



Theses and Dissertations

2009-07-29

Cis-regulatory Sequence and Co-regulatory Transcription Factor Functions in ER α -Mediated Transcriptional Repression

Richard LeRoy Smith
Brigham Young University - Provo

Follow this and additional works at: <https://scholarsarchive.byu.edu/etd>



Part of the [Microbiology Commons](#)

BYU ScholarsArchive Citation

Smith, Richard LeRoy, "Cis-regulatory Sequence and Co-regulatory Transcription Factor Functions in ER α -Mediated Transcriptional Repression" (2009). *Theses and Dissertations*. 2261.
<https://scholarsarchive.byu.edu/etd/2261>

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.

Cis-regulatory Sequence and Co-regulatory Transcription Factor Functions in ER α - Mediated Transcriptional Repression

by

Richard LeRoy Smith

A thesis submitted to the faculty of

Brigham Young University

in partial fulfillment of the requirements for the degree of

Master of Science

Department of Microbiology and Molecular Biology

Brigham Young University

December 2009

Copyright © 2009 Richard LeRoy Smith

All Rights Reserved

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a thesis submitted by

Richard LeRoy Smith

This thesis has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

Date

Chin-Yo Lin, Chair

Date

Laura Bridgewater

Date

Evan Johnson

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the thesis of Richard LeRoy Smith in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

Date

Chin-Yo Lin
Chair, Graduate Committee

Accepted for the Department

Date

Brent Nielsen
Department Chair

Accepted for the College

Date

Rodney J. Brown
College Dean

ABSTRACT

Cis-regulatory Sequence and Co-regulatory Transcription Factor Functions in ER α -Mediated Transcriptional Repression

Richard LeRoy Smith

Department of Microbiology and Molecular Biology

Master of Science

Estrogens exert numerous actions throughout the human body, targeting healthy tissue while also enhancing the proliferative capacity of breast cancers. Estrogen signaling is mediated by the estrogen receptor (ER), which binds DNA and ultimately affects the expression of adjacent genes. Current understanding of ER-mediated transcriptional regulation is mostly limited to genes whose transcript levels increase following estrogen exposure, though recent studies demonstrate that direct down-regulation of estrogen-responsive genes is also a significant feature of ER action. We hypothesized that differences in cis-regulatory DNA was a factor in determining target gene expression and performed computational and experimental studies to test this hypothesis. From our *in silico* analyses, we show that the binding motifs for certain transcription factors are enriched in cis-regulatory sequences adjacent to repressed target genes compared to induced target genes, including the motif for RUNX1. *In silico*

analyses were tested experimentally using dual luciferase reporter assays, which indicate that several ER binding sites are estrogen responsive. Mutagenesis of transcription factor motifs (for ER and RUNX1) reduced the response of reporter gene. Further experiments demonstrated that co-recruitment of ER and RUNX1 is necessary for repression of gene expression at some target genes. These findings highlight a novel interaction between ER and RUNX1 and their role in transcriptional repression in breast cancer.

ACKNOWLEDGEMENTS

This project would have been impossible without the constant support of Dr. Chin-Yo Lin. His efforts in consultation, direction, and assistance go beyond scientific research: he truly has been a coach and mentor as well as an intellectual and philosophical support. Additionally, much of the statistical data in this work have been contributions of Dr. Evan Johnson and his laboratory, and have truly assisted in furthering the integration of computational and experimental approaches. Dr. Laura Bridgewater has also been a sounding board and a source of encouragement throughout this project. I am also grateful for members of the Bridgewater lab, especially Trina Loos for assistance in experimental protocols and procedures. I extend my thanks to the members of the Lin Lab who have helped with cell culture and other aspects of my experimental work: Justin Crofts, Adam Spjute, Garrett Coles, and Nick Candelaria. Finally, I am deeply grateful for continual support, love, patience, and encouragement of my wife Lisa, without whom this project would have little meaning or purpose.

Table of Contents

Chapter 1: Estrogen Receptors and Transcriptional Repression	11
Estrogen Biology	11
Breast Cancer and the Estrogen Receptor.....	13
Estrogen Receptor and Transcriptional Regulation	15
Transcriptional Repression by Estrogen Receptor.....	19
Causative and Effective Mechanisms	19
Chapter 2: In Silico Analysis of Regulatory Sequence	22
Introduction.....	22
Methods.....	25
Results.....	28
Conclusions and Discussion	36
Chapter 3: Functional Analysis of Putative Regulatory Sequence	39
Introduction.....	39
Methods.....	42
Results.....	48
Conclusions and Discussion	50
Chapter 4: RUNX1 Co-recruitment is Necessary for Estrogen Receptor Mediated Transcriptional Repression of SLC35A1	52
Introduction.....	52
Methods.....	55
Results.....	60
Conclusions and Discussion	63
Chapter 5: Summary and Conclusions	66
Bibliography	73

List of Figures

Figure 1 - Experimental Design	24
Figure 2 - Distance of Binding Regions from TSS	29
Figure 3 - Composition of ERE	31
Figure 4 - Distance of ERE to RUNX1 Motifs	35
Figure 5 - ERE Composition Adjacent to RUNX1 Motifs	38
Figure 6 - TF Motifs in Select ER Binding Sites	41
Figure 7 - Luciferase Assays	49
Figure 8 - RUNX1 ChIP	61
Figure 9 - ER-RUNX1 re-ChIP	62
Figure 10 - RUNX1 siRNA Knockdown	64
Figure 11 - RUNX1-ER Co-recruitment to SLC35A1 ER Binding Region	70
Figure 12 - RUNX1-ER Repressive Complex	72

List of Tables

Table 1 - TRANSFAC Analysis	33
Table 2 - Cloning Primers	46
Table 3 - Mutagenesis Primers	47
Table 4 - ChIP Primers	58
Table 5 - siRNA Primers	59

Chapter 1: Estrogen Receptors and Transcriptional Repression

Estrogen Biology

Estrogens comprise a class of eight small, polycyclic steroid hormones, the most common of which is 17- β Estradiol (E_2). As steroid hormones, estrogens are cholesterol derivatives that must be first synthesized at sites far from their target tissues. The primary site of synthesis is within theca and granulosa cells of the ovaries in females, and leydig cells in males. The final, rate-limiting step of estrogen synthesis involves aromatization of an androgen precursor via cytochrome P450¹. Following synthesis, estrogens are transported through the blood in a dissociated state or are bound to sex-hormone binding globulin and other transport molecules. Due to their hydrophobic nature, estrogens readily diffuse across cell membranes once they arrive at target tissues.

After arriving at target tissues, estrogens and their cellular receptors (estrogen receptors or ERs), regulate many aspects of healthy physiology including reproduction, bone deposition and metabolism, and neural development². In females, puberty is initiated by gradual increases in serum concentration of estrogens in response to gonadotropin signaling. Serum estrogen levels fluctuate most significantly during menstruation, reaching the highest concentrations prior to ovulation². Endogenous estrogens have also been shown to be involved in aspects of male physiology, including spermatogenesis, and cardiovascular development, and may be protective against coronary heart disease³.

Tissues such as kidney, intestines, and lung exhibit moderate regulation by estrogens, but they primarily target breast, uterus, brain, bone, endothelial cells, and the prostate gland. Because estrogens affect a wide variety of tissue types, these studies necessitate work in animal and cell culture models. Knockout studies in mice against ER or the P450 aromatase indicate that loss of estrogen signaling results in reduced bone deposition, reduced male libido, and significantly reduced fertility in both males and females⁴. Very few cases of human mutant variants of ER exist, presumably because the majority of embryos possessing these mutations do not survive the first stages of development. In the few recorded surviving cases, individuals present with decreased fertility and early onset osteoporosis⁵.

Estrogen signaling is essential for sexual development and other aspects of healthy physiology; however, estrogens and ERs are also involved in human diseases. In many cases, a lack of estrogen signaling or reduced responsiveness to estrogen signaling play a role in disease processes. Risk for cardiovascular disease in both men and women improves in patients with increased estrogen signaling. In bone, a lack of estrogen responsiveness leads to higher risk for osteoporosis. Disease phenotypes exhibiting reduced estrogen responsiveness are also associated with stroke, Parkinson's, and Alzheimer's disease all exhibit reduced estrogen responsiveness. In mouse models, supplemental estrogens have been shown to decrease the risk of both stroke and Alzheimer's, by increasing the production of essential neurotransmitters⁶.

While decreased estrogen responsiveness is associated with many disease phenotypes, increased estrogen signaling also promotes disease phenotypes. Estrogens enhance cellular proliferation, thus overexposure leads to an increased risk of DNA mutation. Clinical and model organism studies indicate that this process of unchecked estrogen-responsive cell division increases the risk of developing cancer of the endometrium, ovaries, large intestine, prostate, and breast⁶.

Several cancers are responsive to estrogen signaling; nevertheless, breast cancer is by far the most common and best studied. In the US alone, breast cancer is diagnosed in one out of every seven new cases of cancer. It is the third most common cancer overall and the second most common cancer in women after skin cancer, and some 1500 men are also diagnosed annually. According to the American Cancer Society⁷, in 2009 there will be an estimated 194,280 new cases of mammary carcinoma and 40,610 associated deaths. The healthcare and economic burden of breast cancer is also large, particularly in developed nations where breast cancer is more common⁸. Nevertheless, as developing nations adopt more affluent, Western lifestyles, breast cancers are increasing in frequency and cost of treatment⁹⁻¹⁰. Understanding the link between estrogen signaling and breast cancer is of major importance in both public and global health concerns¹¹⁻¹².

Breast Cancer and the Estrogen Receptor

The first direct indication of estrogen's role in breast cancer was recorded in 1896 when Dr. George Beatson of the Glasgow Cancer Hospital observed that bilateral oophorectomy

in patients with inoperable neoplasia reduced the aggressiveness of these tumors¹³. This pioneering procedure presented not only a viable treatment for aggressive breast cancers, but also paved the way for the discovery that the ovaries regulate sexual development and ultimately breast cancer¹⁴. Since then, it has been established that estrogens enhance cell proliferation in breast tissue, a major feature of female reproductive development. Additionally, estrogens have been found to foster development of neoplastic lesions, and encourage progression from benign colonies into solid tumors and metastases¹⁵⁻¹⁷. Because estrogens promote tumor development and progression, elevated blood estrogen levels and increased duration from menarche to menopause are both considered significant risk factors in breast cancer¹⁸⁻¹⁹. For this reason, most breast cancer therapies focus on disruption or modulation of the effects of estrogen signaling.

The estrogens are often targeted for disruption using selective estrogen receptor modulators or SERMs. SERMs are members of a class of pharmacological molecules which compete with estrogen for occupancy of ERs, thus affecting downstream signaling of estrogens. Unlike estrogens, SERMs do not purely act as agonists for ERs, nor are they pure antagonists; instead, they exhibit tissue-specific modulation of ER signaling, activating genes in some tissues which they inhibit in others²⁰. For example Tamoxifen, a SERM, is a common breast cancer drug. Tamoxifen signaling also exerts an overall protective effect against osteoporosis, similar to E₂. However, Tamoxifen treatment is associated with a 2.5-fold increase in cervical and uterine carcinoma, complicating its use in clinical practice. In contrast, Raloxifene, while exhibiting many of the beneficial effects of Tamoxifen treatment in breast cancer and osteoporosis, exhibits none of the

negative side effects in endometrial cancers²¹. In cardiovascular tissue, some SERMs also appear to mimic some of the beneficial effects of endogenous estrogens. One reason for this tissue specificity is tissue-specific expression of ERs and other interacting proteins. Furthermore, SERMs alter ER conformation, which in turn affects the recruitment of various co-regulators²⁰. ER α directs breast tumor progression and aggressiveness, and it follows that the majority of breast cancer treatments focus on modulating or antagonizing ER α signaling.

The majority of estrogen functionality at the cellular level is mediated by the ER²², which serves as the primary diagnostic and prognostic marker for breast cancer screening and staging. Two known ER subtypes exist – ER α and ER β – and ER α plays the more clinically significant role in breast cancer; additionally, basic clinical classification of breast cancers is a measure of ER α presence or absence²³. In general, tumor ER α status reflects the dependence of the tumor on estrogen signals for growth and proliferation. Tumors which have lost or never exhibited estrogen responsiveness are usually more invasive, do not respond to ER α -targeted treatments, and more difficult to treat²⁴. In general, breast cancers that lack ER α are associated with unfavorable disease outcomes and poor patient survival²³.

Estrogen Receptor and Transcriptional Regulation

The estrogen receptor was first described as a protein that bound ³H-Estradiol, and was localized primarily in the nucleus^{22,25}. Due to its subcellular localization and the increase of cellular RNA content following E₂ treatment, it was proposed that this receptor

molecule was responsible for affecting gene expression, thus mediating the cellular effects of known estrogen physiology²². As study of the molecule intensified, researchers established that the tissue specificity of estrogen signaling is a result of the tissue-specific expression of ER. Additionally, a second protein of similar structure was discovered in 1996, now known as estrogen receptor beta – ER β ²⁶. The original molecule has since been renamed ER α .

ER α is a member of the nuclear receptor superfamily of transcription factors, comprised of hundreds of proteins. Most proteins in this superfamily respond to hormone signals, alter gene expression, and possess similar structural features²⁷. Structural regions are divided into regions (A-E), each containing specific protein domains. The primary transactivating domain, AF-1 resides near the N-terminus of the protein, occupying the A/B region. The DNA binding domain (also known as the DBD) is located in region C, close to the center of the protein, and is flanked on the C-terminal end by an unstructured region D, also known as the hinge region. The ligand-binding domain (LBD) resides in region E, following on the C-terminal side of the hinge region. In ER α , the LBD is responsible for the majority of dimer stabilization following E₂ binding and consists of 12 alpha helices which form a ligand-binding pocket²⁸. This pocket binds estrogens, SERMs, and several estrogen-like polycyclic compounds²⁰. A unique feature of this region is helix 12, the second transactivating domain also known as AF-2. The domain exhibits estrogen sensitivity and is only active once the ligand is bound. Once activated by ligand binding, the conformation of AF-2 shifts, enabling ER α to mediate regulation of gene expression.

ER α regulation of gene expression is ligand-dependent. Upon binding E₂, ER α forms dimers and moves to the nucleus. Once inside the nucleus, dimerized ER α binds DNA sequences encoded by a 13-nucleotide motif (GGTCAnnnTGACC) also known as an estrogen response element or ERE²⁹⁻³⁰. Recruitment of ER α depends on several factors including number and location of EREs³¹, and the nucleotide composition of the ERE³². ERE composition which has been shown to directly affect both affinity and activity of ER binding suggesting that ERE sequence differences may result in alternate structural conformations³³. These conformational differences enable chromatin-bound ER to physically interact with regulatory proteins and affect transcription of target genes.

The chromatin which ER α binds is dynamic, and necessitates modification in order to regulate gene expression. Modifications are mediated by co-regulator proteins which physically interact with ER α . These co-regulators include histone acetylase complexes, chromatin remodeling complexes, and basal transcriptional machinery. Work done by Métiévier and colleagues demonstrates the overall cyclical, sequential pattern whereby ER α and co-activators are recruited to the pS2/TFF1 promoter, including the subsequent steps of ER degradation and recycling³⁴. It appears that ER α binds to the pS2 promoter four times in sequence, each time recruiting more protein complexes and further modifying target chromatin. During the first sequence of events, ER α binds the ERE adjacent to pS2, and recruits the SWI/SNF complex to transcriptionally silent DNA. This is followed by binding of TBP and TFIIA. After this first round of recruitment, p68 associates with the ER-TBP complex, recruiting p300 and other proteins possessing HAT (histone acetyltransferase) activity, which assemble to and unwind nascent chromatin. A

third round of recruitment follows, re-recruiting components of SWI/SNF, further modifying target chromatin. The final cycle of assembly further modifies chromatin via acetylation, and culminates in the recruitment of mediator, polymerase, and initiation/elongation factors. Following these cyclical rounds of protein recruitment, transcription of pS2/TFF1 begins. Additionally, co-repressor molecules possessing histone deacetylase (HDAC) activity are also cyclically recruited, silencing chromatin in between activation cycles. Much is known about co-activator assembly following ER α recruitment and is generally accepted as the general model of ER action on the molecular level; however, recent data indicates that repression is also a major component of ER α -mediated regulation of expression. This model of transcriptional repression runs counter to the generally accepted model of activation by ER α .

