TITLE PAGE

Title

ER α -LBD, a novel isoform of estrogen receptor alpha, promotes breast cancer proliferation and endocrine resistance.

Authors and affiliations

Antonio Strillacci^{1,2,12}, Pasquale Sansone^{1,3}, Vinagolu K Rajasekhar⁴, Mesruh Turkecul⁵, Vitaly Boyko⁵, Fanli Meng⁶, Brian Houck-Loomis⁶, David Brown⁶, Michael F Berger⁶, Ronald C Hendrickson⁷, Qing Chang⁸, Elisa de Stanchina⁸, Fresia Pareja⁹, Jorge S Reis-Filho⁹, Ramya Segu Rajappachetty^{1,2}, Bo Liu^{1,2}, Alex Penson², Chiara Mastroleo¹⁰, Marjan Berishaj¹⁰, Francesca Borsetti¹¹, Enzo Spisni¹², David Lyden³, Sarat Chandarlapaty^{1,2,13,*}, Jacqueline Bromberg^{1,13,*}.

¹ Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

² Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

³ Children's Cancer and Blood Foundation Laboratories, Weill Cornell Medicine, New York, NY, USA.

⁴ Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

⁵ Molecular Cytology Core Facility, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

⁶ Center for Molecular Oncology, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

⁷ Microchemistry and Proteomics, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

⁸ Antitumor Assessment Core Facility, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

⁹ Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

¹⁰ Cancer Biology and Genetics Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

¹¹ Department of Pharmacy and Biotechnology, University of Bologna, Bologna, Italy.

¹² Department of Biological, Geological and Environmental Sciences, University of Bologna, Bologna, Italy.

¹³ Department of Medicine, Weill Cornell Medicine, New York, NY, USA.

* Co-corresponding authors

Classification: Biological Sciences, Medical Sciences

Keywords: Estrogen receptor alpha, Breast Cancer, Endocrine Resistance, Stem Cells, Mitochondria

ABSTRACT

Estrogen receptor alpha (ER α) drives mammary gland development and breast cancer (BC) growth through an evolutionarily conserved linkage of DNA binding and hormone activation functions. Therapeutic targeting of the hormone binding pocket is a widely utilized and successful strategy for breast cancer prevention and treatment. However, resistance to this endocrine therapy is frequently encountered and may occur through bypass or reactivation of ER-regulated transcriptional programs. We now identify the induction of a novel ER α isoform, ER α -LBD, that is encoded by an alternative *ESR1* transcript and lacks the activation function and DNA binding domains. Despite lacking the transcriptional activity, ER α -LBD is found to promote breast cancer growth and resistance to the ER α antagonist fulvestrant. ER α -LBD is predominantly localized to the cytoplasm and mitochondria of BC cells and leads to enhanced glycolysis, respiration and stem-like features. Intriguingly, ER α -LBD expression and function does not appear to be restricted to cancers that express full length ER α but also promotes growth of triple negative breast cancers and ER α -LBD transcript (*ESR1*-LBD) is also present in BC samples from both ER α (+) and ER α (-) human tumors. These findings point to ER α -LBD as a potential mediator of breast cancer progression and therapy resistance.

SIGNIFICANCE STATEMENT

Endocrine resistant and metastatic breast cancer (BC) is a clinically significant problem. Our study of fulvestrant resistant cancer cells led to the discovery of a novel ER α isoform which we call ER α -LBD. Encoded by a truncated transcript variant (*ESR1*-LBD) and lacking the N-terminal domains (activation of transcription and DNA binding), ER α -LBD displays a unique role in BC tumorigenesis and progression by mechanisms that may involve metabolic and cell growth advantages, stemness and therapy resistance. Importantly, *ESR1*-LBD is preferentially expressed in human breast tumor tissues and may be used as prognostic marker in BC.

INTRODUCTION

The nuclear receptor estrogen receptor alpha (ER α) is the major therapeutic target for the ~75% of breast cancers (BC) where its expression is detected (1). ER α is a multifunctional protein that contributes to cellular processes via transcriptional regulation, participation in signaling complexes (2, 3) and regulation of mitochondrial function (4, 5). Its function in primary BC growth depends mainly on its activation of transcription in response to hormone binding-mediated conformational changes in the receptor. Thus, therapeutic targeting of ER α has been directed at inhibiting this hormone-receptor interaction. The major forms of therapy include suppressors of estrogen biosynthesis (*e.g.* aromatase inhibitors, GNRH antagonists) and direct ER α antagonists (*e.g.* tamoxifen, fulvestrant) (6-7). Cancers develop resistance to these forms of therapy over time, often through alterations in the *ESR1* gene including activating point mutations in the ligand binding domain and gene fusions (8-10) that restore ER α transcriptional activity. In addition, cancers can bypass ER α signaling and activate oncogenic functions through genetic alterations that involve other growth factor signaling pathways, modification of cell cycle regulators and increases in stem cell activity (11-14). Additionally, epigenetic mechanisms of resistance may also be contributory and whether these can affect ER α or these other growth pathways is a key question, particularly given that up to half of all resistant tumors harbor no mutation known to cause resistance (10, 15).

About 15% of BC cases are triple-negative breast cancers (TNBCs), lacking the expression of ER α , PR and HER2-amplification and are thus not responsive to targeted therapies against ER α or HER2. These cancers are associated with a poorer prognosis due to a higher rate of metastatic progression and resistance to treatment (16, 17). Notably, TNBC may be responsive to estrogen stimulation via ER α -independent pathways, promoting tumor formation and progression via different molecular mechanisms (18).

In this work, we investigated ER α protein and transcripts in *ESR1* wildtype breast cancers that were resistant to fulvestrant and in TNBC. We identified an isoform that maintains the ligand binding domain but lacks the DNA binding domain, and promotes breast cancer growth and endocrine resistance through non-canonical functions outside the nucleus. These data reinforce the importance of the broader biological functions of ER protein family outside of its transcriptional activation role.

RESULTS

Fulvestrant resistant cells express a truncated estrogen receptor α isoform.

In order to investigate novel mechanisms of hormonal therapy resistance (HTR) in breast cancer, we examined the expression of ER α protein in different BC cell lines, either in the absence or presence of the selective estrogen receptor degrader (SERD) fulvestrant (ICI 182,780) which reduces ER α full-length (ER α -FL, 66 kDa) protein levels as a consequence of reduced stability and dimerization (19). We postulated that changes in fulvestrant-mediated suppression of ER α -FL levels may promote drug resistance and investigated this response in several models selected for fulvestrant resistance (MCF-7 FulvRes, MCF-7 Y537S and two PDX models) (14, 20), compared to fulvestrant-sensitive MCF-7 cells (Fig.

1A; Supplementary Fig. S1A and B). After 24 hours exposure to 1 mM fulvestrant, MCF-7 cells and MCF-7 FulvRes showed ~80% and ~60% loss of ER α -FL expression, respectively; MCF-7 Y537S and PDX-ER α (+) showed a slight decrease (~20% and ~30%) and PDX-ER α (-) showed no expression at all. Notably, we observed a faster migrating band (~37 kDa) in all four resistant models that appeared to increase in response to fulvestrant. As this protein was detected in both ER α positive and negative models, we analyzed by immunoblotting two triple-negative breast cancer (TNBC) cell lines, MDA-MB-453 and MDA-MB-231, also lacking ER α -FL expression and therefore resistant to fulvestrant treatment (**Supplementary Fig. S1A**). Notably, both TNBC cell lines displayed this lower molecular weight protein whose expression increased after fulvestrant treatment (**Fig. 1B**). To establish the identity of the novel protein, we performed immunoprecipitation and mass spectrometry experiments from MCF-7 FulvRes and MDA-MB-231 cells and determined that the ER α -related protein is a truncated isoform of ER α -FL, lacking the N-terminal domains AF1 (transcription Activation Function-1), DBD (DNA Binding Domain) and a portion of the hinge domain, and it is composed principally by the C-terminal domains LBD (Ligand Binding Domain) and AF2 (**Fig. 1C; Supplementary Fig. S1C**). Based on our observation, we named this novel ER α protein variant 'ER α -LBD'. ER α -LBD has 332 amino acids, a MW of 37.3 kDa and a predicted 3-D structure that is similar to the predicted ER α -FL (**Fig. 1C; Supplementary Fig. S1D**). As expected, we failed to detect a 37.3 kDa protein using an ER α antibody raised against the amino-terminus (**Supplementary Fig. S1E**).

