Estrogen-independent molecular actions of mutant estrogen receptor alpha in endometrial cancer

Zannel Blanchard[1,2], Jeffery M. Vahrenkamp[1,2], Kristofer C. Berrett[1,2], Spencer Arnesen[1,2], Jason Gertz[1,2,*]

[1]Department of Oncological Sciences, University of Utah, Salt Lake City, UT 84112, USA
[2]Huntsman Cancer Institute, University of Utah, Salt Lake City, UT 84112, USA
*Corresponding author: jay.gertz@hci.utah.edu

Abstract

Estrogen receptor α (ER) mutations have been identified in hormone therapy resistant breast cancer and primary endometrial cancer. Analysis in breast cancer suggest that mutant ER exhibits estrogen independent activity. In endometrial cancer, ER mutations are associated with worse outcome and less obesity, however experimental investigation of these mutations has not been performed. Using a unique CRISPR/Cas9 strategy to introduce the most common mutation (D538G) while epitope tagging the endogenous locus, we discovered estrogen-independent mutant ER genomic binding that includes an expanded set of sites. The D538G mutation altered expression, including a large set of non-estrogen regulated genes, and chromatin accessibility, with most affected loci bound by mutant ER. Mutant ER is unique from constitutive ER activity as mutant-specific changes are not recapitulated with prolonged estrogen exposure. Overall, D538G mutant ER confers estrogen-independent activity while causing additional regulatory changes in endometrial cancer cells that are distinct from breast cancer cells.

Introduction

Estrogen receptor α (ER) is a ligand-inducible steroid hormone receptor that acts as an oncogene in many breast and endometrial tumors. In these diseases, hormone therapies can be used to reduce estrogen signaling either through a reduction in estrogen production or a reduction in ER activity. Mutations in the ligand binding domain (LBD) of ER have been associated with hormone therapy resistance in breast cancer (Fuqua et al., 2014; Fuqua, 1993; Jeselsohn et al., 2014; Osbourne, 2011; Robinson et al., 2013; Toy et al., 2013; Toy et al., 2017) and a recent large-scale analysis of ER mutations found that 14% of metastatic breast cancers harbor an LBD mutation (Toy et al., 2017). ER LBD mutations were not identified in primary tumors in The Cancer Genome Atlas' study on breast cancer (Network, 2012), indicating that ER mutations are not observed at near clonal frequencies and are unlikely to play a role in tumor initiation; however, ER LBD mutations can be found at low mutation frequencies in primary breast tumors (Toy et al., 2017). In contrast, heterozygous ER LBD mutations are found in 3.2% of primary endometrial cancers with endometrioid histology (Backes et al., 2016; Cancer Genome Atlas Research et al., 2013; Gibson et al., 2016), representing approximately 1,500 new uterine cancer diagnoses with an ER LBD mutation in the United States each year. The presence of ER mutations is associated with obesity-independent endometrial cancer and patients with ER LBD mutations trend towards worse prognosis when compared to patients with wildtype ER tumors (Backes et al., 2016).

The ER LBD mutations occur in a region of the protein essential for ligand binding and interactions with co-regulatory proteins, with the majority of mutations found at residues D538 and Y537. Studies into the molecular and phenotypic consequences of ER LBD mutations have been performed in breast cancer, revealing that the mutations confer estrogen-independent ER activity, which drives gene regulation and cell proliferation in the absence of estrogens (Bahreini, 2017; Jeselsohn et al., 2014; Jeselsohn, 2018; Merenbakh-Lamin et al., 2013; Robinson et al., 2013;

Toy et al., 2013; Toy et al., 2017; Zhao, 2017). Biochemical characterization of the mutations suggest that mutant ER favors the activated conformation of the receptor irrespective of ligand, causing constitutive receptor activity (Fanning, 2016; Katzenellenbogen, 2018; Merenbakh-Lamin et al., 2013; Toy et al., 2017; Zhao, 2017). Gene expression analyses highlight the ability of mutant ER to regulate canonical ER target genes in the absence of estrogens (Bahreini, 2017; Jeselsohn et al., 2014; Jeselsohn, 2018; Katzenellenbogen, 2018; Merenbakh-Lamin et al., 2013; Robinson et al., 2013; Toy et al., 2013; Toy et al., 2017). In addition to ligand-independent regulation of genes that are normally impacted by 17β-estradiol (E2), novel non-E2 regulated genes are also affected by mutant ER (Bahreini, 2017; Jeselsohn, 2018), suggesting that these mutations may confer additional functionality to ER than just constitutive activity. While these studies have uncovered important features of the molecular and phenotypic consequences of ER mutations in breast cancer, similar analyses have not been performed in endometrial cancer cells. Because gene expression responses to estrogens and ER genomic binding are highly dissimilar between breast and endometrial cancer (Droog, 2017; Gertz et al., 2013; Gertz, 2012), the impact of ER LBD mutations in endometrial cancer cells could be different than the effects observed in breast cancer cells.

In this study, we sought to gain an understanding of the molecular consequences of ER LBD mutations in endometrial cancer. Utilizing a CRISPR/Cas9 mediated epitope tagging strategy, we created endometrial cancer cells that are heterozygous for the D538G ER mutation (or wildtype ER for controls) coupled with a FLAG epitope tag incorporated at the C-terminus of the endogenous locus. The addition of an epitope tag to the endogenous gene allowed us to specifically analyze binding of the mutant form of ER by Chromatin immunoprecipitation followed by next-generation sequencing (ChIP-seq), which uncovered ligand-independent ER binding as well as mutant-enriched binding sites. We observed mutation-specific gene expression effects, with changes in E2-regulated genes and unanticipated E2-independent genes. An analysis of the chromatin landscape revealed significant changes to accessibility with novel regions becoming more open coincident with mutant ER binding, indicating a possible pioneering role for mutant ER. We also found that most of the regulatory effects of mutant ER are not recapitulated by prolonged exposure to E2. Together, our results indicate that mutant ER behaves like a constitutively active form of ER with some novel regulatory properties and that these models will enable future phenotypic and mechanistic investigation into ER mutant endometrial cancer.

Results

Generation of D538G ER mutant and wildtype cell lines

The molecular consequences of ER LBD mutations have not been explored in endometrial cancer and warrant investigation. We utilized a CRISPR/Cas9-mediated epitope tagging strategy called CETCH-seq (Savic et al., 2015) in Ishikawa cells, an endometrial adenocarcinoma cell line, to model the most common ER LBD mutation, D538G. This technique combines guide RNAs that target Cas9 to the C-terminus of ER and a donor plasmid that leads to the incorporation of a 3X FLAG epitope tag and neomycin resistance gene at ER's endogenous locus (Figure 1A). Ishikawa cells were first transfected with plasmids, treated with G418 to select for resistant cells and subjected to limiting dilution plating to generate single-cell clones. Using this technique, we generated three cell lines that were heterozygous for the D538G ER mutation and two FLAG-tagged wildtype cell lines. FLAG and ER protein expression were established via western blot (Figure 1B) and D538G mutations were confirmed by Sanger sequencing. Additionally, we verified equal expression frequency of the wildtype and mutant alleles in the D538G clonal cell lines by RNA sequencing (RNA-seq) (Figure 1C). The creation of these isogenic cell lines enabled further studies into the gene regulatory changes caused by the D538G ER mutation.

