients was able to maintain cell proliferation through estrogen-independent transcription of ERα target genes. This persistence was accomplished through the tyrosine-kinase dependent interaction of FGFR1 and ERα at transcription start sites; again, we find that transcription of ERα targets can be accomplished by modifying the activity of coregulators, here FGFR1, in an estrogen-independent setting. The proliferative capacity of these FGFR1 amplified, ER+ breast cancer cells was successfully diminished with dual pharmacologic inhibition of FGFR1 and ERα. While few clinical trials have been completed with this dual pharmacologic inhibition,[303,304] a more recent study shows that the resistance to ER, PI3K, and CDK4/6 inhibitors in FGFR1 amplified ER+ breast cancers can be ameliorated with mTOR inhibition.[305]

The transcriptional targets of ERα are often those most important to the survival and proliferation of ER+ breast cancer cells. However when faced with antiestrogen therapy, breast cancer cells can utilize signaling pathways outside estrogen-ERα to promote the transcription of these targets and ultimately survive. This is evidenced above with FGFR1 amplification decoupling ERα target transcription and breast cancer cell survival from dependence on estrogen. Another of these pathways implicated in ETR involves CDK4/6 activation; accordingly endocrine therapies can be paired with CDK4/6 inhibitor palbociclib to treat ER+ breast cancers in the clinic. As is the case with most therapies employed against breast cancer, eventually the tumor becomes resistant and the patient relapses. In our investigation of ER+/HER2- breast cancer cells, a transcriptional signature featuring enrichment of PI3K activity and RB1-loss pathways was found in cells resistant to combinatorial FulvPalb treatment compared to parental cells. The enrichment of these pathways point to possible weaknesses to target in ER+ breast cancers that have displayed resistance to FulvPalb treatment, including PI3K inhibitors and standard chemotherapeutics. While we have not performed ERα ChIP-seq for this study, based on the current literature I believe that ERα transcriptional activity may be reliant on activation by the PI3K cascade in the setting of FulvPalb resistance.

While the successful promotion or inhibition of ERα transcriptional targets involves several categories of effectors, including histone modifiers, coactivators, coregulators, and signaling pathways outside the classical E2-ERα cascade, availability of estrogen and ERα themselves plays a vital role in breast development, homeostasis, and dysregulation leading to cancer. Scenarios in which availability of estrogen is perturbed are incredibly relevant to women who take hormonal contraceptives and hormone replacement therapy. Usage of either has been associated with an increased risk in breast cancer,[291,306] and the findings of our RSK2 study in chapter 5 give clues as to why this may be. Disruption of normal levels of estrogen resulting from either hormonal contraceptives or hormone replacement therapy may decrease RSK2 expression, thereby leading to greater activation of ERK1/2 signaling. Activation of ERK1/2 signaling may then lead to enhanced estrogen-responsive gene expression, driving hyperproliferation of the breast tissue and higher levels of DNA damage. Both phenomena could easily set tumorigenesis in motion.

Utilizing Genomic Information in a Predictive and Therapeutic Capacity

Generation, Collection, and Sharing of Cancer Genomic Information

The samples and models that we employ to study the ERα cistrome and its reflection on response to antiestrogen treatment must be well-chosen to glean as much applicable insight into the biology of ER as possible. So how do we best collect and analyze data about the ER binding profile and transcriptional output to delineate the mechanisms behind ETR? One clinical setting to take advantage of is neoadjuvant endocrine therapy trials in primary breast cancer. These patients have ER+ breast cancer, are already set to have surgical resection of their tumors, are available to be asked for consent for genomic analyses of

their tumor biopsy and resection, and will have documented responses to antiestrogen therapy.[307] Longitudinal collection of biospecimens, especially pre- and post-treatment, gives us the opportunity to identify the molecular underpinnings of *de novo* as well as acquired resistance. Several such studies have identified MLL3 as a risk factor for endocrine therapy resistance. One study done with dual hybrid-capture DNA/RNA sequencing in 12 primary and local recurrence patient-matched samples found that continuous endocrine therapy over several years led to enrichment of pre-existing MLL3 SNVs in the local recurrence across multiple patients.[308] Another study sequenced 11 pairs of primary and metastatic lesions to discover that an alteration in any three of MYC, MLL3, or EPHA7 correlated with early relapse in adjuvant antiestrogen therapy, taking the average overall survival of patients from 144.5 months to 90.7 months.[57]

Several longitudinal studies such as these speak to the heterogeneity of not only ER+ breast tumor response to antiestrogen therapy, but also the molecular mechanisms in which cancer evades death on a tumor-level and a singular cell level. One such multi-platform study of pre-treated and post- 4 months treated tumors revealed that while clonal outgrowth is seen in some breast tumors after antiestrogen treatment, not all tumors evade treatment via drastically changing their clonal dynamics.[309] Another, shorter-term study found that among 58 ER+ breast cancer tumor paired samples pre- and post-neoadjuvant antiestrogen therapy, there was no one single gene that was altered in all cases.[310] Importantly, Razavi et al. found in their large-scale study of 1,501 HR+ breast tumors that this heterogeneity can even expand to the molecular mechanisms breast cancer may employ within one patient. They found that one patient with sequencing of multiple metastatic lesions harbored an ESR1 mutation in one metastatic site and a ERBB2 mutation in another site.[311] This multi-resistance mechanism phenomenon was also seen in a study with matched primary tissue, metastatic tissue, and circulating tumor DNA (ctDNA). Although generally only one ESR1 mutation was detected at a metastatic site, 40% of patients had more than one ESR1 mutation in their ctDNA sample.[312] Other lines of evidence that ctDNA captures mutations facilitating treatment resistance within patients[313–315] support the use of ctDNA in clinic to survey what possible targeted therapies would be useful in an evolving treatment program.

The types of genomic techniques employed in the pursuit of prognostic biomarkers and druggable targets for ER+ breast cancer must also be considered when accruing precious clinical samples. While my research detailed here utilizes RNA-seq and ChIP-seq data, there are several alternate or additional techniques that provide different types of valuable information about ER+ breast cancer genomics. One technique that can provide information about direct interactions between ERα and epigenetic marks is fluorescence lifetime imaging-based Förster resonance energy transfer (FLIM-FRET). This technique works by utilizing energy transfer of fluorophore molecules attached to proteins or receptor and ligands that interact. In FLIM-FRET when the two tagged molecules are in 2-10 nm of proximity to each other, the donor fluorescence is quenched.[316] A study using this technique found that ERα interacts with H3K27ac and H4K12ac. This interaction led the authors to utilize histone acetyltransferase inhibitor anacardic acid (AA) in combination with tamoxifen, which not only reduced H4K12ac at EREs but suppressed MCF7 cell growth in vitro and in mice xenografts.[317] While this study subjected only MCF7 and T47D breast cancer cells to FLIM-FRET, there remain many ER+ breast cancer cell lines with different mutational profiles that could be utilized in combination with FLIM-FRET to provide clues about potential therapeutic targets. In addition, this technique would be useful in the investigation of how mutation or KD or recurrently mutated chromatin modifier genes such as MLL3 affect the direct interaction of ERα with histone modifications.

Another genomic technique that has proven fruitful in terms of information gleaned about transcriptional regulation in ER+ breast cancer cells is single cell RNA-seq (scRNA-seq). This differs from bulk RNA-seq in that individual cells are isolated, lysed

to preserve mRNA, and then mRNA is primed and reverse transcribed into cDNA. The libraries made from this cDNA are then pooled for sequencing.[318] Thus this technique elicits a clearer picture of the heterogeneity that exists within breast tumors, which can be paired with phenotypic data to determine causes of differential response to therapy within one population of cancer cells. A recent paper explored the mechanisms behind long-latency relapse in ER+ breast cancer, as patients relapse on endocrine therapy at a rate of 3% of patients per year up to 20 years after surgery.[319] They identified a rare subpopulation of cells they termed "pre-adapted" that transcribe genes related to dormancy as well as mixed epithelial and mesenchymal features. While these cells can endure short term antiestrogen therapy, they require additional genetic mutations and transcriptional reprogramming to fully reconstitute a tumor cell population in vivo.[320] These findings obtained using scRNA-seq help explain how ER+ breast cancers are able to reappear after decades of endocrine therapy, which gives hope for biomarkers of the process of recurrence and therefore better managed treatment plans.

The sharing and storage of all the genomic information generated from breast cancer studies must be considered as well. This enables scientific peers to not only recreate analyses in the name of responsible conduct of research, but to use genomic data generated by their peers in their own research. As an example, within my own graduate work I utilized publicly available TCGA data to investigate my hypotheses about the transcriptional consequences of MLL3 mutation. There are many online databases for this exact purpose. Among those specific to cancer and cancer cell lines are the International Cancer Genome Consortium, TCGA, the BROAD Tumor Portal, Therapeutically Applicable Research To Generate Effective Treatments (TARGET), and The Cancer Cell Line Encyclopedia (CCLE). Until recently, researchers utilizing these large databases like the TCGA would typically have to provision for several days to download large sets of genomic data and storage on local servers. However in 2016 the Cancer Genomics Cloud (CGC) project addressed these obstacles to collaborative progress and launched the cloud computing structures made by three groups for the National Cancer Institute: the Broad institute, the Institute for Systems Biology, and Seven Bridges Genomics.[321] This project lowers the barrier to entry for many scientists who may not have the computational background or resources to utilize large scale genomic data in their research, while still retaining the flexibility to utilize their own programming for analyses.

Another category of efforts making genomic data available for use in studies toward precision medicine are platforms aimed at identifying relevant genomic resources and integrating them for faster, easier use. Some of these data mining tools come in the form of packages for programming in R, such as the R/Bioconductor CuratedTCGAData and cBioPortalData packages.[322] Web-based platform Precision Medicine Knowledgebase (PreMedKB)[323] intends to find documentation integrating search terms for diseases, genes, variants, and drugs, and create semantic networks based on relationships between search terms. While I found the website to be glitchy, it is still relatively new and hopefully will be optimized soon to become more user-friendly. Application Programming Interface (API) Mastermind[324] accesses and bundles associations between diseases, phenotypes, genes, variants, and therapies into a genomic search engine. This platform accounts for the disparateness of biological nomenclatures (ex. MLL3 is also known as KMT2C), and its use in comparison of commercially available heredity cancer panels highlighted discordance between the panels themselves and the evidence for cancer-risk genes in the current literature.

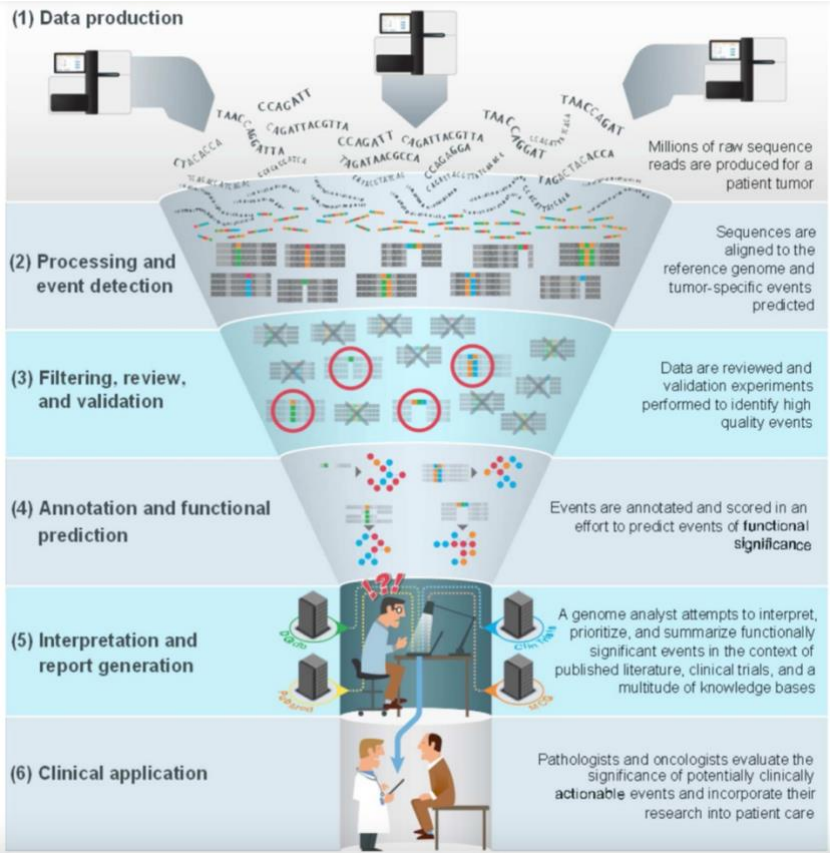


Figure 6-1. Cancer Genomics Workflow. The bottleneck of clinical interpretation is pictured at the bottom of the workflow. This figure was taken from Good et al.[325]

Analysis and Utilization of Cancer Genomic Information

The biologic samples used to generate genomic data, the type of technology used to collect data, and the sharing and storage of the data are all important components leading to the end goal of data analysis and subsequent interpretation of genomic data. Identification of variants that drive cancer must be robust, as it leads to inclusion or exclusion of patients from clinical trials, decisions regarding therapy management plans, and creation of prognostic gene panels. The bottleneck for these facets of cancer precision medicine is often clinical interpretation of genomic variants (**Fig. 6-1**).[325] The research detailed in this dissertation suggests that MLL3 mutation, FGFR1 amplification, and RB1 signatures should be included on ER+ breast cancer gene panel. However, future studies of the breast cancer transcriptome may employ more recent analytic methods.

One tool available for analysis of cancer genomic data is machine learning (ML). As a subtype of artificial intelligence (AI), ML is an analysis method that identifies patterns and make predictions based on provided data. The algorithms that result from ML models are fueled by the data itself, with minimal human programming. While single types of data can be subjected to ML methods, multi-omics, or a combination of several types of data such as transcriptomics, epigenomics, and proteomics, is often analyzed with ML. For example the package DeepProg utilizes ML and deep learning, a more complex and stacked form of AI, on transcriptome, DNA methylation, and miRNA data from the TCGA to more accurately predict patient prognosis than the standard Cox-proportional hazards method.[326] Even within the realm of gene regulation exclusively, ML can be employed to identify molecular subtypes, gene regulatory networks, biomarker discovery, prognostic predictions, re-purposing of drugs, and therapy response predictions.[327–330]

These new tools with which to analyze the growing compendium of genomic data lead to the possibility of more precise personalized medicine in the future. Predictive markers will more accurately reflect response to treatment and prognosis, especially as more clinical data about response to targeted therapies is collected. While submitting biopsies to RNA-seq or ERα ChIP-seq for each breast cancer patient to determine their treatment plan is not feasible due to cost and time, I believe that within the next several decades AI techniques will be able to reduce multi-omics information about breast cancer to a manageable number of dimensions suitable for updated prognostic panels. For example, this process could be a study utilizing phosphoproteomics, transcriptomics, mutation information from exome sequencing, and treatment response in the form of progression free survival (PFS) and overall survival (OS) with ML algorithms to elucidate a 20-gene expression panel in combination with IHC for 3 phosphorylated proteins, and mutation status of 10 genes to accurately predict response to

specific forms of antiestrogen therapy. Discovery of circulating tumor cell (CTC) or ctDNA biomarkers for relapse during or after adjuvant therapy is a possibility as well, as recent papers have shown transcriptional changes and acquired genetic mutations over several years may be the mechanism behind the latency of ER+ breast cancer cells that eventually reconstitute tumors.[320,331]

Treatment management for ER+ breast cancer has possibility for improvement based on genomic research on many fronts. As I mentioned before, research employing ML with combined multi-omics and clinical data may be able to identify drugs that can be repurposed in a clinical setting. For instance, a recent study found that MLL3 or MLL4 inactivation disrupted homologous recombination-mediated DNA repair, and thus sensitized lung cancer cells to Poly ADP Ribose Polymerase inhibitors (PARPi) that are normally only administered in the setting of BRCA1/2 mutations.[332] This kind of re-purposing of currently available drugs may be more frequently possible, and successful, with the new technology we have at hand.

Intratumor heterogeneity presents a major dilemma to be addressed in the clinic, as one type of therapy may not elicit the same level of effect on all cells within a tumor. One possible approach is to utilize drugs targeting epigenetic dysregulation in combination with traditional therapies so that diverse pathways contributing to the proliferation and survival of the cell can be disrupted.[333] This would also potentially impede cancer cells from evading antiestrogen therapy through specific epigenetic remodeling, which has led to the recent studies of "persister cells." In fact HDAC inhibitors act synergistically with antiestrogen therapies.[334,335] Additionally, BRD4 inhibitor JQ1 can be utilized in combination with fulvestrant to successfully inhibit tumor growth in MCF7 xenografts resistant to tamoxifen, as BRD3/4 was found to activate ERα transcription through recruitment of NSD1 to methylate histone H3K36.[336] Within my own research, targeting MLL4 in MLL3-mutant breast tumors may lead to ablation of ERα binding, as MLL4 may take over the place of mutant MLL3 in the ASCOM complex. Interestingly, KDM6A inhibitor GSKJ4 inhibits expansion of the breast cancer stem cell compartment induced by paclitaxel treatment.[337,338] KDM6A is part of the ASCOM complex that MLL3 belongs to, and coordinates demethylation of H3K27 in parallel with methylation of H3K4 for activation of enhancers. Thus targeting KDM6A in MLL3-mutant tumors may ablate endocrine therapy resistance in cases where MLL4 and KDM6A cooperate to activate stem-cell like transcription.

Additional Considerations for MLL3 Studies in ER+ Breast Cancer

Numerous published studies show MLL3 is one of the regulators of ERα binding. However, not all these studies account for additional factors that may influence the interplay between MLL3 and ERα. One of these additional factors is mutational status of PI3K pathway components. As stated in Chapter 2, MLL4 can be regulated by AKT1, leading to ERα-driven therapeutic resistance to PIK3CA inhibition. This is probably accomplished by MLL4 recruiting ERα to shifted binding locations on the breast cancer genome, driving transcription of genes that aid in resistance to targeted therapy. A recent study using a somatic mammary stem cell-based organoid model showed that concomitant MLL3 inactivation and PIK3CA overexpression led to stem cell self-renewal instead of differentiation of the cells.[339] This suggests again that the interplay between PI3K pathway components and MLL3 mutation play a large role in the transcription of breast cancer cells.

TP53 mutation status should also be considered when interpreting the results of studies on MLL3 mutation. MLL3 has been shown cooperate with p53 to recruit MRE11, a DNA replication restart nuclease, to stalled forks. However, this recruitment is impaired in p53 depleted cells.[340] MLL3/4 have also been shown to coactivate p53 transcriptional targets.[165(p4)] Thus MLL3-mutant cancers may display a higher level of genomic instability.

Recently MLL3 and MLL4 have also been implicated in anti-tumor immunity, through direct interactions with epithelial transcriptional factor GRHL2. MLL3/4-GRHL2 interactions increase an epithelial gene expression program including ICAM-1 and multiple IFN response genes.[341] This sensitizes tumor cells to natural killer (NK) cells, as well as prevents epithelial to mesenchymal transition (EMT). Given this information, MLL3 mutant cells are at a higher risk for evading NK cells and undergoing EMT.

Lastly, there are conflicting reports about whether MLL3 is a tumor suppressor or an oncogene. Kim et al. found that while MLL family proteins did indeed help regulate the transcriptional activity of ERα, depletion of MLL3 in tamoxifen-resistant breast cancer cells inhibited ERα target gene expression as well as cell proliferation[94]. Treatment with fulvestrant exacerbated these effects. This study suggests that MLL3 would make a good target for therapeutics. However, the authors used MCF7 and T47D cell lines; MCF7 cells were once believed to harbor an MLL3 mutation, and T47D harbors a missense G892E in MLL3. In addition, cBioPortal records T47D cells as harboring a TP53 L194F missense mutation and the PIK3CA H1047R mutation. MCF7 cells harbor a missense mutation E545K in PIK3CA. Consequently, I do not believe these are the best ER+ breast cancer cell lines with which to study the effects of MLL3 mutation, unless the involvement of PIK3CA is considered.

Future Directions

There are several future directions for the study of how MLL3 affects the transcriptional regulation of ERα in ER+ breast cancer with respect to endocrine therapy resistance. The first is to analyze the H3K27ac ChIP-seq we have already performed in ZR751shLucif and ZR751shMLL3 cells. This will give us a clearer picture of the enhancer landscape in ER+ breast cancer cells with loss of MLL3, as KDM6A demethylates H3K27me3 to allow subsequent acetylation.[342,343(p4)] The combination of H3K4me1 and H3K27me3 marks poised enhancers, whereas H3K4me1 in combination with H3K27ac marks active enhancers. This information will further refine our model for how MLL3 loss alters the epigenetic landscape, and consequently the ERα binding profile, especially for cases of MLL3 mutations that affect the association of MLL3 with KDM6A.

Another incomplete picture that needs to be completed is parsing apart the separate roles that MLL3 and its paralogue MLL4 in regulation of ERα transcriptional activity in breast cancer, especially in the context of antiestrogen therapy. Several lines of evidence point to separate functions. First, KO of MLL3 produces a different phenotype than KO of MLL4 in mice. MLL3 KO mice die around birth due in part to failure of normal lung development, but MLL4 KO mice die at 9.5 days as an embryo due to gastrulation defects.[91,344,345] Heterozygous variants of MLL4, but not MLL3, cause the rare congenital disorder Kabuki syndrome.[346] MLL3 is mutated in ~7-10% of ER+ luminal breast cancers whereas MLL4 is mutated in less than 2% of ER+ luminal breast cancers. One of these cases also has a TP53 mutation (I255S) as well. In fact, cBioPortal shows MLL4 is amplified in two of the ER+ TCGA breast cancers where MLL3 has a truncating mutation (N621 frameshift insertion and E1486 nonsense), and PIK3CA is also mutated (E545K in both). Notably, MLL4 is able to directly interact with the common ERα mutant Y537S implicated in ETR.[347] Knocking down MLL4 instead of MLL3 in an ER+ breast cancer line WT for MLL4 but mutant for MLL3, and then generating and analyzing RNA-seq and ERα, H3K4me1, and H3K27ac ChIP-seq would provide information about compensatory action of MLL4 for mutant MLL3 in ER+ breast cancer. Subjecting these MLL4 KD, MLL3-mutant cells to tamoxifen and fulvestrant therapy in a proliferation assay may show increased sensitivity to antiestrogen therapy in comparison to the shLucif control cell line.

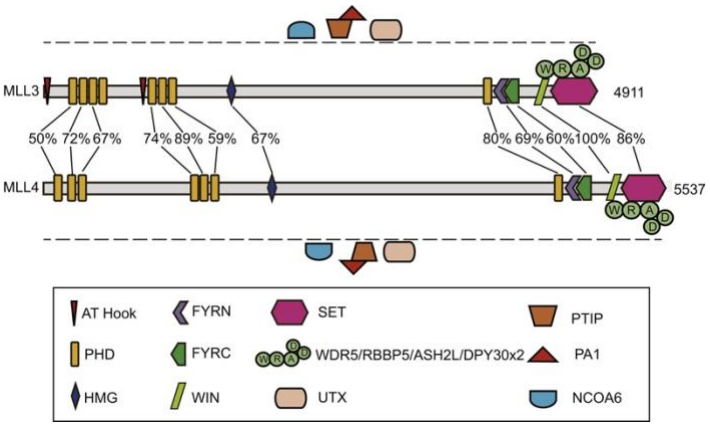


Figure 6-2. Schematic of MLL3/4 domains. While the two genes share similar domains, amino acid sequence similarity is not completely preserved within those functional domains. Abbreviations are as follows: F/Y-rich C- terminus (FYRC); F/Y-rich N- terminus (FYRC); high-mobility group (HMG); plant homeotic domain (PHD); SU(VAR)3–9, E(Z) and TRX (SET); WDR5 interaction (WIN). This figure was taken from Zheng et al.[348]

Cancer studies examining the roles of MLL3/4 have shown that reintroduction of a functional MLL3 and homozygous deletion of MLL4 cause the same negative effect on proliferation.[349,350] The differences in functional domain amino acid sequence (**Fig. 6-2**) may account for this, as several domains aid in interaction with the chromatin and other coregulators. For example, MLL3 is recruited by BAP1 to enhancers of BAP1 targets to regulate tumor suppressors; this recruitment and tumor suppressor expression was reduced in cells that had either a BAP1 deletion or MLL3 mutations in PHD domains.[351] Furthermore, BAP1-MLL4 interactions were not detected in this study. Interestingly, many mutations occur in PHD domains of MLL3, but not in MLL4, in TCGA breast cancer.

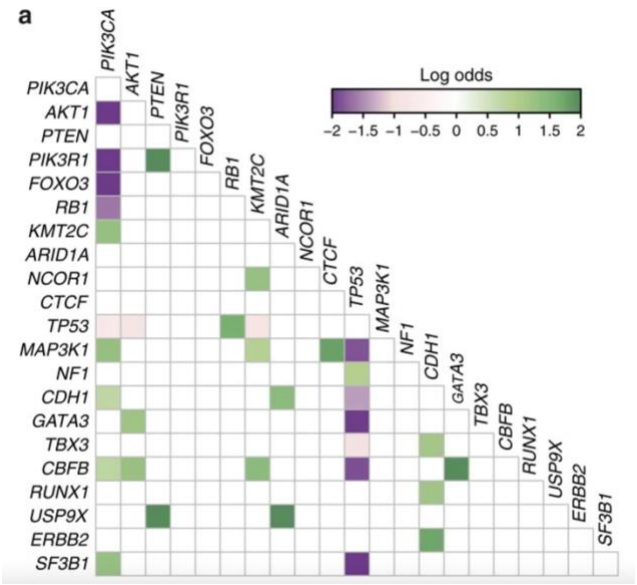


Figure 6-3. Mutations between MLL3 (here, KMT2C) and PIK3CA are often co-occurring. This pairwise association plot was constructed using mutation status of genes mutated in at least 0.5% of 2,433 mixed ER+ and ER- breast cancer samples. Only pairwise associations with FDR<=0.1 were plotted (Fisher's exact test). Magnitude in terms of association by log odds is represented in the color scale. This figure was taken from Pereira et al.[352]

The involvement of the PI3K pathway in the function of MLL3/4 in ER+ breast cancer must be investigated in depth to characterize its effect on ERα transcriptional regulation. Of note, MLL3 and PIK3CA are recurrently co-mutated together in breast cancer (**Fig. 6-3**). While MLL4 is known to be phosphorylated by AKT1 and SGK1 for inactivation[353,354], MLL3 does not possess the AGC kinase consensus sequence (RXRXXS/T) adjacent to the PHD finger cluster utilized by AKT1 and SGK1. These details paint a compelling picture when paired with the fact that targeted PI3K inhibition of PIK3CA mutant-breast cancers exhibit an MLL4-dependent increase in ER-regulated transcription and shift in ERα binding locations. As such, the relationship between loss of MLL3, inhibition of PIK3CA, and anti-estrogen therapies may provide new avenues for therapy of endocrine resistant tumors. Knocking down MLL4 in a PIK3CA-mutant ER+ breast cancer line WT for MLL4 but mutant for MLL3 and then generating and analyzing RNA-seq and ERα, H3K4me1, and H3K27ac ChIP-seq would provide information about compensatory action of MLL4 for mutant MLL3 in ER+ breast cancer. Proliferation assays of these MLL4 KD, MLL3-mutant cells may highlight an increased sensitivity to antiestrogen therapy in comparison to the shLucif control cell line. PI3K inhibitors should be used in combination with the antiestrogen therapies to examine response to targeted treatment with double hits to MLL3 and MLL4, with the hypothesis that targeting MLL4 or perhaps epigenetic marks at enhancers will re-sensitize the cells to therapy.

APPENDIX

geneSet	enrichment Score	normalized Enrichment Score	pValue	FDR	Genes
CREIGHTON_AKT1_SIGNALING_VIA_MTOR_DN	0.654185888	2.027288832	0	0.00264466	ALDOA;ATP6AP1;ATP6V0C;ATP6V1F;CTSA;DHCR7;GPI;KRT8;MRPS7;PAFAH1B3;PFKL;PPP2R1A;PPP4C;RGL2;TOM1;YWHAB
BENPORATH_PROLIFERATION	-0.496941049	-2.026052775	0	0.003326382	AGFG1;ANLN;ARF1;ASF1A;ASPM;ATAD2;AVL9;C1orf112;CCNB2;CDC123;CDCA7;CDK1;CDKN3;CEBPG;CENPA;CENPE;CENPF;CENPN;CHAF1B;CKS1B;CKS2;CNIH4;COX5A;CSE1L;CSNK1G1;DAP3;DBF4;DEK;DTL;EXO1;EZH2;GARS;GART;GDI2;GGCT;GNB4;GPSM2;GTPBP4;H2AFV;HDAC2;HRASLS;HSPA14;ILF2;KDELR2;LBR;LGALS8;MAD2L1;MAGOHB;MND1;MRPL9;NCAPG;NDC80;NDUFB5;NEK2;NFE2L3;NUDT5;NUF2;PBK;PCNA;PDCD10;PFDN2;PPIL1;PRC1;PRIM2;PRPF18;PSMA3;PTS;PURB;RACGAP1;RAD51AP1;RANBP1;RBM8A;RDX;RFC4;RIT1;RRM2;SLC25A5;SNRPD1;SNRPG;SRPK1;SUV39H2;TFRC;TOP2A;TP53BP2;TTK;TYMS;UBE2T;UGGT1
CREIGHTON_AKT1_SIGNALING_VIA_MTOR_UP	0.612966948	2.032512976	0	0.005289321	AKT1;ARHGEF16;BRMS1;BSG;CDC34;CLDN3;CLSTN1;CORO1B;DDR1;KCTD5;MMP15;MVK;NEU1;PMPCA;POR;PRKCD;RNF126;SPINT1;TJP3;TOLLIP;UBE2M
BHAT_ESR1_TARGETS_VIA_AKT1_UP	0.392669837	1.91260383	0	0.010578642	A4GALT;AATF;ACIN1;ADAMTSL5;ADAP1;ADCY3;ADCY9;ALDH3B1;AMZ1;AP1B1;ARAP3;ARHGEF18;ASB13;ATRIP;AXIN1;BEGAIN;C16orf74;C1QTNF6;C1orf159;CA12;CBFA2T3;CCND1;CDC34;CDC42EP1;CHST8;DEGS2;DHRS3;DOK7;DUSP2;EFHD2;EIF3B;FAM102A;FAM207A;FGFRL1;FLAD1;GATAD2A;GPRIN1;GSG1L;HDAC4;HLA-DRB1;HPCAL1;HR;HSPB8;IL20;IMP4;ISG20L2;KCNK15;KCNK6;KDM4B;KHK;KRT13;LETM1;LHX4;LMO1;LONRF2;LRFN4;LTBP3;MAG;MANEAL;MAPT;MED24;MFSD2A;MTFP1;MYBBP1A;NADSYN1;NCOR2;NCS1;NEIL2;NOL6;NT5DC3;PADI3;PAK4;PARP12;PCYT2;PITX1;PKIB;PODXL;PTGES;PTH1R;PTRH2;RAB11FIP3;RAPGEFL1;RARA;RECQL4;RET;RIMS4;RNF144A;RTKN;SBNO2;SCARB1;SCNN1B;SEC14L2;SHB;SLC10A3;SLC25A25;SLC2A8;SLC6A6;SLC7A2;SLC7A5;SLC9A3R1;SULT2B1;SUSD3;SYNDIG1;SYT12;TBC1D16;TBKBP1;TBX2;TJP3;TMEM104;TMEM120B;TMEM51;TOE1;TRMT61A;TSKU;UNC119;UNC5A;UST;VGF;WFS1;WNT4
MALIK_REPRESSED_BY_ESTROGEN	0.702769294	1.854970337	0	0.015206798	BMP7;CERK;CLDN4;EFEMP1;MUC1;MXD4;NDRG1
HUANG_DASATINIB_RESISTANCE_UP	0.432978901	1.732599841	0	0.037245635	BTN3A2;CAST;CDC42EP3;COL5A1;EGFR;EPHB2;F2RL1;FXYD5;GBP3;IFIT3;IL15RA;INPP1;ITGA5;JAG1;KCTD12;LARP6;LYN;MAP7D1;MSN;PCDH7;PRNP;PSMB8;PSMB9;RAC2;SAMD9L;TFPI
DAIRKEE_TERT_TARGETS_UP	0.353794914	1.748012523	0	0.038347577	ACIN1;ADAP2;AIP;AP4B1;ARHGEF4;ASB8;ATP6V0D1;BAX;BCAT2;BET1L;C11orf68;CA12;CAPN1;CAPZB;CC2D1B;CCS;CDC34;CELF6;CES2;CHRD;CHTF8;CKMT1A;CLPB;CLTB;CNPY3;COL13A1;COL7A1;CORO1A;D2HGDH;DCTPP1;DDOST;DDX41;DGAT1;DGCR6L;DHRS1;DRAP1;ECHS1;EPAS1;FAM3A;FARSA;GALNT6;GEMIN4;GNB2;GPC1;GSTP1;HCFC1R1;HDAC3;HIGD2A;IDH3G;IPO13;IRF2BPL;ISOC2;JMJD4;LAMB2;LDB3;LHB;LMAN2;LRP4;LTBP4;LY6E;MAF1;MAPK11;MECR;MLF2;MLLT6;MMAB;MRPL12;MRPL37;MRPL38;MRPS2;MTX1;NAA10;NDUFA13;NDUFA2;NDUFB11;NFKBIL1;NMT1;NOMO1;NPRL3;NQO2;NSMCE1;NTHL1;NUBP2;OGDH;OGFR;PBX2;PCYT2;PDIA3;PEMT;PIP4K2B;PITPNM1;PLD3;PLOD3;POLD1;POLR2E;PPM1F;PPP1R11;PPP4C;PQBP1;PRNP;PRPF31;QARS;QPRT;RAB40C;RAC2;RASL12;RHOT2;RILP;RIN1;RNF126;ROBO3;RPS5;RTN4RL1;SAMD11;SGTA;SHARPIN;SLC1A5;SLC26A6;SLC2A3;SLC4A3;SPNS2;SRRM2;SSNA1;STX8;TELO2;TERT;THOC3;THRA;TK1;UBXN6;UQCRC1;WDR24;ZNF618;ZNF777

Supplemental Table 1 – ZR751 DEG Webgestalt Gene sets explored had "estrogen" in the gene-set description or title.

geneSet	enrichmentScore	normalizedEnrichmentScore	pValue	FDR	Genes
YANG_BREAST_CANCER_ESR1_DN	-0.763282683	-1.785460168	0	9.81E-04	ARHGEF9;BCL11A;BTG3;CDH3;FABP5;GABRP;LDHB;PROM1;RARRES1;SFRP1;SLC9A6;TLE4;TRIM2;TUBB6;YEATS2
DOANE_BREAST_CANCER_ESR1_DN	-0.679357432	-1.718292085	0	0.00637955	BBOX1;BCL11A;CHST3;CRYAB;DSC2;EGFR;FABP7;FOXC1;GABRP;MID1;MMP7;RARRES1;ROPN1B;SCRG1;SERPINB5;SFRP1;TTYH1;VGLL1
DOANE_BREAST_CANCER_ESR1_UP	0.393700473	1.804795743	0	0.02581214	AZGP1;C1orf21;CELSR1;CFB;DNALI1;EPS8L1;ERBB4;ESR1;EVL;FOXA1;GAMT;GATA3;GDF15;MLPH;MUC1;MYB;MYO6;NAT1;PDZK1;PIP;RND1;SCCPDH;SLC44A4;TJP3;TMC5;TSPAN1;TTC39A
CREIGHTON_AKT1_SIGNALING_VIA_MTOR_DN	0.541431727	1.875717268	0.009709	0.0264417	ATP6V0B;CIB1;PAFAH1B3;TNFRSF12A;TOM1;TSPAN1;YWHAB

Supplemental Table 2 – TCGA DEG WebGestalt. Gene sets explored had “estrogen” in the gene-set description or title.

ZR751 Enrichment	
Size=23; leadingEdgeNum=16; enrichmentScore=0.65; normalizedEnrichmentScore=2.03; PValue=0.000e+0; FDR=2.645e-3.	
Gene Symbol	**Score**
ALDOA	0.701
ATP6AP1	2.9927
ATP6V0C	1.1948
ATP6V1F	0.6892
CTSA	1.4395
DHCR7	0.975
GPI	0.657
KRT8	3.0144
MRPS7	2.0628
PAFAH1B3	1.7336
PFKL	1.0281
PPP2R1A	1.6296
PPP4C	2.4504
RGL2	0.954
TOM1	1.2948
YWHAB	2.024

TCGA Enrichment	
Size=23; leadingEdgeNum=7; enrichmentScore=0.54; normalizedEnrichmentScore=1.88; PValue=9.709e-3; FDR=2.644e-2	
Gene Symbol	**Score**
ATP6V0B	2.9707
CIB1	1.7131
PAFAH1B3	3.3718
TNFRSF12A	1.4081
TOM1	1.2041
TSPAN1	1.5901
YWHAB	1.4634

Supplemental Table 3 – GSEA Creighton_AKT1_Signlaling_Via_MTOR_DN for ZR751 & TCGA

ZR751shLucif ERα - Centrimo					ZR751shLucif ERα – DREME TOMTOM		
ID	***E*-value**	**Region Width**	**Region Matches**		**Query Motif**	**Matches**	**Top Target Motifs**
ESR1_HUMAN.H11MO.0.A	5.80E-42	80	712		AAAATANW	7	MEF2C_HUMAN.H11MO.0.A, MEF2A_HUMAN.H11MO.0.A, HMGA1_HUMAN.H11MO.0.D, ONEC3_HUMAN.H11MO.0.D, MEF2B_HUMAN.H11MO.0.A, MEF2D_HUMAN.H11MO.0.A, FOXJ3_HUMAN.H11MO.1.B
ESR2_HUMAN.H11MO.0.A	1.90E-35	80	727		AAAAVAAA	26	FOXL1_HUMAN.H11MO.0.D, FOXJ2_HUMAN.H11MO.0.C, FOXF1_HUMAN.H11MO.0.D, FOXJ3_HUMAN.H11MO.0.A, FOXJ3_HUMAN.H11MO.1.B, STAT2_HUMAN.H11MO.0.A, FOXC2_HUMAN.H11MO.0.D, PRDM6_HUMAN.H11MO.0.C, CPEB1_HUMAN.H11MO.0.D, FUBP1_HUMAN.H11MO.0.D, STAT1_HUMAN.H11MO.1.A, ZIM3_HUMAN.H11MO.0.C, SRY_HUMAN.H11MO.0.B, OLIG1_HUMAN.H11MO.0.D, IRF9_HUMAN.H11MO.0.C, FOXP1_HUMAN.H11MO.0.A, ZN384_HUMAN.H11MO.0.C, FOXO4_HUMAN.H11MO.0.C, SOX4_HUMAN.H11MO.0.B, FOXQ1_HUMAN.H11MO.0.C, IRF7_HUMAN.H11MO.0.C, SOX2_HUMAN.H11MO.0.A, IRF5_HUMAN.H11MO.0.D, SOX10_HUMAN.H11MO.1.A, ZFP28_HUMAN.H11MO.0.C, BPTF_HUMAN.H11MO.0.D
RARG_HUMAN.H11MO.0.B	8.70E-27	95	919		AAATGTV	2	ZN232_HUMAN.H11MO.0.D, NKX61_HUMAN.H11MO.1.B
ESR2_HUMAN.H11MO.1.A	1.40E-13	108	948		ACASWG	6	SOX9_HUMAN.H11MO.0.B, SOX15_HUMAN.H11MO.0.D, GCR_HUMAN.H11MO.1.A, ANDR_HUMAN.H11MO.2.A, ZN322_HUMAN.H11MO.0.B, DMRT1_HUMAN.H11MO.0.D
ERR1_HUMAN.H11MO.0.A	3.90E-12	86	604		AGACRGGG	1	SMAD1_HUMAN.H11MO.0.D
ESR1_HUMAN.H11MO.1.A	6.90E-12	71	619		ATATWY	1	P5F1B_HUMAN.H11MO.0.D
RARB_HUMAN.H11MO.0.D	1.60E-09	62	409		ATGTTTDC	30	OLIG1_HUMAN.H11MO.0.D, FOXO6_HUMAN.H11MO.0.D, BHE22_HUMAN.H11MO.0.D, FOXO3_HUMAN.H11MO.0.B, FOXP2_HUMAN.H11MO.0.C, FOXC1_HUMAN.H11MO.0.C, FOXA2_HUMAN.H11MO.0.A, FOXO4_HUMAN.H11MO.0.C, FOXK1_HUMAN.H11MO.0.A, FOXF2_HUMAN.H11MO.0.D, FOXA1_HUMAN.H11MO.0.A, FOXD1_HUMAN.H11MO.0.D, FOXF1_HUMAN.H11MO.0.D, FOXA3_HUMAN.H11MO.0.B, FOXM1_HUMAN.H11MO.0.A, FOXJ3_HUMAN.H11MO.1.B, FOXO1_HUMAN.H11MO.0.A, FOXP1_HUMAN.H11MO.0.A, FOXB1_HUMAN.H11MO.0.D, FOXD2_HUMAN.H11MO.0.D, FOXD3_HUMAN.H11MO.0.D, ANDR_HUMAN.H11MO.0.A, ZN586_HUMAN.H11MO.0.C, FOXJ3_HUMAN.H11MO.0.A, FOXJ2_HUMAN.H11MO.0.C, OLIG3_HUMAN.H11MO.0.D, ONEC3_HUMAN.H11MO.0.D, IRX3_HUMAN.H11MO.0.D, FOXQ1_HUMAN.H11MO.0.C, FOXG1_HUMAN.H11MO.0.D
THA_HUMAN.H11MO.1.D	3.30E-08	117	990		BCCAGGM	16	HAND1_HUMAN.H11MO.1.D, ZN436_HUMAN.H11MO.0.C, ZEB1_HUMAN.H11MO.0.A, SNAI1_HUMAN.H11MO.0.C, ZN680_HUMAN.H11MO.0.C, NFIB_HUMAN.H11MO.0.D, ZN335_HUMAN.H11MO.1.A, NR5A2_HUMAN.H11MO.0.B, ZF64A_HUMAN.H11MO.0.D, ETV4_HUMAN.H11MO.0.B, HAND1_HUMAN.H11MO.0.D, ELK3_HUMAN.H11MO.0.D, HIC2_HUMAN.H11MO.0.D, STF1_HUMAN.H11MO.0.B, NFIC_HUMAN.H11MO.1.A, ETV6_HUMAN.H11MO.0.D
NR1I3_HUMAN.H11MO.1.D	1.30E-07	71	632		BCTGKG	9	THA11_HUMAN.H11MO.0.B, RUNX3_HUMAN.H11MO.0.A, RUNX1_HUMAN.H11MO.0.A, ZIC3_HUMAN.H11MO.0.B, ZN143_HUMAN.H11MO.0.A, GLI1_HUMAN.H11MO.0.D, ZNF76_HUMAN.H11MO.0.C, HAND1_HUMAN.H11MO.1.D, HTF4_HUMAN.H11MO.0.A
RXRB_HUMAN.H11MO.0.C	1.90E-07	97	878		CACATGBA	67	SNAI1_HUMAN.H11MO.0.C, BMAL1_HUMAN.H11MO.0.A, TWST1_HUMAN.H11MO.1.A, MLX_HUMAN.H11MO.0.D, SNAI2_HUMAN.H11MO.0.A, BHE22_HUMAN.H11MO.0.D, MYC_HUMAN.H11MO.0.A, TFEB_HUMAN.H11MO.0.C, USF1_HUMAN.H11MO.0.A, FIGLA_HUMAN.H11MO.0.D, MITF_HUMAN.H11MO.0.A, HEY1_HUMAN.H11MO.0.D, MESP1_HUMAN.H11MO.0.D, HES5_HUMAN.H11MO.0.D, MXI1_HUMAN.H11MO.1.A, MAX_HUMAN.H11MO.0.A, ID4_HUMAN.H11MO.0.D, BHE40_HUMAN.H11MO.0.A, MYCN_HUMAN.H11MO.0.A, OLIG1_HUMAN.H11MO.0.D, CR3L1_HUMAN.H11MO.0.D, MXI1_HUMAN.H11MO.0.A, MYF6_HUMAN.H11MO.0.C, IRX3_HUMAN.H11MO.0.D, USF2_HUMAN.H11MO.0.A, OLIG3_HUMAN.H11MO.0.D, ZEB1_HUMAN.H11MO.0.A, TFE2_HUMAN.H11MO.0.A, NDF1_HUMAN.H11MO.0.A, TWST1_HUMAN.H11MO.0.A, PO2F2_HUMAN.H11MO.0.A, TFE3_HUMAN.H11MO.0.B, ATOH1_HUMAN.H11MO.0.B, PO2F1_HUMAN.H11MO.0.C, NGN2_HUMAN.H11MO.0.D, HES7_HUMAN.H11MO.0.D, HTF4_HUMAN.H11MO.0.A, MLXPL_HUMAN.H11MO.0.D, NDF2_HUMAN.H11MO.0.B, PO5F1_HUMAN.H11MO.1.A, ASCL2_HUMAN.H11MO.0.D, RUNX2_HUMAN.H11MO.0.A, HIF1A_HUMAN.H11MO.0.C, EPAS1_HUMAN.H11MO.0.B, CLOCK_HUMAN.H11MO.0.C, BHE41_HUMAN.H11MO.0.D, ZBT18_HUMAN.H11MO.0.C, ASCL1_HUMAN.H11MO.0.A, ITF2_HUMAN.H11MO.0.C, PIT1_HUMAN.H11MO.0.C, P5F1B_HUMAN.H11MO.0.D, PO5F1_HUMAN.H11MO.0.A, MYOD1_HUMAN.H11MO.1.A, TAL1_HUMAN.H11MO.1.A, TBX5_HUMAN.H11MO.0.D, PO3F2_HUMAN.H11MO.0.A, NKX23_HUMAN.H11MO.0.D, MYOG_HUMAN.H11MO.0.B, RUNX1_HUMAN.H11MO.0.A, BHE23_HUMAN.H11MO.0.D, SCRT1_HUMAN.H11MO.0.D, TBR1_HUMAN.H11MO.0.D, RUNX3_HUMAN.H11MO.0.A, EOMES_HUMAN.H11MO.0.D, ARNT_HUMAN.H11MO.0.B, ZN816_HUMAN.H11MO.0.C, SCRT2_HUMAN.H11MO.0.D
NR1H4_HUMAN.H11MO.1.B	9.50E-07	45	443		CACGB	36	ARNT_HUMAN.H11MO.0.B, TFEB_HUMAN.H11MO.0.C, EPAS1_HUMAN.H11MO.0.B, MLX_HUMAN.H11MO.0.D, CR3L1_HUMAN.H11MO.0.D, HIF1A_HUMAN.H11MO.0.C, HES5_HUMAN.H11MO.0.D, HES7_HUMAN.H11MO.0.D, CR3L2_HUMAN.H11MO.0.D, AHR_HUMAN.H11MO.0.B, ATF6A_HUMAN.H11MO.0.B, HEY1_HUMAN.H11MO.0.D, MAX_HUMAN.H11MO.0.A, ARNT2_HUMAN.H11MO.0.D, CREB3_HUMAN.H11MO.0.D, BHE40_HUMAN.H11MO.0.A, TFE3_HUMAN.H11MO.0.B, USF1_HUMAN.H11MO.0.A, MITF_HUMAN.H11MO.0.A, MTF1_HUMAN.H11MO.0.C, MYCN_HUMAN.H11MO.0.A, BMAL1_HUMAN.H11MO.0.A, USF2_HUMAN.H11MO.0.A, MXI1_HUMAN.H11MO.0.A, MYC_HUMAN.H11MO.0.A, HEY2_HUMAN.H11MO.0.D, MXI1_HUMAN.H11MO.1.A, KLF14_HUMAN.H11MO.0.D, XBP1_HUMAN.H11MO.0.D, MLXPL_HUMAN.H11MO.0.D,

