

## ARTICLE OPEN



# ER $\alpha$ -LBD, an isoform of estrogen receptor alpha, promotes breast cancer proliferation and endocrine resistance

Antonio Strillacci<sup>1,2,3</sup>, Pasquale Sansone<sup>1,4</sup>, Vinagolu K. Rajasekhar<sup>5</sup>, Mesruh Turkecul<sup>6</sup>, Vitaly Boyko<sup>6</sup>, Fanli Meng<sup>7</sup>, Brian Houck-Loomis<sup>7</sup>, David Brown<sup>7</sup>, Michael F. Berger<sup>7</sup>, Ronald C. Hendrickson<sup>8</sup>, Qing Chang<sup>9</sup>, Elisa de Stanchina<sup>9</sup>, Fresia Pareja<sup>10</sup>, Jorge S. Reis-Filho<sup>10</sup>, Ramya Segu Rajappachetty<sup>1,2</sup>, Isabella Del Priore<sup>1,2</sup>, Bo Liu<sup>1,2</sup>, Yanyan Cai<sup>2</sup>, Alex Penson<sup>2</sup>, Chiara Mastroleo<sup>11</sup>, Marjan Berishaj<sup>11</sup>, Francesca Borsetti<sup>12</sup>, Enzo Spisni<sup>3</sup>, David Lyden<sup>4</sup>, Sarat Chandarlapaty<sup>1,2,13</sup> and Jacqueline Bromberg<sup>1,13</sup>

Estrogen receptor alpha (ER $\alpha$ ) 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 an ER $\alpha$  isoform, ER $\alpha$ -LBD, that is encoded by an alternative *ESR1* transcript and lacks the activation function and DNA binding domains. Despite lacking the transcriptional activity, ER $\alpha$ -LBD is found to promote breast cancer growth and resistance to the ER $\alpha$  antagonist fulvestrant. ER $\alpha$ -LBD is predominantly localized to the cytoplasm and mitochondria of BC cells and leads to enhanced glycolysis, respiration and stem-like features. Intriguingly, ER $\alpha$ -LBD expression and function does not appear to be restricted to cancers that express full length ER $\alpha$  but also promotes growth of triple-negative breast cancers and ER $\alpha$ -LBD transcript (*ESR1*-LBD) is also present in BC samples from both ER $\alpha$ (+) and ER $\alpha$ (-) human tumors. These findings point to ER $\alpha$ -LBD as a potential mediator of breast cancer progression and therapy resistance.

npj Breast Cancer (2022)8:96; <https://doi.org/10.1038/s41523-022-00470-6>

## INTRODUCTION

The nuclear receptor estrogen receptor alpha (ER $\alpha$ ) is the major therapeutic target for the ~75% of breast cancers (BC) where its expression is detected<sup>1</sup>. ER $\alpha$  is a multifunctional protein that contributes to cellular processes via transcriptional regulation, participation in signaling complexes<sup>2,3</sup> and regulation of mitochondrial function<sup>4,5</sup>. 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 $\alpha$  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 $\alpha$  antagonists (e.g. tamoxifen, fulvestrant)<sup>6–8</sup>. 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<sup>9–12</sup> that restore ER $\alpha$  transcriptional activity. In addition, cancers can bypass ER $\alpha$  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<sup>13–17</sup>. In addition, epigenetic mechanisms of resistance may also be contributory and whether these can affect ER $\alpha$  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<sup>13,18</sup>.

About 15% of BC cases are triple-negative breast cancers (TNBCs), lacking the expression of ER $\alpha$ , PR and HER2-amplification and are thus not responsive to targeted therapies against ER $\alpha$  or HER2. These cancers are associated with a poorer prognosis due to a higher rate of metastatic progression and resistance to treatment<sup>19,20</sup>. Notably, TNBC may be responsive to estrogen stimulation via ER $\alpha$ -independent pathways, promoting tumor formation and progression via different molecular mechanisms<sup>21</sup>.

In this work, we investigated ER $\alpha$  protein and transcripts in *ESR1* wild-type 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 an estrogen receptor $\alpha$ isoform

In order to investigate novel mechanisms of hormonal therapy resistance (HTR) in breast cancer, we examined the expression of ER $\alpha$  protein in different BC cell lines, either in the absence or

<sup>1</sup>Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>2</sup>Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>3</sup>Department of Biological, Geological and Environmental Sciences, University of Bologna, Bologna, Italy. <sup>4</sup>Children's Cancer and Blood Foundation Laboratories, Weill Cornell Medicine, New York, NY, USA. <sup>5</sup>Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>6</sup>Molecular Cytology Core Facility, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>7</sup>Center for Molecular Oncology, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>8</sup>Microchemistry and Proteomics, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>9</sup>Antitumor Assessment Core Facility, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>10</sup>Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>11</sup>Cancer Biology and Genetics Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>12</sup>Department of Pharmacy and Biotechnology, University of Bologna, Bologna, Italy. <sup>13</sup>Department of Medicine, Weill Cornell Medicine, New York, NY, USA.

<sup>✉</sup>email: [chandars@mskcc.org](mailto:chandars@mskcc.org); [bromberj@mskcc.org](mailto:bromberj@mskcc.org)

presence of the selective estrogen receptor degrader (SERD) fulvestrant (ICI 182,780) which reduces ER $\alpha$ -FL (ER $\alpha$ -FL, 66 kDa) protein levels as a consequence of reduced stability and dimerization<sup>22,23</sup>. We postulated that changes in fulvestrant-mediated suppression of ER $\alpha$ -FL levels may promote drug resistance and investigated this response in several models selected for fulvestrant resistance (MCF-7 FulvRes, MCF-7 Y537S, two breast cancer PDX models, and T-47D FulvRes)<sup>17,24</sup>, compared to fulvestrant-sensitive cells (Fig. 1a and Supplementary Fig. 1a, b). After 24 h exposure to 1  $\mu$ M fulvestrant, MCF-7 cells, and MCF-7 FulvRes showed ~80% and ~60% loss of ER $\alpha$ -FL expression, respectively; T-47D and T-47D FulvRes showed a ~60% and ~30% reduction respectively; MCF-7 Y537S and PDX-ER $\alpha$ (+) showed no decrease or slight decrease (~30%), respectively. PDX-ER $\alpha$ (-) 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 $\alpha$  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 $\alpha$ -FL expression and therefore resistant to fulvestrant treatment (Supplementary Fig. 1a). Importantly, both TNBC cell lines displayed this lower molecular weight protein whose expression increased after fulvestrant treatment (Fig. 1b). The specificity of the ER $\alpha$ -related ~37 kDa band was confirmed by using two different antibodies both raised against the C-terminus of ER $\alpha$  protein (Fig. 1a, b and Supplementary Fig. 1c). To establish the identity of the protein, we performed immunoprecipitation and mass spectrometry experiments from MCF-7 FulvRes and MDA-MB-231 cells and determined that the ER $\alpha$ -related protein is a truncated isoform of ER $\alpha$ -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 and Supplementary Fig. 1d). Based on our observation, we named this ER $\alpha$  protein isoform 'ER $\alpha$ -LBD'. ER $\alpha$ -LBD has 332 amino acids, a MW of 37.3 kDa, and a predicted 3-D structure that is similar to the predicted ER $\alpha$ -FL (Fig. 1c and Supplementary Fig. 1e). As expected, we failed to detect a 37.3 kDa protein using an ER $\alpha$  antibody raised against the N-terminus (Supplementary Fig. 1f).

Across all the models we investigated, we found that fulvestrant increased expression of ER $\alpha$ -LBD and in some cases was necessary for detection. These data suggested that fulvestrant was serving to stabilize ER $\alpha$ -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 $\alpha$ -LBD and ER $\alpha$ -FL expression upon HSP90 inhibition using an ATP inhibitor of HSP90 function (SNX-2112)<sup>25,26</sup>. Following SNX-2112 treatment, both ER $\alpha$ -FL and ER $\alpha$ -LBD underwent degradation, either in the presence (lane 2, media supplemented with regular 10% FBS) or absence (lane 3, media supplemented with 10% charcoal-stripped FBS) of physiological E2 levels (pM range). Although, ER $\alpha$ -LBD degradation was more rapid, especially in the absence of E2 (lane 3, CSS). Conversely, high levels of E2, fulvestrant or tamoxifen treatment (1  $\mu$ M) led to increased stability of ER $\alpha$ -LBD. (Fig. 1d and Supplementary Fig. 1g, h).

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

### ER $\alpha$ -LBD is encoded by an ESR1 transcript variant identified in both cell lines and primary breast cancers

To understand the basis of ER $\alpha$ -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 $\alpha$ -FL