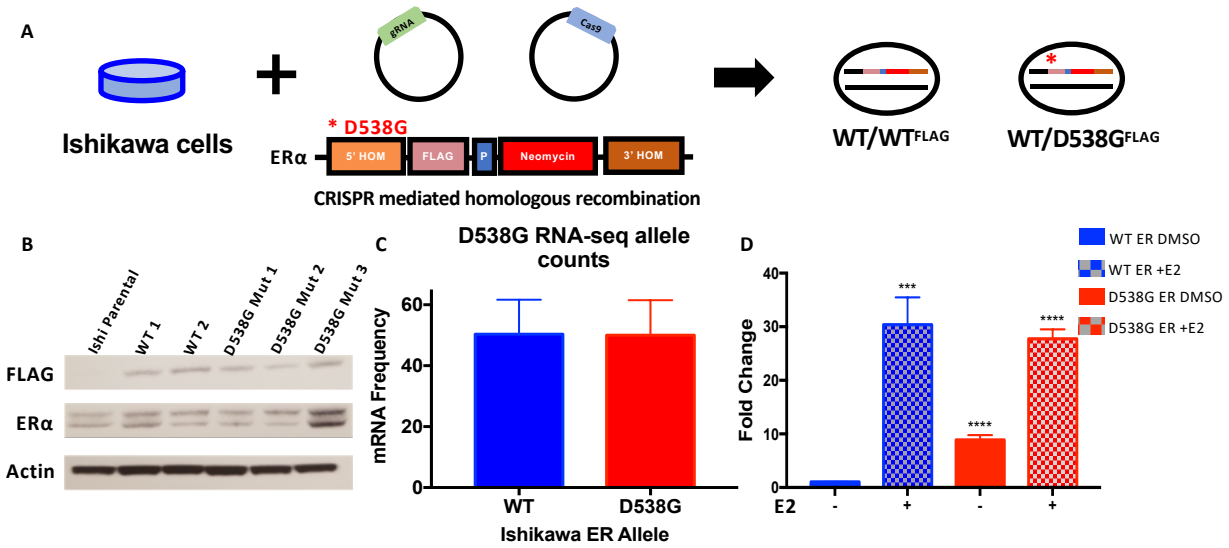

Figure 1. Generation and characterization of ER wildtype and D538G mutant models. A. CRISPR-mediated epitope tagging strategy was used to generate heterozygous FLAG-tagged wildtype ER and D538G mutant ER Ishikawa cell lines. **B**. Immunoblotting for FLAG and ER in Ishikawa parental cells, two heterozygous FLAG-tagged wildtype and three heterozygous FLAG-tagged D538G mutant cell lines show protein expression of epitope-tagged ER and total ER. **C**. The average ER wildtype and mutant allele expression levels based on RNA-seq data is shown across three D538G clonal cell lines. **D**. ERE reporter activity as measured by luciferase activity was assayed in DMSO and E2 induced conditions. Experiments were done in triplicate and the figure shows the average luciferase activity for two wildtype and three D538G mutant clones. (*** $p = 0.0002$, ****$p < 0.0001$). Error bars represent ± SEM.

D538G mutant ER displays ligand-independent regulatory activity

To assess ER's transcriptional activity in our models, we cultured wildtype and D538G mutant cells in hormone deprived media for 5 days and then transfected these cell lines with a luciferase estrogen response element (ERE) reporter assay. One day post transfection, cells were treated with either vehicle (DMSO) or 10 nM E2 for 24 hours before measuring luciferase activity. We observed negligible luciferase expression in the wildtype cells treated with DMSO (Figure 1D). However, there was a significant 30-fold increase in expression in wildtype cells induced with E2. We detected ER activity in the D538G mutant clones treated with DMSO at 30% of the wildtype E2 levels. Activity was significantly increased in the D538G mutant lines treated with E2, with levels similar to the wildtype lines treated with E2. These results suggest ligand-independent transcriptional activity of the D538G mutant in endometrial cancer cell lines.

Introduction of the D538G ER mutation causes large transcriptional changes

The ligand-independent transcriptional activity of mutant ER in the reporter assay suggested that transcription in these mutant lines could be altered, leading to aberrant gene expression. To uncover gene expression changes caused by the mutation, we performed RNA-seq on the wildtype and D538G mutant cell lines following an 8 hour 10 nM E2 induction in hormone deprived media. Principal component analysis of the isogenic lines clustered wildtype and mutant lines separately, with the first principal component accounting for 46% of the variance in our datasets while separating samples based on the presence of the mutation (Figure S1A). Treatment with E2 along with differences between clones with the same genotype accounted for 29% of the variance in these lines and is represented in the second principal component, highlighting the importance of analyzing multiple clones in order to capture the variance in clone derivation. Our analysis indicates that the gene expression profile of wildtype cells supplemented with E2, did not recapitulate the expression changes seen in the D538G mutant lines.

Analysis of differentially expressed genes across clonal cell lines revealed multiple expression patterns (Figure 2A). We identified 119 genes that were upregulated and 48 genes that were downregulated in response to an estrogen induction in wildtype cells (false discovery rate < 0.05). A comparison of these E2 responsive genes to genes regulated in the D538G DMSO treated lines, revealed 47 upregulated and 10 downregulated genes that were responsive to estrogen in wildtype cells and also regulated by the D538G mutation independent of E2. Examples of mutant ER ligand-independent regulation of estrogen responsive genes include progesterone receptor (*PGR*) known to play critical roles in reproductive function (Figure 2B) and an adhesion G-protein coupled receptor (*ADGRF1*) which functions in transmembrane signaling (Figure 2C). The estrogen-independent regulation of these genes by mutant ER is consistent with the reporter assays and the hypothesis that mutant ER has ligand-independent activity.