							PAX1_HUMAN.H11MO.0.D, HES1_HUMAN.H11MO.0.D, BHE41_HUMAN.H11MO.0.D, GMEB2_HUMAN.H11MO.0.D, HIC2_HUMAN.H11MO.0.D, ATF3_HUMAN.H11MO.0.A
NR4A3_HUMAN.H11MO.0.D	4.00E-06	55	313		CAGARA	3	ZIM3_HUMAN.H11MO.0.C, UBIP1_HUMAN.H11MO.0.D, ZN768_HUMAN.H11MO.0.C
GCR_HUMAN.H11MO.0.A	1.00E-05	84	689		CAGSSTGG	1	ZN250_HUMAN.H11MO.0.C
NR1I2_HUMAN.H11MO.1.D	1.20E-05	97	856		CTTCCWG	59	ZN436_HUMAN.H11MO.0.C, ETV1_HUMAN.H11MO.0.A, ETV7_HUMAN.H11MO.0.D, ELK1_HUMAN.H11MO.0.B, ELK3_HUMAN.H11MO.0.D, ELF1_HUMAN.H11MO.0.A, FLI1_HUMAN.H11MO.1.A, EHF_HUMAN.H11MO.0.B, ETS1_HUMAN.H11MO.0.A, ELF2_HUMAN.H11MO.0.C, ERG_HUMAN.H11MO.0.A, ELK4_HUMAN.H11MO.0.A, ELF5_HUMAN.H11MO.0.A, ETV2_HUMAN.H11MO.0.B, GABPA_HUMAN.H11MO.0.A, ETV6_HUMAN.H11MO.0.D, FLI1_HUMAN.H11MO.0.A, ETV4_HUMAN.H11MO.0.B, ELF3_HUMAN.H11MO.0.A, TEAD2_HUMAN.H11MO.0.D, HSF1_HUMAN.H11MO.1.A, ZN680_HUMAN.H11MO.0.C, SPI1_HUMAN.H11MO.0.A, STAT6_HUMAN.H11MO.0.B, NFAT5_HUMAN.H11MO.0.D, SMCA1_HUMAN.H11MO.0.C, ETV5_HUMAN.H11MO.0.C, SPIB_HUMAN.H11MO.0.A, STAT1_HUMAN.H11MO.0.A, TEAD1_HUMAN.H11MO.0.A, ZN394_HUMAN.H11MO.1.D, ETS2_HUMAN.H11MO.0.B, STAT3_HUMAN.H11MO.0.A, TEAD4_HUMAN.H11MO.0.A, FEV_HUMAN.H11MO.0.B, SPDEF_HUMAN.H11MO.0.D, NFAC2_HUMAN.H11MO.0.B, ZN528_HUMAN.H11MO.0.C, NFAC1_HUMAN.H11MO.1.B, E2F3_HUMAN.H11MO.0.A, KAISO_HUMAN.H11MO.2.A, STAT4_HUMAN.H11MO.0.A, STA5A_HUMAN.H11MO.0.A, NFAC4_HUMAN.H11MO.0.C, HSF4_HUMAN.H11MO.0.D, NFAC1_HUMAN.H11MO.0.B, WT1_HUMAN.H11MO.1.B, HXB2_HUMAN.H11MO.0.D, NFAC3_HUMAN.H11MO.0.B, IRF4_HUMAN.H11MO.0.A, IRF8_HUMAN.H11MO.0.B, E2F1_HUMAN.H11MO.0.A, VEZF1_HUMAN.H11MO.1.C, TF65_HUMAN.H11MO.0.A, HSF1_HUMAN.H11MO.0.A, TFDP1_HUMAN.H11MO.0.C, ZF64A_HUMAN.H11MO.0.D, ZN341_HUMAN.H11MO.1.C, BC11A_HUMAN.H11MO.0.A
RARG_HUMAN.H11MO.1.B	2.30E-05	71	618		CWKCCTC	15	ZN770_HUMAN.H11MO.1.C, ZN770_HUMAN.H11MO.0.C, WT1_HUMAN.H11MO.1.B, ZN436_HUMAN.H11MO.0.C, SPI1_HUMAN.H11MO.0.A, EGR4_HUMAN.H11MO.0.D, ZN263_HUMAN.H11MO.1.A, SALL4_HUMAN.H11MO.0.B, SPIB_HUMAN.H11MO.0.A, ZN263_HUMAN.H11MO.0.A, ETV7_HUMAN.H11MO.0.D, ZSC22_HUMAN.H11MO.0.C, WT1_HUMAN.H11MO.0.C, ETV6_HUMAN.H11MO.0.D, ERG_HUMAN.H11MO.0.A
ERR3_HUMAN.H11MO.0.B	6.70E-05	66	501		GAGATGGR	4	NDF1_HUMAN.H11MO.0.A, ATOH1_HUMAN.H11MO.0.B, NDF2_HUMAN.H11MO.0.B, NGN2_HUMAN.H11MO.0.D
ANDR_HUMAN.H11MO.1.A	1.20E-04	91	745		GGSAGR	4	WT1_HUMAN.H11MO.1.B, PATZ1_HUMAN.H11MO.1.C, ZN263_HUMAN.H11MO.1.A, IKZF1_HUMAN.H11MO.0.C
NR1H3_HUMAN.H11MO.1.B	1.30E-04	89	802		GRGAAR	23	NFKB1_HUMAN.H11MO.0.A, STAT4_HUMAN.H11MO.0.A, SUH_HUMAN.H11MO.0.A, ZEP1_HUMAN.H11MO.0.D, STA5A_HUMAN.H11MO.0.A, STA5B_HUMAN.H11MO.0.A, PRDM1_HUMAN.H11MO.0.A, ETS2_HUMAN.H11MO.0.B, NFAT5_HUMAN.H11MO.0.D, PRDM6_HUMAN.H11MO.0.C, STAT3_HUMAN.H11MO.0.A, BRAC_HUMAN.H11MO.1.B, E2F3_HUMAN.H11MO.0.A, E2F1_HUMAN.H11MO.0.A, STAT1_HUMAN.H11MO.0.A, HSF1_HUMAN.H11MO.0.A, E2F4_HUMAN.H11MO.1.A, E2F2_HUMAN.H11MO.0.B, TBR1_HUMAN.H11MO.0.D, TFDP1_HUMAN.H11MO.0.C, SMCA1_HUMAN.H11MO.0.C, E2F6_HUMAN.H11MO.0.A, IKZF1_HUMAN.H11MO.0.C
NR2C1_HUMAN.H11MO.0.C	1.90E-04	66	378		RAAGRAAA	2	IRF5_HUMAN.H11MO.0.D, PRDM6_HUMAN.H11MO.0.C
ERR2_HUMAN.H11MO.0.A	4.50E-04	62	470		RKAAATA	18	MEF2D_HUMAN.H11MO.0.A, MEF2C_HUMAN.H11MO.0.A, FOXJ2_HUMAN.H11MO.0.C, MEF2A_HUMAN.H11MO.0.A, FOXL1_HUMAN.H11MO.0.D, FOXF1_HUMAN.H11MO.0.D, MEF2B_HUMAN.H11MO.0.A, FOXP1_HUMAN.H11MO.0.A, DLX2_HUMAN.H11MO.0.D, ZFP28_HUMAN.H11MO.0.C, CPEB1_HUMAN.H11MO.0.D, FOXO4_HUMAN.H11MO.0.C, HMGA1_HUMAN.H11MO.0.D, FOXJ3_HUMAN.H11MO.1.B, FOXQ1_HUMAN.H11MO.0.C, FOXG1_HUMAN.H11MO.0.D, FOXJ3_HUMAN.H11MO.0.A, HXA9_HUMAN.H11MO.0.B
RARA_HUMAN.H11MO.1.A	5.90E-04	73	609		SAGGAAA	58	NFAC4_HUMAN.H11MO.0.C, ETV6_HUMAN.H11MO.0.D, ZN586_HUMAN.H11MO.0.C, ETV7_HUMAN.H11MO.0.D, ELF3_HUMAN.H11MO.0.A, ETV4_HUMAN.H11MO.0.B, NFAC2_HUMAN.H11MO.0.B, ERG_HUMAN.H11MO.0.A, NFAT5_HUMAN.H11MO.0.D, ETV2_HUMAN.H11MO.0.B, FLI1_HUMAN.H11MO.1.A, ETV5_HUMAN.H11MO.0.C, NFAC1_HUMAN.H11MO.1.B, EHF_HUMAN.H11MO.0.B, NFAC3_HUMAN.H11MO.0.B, ELF5_HUMAN.H11MO.0.A, PRDM6_HUMAN.H11MO.0.C, STAT6_HUMAN.H11MO.0.B, FLI1_HUMAN.H11MO.0.A, NFAC1_HUMAN.H11MO.0.B, ZN257_HUMAN.H11MO.0.C, SPI1_HUMAN.H11MO.0.A, ZFP28_HUMAN.H11MO.0.C, ETS1_HUMAN.H11MO.0.A, ETS2_HUMAN.H11MO.0.B, STAT1_HUMAN.H11MO.0.A, TF65_HUMAN.H11MO.0.A, SPIB_HUMAN.H11MO.0.A, STA5A_HUMAN.H11MO.0.A, STAT4_HUMAN.H11MO.0.A, IRF5_HUMAN.H11MO.0.D, BCL6_HUMAN.H11MO.0.A, BC11A_HUMAN.H11MO.0.A, ELK3_HUMAN.H11MO.0.D, ELF1_HUMAN.H11MO.0.A, ZN384_HUMAN.H11MO.0.C, ELF2_HUMAN.H11MO.0.C, NFKB2_HUMAN.H11MO.0.B, IRF4_HUMAN.H11MO.0.A, NFKB1_HUMAN.H11MO.0.A, ZN394_HUMAN.H11MO.1.D, GABPA_HUMAN.H11MO.0.A, DDIT3_HUMAN.H11MO.0.D, FOXA2_HUMAN.H11MO.0.A, IRF8_HUMAN.H11MO.0.B, ZN528_HUMAN.H11MO.0.C, ETV1_HUMAN.H11MO.0.A, THB_HUMAN.H11MO.0.C, E2F3_HUMAN.H11MO.0.A, ELK1_HUMAN.H11MO.0.B, ELK4_HUMAN.H11MO.0.A, FOXM1_HUMAN.H11MO.0.A, STAT3_HUMAN.H11MO.0.A, TEAD1_HUMAN.H11MO.0.A, OZF_HUMAN.H11MO.0.C, STA5B_HUMAN.H11MO.0.A, TEAD2_HUMAN.H11MO.0.D, TEAD4_HUMAN.H11MO.0.A
RXRA_HUMAN.H11MO.1.A	8.50E-04	95	845		TGACCTB	92	RARB_HUMAN.H11MO.0.D, RARA_HUMAN.H11MO.1.A, RARG_HUMAN.H11MO.1.B, NR2C1_HUMAN.H11MO.0.C, NR1H2_HUMAN.H11MO.0.D, NR1H3_HUMAN.H11MO.1.B, NR4A3_HUMAN.H11MO.0.D, ESR1_HUMAN.H11MO.1.A, THA_HUMAN.H11MO.1.D, ERR1_HUMAN.H11MO.0.A,

							ERR2_HUMAN.H11MO.0.A, ESR2_HUMAN.H11MO.1.A, NR4A2_HUMAN .H11MO.0.C, RXRA_HUMAN.H11MO.1.A, COT2_HUMAN.H11MO.0.A, THB_HUMAN.H11MO.1.D, PPARA_HUMAN.H11MO.1.B, RXRB_HUMAN. H11MO.0.C, RXRG_HUMAN.H11MO.0.B, ERR3_HUMAN.H11MO.0.B, C OT1_HUMAN.H11MO.1.C, RARG_HUMAN.H11MO.2.D, NR1D1_HUMAN. H11MO.1.D, THB_HUMAN.H11MO.0.C, NR1I3_HUMAN.H11MO.1.D, N R1H4_HUMAN.H11MO.1.B, NR6A1_HUMAN.H11MO.0.B, PPARD_HUMA N.H11MO.0.D, COT2_HUMAN.H11MO.1.A, NR1I2_HUMAN.H11MO.1.D, COT1_HUMAN.H11MO.0.C, RARA_HUMAN.H11MO.0.A, PPARG_HUMA N.H11MO.1.A, STF1_HUMAN.H11MO.0.B, NR5A2_HUMAN.H11MO.0.B, RORA_HUMAN.H11MO.0.C, THA_HUMAN.H11MO.0.C, ATF2_HUMAN.H 11MO.2.C, CREB3_HUMAN.H11MO.0.D, RXRA_HUMAN.H11MO.0.A, A TF7_HUMAN.H11MO.0.D, NR2E1_HUMAN.H11MO.0.D, CREB5_HUMAN. H11MO.0.D, CREM_HUMAN.H11MO.0.C, JDP2_HUMAN.H11MO.0.D, E SR1_HUMAN.H11MO.0.A, VDR_HUMAN.H11MO.1.A, NR1I2_HUMAN.H1 1MO.0.C, NR4A1_HUMAN.H11MO.0.A, ATF6A_HUMAN.H11MO.0.B, N R2C2_HUMAN.H11MO.0.B, NR1I3_HUMAN.H11MO.0.C, RARG_HUMAN. H11MO.0.B, NR1H3_HUMAN.H11MO.0.B, PPARG_HUMAN.H11MO.0.A, ATF2_HUMAN.H11MO.1.B, NR2E3_HUMAN.H11MO.0.C, RARA_HUMAN .H11MO.2.A, USF1_HUMAN.H11MO.0.A, CREB1_HUMAN.H11MO.0.A, TFEB_HUMAN.H11MO.0.C, PPARA_HUMAN.H11MO.0.B, USF2_HUMAN. H11MO.0.A, ATF2_HUMAN.H11MO.0.B, E4F1_HUMAN.H11MO.0.D, M ITF_HUMAN.H11MO.0.A, TFE3_HUMAN.H11MO.0.B, RORG_HUMAN.H1 1MO.0.C, ATF1_HUMAN.H11MO.0.B, VDR_HUMAN.H11MO.0.A, ZEB1 _HUMAN.H11MO.0.A, ESR2_HUMAN.H11MO.0.A, ZFX_HUMAN.H11MO. 0.A, CR3L2_HUMAN.H11MO.0.D, NR1H4_HUMAN.H11MO.0.B, MEIS3 _HUMAN.H11MO.0.D, NRL_HUMAN.H11MO.0.D, NF2L1_HUMAN.H11M O.0.C, HNF4G_HUMAN.H11MO.0.B, ZNF18_HUMAN.H11MO.0.C, TF7L 2_HUMAN.H11MO.0.A, HNF4A_HUMAN.H11MO.0.A, TF7L1_HUMAN.H1 1MO.0.B, PRRX1_HUMAN.H11MO.0.D, MEIS1_HUMAN.H11MO.1.B, M EIS2_HUMAN.H11MO.0.B, SNAI1_HUMAN.H11MO.0.C, TGIF2_HUMAN. H11MO.0.D, ZN449_HUMAN.H11MO.0.C, SCRT2_HUMAN.H11MO.0.D, PO4F2_HUMAN.H11MO.0.D, HEY2_HUMAN.H11MO.0.D
FOXA2_HUMAN.H 11MO.0.A	9.60E-04	145	1112				
NR4A2_HUMAN.H 11MO.0.C	6.30E-03	60	503				
VDR_HUMAN.H11 MO.1.A	1.20E-02	110	894				
RORG_HUMAN.H1 1MO.0.C	2.00E-02	85	723				
FOXA3_HUMAN.H 11MO.0.B	8.90E-02	186	1506				
COT2_HUMAN.H1 1MO.0.A	1.80E-01	88	722				
THB_HUMAN.H11 MO.1.D	1.80E-01	95	700				
BACH2_HUMAN.H 11MO.0.A	1.90E-01	30	249				
FOXA1_HUMAN.H 11MO.0.A	1	131	1039				
ATF3_HUMAN.H11 MO.0.A	1.4	30	218				
NFE2_HUMAN.H11 MO.0.A	1.8	128	726				
COT1_HUMAN.H1 1MO.1.C	2	82	630				
NR1D1_HUMAN.H 11MO.1.D	2.6	23	122				
PPARA_HUMAN.H 11MO.1.B	3.9	94	814				
RORA_HUMAN.H1 1MO.0.C	5.6	42	388				
NF2L1_HUMAN.H1 1MO.0.C	6.2	124	1096				
PRGR_HUMAN.H1 1MO.0.A	8.2	181	1400				
FOXM1_HUMAN.H 11MO.0.A	8.4	183	1595				

ZR751shMLL3 ERα - Centrimo					ZR751shMLL3 ERα – DREME TOMTOM		
ID	***E*-value**	**Region Width**	**Region Matches**		**Query Motif**	**Matches**	**Top Target Motifs**
ESR1_HUMAN.H11MO.0.A	4.80E-31	68	342		AAATATTT	2	FOXD2_HUMAN.H11MO.0.D, FOXB1_HUMAN.H11MO.0.D
ESR2_HUMAN.H11MO.0.A	7.10E-28	58	315		ACATRTGT	13	OLIG1_HUMAN.H11MO.0.D, BHE22_HUMAN.H11MO.0.D, TGIF2_HUMAN.H11MO.0.D, OLIG3_HUMAN.H11MO.0.D, TWST1_HUMAN.H11MO.1.A, MEIS2_HUMAN.H11MO.0.B, IRX3_HUMAN.H11MO.0.D, TF2LX_HUMAN.H11MO.0.D, NDF1_HUMAN.H11MO.0.A, TWST1_HUMAN.H11MO.0.A, SCRT2_HUMAN.H11MO.0.D, SCRT1_HUMAN.H11MO.0.D, FIGLA_HUMAN.H11MO.0.D
RARG_HUMAN.H11MO.0.B	6.50E-20	71	377		AGGAAR	57	ETV2_HUMAN.H11MO.0.B, ELF3_HUMAN.H11MO.0.A, ETV6_HUMAN.H11MO.0.D, ETV4_HUMAN.H11MO.0.B, ETV7_HUMAN.H11MO.0.D, FLI1_HUMAN.H11MO.1.A, ERG_HUMAN.H11MO.0.A, EHF_HUMAN.H11MO.0.B, ETV5_HUMAN.H11MO.0.C, FLI1_HUMAN.H11MO.0.A, NFAT5_HUMAN.H11MO.0.D, ELF5_HUMAN.H11MO.0.A, SPI1_HUMAN.H11MO.0.A, ELK3_HUMAN.H11MO.0.D, SPIB_HUMAN.H11MO.0.A, ETS1_HUMAN.H11MO.0.A, ETS2_HUMAN.H11MO.0.B, NFAC4_HUMAN.H11MO.0.C, ZN586_HUMAN.H11MO.0.C, STAT6_HUMAN.H11MO.0.B, STA5A_HUMAN.H11MO.0.A, NFAC2_HUMAN.H11MO.0.B, STAT4_HUMAN.H11MO.0.A, STAT1_HUMAN.H11MO.0.A, ZNF41_HUMAN.H11MO.1.C, BCL6_HUMAN.H11MO.0.A, ELF1_HUMAN.H11MO.0.A, ZN436_HUMAN.H11MO.0.C, BC11A_HUMAN.H11MO.0.A, NFAC1_HUMAN.H11MO.1.B, ETV1_HUMAN.H11MO.0.A, ELK4_HUMAN.H11MO.0.A, ELF2_HUMAN.H11MO.0.C, ELK1_HUMAN.H11MO.0.B, ZN394_HUMAN.H11MO.1.D, ZN680_HUMAN.H11MO.0.C, IRF8_HUMAN.H11MO.0.B, NFAC1_HUMAN.H11MO.0.B, IRF4_HUMAN.H11MO.0.A, NFAC3_HUMAN.H11MO.0.B, GABPA_HUMAN.H11MO.0.A, IRF5_HUMAN.H11MO.0.D, STAT3_HUMAN.H11MO.0.A, PRDM6_HUMAN.H11MO.0.C, SMCA1_HUMAN.H11MO.0.C, NFKB1_HUMAN.H11MO.0.A, STA5B_HUMAN.H11MO.0.A, ZNF41_HUMAN.H11MO.0.C, ZN263_HUMAN.H11MO.1.A, TF65_HUMAN.H11MO.0.A, E2F3_HUMAN.H11MO.0.A, ZFP28_HUMAN.H11MO.0.C, FEV_HUMAN.H11MO.0.B, HSF1_HUMAN.H11MO.1.A, ZN335_HUMAN.H11MO.1.A, TEAD2_HUMAN.H11MO.0.D, ZN263_HUMAN.H11MO.0.A
ESR1_HUMAN.H11MO.1.A	1.20E-12	67	318		AMACAGAR	4	ZIM3_HUMAN.H11MO.0.C, ZN250_HUMAN.H11MO.0.C, ZFP28_HUMAN.H11MO.0.C, FOXP1_HUMAN.H11MO.0.A
NR1H4_HUMAN.H11MO.1.B	1.20E-12	51	280		BCCAGG	18	SNAI1_HUMAN.H11MO.0.C, HAND1_HUMAN.H11MO.1.D, ZEB1_HUMAN.H11MO.0.A, ZN436_HUMAN.H11MO.0.C, SNAI2_HUMAN.H11MO.0.A, FIGLA_HUMAN.H11MO.0.D, ZN680_HUMAN.H11MO.0.C, ZN335_HUMAN.H11MO.1.A, ZF64A_HUMAN.H11MO.0.D, NFIB_HUMAN.H11MO.0.D, ERR2_HUMAN.H11MO.0.A, ZBT18_HUMAN.H11MO.0.C, ID4_HUMAN.H11MO.0.D, ZFX_HUMAN.H11MO.0.A, NR5A2_HUMAN.H11MO.0.B, STF1_HUMAN.H11MO.0.B, MESP1_HUMAN.H11MO.0.D, HAND1_HUMAN.H11MO.0.D
ESR2_HUMAN.H11MO.1.A	4.80E-11	68	328		CACACAYM	5	ZSCA4_HUMAN.H11MO.0.D, GLI2_HUMAN.H11MO.0.D, RUNX2_HUMAN.H11MO.0.A, GLI1_HUMAN.H11MO.0.D, GLI3_HUMAN.H11MO.0.B
ERR1_HUMAN.H11MO.0.A	1.30E-10	58	246		CACGY	39	ARNT_HUMAN.H11MO.0.B, TFEB_HUMAN.H11MO.0.C, EPAS1_HUMAN.H11MO.0.B, MLX_HUMAN.H11MO.0.D, CR3L1_HUMAN.H11MO.0.D, HIF1A_HUMAN.H11MO.0.C, HES5_HUMAN.H11MO.0.D, HES7_HUMAN.H11MO.0.D, CR3L2_HUMAN.H11MO.0.D, ATF6A_HUMAN.H11MO.0.B, HEY1_HUMAN.H11MO.0.D, MAX_HUMAN.H11MO.0.A, CREB3_HUMAN.H11MO.0.D, AHR_HUMAN.H11MO.0.B, BHE40_HUMAN.H11MO.0.A, USF1_HUMAN.H11MO.0.A, TFE3_HUMAN.H11MO.0.B, MITF_HUMAN.H11MO.0.A, MYCN_HUMAN.H11MO.0.A, BMAL1_HUMAN.H11MO.0.A, USF2_HUMAN.H11MO.0.A, MXI1_HUMAN.H11MO.0.A, MYC_HUMAN.H11MO.0.A, MXI1_HUMAN.H11MO.1.A, XBP1_HUMAN.H11MO.0.D, HEY2_HUMAN.H11MO.0.D, BHE41_HUMAN.H11MO.0.D, GMEB2_HUMAN.H11MO.0.D, ARNT2_HUMAN.H11MO.0.D, KLF14_HUMAN.H11MO.0.D, ATF3_HUMAN.H11MO.0.A, PAX1_HUMAN.H11MO.0.D, MTF1_HUMAN.H11MO.0.C, HES1_HUMAN.H11MO.0.D, TBX5_HUMAN.H11MO.0.D, MLXPL_HUMAN.H11MO.0.D, KLF9_HUMAN.H11MO.0.C, SRBP1_HUMAN.H11MO.0.A, E4F1_HUMAN.H11MO.0.D
RARG_HUMAN.H11MO.1.B	5.30E-09	45	235		CAGADA	3	UBIP1_HUMAN.H11MO.0.D, ZN768_HUMAN.H11MO.0.C, ZIM3_HUMAN.H11MO.0.C
RARB_HUMAN.H11MO.0.D	4.30E-08	80	275		CCCAGSM	13	ZN143_HUMAN.H11MO.0.A, THA11_HUMAN.H11MO.0.B, ZNF76_HUMAN.H11MO.0.C, ZN121_HUMAN.H11MO.0.C, IKZF1_HUMAN.H11MO.0.C, HAND1_HUMAN.H11MO.1.D, GLI1_HUMAN.H11MO.0.D, ZF64A_HUMAN.H11MO.0.D, SUH_HUMAN.H11MO.0.A, ZN331_HUMAN.H11MO.0.C, GLI3_HUMAN.H11MO.0.B, ZN449_HUMAN.H11MO.0.C, GLI2_HUMAN.H11MO.0.D
RORA_HUMAN.H11MO.0.C	4.00E-07	36	203		CCTSCCWC	103	WT1_HUMAN.H11MO.1.B, EGR4_HUMAN.H11MO.0.D, MAZ_HUMAN.H11MO.1.A, ZN263_HUMAN.H11MO.1.A, SALL4_HUMAN.H11MO.0.B, WT1_HUMAN.H11MO.0.C, ZN257_HUMAN.H11MO.0.C, ZSC22_HUMAN.H11MO.0.C, ZN281_HUMAN.H11MO.0.A, PATZ1_HUMAN.H11MO.1.C, FLI1_HUMAN.H11MO.0.A, ZN784_HUMAN.H11MO.0.D, SMAD2_HUMAN.H11MO.0.A, ARNT2_HUMAN.H11MO.0.D, GLIS1_HUMAN.H11MO.0.D, VEZF1_HUMAN.H11MO.1.C, EGR2_HUMAN.H11MO.1.A, ZN219_HUMAN.H11MO.0.D, ZN148_HUMAN.H11MO.0.D, EGR2_HUMAN.H11MO.0.A, E2F6_HUMAN.H11MO.0.A, NR0B1_HUMAN.H11MO.0.D, ZN770_HUMAN.H11MO.1.C, SP2_HUMAN.H11MO.1.B, GLIS2_HUMAN.H11MO.0.D, EGR1_HUMAN.H11MO.0.A, GLIS3_HUMAN.H11MO.0.D, ERG_HUMAN.H11MO.0.A, ZN263_HUMAN.H11MO.0.A, GLI2_HUMAN.H11MO.0.D, TBX1_HUMAN.H11MO.0.D, FLI1_HUMAN.H11MO.1.A, MAZ_HUMAN.H11MO.0.A, TBX15_HUMAN.H11MO.0.D, VEZF1_HUMAN.H11MO.0.C, AP2D_HUMAN.H11MO.0.D, KLF5_HUMAN.H11MO.0.A, ETV7_HUMAN.H11MO.0.D, SP4_HUMAN.H11MO.1.A, IKZF1_HUMAN.H11MO.0.C, ZN467_HUMAN.H11MO.0.C, SMAD3_HUMAN.H11MO.0.B, ZBT17_HUMAN.H11

							MO.0.A, SMAD4_HUMAN.H11MO.0.B, ETV2_HUMAN.H11MO.0.B, ZN770_HUMAN.H11MO.0.C, SNAI2_HUMAN.H11MO.0.A, E2F7_HUMAN.H11MO.0.B, CR3L1_HUMAN.H11MO.0.D, ASCL1_HUMAN.H11MO.0.A, E2F3_HUMAN.H11MO.0.A, KLF13_HUMAN.H11MO.0.D, KLF1_HUMAN.H11MO.0.A, KLF16_HUMAN.H11MO.0.D, KLF3_HUMAN.H11MO.0.B, ELF5_HUMAN.H11MO.0.A, ETV4_HUMAN.H11MO.0.B, ETV5_HUMAN.H11MO.0.C, SP3_HUMAN.H11MO.0.B, ZN708_HUMAN.H11MO.0.C, ELF3_HUMAN.H11MO.0.A, HES5_HUMAN.H11MO.0.D, SPI1_HUMAN.H11MO.0.A, ETV6_HUMAN.H11MO.0.D, TFDP1_HUMAN.H11MO.0.C, GLI1_HUMAN.H11MO.0.D, ZFX_HUMAN.H11MO.1.A, KLF4_HUMAN.H11MO.0.A, ZNF41_HUMAN.H11MO.1.C, SP1_HUMAN.H11MO.1.A, AP2C_HUMAN.H11MO.0.A, P63_HUMAN.H11MO.1.A, NFAT5_HUMAN.H11MO.0.D, HES7_HUMAN.H11MO.0.D, THB_HUMAN.H11MO.1.D, SP2_HUMAN.H11MO.0.A, INSM1_HUMAN.H11MO.0.C, KLF15_HUMAN.H11MO.0.A, TGIF1_HUMAN.H11MO.0.A, KLF6_HUMAN.H11MO.0.A, ETS1_HUMAN.H11MO.0.A, KLF12_HUMAN.H11MO.0.C, E2F1_HUMAN.H11MO.0.A, ZN740_HUMAN.H11MO.0.D, ZEB1_HUMAN.H11MO.0.A, ZFX_HUMAN.H11MO.0.A, ZBTB4_HUMAN.H11MO.0.D, ZIC1_HUMAN.H11MO.0.B, SP1_HUMAN.H11MO.0.A, ZN449_HUMAN.H11MO.0.C, SPIB_HUMAN.H11MO.0.A, E2F4_HUMAN.H11MO.0.A, PPARD_HUMAN.H11MO.0.D, SUH_HUMAN.H11MO.0.A, AP2A_HUMAN.H11MO.0.A, MXI1_HUMAN.H11MO.0.A, ZN335_HUMAN.H11MO.1.A, SRBP2_HUMAN.H11MO.0.B, PRDM4_HUMAN.H11MO.0.D, TBX3_HUMAN.H11MO.0.C, ZN524_HUMAN.H11MO.0.D, EHF_HUMAN.H11MO.0.B, GLI3_HUMAN.H11MO.0.B
THA_HUMAN.H11MO.1.D	5.20E-07	113	472		CTCYKCCC	46	ZSC22_HUMAN.H11MO.0.C, MAZ_HUMAN.H11MO.1.A, ZN341_HUMAN.H11MO.0.C, KLF15_HUMAN.H11MO.0.A, ZN263_HUMAN.H11MO.1.A, ZN148_HUMAN.H11MO.0.D, PATZ1_HUMAN.H11MO.1.C, ZN263_HUMAN.H11MO.0.A, ZN281_HUMAN.H11MO.0.A, ZN341_HUMAN.H11MO.1.C, WT1_HUMAN.H11MO.1.B, WT1_HUMAN.H11MO.0.C, SP4_HUMAN.H11MO.1.A, ZN467_HUMAN.H11MO.0.C, TBX1_HUMAN.H11MO.0.D, KLF12_HUMAN.H11MO.0.C, EGR4_HUMAN.H11MO.0.D, PPARD_HUMAN.H11MO.0.D, E2F6_HUMAN.H11MO.0.A, ZN219_HUMAN.H11MO.0.D, NR2C1_HUMAN.H11MO.0.C, SP3_HUMAN.H11MO.0.B, ZN257_HUMAN.H11MO.0.C, PPARG_HUMAN.H11MO.0.A, ZBT17_HUMAN.H11MO.0.A, ZN436_HUMAN.H11MO.0.C, KLF6_HUMAN.H11MO.0.A, EGR1_HUMAN.H11MO.0.A, ZNF76_HUMAN.H11MO.0.C, E2F7_HUMAN.H11MO.0.B, ZN708_HUMAN.H11MO.1.D, VEZF1_HUMAN.H11MO.1.C, KLF3_HUMAN.H11MO.0.B, ZN143_HUMAN.H11MO.0.A, IKZF1_HUMAN.H11MO.0.C, SRBP2_HUMAN.H11MO.0.B, FEV_HUMAN.H11MO.0.B, PPARA_HUMAN.H11MO.0.B, E2F1_HUMAN.H11MO.0.A, EGR2_HUMAN.H11MO.1.A, EGR2_HUMAN.H11MO.0.A, SP1_HUMAN.H11MO.1.A, SP2_HUMAN.H11MO.1.B, HEN1_HUMAN.H11MO.0.C, NFKB1_HUMAN.H11MO.0.A, SP1_HUMAN.H11MO.0.A
VDR_HUMAN.H11MO.1.A	9.20E-07	94	404		DWAAAATA	12	MEF2C_HUMAN.H11MO.0.A, MEF2A_HUMAN.H11MO.0.A, MEF2D_HUMAN.H11MO.0.A, MEF2B_HUMAN.H11MO.0.A, BHE22_HUMAN.H11MO.0.D, ONEC3_HUMAN.H11MO.0.D, OLIG1_HUMAN.H11MO.0.D, HMGA1_HUMAN.H11MO.0.D, CPEB1_HUMAN.H11MO.0.D, FOXJ3_HUMAN.H11MO.0.A, SOX5_HUMAN.H11MO.0.C, FOXL1_HUMAN.H11MO.0.D
NR1H3_HUMAN.H11MO.1.B	9.60E-07	43	228		GDAAAYA	38	FOXA2_HUMAN.H11MO.0.A, FOXO3_HUMAN.H11MO.0.B, FOXA1_HUMAN.H11MO.0.A, FOXB1_HUMAN.H11MO.0.D, FOXL1_HUMAN.H11MO.0.D, FOXA3_HUMAN.H11MO.0.B, FOXK1_HUMAN.H11MO.0.A, FOXM1_HUMAN.H11MO.0.A, FOXP2_HUMAN.H11MO.0.C, FOXO6_HUMAN.H11MO.0.D, FOXF2_HUMAN.H11MO.0.D, FOXD1_HUMAN.H11MO.0.D, FOXO1_HUMAN.H11MO.0.A, FOXP1_HUMAN.H11MO.0.A, FOXO4_HUMAN.H11MO.0.C, FOXJ2_HUMAN.H11MO.0.C, NFAC2_HUMAN.H11MO.0.B, FOXJ3_HUMAN.H11MO.1.B, FOXC1_HUMAN.H11MO.0.C, NFAC4_HUMAN.H11MO.0.C, ANDR_HUMAN.H11MO.0.A, NFAC3_HUMAN.H11MO.0.B, ZN384_HUMAN.H11MO.0.C, FOXD2_HUMAN.H11MO.0.D, FOXC2_HUMAN.H11MO.0.D, PO2F1_HUMAN.H11MO.0.C, NFAC1_HUMAN.H11MO.1.B, FOXF1_HUMAN.H11MO.0.D, ZN586_HUMAN.H11MO.0.C, HMGA1_HUMAN.H11MO.0.D, FOXD3_HUMAN.H11MO.0.D, PRDM6_HUMAN.H11MO.0.C, OLIG1_HUMAN.H11MO.0.D, MEF2C_HUMAN.H11MO.0.A, MEF2B_HUMAN.H11MO.0.A, MEF2D_HUMAN.H11MO.0.A, BHE22_HUMAN.H11MO.0.D, MEF2A_HUMAN.H11MO.0.A
ERR2_HUMAN.H11MO.0.A	4.70E-06	60	247		GGGAAAAR	26	NFAC2_HUMAN.H11MO.0.B, NFAC4_HUMAN.H11MO.0.C, NFAC3_HUMAN.H11MO.0.B, ZN586_HUMAN.H11MO.0.C, NFAC1_HUMAN.H11MO.1.B, NFKB1_HUMAN.H11MO.0.A, NFAT5_HUMAN.H11MO.0.D, ZN384_HUMAN.H11MO.0.C, PRDM6_HUMAN.H11MO.0.C, ETS2_HUMAN.H11MO.0.B, NFAC1_HUMAN.H11MO.0.B, E2F6_HUMAN.H11MO.0.A, TFDP1_HUMAN.H11MO.0.C, SUH_HUMAN.H11MO.0.A, NFKB2_HUMAN.H11MO.0.B, STAT3_HUMAN.H11MO.0.A, STAT1_HUMAN.H11MO.0.A, E2F1_HUMAN.H11MO.0.A, IKZF1_HUMAN.H11MO.0.C, REL_HUMAN.H11MO.0.B, E2F7_HUMAN.H11MO.0.B, E2F5_HUMAN.H11MO.0.B, HXC13_HUMAN.H11MO.0.D, E2F3_HUMAN.H11MO.0.A, ETV7_HUMAN.H11MO.0.D, MZF1_HUMAN.H11MO.0.B
NR1I3_HUMAN.H11MO.1.D	1.40E-05	59	267		RAATAAA	63	FOXQ1_HUMAN.H11MO.0.C, CPEB1_HUMAN.H11MO.0.D, FOXF1_HUMAN.H11MO.0.D, HXC6_HUMAN.H11MO.0.D, FOXL1_HUMAN.H11MO.0.D, FOXJ2_HUMAN.H11MO.0.C, FOXJ3_HUMAN.H11MO.0.A, HXB13_HUMAN.H11MO.0.A, EVX2_HUMAN.H11MO.0.A, HXA10_HUMAN.H11MO.0.C, FOXC1_HUMAN.H11MO.0.C, MYNN_HUMAN.H11MO.0.D, HXD4_HUMAN.H11MO.0.D, HXD9_HUMAN.H11MO.0.D, FOXO4_HUMAN.H11MO.0.C, HXC9_HUMAN.H11MO.0.C, HXA9_HUMAN.H11MO.0.B, HXD12_HUMAN.H11MO.0.D, FOXA1_HUMAN.H11MO.0.A, SOX4_HUMAN.H11MO.0.B, HXB7_HUMAN.H11MO.0.C, HXB8_HUMAN.H11MO.0.C, LHX3_HUMAN.H11MO.0.C, HNF6_HUMAN.H11MO.0.B, HXA13_HUMAN.H11MO.0.C, SOX10_HUMAN.H11MO.1.A, HXD10_HUMAN.H11MO.0.D, EVX1_HUMAN.H11MO.0.D, SOX2_HUMAN.H11MO.0.A, FOXC2_HUMAN.H11MO.0.D, HXD11_HUMAN.H11MO.0.D, CDX2_HUMAN.H11MO.0.A, FOXA2_HUMAN.H11MO.0.A, MEF2D_HUMAN.H11MO.0.A, MEF2C_HUMAN.H11MO.0.A, HXA11_HUMAN.H11MO.0.D, SOX13_HUMAN.H11MO.0.D, SOX3_HUMAN.H11MO.0.B, FOXD2_HUMAN.H11MO.0.D, HXC12_HUMAN.H11MO.0.D, HMGA1_HUMAN.H11MO.0.D, HXC10_HUMAN.H11MO.0.

							D, MEF2A_HUMAN.H11MO.0.A, PO6F1_HUMAN.H11MO.0.D, CDX1_HUMAN.H11MO.0.C, FOXG1_HUMAN.H11MO.0.D, PRRX2_HUMAN.H11MO.0.C, ZFP28_HUMAN.H11MO.0.C, MEIS1_HUMAN.H11MO.0.A, ONEC3_HUMAN.H11MO.0.D, HXD13_HUMAN.H11MO.0.D, FOXD1_HUMAN.H11MO.0.D, PO4F1_HUMAN.H11MO.0.D, ZN384_HUMAN.H11MO.0.C, ZFHX3_HUMAN.H11MO.0.D, PO4F3_HUMAN.H11MO.0.D, FOXA3_HUMAN.H11MO.0.B, PRDM6_HUMAN.H11MO.0.C, FOXO6_HUMAN.H11MO.0.D, ONEC2_HUMAN.H11MO.0.D, SRY_HUMAN.H11MO.0.B, HXC13_HUMAN.H11MO.0.D, RX_HUMAN.H11MO.0.D
GCR_HUMAN.H11MO.0.A	1.60E-05	66	285		RGGTCAS	106	ESR2_HUMAN.H11MO.1.A, ESR1_HUMAN.H11MO.1.A, RARB_HUMAN.H11MO.0.D, RARG_HUMAN.H11MO.1.B, NR4A3_HUMAN.H11MO.0.D, NR1H4_HUMAN.H11MO.1.B, THA_HUMAN.H11MO.1.D, RARA_HUMAN.H11MO.1.A, COT1_HUMAN.H11MO.1.C, NR1D1_HUMAN.H11MO.1.D, COT2_HUMAN.H11MO.0.A, NR4A2_HUMAN.H11MO.0.C, RXRA_HUMAN.H11MO.1.A, RXRB_HUMAN.H11MO.0.C, NR2C1_HUMAN.H11MO.0.C, ERR1_HUMAN.H11MO.0.A, NR1H3_HUMAN.H11MO.1.B, NR1I3_HUMAN.H11MO.1.D, COT2_HUMAN.H11MO.1.A, PPARA_HUMAN.H11MO.1.B, COT1_HUMAN.H11MO.0.C, ERR2_HUMAN.H11MO.0.A, RXRG_HUMAN.H11MO.0.B, ESR1_HUMAN.H11MO.0.A, RARG_HUMAN.H11MO.0.B, RORA_HUMAN.H11MO.0.C, NR1H2_HUMAN.H11MO.0.D, VDR_HUMAN.H11MO.1.A, NR1I2_HUMAN.H11MO.1.D, RARG_HUMAN.H11MO.2.D, RXRA_HUMAN.H11MO.0.A, RORG_HUMAN.H11MO.0.C, THB_HUMAN.H11MO.1.D, ERR3_HUMAN.H11MO.0.B, RARA_HUMAN.H11MO.2.A, NR1I3_HUMAN.H11MO.0.C, ESR2_HUMAN.H11MO.0.A, RARA_HUMAN.H11MO.0.A, USF2_HUMAN.H11MO.0.A, CREM_HUMAN.H11MO.0.C, VDR_HUMAN.H11MO.0.A, THB_HUMAN.H11MO.0.C, PPARD_HUMAN.H11MO.0.D, ATF6A_HUMAN.H11MO.0.B, USF1_HUMAN.H11MO.0.A, PPARG_HUMAN.H11MO.1.A, ATF2_HUMAN.H11MO.2.C, STF1_HUMAN.H11MO.0.B, ATF1_HUMAN.H11MO.0.B, NR5A2_HUMAN.H11MO.0.B, NR6A1_HUMAN.H11MO.0.B, TFE3_HUMAN.H11MO.0.B, CREB1_HUMAN.H11MO.0.A, NR2C2_HUMAN.H11MO.0.B, TFEB_HUMAN.H11MO.0.C, MITF_HUMAN.H11MO.0.A, NR4A1_HUMAN.H11MO.0.A, THA_HUMAN.H11MO.0.C, NR1I2_HUMAN.H11MO.0.C, NR1H3_HUMAN.H11MO.0.B, ATF7_HUMAN.H11MO.0.D, PPARG_HUMAN.H11MO.0.A, CREB5_HUMAN.H11MO.0.D, CREB3_HUMAN.H11MO.0.D, NR1H4_HUMAN.H11MO.0.B, ATF2_HUMAN.H11MO.0.B, NRL_HUMAN.H11MO.0.D, PPARA_HUMAN.H11MO.0.B, NR1D1_HUMAN.H11MO.0.B, FOS_HUMAN.H11MO.0.A, JDP2_HUMAN.H11MO.0.D, JUNB_HUMAN.H11MO.0.A, ZSC31_HUMAN.H11MO.0.C, ZBT7A_HUMAN.H11MO.0.A, NR2F6_HUMAN.H11MO.0.D, GLI2_HUMAN.H11MO.0.D, PRGR_HUMAN.H11MO.1.A, E4F1_HUMAN.H11MO.0.D, MAF_HUMAN.H11MO.1.B, PRRX1_HUMAN.H11MO.0.D, MEIS1_HUMAN.H11MO.1.B, FOSL1_HUMAN.H11MO.0.A, ZEB1_HUMAN.H11MO.0.A, NR2E1_HUMAN.H11MO.0.D, NF2L1_HUMAN.H11MO.0.C, PAX1_HUMAN.H11MO.0.D, JUND_HUMAN.H11MO.0.A, ANDR_HUMAN.H11MO.2.A, GCR_HUMAN.H11MO.1.A, BACH2_HUMAN.H11MO.0.A, JUN_HUMAN.H11MO.0.A, ATF3_HUMAN.H11MO.0.A, SRBP1_HUMAN.H11MO.0.A, MEIS3_HUMAN.H11MO.0.D, TGIF2_HUMAN.H11MO.0.D, ZNF18_HUMAN.H11MO.0.C, FOSL2_HUMAN.H11MO.0.A, CR3L2_HUMAN.H11MO.0.D, FOSB_HUMAN.H11MO.0.A, XBP1_HUMAN.H11MO.0.D, ATF2_HUMAN.H11MO.1.B, MLX_HUMAN.H11MO.0.D, ZN784_HUMAN.H11MO.0.D, MAFK_HUMAN.H11MO.1.A, NFE2_HUMAN.H11MO.0.A, GLIS1_HUMAN.H11MO.0.D
RARA_HUMAN.H11MO.1.A	2.40E-05	47	224		RRAGAAA	36	PRDM6_HUMAN.H11MO.0.C, FOXF1_HUMAN.H11MO.0.D, FOXJ2_HUMAN.H11MO.0.C, FOXQ1_HUMAN.H11MO.0.C, ZFP28_HUMAN.H11MO.0.C, CPEB1_HUMAN.H11MO.0.D, SOX4_HUMAN.H11MO.0.B, STAT2_HUMAN.H11MO.0.A, IRF2_HUMAN.H11MO.0.A, IRF1_HUMAN.H11MO.0.A, ZIM3_HUMAN.H11MO.0.C, SOX10_HUMAN.H11MO.1.A, ZN274_HUMAN.H11MO.0.A, STAT1_HUMAN.H11MO.1.A, SOX2_HUMAN.H11MO.0.A, GATA3_HUMAN.H11MO.0.A, PRDM1_HUMAN.H11MO.0.A, NFAC4_HUMAN.H11MO.0.C, SOX3_HUMAN.H11MO.0.B, STAT4_HUMAN.H11MO.0.A, FOXL1_HUMAN.H11MO.0.D, NFKB1_HUMAN.H11MO.0.A, IRF9_HUMAN.H11MO.0.C, STA5A_HUMAN.H11MO.0.A, BATF3_HUMAN.H11MO.0.B, ETS2_HUMAN.H11MO.0.B, FLI1_HUMAN.H11MO.1.A, NFAC1_HUMAN.H11MO.1.B, ERG_HUMAN.H11MO.0.A, NFAT5_HUMAN.H11MO.0.D, ONEC3_HUMAN.H11MO.0.D, ETV6_HUMAN.H11MO.0.D, FLI1_HUMAN.H11MO.0.A, ETV4_HUMAN.H11MO.0.B, ZN384_HUMAN.H11MO.0.C, GATA4_HUMAN.H11MO.0.A
NR4A2_HUMAN.H11MO.0.C	2.90E-05	58	257		TATDTATR	3	MEF2D_HUMAN.H11MO.0.A, MEF2C_HUMAN.H11MO.0.A, MEF2A_HUMAN.H11MO.0.A
NR1D1_HUMAN.H11MO.1.D	6.70E-05	41	116				
RXRB_HUMAN.H11MO.0.C	9.60E-05	119	501				
RORG_HUMAN.H11MO.0.C	1.10E-04	73	325				
NR2C1_HUMAN.H11MO.0.C	2.30E-04	72	214				
NR1I2_HUMAN.H11MO.1.D	8.60E-04	27	144				
NR4A3_HUMAN.H11MO.0.D	1.30E-03	51	157				
RXRA_HUMAN.H11MO.1.A	6.10E-03	49	233				
PPARA_HUMAN.H11MO.1.B	6.30E-03	48	231				
ERR3_HUMAN.H11MO.0.B	1.90E-02	54	210				

ANDR_HUMAN.H11MO.1.A	2.70E-02	61	255	
RARA_HUMAN.H11MO.2.A	5.80E-02	60	243	
NR5A2_HUMAN.H11MO.0.B	2.40E-01	60	213	
STF1_HUMAN.H11MO.0.B	3.70E-01	44	192	
NR1D1_HUMAN.H11MO.0.B	3.90E-01	36	174	
COT2_HUMAN.H11MO.0.A	9.10E-01	110	425	
NR2C2_HUMAN.H11MO.0.B	1.1	55	207	
COT1_HUMAN.H11MO.0.C	3.8	54	213	
THB_HUMAN.H11MO.1.D	9.7	21	93	
RFX5_HUMAN.H11MO.1.A	9.7	54	202	