Across all the models we investigated, we found that fulvestrant increased expression of ER α -LBD and in some cases was necessary for detection. These data suggested that fulvestrant was serving to stabilize ER α -LBD akin to the way ligands protect full-length ER against proteolytic degradation when unchaperoned by HSP90. To further address this possibility, we analyzed ER α -LBD and ER α -FL expression upon HSP90 inhibition using an ATP inhibitor of HSP90 function (SNX-2112) (**21**). Following SNX-2112 treatment, both ER α -FL and ER α -LBD underwent rapid degradation, either in the presence (lane 2) or absence (lane 3, CSS) of physiological E2 levels. Conversely, high levels of E2, fulvestrant or tamoxifen treatment (1 μ M) led to increased stability of ER α -LBD. (**Fig. 1D; Supplementary Fig. S1F**).

Taken together, our data demonstrate that ER α (+) and (-) BC models characterized by HTR can produce a truncated ER α isoform whose expression can be induced by ER α ligand binding pocket compounds, including fulvestrant.

ER α -LBD is encoded by an ESR1 transcript variant.

To understand the basis of ER α -LBD formation we carried out RNA capture-sequencing and PCR assays to analyze the usage of ESR1 exons and transcripts in different models, including both ER α -FL positive and negative cells. High-depth capture-sequencing data showed that BC ER α -FL(-) cells expressing exclusively ER α -LBD, such as MDA-MB-453/-231 and PDX-ER α , have higher abundance of ESR1 read counts in the C-terminus (CDS exons #4, 5, 6, 7 and 8), compared to the N-terminus (CDS exons #1-3). In contrast, ESR1 exon usage profile was significantly different in ER α -FL(+) cells, such as T-47D, MCF-7 and PDX-ER α (+), where the expression of exons #1-3 was higher than exons #4-8 (**Fig. 2A; Supplementary File S1**). To identify an ESR1 transcript variant for ER α -LBD, we searched the FANTOM CAT (CAGE Associated Transcriptome) database, which displayed an

additional ESR1 exon between exons #3 and #4, with a transcription start site (TSS) at the 5' end (**Supplementary Fig. S2A**) (22). We annotated this region as 5'UTR of a putative ESR1-LBD transcript variant (exon E3a). In this region, a higher number of read counts was found in ER α (-) cells, compared to ER α (+) (**Fig. 2A**). Differential exon usage was determined for ESR1-FL (encoding ER α -FL) and a predicted ESR1-LBD (encoding ER α -LBD) transcript by comparing ER α -FL(+) and ER α (-) groups, revealing a significantly lower usage of exons E1, E2, E3 and E8-3UTR, and higher usage of exons E3a, E5, E6 and E7 for the ER α (-) group (**Fig. 2B**).

Exon data were supported by the analysis of ESR1 transcripts expression, demonstrating higher levels of ESR1-LBD transcript in ER α -FL(-) cells, compared to ESR1-FL (**Fig. 2C** and **D**; **Supplementary File S1**). Also qPCR (exon 'walking') and RT-PCR assays confirmed the RNA-seq results. ER α -FL(+) and ER α -FL(-) cells express an ESR1 mRNA ranging from exon 1 to 8 and from exon 4 to exon 8, respectively (**Supplementary Fig. S2B** and **C**). These results are in accordance with the western blot data in **Fig. 1B**, showing that ER α -FL(-) cells express a truncated ER α protein (LBD) and not full-length ER α . As definitive proof, the stable cloning of the putative ER α -LBD CDS into BC cells (exon 4 to exon 8), led to ER α -LBD overexpression (oe) whereas targeted genomic deletions within exon 4 (by CRISPR technology) induced ER α -LBD knockdown (kd) (**Supplementary Fig. S2D**).

RNA sequencing and qPCR data showed reduced expression of the ESR1 3'UTR sequence in ER α -FL(-) cells as compared to ER α -FL(+) ones (**Fig. 2A** and **B**; **Supplementary Fig. S2B**). It has been observed that shortening of 3'UTRs by alternative cleavage and polyadenylation is associated with increased RNA stability, and that this phenomenon can promote oncogene mediated transformation by enhanced protein production (23). Thus, we hypothesized that the partial/total loss of the 3'UTR sequence in ER α -FL(-) cells may enhance ESR1 mRNA stability. To test this, we treated BC cells with actinomycin-D (inhibitor of RNA synthesis) and indeed the LBD portion of ESR1 transcript was significantly more stable than the FL-specific in MCF-7 FulvRes and TNBC cells (**Supplementary Fig. S2E**).

We note that ER α -LBD is not the only ER α transcript variant that has been observed in breast cancer. An ER α -36 variant, encoded by ESR1 exons #2-6 and an additional exon downstream of the ESR1 gene, encoding a protein expressing the DNA-binding domain, a truncated LBD and a partial dimerization domain has been described in a variety of cancer models including tamoxifen resistant breast cancers (24, 25). Importantly, ER α -LBD isoform is distinct from ER α -36 as specific primers and an antibody for the ER α -36 variant failed to detect ER α -LBD (**Supplementary Fig. S2F** and **G**).

ER α -LBD localizes to the mitochondria and cytoplasm.

ER α -LBD lacks the N-terminal domain of the full-length receptor and we thus hypothesized it may have a distinctive cellular location. Using ER α immunohistochemistry (IHC), full-length expressing MCF-7 cells show robust nuclear staining that is markedly reduced 24 h after fulvestrant treatment. By contrast, ER α -LBD expressing models MCF-7 FulvRes, MDA-MB-453/-231 and PDX-ER α (-) display punctate cytoplasmic foci (**Fig. 3A**). Based on this particular pattern, we hypothesized that ER α -LBD might be localized to the mitochondria and performed confocal analyses examining cells for ER α protein (green), OXPHOS (mitochondrial marker, red) and DAPI (nuclear marker, blue) (**Fig. 3B**). We examined

images from both separate and merged channels, and by 3-D image processing to evaluate the co-localization volume between ER α , nuclei or mitochondria. The overall IF staining confirmed data from the IHC: MCF-7 cells have a strong ER α nuclear localization which is decreased after fulvestrant treatment; MCF-7 FulvRes chronically treated with fulvestrant have both nuclear and mitochondrial ER α staining. In full-length negative cells, ER α -LBD was largely absent from the nucleus and found in the cytoplasm with a substantial fraction co-localized with mitochondria. Detailed co-localization values and statistics are shown in **Supplementary Fig. S3A**.

Given the relative absence of ER α -LBD nuclear localization, we examined its role in canonical ER α signaling and we observed none of the transcriptional activation functions of full-length ER α . For instance, overexpression of ER α -LBD had no effect on Estrogen Response Element (ERE) luciferase expression or ER α target gene transcription in both ER α -FL(+) and TNBC cells (**Supplementary Fig. S2D** and **Fig. 3B-G**) (26).

To further establish the functional significance of cytoplasmic and mitochondrial ER α -LBD, we performed co-IP and mass spectrometry experiments aimed at identifying ER α -LBD protein-protein interactions (PPIs). We designed the experiment comparing MCF-7 cells overexpressing ER α -LBD with parental (NC), in the absence or presence of fulvestrant and we identified a total of 52 peptides associated with ER α -LBD (**Fig. 4A**; **Supplementary Fig. S2D**). Pathway enrichment and neural networking analyses performed on this protein set identified processes involved in carbohydrate metabolism (glycolysis and gluconeogenesis), cell signaling (MYC, mTOR/mTORC1, PIK3C1/AKT, ER α), hypoxia and angiogenesis (HIF1 α and VEGFR signaling) and mitochondrial metabolism (oxphos and respiration) (**Fig. 4B** and **C**; **Supplementary Fig. S4A** and **Supplementary File S2**). Comparable experiments in TNBC models overexpressing or knocking down ER α -LBD identified a total of 93 peptides involved in ER α -LBD PPIs, and 29 of them were shared with those found in MCF-7 cells (**Fig. 4D**; **Supplementary Fig. S2D**). Similarly to the MCF-7 model, enriched pathways included signaling, carbohydrate and mitochondrial metabolism (including also fatty acid β -oxidation) and angiogenesis (**Fig. 4E** and **F**; **Supplementary Fig. S4B** and **Supplementary File S2**).