Transcriptional Repression by Estrogen Receptor

The induction-only paradigm is insufficient to describe the differential genome-wide effects of ER α signaling in breast cancer cells. Microarray studies³⁵, chromatin immunoprecipitation coupled to paired-end ditag (ChIP-PET) experiments³⁶, and other whole genome analyses such as SAGE and ChIP-on-chip³⁷ indicate that repressive effects of estrogen on target genes are widespread. The ChIP-PET and ChIP-Chip studies experiments in particular demonstrate that estrogen receptor is capable of binding to thousands of sites throughout the genome adjacent to putative repressed target genes. In the context of these studies, ER α binding regions have been found to reside as far away as 150kb from target genes, though the majority occur within several kilobases. Each study differs in platform, genes studied, and application, so the number of repressed targets and binding sites varies, as do the sites and genes targeted. Direct causality is difficult to establish in the context of microarray experiments; however, it appears that ER α recruits co-repressors at ER α -binding sites directly repressing many of these E₂-responsive target genes.

Causative and Effective Mechanisms

Nuclear receptor co-repressors or CoRs are any proteins that interact with and lower the transcriptional effects of nuclear receptors (NRs) on their target genes. Though a relatively new discovery, much work has been done to identify these factors and their

target NRs³⁸. CoRs share several structural features, including a NR interacting domain, consisting of an LXXLL sequence or NR Box, a L/IXXI/VI sequence called the CoRNR (“corner”) Box, or other similar sequences³⁹. In the context of ER α transcriptional regulation, co-repressors mediate some of the tissue-specific effects of ER α signaling⁴⁰, binding surfaces usually bound by co-activators⁴¹. Generally, CoRs mediate transcriptional repression by opposing the HAT activity of co-activators with HDAC (histone deacetylase) activity, recruiting chromatin silencing complexes (Mi-2/NuRD⁴²), or through competing with co-activators for ER α binding. Recent thesis research performed by Merrell and colleagues demonstrates that three CoRs (NCoR, NRIP1, and SMRT) associate with ER α at sites adjacent to target genes, deacetylate histones, expel co-activators, and abolish recruitment of polymerase⁴³, thus directly affecting repression of transcriptional activity.

Evidence suggests that ER α is capable of repressing gene expression by opposing mechanisms used to activate gene expression; notwithstanding, it remains to be seen what differences, if any, distinguish induction from inhibition of target genes. At binding sites adjacent to down-regulated genes ER α recruits co-repressors as opposed to co-activators, and it is thought that these differences are due to conformational differences in ER α ^{20,41}. Nevertheless, the mechanisms that cause these conformational differences are unclear. It is likely that the differences responsible for affecting ER α conformation exist at the regulatory sequence level.

We posit that features of ER α binding regions and cis-regulatory DNA adjacent to repressed target genes are likely involved in differential regulation, including the following:

1. Relative distance of the ERE from the transcriptional start site (TSS)
2. Composition of putative EREs,
3. Presence of motifs for other transcription factors
4. Distance of from the ERE to these other TF motifs.
5. Composition of ERE adjacent to these other TF motifs.

We carried out analyses to investigate these hypotheses on a genome-wide scale, using breast cancer specific data. These analyses produced several inferences regarding mechanisms of repression by ER α . Once the sequence differences were established, we validated our findings experimentally.

In order to assess experimentally the cis-regulatory sequence requirements associated with repressed target genes, we cloned several selected ER α binding regions into luciferase gene reporter constructs and performed reporter assays (Chapter 3). We then verified the physical presence of a transcription factor whose motifs occur adjacent to the ERE using ChIP and re-ChIP experiments, RUNX1 (Chapter 4). These studies indicated that RUNX1 binds to a motif adjacent to the SLC35A1 ERE, and interacts with ER α . After validating the recruitment of this transcription factor, we tested its ability to repress target gene expression by knocking down its expression with siRNA (Chapter 4). We present here a mechanism whereby cis-regulatory sequence and transcription factor co-recruitment regulate ER α mediated repression of target gene expression in breast cancer.

Chapter 2: In Silico Analysis of Regulatory Sequence

Introduction

Several new genome-wide analysis of transcription factor binding have allowed for high throughput detection of ER α binding regions. ChIP-PET technology isolates transcription factor-bound DNA from standard ChIP procedures, amplifies the enriched DNA, and uses plasmids and cloning strategies to generate sequencing libraries. These sequencing products are then compared to the host species' genome sequence in order to determine location of transcription factor binding sites and the relative abundance with which those binding sites are occupied by the transcription factor of interest. ChIP-chip technology differs slightly in that it utilizes probe hybridization on a microarray surface to determine the location and abundance of the fragment of interest. ChIP-seq is another variation, similar to ChIP-PET, in which the ChIP DNA fragments are subjected to high throughput sequencing. Recent studies by Edwin Cheung (unpublished) utilized a variation on the ChIP-PET technology, combined with a microarray study. In these experiments, ChIP DNA regions were correlated with their nearest gene. Data from these studies formed the basis of our *in silico* analyses.

Our approach to solving the cis regulatory mechanisms behind ER α -mediated down-regulation of gene expression utilized large-scale data sets of gene expression and ER α binding regions throughout the genome. Specifically, the data sets were composed of

thousands of ER α binding regions from across the genome. Each ER α binding region was then paired with an E₂ responsive gene, obtained from microarray studies. Additionally, each binding region was analyzed for the presence or absence of an ERE. We then proceeded to investigate these data further using computational and statistical tools.

Our *in silico* analyses followed the design in Figure 1. After obtaining the compiled binding region and gene expression data, they were then divided into two categories based on their relative response to estrogen treatment versus a vehicle control. Binding region-target gene pairs whose expression was activated following E₂ treatment comprised one group, while those whose expression was reduced following E₂ treatment comprised the other. Binding region-gene pairs that showed no E₂ response were ignored for our study. Next, sequence analyses were performed on each ER α binding region to assess distinguishing differences between the two estrogen response categories based on five features of repression-associated ER α binding regions:

1. Proximity of ERE to TSS
2. Relative nucleotide composition of ERE
3. Presence of other transcription factor motifs
4. Proximity of ERE to other TF motifs
5. Composition of ERE adjacent to other TF motifs

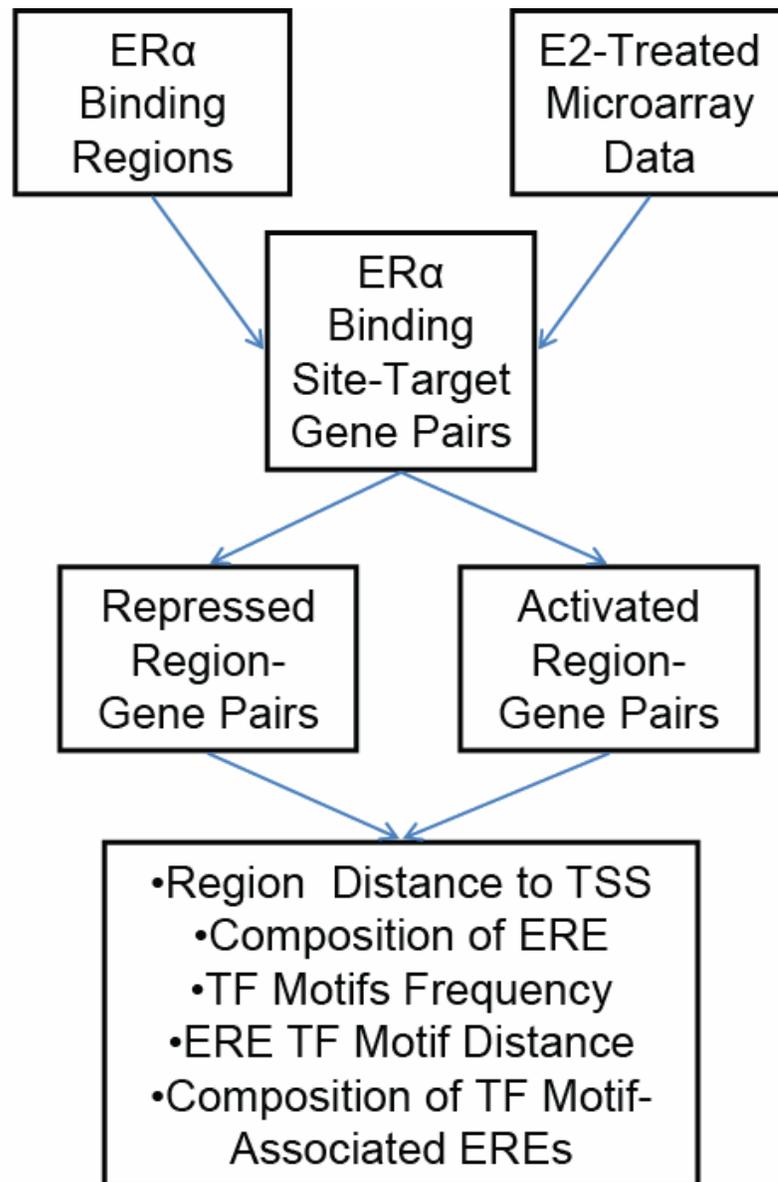


Figure 1 - Analytical Procedure For Detecting and Validating Putative Mechanisms of ER α Mediated Transcriptional Repression

ER α binding regions from MCF-7 cells were correlated with nearest genes and assessed for E2 response from an independent microarray experiment. These binding site-target gene pairs were then divided into two groups based on E2 response. Following compilation of data sets, each of the five hypotheses was tested (see Chapter 2 Methods).

Methods

Global ERE Proximity to TSS

Following compilation of the 1788 binding region-target gene pairs, the proximity of each of the ER α binding region to their putative gene targets was determined and the distances were compared between 850 activated and 938 repressed genes. Distance was calculated as the nucleotide distance in bases from the closest edge of the binding region to the annotated transcriptional start site of the target gene. The mean and median for each distribution was calculated, and the log absolute difference of distributions of these distances was compared for these two groups using a t-test.

Relative Composition of ERE

Putative EREs from each ER α binding region sequence were aligned and compiled into a batch file from each estrogen responsive category and then passed to WebLogo⁴⁴ (<http://weblogo.berkeley.edu/>). WebLogo analyzes each position of the putative ERE for the relative occurrence of each nucleotide and depicts the output in a Sequence Logo format⁴⁵. A sequence logo graphically depicts each position of aligned sequences as a stack of the four nucleotide letters. The height of each nucleotide letter within the stack represents its significance of occurrence at that position. The graphical logos indicate a greater abundance with larger nucleotide letters, and less frequent occurrence with smaller letters. The height of each letter is also indicative of statistical significance. Sequence logos from both estrogen responsive categories were then visually compared to

assess ERE differences between EREs adjacent to activated and repressed target genes.. All 90 putative EREs adjacent to repressed genes were compared to all 95 ERE's from activated target genes, resulting in two sequence logos.

Transcription Factor Motif Occurrence in Regulatory Sequence

The center of each ER α binding region was calculated and sequences of 500 bases flanking either side of these centers (1000 bases total) were compiled into a batch file and categorized based on estrogen responsiveness. Each sequence was then submitted for TRANSFAC⁴⁶ (professional version 12.1) database analysis. The TRANSFAC database is a collection of position-weight matrices representing transcription factor binding sites based on biological and biochemical data. Parameters were optimized for non-redundant vertebrate transcription factors, usage of high-quality matrices, and reduction of false positives. We used the software package MATCH⁴⁷ to analyze each sequence for the presence of known TRANSFAC transcription factor motifs. The parameters used were non-redundant, high quality, vertebrate position weight matrices, and set to minimize false positives. Each motif from this analysis was analyzed for the frequency of occurrence per region, and the frequency of occurrence within the repression versus the induction category.

Following MATCH analysis, the motifs within each estrogen receptor binding nucleotide sequence were counted and then motifs from repressed genes were compared to those from activated genes. A Fisher exact test was performed utilizing the following factors: presence of specific motif and E₂-responsiveness. The resulting table of p-values

indicates the significance of the frequency at which specific motifs occur adjacent to repressed as opposed to activated target genes. In addition to this analysis, each factor was further assessed for its frequency of occurrence in ER α binding regions which contained putative EREs as opposed to those that did not.

Relative Proximity of ERE to Other TF Motifs

Following calculation of TF motif frequency of occurrence, the most highly enriched motif adjacent to E2 repressed genes containing EREs – AML1 or RUNX1 – was selected for further analysis. The nucleotide distances between the two closest edges of each RUNX1 motif and ERE were calculated and averaged across all regions adjacent to activated and repressed target genes. The distances between ERE-RUNX1 motif pairs for activated and repressed genes was then compared and calculated by T-Test of the log absolute difference of distributions.

Relative Composition of EREs Adjacent to RUNX1 Motifs

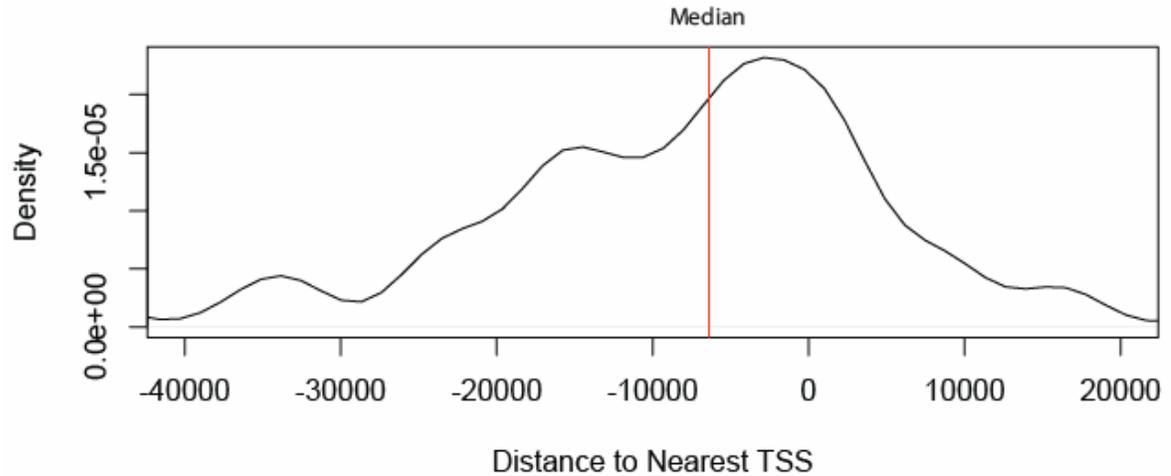
EREs adjacent to RUNX1 motifs were arranged into 4 categories (Repressed, with RUNX1 motifs; activated, with RUNX1; repressed, without RUNX1; activated, without RUNX1) and analyzed using WebLogo according to the procedures described above.