RNA-seq Overlap Downregulated Genes in iRegulon						
Rank	Motif id	AUC	NES	ClusterCode	Transcription factor	Target genes
1	element o-AAAATGGCG	0.0504026	4.66433	M1	YY1	TRA2B,PPP2CA,GABPB1,EIF4G2,PPIA,CLK4,SRSF3,CC2D2A,PTBP2,PDS5B,HNRNPA1,NAA15,TRIM37,FBXL3,RBM23,DEK,CDC14A,RBM5,SMC3,CSE1L,CDK5RAP2,LIN54,ANP32B,SGK1,NPM1,TOP1,ZBED5,PPP6R3,HNRNPK,SOX4,HMGB1,RBMX,EIF1B,RAB14,GDI2,MORF4L1,ABCE1,ZNF280D,NFYB,MMADHC,ARID5B,MARCKS,SIAH1,SET,ATXN7,SRSF1,CCNYL1,SIX4,UQCRC2,PTMA,HNRNPC,BMP4,ANP32A,ZFR,KRAS
2	stark-RCGCMATTW	0.0447813	3.7955	M1		PPIA,ATF7IP,HNRNPK,CC2D2A,PPP2CA,EIF1,GABPB1,PDS5B,CLK4,FERMT2,HNRNPA3,SOX4,CDC14A,RAP1B,PIAS1,RAB21,ATF7,BAMBI,RBM5,SGK1,FBXL3,TOP1,CGGBP1,CSE1L,KANK1,EIF4A2,LSP1,SIX1,RBMX,KRAS,TOB1,OSR1,C11orf95,ZNF146,IQCJ,RCOR1,HNRNPC,CDCA7L,RHBDD1,ZFR,NFYB,CTNNA1,ATXN7,RCN1,VRK1,TMPO,C12orf65,MID1,TMTC2,PLAC9,PPP6R3,TANC1,PPM1D,CTBP2,ARID5B,PTMA,KLF5,TBL1X,CTNBP2NL,MARCKS,NRP2,SOX17,RBM23,PLD1,C3orf38,FGFR1,MTSS1,PPM1B,THNSL1,YME1L1,KCTD20,SIAH1,SNN,SPIN1,EPS8,BHLHE41,NPM1
3	flyfactorsurvey-rn_SOLEXA_5_FBgn0259172	0.0429416	3.51116	M2	ZNF362,ZNF384,PAX4,TCF3,POU5F1,BPTF,PGR,ZNF513,JAZF1,SRF,TBPL2,TBPL1,MEF2D,MEF2B,MEF2BNB-MEF2B,SNAPC4,MYBL2,MYBL1,MYB,ELF3	APOLD1,ZFP36L1,EIF4E,TGIF1,SOX4,STIM2,CTBP2,HMGB1,SP3,MSL1,SGK1,CSE1L,TMTC2,PLEKHA5,GCNT1,SIAH1,EDN1,LSP1,KLF5,MBNL1,CC2D2A,ZNF608,SRSF3,SPRY1,MBOAT1,ATF7,RAB2B,GPN3,PPM1B,CXCR4,TMCC3,MRPL48,TES,PTK2,FGF1,LSM6,SEC61B,GABPB1,TPD52L1,QKI,PLAGL1,RHBDD1,KIAA1217,BAMBI,MTSS1,PPP1R14A
4	hdpi-RBBP9	0.0427521	3.48186	M3	RBBP9	TGIF1,TCF12,QKI,MLLT10,LRIG3,RARB,TRA2B,ZNF608,ZSWIM6,PPP1R12A,KPNA4,CTBP2,EFNA5,KLF5,MEOX2,FRMD6,SKA2,AHR,EIF4G2,RAB14,TES,CGGBP1,TMTC2,PIAS1,SPRY1,EFNB2,TANC1,HNRNPC,CLTC,GABPB1,PPM1B,PKN2,BAMBI,HMGB1,PALMD,SSBP2,PAWR,ZFP36L1,SMNDC1,COL11A1,PIK3R1,ALDH1A1,KLHL13,PHLDA1,ARID5B,RAP1B,ATF7IP,STK38L,PDS5B,PDE7A,ZNF654,C11orf95,FRRS1,CD36,SRSF3,SOX4,AFF1,KANK1,SLTM,TOB1,C3orf58,CSDE1,RBM25,PALLD,GATA6,HNRNPA3,TOP1,MSL2,AXIN2,SP3,MID1,LSM6,ATF7,BTBD10,RFWD3,MIER3,ZNF280D,RHBDD1,FZD4,C12orf65,C11orf58,STIM2,TMCC3,ELL2,BHLHE41,CDK13,SIX1,RBMX,ATXN7,KCTD9,LIMS1,CDC14A,TTC39B,SGCD,SUPT16H,ZNF704,RYBP,UTP6,NCALD,EDN1,ATP2B1,KCMF1,MMADHC,PPP1CC,SIAH1,PPP2CA,G2E3,STK17B,UBE2E2,MPHOSPH9,APOLD1,MYBL1,CSNK1G3,KIAA1217,LIN7C,DCP1A,CLINT1,PTK2,ARRDC3,RCOR1,NRP2,C1GALT1,CDC42SE2,ANP32B,TSPAN13,FAM126A,RHOU,TPP2,VPS29,BMP4,MARCKS,GCNT1,MTSS1,FGFR1,ZEB1,SIX4,IRF2,TBL1X,TANK,FERMT2,CXCR4,ZNF423,IQCJ,EIF3J,SVIL,PTPN12,HNRNPM,TSPAN12,ETV6,POC1B,PLEKHA5,MAT2B,EPS8,TIAL1,SOSTDC1,TMOD3,DAPL1,PAIP2,CCDC88A,RBBP8,FGF13,MBNL1,FKTN,CNN1,PTGES3,C3orf38,EIF4A2,PLAGL1,ISM1,ARHGDIB,MORF4L1,GALNT3,PRPF40A,RNF125,FGF1,QRICH1,SRSF1,MSL1,THNSL1,ZBTB34,CORO1C,KLHL2,TTC8,NFYB,RHOA,PLD1,CAMSAP1,FBXO34,RAB33B,CDYL,MAPK1IP1L,TMPO,ZRANB2,KLF6,PPP4R2,APPBP2,FBXL3,ABI2,ANTXR2,LIFR,ACVR1,EMP1,OXCT1,GTF2A2,RNF38,STAM,LIMCH1,CCNYL1,SMARCA5,CLK4,IQGAP2,UBQLN1,THUMPD1,UGGT2,RAB21,ARFIP1,EIF4E,SOX17,H2AFV,LIN54,CCDC82,SHCBP1,DENR,TNPO1,UBA3,SYNE2,LRRC8D,NMD3,PPM1D,COPB1,SPIN1,KDELC2,TBCA,SSPN,CTSC,KRAS,HCCS,RASL11A,PTBP2,TSPAN5,FAM122A,DOCK11,SNRK,TRIM36,MTX3,XPOT,RPP30,STARD3NL,PDGFC,UACA,UBA2,KDM4C,CLDND1,GNA13,RIOK3,RAB23,VEGFC,GDI2,YME1L1,RND3,UHRF2,FAR1,ROPN1B,CETN3,PARG,DNAJC24,MBOAT1,GOLIM4,DAD1,KCTD20,SNRNP48,SNHG6,TPD52L1,PLRG1,CC2D2A,XRCC5,CBFB,IDH3A,ZDHHC20,ANKS1A,DEK,NUP98,NCBP2,VRK1,WDR61,RCAN1,POLR3B,NAA15,FRMD4A,PIGA,ARIH2
5	yetfasco-1614	0.0422709	3.40749	M4		TBCA,SRSF1,PPP6R3,NAA15,CDC14A,SIAH1,PPP2CA,ABCE1,CC2D2A,ZNF608,GABPB1,CTBP2,LSM6,OSR1,SIX1,AXIN2,LIN54,KANK1,CLTC,ZSWIM6,PDS5B,RNF128,TNPO1,FBXL3,BAMBI,MIER3,ACVR1,MID1,SLC20A1,DEK
6	taipale-NAACCGGTTN-TFCP2-full	0.0416001	3.30382	M4	TFCP2,GRHL1,TFCP2L1	GPCPD1,AFF1,RNF128,HMGN1,ARRDC3,UBE2E2,EIF4E,DSG3,MSL2,BLMH,RND3,LRCH1,LIFR,LIN54,HNRNPC,OSR1,PDGFC,DYRK2,AP3B1,LGALS3,CAMSAP1,DAPL1,TGIF1,SVIL,FKTN,UQCRC2,CTBP2,PPP1R12A,CCDC88A,TOB1,CGGBP1,TBCA,CD36,PHLDA1,ZNF608,PPM1B,CDYL,EDN1
7	taipale-NNCCGCCATNW-YY2-DBD	0.0411849	3.23965	M1	YY2,YY1	HNRNPK,SRSF3,HNRNPA1,EIF4G2,NAA15,ZBED5,LIN54,TRA2B,PPP2CA,PPP6R3,DEK,PHB2,TRIM37,HMGB1,PPM1B,HNRNPA3,ZNF280D,FGF1,KRAS,ABCE1,CSNK1G3,RPS29,UQCRC2,DDX46,PTBP2,MORF4L1,CLK4,CLTC
8	hdpi-ZRSR2	0.0411636	3.23635	M5	ZRSR2	QKI,TMTC2,CGGBP1,ATF7IP,ARID5B,SP3,SIAH1,IRF2,MLLT10,ELL2,ZNF423,MEOX2,FRMD6,TGIF1,GATA6,TCF12,RAB14,TRA2B,CD36,C3orf58,FKTN,PAWR,C1GALT1,HNRNPA3,ZSWIM6,PKN2,AFF1,ZNF608,TSPAN12,ALDH1A1,MID1,ZNF654,VEGFC,LSM6,PDE7A,GCNT1,GOLIM4,RAP1B,SKA2,NCALD,TIAL1,FBXO34,RARB,SRSF1,PDS5B,GABPB1,CSDE1,RYBP,ACVR1,C3orf38,CTNNA1,RHBDD1,NRP2,KCTD9,WDR41,KRAS,CBFB,POC1B,ARHGDIB,MYLK,CTBP2,EIF4E,TANK,ISM1,CTTNBP2NL,PPIA,GTF2A2,TBL1X,RNF125,MBNL1,MPHOSPH9,C12orf65,OXCT1,ATF7,TSPAN13,HMGB1,TIMM23,SPRY1,EMP1,SIX4,AHR,ETV6,RCAN1,SSBP2,PALMD,FRRS1,SRSF3,KIAA1217,CDYL,PSMG1,EPB41L4A,RPP30,UACA,MSL2,SLTM,FGF1,CHST3,UQCRC2,LAMA4,UGGT2,MTX3,FGF13,HNRNPC,CSNK1G3,PTBP2,ATP2B1,EEF1B2,PPP1R12A,COL11A1,TOB1,PLA2G7,UBQLN1,KLHL2,FAM122A,MIER3,HNRNPK,BHLHE41,CLK4,STK17B,NAF1,EFNA5,ZEB1,EFNB2,ORC2,ZFR,RND3,KCMF1,TOP1,PLAC9,CCNYL1,STK38L,MTSS1,TCEA1,OSR1,ZNF280D,UHRF2,ZNHIT6,PALLD,CDC14A,UTP6,TSPAN5,TNPO1,PPM1B,VRK1,MARCKS,RCN1,PPP1CC,DYRK2,CLINT1,ZFP36L1,SGK1,LRIG3,SMC3,NAA15,SVIL,NOC3L,MYBL1,ZBTB34,PLAGL1,PIAS1,NOD1,SUPT16H,EIF4A2,FZD4,RPL32,SNRK,NUP160,ATXN7,KDELC2,PLD1,APOLD1,FBXL3,AKAP2,FAR1,LIMCH1,TANC1,PCBD2,DPY19L1,TYRO3,BMP4,SPARCL1,G2E3,TMCC3,PPP2CA,LIFR,ZRANB1,SSPN,PAPSS1,KLHL13,M6PR,TES,FIBIN,SEC31A,LIMS1,STARD3NL,TMPO,UBE2E2,PIK3R1,LRRC8D,ZFP62,HSPA9,BAMBI,FBLN5,EIF4G2,ANP32B,SMNDC1,PTK2,CCDC88A,GPCPD1,WDR61,CEP120,DOCK11,CDC5L,OLFML1,SGCD,MBOAT1,BZW1,DSG3,IQCJ,ANXA3,KANK1,GALNT3,FAS,EPS8,ARFIP1,PIGY,BTBD10,ABI2,AP3B1,APPL1,TTC39B,DDHD2,FERMT2,MBD2,STIM2,ZNF41,QRICH1,RNF128,TMOD3,PARG,SMC2,RBM23,ANTXR2,FAM126A,RYK,CCDC82,SCFD2,G0S2,KLF6,C11orf95,FGFR1,PIGA,LOC154761,RHOU,HEATR3,DSC2,HCCS,ANP32A,ZNF704,SH3PXD2A,THNSL1,POLR3B,PAIP2,LARS,AXIN2,CORO1C,PHLDA1,SMARCA5,CLTC,RNF38,IARS,SNHG6,CXCL2,ZBTB11,GMNN

9	homer-M00682	0.0411189	3.22944	M3		ZNF608,ZFP36L1,SOX4,CTBP2,HMGB1,EFNB2,TANC1,ANP32B,MTSS1,EDN1,ARID5B,APOLD1,SKA2,QKI,ABI2,EIF4E,HNRNPC,MARCKS,TCF12,APPBP2,CSNK1G3,GATA6,SPIN1,STIM2,MAT2B,STK38L,KLF5,TGIF1,APPL1,LRIG3,RARB,GABPB1,PALLD,GCNT1,EFNA5,C12orf65,C1GALT1,ATF7,TES,LIN54,SPRY1,MLLT10,SRSF1,MORF4L1,COPB1,ANO1,CLK4,ROPN1B,DDX46,ARMC1,SP3,MDM2,NAA25,PLAGL1,RFWD3,ZFP62,EIF4A2,AFF1,RBMX,EIF4G2,TNPO1,C18orf25,ANP32A,SSBP2,ZSWIM6,FGF1,CBFB,PTMA,PTPN12,LRRC8D,TPP2,ZEB1,MSL1,EMP1,IDH3A,CDC42SE2,KIAA1217,KPNA4,RAN,TSPAN12,C3orf58,CD300LG,TBL1X,MPHOSPH9,RAB14,KANK1,MED7,CTNNA1,CSDE1,CCNYL1,PTK2,CHRAC1,RAP1B,TSPAN13,PKN2,ARHGDIB,GMNN,ZBTB34,ZNF146,SVIL,PDS5B,ARRDC3,DHX8,DDHD2,TTC39B,RBBP8,DCP1A,CLINT1,TFAM,XRCC5,FRMD6,SHCBP1,DSG3,PLD1,PTGES3,SIAH1,PLEKHA5,MMADHC,RBM23,FGFR1,ETV6,IPO5,UTP6,KDM4C,PIAS1,NFYB,TOB1,PPM1B,THNSL1,ACVR1,PDE7A,HNRNPM,HAS3,ATF7IP,ACTL6A,CDH5,TTC8,RPL24,CXCR4,NUP160,COQ5,TTC28,PALMD,BHLHE41,SRP54,LIMS1,SUPT16H,MRPL48,H2AFV,CAMSAP1,NAA15,FAR1,RPL23,VPS29,EIF1B,DENR,TRA2B,MSL2,C3orf38,ATXN7,QRICH1,PCNP,DNM1L,FBXO34,BMP4,SRSF3,SEC31A,MID1,SET,PPP4R2,MAPK1IP1L,FKTN
10	wolfe-rn_SOLEXA_F2-4	0.0409571	3.20443	M2	ZNF362,ZNF384,ZNF513	APOLD1,HMGB1,ZFP36L1,SOX4,SGK1,EIF4E,SP3,KLF6,GCNT1,STIM2,TOR1B,PLEKHA5,PTMA,CSE1L,CTBP2,HNRNPC,ZNF608,LSM6,CDYL,SPRY1,FGFR1,RBM5,PPP1R14A,QRICH1,RPL6,GPN3,RAB2B,EDN1,SIAH1,PALLD,STK38L,DHX8,MBNL1,ATF7,TBCA,CDC42SE2,EFNA5,HAS3,MBD2,XPOT,PPM1B,QKI,LRRC8D,ATF7IP,AFF1,C12orf65,RHOA,MTSS1,HSPA9,RAP1B,TMPO,KANK1,TES,FGF1,TBL1X,ARIH2,FAM126A,ZCCHC10,TTC39B,MBOAT1,APPL1,ZBTB34,C11orf95,PIK3R1,ZNF423,GALNT3,BMP4,SEC61B,MSL2,CNBP,PAIP2,ACVR1,MSL1,TSPAN12,SHCBP1,MRPL48,DDX46,SIX1,KDM4C,ANP32B,CD300LG,TMCC3,BAMBI,BTBD10,GABPB1,MTFR1,UBQLN1,RAMP2,RBL1,TMOD3,TANC1,LAMA4,HADH,ARMC1,PDS5B,DENR,ANKRD40,ZNF75A,ZFP62,SMNDC1,PTPN2,TNPO1,C18orf25,CLTC,PLAC9,LIFR,CCDC88A,ATXN7,CLK4,TCF12,RAB23,PKN2,SET,VRK1,LRCH1,ANP32A,GPCPD1,KNTC1,NCALD,PIAS1,CC2D2A,KPNA4,TMTC2,TOP1,RNF34,ESYT1,GMNN,CSDE1,ANKS1A,CDH5,ETV6,CLINT1,GNA13,NAA25,CXCR4,FRMD4B,FUBP3,CTTNBP2NL,RFWD3,CBFB,CORO1C,RPL23,NFYB,C3orf58,COQ5,ARID5B,CSNK1G3,RBM25,EFNB2,ZRANB2,MPHOSPH9,TRA2B,SLC20A1,PLAGL1,KLHL2,MDM2,RCAN1,DPY19L1,DAPL1,CTSC,CDC27,DDHD2,MAPK1IP1L,TGIF1,CTNNA1,DAPK1,DNM1L,KLF5,FGF13,MID1,CDK14,MLLT10,SRSF3,UBE2N,CNN1,RAD52,YME1L1,C11orf58,ATP2B1,SUZ12,MORF4L1,IDH3A,G2E3,TSPAN5,ANAPC1,ERCC4,SNRPC,DDX19B,TPD52L1,KPNB1,BTG3,PDE7A,KIAA1217,PTGES3,EIF4A2,LIMS1,ABI2,ELL2,DOCK11,PPM1D,MARCKS,CDK13,ARL6IP1,ARMC8,KCTD20,CCNYL1,RAB14,SKA2,USP7,KDELR2,APTX,ZFAND3,GDI2,TESK2,COPZ1,RNF128,APPBP2,CEP120
11	homer-M00602	0.0407803	3.17712	M4		PLRG1,SSBP2,CTNNA1,ZFP36L1,PDE7A,CTSC,BHLHE41,DPY19L1,SRSF3,ZNF608,CTBP2,C3orf58,SMARCA5,PKN2,LIMS1,PTMA,CGGBP1,RYBP,CBFB,GABPB1,ZEB1,RAB14,ANTXR2,STK38L,AFF1,ZNF654,GALNT3,CCNYL1,ATF7IP,PDGFC,SYNE2,ATF7,ZNF704,TCF12,HMGB1,MIER3,EIF4E,FZD4,CDK5RAP2,MTSS1,FBXL3,ARID5B,MID1,LRIG3,PDS5B,SIAH1,CDC27,NAA15,ZFR,ACVR1,RAP1B,ARHGDIB,CCDC88A,TRA2B,HNRNPA3,LRRC8D,CDYL,ZNF641,EIF1,SVIL,NCALD,DDX52,MEOX2,MBNL1,PPM1B,AP3B1,ISM1,RCOR1,G2E3,ABI2,APOLD1,RHBDD1,SSPN,KIAA1217,TNPO1,FGF1,TES,PPP1R12A,CSDE1,HNRNPC,RNF38,LEMD3,HAS3,CDC42SE2,ANXA3,MBD2,FGFR1,ATXN7,PPP6R3,SEC31A,FGF13,SPIN1,ZRANB1,TANK,LIFR,FERMT2,UGGT2,KDELR2,MKRN2,FBXW8,ROPN1B,ATP2B1,TSPAN12,TFDP1,TMTC2,PALLD,FKTN,NRP2,SPRY1,TTC39B,RAB2B,GMNN,EPS8,LBR,KPNA4,RND3,PPP4R2,TCEA1,EFNA5,ZC3H15,C2CD2,LAMA4,LIN7C,EHMT1,STARD3NL,KRAS,CXCR4,ARMC1,MMADHC,MTFR1,GPCPD1,UQCRC2,SOX4,AMPH,SMNDC1,CLINT1,TOB1,FRMD4B,CCDC90B,PLD1,VEGFC,CC2D2A,EMP1,FAM126A,EDN1,PIK3R1,CDK14,GCNT1,GTF2A2,SP3,NUP160,HADH,BAMBI,PHLDA1,KLF5,TGIF1,SIX4,SHCBP1,RARB,CSNK1G3,FAR1,LSM6,CHST3,PARG,MORF4L1,FUBP3,PIAS1,PLA2G7,CDK13,ANAPC1,CDC14A,KLHL13,FRRS1
12	hdpi-SF1	0.0407378	3.17053	M3	SF1	APOLD1,ATF7,QRICH1,HNRNPC,QKI,TCF12,SSBP2,SRSF1,ZNF146,EIF4G2,ZNF608,MTSS1,TTC39B,CSNK1G3,GABPB1,MEOX2,CGGBP1,HMGB1,EIF4A2,CTBP2,FRMD6,PPP1R12A,SPRY1,TSPAN13,SKA2,PALMD,CORO1C,ABI2,IRF2,APPBP2,COL11A1,MORF4L1,CD36,EFNA5,SPIN1,SP3,EIF4E,GATA6,ANP32B,ZNF423,ZFP36L1,SIAH1,RAP1B,PTGES3,TOB1,ATF7IP,MIER3,RFWD3,ARID5B,RARB,ZFR,GMNN,MPHOSPH9,MLLT10,ANKS1A,LIMS1,PLAGL1,PAWR,APPL1,CDC42SE2,ZSWIM6,SIX4,RHOA
13	transfac_pro-M02916	0.0407314	3.16955	M2	SRF,ELF3,ZNF362,ZNF384	ZNF608,SOX4,ZFP36L1,APOLD1,TGIF1,PTMA,KLF5,RARB,HMGB1,EDN1,LRIG3,ZSWIM6,DHX8,SIX1,TES,CD36,EIF4E,STIM2,MSL1,PIK3R1,TBL1X,EMP1,PLEKHA5,CTBP2,PPP1R14A,MTSS1,CSE1L,SGK1,ANP32B,MARCKS,EFNA5,SMNDC1,KLF6,CBFB,SP3,ANP32A,TANC1,SVIL,ABI2,HNRNPC,ATF7IP,DDX19B,TTC39B,G2E3,PPM1D,CXCR4,GCNT1,ETV6,KIAA1217,TOP1,ATF7,HAS3,MSL2,NFYB,MRPL48,GATA6,SOX17,BTG3,PDGFC,CCDC88A,TOR1B,LSM6,ZNF280D,LRCH1,SET,CDC42SE2,TCF12,CLINT1,PAWR,PTBP2,C3orf58,LRRC8D,ARMC1,RHOU,STK38L,APPL1,GABPB1,SIX4,AFF1,HNMT,MID1,FRMD4B,NRP2,EFNB2,SSBP2,APPBP2,PLAGL1,GPCPD1,PALLD,MBNL1,MEOX2,DENR,RBM23,BTG2,C12orf65,ANAPC1,RAP1B,CLTC,RFWD3,UACA,FGF1,FGFR1,TBC1D4,FAM126A,TTC8,MBD2,FRMD6,TMOD3,ZKSCAN5,ARFIP1,GDI2,RCN2,LIMS1,DDX46,SUPT16H,RBM5,ZNF506,AFAP1,MTX3,PTK2,KPNA4,BMP4,QKI,ANKS1A,KANK1,FBXO38,SIAH1,GPN3,SRSF3,CCNYL1,ARHGDIB,RBL1,PPM1B,EPS8,RHOA,ZRANB1,SH3PXD2A,RPL6,SUZ12,TSPAN12,MMADHC,MLLT10,ARRDC3,PAIP2,PALMD,RNF128,DAPL1,BAMBI,FRS1,CD300LG,ZNF704,MTFR1,ATP2B1,CDK14,IDH3A,PPP4R2,ERCC4,ZFP62,TMPO,ZNF423,ZEB1,RAB2B,TRA2B,RNF125,VRK1,COPB1,SPRY1,RHBDD1,PLAC9,PIAS1,DSG3,ATXN7,PDS5B,CC2D2A,PKN2,ARL6IP1,HADH,ZBTB34,MDM2,ANKRD40,ZDHHC13,LIMCH1,IRF2,RAMP2,GRASP,TBCA,PLD1,FBXW8,TTC28,RCAN1,PHLDA1,TPP2,FAR1,ARIH2,QRICH1,MAT2B,HSPA9,PTGES3,GNA13,YME1L1,PCBD2,DPY19L1,ACVR1
14	wolfe-Sqz_SOLEXA_F1-3	0.0404695	3.12907	M2	ZNF362,ZNF513	APOLD1,ZFP36L1,SOX4,HMGB1,EIF4E,SGK1,KLF6,SP3,CSE1L,STIM2,GCNT1,TOR1B,PLEKHA5,ZNF608,HNRNPC,CTBP2,QRICH1,PTMA,FGFR1,PPP1R14A,GPN3,LSM6,RPL6,TBCA,CDYL,RBM5,ATF7,EDN1,PALLD,SIAH1,RAB2B,TMPO,DHX8,EFNA5,RHOA,SPRY1,STK38L,HAS3,ATF7IP,CDC42SE2,MTSS1,AFF1,MBD2,ZCCHC10,TBL1X,C12orf65,MBNL1,HSPA9,ZBTB34,QKI,LRRC8D,XPOT,SHCBP1,MSL2,APPL1,RAMP2,MBOAT1,MTFR1,PAIP2,TTC39B,TES,ARIH2,RAP1B,PPM1B,FGF1,UBQLN1,RBL1,MRPL48,TANC1,SIX1,CNBP,KANK1,DDX46,TMCC3,C11orf95,ZFP62,GALNT3,ARMC1,MSL1,BMP4,KDM4C,BAMBI,TMOD3,PTPN2,FAM126A,PIK3R1,TSPAN12,SEC61B,BTBD10,MARCKS,ZNF423,ACVR1,PKN2,DENR,PIAS1,ESYT1,KNTC1,RBM25,GABPB1,CLTC,HADH,C18orf25,ANKRD40,KLHL2,CD300LG,SET,LRCH1,RAB23,VRK1,PLAC9,ANP32B,LAMA4,PDS5B,ANP32A,SMNDC1,CLK4,ZFAND3,FRMD4B,ANAPC1,CDH5,ETV6,TCF12,RCAN1,GPCPD1,SUZ12,COQ5,G2E3,CDC27,PCBD2,TGIF1,TMTC2,CXCR4,CSDE1,CC2D2A,ATXN7,FUBP3,RNF34,EFNB2,GMNN,CTTNBP2NL,ZNF75A,CLINT1,ATP2B1,CD36,ANKS1A,NAA25,GNA13,CCDC88A,LIFR,MID1,NFYB,DNM1L,DPY19L1,UBE2N,CBFB,BTG2,DDX19B,MPHOSPH9,SLC20A1,RFWD3,TNPO1,MDM2,SRSF3,PTGES3,DAPL1,GDI2,ARID5B,MAPK1IP1L,

						TESK2,SNRPC,TSPAN5,NCALD,SOX17,FGF13,ZNF506,FBXL3,KPNA4,CCNYL1,RPL23,ARMC8,KIAA1217,DDHD2,C3orf58,BTG3,TOP1,CDK14,LIMS1
15	selexconsensus-pho	0.0401841	3.08497	M1	YY1,YY2	SAA1,HNRNPA3,PPIA,HNRNPK,RBMX,MTFR1,TANC1,ATF7,UQCRC2,PECAM1,ZNF608,IQCJ,ETV6,CCDC88A,FGF1,RBM25,VRK1,SRSF3,SLC31A1,EFNA5,SP3
16	stark-GCCATT	0.0400585	3.06555	M1	YY1,YY2	SAA1,HNRNPA3,RBMX,PPIA,ZNF608,HNRNPK,MTFR1,EFNA5,UQCRC2,CCDC88A,SOX4,ATF7,PECAM1,ANP32A,TANC1,SLTM,VRK1,SVIL,HNRNPC,MID1,FGF1,IQCJ,FAR1,CLK4,ARMC1,PALLD,TMCC3,CTBP2,SLC31A1,SRSF3,SP3,LIMS1,STK17B,CORO1C,CTTNBP2NL,SEC11A,SIAH1,ZBTB34,RBM25,ATXN7,KLF5,CLDN5,ZNF280D,AQP1,ZRANB1,TES,FIBIN,HMGB1,CHST3,TSPAN12,LRRC8D,ETV6,STK38L,RARB,EIF4G2,MTX3,PLD1,FGF13,CC2D2A
17	flyfactorsurvey-phol_SANGER_5_FBgn0035997	0.0397987	3.0254	M1	YY1,YY2	TRA2B,EIF4G2,PPP2CA,CC2D2A,SRSF3,PTMA,ZNF280D,CDK5RAP2,MORF4L1,CLK4,PDS5B,PPIA,RBM25,NAA15,NPM1,SIX4,RPRD1A,PIAS1,TRIM37,FBXL3,ARMC8,ISCU,RAB14,RBM23,ABCE1,RBM5,SART3,CSE1L,UQCRC2,DEK,SIAH1,HNRNPC,EDN1,SMC3,MMADHC,MRPL42,WBP11,ANP32B,PPP6R3,HNRNPA1,ZNF585B,RBMX,RPL14,HSPA9,RAP1B
18	yetfasco-1820	0.0397029	3.01059	M2		HMGB1,CMKLR1,RBM23,ARID5B,MEOX2,ZBTB34,MID1,FGF1,UQCRC2,SAA1,VEGFC,TANC1,FGF13,C9orf47,VRK1,PPM1B,SPIN1,NRP2,ZC3H15,SSBP2,TMTC2,SSPN,LSP1,MTFR1,PKN2,ZFP36L1
19	transfac_public-M00183	0.0396795	3.00697	M4	MYB,TP53,STAT6,MYBL2	ACVR1,RPS12,CTNNA1,SIAH1,RCOR1,TOB1,SSBP2,CSDE1,FGF13,APOLD1,MDM2,SRSF3,MBNL1,UBE2E2,TSPAN5,GCNT1,PALLD,KLF5,KANK1,AFF1,MBD2,CTBP2,ZEB1,DDX20,TFPI2,CDK14,ESYT1,PIGY,EMP1,UQCRC2,C12orf65,LIMS1,KIAA1217,CTSC,MID1,STARD3NL,TBL1X,TNPO1,ZBTB34,ABCE1,MYLK,PLD1,EFNA5,RPP30,PPP2CA,KCTD9,HNRNPC,RAB14,SNN,ATF7,ISM1,TNFAIP8L3,PDS5B,SLC43A3,TES,RND3,RAP1B,PLA2G7,PPP6R3,DSG3,SIX4,NRP2,BHLHE41,ARMC1,PKN2,CLTC,ZRANB1,UGGT2,HSPA9,SOX4,MTFR1,ANTXR2,EIF4G2,FAM126A,KRAS,SMARCA5,MIER3,GALNT3,PINX1,CGGBP1,ETV6,FKTN,FGF1,RBBP8,ZNHIT6,SVIL,CTTNBP2NL,BMP4,HAS3,CDC14A,LGALS3,NUP160,NCALD
20	jaspar-PF0164.1	0.039671	3.00565	M1	E2F3,E2F2	GABPB1,PPP2CA,PDS5B,CC2D2A,FBXL3,CDC14A,CSE1L,TOP1,RBMX,EIF4G2,RBM5,PPIA,SUPT16H,SRSF1,ZBED5,TRA2B,CLK4,GALNT3,SIX4,NFYB,ZNF608,SRSF3,DEK,TRIM37,PTBP2,HNRNPC,CLDND1,CCNYL1,PPP1CC,NAA15,QKI,KIAA1217,HNRNPA3,CDC42SE2,TIAL1,FGFR1,SLC20A1,ATF7,KLHL2
Rank	Track id	AUC	NES	ClusterCode	Transcription factor	Target genes
1	wgEncodeHaibTfbsMcf7MaxV0422111PkRep1.broadPeak.gz	0.0488866	3.41698	T1	MAX	NOC3L,CNBP,MBLAC2,SNORA9,CAMSAP1,ZNF280D,THNSL1,ANP32B,ENOPH1,HNRNPA1,GPN3,CHRAC1,ING5,APOLD1,TIGD7,ZNF131,ZNF75A,NPM1,PPIA,HSPA9,GDI2,KPNB1,VPS29,RNF130,ZFP36L1,MSL2,XPOT,RHBDD1,RRP1B,IARS,ABCE1,PA2G4,SP3,HNRNPK,TGDS,GOLIM4,UBE2D2,GABPB1,FMNL3,ANAPC7,SCFD2,RPSA,LEMD3,TBL1X,TCERG1,L2HGDH,RPIA,PWP1,RPL35,HNRNPA3,SLC20A1,RAN,MBNL1,TRA2B,RIOK1,MRPL48,PCBD2,SRSF1,FBXW8,EIF4E,TGIF1,PSMG1,NCBP2,PINX1,PRR3
2	GSM1208654_batch1_chrom1_LoVo_MYC_Passed QC_peaks_hg19	0.0473407	3.23186	T2	MYC	RPL4,KIAA0586,PPID,HNRNPA1,GPN3,PWP1,ABCE1,NPM1,IARS,SNORA9,CCT2,GABPB1,PHB2,POLR3B,FAM173B,CCT5,TARS,FARSB,RIOK1,MSL2,SRSF1,PA2G4,PTGES3,HSPA9,RPL7L1,ZNF146,HNRNPK,MPHOSPH9,C12orf65,SET,ANAPC7,LGALS3,MBLAC2,RPLP0,RPL23,RRP9,XPOT,CLTC,GOLIM4,ZC3H15,RSL24D1,EIF3E,RPLP1,EIF1,NAA25,RPL35,RPL12,TRIAP1,ZFAND1,TBL1X,RPS25,RPL24,TCERG1,TGIF1,RIOK2,KDM4C,COPZ1,KPNB1,RPL6,MRPL48,EEF1B2,RPS24,THNSL1,CNBP,TIAL1,ZNHIT6,TRMT11,RPSA,RPS29,CBR1,LARS,MTPAP,RPS12,ARRDC3,SPRY1,HNRNPA3,GTPBP4,SNAPC5,NUP155,KARS,RFC4,PCNP,NCBP2,IQGAP2,MSH3,SMC3,ANAPC1,SKA2,PRPF40A,MSL1,SRP54,UBE2D2,RHBDD1,KNTC1,MRPL42,PPIA,HNRNPC,SCFD2,RNF125,ZNF181,RPL32,C1GALT1,SUPT16H,NUFIP1,EIF4A2,NAF1,EIF4G2,SART3,ISCU,KPNA4,ZNF207,RNF34,ZNF131,RAD52,RBMX,EIF4E,ATF7,HIST1H2BC,NOL11,SOX4,BLMH,AFF1,KDELC2,GDI2,XRCC5,POC5,RPIA
3	wgEncodeHaibTfbsSknshMaxV0422111PkRep2.broadPeak.gz	0.0470469	3.19667	T1	MAX	IARS,NPM1,GPN3,GOLIM4,ENOPH1,RHBDD1,SCFD2,SMC3,GDI2,SNORA9,GABPB1,NOC3L,PTGES3,QRICH1,GTPBP4,KPNB1,KIAA0586,NME6,RNF130,SUPT16H,HNRNPA1,TCERG1,ETV6,VPS29,PWP1,TIGD7,ZNF75A,RRP9,TIAL1,PPIA,MBLAC2,THNSL1,TBL1X,TGDS,PHB2,RIOK2,CLTC,SRSF1,MSL1,HSPA9,RPS18,TARS,RIOK1,ANAPC7,ZNF131,KPNA4,HNRNPK,M6PR,RPL4,ZSWIM6,MBNL1,ANP32B,RPL6,LARS,NUP98,ING5,NSUN2,EFNA5,MRPL48,RPL35,SKA2,RPSA,EEF1B2,MSL2,CHRAC1,ZFR,NUP155,CDC14A,TGIF1,L2HGDH,APOLD1,XPOT,FBXW8,XRCC5,PPID,ZNF181,ARIH2,RRP1B,FARSB,KNTC1,RPL23,EIF4E,CRY1,RPS12,ZKSCAN5,ZFP36L1,ZNF280D,SNRPC,SPRY1,RPL12,ZNF510,MLLT10,NUFIP1,CBR1,PTMA,ABCE1,PINX1,RFC3,RSL24D1,NAA15,TCF12,TRA2B,PPM1B,TMPO,MAT2B,CCDC88A,SRP54,CNBP,FGFR1,RPIA,PRR3,TEX10,ANAPC1,UBE2D2,TSPAN5,NFYB,RBM25,EIF3J,RPLP0,CHST3,CNN1,NUDT15,PA2G4,RPS24,HNRNPA3,ZNF146,POLR3B,RBBP8,CTBP2,PCNP,GTF2H3,ANKRD40,CDCA7L
4	wgEncodeHaibTfbsMcf7MaxV0422111PkRep2.broadPeak.gz	0.046934	3.18315	T1	MAX	GABPB1,NPM1,TRIM37,RHBDD1,SNORA9,IARS,RNF130,CLTC,HSPA9,TIGD7,ZNF75A,ING5,SCFD2,ZNF131,TOP1,PWP1,MBLAC2,TCERG1,SMAGP,ZFP36L1,MSL2,GDI2,LEMD3,SRSF1,SKA2,GOLIM4,RPL35,RPS18,THNSL1,RRP1B,VPS29,SUPT16H,KIAA0586,TBL1X,FAM173B,CCT5,ABCE1,RPL32,ZNF280D,RIOK2,NOC3L,PA2G4,GTF2H3,HNRNPA1,RPSA,EEF1B2,HNRNPK,XPOT,RPS29,CNBP,RIOK1,EIF3E,L2HGDH,FMNL3,PPIA,ANAPC1,RAD17,TGDS,NOL11,GPN3,ISOC1,RRP9,TIAL1,GTPBP4,FBXW8,POLR3B,SLC20A1,APOLD1,ANAPC7,ENOPH1,RPLP0,ANP32B,SRP54,MAPK1IP1L,RPS12,NDUFAF2,FARSB,PTGES3,RPL24,SMC3,MRPL48,PRR3,PINX1,MYLK,RPS25
5	wgEncodeHaibTfbsH1hescPol24h8V0416102PkRep2.broadPeak.gz	0.0468893	3.1778	T3	POLR2A	EIF1B,MRPL48,ZNF721,KLF6,SMARCA5,TFAM,RAD17,ZCCHC9,APTX,RCN1,KPNA4,RPL6,HNRNPC,CLTC,RPL35,TMPO,ZNF131,RPS18,EIF1,SNAPC1,RPL4,ZNF146,HNRNPK,EIF4E,MDM2,ZFP36L1,PCNP,TIMM23,ASH2L,PPP4R2,SNORD83A,CSE1L,SEC31A,PRPF40A,PTMA,NUP160,NCBP2,RBMX,PRR3,CNBP,TRA2B,PPP1R12A,CCDC90B,ENOPH1,RPS24,RPL24,RPLP0,SRP54,FBXO38,MORF4L1,ZC3H15,RBM25,HNRNPA1,CCDC59,C9orf85,PAIP2,CSDE1,HIST1H2BC,RPSA,RAB21,STRAP,MAT2B,GABPB1,DDX46,HNRNPA3,EIF4G2,NUDT15,ARL6IP1,ING5,QRICH1,PARG,SET,RPS29,SRSF3,MAPK1IP1L,CDK13,SNAPC5,PPP2CA,C3orf58,SNHG6,HMGB1,TPP2,DDX19B,RPL7L1,RPL12,CCT2,PCID2,RPS12,EIF4A2,KPNB1,CEP57,RBL1,TOB1,FABP5,RPLP1,NPM1,RPL32,PSMG1,PDE7A,PMAIP1,HAPLN3,WDR48,PPM1D,ZNF397,SNORA9,PPIA,MSL2,MARCKS,ARRDC3,SRSF1,RPL23,TRIAP1,NAF1,SIAH1,CDC14A,C11orf95,FAM173B,CCT5,PALLD,GMNN,A

						CTL6A,KIF20B,GDI2,IARS,UBE2N,UBQLN1,BZW1,LIN7C,GPN3,TEX10,ANAPC1,EIF3J,MMADHC,DSC2,NARS,SLTM,CDCA7L,SLC15A4,TRMT11,RAB33B,RFC4,PAIP1,GTPBP4,TIAL1,TFDP1,TRIM37,UHRF2,SKA2,PKN2,ZBED5,HNRNPM,ANKS1A,TARS,GTF2H3,CCNYL1,MED21,EIF3E,RPL14,RSL24D1,EXOSC9,PCBD2,ABCE1,KDM4C,PTBP2,OXCT1,TOP1,RNF38,DDX20,DCP1A,PPP1R14A,PA2G4,XRCC5,SEC61B,ATF7IP,CETN3,HMGN1,MLLT10,SNRPC,SEC11A,PTGES3,ELP3,ZNF23,CENPK,SOX4,NUFIP1,RHBDD1,TERF1,GUF1,CCDC82,MKRN2,RAD52,UBA2,CLDND1,NFYB,TIGD7,ZNF75A,CGGBP1,C3orf38,SP3,KNTC1,NOL11,RHOA,ZBTB11,RNF34,GLO1,CHRAC1,MRPL42,ARMC8,COQ5,AFF1,PTPN2,IK,EXOSC7,ZNF207,RRP1B,L2HGDH,RND3,RPS25,SMC3,SUPT16H,PHB2,PWP1,ANP32A,TNPO1,MRPL2,RCN2,HSPA9,AHR,M6PR,SGK1,TOMM6,H2AFV,SMNDC1,SUZ12,CTSC,THNSL1,ANXA3,NAA15,TSPAN2,TCERG1,ZSWIM6,PSPC1,ZNF844,RPE,ATF7,LARS,MBLAC2,ZFAND1,CSNK1G3,ACTR8,ORC2,ESYT1,XPOT,RCOR1,MTX3,LRIG2,APPBP2,BTG3
6	wgEncodeHaibTfbsA549Taf1V04022111Etoh02PkRep2.broadPeak.gz	0.0467254	3.15816	T4	TAF1	RPL4,RPLP1,PTMA,AKR1C3,HNRNPA1,ZSWIM6,CLTC,RPSA,NDUFA12,RPLP0,RPL23,RPL6,GABPB1,EIF4A2,SNHG6,KNTC1,NPM1,NOL11,PHB2,RNF34,COPZ1,CCDC88A,MNAT1,RPL12,SNORA9,RPS18,RPL35,C9orf85,ZBED5,C12orf65,SRSF1,STK17B,CCT2,EIF1,TMPO,RPL32,RPL24,PWP1,DPM1,CSDE1,ASH2L,PPM1D,C11orf58,HNRNPC,MRPL48,POLR3B,WBP11,STRAP,TRA2B,ARRDC3,GTF2H3,MMADHC,ANAPC1,RPS24,TNPO1,SLTM,CCDC59,RND3,EIF3E,FAM173B,CCT5,KLF5,ALDH1A1,GTF2A2,DCP1A,CCDC90B,KARS,UBA2,RFWD3,CSNK1G3,EIF4G2,ZNF207,SEC11A,ZCCHC10,RPS12,RPL7L1,PMAIP1,HNRNPM,TOMM6,CENPK,TOB1,CGGBP1,C3orf38,ZFP36L1,HSPA9,EDN1,SEC31A,CDK13,AFF1,CETN3,PARG,PIGY,RPL14,KLF6,HNRNPK,MRPL42,PPP1R12A,RARB,LIN7C,GDI2,SMC2,SART3,RPS25,ELP3,IDH1,SKA2,MKRN2,RBMX,XRCC5,THNSL1,TBCA,RHBDD1,ARL6IP1,APPBP2,TCERG1,HEATR6,TIMM23,DENR,MED21,RPS29,C1GALT1,KRT81,WDR11,WDR48,KPNB1,LRRC59,STAM,TXN,COQ5,MED7,MARS2,SLC15A4,NOC3L,PPID,CDC27,PPIA,NUP160,RSL24D1
7	wgEncodeHaibTfbsSknshTaf1V0416101PkRep2.broadPeak.gz	0.0467168	3.15714	T4	TAF1	UBE2D2,UBA3,ANTXR2,HIST1H2BC,RPS18,EIF1,RPL4,HNRNPK,ANKRD40,EIF4E,DPM1,PCNP,SLTM,KLF6,TMPO,RPL6,PRPF40A,PTMA,SIX1,CCDC88A,CNBP,DCP1A,TRA2B,KNTC1,PHLDA1,PPP1R12A,ENOPH1,RPS24,RPL24,RPLP0,SP3,MMADHC,MORF4L1,RBM25,HNRNPA1,C9orf85,CSDE1,RPSA,GABPB1,HNRNPA3,EIF4G2,ARL6IP1,PHB2,ANP32A,SRSF3,MAPK1IP1L,CDK13,APOLD1,MNAT1,RPL7L1,RPL12,MRPL48,ZSWIM6,RPS12,GDI2,EIF4A2,KPNB1,STK17B,SKA2,TOB1,RPL35,RPLP1,NPM1,NOL11,PMAIP1,ZNF207,PPM1D,ZBED5,AFF1,PPP1CC,SNORA9,PPIA,CLTC,MSL2,MARCKS,ARRDC3,SRSF1,RPL23,TIAL1,FRMD6,RCOR1,CCDC90B,RPS25,TNPO1,CCDC59,RNF34,ZKSCAN5,BZW1,WDR11,SNHG6,ZNF397,CSNK1G3,MKRN2,FBXO21,QRICH1,ASH2L,ELP3,NCBP2,SGK1,HSPA9,CHRAC1,ACTR8,RND3,XRCC5,C1GALT1,CTR9,HMGB1,H2AFV,APPBP2,TCF12,CORO1C,PKN2,DDX20,SEC31A,HNRNPM,TPP2,SENP1,RPL14,SOX4,MLLT10,PIGY,PSPC1,NDUFA12,GTF2H3,RHOA,DDX46,ZRANB2,SLC20A1,EIF3J,XPOT,ZC3H15,TOMM6,SEC11A,KPNA4,CEP57,SMC3,RPL32,C11orf58,STAM,CCT2,LIN7C,PAIP1,PTPN12,DHX8,KIAA0586,TGDS,MBNL1,SNRPC,C12orf65,GTPBP4,CSE1L,SART3,ISCU,POLR3B,TANC1,COQ5,PRR3,FAM173B,CCT5,GNA13,ZNF146,WBP11,PTGES3,EIF3E,RFWD3,GPN3,DNM1L,CETN3,ANXA5,MRPL10,UBE2N,UBQLN1,SRP54,HEATR6,TRMT11,TARS,PCBD2,COMMD2,MTX3,DENR,CENPK,SLC15A4,ELL2,WDR61,ANAPC7,TRIM37,PCID2,RARS,RAD52,PPM1B,AKAP2,RAD17,RANBP6,PPP2CA,VEGFC,NAA15,RBMX,HNRNPC,M6PR,ETV6,TCERG1,KARS,TOP1,PPID,ZFP62,GTF2A2,EPS8,LIMS1,TXN,CGGBP1,C3orf38,C18orf25,SMC2,SSBP2,THNSL1,ZFAND3,MRPL42,ARMC1,RAB33B,SUPT16H,RBM5,CBR1,WARS,PDS5B,CDC42SE2,HSPB6,DDX19B,MDM2,KIF20B,ARMC8,ZNF23,RIOK1,PA2G4,LRRC59,GMNN,SEC61B,UBA2
8	wgEncodeHaibTfbsGm12892Taf1V0416102PkRep2.broadPeak.gz	0.0467041	3.15561	T4	TAF1	DPM1,PTMA,RPLP1,RPL12,TRA2B,RPL6,SNHG6,ZSWIM6,RPS24,RPL35,RPL4,MNAT1,GABPB1,CLTC,RPSA,RPL7L1,C9orf85,FBXO21,SNORA9,ZNF207,EIF4G2,DDX46,STK17B,HNRNPK,CNBP,EIF4A2,ENOPH1,RNF34,SLTM,TOB1,RPS18,CSDE1,RPLP0,NPM1,GTF2H3,FAM126A,RPL23,SRSF1,RPL24,AFF1,MKRN2,ANP32A,ZNF397,HNRNPA1,GDI2,PPP1R12A,XRCC5,RPL32,CCT2,QRICH1,KARS,CCDC59,SEC61B,PKN2,TXN,TPP2,ACTR8,KNTC1,TGDS,SLC15A4,ARRDC3,GLO1,TMPO,HMGB1,MRPL48,TRMT11,ANXA5,FRMD6,UBA3,RBM25,C18orf25,SP3,HNRNPC,RHBDD1,LEMD3,RIOK1,MMADHC,ASH2L,CDK13,PPID,SEC11A,C11orf58,MSL2,TNPO1,PMAIP1,FAM173B,CCT5,TOMM6,PHB2,CHRAC1,EIF1,ZC3H15,CSNK1G3,GMNN,TIMM23,CLK4,RPS12,PCID2,PARG,PCNP,ZBED5,RHOA,SRSF3,NCBP2,APOLD1,NUP155,HNRNPM,RPS25,UBE2D2,MDM2,RFWD3,ZKSCAN5,ARMC8,KLF6,EIF4E,CCDC88A,RARS,UBQLN1,PPIA,SMNDC1,RSL24D1,SART3,C12orf65,RPL14,DCP1A,MBLAC2,NARS,PTGES3,GUF1,UBA2,RBMX,PAIP1,FGFR1,EIF3J,MORF4L1,NDUFA12,CDC42SE2,ZCCHC9,ARMC1,NOL11,MARCKS,ANAPC1,ZNF23,TCERG1,ZFAND1,HSPA9,ZNF721,IARS,PPP1CC,KPNB1,RPS29,PTPN2,HNRNPA3,ZRANB2,RBM5,PIGY,BZW1,CSE1L,PPM1D,DDX20,GTPBP4,RFC5,MED21,PRR3,FAS,SUZ12,TESK2,EIF3E,LRRC59,NAF1,ING5,VPS29,KIF20B,WDR11,MAPK1IP1L,TCF12,NUP160,ZNF280D,RBL1,CORO1C,DNM1L,COPZ1,CCDC90B,WDR61,MBNL1,MBD2,ZNF506,CTR9,CXCR4,ZNF146,ANKRD40,DENR,FBXO38,RPE,TRIAP1,LIN7C,DDX19B,COQ5,NDUFAF2,YME1L1,SOX4,ATF7IP,STK38L,BTG3,MRPS9,NME6,CETN3,PWP1,PSPC1,ZCCHC10,FARSB,UHRF2,ZFP36L1,C3orf58,IRF2,PPP2CA,ANP32B,GTF2A2,CDCA7L,TBCA,RAD17,NMD3,EXOSC7,SEC31A,LOC374443,POC5,CLDND1,SRPK1,ELP3,PYROXD1,TFAM,CEP57,KDM4C,KDELC2,SSBP2,FRG1,RRP1B,PRPF40A,COMMD2,SLC20A1,FAM122A,WDR48,RFC4,ARL6IP1,TIGD7,ZNF75A,LSP1,ZNHIT6,SET,PLAGL1,PDE7A,OXCT1,KCTD9,SNRPC,MRPL42,CBR1,ELL2,PIAS1,MRPL2,FKTN,KCTD20,CEP120,RCN2,CWC27,IDH3A,BTBD10,RANBP6,ERCC4,GPCPD1,MRPL10,CTSC,MED7,RNGTT,FABP5,ZFAND3,EIF1B,CENPK,RCOR1,ACTL6A,SUPT16H,DDHD2,TEX10,EHMT1,M6PR,MTX3,HMGN1,RNF8,GMPS,TOP1,AP3B1,RAP1B,NFYB,RNF125,WARS,CCNYL1,CDC23,NAPSB,CDV3,SNRNP48,TERF1,RIOK3,ETV6,TIAL1,ZNF41,PPM1B,HADH,MARS2,ISOC1,RAN,PTPN12,RAB14,COPB1,ABCE1,NSUN2,BTG2,FUBP3,SLC43A3,APPBP2,POLR3B,L2HGDH,AHR,ZNF131,NUDT15,RPP30,CDC27,TANK,APTX,ANAPC7,NGDN,UQCRC2,STRAP,THNSL1,DNAJC24,SMC3,SRP54,TANC1,DOCK11,SGK1,ARFIP1,TCHP
9	wgEncodeHaibTfbsPfsk1Taf1V04416101PkRep2.broadPeak.gz	0.0458161	3.04928	T4	TAF1	RPS18,RPL7L1,HNRNPA1,RPL4,HNRNPK,GABPB1,RPS12,RPL23,HIST1H2BC,CLTC,ZSWIM6,RPL6,EIF4A2,KPNB1,RPLP1,TOB1,ASH2L,SNORA9,PRPF40A,RPLP0,TRA2B,RPS24,EIF3E,ARRDC3,GPN3,NPM1,KNTC1,SLTM,EIF4G2,SKA2,ENOPH1,ELP3,SRSF1,TARS,HNRNPM,CEP57,WBP11,EIF4E,MRPL48,SRSF3,STK17B,CSNK1G3,CSE1L,PPIA,PHB2,ZNF207,RPL14,RPL24,PTGES3,RPSA,DPM1,TIAL1,CHRAC1,RPL35,CDK13,GTF2H3,HNRNPA3,MMADHC,MNAT1,CSDE1,SRPK1,RPL12,FAM173B,CCT5,PPP1CC,ZC3H15,RNF128,PRR3,ARMC1,MTFR1,SNHG6,RPS25,MARCKS,CCDC59,RND3,UBE2D2,MSL2,RPS29,CCDC88A,ZRANB2,MORF4L1,GDI2,SOX4,UBA2,CCT2,PDE7A,KARS,CDC14A,SP3,KCTD9,XRCC5,TMPO,EIF3J,RIOK1,C9orf85,NOL11,MKRN2,STRAP,TOMM6,EIF1,CETN3,SEC11A,CNBP,TES,ZNF146,UBA3,ANKRD40,ANAPC1,MBLAC2,SNRPC,STAM,PKN2,NARS,TNPO1,UQCRC2,IARS,APOLD1,GTPBP4,RSL24D1,WDR61,KIAA0586,PTPN12,LARS,MLLT10,COQ5,DCP1A,MBNL1,SLC20A1,XPOT,GUF1,SENP1,RBM25,TCERG1,C12orf65,SGK1,APPBP2,SLC15A4,RPL32,CCDC90B,AFF1,PAWR,FAM126A,CGGBP1,C3orf3