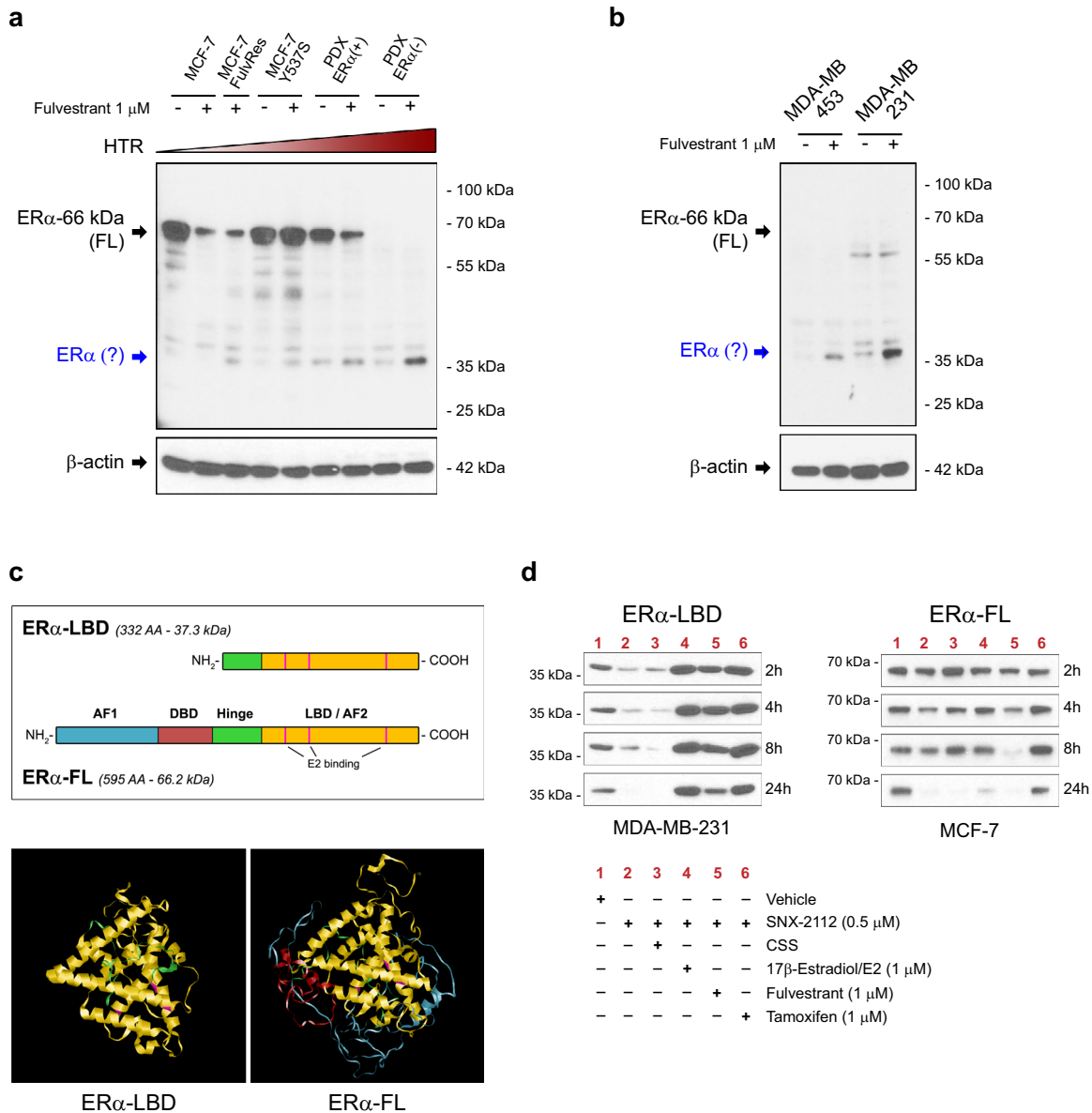
positive and negative cells. High-depth capture-sequencing data showed that BC ER $\alpha$ -FL(-) cells expressing exclusively ER $\alpha$ -LBD, such as MDA-MB-453/-231 and PDX-ER $\alpha$ , have higher abundance of ESR1 read counts in the C-terminus (CDS exons #4, 5, 6, 7 and 8), compared to the N-terminus (CDS exons #1-3). In contrast, ESR1 exon usage profile was significantly different in ER $\alpha$ -FL(+) cells, such as T-47D, MCF-7, and PDX-ER $\alpha$ (+), where the expression of exons #1-3 was higher than exons #4-8 (Fig. 2a and Supplementary Data 1). To identify an ESR1 transcript variant for ER $\alpha$ -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. 2a)<sup>27,28</sup>. 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 $\alpha$ (-) cells, compared to ER $\alpha$ (+) (Fig. 2a). Differential exon usage was determined for ESR1-FL (encoding ER $\alpha$ -FL) and a predicted ESR1-LBD (encoding ER $\alpha$ -LBD) transcript by comparing ER $\alpha$ -FL(+) and ER $\alpha$ (-) 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 $\alpha$ (-) group (Fig. 2b).

Exon data were supported by the analysis of ESR1 transcripts expression, demonstrating higher levels of ESR1-LBD transcript in ER $\alpha$ -FL(-) cells, compared to ESR1-FL (Fig. 2c, d and Supplementary Data 1). Also, qPCR (exon 'walking') and RT-PCR assays confirmed the RNA-seq results. ER $\alpha$ -FL(+) and ER $\alpha$ -FL(-) cells express an ESR1 mRNA ranging from exon 1 to 8 and from exon 4 to exon 8, respectively (Supplementary Fig. 2b, c). These results are in accordance with the western blot data in Fig. 1b, showing that ER $\alpha$ -FL(-) cells express a truncated ER $\alpha$  protein (LBD) and not full-length ER $\alpha$ . As definitive proof, the stable cloning of the putative ER $\alpha$ -LBD CDS into BC cells (exon 4 to exon 8), led to ER $\alpha$ -LBD overexpression (oe) whereas targeted genomic deletions within exon 4 (by CRISPR technology) induced ER $\alpha$ -LBD knockdown (kd) (Supplementary Fig. 2d). Alignment studies performed on public databases (UCSC Genome Browser) showed that the genomic region surrounding exon E3a is characterized by high-density transcription binding sites and candidate cis-regulatory elements (Supplementary Fig. 2e). Thus, we collected preliminary data based on luciferase-reporter assays to support the hypothesis of a putative alternative ESR1 promoter within exon E3a genomic region. Our data demonstrated the presence of transcriptional activity within the ESR1-LBD 5'UTR region (E3a) in both ER $\alpha$ (+) and ER $\alpha$ (-) cell lines (Supplementary Fig. 2f).

RNA sequencing and qPCR data showed reduced expression of the ESR1 3'UTR sequence in ER $\alpha$ -FL(-) cells as compared to ER $\alpha$ -FL(+) ones (Fig. 2a, b and Supplementary Fig. 2b). 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 transfection by enhanced protein production<sup>29</sup>. Thus, we hypothesized that the partial/total loss of the 3'UTR sequence in ER $\alpha$ -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. 2g).

We note that ER $\alpha$ -LBD is not the only ER $\alpha$  transcript variant that has been observed in breast cancer. An ER $\alpha$ -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<sup>30,31</sup>. Importantly, ER $\alpha$ -LBD is distinct from ER $\alpha$ -36 as specific primers and an antibody for the ER $\alpha$ -36 isoform failed to detect ER $\alpha$ -LBD (Supplementary Fig. 2h, i).

In addition, we examined ESR1-FL /-LBD expression by analyzing RNA-seq samples from human BC specimens.



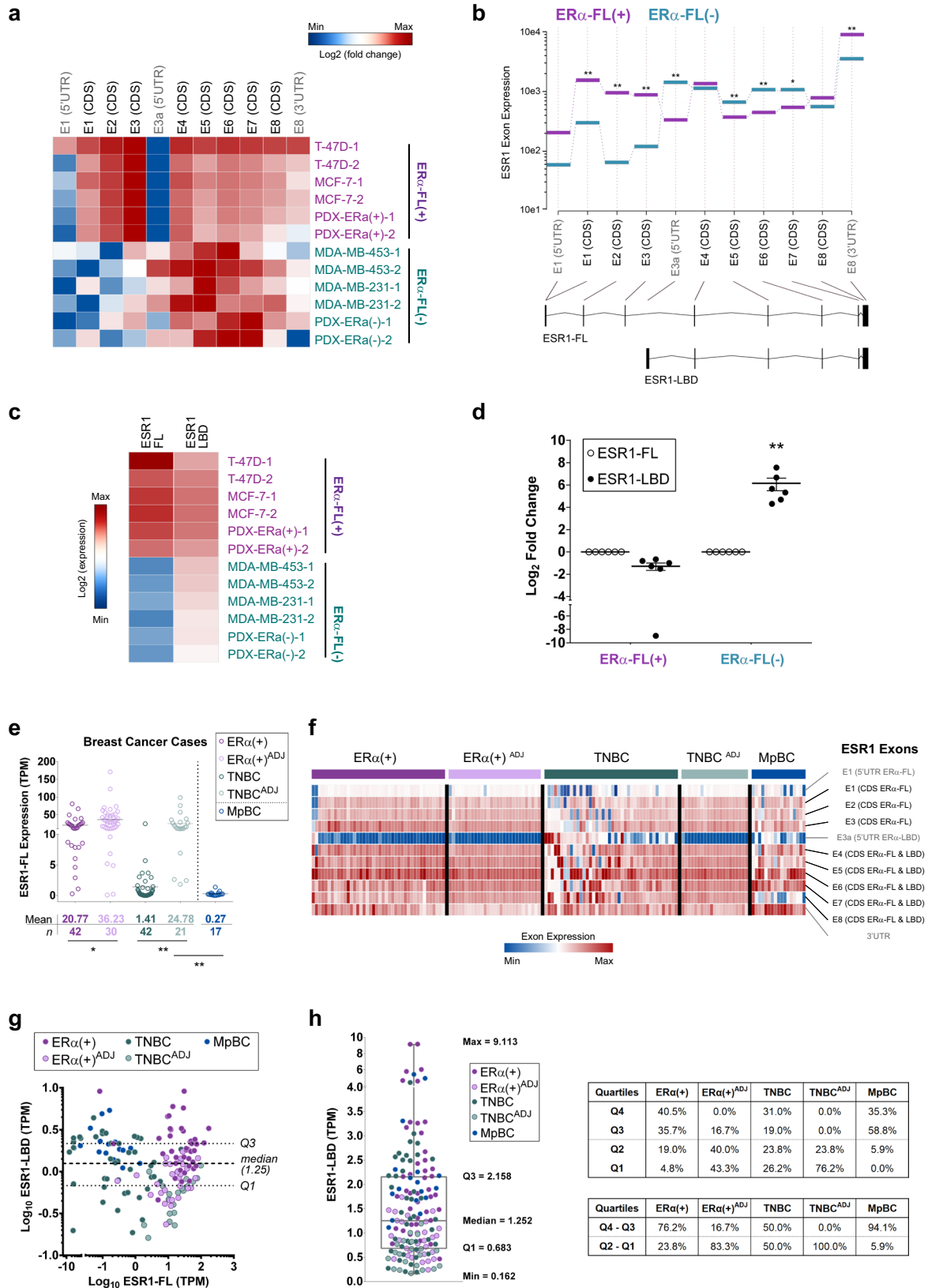
**Fig. 1 Fulvestrant resistant BC cell lines express ER $\alpha$ -LBD, a novel estrogen receptor  $\alpha$  isoform.** **a, b** Western blot analysis of estrogen receptor alpha (ER $\alpha$ ) protein expression in breast cancer (BC) cell lines. Cells were cultured for 24 h in the presence of vehicle (–) or fulvestrant 1  $\mu$ M (+) before lysis. Vehicle = DMSO. ER $\alpha$  protein expression was normalized against  $\beta$ -actin. Increasing levels of hormonal therapy resistance (HTR) are indicated. A potential smaller ER $\alpha$  isoform is indicated in blue. **c** Comparison between ER $\alpha$ -LBD and ER $\alpha$ -FL (full-length): domains (AF-1, transcription Activation Function-1; DBD, DNA Binding Domain; Hinge; LBD, Ligand Binding Domain; AF-2, transcription Activation Function-2), amino acid (AA) number, molecular weight, 3-D structure prediction (Phyre2). **d** ER $\alpha$  protein stability assay. ER $\alpha$ -LBD and ER $\alpha$ -FL protein levels were analyzed by western blot in MDA-MB-231 and MCF-7 cells, respectively. Protein samples were collected at different time points (2–24 h), after treatment. Physiological levels of E2 (pM range) were present in all samples, except samples in lane 3 and 4. Vehicle DMSO, CSS charcoal-stripped serum. Source data are provided as a Source Data file.