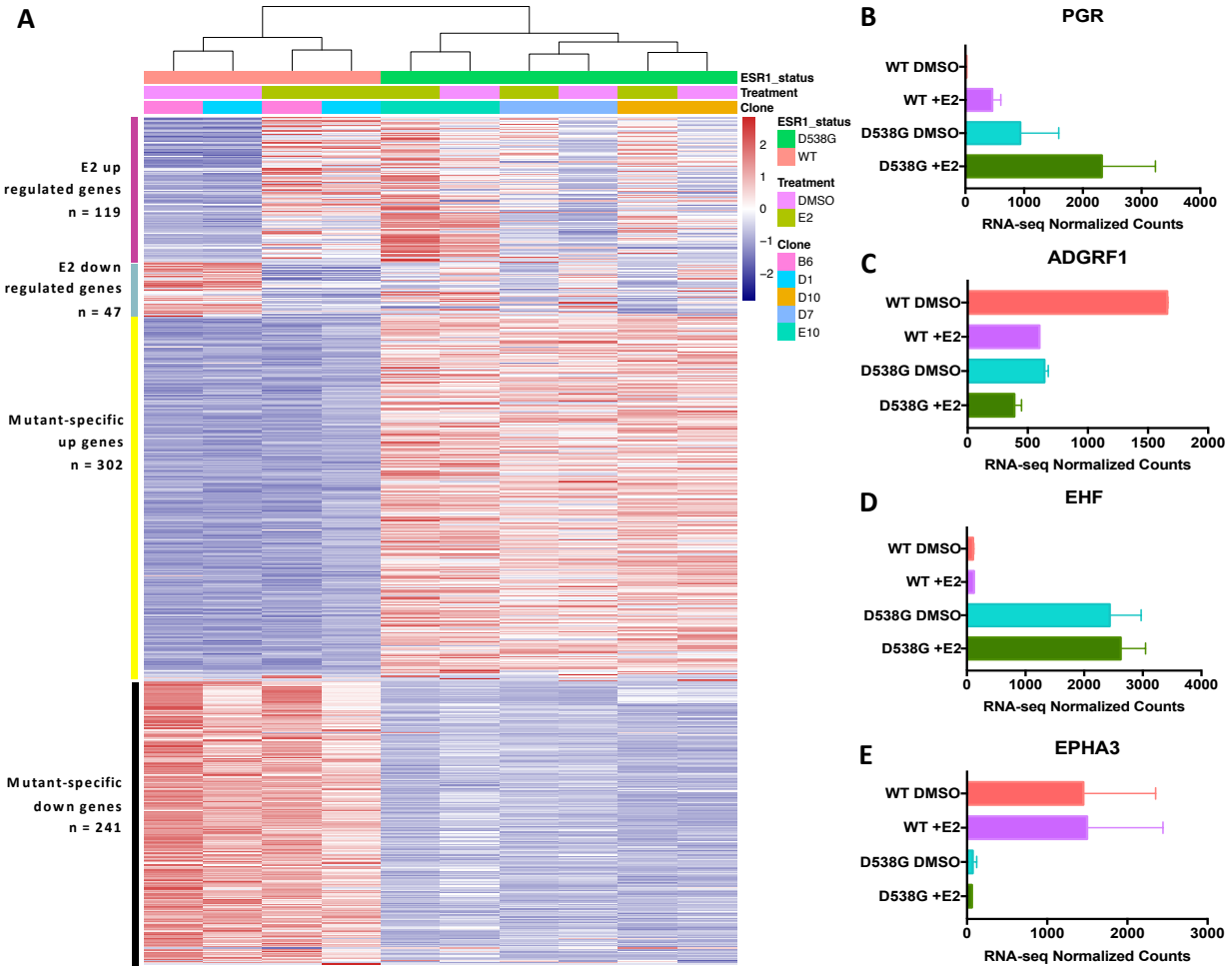

Figure 2. D538G mutant drives a distinct transcriptional program. A. Heatmap shows the relative expression of E2 up- and down-regulated genes as well as novel, mutant-specific differentially expressed genes. Sample types are indicated by the column annotations described in the legend. Examples of ligand-independent E2 upregulated gene *PGR* **(B)** and downregulated gene *ADGRF1* **(C)**. Examples of novel mutant-specific upregulated gene *EHF* **(D)** and downregulated gene *EPHA3* **(E)**. Figures show average RNA-seq normalized counts for two ER LBD wildtype and three D538G mutant clones.

In addition to changes in estrogen regulated genes, the D538G mutation impacted the expression of many genes not normally regulated by estrogen. We identified mutation-specific gene expression changes with 302 novel genes that are upregulated and 241 novel genes that are downregulated due to the mutation (see Table S2 for mutant-specific genes). Examples of novel

genes include *EHF*, an ETS factor not normally expressed in Ishikawa cells (Figure 2D) and *EPHA3*, a receptor tyrosine kinase (Figure 2E). Ingenuity Pathway Analysis (IPA) indicated that novel genes were enriched for pathways associated with more aggressive tumors, which includes cellular growth, proliferation and movement (Figure S1B,C). This pathway enrichment is similar to findings from studies done on ER mutations in MCF-7 and T47-D breast cancer cell lines (Bahreini, 2017). While the pathways are consistent with novel gene regulation observed in breast cancer cells, there is very little overlap in the differentially expressed genes (MCF-7: 29 genes (5.3%); T-47D 20 genes (3.7%)). Collectively, our data indicates that the D538G mutation impacts many non-E2 regulated genes and enables a more expansive and potentially aggressive transcription program.

D538G mutation alters ER genomic binding

To ascertain ER genomic binding sites in Ishikawa cells and the manner in which the D538G mutation alters ER's genomic interactions, we performed ChIP-seq with an antibody that recognizes the FLAG-epitope tag in the two wildtype and three mutant clones after 1 hour treatments with DMSO or E2. We identified 51 and 45 peaks in the two wildtype lines treated with DMSO, indicating negligible ER binding in the absence of E2 treatment. We observed a significant increase in genomic binding, with 5,000 – 8,000 ER binding sites identified in the two wildtype cell lines following E2 induction. An analysis of the overlap of binding sites called in the wildtype clones treated with E2 showed 92% concordance between lines, highlighting the reproducibility of our experimental findings (Figure S2A). To confirm that the FLAG-tag did not significantly alter ER genomic binding, we overlapped the binding sites discovered with the FLAG antibody to binding sites from a ChIP-seq previously done in Ishikawa cells with an antibody that recognizes ER (Gertz et al., 2013; Gertz, 2012). We observed 74% - 83% overlap with this previously generated ER ChIP-seq dataset, indicating that the epitope tag does not drastically affect ER binding in the ER LBD cell lines (Figure S2B,C). In contrast, we observed ligand-independent D538G mutant ER binding at more than 9,800 binding sites in the three mutant lines in the absence of E2. A comparison of D538G mutant ER bound sites indicated more than 70% overlap between lines, suggesting that the clonally derived mutant lines share the majority of their ER binding sites (Figure S2D).

Differential binding analysis established 8,398 constant ER binding sites that were bound by wildtype ER and the D538G mutant, indicating that most loci bound by wildtype ER are also bound by mutant ER (98%) (Figures 3A, S3). However, mutant ER bound to an increased number of genomic sites with 7,670 binding sites that were enriched in the D538G mutant lines. In contrast, only 144 sites were enriched in the wildtype lines, suggesting that the mutation is able to expand ER's genomic targets by binding to thousands of novel sites in the genome (Figure 3B). *De novo* motif analysis of mutant-specific binding sites revealed a significant enrichment for the canonical ERE (p-value = 9.1×10^{-103}, MEME), (Bailey, 2009), indicating that many of these novel sites are direct targets of D538G mutant ER (Figure S2E). Additional motifs were also enriched at these sites including the zic family of transcription factors (zic1/2/3; p-value = 2.0×10^{-65}, MEME) (Bailey, 2009)(Bailey, 2009) and ebox binding proteins (myod1; p-value = 8.4×10^{-23}, MEME) (Bailey, 2009). We examined the strength of EREs found in constant ER binding sites and mutant-enriched ER binding sites to ascertain if mutant ER was capable of binding to sites with sub-optimal/weak EREs, which includes half-sites. Our analysis revealed a two-fold enrichment for strong EREs in the constant sites as opposed to mutant-enriched sites (p-value < 2.2×10^{-16}, Wilcoxon; Figure 3C), suggesting that mutant ER binds to an expanded set of genomic loci, including sites with suboptimal EREs.

To explore the connection between mutant-enriched ER binding and gene expression, we analyzed the distance between mutant-enriched ER binding sites and the transcription start sites (TSS) of genes upregulated, downregulated or not regulated by the mutation. We found that

mutant-enriched ER binding was significantly closer to mutant upregulated (p-value = 4.1×10^{-14}, Wilcoxon) and downregulated genes (p-value = 8.2×10^{-13}, Wilcoxon) than control genes (Figure 3D). This analysis indicates that mutant-specific gene expression appears to be driven in part by direct mutant ER binding.