						8,RFWD3,NAA15,UBE2N,ELL2,RANBP6,SART3,ISCU,RBMX,SNRNP48,DDX46,DDX19B,RHOA,NFYB,PIGY,NDUFA12,ARL6IP1,SMNDC1,PTPN2,ABCE1,KCTD20,TIGD7,ZNF75A,C11orf58,KLF6,MAT2B,RARS,LEMD3,WDR11,ZNF397,TGDS,SEC31A,RCOR1,ZNF23,LIN7C,IRF2,CWC27,FBXW8,DDHD2,PPP1R12A,SEC61B,ZNHIT6,COPB1,SRP54
Rank	Motif id	AUC	NES	ClusterCode	Transcription factor	Target genes
1	element o-AAAATGGCG	0.0504026	4.66433	M1	YY1	TRA2B,PPP2CA,GABPB1,EIF4G2,PPIA,CLK4,SRSF3,CC2D2A,PTBP2,PDS5B,HNRNPA1,NAA15,TRIM37,FBXL3,RBM23,DEK,CDC14A,RBM5,SMC3,CSE1L,CDK5RAP2,LIN54,ANP32B,SGK1,NPM1,TOP1,ZBED5,PPP6R3,HNRNPK,SOX4,HMGB1,RBMX,EIF1B,RAB14,GDI2,MORF4L1,ABCE1,ZNF280D,NFYB,MMADHC,ARID5B,MARCKS,SIAH1,SET,ATXN7,SRSF1,CCNYL1,SIX4,UQCRC2,PTMA,HNRNPC,BMP4,ANP32A,ZFR,KRAS

RNA-seq Overlap Upregulated Genes in iRegulon						
Rank	Motif id	AUC	NES	ClusterCode	Transcription factor	Target genes
1	transfac_pro-M00693	0.0649903	4.53221	M1	TCF3,MYOD1,MYF6,MYOG,TCF4,MYF5,ASCL1,ARID5B,TCF12,SNAI2,ID4,LMO2,SREBF2,SREBF1,NR3C1,ASCL2	SYNPO,DEGS2,ZFYVE28,DUSP8,TDRD5,CRB3,CYB561D1,CST6,TMEM125,C16orf74,ANKRD9,TUBB3,ESPN,STXBP2,C1orf159,SLC9A1,SCAMP2,HPCAL1,PCSK6,CTDSP2
2	transfac_pro-M01669	0.0640661	4.43901	M2		ZBTB7B,CD81,EFHD2,KIAA1522,MNT,C2CD2L,SRRM2,CCND3,ADAM15,DEGS2,BAZ2A,MXI1,SYNPO,SNPH,STK32C,C16orf74,SCAMP2,POMGNT1,VWA1,ZFYVE28,MTMR11,AMOTL2,LYPLA2,S100A13,TMEM102,TDRD5,TRIM62,MLXIP,HPCAL1,CBX8,LYN,IQSEC1,ST14,NAT8L,S100A16,CYB561D1,DUSP23,SLC9A1,ANKRD9,RPS6KA1,PCSK6,PTK6,KLF13,BCKDK,CHD4,MLLT6,AGTRAP,SLC1A1,C1orf159,LIN7A,MSN,ARHGEF10,CELSR1,ST3GAL4,SEC14L1,PLEKHA2,YWHAB,PDE4A,RELB,PQLC2,FOXO1,STIM1,ESPN,SNX1,CTDSP2
3	transfac_public-M00344	0.0620007	4.23073	M1	NHLH2	LLGL2,CYB561D1,SYNPO,C16orf74,CRB3,LRRC8E,C1orf159,ANKRD9,PDE4A,KIAA1522,POM121,DEGS2,HPCAL1,PQLC2,ESPN,ST3GAL4,TDRD5,PCSK6,RPS6KA1,GTF2A1L,CELSR1,SLC9A1
4	jaspar-PF0006.1	0.0601443	4.04351	M3	SP1,SP4,SP2,SP3	TRIM62,MLLT6,C2CD2L,MNT,ZNF687,PHF21A,LLGL2,MXI1,STK32C,ELMO3,SLC44A2,ZNF296,USF1,ZBTB7B,KLF13,FOXO1,CBX8,ACSS2,STIM1,ANKRD9,BAZ2A,CD81,CELSR1,TUBB3,S100A13,MUC1,LRRC8E,SNX1,ESPN,CRB3,TDRD5,DDRGK1,UBE2E1,VWA1,MSN,SYNPO,DEGS2,DUSP23,POM121,NMT2,ORAI2,KLHL36,HPCAL1,ADAM15,PQLC2,EFNA4,C16orf74,PDE4A,POMGNT1,ST3GAL4,HDAC10,FBXW4,CTDSP2,ZFYVE28,RELB,NDST2,SLC1A1,SNAPC2,IQSEC1,HLA-E,PIGV,SCAMP2,ST6GALNAC4,CCND3,TPCN1,GABARAPL1,EPS8L1,PSKH1,FAM43A,TMC4,NRP1,MMP11
5	homer-M00860	0.0582638	3.85387	M4		ZBTB7B,CD81,MTMR11,BAZ2A,KIAA1522,C16orf74,CHD4,VWA1,MSN,EFHD2,RND1,S100A13,TDRD5,EFNA4,ZNF687,ZFYVE28,MXI1,ORAI2,CELSR1,ST3GAL4,SLC9A1,PIGV,PHF21A,TTC39A,LRRC8E,KLF13,SYNPO,TUBB3,TMEM102,SLC22A18,GTF2A1L,BCKDK,PDE4A,PNPLA2,SNPH,FOXO1,TRIM62
6	hdpi-ZNF160	0.0573236	3.75905	M5	ZNF160	BAZ2A,ZNF593,SYNPO,POMGNT1,PDE4A,TMEM125,AMOTL2,CBLC,ZNF687,ZBTB7B,RND1,SNPH,PI4KB,LPIN1,SLC9A1,ORAI2,FBXW4,C2CD2L,ESPN,S100A13,EFHD2,KIAA1522,MSN,CHD4,TDRD5,TRIM62,SLC22A18,ST3GAL4,ZFYVE28,S100A16,CCND3,C16orf74,EHD2,RPS6KA1,FBXO41,SERPING1,ARHGEF10,PHF21A,TMEM102,CASP9,MXI1,MLLT6,LYPLA2,RELB,HPCAL1,PTPRB,LMBR1L,UBE2E1,SRRM2,YWHAB,DDRGK1,IQSEC1,LRRC8E,NMT2,CST6
7	tfdimers-MD00080	0.0568976	3.71609	M6	MAFA,OVOL2	PHF21A,CYB561D1,NDST2,DUSP8,EFHD2,ZFYVE28,ZBTB7B,C16orf74,NRP1,EFEMP1,BAZ2A,ANKRD9,STK32C,SYNPO
8	encode-UW.Motif.0600	0.056544	3.68043	M4		TMC4,SYNPO,C16orf74,ZNF611,ST3GAL4,EPS8L1,CCND3,VWA1,TUBB3,TTC39A,ANKRD9,EFHD2,CBLC,C1orf159,NAT8L,DUSP8,CYB561D1,RND1,EHD2,TSTA3,CELSR1,BAZ2A,ACSS2,ZBTB7B,CD81,MLLT6,ADAM15,GTF2A1L,TMEM102,TRIM62,PNPLA2,CTDSP2,SLC9A1,TWIST2,ZNF593,DEGS2,S100A13,CLSTN1,RAB25,IQSEC1,LPIN1,TDRD5,ZFYVE28,SNPH,PCSK6,PTPRB,ESPN,POM121,KIAA1522,RPS6KA1,MXI1,HPCAL1,ORAI2,SLC1A1,STIM1,PTK6,TPCN1,ST14,MLXIP,FAM173A,S100A14,PDE4A,RELB,FBXO41,IL17RC,EXT2,KLF13,HLA-E,CHD4,KLHL36,DGKQ,EXOC7,LAIR1,FBXW4,AGTRAP,ZNF687,LRRC8E,AMOTL2,STK32C
9	transfac_pro-M01610	0.0565199	3.678	M7		ZBTB7B,C2CD2L,SYNPO,KLF13,MXI1,MNT,SDF4,ORAI2,ANKRD9,CHD4,POM121,CBX8
10	yetfasco-606	0.0562869	3.6545	M2		ZBTB7B,CD81,KIAA1522,C2CD2L,ADAM15,MXI1,EFHD2,SRRM2,CCND3,MNT,DEGS2,BAZ2A,SCAMP2,STK32C,C16orf74,SNPH,TMEM102,CBX8,ZFYVE28,MTMR11,SYNPO,TDRD5,VWA1,PCSK6,AGTRAP,LYN,TRIM62,POMGNT1,CHD4,SNX1,S100A16,SEC14L1,ST14,ANKRD9,DUSP23,FOXO1,IQSEC1,MLXIP,S100A13,HPCAL1,RELB
11	jaspar-MA0310.1	0.0558127	3.60668	M7		ZBTB7B,SYNPO,KLF13,C2CD2L,SDF4,ANKRD9,CHD4,PDE4A,POM121,ORAI2,TRPC4AP,LIN7A,LRRC56,MNT
12	yetfasco-543	0.0558127	3.60668	M7		ZBTB7B,SYNPO,KLF13,C2CD2L,SDF4,ANKRD9,CHD4,PDE4A,POM121,ORAI2,TRPC4AP,LIN7A,LRRC56,MNT
13	homer-M01865	0.0557725	3.60263	M8		TMEM102,TWIST2,EHD2,IQSEC1,EPS8L1,TDRD5,ST3GAL4,DEGS2,SYNPO,CBLC,ARHGEF10,NAT8L,KLHL36,ESPN,C16orf74,MSN,TMC4,POM121,ZBTB7B,PDE4A,PHF21A,ZFYVE28,TTC39A,TUBB3,KIAA1522,VWA1,AMOTL2,ANKRD9
14	element o-AGGAGCTG	0.0557163	3.59696	M9		ZBTB7B,SYNPO,TDRD5,NRP1,ST3GAL4,PHF21A,SLC9A1,ANKRD9,IQSEC1,CCND3,SRRM2,YWHAB,RHBDL1,MTMR11,MXI1,ATP9B,ESPN,DEGS2,DUSP8,MLXIP,HPCAL1,SNPH,PDE4A,TMEM102,RPS6KA1,TMEM79,VWA1,C16orf74,LLGL2,S100A14,STK32C,STIM1,ORAI2,ZFYVE28
15	transfac_pro-M00653	0.0555636	3.58156	M7		PDE4A,ZNF593,CBX8,TDRD5,ANKRD9,KLF13,NAT8L,ZBTB7B,SYNPO,SDF4,ORAI2,C2CD2L,ZFYVE28,LRRC56,STK32C

16	transfac_public-M00016	0.0551939	3.54428	M2	ELF2,ELF1,ELF4,ELF5,EHF,ELF3,SPI1,ETV6,ELK3,ELK1,FLI1,GABPA,ETS1,ELK4,ETV4,ETV1,ETV7,GABPB1	CRB3,ZBTB7B,MLLT6,LYPLA2,CYB561D2,EIF3I,SRRM2,RPS6KA1,RND1,TMC4,PTPRB,ZNF611,CST6,CHD4,KLF13,ADAM15
17	hdpi-TCEAL6	0.0550332	3.52807	M2	TCEAL6	ZBTB7B,VWA1,BAZ2A,S100A13,KIAA1522,ZNF296,CYB561D1,DEGS2,SYNPO,RPS6KA1,CD81,CTDSP2,EFHD2,C16orf74,AMOTL2,MNT,PLEKHA2,LYPLA2,ST3GAL4,CCND3,TDRD5,CBX8,S100A16,ZNF687,RHBDD3,MLLT6,C2CD2L,DUSP8,CELSR1,ANKRD9,SNPH,MXI1,ZFYVE28,PCSK6,HPCAL1,LIN7A,YWHAB,SLC1A1,PHF21A,ST14,LLGL2,IQSEC1,MLXIP,CHD4,SLC9A1,NAT8L,SLC44A2,MSN
18	encode-UW.Motif.0469	0.0546635	3.49079	M4		ZBTB7B,DUSP8,S100A13,S100A16,MLLT6,ST3GAL4,VWA1,KIAA1522,RPS6KA1,CYB561D1,LYPLA2,SYNPO,CHD4,ZNF296,PHF21A,MNT,C16orf74,SERPING1,ANKRD9,DEGS2,CCND3,AMOTL2,BAZ2A,ZFYVE28,TDRD5,NAT8L,MMP11,LLGL2,RHBDD3,MXI1,CBX8,STIM1,PNPLA2,EFHD2,TTC39A,PIGV,ESPN,MSN,PCSK6,IQSEC1,NRIP3,LYN,ZNF687,STXBP2,EHD2,TMEM79,CELSR1,KLF13,CST6,C1orf159,PI4KB,ZNF593,TMEM102,HPCAL1,SNPH,TNFRSF18,ST14,PYGO2,TRIM62,KLHL36,C2CD2L,FBXO41,RND1,PQLC2,CD81,SLC9A1,PDE4A,ZNF7,FAM173A,RHBDL1,PTK6,SLC44A2,GABARAPL1,TMEM125,DGKQ
19	encode-UW.Motif.0634	0.054551	3.47944	M10		LYPLA2,CCND3,PDE4A,EFHD2,EHD2,KLHL36,TRIM62,RPS6KA1,SYNPO,KLF13,SNPH,ST14,TDRD5,ESPN,ZBTB7B,LIN7A,ZFYVE28,SLC44A2,SLC9A1,KCTD2,STIM1,TMEM102,KIAA1522,ANKRD9,CTDSP2,C2CD2L,VWA1,SEC14L1,IQSEC1,SERPING1,C16orf74,ST6GALNAC4,TMEM125,FBXW4,CLSTN1,PNPLA2,PI4KB,PLEKHA2,BAZ2A,HPCAL1,STK32C,ST3GAL4,ARHGEF10,CST6,MLLT6,BCKDK,MXI1,TMEM79,NMT2,GTF2A1L,PCSK6,CYB561D1,DDRGK1,UBE2E1,PRCC,SLC22A18,PGLS,SCAMP2,EFEMP1,ORAI2,C1orf159,TWIST2,TPCN1,PIGV,RHBDD3,CHD4,POMGNT1,TTC39A,PTX3,FOXO1,PHF21A
20	transfac_pro-M01820	0.0544144	3.46567	M11	CREM,CREB1,ATF2,ATF3,ATF6,ATF5,ATF1,ATF7,ATF4,CREB3,JUND,JUN,JUNB	PDE4A,ZNF593,ATP9B,DUSP8,CBX8,AMOTL2,MNT,HPCAL1,MXI1,RAB25,ANKRD9,ARHGEF10,GUK1,PSKH1
21	tfdimers-MD00461	0.0543581	3.45999	M10	PURA,TFAP2C,TFAP2A,TFAP2B	DEGS2,CTDSP2,SYNPO,ZNF296,CD81,ZBTB7B,S100A14,IQSEC1,C16orf74,VWA1,CST6,TDRD5,SLC9A1,ST3GAL4
22	transfac_public-M00141	0.0543501	3.45918	M10	IKZF1,ETS1	PDE4A,AMOTL2,LRRC8E,S100A13,KIAA1522,CCND3,ZBTB7B,ZNF687,STIM1,MXI1,C16orf74,TMEM102,RAB25,BAZ2A,TTC39A,SLC9A1,CLSTN1,LLGL2,DEGS2,UBAP1,EFHD2,CHD4,MSN,S100A16,KLF13,C2CD2L,TUBB3,SNPH,HPCAL1,PNPLA2,MLLT6,SLC44A2,SCAMP2,EHD2,SYNPO,PAFAH1B3,UBE2E1,CBX8,PSKH1,POMGNT1,ORAI2,PI4KB,ZNF611,PYGO2,NRP1,IQSEC1,TDRD5,RELB,ST3GAL4,MLXIP,INPP4A,MNT,LYN,RPS6KA1,ZFYVE28,CTDSP2,MMP11,SEC14L1,KLHL36,SLC1A1,NMT2,GTF2A1L,ST14
23	elemento-CAGGTGG	0.0542858	3.4527	M1	ZNF354C,MEIS1	CYB561D1,ANKRD9,TDRD5,C1orf159,ARHGEF10,VWA1,PCSK6,ZBTB7B,C16orf74,STXBP2,TNFRSF18,ZNF687,EHD2,ZFYVE28,KLF13,EFHD2,CTDSP2,RPS6KA1,SYNPO,KIAA1522,CD81,WIPI2,HPCAL1,ESPN,BAZ2A,DEGS2,INPP4A,BCKDK,PHF21A,PQLC2,SNPH,ST3GAL4,ZNF296,CELSR1,LLGL2,SLC9A1,CRB3,PTK6,ST14,AGTRAP,SERPING1,FAM173A,IQSEC1,YWHAB,STIM1,KLHL36,MXI1,S100A16,RAB25,NMT2,LYN,TTC39A,MNT,LPIN1,TMEM125,DDRGK1,NAT8L,SCAMP2,EFEMP1,LRRC42,PDE4A,PTPRB,DUSP8,TMEM102,ORAI2,GTF2A1L,POM121,PAFAH1B3,CST6,SDF4,CLSTN1
24	yetfasco-633	0.0540286	3.42676	M12		ZBTB7B,NDST2,CCND3,PHF21A,S100A13,VWA1,SYNPO,ZNF296,CBLC,TDRD5,CD81,EFHD2,MLLT6,KIAA1522,TMEM102,RPS6KA1,ST3GAL4,IQSEC1,AMOTL2,PDE4A,LYPLA2,MNT,ST14,ARHGEF10,ESPN,ZFYVE28,POMGNT1,SLC1A1,CHD4,MXI1,DEGS2,TUBB3,NRP1,CELSR1,MTMR11,MUC1,STK32C,SNPH,LIN7A,ZNF593,GABARAPL1,C16orf74,PYGO2,LRRC42,STIM1,MSN,CBX8,STXBP2,C1orf159,BAZ2A,PIGV,YWHAB,ZNF687,FBXW4,PI4KB,TTC39A,CTDSP2,TMEM9,DUSP8,BCKDK,HPCAL1,LLGL2,TMEM125,KLF13,LMBR1L,LRRC8E,SERPING1,RND1,PCSK6,SLC9A1
25	transfac_pro-M01865	0.0539242	3.41623	M10	KLF13	BAZ2A,ZBTB7B,KIAA1522,MLLT6,PHF21A,HPCAL1,CCND3,ZNF687,S100A13,SYNPO,ADAM15,CBX8,CHD4,EPS8L1,AMOTL2,RND1,MUC1,CBLC,SLC9A1,SNPH,PIGV,LYN,EFHD2,S100A16,DEGS2,STIM1,MNT,MXI1,GTF2A1L,NRIP3,EHD2,MTMR11,TDRD5,CTDSP2
26	encode-UW.Motif.0074	0.0536992	3.39354	M4		MNT,MUC1,ZFYVE28,SLC9A1,RPS6KA1,SYNPO,LLGL2,TDRD5
27	selexconsensus-esg	0.0530643	3.32951	M1	SNAI2,SNAI1,SNAI3,ID4,LMO2,TCF3,TCF4,MYOG,MYOD1,MYF6,MESP1,ZEB1,GATA4,BHLHE41,ARID5B,MXD3,HAND1,GATA2,USF2,MXD1,MYF5,TAL2,GATA3,TAL1,MYC,MITF,GATA6,MAX,NHLH1,HAND2,MYCN,MXD4,TCF12,GATA5,GATA1,TFEB,MXI1,ASCL1,TCF21,MSC,TCF24,TCF23,ASCL2	CRB3,C1orf159,TMEM125,TDRD5,SYNPO,CTDSP2,ARHGEF10,MNT,DEGS2,ZFYVE28,C16orf74,SLC9A1,HPCAL1,ORAI2,PCSK6,LLGL2,LRRC8E,ZNF687,PDE4A,SNPH,CELSR1,BCKDK,ESPN,KLHL36,WIPI2,CBX8,SCAMP2,KIAA1522,STXBP2
28	homer-M00179	0.0530482	3.32789	M3	SP1,SP2,SP3,SP4,PATZ1,KLF5,KLF16,SP9,SP6,SP7,SP5,KLF7,KLF17	ACSS2,RELB,STK32C,S100A13,HPCAL1,ZNF7,MNT,SNAPC2,IQSEC1,ELMO3,ADAM15,LRRC42,CRB3,TRIM62,LLGL2,STXBP2,PDE4A,ZNF687,ST6GALNAC4,SNX1,TMC4,PQLC2,TMEM102,PHF21A,PNPLA2,NMT2,CASP9,HDAC10,EFNA4,NDST2,EPS8L1,KLHL36,TUBB3,ESPN,LRRC8E,ANKRD9,HSD11B1L,VWA1,ZBTB7B,ZNF296,DEGS2,SLC44A2,PIGV,ORAI2,SYNPO,CYB561D1,CD81,KLF13,DDRGK1,TPCN1,CTDSP2,MTHFR,LRRC56,NRIP3,DUSP8,RPS6KA1,CBX8,PLEKHA2,PAFAH1B3,DGKQ,GABARAPL1,

						SCAMP2,BAZ2A,TDRD5,MMP11,INPP4A,GTF2A1L,STIM1,ZNF611,ST14,HLA-E,LPIN1,MXI1,ZFYVE28,CELSR1,POMGNT1,C2CD2L,BCKDK,TMEM25,C16orf74,FOXO1,TINF2,KIAA1522,EFHD2,CHD4,ATP9B,DUSP23,TMEM79
29	encode-UW.Motif.0012	0.0528312	3.30601	M6		C16orf74,ANKRD9,SCAMP2,ZBTB7B,KIAA1522,SLC44A2,NAT8L,VWA1,ESPN,LRRC8E,TMEM102,ZNF687,PDE4A,PNPLA2,KLF13,PI4KB,IQSEC1,ZFYVE28,SLC9A1
30	yetfasco-663	0.0528071	3.30358	M8		ZBTB7B,EHD2,EPS8L1,TMEM102,DEGS2,ST3GAL4,SYNPO,ESPN,TMC4,VWA1,TWIST2,IQSEC1,ANKRD9,CBLC,NAT8L,TDRD5,RPS6KA1,KLHL36,ZFYVE28,DUSP8,CBX8,KIAA1522,HPCAL1,PDE4A,SLC9A1,CD81,MNT,AMOTL2,C16orf74,C1orf159,ARHGEF10,EFHD2,TRIM62,RHBDD3,TTC39A,POMGNT1,HDAC10,DGKQ,PQLC2,SNPH,S100A13,CTDSP2,TUBB3,ZNF687,CCND3,PCSK6,ZNF593,STK32C,MMP11,ORAI2,MXI1,PIGV,SLC22A18,SEC14L1,LYPLA2,CHD4,PTK6,S100A16,ZNF296,SERPING1,STIM1,ETAA1,SRRM2,CELSR1,ZNF7,YWHAB,POM121,MLXIP,LRRC8E,LIN7A,TNFRSF18,RND1,TSTA3,SLC44A2,CLSTN1,PTPRB,BAZ2A
31	homer-M01764	0.0527187	3.29466	M2		ZBTB7B,CD81,SRRM2,ADAM15,KIAA1522,MNT,DEGS2,STK32C,EFHD2,C2CD2L
32	homer-M01799	0.0527026	3.29304	M12		ZBTB7B,CCND3,NDST2,VWA1,PHF21A,ST3GAL4,CBLC,SYNPO,CD81,RPS6KA1,LYPLA2,IQSEC1,MLLT6,KIAA1522,TDRD5,ZNF296,EFHD2,ST14,ARHGEF10,C1orf159,PDE4A,TMEM102,S100A13,ZNF593,LRRC42,CELSR1,ZNF687,STK32C,AMOTL2,NRP1,ESPN,TUBB3,ZFYVE28,PCSK6,USF1,DEGS2,LRRC8E,MSN,STIM1,CBX8,LLGL2,TTC39A,YWHAB,C16orf74,SNPH,HPCAL1,MNT,CYB561D1,LMBR1L,PTPRB,CTDSP2,MXI1,KLF13,ZNF7,NRIP3,DGKQ,ORAI2,SLC1A1,TMEM125,LYN,LIN7A,CHD4,DUSP8,KLHL36,ANKRD9,TNFRSF18,SLC9A1,BCKDK,AGTRAP,BAZ2A,SERPING1,CRB3
33	element-CACCTGC	0.0526303	3.28575	M1	TCF3	C1orf159,CRB3,SYNPO,PCSK6,TMEM125,DEGS2,ANKRD9,ZFYVE28,ARHGEF10,HPCAL1,DUSP8,KLF13,SLC9A1,STXBP2,DGKQ,CELSR1,ESPN,TUBB3,ST14,C16orf74,CTDSP2,PHF21A,AGTRAP,TDRD5,EFHD2,AMOTL2,CBX8,STK32C,LRRC8E,CYB561D1,CST6,POMGNT1,SLC1A1,SNPH,CD81,SCAMP2,POM121,KIAA1522,KLHL36,EHD2,ZNF687,PDE4A,IQSEC1,MTMR11,SLC22A18,MMP11,DDRGK1,ST3GAL4,LPIN1,VWA1,RND1
34	jaspar-MA0105.1	0.0524133	3.26386	M5	NFKB1,NFKB2,OVOL2,RELA,AP3B1,CHURC1	BAZ2A,TMEM125,SYNPO,MLLT6,LYPLA2,PIGV,FBXW4,PDE4A,ZBTB7B,FBXO41,PI4KB,ZFYVE28,ZNF296,DUSP23,EFHD2,PCSK6
35	encode-UW.Motif.0042	0.0521722	3.23955	M4		TMEM102,TUBB3,SYNPO,PSKH1,C1orf159,EFHD2,DEGS2,ST3GAL4,TDRD5,CST6,RPS6KA1,ZBTB7B,ANKRD9,ZFYVE28,ESPN,LPIN1,PIGV,EHD2,C2CD2L,C16orf74,PHF21A,MLLT6,MLXIP,CELSR1,SCAMP2,ZNF593,TMEM125,SLC9A1,PDE4A,FBXW4,VWA1,LGMN,AGTRAP,NAT8L
36	flyfactorsurvey-sna_SOLEXA_5_FBgn0003448	0.0519311	3.21524	M1	SNAI2,SNAI1,SNAI3,MEIS1,ZNF354C,TCF3,MYOD1,MYF6,MYOG,TCF21,TCF24,TCF23,NEUROD2,NEUROD4,NEUROD6	CRB3,SYNPO,TMEM125,ARHGEF10,C1orf159,ZBTB7B,LLGL2,HPCAL1,WIPI2,ORAI2,ZNF687,PSKH1
37	yetfasco-598	0.0518749	3.20956	M5		EPS8L1,TMC4,C16orf74,ZNF593,CBLC,DUSP8,S100A13,CD81,PNPLA2,ZBTB7B,GUK1,EHD2,VWA1,MXI1,ZFYVE28,STK32C,TUBB3,ST3GAL4,NAT8L,TNFRSF18,EFHD2,LYN,PDE4A,ACSS2,SYNPO,ORAI2,LLGL2,DEGS2,MSN,SERPING1
38	transfac_public-M00177	0.0517945	3.20146	M11	CREB1,ATF4,CREM,ATF2,ATF3,ATF7,ATF6,ATF5,ATF1	PDE4A,ZNF593,CBX8,DUSP8,MXI1,AMOTL2,UBAP1,ZNF687,MNT,ATP9B,GUK1,ANKRD9,MUC1
39	transfac_pro-M01995	0.0517463	3.1966	M7		ZBTB7B,CBX8,PDE4A,TDRD5,ANKRD9,SDF4,NAT8L,SYNPO,TRPC4AP,KLF13,ORAI2
40	transfac_pro-M01207	0.0516499	3.18687	M2	ETS2,GATA6,GATA4,GATA3,GATA2,ELK1,GATA5,GATA1,ERG,ELF2,ERF,ELF1,ELK4,ETV7,ETS1,FLI1	BAZ2A,SYNPO,EFHD2,KIAA1522,CCND3,ZBTB7B,CD81,HPCAL1,VWA1,PCSK6,AMOTL2,RPS6KA1,S100A16,POMGNT1,CYB561D1,C16orf74,LYPLA2,CHD4,MNT,C2CD2L,TDRD5,SNPH,DEGS2,SLC9A1,CBX8,ST3GAL4,S100A13,MSN,TTC39A,SRRM2,NRIP3,IQSEC1,NRP1,CTDSP2,INPP4A,ZNF593,ST14,MXI1,KLHL36,SLC44A2,ZNF687,ADAM15,RHBDD3,TPCN1,LLGL2,ANKRD9,AGTRAP,LYN,PIGV,PQLC2,DUSP23,ORAI2,UBE2E1,TRIM62,ZFYVE28,SCAMP2,FAM173A,MLXIP,PHF21A,PLEKHA2,MTMR11,CELSR1,ESPN,SEC14L1,TMC4,STXBP2,MLLT6,MUC1,PTK6,RND1,DUSP8,RELB,FAM43A,C1orf159,FBXW4,FOXO1,DDRGK1,USF1
41	transfac_pro-M02088	0.0515775	3.17958	M1	TCF3,TCF12,MYOG,MYF6,MYOD1,ATOH1,LMO2,ASCL2	CYB561D1,CRB3,ANKRD9,CST6,SYNPO,ZFYVE28,C16orf74,C1orf159,TDRD5,BAZ2A,DEGS2
42	taipale-GATGACGTCATC-XBP1-DBD	0.0514088	3.16256	M11	XBP1,JDP2,JUNB,JUND,JUN,BATF3,ATF7,CREB3,ATF1,ATF3,CREM,CREB1,ATF2,ATF4,NPDC1	ZNF593,DUSP8,CBX8,PDE4A,RAB25,AMOTL2,MXI1,MNT,ATP9B,ZNF687
43	encode-UW.Motif.0580	0.0508543	3.10664	M4		ZNF587,S100A13,S100A14,ESPN,ZFYVE28,DEGS2,PAFAH1B3,GTF2A1L,PIGV,S100A16,ZNF593,ZBTB7B,SYNPO,AGTRAP,ANKRD9,STIM1,ST3GAL4,EFHD2,NRP1,HPCAL1,IQSEC1,TDRD5,NRIP3,LLGL2,ARHGEF10,ST14,CELSR1,KIAA1522,ORAI2,PRCC,STXBP2,ZNF687,RND1,SLC44A2,AMOTL2,PDE4A,UBE2E1,PHF21A,MNT,C16orf74,MLXIP,CHD4,EHD2,FBXW4,MXI1,TCN2,KLF13,CTDSP2,FOXO1
44	transfac_pro-M00973	0.05079	3.10015	M1	MYF6,ASCL1,MYOD1,TCF3,TCF4,MYF5,ARID5B,MYOG,TCF12,ID4,SREBF1,SREBF2,NR3C1,MXD1,TFEB,MXI1,MAX,MXD3,TAL1,TAL2,HAND1,MYC,BHLHE41,MXD4,MITF,HAND2,TP53,NHLH1,USF2,MYC	C1orf159,CRB3,ANKRD9,DEGS2,IQSEC1,CYB561D1,SYNPO,ZFYVE28,C16orf74,C2CD2L,CST6,TUBB3,ARHGEF10,CTDSP2,TMEM125,RPS6KA1,TDRD5,PCSK6,ESPN,SCAMP2,DUSP8,ZNF7,EFHD2,SNPH,WIPI2,ZBTB7B,MNT,SLC9A1,VWA1,KLHL36,DGKQ,FAM173A,PDE4A,HPCAL1,BAZ2A,ST3GAL4,TNFRSF18,PLEKHA2,TTC39A,LLGL2,S100A16,MLLT6,STXBP2,CD81,EXT2,CELSR1,LYN,ZNF687,TRPC4AP,ORAI2,NMT2,ZNF593,KIAA1522,ST14,NRIP3,AMOTL2,INPP4A

					N,SNAI2,GATA1,GATA5,GATA4,GATA6,GATA2,GATA3,ZNF146,ZNF260	
45	flyfactorsurvey-l_1_sc_da_SANGER_5_FBgn0002561	0.0507819	3.09934	M1	ASCL2,ASCL1,TCF3	CRB3,C1orf159,TMEM125,SYNPO,PCSK6,DEGS2,ARHGEF10,HPCAL1,CTDSP2,STXBP2,ANKRD9,ST14,ZFYVE28,TDRD5,ESPN,MUC1,ZNF593,KLHL36,DUSP8,KLF13,LRRC8E,SLC9A1,ZNF687,TUBB3,EFHD2,CELSR1,PHF21A,STK32C,CBX8,MNT,AMOTL2,C16orf74,EHD2,CD81,SCAMP2,IQSEC1,SLC1A1,SNPH,PSKH1,DGKQ,POMGNT1,SLC22A18,POM121,GTF2A1L,TMEM79,C2CD2L,WIPI2,DDRGK1,FAM173A,CST6,STIM1,RND1,LPIN1,NRP1,PDE4A,MTMR11,MLLT6,ZBTB7B,ZNF7,ORAI2,MMP11,ST3GAL4,LRRC42,PNPLA2,VWA1,AGTRAP,YWHAB,LAIR1,TRIM62,NRIP3,PLEKHA2,FBXW4,UBE2E1,TNFRSF18
46	swissregulon-MAFB.p2	0.0507418	3.09529	M11	MAFB	SRRM2,MXI1,FOXO1,PDE4A,TUBB3,PHF21A,TDRD5,CD81,SYNPO
47	transfac_pro-M03831	0.0507337	3.09448	M1	MYF6,MYOG,ASCL2	CRB3,C1orf159,ANKRD9,IQSEC1,SYNPO,DUSP8,TUBB3,DEGS2,ZFYVE28,PCSK6,SNPH,ARHGEF10,C16orf74,SLC9A1,CYB561D1,TMEM125,TDRD5,HPCAL1,EFHD2,ESPN,AGTRAP,KLHL36,CELSR1,SCAMP2,LRRC42,ZNF7,TNFRSF18,DGKQ,VWA1,CD81,TMEM79,STK32C,CTDSP2,SLC1A1,NAT8L,AMOTL2,KLF13,CST6,FAM173A,ZNF593,KIAA1522,PHF21A,ST3GAL4,ZBTB7B,STXBP2,TXNRD2,LLGL2,CBX8,SLC44A2,PDE4A,LPIN1,PNPLA2,ZNF687,ORAI2,EXT2,SLC22A18,MLLT6,RPS6KA1,TRPC4AP,ST14,FBXW4,EHD2,SDF4,MXI1,TTC39A,MTMR11,MMP11,STIM1,BCKDK,MNT,LRRC8E,INPP4A,PTK6
48	transfac_pro-M01186	0.0506855	3.08962	M11	CREM,CREB1,XBP1,ATF4,JUND,JUNB,JUN,ATF1,ATF3,ATF7,ATF6,ATF2,ATF5,JDP2,CREB3,BATF3,NPDC1	ZNF593,PDE4A,DUSP8,AMOTL2,MXI1,RAB25,ATP9B,ZNF687,CBX8
49	yetfasco-1995	0.050581	3.07908	M4	JAZF1,FOXN4,FOXN2,FOXN3,FOXN1,FOXH1	TMEM102,C1orf159,CLSTN1,ZBTB7B,ST3GAL4,EPS8L1,TDRD5,FBXO41,EHD2,ZNF593,IQSEC1,HPCAL1,PCSK6,TNFRSF18,CHD4
50	homer-M00215	0.0504123	3.06206	M8	ZNF711,ZFX	TMEM102,ZBTB7B,IQSEC1,POMGNT1,C16orf74,ARHGEF10,ANKRD9,STK32C,MNT,DEGS2,SYNPO,CHD4,KIAA1522,KLHL36,ZFYVE28,TRIM62,PHF21A,VWA1,NAT8L,ST3GAL4,ZNF296,RPS6KA1,C1orf159,TDRD5,TTC39A,EHD2,DUSP8,SERPING1,KLF13,TUBB3,PRCC,LYPLA2,AGTRAP,CTDSP2,ESPN,CD81,MMP11,ZNF687,GUK1
51	jaspar-MA0128.1	0.0503882	3.05963	M7		ZBTB7B,KLF13,ORAI2,SDF4,SYNPO,C1orf159,C2CD2L,ZFYVE28,CHD4,ANKRD9
52	transfac_public-M00277	0.050364	3.0572	M1	LMO2,ZEB1,TCF4,TCF3,MESP1,MYOG,MYOD1,MYF6,ID4,TCF12	CYB561D1,C1orf159,IQSEC1,CRB3,TDRD5,SYNPO,ANKRD9,CST6,C16orf74,DEGS2
53	encode-UW.Motif.0131	0.0503399	3.05477	M4		ZBTB7B,SYNPO,DEGS2,ANKRD9,MNT,MSN,TMC4,ZFYVE28,NAT8L,EFHD2,ORAI2,ARHGEF10,ST3GAL4,C16orf74,SLC9A1,TDRD5,INPP4A,ESPN,FBXW4,CBLC,EPS8L1,HPCAL1,NMT2,RND1,IQSEC1,DUSP8,RPS6KA1,STIM1
54	transfac_pro-M01778	0.0503319	3.05396	M4	PLAG1	ZBTB7B,SYNPO,S100A13,NDST2,RPS6KA1,ST3GAL4,LLGL2,VWA1,TDRD5,DUSP8,AMOTL2,CD81
55	iDMMPMM-SNA	0.0502274	3.04342	M1	SNAI2,SNAI1,SNAI3,MYOD1,SCRT2,TCF3,MYOG,MYF6,SCRT1,TCF21,TCF24,TCF23	CRB3,ZBTB7B,SYNPO,TMEM125,C1orf159,TMEM79,ARHGEF10,DEGS2,RPS6KA1,HPCAL1,TNFRSF18,SLC9A1,ANKRD9,TDRD5,ST14,UBE2E1,ORAI2,EFHD2,FAM173A,KLF13,NRIP3,GUK1,PCSK6,STK32C,DUSP8,STIM1,CTDSP2,PHF21A,KLHL36,PLEKHA2,CD81,RAB25,NRP1,LLGL2,POMGNT1,LRRC8E,TFPI,ZFYVE28,PSKH1,TTC39A,CBX8,ZNF593,MNT,SNPH,GTF2A1L,SCAMP2,IQSEC1,S100A16,ESPN,ZNF687,TMEM102,EHD2
56	jaspar-MA0080.1	0.0502194	3.04261	M2	SPI1	BAZ2A,ZBTB7B,LYPLA2,CCND3,ST14,EFHD2,SRRM2,KIAA1522,MLLT6,TMEM102,MNT,CD81,PCSK6,SNPH,S100A13,AMOTL2,CBX8,MSN,YWHAB,ADAM15,MLXIP,HPCAL1,SYNPO,ACSS2,MXI1,MTMR11,NRP1,AGTRAP,RPS6KA1,S100A16,LYN,TRIM62,C16orf74,DUSP23,VWA1,CYB561D1,DEGS2,PLEKHA2,SNX1,C1orf159,CHD4,ZNF687,C2CD2L,PIGV,SLC9A1,CTDSP2,SLC44A2,ZFYVE28,TDRD5
57	transfac_pro-M00712	0.0502194	3.04261	M1	MYOG,TCF3,MYOD1,MYF6,ASCL1,ARID5B,MYF5,TCF12,TCF4,ASCL2	IQSEC1,C1orf159,ANKRD9,SYNPO,CYB561D1,DEGS2,VWA1,ESPN,C16orf74,KIAA1522,TUBB3,CRB3,CST6,TDRD5,ZFYVE28,ARHGEF10,MLLT6,ST3GAL4,PCSK6,DUSP8,DGKQ,CTDSP2,RPS6KA1,C2CD2L,EFHD2,CD81,ZBTB7B,PDE4A,AMOTL2
58	yetfasco-1036	0.0502033	3.04099	M12		RPS6KA1,ST3GAL4,SYNPO,CD81,VWA1,CCND3,ZBTB7B,LYPLA2,ST14,IQSEC1,ZFYVE28,MTMR11,LRRC42,KIAA1522,TDRD5,DUSP8,BCKDK,ARHGEF10,ORAI2,ZNF687,USF1,TUBB3,CRB3,AMOTL2,LLGL2,C1orf159,NDST2,HPCAL1,CELSR1,PTPRB,DGKQ,C16orf74,KLF13,MXI1,TMEM102,PHF21A,ZNF296,PLEKHA2,LYN,CTDSP2,SLC9A1,ESPN,ZNF7,CBX8,YWHAB,STIM1,SEC14L1,BAZ2A,DEGS2,SLC1A1,KLHL36,FBXW4,CYB561D1,UBE2E1,S100A13,MSN,TTC39A,STK32C,ZNF593,MLLT6,TMEM25,EHD2,ANKRD9,POM121,LIN7A,PDE4A,TRIM62,SNPH,ETAA1
59	swissregulon-SP1.p2	0.0501069	3.03127	M3	SP1,SP3,SP2,SP4,PATZ1,KLF5,ZBTB14,KLF4,KLF16,PAX5,EGR1,SP9,SP6,SP7,SP5,KLF17,KLF7,SMAD3,ZNF410	RELB,ACSS2,ZBTB7B,PHF21A,BAZ2A,STK32C,ZNF687,TRIM62,MNT,CRB3,KIAA1522,ELMO3,LLGL2,HPCAL1,ST6GALNAC4,ZNF7,TMC4,LRRC42,KLF13,SNAPC2,IQSEC1,ADAM15,PIGV,ANKRD9,CHD4,MUC1,TDRD5,SNX1,NMT2,S100A13,TPCN1,TUBB3,CASP9,MXI1,SYNPO,CD81,EFNA4,DUSP8,KLHL36,STXBP2,CELSR1,PQLC2,ORAI2,EPS8L1,CBX8,TMEM102,PDE4A,LYN,PNPLA2,HSD11B1L,DDRGK1,VWA1,MTMR11,FOXO1,NRIP3,BCKDK,EHD2,ST3GAL4,ESPN,LRRC8E,C2CD2L,TMEM25,MLLT6,PAFAH1B3,AMOTL2,TMEM79,C16orf74,EFHD2,NDST2,FAM43A,SLC44A2,UBE2E1,CCND3,PTPRB,PCSK6,RPS6KA1,PLEKHA2,DEGS2,HLA-E