ER α -LBD regulates cell metabolism.

The localization and protein-protein interaction analyses point to a potential metabolic function for ER α -LBD. We speculated ER α -LBD may play a role regulating glycolysis and mitochondrial respiration and so assessed the effect of ER α -LBD overexpression or knockdown on key metabolic parameters. In ER α -FL expressing MCF-7, control (NC) and ER α -LBDoe cells showed similar OCR (oxygen consumption rate, index for respiration) and ECAR (extracellular acidification rate, index for glycolysis) under basal conditions. However, following fulvestrant-mediated depletion of ER α -FL and stabilization of ER α -LBD, we observed significantly higher respiratory parameters (ATP production, basal and maximal respiration and spare respiratory capacity) and higher glycolytic parameters (glycolytic capacity and glycolytic reserve) (**Fig. 5A**; **Supplementary Fig. S5A**). In the TNBC models featuring knockdown of ER α -LBD, ER α -LBDkd cells were characterized by reduced levels of respiratory parameters (ATP production, basal and maximal respiration), compared to controls (NC). Notably, ER α -LBDkd led to the impairment of glycolysis only in MDA-MB-231 (**Fig. 5B** and **C**; **Supplementary Fig. S5B** and **C**).

ER α -LBD promotes growth and fulvestrant resistance, *in vitro* and *in vivo*.

Given the impact of ER α -LBD on breast cancer cell metabolism, we investigated the effects of ER α -LBD on cell proliferation. *In vitro*, MCF-7 cells overexpressing ER α -LBD, compared to control cells (NC), showed no significant differences in growth. However, upon treatment with fulvestrant, control cells were growth inhibited while ER α -LBDoe cells continued to proliferate establishing a role for ER α -LBD in mediating fulvestrant resistance (**Fig. 6A**). Moreover, growth of MCF-7 FulvRes cells was potently impaired after ER α -LBD knockdown, especially under fulvestrant treatment (**Fig. 6B**). Beyond the effects in ER α -FL(+) cells, we also observed effects of ER α -LBD in TNBC cells. In both MDA-MB-453 and MDA-MB-231 we observed ER α -LBD knockdown to slow cell proliferation (~50%) (**Fig. 6C and D**).

To further verify the biologic significance of these effects *in vivo*, these models were grown orthotopically as xenograft models. As observed *in vitro*, MCF-7 ER α -LBDoe tumors continued to grow upon fulvestrant treatment while tumors from NC cells were expectedly growth inhibited with fulvestrant (**Fig. 6E and F**). Moreover, *in vivo* tumor growth for MDA-MB-231 ER α -LBDkd mice group was markedly reduced, when compared to NC group (**Fig. 6G and H**).

MCF-7 ER α -LBDoe cells showed a growth advantage when cultured in low-attachment (3-D growth) either in the presence or absence of fulvestrant treatment (**Supplementary Fig. S6A**), whereas ER α -LBD knockdown led to reduced 3-D growth in both MDA-MB-453 and -231 cells (**Supplementary Fig. S6B**). Similarly, the migratory capacity of BC models was enhanced in ER LBD overexpression models and impaired in knockdown models (**Supplementary Fig. S6C and D**).

ER α -LBD expression is associated with proliferation, endocrine resistance, stemness and metabolism of breast cancer cells.

Given the growth advantage, drug resistance and metabolic phenotypes afforded by ER α -LBD, we further evaluated its impact by gene expression analysis. We considered fulvestrant treated MCF-7 ER α -LBDoe and MCF-7 FulvRes cells as ‘gain-of-function’ models compared to MCF-7 control cells (NC). RNA-seq analysis led to the identification of ~17K differentially regulated genes (**Fig. 7A; Supplementary File S3**). Gene set enrichment analysis (GSEA) was performed on these same samples using specific gene set collections related to breast cancer, cell proliferation and cell metabolism (**Supplementary File S4**). Interestingly, gene sets involving cell signaling, breast cancer malignancy/endocrine resistance, EMT and stemness, carbohydrate and mitochondrial metabolism were found to be enriched in the ‘gain-of-function’ models (**Fig. 7B and C; Supplementary File S3**). Conversely, RNA-seq analysis of MDA-MB-453/-231 ER α -LBDkd “loss-of-function” models identified ~16K differentially regulated genes in TNBC ER α -LBDkd cells compared to NC cells (**Fig. 7d; Supplementary File S3**). Similarly to the MCF-7 model, gene sets associated with cell proliferation and survival, cell stemness and metabolism were found to be enriched in TNBC NC cells compared to TNBC ER α -LBDkd (**Fig. 7E and F; Supplementary File S3**).

We further explored the role of ER α -LBD in breast cancer stemness and endocrine resistance. Cells from MCF-7 and MCF-7 FulvRes xenografts tumors were screened by flow cytometry and qPCR to measure the expression of CD44, SOX2 and SOX9, markers associated with stemness and endocrine resistance (**27, 28**). MCF-7 FulvRes tumors were

characterized by a higher percentage of CD44^{High} cells which showed higher expression of ESR1-LBD, SOX2 and SOX9, suggesting a correlation between ER α -LBD expression and stemness (**Supplementary Fig. S7B and C**).

ESR1-LBD expression is identified in human breast cancers.

Our data supports a role for ER α -LBD in fulvestrant resistance and proliferation of BC cells. To determine the clinical relevance of these observations, we examined ESR1-LBD expression by analyzing RNA-seq samples from human BC specimens.

ESR1-FL overall expression was assessed on the following collection of samples: ER α (+) BC primary tumors ($n = 42$); uninvolved breast tissue adjacent to ER α (+) primary tumors ($n = 30$); TNBC primary tumors ($n = 42$); uninvolved breast tissue adjacent to TNBC primary tumors ($n = 21$); metaplastic BC primary tumors (MpBC, $n = 17$) (**29, 30**). As expected, ESR1-FL transcript was found to be markedly reduced in TNBC and MpBC samples (**Fig. 8A**). Notably, ER α expression is known to be present in normal breast epithelium (**31**). For each cancer sample, ESR1 exon expression was also calculated and, similarly to what we observed with BC cell lines (**Fig. 2A**), TNBC and MpBC samples showed higher amounts of 5'UTR exon 3a and CDS exons #4-8, compared to #1-3 (**Fig. 8B**). We also examined the levels of the ESR1-LBD transcript in BC groups, relative to ESR1-FL (**Fig. 8C**). 76% of ER α (+), 50% of TNBC and 81% of MpBC were characterized by high ESR1-LBD levels (Q3+Q4 quartiles, above the median calculated on all samples). Interestingly, only 17% of ER α (+)^{ADJ} and none of TNBC^{ADJ} samples were found to have high ESR1-LBD expression (**Fig. 8D; Supplementary Fig. S8A**).

Our second investigation was performed on TCGA-BRCA (The Cancer Genome Atlas Breast Invasive Carcinoma) RNA-seq samples (**32**). Based on the available IHC data, samples were grouped as: ER α (+)/HER2(+) ($n = 123$), ER α (+)/HER2(-) ($n = 436$), ER α (-)/HER2(+) ($n = 40$) and ER α (-)/HER2(-) ($n = 126$). The analysis of ESR1-FL expression showed low levels in ER α (-) samples but no significant differences between HER2(+) and HER2(-) cancers, in both ER α (+) and ER α (-) groups (**Fig. 8E**). As above, ESR1 exon expression is shown by using heat maps revealing a slightly higher abundance of exons #3a and #4-8, compared to #1-3 in the ER α (-) samples (**Fig. 8F**). ESR1-LBD transcript expression was also assessed in all groups and correlated to ESR1-FL (**Fig. 8G**). Higher ESR1-LBD levels (Q3+Q4) were identified in 60% of ER α (+)/HER2(+), 58% of ER α (+)/HER2(-), 13% of ER α (-)/HER2(+) and 25% of ER α (-)/HER2(-) (**Fig. 8H; Supplementary Fig. S8B**). All expression data are summarized in **Supplementary Files S5 and S6**.