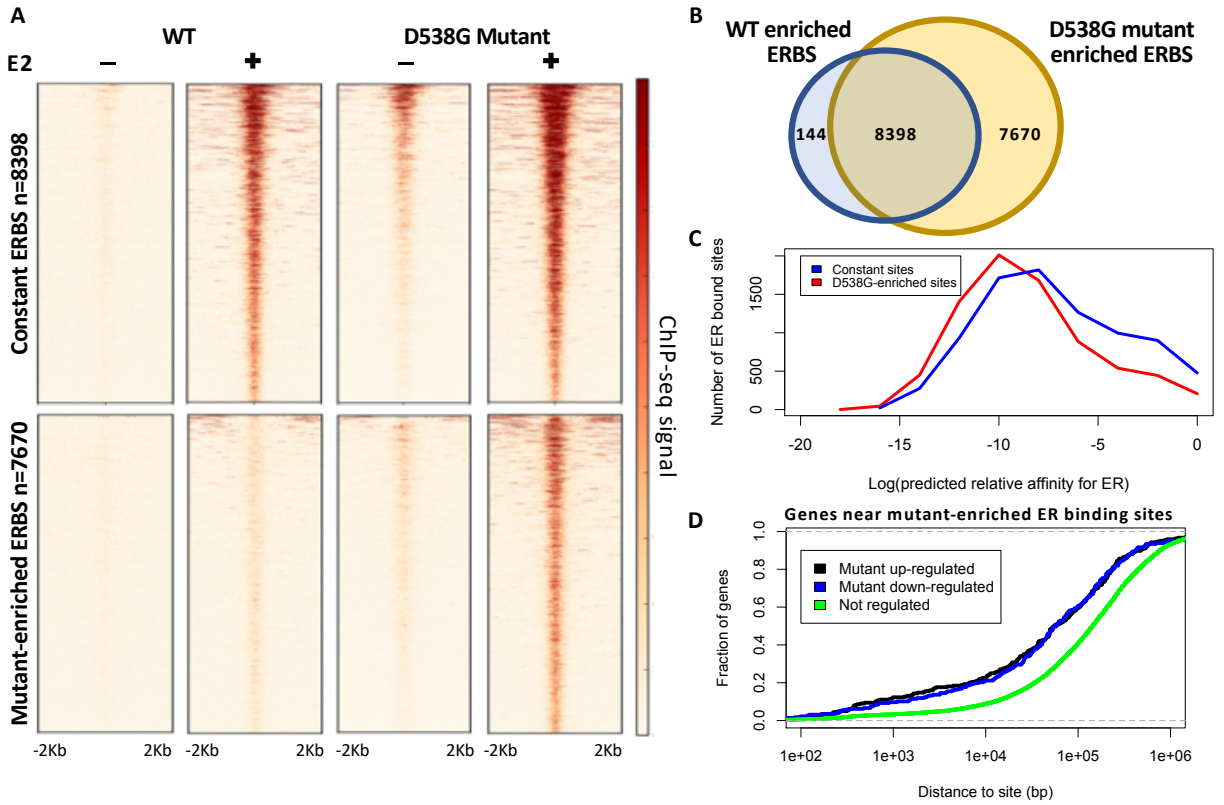

Figure 3. D538G mutation alters ER genomic binding. A. Heatmap displays ER binding in representative wildtype and D538G mutant clones, where each row is a site. The heatmaps include ER binding sites that are constant in wildtype and mutant lines (top panels) and sites that are enriched in the mutant lines (bottom panels). **B.** Venn diagram shows the overlap between FLAG-tagged ER binding sites (ERBS) called in ER LBD wildtype and D538G mutant lines. **C.** The distribution of the predicted relative affinity for ER, based on the best match to an ERE, is higher in constant ER binding sites (blue) than D538G mutant-enriched ER binding sites (red). **D.** Mutant-enriched ER binding sites are found within a 100 kb window of the TSS of 60% of mutant-specific upregulated (black line) and downregulated (blue line) genes. This is in comparison to only 30% of genes not regulated by the D538G mutation (green line) that have mutant-enriched ER binding nearby.

D538G ER mutation leads to changes in accessible chromatin

The mutant-specific gene expression and mutant-enriched ER binding at suboptimal EREs led us to hypothesize that the D538G mutation may affect chromatin accessibility. To test this hypothesis, we performed the Assay for Transposase Accessible Chromatin followed by sequencing (ATAC-seq) in the wildtype and D538G mutant cell lines, after a 1 hour treatment with DMSO or E2. Similar to the distinct RNA-seq profiles, principal component analysis of the isogenic lines once again clustered wildtype and mutant lines separately (Figure 4A). The first principal component accounted for 60% of the variance in the ATAC-seq data and separated wildtype and D538G mutant lines accordingly. The second principal component, which accounted for 18% of the variance, separated clones, while E2 treatment was not a large contributor to the variance. This analysis suggests that a 1 hour E2 induction does not have major genome-wide effects on chromatin accessibility in comparison to the D538G ER mutation.

Analysis of ATAC-seq signal identified 881 regions that were more accessible and 161 regions that were less accessible in the D538G mutant cell lines. Examples of these mutant-enriched and mutant-depleted ATAC-seq regions are shown in Figures 4C and 4D. We found that ER bound sites were over-represented in mutant-enriched ATAC-seq sites. In total, 17% of open chromatin regions are associated with ER binding. However, 80% of the mutant-enriched ATAC-seq sites are associated with ER binding and this includes both constant (42%) as well as mutant-enriched ER binding sites (39%) (Figure 4B). These results suggest that mutant ER plays a major role in mutant-enriched chromatin accessibility, possibly owing to mutant ER's constitutive activity. The presence of mutant-enriched accessible chromatin that is ER bound is consistent with a model where mutant ER has some pioneering activity.

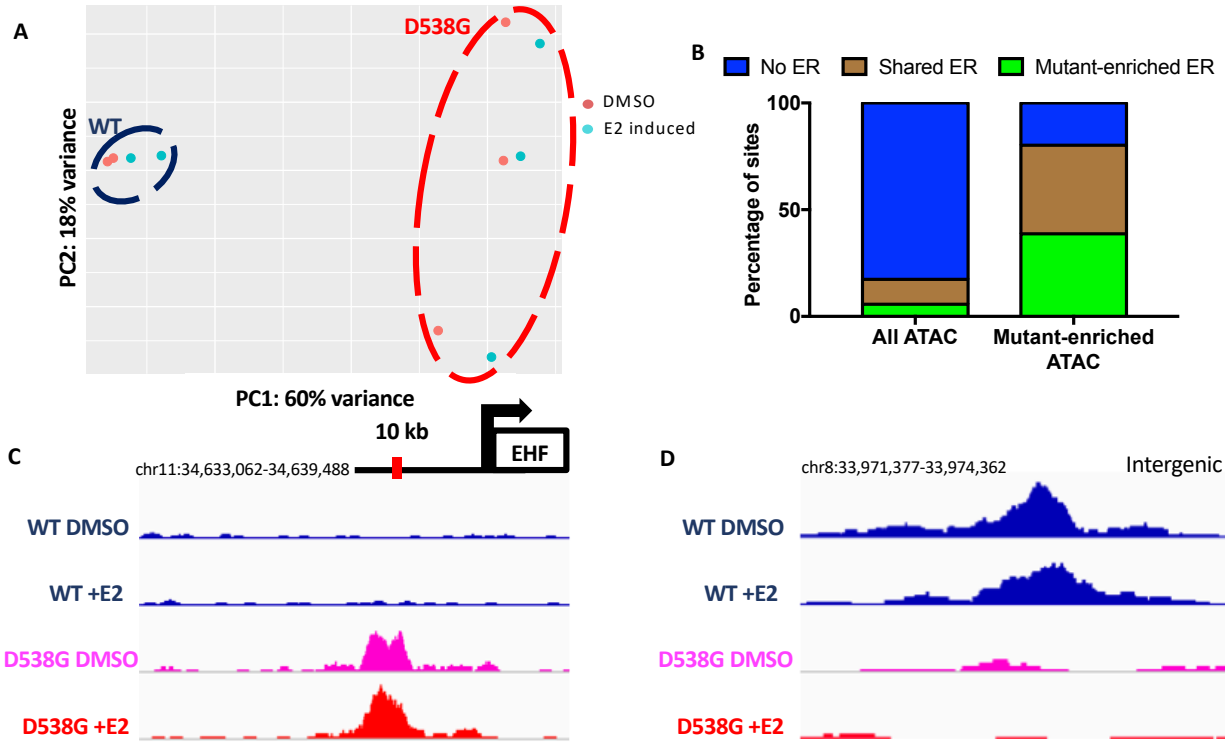

Figure 4. ER D538G mutation alters chromatin accessibility at multiple loci. A. Principal component analysis shows the relationship between ATAC-seq signal of ER wildtype (blue circle) and D538G mutant cell lines (red circle). **B.** 80% of mutant-enriched ATAC-seq sites overlap ER binding sites, including shared ER binding (brown) and mutant-enriched ER binding (green). Representative browser tracks indicating ATAC-seq signal increase with the D538G mutation at a region near *EHF* **(C)** and ATAC-seq signal decrease due to the D538G mutation at an intergenic region on chromosome 8 **(D)**. Wildtype ATAC-seq signal DMSO/+E2 (blue), D538G DMSO (pink) and D538G +E2 (red) are scaled to the same value at each locus.