60	jaspar-MA0305.1	0.0498417	3.00452	M2		ZBTB7B,CD81,KIAA1522,ADAM15,SRRM2,MNT
Rank	Track id	AUC	NES	ClusterCode	Transcription factor	Target genes
1	wgEncodeSydhTfbsHepg2JundIggrabPk.narrowPeak.gz	0.0638009	3.89125	T1	JUND	PAFAH1B3,LMBR1L,SYNPO,UBAP1,DUSP8,DEGS2,ZNF593,MTMR11,ATP9B,AMOTL2,MXI1,CTDSP2,HPCAL1,ZNF611,RAB25,GABARAPL1,ST6GALNAC4,TTC39A,TDRD5,PLEKHA2,MLXIP
2	wgEncodeHaibTfbsEcc1Tcf12V0422111PkRep1.broadPeak.gz	0.0623142	3.75719	T2	TCF12	AMOTL2,ESPN,NT5E,PTK6,ZBTB7B,SLC22A18,KIAA1522,KLF13,SLC44A2,LLGL2,CELSR1,CST6,MTHFR,MNT,LRRC56,SYNPO,C16orf74,MLLT6,RHBDD3,FAM173A,CTDSP2
3	wgEncodeSydhTfbsHelas3Mxi1af4185IggrabPk.narrowPeak.gz	0.0605622	3.59923	T3	MXI1	FOXO1,ANKRD9,POM121,C1orf159,NRP1,TUBB3,CLSTN1,NAT8L,DGKQ,VWA1,PLEKHA2,ORAI2,ZFYVE28,LRRC56,KCTD2,CTDSP2,DUSP8,MUC1,C16orf74,ESPN,GUK1,HERC5,HDAC10,CHD4,WIPI2,ZNF296,F12,MLXIP,TXNRD2,MMP11,KLF13,ZNF593,MNT,MXI1,NMT2,STK32C,EXT2,ARHGEF10,LRRC42,SDF4,FAM173A,BAZ2A,IQSEC1,TPCN1
4	wgEncodeSydhTfbsHelas3Znf143IggrabPk.narrowPeak.gz	0.060064	3.5543	T4	ZNF143	CLSTN1,LPIN1,C1orf159,ZFYVE28,TUBB3,NAT8L,TSTA3,BAZ2A,KLHL36,FBXO41,POM121,DUSP8,ANKRD9,FOXO1,CYB561D2,BCKDK,SRRM2,LLGL2,FAM189B,TMEM102,HSD11B1L
5	wgEncodeHaibTfbsEcc1Zbtb7aV0422111PkRep1.broadPeak.gz	0.0599274	3.54198	T5	ZBTB7A	ZBTB7B,ARHGEF10,DUSP8,PDE4A,SLC44A2,KIAA1522,MLLT6,BAZ2A,MXI1,PLEKHA2,RPS6KA1,CTDSP2,CELSR1,VWA1,CBLC,ANKRD9,LLGL2,C16orf74,ORAI2,LRRC56,CRB3,HPCAL1,EFHD2,INPP4A,UBE2E1,C1orf159,CD81,ESPN,TNFRSF18,FOXO1,HDAC10,MMP11,PNPLA2,MUC1,AMOTL2,F12,PTK6,KLF13,BCKDK,PTPRB,EFNA4,NAT8L,PYGO2,HMGCL,KLHL36,CHD4,DDRGK1,SLC22A18,MTHFR,EXOC7,CBX8,ST3GAL4,PRSS8,TUBB3,SEC14L1,AGTRAP,RHBDD3,TMEM25,CLSTN1,MNT,KCTD2,TRIM62,SCAMP2,WIPI2,SCAMP3,FAM189B,S100A13,DEGS2,ZNF593,LRRC8E,ZNF7,ZFYVE28,STK32C,LIN7A,ZNF296,TRPC4AP,ADAM15,TMEM9,C2CD2L,NDST2,LRRC42,NRP1,FBXO6,DGKQ,SLC9A1
6	wgEncodeSydhTfbsHelas3Zkscan1hpa006672IggrabPk.narrowPeak.gz	0.0581031	3.3775	T6	ZKSCAN1	DGKQ,LYN,DUSP8,FOXO1,RPS6KA1,FAM189B,PNPLA2,ST3GAL4,POM121,IQSEC1,DDRGK1,TUBB3,EFHD2,S100A14,SYNPO,KLF13,EXT2,ZFYVE28,ZNF593,MSN,GUK1,C1orf159,MLLT6,C16orf74,NRP1,CBX8,HLA-E,HPCAL1,LLGL2,C2CD2L,FAM173A,SNX1,STK32C,TSTA3,ARHGEF10,WIPI2,RELB,EHD2,PYGO2,VWA1,ANKRD9,PLEKHA2,HSD11B1L,SEC14L1,SDF4,F12,CTDSP2,NENF,BAZ2A,LYPLA2,ESPN,ATP9B,TMEM79,MNT,FBXW4,TXNRD2,CELSR1,KCTD2,PAFAH1B3,S100A16,FBXO41,HERC5,ZBTB7B,STXBP2,LPIN1,MLXIP,MXI1,UBE2E1,ADAM15,AMOTL2,TMEM9,IL17RC,PQLC2,ST6GALNAC4,SLC22A18,MTHFR,CYB561D1,SNAPC2,FBXO6,NDST2,CHD4,ST14,LRRC56,TTC39A,LMBR1L,STIM1,USF1
7	wgEncodeSydhTfbsGm12878MaxIggmusPk.narrowPeak.gz	0.0578138	3.35141	T7	MAX	ST6GALNAC4,FAM43A,RPS6KA1,ANKRD9,MSN,DGKQ,POM121,MTMR11,TNFRSF18,CBX8,LYN,HSD11B1L,INPP4A,C4orf3,CD81,MUC1,UBE2E1,ZNF593,EXT2
8	wgEncodeSydhTfbsK562Tblr1ab24550IggrabPk.narrowPeak.gz	0.0559413	3.18258	T8	TBL1XR1	CD81,TMC4,HDAC10,PAFAH1B3,ANKRD9,LRRC8E,POM121,LRRC42,MNT,UBE2E1,NAT8L,ECM1,PDE4A,TMEM9,TUBB3,SNPH,FOXO1,LRRC56,UBAP1,NDST2,YWHAB,VWA1,PNPLA2,DGKQ,STK32C,TSTA3,GNA15,ADAM15,FAM173A,BCKDK,KLF13,WIPI2,RELB,MXI1,MMP11,C1orf159,HPCAL1,PHF21A,GUK1,FBXO6,RPS6KA1,ZBTB7B,ESPN,GABARAPL1,SYNPO,NENF,KCTD2,PSKH1,EPS8L1,PQLC2,STXBP2,ATP9B,LMBR1L,LPIN1,EFHD2,IL17RC,CBX8,MLLT6,KIAA1522,CLSTN1,ARHGEF10,LLGL2,TTC39A,MLXIP,HLA-E,CTDSP2,PROS1,NMT2,IQSEC1,CHD4,ZNF593,SLC44A2,TINF2,HSD11B1L,DUSP8,LYN,BAZ2A,FAM43A,ORAI2,TMEM25,ST6GALNAC4,ZFYVE28
9	wgEncodeSydhTfbsHelas3Mazab85725Iggrab	0.0546153	3.06302	T9	MAZ	LYPLA2,LRRC42,MMP11,C1orf159,CBX8,ELMO3,POM121,FOXO1,DGKQ,EXT2,NAT8L,SLC22A18,RHBDL1,TUBB3,HDAC10,MNT,ORAI2,KLF13,CD81,ESPN,TNFRSF18,GUK1,DUSP8,LLGL2,LRRC56,ANKRD9,ARHGEF10,STK32C,CHD4,CST6,BAZ2A,FAM43A,ZFYVE28,TWIST2,TSTA3,NMT2,ZBTB7B,TRIM62,WIPI2,TTC39A,MXI1,CTDSP2,NDST2,CELSR1,PHF21A,SEC14L1,CASP9,F12,SDF4,PYGO2,VWA1,PAFAH1B3,MLLT6,CYB561D1,KIAA1522,YWHAB,C16orf74,FAM189B,SYNPO,PQLC2,PLEKHA2,ATP9B,ZNF593,LMBR1L,MLXIP,MUC1,AMOTL2,EFHD2,FAM173A,IL17RC,S100A14,GNA15,

	Pk.narrowPeak.gz					KLHL36,HERC5,PRSS8,STIM1,TXNRD2,PNPLA2,TMC4,ZNF687,EHD2,IQSEC1,BCKDK,MTHFR,UBAP1,CLSTN1,RPS6KA1,FBXW4,HPCAL1,KCTD2,ZNF296,LYN

Supplemental Table 4 – GSEA Creighton_AKT1_Signlaling_Via_MTOR_DN for ZR751 & TCGA

geneSet	size	overlap	enrichment Ratio	pValue	FDR	userId
XIE_ST_HSC_S1PR3_OE_UP	180	25	5.276019691	1.07E-11	3.62E-08	AVPR1A;C1orf116;DDX58;DDX60;GATA3;GBP2;HDAC9;HERC5;HERC6;HMCN1;IL24;OAS1;OAS2;PARP9;PLSCR1;RDH10;RSAD2;SAMD9;SMAD6;SP110;ARHGAP20;CD276;IFIH1;MYL9;RTP4
GOZGIT_ESR1_TARGETS_DN	720	52	2.743530239	4.11E-11	6.92E-08	ADCY9;CACNG4;CASK;CD109;COL4A5;CTHRC1;DEGS2;EFEMP1;EPB41L1;GABARAPL1;GATA3;GBP2;GULP1;HDAC9;HERC6;HMCN1;ITGB8;LDB3;LIN7A;NELL2;NF2;NNT;NPAS3;OAS2;PKIB;RAB30;RARA;RGS22;SAMD9;SDK1;SH3BGRL;SLC1A1;SLC7A2;SP110;SYT1;TFPI;TIMP2;TMTC1;TNFRSF11B;TTC39A;ZNRF1;BAG3;CYP2C8;IFITM10;JAG1;LYPD1;MVP;NAB1;PCDHB16;PCDHB7;PTTG1IP;TCF4
NUYTTEN_EZH2_TARGETS_UP	1015	60	2.245557149	3.37E-09	3.78E-06	ADIPOR2;ANPEP;APOL6;APP;BAIAP2;C1orf116;CD83;CITED4;CYB5R1;DDX58;DDX60;DENND3;EFEMP1;ERAP1;FLVCR2;FOXO1;GBP2;GBP3;GBP4;GJA1;GULP1;HERC5;HERC6;HIVEP3;IL15RA;IL24;ITGB5;LARP6;NT5E;OAS1;PAM;PARP9;PKIB;PLSCR1;PSKH1;QDPR;QSOX1;RHOC;SAMD9;SNX21;ST3GAL6;TK2;TMEM45A;TNIK;TNKS1BP1;TTC7B;WNT4;BLCAP;CLN5;CPQ;FDX1;IFIH1;JAZF1;MVP;MXRA7;PBX3;PLEKHA3;PLXNA1;RTP4;TCF4
MOSERLE_IFNA_RESPONSE	32	10	11.8710443	5.58E-09	4.70E-06	DDX58;DDX60;HERC5;OAS1;OAS2;PLSCR1;RSAD2;SAMD9;IFIH1;RTP4
BLANCO_MELO_HUMAN_PARAINFLUENZA_VIRUS_3_INFECTION_A594_CELLS_UP	194	21	4.112031841	4.55E-08	3.06E-05	APOL6;DDX58;DDX60;ERAP1;GBP3;GBP4;HERC5;HERC6;IL15RA;NLRC5;NT5E;OAS1;OAS2;PARP9;PLSCR1;RSAD2;SAMD9;SP110;BMPER;IFIH1;RTP4
LIU_PROSTATE_CANCER_DN	474	33	2.644688351	4.09E-07	2.30E-04	ADCY9;AKR1B1;CLIP4;EFEMP1;GATA3;GBP2;GJA1;GPC6;GPR161;INPP1;KLHL29;KRT19;LDHB;MRC2;MYH11;NELL2;NHSL2;NNT;PGF;SGPP2;SH3BGRL;SLC14A1;SMOC1;TIMP2;TNS1;FCHSD2;JAZF1;MXRA7;MYL9;PDGFD;PRRT2;SH3PXD2B;SLCO3A1
PAPASPYRIDONOS_UNSTABLE_ATEROSCLEROTIC_PLAQUE_DN	41	9	8.338684779	9.32E-07	4.33E-04	ADCY9;EGFR;GJA1;MYH11;PAM;PKIG;SUSD5;TNS1;MXRA7
GRAESSMANN_APOPTOSIS_BY_DOXORUBICIN_UP	1139	58	1.934386148	1.03E-06	4.33E-04	AEN;AKR1B1;BAIAP2;CABYR;CITED4;CROT;DDX60;ECM1;FAH;FBXW4;GATA3;GBP2;GFOD2;GGA2;IKBKE;IL24;LCMT1;LDB3;LIF;MMD;OAS1;OAS2;PGF;PHLDA3;PRCP;PYGO2;QSOX1;RALGPS1;RDH10;RNF135;RSAD2;S100A13;SP110;SPR;ST14;TK2;TSPAN33;UBE2E1;UBE2F;UNC5C;ZNF622;ZNRF1;ACSL1;ATOX1;BAG1;BLCAP;CARHSP1;CDK18;HS6ST1;HYAL1;IFIH1;IFITM10;LPIN1;NCOA1;PLEKHA3;RTP4;SLC45A3;TCIRG1
LEE_BMP2_TARGETS_UP	751	43	2.175040874	1.54E-06	5.77E-04	ANPEP;BMP7;C14orf132;CITED4;CTDSP2;DEGS2;EFEMP1;EXTL3;FOXO1;GABARAPL1;GBP2;GBP4;HDAC11;ITGB5;ITGB8;KCND2;KRT19;LARP6;LDHB;LIX1L;LTF;NENF;PAM;PBXIP1;PIGR;PLSCR1;PLSCR4;PTPRB;RND2;SGPP2;SLC7A2;SNX21;ST14;SV2B;TIMP2;CKB;CPQ;CYP39A1;FAM124A;IFIH1;JAG1;RGMA;SERINC2
DAUER_STAT3_TARGETS_DN	48	9	7.122626582	3.79E-06	0.00127576	DDX58;DDX60;HERC5;HERC6;OAS1;OAS2;SAMD9;SP110;IFIH1
TAKEDA_TARGETS_OF_NUP98_HOXA9_FUSION_3D_UP	177	17	3.648501752	4.58E-06	0.00140068	DDX58;DDX60;GBP3;HERC5;HERC6;HLX;ITGB8;OAS1;OAS2;PARP9;PLSCR1;RSAD2;SAMD9;SP110;IFIH1;MVP;PBX3
HECKER_IFNB1_TARGETS	92	12	4.954870666	5.19E-06	0.00145516	DDX58;DDX60;HERC5;HERC6;OAS1;OAS2;PARP9;PLSCR1;RSAD2;SAMD9;MYL9;RTP4
BLANCO_MELO_COVID19_SARS_COV_2_INFECTION_CALU3_CELLS_UP	318	23	2.747512141	1.29E-05	0.00313539	APOL6;DDX58;DDX60;EDN2;GBP4;HDAC9;HERC5;IL15RA;IRS2;KCNV1;NLRC5;NR4A3;OAS1;OAS2;PARP9;RSAD2;SAMD9;SP110;SYT1;TNFRSF11B;IFIH1;KIF6;RTP4
MEISSNER_BRAIN_HCP_WITH_H3K4ME3_AND_H3K27ME3	1066	52	1.853041062	1.30E-05	0.00313539	ACE;ADAP2;AVPR1A;BMP7;CACNG4;CBLN1;CBLN2;CCND1;CD83;DPP10;EGFR;EMX1;EPB41L1;ETNK2;EYA1;FLVCR2;GULP1;HLX;HMCN1;IGFBP7;ITGB5;LIN7A;LRP5;MOB3B;NRG3;PGF;PKIB;PLXNA4;PTPRB;RHOC;ROR2;SLC7A2;SMAD6;SMOC1;SMOC2;SOWAHB;SOX13;ST14;SULF2;TEAD4;WNT4;WNT9A;BAG3;CARHSP1;CDK18;CISH;EBF1;GLIS3;JAG1;LYPD1;PDGFD;SERINC2
BLANCO_MELO_RESPIRATORY_SYNCYTIAL_VIRUS_INFECTION_A594_CELLS_UP	279	21	2.859262284	1.69E-05	0.00379468	ANPEP;APOL6;DDX58;DDX60;GBP2;GBP3;GBP4;HERC5;HERC6;HIVEP3;IL15RA;NLRC5;NT5E;OAS1;PARP9;PGF;RSAD2;SAMD4A;SAMD9;SP110;IFIH1
SMID_BREAST_CANCER_BASAL_DN	677	37	2.076117645	2.36E-05	0.00496847	ADCY9;BCAS4;C14orf132;CACNG1;CACNG4;CCND1;COL4A5;CROT;DLG5;DNAJC1;ECM1;FAH;GATA3;GJA1;HDAC11;KRT19;LIN7A;MCCC2;MYH11;NELL2;POLD4;QDPR;RALGPS1;RARA;SH3BGRL;SLC1A1;SSH3;SYT1;TK2;TNIK;TTC39A;KIF5C;NADSYN1;PDGFD;PTPRT;SNX1;SPATA7
REN_ALVEOLAR_RHABDOMYOSARCOMA_DN	407	26	2.426709794	3.14E-05	0.00622588	AP2M1;CCND1;ECM1;EFEMP1;EGFR;GABARAPL1;GJA1;GULP1;IGFBP7;MRC2;NRG1;NT5E;PAM;QSOX1;RHOC;S100A13;SMAD7;TIMP2;TMEM45A;TNFRSF11B;TNFRSF1A;TUSC3;ATOX1;BAG3;MXRA7;XYLT1
ENK_UV_RESPONSE_EPIDERMIS_DN	511	30	2.230176621	3.87E-05	0.00724778	ADAP2;CCND1;CD83;DNAJC1;EML1;FAM117A;GATA3;GBP2;GNAI1;IRS2;NCOR2;NELL2;PBXIP1;PKIG;PLXNA2;PPP3CA;QDPR;SH3BGRL;SMAD7;ST14;TIMP2;WNT4;ACSL1;CKB;CYP39A1;NCOA1;PBX3;SLCO3A1;TCF4;UVRAG
WONG_ADULT_TISSUE_STEM_MODULE	712	37	1.9740613	6.83E-05	0.012068	ADCY9;APP;ARHGEF5;CCND1;CTSZ;GABARAPL1;GATA3;GBP2;GJA1;HDAC11;HLX;IGFBP7;ITGB5;KCND2;PBXIP1;PIGR;PLXNA2;RASL10A;S100A13;SLC1A1;SLC41A1;SLC9B2;SMAD7;SMOC2;TNFRSF11B;UNC5C;BAG3;CISH;CPQ;IFIH1;LPIN1;MYL9;PBX3;PTTG1IP;SLCO3A1;TCF4;ZRSR2
WANG_SMARCE1_TARGETS_UP	285	20	2.66577837	7.17E-05	0.012068	AVIL;CASK;CD109;CTDSP2;CTHRC1;FOXO1;GABRA2;GATA3;GULP1;HMCN1;ITGB5;ITGB8;JDP2;NLGN4X;PLSCR4;SULF2;TIMP2;TSHZ2;JAG1;MYL9

TCGA_GLIOBLASTOMA_COPY_NUMBER_UP	70	9	4.884086799	8.80E-05	0.01411679	AVIL;CTDSP2;EGFR;ETNK2;HS1BP3;LRRN2;PIK3C2B;SAMD9;SOX13
SMID_BREAST_CANCER_LUMINAL_B_UP	164	14	3.242821859	1.14E-04	0.01749811	CACNG4;CCND1;CGA;DNAJC1;GATA3;HDAC11;IL24;MCCC2;NELL2;QDPR;SLC1A1;SYT1;KIF5C;PTPRT
SANA_RESPONSE_TO_IFNG_UP	73	9	4.683370903	1.23E-04	0.01794724	DDX60;GBP3;GBP4;NLRC5;OAS1;OAS2;PARP9;PPP3CA;IFIH1
GRAESSMANN_RESPONSE_TO_MC_AND_DOXORUBICIN_UP	604	32	2.012574399	1.49E-04	0.02095121	AKR1B1;BAIAP2;CABYR;CITED4;CROT;FAH;FBXW4;GATA3;GGA2;IKBKE;LDB3;LIF;MMD;OAS1;PGF;PHLDA3;PRCP;PYGO2;RALGPS1;RNF135;S100A13;SP110;ST14;TSPAN33;ZNRF1;ATOX1;CARHSP1;CDK18;HS6ST1;IFITM10;LPIN1;NCOA1
BLANCO_MELO_BRONCHIAL_EPITHELIAL_CELLS_INFLUENZA_A_DEL_NS1_INFECTION_UP	614	32	1.979796314	2.01E-04	0.0265596	APOL6;AVIL;CARD16;CD83;DDX58;DENND3;EDN2;GATA3;GBP2;GBP3;HERC5;HERC6;IKBKE;IL15RA;LIF;LYSMD2;NLRC5;OAS1;OAS2;PARP9;PKIB;PLSCR1;PLSCR4;RDH10;RSAD2;SOWAHB;SOX13;SP110;WNT4;IFIH1;RTP4;STOML1
BASSO_HAIRY_CELL_LEUKEMIA_UP	78	9	4.38315482	2.05E-04	0.0265596	ADCY9;CCND1;CYB5R1;S100A13;SUSD5;SYT1;TNFRSF1A;CPVL;PTTG1IP
SMID_BREAST_CANCER_RELAPSE_IN_BONE_UP	96	10	3.957014768	2.17E-04	0.02704058	GATA3;IL24;LIN7A;MB;NELL2;PAH;SLC1A1;SYT1;KIF5C;PTPRT
ZHANG_RESPONSE_TO_IKK_INHIBITOR_AND_TNF_UP	221	16	2.750214789	2.62E-04	0.03121863	CD83;CITED4;CLIP4;DDX58;GBP3;HDAC9;IGFL1;ITGB8;KDM2A;LIF;SAMD4A;SGPP2;SLC7A2;CLN5;IFIH1;SLCO3A1
SMIRNOV_CIRCULATING_ENDOTHELIOCYTES_IN_CANCER_UP	157	13	3.145448682	2.69E-04	0.03121863	ADAP2;ANPEP;NRG1;QSOX1;RARA;SLC2A3;SMAD7;TIMP2;TNFRSF1A;TNS1;ACSL1;CKB;SCML1
BROWNE_INTERFERON_RESPONSIVE_GENES	65	8	4.675365141	2.94E-04	0.03223629	GBP2;IL15RA;OAS1;OAS2;PLSCR1;RSAD2;SP110;KIF5C
BENPORATH_ES_WITH_H3K27ME3	1085	48	1.680545995	2.97E-04	0.03223629	ADAP2;C11orf45;C14orf132;CBLN1;CLIP4;COL4A5;DLX4;DOK6;GABRA2;GATA3;GNAO1;GSC;HLX;HOXA4;HOXD1;KCNV1;KL;LHX4;LYSMD2;MESP1;NKX2-5;NR4A3;NRG1;PLXNA2;POLR3GL;RARA;RASL10A;RNPEPL1;SGPP2;SMOC2;SORCS3;SV2B;UNC5C;WNT3A;ARHGAP20;CYP39A1;DACH2;EBF1;FFAR4;HPSE2;MXRA7;PITX1;PTPRT;RBBP7;SKAP1;SLC6A20;SLCO3A1;XYLT1
GRYDER_PAX3FOXO1_ENHANCERS_INTADS	998	45	1.712856092	3.06E-04	0.03223629	ADCY9;AP1B1;APP;ATP9B;B4GALNT3;BAIAP2;BFAR;CASK;COX6C;CTDSP2;DIS3L;DLG5;EXTL3;EYA1;FOXO1;GGA2;GNA12;HS1BP3;ITGB8;KLHL29;NCOR2;NRG1;PCCB;PGF;PLCE1;PLXNA2;PRCP;RAE1;RRAGA;SAMD4A;SDK1;SLC25A26;SLC7A2;SULF2;TFPI;ZNF622;NCOA1;NOSIP;PLCG1;PLXNA1;PTTG1IP;RGMA;RNF216;SSB;STOML1
TAKEDA_TARGETS_OF_NUP98_HOXA9_FUSION_10D_UP	184	14	2.890341222	3.79E-04	0.03867024	APOL6;DDX58;DDX60;GNAI1;HERC5;OAS1;OAS2;RSAD2;SAMD9;SP110;TMEM45A;CKB;IFIH1;RTP4
OUELLET_CULTURED_OVARIAN_CANCER_INVASIVE_VS_LMP_UP	68	8	4.469099032	4.02E-04	0.0398514	KRT19;LDHB;MMD;RAE1;TNFRSF1A;NDUFS8;RBBP7;SALL2
VANTVEER_BREAST_CANCER_ESR1_UP	145	12	3.143780009	4.61E-04	0.04430031	COX6C;CYB5R1;GATA3;HDAC11;ITPK1;MCCC2;QDPR;RARA;BAG1;CISH;PTPRT;SNX1
BLALOCK_ALZHEIMERS_DISEASE_UP	1634	65	1.511124367	5.00E-04	0.04499025	ADAP1;AP1B1;APP;AVPR1A;BFAR;CASK;COG4;COL4A5;CTDSP2;DDX27;DLG5;DLX4;EFEMP1;EGFR;ENGASE;ERAP1;FBXW4;FOXO1;GBP2;GJA1;GNA12;HERC5;HS1BP3;INPP1;IRS2;ITGB5;ITGB8;ITPK1;KDM2A;MB;MRC2;NCOR2;NPAS3;NR4A3;PAH;PBXIP1;PHLDA3;PIK3C2B;PLCE1;PRCP;PSD4;PSKH1;RALBP1;RALGPS1;RNPEPL1;SLC14A1;SOX13;SP110;SSH3;TNFRSF11B;TNFRSF1A;ACSL1;BAG1;CPQ;CYP39A1;LPIN1;NOSIP;PTTG1IP;RNF216;SALL2;SLCO3A1;SNX11;UVRAG;ZNF133;ZRSR2
HEIDENBLAD_AMPLICON_8Q24_DN	39	6	5.844206426	5.03E-04	0.04499025	GPC1;HIVEP3;RAE1;SLC14A1;HPS1;NAB1
GRYDER_PAX3FOXO1_ENHANCERS_KO_DOWN	432	24	2.110407876	5.08E-04	0.04499025	AP1B1;APP;BFAR;CASK;CTDSP2;DIS3L;DLG5;EXTL3;EYA1;HS1BP3;ITGB8;NRG1;PCCB;PGF;PLCE1;RRAGA;SDK1;SLC25A26;SLC7A2;SULF2;PLCG1;PTTG1IP;RGMA;RNF216
SEITZ_NEOPLASTIC_TRANSFORMATION_BY_8P_DELETION_UP	72	8	4.220815752	5.95E-04	0.05133561	DDX58;DDX60;HERC5;HERC6;OAS1;OAS2;SAMD9;IFIH1
CHIARADONNA_NEOPLASTIC_TRANSFORMATION_CDC25_DN	151	12	3.018861598	6.63E-04	0.0557991	BAIAP2;CYB5R1;GALK1;GPC1;NAA38;PHLDA3;POLD4;RNPEPL1;S100A13;TEAD4;TMEM45A;CARHSP1
HORIUCHI_WTAP_TARGETS_UP	289	18	2.365993605	6.85E-04	0.056272	CD109;CROT;CYB5R1;ERAP1;HMCN1;INPP1;KDM2A;PIK3C2B;PRCP;RSAD2;SLC2A3;TFPI;TMEM45A;CPQ;GLIS3;IFIH1;JAG1;MYL9
DER_IFN_ALPHA_RESPONSE_UP	74	8	4.106739651	7.15E-04	0.05734347	OAS1;OAS2;PLSCR1;PPP3CA;RHOC;SP110;TEAD4;BAG1
KORKOLA_EMBRYONAL_CARCINOMA_UP	42	6	5.42676311	7.57E-04	0.05925993	GABARAPL1;LDHB;NANOG;SLC2A3;TEAD4;TNFRSF1A
JINESH_BLEBBISHIELD_VS_LIVE_CONTROL_DN	269	17	2.400687026	8.14E-04	0.06230841	ACOT4;C1orf116;GBP2;GBP4;HERC5;HOXD1;IKBKE;NNT;NT5E;OAS1;OAS2;PAM;SAMD9;SLC7A2;SMAD6;SP110;SALL2
KORKOLA_SEMINOMA_UP	43	6	5.300559317	8.60E-04	0.06436675	ATN1;GABARAPL1;LDHB;NANOG;SLC2A3;TEAD4
MYLLYKANGAS_AMPLIFICATION_HOT_SPOT_29	8	3	14.24525316	0.00091975	0.06710238	EGFR;ETV1;JAZF1
ISSAEVA_MLL2_TARGETS	60	7	4.43185654	9.70E-04	0.06710238	DENND3;HDAC9;INSL4;NT5E;OAS1;SMAD6;SLCO3A1
FORTSCHEGGER_PHF8_TARGETS_DN	759	35	1.751721953	9.71E-04	0.06710238	ACTR1A;ADCY9;BAIAP2;CASK;DDX58;DLG5;EML1;EPB41L1;GBP3;GNA12;GULP1;HS1BP3;IGFBP7;ITGB8;ITPK1;KRT19;LRP12;MOB3B;NF2;NT5E;OBSCN;PPOX;RAE1;RASL10A;RRAGA;SAMD9;SSH3;TMEM139;BLCAP;CLN5;HYAL1;IFIH1;MVP;PBX3;PLEKHA3
KINSEY_TARGETS_OF_EWSR1_FLII_FUSION_DN	323	19	2.234549516	9.77E-04	0.06710238	BAIAP2;EYA1;FOXO1;GNAI1;GULP1;HMCN1;IGFBP7;NCOR2;NT5E;PBXIP1;QSOX1;SMAD6;SULF2;TFPI;TIMP2;ZFYVE1;BLCAP;EBF1;MVP

TAVOR_CEBPA_TARGETS_DN	31	5	6.126990608	0.001194 34	0.07958409	AKR1B1;FOXO1;GJA1;ITGB5;OAS2

Supplemental Table 5 – Module 1 C2_CGP WebGestalt ORA

geneSet	size	overlap	enrichmentRatio	pValue	FDR	userId
JOHNSTONE_PARVB_TARGETS_3_DN	833	82	2.49647803	1.40E-14	4.71E-11	ALG10;ASPH;ATAD2;BARD1;BLZF1;BRIP1;BROX;C1GALT1;C6orf52;CDC42SE2;CDCA7L;CENPE;CENPF;CEP128;CEP76;CETN3;CMC2;CPNE8;CRNDE;CSE1L;CSNK1A1;DTL;ELL2;FANCL;FH;FUBP1;G2E3;GTF2H3;GTPBP4;HMGB1;IMMP1L;IVNS1ABP;KANSL1L;KBTBD6;KLHL20;LSM5;MAD2L1;MDM1;MDM4;MGP;MLLT10;MPHOSPH6;MRPS9;MTBP;MYBL1;MZT1;NAA50;NINJ2;NUCKS1;PAWR;POC1B;PPP4R2;PTK2;RPE;SEH1L;SIX1;SLC7A1;SMCHD1;SNX16;STARD3NL;TCF12;TFAM;TIPRL;TMPO;TSHZ1;UAP1;ZNF670;AP1S2;ARHGAP11A;CTSC;EIF4E;HERPUD2;KIF20B;LSM6;MBD2;MBNL1;SLC25A24;SMIM15;STYK1;TAX1BP1;TCEA1;TSEN15
RODRIGUES_THYROID_CARCINOMA_POORLY_DIFFERENTIATED_UP	619	63	2.58112358	4.78E-12	8.05E-09	ABCE1;AP1AR;ATAD2;BBIP1;CACYBP;CBX3;CD58;CENPF;CMC2;COA1;CRNDE;CSE1L;CSNK1A1;CSNK1G3;DHX9;DNM1L;DTL;EIF4G2;FH;GTPBP4;HEATR1;IDE;IGF2BP3;IL1RAP;INTS7;KIAA0586;LACTB2;LIN9;MAD2L1;MIPOL1;MRPS9;MYBL1;NMD3;NUP155;PMAIP1;PNPT1;RALGPS2;RPAP3;RPE;RPRD1A;SEH1L;STAM;TFAM;TIPRL;TMEM161B;TMPO;TPR;UCHL5;WDR3;ZC3H15;ZC3H7A;AIMP1;CTSC;DPY19L1;KIF20B;KIF5B;MBNL1;ME2;OSBPL3;RBBP8;SNAPC1;TSEN15;ZEB1
RODRIGUES_THYROID_CARCINOMA_ANAPLASTIC_UP	692	65	2.38213384	7.16E-11	8.04E-08	API5;ATAD2;ATP2B1;BLZF1;BTAF1;C3orf80;CBX3;CCT2;CD58;CDK7;CDKN2A;CENPE;CENPF;CHML;CSE1L;CSNK1A1;CSNK1G3;DHX9;DNM1L;DPM1;EIF4G2;EREG;FAM126A;FIGN;FN1;G2E3;GTPBP4;IGF2BP3;IL1RAP;KCNMA1;LCORL;LGALS8;LIN9;MAD2L1;MIOS;MIPOL1;MYBL1;NAA50;NMD3;PGM3;PMAIP1;PNPT1;RALGPS2;RBM34;RPAP3;RPE;SIX4;SLC1A4;SLC7A1;STK17A;TBL1XR1;TMOD3;TMPO;TPR;VTA1;XPOT;AIMP1;ARHGAP11A;FBXO38;KIF20B;KIF5B;MBNL1;RBBP8;RFX3;TOP1
DACOSTA_UV_RESPONSE_VIA_ERCC3_DN	847	71	2.12585596	1.44E-09	1.21E-06	ABCE1;AFF1;APPBP2;ATP2B1;BARD1;BTAF1;C2CD5;CENPE;CENPF;CSE1L;CSNK1A1;CTBP2;DUSP4;EFNA5;ELL2;EPS8;FUBP1;HOMER1;IGF2BP3;IL1RAP;KDM4C;KLF6;KLHL20;LSM5;MSH3;MYBL1;NAV3;NVL;PAIP1;PAWR;PDLIM5;PPFIA1;PPP2R5A;PPP2R5E;PTK2;PTPN2;RAB21;RALA;RNF2;SFMBT1;SLC7A1;SMCHD1;STAM;STK24;SWAP70;SYNE2;TCF12;TERF1;TGIF1;TIPRL;TLE1;TP53BP2;TRIM37;UBE2D2;WDR37;AMPH;ARHGAP11A;ARID5B;EIF4E;FGF5;HIF1A;KIF20B;MBD2;MBNL1;ME2;PRKCA;SLC25A24;TOP1;USP7;VLDLR;ZEB1
GEORGES_TARGETS_OF_MIR192_AND_MIR215	874	72	2.08919973	2.31E-09	1.55E-06	ALG10;ASPH;ATAD2;ATF1;BARD1;BCL2;BRIP1;CDKN2A;CENPE;CENPF;CEP128;CGNL1;CHML;COBLL1;DEGS1;DTL;ELL2;EPS8;ERCC4;FAM126A;FBN1;FUBP1;G2E3;GABPB2;GPR19;HADH;HAS3;HOXA13;ID2;IL1RAP;LIN9;MAD2L1;MDM4;MIPOL1;MTSS1;MZT1;NAA50;NUCKS1;PAWR;PGM3;POU2F1;RAB23;RAD54B;RPAP3;SFR1;SIX4;SLC19A2;SLC1A4;SP4;SPTBN1;STAM;TICRR;TMEM170B;TMPO;TMTC2;TTPA;UBE2D1;UBE2D2;VASH2;ZNF704;AP1S2;ARHGAP11A;CAMK4;FHDC1;GINM1;KIF20B;KIF5B;NHSL1;OSBPL6;PERP;PTS;TOP1
JOHNSTONE_PARVB_TARGETS_2_DN	330	38	2.9203073	3.31E-09	1.86E-06	ADM;ADO;ASPH;ATF1;AXIN2;C1GALT1;CHML;CPNE8;CSNK1A1;CTBP2;DERA;EFNA5;EMP1;FAM126A;FIGN;FUBP1;GCNT2;GTF2A1;HADH;IPO5;LIN7C;PLEKHA5;PPP2R5E;PTK2;RIOK1;SPRY1;STK17A;TBL1XR1;TMPO;TOMM20;UCHL5;ISOC1;LEMD3;MBNL1;PRKCA;RAP2B;SLC44A1;SRSF1
HAMAI_APOPTOSIS_VIA_TRAIL_UP	647	58	2.27343536	4.58E-09	2.20E-06	AP1AR;ATAD2;BHLHE41;BLZF1;C1GALT1;CBX3;CCDC59;CDK5RAP2;CENPE;CETN3;DNM1L;EIF3E;EMP1;ESRRG;FBLN5;G0S2;GPR19;HMGB1;HSPH1;IPO5;IQGAP2;LACTB2;LGALS8;MSH3;MYBL1;NID1;NMD3;PAIP1;PALLD;PPFIA1;PPP4R2;RAD54B;RALA;RPAP3;SMCHD1;SNX16;SSPN;SYNE2;TTLL7;UCHL5;YEATS4;AIMP1;AP1S2;ARHGAP11A;DOCK10;FBXL3;GLIPR1;KIF20B;KIF5B;PPIL4;RBBP8;SNAPC1;SPINK1;ST6GALNAC1;TAX1BP1;TCEA1;VLDLR;ZNF397
SENGUPTA_NASOPHARYNGEAL_CARCINOMA_WITH_LMP1_UP	391	41	2.65929181	1.24E-08	5.21E-06	ADCY10;ANO1;BCL2;BLVRA;CALB1;CBX3;CEP76;CFH;DHX9;ELL2;ESRRG;F5;G2E3;HNMT;IGF2BP3;KRAS;LACTB2;LIFR;LSM5;MDM2;MEOX2;PMAIP1;RPE;RPRD1A;SEH1L;SERPINI1;SLC12A2;SMCHD1;TBL1XR1;TC2N;TFAM;TPR;VASH2;WDR3;ZC3H11A;ZNF678;DYRK2;MBNL1;SLCO1A2;SPINK1;TOP1
ZHENG_BOUND_BY_FOXP3	485	46	2.40533178	3.59E-08	1.34E-05	AHR;ARHGAP15;ATP2B1;BCL2;C2CD5;CC2D2A;CDC42SE2;CLEC2D;COBLL1;COL11A1;DAPL1;FBLN5;GABRG1;ICA1;KLF6;LSM5;MDM2;NINJ2;NR2E3;NSMCE2;NT5C3A;PDE7A;PELI1;POLI;RAB3IP;SAMSN1;SPTBN1;ST8SIA1;STK24;SYNE2;TCF12;TOX;TSPAN13;USP3;WDR37;ZNF608;AMPH;ARID5B;CAMK4;DOCK10;HERPUD2;HIF1A;MBNL1;NR5A2;RFX3;TMEM67
MIYAGAWA_TARGETS_OF_EWSR1_ETS_FUSIONS_UP	258	30	2.94890272	1.21E-07	4.08E-05	ATP1A1;BHLHE41;CD58;ETV6;FAT3;FBLN5;FGFR1;G0S2;GALNT3;HOXA13;HSD17B2;ID2;IL1RAP;LGALS8;LRIG3;NID1;OPN3;RALGPS2;SIX1;SLC1A4;SPTBN1;SSPN;SYNE2;TBL1XR1;TCF12;TSPAN13;ZDHHC21;ARHGDIB;TM4SF1;ZSWIM6
DODD_NASOPHARYNGEAL_CARCINOMA_DN	1341	91	1.72096291	2.12E-07	6.48E-05	ABCE1;ADO;AIDA;ATAD2;BRIP1;C1orf131;CACYBP;CBX3;CCDC59;CCT2;CENPE;CENPF;COA1;COL6A3;CSE1L;DEGS1;DNM1L;DTL;ETV6;FANCL;FBN1;FN1;FUBP1;GATA6;GTF2H3;GTPBP4;HEATR1;ID2;INTS7;IPO5;KCNMA1;LIN9;LSM5;MAD2L1;MARK1;MBTPS2;MRPL32;MRPL9;NAA50;NID1;NMD3;NUCKS1;NUP155;PAIP1;PAWR;PMAIP1;PNPT1;POP1;RAD54B;RBM34;RFC3;RPE;RPRD1A;SEH1L;SLCO5A1;STARD3NL;STRAP;TARBP1;TCF12;TDO2;TFAM;TICRR;TIPRL;TLE1;TMPO;UCHL5;VASH2;WBP11;WDR3;XPOT;ZC3H11A;ZC3H15;ZNF124;ZNF678;ZNRF3;AIMP1;ARHGAP11A;CTSC;DPY19L1;KIF20B;LSM6;ME2;MED21;PANK1;RBBP8;SLC44A1;SPARC;SRSF1;TGIF2;TSEN15;ZNF697
SENESE_HDAC3_TARGETS_UP	472	43	2.31039031	3.10E-07	8.69E-05	AKR1C3;AP1AR;ASPH;ATP2B1;BTG2;CALB1;DHX9;EMP1;FAM126A;FAT3;FN1;GNA13;HEATR1;HNMT;IGF2BP3;IL1RAP;IQGAP2;KCNMA1;KLF6;KLRC1;MIER3;NAV3;OPN3;PMAIP1;RPAP3;SAA1;SLC7A1;SPTBN1;TFAM;TIMP3;TMOD3;TPR;UEVLD;ZC3H11A;ANTXR2;B4GALT1;CEP120;MBNL1;RFX3;SERPINE2;TM4SF1;TOP1;ZNF697
ENK_UV_RESPONSE_KERATINOCYTE_DN	474	43	2.30064183	3.48E-07	9.01E-05	ABCE1;AFF1;ASH2L;CBX3;CCNG2;CD58;CETN3;CLTC;CSE1L;CSNK1A1;DHX9;DUSP4;FH;GCSH;GNPAT;IGF2BP3;IPO5;KRAS;LBR;LYST;MAD2L1;MPHOSPH6;PAWR;PPP2R5E;RAPGEF5;RNF2;SCFD1;SVIL;TARBP1;TERF1;TGDS;TOMM20;TPR;UGDH;DYRK2;ENOSF1;MBD2;MBNL1;ME2;RBBP8;SERPINE2;SRSF1;USP7