Results

ERE Proximity to Target Genes

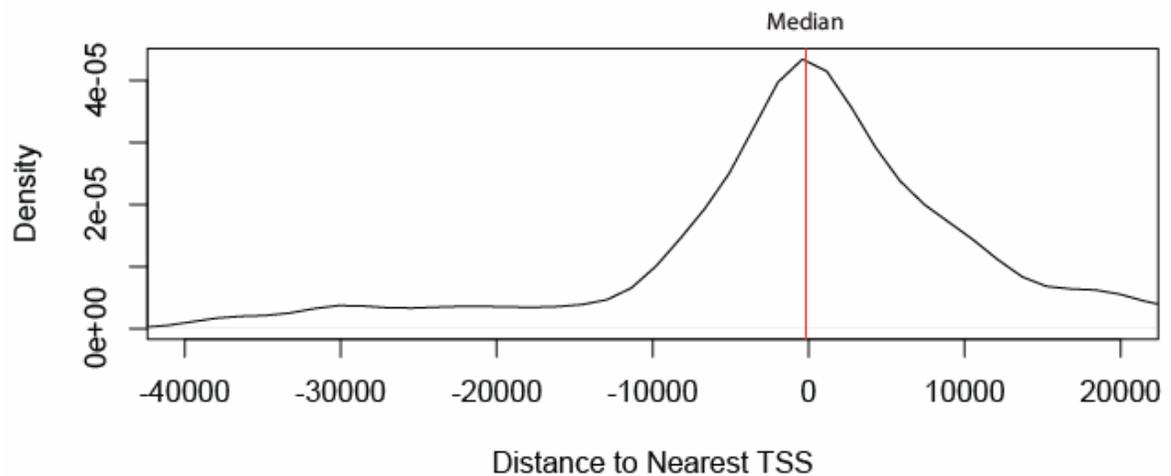
After calculating distances from ER α binding regions to target gene TSS, distributions were compiled and compared between repressed (Figure 2a) and activated (Figure 2b) target genes. The independent variable in the analysis was established as the relative distance of the ER α binding region to the TSS, and the dependent variable was the number of binding sites (the graph displays this as density, which assumes the area under the distribution curve equals 1 and adjusts each value to yield a smoothed histogram). The median distance was -6,403 adjacent to repressed target genes and -176 adjacent to activated target genes. Negative numbers indicate relative upstream (5') locations and positive numbers indicate relative downstream (3') locations. Statistical comparison (T-test) of the log absolute differences in distributions of repressed vs. activated genes yielded a significant difference ($p=.001031$), indicating that distance from ER α binding regions to target gene TSS is a distinguishing characteristic of estrogen responsiveness in the studied genes.

A Repressed Genes



Mean = -24,270 Median = -6,403

B Activated Genes



Mean = -1,948.0 Median = -176.8

Figure 2 - Relative Distributions of ERα Binding Site Distances to Target Gene TSS Show Significant Differences Between Up and Down Regulated Genes

Distances calculated based on location relative to database annotation of TSS. Negative numbers indicate binding sites which are located 5', positive numbers show binding sites located 3'. **A** Distribution for sites adjacent to repressed target genes. **B** Distribution of sites adjacent to activated target genes. Note the more uniform distribution and shape. Calculation of log absolute difference in distributions t-test yielded a p-value of $p=.001031$

ERE Composition Near Repressed Target Genes

In addition to our hypothesis that proximity of ER α binding regions to the TSS was a significant feature in distinguishing activated and repressed genes, we postulated that the composition of the ERE might be a distinguishing factor between activated and repressed genes (Figure 3). An online sequence analysis tool, WebLogo, produced sequence logos of EREs from both activated and repressed genes. The nucleotide composition differences between activated and repressed target gene regulatory elements are slight, and similar to the composition of the canonical ERE.

Frequency of Transcription Factor Motif Occurrence in Regulatory Regions

The nature of chromatin is such that no one state is fully “off” nor fully “on”. Instead, the state of chromatin is dynamic, fluctuating according to the presence or absence of various proteins and covalently linked molecules. Composition of the ERE appeared to be an insignificant factor distinguishing repression from induction of target genes, suggesting that ER α might not take on alternative conformations at activated genes versus repressed genes.

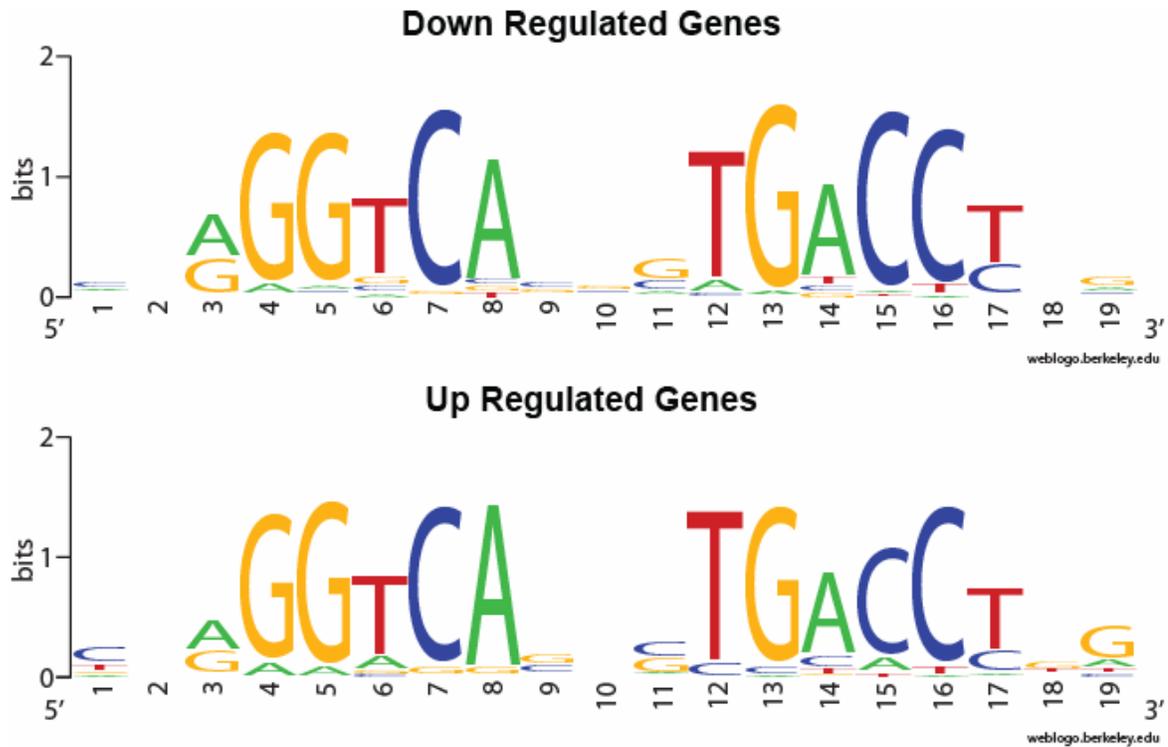


Figure 3 - ERE Are Not Significantly Distinct Between Up and Down Regulated Genes
 Weblogo analysis of EREs from down and up regulated genes. Each stack of letters represents a nucleotide position within the ERE. Letter heights indicate the relative frequency with which that nucleotide occurs at that position. Note that despite slight differences at position 15, both nucleotide distributions reflect the consensus ERE - GGTCAnnnTGACC.

In addition to ERE composition, we questioned whether the presence of additional proteins bound to regulatory motifs adjacent to the ERE would affect the interaction of ER α with target genes. In other words, we postulated that transcription factors outside of ER α might modulate the direction of ER α transcriptional regulation. This was accomplished by searching for transcription factor specific motifs using TRANSFAC. After comparisons, it was found (Table 1a) that 14 motifs occurred more frequently within ER α binding regions containing EREs adjacent to repressed target genes ($p < .05$). An additional 21 transcription factor motifs (Table 1b) occur with greater frequency within ER α binding regions where the ERE is absent adjacent to repressed genes ($p < .05$). Of those factors occurring in regions possessing EREs, motifs corresponding to AML1 (RUNX1) occurred most frequently overall with 196 occurrences adjacent to repressed target genes and 112 occurrences adjacent to activated genes.

A

Transcription Factor Name	Motifs - Down Regulated	Motifs - Up Regulated	Regions With Motifs Down	Regions With Motifs Up	Motif Frequency Down Regulated	Motif Frequency Up Regulated	Fishers Test p-value
AML	79	46	35	19	39%	21%	0.0141
AML1	78	44	35	19	39%	21%	0.0141
Pax-5	54	35	40	25	44%	28%	0.0289
CDP	47	30	39	25	43%	28%	0.0254
TEF-1	47	29	30	16	33%	18%	0.0419
AML1a	39	22	35	19	39%	21%	0.0141
Sn	31	14	27	11	30%	12%	0.0056
AP-4	24	12	21	10	23%	11%	0.0469
Tal-1alpha:E47	19	5	17	5	19%	6%	0.0109
SREBP2	18	9	17	7	19%	8%	0.0465
IRF	14	6	13	3	14%	3%	0.0159
Tal-1beta:E47	13	1	13	1	14%	1%	0.0012
COUPTF	12	2	11	2	12%	2%	0.0179
Tal-1beta:ITF-2	12	2	11	2	12%	2%	0.0179
CBF	9	1	8	1	9%	1%	0.0344

B

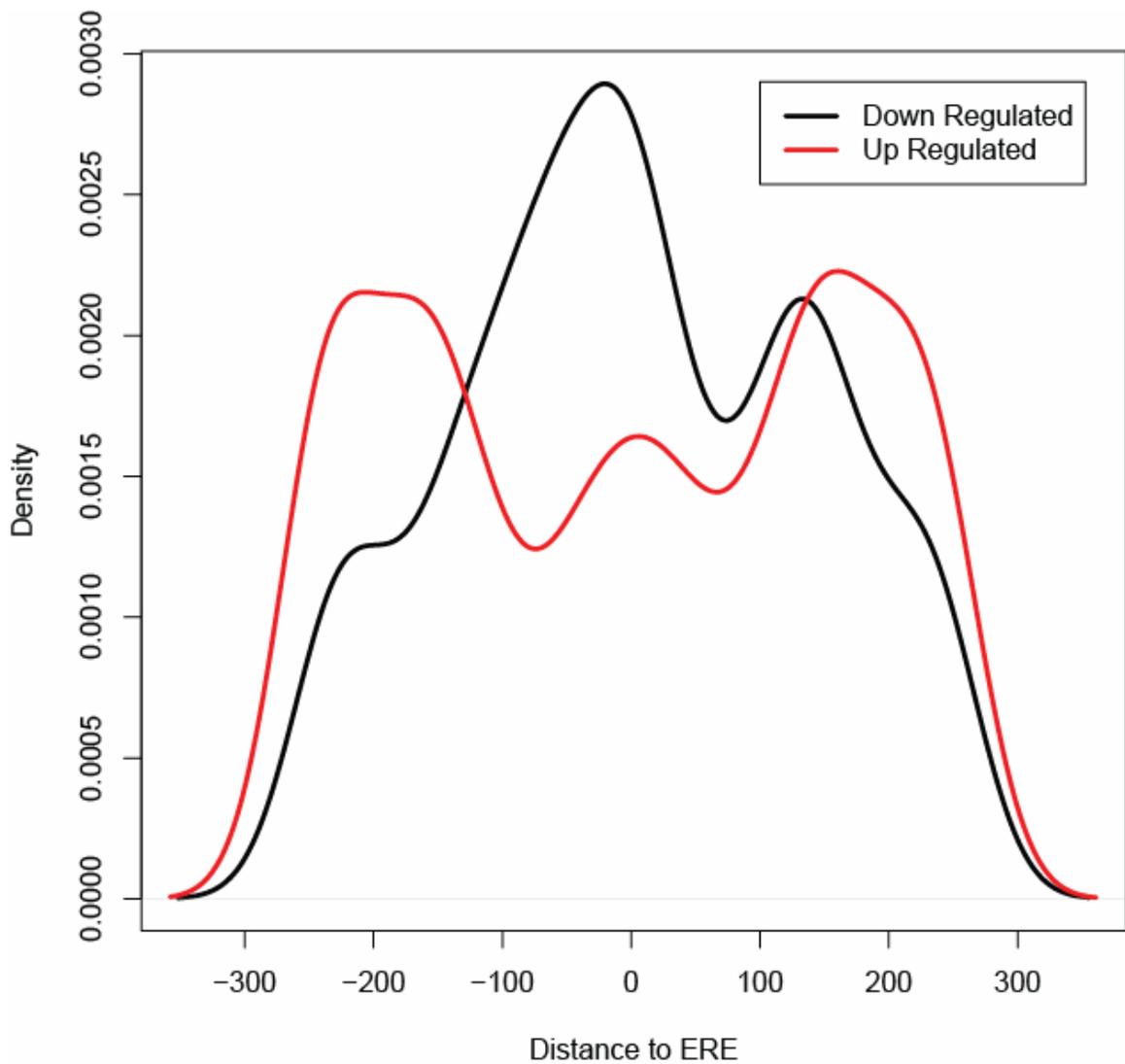
Transcription Factor Name	Motifs - Down Regulated	Motifs - Up Regulated	Regions With Motifs Down	Regions With Motifs Up	Motif Frequency Down Regulated	Motif Frequency Up Regulated	Fishers Test p-value
E2	770	598	370	291	44%	43%	0.0412
NF-1	669	515	377	286	44%	42%	0.008
ZF5	646	431	247	166	29%	25%	0.0011
Ets	639	456	244	173	29%	26%	0.0085
SREBP	618	519	330	251	39%	37%	0.0188
myogenin / NF-1	530	380	362	264	43%	39%	0.0017
TEF	426	297	311	237	37%	35%	0.0264
TEF-1	416	291	304	235	36%	35%	0.0493
ADR1	379	284	292	219	34%	33%	0.0206
SREBP-1	284	213	213	157	25%	23%	0.0431
Zic2	255	173	204	149	24%	22%	0.0398
CACD	213	157	185	130	22%	19%	0.0231
Tax/CREB	201	152	176	125	21%	19%	0.0342
T3R (half-site)	171	102	146	93	17%	14%	0.0061
c-Ets-2	141	97	130	88	15%	13%	0.0341
Freac-4	135	84	121	79	14%	12%	0.023
ABI4	84	47	74	39	9%	6%	0.006
ARF (half-site)	32	14	31	12	4%	2%	0.0125
OTX	23	8	23	8	3%	1%	0.0179
DMRT1	21	5	19	5	2%	1%	0.012
EBF	8	1	8	1	1%	0%	0.0414

Table 1 - Motifs Occuring More Frequently Adjacent to Down Regulated Genes Detected By TRANSFAC Analysis

Motifs are sorted according to p-value. Fisher exact calculated based on frequency of occurrence adjacent to down vs up regulated genes. **A** Regions containing an ERE. **B** Sites not containing an ERE.

Proximity of EREs to Nearby RUNX1/AML1 Motifs

Following the discovery that AML1 elements occur most frequently, we hypothesized that if ER α interacted with RUNX1, the relative distances between ERE and RUNX1 motifs might be distinct in sites adjacent to repressed genes as opposed to activated genes. This would indicate potential interactions between ER α and RUNX1. The distribution of distances in sites adjacent to repressed genes was found to be distinct from the distribution of distances from activated target genes (Figure 4). It appears that adjacent to repressed target genes, a more normal distribution exists, with the average distances being relatively close (mean=112, median=112). Adjacent to activated genes, however, the distributions were bimodal and distributed further away (mean=156, median=147). Statistical comparison of the distributions resulted in a p-value of p=.005876.



Distance between ERE/RUNX in Down: Median = 112.0; Mean = 112.5
Distance between ERE/RUNX in Up: Median = 156.0; Mean = 147.1

Figure 4 - Distribution of Distances From AML/RUNX1 Binding Sites to EREs Show Significant Differences Between Down and Up Regulated Genes

Distances were calculated from closest end of RUNX1 motifs to EREs within ER α Binding Regions. Distributions from down-regulated gene-associated distances are compared to those from up-regulated genes. T-test of log absolute difference in distributions yields a p-value of $p = 0.005876$

ERE Composition in ER α binding regions Containing RUNX1 Motifs

Evidence suggests ERE composition may affect ER conformation³³, and since ER α -RUNX1 co-occurrence may be associated with repression of target gene transcription, we hypothesized that the ERE composition in ER α binding regions containing RUNX1 motifs might be distinct. Data were divided into four categories based on estrogen response and presence of the RUNX1 motif (Down, with; Up, with; Down, without; Up, without – Figure 5). Utilizing analyses used to assess ERE composition globally, it was found that the EREs within the ER α binding regions not containing RUNX1 did not differ significantly from one another, taking on the appearance of analyses in Figure 3.

Conclusions and Discussion

In the first of our *in silico* analyses, comparison of the distributions of distances between activated and repressed genes showed distinct differences. From these comparisons, we identified the median distance of the ERE from active genes as being 176 bases upstream, reasonably within the target gene promoter. However, median repression-associated EREs fall significantly further upstream (approximately 6.4kb). Although it might be likely that distance would prevent accessibility of E₂-bound receptor from accessing a target gene promoter, some gene enhancer regions may reside at sites more than 60kb away⁴⁸, thus, it may not follow that this would lead ultimately to repression of the target gene. Instead, it is more likely that the effect of this distance is due to the spatial

arrangement of chromatin⁴⁹, suggesting that the distance of repressor sites from TSSs and transcriptional factories may be an important feature of ER α function.