Prolonged exposure to estrogen does not recreate mutant-specific regulatory effects

The mechanism by which mutant ER regulates a novel set of genes may be explained by constitutive ER activity or neomorphic functions conferred by the D538G mutation. To determine how much of the gene regulatory effects of mutant ER are due to constitutive ER activity, we grew the two wildtype clones in the presence of 10 nM E2 for a 25 day period and performed RNA-seq and ATAC-seq on cells collected at 5 day intervals during this prolonged exposure. Differential gene expression analysis identified 658 genes that were upregulated and 1,138 genes that were downregulated in prolonged E2 treated wildtype cells compared to untreated wildtype cells (Figure 5A). These numbers suggest there are more genes regulated by prolonged exposure to E2 than the transient 8 hour induction; however, the overlap with mutant-specific gene expression changes was minimal. Only 25 genes (8.3%) that were upregulated in response to prolonged E2

were also upregulated by the mutation (Figure 5B, examples in Figure S4). Additionally, only 34 downregulated genes (14%) showed overlap with mutant-specific downregulated genes (Figure 5C). These findings suggest that the D538G mutant regulates a ligand-independent transcriptional program that is dissimilar to prolonged E2 exposure in endometrial cancer cells.

To determine if constitutive ER activity could re-create the mutant-specific chromatin accessibility patterns, we compared the ATAC-seq results from the prolonged E2 experiment in wildtype cells to the results from the D538G mutant clones. Principal component analysis of all ATAC-seq samples identified three distinct clusters: DMSO controls along with short-term E2 treated wildtype cells (1 hour), prolonged E2 treated wildtype cells (10 days, 15 days, 20 days, 25 days) and the D538G mutant lines (Figure 5D). While Figure 4A indicates that the 1 hour E2 induction did not significantly affect chromatin accessibility globally, prolonged E2 exposure is able to affect chromatin accessibility, with some features being shared with the mutation. However, since prolonged E2 cells do not cluster with the mutant lines, we can conclude that these features do not recapitulate most of the mutant-specific genome-wide effects on chromatin accessibility.

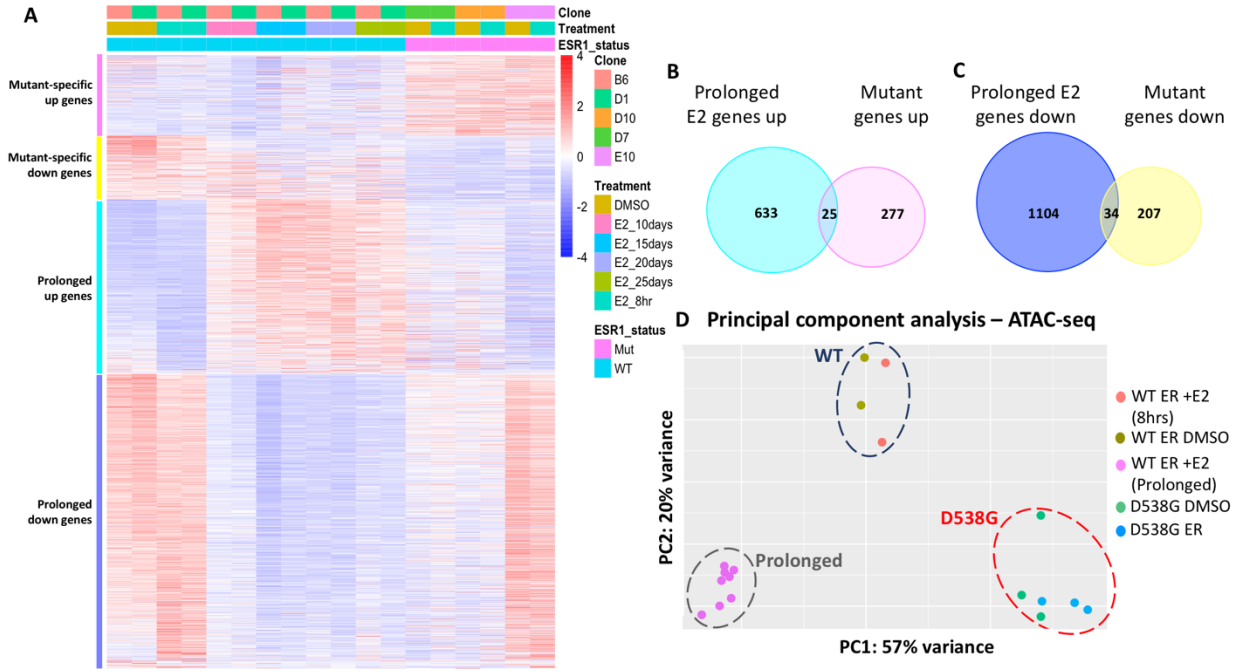

Figure 5. Prolonged E2 exposure does not recapitulate D538G mutant regulatory consequences. A. Heatmap shows the relative expression of novel, mutant-specific differentially expressed genes as well as genes up- and down-regulated in response to prolonged E2. Samples are indicated by the column annotations described in the legend. **B.** Venn diagram shows the overlap between genes upregulated in wildtype lines exposed to prolonged E2 and mutant-specific upregulated genes. **C.** Venn diagram shows the overlap between genes downregulated in wildtype lines exposed to prolonged E2 and mutant-specific downregulated genes. **D.** Principal component analysis of ATAC-seq signal exhibits three distinct sample groups: wildtype lines with short-term or no E2 treatment (blue circle), wildtype lines with prolonged E2 exposure (greys circle), and D538G mutant lines (red circle).

The ATAC-seq analysis revealed 1,488 regions that became more accessible and 2,915 regions that were less accessible following prolonged E2 exposure in wildtype cells. We overlapped regions that are more accessible after prolonged E2 exposure and regions that are more accessible in the D538G mutant clones, as described above, and identified only 77 regions that were common between datasets, which represented 8.7% of the mutant-enriched regions. Similarly, there were 44 loci that overlapped between chromatin that becomes less accessible in response to prolonged E2 treatment and mutant-depleted accessible chromatin, representing 27% of the mutant-depleted regions (examples in Figure S4). Our findings indicate that prolonged

E2 is unable to recapitulate the observed ER mutant driven changes to chromatin accessibility. Collectively, the effects on gene expression and chromatin accessibility both suggest that the D538G mutation confers distinct, neomorphic properties to the receptor which cannot be adequately explained by constitutive estrogen signaling.

Discussion

Through the creation of isogenic models of the most common ER mutation, D538G, we have found that mutant ER exhibits estrogen-independent activity in endometrial cancer cells. Similar to findings in breast cancer (Bahreini, 2017; Jeselsohn et al., 2014; Jeselsohn, 2018; Merenbakh-Lamin et al., 2013; Robinson et al., 2013; Toy et al., 2013), mutant ER is able to bind the genome and drive the transcription of estrogen-responsive genes in the absence of estrogens. The estrogen-independent activity of mutant ER that we observed is consistent with the clinical observation that endometrial cancer patients with ER mutations have lower body mass index than patients without ER mutations (Backes et al., 2016). Adipose tissue, which is more prevalent in obese patients, is capable of peripheral estrogen production (Siiteri, 1987), but endometrial tumors with ER mutations appear not to rely on this excess estrogen, presumably due to constant ER activity.