DISCUSSION

Fulvestrant resistant breast cancer is frequently observed in the metastatic setting however little data exists on the molecular events underlying resistance (**13, 20, 33, 34**). Our analysis of fulvestrant resistant BC models led to the identification of a novel ER α protein variant (ER α -LBD, 37.3 kDa MW) lacking the N-terminal transcriptional activation (AF1) and DNA binding domains (DBD) but including a portion of the hinge domain followed by the C-

terminal domains LBD (Ligand Binding Domain) and AF2. Not only was ER α -LBD identified in fulvestrant resistant cell lines but also in TNBC models. Unlike full-length ER α which is degraded by fulvestrant, ER α -LBD is stabilized by this agent along with other ER α ligands such as estradiol and tamoxifen, leading to increased ER α -LBD expression (**Fig. 1**). We hypothesize that low expression of ER α -LBD in the absence of high levels of ligands may explain why this isoform has not been previously observed in BC cells by others. We also suggest that this feature should be exploited to better understand ER α protein stability and discover novel pharmacologic inhibitors of ER α and its relevant isoforms. A 37 kDa ER species expressed at low levels was previously isolated from the mouse uterus. Interestingly, this ER variant appeared to have the ligand-binding region and a portion of the hinge domain, but lacked a DNA-binding region (**35**).

The identity of ESR1-LBD transcript was determined by analyzing available sequence databases (specifically, FANTOM CAT) and by RNA-seq experiments, suggesting that the ESR1-LBD transcript might utilize a novel transcription start site and 5'UTR rather than processed by alternative splicing of the full-length transcript (ESR1-FL). The hypothesis of a new putative *ESR1* promoter was supported by preliminary data (not shown), describing potential transcriptional activity in the upstream genomic region from exon E3a. Additionally, we noted reduced expression of the 3'UTR sequence of ESR1-LBD in the ER α (-)/TNBC models, a phenomenon that could alter mechanisms of post-transcriptional inhibition mediated for example by microRNAs. Indeed, we demonstrated that ESR1-LBD is more stable than ESR1-FL and this may explain how a relatively low abundant transcript is translated into a detectable protein (**Fig. 2**). ESR1-LBD variant has not been previously described and we suggest that the shortening of the 3'UTR might account for this, as it might limit its detection in experiments based on poly(A)-enriched RNA samples.

The putative function of ER α -LBD was first investigated by examining its localization demonstrating its absence from the nucleus and enrichment in the cytoplasm and mitochondria of breast cancer cells (**Fig. 3**). We examined ER α -LBD protein-protein interactions by co-IP/MS using overexpression and knockdown models and determined a role in glycolysis/gluconeogenesis, mitochondrial metabolism, signaling and angiogenesis (**Fig. 4**). These observations were further supported by functional analyses revealing metabolic and cell/tumor growth advantages associated with ER α -LBD (**Figs. 5 and 6**) and RNA-seq analysis (**Fig. 7**).

An extra-nuclear role for ER α and its isoform ER α -36 has been well established most notably by their presence in the plasma membrane and their association with receptor tyrosine kinases and MAPK (**36, 37**). Our data would suggest a distinct role for ER α -LBD as no apparent association with RTKs nor MAPK/ERK was identified. The role of mitochondrial ER α has also been described functioning as a transcription factor interacting with mtDNA and preventing UV-induced apoptosis as well as directly interacting with the mitochondrial protein HADHB (**4, 38**). Although ER α -LBD has no DNA binding capacity, we also found HADHB among its protein partners, suggesting both unique and overlapping functions for ER α and ER α -LBD in the mitochondria (**Fig. 4**).

Additionally, a role of ER α -LBD in breast cancer stemness was suggested by RNA-seq analysis and qPCR demonstrating an enrichment of the ESR1-LBD transcript in fulvestrant resistant CD44^{High} tumor cells (**Fig. 7** and **Supplementary Fig. S7**). Importantly, ER α -LBD's role in fulvestrant resistance was also demonstrated by overexpression studies on BC cell

growth (**Fig. 6**) and supported by RNA-seq analysis (**Fig. 7**). Given the role of stemness with endocrine resistance, we suggest that ER α -LBD may provide a link between these phenomena (**13, 14, 39, 40**). Furthermore, the higher abundance of this transcript and protein in cancer stem cells may also explain its relatively low detection in bulk tumor and cell line analyses.

Notably, some of ER α -LBD features are shared with the more recently discovered ER α isoform, ER α -36 (**24, 25, 41**). No cross-detection between ER α -LBD and ER α -36 was found in our models. Specifically, the C-terminal antibody we used for IF, IHC and co-IP studies does not recognize ER α -36. Additionally, the use of primers and an antibody specific for ER α -36 did not lead to its detection in our hands (**Supplementary Fig. S2**). We suggest that isoforms of ER α may play unique and/or overlapping critical roles in endocrine resistant disease, TNBC and prognosis. Identifying the unique contributions of these isoforms in breast cancer will be of interest.

Lastly, we demonstrated the presence of the ESR1-LBD transcript in primary BC specimens using publicly available RNA-Seq data (**29, 30, 32**). ESR1-LBD was absent or poorly expressed in non-malignant breast tissue but found to variable levels in ER α (+) tumors, TNBC and metaplastic breast cancers (MpBC). Remarkably, high levels of ESR1-LBD were identified in 16/17 of metaplastic breast cancers which is one of the most aggressive and chemo-resistant subtypes. The TCGA-BRCA study allowed us to examine the contribution of HER2 expression on ESR1-LBD expression. We observed higher levels of ESR1-LBD in TNBC (25%) as compared to ER-/HER2+ samples (13%), demonstrating differential expression of this transcript in the different subtypes of breast cancer (**Fig. 8**). Unfortunately, data from BC cases with acquired fulvestrant resistance were not available but we hypothesize that expression of ESR1/ER α -LBD would be found in a subset of these cancers. A limitation of our study is the lack of a selective ER α -LBD protein assay to test for this isoform in patient samples and thus cannot ascertain its protein expression and prevalence.

In summary, we have identified a novel ER α isoform “ER α -LBD” in ER α (+) fulvestrant resistant cancer cells and TNBCs. The development of specific strategies for ER α -LBD detection (*e.g.* antibodies, probes) is essential. Further studies are ongoing, to unravel a broader presence and the unique role that ER α -LBD plays in tumorigenesis and cancer progression, as well as its prognostic significance.

METHODS

Cell culture and reagents

All cell lines were obtained from ATCC (American Type Culture Collection) and authenticated using short tandem repeat (STR) analysis. Cells were cultured in MEM or RPMI-1640 medium (MSKCC Media Prep, USA) supplemented with 10% heat-inactivated fetal bovine serum (Gibco, USA) and 1% penicillin/streptomycin (Gibco, USA) and maintained at 37°C and 5% CO₂ in humidified atmosphere. Reagents used for cell culture and treatments: fulvestrant (Selleckchem, USA), SNX-2112 (Selleckchem, USA), charcoal stripped fetal bovine serum (CSS, Gibco, USA), 17 β -estradiol (E2, Sigma-Aldrich, USA), 4-

Hydroxytamoxifen (4-OHT, Sigma-Aldrich, USA). Further information on cell lines and clones in **Supplementary Data**.

Western blot

Cells were homogenized in RIPA lysis buffer for protein extraction, supplemented with protease inhibitors (Thermo Scientific, USA). Denatured proteins were separated in 12% SDS–PAGE and then transferred onto nitrocellulose papers (Pall, USA). After blotting, nitrocellulose papers were incubated with specific antibodies. Primary antibodies: anti-ER α F-10 (Santa Cruz Biotechnology, USA, 1:500); anti- β -actin 13E5 (Cell Signaling Technology, USA, 1:2000). Secondary antibodies (HRP conjugated): anti-mouse (Cytiva, USA, 1:5000). Immunolabelling was visualized using ECL procedure (PerkinElmer, USA).

Protein structure analysis

Protein sequence and features were analyzed using Sequence Analysis software (Informagen, USA). Three-dimensional protein models were generated by Phyre² web portal (www.sbg.bio.ic.ac.uk) and visualized by using RasMol software (<http://www.openrasmol.org>).