Compositional analysis of the ERE via Weblogo revealed very little differences between EREs associated with repression and those associated with activation. This may be in part due to the fact that the model for detection of the ERE is based on a consensus sequence whose constituents were derived primarily from activated genes. This introduces a caveat in utilizing such a model: if repression-associated EREs are actually distinct from activation-associated EREs, differences are likely to be difficult to detect.

In addition to investigating the distance of ERE from target gene TSS, we hypothesized that transcription factors other than ER α might bind sequence adjacent to the ERE, and we investigated their presence TRANSFAC. The resulting list of transcription factors, like any bioinformatics-based lists, was composed of statistically enriched TF motifs, some of which had known interaction with cancer, and some of which had little known biology at all.

In order to determine the functionality of selected regions and test the potential involvement of statistically enriched TF motifs in ER α mediated transcriptional regulation, we decided to proceed with functional analysis via reporter assays.

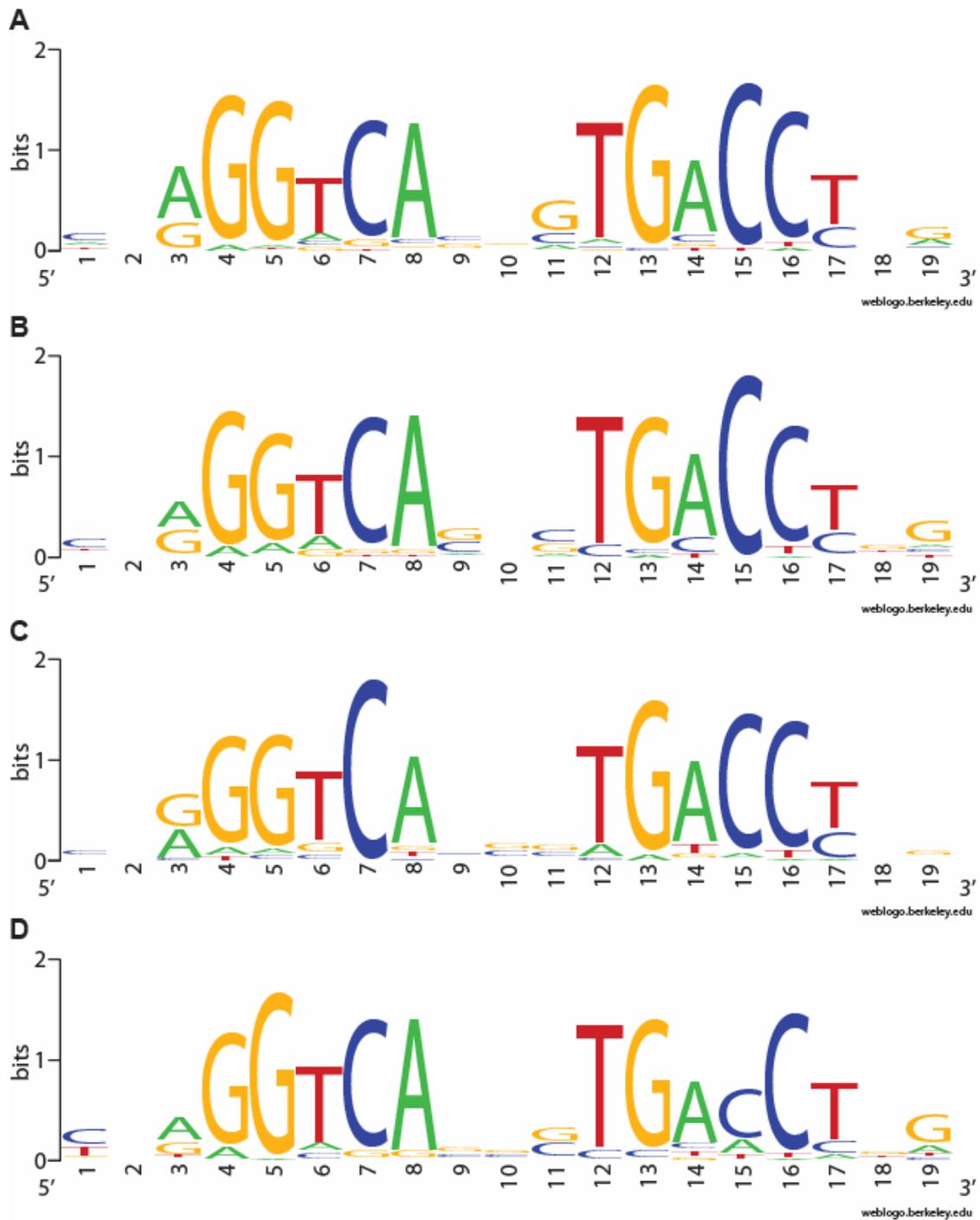


Figure 5 - EREs in Regulatory Regions Adjacent to Down Regulated Genes Containing RUNX1 Motifs Are Distinct From Those Without RUNX1 Motifs and From Up Regulated Genes Containing RUNX1 Motifs

Sequence logos of EREs not adjacent to RUNX1 motifs associated with down-regulated genes **A** and up-regulated genes **B** do not differ significantly from each other. EREs adjacent to RUNX1 motifs associated with down-regulated genes **C** and up-regulated genes **D** show significant differences.

Chapter 3: Functional Analysis of Putative Regulatory Sequence

Introduction

Selection of Binding Sites Adjacent to E₂-Repressed Target Genes

Hypotheses arising from computational and statistical analyses require further testing experimentally to establish their validity. In the context of our *in silico* analyses, we decided to first confirm that ER α binding regions actually function as repressive motifs for associated target genes. Five binding sites adjacent to estrogen repressed target genes ZNRF3, PSCA, NFIA, HES1, and SLC35A1 were chosen from the Merrell et al study as candidates based on their relative proximity to the TSS 5' distal, 5' proximal, intragenic, 3' proximal and 3' distal, respectively and represented a logical starting point for validation of our *in silico* findings. ZNRF3, HES1, and NFIA were discarded during preliminary studies due to performance inconsistencies. An additional site adjacent to the E₂ repressed gene MME was also added based on relevance to patient survival outcomes in an independent, unpublished study (Chin-Yo Lin).

In previously conducted studies, Merrell and colleagues showed that each binding site exhibited E₂-responsive recruitment of ER α and co-repressors, de-recruitment of co-activators, deacetylation of histones, and expulsion of polymerase. Each site exhibited similar repression-associated characteristics in terms of co-factor recruitment and reduced

expression of target gene, nevertheless, each exhibited these characteristics in a slightly different fashion. Moreover, TRANSFAC analysis of the binding sites adjacent to MME, PCSA, and SLC35A1 (Figure 6) revealed the presence of several TF motifs found to be significantly enriched adjacent to repressed genes. Even though the mechanisms for involvement of these three genes in breast cancer remain unclear, they still represent a sample of the potential mechanisms whereby ER α represses target gene expression in breast cancer. Therefore, we proceeded to test the functionality of these DNA regions by cloning the ER α binding regions for MME, PCSA, and SLC35A1 into reporter constructs and performing reporter assays. Theoretically, if the region of DNA actually recruits ER α in response to E₂ treatment and in turn represses target gene expression, this would be indicated by a reduction of luciferase activity in comparison to the control.

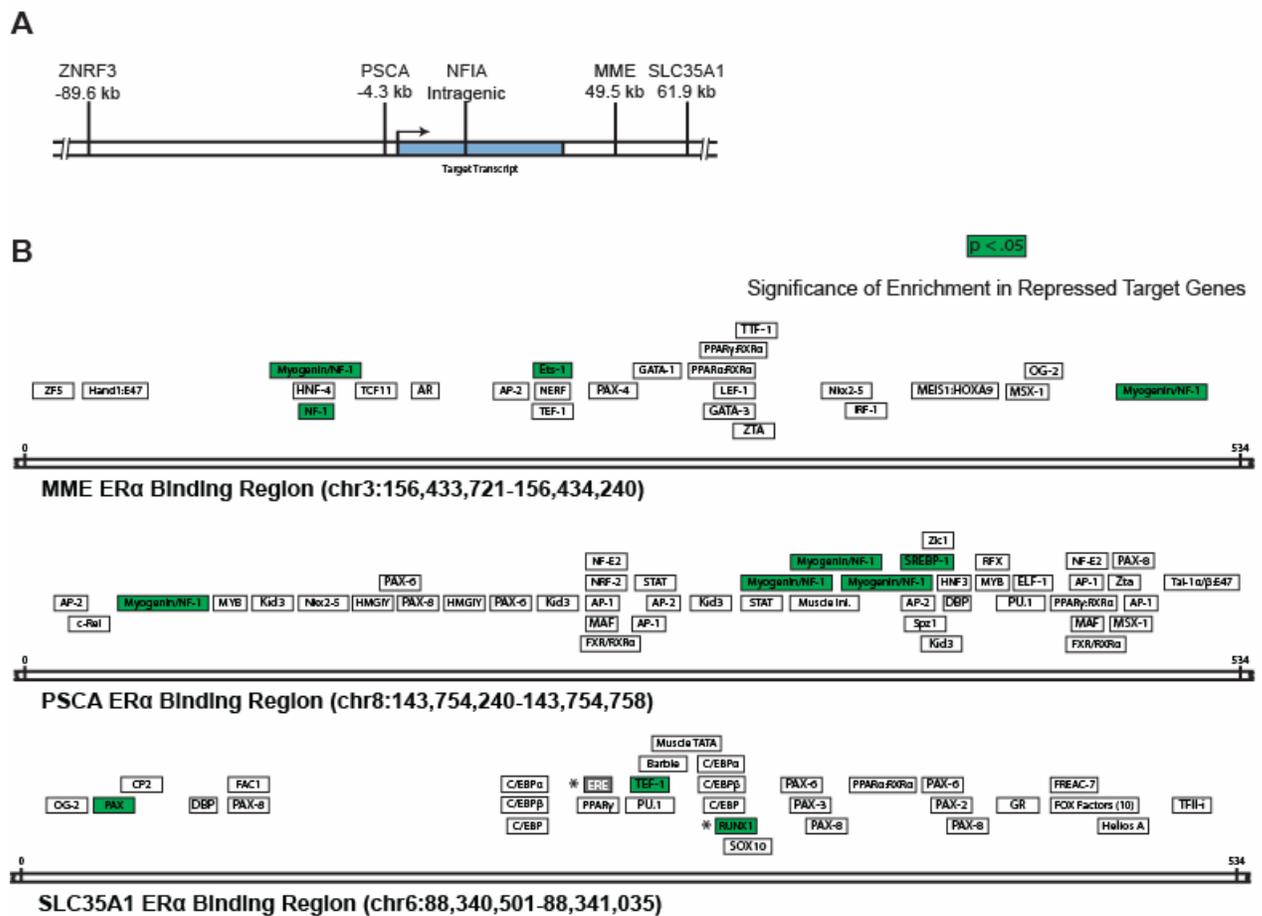


Figure 4 - Transcription Factor Motifs Within Selected ERα Binding Regions Adjacent to MME, PSCA, and SLC35A1

Depiction of TRANSFAC analysis and T-test p-values associated with detected binding site frequencies. A Relative location of original selected binding regions to associated target genes. B Transcription factor motifs mapped to binding region sequences. Size of boxes reflect only the size of the TF name and do not indicate relative size of TF motif. Green colored boxes indicate motifs whose occurrence is associated with repressed target genes at a p-value of $< .05$. RUNX1 and ERE marked with *.

Methods

Cell culture

MCF-7 cells are derived from a malignant adenocarcinoma of the breast, express estrogen receptor, and are considered to be an effective model system for studying E2-responsive breast cancer⁵⁰. Adherent MCF-7 adenocarcinoma cells were obtained from ATCC (ATCC Number HTB-22) and maintained at 37°C, 5% CO₂, with 10% fetal bovine serum suspended in phenol-red Dulbecco's Modified Eagle Media (DMEM). Cell passages were maintained below passage 20. Hormone starvation was maintained in 5% charcoal-stripped fetal bovine serum suspended in phenol-red free DMEM. Though each protocol's serum starve was unique, all protocols require starve after cells reach at least 50-60% confluence. Following protocol-specific serum starve, cells were exposed to 10nM 17-β Estradiol (E₂) for 45 min or 24 hours, depending on experiment. Passaging was performed when cells were approximately 80% confluent.

Overview of Cloning, Luciferase Assay, and Analysis of ER α -Binding Site

DNA

In general, reporter assays test the function of a section of DNA in regulating the expression of an easily measurable target gene. First, the target regions of DNA are amplified via PCR and cloned into plasmids upstream of the reporter gene – in this case, luciferase. The plasmids are transfected into the cell line of choice, MCF-7. Transfections are followed by treatments such as estradiol and treatment-appropriate controls.

Luciferase expression is assayed by adding luciferin and ATP, necessary substrates for luciferase luminescence. Luminescence is then measured using a luminometer and compared to control. In the context of the dual luciferase assay, secondary plasmids containing a constitutively expressed *Renilla* luciferase gene are co-transfected with individual firefly luciferase plasmids to provide a transfection efficiency background control.

Cloning

To assess the role of ER α binding regions in regulating expression of associated target genes, the regions adjacent to the remaining genes, MME, PSCA, and SLC35A1 were cloned into luciferase gene reporter constructs amplified by conventional PCR using DNA from MCF7 cells and primers from Table 2. PCR reagents were purchased from Roche (FastStart High Fidelity PCR System, dNTPack, Cat. #04738284001). Cloning enzymes were order from NEB (NheI and XhoI, Cat. #R0131S and #R0146S) and Promega (T4 DNA Ligase Cat. #M1801). Following cloning, plasmids were amplified using kits purchased from Qiagen (QIAprep Spin Miniprep Kit Cat. #27106, QIAGEN Plasmid Midi Kit, Cat. #12145), and cloning was confirmed via cycle sequencing (ABI 3730xl DNA Analyzer).

Transfection

Following a 72-hr serum starve, cells in a 150mm cell culture dish at $\geq 80\%$ confluence were washed once with 1X PBS and removed by trypsin digest (2ml .05% Trypsin EDTA

with 2ml 1X PBS). Digested cells were suspended in 25ml (final volume) starve media. 500µl cell suspension was then aliquoted into 1ml starve media per well of a 24-well plate and incubated for 24hrs at 37°C. The following day, cells were ~80% confluent. Individual transfection reactions were then prepared as follows: 6µl of firefly luciferase plasmid (Promega, pGL3-Control, Cat. #E1714 and 2ERE-pGL4 gift from Edwin Cheung) 250ng/µl and 6µl of *Renilla* luciferase plasmid (Promega, pGL4.75 hRluc-CMV, Cat. #E6931 gift from Edwin Cheung) 250ng/µl were combined in 38µl of phenol-red free DMEM with no additives. Another tube was prepared with 5µl Fugene HD transfection reagent (Roche, Cat. #04709705001) in 45µl additive-free DMEM. After 5min, both tubes were combined into a 15ml conical vial incubating at 25°C for 15-30min. After incubation, 3ml starve media were added to each 15ml conical vial and mixed. The media was then removed from each well of the 24-well plate and 500µl of transfection reaction was added to each well of one row of the 24-well plate. The 24-well plate is then returned to incubation at 37°C for 24hrs. After 24hr. incubation, starvation media containing 10nM E₂ or an equal volume of 100% EtOH is added to each well of the 24-well plate and incubated at 37°C for 24hrs. Following incubation, media is removed and each well is washed once with 1ml 1X PBS before proceeding to Dual-Glo Luciferase assay.

Dual-Glo Luciferase assay

The Promega Dual-Glo Luciferase Assay Protocol (Technical Manual TM058) was used with modifications: 200µl of Dual-Glo Luciferase Reagent (Dual-Glo Luciferase Assay System, Promega, Cat. #E2920) was added to each well of a 24-well plate, and allowed

to incubate at 25°C for 10min and luciferase activity is measured with a FLUOstar Optima Plate Reader using the Luminescence setting. This procedure was repeated with 200µl Dual-Glo Stop & Glo Reagent. Quantified firefly luminescence was then normalized against quantified *Renilla* luminescence levels. The resulting ratios were then averaged across replicates and the fold induction of luciferase luminescence was calculated. Each set of transfection and luciferase assay experiments was fully replicated three times, and averaged.

