Estrogen receptor alpha mutations in breast cancer cells cause gene expression changes through constant activity and through secondary effects

Spencer Arnesen[1,2], Zannel Blanchard[1,2], Michelle M Williams[3], Kristofer C Berrett[1,2], Zheqi Li[4,5], Steffi Oesterreich[4,5], Jennifer K Richer[3], and Jason Gertz[1,2,*]

[1]Department of Oncological Sciences, University of Utah, Salt Lake City, Utah, USA
[2]Huntsman Cancer Institute, University of Utah, Salt Lake City, Utah, USA
[3]Department of Pathology, Anschutz Medical Campus, University of Colorado, Aurora, Colorado, USA
[4]Women's Cancer Research Center, University of Pittsburgh Medical Center (UPMC) Hillman Cancer Center (HCC), Magee-Womens Research Institute, Pittsburgh, Pennsylvania, USA
[5]Department of Pharmacology & Chemical Biology, University of Pittsburgh, Pittsburgh, Pennsylvania, USA
*Corresponding author: jay.gertz@hci.utah.edu

Abstract

While breast cancer patients with tumors that express estrogen receptor α (ER) generally respond well to hormone therapies that block ER's actions, a significant number relapse. Approximately 30% of these recurrences harbor activating mutations in ER's ligand binding domain (LBD). ER mutations have been shown to confer ligand-independent function to ER; however, much is still unclear regarding the effect of mutant ER beyond its estrogen independence. To investigate mutant ER's molecular effects, we developed multiple isogenic ER mutant cell lines for the most common ER LBD mutations, Y537S and D538G. We found that these mutations induce differential expression of thousands of genes, the majority of which are mutant allele-specific and are not observed upon estrogen treatment of wildtype cells. The mutant-specific genes show consistent differential expression across ER mutant lines developed in other laboratories. The observed gene expression changes cannot be explained by constitutive ER activity alone, as wildtype cells with long-term estrogen exposure only exhibit some of these transcriptional changes, suggesting that mutant ER causes novel regulatory effects that are not simply due to constant activity. While ER mutations have minor effects on ER genomic binding, with the exception of ligand independence, we observed substantial differences in chromatin accessibility due to ER mutations. Mutant ER is bound to approximately a quarter of mutant-enriched accessible regions that are enriched for other DNA binding factors including FOXA1, CTCF, and OCT1. Our findings indicate that mutant ER causes several consistent effects on gene expression both indirectly and through constant activity.

Introduction

Estrogen receptor α (ER) is a ligand-inducible nuclear hormone receptor that binds estrogens. ER is expressed in roughly 70% of breast cancers and plays a key role in the development and progression of these tumors (1). Due to ER's role in the growth of ER+ breast cancer, these tumors are generally treated with endocrine therapies, including aromatase inhibitors (AIs), selective estrogen receptor modulators (SERMs), and selective estrogen receptor degraders (SERDs). While hormone therapies have been effective in preventing recurrence, approximately 20% of these cancers develop resistance to hormone therapies and will eventually recur (2, 3). Recent genome sequencing efforts have established mutations in the ligand binding domain (LBD) of ER as a common path to hormone therapy resistance that are found in approximately 30% of metastatic ER-positive tumors (4-6).

ER mutations are rarely found in primary ER+ breast cancers but have been observed often in metastatic tumors and in circulating cell free DNA particularly after treatment with AIs (4-10). ER mutations occur almost exclusively in the receptor's LBD, which is responsible for binding to estrogens and recruiting cofactors, with the majority of mutations occurring at residues Y537 and D538. Crystal structures of the mutant LBD revealed that mutant ER adopts an active conformation in the absence of estrogen binding and several studies have observed ligand-independent transcriptional regulation, growth, and proliferation of ER mutant breast cancer cells both *in vitro* and *in vivo* (4-6, 10-16). Large-scale alterations in transcription have also been observed in cells expressing mutant ER (17-19). RNA-seq experiments identified hundreds of ER target genes that are differentially expressed in mutant ER cell lines in the absence of estrogens. A number of genes were also observed as potential "novel" targets of mutant ER. These genes are not regulated by wildtype (WT) ER, but are differentially regulated in mutant ER cells compared to WT regardless of estrogen treatment. Mutant ER also exhibits ligand-independent activity in regard to DNA binding. Recent ChIP-seq experiments showed that mutant ER is capable

of binding at many ER binding sites (ERBS) in the absence of estrogens (12, 18, 20). Additionally, mutant ER may differentially bind at certain sites compared to WT ER (18).

While much has been elucidated regarding the molecular consequences of ER mutations in breast cancer, what drives these effects is less clear. Multiple studies have reported that ER mutations alter the expression of a "novel" or mutant-specific set of genes; however, the consistency of these reported gene expression changes has not been evaluated. Establishing which gene expression changes are found across multiple mutant ER models will assist in identifying molecular or cellular pathways consistently altered when ER is mutated, and reveal potential factors driving mutant ER's transcriptional impact. In addition, how these mutant-specific transcriptional changes are achieved remains unclear. One model is that mutant-specific gene expression changes are caused by mutant ER's constant activity. The long-term constitutive activation achieved by mutant ER likely results in the regulation of genes which are not typically regulated by WT ER with short-term estrogen exposure. These long-term constitutively active ER effects could account for many of the mutant-specific ER "novel" gene expression changes. A second model is that mutations confer neomorphic properties to ER, which could induce new ER genomic binding sites or alter its ability to regulate gene expression. A third model is that ER mutations cause indirect changes in gene expression by changing the activity of other transcription factors and/or epigenetic regulators which in turn impact gene expression. It is unclear which of these models contribute to the novel gene expression observed in breast cancer cells harboring ER mutations.

In this study, we set out to answer these outstanding questions concerning ER mutations in breast cancer cells. Using a unique CRISPR/Cas9 strategy that allowed us to introduce the Y537S and D538G mutations and an epitope tag at ER's endogenous locus, we examined the effects of mutant ER on gene expression and found thousands of genes consistently impacted by ER mutations across models from different studies. By analyzing WT cells exposed to estrogens for several weeks, we discovered that approximately half of the mutant ER-specific gene expression changes can be attributed to constant ER activity. Mutant-specific gene expression changes could be partially explained by alterations in ER genomic binding; however, chromatin accessibility alterations not involving ER were more commonly found nearby mutant-specific genes. Motif analysis of differentially accessible regions helped to identify transcription factors bound at these loci. Together, our results suggest that ER mutations consistently impact the expression of thousands of novel genes partially through constant ER activity and partially by indirectly altering regulatory elements through additional transcription factors, including OCT1, FOXA1, and CTCF.

Results

Establishment of endogenously expressed FLAG tagged ER mutant models

Investigation of mutant ER's molecular and phenotypic consequences relies on the development of models that faithfully recapitulate mutant ER in tumors. Many previous studies investigating mutant ER's molecular and phenotypic consequences used ectopic expression of mutant ER (4-6, 10, 14-16, 18, 21). In order to characterize endogenous ER mutations, we created multiple isogenic clonal lines that heterozygously express FLAG tagged mutant or WT (control) ER from the endogenous locus (22) (Supplemental Fig. S1a). We developed two clones each for WT and the two most common ER LBD mutations (Y537S and D538G) in both T-47D and MCF-7 breast cancer cell lines (Supplemental Fig. S1b). Engineered lines included a FLAG epitope tag at the C-terminus to allow for downstream analyses. The heterozygous expression of FLAG tagged mutant or WT ER and the availability of multiple clones per genotype provided a robust system in which to investigate mutant ER's molecular effects.

Previous studies have shown estrogen-independent expression of known ER target genes in cells expressing mutant ER (5, 12, 14, 17). To determine whether our mutant clones demonstrated similar gene regulatory effects, we measured the expression of known ER target genes using quantitative PCR in cells treated with 17β-estradiol (E2) or vehicle (DMSO) (Supplemental Fig.S1c). In WT clones, ER regulated genes showed significant changes in expression upon E2 treatment, as expected. In mutant ER clones, these genes were significantly differentially expressed compared to WT controls even with no E2 treatment, while the addition of E2 in some cases further increased the expression of these genes. These results indicate that our mutant clones have ligand-independent function similar to that observed in previous studies (4, 5, 12, 14, 17).

Expression levels of thousands of genes are consistently affected by ER mutations

To investigate whether mutant ER exhibits ligand-independent gene regulation on a genome-wide scale, we performed RNA-seq on WT and mutant clones grown in hormone depleted media and treated with E2 or DMSO for 8 hours. Principal component analysis (PCA) of RNA-seq results indicated that T-47D mutant clones, regardless of E2 treatment, differ distinctly from WT clones in their gene expression profiles (Fig. 1a). Further analysis of T-47D RNA-seq data identified a set of genes that are differentially regulated (adjusted p-value < 0.05) in the WT clones with the addition of E2 (E2-regulated genes). More than half (472 of 771) of the E2-regulated genes in the WT clones were similarly up- or down-regulated in clones from one or both mutants independently of E2 (ligand-independent genes) (Fig. 1b, Supplemental Table S1). A relatively small number of E2-regulated genes were differentially expressed in the opposite direction in ER mutant cells (85 in D538G clones and 32 in Y537S clones). Known E2-regulated genes including *CISH*, *SUSD3*, *KCNK5*, and *PGR* were included in the ligand-independent gene set (example of *CISH* in Fig. 1c). MCF-7 mutants also distinctly clustered apart from MCF-7 WT cells based on RNA-seq data as observed via PCA (Supplemental Fig. S2a). MCF-7 mutant clones exhibited similar ligand-independent effects for both the Y537S and D538G mutations. In these clones, 30% of E2-regulated genes (85 of 289) were regulated in a ligand-independent fashion in one or both mutants (Supplemental Fig. S2b, example of *CISH* in Fig. S2c). These results support previous data showing the ligand-independent function of mutant ER.