RNA capture-sequencing and ESR1 transcripts analysis

Total RNA was extracted as described elsewhere (42) and samples (100 ng) were input in the RNA library construction using the KAPA RNA Hyper library prep kit (Roche, Switzerland). Customized adapters with unique molecular indexes (UMI) (Integrated DNA Technologies, USA) and Sample-specific dual-indexes primers (Integrated DNA Technologies, US) were added to each library. Each RNA library was pooled for hybridization capture with customized ESR1 Panel (Integrated DNA Technologies, USA) using a capture protocol modified from NimbleGen SeqCap Target Enrichment system (Roche, Switzerland). Libraries were then sequenced on an Illumina MiniSeq with paired-end reads (x150 cycles, 1.4millions reads/sample). Raw sequencing data output was processed for expression analysis using STAR Aligner (43), DEXSeq-count (44), Cluster 3.0 and Java TreeView (45, 46), DEXSeq and plotDEXSeq (44), Kallisto (47) (**Supplementary Fig. S2H**).

Immunohistochemistry and confocal microscopy

ER α immunostaining (IHC) was performed on Benchmark Ultra using the ultraView DAB Detection kit (Ventana, USA). Antigen retrieval was performed onboard with UltraCC1 buffer (pH 8.2–8.5) at 95°C for 52 min. Primary antibody (anti-ER α F-10, Santa Cruz Biotechnology, USA) 1:100 for 28 min at 37°C. Secondary antibody 1:100 for 1 h. Images were obtained using a Zeiss Axiovert Widefield Microscope and Zeiss ZEN software (Carl Zeiss, Germany). Immunofluorescence was performed using Leica Bond RX stainer and the Bond Polymer Refine Detection kit (Leica, Germany). Antibodies and detection: OXPHOS (2 μ g/ml, Invitrogen, USA); ER α D8H8 (1 μ g/ml, Cell Signaling Technology, USA); Alexa Fluor

Tyramide signal amplification reagents (Life Technologies, USA); 4',6-diamidino-2-phenylindole (DAPI, Sigma-Aldrich, USA). Slides were mounted in Mowiol 4-88 (Calbiochem, USA). Confocal imaging was performed on a Leica SP8 inverted microscope (Leica, Germany). Image processing and analysis (2-D and 3-D) was performed using Imaris software (Bitplane, CH).

ERα Co-Immunoprecipitation and LC/MS

Protein were extracted using Co-IP buffer (NaCl 150 mM, EDTA 0.5 mM, NP-40 0.5% (v/v), Tris-HCl 10 mM pH 7.4, PMSF 1 mM). Immunoprecipitation (IP): anti-ERα antibody D8H8 1:75 (Cell Signaling Technology, USA) and Dynabeads™ Protein A (Invitrogen, USA). IP samples were washed 4 times with 50 mM ammonium bicarbonate buffer and collected by centrifugation. Samples were then digested overnight with 2 µg trypsin in 80 µl of 50 mM ammonium bicarbonate at 37°C. Desalted and dried peptides were reconstituted in 10 µl 0.1% (vol/vol) formic acid and analyzed (4 µl) by microcapillary liquid chromatography with tandem mass spectrometry. Further details in **Supplementary Methods**.

Pathway enrichment analysis was performed by using Enrichr platform (<https://maayanlab.cloud/Enrichr>). Enriched pathways network, clustering and protein–protein interaction (PPI) network analyses were carried out using Cytoscape, EnrichmentMap, AutoAnnotate and stringApp software (48).

Metabolic assay

Mitochondrial respiration and glycolysis were assessed using Seahorse extracellular flux analyzer (XFe96) and Seahorse XF Cell Mito Stress Test Kit (Agilent Technologies, USA). Cells were seeded in 96-well plates (2 x 10⁴ cells/well), with or without Fulvestrant 1 µM depending on the experiments. Raw data output was collected and analyzed using Wave Software (Agilent Technologies). Procedures and OCR/ECAR data interpretation were carried out accordingly to manufacturer's guidance.

Animal models

NOD scid gamma (NSG) female mice at age of 6-8 weeks were obtained from Jackson Laboratory (USA) and maintained in pressurized ventilated caging. To sustain tumor growth in MCF-7 models, 17β-estradiol pellets (0.18 mg) were implanted subcutaneously 3 days before BC cells injection. For both MCF-7 and MDA-MB-231 models, cancer cells were injected in the mammary fat pads (MFPs) of mice. For each experimental sample, cell suspensions were mixed with an equal volume of Matrigel (BD Biosciences, USA). For the MCF-7 model, injectable fulvestrant (Faslodex®, AstraZeneca, UK) was given intramuscularly in the tibialis posterior/popliteal muscles (100 mg/injection/once a week) for 20 days. Control mice received isotype control (placebo) or PBS injection. Tumor volumes were measured with vernier calipers starting from 14 days after cell implantation. Animals were sacrificed as they reached experimental endpoint. All procedures and experiments were completed in accordance with the Guidelines for the Care and Use of Laboratory

Animals and were approved by the Institutional Animal Care and Use Committees at MSKCC (MSKCC#12-10-016).

RNA sequencing

Total RNA was extracted from cell lines as described elsewhere (42) and sequenced on an Illumina HiSeq instrument (2x150 paired-end, 100 million reads/sample) by GENEWIZ, LLC. (USA). Raw sequencing data output was processed for downstream analyses using the following bioinformatic tools: STAR (reads alignment) (43), HTSeq (gene expression) (49), Cluster 3.0 and Java TreeView (hierarchical clustering and heat map) (45, 46), GSEA (<https://www.gsea-msigdb.org>) (Supplementary Fig. S7A; Supplementary File S4).

Breast cancer patients data

RNA-seq FASTQ files of 135 breast cancer were obtained from the European Nucleotide Archive, ENA (Study Accession: PRJNA251383; Supplementary File S5) (29). RNA-seq FASTQ files of 17 metaplastic breast cancer were kindly provided by Dr. Jorge S Reis-Filho (Supplementary File S5) (30). RNA-seq bam files of 1097 breast IDC patients were obtained from TCGA breast cancer (BRCA) cohort (TCGA dbGaP Accession ID: phs000178.v11.p8; Supplementary File S6) (32). Raw data were processed by using STAR (reads alignment) (43), DEXSeq-count (44), Kallisto (47) and Morpheus (<https://software.broadinstitute.org/morpheus>).

Statistics

Data were expressed as mean $\pm$ s.e.m. Student's t test (unpaired/paired; one/two-tailed) or Analysis of Variance (two-way ANOVA, followed by Fisher's test or Sidak's/Tukey's correction for multiple comparisons) were used to assess the statistical significance of the differences. For DEXSeq analysis, *P* values were adjusted for multiple testing using Benjamini & Hochberg (BH) correction (FDR). For Enrichr analysis, *P* values were computed from the Fisher exact test and FDR adjusted *P* values (*Q* values) were used to filter enriched pathways for Cytoscape and Enrichment Map analysis. For statistics on GSEA analysis, please refer to official website (www.gsea-msigdb.org). Differences were considered statistically significant at $P < 0.01$ and $P < 0.05$.

ACKNOWLEDGMENTS AND FUNDING SOURCES

This work was supported by grants from National Institute of Health to MSKCC's Cancer Center (P30CA008748) and to S.C. (R01CA204999), and a grant from University of Bologna to E.S. (RFO 2015). J.B. acknowledges Charles and Marjorie Holloway Foundation, Sussman Family Fund, Lerner Foundation and Beth C. Tortolani Foundation. S.C. also acknowledges support by the Breast Cancer Research Foundation. F.P. is partially funded by a National Institutes of Health/ National Cancer Institute K12CA184746 grant.