ER binds to different loci in breast cancer and endometrial cancer cells (Gertz et al., 2013) and primary tumors (Droog, 2017) leading to different transcriptional responses to E2. Mutant ER exhibits a similar cell type-specific pattern in which the genes regulated by mutant ER are different between endometrial cancer cells and breast cancer cells. This is true for both estrogen independent regulation of normally estrogen responsive genes as well as novel regulation by mutant ER of non-estrogen responsive genes. These results suggest that mutant ER binding site and target gene selection is still constrained by the other transcription factors and co-factors expressed in the cell as well as the chromatin landscape. Although different genes are regulated by mutant ER in breast and endometrial cancer cells, similar pathways including cellular growth, proliferation, and movement, are affected by the mutations suggesting that ER mutations might cause similar phenotypes in breast and endometrial tumors.

Mutant ER does not just confer ligand-independent estrogen signaling. In fact, most of the genes differentially expressed between the mutant lines and the wildtype lines are genes that do not normally respond to E2. This mutant-specific gene regulation appears to be directed by mutant ER with the D538G mutation causing a large increase in the number of ER genomic binding sites that are significantly enriched near mutant-specific regulated genes. Chromatin accessibility also increases at specific sites across the genome due to the D538G mutation and a large majority of these sites are bound by mutant ER. These results suggest a small pioneering role for mutant ER; although, we have not shown this directly. Mutant ER could be increasing chromatin accessibility in several ways, including changes in co-factor recruitment, as found in breast cancer (Jeselsohn, 2018), or because of its constant activity and binding.

In order to determine how much of mutant ER's ability to regulate a new set of genes is related to its constant activity, we treated wildtype clones with continuous saturating doses of E2 for 25 days. Prolonged exposure to E2 changed the expression of thousands of genes and altered chromatin accessibility at thousands of loci; however, there was little overlap with the gene regulatory changes caused by the D538G mutation. These results suggest that the D538G ER mutation is neomorphic/gain-of-function and does not simply cause hyperactivity. It is unclear how the mutation changes ER's gene regulatory role, but alterations to the placement of helix 12 of the ligand binding domain (Fanning, 2016; Merenbakh-Lamin et al., 2013) could cause changes in binding affinities to transcription factors or cofactors that bind to this region. Determining how mutant ER causes novel gene regulation could provide valuable insights into treatment strategies

aimed at blocking mutant ER's activity. In summary, our study has led to the creation of isogenic models of mutant ER in endometrial cancer cells, the confirmation of estrogen-independent mutant ER activity, and the discovery of novel gene regulation through mutant ER that cannot be explained by constant activity alone.

Materials and Methods

Plasmid construction for ER LBD mutant generation

Mutant cell lines were created using the CETCH-seq method (Savic et al., 2015) in which a pFETCH plasmid is the homology donor (containing the mutation, 3x FLAG tag, P2A linker, and Neomycin resistance gene) and Cas9 is targeted proximal to the stop codon by guide RNAs. To create the pFETCH homology donor plasmid with the D538G ER LBD mutation, we used primers (see Table S1 for sequences) to PCR amplify ER homology arms, using a gBlock (IDT, Table S1) that encompassed 1000 bp surrounding the D538G mutation as a template for the amplification of homology arm 1 and genomic DNA from Ishikawa cells (Sigma-Aldrich) as the template for homology arm 2. For amplification, Phusion high fidelity master mix (New England BioLabs) was used with 10 μM of each primer and 1 ng gBlock or 50 ng genomic DNA and amplified for 25 cycles. To generate the wildtype pFETCH donor plasmid, we repeated this technique with Ishikawa genomic DNA as template for arm 1 amplification. Using BsaI and BbsI (New England BioLabs), we double digested the destination pFETCH vector (Addgene 63934, a gift from Eric Mendenhall & Richard M. Myers) and used Gibson assembly HiFi master mix (New England BioLabs) to clone homology arms 1 and 2 simultaneously to create D538G or wildtype pFETCH plasmids. Clones for each pFETCH vector underwent minipreps (Zymo Research) and were verified by Sanger sequencing (Genewiz). In order to create a Cas9 and guide RNA expressing plasmid that targeted near the stop codon of ER, a Cas9 and guide RNA expression vector (Addgene 62988, a gift from Feng Zhang) was digested with BbsI (New England BioLabs). Guide RNA oligos (Table S1) were annealed and then ligated into the Cas9 guide RNA vector. Clones for each guide RNA underwent minipreps (Zymo Research) and were verified by Sanger sequencing (Genewiz). The pFETCH mutant and wildtype plasmids were SalI (New England BioLabs) digested prior to transfection to linearize the vector.

Cell culture and transfection for generation of ER LBD mutant and wildtype lines

Ishikawa cells (Sigma-Aldrich) were seeded in 6-well plates at a density of 300,000 cells per well in RPMI-1640 (Thermo-Fisher Scientific) with 10% fetal bovine serum (Thermo-Fisher Scientific) and 1% penicillin-streptomycin (Thermo-Fisher Scientific). At approximately 50% confluency, 250 ng of each of the two Cas9 guide RNA vectors and 250 ng of either the mutant or wildtype pFETCh vectors were transfected into cells with Lipofectamine 3000 (Thermo-Fisher Scientific). Cells were also treated with 1 μM SCR7 (Xcessbio) to inhibit non-homologous end joining (NHEJ) for 3 days post-transfection. 72 hours post transfection, the media was changed and G418 (Thermo-Fisher Scientific) was added to a final concentration of 200 μg/mL. RPMI media and G418 were replaced every 2 days until resistant cells remained. To generate single-cell colonies, transfected cells were plated at limiting dilution and cultured until colonies were large enough to identify visually. Individual colonies were picked and transferred to a 24-well plate and grown until they reached confluency. At this time, genomic DNA was extracted from individual colonies using the DNeasy Blood and Tissue Kit (Qiagen). To genotype clones, we amplified a 700 bp region within the ligand binding domain of ER (Table S1). PCR products were purified with Ampure XP beads (Beckman Coulter) and Sanger sequenced (Genewiz) to confirm the D538G mutation or wildtype sequence. Positive clones then underwent immunoblotting (see below) to verify protein expression.

Immunoblotting

Ishikawa ER LBD mutant and wildtype cell lines were seeded in 100mm dishes with RPMI-1640 with 10% fetal bovine serum and 1% penicillin-streptomycin until they reached confluency. At confluency, cells were washed with cold phosphate buffered saline (PBS), adherent cells were scraped, lysed with RIPA buffer (1X PBS, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) supplemented with protease inhibitors (Thermo-Fisher Scientific) and sonicated with an Active Motif EpiShear probe-in sonicator with 3 cycles of 10 seconds on, 10 seconds of rest at 40% amplitude. 80 μg of protein was loaded onto a 4-12% Bis-Tris gel (Thermo-Fisher Scientific) and electrophoresed for 90 minutes at 130V in 1X MOPS buffer (Thermo-Fisher Scientific). Proteins were transferred onto a polyvinylidene difluoride (PVDF) membrane with the iBlot system (Thermo-Fisher Scientific). Membranes were blocked in 5% milk/PBST for 1 hour at room temperature. Blots were probed with primary antibodies to Anti-FLAG 1:1,000 (Sigma Aldrich M2), ER α 1:200 (Santa Cruz HC-20) and β-Actin 1:1,000 (Santa Cruz C4) in 2.5% milk/PBST overnight at 4 °C. Membranes were washed with PBST and incubated in secondary antibody (Goat anti-Mouse IgG, Goat anti-Rabbit IgG; Thermo-Fisher Scientific) in PBST at 1:5,000 dilution. Protein signal was detected with SuperSignal West Femto Maximum Sensitivity Substrate (Thermo-Fisher Scientific).