MILI_PSEUDOPODIA_HAPTOTAXIS_UP	512	45	2.22895577	4.53E-07	1.07E-04	AP4S1;BBIP1;C9orf85;CACYBP;CETN3;COMMD3-BMI1;CSNK1G3;CTBP2;DNM1L;EIF3E;EIF4G2;GTPBP4;IVNS1ABP;KRAS;NAA50;PAIP1;PALLD;PPFIA1;PPP4R2;PTPN2;RALA;RBM34;RPE;SMCHD1;SNX16;STRAP;TCF12;TXNL1;UBA3;UBE2D2;UCHL5;ZNRF2;DTNA;EIF2A;HIF1A;KIF5B;MBNL1;RBBP8;SLTM;SMIM15;SRSF1;TAX1BP1;TCEA1;TRIP4;ZEB1
GRYDER_PAX3FOXO1_ENHANCERS_IN_TADS	998	72	1.8296198	4.79E-07	1.07E-04	ADO;ARHGAP15;BBIP1;BCL2;BTAF1;C1GALT1;C1orf131;CAAP1;CASQ2;CEP128;CGNL1;CSE1L;DHX9;EEF2K;EFNA5;FGF7;FUBP1;GLUL;GNA13;GNPAT;HACL1;HSPH1;IL1RAP;KCNMA1;LRIG3;M6PR;MBTPS2;MIOS;MSH3;PPP4R2;RALA;RBM20;RFC3;RIMKLB;RXRG;SETBP1;SLC7A1;SPOCK3;SPRY1;SRP9;STIM2;STK24;SVIL;TBL1XR1;TCF12;THSD4;TLE1;TNNT2;TOX3;TRIM45;TTLL7;UBE2E2;VASH2;WNT5B;ZDHHC21;ZFP36L1;ANTXR2;ATF6;DOCK10;DYRK2;KIAA1614;KIF5B;MSX2;PDHB;PPIL4;PTS;RBBP8;SERPINE2;SNN;SYNE3;TOP1;VLDLR
ZHANG_BREAST_CANCER_PROGENITORS_UP	426	39	2.32174172	9.77E-07	2.06E-04	ABHD13;AHR;APPBP2;ATAD2;BBIP1;BCL2;C1orf131;CD38;CENPF;CETN3;COMMD3-BMI1;EDARADD;EPS8;FH;IMMP1L;KANSL1L;LCORL;LSM5;NUP155;NVL;RPE;RPRD2;SLC7A1;SMCHD1;STRAP;TMPO;TRIM37;TSPAN13;UBA3;UBE2D2;UCHL5;UGDH;WDR3;PANK1;RBBP8;RFX3;SMIM15;TAX1BP1;TCEA1
DIAZ_CHRONIC_MEYLOGENOUS_LEUKEMIA_UP	1378	90	1.65635029	1.29E-06	2.56E-04	ABCE1;AHR;AKR1C3;ALDH1A1;API5;ASH2L;ATP1A1;CACYBP;CBX3;CCT2;CD36;CDK7;CETN3;CFH;CSDE1;CSNK1A1;CSNK1G3;DEGS1;DHX9;DNM1L;DPM1;EIF4G2;FH;GCSH;GTPBP4;HMGB1;HSPH1;IDE;IL1RAP;IPO5;IQGAP2;KLHDC2;KRAS;LBR;LIN7C;LSM5;MAD2L1;MAT2B;MDM1;MLLT10;NMD3;PAIP1;PDCD6IP;PDLIM5;PIP5K1B;PPP2R5A;PPP2R5E;PPP4R1;PRPS1;RAB4A;RALA;RFC3;RTN4;SAMSN1;SCFD1;STAM;STK24;STRAP;TCF12;TERF1;TMOD3;TOMM20;TP53BP2;TXNL1;UBE2D2;UCHL5;WBP11;WDR3;XPOT;YEATS4;ZFP36L1;AIMP1;ARCN1;EIF4E;ENOSF1;KCMF1;MED21;MGST2;PAPSS1;PDHB;PPM1D;RAB38;RAP2B;RBBP8;SERPINE2;SRSF1;TAX1BP1;TCEA1;TOP1;VAV3
DACOSTA_UV_RESPONSE_VIA_ERCC3_COMMON_DN	459	40	2.21007088	2.45E-06	4.37E-04	AFF1;APPBP2;ATP2B1;BARD1;C2CD5;CENPE;CENPF;CTBP2;EPS8;FUBP1;HOMER1;IGF2BP3;IL1RAP;KLF6;MYBL1;NPIPB5;NVL;PAIP1;PAWR;PPFIA1;PPP2R5E;PTK2;PTPN2;SLC7A1;SMCHD1;STAM;STK24;SWAP70;TCF12;TLE1;TP53BP2;TRIM37;UBE2D2;ARID5B;EIF4E;FGF5;KIF20B;MBD2;MBNL1;PRKCA
TONKS_TARGETS_OF_RUNX1_RUNX1T1_FUSION_HSC_UP	180	22	3.09962441	2.54E-06	4.37E-04	AHR;ALDH1A1;BAMBI;BBIP1;CA2;CCNG2;DEPTOR;EMP1;EPS8;HLF;KLF6;NID1;PALLD;PDLIM5;SAMSN1;SPTBN1;THSD7A;TOX;ARG2;ARID5B;SPARC;TM4SF1
KOINUMA_TARGETS_OF_SMAD2_OR_SMAD3	817	60	1.86246487	2.60E-06	4.37E-04	ADM;AHR;ANO1;AP3S1;ATAD2;BLZF1;BTG2;CBX3;CLTC;COBLL1;DMTF1;DTL;EDN1;ELL2;EMP1;FUBP1;INTS7;IRF2BP2;IVNS1ABP;KLF6;LGALS8;MAL2;MIOS;NUCKS1;PAWR;PKP2;PRR15;PTK2;RAB4A;RTN4;RTTN;SERPINB8;SETBP1;SFR1;SLC7A1;SPTBN1;STK17A;SVIL;SYNE2;TBL1XR1;TGIF1;TLE1;TMOD3;UBE2D1;ZFP36L1;ANTXR2;AP1S2;ARHGDIB;EIF4E;HERPUD2;KCTD1;MBNL1;OSBPL3;PERP;RAB38;RAP2B;SLC25A24;TAX1BP1;TCEA1;TM4SF1
KINSEY_TARGETS_OF_EWSR1_FLII_FUSION_UP	1257	82	1.65438838	4.22E-06	6.77E-04	ATAD2;ATF1;BARD1;BRIP1;CACYBP;CCDC117;CCT2;CDC42SE2;CENPF;CLGN;CRNDE;DTL;DUS4L;ERCC4;FH;FUBP1;G2E3;GTF2A1;HADH;HMGB1;HSD17B2;HSPH1;IL1RAP;IMMP1L;LIN9;MAB21L3;MAD2L1;MAT2B;METTL4;MYBL1;MZT1;NAA50;NSG2;NUCKS1;NUP155;PAWR;PKP2;PMAIP1;POC1B;POP1;PPFIA1;PPP2R5E;PRPS1;RAB23;RAB3IP;RAD54B;RALGPS2;RFC3;RPE;RPRD1A;RTTN;SEC11C;SEH1L;SLC1A4;SLC25A21;SNAPC5;SOX2;TICRR;TIMP3;TMEM19;TMPO;UAP1;UBE3D;UCHL5;WBP11;YEATS4;ZNF496;ZNRF3;AIMP1;ARG2;CDS1;CTSC;EIF2A;HSF2;ISOC1;KIF20B;KIF5B;ME2;OSBPL3;SRSF1;TCEA1;TSEN15
FISCHER_DREAM_TARGETS	911	64	1.78164221	5.16E-06	7.90E-04	ABCE1;ALG10;ATAD2;BARD1;BRIP1;CACYBP;CBX3;CDCA7L;CDK5RAP2;CENPE;CENPF;CMC2;COX20;CSE1L;DTL;FANCL;FIGN;G2E3;GABPB2;GPR19;HADH;HEATR1;HMGB1;INTS7;LBR;LCORL;LIN9;LSM5;MAD2L1;MDM1;METTL4;MSH3;MTBP;MYBL1;MZT1;NT5C3A;NUCKS1;NUP155;PMAIP1;POU2F1;RAD54B;RFC3;RFT1;RNF2;RTTN;SFR1;SLC1A4;SMCHD1;SP4;TICRR;TMPO;TRIM37;TRIM45;UCHL5;WBP11;YEATS4;ARHGAP11A;DCLRE1A;EIF4E;KIF20B;SRSF1;TCEA1;TOP1;TSEN15
TURASHVILI_BREAST_DUCTAL_CARCINOMA_VS_LOBULAR_NORMAL_UP	74	13	4.45523411	5.67E-06	8.30E-04	ATP2B1;COL11A1;DTL;FBN1;FN1;GDPD1;INTS7;MLLT10;RALGPS2;RBM34;UBE2D1;UBE2D2;KCMF1
NIKOLSKY_BREAST_CANCER_12Q13_Q21_AMPLICON	44	10	5.7637644	6.27E-06	8.80E-04	CCT2;MDM1;MDM2;RAB3IP;RAP1B;SLC35E3;TMEM19;XPOT;YEATS4;DYRK2
GILDEA_METASTASIS	29	8	6.99601748	1.15E-05	0.00155505	AKR1C3;DUSP4;FN1;HNMT;TIMP3;ARHGDIB;GLIPR1;SPARC
RICKMAN_TUMOR_DIFFERENTIATED_WELL_VS_POORLY_UP	234	24	2.60108342	1.90E-05	0.00246656	ALDH1A1;C5;CACYBP;CENPF;CGNL1;FUBP1;GATAD2B;GCNT2;IRF2BP2;LBR;LIFR;MDM4;MSL2;NUCKS1;SMCHD1;TARBP1;TERF1;TMPO;TOMM20;ZBED5;ZNF496;ZNF704;SRSF1;TGIF2
DUTERTRE_ESTRADIOL_RESPONSE_6HR_UP	221	23	2.63933465	2.24E-05	0.00278972	ABCE1;BCL2;BRIP1;DEPTOR;DTL;JAK2;MYBL1;OPN3;RFC3;SEH1L;SLC19A2;SLC1A4;SLC7A1;SVIL;TICRR;WDR3;HIF1A;ISOC1;PGR;PHLDA1;PLEKHH1;RBBP8;SLC25A24
TAKEDA_TARGETS_OF_NUP98_HOXA9_FUSION_10D_UP	184	20	2.75658298	4.07E-05	0.00489482	ALDH1A1;CALB1;CD1D;CD38;DUSP4;FN1;PALLD;SCN2A;SLC8A1;TEX9;TIMP3;TOX;ADRB1;ARG2;ASB2;OSBPL6;SERPINE2;STS;TM4SF1;ZEB1
MARTINEZ_RESPONSE_TO_TRABECTEDIN_DN	267	25	2.37458459	5.93E-05	0.00688909	ATF1;BLZF1;BTAF1;CDK7;DTL;EPS8;GABPB2;GGPS1;HSPH1;LBR;PPP2R5E;RAD54B;RAP1B;STK17A;TARBP1;TBL1XR1;TGIF1;TMPO;UBE2D2;USP3;WNT5B;XBP1;FBXL3;TOP1;TRIP4
PUJANA_BRCA1_PCC_NETWORK	1570	92	1.48609671	7.22E-05	0.007931	ABCE1;AP3S1;API5;ASH2L;BARD1;BTAF1;CBX3;CCT2;CD38;CDK7;CENPE;CENPF;CETN3;COA1;CSE1L;DHX9;DPM1;EIF3E;FANCL;FH;FUBP1;GCSH;GNA13;GNPAT;GTF2A1;GTF2H3;HINT1;HSPH1;IGF2BP3;IPO5;KRAS;LBR;LSM5;M6PR;MAD2L1;MDM1;MDM2;MDM4;MLLT10;MPHOSPH6;NUP155;PMAIP1;PPP2R5E;PRPS1;PTPN2;RAP1B;RBM34;RFC3;RIMS1;RNF2;RPL35;SCFD1;SLC1A4;SLC6A7;SLC7A1;SNA

						PC5;SRP9;STK24;SWAP70;TACR3;TARBP1;TFAM;TGDS;TIPRL;TMEM183A;TMPO;TPR;TXNL1;UBE2D2;XPOT;ZC3H15;AIMP1;ARCN1;ARHGAP11A;ARHGDIB;GLIPR1;HSF2;KIF20B;LSM6;MBD2;MBNL1;ME2;MED21;OSBPL3;PDHB;PPM1D;RBBP8;SNAPC1;SRSF1;TCEA1;TGIF2;TOP1
NELSON_RESPONSE_TO_ANDROGEN_UP	81	12	3.7571205	7.51E-05	0.007931	APPBP2;ELL2;ID2;IQGAP2;LIFR;PDLIM5;PGM3;RAB4A;UAP1;UGDH;B4GALT1;HOMER2
TURASHVILI_BREAST_DUCTAL_CARCINOMA_VS_DUCTAL_NORMAL_UP	47	9	4.85627809	7.77E-05	0.007931	CENPF;CLHC1;COL11A1;DTL;FBN1;FN1;GDPD1;MBTPS2;UBE2D1
MILI_PSEUDOPODIA	47	9	4.85627809	7.77E-05	0.007931	BBIP1;COMMD3-BMI1;EIF3E;NAA50;PALLD;PPFIA1;EIF2A;KIF5B;ZEB1
KARLSSON_TGFB1_TARGETS_UP	121	15	3.14387149	8.33E-05	0.00824798	ABCE1;CACYBP;CCT2;GCSH;GTPBP4;HSPH1;IPO5;KLF6;MPHOSPH6;PTK2;SLC8A1;TOMM20;ZC3H15;DPY19L1;SLTM
BENPORATH_CYCLING_GENES	629	45	1.81434873	8.80E-05	0.00846684	ASIP;ATAD2;BARD1;CBX3;CDCA7L;CDK7;CENPE;CENPF;CHML;DMTF1;DTL;DUSP4;G2E3;GCSH;HOXB4;INTS7;IVNS1ABP;KIAA0586;KLF6;KRAS;LBR;MAD2L1;MDM1;MDM2;NUCKS1;PPP6R3;PRR16;RAB23;TGIF1;TMPO;TRIM45;VTA1;ZBED5;ARHGAP11A;ARHGDIB;EIF4E;ENOSF1;HERPUD2;HIF1A;HSF2;KIF20B;KIF5B;OSBPL6;RBBP8;TOP1
GROSS_HYPOXIA_VIA_ELK3_DN	152	17	2.8363788	1.06E-04	0.009922	ADM;AFF1;ATF1;DUSP4;EDN1;KLF6;PDLIM5;RAD54B;RAMP3;RIOK1;SAMSN1;TGIF1;TSHZ1;TXNL1;XBP1;TM4SF1;VLDLR
IGARASHI_ATF4_TARGETS_DN	97	13	3.39883839	1.10E-04	0.01000781	AGR2;AKR1C3;ASPH;DDC;DDHD2;OPN3;PRR15;SLC1A4;TMTC2;TOMM20;TRIM37;APOH;CDS1
BOQUEST_STEM_CELL_CULTURED_VS_FRESH_UP	420	33	1.99261569	1.40E-04	0.01237152	ANO1;BTG2;CDK7;CDO1;CFH;COL11A1;FN1;GREM2;HLF;KLF6;LRP1B;MARK1;MGP;MTSS1;MYBL1;NINJ2;NOVA1;PDLIM5;PELI1;PGM3;PIK3R1;PMAIP1;SLC19A2;SLC1A4;SLC7A1;SPRY1;SPTBN1;SSPN;TMPO;TP53BP2;ARHGDIB;GLIPR1;SERPINE2
CERVERA_SDHB_TARGETS_2	113	14	3.1420167	1.43E-04	0.01237152	CTBP2;DDC;F5;FAT3;GALNT3;IL1RAP;PKHD1;SERPINI1;SIPA1L2;SLC12A2;SOSTDC1;STK17A;HOMER2;SERPINE2
FEVR_CTNNB1_TARGETS_DN	549	40	1.84776418	1.47E-04	0.01237691	APPBP2;AQP1;AQP4;AXIN2;CA2;CARD11;CCT2;CDO1;CENPE;CETN3;CTBP2;EIF3E;ESRRG;FGF1;G2E3;GTPBP4;HADH;HMGB1;IDE;LBR;LSM5;MSH3;MTBP;NUP155;PIK3R1;PPP2R5E;RFC3;SLC12A2;STRAP;TCF12;TERF1;TFAM;TMPO;TOX;TRIM37;UCHL5;EIF2A;KIF5B;SERPINE2;SRSF1
PATIL_LIVER_CANCER	628	44	1.77685476	1.68E-04	0.01379389	AKR1C3;AP3S1;ATAD2;BLZF1;CCDC117;CDK7;CENPF;CHML;CLGN;DTL;DUS4L;EIF3E;GNPAT;HHAT;INTS7;IRF2BP2;LBR;MAD2L1;MAL2;MTSS1;MZT1;NUCKS1;NUP155;PPP2R5A;PTK2;RPRD2;SLC1A4;STK24;TADA1;TARBP1;TBL1XR1;TMPO;TOB1;TP53BP2;TPR;UCHL5;XPOT;ZNF704;HSF2;KIF20B;PRKCA;RNF187;SPARC;TCEA1
MARSON_BOUND_BY_FOXP3_UNSTIMULATED	1170	71	1.53897436	1.86E-04	0.01490451	ARHGAP15;C12orf60;CAGE1;CCNG2;CLEC12A;CLEC2D;CLTC;COX20;CSNK1A1;EDARADD;EMP1;FUBP1;GGPS1;GNA13;HHAT;HMGB1;HSPH1;ID2;JAK2;KLF6;LBR;LGALS8;LSM5;MAT2B;MBTPS2;MDM4;MSH3;NAA16;PDCD6IP;PPP2R5A;PRPS1;PTK2;PTPN2;RIOK1;SLC17A6;TERF1;TMOD3;TMPO;TSHZ1;UBA3;USP3;VTA1;WBP11;XBP1;ZBED5;ZFP36L1;ZNF280D;ZNRF2;ACTRT3;AMPH;ARCN1;ARG2;ARHGAP11A;ARHGDIB;ARID5B;B4GALT1;CEP120;DCLRE1A;EFCAB9;FBXL3;GLIPR1;KIF5B;MBNL1;ME2;RFX3;SLC25A24;SRSF1;ST6GALNAC1;SYNE3;TAX1BP1;ZEB1
MORI_SMALL_PRE_BII_LYMPHOCYTE_UP	77	11	3.62293763	2.07E-04	0.01619457	BTG2;C1GALT1;CCNG2;CD1D;EPS8;IL12A;KLHDC2;SEC11C;TCF12;B4GALT1;DTNA
CHICAS_RB1_TARGETS_GROWING	240	22	2.32471831	2.19E-04	0.01674631	ANO1;C2CD5;CACYBP;CCBE1;CDCA7L;CDKN2A;CENPF;FANCL;FGF1;KCNMA1;RALGPS2;SLC1A4;TBL1XR1;THSD4;TIFA;TMPO;TOX;TPR;VAT1L;DOCK10;RBBP8;SYNE3
HIRSCH_CELLULAR_TRANSFORMATION_SIGNATURE_UP	241	22	2.31507218	2.32E-04	0.01736734	CAAP1;CFH;CFHR3;EREG;IL1RAP;IVNS1ABP;LGALS8;MRPL9;PELI1;PMAIP1;PTPN2;SLC1A4;STAM;TGIF1;UAP1;ZFAND1;HIF1A;PHLDA1;PPM1D;SERPINE2;SNAPC1;TM4SF1
LEE_NEURAL_CREST_STEM_CELL_DN	119	14	2.98359569	2.49E-04	0.01820465	BCL2;CA2;GDF7;NAV3;PALLD;PDLIM5;PMAIP1;PRR16;SLC1A4;SPRY1;SYNE2;WNT5B;DTNA;LGI1
PEDERSEN_METASTASIS_BY_ERBB2_ISOFORM_4	106	13	3.11025777	2.71E-04	0.01944959	AGR2;ASPH;BRIP1;CA2;CDCA7L;CGNL1;ELL2;EMP1;FN1;MYBL1;OPN3;RAPGEF5;ISOC1
MARSON_BOUND_BY_FOXP3_STIMULATED	965	60	1.57682259	3.18E-04	0.02233435	ABHD13;ADM;ASPH;C12orf60;C1GALT1;CBX3;CCDC117;CCNG2;CLEC2D;COBLL1;COMMD3-BMI1;DTL;FGF7;FUBP1;HHAT;HSPH1;JAK2;KLF6;LCORL;LGALS8;MAT2B;MDM1;MDM2;MSH3;PPP2R5A;PPP4R1;SAMSN1;SFR1;SLC17A6;TBL1XR1;TGIF1;TMEM19;TMOD3;TSHZ1;UBA3;UBE2D1;USP3;WBP11;XBP1;ZBED5;ZFP36L1;ZNF280D;ZNRF2;ACTRT3;ADIPOQ;ANTXR2;ARG2;ARHGAP11A;ARHGDIB;ARID5B;CEP120;CTSC;EIF4E;GLIPR1;HERPUD2;KIF5B;MBNL1;SLC25A24;SRSF1;SYNE3
TIEN_INTESTINE_PROBIOTICS_6HR_DN	168	17	2.56624748	3.55E-04	0.02335125	ADM;BAMBI;CEP76;CETN3;DERA;ID2;IQGAP2;PLEKHA5;PPP2R5A;RBM34;SLC19A2;TMPO;TRIM37;UBE2D1;YEATS4;MSX2;PTS
HUTTMANN_B_CLL_POOR_SURVIVAL_DN	57	9	4.00429948	3.61E-04	0.02335125	ATP2B1;BANK1;COBLL1;MTSS1;MYBL1;PIP5K1B;PTK2;ZNF280D;GLIPR1
SAKAI_CHRONIC_HEPATITIS_VS_LIVER_CANCER_UP	82	11	3.40202679	3.63E-04	0.02335125	AXIN2;CBX3;CSE1L;MPHOSPH6;UBE2D1;UBE2D2;ATF6;HIF1A;HSF2;SRSF1;USP7
ONDER_CDH1_TARGETS_2_UP	249	22	2.24069235	3.66E-04	0.02335125	CDK5RAP2;COL6A3;EPS8;FBLN5;FBN1;FGFR1;FN1;KCNMA1;MIOS;MYBL1;NID1;PRR16;RIMS1;TOX;UGDH;WNT5B;AP1S2;ARID5B;DOCK10;PRKCA;SPARC;ZEB1
SENGUPTA_NASOPHARYNGEAL_CARCINOMA_UP	300	25	2.11338028	3.68E-04	0.02335125	ATAD2;BRIP1;CBX3;CENPE;CENPF;DTL;FBN1;FN1;GATA6;ID2;IPO5;KCNMA1;MAD2L1;NID1;PMAIP1;RFC3;SLCO5A1;TMPO;VASH2;ZNRF3;CTSC;DPY19L1;RBBP8;SLC44A1;SPARC
KORKOLA_TERATOMA_UP	17	5	7.45898923	3.92E-04	0.02441732	EMP1;ETV6;MGP;WBP11;ARHGDIB

AMIT_SERUM_RESPONSE_60_MCF10A	58	9	3.93525983	4.13E-04	0.024897 59	COBLL1;ELL2;EREG;FN1;PDLIM5;PMAIP1;STK38L;TGIF1;TM4SF1
WANG_SMARCE1_TARGETS_DN	373	29	1.97173281	4.14E-04	0.024897 59	ABCE1;ATP2B1;BAMBI;BBIP1;CACYBP;CSE1L;DUSP4;FH;HADH;HSPH1;IGF2BP3;LBR;LSM5;MAD2L1;MIOS;MYBL1;OPN3;PDLIM5;PRR16;STAM;TSPAN13;UCHL5;WNT5B;AIMP1;CTSC;EIF4E;FGF5;TM4SF1;ZNF385D
SWEET_LUNG_CANCER_KRAS_DN	411	31	1.91284055	4.47E-04	0.025936 7	AHR;ALDH1A1;AQP1;CA2;CDO1;CFH;EDN1;FGF7;G0S2;GLUL;GREM2;INMT;LIFR;LIN9;MEOX2;MGP;NID1;PRKCE;SOX11;SOX2;SPTBN1;SSPN;TIMP3;TNNT2;TSPAN13;ADIPOQ;ANTXR2;CAV3;SOX17;SPARC;ZEB1
SMID_BREAST_CANCER_BASAL_DN	677	45	1.68570953	4.51E-04	0.025936 7	AGR2;ALDH1A1;ANO1;APPBP2;ASPH;BCL2;BLVRA;BTG2;CAPN9;CCNG2;CD36;CHN2;CLGN;DEPTOR;DUSP4;ESRRG;HNMT;ICA1;IQGAP2;KCNMA1;LGALS8;MDM1;MEOX2;NAV3;NOVA1;RALGPS2;SERPINI1;SIX1;SLC19A2;SLC1A4;SLC35E3;TOB1;TOX3;TSPAN13;UGDH;ADIPOQ;BRINP2;CDS1;HMGCS2;MSX2;PGR;RAB38;STS;VAV3;ZNF385D
DEBIASI_APOPTOSIS_BY_REOVIRUS_INFECTION_UP	287	24	2.12074398	4.54E-04	0.025936 7	ADM;ATF1;BLZF1;CETN3;DHX9;GNA13;KLF6;LSM5;MBTPS2;PAWR;PMAIP1;PTPN2;RAB21;RPE;SPRY1;TGDS;UBA3;UBE2D1;ATF6;DTNA;EIF4E;LSM6;MED21;SNAPC1
JOSEPH_RESPONSE_TO_SODIUM_BUTYRATE_DN	59	9	3.86856052	4.70E-04	0.026395 13	AFF1;ALDH1A1;AQP1;COL6A3;EIF4G2;HMGB1;THSD7A;TIMP3;USP7
RIGGI_EWING_SARCOMA_PROGENITOR_UP	416	31	1.88984967	5.48E-04	0.030238 67	ATP1A1;BAMBI;BHLHE41;BTG2;C5;CD36;COL11A1;FAT3;G0S2;GALNT3;GDF7;HSD17B2;ID2;IL1RAP;LGALS8;OPN3;PIK3R1;POC1B;RIMKLB;SLCO5A1;SOX2;TCF12;TDO2;TSPAN11;TSPAN13;UBE3D;ZDHHC21;ZNF704;ARHGDIB;CDS1;ZSWIM6
YAGI_AML_FAB_MARKERS	191	18	2.39000074	5.72E-04	0.030741 13	ALDH1A1;BAMBI;CD58;CFH;COA1;COL11A1;DDHD2;ELL2;RFC3;SERPINI1;SLC7A1;TGIF1;TIMP3;AOAH;AP1S2;ARID5B;MED21;SERPINE2
REN_MIF_TARGETS_DN	5	3	15.216338	5.75E-04	0.030741 13	RAB4A;TDO2;FGF5
SANSOM_APC_TARGETS	208	19	2.31658992	6.00E-04	0.031546 28	ABCE1;AQP4;AXIN2;CAGE1;CDK5RAP2;DTL;DUSP4;EDN1;ERCC4;FGF1;HHAT;NR2E3;RPRD1A;SLC12A2;SLC1A4;TC2N;TXNL1;HMGCS2;SOX17
LEE_AGING_MUSCLE_UP	38	7	4.67168273	6.24E-04	0.032324 21	ALDH1A1;AMY2A;HINT1;RAB21;TGIF1;ARHGDIB;SOX17
WILLIAMS_ESR2_TARGETS_DN	11	4	9.22202305	6.34E-04	0.032324 21	ADM;EDN1;SIPA1L2;TBL1XR1
DAZARD_RESPONSE_TO_UV_SCC_UP	102	12	2.98359569	6.76E-04	0.033956 73	AP4S1;ID2;PMAIP1;RTN4;SCFD1;TOB1;UPK1B;AIMP1;AP1S2;ARHGDIB;ME2;SRSF1
THUM_SYSTOLIC_HEART_FAILURE_UP	404	30	1.88321015	7.08E-04	0.035051 45	ADM;ANGPTL1;CCNG2;CCT2;EDN1;FGF1;FH;HNMT;JAK2;KLHL20;M6PR;MIER3;NUCKS1;PKP2;PRPS1;PTK2;SFR1;SLC8A1;TATDN3;TBL1XR1;TPR;UEVLD;ZBED5;ZNF608;ARHGDIB;CTSC;NHSL1;SERPINE2;TM4SF1;ZEB1
ZHENG_FOXP3_TARGETS_IN_THYMUS_UP	195	18	2.34097508	7.31E-04	0.035678 01	AP4S1;ARHGAP15;ATP2B1;C2CD5;KLF6;NSMCE2;PELI1;PIK3R1;TCF12;TOX;WDR37;ARID5B;DOCK10;HERPUD2;HIF1A;MBNL1;RFX3;TOP1
XIE_LT_HSC_S1PR3_OE_UP	29	6	5.24701311	8.04E-04	0.038196 68	CA2;CD36;CD38;IL1RAP;NID1;ZFP36L1
DELYS_THYROID_CANCER_DN	230	20	2.20526638	8.13E-04	0.038196 68	ALDH1A1;ASPH;BCL2;CASQ2;CD36;ESRRG;FBLN5;HLF;IQGAP2;LIFR;LRP1B;MDM1;MEOX2;MGP;RERGL;SETBP1;SPTBN1;TLE1;TOB1;UCHL5
SCHLOSSER_MYC_TARGETS_REPRESSED_BY_SERUM	149	15	2.55307685	8.17E-04	0.038196 68	API5;CBX3;CCT2;DHX9;FH;IPO5;MAD2L1;PRPS1;TFAM;TGDS;TOMM20;XPOT;ZC3H15;EIF4E;SRSF1
FISCHER_G1_S_CELL_CYCLE	182	17	2.36884383	8.90E-04	0.040306 79	ASPH;ATAD2;BARD1;BRIP1;C1GALT1;DTL;IVNS1ABP;MDM1;RIMKLB;TIFA;TRIM45;TTLL7;YEATS4;ARID5B;DCLRE1A;OSBPL6;RBBP8
ACEVEDO_LIVER_CANCER_WITH_H3K9ME3_DN	91	11	3.06556261	8.90E-04	0.040306 79	CEP128;CFH;CFHR3;HOXA1;HOXA13;MYBL1;NUP210L;TOX3;C4orf33;KIF20B;ST6GALNAC1
KRIEG_HYPOXIA_VIA_KDM3A	52	8	3.90162514	9.10E-04	0.040306 79	ADM;ASPH;EDN1;IL1RAP;SERPINB8;SLC7A1;SPINK1;TM4SF1
ROVERSI_GLIOMA_COPY_NUMBER_DN	52	8	3.90162514	9.10E-04	0.040306 79	ARHGAP15;CDKN2A;CTBP2;HNMT;KCNMA1;LRP1B;SFTPA1;SFTPA2
ZWANG_CLASS_3_TRANSIENTLY_INDUCED_BY_EGF	216	19	2.2307903	9.47E-04	0.040881 45	ADM;AFF1;BLZF1;EDN1;ELL2;EMP1;GATA6;HOXA1;IL1RAP;KRT12;PELI1;RIMKLB;STK38L;TGIF1;TOB1;ZFP36L1;RAP2B;TMEM67;ZSWIM6
PLASARI_TGFB1_TARGETS_10HR_DN	250	21	2.13028732	9.51E-04	0.040881 45	ADM;AQP1;CD36;DDC;DTL;EPHX1;FGF7;LIFR;PIK3R1;RFC3;RSPO2;SIX1;SLC12A2;SVIL;THSD7A;TOX;TRIM37;TSHZ1;ZFAND1;ZNF608;ARID5B
BRUINS_UVC_RESPONSE_VIA_TP53_GROUP_B	545	37	1.72172632	9.70E-04	0.040881 45	ANO1;BTG2;DDC;EPHX1;EREG;FGF1;GPR19;INMT;KLF6;LGALS8;LRRIQ4;MGP;PGM3;PMAIP1;PRPS1;RAB21;RAB23;RALGPS2;RIMS1;RPAP3;RPE;SFMBT1;SLC35E3;TSPAN13;UBE3D;ARG2;CRCT1;FBXL3;GABRB2;GINM1;HERPUD2;MSX2;SERPINE2;SLC10A6;SMIM15;SOX17;SPARC
HOLLERN_SOLID_NODULAR_BREAST_TUMOR_DN	30	6	5.07211268	9.71E-04	0.040881 45	AQP1;NID1;RALGPS2;TSPAN11;VASH2;SPARC
KORKOLA_TERATOMA	41	7	4.32985228	0.001003 79	0.040885 22	AGR2;EMP1;EPS8;MGP;SSPN;MBD2;TM4SF1
GABRIELY_MIR21_TARGETS	286	23	2.03948587	0.001008 73	0.040885 22	COBLL1;DDHD2;EXOC8;FIGN;GTF2A1;KBTBD6;LCORL;LIFR;LIN7C;PALLD;PELI1;PIK3R1;RALGPS2;SOX2;SYNE2;TBL1XR1;TIMP3;VASH2;ZRANB1;DOCK10;HERPUD2;OSBPL3;USP7
GRAESSMANN_APOPTOSIS_BY_DOXORUBICIN_DN	1742	94	1.36848046	0.001017 95	0.040885 22	ABCE1;ASPH;BCL2;CBX3;CDCA7L;CDK5RAP2;CETN3;CGNL1;CHML;COX20;CSDE1;DEPTOR;DHX9;DNM1L;DTL;EEF2K;EFNA5;EMP1;ERCC4;FANCL;FGFR1;FUBP1;GATAD2B;GCSH;GDPD1;HADH;HEATR1;HLF;HMGB1;IDE;INTS7;IVNS1ABP;KLF6;KLHL20;KRAS;LACTB2;LBR;LYST;MIOS;MLLT10;MRPL32;MSH3;MTBP;MTSS1;NSMCE2;PDE7A;PIK3R1;PLEKHA5;POLI;POU2F1;RAB4A;RCAN3;RHBDD1;RTTN;S

						EH1L;SFR1;SPRY1;SPTBN1;SSPN;SVIL;SWAP70;TBL1XR1;TCF12;TIFA;TIMP3;TMPO;TPR;UBA3;UBE2E2;XBP1;YEATS4;ZNF280D;ZNF704;ATF6;FKTN;GABRB2;HIF1A;HOMER2;ISOC1;KCTD1;KIF5B;MBD2;ME2;OSBPL6;PANK1;PAPSS1;PPIL4;PTS;SLC25A24;SLC44A1;SRSF1;TRIP4;TSEN15;VAV3
VECCHI_GASTRIC_CANCER_EARLY_DN	358	27	1.91266819	0.00102001	0.04088522	ALDH1A1;ANGPTL1;AQP4;CASQ2;CD36;CFH;CGNL1;ELL2;ESRRG;FCRL5;FIGN;GCNT2;GLUL;GREM2;HLF;LIFR;MYRIP;SERPINI1;SETBP1;SOSTDC1;SOX2;SSTR1;THSD4;TTLL7;ADIPOQ;DTNA;VLDLR
JI_RESPONSE_TO_FSH_DN	53	8	3.82800957	0.00103515	0.04100425	ASPH;EMP1;FN1;GNA13;PRPS1;FGF5;KIF5B;TOP1
PUJANA_CHEK2_PCC_NETWORK	746	47	1.59778348	0.00106762	0.04179847	ABCE1;AP3S1;BARD1;CBX3;CCT2;CENPE;CENPF;CSE1L;DHX9;DPM1;FUBP1;GCSH;GNA13;GNPAT;HSPH1;IPO5;KRAS;LBR;LSM5;MAD2L1;MPHOSPH6;PPP2R5E;PRPS1;PTPN2;RBM34;RFC3;RIMS1;RNF2;SLC1A4;SLC7A1;SRP9;STK24;TACR3;TFAM;TGDS;TMPO;ZC3H15;AIMP1;ARHGAP11A;EIF4E;KIF20B;LSM6;MED21;RBBP8;SRSF1;TCEA1;TOP1
KAPOSI_LIVER_CANCER_MET_DN	6	3	12.6802817	0.00111665	0.04321554	ALDH1A1;EPHX1;PIK3R1
PYEON_CANCER_HEAD_AND_NECK_VS_CERVICAL_UP	186	17	2.31790095	0.00113244	0.04332862	ATAD2;BARD1;BRIP1;C2CD5;CDKN2A;CENPF;CHML;DTL;FANCL;INTS7;MSL2;PMAIP1;RIMKLB;TMPO;ZNF678;ENOSF1;STS
HAHTOLA_SEZARY_SYNDROM_UP	95	11	2.93648629	0.00127496	0.04823363	BARD1;CA1;CDK7;F5;GLUL;LGALS8;NINJ2;SAMSN1;GLIPR1;PAPSS1;TOP1
MAHADEVAN_IMATINIB_RESISTANCE_UP	22	5	5.7637644	0.00141497	0.05293575	ALDH1A1;BAMBI;HSD17B2;SCN3A;SERPINE2
NIKOLSKY_BREAST_CANCER_8P12_P11_AMPLICON	56	8	3.62293763	0.00149582	0.05534265	ADAM2;ASH2L;C8orf86;DDHD2;DKK4;FGFR1;NKX6-3;SMIM19
RAO_BOUND_BY_SALL4	225	19	2.14155869	0.0015315	0.05534265	CACYBP;CASQ2;CFH;FGFR1;FN1;GCNT2;GNA13;GPM6A;MBTPS2;SLCO5A1;SOX2;UHRF2;ZFP36L1;ZNF608;ZNF704;KCMF1;PDHB;RAB38;SLC44A1
ONDER_CDH1_SIGNALING_VIA_CTNNB1	83	10	3.05548956	0.00154506	0.05534265	CCNG2;COL6A3;FBLN5;FBN1;KCNMA1;MIOS;PELI1;SAA1;TSPAN13;DOCK10
MILI_PSEUDOPODIA_CHEMOTAXIS_UP	83	10	3.05548956	0.00154506	0.05534265	BBIP1;COMMD3-BMI1;COX20;EIF3E;NAA50;PALLD;PPFIA1;EIF2A;KIF5B;ZEB1
GOZGIT_ESR1_TARGETS_DN	720	45	1.58503521	0.0015857	0.05620039	AGR2;AHR;AKR1C3;BAMBI;BLVRA;CAPN9;CDCA7L;COA1;COBLL1;DEPTOR;EDN1;EPHX1;FIGN;GATA6;LGALS8;MTSS1;NMD3;NUP210L;PDLIM5;PGM3;PLEKHA5;PTK2;SIPA1L2;SLC12A2;SLC19A2;SNX16;SOX2;SPTSSB;TC2N;THSD4;TLE1;TMTC2;TRIM37;TSPAN13;UGDH;ARG2;ARID5B;DTNA;NR5A2;PGR;PRKCA;TGIF2;VAV3;VLDLR;ZSWIM6
SCHRAETS_MLL_TARGETS_UP	33	6	4.61101152	0.0016388	0.05747736	HOXA1;TNNT2;TOB1;CTSC;MSX2;SERPINE2
HAHTOLA_CTCL_CUTANEOUS	23	5	5.51316595	0.00174999	0.06047568	BARD1;G0S2;GLUL;GLIPR1;TOP1
MONNIER_POSTRADIATION_TUMOR_ESCAPE_DN	372	27	1.84068605	0.00178123	0.06047568	ABHD13;BCL2;BLVRA;CCNG2;CHML;COL6A3;COX20;EEF2K;FGF19;GCNT2;GPR19;HACL1;HOXA1;IL12A;IMMP1L;KANSL1L;KLF6;METTL4;MZT1;PDE7A;PKP2;RCOR3;STIM2;XBP1;ARID5B;FBXL3;SGK2
CREIGHTON_ENDOCRINE_THERAPY_RESISTANCE_5	467	32	1.7377688	0.00179756	0.06047568	AGR2;AHR;ATP2B1;BAMBI;CAPN9;CLTC;COBLL1;ELL2;EMP1;GALNT3;GATA6;ICA1;ID2;KCNMA1;KRAS;MAL2;MIOS;NXPH1;PAIP1;SLC12A2;STK38L;SWAP70;TC2N;TIMP3;TSPAN13;UGDH;WDR37;ZNF704;ARID5B;CDS1;DYRK2;MBNL1
NUYTTEN_EZH2_TARGETS_UP	1015	59	1.47416083	0.00181035	0.06047568	ABHD13;ADM;AP1AR;AXIN2;BANK1;BLZF1;CCDC117;CD58;CDC42SE2;CFHR3;CNIH3;COBLL1;CPNE8;EDN1;EFNA5;FGF1;FGF7;FMN2;FN1;G0S2;GALNT3;GATA6;GCNT2;GNA13;ID2;JAK2;KDM4C;KLF6;KLHL20;LGALS8;LYST;MAT2B;NT5C3A;PMAIP1;PNPT1;POC1B;PRKCE;RNF2;SAA1;SERPINI1;SETBP1;TMTC2;TP53BP2;UAP1;UBE2E2;ZDHHC21;ZFP36L1;ZNF704;AP1S2;ARID5B;ASB2;CTSC;FGF5;HERPUD2;KCTD1;SERPINE2;SPINK1;STYK1;SYNE3

Supplemental Table 6 – Module 2 C2_CGP WebGestalt ORA

Category	H3K4me1	ERα	SP1	Direction of Effect	Overlap	P-value	Jaccard Index	Odds Ratio
1	NA	Loss - 494	Loss - 68	Upreg	20	4.80E-08	0	5.4
2	NA	Loss - 1170	Loss - 112	Downreg	58	8.50E-17	0	5.3
3	**NA**	**Loss - 494**	**Gain - 886**	**Upreg**	**203**	**4.80E-59**	**0.2**	**5.6**
4	**NA**	**Loss - 1170**	**Gain - 1015**	**Downreg**	**470**	**1.20E-123**	**0.3**	**6.1**
5	NA	Gain - 159	Loss - 68	Upreg	4	0.078	0	2.6
6	NA	Gain - 333	Loss - 112	Downreg	9	0.11	0	1.7
7	**NA**	**Gain - 159**	**Gain -886**	**Upreg**	**56**	**7.60E-13**	**0.1**	**3.7**
8	**NA**	**Gain - 333**	**Gain - 1015**	**Downreg**	**98**	**8.80E-12**	**0.1**	**2.5**
9	Loss - 1356	NA	Loss - 68	Upreg	44	1.70E-15	0	7.4
10	Loss - 0	NA	Loss -112	Downreg	0	1.00E+00	0	0
11	**Loss - 1356**	**NA**	**Gain - 884**	**Upreg**	**503**	**1.10E-151**	**0.3**	**7.6**
12	Loss - 0	NA	Gain - 1015	Downreg	0	1.00E+00	0	0
13	Gain - 53	NA	Loss - 68	Upreg	1	0.42	0	1.9
14	**Gain - 1150**	**NA**	**Loss - 112**	**Downreg**	**65**	**1.10E-22**	**0.1**	**7**
15	Gain - 53	NA	Gain - 884	Upreg	15	2.80E-03	0	2.6
16	**Gain - 1150**	**NA**	**Gain - 1015**	**Downreg**	**501**	**3.00E-153**	**0.3**	**7.5**

Supplemental Table 7 – Peak to DEG Group Significance with SP1. Comparisons of changes in H3K4me1, ERα, SP1, and gene expression. Groups of genes in each category were overlapped (i.e. upregulated genes with a loss in SP1 peaks vs upregulated genes with a gain in H3K4me1) and tested with Fisher's exact test. Categories in bold had significant overlaps. Significant categories sharing two characteristic changes were collapsed into groups.
Genome Size = the # of ZR751 DEG: 6,677 genes.
Statistical Test Used = One Sided Fisher's exact test
The null hypothesis is that the odds ratio is no larger than 1. The alternative is that the odds ratio is larger than 1.0.
GeneOverlap in R, by Li Shen
https://www.bioconductor.org/packages/release/bioc/vignettes/GeneOverlap/inst/doc/GeneOverlap.pdf

Experiment	Sample_ID	Name	Total_Mass
RNASeq	3040-TPS-215	ZR751shLucifP4	17,468,400
RNASeq	3040-TPS-206	ZR751shLucifP3	28,396,800
RNASeq	3040-TPS-213	ZR751shMLL3P5	26,836,400
RNASeq	3040-TPS-214	ZR751shMLL3P7	21,714,800
ChIPSeq	3040-TPS-124	ZR751shMLL3_ERa	16,856,209
ChIPSeq	3040-TPS-135	ZR751shMLL3_ERa	9,157,047
ChIPSeq	2501-KS-8	ZR751shMLL3_SP1	6,156,597
ChIPSeq	2501-KS-15	ZR751shMLL3_SP1	10,803,852
ChIPSeq	3040-TPS-125	ZR751shMLL3_H3K4me1	15,748,940
ChIPSeq	3040-TPS-119	ZR751shMLL3_H3K4me1	8,340,487
ChIPSeq	3040-TPS-137	ZR751shLucif_ERa	5,989,465
ChIPSeq	2501-KS-4	ZR751shLucif_ERa	8,835,241
ChIPSeq	2501-KS-6	ZR751shLucif_SP1	10,435,542
ChIPSeq	2501-KS-7	ZR751shLucif_SP1	12,602,784
ChIPSeq	3040-TPS-121	ZR751shLucif_H3K4me1	5,664,688
ChIPSeq	3040-TPS-126	ZR751shLucif_H3K4me1	18,841,450

Supplemental Table 8 – Read Counts of Sequencing Data. Read Counts of RNA-seq and ChIP-seq data.

True Replicates		Rep 1
shMLL3	ERa – 0.7	1,736
	H3K4me1 – 0.1	4,541
	SP1 – 0.5	2,489
shLucif	ERa – 0.7	1,483
	H3K4me1 – 0.1	17,242
	SP1 – 0.5	473

Self PseudoReplicates		Rep 1	Rep 2
shMLL3	ERa - 0.7	3,367	645
	H3K4me1 – 0.25	3,654	3,730
	SP1 – 0.5	202	14,336
shLucif	ERa – 0.7	385	5,305
	H3K4me1 – 0.25	22,345	7,998
	SP1 – 0.5	270	189

Pooled PseudoReplicates		Rep 1
shMLL3	ERa – 0.5	1,511
	H3K4me1 – 0.05	865
	SP1 – 0.001	1,774
shLucif	ERa – 0.5	3,986
	H3K4me1 – 0.05	24,108
	SP1 – 0.001	118

N1 and N2 = No. of peaks passing IDR threshold by comparing self-pseudoReplicates for Rep1 and Rep2 respectively
Np = No. of peaks passing IDR threshold by comparing pooled pseudo-replicates
Nt = Best no. of peaks passing IDR threshold by comparing true replicates
Optimal Peak set = Longest of the Nt and Np peak lists
Rescue Ratio = max(Np,Nt) / min(Np,Nt)
Nt and Np should be within a factor of 2 of each other
Self-consistency Ratio = max(N1,N2) / min(N1,N2)
N1 and N2 should be within a factor of 2 of each other
If Rescue Ratio AND self-consistency Ratio are both > 2, Flag the file for reproducibility FAIL (-1)
If Rescue Ratio OR self-consistency Ratio are > 2, Flag the file for reproducibility Borderline (0)

Number Peaks		Np	Nt	N1	N2
shMLL3	ERa	1,511	1,736	3,367	645
	H3K4me1	865	4,541	3,654	3,730
	SP1	1,774	2,489	202	14,336
shLucif	ERa	3,986	1,483	385	5,305
	H3K4me1	24,108	17,242	22,345	7,998
	SP1	118	473	270	189

Ratios		Rescue	Self-Consistency
shMLL3	ERa - Borderline	1.14	5.22
	H3K4me1 - Borderline	5.24	1.02
	SP1 - Borderline	1.40	75.852
shLucif	ERa – Borderline	2.68	13.77
	H3K4me1 - Borderline	1.39	2.79
	SP1 - Borderline	4.1	1.429

Supplemental Table 9 – ChIP-seq IDR process. Peak thresholds from IDR process for ChIP-seq samples.