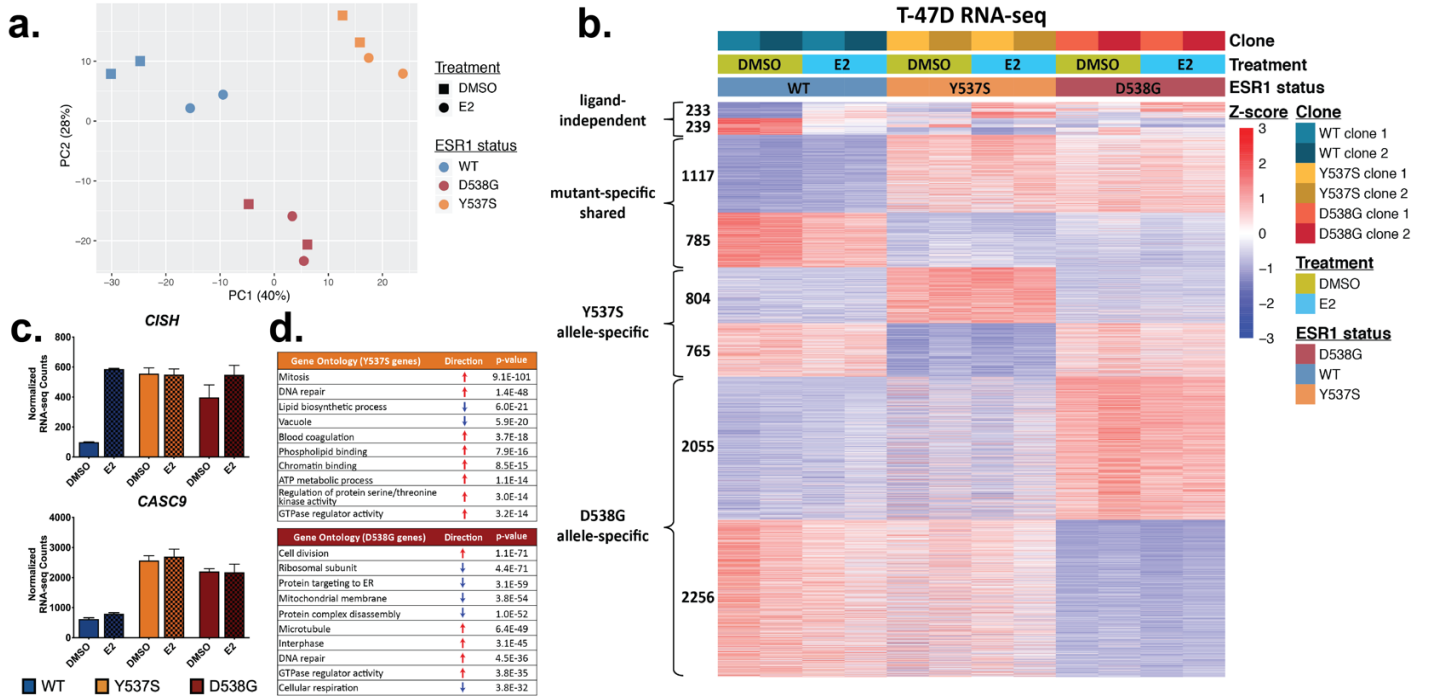
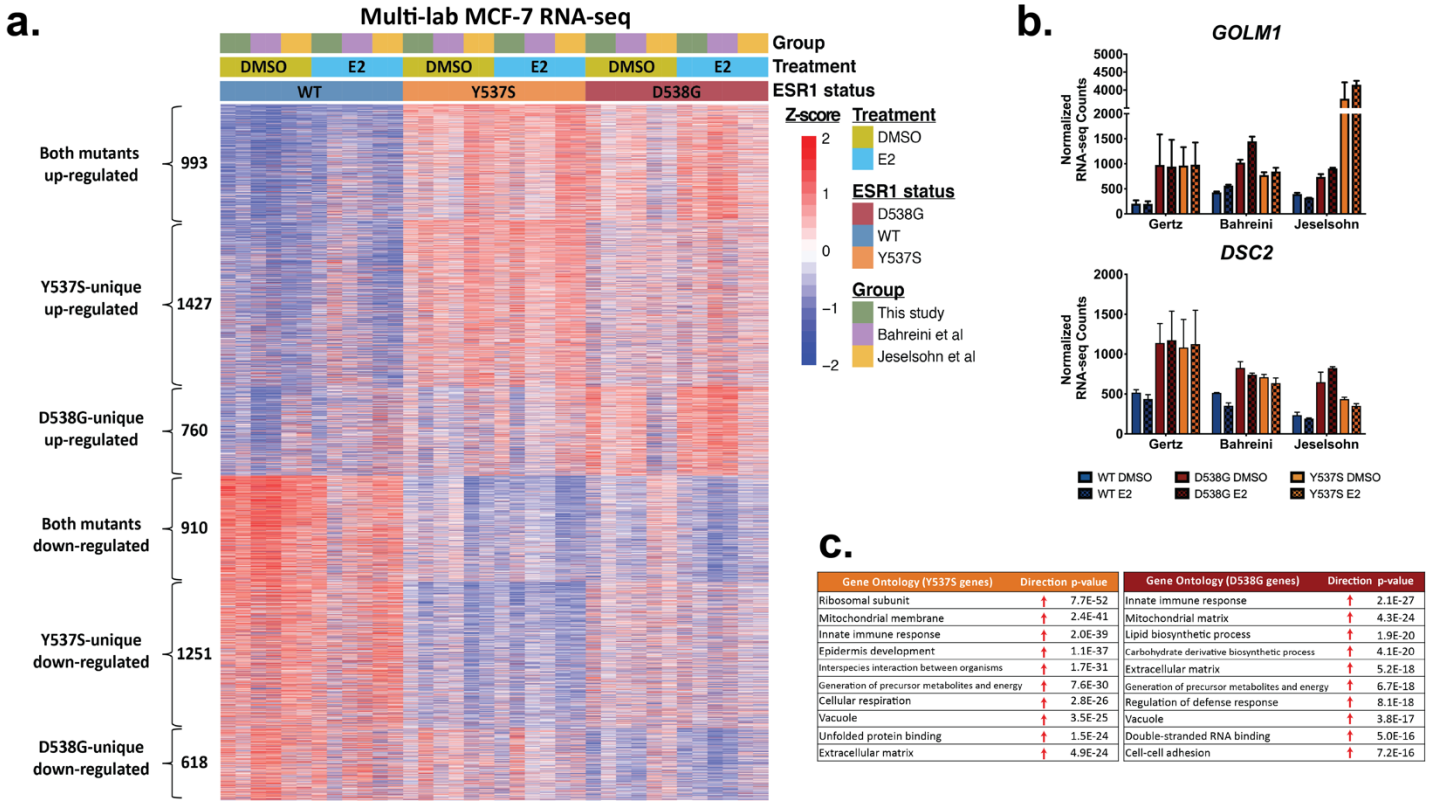


Figure 1. Y537S and D538G ER mutations exhibit a mutation-specific gene expression profile. (a) Principal component analysis of RNA-seq data displays the relationships between WT (blue), Y537S (yellow), and D538G (red) T-47D clones. (b) Heatmap shows gene expression levels for ligand-independent, mutant-specific shared, and allele-specific genes. (c) Bar graphs show examples of ligand-independent (*CISH*) and mutant-specific (*CASC9*) gene expression changes. Graphs show average normalized RNA-seq counts. Error bars show standard deviation for two clones for each genotype and each treatment. (d) Enriched gene ontology terms for Y537S- and D538G-specific differentially regulated genes with Fisher's exact test p-values are shown.

In addition to ligand-independent gene regulation, ER mutant cells also exhibit differential expression of a large number of genes that are not differentially regulated upon an 8-hour E2 treatment in WT ER cells. In T-47D clones, over 1900 genes were up- (1117) or down-regulated (785) in both the Y537S and the D538G mutations (shared mutant-regulated genes) (Fig. 1b, example of *CASC9* in Fig. 1c). The majority of mutant-regulated genes were allele-specific with 1569 Y537S-specific and 4311 D538G-specific up- or down-regulated genes (Fig. 1b). Similar results were seen in our MCF-7 mutant clones, with over a thousand shared mutant-specific genes and thousands of genes regulated in an allele-specific fashion (Supplemental Fig. S2b, examples in Supplemental Fig. S2c). Overall, mutant-specific genes made up the vast majority of differentially regulated genes in our mutant clones, indicating that a major result of ER mutation, in addition to ligand-independent regulation of WT ER target genes, is the differential regulation of a novel gene expression program. Of these mutant-specific genes, nearly 30% of D538G regulated genes (up-regulated genes: p-value = $5.1e^{-55}$; down-regulated genes: p-value = $1.8e^{-59}$) and 15% of Y537S differentially regulated genes (up-regulated genes: p-value = $2.3e^{-28}$; down-regulated genes: p-value = $7.04e^{-17}$) were shared across both cell lines (T-47D and MCF-7, examples in Supplemental Fig.

S2d) suggesting that while some mutant-specific genes are consistently observed in multiple cell lines, many may be cell line-specific.

To investigate enriched pathways in mutation-regulated genes, we performed gene ontology and pathway analysis and found that mutant-specific genes in T-47D cells were highly enriched for positive roles in cell cycle and cell division or protein synthesis and processing (Fig. 1d). Mutant-specific genes in MCF-7 cells were enriched for genes involved in increased cellular migration and motility but decreased cell division (Supplemental Fig. S2e). Using reverse phase protein array (RPPA) analysis, we validated that a number of gene expression changes that occur in our T-47D mutant clones also occur at the protein level (Supplemental Fig. S3a). Many proteins that were significantly different between mutant and WT cells were also differentially regulated at the transcriptional level in our mutant clones according to our RNA-seq results (Supplemental Fig. S3b). These proteins included cyclins and cyclin-dependent kinases as well as DNA-repair proteins which are key to successful cell-cycle progression. In addition to cell cycle-related genes, we also validated increased expression of *IGF1R* at the protein level and decreased expression of *CDH1* which codes for E-cadherin (Supplemental Fig. S3b). These proteins play important roles in cell growth and cell motility, and IGF1R has been shown to be activated by ER mutations (23, 24). These results suggest that mutant ER's expanded transcriptional impact may promote cell division, growth, and migration, phenotypes that have been previously observed in ER mutant breast cancer cells (4-6, 10, 12, 14). This expanded transcriptional impact may also affect other properties such as anchorage independent growth and metastatic outgrowth as observed in Williams et al. (submitted).



Gene Ontology (Y537S genes)	Direction	p-value
Ribosomal subunit	↑	7.7E-52
Mitochondrial membrane	↑	2.4E-41
Innate immune response	↑	2.0E-39
Epidermis development	↑	1.1E-37
Interspecies interaction between organisms	↑	1.7E-31
Generation of precursor metabolites and energy	↑	7.6E-30
Cellular respiration	↑	2.8E-26
Vacuole	↑	3.5E-25
Unfolded protein binding	↑	1.5E-24
Extracellular matrix	↑	4.9E-24

Gene Ontology (D538G genes)	Direction	p-value
Innate immune response	↑	2.1E-27
Mitochondrial matrix	↑	4.3E-24
Lipid biosynthetic process	↑	1.9E-20
Carbohydrate derivative biosynthetic process	↑	4.1E-20
Extracellular matrix	↑	5.2E-18
Generation of precursor metabolites and energy	↑	6.7E-18
Regulation of defense response	↑	8.1E-18
Vacuole	↑	3.8E-17
Double-stranded RNA binding	↑	5.0E-16
Cell-cell adhesion	↑	7.2E-16

Figure 2. Gene expression analysis of RNA-seq data from three independent groups reveals consistent ER mutant-specific gene expression patterns. (a) Heatmap shows levels of consistent mutant-specific genes. Heatmap columns indicate group of origin, sample type, and treatment for each sample. (b) Examples of genes exhibiting consistent mutant-specific gene expression are shown. Y-axis shows normalized RNA-seq counts from DESeq2; x-axis groups gene expression by study of origin and *ESR1* mutational status. (c) Significantly enriched gene ontology terms are displayed for Y537S-specific and D538G-specific gene sets from the multi-lab comparison.

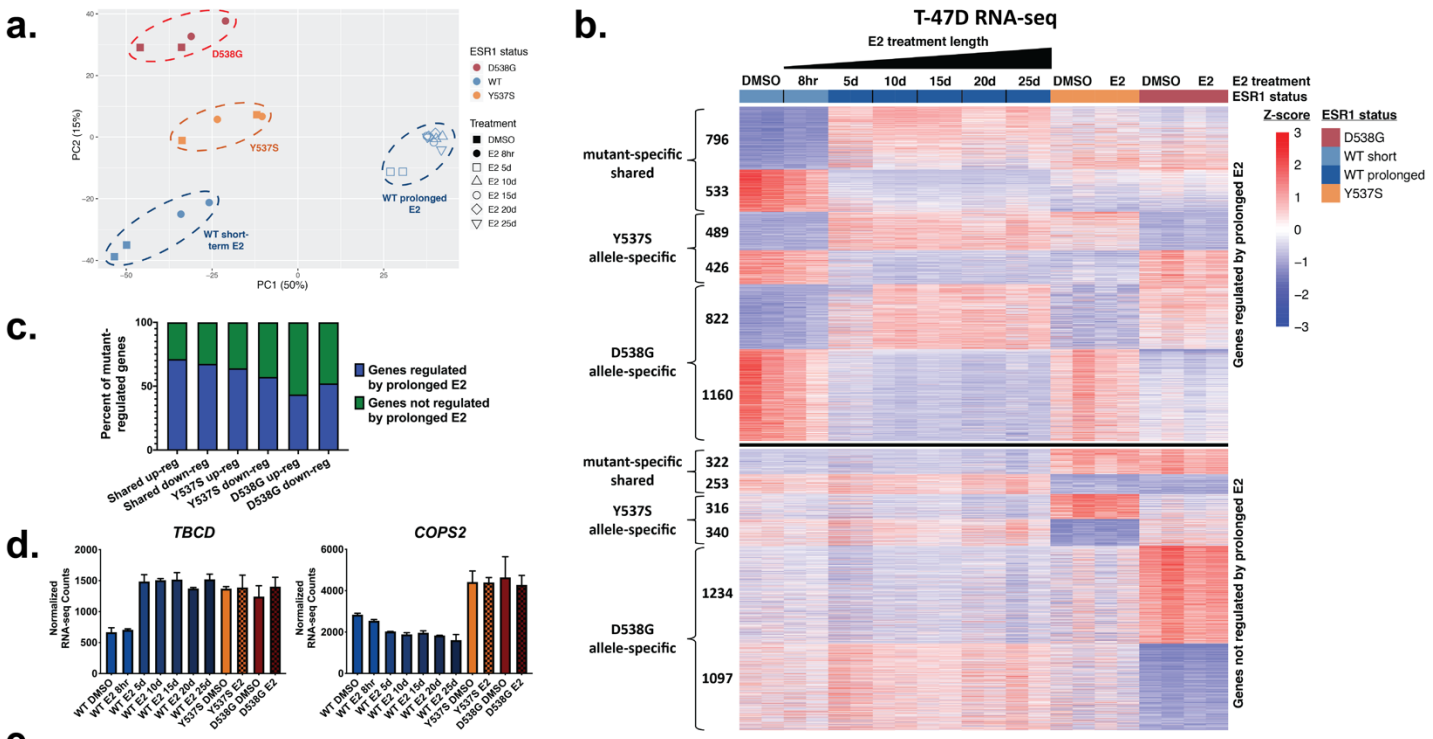
Previous studies have described T-47D and MCF-7 ER mutant lines made using CRISPR or lentiviral techniques and have reported similar E2-independent and mutant-specific gene regulation (12, 17-19). In an effort to define which genes are consistently regulated by ER mutations, we compared RNA-seq data from the clones described above with RNA-seq data collected from ER mutant lines described in two other studies (17, 18). A PCA plot of these multi-lab data from MCF-7 lines clustered them by study than by ER mutational status (Supplemental Fig. S4a). This is likely due to multiple factors including differences in mutation introduction techniques, duration of hormone depletion, E2 induction time, and sequencing library generation protocol. However, within PCA study clusters, clones segregated by mutational status and E2 treatment. Multivariate analysis of these data to account

for variation between studies identified genes that were consistently differentially regulated by ER mutations across all datasets. Over 1900 genes were up- (993) or down-regulated (910) by both mutations across all MCF-7 lines (Fig. 2a, examples in Fig. 2b, Supplemental Table S2). Thousands of genes were specific to each mutation (1378 D538G-specific genes and 2678 Y537S-specific genes). Genes that were consistently differentially expressed in mutant lines from all three studies were enriched for a variety of gene ontology terms shared between both mutations including innate immune response, mitochondrial matrix, and extracellular matrix. Allele specific terms were also observed including cellular respiration and ribosomal subunit for the Y537S mutation and cell-cell adhesion and lipid biosynthesis for the D538G mutation (Fig. 2c). Although these terms differ somewhat from those identified in individual studies' lines, they represent underlying changes that consistently occur in response to the expression of ER mutations and therefore may provide direction for future studies regarding mutant ER action and effects.