REFERENCES

1. Bernstein L, Lacey JV Jr. Receptors, associations, and risk factor differences by breast cancer subtypes: positive or negative? *J Natl Cancer Inst* **2011** 103:451-3.
2. Hennighausen L, Robinson GW. Signaling pathways in mammary gland development. *Dev Cell* **2001** 1:467-75.
3. Siersbæk R, Kumar S, Carroll JS. Signaling pathways and steroid receptors modulating estrogen receptor α function in breast cancer. *Genes Dev* **2018** 32:1141-54.
4. Pedram A, Razandi M, Wallace DC, Levin ER. Functional estrogen receptors in the mitochondria of breast cancer cells. *Mol Biol Cell* **2006** 17:2125-37.
5. Hamilton DJ, Minze LJ, Kumar T, Cao TN, Lyon CJ, Geiger PC, Hsueh WA, *et al.* Estrogen receptor alpha activation enhances mitochondrial function and systemic metabolism in high-fat-fed ovariectomized mice. *Physiol Rep* **2016** 4:e12913.
6. Osborne CK. Aromatase inhibitors in relation to other forms of endocrine therapy for breast cancer. *Endocr Relat Cancer* **1999** 6:271-6.
7. Johnston SJ, Cheung KL. Fulvestrant - a novel endocrine therapy for breast cancer. *Curr Med Chem* **2010** 17:902-14.
8. Toy W, Shen Y, Won H, Green B, Sakr RA, Will M, *et al.* ESR1 ligand-binding domain mutations in hormone-resistant breast cancer. *Nat Genet* **2013** 45:1439-45.
9. Jeselsohn R, Yelensky R, Buchwalter G, Frampton G, Meric-Bernstam F, Gonzalez-Angulo AM, *et al.* Emergence of constitutively active estrogen receptor- α mutations in pretreated advanced estrogen receptor-positive breast cancer. *Clin Cancer Res* **2014** 20:1757-67.
10. Pejerrey SM, Dustin D, Kim JA, Gu G, Rechoum Y, Fuqua SAW. The Impact of ESR1 Mutations on the Treatment of Metastatic Breast Cancer. *Horm Cancer* **2018** 9:215-28.
11. Razavi P, Chang MT, Xu G, Bandlamudi C, Ross DS, Vasan N, *et al.* The Genomic Landscape of Endocrine-Resistant Advanced Breast Cancers. *Cancer Cell* **2018** 34:427-38.
12. Finlay-Schultz J, Sartorius CA. Steroid hormones, steroid receptors, and breast cancer stem cells. *J Mammary Gland Biol Neoplasia* **2015** 20:39-50.
13. Sansone P, Ceccarelli C, Berishaj M, Chang Q, Rajasekhar VK, Perna F, *et al.* Self-renewal of CD133(hi) cells by IL6/Notch3 signalling regulates endocrine resistance in metastatic breast cancer. *Nat Commun* **2016** 7:10442.
14. Sansone P, Berishaj M, Rajasekhar VK, Ceccarelli C, Chang Q, Strillacci A, *et al.* Evolution of Cancer Stem-like Cells in Endocrine-Resistant Metastatic Breast Cancers Is Mediated by Stromal Microvesicles. *Cancer Res* **2017** 77:1927-41.
15. Garcia-Martinez L, Zhang Y, Nakata Y, Chan HL, Morey L. Epigenetic mechanisms in breast cancer therapy and resistance. *Nat Commun* **2021** 12:1786.
16. Foulkes WD, Smith IE, Reis-Filho JS. Triple-negative breast cancer. *N Engl J Med* **2010** 363:1938-48.
17. Shen M, Pan H, Chen Y, Xu YH, Yang W, Wu Z. A review of current progress in triple-negative breast cancer therapy. *Open Med (Wars)* **2020** 15:1143-49.
18. Treeck O, Schöler-Toprak S, Ortmann O. Estrogen Actions in Triple-Negative Breast Cancer. *Cells* **2020** 9:2358.
19. Howell A, Osborne CK, Morris C, Wakeling AE. ICI 182,780 (Faslodex): development of a novel, "pure" antiestrogen. *Cancer* **2000** 89:817-25.

20. Toy W, Weir H, Razavi P, Lawson M, Goeppert AU, Mazzola AM, *et al.* Activating ESR1 Mutations Differentially Affect the Efficacy of ER Antagonists. *Cancer Discov* **2017** 7:277-87.
21. Chandarlapaty S, Sawai A, Ye Q, Scott A, Silinski M, Huang K, *et al.* SNX2112, a synthetic heat shock protein 90 inhibitor, has potent antitumor activity against HER kinase-dependent cancers. *Clin Cancer Res* **2008** 14:240-8.
22. Hon CC, Ramiłowski JA, Harshbarger J, Bertin N, Rackham OJ, Gough J, *et al.* An atlas of human long non-coding RNAs with accurate 5' ends. *Nature* **2017** 543:199-204.
23. Mayr C, Bartel DP. Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* **2009** 138:673-84.
24. Wang Z, Zhang X, Shen P, Loggie BW, Chang Y, Deuel TF. Identification, cloning, and expression of human estrogen receptor- α 36, a novel variant of human estrogen receptor- α 66. *Biochem Biophys Res Commun* **2005** 336:1023-7.
25. Wang Q, Jiang J, Ying G, Xie XQ, Zhang X, Xu W, *et al.* Tamoxifen enhances stemness and promotes metastasis of ER α 36+ breast cancer by upregulating ALDH1A1 in cancer cells. *Cell Res* **2018** 28:336-58.
26. Lin CY, Ström A, Vega VB, Kong SL, Yeo AL, Thomsen JS, *et al.* Discovery of estrogen receptor α target genes and response elements in breast tumor cells. *Genome Biol* **2004** 5:R66.
27. Bellerby R, Smith C, Kyme S, Gee J, Günthert U, Green A, *et al.* Overexpression of Specific CD44 Isoforms Is Associated with Aggressive Cell Features in Acquired Endocrine Resistance. *Front Oncol* **2016** 6:145.
28. Domenici G, Aurrekoetxea-Rodríguez I, Simões BM, Rábano M, Lee SY, Millán JS, *et al.* A Sox2-Sox9 signalling axis maintains human breast luminal progenitor and breast cancer stem cells. *Oncogene* **2019** 38:3151-69.
29. Varley KE, Gertz J, Roberts BS, Davis NS, Bowling KM, Kirby MK, *et al.* Recurrent read-through fusion transcripts in breast cancer. *Breast Cancer Res Treat* **2014** 146:287-97.
30. Piscuoglio S, Ng CKY, Geyer FC, Burke KA, Cowell CF, Martelotto LG, *et al.* Genomic and transcriptomic heterogeneity in metaplastic carcinomas of the breast. *NPJ Breast Cancer* **2017** 3:48.
31. Khan SA, Rogers MA, Khurana KK, Meguid MM, Numann PJ. Estrogen receptor expression in benign breast epithelium and breast cancer risk. *J Natl Cancer Inst* **1998** 90:37-42.
32. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **2012** 490:61-70.
33. Kaminska K, Akrap N, Staaf J, Alves CL, Ehinger A, Ebbesson A, *et al.* Distinct mechanisms of resistance to fulvestrant treatment dictate level of ER independence and selective response to CDK inhibitors in metastatic breast cancer. *Breast Cancer Res* **2021** 23:26.
34. Giessrigl B, Schmidt WM, Kalipciyan M, Jeitler M, Bilban M, Gollinger M, *et al.* Fulvestrant induces resistance by modulating GPER and CDK6 expression: implication of methyltransferases, deacetylases and the hSWI/SNF chromatin remodelling complex. *Br J Cancer* **2013** 109:2751-62.
35. Horigome T, Golding TS, Quarmby VE, Lubahn DB, McCarty K Sr, Korach KS. Purification and characterization of mouse uterine estrogen receptor under conditions of varying hormonal status. *Endocrinology* **1987** 121:2099-111.

36. Boonyaratanakornkit V, Hamilton N, Márquez-Garbán DC, Pateetin P, McGowan EM, Pietras RJ. Extranuclear signaling by sex steroid receptors and clinical implications in breast cancer. *Mol Cell Endocrinol* **2018** 466:51-72.
37. Omarjee S, Jacquemetton J, Poulard C, Rochel N, Dejaegere A, Chebaro Y, *et al.* The molecular mechanisms underlying the ER α -36-mediated signaling in breast cancer. *Oncogene* **2017** 36:2503-14.
38. Zhou Z, Zhou J, Du Y. Estrogen receptor alpha interacts with mitochondrial protein HADHB and affects beta-oxidation activity. *Mol Cell Proteomics* **2012** 11:M111.011056.
39. O'Brien CS, Farnie G, Howell SJ, Clarke RB. Breast cancer stem cells and their role in resistance to endocrine therapy. *Horm Cancer* **2011** 2:91-103.
40. Leung EY, Askarian-Amiri ME, Sarkar D, Ferraro-Peyret C, Joseph WR, Finlay GJ, *et al.* Endocrine Therapy of Estrogen Receptor-Positive Breast Cancer Cells: Early Differential Effects on Stem Cell Markers. *Front Oncol* **2017** 7:184.
41. Yan Y, Yu L, Castro L, Dixon D. ER α 36, a variant of estrogen receptor α , is predominantly localized in mitochondria of human uterine smooth muscle and leiomyoma cells. *PLoS One* **2017** 12:e0186078.
42. Strillacci A, Griffoni C, Sansone P, Paterini P, Piazzzi G, Lazzarini G, *et al.* MiR-101 downregulation is involved in cyclooxygenase-2 overexpression in human colon cancer cells. *Exp Cell Res* **2009** 315:1439-47.
43. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **2013** 29:15-21.
44. Anders S, Reyes A, Huber W. Detecting differential usage of exons from RNA-seq data. *Genome Res* **2012** 22:2008-17.
45. de Hoon MJ, Imoto S, Nolan J, Miyano S. Open source clustering software. *Bioinformatics* **2004** 20:1453-4.
46. Saldanha AJ. Java Treeview--extensible visualization of microarray data. *Bioinformatics* **2004** 20:3246-8.
47. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* **2016** 34:525-7.
48. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **2003** 13:2498-504.
49. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **2015** 31:166-9.