Site-Directed Mutagenesis

Site-specific mutants were generated to assess the functionality of relevant ER α binding regions in modulation of reporter gene expression. Mutagenesis reagents were purchased from Stratagene (QuikChange II Site-Directed Mutagenesis Kit, Cat. #200524), and manufacturer's protocols followed with modifications: 1µl each primer at 10µM concentration, elongation for 15min per cycle, and use of LB broth during transformation of *E. coli* instead of NZY⁺. Mutation specific primers were designed using Stratagene Quick Change Primer Design Program (<http://www.stratagene.com/tradeshows/feature.aspx?fpId=118>) and may be found in Table 3. Following mutagenesis, colonies were picked, plasmids sequenced and amplified following protocol from cloning.

Primer		Sequence
MME	For	GATTAG CTAG CCCAAGCTCGTTCATGTCACC
	Rev	GATT ACTCGAG CCTGTCACCATCAGACTGCT
PSCA	For	GATTAG CTAGC GTGCATCAGTGGGTTTTCTC
	Rev	GATT ACTCGAG AGCGTCCTCCACTGTGAAAC
SLC35A1	For	GATTAG CTAGC CTAAGAGATGCTGCCCAAGC
	Rev	GATT ACTCGAG TGCCTCCCTCCCTATTCTCTC

Table 2 - Primers for Cloning ER α Binding Regions

Primers written 5'-3' orientation. Bolded nucleotides correspond to restriction sites for NheI (Forward primers) and XhoI (Reverse primers). Additional nucleotides (GATTA) added to the 5' end as a stabilizer.

Primer		Sequence
ERE ^{mut}	For	GAGGAATAATGTGTACAG t CACCAT t ACCTACCAGAGAGAGAGA
	Rev	TCTCTCTCTCTGGTAGGT a ATGGTGA a CTGTACACATTATTCCTC
RUNX1 ^{mut}	For	GGGGAGGAGTTTATATTTTTGCAA a CA a AAAGGTGAGGTAAGCCAAAA
	Rev	TTTTGGCTTACCTCACCTTT t TG t TGCAAAAATATAAACTCCTCCCC

Table 3 - Primers for Site-Directed Mutagenesis

Primer written 5'-3' orientation. Bolded, lower case nucleotides indicate deviations from parental sequence. Reverse primers are reverse complements of the forward primers. In designing primers, T was exchanged for G, and A for C.

Results

MME, PSCA, SLC35A1

Of the five binding sites initially selected, only PSCA and SLC35A1 were retained for further analysis. The PSCA (prostate stem cell antigen) gene codes for a membrane-anchored glycoprotein and is overexpressed in many prostate cancers⁵¹. SLC35A1 (solute carrier family 35 member A1) is a membrane-spanning transport protein responsible for moving sialic acid across the Golgi membrane. A third gene and associated ER α binding region was selected, MME (membrane metallo-endopeptidase), as an independent study (Unpublished data, Chin-Yo Lin, 2007) demonstrated its association with positive outcomes during adjuvant therapy for breast cancer. Like PSCA, MME is also highly expressed in prostate neoplasia, and encodes a membrane bound peptidase⁵².

Establishment of reporter gene assays allowed for assessment of ER α binding regions' ability to suppress target gene expression (Figure 7a). In comparison to positive and negative controls, inserts corresponding to the binding sites adjacent to MME and PSCA did not show significant E₂-responsiveness as indicated by the ≤ 1 fold difference from negative controls. The binding site for SLC35A1 did show E₂-responsiveness, indicating that ER α binding at this site does occur and does interact with transcriptional machinery. An unexpected finding was that the response reflected an approximately 4-fold increase in luciferase activity as opposed to a reduction as was expected. No other phenotypic differences were noted.

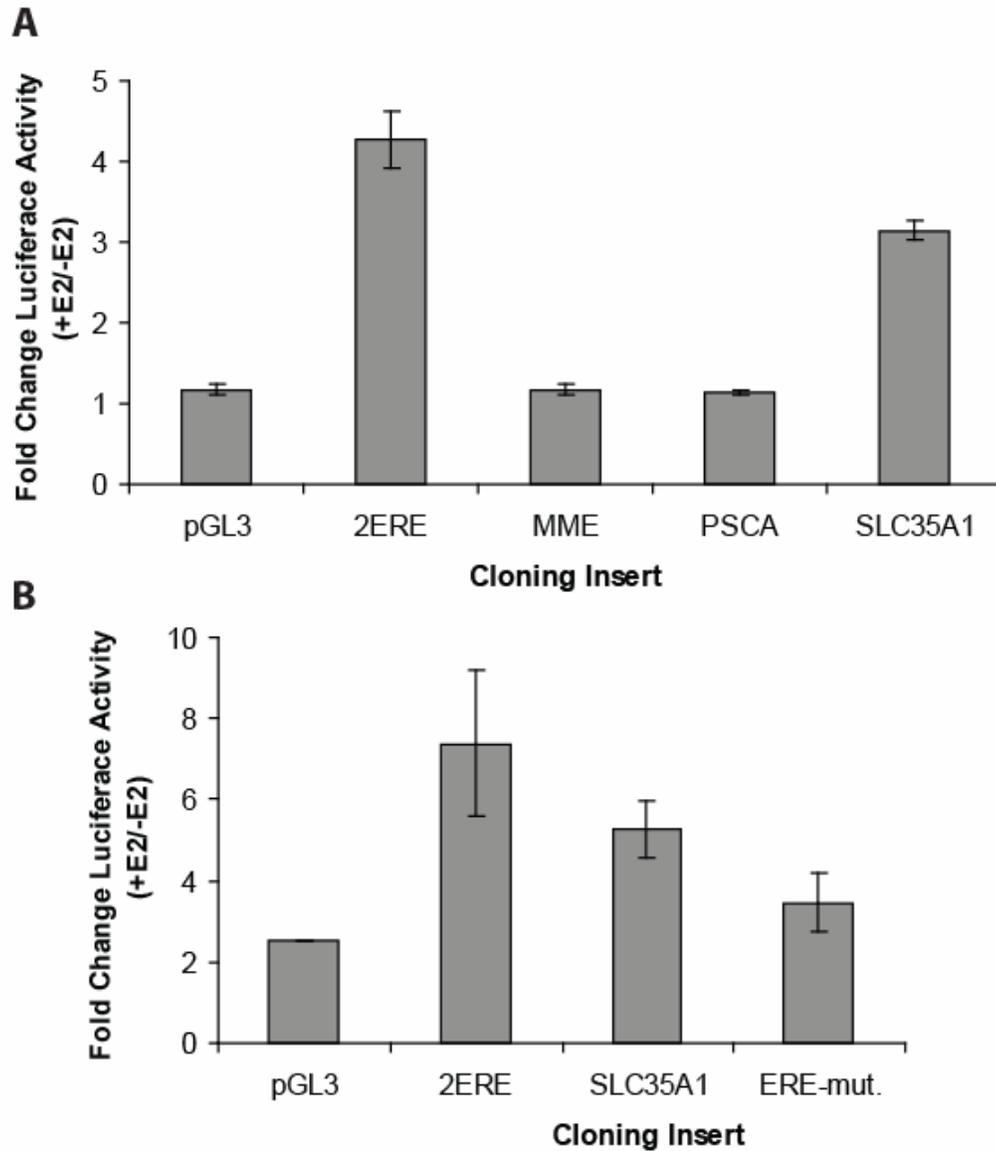


Figure 5 - ERE and RUNX1 Motifs Direct E2 Responsive Activity In The ER α Binding Region Adjacent to SLC35A1

Quantification of fold change in luciferase expression following treatment with 10nM estrogen treatment. Luciferase assays calculated as a ratio of firefly to *Renilla* luciferase expression, estrogen treated samples divided by ethanol treated samples. **A** Effect of ER α binding regions adjacent to MME, PSCA, and SLC35A1 cloned into pGL3 vector on luciferase reporter activity. Only the binding site adjacent to SLC35A1 exhibits E2 responsiveness. **B** Effect of site-specific mutagenesis on ER α binding region adjacent to SLC35A1. Mutation of ERE resulted in diminished activity in response to estrogen exposure.

As mentioned previously, regions surrounding genomic EREs contain motifs for other transcription binding. To extend the findings of the dual luciferase assays, we performed site-directed mutagenesis to reduce the ability of target DNA to recruit putative transcription factors involved in repression (Figure 7b).

We decided to investigate the binding sites for RUNX1 because they occurred more frequently than any other TF motifs and had already been described as involved in cancer mechanisms. RUNX1 motifs not only occurred more frequently adjacent to repressed target genes, but also were associated with EREs whose composition was distinct between activated and repressed genes. These findings suggest that while other transcription factors may play a role in regulating ER α -mediated transcriptional repression, RUNX1 was the most statistically sound factor to investigate.

Mutation of the ERE diminished luciferase activity, indicating that E₂-responsive induction of luciferase activity depends on both of these factors. The RUNX1 mutant plasmid was not consistently lower than the unmutated control, but exhibited a trend of reduced activity.

Conclusions and Discussion

The ER α binding region adjacent to SLC35A1 has been shown by Merrell and colleagues to recruit ER α , co-repressors, allow for histone deacetylation, and expel polymerase – all

within the context of a breast cancer cell line. Aside from this, SLC35A1 has little known involvement in breast cancer. SLC35A1 is a sialic acid transporter commonly found in the Golgi membrane, and is involved in sugar-nucleotide import for protein modifications⁵³⁻⁵⁴. SLC35A1 expression in breast cancer has not been studied; however, it has been shown to be a marker for osteosarcoma in mice and humans⁵⁵. This finding is intriguing and indicates a potential for SLC35A1/ER α involvement - ER α is responsible for many processes in bone formation⁵⁶⁻⁵⁷ and breast cancer metastases often preferentially colonize bone⁵⁸⁻⁵⁹.

Our findings indicate that while downstream events proceeding from ER α recruitment may direct the effective mechanisms of transcriptional regulation, cis-regulatory DNA sequence represents a dynamic environment which affects all downstream events. Despite these findings, the reporter gene context proved insufficient to validate repression. The binding site adjacent to SLC35A1 exhibited induction instead of repression of the target gene. This anomaly may be due to the absence of histones and the dynamic chromatin context present *in vivo*, indicating the importance of chromatin remodeling in transcriptional repression. This hypothesis is strengthened by findings that indicate that histone deacetylation is a putative hallmark of repression associated ER α binding regions, specifically at the regulatory sequence adjacent to SLC35A1. To better assess the role of cis-regulatory DNA elements in mediating transcriptional repression, we chose to perform additional experiments that monitor the *in vivo* binding of DNA by ER α .

Chapter 4: RUNX1 Co-recruitment is Necessary for Estrogen Receptor Mediated Transcriptional Repression of SLC35A1

Introduction

ER Interactions with Other Transcription Factors

Most studies of ER α signaling focus on ER α homodimers. Additionally, extensive evidence indicates that ER α forms functional heterodimers with ER β , as well as other proteins²³. Specifically, ER α has been shown to heterodimerize with transcription factors such as Fos⁶⁰, Jun⁶¹⁻⁶², and NF-kB⁶³⁻⁶⁴. The effects of these interactions are further complicated by the fact that at some regulatory sites, ER α does not participate directly in DNA binding at all, but is tethered to target chromatin via interactions with FOXA1 and SP1⁶⁵⁻⁶⁷. In short, the activity of ER α is dynamic, affected by various functional elements and the chromatin context as a whole. Taken together, these observations indicate that sequence adjacent to putative EREs may be responsible for recruiting various transcription factors capable of affecting ER α regulation of target gene expression.

RUNX1 Biology and Carcinogenesis

There were a large number of potential transcription factors, so preliminary work was necessary to reduce the list of candidates. As previously described, TRANSFAC analysis of regions adjacent to E₂ target genes revealed that RUNX1 motifs occurred more frequently than any other motif adjacent to repressed target sites, and thus was a prime candidate for further screening. From a biological perspective, RUNX1 is an interesting protein – it has known association with breast cancers and leukemias.

RUNX1 also known as AML-1, is a member of the core-binding factor alpha (CBF α) class of transcription factors, and was first described in leukemia. CBFs usually consist of heterodimers of a DNA-binding α subunit (RUNX1, RUNX2, or RUNX3) and a non-DNA-binding β subunit (CBFB). CBFs are necessary factors for hematopoiesis⁶⁸, and their mutation is one of the most frequent events resulting in leukemogenesis⁶⁹⁻⁷⁰. Despite its unclear role in breast cancer, RUNX1 expression is enriched in mouse mammary tissue⁷¹, over expressed in many breast cancer cell lines⁷², and exhibits estrogen-responsiveness⁷³⁻⁷⁵, and has been shown to be a marker for breast cancer⁷⁶⁻⁷⁷. These preliminary findings suggest that the motifs flanking the ERE, and the downstream recruitment of transcription factors may be influential on overall regulatory effects, and are therefore key to uncovering the underlying mechanisms of E₂-mediate repression. The most serendipitous finding was that a RUNX1 motif occurs within 80bp of the 3' end of the ERE on the ER binding site adjacent to SLC35A1. This discovery enabled functional characterization of RUNX1 involvement at a site which was already being studied.

During our functional characterization of MME, PSCA, and SLC35A1, it became clear that with respect to the binding sites in question, the dual luciferase assays did not have the capacity to adequately recapitulate the deacetylation of histones associated with target gene repression. Thus, it was determined that an experimental system would have to be established wherein an *in vivo* environment was more closely represented.

In order to select a more *in vivo* approach to monitoring the proteins associated with chromatin, we selected chromatin immunoprecipitation (ChIP). ChIP does not assay the level of gene expression, but rather the relative binding to DNA of a DNA binding protein. In this case, the ChIP assay would focus on binding of RUNX1 to the SLC35A1 ER-binding site, and whether this binding occurs in combination with ER α . To confirm the role of RUNX1 plays in ER α -mediated repression of SLC35A1, we used RNA interference (RNAi) to disrupt RUNX1 expression. If RUNX1 actually participates in the ER α -mediated repression of SLC35A1, RNAi knockdown of RUNX1 mRNA should lead to a reversal of this repression.

Methods

Cell culture

Cellular cultures were maintained as outlined in the Reporter Assay section. E2 treatments were performed for 45min (ChIP and re-ChIP) or 24hrs (siRNA).

ChIP

All chromatin immunoprecipitation assays followed previously described procedures³⁴.

In summary, following a 72-hour serum starve, cells were treated with 10nM E₂ or vehicle for 45 min. After E₂ treatment, cells were washed twice with 1X PBS, and crosslinking performed with 1.5% Formaldehyde in PBS. Following crosslinking, cells were washed, collected with cell scrapers and suspended in collection buffer. The cell pellet was then washed once each with nuclear collection buffers - 1X PBS, Nuclear/Chromatin Preparation Buffers I and II - in order to isolate cell nuclei. Following nuclear isolation, nuclear suspensions were lysed in Lysis Buffer, sonicated on ice, and centrifuged at 10,000-x g for 10min at 4°C, and DNA-protein complexes collected from supernatant. Aliquots were removed from these supernatant samples, and set aside as inputs. The remaining supernatant samples were incubated in IP Buffer with antibodies overnight with rocking at 4°C. After antibody incubation, IP samples were treated with 40µl of a 50% Protein A/Beads Buffer slurry and incubated further ≥3 hours. Following incubation, beads were washed as follows: once each with Washing Buffers I-III, and three times each with Washing Buffer IV. Following beads pellet washes, DNA-protein

complexes were extracted from beads solution with Extraction Buffer. Input and IP samples were then de-crosslinked overnight at 65°C to evaporate formaldehyde. DNA was then purified using the Qiagen PCR Cleanup Kit, eluted in 25µl ddH₂O, and stored at -20°C. Relative abundances of DNA were quantified via qPCR in 96-well plates utilizing a Roche Light Cycler 480. Each reaction consisted of 1µl DNA extraction, 5µl SYBR Green, 5pmoll each forward and reverse primers, and 3µl ddH₂O. Primer sets may be seen in Table 4. Experiments were further visualized by running PCR products on a 1% agarose gel.