We used WT and mutant T-47D RNA-seq data from the same three studies to perform an identical multivariate analysis as above with the exception that one study lacked the D538G mutation in T-47D cells. The T-47D multivariate analysis yielded similar results to the MCF-7 analysis with over 1700 up- or down-regulated genes shared between both mutations and thousands of additional genes differentially regulated by each mutant allele uniquely (Supplemental Fig. S4b, examples in Supplemental S4c). These genes were heavily enriched for cell cycle-related genes (Supplemental Fig. S4d). Overall, the comparison of multiple lines from multiple individual groups provides evidence of consistent regulation of novel genes by ER mutations.

Constant ER activity can explain some of mutant ER's impact on gene expression

The constant activity of ER brought about by LBD mutations could explain some of the mutant-specific gene expression changes observed in our RNA-seq analyses. Long-term constitutive ER activity may lead to the differential expression of genes that are in fact regulated by WT ER but do not change expression with short-term E2 treatment. To determine how much of mutant ER's gene regulatory changes can be attributed to mutant ER's constitutive activity, we treated T-47D and MCF-7 WT clones with E2 for 25 days. WT ER cells were collected at 5 day intervals and RNA was harvested for RNA-seq analysis. A PCA of RNA-seq data compared WT T-47D cells treated with long-term (5, 10, 15, 20, 25 days) E2 to WT ER short-term (0 and 8hr) E2 treatments and mutant T-47D cells. PC1, which accounted for 50% of the variance between samples, revealed a distinct separation between long-term E2 treated WT ER cells and both WT ER short-term E2 treatments and mutant ER cells. (Fig. 3a). PC2 grouped long-term E2-treatments apart from short-term E2 treatments and instead clustered them closer to Y537S and D538G mutant cells. Together, these observations indicate that while long-term ER activity may account for some of the differences between mutant and WT cells, it may not account for all mutant ER gene expression changes. Further analysis of RNA-seq data revealed that 54% of all mutant-regulated genes were similarly up- or down-regulated by long-term E2 treatment (adjusted p-value < 0.05) (Fig. 3b,c, example of *TBCD* in Fig. 3d). However, a considerable number of mutant ER regulated genes were not differentially regulated with long-term E2 treatment in WT cells, but remained specific to one or both mutations (example of *COPS2* in Fig. 3d). Similar results were observed in MCF-7 WT cells treated with long-term E2 with 38% of all mutant-specific genes being explained by constant ER activity (Supplemental Fig. S5a-d). These results suggest that while the constitutive activity of the mutant receptor may account for approximately half of the gene expression changes observed in mutant ER breast cancer cells, a large percentage of mutant-specific gene expression changes cannot be attributed to constant ER activity. These unexplained mutant-specific gene expression changes may instead result from additional properties of mutant ER which allow it to directly or indirectly alter the expression of these genes.



e.

Gene Ontology (Y537S genes regulated by constant E2)	Direction	p-value
Mitosis	↑	8.7E-94
DNA repair	↑	4.0E-44
Microtubule	↑	1.3E-36
G2/M transition of mitotic cell cycle	↑	1.9E-33
Kinase binding	↑	6.6E-26
G1/S transition of mitotic cell cycle	↑	5.2E-20
Cell morphogenesis involved in neuron differentiation	↑	6.6E-17
Phospholipid binding	↑	9.4E-17
Hemostasis	↑	3.8E-16
GTPase regulator activity	↑	4.6E-15

Gene Ontology (Y537S genes not regulated by constant E2)	Direction	p-value
Lipid biosynthetic process	↓	3.4E-13
Ligase activity	↓	1.8E-12
Regulation of MAPK cascade	↓	3.5E-11
Regulation of cell adhesion	↓	8.7E-11
Extracellular matrix organization	↓	2.2E-10
Negative regulation of protein metabolic process	↓	3.0E-10
Microtubule organizing center	↑	6.8E-10
Endoplasmic reticulum unfolded protein response	↓	8.2E-10
Mitosis	↑	9.1E-10
Regulation of cell morphogenesis	↓	1.9E-9

Gene Ontology (D538G genes regulated by constant E2)	Direction	p-value
Mitosis	↑	2.2E-75
Ribosomal subunit	↓	6.1E-73
Protein targeting to ER	↓	1.6E-60
Viral infectious cycle	↓	3.0E-53
Cytosolic part	↓	5.7E-48
Mitochondrial membrane	↓	1.0E-45
Cellular respiration	↓	4.7E-38
DNA repair	↑	3.4E-32
Cilium	↑	6.0E-28
GTPase regulator activity	↑	8.0E-25

Gene Ontology (D538G genes not regulated by constant E2)	Direction	p-value
Endosomal part	↑	7.7E-18
Ligase activity, forming carbon-nitrogen	↑	2.6E-16
Cell division	↑	1.4E-14
Negative regulation of protein metabolic process	↓	3.7E-13
Lipid biosynthetic process	↓	4.8E-13
Ubiquitin-protein ligase activity	↑	6.7E-13
Regulation of protein serine/threonine kinase activity	↓	1.2E-12
Protein binding transcription factor activity	↓	1.5E-11
Cell leading edge	↑	2.4E-11
Regulation of secretion	↓	3.3E-11

Figure 3. Constant activity of ER explains approximately half of T-47D mutant-specific genes. (a) Principle component analysis of RNA-seq data captures the relationships between T-47D ER mutant clones and WT clones with short- or long-term E2 treatment. Mutational status is indicated by color and E2 treatment length by shape of symbol. (b) Heatmap shows all mutant-specific genes separated by expression pattern of long-term E2-treatment in WT clones (top: genes regulated by long-term E2 treatment; bottom: genes not regulated by long-term E2 treatment). Mutational status is indicated by color bar and legend. E2 treatment time is indicated above the heatmap. (c) Percent of mutant-specific genes regulated by long-term ER activation in WT clones is shown. (d) Bar graphs show examples of genes that are regulated by long-term E2 treatment (*TBCD*) and that are differentially regulated by mutation only (*COPS2*). (e) Significantly enriched gene ontology terms are displayed for Y537S-specific and D538G-specific gene sets partitioned based on overlap with long-term E2 regulated genes.

ER mutations do not cause broad reprogramming of genomic binding

Mutant-specific gene expression changes could result from altered genomic binding by mutant ER compared to WT ER. To investigate this possibility, we performed ChIP-seq experiments on WT and mutant T-47D and MCF-7 clones, grown for 5 days in hormone depleted media followed by 1 hour E2 or DMSO treatments, using an antibody that recognizes the FLAG epitope tag. Results from these ChIP-seq experiments compared well with previously performed ER ChIP-seq experiments performed in another lab, with 68-73% binding site overlaps observed between samples of the same *ESR1* genotype (18). In WT T-47D clones, ER binding by WT ER is absent at thousands of genomic regions in the absence of E2 but is gained with the addition of E2 (Fig. 4a, example in Fig. 4b). ChIP-seq of mutant ER showed that for both mutations ER bound to the majority of WT ERBS independent of E2 treatment and mutant ER binding at these sites increased with the addition of E2. The majority of ER bound sites were identified in both mutants and in E2-treated WT clones (34920 constant sites); only ~10% (3837 sites) of ER-bound regions exhibited significant differential binding between WT and mutant with the large majority of these sites being differentially bound in Y537S mutant clones (Fig. 4c, example in Fig. 4d). In both constant and mutant-enriched and -depleted sites the ER binding motif (ERE) was the most highly enriched motif.

To determine how much of the mutant-specific expression differences that we observed could be explained by ER binding, we analyzed the distance between mutant-specific genes and constant, mutant-enriched, or mutant-depleted ERBS. Both Y537S and D538G mutant-specific up- and down-regulated genes are significantly closer

to constant ERBS than are all other genes (background). Using a previously described distance threshold of 100 kilobasepairs (kb) (25-27), we determined the number of mutant-specific genes that contained ERBS near their transcription start site (TSS). Of both Y537S and D538G mutant-specific genes, 55-80% contain nearby constant ERBS, which is 13-16% more than background genes. Mutant-specific up- and down-regulated genes are also significantly closer to mutant-enriched and -depleted sites respectively (Fig. 4e, Supplemental Table S3). 10% of Y537S up- and down-regulated genes contain at least one nearby Y537S-enriched or -depleted ERBS (Supplemental Table S3). Enriched and depleted ERBS were found to a much lesser extent in D538G mutant clones; less than 1% of D538G-specific up-regulated genes contain a nearby D538G-enriched ERBS and less than 2% of D538G-specific down-regulated genes contain a nearby D538G-depleted ERBS. These results indicate that while constant ligand-independent ER genomic binding likely affects the expression of a portion of mutant-specific genes, differential ER binding, particularly in D538G mutant cells, may only have a subtle effect on gene expression.