Figure 1

Figure 2

Figure 2.

ERα-LBD is encoded by an ESR1 transcript variant. **A-D**, Total RNA was extracted from different BC cell lines and analyzed by RNA capture sequencing (capture-seq). Data output was used to investigate differential expression (DE) of ESR1 exons and transcript variants among samples. **A**, Heat map showing ESR1 exon expression calculated by using DEXSeq-count tool (log2, mean-centered). Exons encoding ERα-FL protein are indicated in black (CDS). BC cell lines were ranked based on their known ERα-FL protein expression and grouped as ERα-FL(+), including T-47D, MCF-7, PDX-ERα(+) (in purple), and ERα-FL(-), including MDA-MB-453, MDA-MB-231, PDX-ERα(-) (in teal). **B**, Differential ESR1 exon expression between ERα-FL(+) and ERα-FL(-) groups was calculated using DEXSeq and plotted using plotDEXSeq tools. Exons associated with two different ESR1 transcript variants were taken under consideration for the analysis: ESR1-FL (main variant encoding ERα-FL) and ESR1-LBD (predicted, encoding ERα-LBD). Expression values for ERα-FL(+) and ERα-FL(-) groups are plotted in purple and in teal, respectively. BH adjusted *P*-values: **P* < 0.05, ***P* < 0.01 (*n* = 6/group). **C**, Heat map showing ESR1 transcript (ESR1-FL and ESR1-LBD) expression among BC cell lines, ranked as ERα-FL(+) (purple) and ERα-FL(-) (teal). Transcript expression was calculated by using Kallisto tool and quantified as counts / transcript length. **D**, ESR1-LBD transcript expression plotted as log2 fold change, relative to ESR1-FL. Color code: grey, ESR1-FL; black, ESR1-LBD. Mean ± s.e.m. is shown (*n* = 6); ** *P* < 0.01, ratio paired *t* test, two-sided.

Figure 3

Figure 3.

ER α -LBD localizes in the cytoplasm and mitochondria of breast cancer cells. **A**, Representative images of ER α staining by immunohistochemistry (IHC) in six different BC cell lines. All histological sections were counterstained with hematoxylin. Scale bar representing 50 μ m is shown. Magnification: 40X. **B**, Representative confocal images of BC cells stained for ER α (green), mitochondria/OXPHOS (red) and nuclei/DAPI (blue). For each cell line, merged images and colocalization images are also shown. Color code for colocalization: cyan (ER α -Nuclei, N); yellow (ER α -Mitochondria, M). A scale bar for each image representing 40 μ m is shown. Magnification: 63X.

Figure 4

A

B

C

MCF-7	
Enriched Pathways / Clustering	Proteins
Glycolysis and Gluconeogenesis	ALDOA, ATP5A1, ATP5O, CALM1, ENO1, ESR1, G6PD, GAPDH, HSD17B4, HSP90AA1, HSPB1, LDHA, MTHFD1, NME1, NPM1, NUP210, PGK1, PKM, SLC25A1
Signaling	ALDOA, CALM1, CANX, COL1A1, EIF2S2, ESR1, FLNA, G6PD, HSP90AA1, HSP90AB1, HSPB1, HSPD1, LDHA, NME1, NPM1, NUP210, PFN1, PGK1, PHB2, PPIA, RAB10, VDAC1, YWHAQ, YWHAQ
Mitochondrial Metabolism and Respiration	ATP5A1, ATP5O, HSPD1, LDHA, NPM1, PHB2, PKM, VDAC1, VDAC2
Hypoxia and Angiogenesis	CALM1, COL1A1, EIF2S2, ESR1, FLNA, HSD17B4, HSP90AA1, HSP90AB1, LDHA, NRIP1
ERα Signaling	ESR1, GRN, HSPB1, NRIP1, PHB2
MYC Signaling	ALDOA, CANX, DSP, EIF2S2, ENO1, G6PD, GAPDH, HSP90AA1, HSPD1, LDHA, NME1, NPM1, PGK1, PPIA

D

E

F

TNBC	
Enriched Pathways / Clustering	Proteins
Signaling	ALB, CD59, COL1A1, COL1A2, COL2A1, COL3A1, ESR1, HIST1H2BC, HSP90AA1, HSP90AB1, HSPA5, HSPA8, RAB10, RAB11A, RAB1B, TMED9, YWHAQ, YWHAQ
Integrins, Angiogenesis, and EMT	CD59, COL1A1, COL1A2, COL2A1, COL3A1, COL3A1, COL8A1, FLNA, HSP90AA1, HSP90AB1, HSPA8, PPIA, RAB10, SLC25A3, TPM1, VIM
Mitochondrial Metabolism and Respiration	ATP5A1, ATP5B, ATP5L, CANX, CD59, DDOST, DLAT, EE1G, EIF2S3, HIST1H2BC, HSPA5, HSPD1, IMPDH2, LDHB, LMNA, PABPC1, PARP1, PKM, RAB1B, RPN1, RPN2, TMED9, VDAC1, VDAC2
Glycolysis and Gluconeogenesis	ALB, ATP5A1, ATP5B, ATP5L, CYB5R3, DDOST, DLAT, EPHX1, EPRS, ESR1, G6PD, GAPDH, HADHB, HSD17B4, HSP90AA1, IMPDH2
Fatty Acid Metabolism	ALB, CRABP2, EPHX1, ESR1, HADHB, HSD17B4, HSP90AA1, NRIP1, PABPC1, RETSAT, SCF2, TSPO
ERα Signaling	ESR1, NRIP1, PHB2, VIM
mRNA Splicing	DDOST, DHX9, EE1G, EIF2S3, EPRS, ESR1, G6PD, HIST1H2BC, HNRNPA3, HNRNP, HSP90AA1, HSPA8, NPM1, PABPC1, PARP1, PRDX2, RPN1, RPN2, SRSF11, U2AF1, YBX1, YWHAQ, YWHAQ

Figure 4.

Predicted ERα-LBD network: protein-protein interactions (PPIs) and biological pathways. **A**, Heat map showing the enrichment level of proteins that preferentially bind ERα-LBD, based on ERα co-immunoprecipitation (Co-IP) experiments followed by LC/MS analysis. Samples: NC and NC +Fulv (negative controls); ERα-LBDoe, ERα-LBDoe +Fulv (overexpression). Fulv = 1 μM fulvestrant treatment, 24 h. Different comparative criteria between samples were used to identify groups of ERα-LBD protein interactors, color-coded as follow: ERα (blue); proteins present only in ERα-LBD samples (green, cyan and yellow); proteins enriched more in ERα-LBD samples compared to NC, in the absence or presence of fulvestrant treatment (red and purple, respectively). Protein names labelled in dark red color are shared between MCF-7 and TNBC models (see below). **B**, 'Enrichr' platform based on multiple libraries was used to discover cell pathways in which ERα-LBD PPIs are significantly involved (FDR Adj-P (Q) < 0.05), then clustered and annotated by using Cytoscape software and MCL algorithm. **C**, ERα-LBD PPIs contribution to clustered pathways. **D**, Same as in (A). Samples: MDA-MB-453/-231 ERα-LBDoe (overexpression); MDA-MB-453/-231 ERα-LBDkd (knockdown). ERα PPIs and protein groups, color-coded as follow: ERα (blue); proteins enriched in both TNBC cell lines overexpressing ERα-LBD, compared to ERα-LBD silenced cells (green and cyan); enriched in MDA-MB-453 ERα-LBDoe cells only (yellow); enriched in MDA-MB-231 ERα-LBDoe cells only (red). **E** and **F**, Same as in (B and C).