Figure 5-1B				
ANOVA summary				
F	18.62		R square	0.867
P value	< 0.0001		Number of families	1
P value summary	****		Number of comparisons per family	4
Are differences among means statistically significant? (P < 0.05)	Yes		Alpha	0.05
Holm-Sidak's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*
WT PE vs. KO PE	0.5386	Yes	*	0.0265
WT E vs. KO E	0.3825	Yes	*	0.0376
WT ME vs. KO ME	1.081	Yes	****	< 0.0001
WT DE vs. KO DE	0.6243	Yes	**	0.0065
ANOVA summary				
F	215.3		R square	0.5039
P value	< 0.0001		Number of families	1
P value summary	****		Number of comparisons per family	3
Are differences among means statistically significant? (P < 0.05)	Yes		Alpha	0.05
Dunnett's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*
WT E vs. WT PE	-0.1718	Yes	**	0.0042
WT E vs. WT ME	-0.8292	Yes	****	< 0.0001
WT E vs. WT DE	-1.132	Yes	****	< 0.0001
Figure 5-1C				
ANOVA summary				
F	11.09		R square	0.7281
P value	< 0.0001		Number of families	1
P value summary	****		Number of comparisons per family	4
Are differences among means statistically significant? (P < 0.05)	Yes		Alpha	0.05
Holm-Sidak's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*
WT 12 PE vs. KO 12 PE	22.28	Yes	**	0.0023
WT 12 E vs. KO 12 E	20.57	Yes	**	0.0023
WT 12 ME vs. KO 12 ME	22.14	Yes	***	0.0002
WT 12 DE vs. KO 12 DE	15.79	Yes	**	0.0023
Figure 5-3B				
ANOVA summary				
F	106.4		R square	0.8944
P value	< 0.0001		Number of families	1
P value summary	****		Number of comparisons per family	4

Are differences among means statistically significant? (P < 0.05)	Yes		Alpha	0.05
Holm-Sidak's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*
ML WT (E) vs. ML KO (E)	19.18	Yes	****	< 0.0001
LP WT vs. LP KO	0.9114	No	ns	0.9613
AP WT vs. AP KO	-0.7909	No	ns	0.9613
U WT vs. U KO	-19.2	Yes	****	< 0.0001
Figure 5-3C				
ANOVA summary				
F	13.56		R square	0.655
P value	< 0.0001		Number of families	1
P value summary	****		Number of comparisons per family	4
Are differences among means statistically significant? (P < 0.05)	Yes		Alpha	0.05
Holm-Sidak's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*
ML WT (PE) vs. ML KO (PE)	24.18	Yes	**	0.0031
ML WT (E) vs. ML KO (E)	19.1	Yes	****	< 0.0001
ML WT (ME) vs. ML KO (ME)	18.64	Yes	*	0.0197
ML WT (DE) vs. ML KO (DE)	11.79	Yes	*	0.0461
Figure 5-3F				
ANOVA summary				
F	1.881		R square	0.116
P value	0.1471		Number of families	1
P value summary	ns		Number of comparisons per family	2
Are differences among means statistically significant? (P < 0.05)	No		Alpha	0.05
Holm-Sidak's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*
6wk WT vs. 6 KO	0.2369	No	ns	0.8948
WT 8wk vs. KO 8wk	4	No	ns	0.3524
Figure 5-3H				
ANOVA summary				
F	8.896		R square	0.6402
P value	< 0.0001		Number of families	1
P value summary	****		Number of comparisons per family	3
Are differences among means statistically significant? (P < 0.05)	Yes		Alpha	0.05
Holm-Sidak's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*
WT6 vs. KO6	0.9722	No	ns	0.6945
WT8 vs. KO8	4	No	ns	0.583

WT10 vs. KO10	23.32	Yes	****	< 0.0001
Figure 5-5A				
ANOVA summary				
F	6.893		R square	0.7338
P value	0.0012		Number of families	1
P value summary	**		Number of comparisons per family	4
Are differences among means statistically significant? (P < 0.05)	Yes		Alpha	0.05
Holm-Sidak's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*
8wkWT vs. 6wkWT	0.2	No	ns	0.4754
10wkWTE vs. 6wkWT	0.4374	No	ns	0.2435
12wk WTE vs. 6wkWT	1.196	Yes	***	0.0003
12wkKOE vs. 6wkKO	0.7278	Yes	*	0.0379
Figure 5-5C				
ANOVA summary				
F	10.98		R squared	0.6281
P value	0.0016		Number of families	1
P value summary	**		Number of comparisons per family	3
Significant diff. among means (P < 0.05)?	Yes		Alpha	0.05
Tukey's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*
WTE vs. WTDE	0.7499	Yes	**	0.0042
WTE vs. WTPE	0.9749	Yes	**	0.0096
WTDE vs. WTPE	0.2251	No	ns	0.7157
Figure 5-7C				
ANOVA summary				
F	6.587		R square	0.7671
P value	0.0251		Number of families	1
P value summary	*		Number of comparisons per family	3
Are differences among means statistically significant? (P < 0.05)	Yes		Alpha	0.05
Holm-Sidak's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*
KO_E vs. WT_E	181.5	Yes	*	0.0329
KO_E vs. KO_DE	239.8	Yes	*	0.024
KO_E vs. WT_DE	196	Yes	*	0.0329
Figure 5-7D				
ANOVA summary				
F	11.53		R square	0.8122
P value	0.0028		Number of families	1
P value summary	**		Number of comparisons per family	3
Are differences among means statistically significant? (P < 0.05)	Yes		Alpha	0.05
Holm-Sidak's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*

WTV vs. WTPS	-0.05347	No	ns	0.6881
KOV vs. KOPS	-0.7145	Yes	**	0.0016
WTV vs. KOV	0.4995	Yes	**	0.0092
Figure 5-12A				
ANOVA summary				
F	7.01		R squared	0.5126
P value	0.0021		Number of families	1
P value summary	**		Number of comparisons per family	3
Significant diff. among means (P < 0.05)?	Yes		Alpha	0.05
Sidak's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*
WT E vs. WT DE	0.7597	Yes	**	0.002
KO E vs. KO DE	0.2441	No	ns	0.6895
WT DE vs. KO DE	-0.6181	Yes	*	0.0437
Figure 5-12B				
ANOVA summary				
F	574.1		R square	0.9957
P value	< 0.0001		Number of families	1
P value summary	****		Number of comparisons per family	4
Are differences among means statistically significant? (P < 0.05)	Yes		Alpha	0.05
Dunnett's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*
V vs. 216	0.6396	Yes	****	< 0.0001
V vs. BID	0.4211	Yes	****	< 0.0001
V vs. Tram	0.1658	Yes	****	< 0.0001
V vs. U0126	0.2126	Yes	****	< 0.0001
Figure 5-12E				
ANOVA summary				
F	25.58		R square	0.895
P value	< 0.0001		Number of families	1
P value summary	****		Number of comparisons per family	3
Are differences among means statistically significant? (P < 0.05)	Yes		Alpha	0.05
Sidak's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*
WT-V vs. WT-216	100	Yes	****	< 0.0001
167-V vs. 167-216	53.26	Yes	**	0.0013
118-V vs. 118-216	-22.7	No	ns	0.2554
Figure 5-12I				
ANOVA summary				
F	42.55		R square	0.9466
P value	< 0.0001		Number of families	1
P value summary	****		Number of comparisons per family	5
Are differences among means statistically significant? (P < 0.05)	Yes		Alpha	0.05

Dunnett's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*
V vs. 216	100	Yes	****	< 0.0001
V vs. Ebs	-56.73	Yes	**	0.0011
V vs. NAC	-1.255	No	ns	0.9999
V vs. 216+Ebs	0.1907	No	ns	> 0.9999
V vs. 216+NAC	-2.203	No	ns	0.9997
Figure 5-12J				
ANOVA summary				
F	11.29		R square	0.9039
P value	0.0052		Number of families	1
P value summary	**		Number of comparisons per family	5
Are differences among means statistically significant? (P < 0.05)	Yes		Alpha	0.05
Dunnett's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*
veh vs. 216	-100	Yes	**	0.0046
veh vs. NAC	4.143	No	ns	0.9988
veh vs. Ebs	7.899	No	ns	0.9856
veh vs. NAC+216	2.002	No	ns	0.9999
veh vs. Ebs+216	-4.961	No	ns	0.9981
Figure 5-13C				
ANOVA summary				
F	4483		R square	0.9989
P value	< 0.0001		Number of families	1
P value summary	****		Number of comparisons per family	2
Are differences among means statistically significant? (P < 0.05)	Yes		Alpha	0.05
Dunnett's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*
v vs. 216	100	Yes	****	< 0.0001
v vs. no OPP	110.7	Yes	****	< 0.0001
Figure 5-13E				
ANOVA summary				
F	53.24		R square	0.8331
P value	< 0.0001		Number of families	1
P value summary	****		Number of comparisons per family	3
Are differences among means statistically significant? (P < 0.05)	Yes		Alpha	0.05
Holm-Sidak's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*
WT WT vs. KO WT	2.2	Yes	****	< 0.0001
WT KO vs. KO KO	2.495	Yes	****	< 0.0001
WT WT vs. KO KO	3.167	Yes	****	< 0.0001
Figure 5-13F				
ANOVA summary				
F	8.231		R square	0.5784
P value	0.0056		Number of families	1

P value summary	**		Number of comparisons per family	3
Are differences among means statistically significant? (P < 0.05)	Yes		Alpha	0.05
Tukey's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*
F vs. L	1.323	Yes	*	0.0023
F vs. OC	1.565	Yes	**	0.0007
L vs. OC	0.2425	No	ns	0.8481
Figure 5-13G				
ANOVA summary				
F	16.2		R square	0.7137
P value	0.0003		Number of families	1
P value summary	***		Number of comparisons per family	3
Are differences among means statistically significant? (P < 0.05)	Yes		Alpha	0.05
Tukey's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*
F vs. L	-370.8	Yes	**	0.0403
F vs. OC	-425.5	Yes	***	0.0086
L vs. OC	-54.69	No	ns	0.8938
Figure 5-4B				
ANOVA summary				
F	15.15		R square	0.6152
P value	< 0.0001		Number of families	1
P value summary	****		Number of comparisons per family	10
Are differences among means statistically significant? (P < 0.05)	Yes		Alpha	0.05
Holm-Sidak's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*
Lum WT (PE) vs. Lum KO (PE)	2.92	No	ns	0.9759
Lum WT (E) vs. Lum KO (E)	-3.972	No	ns	0.9001
Lum WT (ME) vs. Lum KO (ME)	11.03	No	ns	0.3438
Lum WT (DE) vs. Lum KO (DE)	-0.8083	No	ns	0.9883
6wk WT Lum vs. 6wk KO Lum	-0.7805	No	ns	0.9883
6wk WT Basal vs. 6wk KO Basal	2.215	No	ns	0.9759
Basal WT (PE) vs. Basal KO (PE)	-3.224	No	ns	0.9759
Basal WT (E) vs. Basal KO (E)	5.276	No	ns	0.7059
Basal WT (ME) vs. Basal KO (ME)	-4.4	No	ns	0.9759
Basal WT (DE) vs. Basal KO (DE)	0.8421	No	ns	0.9883

Figure 5-4D (upper panel)				
ANOVA summary				
F	14.08		R square	0.6827
P value	< 0.0001		Number of families	1
P value summary	****		Number of comparisons per family	6
Are differences among means statistically significant? (P < 0.05)	Yes		Alpha	0.05
Holm-Sidak's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*
LP WT (PE) vs. LP KO (PE)	0.6	No	ns	0.6675
LP WT (E) vs. LP KO (E)	0.7409	No	ns	0.4535
LP WT (ME) vs. LP KO (ME)	0.685	No	ns	0.6675
LP WT (DE) vs. LP KO (DE)	1.354	No	ns	0.2605
LP WT 6 wk vs. LP KO 6 wk	1.357	No	ns	0.2605
LP WT 8wk vs. LP KO 8wk	0.8333	No	ns	0.5929
Figure 5-4D (middle panel)				
ANOVA summary				
F	25.68		R square	0.788
P value	< 0.0001		Number of families	1
P value summary	****		Number of comparisons per family	6
Are differences among means statistically significant? (P < 0.05)	Yes		Alpha	0.05
Holm-Sidak's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*
AP WT (PE) vs. AP KO (PE)	-1	No	ns	0.9127
AP WT (E) vs. AP KO (E)	-0.8557	No	ns	0.8246
AP WT (ME) vs. AP KO (ME)	-0.95	No	ns	0.9127
AP WT (DE) vs. AP KO (DE)	-0.2083	No	ns	0.9127
AP WT 6 wk vs. AP KO 6 wk	-0.9667	No	ns	0.8246
AP WT 8wk vs. AP KO 8 wk	0.4833	No	ns	0.9127
Figure 5-4D (lower panel)				
ANOVA summary				
F	10.23		R square	0.597
P value	< 0.0001		Number of families	1
P value summary	****		Number of comparisons per family	6

Are differences among means statistically significant? (P < 0.05)	Yes		Alpha	0.05
Holm-Sidak's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*
U WT 6 wk vs. U KO 6 wk	-4.556	No	ns	0.4122
U WT 8wk vs. U KO 8wk	-4.417	No	ns	0.4122
U WT (PE) vs. U KO (PE)	-22.25	Yes	***	0.001
U WT (E) vs. U KO (E)	-18.63	Yes	****	< 0.0001
U WT (ME) vs. U KO (ME)	-18.46	Yes	**	0.004
U WT (DE) vs. U KO (DE)	-10.75	Yes	*	0.046
Figure 5-4F (upper panel)				
ANOVA summary				
F	48.26		R square	0.9013
P value	< 0.0001		Number of families	1
P value summary	****		Number of comparisons per family	4
Are differences among means statistically significant? (P < 0.05)	Yes		Alpha	0.05
Holm-Sidak's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*
4wt vs. 4ko	-0.01537	No	ns	0.9784
6wt vs. 6ko	-0.1672	No	ns	0.9784
8wt vs. 8ko	-0.2023	No	ns	0.9784
10wt vs. 10ko	0.32	No	ns	0.9784
Figure 5-4F (lower panel)				
ANOVA summary				
F	10.14		R square	0.6827
P value	< 0.0001		Number of families	1
P value summary	****		Number of comparisons per family	4
Are differences among means statistically significant? (P < 0.05)	Yes		Alpha	0.05
Holm-Sidak's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*
4wt vs. 4ko	2	No	ns	0.8956
6wt vs. 6ko	-3.8	No	ns	0.8956
8wt vs. 8ko	-2.214	No	ns	0.8956
10wt vs. 10ko	-1.5	No	ns	0.8956
Figure 5-8B				
ANOVA summary				
F	6.587		R square	0.7671
P value	0.0251		Number of families	1
P value summary	*		Number of comparisons per family	3

Are differences among means statistically significant? (P < 0.05)	Yes		Alpha	0.05
Holm-Sidak's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*
KO_E vs. WT_E	181.5	Yes	*	0.0329
KO_E vs. KO_DE	239.8	Yes	*	0.024
KO_E vs. WT_DE	196	Yes	*	0.0329
Figure 5-8D				
ANOVA summary				
F	32.99		R square	0.9428
P value	0.0004		Number of families	1
P value summary	***		Number of comparisons per family	3
Are differences among means statistically significant? (P < 0.05)	Yes		Alpha	0.05
Holm-Sidak's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*
WT-sham vs. WT-OVX	-0.3735	No	ns	0.0523
KO-sham vs. KO-OVX	-1.381	Yes	***	0.0003
WT-sham vs. KO-sham	0.9988	Yes	**	0.0021
Figure 5-14F				
ANOVA summary				
F	16.37		R square	0.7318
P value	0.0004		Number of families	1
P value summary	***		Number of comparisons per family	3
Are differences among means statistically significant? (P < 0.05)	Yes		Alpha	0.05
Holm-Sidak's multiple comparisons test	*Mean Diff.*	*Significant?*	*Summary*	*Adjusted P Value*
F vs. L	-307.9	Yes	**	0.0032
F vs. HC	-349.3	Yes	***	0.0008
L vs. HC	-41.37	No	ns	0.6384

Supplemental Table 10. ANOVA table for "RSK2 maintains adult estrogen homeostasis by inhibiting ERK1/2-mediated degradation of estrogen receptor alpha" figures.

REFERENCES

1. Cancer statistics, 2019 - Siegel - 2019 - CA: A Cancer Journal for Clinicians - Wiley Online Library. Accessed October 22, 2021. https://acsjournals-onlinelibrary-wiley-com.proxy.library.vanderbilt.edu/doi/10.3322/caac.21551

2. Watson CJ, Khaled WT. Mammary development in the embryo and adult: a journey of morphogenesis and commitment. *Development*. 2008;135(6):995-1003. doi:10.1242/dev.005439

3. Macias H, Hinck L. Mammary gland development. *WIREs Dev Biol*. 2012;1(4):533-557. doi:10.1002/wdev.35

4. Sreekumar A, Roarty K, Rosen JM. The Mammary Stem Cell Hierarchy: A Looking Glass into Heterogeneous Breast Cancer Landscapes. *Endocr Relat Cancer*. 2015;22(6):T161-T176. doi:10.1530/ERC-15-0263

5. Dall GV, Britt KL. Estrogen Effects on the Mammary Gland in Early and Late Life and Breast Cancer Risk. *Front Oncol*. 2017;7:110. doi:10.3389/fonc.2017.00110

6. Horn J, Vatten LJ. Reproductive and hormonal risk factors of breast cancer: a historical perspective. *Int J Womens Health*. 2017;9:265-272. doi:10.2147/IJWH.S129017

7. Washbrook E. Risk factors and epidemiology of breast cancer. *Womens Health Med*. 2006;3(1):8-14. doi:10.1383/wohm.2006.3.1.8

8. Islami F, Ward EM, Sung H, et al. Annual Report to the Nation on the Status of Cancer, Part 1: National Cancer Statistics. *JNCI J Natl Cancer Inst*. 2021;(djab131). doi:10.1093/jnci/djab131

9. Sukocheva OA, Lukina E, Friedemann M, Menschikowski M, Hagelgans A, Aliev G. The crucial role of epigenetic regulation in breast cancer anti-estrogen resistance: Current findings and future perspectives. *Semin Cancer Biol*. Published online December 7, 2020. doi:10.1016/j.semcancer.2020.12.004

10. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. *The Lancet*. 2011;378(9793):771-784. doi:10.1016/S0140-6736(11)60993-8

11. Nass N, Kalinski T. Tamoxifen resistance: From cell culture experiments towards novel biomarkers. *Pathol - Res Pract*. 2015;211(3):189-197. doi:10.1016/j.prp.2015.01.004

12. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin*. 2020;70(1):7-30. doi:10.3322/caac.21590

13. Fuentes N, Silveyra P. Chapter Three - Estrogen receptor signaling mechanisms. In: Donev R, ed. *Advances in Protein Chemistry and Structural Biology*. Vol 116. Intracellular Signalling Proteins. Academic Press; 2019:135-170. doi:10.1016/bs.apcsb.2019.01.001

14. Cagnacci A, Soldani R, Carriero PL, Paoletti AM, Fioretti P, Melis GB. Effects of low doses of transdermal 17 beta-estradiol on carbohydrate metabolism in postmenopausal women. *J Clin Endocrinol Metab*. 1992;74(6):1396-1400. doi:10.1210/jcem.74.6.1317387

15. Eliassen AH, Hankinson SE. Endogenous hormone levels and risk of breast, endometrial and ovarian cancers: prospective studies. *Adv Exp Med Biol*. 2008;630:148-165.

16. Acconcia F, Ascenzi P, Fabozzi G, Visca P, Marino M. S-palmitoylation modulates human estrogen receptor-α functions. *Biochem Biophys Res Commun*. 2004;316(3):878-883. doi:10.1016/j.bbrc.2004.02.129

17. Marino M, Ascenzi P, Acconcia F. S-palmitoylation modulates estrogen receptor α localization and functions. *Steroids*. 2006;71(4):298-303. doi:10.1016/j.steroids.2005.09.011

18. Marino M, Ascenzi P. Membrane association of estrogen receptor α and β influences 17β-estradiol-mediated cancer cell proliferation. *Steroids*. 2008;73(9):853-858. doi:10.1016/j.steroids.2007.12.003

19. Pedram A, Razandi M, Sainson RCA, Kim JK, Hughes CC, Levin ER. A Conserved Mechanism for Steroid Receptor Translocation to the Plasma Membrane *. *J Biol Chem*. 2007;282(31):22278-22288. doi:10.1074/jbc.M611877200

20. Razandi M, Pedram A, Merchenthaler I, Greene GL, Levin ER. Plasma Membrane Estrogen Receptors Exist and Functions as Dimers. *Mol Endocrinol*. 2004;18(12):2854-2865. doi:10.1210/me.2004-0115

21. ChIP-Seq of ERα and RNA polymerase II defines genes differentially responding to ligands. *EMBO J*. 2009;28(10):1418-1428. doi:10.1038/emboj.2009.88

22. Carroll JS, Meyer CA, Song J, et al. Genome-wide analysis of estrogen receptor binding sites. *Nat Genet*. 2006;38(11):1289-1297. doi:10.1038/ng1901

23. Lin CY, Vega VB, Thomsen JS, et al. Whole-Genome Cartography of Estrogen Receptor α Binding Sites. *PLOS Genet*. 2007;3(6):e87. doi:10.1371/journal.pgen.0030087

24. Klinge CM. Estrogen receptor interaction with estrogen response elements. *Nucleic Acids Res*. 2001;29(14):2905-2919. doi:10.1093/nar/29.14.2905

25. Yi P, Wang Z, Feng Q, et al. Structural and Functional Impacts of ER Coactivator Sequential Recruitment. *Mol Cell*. 2017;67(5):733-743.e4. doi:10.1016/j.molcel.2017.07.026

26. Murakami S, Nagari A, Kraus WL. Dynamic assembly and activation of estrogen receptor α enhancers through coregulator switching. *Genes Dev*. 2017;31(15):1535-1548. doi:10.1101/gad.302182.117

27. 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. doi:10.1038/ng.730

28. Magnani L, Ballantyne EB, Zhang X, Lupien M. PBX1 Genomic Pioneer Function Drives ERα Signaling Underlying Progression in Breast Cancer. *PLOS Genet*. 2011;7(11):e1002368. doi:10.1371/journal.pgen.1002368

29. 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. doi:10.1101/gr.139469.112

30. Ali S, Coombes RC. Estrogen Receptor Alpha in Human Breast Cancer: Occurrence and Significance. :11.

31. Dubik D, Shiu RP. Transcriptional regulation of c-myc oncogene expression by estrogen in hormone-responsive human breast cancer cells. *J Biol Chem*. 1988;263(25):12705-12708. doi:10.1016/S0021-9258(18)37810-4

32. Altucci L, Addeo R, Cicatiello L, et al. 17beta-Estradiol induces cyclin D1 gene transcription, p36D1-p34cdk4 complex activation and p105Rb phosphorylation during mitogenic stimulation of G(1)-arrested human breast cancer cells. *Oncogene*. 1996;12(11):2315-2324.

33. Carroll JS, Hickey TE, Tarulli GA, Williams M, Tilley WD. Deciphering the divergent roles of progestogens in breast cancer. *Nat Rev Cancer*. 2017;17(1):54-64. doi:10.1038/nrc.2016.116

34. Palaniappan M, Nguyen L, Grimm SL, et al. The genomic landscape of estrogen receptor α binding sites in mouse mammary gland. *PLOS ONE*. 2019;14(8):e0220311. doi:10.1371/journal.pone.0220311

35. Lin Z, Reierstad S, Huang CC, Bulun SE. Novel Estrogen Receptor-α Binding Sites and Estradiol Target Genes Identified by Chromatin Immunoprecipitation Cloning in Breast Cancer. *Cancer Res*. 2007;67(10):5017-5024. doi:10.1158/0008-5472.CAN-06-3696

36. Gruber CJ, Tschugguel W, Schneeberger C, Huber JC. Production and Actions of Estrogens. *N Engl J Med*. 2002;346(5):340-352. doi:10.1056/NEJMra000471

37. Meisel JL, Venur VA, Gnant M, Carey L. Evolution of Targeted Therapy in Breast Cancer: Where Precision Medicine Began. *Am Soc Clin Oncol Educ Book*. 2018;(38):78-86. doi:10.1200/EDBK_201037

38. Nilsson S, Koehler KF. Oestrogen receptors and selective oestrogen receptor modulators: molecular and cellular pharmacology. *Basic Clin Pharmacol Toxicol*. 2005;96(1):15-25. doi:10.1111/j.1742-7843.2005.pto960103.x

39. Fisher B, Costantino JP, Wickerham DL, et al. Tamoxifen for Prevention of Breast Cancer: Report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *JNCI J Natl Cancer Inst*. 1998;90(18):1371-1388. doi:10.1093/jnci/90.18.1371

40. McDonnell DP, Wardell SE, Norris JD. Oral Selective Estrogen Receptor Downregulators (SERDs), a Breakthrough Endocrine Therapy for Breast Cancer. *J Med Chem*. 2015;58(12):4883-4887. doi:10.1021/acs.jmedchem.5b00760

41. Buzdar AU, Robertson JFR, Eiermann W, Nabholtz JM. An overview of the pharmacology and pharmacokinetics of the newer generation aromatase inhibitors anastrozole, letrozole, and exemestane. *Cancer*. 2002;95(9):2006-2016. doi:10.1002/cncr.10908

42. Hamadeh IS, Patel JN, Rusin S, Tan AR. Personalizing aromatase inhibitor therapy in patients with breast cancer. *Cancer Treat Rev*. 2018;70:47-55. doi:10.1016/j.ctrv.2018.07.014

43. Cardoso F, Costa A, Senkus E, et al. 3rd ESO-ESMO International Consensus Guidelines for Advanced Breast Cancer (ABC 3). *Ann Oncol Off J Eur Soc Med Oncol*. 2017;28(1):16-33. doi:10.1093/annonc/mdw544

44. Fulvestrant, Formerly ICI 182,780, Is as Effective as Anastrozole in Postmenopausal Women With Advanced Breast Cancer Progressing After Prior Endocrine Treatment | Journal of Clinical Oncology. Accessed November 2, 2021. https://ascopubs-org.proxy.library.vanderbilt.edu/doi/full/10.1200/JCO.2002.10.057

45. Osborne C k., Pippen J, Jones S e., et al. Double-Blind, Randomized Trial Comparing the Efficacy and Tolerability of Fulvestrant Versus Anastrozole in Postmenopausal Women With Advanced Breast Cancer Progressing on Prior Endocrine Therapy: Results of a North American Trial. *J Clin Oncol*. 2002;20(16):3386-3395. doi:10.1200/JCO.2002.10.058

46. Robertson JFR, Osborne CK, Howell A, et al. Fulvestrant versus anastrozole for the treatment of advanced breast carcinoma in postmenopausal women. *Cancer*. 2003;98(2):229-238. doi:10.1002/cncr.11468

47. Vergote I, Robertson JFR. Fulvestrant is an effective and well-tolerated endocrine therapy for postmenopausal women with advanced breast cancer: results from clinical trials. *Br J Cancer*. 2004;90(1):S11-S14. doi:10.1038/sj.bjc.6601631

48. Hanker AB, Sudhan DR, Arteaga CL. Overcoming Endocrine Resistance in Breast Cancer. *Cancer Cell*. 2020;37(4):496-513. doi:10.1016/j.ccell.2020.03.009

49. Malhotra GK, Zhao X, Band H, Band V. Histological, molecular and functional subtypes of breast cancers. *Cancer Biol Ther*. 2010;10(10):955-960. doi:10.4161/cbt.10.10.13879

50. Perou CM, Sørlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature*. 2000;406(6797):747-752. doi:10.1038/35021093

51. Prat A, Parker JS, Karginova O, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res BCR*. 2010;12(5):R68. doi:10.1186/bcr2635

52. Sørlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci*. 2003;100(14):8418-8423. doi:10.1073/pnas.0932692100

53. Herschkowitz JI, Simin K, Weigman VJ, et al. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome Biol*. 2007;8(5):R76. doi:10.1186/gb-2007-8-5-r76

54. Sestak I, Cuzick J, Dowsett M, et al. Prediction of Late Distant Recurrence After 5 Years of Endocrine Treatment: A Combined Analysis of Patients From the Austrian Breast and Colorectal Cancer Study Group 8 and Arimidex, Tamoxifen Alone or in Combination Randomized Trials Using the PAM50 Risk of Recurrence Score. *J Clin Oncol*. 2015;33(8):916-922. doi:10.1200/JCO.2014.55.6894

55. Burstein HJ, Lacchetti C, Anderson H, et al. Adjuvant Endocrine Therapy for Women With Hormone Receptor–Positive Breast Cancer: ASCO Clinical Practice Guideline Focused Update. *J Clin Oncol*.:19.

56. Schaafsma E, Zhang B, Schaafsma M, Tong CY, Zhang L, Cheng C. Impact of Oncotype DX testing on ER+ breast cancer treatment and survival in the first decade of use. *Breast Cancer Res*. 2021;23(1):74. doi:10.1186/s13058-021-01453-4

57. Manso L, Mourón S, Tress M, et al. Analysis of Paired Primary-Metastatic Hormone-Receptor Positive Breast Tumors (HRPBC) Uncovers Potential Novel Drivers of Hormonal Resistance. *PLOS ONE*. 2016;11(5):e0155840. doi:10.1371/journal.pone.0155840

58. van Geelen CT, Savas P, Teo ZL, et al. Clinical implications of prospective genomic profiling of metastatic breast cancer patients. *Breast Cancer Res*. 2020;22(1):91. doi:10.1186/s13058-020-01328-0

59. Prognostic Impact of Discordance in Hormone Receptor Status Between Primary and Recurrent Sites in Patients With Recurrent Breast Cancer - PubMed. Accessed November 10, 2021. https://pubmed-ncbi-nlm-nih-gov.proxy.library.vanderbilt.edu/27268749/

60. Jeselsohn R, Buchwalter G, De Angelis C, Brown M, Schiff R. ESR1 mutations—a mechanism for acquired endocrine resistance in breast cancer. *Nat Rev Clin Oncol*. 2015;12(10):573-583. doi:10.1038/nrclinonc.2015.117

61. Jeselsohn R, Yelensky R, Buchwalter G, et al. Emergence of constitutively active estrogen receptor-α mutations in pretreated advanced estrogen receptor-positive breast cancer. *Clin Cancer Res Off J Am Assoc Cancer Res*. 2014;20(7):1757-1767. doi:10.1158/1078-0432.CCR-13-2332

62. Magnani L, Frige G, Gadaleta RM, et al. Acquired CYP19A1 amplification is an early specific mechanism of aromatase inhibitor resistance in ERα metastatic breast cancer. *Nat Genet*. 2017;49(3):444-450. doi:10.1038/ng.3773

63. Razavi P, Chang MT, Xu G, et al. The Genomic Landscape of Endocrine-Resistant Advanced Breast Cancers. *Cancer Cell*. 2018;34(3):427-438.e6. doi:10.1016/J.CCELL.2018.08.008

64. Fu X, Jeselsohn R, Pereira R, et al. FOXA1 overexpression mediates endocrine resistance by altering the ER transcriptome and IL-8 expression in ER-positive breast cancer. *Proc Natl Acad Sci U S A*. 2016;113(43):E6600-E6609. doi:10.1073/pnas.1612835113

65. The Lineage Determining Factor GRHL2 Collaborates with FOXA1 to Establish a Targetable Pathway in Endocrine Therapy-Resistant Breast Cancer - PubMed. Accessed November 10, 2021. https://pubmed-ncbi-nlm-nih-gov.proxy.library.vanderbilt.edu/31644911/

66. Formisano L, Stauffer KM, Young CD, et al. Association of FGFR1 with ERα Maintains Ligand-Independent ER Transcription and Mediates Resistance to Estrogen Deprivation in ER+ Breast Cancer. *Clin Cancer Res*. 2017;23(20):6138-6150. doi:10.1158/1078-0432.CCR-17-1232

67. O'Leary B, Cutts RJ, Liu Y, et al. The Genetic Landscape and Clonal Evolution of Breast Cancer Resistance to Palbociclib plus Fulvestrant in the PALOMA-3 Trial. *Cancer Discov*. 2018;8(11):1390-1403. doi:10.1158/2159-8290.CD-18-0264

68. Inhibition of HER2/neu (erbB-2) and mitogen-activated protein kinases enhances tamoxifen action against HER2-overexpressing, tamoxifen-resistant breast cancer cells - PubMed. Accessed November 10, 2021. https://pubmed-ncbi-nlm-nih-gov.proxy.library.vanderbilt.edu/11059787/

69. Croessmann S, Formisano L, Kinch LN, et al. Combined Blockade of Activating ERBB2 Mutations and ER Results in Synthetic Lethality of ER+/HER2 Mutant Breast Cancer. *Clin Cancer Res Off J Am Assoc Cancer Res*. 2019;25(1):277-289. doi:10.1158/1078-0432.CCR-18-1544

70. Smyth LM, Piha-Paul SA, Won HH, et al. Efficacy and Determinants of Response to HER Kinase Inhibition in HER2-Mutant Metastatic Breast Cancer. *Cancer Discov*. 2020;10(2):198-213. doi:10.1158/2159-8290.CD-19-0966

71. Griffith OL, Spies NC, Anurag M, et al. The prognostic effects of somatic mutations in ER-positive breast cancer. *Nat Commun*. 2018;9(1):3476. doi:10.1038/s41467-018-05914-x

72. Pearson A, Proszek P, Pascual J, et al. Inactivating NF1 Mutations Are Enriched in Advanced Breast Cancer and Contribute to Endocrine Therapy Resistance. *Clin Cancer Res Off J Am Assoc Cancer Res*. 2020;26(3):608-622. doi:10.1158/1078-0432.CCR-18-4044

73. Magnani L, Stoeck A, Zhang X, et al. Genome-wide reprogramming of the chromatin landscape underlies endocrine therapy resistance in breast cancer. *Proc Natl Acad Sci U S A*. 2013;110(16):E1490-9. doi:10.1073/pnas.1219992110

74. Ross-Innes CS, Stark R, Teschendorff AE, et al. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature*. Published online 2012. doi:10.1038/nature10730

75. Ellis MJ, Ding L, Shen D, et al. Whole Genome Analysis Informs Breast Cancer Response to Aromatase Inhibition. *Nature*. 2012;486(7403):353-360. doi:10.1038/nature11143

76. Mooney SD, Klein TE. The functional importance of disease-associated mutation. *BMC Bioinformatics*. 2002;3(1):24. doi:10.1186/1471-2105-3-24

77. Mutational landscape and significance across 12 major cancer types | Nature. Accessed July 10, 2021. https://www.nature.com/articles/nature12634

78. Calo E, Wysocka J. Modification of Enhancer Chromatin: What, How, and Why? *Mol Cell*. 2013;49(5):825-837. doi:10.1016/j.molcel.2013.01.038

79. Rickels R, Shilatifard A. Enhancer Logic and Mechanics in Development and Disease. *Trends Cell Biol*. 2018;28(8):608-630. doi:10.1016/j.tcb.2018.04.003

80. Jones PA, Issa JPJ, Baylin S. Targeting the cancer epigenome for therapy. *Nat Rev Genet*. 2016;17(10):630-641. doi:10.1038/nrg.2016.93

81. Messier TL, Gordon JAR, Boyd JR, et al. Histone H3 lysine 4 acetylation and methylation dynamics define breast cancer subtypes. *Oncotarget*. 2016;7(5):5094-5109. doi:10.18632/oncotarget.6922

82. Dynan WS, Tjian R. The promoter-specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter. *Cell*. 1983;35(1):79-87. doi:10.1016/0092-8674(83)90210-6

83. Anthony Weil P, Luse DS, Segall J, Roeder RG. Selective and accurate initiation of transcription at the ad2 major late promotor in a soluble system dependent on purified rna polymerase ii and dna. *Cell*. 1979;18(2):469-484. doi:10.1016/0092-8674(79)90065-5

84. Nacev BA, Feng L, Bagert JD, et al. The expanding landscape of "oncohistone" mutations in human cancers. *Nature*. 2019;567(7749):473-479. doi:10.1038/s41586-019-1038-1

85. Ciriello G, Gatza ML, Beck AH, et al. Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. *Cell*. 2015;163(2):506-519. doi:10.1016/j.cell.2015.09.033

86. Chen X, Zhang G, Chen B, et al. Association between histone lysine methyltransferase KMT2C mutation and clinicopathological factors in breast cancer. *Biomed Pharmacother*. 2019;116:108997. doi:10.1016/j.biopha.2019.108997

87. Liu L, Kimball S, Liu H, Holowatyj A, Yang ZQ. Genetic alterations of histone lysine methyltransferases and their significance in breast cancer. *Oncotarget*. 2014;6(4):2466-2482. doi:10.18632/oncotarget.2967

88. Melloni GEM, Mazzarella L, Bernard L, et al. A knowledge-based framework for the discovery of cancer-predisposing variants using large-scale sequencing breast cancer data. *Breast Cancer Res BCR*. 2017;19(1):63. doi:10.1186/s13058-017-0854-1

89. Hu D, Gao X, Morgan MA, Herz HM, Smith ER, Shilatifard A. The MLL3/MLL4 branches of the COMPASS family function as major histone H3K4 monomethylases at enhancers. *Mol Cell Biol*. 2013;33(23):4745-4754. doi:10.1128/MCB.01181-13

90. Lai B, Lee JE, Jang Y, Wang L, Peng W, Ge K. MLL3/MLL4 are required for CBP/p300 binding on enhancers and super-enhancer formation in brown adipogenesis. *Nucleic Acids Res*. 2017;45(11):6388-6403. doi:10.1093/nar/gkx234

91. Wang C, Lee JE, Lai B, et al. Enhancer priming by H3K4 methyltransferase MLL4 controls cell fate transition. *Proc Natl Acad Sci*. 2016;113(42):11871-11876. doi:10.1073/pnas.1606857113

92. Wang SP, Tang Z, Chen CW, et al. A UTX-MLL4-p300 Transcriptional Regulatory Network Coordinately Shapes Active Enhancer Landscapes for Eliciting Transcription. *Mol Cell*. 2017;67(2):308-321.e6. doi:10.1016/j.molcel.2017.06.028

93. Miranda-Rojas S, Blanco-Esperguez K, Tuñón I, Kästner J, Mendizábal F. Exploration of the Activation Mechanism of the Epigenetic Regulator MLL3: A QM/MM Study. *Biomolecules*. 2021;11(7):1051. doi:10.3390/biom11071051

94. Kim SS, Lee MH, Lee MO. Histone methyltransferases regulate the transcriptional expression of ERα and the proliferation of tamoxifen-resistant breast cancer cells. *Breast Cancer Res Treat*. 2020;180(1):45-54. doi:10.1007/s10549-019-05517-0

95. 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. doi:10.1016/j.celrep.2016.11.028

96. Gala K, Li Q, Sinha A, et al. KMT2C mediates the estrogen dependence of breast cancer through regulation of ERα enhancer function. *Oncogene*. 2018;37(34):4692-4710. doi:10.1038/s41388-018-0273-5

97. Angus L, Smid M, Wilting SM, et al. The genomic landscape of metastatic breast cancer highlights changes in mutation and signature frequencies. *Nat Genet*. 2019;51(10):1450-1458. doi:10.1038/s41588-019-0507-7

98. Turner N, Pearson A, Sharpe R, et al. FGFR1 Amplification Drives Endocrine Therapy Resistance and Is a Therapeutic Target in Breast Cancer. *Cancer Res*. 2010;70(5):2085-2094. doi:10.1158/0008-5472.CAN-09-3746

99. Badia E, Oliva J, Balaguer P, Cavaillès V. Tamoxifen resistance and epigenetic modifications in breast cancer cell lines. *Curr Med Chem*. 2007;14(28):3035-3045. doi:10.2174/092986707782794023

100. Fiorillo M, Sanchez-Alvarez R, Sotgia F, Lisanti MP. The ER-alpha mutation Y537S confers Tamoxifen-resistance via enhanced mitochondrial metabolism, glycolysis and Rho-GDI/PTEN signaling: Implicating TIGAR in somatic resistance to endocrine therapy. *Aging*. 2018;10(12):4000-4023. doi:10.18632/aging.101690

101. Hao L, Rizzo P, Osipo C, et al. Notch-1 activates estrogen receptor-alpha-dependent transcription via IKKalpha in breast cancer cells. *Oncogene*. 2010;29(2):201-213. doi:10.1038/onc.2009.323

102. Lee S, Lee DK, Dou Y, et al. Coactivator as a target gene specificity determinant for histone H3 lysine 4 methyltransferase. *Proc Natl Acad Sci U S A*. 2006;103(42):15392-15397. doi:10.1073/pnas.0607313103

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

104. Ansari KI, Hussain I, Kasiri S, Mandal SS. HOXC10 is overexpressed in breast cancer and transcriptionally regulated by estrogen via involvement of histone methylases MLL3 and MLL4. *J Mol Endocrinol*. 2012;48(1):61-75. doi:10.1530/JME-11-0078

105. Chen C, Liu Y, Rappaport AR, et al. MLL3 Is a Haploinsufficient 7q Tumor Suppressor in Acute Myeloid Leukemia. *Cancer Cell*. 2014;25(5):652-665. doi:10.1016/J.CCR.2014.03.016

106. Gellert P, Segal C V., Gao Q, et al. Impact of mutational profiles on response of primary oestrogen receptor-positive breast cancers to oestrogen deprivation. *Nat Commun*. 2016;7:13294. doi:10.1038/ncomms13294

107. Martincorena I, Raine KM, Gerstung M, et al. Universal Patterns of Selection in Cancer and Somatic Tissues. *Cell*. 2017;171(5):1029-1041.e21. doi:10.1016/j.cell.2017.09.042

108. Gonzalez-Perez A, Jene-Sanz A, Lopez-Bigas N. The mutational landscape of chromatin regulatory factors across 4,623 tumor samples. *Genome Biol*. 2013;14(9):r106. doi:10.1186/gb-2013-14-9-r106

109. Wang L, Zhao Z, Ozark PA, et al. Resetting the epigenetic balance of Polycomb and COMPASS function at enhancers for cancer therapy. *Nat Med*. 2018;24(6):758-769. doi:10.1038/s41591-018-0034-6

110. Zhang Z, Christin JR, Wang C, Ge K, Oktay MH, Guo W. *Mammary-Stem-Cell-Based Somatic Mouse Models Reveal Breast Cancer Drivers Causing Cell Fate Dysregulation*. Vol 16.; 2016. doi:10.1016/j.celrep.2016.08.048

111. Hu ZY, Xie N, Tian C, et al. Identifying Circulating Tumor DNA Mutation Profiles in Metastatic Breast Cancer Patients with Multiline Resistance. *EBioMedicine*. 2018;32:111-118. doi:10.1016/j.ebiom.2018.05.015

112. Cooper GM, Stone EA, Asimenos G, et al. Distribution and intensity of constraint in mammalian genomic sequence. *Genome Res*. 2005;15(7):901-913. doi:10.1101/gr.3577405

113. Ludwik KA, McDonald OG, Brenin DR, Lannigan DA. ERα-Mediated Nuclear Sequestration of RSK2 Is Required for ER+ Breast Cancer Tumorigenesis. *Cancer Res*. 2018;78(8):2014-2025. doi:10.1158/0008-5472.CAN-17-2063

114. Wiznerowicz M, Trono D. Conditional suppression of cellular genes: lentivirus vector-mediated drug-inducible RNA interference. *J Virol*. 2003;77(16):8957-8961. doi:10.1128/jvi.77.16.8957-8951.2003

115. Yang X, Boehm JS, Yang X, et al. A public genome-scale lentiviral expression library of human ORFs. *Nat Methods*. 2011;8(8):659-661. doi:10.1038/nmeth.1638

116. Groehler AL, Lannigan DA. A chromatin-bound kinase, ERK8, protects genomic integrity by inhibiting HDM2-mediated degradation of the DNA clamp PCNA. *J Cell Biol*. 2010;190(4):575-586. doi:10.1083/jcb.201002124

117. Trapnell C, Roberts A, Goff L, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc*. 2012;7(3):562-578. doi:10.1038/nprot.2012.016

118. Leek JT, Storey JD. Capturing Heterogeneity in Gene Expression Studies by Surrogate Variable Analysis. *PLOS Genet*. 2007;3(9):e161. doi:10.1371/journal.pgen.0030161

119. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinforma Oxf Engl*. 2009;25(14):1754-1760. doi:10.1093/bioinformatics/btp324

120. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinforma Oxf Engl*. 2009;25(16):2078-2079. doi:10.1093/bioinformatics/btp352

121. Landt SG, Marinov GK, Kundaje A, et al. ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res*. 2012;22(9):1813-1831. doi:10.1101/gr.136184.111

122. Kharchenko PV, Tolstorukov MY, Park PJ. Design and analysis of ChIP-seq experiments for DNA-binding proteins. *Nat Biotechnol*. 2008;26(12):1351-1359. doi:10.1038/nbt.1508

123. Li Q, Brown JB, Huang H, Bickel PJ. Measuring reproducibility of high-throughput experiments. *Ann Appl Stat*. 2011;5(3):1752-1779. doi:10.1214/11-AOAS466

124. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010;26(6):841-842. doi:10.1093/bioinformatics/btq033

125. Lettice LA, Heaney SJH, Purdie LA, et al. A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Hum Mol Genet*. 2003;12(14):1725-1735. doi:10.1093/hmg/ddg180

126. McLean CY, Bristor D, Hiller M, et al. GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol*. 2010;28(5). doi:10.1038/nbt.1630

127. Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res*. 2019;47(W1):W199-W205. doi:10.1093/nar/gkz401

128. Otasek D, Morris JH, Bouças J, Pico AR, Demchak B. Cytoscape Automation: empowering workflow-based network analysis. *Genome Biol*. 2019;20(1):185. doi:10.1186/s13059-019-1758-4

129. Bailey TL, Boden M, Buske FA, et al. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res*. 2009;37(Web Server):W202-W208. doi:10.1093/nar/gkp335

130. Tsherniak A, Vazquez F, Montgomery PG, et al. Defining a Cancer Dependency Map. *Cell*. 2017;170(3):564-576.e16. doi:10.1016/j.cell.2017.06.010

131. Robinson JT, Thorvaldsdóttir H, Winckler W, et al. Integrative Genomics Viewer. *Nat Biotechnol*. 2011;29(1):24-26. doi:10.1038/nbt.1754

132. Leek J, Johnson W, Parker H, et al. sva: Surrogate Variable Analysis. Published online 2021.

133. Storey JD. A direct approach to false discovery rates. *J R Stat Soc Ser B Stat Methodol*. 2002;64(3):479-498. doi:10.1111/1467-9868.00346

134. Shen L, Sinai ISoMaM. *GeneOverlap: Test and Visualize Gene Overlaps*. Bioconductor version: Release (3.13); 2021. doi:10.18129/B9.bioc.GeneOverlap

135. Cerami E, Gao J, Dogrusoz U, et al. The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. *Cancer Discov*. 2012;2(5).

136. Gao J, Aksoy BA, Dogrusoz U, et al. Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. *SciSignal*. 2013;6(269):l1. doi:10.1016/j.immuni.2010.12.017.Two-stage

137. Aran D, Sirota M, Butte AJ. Systematic pan-cancer analysis of tumour purity. *Nat Commun*. 2015;6(1):8971. doi:10.1038/ncomms9971

138. Cunningham SC, Gallmeier E, Hucl T, et al. Theoretical Proposal: Allele Dosage of MAP2K4/MKK4 Could Rationalize Frequent 17p Loss in Diverse Human Cancers. *Cell Cycle*. 2006;5(10):1090-1093. doi:10.4161/cc.5.10.2805

139. Wang L, Pan Y, Dai J Le. Evidence of MKK4 pro-oncogenic activity in breast and pancreatic tumors. *Oncogene*. 2004;23(35):5978-5985. doi:10.1038/sj.onc.1207802

140. Cooper GM, Goode DL, Ng SB, et al. Single-nucleotide evolutionary constraint scores highlight disease-causing mutations. *Nat Methods*. 2010;7(4):250-251. doi:10.1038/nmeth0410-250

141. Goo YH, Sohn YC, Kim DH, et al. Activating signal cointegrator 2 belongs to a novel steady-state complex that contains a subset of trithorax group proteins. *Mol Cell Biol*. 2003;23(1):140-149.

142. Hu D, Gao X, Morgan MA, Herz HM, Smith ER, Shilatifard A. The MLL3/MLL4 branches of the COMPASS family function as major histone H3K4 monomethylases at enhancers. *Mol Cell Biol*. 2013;33(23):4745-4754. doi:10.1128/MCB.01181-13

143. Diaz A, Park K, Lim DA, Song JS. Normalization, bias correction, and peak calling for ChIP-seq. *Stat Appl Genet Mol Biol*. 2012;11(3). doi:10.1515/1544-6115.1750

144. Stark R, Brown G. DiffBind: Differential binding analysis of ChIP-Seq peak data. :75.

145. Creighton CJ. A gene transcription signature of the Akt/mTOR pathway in clinical breast tumors. *Oncogene*. 2007;26(32):4648-4655. doi:10.1038/sj.onc.1210245

146. Leiserson MDM, Vandin F, Wu HT, et al. Pan-cancer network analysis identifies combinations of rare somatic mutations across pathways and protein complexes. *Nat Genet*. 2015;47(2):106-114. doi:10.1038/ng.3168

147. Sonnenblick A, Venet D, Brohée S, Pondé N, Sotiriou C. pAKT pathway activation is associated with PIK3CA mutations and good prognosis in luminal breast cancer in contrast to p-mTOR pathway activation. *NPJ Breast Cancer*. 2019;5:7. doi:10.1038/s41523-019-0102-1

148. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2017. *CA Cancer J Clin*. 2017;67(1):7-30. doi:10.3322/caac.21387

149. Cuzick J, Sestak I, Baum M, et al. Effect of anastrozole and tamoxifen as adjuvant treatment for early-stage breast cancer: 10-year analysis of the ATAC trial. *Lancet Oncol*. 2010;11(12):1135-1141. doi:10.1016/S1470-2045(10)70257-6

150. Howell A, Cuzick J, Baum M, et al. Results of the ATAC (Arimidex, Tamoxifen, Alone or in Combination) trial after completion of 5 years' adjuvant treatment for breast cancer. *Lancet Lond Engl*. 2005;365(9453):60-62. doi:10.1016/S0140-6736(04)17666-6

151. Dowsett M, Smith IE, Ebbs SR, et al. Prognostic value of Ki67 expression after short-term presurgical endocrine therapy for primary breast cancer. *J Natl Cancer Inst*. 2007;99(2):167-170. doi:10.1093/jnci/djk020

152. Giltnane JM, Hutchinson KE, Stricker TP, et al. Genomic profiling of ER+ breast cancers after short-term estrogen suppression reveals alterations associated with endocrine resistance. *Sci Transl Med*. 2017;9(402):eaai7993. doi:10.1126/scitranslmed.aai7993

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

154. Johnston S, Pippen J, Pivot X, et al. Lapatinib combined with letrozole versus letrozole and placebo as first-line therapy for postmenopausal hormone receptor-positive metastatic breast cancer. *J Clin Oncol Off J Am Soc Clin Oncol*. 2009;27(33):5538-5546. doi:10.1200/JCO.2009.23.3734

155. Arpino G, Green SJ, Allred DC, et al. HER-2 amplification, HER-1 expression, and tamoxifen response in estrogen receptor-positive metastatic breast cancer: a southwest oncology group study. *Clin Cancer Res Off J Am Assoc Cancer Res*. 2004;10(17):5670-5676. doi:10.1158/1078-0432.CCR-04-0110

156. Curigliano G, Bagnardi V, Viale G, et al. Should liver metastases of breast cancer be biopsied to improve treatment choice? *Ann Oncol Off J Eur Soc Med Oncol*. 2011;22(10):2227-2233. doi:10.1093/annonc/mdq751

157. Gong Y, Han EY, Guo M, Pusztai L, Sneige N. Stability of estrogen receptor status in breast carcinoma: a comparison between primary and metastatic tumors with regard to disease course and intervening systemic therapy. *Cancer*. 2011;117(4):705-713. doi:10.1002/cncr.25506

158. Perey L, Paridaens R, Hawle H, et al. Clinical benefit of fulvestrant in postmenopausal women with advanced breast cancer and primary or acquired resistance to aromatase inhibitors: final results of phase II Swiss Group for Clinical Cancer Research Trial (SAKK 21/00). *Ann Oncol*. 2007;18(1):64-69. doi:10.1093/annonc/mdl341

159. Ingle JN, Suman VJ, Rowland KM, et al. Fulvestrant in women with advanced breast cancer after progression on prior aromatase inhibitor therapy: North Central Cancer Treatment Group Trial N0032. *J Clin Oncol Off J Am Soc Clin Oncol*. 2006;24(7):1052-1056. doi:10.1200/JCO.2005.04.1053

160. Davies C, Pan H, Godwin J, et al. Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial. *Lancet Lond Engl*. 2013;381(9869):805-816. doi:10.1016/S0140-6736(12)61963-1

161. Goss PE, Ingle JN, Martino S, et al. A randomized trial of letrozole in postmenopausal women after five years of tamoxifen therapy for early-stage breast cancer. *N Engl J Med*. 2003;349(19):1793-1802. doi:10.1056/NEJMoa032312

162. Rao RC, Dou Y. Hijacked in cancer: the MLL/KMT2 family of methyltransferases. *Nat Rev Cancer*. 2015;15(6):334-346. doi:10.1038/nrc3929

163. Heintzman ND, Stuart RK, Hon G, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet*. 2007;39(3):311-318. doi:10.1038/ng1966

164. Shilatifard A. The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annu Rev Biochem*. 2012;81:65-95. doi:10.1146/annurev-biochem-051710-134100

165. Lee J, Kim DH, Lee S, et al. A tumor suppressive coactivator complex of p53 containing ASC-2 and histone H3-lysine-4 methyltransferase MLL3 or its paralogue MLL4. *Proc Natl Acad Sci*. 2009;106(21):8513-8518. doi:10.1073/pnas.0902873106

166. Bhan A, Hussain I, Ansari KI, Bobzean SAM, Perrotti LI, Mandal SS. Bisphenol-A and diethylstilbestrol exposure induces the expression of breast cancer associated long noncoding RNA HOTAIR in vitro and in vivo. *J Steroid Biochem Mol Biol*. 2014;141:160-170. doi:10.1016/j.jsbmb.2014.02.002

167. Kim JH, Sharma A, Dhar SS, et al. UTX and MLL4 Coordinately Regulate Transcriptional Programs for Cell Proliferation and Invasiveness in Breast Cancer Cells. *Cancer Res*. 2014;74(6):1705-1717. doi:10.1158/0008-5472.CAN-13-1896

168. Otto T, Sicinski P. Cell cycle proteins as promising targets in cancer therapy. *Nat Rev Cancer*. 2017;17(2):93-115. doi:10.1038/nrc.2016.138

169. Musgrove EA, Sutherland RL. Biological determinants of endocrine resistance in breast cancer. *Nat Rev Cancer*. 2009;9(9):631-643. doi:10.1038/nrc2713

170. Thangavel C, Dean JL, Ertel A, et al. Therapeutically activating RB: reestablishing cell cycle control in endocrine therapy-resistant breast cancer. *Endocr Relat Cancer*. 2011;18(3):333-345. doi:10.1530/ERC-10-0262

171. Finn RS, Dering J, Conklin D, et al. PD 0332991, a selective cyclin D kinase 4/6 inhibitor, preferentially inhibits proliferation of luminal estrogen receptor-positive human breast cancer cell lines in vitro. *Breast Cancer Res*. 2009;11(5):R77. doi:10.1186/bcr2419

172. DeMichele A, Clark AS, Heitjan D, et al. A phase II trial of an oral CDK 4/6 inhibitor, PD0332991, in advanced breast cancer. *J Clin Oncol*. 2013;31(15_suppl):519-519. doi:10.1200/jco.2013.31.15_suppl.519

173. Finn RS, Crown JP, Lang I, et al. The cyclin-dependent kinase 4/6 inhibitor palbociclib in combination with letrozole versus letrozole alone as first-line treatment of oestrogen receptor-positive, HER2-negative, advanced breast cancer (PALOMA-1/TRIO-18): a randomised phase 2 study. *Lancet Oncol*. 2015;16(1):25-35. doi:10.1016/S1470-2045(14)71159-3

174. Finn RS, Martin M, Rugo HS, et al. Palbociclib and Letrozole in Advanced Breast Cancer. *N Engl J Med*. 2016;375(20):1925-1936. doi:10.1056/NEJMoa1607303

175. Nouri M, Massah S, Caradec J, et al. Transient Sox9 Expression Facilitates Resistance to Androgen-Targeted Therapy in Prostate Cancer. *Clin Cancer Res*. 2020;26(7):1678-1689. doi:10.1158/1078-0432.CCR-19-0098

176. Gao J, Zhang JY, Li YH, Ren F. Decreased expression of SOX9 indicates a better prognosis and inhibits the growth of glioma cells by inducing cell cycle arrest. *Int J Clin Exp Pathol*. 2015;8(9):10130-10138.