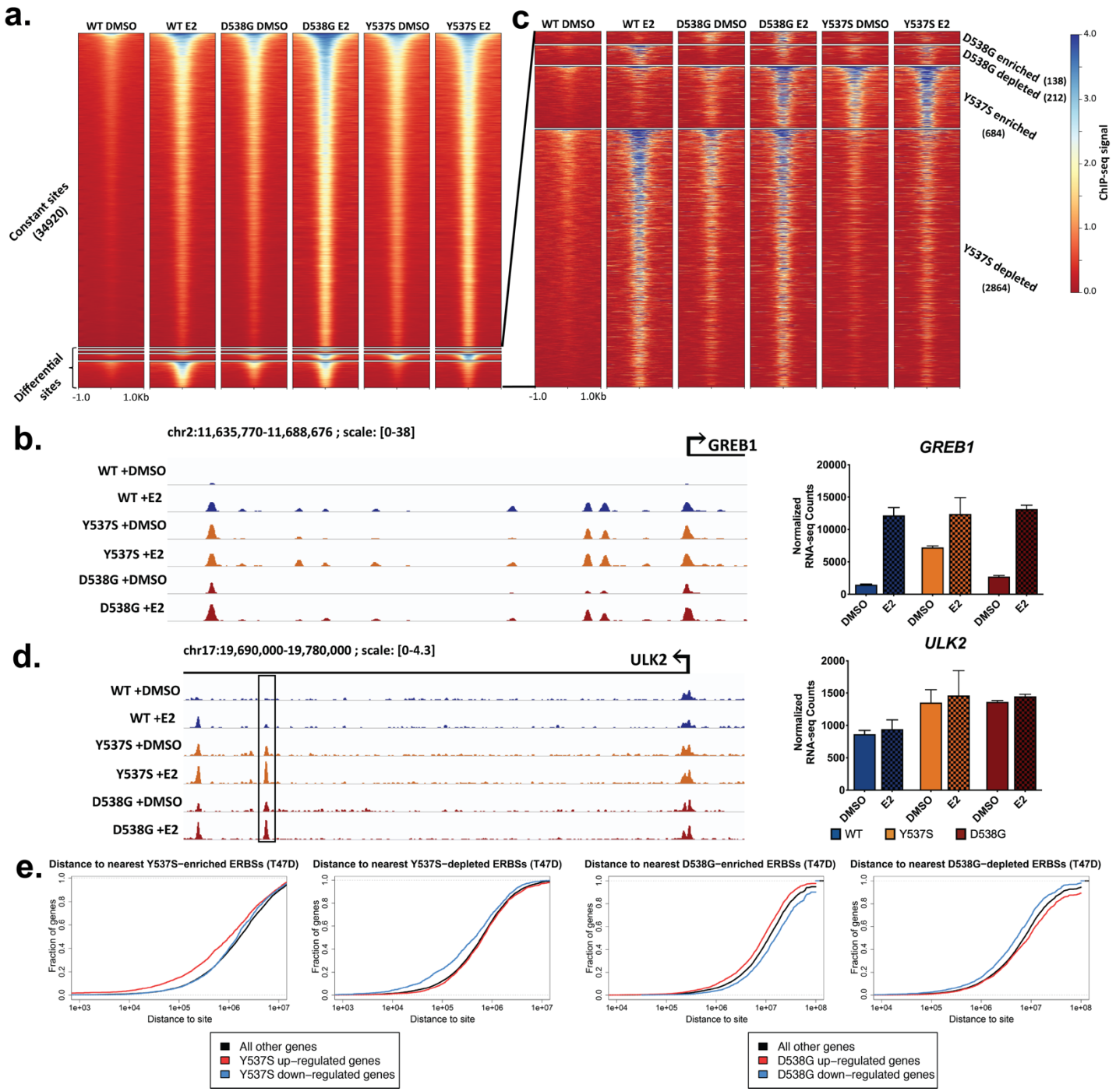


Figure 4. Limited changes in ER's genomic binding are observed in mutants compared to WT and these changes may explain a small portion of mutant-specific gene expression. (a) Heatmap shows the binding signal of constant and mutant-specific ERBS

in T-47D WT and mutant clones with addition of DMSO or E2. (b) Example locus of ligand-independent (constant) ER-binding, where arrow indicates the TSS of the *GREB1* gene. All tracks are normalized to the same scale. *GREB1* ligand-independent gene expression is shown on the right. (c) Heatmap displays the signal of mutant-enriched or depleted sites only. (d) Example locus of mutant-enriched ER-binding, where arrow indicates the TSS of the *ULK2* gene. All tracks are normalized to the same scale. *ULK2* gene expression is shown on the right. (e) Distance from mutant-specific genes to mutant-enriched and -depleted ERBS is shown as a cumulative distribution. Line colors indicate up- or down-regulated mutant-specific genes (red and blue) or all other genes (black). p-values from Wilcoxon tests can be found in Supplemental Table S3.

Similar results were seen in MCF-7 WT and mutant clones where 36606 sites were bound in WT cells treated with E2 that were also bound in mutant cells regardless of E2 treatment (Supplemental Fig. S6a-c). Differential ER binding in MCF-7 clones also follows similar patterns to those observed in T-47D clones with fewer D538G differentially bound ERBS (1161) than Y537S differentially bound ERBS (3020). Similar numbers of mutant-specific genes harboring an ERBS within 100 kb of the TSS were observed in MCF-7 cells compared to T-47D cells (Supplemental Fig. S6d, Supplemental Table S3). Approximately 13% of Y537S-enriched and -depleted genes contain at least one nearby enriched or depleted Y537S ERBS while only 1-6% of D538G genes contain nearby D538G-enriched or -depleted ERBS. While ligand-independent genomic binding by mutant ER is apparent, the low percentage of ERBS that exhibited significant differential binding suggests that changes in ER binding may not have a broad impact on the regulation of mutant-specific genes. However, mutant-specific genes are significantly closer to mutant-altered ERBS and therefore may account for a small portion (2-14%) of mutant-specific differential gene expression.

ER mutations cause widespread changes in chromatin accessibility

Because *ESR1* mutations do not appear to drastically reprogram ER genomic binding, we hypothesized that additional changes might occur at genomic loci bound by ER, or other transcription factors, that contribute to mutant ER's unique gene expression program. In order to identify genomic regions that could contribute to the observed mutant ER gene expression changes, we analyzed genome-wide chromatin accessibility using ATAC-seq in our WT and mutant T-47D clones with a 1-hour E2 or DMSO treatment. Minimal changes were observed in chromatin accessibility in response to E2 treatment with only 169 regions exhibiting significantly increased accessibility and no regions of decreased accessibility in WT clones. However, in mutant ER cells, thousands of genomic regions exhibit increased (1808 in Y537S clones and 1563 in D538G clones; mutant-enriched sites, Fig. 5a, example in Fig. 5b) or decreased accessibility (2060 in Y537S clones and 1587 in D538G clones; mutant-depleted sites, Fig. 5a) when compared to WT clones. The changes in chromatin accessibility were mutation-allele specific with only 21.6% of mutant-enriched and 23.7% of mutant-depleted regions shared between both mutations. To determine the extent to which ER could be directly involved at regions of altered chromatin accessibility, we overlapped ER bound sites identified from our FLAG tag ChIP-seq experiments with all regions of differential chromatin accessibility. Of the mutant-enriched ATAC-seq regions, only 30-50% overlap with ERBS while only 20-30% of mutant-depleted ATAC-seq regions are bound by ER (Fig. 5d), indicating that many of these chromatin changes are not direct effects of mutant ER. To determine if the results seen in our mutant cells are recapitulated in mutant lines described in previous studies, we performed ATAC-seq on T-47D WT and mutant D538G and Y537S cells first described in Bahreini et al (17). We performed multivariate differential accessibility analysis on the combined ATAC-seq data from these sets of mutant ER models and identified thousands of genomic regions that show consistent differential chromatin accessibility between mutant ER and WT lines (Supplemental Fig. S7a). ER binding again occurs at a minority (20-30%) of the differentially accessible regions from the multivariate analysis (Supplemental Fig. S7b). These findings suggest that considerable changes in chromatin accessibility are consistently seen at thousands of sites in multiple independent ER mutant models, and that they are largely mutation-allele specific and mostly lack ER binding.

We explored the relationship between changes in chromatin accessibility and mutant-specific gene expression by comparing ATAC-seq and RNA-seq results from our T-47D mutant clones. We found that mutant-specific up- and down-regulated genes were significantly enriched near regions of mutant altered chromatin accessibility (Fig. 5c, Supplemental Fig. S8a). Of mutant-specific up-regulated genes 17% above background (all other genes) had at least one mutant-enriched region within 100 kb. Of mutant-specific down-regulated genes, more than 15% above background contained at least one nearby mutant-depleted region (Supplemental Table S3). To illustrate the effect that altered chromatin accessibility could have on gene expression, two regions of increased chromatin accessibility near the mutant-specific *CEP78* gene are shown. These regions correlate positively to the increased expression of the *CEP78* gene and may function as regulatory elements controlling its expression (Fig. 5b). We next compared genes that had nearby mutant-altered ERBS with those harboring nearby mutant-altered accessible chromatin regions and found that many Y537S-specific genes harbored both differential ER binding and accessible chromatin sites within 100 kb of their TSS (Supplemental Fig S8b).

However, chromatin accessibility changes were found much more often nearby mutant-specific genes than changes in ER binding, especially in the context of the D538G mutation. These findings suggest that while some mutant-specific gene expression changes may be a direct consequence of altered ER binding, more often these changes appear to be associated with regions of altered chromatin accessibility and potentially other transcription factors binding these regions.

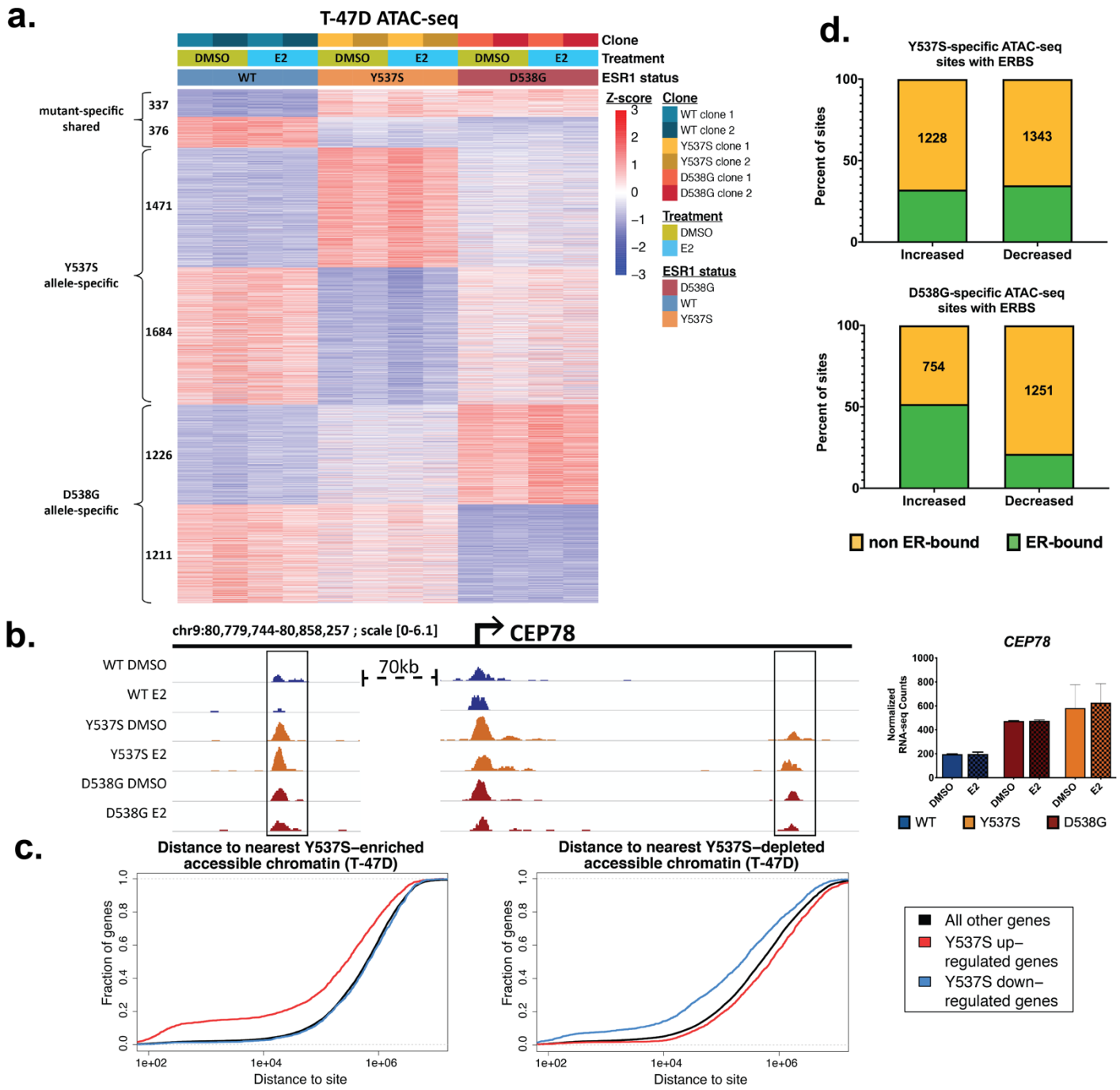


Figure 5. ER mutant cells exhibit large-scale alterations in chromatin accessibility. (a) Heatmap displays ATAC-seq signal of T-47D mutant-enriched and -depleted chromatin accessibility sites. Differential chromatin accessibility regions are grouped by those shared across both mutations and those specific to one mutation. Clone, treatment, and mutational status are indicated by the header bars and legend. (b) Example locus of shared mutant-enriched accessible chromatin region is shown. Arrow indicates the TSS for the *CEP78* gene. All tracks are normalized to the same scale. *CEP78* gene expression is shown on the right. (c) Cumulative distribution graphs show distance from Y537S mutant-specific genes to Y537S mutant-enriched accessible chromatin. Line colors indicate up- or down-regulated mutant-specific genes (red and blue) or all other genes (black). p-values from Wilcoxon tests can be found in Supplemental Table S3. (d) Percent of mutant-enriched or -depleted accessible chromatin regions that also exhibit ER binding is shown.