Figure 5

A

B

C

Figure 5.

Effect of ERα-LBD 'gain/ loss-of-function' on BC cell metabolism. **A**, Mitochondrial respiration and glycolysis levels in MCF-7 cell clones. Fulvestrant 1 μM (24 h pre-treatment) or vehicle (DMSO) was added to cells. ERα-LBD overexpressing (ERα-LBDoe; in red) cells were compared to controls (NC; in blue). Analyses were carried out using XF Cell Mito Stress Kit (Agilent). Respiration was evaluated as oxygen consumption rate (OCR, pmol/min), normalized on cell number. Glycolysis was evaluated as extracellular acidification rate (ECAR, mpH/min), normalized on cell number. Calculation of metabolic parameters was based on ΔOCR and ΔECAR values, following manufacturer's guidelines. **B** and **C**, Mitochondrial respiration and glycolysis levels in MDA-MB-453 and MDA-MB-231 cells. Cells with ERα-LBD knockdown (kd; in red) were compared to controls (NC; in blue). Analyses were carried out as described above. FCCP = p-trifluoromethoxy-phenylhydrazine; Rot = rotenone; ActmA = antimycinA. Data in the figures are presented as mean ± s.e.m. (*n* = 2 independent experiments). * *P* < 0.05, ** *P* < 0.01; unpaired *t* test, one-sided.

Figure 6

Figure 6.

ERα-LBD promotes *in vitro* and *in vivo* growth and fulvestrant resistance. **A-D**, *In vitro* proliferation of BC cell clones, using resazurin reagent and expressed as fluorescence intensity (absorbance at 590 nm). ERα-LBD overexpression (**A**) or knockdown (**B-D**) was compared to controls (NC). In (**A**) and (**B**), cell proliferation was also tested in the presence or absence of fulvestrant 1 μM treatment ($n = 3$ independent experiments). **E-H**, Xenografts of BC cells implanted into the mammary fat pads of NOD scid gamma (NSG) mice. Plots in (**E**) and (**F**) show mean tumor volumes (mm³) as a function of time (days). MCF-7 model: when tumors reached 100 mm³, mice were randomized ($n = 6$ /group) to weekly treatments with 200 mg/kg fulvestrant injected intramuscularly, and compared to control groups ($n = 6$ /group). MDA-MB-231 model: $n = 5$ /group. **F-H**, Images and size of tumors, isolated from animals after reaching endpoint. In (**A**) and (**E**), NC is in blue and ERα-LBD overexpression in red. In (**B-D**) and (**G**), NC is in blue and ERα-LBD knockdown in red. All data in the figure are presented as mean ± s.e.m., ns = not significant, * $P < 0.05$, ** $P < 0.01$, two-way ANOVA.

Figure 7

Figure 7.

ER α -LBD expression is associated with proliferation, endocrine resistance, stemness and metabolism of breast cancer cells. **A-F**, Two different BC cell models were used to investigate cell phenotypes associated with ER α -LBD 'gain/ loss-of-function'. **A-C** MCF-7 model, with 24 h fulvestrant 1 μ M treatment (+Fulv): MCF-7 ER α -LBDoe, MCF-7 FulvRes and MCF-7 NC (control). **D-F** TNBC model: MDA-MB-453/-231 ER α -LBDkd (knockdown) and MDA-MB-453/-231 NC (controls). RNA samples were extracted from each cell line and analyzed by RNA-Seq ($n = 3$ /sample). **A**, Heat map showing differentially expressed genes in MCF-7 ER α -LBDoe (+Fulv) and MCF-7 FulvRes (+Fulv), compared to MCF-7 NC. Gene expression (counts) was determined by using HTSeq and differential expression was calculated as $\log_2$ (fold change). Only genes with >1 fold change were taken under consideration for the analysis. Hierarchical clustering based on Euclidean Distance was carried out. **B** GSEA analysis based on gene counts and showing gene sets that are both highly and significantly ($P < 0.05$) enriched in MCF-7 ER α -LBDoe (+Fulv) and MCF-7 FulvRes (+Fulv), compared to controls. NES = normalized enrichment score. Enriched gene sets were grouped and color-coded considering their association with specific gene set collections. Please refer to Methods section for details. **C**, Enrichment plots of interest, derived from GSEA analysis described in (**B**). **D**, Heat map showing differentially expressed genes in TNBC ER α -LBDkd cells, compared to controls (NC). Calculation and analysis was carried out as described in (**A**). **E**, GSEA analysis showing gene sets that are both highly and significantly ($P < 0.05$) enriched in TNBC NC samples compared to TNBC ER α -LBDkd, grouped/color-coded as in (**B**). **F**, Enrichment plots of interest, derived from GSEA analysis described in (**E**).

Figure 8

A

B

C

D

Quartiles	ERα(+)	ERα(+) ^{ADJ}	TNBC	TNBC ^{ADJ}	MpBC
Q4	40.5%	0.0%	31.0%	0.0%	35.3%
Q3	35.7%	16.7%	19.0%	0.0%	58.8%
Q2	19.0%	40.0%	23.8%	23.8%	5.9%
Q1	4.8%	43.3%	26.2%	76.2%	0.0%

Quartiles	ERα(+)	ERα(+) ^{ADJ}	TNBC	TNBC ^{ADJ}	MpBC
Q4 - Q3	76.2%	16.7%	50.0%	0.0%	94.1%
Q2 - Q1	23.8%	83.3%	50.0%	100.0%	5.9%

E

F

G

H

Quartiles	ERα(+)		ERα(-)	
	HER2(+)	HER2(-)	HER2(+)	HER2(-)
Q4	26.8%	31.9%	2.5%	7.1%
Q3	33.3%	26.1%	10.0%	17.5%
Q2	23.6%	24.3%	27.5%	27.8%
Q1	16.3%	17.7%	60.0%	47.6%

Quartiles	ERα(+)		ERα(-)	
	HER2(+)	HER2(-)	HER2(+)	HER2(-)
Q4 - Q3	60.2%	58.0%	12.5%	24.6%
Q2 - Q1	39.8%	42.0%	87.5%	75.4%

Figure 8.

ESR1-LBD expression is identified in human breast cancers. **A**, RNA-seq samples from BC patients were grouped as follow: ER α (+) BC primary tumors ($n = 42$; in purple); uninvolved breast tissue adjacent to ER α (+) primary tumors (ER α ^{ADJ}, $n = 30$; in light purple); TNBC primary tumors ($n = 42$; in teal); uninvolved breast tissue adjacent to TNBC primary tumors (TNBC^{ADJ}, $n = 21$; in light teal); metaplastic BC primary tumors (MpBC, $n = 17$; in blue). For each group/sample, ESR1-FL transcript expression was calculated by using Kallisto and plotted as Transcripts Per Kilobase Million (TPM). **B**, Heat map showing ESR1 exon expression (including 5'/3'-UTR regions) in BC patient groups. Exon expression was calculated by using DEXSeq-count tool [# of read counts/exon length] and scaled by column. For each group, columns represent BC samples and were sorted by ESR1-FL expression (calculated as mean of all CDS exons), increasing from the left to the right side of the maps. CDS exons are indicated in black and untranslated regions (UTR) in grey. **C**, Dispersion plot for ESR1 transcript variants expression (ESR1-FL vs. ESR1-LBD), calculated as TPM. Dashed lines corresponding to ESR1-LBD expression median and Q1-Q3 quartiles are shown. Color code as described in **(A)**. **D**, Percentage of samples from each group, distributed into quartiles. Percentage of samples above the median (Q4+Q3) and below the median (Q2+Q1) is also shown. **E**, RNA-seq samples from Breast Invasive Carcinoma study (TCGA-BRCA) were assigned to four different groups, based on ER α and HER2 IHC staining: ER α (+)/HER2(+) ($n = 123$; in light purple), ER α (+)/HER2(-) ($n = 436$; in purple), ER α (-)/HER2(+) ($n = 40$; in light teal) and ER α (-)/HER2(-) ($n = 126$; in teal). ESR1-FL transcript levels were calculated and plotted as described in **(A)**. **F**, **G** and **H**, As described in **B**, **C** and **D** (respectively). * $P < 0.05$, ** $P < 0.01$, unpaired t test, two-sided.