Re-ChIP

Re-ChIP deviated from standard ChIP procedures during the first immunoprecipitation (IP). After incubating ChIP samples with ERα antibody overnight, 40µl beads slurry was added and incubated as usual. Beads were then washed once with Washing Buffer I and twice with Washing Buffer II, extracted three times with 100mM DTT, and resuspended in Re-ChIP Buffer. Secondary (RUNX1) antibodies were then added, and the remaining procedures follow standard ChIP protocol. Experiments were further visualized by running PCR products on a 1% agarose gel.

SiRNA knockdown of RUNX1

Cells at 80% confluency were seeded ~1 part in 25 to 6-well plates. After reaching 50% confluency in 6-well plates, media was removed and transfection reagents added following protocol from Dharmacon. In summary, 3µl Dharmafect A reagent in 197µl

additive-free DMEM were combined with 2.5µl SMART pool siRNA (Dharmacon, Cat. # LQ-003926-00) in 197.5 µl additive-free media, incubating at room temperature for ≥ 40 min. Following incubation, 400µl transfection reaction was added to each well, and treatment lasted for 24 hours. After 24-hour incubation, each well was treated with 10nM final concentration E₂ or ethanol vehicle and incubated for another 24 hours. Upon completion of E₂ treatment, cells were lysed and RNA extracted with 1ml Trizol (TRIzol reagent Invitrogen, Cat. #15596-026) following manufacturer's protocol. Trizol-extracted RNA aliquots were quantified with Nanodrop mass spectrometry, and diluted (if necessary) to 1µg/µl concentration and stored at -80°C.

Reverse transcription of RNA into cDNA was performed on RNA extraction samples using Promega ImProm II (Promega, Cat. #A3800) reverse transcription reagents and protocols using 20µl reactions. Equal quantities (1µg) of RNA were reverse transcribed in order to maintain consistency between samples and control for variation in cell count and confluency. Final cDNA products were suspended 1:5 in 100µl ddH₂O and stored at -20°C. Relative RNA expression levels were quantified using SYBR Green qPCR (Lightcycler 480 SYBRGreen I Master, Roche, Cat. #04707516001) reagents performed in 96-well plates in a Roche Light Cycler 480 PCR machine. Each PCR consisted of 4µl diluted cDNA, 5µl SYBR Green, 5pmol each forward and reverse primers. Primer sets used are shown in Table 7. Experiments were further visualized by running PCR products on a 1% agarose gel.

Primer		Sequence
SLC35A1 ERE	For	GCAACCACAAAGGTGAGGTA
	Rev	CACCTATTGGTGATGCCTGT
SLC35A1 Exon 4	For	GACCTACCAGTTGAAGATTCCG
	Rev	AAGCGTAACTCCAGCACACA

Table 4 - Primers for ChIP and re-ChIP

Primers written in 5'-3' orientation. Primers were designed to amplify a region 100-200bp in length, so the conditions differ slightly from those for cloning primers. Exon 4 contains no ERE or RUNX1 sites and acts as a negative control

Primer		Sequence
36B4/RPLPO	For	GTGTTTCGACAATGGCAGCAT
	Rev	GACACCCTCCAGGAAGCGA
pS2/TFF1	For	GACAGAGACGTGTACAGTGG
	Rev	ACAACAGTGGCTCACGGGCT
RUNX1	For	TCCATTGCCTCTCCTTCTGT
	Rev	GGGTCGCTGAACGCTGTC
SLC35A1	For	GTGGCCTCTACACTTCTGTT
	Rev	GGATGTAGTGTCTTGTCTGG

Table 5 - Primers for siRNA and Expression Studies

Primers written in 5'-3' orientation. 36B4 corresponds to the sequence for the large ribosomal protein mRNA, and serves as a housekeeping gene/negative control. pS2/TFF1 is a classical estrogen induced gene and serves as a positive control.

Results

In order to determine the binding of RUNX1 to the ER α binding region, we performed ChIP assays (Figure 8). ChIP primarily allows relative quantification of DNA bound by transcription factors or other DNA binding proteins. RUNX1 binding was first assayed and quantified with qPCR using SYBR Green reagents. At the SLC35A1 ER α binding region, RUNX1 associates with target DNA, binding ~40 fold over the SLC35A1 exon 4 negative control (Figure 8a-b). It also appears that this association is E₂-independent, as indicated by the overlapping ranges in standard errors on the RUNX1 binding results distribution and results from a PCR run on agarose gel (Figure 8c).

We followed the RUNX1 ChIP with a re-ChIP to test the hypothesis that at regulatory DNA adjacent to repressed target genes, RUNX1 and ER α are in the same molecular complex. The ER α ChIP product was immunoprecipitated with a secondary ChIP against RUNX1 and found that RUNX1 and ER α do co-recruit to the SLC35A1 regulatory sequence 50% (E₂-) and 80% (E₂+) greater than inputs while SLC35A1 exon 4 negative control recruitment is negligible (Figure 9a). In addition to this finding, it appears that though RUNX1 recruitment is E₂-independent, E₂ signaling is still necessary for co-recruitment of ER α , indicated by a ~1.5-fold increase in binding of SLC35A1 ERE following E₂ treatment (Figure 9b). PCR products were also run on an agarose gel (Figure 9c)

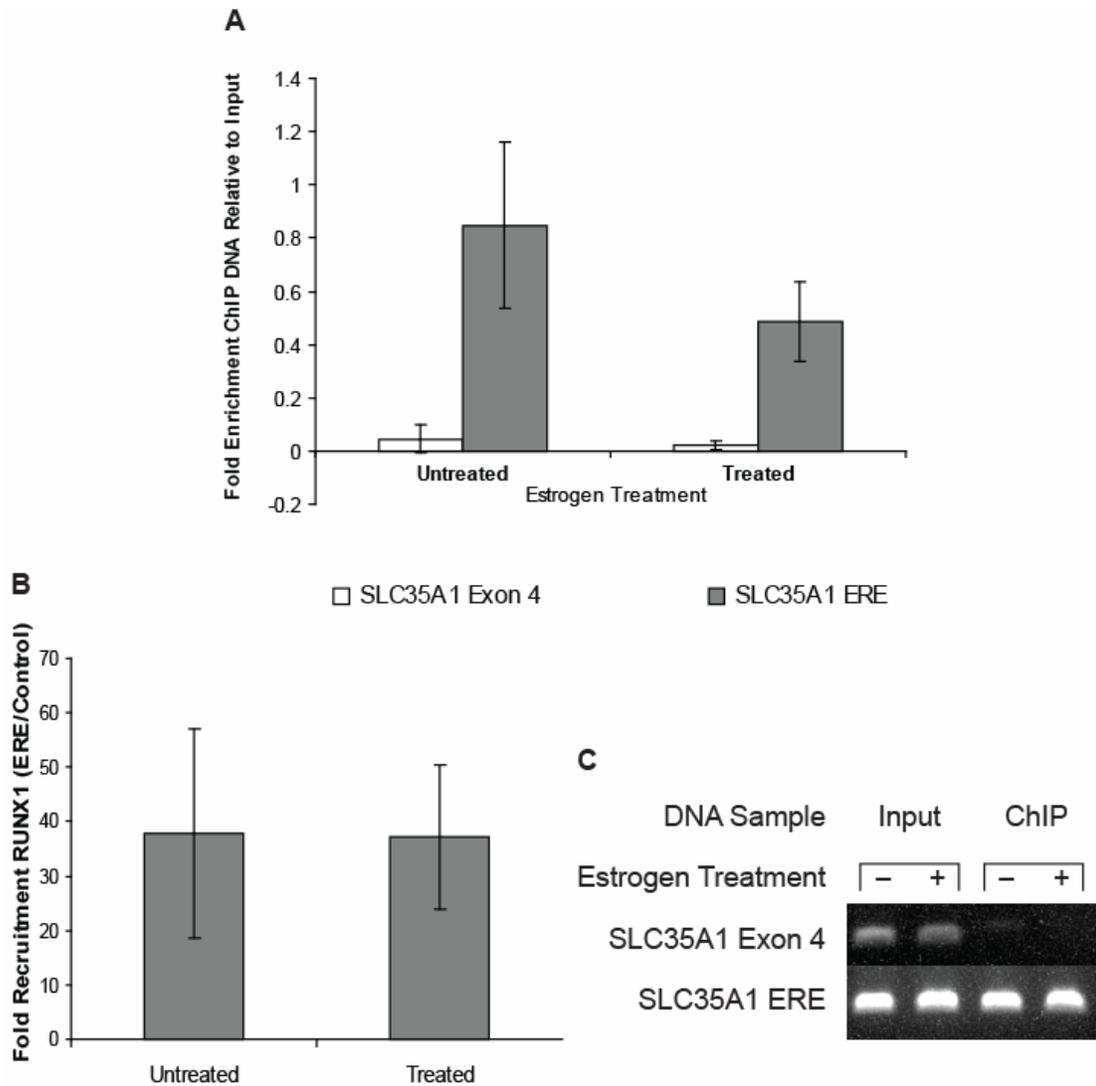


Figure 8 - RUNX1 Is Recruited to the SLC35A1 Adjacent ERE and Binds Independent of Estrogen Treatment

A ChIP against RUNX1 quantified via qPCR. Primers were specific to the SLC35A1 ER α binding site and a non-target, SLC35A1 primer. **B** After normalizing against controls, estrogen response is shown to be negligible. **C** Gel picture of PCR on ChIP and Input DNA samples. Enrichment of SLC35A1 DNA following RUNX1 ChIP indicates that RUNX1 binds this sequence in the presence or absence of estradiol.

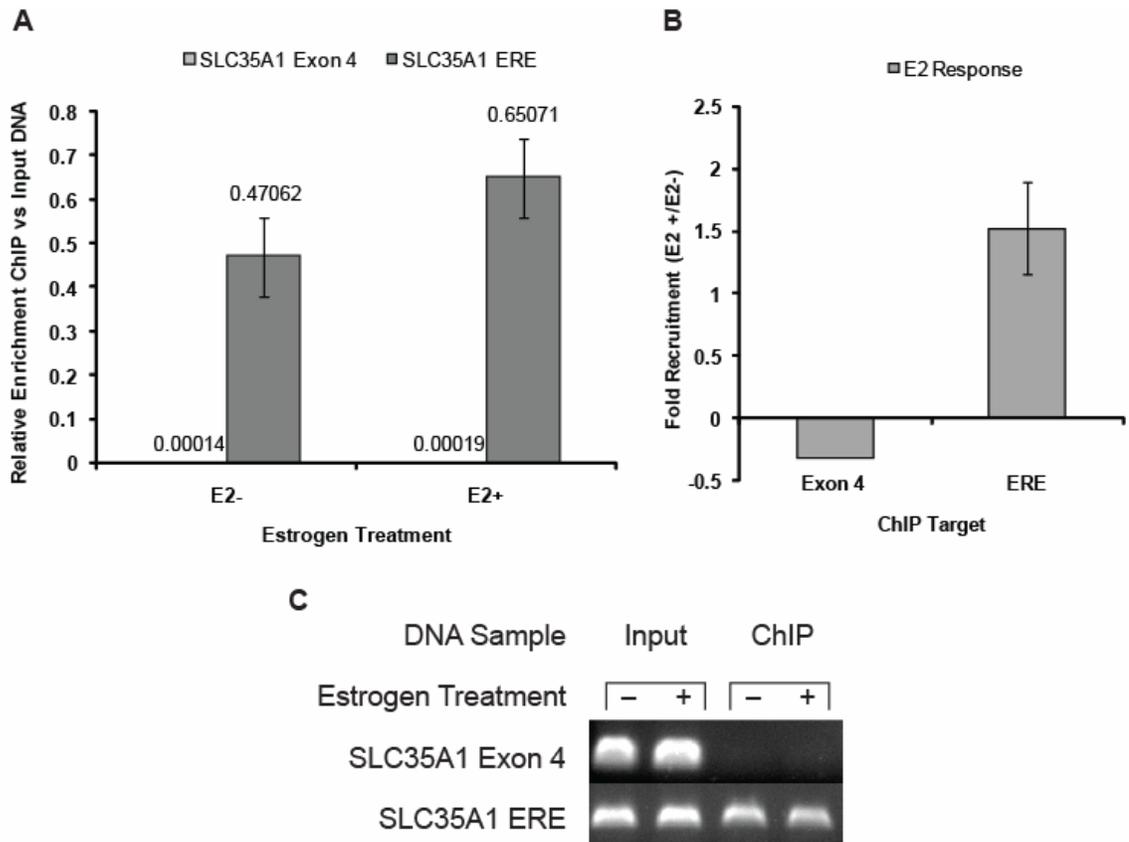


Figure 9 - RUNX1 and ER α Co-Recruit to the SLC35A1 Adjacent ERE and Exhibits Moderate Estrogen Response

A ChIP against ER α followed by re-ChIP against RUNX1 quantified via qPCR. Binding is depicted as relative enrichment of ChIP DNA over inputs from the same experiment. Exon 4 enrichment was too low to show with bars, so numbers were chosen to indicate relative enrichment. **B** Fold recruitment of ER α -RUNX1 complex to the SLC35A1 ERE and SLC35A1 Exon 4 in response to E2 treatment. **C** Gel picture of PCR on ChIP and Input DNA samples. Response to estrogen was sufficient for detection by LightCycler, though the cycle differences are too slight to distinguish on the gel

In order to test the role of RUNX1 in regulating SLC35A1 expression, we performed RNAi knockdown against RUNX1 and quantified the results with qPCR (Figure 10). Following knockdown, RUNX1 is knocked down approximately 5-fold following siRNA treatment. This knockdown appears to be E₂-independent (Figure 10a), suggesting that RUNX1 binds the ER α binding region independent of ER α recruitment. The typical ~2 fold repression of SLC35A1 is reversed following RUNX1 knockdown (Figure 8B), indicated by a ≤ 1 fold difference from control treatments and PCR products were run on an agarose gel (Figure 10c).