Transcription factors associated with regions of differential accessibility in ER mutant cells

To identify additional factors associated with mutant-enriched accessible chromatin, we performed motif enrichment analysis on mutant-enriched ATAC-seq sites, identified above, that were negative for ER binding (ERBS-neg). We also excluded any sites that were proximal to gene TSS (within 2 kb) from our analysis due to the heavy over representation of GC rich regions that could bias motif analysis results. Motif analysis of mutant-enriched ATAC-seq sites resulted in a number of motifs that were significantly enriched in one or both of the mutant genotypes (Fig. 6a, Supplemental Fig. S7c). In D538G-enriched sites, the forkhead factor (FOX) motif was most highly enriched (p-value = 1.5e-31). Y537S-enriched sites were highly enriched for POU (OCT) factor (p-value = 8.0e-78) and CTCF (3.3e-50) motifs. Since the FOX and OCT motifs can bind to many related transcription factors, we analyzed expression data and found that 4 FOX factors and 3 OCT factors had detectable mRNA (Fig. 6b); however, FOXA1 exhibited much higher expression than other FOX factors and OCT1 was the highest expressed OCT factor. OCT1 was also more highly expressed in ER mutant cells compared to WT. These observations led us to examine OCT1, FOXA1, and CTCF in more depth.

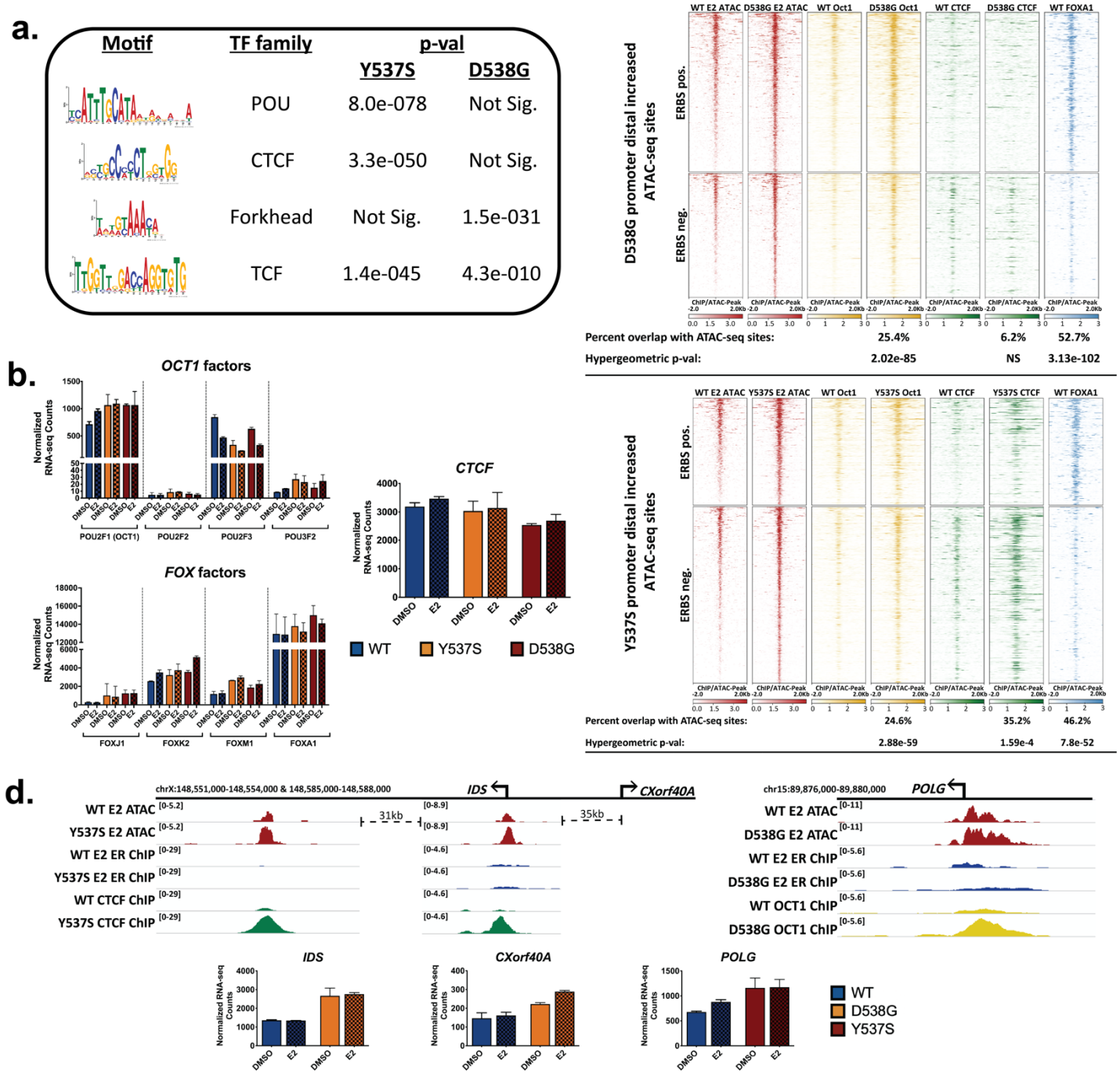


Figure 6. Transcription factors are observed at T-47D mutant-enriched accessible chromatin regions. (a) Enriched transcription factor DNA binding motifs are found at T-47D mutant-enriched accessible chromatin. Identified motifs are shown (left) along with corresponding adjusted p-values (MEME-Suite E-value) for each mutation. (b) Gene expression is shown for *CTCF*, *FOX* family genes, and *POU* family genes. Error bars show SEM for two clones for each genotype and each treatment. (c) Heatmaps display the intensity

(depth) of accessible chromatin (red) or DNA-binding of three factors: OCT1 (yellow), CTCF (green), and FOXA1 (blue) at mutant-enriched accessible chromatin regions. Percent overlap of transcription factor binding with mutant-enriched accessible chromatin regions is indicated below each respective heatmap. p-values are from hypergeometric tests. (d) Example loci show mutant-enriched chromatin accessibility regions with corresponding increased CTCF or OCT1 binding and correlated expression of nearby genes (*IDS, CXorf40A,* and *POLG*; bottom panels). All tracks are normalized to the same scale. Bar graphs show gene expression in normalized RNA-seq count. Error bars show SEM for two clones for each genotype and each treatment.

To investigate the potential role of these factors in mutant ER cells, we determined the extent to which these factors bind to mutant-enriched accessible chromatin. We performed ChIP-seq for OCT1 and CTCF in the T47-D WT and ER mutant cells in hormone-depleted conditions. We also analyzed FOXA1 ChIP-seq data collected previously in T-47D WT cells (28). OCT1 binding was highly enriched in TSS-distal, mutant-enriched accessible chromatin regions in both the Y537S (p-value = 2.88e-59; hypergeometric test) and D538G (p-value = 2.02e-85) mutant clones. Significant OCT1 binding was found at 25% of both ERBS-negative and ERBS-positive accessible chromatin regions (Fig. 6c). CTCF showed significant enrichment in Y537S mutant-enriched, TSS-distal accessible chromatin (p-value = 1.59e-4) and overlapped with 35% of these sites, but was not enriched in D538G-enriched accessible regions, matching the motif pattern. CTCF was more prevalent at ERBS-negative than at ERBS-positive regions. FOXA1 was enriched in TSS-distal accessible chromatin regions in both Y537S (p-value = 7.8e-52) and D538G (p-value = 3.13e-102) mutant clones and was found at approximately half of both mutants' TSS-distal accessible regions, but was more evident at ERBS-positive regions than ERBS-negative regions. The enrichment for FOXA1 in ERBS-positive regions is not surprising as it is a known pioneer factor for ER (29) and could be functioning in this role to increase chromatin accessibility in ER mutant breast cancer cells. The binding of these factors at regions of mutant-enriched chromatin accessibility suggests that they could function in altering the expression of mutant-specific genes. Increased CTCF or OCT1 binding in mutant compared to WT cells near mutant-specific genes positively correlates with increased mutant-specific gene expression and suggests a role for these factors in altering gene expression in these cells (Examples in Fig 6d). Overall, the chromatin accessibility analysis has led us to three transcription factors (OCT1, CTCF, and FOXA1) that bind to regions of the genome that become more accessible in ER mutant cells, indicating that these factors may contribute to the unique gene expression patterns attributed to mutant ER.

Discussion

In breast cancer, ER mutations arise under the challenge of hormone therapy that aims to block estrogen production or disrupt ER activity through binding to the LBD. From previous studies it is clear that mutations in ER's LBD confer ligand independent activity to the receptor (4, 5, 10, 12, 14-16), which in large part explains the clinical observation of ER mutations in metastatic breast cancer. However, it has been reported that ER mutations cause additional gene expression changes that appear unrelated to ER's usual target genes (17-19). In this study, we aimed to confirm ER mutant's ability to regulate genes that are not normally E2 responsive. By creating isogenic models of the Y537S and D538G mutations in MCF-7 and T-47D breast cancer cell lines, we found that thousands of non-E2 responsive genes changed expression due to the mutation in addition to hundreds of E2 responsive genes that exhibited ligand independent regulation in mutant cells. To determine if these gene expression effects were specific to our clones or consistently observed across multiple models from different studies, we performed an aggregate gene expression analysis with a total of 16 lines for T-47D and 18 lines for MCF-7 from 3 studies. When taking into account study-to-study variation, we identified thousands of genes that consistently change expression due to ER mutation and do not change expression in response to 8-24 hours of E2 treatment. We observed that most genes have different responses to Y537S and D538G, confirming that allele-specific changes in gene expression are consistent across models from different studies.

One possible explanation for the effect of mutant ER on genes that are not normally E2 responsive is that constant ER activity could have different gene expression consequences than short-term E2 treatments. In an effort to understand the contribution of constant ER activity to the unique ER mutant gene expression patterns, we treated WT cells with E2 for up to 25 days and performed RNA-seq at 5 day intervals. Prolonged exposure to E2 explained approximately half of the mutant-specific gene expression effects with Y537S being more similar to long-term E2 treatment than D538G reiterating the differences unique to each mutant allele. Our findings indicate that approximately half of the ER mutant gene expression effects that appear unrelated to normal ER regulation are in fact estrogen target genes that rely on long-term E2 exposure.

Another way to interpret our prolonged E2 treatment experiment is that half of ER mutant regulated genes cannot be attributed to long-term E2 exposure and must involve novel function of the mutant receptor. We first analyzed genomic binding of mutant and WT ER taking advantage of the FLAG epitope tag that we introduced in the native

ER locus. While we observed ligand-independent ER binding in mutant ER cells at the majority of ERBS, we did not observe large-scale reprogramming of ER genomic binding, where approximately 10% of ER bound loci were significantly altered by the mutations. These findings were in contrast to our recent studies on the D538G mutation in endometrial cancer cells that revealed more substantial changes in ER binding (46% of ER bound sites were significantly affected) (22). This indicates that the manner in which mutations influence ER could be unique to different cell types or that ER genomic binding is more robust to mutation in breast cancer cell lines. Both mutant-specific alterations in ER genomic binding (18) and lack of alterations (12) have been reported and our study supports the conclusion that mutations have a relatively small impact on ER genomic binding. While the overall picture of ER genomic binding does not dramatically change with mutation, we still found that approximately 10% of Y537S expression changes could be explained by significantly altered ER binding in ER mutant cells. D538G changes in ER binding were minimal in both T-47D and MCF-7 lines, suggesting that the Y537S mutation may have more of an impact on ER's genomic binding site selection. Looking at constant ERBS, we found a 10% enrichment near mutant specific genes above background; however, this may be an underestimate of constant ERBS impact on mutant-specific expression due to the large number of constant ERBS, which leads to a high likelihood of a gene having an ERBS nearby simply by chance.

Considering the modest changes in ER genomic binding, we looked for other regulatory regions and factors that could contribute to the unexpected mutant-specific gene expression patterns. By performing ATAC-seq in our WT and ER mutant clones, we found that thousands of loci exhibited differential chromatin accessibility in ER mutant cells. We were able to show consistency in chromatin accessibility changes by analyzing another set of ER mutant clones from a different study (17). Some of these changes in chromatin could be a direct result of mutant ER binding; however, the majority of differentially accessible regions did not harbor ER binding. In addition, differentially accessible regions were at least twice as likely to be found near mutant-specific genes as differential ERBS and could explain approximately 15% of mutant-specific up- and down-regulated genes. These observations led us to the conclusion that other transcription factors contribute to the mutant-specific chromatin landscape and gene expression program.