Cell culture

Ishikawa ER LBD mutant and wildtype cells were cultured in RPMI-1640 with 10% fetal bovine serum and 1% penicillin-streptomycin. Cells were incubated at 37°C with 5% CO2 for the duration of all experiments. At least 5 days prior to inductions, cells were placed in phenol-red free RPMI-1640 media (Thermo-Fisher Scientific), since phenol-red is estrogenic, with 10% charcoal-dextran stripped fetal bovine serum (Thermo-Fisher Scientific) and 1% penicillin-streptomycin. Media was changed 1 day prior to inductions and cells were treated with either DMSO (vehicle) or 10 nM E2 for 1 hour for ChIP-seq and ATAC-seq experiments, or 8 hours for RNA-seq experiments. For prolonged estrogen RNA-seq and ATAC-seq experiments, mutant and wildtype cell lines were cultured in phenol-red free RPMI with 10% charcoal-dextran stripped fetal bovine serum and 1% pen-strep for up to 25 days. Cells were treated with 10 nM E2 every 2 days, and cell lysates were collected at 5 day intervals, starting at day 10.

ERE luciferase reporter assay

Ishikawa ER LBD mutant and wildtype cells were cultured in phenol red-free RPMI with 10% charcoal-dextran stripped fetal bovine serum and 1% pen-strep for 4 days. Approximately 15,000 cells per well for each cell line were seeded in a 96-well plate in hormone-deprived media on day 5. Cells were then transfected, according to the manufacturer's protocol with inducible dual-luciferase and renilla estrogen response element (ERE) constructs (Qiagen) using FuGENE HD transfection reagent (Promega). One day post transfection, media was changed and cells were treated with either DMSO or 10 nM E2. 24 hours post treatment, luciferase activity was measured using the Dual-Glo luciferase assay system (Promega) as per manufacturer's instructions. All experiments were performed in triplicate with three biological replicates. Statistical analysis was performed using Student's t-test.

ChIP-seq

Cells were induced with DMSO or 10 nM E2 for one hour, followed by fixation with 1% formaldehyde for 10 minutes at room temperature to cross-link cells. The cross-linking reaction was stopped with the addition of glycine to a final concentration of 0.125 M. Cells were washed with cold PBS and harvested via cell scraping in Farnham lysis buffer supplemented with protease inhibitors. Chromatin immunoprecipitation was performed as previously described (Reddy, 2009) with an Anti-FLAG (Sigma Aldrich M2) antibody. ChIP-seq libraries were sequenced on an

Illumina HiSeq 2500 and sequencing reads were aligned to the hg19 build of the human genome using bowtie (Kim, 2015) with the following parameters: -m 1 -t --best -q -S -l 32 -e 80 -n 2. Macs2 (Zhang, 2008) was used to call peaks with a p-value cutoff of 1e-10 and a mfold parameter between 15 and 100. Input control libraries from both wildtype and mutant cell lines were used as controls for each ChIP experiment. To identify mutant specific novel binding sites, regions needed to be bound by ER in two out of three mutant FLAG ChIPs treated with E2 but none of the wildtype E2 FLAG ChIP experiments. Motif finding was performed on 500bp fragments surrounding the summit of identified peaks. Motifs between 6 and 30 bases in length were identified by MEME Suite (Bailey, 2009), with a motif distribution of zero to one occurrence per sequence. To identify strong and weak EREs in mutant-specific and constant ER binding sites, ERE scores were calculated by Patser (Hertz, 1999), within 100 bp of an individual peak's summit.

RNA-seq

Following eight hour treatments in hormone deprived media with either 10 nM E2 or DMSO, cells were washed with PBS and harvested with buffer RLT plus (Qiagen) containing 1% beta-mercaptoethanol. Cells were passed through a 21-gauge needle and syringe (Sigma-Aldrich) to lyse genomic DNA before RNA was extracted and purified using a Quick RNA Mini Prep kit (Zymo Research). Using the KAPA Stranded mRNA-Seq kit (KAPA Biosystems), poly(A) selected libraries were created with 500 ng of input RNA per sample. RNA-seq libraries were sequenced on an Illumina HiSeq 2500 and sequencing reads were aligned to the hg19 build of the genome using HISAT2 (Kim, 2015), with RefSeq gene definitions used to build indexes. Sam files were converted to bam files and sorted with Samtools (Li, 2009). To quantify reads that mapped to genes, we used featureCounts (Liao, 2014). Reads were normalized and analyzed for differential enrichment using the DEseq2 package for R (Love, 2014). To parse differentially regulated genes, statistically significant E2 upregulated and downregulated genes in wildtype cells were first identified. Ligand independent E2 regulated candidate genes were defined as genes that are differentially expressed in response to the D538G mutant as well as responsive to E2 inductions in wildtype cells. To identify novel genes not normally regulated by E2, we compared gene expression in the three mutant lines to gene expression in the two wildtype cell lines and then subtracted genes that were regulated by E2 in wildtype cells. The adjusted p-value cutoff for significant genes was <=0.05. Differentially expressed novel upregulated (n=302) and downregulated (n=241) genes were analyzed via IPA (QIAGEN). The five statistically significant molecular and cellular functions identified in this analysis are included in the supplemental data (Figure S1B,C).

ATAC-seq

After a one hour treatment in hormone deprived media with either 10 nM E2 or DMSO, 250,000 cells were isolated from ER LBD wildtype or mutant cell lines and ATAC-seq was performed as previously described (Buenrostro, 2013). ATAC-seq libraries were sequenced on an Illumina HiSeq 2500 and sequencing reads were aligned to hg19 using bowtie (Kim, 2015) with the following parameters: -m 1 -t --best -q -S -l 32 -e 80 -n 2. Sam files were converted to bam files and sorted with Samtools (Li, 2009). Macs2 (Zhang, 2008) was used to call peaks without an input control but with a p-value cutoff of 1e-10. We used featureCounts (Liao, 2014) to quantify reads that aligned to each ATAC-seq peak. Reads were normalized and analyzed for differential enrichment using the DEseq2 package for R (Love, 2014). Mutant-enriched ATAC-seq sites were identified by comparing ATAC-seq signal between all wildtype and all mutant samples.

Author Contributions

Conceptualization, Z.B. and J.G.; Methodology, Z.B. and J.G.; Investigation, Z.B., K.C.B., S.A. and J.G.; Formal Analysis, Z.B., J.M.V. and J.G.; Writing – Original Draft, Z.B. and J.G.; Writing – Review & Editing, Z.B., J.M.V., K.C.B., S.A., and J.G.; Funding Acquisition, J.G.

Acknowledgments

This work was supported by a DOD Breast Cancer Research Program Breakthrough Award to J.G., 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 Ed Grow for providing reagents, Margit Janat-Amsbury, Jennifer Richer, K-T Varley, as well as Gertz and Varley laboratory members for their helpful comments on the study and the manuscript.

References

Backes, F.J., Walker, C.J., Goodfellow, P.J., Hade, E.M., Agarwal, G., Mutch, D., Cohn, D.E., and Suarez, A.A. (2016). Estrogen receptor-alpha as a predictive biomarker in endometrioid endometrial cancer. Gynecol Oncol *141*, 312-317.

Bahreini, A.Z., L.; Wang, P.; Levine, K.M.; Tasdemir, N.; Cao, L.; Weir, H.; Puhalla, S.; Davidson, N.E.; Stern, A.M.; Chu, D.; Park, B.H.; Lee, A.V.; Oesterreich, S. (2017). Mutation site and context dependent effects of ESR1 mutation in genome-edited breast cancer cell models. Breast Cancer Research *19*.