ESR1-FL overall expression was assessed on the following collection of samples: ER $\alpha$ (+) BC primary tumors ( $n = 42$ ); uninvolved breast tissue adjacent to ER $\alpha$ (+) primary tumors ( $n = 30$ ); TNBC primary tumors ( $n = 42$ ); uninvolved breast tissue adjacent to TNBC primary tumors ( $n = 21$ ); metaplastic BC primary tumors (MpBC,  $n = 17$ )<sup>32,33</sup>. As expected, ESR1-FL transcript was found to be markedly reduced in TNBC and MpBC samples (Fig. 2e). Notably, ER $\alpha$  expression is known to be present in normal breast epithelium<sup>34</sup>. 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. 2f). We then examined the levels of the

ESR1-LBD transcript in BC groups, relative to ESR1-FL (Fig. 2g). 76% of ER $\alpha$ (+), 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 $\alpha$ (+)<sup>ADJ</sup> and none of TNBC<sup>ADJ</sup> samples were found to have high ESR1-LBD expression (Fig. 2h). All expression data are summarized in Supplementary Data 2.

#### ER $\alpha$ -LBD localizes to the mitochondria and cytoplasm of BC cells

ER $\alpha$ -LBD lacks the N-terminal domain of the full-length receptor and we thus hypothesized it may have a distinctive cellular



location. Using ER $\alpha$  immunohistochemistry (IHC), full-length expressing MCF-7 cells show robust nuclear staining that is markedly reduced 24 h after fulvestrant treatment. By contrast, ER $\alpha$ -LBD expressing models MCF-7 FulvRes, MDA-MB-453-231,

and PDX-ER $\alpha$ (-) display punctate cytoplasmic foci (Fig. 3a). Based on this particular pattern, we hypothesized that ER $\alpha$ -LBD might be localized to the mitochondria and performed confocal analyses examining cells for ER $\alpha$  protein (green), OXPHOS (mitochondrial



**Fig. 2 ER $\alpha$ -LBD is encoded by an ESR1 transcript variant.** **a** Heat map showing ESR1 exon expression in BC cell lines, calculated by using DEXSeq-count tool (log<sub>2</sub>, mean-centered). Exons encoding ER $\alpha$ -FL protein are indicated in black (CDS). BC cell lines were grouped as ER $\alpha$ -FL(+), including T-47D, MCF-7, PDX-ER $\alpha$ (+) (in purple), and ER $\alpha$ -FL(-), including MDA-MB-453, MDA-MB-231, PDX-ER $\alpha$ (-) (in teal). **b** Differential ESR1 exon expression between ER $\alpha$ -FL(+) and ER $\alpha$ -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 ER $\alpha$ -FL) and ESR1-LBD (predicted, encoding ER $\alpha$ -LBD). Expression values for ER $\alpha$ -FL(+) and ER $\alpha$ -FL(-) groups are plotted in purple and in teal, respectively. BH adjusted *P*-values (FDR test): \**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, grouped as ER $\alpha$ -FL(+) (purple) and ER $\alpha$ -FL(-) (teal). Transcript expression was calculated by using Kallisto tool and quantified as counts/transcript length. **d** ESR1-LBD transcript expression plotted as log<sub>2</sub> fold change, relative to ESR1-FL. ESR1-FL, empty circles; ESR1-LBD, solid circles. Mean  $\pm$  s.e.m. is shown (*n* = 6); \*\**P* < 0.01, ratio paired *t* test, two-sided. **e** ESR1-FL transcript expression was calculated in BC samples by using Kallisto and plotted as Transcripts Per Kilobase Million (TPM). ER $\alpha$ (+) BC primary tumors (*n* = 42); uninvolved breast tissue adjacent to ER $\alpha$ (+) 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). **f** Heat map showing ESR1 exon expression in BC patient groups. Exon expression was calculated by using DEXSeq-count tool [# of read counts/exon length] and scaled by column. **g** 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 in **e**. **h** Box and whiskers plot representing the distribution of ESR1-LBD expression values (TPM) from BC samples. Median, Q1/Q3 and Min/Max values are shown. Color code as in **e**. Percentage of samples from each group is also shown, distributed into quartiles. Source data are provided as Supplementary Data 2 file.

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 $\alpha$ , nuclei or mitochondria. The overall IF staining confirmed data from the IHC: MCF-7 cells have a strong ER $\alpha$  nuclear localization which is decreased after fulvestrant treatment; MCF-7 FulvRes chronically treated with fulvestrant have both nuclear and mitochondrial ER $\alpha$  staining. In ER $\alpha$  full-length negative cells, ER $\alpha$ -LBD was largely absent from the nucleus and found in the cytoplasm with a substantial fraction co-localized with mitochondria. Mitochondrial ER $\alpha$ -LBD was also observed by western blot analysis of BC cell fractions confirming the association of ER $\alpha$ -LBD with mitochondria (associated to outer membrane or residing inside mitochondria) (Supplementary Fig. 3a, b).

ER $\alpha$ -LBD expression and localization were further supported by IHC analysis of a human breast cancer tissue microarray (TMA) which included primarily TNBC. To perform this work, we utilized two different antibodies: one targeting the N-terminus of ER $\alpha$  protein which is commonly used in the clinical setting and one targeting the C-terminus, which would detect both ER $\alpha$ -FL and ER $\alpha$ -LBD. We observed that cytosolic staining was much more apparent using the C-terminal antibody (43.7% of total cores; 36.8% of ER $\alpha$  nuclear null/low cores) as compared to the N-terminal antibody (9.4% of total cores; 2.1% of ER $\alpha$  nuclear null/low cores). Notably, levels of cytoplasmic staining were weak compared to nuclear staining of ER $\alpha$ (+) cores and heterogeneous in distribution (Supplementary Fig. 3c–e and Supplementary Data 3). This expression pattern is consistent with our observations by western blot and RNA analyses in ER $\alpha$ (-) models (Figs. 1 and 2).

Given the relative absence of ER $\alpha$ -LBD nuclear localization, we examined its role in canonical ER $\alpha$  signaling, and we observed none of the transcriptional activation functions of full-length ER $\alpha$ . For instance, overexpression of ER $\alpha$ -LBD had no effect on Estrogen Response Element (ERE) luciferase expression or ER $\alpha$  target gene transcription in both ER $\alpha$ -FL(+) and TNBC cells (Supplementary Fig. 3f–k)<sup>35</sup>.

To further establish the functional significance of cytoplasmic and mitochondrial ER $\alpha$ -LBD, we performed co-IP and mass spectrometry experiments aimed at identifying ER $\alpha$ -LBD protein–protein interactions (PPIs). We designed the experiment comparing MCF-7 cells overexpressing ER $\alpha$ -LBD with parental (NC), in the absence or presence of fulvestrant and we identified a total of 52 peptides associated with ER $\alpha$ -LBD (Fig. 4a). 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 $\alpha$ ), hypoxia and angiogenesis

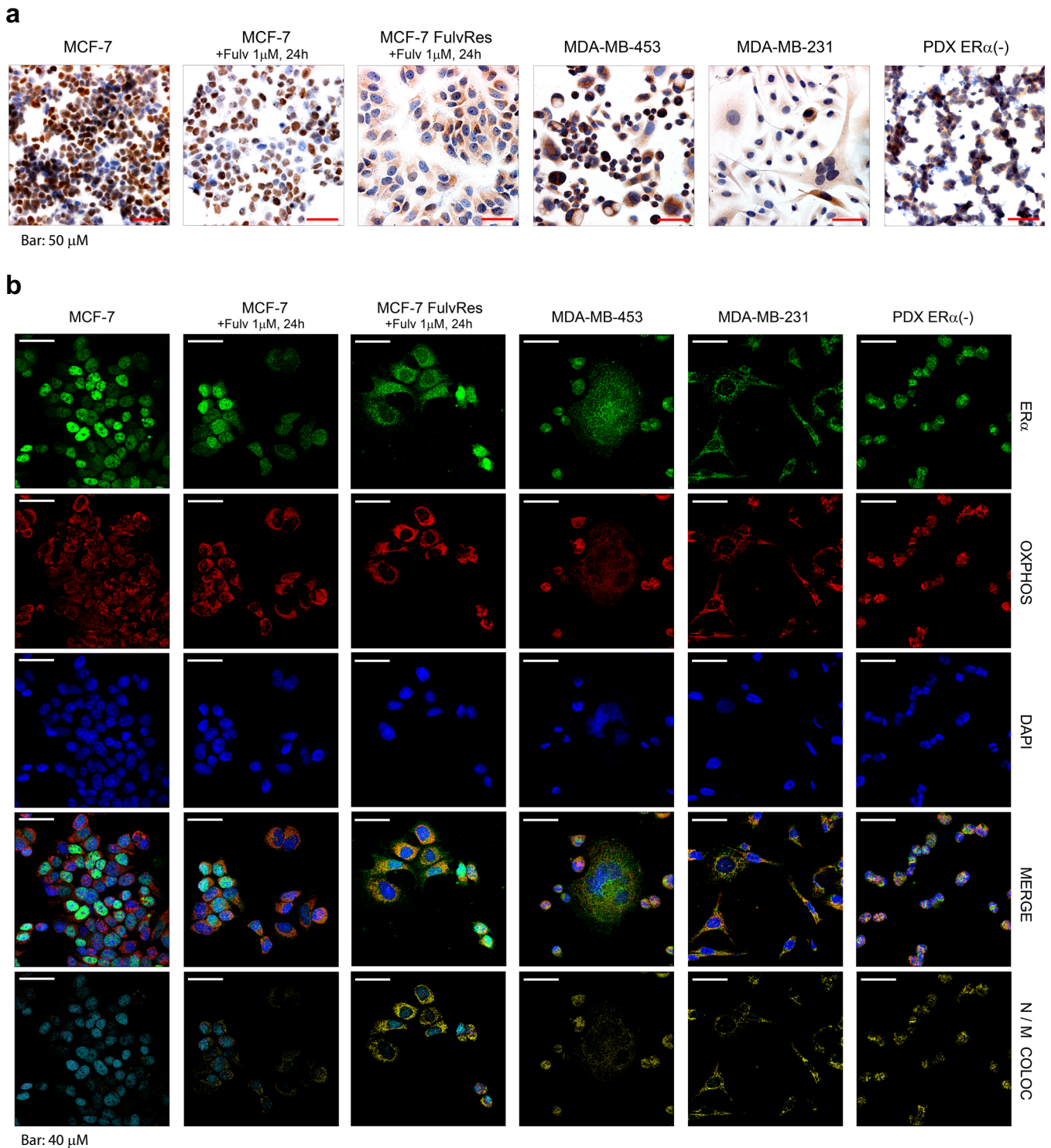
(HIF1 $\alpha$  and VEGFR signaling) and mitochondrial metabolism (oxphos and respiration) (Fig. 4b, c; Supplementary Data 4 and Supplementary Fig. 4a). Comparable experiments in TNBC models overexpressing or knocking down ER $\alpha$ -LBD identified a total of 93 peptides involved in ER $\alpha$ -LBD PPIs, and 29 of them were shared with those found in MCF-7 cells (Fig. 4d). Similarly to the MCF-7 model, enriched pathways included signaling, carbohydrate and mitochondrial metabolism (including also fatty acid  $\beta$ -oxidation) and angiogenesis (Fig. 4e, f; Supplementary Data 4 and Supplementary Fig. 4b).