Conclusions and Discussion

In the context of SLC35A1, RUNX1 binding is key to E₂- mediated repression. RUNX1 binds the SLC35A1 ER α binding region prior to E₂ treatment, and may thus be responsible for establishing an appropriate context for ER α recruitment of co-repressors and other events necessary for transcriptional inhibition. In the absence of RUNX1, ER α is no longer sufficient to repress SLC35A1 expression, perhaps due to conformational differences. The physical differences in repressive and activating conformations may be slight, though they are not investigated in this study. Regardless of the extent to which ER α conformational shifts take place following RUNX1 co-recruitment, this event is sufficient to promote co-repressor recruitment and activity and inhibit co-activator activity, perhaps by eliminating the interface for co-activator action. It is also unclear whether ER α and RUNX1 physically interact, or whether they are

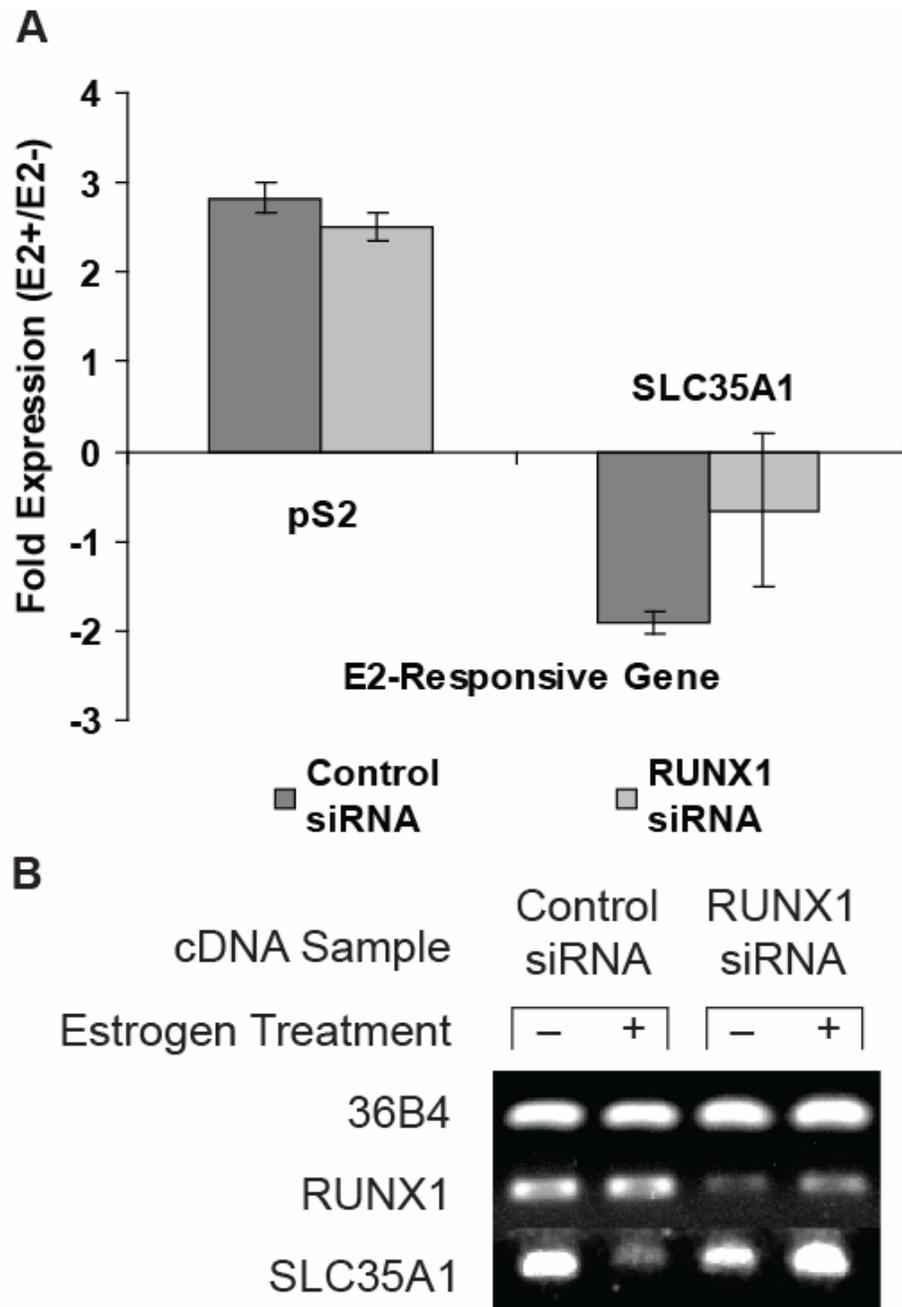


Figure 10 - siRNA Mediated RUNX1 Knockdown Reverses Transcriptional Repression of SLC35A1

A Quantification of cDNA amplification with qPCR shows that when siRNA against RUNX1 is transfected, control gene expression is unaffected, while SLC35A1 repression is reversed. B Gel on PCR products. A reduction in RUNX1 expression is shown and SLC35A1 repression is relieved. RUNX1 knockdown is approximately 5 fold according to qPCR cycle calculations.

simply present as a part of a larger repressor complex. Future extensions to investigate RUNX1 and ER α interactions and to identify other components essential for repression may include co-immunoprecipitation (Co-IP), competitive binding assays, and alterations of ER α binding interfaces with SERMs. Additionally, siRNA titrations against ER α and RUNX1 may indicate threshold levels necessary for transcriptional repression.

Chapter 5: Summary and Conclusions

Mechanisms of ER α -Mediated Transcriptional Repression

Estrogen receptor action in response to E₂ exposure is necessary for healthy physiology, but is also a hallmark of malignant breast cancer. Breast cancer staging and aggressiveness are commonly assessed based on ER α expression status. ER α binds E₂ and thus mediates its effects on healthy physiology as well as breast cancer. By binding to response elements throughout the genome, E₂-bound ER α recruits various chromatin remodeling and transcriptional machinery, thereby modulating the expression of many genes. Though the majority of known E₂-responsive genes from previous studies are activated following exposure to E₂, our most recent findings indicate that ER α binds to more binding regions adjacent to repressed targets than those associated with activation. Cis-regulatory sequence differences are likely the cause of the characteristics that distinguish mechanisms whereby ER α possess the ability to both activate and repress gene activity. In order to assess the role that these cis-regulatory sequences play in affecting ER α regulation of gene expression, we performed computational analyses of ER α binding and effects on gene expression on genome-wide scale.

Our *in silico* analyses indicate that binding sites associated with repressed target genes are located significantly more distal to the TSS than those associated with activated

target genes. This suggests that accessibility to the target gene promoter is an important factor in determining level and direction of gene expression. We also found that ERE composition does not significantly distinguish repression from activation, suggesting that it is not a major determinant of ER α -mediated transcriptional control. Furthermore, many motifs associated with the recruitment of other transcription factors are present at greater frequency in ER α binding regions adjacent to repressed target genes as opposed to activated target genes. The most frequently occurring of these transcription factor motifs corresponds to AML1/RUNX1, a transcription factor involved in leukemias. Given the frequency with which RUNX1 motifs occur adjacent to repressed target genes, it is possible that RUNX1 is a common mediator of ER α repression of target genes. This hypothesis is strengthened by the fact that although ERE composition is not significantly different between activated and repressed genes, EREs associated with RUNX1 motifs do exhibit differences at more than one position, which may suggest that when in the presence of RUNX1, ER α adopts a repressive conformation resulting in recruitment of co-repressors and repression of target genes.

Following our *in silico* analyses, we performed functional *in vitro* and *in vivo* assays to validate and extend our findings. First, we utilized dual luciferase assays combined with mutagenesis studies of binding sites adjacent to confirmed E₂-repressed target genes. These binding sites did not adequately recapitulate the *in vivo* chromatin context, still they indicated that the ERE and the RUNX1 motif are both necessary for the estrogen mediated response of the binding site adjacent to SLC35A1, a known marker for

osteosarcoma. It appears that in the absence of either factor, target gene regulation is diminished.

After finding that the reporter assay system did not mimic the *in vivo* context of ER α binding, we performed ChIP assays to determine RUNX1 and ER α -RUNX1 binding to repression-associated ER α binding regions. Our ChIP studies indicate that RUNX1 binds to the SLC35A1 ER binding site approximately 70 bases from the ERE. Though RUNX1 binding is E₂-independent, re-ChIP indicated that E₂-responsive ER α co-assembles with RUNX1 at the SLC35A1 ER α binding region. More importantly, siRNA knockdown of RUNX1 reverses the E₂-responsive repression of SLC35A1, demonstrating that RUNX1 recruitment is necessary for SLC35A1 repression by ER α (Figure 11). This supports the findings of the reporter gene assays, which indicate that removal of either ER α or RUNX1 signaling abolishes their combined effect on target gene expression.

Given that SLC35A1 involvement in breast cancer is unknown, assessing the direct effect of RUNX1-ER α co-recruitment in breast cancer is difficult. RUNX1 elements occur with greater frequency than any other transcription factor motif in regulatory DNA adjacent to repressed target genes, and it is likely that RUNX1 plays a role in modulating ER α signaling at many sites in the genome. Additionally, the presence of both ER α and RUNX1 is required in order to mediate transcriptional repression of SLC35A1, suggesting that if similar binding sites exist, they may potentially be regulated by the same mechanisms. E₂ signaling is responsible for cellular proliferation, and in the context of RUNX1-ER α co-recruitment, it may be responsible for the repression of tumor

suppressor-like anti-growth genes. Thus, RUNX1 may be a potential target for future pharmacological studies, whose knockdown may allow the expression of suppressed growth-inhibitory genes in breast tumors.

Future Directions

Unpublished data from Merrell and colleagues suggest that siRNA knockdown of a single co-repressor is insufficient to completely abolish transcription of SLC35A1, but instead requires simultaneous knockdown of SMRT, NCoR, and NRIP1. In the case of RUNX1, single knockdown is sufficient to abolish transcriptional repression, suggesting that RUNX1 may play a key role in the assembly of a repressor complex at the SLC35A1 regulatory region. The existence of such a complex has already been proposed⁷⁸, and due to widespread occurrence of the RUNX1 motif adjacent to 39% of E2-repressed genes also possessing EREs throughout the genome, a RUNX1-repressor complex may play a similar role at other ER α binding regions. Future studies may pursue this further through co-immunoprecipitation (Co-IP) of RUNX1 and ER α , as well as other components of the hypothetical repressor complex.

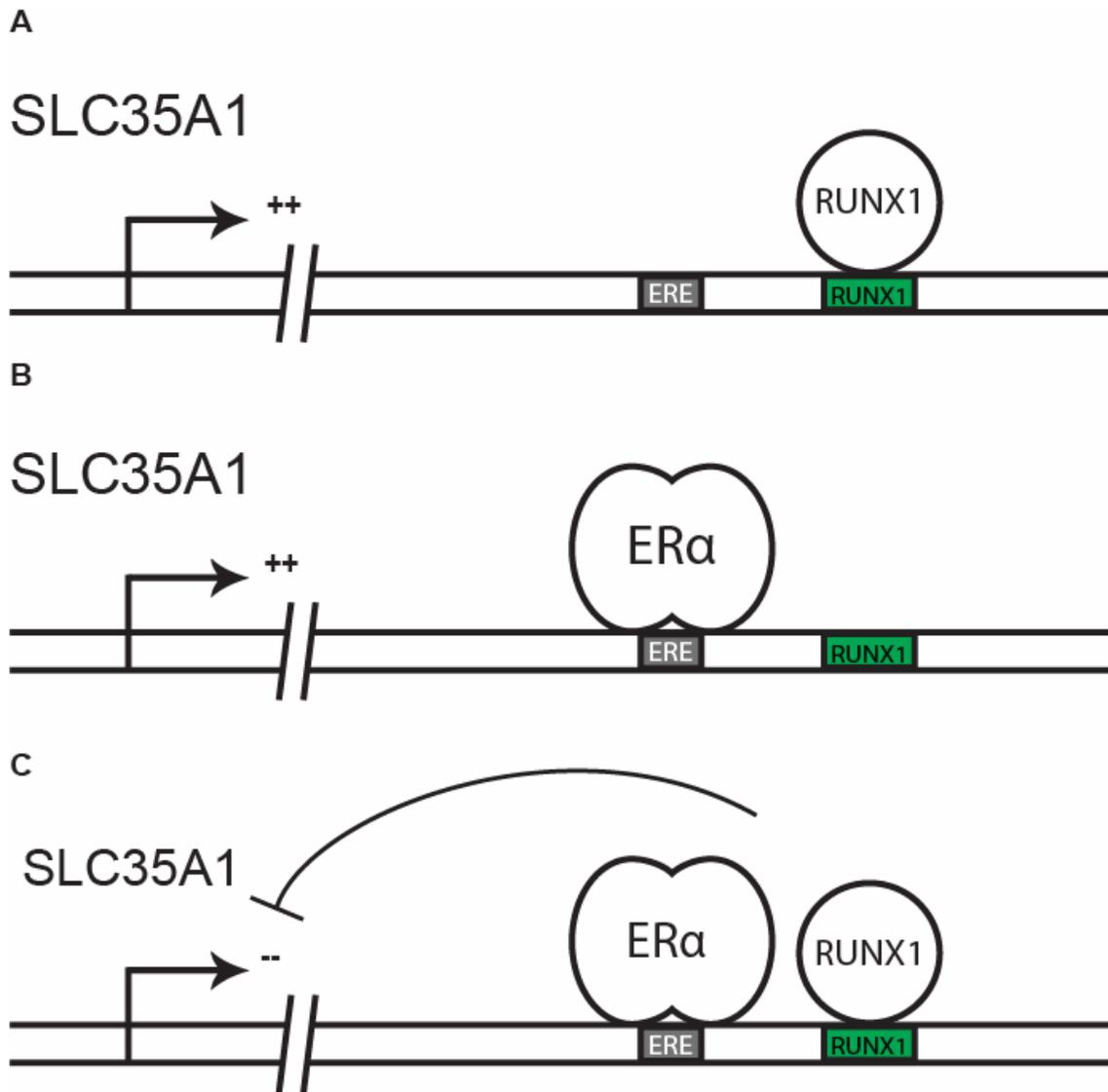


Figure 11 - RUNX1 Co-recruitment is Necessary for ER α -Mediated Transcriptional Repression of SLC35A1

A In absence of ER α signaling (E2-), RUNX1 (CBF) binds, but does not affect transcription.

B In the absence of RUNX1 (RUNX1 knockdown) ER α binds, but is unable to affect SLC35A1 expression.

C Co-recruitment of RUNX1 and ER α to the SLC35A1 ER α binding site leads to transcriptional repression of SLC35A1.

Given that the ER α -responsive assembly and recruitment of co-factors and transcriptional machinery at the pS2 promoter is not static, but rather a dynamic, cyclical, array of sequential assembly and disassembly, the role of RUNX1 may be less integral to the formation of an overall complex. Instead, RUNX1 may represent a key step within a similar ER α -responsive, dynamic assembly of co-repressors and other chromatin remodeling machinery, leading to the winding up of active chromatin, and perhaps the modification of histones through methylation or other covalent modifications (Figure 12). Elucidation of these and similar mechanisms necessitate extensive ChIP and re-ChIP of putative co-factors and other proteins. In addition to ChIP and re-ChIP experiments, identification and screening of other components via yeast-2-hybrid assays may also be pursued to screen multiple interacting partners in a high-throughput fashion.

Elucidation of the mechanisms behind RUNX1-ER α co-recruitment and interaction represents a major break from previous understanding of ER α function in breast cancer. Unlike preceding studies, we have demonstrated that target gene activation is not ER α 's sole function, and perhaps not even its major function. In this context, ER α functions as an effector of proliferative signals, repressing the expression of hundreds of genes possibly involved in regulating cell growth. Many of these repression-associated ER α target genes may even be tumor suppressors. Perhaps, if newer treatments are more focused on co-recruited factors like RUNX1, or used in combination with current adjuvant therapies, the mechanism behind ER α -RUNX1 transcriptional repression may ultimately lead to more targeted and effective breast cancer therapies with fewer side effects on the wide range of ER α target tissues.

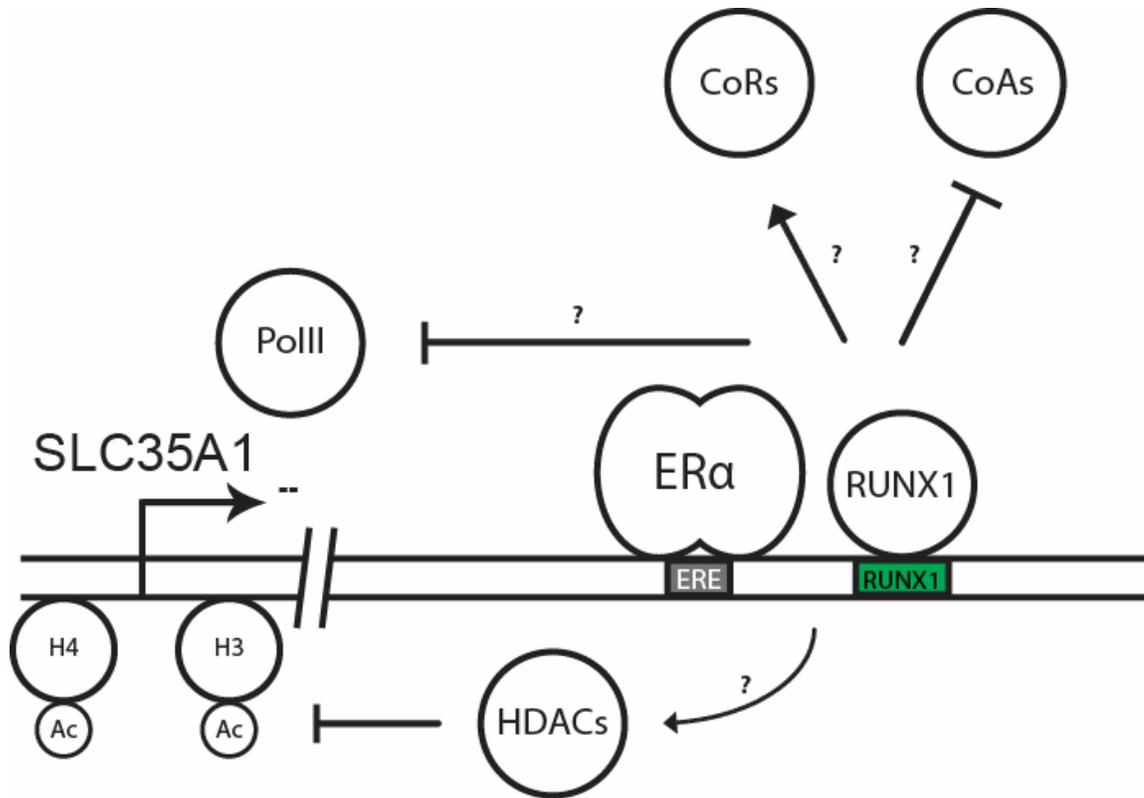


Figure 12 - RUNX1-ERα As Necessary Components of a Repressive Complex

Knockdown of RUNX1 leads to Repressive Incompetence of ERα, suggesting that RUNX1 is a key feature in the recruitment of co-repressors (CoRs), co-activators (CoAs), HDACs and deacetylation of histones H3 and H4, and the expulsion of polymerase (PolII). Further work is needed to validate these hypotheses.