177. Knudsen ES, Pruitt SC, Hershberger PA, Witkiewicz AK, Goodrich DW. Cell Cycle and Beyond: Exploiting New RB1 Controlled Mechanisms for Cancer Therapy. *Trends Cancer*. 2019;5(5):308-324. doi:10.1016/j.trecan.2019.03.005

178. Iaquinta PJ, Aslanian A, Lees JA. Regulation of the Arf/p53 tumor surveillance network by E2F. *Cold Spring Harb Symp Quant Biol*. 2005;70:309-316. doi:10.1101/sqb.2005.70.050

179. Herschkowitz JI, He X, Fan C, Perou CM. The functional loss of the retinoblastoma tumour suppressor is a common event in basal-like and luminal B breast carcinomas. *Breast Cancer Res*. 2008;10(5):R75. doi:10.1186/bcr2142

180. Hess KR, Anderson K, Symmans WF, et al. Pharmacogenomic predictor of sensitivity to preoperative chemotherapy with paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide in breast cancer. *J Clin Oncol Off J Am Soc Clin Oncol*. 2006;24(26):4236-4244. doi:10.1200/JCO.2006.05.6861

181. Chen HZ, Tsai SY, Leone G. Emerging roles of E2Fs in cancer: an exit from cell cycle control. *Nat Rev Cancer*. 2009;9(11):785-797. doi:10.1038/nrc2696

182. O'Brien NA, McDermott MSJ, Conklin D, et al. Targeting activated PI3K/mTOR signaling overcomes acquired resistance to CDK4/6-based therapies in preclinical models of hormone receptor-positive breast cancer. *Breast Cancer Res*. 2020;22(1):89. doi:10.1186/s13058-020-01320-8

183. Courjal F, Cuny M, Simony-Lafontaine J, et al. Mapping of DNA Amplifications at 15 Chromosomal Localizations in 1875 Breast Tumors: Definition of Phenotypic Groups. *Cancer Res*. 1997;57(19):4360-4367.

184. Jacquemier J, Adelaide J, Parc P, et al. Expression of the FGFR1 gene in human breast-carcinoma cells. *Int J Cancer*. 1994;59(3):373-378. doi:10.1002/ijc.2910590314

185. Reis-Filho JS, Simpson PT, Turner NC, et al. FGFR1 Emerges as a Potential Therapeutic Target for Lobular Breast Carcinomas. *Clin Cancer Res*. 2006;12(22):6652-6662. doi:10.1158/1078-0432.CCR-06-1164

186. Gorringe KL, Jacobs S, Thompson ER, et al. High-Resolution Single Nucleotide Polymorphism Array Analysis of Epithelial Ovarian Cancer Reveals Numerous Microdeletions and Amplifications. *Clin Cancer Res*. 2007;13(16):4731-4739. doi:10.1158/1078-0432.CCR-07-0502

187. Simon R, Richter J, Wagner U, et al. High-Throughput Tissue Microarray Analysis of 3p25 (RAF1) and 8p12 (FGFR1) Copy Number Alterations in Urinary Bladder Cancer. *Cancer Res*. 2001;61(11):4514-4519.

188. Missiaglia E, Selfe J, Hamdi M, et al. Genomic imbalances in rhabdomyosarcoma cell lines affect expression of genes frequently altered in primary tumors: An approach to identify candidate genes involved in tumor development. *Genes Chromosomes Cancer*. 2009;48(6):455-467. doi:10.1002/gcc.20655

189. Jiang T, Gao G, Fan G, Li M, Zhou C. FGFR1 amplification in lung squamous cell carcinoma: a systematic review with meta-analysis. *Lung Cancer Amst Neth*. 2015;87(1):1-7. doi:10.1016/j.lungcan.2014.11.009

190. Schlessinger J, Plotnikov AN, Ibrahimi OA, et al. Crystal structure of a ternary FGF-FGFR-heparin complex reveals a dual role for heparin in FGFR binding and dimerization. *Mol Cell*. 2000;6(3):743-750. doi:10.1016/s1097-2765(00)00073-3

191. Turner N, Grose R. Fibroblast growth factor signalling: from development to cancer. *Nat Rev Cancer*. 2010;10(2):116-129. doi:10.1038/nrc2780

192. Bryant DM, Stow JL. Nuclear translocation of cell-surface receptors: lessons from fibroblast growth factor. *Traffic Cph Den*. 2005;6(10):947-954. doi:10.1111/j.1600-0854.2005.00332.x

193. Zammit C, Barnard R, Gomm J, et al. Altered intracellular localization of fibroblast growth factor receptor 3 in human breast cancer. *J Pathol*. 2001;194(1):27-34. doi:10.1002/path.846

194. Cerliani JP, Guillardoy T, Giulianelli S, et al. Interaction between FGFR-2, STAT5, and Progesterone Receptors in Breast Cancer. *Cancer Res*. 2011;71(10):3720-3731. doi:10.1158/0008-5472.CAN-10-3074

195. Stachowiak MK, Birkaya B, Aletta JM, et al. Nuclear FGF receptor-1 and CREB binding protein: an integrative signaling module. *J Cell Physiol*. 2015;230(5):989-1002. doi:10.1002/jcp.24879

196. Nuclear translocation of FGFR1 and FGF2 in pancreatic stellate cells facilitates pancreatic cancer cell invasion. *EMBO Mol Med*. 2014;6(4):467-481. doi:10.1002/emmm.201302698

197. Chioni AM, Grose R. FGFR1 cleavage and nuclear translocation regulates breast cancer cell behavior. *J Cell Biol*. 2012;197(6):801-817. doi:10.1083/jcb.201108077

198. Myers JM, Martins GG, Ostrowski J, Stachowiak MK. Nuclear trafficking of FGFR1: a role for the transmembrane domain. *J Cell Biochem*. 2003;88(6):1273-1291. doi:10.1002/jcb.10476

199. Stachowiak MK, Fang X, Myers JM, et al. Integrative nuclear FGFR1 signaling (INFS) as a part of a universal "feed-forward-and-gate" signaling module that controls cell growth and differentiation. *J Cell Biochem*. 2003;90(4):662-691. doi:10.1002/jcb.10606

200. Terranova C, Narla ST, Lee YW, et al. Global Developmental Gene Programing Involves a Nuclear Form of Fibroblast Growth Factor Receptor-1 (FGFR1). *PloS One*. 2015;10(4):e0123380. doi:10.1371/journal.pone.0123380

201. Katoh M. WNT and FGF gene clusters (review). *Int J Oncol*. 2002;21(6):1269-1273.

202. Brady NJ, Chuntova P, Bade LK, Schwertfeger KL. The FGF/FGF receptor axis as a therapeutic target in breast cancer. *Expert Rev Endocrinol Metab*. 2013;8(4):391-402. doi:10.1586/17446651.2013.811910

203. Cuny M, Kramar A, Courjal F, et al. Relating Genotype and Phenotype in Breast Cancer: An Analysis of the Prognostic Significance of Amplification at Eight Different Genes or Loci and of p53 Mutations. *Cancer Res*. 2000;60(4):1077-1083.

204. Balko JM, Mayer IA, Sanders ME, et al. Discordant Cellular Response to Presurgical Letrozole in Bilateral Synchronous ER+ Breast Cancers with a KRAS Mutation or FGFR1 Gene Amplification. *Mol Cancer Ther*. 2012;11(10):2301-2305. doi:10.1158/1535-7163.MCT-12-0511

205. Allred DC, Harvey JM, Berardo M, Clark GM. Prognostic and predictive factors in breast cancer by immunohistochemical analysis. *Mod Pathol Off J U S Can Acad Pathol Inc*. 1998;11(2):155-168.

206. Miller TW, Balko JM, Fox EM, et al. ERα-Dependent E2F Transcription Can Mediate Resistance to Estrogen Deprivation in Human Breast Cancer. *Cancer Discov*. 2011;1(4):338-351. doi:10.1158/2159-8290.CD-11-0101

207. Bhola NE, Jansen VM, Bafna S, et al. Kinome-wide Functional Screen Identifies Role of PLK1 in Hormone-Independent, ER-Positive Breast Cancer. *Cancer Res*. 2015;75(2):405-414. doi:10.1158/0008-5472.CAN-14-2475

208. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*. 2013;14(4):R36. doi:10.1186/gb-2013-14-4-r36

209. DeLuca DS, Levin JZ, Sivachenko A, et al. RNA-SeQC: RNA-seq metrics for quality control and process optimization. *Bioinforma Oxf Engl*. 2012;28(11):1530-1532. doi:10.1093/bioinformatics/bts196

210. Debnath J, Muthuswamy SK, Brugge JS. Morphogenesis and oncogenesis of MCF-10A mammary epithelial acini grown in three-dimensional basement membrane cultures. *Methods San Diego Calif*. 2003;30(3):256-268. doi:10.1016/s1046-2023(03)00032-x

211. Miller TW, Hennessy BT, González-Angulo AM, et al. Hyperactivation of phosphatidylinositol-3 kinase promotes escape from hormone dependence in estrogen receptor-positive human breast cancer. *J Clin Invest*. 2010;120(7):2406-2413. doi:10.1172/JCI41680

212. Dixon JM, Renshaw L, Young O, et al. Letrozole suppresses plasma estradiol and estrone sulphate more completely than anastrozole in postmenopausal women with breast cancer. *J Clin Oncol Off J Am Soc Clin Oncol*. 2008;26(10):1671-1676. doi:10.1200/JCO.2007.13.9279

213. Wolff B, Sanglier JJ, Wang Y. Leptomycin B is an inhibitor of nuclear export: inhibition of nucleo-cytoplasmic translocation of the human immunodeficiency virus type 1 (HIV-1) Rev protein and Rev-dependent mRNA. *Chem Biol*. 1997;4(2):139-147. doi:10.1016/s1074-5521(97)90257-x

214. Kwek SS, Roy R, Zhou H, et al. Co-amplified genes at 8p12 and 11q13 in breast tumors cooperate with two major pathways in oncogenesis. *Oncogene*. 2009;28(17):1892-1903. doi:10.1038/onc.2009.34

215. Győrffy B, Lanczky A, Eklund AC, et al. An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. *Breast Cancer Res Treat*. 2010;123(3):725-731. doi:10.1007/s10549-009-0674-9

216. Bello E, Colella G, Scarlato V, et al. E-3810 Is a Potent Dual Inhibitor of VEGFR and FGFR that Exerts Antitumor Activity in Multiple Preclinical Models. *Cancer Res*. 2011;71(4):1396-1405. doi:10.1158/0008-5472.CAN-10-2700

217. Dunham-Ems SM, Lee YW, Stachowiak EK, et al. Fibroblast Growth Factor Receptor-1 (FGFR1) Nuclear Dynamics Reveal a Novel Mechanism in Transcription Control. *Mol Biol Cell*. 2009;20(9):2401-2412. doi:10.1091/mbc.e08-06-0600

218. Sosa L del V, Gutiérrez S, Petiti JP, Vaca AM, De Paul AL, Torres AI. Cooperative effect of E2 and FGF2 on lactotroph proliferation triggered by signaling initiated at the plasma membrane. *Am J Physiol-Endocrinol Metab*. 2013;305(1):E41-E49. doi:10.1152/ajpendo.00027.2013

219. Liu PC, Wu L, Koblish H, et al. Abstract 771: Preclinical characterization of the selective FGFR inhibitor INCB054828. *Cancer Res*. 2015;75(15 Supplement):771-771. doi:10.1158/1538-7445.AM2015-771

220. Somanathan S, Stachowiak EK, Siegel AJ, Stachowiak MK, Berezney R. Nuclear matrix bound fibroblast growth factor receptor is associated with splicing factor rich and transcriptionally active nuclear speckles. *J Cell Biochem*. 2003;90(4):856-869. doi:10.1002/jcb.10672

221. Romisch K. Surfing the Sec61 channel: bidirectional protein translocation across the ER membrane. *J Cell Sci*. 1999;112(23):4185-4191. doi:10.1242/jcs.112.23.4185

222. Bryant DM, Wylie FG, Stow JL. Regulation of Endocytosis, Nuclear Translocation, and Signaling of Fibroblast Growth Factor Receptor 1 by E-Cadherin. *Mol Biol Cell*. 2005;16(1):14-23. doi:10.1091/mbc.e04-09-0845

223. Reilly JF, Maher PA. Importin β–Mediated Nuclear Import of Fibroblast Growth Factor Receptor: Role in Cell Proliferation. *J Cell Biol*. 2001;152(6):1307-1312. doi:10.1083/jcb.152.6.1307

224. Więdłocha A, Falnes PØ, Madshus IH, Sandvig K, Olsnes S. Dual mode of signal transduction by externally added acidic fibroblast growth factor. *Cell*. 1994;76(6):1039-1051. doi:10.1016/0092-8674(94)90381-6

225. Barros RPA, Gustafsson JÅ. Estrogen Receptors and the Metabolic Network. *Cell Metab*. 2011;14(3):289-299. doi:10.1016/j.cmet.2011.08.005

226. Lovett JL, Chima MA, Wexler JK, et al. Oral contraceptives cause evolutionarily novel increases in hormone exposure: A risk factor for breast cancer. *Evol Med Public Health*. 2017;2017(1):97-108. doi:10.1093/emph/eox009

227. Walmer DK, Wrona MA, Hughes CL, Nelson KG. Lactoferrin expression in the mouse reproductive tract during the natural estrous cycle: correlation with circulating estradiol and progesterone. *Endocrinology*. 1992;131(3):1458-1466. doi:10.1210/en.131.3.1458

228. Vogel PM, Georgiade NG, Fetter BF, Vogel FS, McCarty KS. The correlation of histologic changes in the human breast with the menstrual cycle. *Am J Pathol*. 1981;104(1):23-34.

229. Jia M, Dahlman-Wright K, Gustafsson JÅ. Estrogen receptor alpha and beta in health and disease. *Best Pract Res Clin Endocrinol Metab*. 2015;29(4):557-568. doi:10.1016/j.beem.2015.04.008

230. Feng Y, Manka D, Wagner KU, Khan SA. Estrogen receptor-α expression in the mammary epithelium is required for ductal and alveolar morphogenesis in mice. *Proc Natl Acad Sci*. 2007;104(37):14718-14723. doi:10.1073/pnas.0706933104

231. Förster C, Mäkela S, Wärri A, et al. Involvement of estrogen receptor β in terminal differentiation of mammary gland epithelium. *Proc Natl Acad Sci*. 2002;99(24):15578-15583. doi:10.1073/pnas.192561299

232. 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. *Mol Cell*. 2000;5(6):939-948. doi:10.1016/S1097-2765(00)80259-2

233. Nawaz Z, Lonard DM, Dennis AP, Smith CL, O'Malley BW. Proteasome-dependent degradation of the human estrogen receptor. *Proc Natl Acad Sci*. 1999;96(5):1858-1862. doi:10.1073/pnas.96.5.1858

234. Reid G, Hübner MR, Métivier R, et al. Cyclic, Proteasome-Mediated Turnover of Unliganded and Liganded ERα on Responsive Promoters Is an Integral Feature of Estrogen Signaling. *Mol Cell*. 2003;11(3):695-707. doi:10.1016/S1097-2765(03)00090-X

235. Silberstein GB, Horn KV, Hrabeta-Robinson E, Compton J. Estrogen-triggered delays in mammary gland gene expression during the estrous cycle: evidence for a novel timing system. *J Endocrinol*. 190(2):225-239. doi:10.1677/joe.1.06725

236. The catalytic subunit of the proteasome is engaged in the entire process of estrogen receptor-regulated transcription. *EMBO J*. 2006;25(18):4223-4233. doi:10.1038/sj.emboj.7601306

237. Blair RM, Fang H, Branham WS, et al. The Estrogen Receptor Relative Binding Affinities of 188 Natural and Xenochemicals: Structural Diversity of Ligands. *Toxicol Sci*. 2000;54(1):138-153. doi:10.1093/toxsci/54.1.138

238. Valley CC, Solodin NM, Powers GL, Ellison SJ, Alarid ET. Temporal variation in estrogen receptor-α protein turnover in the presence of estrogen. *J Mol Endocrinol*. 2008;40(1):23-34. doi:10.1677/JME-07-0067

239. Zhou W, Slingerland JM. Links between oestrogen receptor activation and proteolysis: relevance to hormone-regulated cancer therapy. *Nat Rev Cancer*. 2014;14(1):26-38. doi:10.1038/nrc3622

240. Deutsch MB, Bhakri V, Kubicek K. Effects of Cross-Sex Hormone Treatment on Transgender Women and Men. *Obstet Gynecol*. 2015;125(3):605-610. doi:10.1097/AOG.0000000000000692

241. Xie W, Paterson AJ, Chin E, Nabell LM, Kudlow JE. Targeted Expression of a Dominant Negative Epidermal Growth Factor Receptor in the Mammary Gland of Transgenic Mice Inhibits Pubertal Mammary Duct Development. *Mol Endocrinol*. 1997;11(12):1766-1781. doi:10.1210/mend.11.12.0019

242. Pasic L, Eisinger-Mathason TSK, Velayudhan BT, et al. Sustained activation of the HER1–ERK1/2–RSK signaling pathway controls myoepithelial cell fate in human mammary tissue. *Genes Dev*. 2011;25(15):1641-1653. doi:10.1101/gad.2025611

243. Maennling AE, Tur MK, Niebert M, et al. Molecular Targeting Therapy against EGFR Family in Breast Cancer: Progress and Future Potentials. *Cancers*. 2019;11(12):1826. doi:10.3390/cancers11121826

244. 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. doi:10.1128/MCB.18.4.1978

245. Kato S, Endoh H, Masuhiro Y, et al. Activation of the Estrogen Receptor Through Phosphorylation by Mitogen-Activated Protein Kinase. *Science*. 1995;270(5241):1491-1494. doi:10.1126/science.270.5241.1491

246. Yang X, Matsuda K, Bialek P, et al. ATF4 Is a Substrate of RSK2 and an Essential Regulator of Osteoblast Biology: Implication for Coffin-Lowry Syndrome. *Cell*. 2004;117(3):387-398. doi:10.1016/S0092-8674(04)00344-7

247. Wood GA, Fata JE, Watson KLM, Khokha R. Circulating hormones and estrous stage predict cellular and stromal remodeling in murine uterus. *Reproduction*. 2007;133(5):1035-1044. doi:10.1530/REP-06-0302

248. Byers SL, Wiles MV, Dunn SL, Taft RA. Mouse Estrous Cycle Identification Tool and Images. *PLOS ONE*. 2012;7(4):e35538. doi:10.1371/journal.pone.0035538

249. Brill B, Boecher N, Groner B, Shemanko CS. A sparing procedure to clear the mouse mammary fat pad of epithelial components for transplantation analysis. *Lab Anim*. 2008;42(1):104-110. doi:10.1258/la.2007.06003e

250. Giraddi RR, Shehata M, Gallardo M, Blasco MA, Simons BD, Stingl J. Stem and progenitor cell division kinetics during postnatal mouse mammary gland development. *Nat Commun*. 2015;6(1):8487. doi:10.1038/ncomms9487

251. Pardo I, Lillemoe HA, Blosser RJ, et al. Next-generation transcriptome sequencing of the premenopausal breast epithelium using specimens from a normal human breast tissue bank. *Breast Cancer Res*. 2014;16(2):R26. doi:10.1186/bcr3627

252. Fata JE, Chaudhary V, Khokha R. Cellular Turnover in the Mammary Gland Is Correlated with Systemic Levels of Progesterone and Not 17β-Estradiol During the Estrous Cycle1. *Biol Reprod*. 2001;65(3):680-688. doi:10.1095/biolreprod65.3.680

253. Shehata M, Teschendorff A, Sharp G, et al. Phenotypic and functional characterisation of the luminal cell hierarchy of the mammary gland. *Breast Cancer Res*. 2012;14(5):R134. doi:10.1186/bcr3334

254. Mehta RG, Hawthorne M, Mehta RR, et al. Differential Roles of ERα and ERβ in Normal and Neoplastic Development in the Mouse Mammary Gland. *PLOS ONE*. 2014;9(11):e113175. doi:10.1371/journal.pone.0113175

255. Brisken C, Ataca D. Endocrine hormones and local signals during the development of the mouse mammary gland. *WIREs Dev Biol*. 2015;4(3):181-195. doi:10.1002/wdev.172

256. Hynes NE, Watson CJ. Mammary Gland Growth Factors: Roles in Normal Development and in Cancer. *Cold Spring Harb Perspect Biol*. 2010;2(8):a003186. doi:10.1101/cshperspect.a003186

257. Eisinger-Mathason TSK, Andrade J, Lannigan DA. RSK in tumorigenesis: Connections to steroid signaling. *Steroids*. 2010;75(3):191-202. doi:10.1016/j.steroids.2009.12.010

258. Pinter O, Beda Z, Csaba Z, Gerendai I. Differences in the onset of puberty in selected inbred mouse strains. *Endocr Abstr*. 2007;14. Accessed September 12, 2021. https://www.endocrine-abstracts.org/ea/0014/ea0014p617

259. Dalby KN, Morrice N, Caudwell FB, Avruch J, Cohen P. Identification of Regulatory Phosphorylation Sites in Mitogen-activated Protein Kinase (MAPK)-activated Protein Kinase-1a/p90 rsk That Are Inducible by MAPK *. *J Biol Chem*. 1998;273(3):1496-1505. doi:10.1074/jbc.273.3.1496

260. Dutertre M, Gratadou L, Dardenne E, et al. Estrogen Regulation and Physiopathologic Significance of Alternative Promoters in Breast Cancer. *Cancer Res*. 2010;70(9):3760-3770. doi:10.1158/0008-5472.CAN-09-3988

261. Ludwik KA, Campbell JP, Li M, et al. Development of a RSK Inhibitor as a Novel Therapy for Triple-Negative Breast Cancer. *Mol Cancer Ther*. 2016;15(11):2598-2608. doi:10.1158/1535-7163.MCT-16-0106

262. Li M, Li Y, Ludwik KA, Sandusky ZM, Lannigan DA, O'Doherty GA. Stereoselective Synthesis and Evaluation of C6″-Substituted 5a-Carbasugar Analogues of SL0101 as Inhibitors of RSK1/2. *Org Lett*. 2017;19(9):2410-2413. doi:10.1021/acs.orglett.7b00945

263. Regulation of elongation factor 2 kinase by p90RSK1 and p70 S6 kinase. *EMBO J*. 2001;20(16):4370-4379. doi:10.1093/emboj/20.16.4370

264. Robertshaw I, Bian F, Das SK. Mechanisms of uterine estrogen signaling during early pregnancy in mice: an update. *J Mol Endocrinol*. 2016;56(3):R127-R138. doi:10.1530/JME-15-0300

265. Winuthayanon W, Hewitt SC, Korach KS. Uterine Epithelial Cell Estrogen Receptor Alpha-Dependent and -Independent Genomic Profiles That Underlie Estrogen Responses in Mice1. *Biol Reprod*. 2014;91(5). doi:10.1095/biolreprod.114.120170

266. Eeckhoute J, Keeton EK, Lupien M, Krum SA, Carroll JS, Brown M. Positive Cross-Regulatory Loop Ties GATA-3 to Estrogen Receptor α Expression in Breast Cancer. *Cancer Res*. 2007;67(13):6477-6483. doi:10.1158/0008-5472.CAN-07-0746

267. Uhlen M, Zhang C, Lee S, et al. A pathology atlas of the human cancer transcriptome. *Science*. 2017;357(6352):eaan2507. doi:10.1126/science.aan2507

268. 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. doi:10.1096/fj.14-265637

269. 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. doi:10.1074/jbc.273.21.13317

270. Perillo B, Ombra MN, Bertoni A, et al. DNA Oxidation as Triggered by H3K9me2 Demethylation Drives Estrogen-Induced Gene Expression. *Science*. 2008;319(5860):202-206. doi:10.1126/science.1147674

271. Wentworth CC, Alam A, Jones RM, Nusrat A, Neish AS. Enteric Commensal Bacteria Induce Extracellular Signal-regulated Kinase Pathway Signaling via Formyl Peptide Receptor-dependent Redox Modulation of Dual Specific Phosphatase 3 *. *J Biol Chem*. 2011;286(44):38448-38455. doi:10.1074/jbc.M111.268938

272. Tanaka T, Halicka HD, Huang X, Traganos F, Darzynkiewicz Z. Constitutive Histone H2AX Phosphorylation and ATM Activation, the Reporters of DNA Damage by Endogenous Oxidants. *Cell Cycle*. 2006;5(17):1940-1945. doi:10.4161/cc.5.17.3191

273. Roux PP, Topisirovic I. Signaling Pathways Involved in the Regulation of mRNA Translation. *Mol Cell Biol*. 38(12):e00070-18. doi:10.1128/MCB.00070-18

274. Scharfman HE, MacLusky NJ. Estrogen–Growth Factor Interactions and Their Contributions to Neurological Disorders. *Headache J Head Face Pain*. 2008;48(s2):S77-S89. doi:10.1111/j.1526-4610.2008.01200.x

275. Filipčík P, Curry JR, Mace PD. When Worlds Collide—Mechanisms at the Interface between Phosphorylation and Ubiquitination. *J Mol Biol*. 2017;429(8):1097-1113. doi:10.1016/j.jmb.2017.02.011

276. Kelleher AM, Burns GW, Behura S, Wu G, Spencer TE. Uterine glands impact uterine receptivity, luminal fluid homeostasis and blastocyst implantation. *Sci Rep*. 2016;6(1):38078. doi:10.1038/srep38078

277. Zhang P, Fu Y, Ju J, et al. Estradiol inhibits fMLP-induced neutrophil migration and superoxide production by upregulating MKP-2 and dephosphorylating ERK. *Int Immunopharmacol*. 2019;75:105787. doi:10.1016/j.intimp.2019.105787

278. Denu JM, Tanner KG. Specific and Reversible Inactivation of Protein Tyrosine Phosphatases by Hydrogen Peroxide: Evidence for a Sulfenic Acid Intermediate and Implications for Redox Regulation. *Biochemistry*. 1998;37(16):5633-5642. doi:10.1021/bi973035t

279. Meng TC, Fukada T, Tonks NK. Reversible Oxidation and Inactivation of Protein Tyrosine Phosphatases In Vivo. *Mol Cell*. 2002;9(2):387-399. doi:10.1016/S1097-2765(02)00445-8

280. Jiang J, Sarwar N, Peston D, et al. Phosphorylation of Estrogen Receptor-α at Ser167 Is Indicative of Longer Disease-Free and Overall Survival in Breast Cancer Patients. *Clin Cancer Res*. 2007;13(19):5769-5776. doi:10.1158/1078-0432.CCR-07-0822

281. Ludwik KA, McDonald OG, Brenin DR, Lannigan DA. ERα-Mediated Nuclear Sequestration of RSK2 Is Required for ER+ Breast Cancer Tumorigenesis. *Cancer Res*. 2018;78(8):2014-2025. doi:10.1158/0008-5472.CAN-17-2063

282. Moon HG, Yi JK, Kim HS, et al. Phosphorylation of p90RSK is associated with increased response to neoadjuvant chemotherapy in ER-positive breast cancer. *BMC Cancer*. 2012;12(1):585. doi:10.1186/1471-2407-12-585

283. Chen D, Riedl T, Washbrook E, et al. Activation of Estrogen Receptor α by S118 Phosphorylation Involves a Ligand-Dependent Interaction with TFIIH and Participation of CDK7. *Mol Cell*. 2000;6(1):127-137. doi:10.1016/S1097-2765(05)00004-3

284. González L, Zambrano A, Lazaro-Trueba I, et al. Activation of the unliganded estrogen receptor by prolactin in breast cancer cells. *Oncogene*. 2009;28(10):1298-1308. doi:10.1038/onc.2008.473

285. Park J, Lee Y. Hypoxia induced phosphorylation of estrogen receptor at serine 118 in the absence of ligand. *J Steroid Biochem Mol Biol*. 2017;174:146-152. doi:10.1016/j.jsbmb.2017.08.013

286. 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. doi:10.1007/s10549-008-0221-0

287. Yang XR, Figueroa JD, Hewitt SM, et al. Estrogen receptor and progesterone receptor expression in normal terminal duct lobular units surrounding invasive breast cancer. *Breast Cancer Res Treat*. 2013;137(3):837-847. doi:10.1007/s10549-012-2380-2

288. Sigurgeirsson B, Åmark H, Jemt A, et al. Comprehensive RNA sequencing of healthy human endometrium at two time points of the menstrual cycle†. *Biol Reprod*. 2017;96(1):24-33. doi:10.1095/biolreprod.116.142547

289. Beral V. Breast cancer and hormonal contraceptives: further results. Collaborative Group on Hormonal Factors in Breast Cancer. *Contraception*. 1996;54(3 Suppl). Accessed September 12, 2021. https://ora.ox.ac.uk/objects/uuid:f7035078-e9c9-4a4c-b155-c31409b973b1

290. Hunter DJ, Colditz GA, Hankinson SE, et al. Oral Contraceptive Use and Breast Cancer: A Prospective Study of Young Women. *Cancer Epidemiol Prev Biomark*. 2010;19(10):2496-2502. doi:10.1158/1055-9965.EPI-10-0747

291. Mørch LS, Skovlund CW, Hannaford PC, Iversen L, Fielding S, Lidegaard Ø. Contemporary Hormonal Contraception and the Risk of Breast Cancer. *N Engl J Med*. 2017;377(23):2228-2239. doi:10.1056/NEJMoa1700732

292. Romieu I, Hernandez-Avila M, Liang MH. Oral Contraceptives and the Risk of Rheumatoid Arthritis: A Meta-Analysis of a Conflicting Literature. *Rheumatology*. 1989;XXVIII(suppl_1):13-17. doi:10.1093/rheumatology/XXVIII.suppl_1.13

293. Simin J, Tamimi R, Lagergren J, Adami HO, Brusselaers N. Menopausal hormone therapy and cancer risk: An overestimated risk? *Eur J Cancer*. 2017;84:60-68. doi:10.1016/j.ejca.2017.07.012

294. Zhang X, Diaz MR, Yee D. Fulvestrant regulates epidermal growth factor (EGF) family ligands to activate EGF receptor (EGFR) signaling in breast cancer cells. *Breast Cancer Res Treat*. 2013;139(2):351-360. doi:10.1007/s10549-013-2541-y

295. Fu X, Creighton CJ, Biswal NC, et al. Overcoming endocrine resistance due to reduced PTEN levels in estrogen receptor-positive breast cancer by co-targeting mammalian target of rapamycin, protein kinase B, or mitogen-activated protein kinase kinase. *Breast Cancer Res BCR*. 2014;16(5):430. doi:10.1186/s13058-014-0430-x

296. Butt AJ. Predicting response: identifying molecular determinants of endocrine response and resistance in breast cancer. *Expert Rev Endocrinol Metab*. 2011;6(5):661-663. doi:10.1586/eem.11.53

297. Fox EM, Arteaga CL, Miller TW. Abrogating endocrine resistance by targeting ERα and PI3K in breast cancer. *Front Oncol*. 2012;2:145. doi:10.3389/fonc.2012.00145

298. Ghayad SE, Vendrell JA, Larbi SB, Dumontet C, Bieche I, Cohen PA. Endocrine resistance associated with activated ErbB system in breast cancer cells is reversed by inhibiting MAPK or PI3K/Akt signaling pathways. *Int J Cancer*. 2010;126(2):545-562. doi:10.1002/ijc.24750

299. Parsons J, Francavilla C. 'Omics Approaches to Explore the Breast Cancer Landscape. *Front Cell Dev Biol*. 2020;7:395. doi:10.3389/fcell.2019.00395

300. Dorighi KM, Swigut T, Henriques T, et al. Mll3 and Mll4 Facilitate Enhancer RNA Synthesis and Transcription from Promoters Independently of H3K4 Monomethylation. *Mol Cell*. 2017;66(4):568-576.e4. doi:10.1016/j.molcel.2017.04.018

301. Local A, Huang H, Albuquerque CP, et al. Identification of H3K4me1-associated proteins at mammalian enhancers. *Nat Genet*. 2018;50(1):73-82. doi:10.1038/s41588-017-0015-6

302. A small UTX stabilization domain of Trr is conserved within mammalian MLL3-4/COMPASS and is sufficient to rescue loss of viability in null animals. Accessed November 14, 2021. http://genesdev.cshlp.org.proxy.library.vanderbilt.edu/content/34/21-22/1493.short

303. Musolino A, Campone M, Neven P, et al. Phase II, randomized, placebo-controlled study of dovitinib in combination with fulvestrant in postmenopausal patients with HR+, HER2- breast cancer that had progressed during or after prior endocrine therapy. *Breast Cancer Res BCR*. 2017;19(1):18. doi:10.1186/s13058-017-0807-8

304. Campone M, Bachelot T, Penault-Llorca F, et al. A phase Ib dose allocation study of oral administration of lucitanib given in combination with fulvestrant in patients with estrogen receptor-positive and FGFR1-amplified or non-amplified metastatic breast cancer. *Cancer Chemother Pharmacol*. 2019;83(4):743-753. doi:10.1007/s00280-018-03765-3

305. Drago JZ, Formisano L, Juric D, et al. FGFR1 Amplification Mediates Endocrine Resistance but Retains TORC Sensitivity in Metastatic Hormone Receptor–Positive (HR+) Breast Cancer. *Clin Cancer Res*. 2019;25(21):6443-6451. doi:10.1158/1078-0432.CCR-19-0138

306. Narod SA. Hormone replacement therapy and the risk of breast cancer. *Nat Rev Clin Oncol*. 2011;8(11):669-676. doi:10.1038/nrclinonc.2011.110

307. Chia YH, Ellis MJ, Ma CX. Neoadjuvant endocrine therapy in primary breast cancer: indications and use as a research tool. *Br J Cancer*. 2010;103(6):759-764. doi:10.1038/sj.bjc.6605845

308. Priedigkeit N, Ding K, Horne W, et al. Acquired mutations and transcriptional remodeling in long-term estrogen-deprived locoregional breast cancer recurrences. *Breast Cancer Res*. 2021;23(1):1. doi:10.1186/s13058-020-01379-3

309. Miller CA, Gindin Y, Lu C, et al. Aromatase inhibition remodels the clonal architecture of estrogen-receptor-positive breast cancers. *Nat Commun*. 2016;7(1):12498. doi:10.1038/ncomms12498

310. Miller WR, Larionov AA, Renshaw L, et al. Changes in breast cancer transcriptional profiles after treatment with the aromatase inhibitor, letrozole. *Pharmacogenet Genomics*. 2007;17(10):813-826. doi:10.1097/FPC.0b013e32820b853a

311. Razavi P, Chang MT, Xu G, et al. The Genomic Landscape of Endocrine-Resistant Advanced Breast Cancers. *Cancer Cell*. 2018;34(3):427-438.e6. doi:10.1016/j.ccell.2018.08.008

312. Spoerke JM, Gendreau S, Walter K, et al. Heterogeneity and clinical significance of ESR1 mutations in ER-positive metastatic breast cancer patients receiving fulvestrant. *Nat Commun*. 2016;7(1):11579. doi:10.1038/ncomms11579

313. Coombes RC, Page K, Salari R, et al. Personalized Detection of Circulating Tumor DNA Antedates Breast Cancer Metastatic Recurrence. *Clin Cancer Res*. 2019;25(14):4255-4263. doi:10.1158/1078-0432.CCR-18-3663

314. Hrebien S, Citi V, Garcia-Murillas I, et al. Early ctDNA dynamics as a surrogate for progression-free survival in advanced breast cancer in the BEECH trial. *Ann Oncol*. 2019;30(6):945-952. doi:10.1093/annonc/mdz085

315. Ma F, Guan Y, Yi Z, et al. Assessing tumor heterogeneity using ctDNA to predict and monitor therapeutic response in metastatic breast cancer. *Int J Cancer*. 2020;146(5):1359-1368. doi:10.1002/ijc.32536

316. Rajoria S, Zhao L, Intes X, Barroso M. FLIM-FRET for Cancer Applications. *Curr Mol Imaging*. 2014;3(2):144-161. doi:10.2174/22115552036661411172211l

317. Liu W, Cui Y, Ren W, Irudayaraj J. Epigenetic biomarker screening by FLIM-FRET for combination therapy in ER+ breast cancer. *Clin Epigenetics*. 2019;11(1):16. doi:10.1186/s13148-019-0620-6

318. Haque A, Engel J, Teichmann SA, Lönnberg T. A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. *Genome Med*. 2017;9(1):75. doi:10.1186/s13073-017-0467-4

319. Pan H, Gray R, Braybrooke J, et al. 20-Year Risks of Breast-Cancer Recurrence after Stopping Endocrine Therapy at 5 Years. *N Engl J Med*. 2017;377(19):1836-1846. doi:10.1056/NEJMoa1701830

320. Hong SP, Chan TE, Lombardo Y, et al. Single-cell transcriptomics reveals multi-step adaptations to endocrine therapy. *Nat Commun*. 2019;10:3840. doi:10.1038/s41467-019-11721-9

321. Lau JW, Lehnert E, Sethi A, et al. The Cancer Genomics Cloud: Collaborative, Reproducible, and Democratized-A New Paradigm in Large-Scale Computational Research. *Cancer Res*. 2017;77(21):e3-e6. doi:10.1158/0008-5472.CAN-17-0387

322. Multiomic Integration of Public Oncology Databases in Bioconductor. *JCO Clin Cancer Inform*. Accessed November 16, 2021. http://ascopubs.org/doi/full/10.1200/CCI.19.00119

323. Yu Y, Wang Y, Xia Z, et al. PreMedKB: an integrated precision medicine knowledgebase for interpreting relationships between diseases, genes, variants and drugs. *Nucleic Acids Res*. 2019;47(D1):D1090-D1101. doi:10.1093/nar/gky1042

324. Chunn LM, Nefcy DC, Scouten RW, et al. Mastermind: A Comprehensive Genomic Association Search Engine for Empirical Evidence Curation and Genetic Variant Interpretation. *Front Genet*. 2020;0. doi:10.3389/fgene.2020.577152

325. Good BM, Ainscough BJ, McMichael JF, Su AI, Griffith OL. Organizing knowledge to enable personalization of medicine in cancer. *Genome Biol*. 2014;15(8):438. doi:10.1186/s13059-014-0438-7

326. Poirion OB, Jing Z, Chaudhary K, Huang S, Garmire LX. DeepProg: an ensemble of deep-learning and machine-learning models for prognosis prediction using multi-omics data. *Genome Med*. 2021;13(1):112. doi:10.1186/s13073-021-00930-x

327. Nicora G, Vitali F, Dagliati A, Geifman N, Bellazzi R. Integrated Multi-Omics Analyses in Oncology: A Review of Machine Learning Methods and Tools. *Front Oncol*. 2020;0. doi:10.3389/fonc.2020.01030

328. Reel PS, Reel S, Pearson E, Trucco E, Jefferson E. Using machine learning approaches for multi-omics data analysis: A review. *Biotechnol Adv*. 2021;49:107739. doi:10.1016/j.biotechadv.2021.107739

329. Way GP, Sanchez-Vega F, La K, et al. Machine Learning Detects Pan-cancer Ras Pathway Activation in The Cancer Genome Atlas. *Cell Rep*. 2018;23(1):172-180.e3. doi:10.1016/j.celrep.2018.03.046

330. Malta TM, Sokolov A, Gentles AJ, et al. Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. *Cell*. 2018;173(2):338-354.e15. doi:10.1016/j.cell.2018.03.034

331. Oren Y, Tsabar M, Cuoco M, et al. Cycling cancer persister cells arise from lineages with distinct programs. *Nature*. 2021;596:1-7. doi:10.1038/s41586-021-03796-6

332. Chang A, Liu L, Ashby JM, et al. Recruitment of KMT2C/MLL3 to DNA Damage Sites Mediates DNA Damage Responses and Regulates PARP Inhibitor Sensitivity in Cancer. *Cancer Res*. 2021;81(12):3358-3373. doi:10.1158/0008-5472.CAN-21-0688

333. Li W, Wu H, Sui S, Wang Q, Xu S, Pang D. Targeting Histone Modifications in Breast Cancer: A Precise Weapon on the Way. *Front Cell Dev Biol*. 2021;9:2463. doi:10.3389/fcell.2021.736935

334. Hodges-Gallagher L, Valentine CD, Bader SE, Kushner PJ. Inhibition of histone deacetylase enhances the anti-proliferative action of antiestrogens on breast cancer cells and blocks tamoxifen-induced proliferation of uterine cells. *Breast Cancer Res Treat*. 2007;105(3):297-309. doi:10.1007/s10549-006-9459-6

335. Lee YJ, Won AJ, Lee J, et al. Molecular Mechanism of SAHA on Regulation of Autophagic Cell Death in Tamoxifen-Resistant MCF-7 Breast Cancer Cells. *Int J Med Sci*. 2012;9(10):881-893. doi:10.7150/ijms.5011

336. Feng Q, Zhang Z, Shea MJ, et al. An epigenomic approach to therapy for tamoxifen-resistant breast cancer. *Cell Res*. 2014;24(7):809-819. doi:10.1038/cr.2014.71

337. Yan N, Xu L, Wu X, et al. GSKJ4, an H3K27me3 demethylase inhibitor, effectively suppresses the breast cancer stem cells. *Exp Cell Res*. 2017;359(2):405-414. doi:10.1016/j.yexcr.2017.08.024

338. Lu H, Xie Y, Tran L, et al. Chemotherapy-induced S100A10 recruits KDM6A to facilitate OCT4-mediated breast cancer stemness. *J Clin Invest*. 2020;130(9):4607-4623. doi:10.1172/JCI138577

339. Zhang Z, Christin JR, Wang C, Ge K, Oktay MH, Guo W. Mammary-Stem-Cell-Based Somatic Mouse Models Reveal Breast Cancer Drivers Causing Cell Fate Dysregulation. *Cell Rep*. 2016;16(12):3146-3156. doi:10.1016/j.celrep.2016.08.048

340. Roy S, Tomaszowski KH, Luzwick JW, et al. p53 orchestrates DNA replication restart homeostasis by suppressing mutagenic RAD52 and POLθ pathways. Powell S, ed. *eLife*. 2018;7:e31723. doi:10.7554/eLife.31723

341. MacFawn I, Wilson H, Selth LA, et al. Grainyhead-like-2 confers NK-sensitivity through interactions with epigenetic modifiers. *Mol Immunol*. 2019;105:137-149. doi:10.1016/j.molimm.2018.11.006

342. Lee MG, Villa R, Trojer P, et al. Demethylation of H3K27 Regulates Polycomb Recruitment and H2A Ubiquitination. *Science*. 2007;318(5849):447-450. doi:10.1126/science.1149042

343. Herz HM, Mohan M, Garruss AS, et al. Enhancer-associated H3K4 monomethylation by Trithorax-related, the Drosophila homolog of mammalian Mll3/Mll4. *Genes Dev*. 2012;26(23):2604-2620. doi:10.1101/gad.201327.112

344. Lee JE, Wang C, Xu S, et al. H3K4 mono- and di-methyltransferase MLL4 is required for enhancer activation during cell differentiation. Glass C, ed. *eLife*. 2013;2:e01503. doi:10.7554/eLife.01503

345. Ashokkumar D, Zhang Q, Much C, et al. MLL4 is required after implantation, whereas MLL3 becomes essential during late gestation. *Development*. 2020;147(12). doi:10.1242/dev.186999

346. Wang YR, Xu NX, Wang J, Wang XM. Kabuki syndrome: review of the clinical features, diagnosis and epigenetic mechanisms. *World J Pediatr WJP*. 2019;15(6):528-535. doi:10.1007/s12519-019-00309-4

347. Gates LA, Gu G, Chen Y, et al. Proteomic profiling identifies key coactivators utilized by mutant ERα proteins as potential new therapeutic targets. *Oncogene*. 2018;37(33):4581-4598. doi:10.1038/s41388-018-0284-2

348. Zheng Y, Huang Y, Mencius J, et al. Distinct kinetic mechanisms of H3K4 methylation catalyzed by MLL3 and MLL4 core complexes. *J Biol Chem*. 2021;296:100635. doi:10.1016/j.jbc.2021.100635

349. Guo C, Chen LH, Huang Y, et al. KMT2D maintains neoplastic cell proliferation and global histone H3 lysine 4 monomethylation. *Oncotarget*. 2013;4(11):2144-2153. doi:10.18632/oncotarget.1555

350. Larsson C, Cordeddu L, Siggens L, et al. Restoration of KMT2C/MLL3 in human colorectal cancer cells reinforces genome-wide H3K4me1 profiles and influences cell growth and gene expression. *Clin Epigenetics*. 2020;12:74. doi:10.1186/s13148-020-00863-z

351. Wang L, Zhao Z, Ozark PA, et al. Resetting the epigenetic balance of Polycomb and COMPASS function at enhancers for cancer therapy. *Nat Med*. 2018;24(6):758-769. doi:10.1038/s41591-018-0034-6

352. Pereira B, Chin SF, Rueda OM, et al. The somatic mutation profiles of 2,433 breast cancers refine their genomic and transcriptomic landscapes. *Nat Commun*. 2016;7(1):11479. doi:10.1038/ncomms11479

353. New Insights into mTOR Signaling: mTORC2 and Beyond. Accessed November 17, 2021. https://www-science-org.proxy.library.vanderbilt.edu/doi/10.1126/scisignal.267pe27?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%20%200pubmed

354. Pearce LR, Komander D, Alessi DR. The nuts and bolts of AGC protein kinases. *Nat Rev Mol Cell Biol*. 2010;11(1):9-22. doi:10.1038/nrm2822