### ER $\alpha$ -LBD regulates cell metabolism

The localization and protein–protein interaction analyses point to a potential metabolic function for ER $\alpha$ -LBD. We speculated ER $\alpha$ -LBD may play a role in regulating glycolysis and mitochondrial respiration, and so assessed the effect of ER $\alpha$ -LBD overexpression or knockdown on key metabolic parameters. In ER $\alpha$ -FL expressing MCF-7, control (NC) and ER $\alpha$ -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 $\alpha$ -FL and stabilization of ER $\alpha$ -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 and Supplementary Fig. 5a). In the TNBC models featuring knockdown of ER $\alpha$ -LBD, ER $\alpha$ -LBDkd cells were characterized by reduced levels of respiratory parameters (ATP production, basal and maximal respiration, spare respiratory capacity) and glycolysis (glycolytic capacity and reserve; basal glycolysis only in MDA-MB-231), compared to controls (NC) (Fig. 5b, c and Supplementary Fig. 5b, c). In TNBC models, OCR and ECAR were also tested in the presence of fulvestrant treatment and, despite the overall reduction of respiration (MDA-MB-453 and -231) and glycolysis (MDA-MD-231), metabolic parameters were significantly reduced in ER $\alpha$ -LBDkd cells (Fig. 5b, c and Supplementary Fig. 5b, c).

### ER $\alpha$ -LBD promotes growth and endocrine resistance, in vitro and in vivo

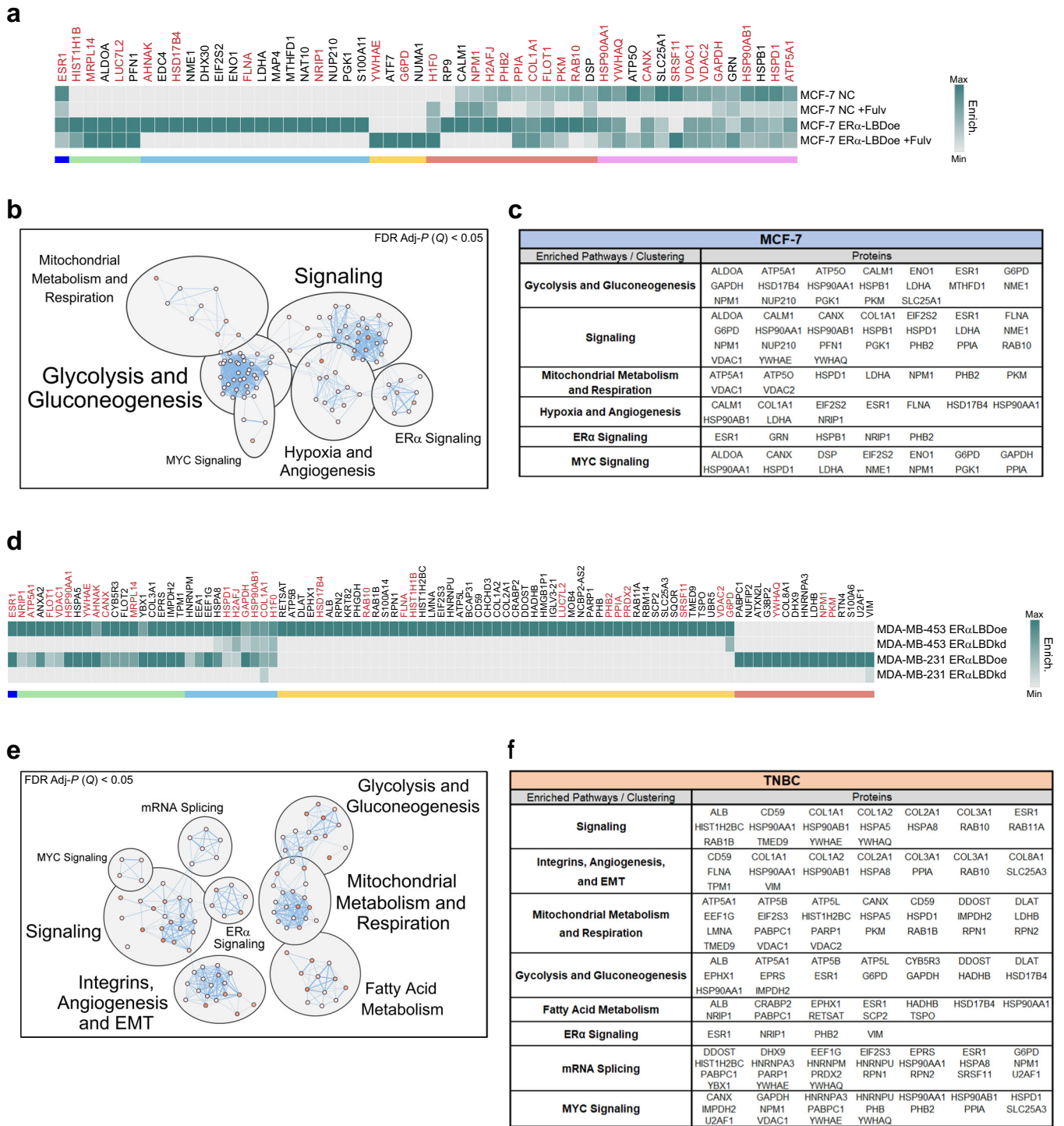
Given the impact of ER $\alpha$ -LBD on breast cancer cell metabolism, we investigated the effects of ER $\alpha$ -LBD on cell proliferation and response to endocrine therapy. In vitro, MCF-7 cells overexpressing ER $\alpha$ -LBD, compared to control cells (NC), showed no significant differences in growth. However, upon treatment with fulvestrant, NC cells were potently growth-inhibited (~80%



**Fig. 3 ER $\alpha$ -LBD localizes in the cytoplasm and mitochondria of breast cancer cells. a** Representative images of ER $\alpha$  staining by immunohistochemistry (IHC) in six different BC cell lines. All histological sections were counterstained with hematoxylin. Scale bar representing 50  $\mu$ m is shown. Magnification:  $\times 40$ . **b** Representative confocal images of BC cells stained for ER $\alpha$  (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 $\alpha$ -Nuclei, N); yellow (ER $\alpha$ -Mitochondria, M). A scale bar for each image representing 40  $\mu$ m is shown. Magnification:  $\times 63$ .

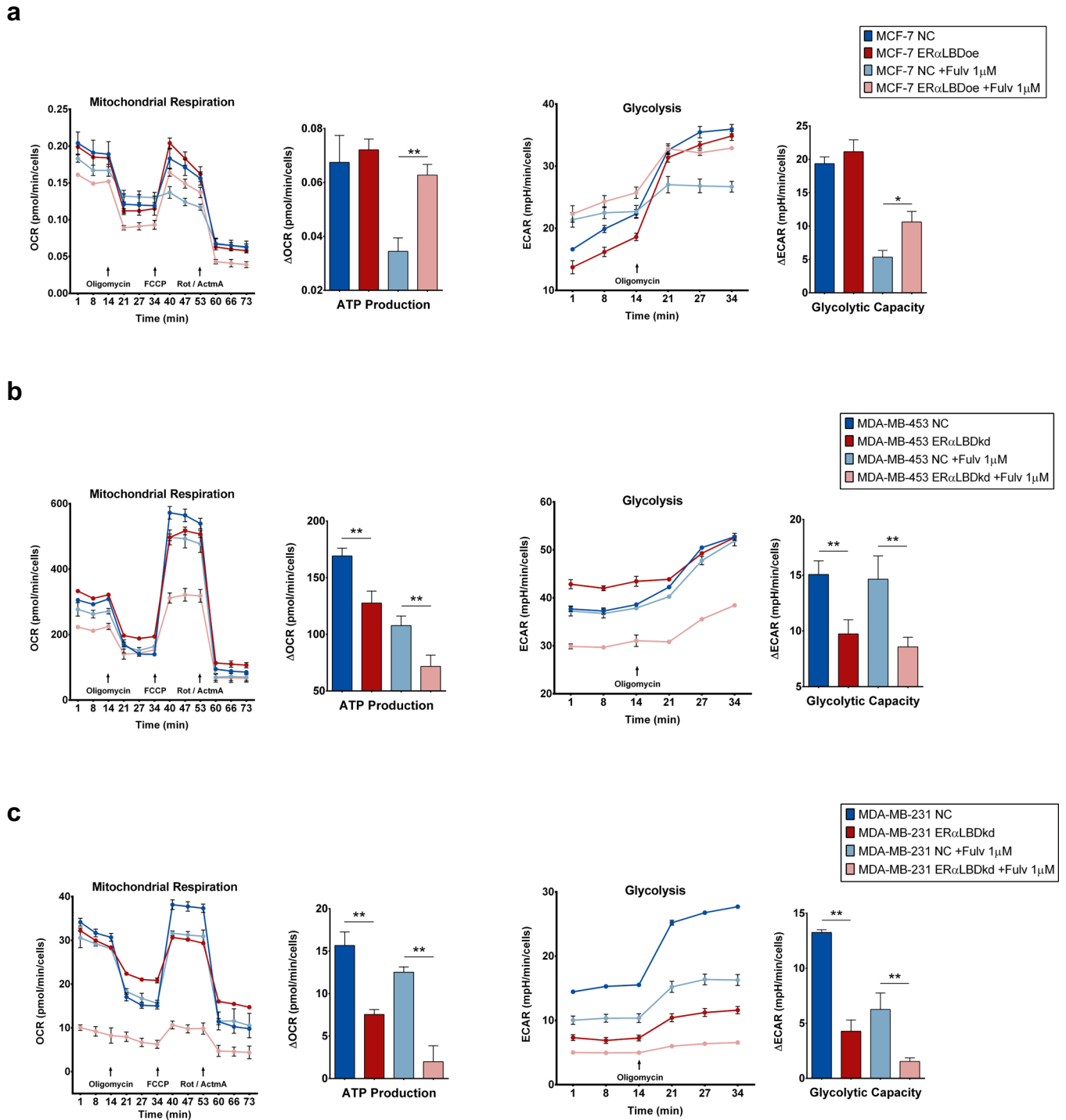
inhibition,  $P > 0.0001$ , vs. untreated, two-way ANOVA) while ER $\alpha$ -LBDoe cells were only partially growth-inhibited ( $\sim 40\%$  inhibition,  $P = 0.003$ , vs. untreated, two-way ANOVA), establishing a role for ER $\alpha$ -LBD in mediating endocrine resistance (Fig. 6a) that was confirmed also by treating MCF-7 NC and ER $\alpha$ -LBDoe cells with