Motif analysis of differential ATAC-seq sites uncovered OCT, forkhead, and CTCF motifs. Additional experiments confirmed that OCT1, FOXA1, and CTCF were bound at a significant fraction of differentially accessible regions, suggesting that they may be playing a role in increasing chromatin accessibility or possibly taking advantage of new open regions. FOXA1 is a critical factor for ER genomic binding in breast cancer (29) and previous reports have suggested that FOXA1 may not bind to mutant-specific ER bound sites (18). However, our data reveals that FOXA1 could be playing a role in mutant-specific gene expression. OCT1 has been shown to both negatively and positively impact gene expression depending on its associating factors. It can both maintain a poised state of gene expression by facilitating the removal of repressive histone marks and drive a repressed gene expression state by failing to facilitate the removal of these marks (30). In its role as a transcription factor, OCT1 has been shown to contribute to increased cell proliferation, cellular stress response, altered metabolism, regulation of the immune response, and invasion and metastasis (31, 32). OCT1 also regulates somatic and cancer stem cell phenotypes in normal and tumor cells (33). OCT1's presence at mutant-enriched accessible chromatin and increased expression in ER mutant cells suggest that it may be playing a regulatory role and contributing to the expression changes caused by mutant ER. CTCF was found specifically at the Y537S enriched ATAC-seq sites. CTCF plays a key role in the 3D organization of the genome, as reviewed in Ong et al (34), and the presence of CTCF at differential ATAC-seq regions indicates that 3D genome architecture could be altered in ER mutant cells. Overall, our analyses of ER genomic binding and chromatin accessibility can explain approximately a quarter of mutant-specific gene expression effects and prolonged ER activity can explain an overlapping 50% of mutant-specific genes. This indicates that additional gene regulation mechanisms are contributing to the ER mutant gene expression program, which may include changes in the 3D genome architecture, post-transcriptional changes, and differences at ERBS, possibly through differential recruitment of cofactors.

Our findings show that the majority of gene expression patterns as well as a small portion of ER genome binding sites are unique to either the Y537S or D538G mutations. Additionally, we see clear allele-specific changes in chromatin accessibility and accompanying transcription factors. Previous studies have also demonstrated allele-specific differences between ER mutations that appear in the structures and molecular effects of mutant ER including differences in their ability to bind cofactors and in their response to hormone therapies (5, 6, 11, 13, 17, 18, 35). These unique differences between ER mutants have significant clinical relevance as *ESR1* mutational status could be used to determine treatment strategies that would best serve the patient. This is particularly true as novel therapies are identified that more effectively treat breast cancers harboring specific *ESR1* mutations. Overall, our study shows that ER mutations consistently impact the expression of thousands

of genes that are not normally estrogen regulated and do so partially through constant ER activity and partially through the use of novel regulatory elements.

Materials and Methods:

Cell Culture

WT and mutant T-47D cells were grown in RPMI 1640 media (Thermo Fisher Scientific) supplemented with 10% fetal bovine serum (Sigma-Aldrich) and 1% penicillin-streptomycin (Thermo Fisher Scientific) (full media). MCF-7 WT and mutant cells were grown in Minimum Essential Media (MEM, Thermo Fisher Scientific) with 5% fetal bovine serum, 1% penicillin-streptomycin, 1x final concentration non-essential amino acids (NEAA, Thermo Fisher Scientific), and 1nM final concentration insulin (Sigma-Aldrich) (full media). For all experiments, cells were grown at 37°C with 5% CO_2. Five days before all experiments requiring an E2 induction, cells were placed in hormone-depleted media. Hormone-depleted media consisted of phenol-red free MEM (Thermo Fisher Scientific, 5% charcoal-stripped fetal bovine serum (Thermo Fisher Scientific), 1% penicillin-streptomycin, and NEAA and insulin as described above for MCF-7 cells and phenol-red free RPMI (Thermo Fisher Scientific), 10% charcoal-stripped fetal bovine serum, and 1% penicillin-streptomycin for T-47D cells. Hormone-depleted media was changed twice during each 5-day estrogen-deprivation period in order to ensure complete removal of estrogens from the media. After 5 days of estrogen deprivation, cells were treated with DMSO as a vehicle control (Fisher Scientific) or 10nM E2 (Sigma-Aldrich) for 1 hour for ChIP-seq and ATAC-seq experiments and 8 hours for qPCR and RNA-seq experiments.

Generation of mutant clones

Mutant clones were generated using the CETCH-seq method described by Savic et al (36). We followed procedures outlined by Blanchard et al. (22) (Supplemental Fig. S1a) using the same guide RNA/Cas9 vector and WT and D538G pFETCH vectors for mutant and WT ER clone generation. For this study, we generated a Y537S pFETCH vector using the same approach as in Blanchard et al., substituting the D538G gBlock with a Y537S gBlock (Supplemental Table S4). Transfections were performed on cells in full media using either Lipofectamin (Thermo Fisher Scientific) or Fugene (Promega) transfection reagents both of which yielded similar success rates. Cells were treated with 1uM SCR7 (Xcessbio) for 3 days post-transfection to block non-homologous end joining. 72 hours post-transfection, cells were treated with G418 (Thermo Fisher Scientific) at 300ug/ml final concentration, which was applied every two days in conjunction with media changes. Single-cell clones were picked and validated using the same procedures as described by Blanchard et al. (22) including limiting dilution plating, colony picking, sanger sequencing and FLAG and ER immunoblots. In all, two clones for each genotype (WT, Y537S, and D538G) were validated for both T-47D and MCF-7 cell lines.

Quantitative PCR

T-47D and MCF-7 WT or mutant clones were grown in hormone-depleted media for 5 days and then treated for 8 hours with DMSO or 10nM E2. Cell lysates were harvested using buffer RLT plus (Qiagen) containing 1% beta-mercaptoethanol (Sigma-Aldrich). Total RNA was harvested using a Quick RNA Miniprep Kit (Zymo Research). 50ng of RNA was used per sample to perform qPCR using a Power SYBR Green RNA-to-CT 1-Step Kit (Thermo Fisher Scientific). A CFX Connect Real-Time light cycler (Bio-Rad) was used for thermocycling and fluorophore detection. Expression levels were calculated using the $\Delta\Delta C_t$ method with *CTCF* measurements as the control. Duplicates were measured for each sample. Primers for *CISH*, *SUSD3*, *PGR*, and *CTCF* are found in Supplemental Table S4. Significance of relative expression changes were determined using Student's Two Sample t-test, as implemented in R using t.test, with a p-value cutoff of <0.05.

RNA-seq

Cells from T-47D and MCF-7 WT and mutant clones were plated in 100mm dishes and grown in hormone-depleted media for 5 days followed by treatment with DMSO or 10nM E2 for 8 hours. Cells were washed with phosphate buffered saline (PBS; Thermo Fisher Scientific) and treated with buffer RLT plus (Qiagen) containing 1% beta-mercaptoethanol (Sigma-Aldrich) for cell lysis. Cells were passed through a 21-guage needle and syringe (Sigma-Aldrich) several times to facilitated lysis of genomic DNA. RNA was then extracted using a Quick-RNA Miniprep kit (Zymo Research). RNA was quantified and poly(A) selected RNA-seq libraries were generated using a KAPA Stranded mRNA-Seq kit (KAPA Biosystems) with 500ng RNA per sample as starting material. Libraries were sequenced on the Illumina HiSeq 2500 platform. Resulting RNA-seq data were analyzed to call differential gene expression between WT and mutant clones. Fastq files were aligned to the hg19 human genome

build using the HISAT2 spliced aligner (37). SAM files were converted to BAM files using SAMtools (38). Counts for each gene from the University of California Santa Cruz's (UCSC) Known Genes table were assigned using the Subread package's featureCounts program with BAM files as inputs (39). Read counts were normalized and analyzed for differential expression analysis using the DESeq2 package in R (40). RNA-seq was performed for each T-47D and MCF-7 WT, Y537S, and D538G clone. Clones from the same parental cell line and of the same genotype were used as biological replicates in the differential expression analysis between each mutant genotype and the WT clones. PCA plots were generated using the top 5000 most variable genes for each cell line (T-47Ds or MCF-7s). Genes that were significantly up- or down-regulated (adjusted p-value <0.05) in response to E2 treatment (E2-regulated genes) were identified by comparing WT clones treated with E2 to WT clones treated with DMSO. Next, genes that were significantly up- or down-regulated in either the Y537S or D538G mutants were identified by comparing all WT samples to all D538G or all Y537S samples from the same parental cell line regardless of E2 or DMSO treatment. From the resulting differentially expressed genes, E2-regulated genes were removed and the remaining genes were defined as mutant-specific genes. Mutant-specific genes sets were analyzed using Illumina's BaseSpace Correlation Engine to identify pathways and molecular and cellular processes that are significantly associated with mutant-specific gene expression changes.

For the long-term E2 treatment RNA-seq experiments, cells from T-47D and MCF-7 WT clones were grown in 100mm dishes in hormone-depleted media and were treated with 10nM E2 for 25 days. E2 was added with media changes every 2 days to maintain E2 levels. Cells were taken at 5 day time points (5, 10, 15, 20, and 25 days). RNA purification, RNA-seq library generation, sequencing, and RNA-seq data analysis were all performed as described above. Genes that were significantly differentially expressed in response to long-term estrogen treatment were identified by comparing RNA-seq data from WT clones with DMSO and 8 hour E2 treatments to all WT long-term E2 treatment RNA-seq data. The cutoff for all significant genes was an adjusted p-value of <0.05.

Multi-lab RNA-seq analyses were performed by utilizing publicly-available RNA-seq data from T-47D and MCF-7 WT and Y537S and D538G mutant clones generated by two other groups (17, 18). GEO accession numbers and specific RNA-seq data sets used can be found in Supplemental Table S2. RNA-seq data alignment and count assignment were performed as described above using HISAT2 and featureCounts. PCA plots were generated using the top 10000 most variable genes for each cell line. Differential gene expression analysis was then performed using DESeq2 with a modified multivariate negative binomial regression model to account for variation due to the study in which the RNA-seq data was collected. To implement this multivariate approach, we used the DESeq2 design formula: ~group + E2TreatmentLength + ESR1Status. Y537S and D538G mutant up- and down-regulated genes were identified using an adjusted p-value cutoff of <0.05, which represented significance of the ESR1Status variable.

RPPA analysis

Cells from T-47D WT and mutant clones were plated at 2 million cells per clone on 100mm plates and were grown in hormone-depleted media for 5 days. Cells were then trypsinized using phenol-red free trypsin and washed twice with PBS. 2 million cells were pelleted and shipped on dry ice for RPPA analysis. RPPA was performed at the MD Anderson Cancer Center Functional Proteomics RPPA Core Facility as described previously (41). In short, protein lysates were serially diluted then spotted onto nitrocellulose-coated slides and probed with 445 antibodies using a tyramide-based signal amplification approach. Slides were stained using a DAB colorimetric reaction and were scanned using a Huron TissueScope scanner. Relative protein levels were determined using a standard curve approach. The resulting normalized log2 protein concentrations were analyzed to identify proteins that exhibited significant increased or decreased expression in either mutant compared to WT. Significance was determined using the Welch t-test, using t.test in R, with a p-value cutoff of <0.05.

ChIP-seq

Cells from WT and mutant clones were plated in 150mm dishes and grown in hormone-depleted media for 5 days. Cells were then treated with either DMSO or 10nM E2 for 1 hour. Following DMSO or E2 treatments, cells were fixed for 10 minutes at room temperature with 1% formaldehyde (final concentration) added directly to the media (Sigma-Aldrich). Fixation was ended by briefly treating media with glycine to a final concentration of 125mM to quench the remaining formaldehyde. Media was removed, and fixed cells were washed once in cold PBS. Fixed chromatin was then harvested by scraping in cold Farnham lysis buffer (5 mM PIPES pH 8.0; 85 mM KCl; 0.5% NP-40) supplemented with protease inhibitor (Thermo Fisher Scientific). Chromatin immunoprecipitation was performed on fixed chromatin as described previously (42) using antibodies against

the FLAG epitope tag (Anti-FLAG M2 antibody; Sigma-Aldrich F1804), CTCF (sc-5916, Santa Cruz Biotechnology) or two antibodies against OCT1 combined at equal ratios (A301-716A and A301-717A, Bethyl Laboratories Inc.). Sequencing was performed using the Illumina HiSeq 2500 platform. FOXA1 binding was analyzed using previously obtained FOXA1 ChIP-seq data (28). Resulting reads in Fastq format were aligned to the hg19 human genome build using bowtie2 (43) with the following parameters: --chunkmbs 512 -m 1 -t –best -q -S -l 32 -e 80 -n 2 -p. Output SAM files were sorted and converted to BAM files using SAMtools (38). MACS2 (44) was used to call peaks using a p-value cutoff of 1×10^{-10} and a mfold parameter between 15 and 100. Control input libraries from WT and mutant lines were used as controls for MACS2 peak calling in each ChIP-seq experiment. Clones were used as replicates for each experiment. Significantly differentially bound FLAG tagged ERBS sites were determined using DESeq2 (40) to compare counts per million at each ER binding site called. FLAG tag ER ChIP-seq results for E2-treated WT clones were compared to either Y537S or D538G DMSO- and E2-treated FLAG tag ER ChIP-seq results. An adjusted p-value cutoff of <0.05 was used to identify significantly mutant-enriched and -depleted ER-bound sites. Constant ER-bound sites were sites that were bound in at least one replicate from any genotype that were not differentially bound according to our DESeq2 analysis. Enrichment of CTCF, OCT1, and FOXA1 binding at mutant-enriched accessible chromatin was assessed using the BEDTools (45) and deepTools (46) packages.