Bibliography

1. Gruber, C. J., Tschugguel, W., Schneeberger, C. & Huber, J. C. Production and actions of estrogens. *N Engl J Med* **346**, 340-352,5, (2002)
2. McDonnell, D. P. & Norris, J. D. Connections and regulation of the human estrogen receptor. *Science* **296**, 1642-1644,5573, (2002)
3. Sudhir, K. & Komesaroff, P. A. Clinical review 110: Cardiovascular actions of estrogens in men. *J Clin Endocrinol Metab* **84**, 3411-3415,10, (1999)
4. Couse, J. F. & Korach, K. S. Estrogen receptor null mice: what have we learned and where will they lead us? *Endocr Rev* **20**, 358-417,3, (1999)
5. Smith, E. P. *et al.* Estrogen resistance caused by a mutation in the estrogen-receptor gene in a man. *N Engl J Med* **331**, 1056-1061,16, (1994)
6. Deroo, B. J. & Korach, K. S. Estrogen receptors and human disease. *J Clin Invest* **116**, 561-570,3, (2006)
7. A.C.S. Cancer Facts & Figures, 2009. *Atlanta: American Cancer Society*, (2009)
8. Barron, J. J., Quimbo, R., Nikam, P. T. & Amonkar, M. M. Assessing the economic burden of breast cancer in a US managed care population. *Breast Cancer Res Treat* **109**, 367-377,2, (2008)
9. Anderson, B. O. *et al.* Breast cancer in limited-resource countries: an overview of the Breast Health Global Initiative 2005 guidelines. *Breast J* **12 Suppl 1**, S3-15, (2006)
10. Gerber, B., Muller, H., Reimer, T., Krause, A. & Friese, K. Nutrition and lifestyle factors on the risk of developing breast cancer. *Breast Cancer Res Treat* **79**, 265-276,2, (2003)
11. Parkin, D. M., Bray, F. I. & Devesa, S. S. Cancer burden in the year 2000. The global picture. *Eur J Cancer* **37 Suppl 8**, S4-66, (2001)
12. Yabroff, K. R., Warren, J. L. & Brown, M. L. Costs of cancer care in the USA: a descriptive review. *Nat Clin Pract Oncol* **4**, 643-656,11, (2007)
13. Beatson, G. T. On The Treatment of Inoperable Cases of Carcinoma of the Mamma: Suggestions for a New Method of Treatment, with Illustrative Cases. *Lancet* **2**, 104-107, (1896)
14. Jensen, E. V. & Jordan, V. C. The estrogen receptor: a model for molecular medicine. *Clin Cancer Res* **9**, 1980-1989,6, (2003)

15. Ariazi, E. A., Ariazi, J. L., Cordera, F. & Jordan, V. C. Estrogen receptors as therapeutic targets in breast cancer. *Curr Top Med Chem* **6**, 181-202,3, (2006)
16. Harrell, J. C. *et al.* Estrogen receptor positive breast cancer metastasis: altered hormonal sensitivity and tumor aggressiveness in lymphatic vessels and lymph nodes. *Cancer Res* **66**, 9308-9315,18, (2006)
17. Yager, J. D. & Davidson, N. E. Estrogen carcinogenesis in breast cancer. *N Engl J Med* **354**, 270-282,3, (2006)
18. Chen, W. Y. Exogenous and endogenous hormones and breast cancer. *Best Pract Res Clin Endocrinol Metab* **22**, 573-585,4, (2008)
19. Parl, F. F., Dawling, S., Roodi, N. & Crooke, P. S. Estrogen metabolism and breast cancer: a risk model. *Ann N Y Acad Sci* **1155**, 68-75, (2009)
20. Riggs, B. L. & Hartmann, L. C. Selective estrogen-receptor modulators -- mechanisms of action and application to clinical practice. *N Engl J Med* **348**, 618-629,7, (2003)
21. D.U. Boehm *et al.* Quality of life and adjuvant tamoxifen treatment in breast cancer patients. *European Journal of Cancer Care* **9999**,9999, (2009)
22. Jensen, E. V. & DeSombre, E. R. Mechanism of action of the female sex hormones. *Annu Rev Biochem* **41**, 203-230, (1972)
23. Pearce, S. T. & Jordan, V. C. The biological role of estrogen receptors alpha and beta in cancer. *Crit Rev Oncol Hematol* **50**, 3-22,1, (2004)
24. Cianfrocca, M. & Goldstein, L. J. Prognostic and predictive factors in early-stage breast cancer. *Oncologist* **9**, 606-616,6, (2004)
25. Toft, D. & Gorski, J. A receptor molecule for estrogens: isolation from the rat uterus and preliminary characterization. *Proc Natl Acad Sci U S A* **55**, 1574-1581,6, (1966)
26. Kuiper, G. G., Enmark, E., Pelto-Huikko, M., Nilsson, S. & Gustafsson, J. A. Cloning of a novel receptor expressed in rat prostate and ovary. *Proc Natl Acad Sci U S A* **93**, 5925-5930,12, (1996)
27. Robinson-Rechavi, M., Escriva Garcia, H. & Laudet, V. The nuclear receptor superfamily. *J Cell Sci* **116**, 585-586,Pt 4, (2003)
28. Aranda, A. & Pascual, A. Nuclear hormone receptors and gene expression. *Physiological reviews* **81**, 1269-1304,3, (2001)
29. Vega, V. B. *et al.* Multi-platform genome-wide identification and modeling of functional human estrogen receptor binding sites. *Genome Biol* **7**, R82,9, (2006)

30. Gruber, C. J., Gruber, D. M., Gruber, I. M., Wieser, F. & Huber, J. C. Anatomy of the estrogen response element. *Trends Endocrinol Metab* **15**, 73-78,2, (2004)
31. Sathya, G. *et al.* Effects of multiple estrogen responsive elements, their spacing, and location on estrogen response of reporter genes. *Mol Endocrinol* **11**, 1994-2003,13, (1997)
32. Kulakosky, P. C., McCarty, M. A., Jernigan, S. C., Risinger, K. E. & Klinge, C. M. Response element sequence modulates estrogen receptor alpha and beta affinity and activity. *J Mol Endocrinol* **29**, 137-152,1, (2002)
33. Klinge, C. M. Estrogen receptor interaction with estrogen response elements. *Nucl. Acids Res.* **29**, 2905-2919,14, (2001)
34. Metivier, R. *et al.* Estrogen receptor-alpha directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter. *Cell* **115**, 751-763,6, (2003)
35. Lin, C. Y. *et al.* Discovery of estrogen receptor target genes and response elements in breast tumor cells. *Genome biology* **5**, -,9, (2004)
36. Lin, C. Y. *et al.* Whole-Genome Cartography of Estrogen Receptor alpha Binding Sites. *PLoS Genet* **3**, e87,6, (2007)
37. Carroll, J. S. *et al.* Genome-wide analysis of estrogen receptor binding sites. *Nat Genet* **38**, 1289-1297,11, (2006)
38. Klinge, C. M. Estrogen receptor interaction with co-activators and co-repressors. *Steroids* **65**, 227-251,5, (2000)
39. Dobrzycka, K., Townson, S., Jiang, S. & Oesterreich, S. Estrogen receptor corepressors -- a role in human breast cancer? *Endocr Relat Cancer* **10**, 517-536,4, (2003)
40. Shang, Y. & Brown, M. Molecular Determinants for the Tissue Specificity of SERMs. *Science* **295**, 2465-2468,5564, (2002)
41. Huang, H.-J., Norris, J. D. & McDonnell, D. P. Identification of a Negative Regulatory Surface within Estrogen Receptor {alpha} Provides Evidence in Support of a Role for Corepressors in Regulating Cellular Responses to Agonists and Antagonists. *Mol Endocrinol* **16**, 1778-1792,8, (2002)
42. Denslow, S. A. & Wade, P. A. The human Mi-2/NuRD complex and gene regulation. *Oncogene* **26**, 5433-5438,37, (2007)
43. Merrell, K. W. *Mechanisms of Estrogen Receptor Alpha Mediated Transcriptional Suppression* Master of Science thesis, Brigham Young University, (2007).

44. Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res* **14**, 1188-1190,6, (2004)
45. Schneider, T. D. & Stephens, R. M. Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res* **18**, 6097-6100,20, (1990)
46. Wingender, E., Dietze, P., Karas, H. & Knuppel, R. TRANSFAC: a database on transcription factors and their DNA binding sites. *Nucleic Acids Res* **24**, 238-241,1, (1996)
47. Kel, A. E. *et al.* MATCH: A tool for searching transcription factor binding sites in DNA sequences. *Nucleic Acids Res* **31**, 3576-3579,13, (2003)
48. Blackwood, E. M. & Kadonaga, J. T. Going the distance: A current view of enhancer action. *Science* **281**, 60,5373, (1998)
49. Cook, P. R. Nongenic transcription, gene regulation and action at a distance. *J Cell Sci* **116**, 4483-4491,22, (2003)
50. Brooks, S. C., Locke, E. R. & Soule, H. D. Estrogen receptor in a human cell line (MCF-7) from breast carcinoma. *J Biol Chem* **248**, 6251-6253,17, (1973)
51. Lam, J. S. *et al.* Prostate stem cell antigen is overexpressed in prostate cancer metastases. *Clin Cancer Res* **11**, 2591-2596,7, (2005)
52. Fleischmann, A. *et al.* Distinct subcellular expression patterns of neutral endopeptidase (CD10) in prostate cancer predict diverging clinical courses in surgically treated patients. *Clin Cancer Res* **14**, 7838-7842,23, (2008)
53. Eckhardt, M., Muhlenhoff, M., Bethe, A. & Gerardy-Schahn, R. Expression cloning of the Golgi CMP-sialic acid transporter. *Proc Natl Acad Sci U S A* **93**, 7572-7576,15, (1996)
54. Ishida, N., Ito, M., Yoshioka, S., Sun-Wada, G. H. & Kawakita, M. Functional expression of human golgi CMP-sialic acid transporter in the Golgi complex of a transporter-deficient Chinese hamster ovary cell mutant. *J Biochem* **124**, 171-178,1, (1998)
55. Nathrath, M. H. *et al.* Two novel tumor suppressor gene loci on chromosome 6q and 15q in human osteosarcoma identified through comparative study of allelic imbalances in mouse and man. *Oncogene* **21**, 5975-5980,38, (2002)
56. Nakamura, T. *et al.* Estrogen prevents bone loss via estrogen receptor alpha and induction of Fas ligand in osteoclasts. *Cell* **130**, 811-823,5, (2007)
57. Zallone, A. Direct and indirect estrogen actions on osteoblasts and osteoclasts. *Ann N Y Acad Sci* **1068**, 173-179, (2006)

58. Kang, Y. *et al.* A multigenic program mediating breast cancer metastasis to bone. *Cancer Cell* **3**, 537-549,6, (2003)
59. Wei, B. *et al.* Bone metastasis is strongly associated with estrogen receptor-positive/progesterone receptor-negative breast carcinomas. *Hum Pathol* **39**, 1809-1815,12, (2008)
60. Cherlet, T. & Murphy, L. C. Estrogen receptors inhibit Smad3 transcriptional activity through Ap-1 transcription factors. *Mol Cell Biochem* **306**, 33-42,1-2, (2007)
61. Safe, S. & Kim, K. Non-classical genomic estrogen receptor (ER)/specificity protein and ER/activating protein-1 signaling pathways. *J Mol Endocrinol* **41**, 263-275,5, (2008)
62. Teyssier, C., Belguise, K., Galtier, F. & Chalbos, D. Characterization of the physical interaction between estrogen receptor alpha and JUN proteins. *J Biol Chem* **276**, 36361-36369,39, (2001)
63. Feldman, I., Feldman, G. M., Mobarak, C., Dunkelberg, J. C. & Leslie, K. K. Identification of proteins within the nuclear factor-kappa B transcriptional complex including estrogen receptor-alpha. *Am J Obstet Gynecol* **196**, 394 e391-311; discussion 394 e311-393,4, (2007)
64. Kalaitzidis, D. & Gilmore, T. D. Transcription factor cross-talk: the estrogen receptor and NF-kappaB. *Trends Endocrinol Metab* **16**, 46-52,2, (2005)
65. Carroll, J. S. *et al.* Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. *Cell* **122**, 33-43,1, (2005)
66. Itoh, Y. *et al.* 17beta-estradiol activates IL-1alpha gene expression in rheumatoid fibroblast-like synovial cells through estrogen receptor alpha (ERalpha) and augmentation of transcriptional activity of Sp1 by dissociating histone deacetylase 2 from ERalpha. *J Immunol* **178**, 3059-3066,5, (2007)
67. Safe, S. Transcriptional activation of genes by 17 beta-estradiol through estrogen receptor-Sp1 interactions. *Vitam Horm* **62**, 231-252, (2001)
68. de Bruijn, M. F. & Speck, N. A. Core-binding factors in hematopoiesis and immune function. *Oncogene* **23**, 4238-4248,24, (2004)
69. Hart, S. M. & Foroni, L. Core binding factor genes and human leukemia. *Haematologica* **87**, 1307-1323,12, (2002)
70. Michaud, J., Scott, H. S. & Escher, R. AML1 interconnected pathways of leukemogenesis. *Cancer Invest* **21**, 105-136,1, (2003)

71. Kouros-Mehr, H., Slorach, E. M., Sternlicht, M. D. & Werb, Z. GATA-3 maintains the differentiation of the luminal cell fate in the mammary gland. *Cell* **127**, 1041-1055,5, (2006)
72. Ramaswamy, S., Ross, K. N., Lander, E. S. & Golub, T. R. A molecular signature of metastasis in primary solid tumors. *Nat Genet* **33**, 49-54,1, (2003)
73. Dong, S. & Kiyama, R. Characterisation of oestrogenic activity of ginsenosides in MCF-7 cells using a customised DNA microarray. *Food Chemistry* **113**, 672-678,2, (2009)
74. Inoue, A. *et al.* Comparative profiling of the gene expression for estrogen responsiveness in cultured human cell lines. *Toxicol In Vitro* **21**, 741-752,4, (2007)
75. Terasaka, S. *et al.* Using a customized DNA microarray for expression profiling of the estrogen-responsive genes to evaluate estrogen activity among natural estrogens and industrial chemicals. *Environ Health Perspect* **112**, 773-781,7, (2004)
76. Dairkee, S. H. *et al.* A molecular 'signature' of primary breast cancer cultures; patterns resembling tumor tissue. *BMC Genomics* **5**, 47,1, (2004)
77. Ma, Y. *et al.* Population-based molecular prognosis of breast cancer by transcriptional profiling. *Clin Cancer Res* **13**, 2014-2022,7, (2007)
78. Stossi, F., Likhite, V. S., Katzenellenbogen, J. A. & Katzenellenbogen, B. S. Estrogen-occupied estrogen receptor represses cyclin G2 gene expression and recruits a repressor complex at the cyclin G2 promoter. *J Biol Chem* **281**, 16272-16278,24, (2006)