tamoxifen (Supplementary Fig. 6a). Moreover, growth of MCF-7 FulvRes cells was significantly impaired after ER $\alpha$ -LBD knockdown (especially under fulvestrant treatment;  $P > 0.0001$ , two-way ANOVA) whereas no growth inhibition was observed in parental MCF7 FulvRes cells treated with fulvestrant, compared to



**Fig. 4 Predicted ER $\alpha$ -LBD network: protein–protein interactions (PPIs) and biological pathways.** **a** Heat map showing the enrichment level of proteins that preferentially bind ER $\alpha$ -LBD, based on ER $\alpha$  co-immunoprecipitation (Co-IP) experiments followed by LC/MS analysis. Samples: NC and NC + Fulv (negative controls); ER $\alpha$ -LBDoe, ER $\alpha$ -LBDoe +Fulv (overexpression). Fulv = 1  $\mu$ M fulvestrant treatment, 24 h. Different comparative criteria between samples were used to identify groups of ER $\alpha$ -LBD protein interactors, color-coded as follows: ER $\alpha$  (blue); proteins present only in ER $\alpha$ -LBD samples (green, cyan and yellow); proteins enriched more in ER $\alpha$ -LBD samples compared to NC, in the absence or presence of fulvestrant treatment (red and purple, respectively). Protein names labeled 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 ER $\alpha$ -LBD PPIs are significantly involved (Adj- $P$  ( $Q$ ) < 0.05, FDR test), then clustered and annotated by using Cytoscape software and MCL algorithm. **c** ER $\alpha$ -LBD PPIs contribution to clustered pathways. **d** Same as in **a**. Samples: MDA-MB-453/-231 ER $\alpha$ -LBDoe (overexpression); MDA-MB-453/-231 ER $\alpha$ -LBDkd (knockdown). ER $\alpha$  PPIs and protein groups, color-coded as follow: ER $\alpha$  (blue); proteins enriched in both TNBC cell lines overexpressing ER $\alpha$ -LBD, compared to ER $\alpha$ -LBD silenced cells (green and cyan); enriched in MDA-MB-453 ER $\alpha$ -LBDoe cells only (yellow); enriched in MDA-MB-231 ER $\alpha$ -LBDoe cells only (red). **e, f** Same as in **b, c**. Source data are provided as Supplementary Data 4 file.



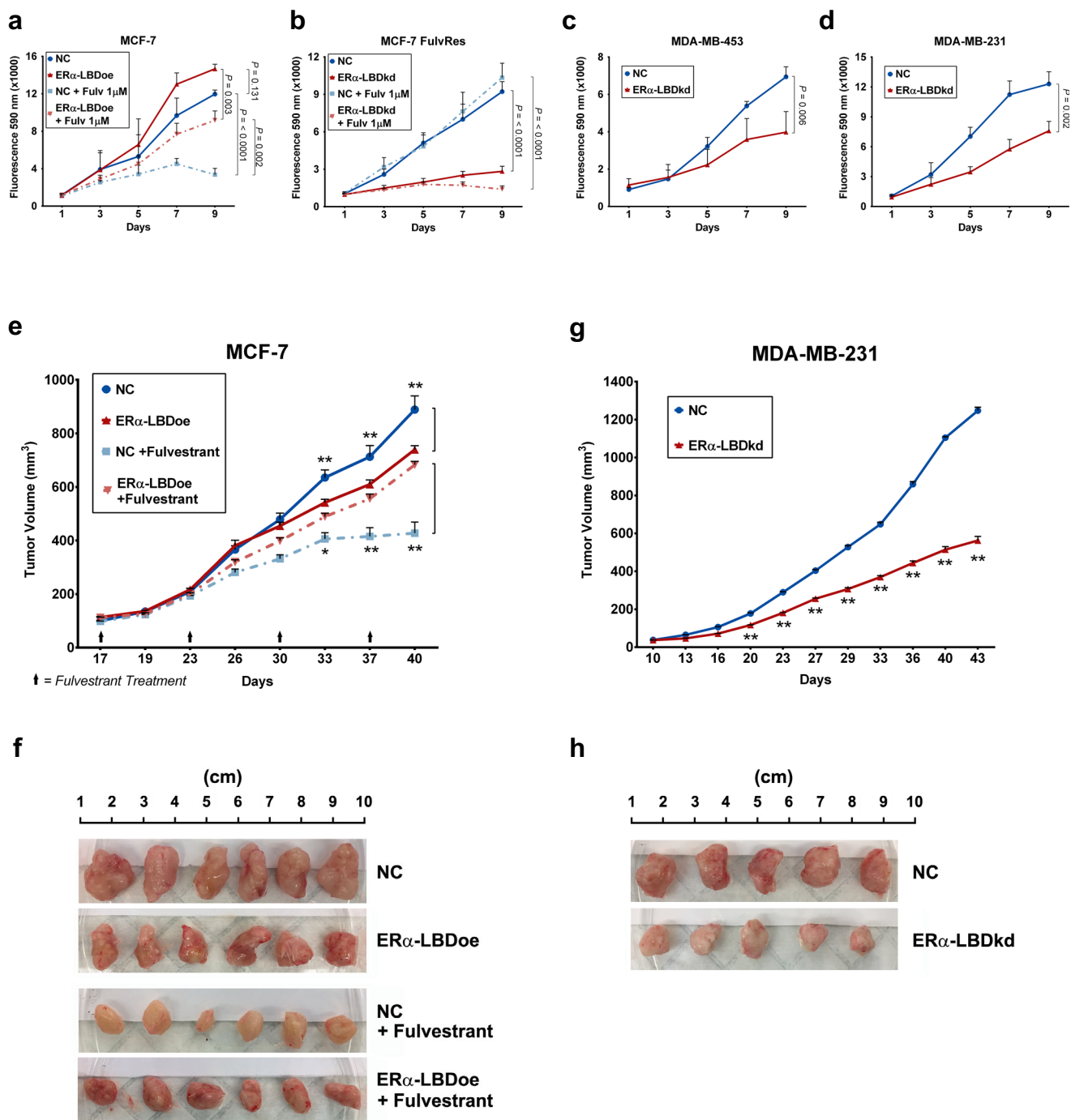


**Fig. 5 Effect of ER $\alpha$ -LBD ‘gain/loss-of-function’ on BC cell metabolism.** **a** Mitochondrial respiration and glycolysis levels in MCF-7 cell clones. Fulvestrant 1  $\mu$ M (24 h pre-treatment) or vehicle (DMSO) was added to cells. ER $\alpha$ -LBD overexpressing (ER $\alpha$ -LBDoe; in dark/light red) cells were compared to controls (NC; in dark/light blue). Analyses were carried out using XF Cell Mito Stress Kit (Agilent). Respiration was evaluated as oxygen consumption rate (OCR, pmol/min/cells), normalized on cell number. Glycolysis was evaluated as extracellular acidification rate (ECAR, mpH/min), normalized on cell number. The calculation of metabolic parameters was based on  $\Delta$ OCR and  $\Delta$ ECAR values, following manufacturer’s guidelines. **b, c** Mitochondrial respiration and glycolysis levels in MDA-MB-453 and MDA-MB-231 cells. Treatments as in (a). Cells with ER $\alpha$ -LBD knockdown (kd; in dark/light red) were compared to controls (NC; in dark/light blue). Analyses were carried out as described above. FCCP p-trifluoromethoxy-phenylhydrazone, Rot rotenone, AtmA antimycin A. Data in the figures are presented as mean  $\pm$  s.e.m. ( $n = 2$  independent experiments). \* $P < 0.05$ , \*\* $P < 0.01$ ; unpaired  $t$  test, one-sided. Source data are provided as a Source Data file.

untreated cells (Fig. 6b). Beyond the effects in ER $\alpha$ -FL(+) cells, we also assessed the effects of ER $\alpha$ -LBD on the growth of TNBC cells. In both MDA-MB-453 and MDA-MB-231 we observed ER $\alpha$ -LBD knockdown to reduce cell proliferation (~50%) (Fig. 6c, 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 $\alpha$ -LBDoe tumors continued to grow upon fulvestrant treatment while tumors from NC cells were