ATAC-seq

Cells from T-47D WT and mutant clones, including clones derived in this study and clones reported in Bahreini et al. (17), were plated in 6-well plates and grown in estrogen-depleted media for 5 days. Following estrogen deprivation, cells were treated with DMSO vehicle control or 10nM E2 for 1 hour, then washed with PBS and removed from the plates using phenol-red free trypsin. Cells were pelleted and resuspended in hormone-depleted media. 250,000 cells for each sample were used for each ATAC-seq experiment. ATAC-seq libraries were prepared as previously described (47) and were sequenced on an Illumina HiSeq 2500 sequencer. Sequenced reads were aligned to the hg19 human genome build using bowtie2 (43) with the following parameters: --chunkmbs 512 -m 1 -t --best -q -S -l 32 -e 80 -n 2. Resulting SAM files were converted to BAM files using SAMtools (38). Regions of accessible chromatin were identified by calling ATAC-seq peaks using MACS2 with an adjusted p-value cutoff of <0.05 and a mfold parameter between 15 and 100. Subread's featureCounts (39) was used to quantify reads that align to all peaks called in any sample. Aligned reads were normalized and mutant-enriched and -depleted chromatin accessibility regions were identified using the DESeq2 package in R (40). Clones developed by our lab were used to identify regions of differentially accessible chromatin. Clones with the same ER genotype were grouped as biological replicates, and all WT samples were compared to all Y537S or all D538G mutant samples to identify mutant-enriched and -depleted chromatin accessibility regions. Overlap of ATAC-seq sites and ERBS identified through FLAG epitope tag ChIP-seq was performed using the BEDTools intersect tool (45) with an overlap size of at least 1 base pair. 100-bp regions surrounding the summits of mutant-enriched non-ER bound ATAC-seq regions >2kb from gene TSS were used to perform motif analysis using MEME-Suite's MEME-ChIP tool (48). Enriched motifs between 6 and 20 bps in length were reported, with a motif site distribution of 0 to 1 occurrence per sequence.

Multi-lab ATAC-seq was performed using ATAC-seq data from T-47D WT and mutant clones developed in this study along with those reported in Bahreini et al. (17). Differential chromatin accessibility was assessed using DESeq2 with a modified multivariate negative binomial regression model to account for variation due to the study in which the clones were developed. A multivariate negative binomial regression model was implemented by using the DESeq2 design formula: ~group + ESR1Status. Y537S and D538G mutant enriched- and depleted-accessible chromatin regions were identified using an adjusted p-value cutoff of <0.05, which represented significance of the ESR1Status variable.

Statistical and Graphical Packages

All statistical analyses were performed in R versions 3.5.2 or 3.5.3 (49) except p-values for gene ontology and pathway enrichments which were calculated by Illumina's BaseSpace Correlation Engine and p-values for motif enrichments which were calculated by MEME suite (48). P-values and statistical tests used can be found throughout the text. Heatmaps for gene expression and differential chromatin accessibility were generated using the pheatmap package in R. Heatmaps display z-scores based on reads per million that align to each gene or region. ChIP-seq results were visualized using the deepTools package (46) to compare constant and mutant-enriched and -depleted ERBS. Distances between mutant up- or down-regulated genes and mutant-enriched or -depleted ERBS or ATAC-seq sites were calculated using a Perl script. Distance plots were generated using the R plot function.

Data Access

Raw and processed data is available at the Gene Expression Omnibus (GEO) under the following accession: GSE148279

Acknowledgments

This work was supported by a DOD Breast Cancer Research Program Breakthrough Award to J.G. and J.K.R., a DOD Breast Cancer Research Program Expansion Award to J.G., NIH R01CA221303 to S.O., and the Huntsman Cancer Institute. Research reported in this publication utilized the High-Throughput Genomics Shared Resource at the University of Utah and was supported by NIH/NCI award P30 CA042014. We thank K-T Varley and Dan Savic as well as Gertz, Varley, Richer, and Oesterreich laboratory members for their helpful comments on the study and the manuscript.

References

1. Blows FM, Driver KE, Schmidt MK, Broeks A, van Leeuwen FE, Wesseling J, Cheang MC, Gelmon K, Nielsen TO, Blomqvist C, Heikkila P, Heikkinen T, Nevanlinna H, Akslen LA, Begin LR, Foulkes WD, Couch FJ, Wang X, Cafourek V, Olson JE, Baglietto L, Giles GG, Severi G, McLean CA, Southey MC, Rakha E, Green AR, Ellis IO, Sherman ME, Lissowska J, Anderson WF, Cox A, Cross SS, Reed MW, Provenzano E, Dawson SJ, Dunning AM, Humphreys M, Easton DF, Garcia-Closas M, Caldas C, Pharoah PD, Huntsman D. Subtyping of breast cancer by immunohistochemistry to investigate a relationship between subtype and short and long term survival: a collaborative analysis of data for 10,159 cases from 12 studies. PLoS Med. 2010;7(5):e1000279. Epub 2010/06/04. doi: 10.1371/journal.pmed.1000279. PubMed PMID: 20520800; PMCID: PMC2876119.

2. Group EBCTC. Tamoxifen for early breast cancer: an overview of the randomised trials. Lancet. 1998;351(9114):1451-67. Epub 1998/05/30. PubMed PMID: 9605801.

3. Group EBCTC. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet. 2005;365(9472):1687-717. Epub 2005/05/17. doi: 10.1016/s0140-6736(05)66544-0. PubMed PMID: 15894097.

4. Robinson DR, Wu YM, Vats P, Su F, Lonigro RJ, Cao X, Kalyana-Sundaram S, Wang R, Ning Y, Hodges L, Gursky A, Siddiqui J, Tomlins SA, Roychowdhury S, Pienta KJ, Kim SY, Roberts JS, Rae JM, Van Poznak CH, Hayes DF, Chugh R, Kunju LP, Talpaz M, Schott AF, Chinnaiyan AM. Activating ESR1 mutations in hormone-resistant metastatic breast cancer. Nat Genet. 2013;45(12):1446-51. Epub 2013/11/05. doi: 10.1038/ng.2823. PubMed PMID: 24185510; PMCID: PMC4009946.

5. Toy W, Shen Y, Won H, Green B, Sakr RA, Will M, Li Z, Gala K, Fanning S, King TA, Hudis C, Chen D, Taran T, Hortobagyi G, Greene G, Berger M, Baselga J, Chandarlapaty S. ESR1 ligand-binding domain mutations in hormone-resistant breast cancer. Nat Genet. 2013;45(12):1439-45. Epub 2013/11/05. doi: 10.1038/ng.2822. PubMed PMID: 24185512; PMCID: PMC3903423.

6. Toy W, Weir H, Razavi P, Lawson M, Goeppert AU, Mazzola AM, Smith A, Wilson J, Morrow C, Wong WL, De Stanchina E, Carlson KE, Martin TS, Uddin S, Li Z, Fanning S, Katzenellenbogen JA, Greene G, Baselga J, Chandarlapaty S. Activating ESR1 Mutations Differentially Affect the Efficacy of ER Antagonists. Cancer Discov. 2017;7(3):277-87. Epub 2016/12/18. doi: 10.1158/2159-8290.CD-15-1523. PubMed PMID: 27986707; PMCID: PMC5340622.

7. Guttery DS, Page K, Hills A, Woodley L, Marchese SD, Rghebi B, Hastings RK, Luo J, Pringle JH, Stebbing J, Coombes RC, Ali S, Shaw JA. Noninvasive detection of activating estrogen receptor 1 (ESR1) mutations in estrogen receptor-positive metastatic breast cancer. Clin Chem. 2015;61(7):974-82. Epub 2015/05/17. doi: 10.1373/clinchem.2015.238717. PubMed PMID: 25979954.

8. Allouchery V, Beaussire L, Perdrix A, Sefrioui D, Augusto L, Guillemet C, Sarafan-Vasseur N, Di Fiore F, Clatot F. Circulating ESR1 mutations at the end of aromatase inhibitor adjuvant treatment and after relapse in breast cancer patients. Breast Cancer Res. 2018;20(1):40. Epub 2018/05/18. doi: 10.1186/s13058-018-0968-0. PubMed PMID: 29769099; PMCID: PMC5956618.

9. Li S, Shen D, Shao J, Crowder R, Liu W, Prat A, He X, Liu S, Hoog J, Lu C, Ding L, Griffith OL, Miller C, Larson D, Fulton RS, Harrison M, Mooney T, McMichael JF, Luo J, Tao Y, Goncalves R, Schlosberg C, Hiken JF, Saied L, Sanchez C, Giuntoli T, Bumb C, Cooper C, Kitchens RT, Lin A, Phommaly C, Davies SR, Zhang J, Kavuri MS, McEachern D, Dong YY, Ma C, Pluard T, Naughton M, Bose R, Suresh R, McDowell R, Michel L, Aft R, Gillanders W, DeSchryver K, Wilson RK, Wang S, Mills GB, Gonzalez-Angulo A, Edwards JR, Maher C, Perou CM, Mardis ER, Ellis MJ. Endocrine-therapy-resistant ESR1 variants revealed by genomic characterization of breast-cancer-derived xenografts. Cell Rep. 2013;4(6):1116-30. Epub 2013/09/24. doi: 10.1016/j.celrep.2013.08.022. PubMed PMID: 24055055; PMCID: PMC3881975.

10. Jeselsohn R, Yelensky R, Buchwalter G, Frampton G, Meric-Bernstam F, Gonzalez-Angulo AM, Ferrer-Lozano J, Perez-Fidalgo JA, Cristofanilli M, Gomez H, Arteaga CL, Giltnane J, Balko JM, Cronin MT, Jarosz M, Sun J, Hawryluk M, Lipson D, Otto G, Ross JS, Dvir A, Soussan-Gutman L, Wolf I, Rubinek T, Gilmore L, Schnitt S, Come SE, Pusztai L, Stephens P, Brown M, Miller VA. Emergence of constitutively active estrogen receptor-alpha mutations in pretreated advanced estrogen receptor-positive breast cancer. Clin Cancer Res. 2014;20(7):1757-67. Epub 2014/01/09. doi: 10.1158/1078-0432.CCR-13-2332. PubMed PMID: 24398047; PMCID: PMC3998833.

11. Fanning SW, Mayne CG, Dharmarajan V, Carlson KE, Martin TA, Novick SJ, Toy W, Green B, Panchamukhi S, Katzenellenbogen BS, Tajkhorshid E, Griffin PR, Shen Y, Chandarlapaty S, Katzenellenbogen JA, Greene GL. Estrogen receptor alpha somatic mutations Y537S and D538G confer breast cancer endocrine resistance by stabilizing the activating function-2 binding conformation. Elife. 2016;5. Epub 2016/02/03. doi: 10.7554/eLife.12792. PubMed PMID: 26836308; PMCID: PMC4821807.

12. Harrod A, Fulton J, Nguyen VTM, Periyasamy M, Ramos-Garcia L, Lai CF, Metodieva G, de Giorgio A, Williams RL, Santos DB, Gomez PJ, Lin ML, Metodiev MV, Stebbing J, Castellano L, Magnani L, Coombes RC, Buluwela L, Ali S. Genomic modelling of the ESR1 Y537S mutation for evaluating function and new therapeutic approaches for metastatic breast cancer. Oncogene. 2017;36(16):2286-96. Epub 2016/10/18. doi: 10.1038/onc.2016.382. PubMed PMID: 27748765; PMCID: PMC5245767.

13. Katzenellenbogen JA, Mayne CG, Katzenellenbogen BS, Greene GL, Chandarlapaty S. Structural underpinnings of oestrogen receptor mutations in endocrine therapy resistance. Nat Rev Cancer. 2018;18(6):377-88. Epub 2018/04/18. doi: 10.1038/s41568-018-0001-z. PubMed PMID: 29662238; PMCID: PMC6252060.