Bailey, T.L.B., M.; Buske, F.A.; Frith, M.; Grant, C.E.; Clementi, L.; Ren, J.; Li, W.W.; Noble, W.S. (2009). MEME Suite: tools for motif discovery and searching. Nucleic Acids Res *37*, W202-208.

Buenrostro, J.D.G., P.G.; Zaba, L.C.; Chang, H.Y.; Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods *10*, 1213-1218.

Cancer Genome Atlas Research, N., Kandoth, C., Schultz, N., Cherniack, A.D., Akbani, R., Liu, Y., Shen, H., Robertson, A.G., Pashtan, I., Shen, R.*, et al.* (2013). Integrated genomic characterization of endometrial carcinoma. Nature *497*, 67-73.

Droog, M.N., E.; Dackus, G.M.; Fles, R.; Kim, Y.; Hollema, H.; Mourits, M.J.; Nederlof, P.M.; Van Boven, H.H.; Linn, S.C.; Van Leeuwen, F.E.; Wessels, L.F.A.; Zwart, W. (2017). Estrogen receptor α wields treatment-specific enhancers between morphologically similar endometrial tumors. Proc Natl Acad Sci USA *114*, E1316–E1325.

Fanning, S.W.M., C.G.; Dharmarajan, V.; Carlson, K.E.; Martin, T.A.; Novick, S.J.; Toy, W.; Green, B.; Panchamukhi, S.; Katzenellenbogen, B.S.; Tajkhorshid, E.; Griffin, P.R.; Shen. Y.; Chandarlapaty, S.; Katzenellenbogen, J.A.; Greene, G.L. (2016). Estrogen receptor alpha somatic mutations Y537S and D538G confer breast cancer endocrine resistance by stabilizing the activating function-2 binding conformation. Elife *5*.

Fuqua, S.A., Gu, G., and Rechoum, Y. (2014). Estrogen receptor (ER) alpha mutations in breast cancer: hidden in plain sight. Breast Cancer Res Treat *144*, 11-19.

Fuqua, S.A.C., G.C.; McGuire, W.L. (1993). Estrogen receptor mutations in breast cancer. J Cell Biochem *51*, 135-139.

Gertz, J., Savic, D., Varley, K.E., Partridge, E.C., Safi, A., Jain, P., Cooper, G.M., Reddy, T.E., Crawford, G.E., and Myers, R.M. (2013). Distinct properties of cell-type-specific and shared transcription factor binding sites. Mol Cell *52*, 25-36.

Gertz, J.R., T.E.; Varley, K.E.; Garabedian, M.J.; Myers, R.M. (2012). Genistein and bisphenol A exposure cause estrogen receptor 1 to bind thousands of sites in a cell type-specific manner. Genome Research *22*, 2153-2162.

Gibson, W.J., Hoivik, E.A., Halle, M.K., Taylor-Weiner, A., Cherniack, A.D., Berg, A., Holst, F., Zack, T.I., Werner, H.M., Staby, K.M.*, et al.* (2016). The genomic landscape and evolution of endometrial carcinoma progression and abdominopelvic metastasis. Nat Genet *48*, 848-855.

Hertz, G.Z.S., G.D. (1999). Identifying DNA and protein patterns with statistically significant alignments of multiple sequences. Bioinformatics *15*, 563-577.

Jeselsohn, R., Yelensky, R., Buchwalter, G., Frampton, G., Meric-Bernstam, F., Gonzalez-Angulo, A.M., Ferrer-Lozano, J., Perez-Fidalgo, J.A., Cristofanilli, M., Gomez, H.*, et al.* (2014). Emergence of constitutively active estrogen receptor-alpha mutations in pretreated advanced estrogen receptor-positive breast cancer. Clin Cancer Res *20*, 1757-1767.

Jeselsohn, R.B., J.S.; 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.; Brown, M. (2018). Allele-Specific Chromatin Recruitment and Therapeutic Vulnerabilities of ESR1 Activating Mutations. Cancer Cell *33*, 173-186.

Katzenellenbogen, J.A.M., C.; Katzenellenbogen, B.S.; Greene, G.L.; Chandarlapaty, S. (2018). Structural underpinnings of oestrogen receptor mutations in endocrine therapy resistance. Nature Reviews *18*, 377-388.

Kim, D.L., B.; Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low memory requirements. Nat Methods *12*, 357-360.

Li, H.H., B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools Bioinformatics *25*, 2078-2079.

Liao, Y.S., G.K.; Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923-930.

Love, M.I.H., W.; Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology *15*.

Merenbakh-Lamin, K., Ben-Baruch, N., Yeheskel, A., Dvir, A., Soussan-Gutman, L., Jeselsohn, R., Yelensky, R., Brown, M., Miller, V.A., Sarid, D.*, et al.* (2013). D538G mutation in estrogen receptor-alpha: A novel mechanism for acquired endocrine resistance in breast cancer. Cancer Res *73*, 6856-6864.

Network, T.C.G.A. (2012). Comprehensive Molecular Portraits of Human Breast Tumors. Nature *490*, 61-70.

Osbourne, K.C.S., R. (2011). Mechanisms of Endocrine Resistance in Breast Cancer. Annu Rev Med. *62*, 233-247.

Reddy, T.E.P., F.; Sprouse, R.O.; Neff, N.F,; Newberry, K.M.; Garabedian, M.J.; Myers, R.M. (2009). Genomic determination of the glucocorticoid response reveals unexpected mechanisms of gene regulation. Genome Res *19*, 2163-2171.

Robinson, D.R., Wu, Y.M., Vats, P., Su, F., Lonigro, R.J., Cao, X., Kalyana-Sundaram, S., Wang, R., Ning, Y., Hodges, L.*, et al.* (2013). Activating ESR1 mutations in hormone-resistant metastatic breast cancer. Nat Genet *45*, 1446-1451.

Savic, D., Partridge, E.C., Newberry, K.M., Smith, S.B., Meadows, S.K., Roberts, B.S., Mackiewicz, M., Mendenhall, E.M., and Myers, R.M. (2015). CETCh-seq: CRISPR epitope tagging ChIP-seq of DNA-binding proteins. Genome Res *25*, 1581-1589.

Siiteri, P.K. (1987). Adipose tissue as a source of hormones. The American Journal of Clinical Nutrition *45*, 277-282.

Toy, W., Shen, Y., Won, H., Green, B., Sakr, R.A., Will, M., Li, Z., Gala, K., Fanning, S., King, T.A.*, et al.* (2013). ESR1 ligand-binding domain mutations in hormone-resistant breast cancer. Nat Genet *45*, 1439-1445.

Toy, W., Weir, H., Razavi, P., Lawson, M., Goeppert, A.U., Mazzola, A.M., Smith, A., Wilson, J., Morrow, C., Wong, W.L.*, et al.* (2017). Activating ESR1 Mutations Differentially Affect the Efficacy of ER Antagonists. Cancer Discov *7*, 277-287.

Zhang, Y.L., T.; Meyer, C.A.; Eeckhoute, J.; Johnson, D.S.; Bernstein, B.E.; Nusbaum, C.; Myers, R.M.; Brown, M.; Li, W.; Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biology *9*, R137.

Zhao, Y.L., M.J.; Guillen, V.S.; Ziegler, Y.; Min, J.; Sharma, A.; Kim, A.H.; Chu, D.; Park, B.H.; Oesterreich, S.; Mao, C.; Shapiro, D.; Nettles, K.W.; Katzenellenbogen, J.A.; Katzenellenbogen, B.S. (2017). Structurally Novel Antiestrogens Elicit Differential Responses from Constitutively Active Mutant Estrogen Receptors in Breast Cancer Cells and Tumors. Cancer Research *77*, 5602-5613.