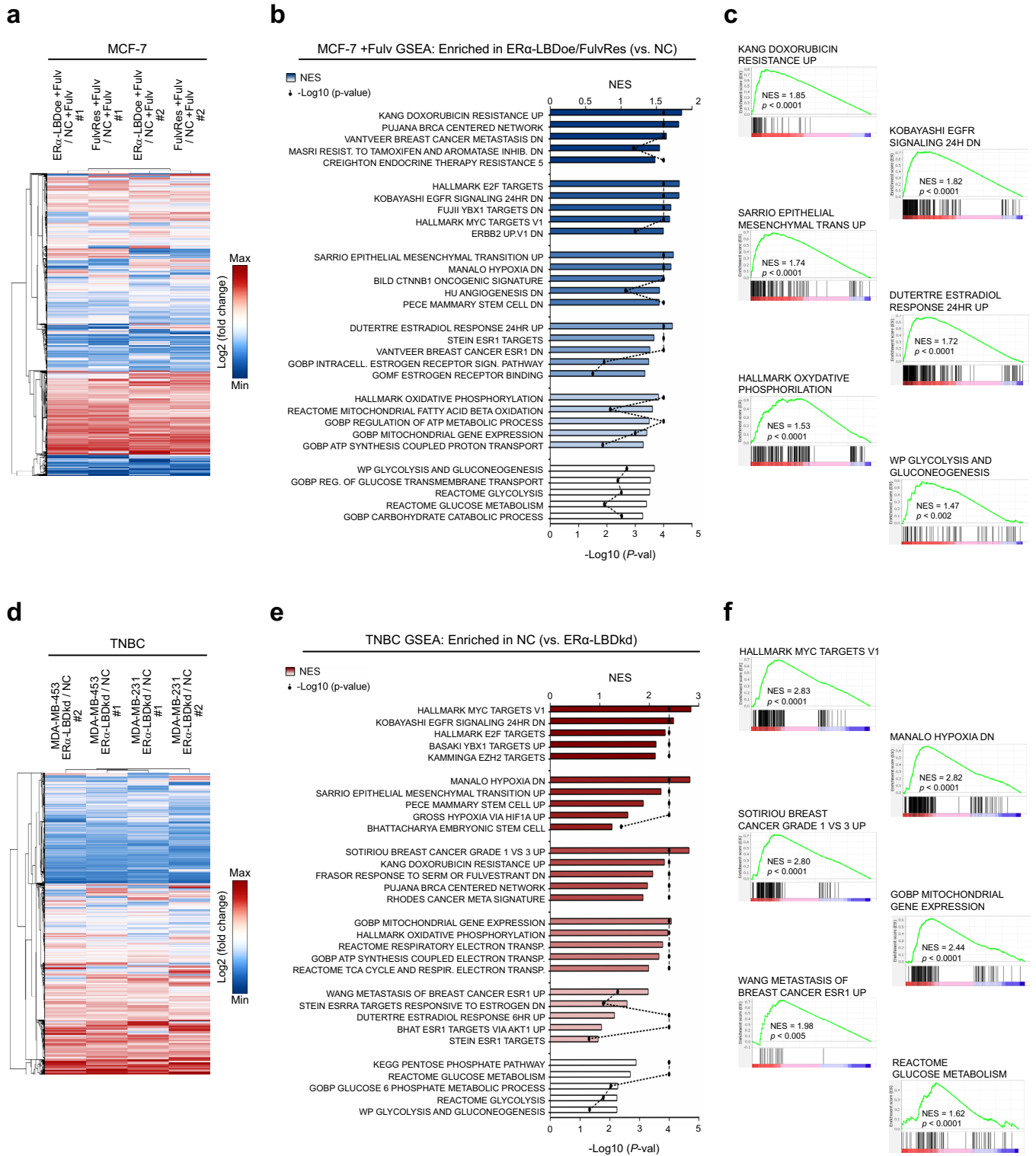




**Fig. 6** ER $\alpha$ -LBD promotes *in vitro* and *in vivo* growth and fulvestrant resistance. **a–d** *In vitro* proliferation of BC cell clones, using resazurin reagent and expressed as fluorescence intensity (absorbance at 590 nm). ER $\alpha$ -LBD overexpression (**a**) or knockdown (**b–d**) was compared to controls (NC). In **a** and **b**, cell proliferation was also tested in the presence or absence of fulvestrant 1  $\mu$ M treatment ( $n = 3$  independent experiments). All  $P$  values describing statistical differences between samples are shown. **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 (mm<sup>3</sup>) as a function of time (days). MCF-7 model: when tumors reached 100 mm<sup>3</sup>, 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 ER $\alpha$ -LBD overexpression in red. In **b–d** and **g**, NC is in blue and ER $\alpha$ -LBD knockdown in red. All data in the figure are presented as mean  $\pm$  s.e.m., ns not significant, \* $P < 0.05$ , \*\* $P < 0.01$ , two-way ANOVA. Source data are provided as a Source Data file.

expectedly growth inhibited with fulvestrant (Fig. 6e, f). Moreover, *in vivo* tumor growth for MDA-MB-231 ER $\alpha$ -LBDkd mice group was markedly reduced, when compared to NC group (Fig. 6g, h). MCF-7 ER $\alpha$ -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. 6b), whereas ER $\alpha$ -LBD knockdown led to reduced 3-D growth in both MDA-MB-453 and –231 cells (Supplementary Fig. 6c). Similarly, the migratory capacity



of BC models was enhanced in ER LBD overexpression models and impaired in knockdown models (Supplementary Fig. 6d, e).

### ER $\alpha$ -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 $\alpha$ -LBD, we further evaluated its impact

by gene expression analysis. We considered fulvestrant treated MCF-7 ER $\alpha$ -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 ~17 K differentially regulated genes (Fig. 7a and Supplementary Data 5). 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 Data 6). Interestingly, gene sets

**Fig. 7 ER $\alpha$ -LBD expression is associated with proliferation, endocrine resistance, stemness, and metabolism of breast cancer cells.** Two different BC cell models were used to investigate cell phenotypes associated with ER $\alpha$ -LBD 'gain/loss-of-function'. **a–c** MCF-7 model, with 24 h fulvestrant 1  $\mu$ M treatment (+Fulv): MCF-7 ER $\alpha$ -LBDoe, MCF-7 FulvRes and MCF-7 NC (control). **d–f** TNBC model: MDA-MB-453/-231 ER $\alpha$ -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 $\alpha$ -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 ( $P < 0.05$ , FDR test) enriched in MCF-7 ER $\alpha$ -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 the "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 $\alpha$ -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 ( $P < 0.05$ , FDR test) enriched in TNBC NC samples compared to TNBC ER $\alpha$ -LBDkd, grouped/color-coded as in **b**. **f** Enrichment plots of interest, derived from GSEA analysis described in **e**. Source data are provided as Supplementary Data 5 file.

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, c and Supplementary Data 5). Conversely, RNA-seq analysis of MDA-MB-453/-231 ER $\alpha$ -LBDkd "loss of function" models identified ~16K differentially regulated genes in TNBC ER $\alpha$ -LBDkd cells compared to NC cells (Fig. 7d and Supplementary Data 5). Similar 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 $\alpha$ -LBDkd (Fig. 7e, f and Supplementary Data 5).

We further explored the role of ER $\alpha$ -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<sup>36–39</sup>. MCF-7 FulvRes tumors were characterized by a higher percentage of CD44<sup>High</sup> cells which showed higher expression of ESR1-LBD, SOX2 and SOX9, suggesting a correlation between ER $\alpha$ -LBD expression and stemness (Supplementary Fig. 7a, b).

## DISCUSSION

Fulvestrant-resistant breast cancer is frequently observed in the metastatic setting however little data exists on the molecular events underlying resistance<sup>16,24,40–42</sup>. Our analysis of fulvestrant resistant BC models led to the identification of an ER $\alpha$  protein isoform (ER $\alpha$ -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 $\alpha$ -LBD identified in fulvestrant resistant cell lines but also in TNBC models. Unlike full-length ER $\alpha$  which is degraded by fulvestrant, ER $\alpha$ -LBD is stabilized by this agent along with other ER $\alpha$  ligands such as estradiol and tamoxifen, leading to increased ER $\alpha$ -LBD expression (Fig. 1). We hypothesize that the lack of fulvestrant-mediated degradation of ER $\alpha$ -LBD is likely dependent on its different protein sequence/structure, compared to ER $\alpha$ -FL. In fact, many protein elements involved in the regulation of the ubiquitin/proteasome system are missing in ER $\alpha$ -LBD or might be differently exposed/accessible<sup>43</sup>. We also hypothesize that low expression of ER $\alpha$ -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 $\alpha$  protein stability and discover novel pharmacologic inhibitors of ER $\alpha$  and its relevant isoforms. Interestingly, a 37 kDa ER species expressed at low levels was previously isolated from the mouse uterus. In accordance to our finding, this ER isoform appeared to have the ligand-binding region and a portion of the hinge domain, but lacked a DNA-binding region<sup>44</sup>.

The identity of ESR1-LBD transcript was determined by analyzing available sequence databases (specifically, FANTOM CAT) and by RNA-seq experiments (Fig. 2). We suggest that the ESR1-LBD transcript ranges from *ESR1* exons E4 to E8 and might utilize an alternative 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, describing potential transcriptional activity in the upstream genomic region between E3 and E4 (exon E3a). In addition, we noted reduced expression of the 3'UTR sequence of ESR1-LBD in the ER $\alpha$ (-)/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.

We demonstrated the presence of the ESR1-LBD transcript in primary BC specimens, examining publicly available RNA-Seq data<sup>32,33</sup>. ESR1-LBD was absent or poorly expressed in non-malignant breast tissue but found to variable levels in ER $\alpha$ (+) 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 (Fig. 2). A limitation of this study is the lack of ER $\alpha$ -LBD protein data and thus we cannot ascertain its expression and prevalence in these patient samples. Moreover, data from BC cases with acquired fulvestrant resistance were not available but we hypothesize that expression of ESR1/ER $\alpha$ -LBD would be found in a subset of these cancers. Nevertheless, IHC analyses on a breast cancer tissue microarray allowed us to detect cytoplasmic ER $\alpha$  protein primarily in those specimens lacking nuclear ER $\alpha$  expression, by using an ER $\alpha$  C-terminal antibody (36.8% of samples). Conversely, using an N-terminal antibody (that is typically used to assess the expression of ER $\alpha$  in the clinical setting and which cannot detect ER $\alpha$ -LBD), we identified cytoplasmic staining in only 2.1% of samples. These data support our findings on a truncated ER $\alpha$  isoform expression in BC cells and may account for its lack of detection until now.

The putative function of ER $\alpha$ -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 $\alpha$ -LBD protein–protein interactions by co-IP/MS using overexpression and knockdown models and determined an involvement with 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 $\alpha$ -LBD (Figs. 5 and 6). Moreover, RNA-seq analyses confirmed the association of ER $\alpha$ -LBD

expression with gene signatures related to breast cancer and signaling, drug resistance and cell metabolism, stemness, and adaptation to hypoxia (Fig. 7).

An extra-nuclear role for ER $\alpha$  isoforms has been well established most notably by their presence in the plasma membrane and their association with receptor tyrosine kinases and MAPK<sup>45</sup>. However, our data would suggest a distinct role for ER $\alpha$ -LBD as no apparent association with RTKs nor MAPK/ERK was identified. The role of mitochondrial ER $\alpha$  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<sup>4,46</sup>. Although ER $\alpha$ -LBD has no DNA binding capacity, we also found HADHB among its protein partners, suggesting both unique and overlapping functions for ER $\alpha$  and ER $\alpha$ -LBD in the mitochondria (Fig. 4).