14. Merenbakh-Lamin K, Ben-Baruch N, Yeheskel A, Dvir A, Soussan-Gutman L, Jeselsohn R, Yelensky R, Brown M, Miller VA, Sarid D, Rizel S, Klein B, Rubinek T, Wolf I. D538G mutation in estrogen receptor-alpha: A novel mechanism for acquired endocrine resistance in breast cancer. Cancer research. 2013;73(23):6856-64. Epub 2013/11/13. doi: 10.1158/0008-5472.CAN-13-1197. PubMed PMID: 24217577.

15. Weis KE, Ekena K, Thomas JA, Lazennec G, Katzenellenbogen BS. Constitutively active human estrogen receptors containing amino acid substitutions for tyrosine 537 in the receptor protein. Mol Endocrinol. 1996;10(11):1388-98. Epub 1996/11/01. doi: 10.1210/mend.10.11.8923465. PubMed PMID: 8923465.

16. Zhang QX, Borg A, Wolf DM, Oesterreich S, Fuqua SA. An estrogen receptor mutant with strong hormone-independent activity from a metastatic breast cancer. Cancer research. 1997;57(7):1244-9. Epub 1997/04/01. PubMed PMID: 9102207.

17. Bahreini A, Li Z, Wang P, Levine KM, Tasdemir N, Cao L, Weir HM, Puhalla SL, Davidson NE, Stern AM, Chu D, Park BH, Lee AV, Oesterreich S. Mutation site and context dependent effects of ESR1 mutation in genome-edited breast cancer cell models. Breast Cancer Res. 2017;19(1):60. Epub 2017/05/26. doi: 10.1186/s13058-017-0851-4. PubMed PMID: 28535794; PMCID: PMC5442865.

18. Jeselsohn R, Bergholz JS, Pun M, Cornwell M, Liu W, Nardone A, Xiao T, Li W, Qiu X, Buchwalter G, Feiglin A, Abell-Hart K, Fei T, Rao P, Long H, Kwiatkowski N, Zhang T, Gray N, Melchers D, Houtman R, Liu XS, Cohen O, Wagle N, Winer EP, Zhao J, Brown M. Allele-Specific Chromatin Recruitment and Therapeutic Vulnerabilities of ESR1 Activating Mutations. Cancer Cell. 2018;33(2):173-86 e5. Epub 2018/02/14. doi: 10.1016/j.ccell.2018.01.004. PubMed PMID: 29438694; PMCID: PMC5813700.

19. Yu L, Wang L, Mao C, Duraki D, Kim JE, Huang R, Helferich WG, Nelson ER, Park BH, Shapiro DJ. Estrogen-independent Myc overexpression confers endocrine therapy resistance on breast cancer cells expressing ERalphaY537S and ERalphaD538G mutations. Cancer Lett. 2019;442:373-82. Epub 2018/11/13. doi: 10.1016/j.canlet.2018.10.041. PubMed PMID: 30419347; PMCID: PMC6351074.

20. Martin LA, Ribas R, Simigdala N, Schuster E, Pancholi S, Tenev T, Gellert P, Buluwela L, Harrod A, Thornhill A, Nikitorowicz-Buniak J, Bhamra A, Turgeon MO, Poulogiannis G, Gao Q, Martins V, Hills M, Garcia-Murillas I, Fribbens C, Patani N, Li Z, Sikora MJ, Turner N, Zwart W, Oesterreich S, Carroll J, Ali S, Dowsett M. Discovery of naturally occurring ESR1 mutations in breast cancer cell lines modelling endocrine resistance. Nat Commun. 2017;8(1):1865. Epub 2017/12/02. doi: 10.1038/s41467-017-01864-y. PubMed PMID: 29192207; PMCID: PMC5709387.

21. Fiorillo M, Sanchez-Alvarez R, Sotgia F, Lisanti MP. The ER-alpha mutation Y537S confers Tamoxifen-resistance via enhanced mitochondrial metabolism, glycolysis and Rho-GDI/PTEN signaling: Implicating TIGAR in somatic resistance to endocrine therapy. Aging (Albany NY). 2018;10(12):4000-23. Epub 2018/12/24. doi: 10.18632/aging.101690. PubMed PMID: 30573703; PMCID: PMC6326696.

22. Blanchard Z, Vahrenkamp JM, Berrett KC, Arnesen S, Gertz J. Estrogen-independent molecular actions of mutant estrogen receptor 1 in endometrial cancer. Genome Res. 2019;29(9):1429-41. Epub 2019/08/01. doi: 10.1101/gr.244780.118. PubMed PMID: 31362937; PMCID: PMC6724674.

23. Gelsomino L, Gu G, Rechoum Y, Beyer AR, Pejerrey SM, Tsimelzon A, Wang T, Huffman K, Ludlow A, Ando' S, Fuqua SA. ESR1 Mutations Affect Anti-proliferative Responses to Tamoxifen through Enhanced Cross-Talk with IGF Signaling. Breast Cancer Res Treat. 2016;157(2):253-65. doi: 10.1007/s10549-016-3829-5. PubMed PMID: 27178332; PMCID: 5510243.

24. Li Z, Levine KM, Bahreini A, Wang P, Chu D, Park BH, Oesterreich S, Lee AV. Upregulation of IRS1 Enhances IGF1 Response in Y537S and D538G ESR1 Mutant Breast Cancer Cells. Endocrinology. 2018;159(1):285-96. Epub 2017/10/14. doi: 10.1210/en.2017-00693. PubMed PMID: 29029116; PMCID: PMC5761602.

25. Chepelev I, Wei G, Wangsa D, Tang Q, Zhao K. Characterization of genome-wide enhancer-promoter interactions reveals co-expression of interacting genes and modes of higher order chromatin organization. Cell Res. 2012;22(3):490-503. Epub 2012/01/25. doi: 10.1038/cr.2012.15. PubMed PMID: 22270183; PMCID: PMC3292289.

26. Jin F, Li Y, Dixon JR, Selvaraj S, Ye Z, Lee AY, Yen CA, Schmitt AD, Espinoza CA, Ren B. A high-resolution map of the three-dimensional chromatin interactome in human cells. Nature. 2013;503(7475):290-4. Epub 2013/10/22. doi: 10.1038/nature12644. PubMed PMID: 24141950; PMCID: PMC3838900.

27. Carleton JB, Berrett KC, Gertz J. Multiplex Enhancer Interference Reveals Collaborative Control of Gene Regulation by Estrogen Receptor alpha-Bound Enhancers. Cell Syst. 2017;5(4):333-44 e5. Epub 2017/10/02. doi: 10.1016/j.cels.2017.08.011. PubMed PMID: 28964699; PMCID: PMC5679353.

28. Gertz J, Savic D, Varley KE, Partridge EC, Safi A, Jain P, Cooper GM, Reddy TE, Crawford GE, Myers RM. Distinct properties of cell-type-specific and shared transcription factor binding sites. Mol Cell. 2013;52(1):25-36. Epub 2013/10/01. doi: 10.1016/j.molcel.2013.08.037. PubMed PMID: 24076218; PMCID: PMC3811135.

29. Carroll JS, Liu XS, Brodsky AS, Li W, Meyer CA, Szary AJ, Eeckhoute J, Shao W, Hestermann EV, Geistlinger TR, Fox EA, Silver PA, Brown M. Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. Cell. 2005;122(1):33-43. Epub 2005/07/13. doi: 10.1016/j.cell.2005.05.008. PubMed PMID: 16009131.

30. Shakya A, Kang J, Chumley J, Williams MA, Tantin D. Oct1 is a switchable, bipotential stabilizer of repressed and inducible transcriptional states. J Biol Chem. 2011;286(1):450-9. Epub 2010/11/06. doi: 10.1074/jbc.M110.174045. PubMed PMID: 21051540; PMCID: PMC3013004.

31. Magne S, Caron S, Charon M, Rouyez MC, Dusanter-Fourt I. STAT5 and Oct-1 form a stable complex that modulates cyclin D1 expression. Mol Cell Biol. 2003;23(24):8934-45. Epub 2003/12/04. doi: 10.1128/mcb.23.24.8934-8945.2003. PubMed PMID: 14645506; PMCID: PMC309603.

32. Vazquez-Arreguin K, Tantin D. The Oct1 transcription factor and epithelial malignancies: Old protein learns new tricks. Biochim Biophys Acta. 2016;1859(6):792-804. Epub 2016/02/16. doi: 10.1016/j.bbagrm.2016.02.007. PubMed PMID: 26877236; PMCID: PMC4880489.

33. Maddox J, Shakya A, South S, Shelton D, Andersen JN, Chidester S, Kang J, Gligorich KM, Jones DA, Spangrude GJ, Welm BE, Tantin D. Transcription factor Oct1 is a somatic and cancer stem cell determinant. PLoS Genet. 2012;8(11):e1003048. Epub 2012/11/13. doi: 10.1371/journal.pgen.1003048. PubMed PMID: 23144633; PMCID: PMC3493455.

34. Ong CT, Corces VG. CTCF: an architectural protein bridging genome topology and function. Nat Rev Genet. 2014;15(4):234-46. Epub 2014/03/13. doi: 10.1038/nrg3663. PubMed PMID: 24614316; PMCID: PMC4610363.

35. Chandarlapaty S, Chen D, He W, Sung P, Samoila A, You D, Bhatt T, Patel P, Voi M, Gnant M, Hortobagyi G, Baselga J, Moynahan ME. Prevalence of ESR1 Mutations in Cell-Free DNA and Outcomes in Metastatic Breast Cancer: A Secondary Analysis of the BOLERO-2 Clinical Trial. JAMA Oncol. 2016;2(10):1310-5. Epub 2016/08/18. doi: 10.1001/jamaoncol.2016.1279. PubMed PMID: 27532364; PMCID: PMC5063698.

36. Savic D, Partridge EC, Newberry KM, Smith SB, Meadows SK, Roberts BS, Mackiewicz M, Mendenhall EM, Myers RM. CETChseq: CRISPR epitope tagging ChIP-seq of DNA-binding proteins. Genome Res. 2015;25(10):1581-9. Epub 2015/09/12. doi: 10.1101/gr.193540.115. PubMed PMID: 26355004; PMCID: PMC4579343.

37. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015;12(4):357-60. Epub 2015/03/10. doi: 10.1038/nmeth.3317. PubMed PMID: 25751142; PMCID: PMC4655817.

38. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078-9. Epub 2009/06/10. doi: 10.1093/bioinformatics/btp352. PubMed PMID: 19505943; PMCID: PMC2723002.

39. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30(7):923-30. Epub 2013/11/15. doi: 10.1093/bioinformatics/btt656. PubMed PMID: 24227677.

40. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550. Epub 2014/12/18. doi: 10.1186/s13059-014-0550-8. PubMed PMID: 25516281; PMCID: PMC4302049.

41. Tibes R, Qiu Y, Lu Y, Hennessy B, Andreeff M, Mills GB, Kornblau SM. Reverse phase protein array: validation of a novel proteomic technology and utility for analysis of primary leukemia specimens and hematopoietic stem cells. Mol Cancer Ther. 2006;5(10):2512-21. Epub 2006/10/17. doi: 10.1158/1535-7163.MCT-06-0334. PubMed PMID: 17041095.

42. Reddy TE, Pauli F, Sprouse RO, Neff NF, Newberry KM, Garabedian MJ, Myers RM. Genomic determination of the glucocorticoid response reveals unexpected mechanisms of gene regulation. Genome Res. 2009;19(12):2163-71. Epub 2009/10/06. doi: 10.1101/gr.097022.109. PubMed PMID: 19801529; PMCID: PMC2792167.

43. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10(3):R25. Epub 2009/03/06. doi: 10.1186/gb-2009-10-3-r25. PubMed PMID: 19261174; PMCID: PMC2690996.

44. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008;9(9):R137. Epub 2008/09/19. doi: 10.1186/gb-2008-9-9-r137. PubMed PMID: 18798982; PMCID: PMC2592715.

45. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26(6):841-2. Epub 2010/01/30. doi: 10.1093/bioinformatics/btq033. PubMed PMID: 20110278; PMCID: PMC2832824.

46. Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dundar F, Manke T. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. 2016;44(W1):W160-5. Epub 2016/04/16. doi: 10.1093/nar/gkw257. PubMed PMID: 27079975; PMCID: PMC4987876.

47. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods. 2013;10(12):1213-8. Epub 2013/10/08. doi: 10.1038/nmeth.2688. PubMed PMID: 24097267; PMCID: PMC3959825.

48. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 2009;37(Web Server issue):W202-8. Epub 2009/05/22. doi: 10.1093/nar/gkp335. PubMed PMID: 19458158; PMCID: PMC2703892.

49. Team RC. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna.2019.