In addition, a role of ER $\alpha$ -LBD in breast cancer stemness and malignancy was suggested by RNA-seq analysis and qPCR demonstrating an enrichment of the ESR1-LBD transcript in fulvestrant resistant CD44<sup>High</sup> tumor cells (Supplementary Fig. 7). Importantly, ER $\alpha$ -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 in endocrine resistance, we suggest that ER $\alpha$ -LBD may provide a link between these phenomena<sup>15–17,47–49</sup>. 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 $\alpha$ -LBD features are shared with ER $\alpha$  isoform, ER $\alpha$ -36<sup>30,31,50,51</sup>. No cross-detection between ER $\alpha$ -LBD and ER $\alpha$ -36 was found in our models. Specifically, the C-terminal antibody we used for IF, IHC, and co-IP studies does not recognize ER $\alpha$ -36. In addition, the use of primers and an antibody specific for ER $\alpha$  –36 did not lead to its detection in our hands (Supplementary Fig. 2). We suggest that isoforms of ER $\alpha$  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.

In summary, we have identified an ER $\alpha$  isoform “ER $\alpha$ -LBD” in ER $\alpha$ (+) fulvestrant resistant cancer cells, TNBCs and in a subset of human BC specimens. The development of specific strategies for ER $\alpha$ -LBD detection (*e.g.* antibodies, probes) is essential. Further studies are ongoing, to unravel a broader presence and the unique role that ER $\alpha$ -LBD plays in tumorigenesis and cancer progression.

## 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% CO<sub>2</sub> 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 $\beta$ -estradiol (E2, Sigma-Aldrich, USA), 4-Hydroxytamoxifen (4-OHT, Sigma-Aldrich, USA).

### Cell proliferation and manipulation

Cells (1000/well) were seeded into 96-well plates and, accordingly to experimental design, vehicle (DMSO 0.01%), fulvestrant 1  $\mu$ M or tamoxifen 1  $\mu$ M was added to the medium. Media and drug were replaced every 3 days. At different time points (day 1, 3, 5, 7, and 9), 20  $\mu$ l of Resazurin (R&D Systems, USA) was added to 180  $\mu$ l of medium in each well and incubated for 4 h. Fluorescence (absorbance 590 nm) was measured for each well using a SpectraMax M5 microplate reader (Molecular Devices, USA) and correlated to cell growth. To generate fulvestrant-resistant MCF-7 (MCF-7 Fulv/Res) and T-47D (T-47D Fulv/Res) cells, cells were cultured in the presence of increasing concentrations of fulvestrant (0.1–1  $\mu$ M). Cells were

deemed resistant when able to actively proliferate in 1  $\mu$ M fulvestrant. MCF-7 Y537S CRISPR knock-in cells were generated as described elsewhere<sup>24</sup>. PDX models of ER $\alpha$ (+)/(–) metastatic BC were generated and maintained as described elsewhere<sup>17</sup>. BC cell clones overexpressing ER $\alpha$ -LBD (ER $\alpha$ -LBDoe) were generated using stable retroviral transduction. Briefly, ER $\alpha$ -LBD coding sequence was cloned into pBABE-Puro vector (Addgene, USA; Supplementary Data 7). To generate retroviral particles, three plasmids were co-transfected into 293T packaging cells using Lipofectamine 2000 (Invitrogen, USA); pBABE-Puro-ER $\alpha$ LBDoe, pCMV-VSV-G and pUMVC (Addgene, USA). Viral supernatants were collected 48 h and 72 h later, centrifuged to remove cell debris, filtered through 0.45- $\mu$ m filters (Millipore, USA) then used with polybrene 8  $\mu$ g/ml polybrene (Santa Cruz Biotechnology, USA) to transduce BC cell lines. After 2 days, stable clones were selected with puromycin 2  $\mu$ g/ml. BC cell clones with ER $\alpha$ -LBD knockdown (ER $\alpha$ -LBDkd) were generated using stable lentiviral transduction and CRISPR/CAS9 technology for genome editing. Briefly, three different guide RNA sequences targeting ESR1 gene (exon #6) were cloned in the BsmBI site of lentiCRISPRv2 vector (Addgene, USA; Supplementary Data 7) to create a pool of pCR-ER $\alpha$ LBDkd lentivectors. To generate lentiviral particles, the following plasmids were co-transfected into 293T packaging cells: pCR-ER $\alpha$ LBDkd (pool), pCMV-VSV-G and psPAX2 (Addgene, USA). Infection and selection procedures: same as retroviral approach (see above).

### 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 $\alpha$  antibody D8H8, 1:2000 (Cell Signaling Technology, USA, #8644); anti-ER $\alpha$  F-10, 1:1000 (Santa Cruz Biotechnology, USA, #sc-8002), anti-ER $\alpha$  1D5, 1:1000 (Invitrogen, USA, #MA5-13191), anti-ER $\alpha$ -36, 1:200 (Alpha Diagnostic International, USA, #ERA361-A); anti- $\beta$ -actin 13E5, 1:2000 (Cell Signaling Technology, USA, #8457); anti-HDAC2, 1:1000 (Cell Signaling Technology, USA, #2540); anti-VDAC1 N-18 (Santa Cruz Biotechnology, USA, #sc-8828). Secondary antibodies (HRP conjugated): anti-rabbit, 1:5000 (Cell Signaling Technology, USA, #7074); anti-mouse, 1:2000 (Cytiva, USA, #RPN4201). Immunolabelling was visualized using ECL procedure (PerkinElmer, USA). Uncropped and unprocessed scans of the most important blots are supplied in the Source Data file. All blots derive from the same experiment and they were processed in parallel.

### Gel extraction, mass spectrometry, and protein structure analysis

Protein lysates from MCF-7 FR and MDA-MB-231 (both treated with fulvestrant 1  $\mu$ M for 24 h) were obtained using 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) and immunoprecipitated using anti-ER $\alpha$  antibody (1:100, Cell Signaling Technology, USA, #8644). IP samples were run on two electrophoresis gels, one used for staining with SimplyBlue Safe Stain (Invitrogen, USA) and band extraction, the other for WB sample check. In-gel digestion was performed using the method by Shevchenko et al.<sup>52</sup>. Briefly, gel bands were excised, washed with acetonitrile and 100 mM ammonium bicarbonate solution (1:1) for 30 min, dehydrated with 100% acetonitrile for 10 min and dried in a speed-vac for 10 min without heat. Gel slices were reduced with 5 mM DTT for 30 min at 56 °C in a thermo-mixer with gentle mixing, removed, allowed to cool to room temperature then alkylated with 11 mM IAA for 30 min in the dark. Gel slices were washed with 100 mM ammonium bicarbonate and 100% acetonitrile for 10 min each. Excess acetonitrile was removed and the slices dried in a speed-vac for 10 min without heating. Gel slices were then rehydrated in a solution of 25 ng/ml trypsin in 50 mM ammonium bicarbonate on ice for 30 min. Digestions were performed overnight at 37 °C in a thermo-mixer with gentle mixing. Digested peptides were collected and further extracted from gel slices in extraction buffer (5% formic acid and 50% acetonitrile, 1:2 vol/vol) at high-speed mixing. Extractions were combined and dried down in a vacuum centrifuge. Peptides were desalted with C18 resin-packed stage-tips, lyophilized to dryness, then re-constituted in 3% acetonitrile/0.1% formic acid for LC-MS/MS analysis. LC-MS/MS Analysis LC-MS/MS was performed using a Waters NanoAcquity LC system (with a 100 mm inner diameter  $\times$  10 cm length C18 column (1.7 mm BEH130; Waters, USA) configured with a 180 mm  $\times$  2 cm trap column coupled to a Thermo



Q-Exactive Plus orbitrap mass spectrometer (Thermo Scientific, USA). Trapping was performed at 15 ml/min 0.1% formic acid (Buffer A) for 1 min. The LC gradient was 0.5 to 50% B (100% acetonitrile; 0.1% formic acid) over 90 min at 300 nl/min. MS data were collected in data-dependent acquisition (DDA) mode utilizing a top ten precursor ion selection for HCD fragmentation. Full MS scans were performed with the following parameters: Resolution: 70,000; AGC target: 1e6; Maximum IT: 50 ms; Scan Range: 400 to 1600 m/z. DDA parameters were as follows: Resolution: 17,500; AGC target 5e4; Maximum IT: 50 ms; Isolation window: 1.5 m/z; NCE: 27; Minimum AGC target: 2e3; Intensity Threshold: 4e4; Dynamic Exclusion: 15 s; Charge exclusion: unassigned, 1, 6–8, >8. All MS/MS samples were analyzed using Mascot (Matrix Science, UK). Mascot was set up to search the SwissProt\_sprot\_20170705\_20180523 database (selected for Homo sapiens, unknown version, 20215 entries) assuming the digestion enzyme trypsin. Mascot was searched with a fragment ion mass tolerance of 0.080 Da and a parent ion tolerance of 10.0 PPM. Carbamidomethyl of cysteine was specified in Mascot as a fixed modification. Deamidated of asparagine and glutamine, oxidation of methionine, acetyl of the N-terminus and phosphorylation of serine, threonine and tyrosine were specified in Mascot as variable modifications. Mass spectrometry data were further processed using Scaffold software (Proteome Software Inc., USA). Peptide identifications were accepted if they could be established at >90.0% probability to achieve an FDR <1.0% by the Peptide Prophet algorithm with Scaffold delta-mass correction. Protein identifications were accepted if they could be established at >89.0% probability to achieve an FDR <1.0% and contained at least 1 identified peptide. Protein probabilities were assigned by the Protein Prophet algorithm. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Proteins sharing significant peptide evidence were grouped into clusters. Protein sequence and features were analyzed using Sequence Analysis software (Informagen, USA). Three-dimensional protein models were generated by Phyre<sup>2</sup> web portal ([www.sbg.bio.ic.ac.uk](http://www.sbg.bio.ic.ac.uk))<sup>53</sup>.

### ESR1-LBD transcript variant prediction

Comprehensive ESR1 transcript variants annotation was obtained from Ensembl Genome Browser ([www.ensembl.org](http://www.ensembl.org)). We used ZEMBU Genome Browser (<https://fantom.gsc.riken.jp/zenbu/>) for variant prediction study, showing data collected on ESR1 gene, based on FANTOM5 and FANTOM CAT analyses<sup>27,28</sup>. For ESR1 gene and transcripts details please refer to this link: <https://fantom.gsc.riken.jp/zenbu/glyphs/#config=eljPd00TPhloifL81WStGC;loc=hg19:chr6:152118533..152432431+>.

### RNA capture-sequencing and ESR1 transcripts analysis

Total RNA was extracted as described elsewhere<sup>54</sup> 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 (×150 cycles, 1.4 millions reads/sample). Raw sequencing data output was processed for expression analysis using STAR Aligner<sup>55</sup>, DEXSeq-count<sup>56</sup>, Cluster 3.0 and Java TreeView<sup>57,58</sup>, DEXSeq and plotDEXSeq<sup>56</sup>, Kallisto<sup>59</sup> (Supplementary Fig. 8a).

### RNA extraction and real-time qPCR

Total RNA was extracted from cells using Trizol reagent (Invitrogen, USA), according to the manufacturer's instructions. Extracted RNA samples were quantified and then treated with DNase I to remove any genomic DNA contamination, using Ambion DNase I kit (Invitrogen, USA). Reverse transcription was carried out using iScript™ Select cDNA Synthesis Kit (Bio-Rad, USA). cDNA levels were analyzed by real-time PCR using TaqMan Universal PCR Master Mix or SYBR Select Master Mix reagents and ViiA 7 Real-Time PCR system, according to the manufacturer's instructions (Applied Biosystems, USA). Melting curve data were collected to check PCR specificity. Samples were run in triplicate and mRNA levels were normalized against those of RPLP0 or  $\beta$ -actin. Relative expressions were calculated using the formula  $2^{-\Delta\Delta Ct}$  values. Expression data were obtained by using QuantStudio™ Real-Time PCR Software (Applied Biosystem, USA).

RT-PCR was carried out by using OneTaq Hot Start 2X Master Mix (New England Biolabs, USA). PCR samples were separated on a 1.5% agarose gel and results visualized using Gel Doc XR + Imaging System (Bio-Rad, USA). For RNA Decay Assay, BC cells were seeded in 6-well plates and treated with actinomycin-D 5  $\mu$ g/ml (Sigma-Aldrich, USA). RNA samples were collected at different time points (0, 4, 8, 24, 48 h). All PCR primers are summarized in Supplementary Data 7.

### Transfection and luciferase assay

Cells were seeded in 24-well plate ( $2 \times 10^5$  cells/well) 1 day before transfection and treatments were added accordingly to the experimental design (not treated, vehicle/DMSO 0.01% or fulvestrant 1  $\mu$ M). For the putative promoter experiment, three different human genomic DNA regions representing putative ESR1-LBD promoter (pESR1-LBD-1/2/3) were cloned into pGL3.basic plasmid (Promega, USA; Supplementary Data 7). BC cells were co-transfected with pGL3.basic (NC) or p-ESR1-LBD reporter plasmids (750 ng) and pRL-TK Renilla Luciferase plasmid (75 ng) (Addgene, USA) by using Lipofectamine 2000 (Invitrogen, USA) and following reagent protocol. For ER $\alpha$ -driven transcriptional activity experiment, BC cells were co-transfected with 3x-ERE-TATA-Luciferase reporter plasmid (750 ng) and pRL-TK Renilla Luciferase plasmid (75 ng) (Addgene, USA) by using the same procedure described above. Cells were harvested 24 h after transfection and cell lysates were used for Dual-Luciferase® Reporter Assay System analysis, according to the manufacturer's instructions (Promega, USA). Luciferase bioluminescence measurements were performed with the Veritas™ Microplate Luminometer (Promega, USA). For each sample, Firefly luciferase activity was normalized against Renilla luciferase activity.

### Breast cancer patients data

RNA-seq FASTQ files of 135 breast cancer were obtained from the European Nucleotide Archive, ENA (Study Accession: PRJNA251383; Supplementary Data 2)<sup>32</sup>. RNA-seq FASTQ files of 17 metaplastic breast cancer were kindly provided by Dr. Jorge S Reis-Filho (Supplementary Data 2)<sup>33</sup>. Raw data were processed by using DEXSeq-count<sup>53</sup>, Kallisto<sup>54</sup> and Morpheus (<https://software.broadinstitute.org/morpheus>).

### Immunohistochemistry and confocal microscopy

ER $\alpha$  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 $\alpha$  antibody D8H8, Cell Signaling Technology, USA, #8644) 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  $\mu$ g/ml, Invitrogen, USA, #A-21347); ER $\alpha$  D8H8 (1  $\mu$ g/ml, Cell Signaling Technology, USA, #8644); 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).

### Cell fractionation for western blot

Whole-cell lysate. Cells were pelleted and homogenized in lysis buffer (50 mM Tris-HCl, pH 7.5, 2 mM EDTA, 100 mM NaCl, 1% Triton X-100, protease inhibitors). Cell lysates were incubated 1 h on ice and centrifuged at 12,000 g at 4 °C for 20 min to collect supernatants (sample W). Nuclear and cytosolic fractions. Pelleted cells were resuspended in nuclear lysis buffer (10 mM HEPES; pH 7.5, 10 mM KCl, 0.1 mM EDTA, 0.5% Nonidet-40, 0.5 mM PMSF, protease inhibitors). Cells were allowed to swell on ice for 15–20 min with intermittent mixing to disrupt cell membranes and then centrifuged at 12,000  $\times$  g at 4 °C for 10 min. Supernatant were transferred into a new tube and used as cytosolic fraction (sample C). The pelleted nuclei were washed with nuclear lysis buffer and resuspended in nuclear extraction buffer (20 mM HEPES pH 7.5, 400 mM NaCl, 1 mM EDTA, 1 mM DTT, 1 mM PMSF, protease inhibitors) and incubated in ice for 30–60 min. Nuclear extract (supernatant) was collected by centrifugation at 16,000  $\times$  g for 15 min at 4 °C and transferred into a new tube (sample N). Mitochondrial fraction. Cells were collected by centrifugation at  $\sim$ 500  $\times$  g

for 10 min then resuspended in 10 packed cell volumes of NKM buffer, x2 (Tris-HCl 1 M pH 7.4, NaCl 2 M, KCl 1 M, MgCl<sub>2</sub> 0.5 M). Cells were pelleted and resuspended in 6 packed cell volumes of homogenization buffer (Tris-HCl 1 M pH 6.7, KCl 1 M, MgCl<sub>2</sub> 0.5 M PMSF 200 mM). Cell suspension was transferred to a glass homogenizer, incubate for 10 min on ice, then homogenized using a tight pestle. The level of cell breakage was checked under the microscope. Homogenate was poured into a conical centrifuge tube containing 2 packed cell volume of 2 M sucrose solution and mixed gently. Unbroken cells, nuclei and large debris were pelleted at 1200 g for 5 min. The supernatant was transferred to another tube containing 1.5 packed cell volume of 2 M sucrose and mixed gently. Mitochondrial fraction was collected by centrifuging at 12,000 × g for 15 min and resuspended in isotonic mitochondrial buffer (10 mM Hepes buffer pH 8.0, 250 mM sucrose, 0.5 mM EGTA). The suspension was split 1:2. One half, centrifuged at 12,000 × g for 15 min and resuspended in lysis buffer. After incubation of 30–60 min on ice, mitochondrial fraction (supernatant) was collected by centrifugation at 12,000 × g for 15 min (sample M). The other half was incubated with proteinase K (50 µg/ml, Thermo Fisher) for 30 min at 25 °C. PMSF was added to a final concentration of 2 mM to terminate the proteinase K activity, followed by 10 min of incubation. PK-treated mitochondria were collected by centrifugation at 12,000 × g for 10 min (sample M\*).

### TMA immunohistochemistry

ERα 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. The primary anti-ERα antibody (D8H8, Cell Signaling Technology, USA, #8644) was incubated 1:100 for 28 min at 37 °C. Secondary antibody 1:100 for 1 h. Hematoxylin/eosin counterstain was also carried out. Slides were scanned by Leica Aperio AT2 (Leica Biosystems, USA) whole slide scanners. Images from ERα staining by using 6F11 antibody were kindly obtained by US BioMax Inc., USA. For each sample/core, IHC grading was assessed considering the intensity of the staining and the number of stained cells.

### 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, #8644) and Dynabeads™ Protein A (Invitrogen, USA). IP samples were washed four 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. Pathway enrichment analysis was performed by using Enrichr platform (<https://maayanlab.cloud/Enrichr/>)<sup>60</sup>. Enriched pathways network, clustering, and protein–protein interaction (PPI) network analyses were carried out using Cytoscape, EnrichmentMap, AutoAnnotate and stringApp softwares<sup>51</sup>.

### 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 × 10<sup>4</sup> cells/well), with or without Fulvestrant 1 µM 24 h pre-treatment. Raw data output was collected and analyzed using Wave Software (Agilent Technologies). Procedures and OCR/ECAR data interpretation were carried out accordingly to the 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 (200 mg/kg 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 an 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).

### Microscopy, 3-D growth, and wound healing assay

MCF-7 cells (WT and FulvRes) were seeded into six-well plate (1 × 10<sup>5</sup> cells/well) and treated accordingly to the experimental design. After controls reached 100% confluence, images of cells were captured with Zeiss Axiovert Microscope and processed with Zeiss ZEN software (Carl Zeiss, Germany). For 3-D growth, BC cell clones were suspended sparsely and plated in 24-well ultra-low attachment plates (Corning, USA), 3 × 10<sup>3</sup> cells/well. Treatments (DMSO 0.01% or fulvestrant 1 µM) were added to medium accordingly to experimental design. After 5 days, representative images of cells were acquired with a Zeiss Axiovert Microscope. Quantification of cell 3-D proliferation was based on the optical density (OD 600 nm) of cell suspensions and determined using a SpectraMax M5 microplate reader (Molecular Devices, USA). For the wound healing assay, about 1 × 10<sup>5</sup> cells were seeded into each well of 12-well plates and cultured until 100% confluence. The 'scratch' was created with a p200 pipet tip on the cell monolayer through the center of the well. The debris was removed by washing the well with 1 ml of culture medium and then 1 ml of medium (with 1% serum) was added into each well. Treatments (DMSO 0.01% or fulvestrant 1 µM) were also added accordingly to the experimental design. The plate was incubated at 37 °C, and images of the 'scratches' were captured at various time points with a Zeiss Axiovert Microscope. The width of scratch was measured with FIJI software.

### RNA sequencing

Total RNA was extracted from cell lines as described elsewhere<sup>54</sup> and sequenced on an Illumina HiSeq instrument (2 × 150 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)<sup>55</sup>, HTSeq (gene expression)<sup>62</sup>, Cluster 3.0 and Java TreeView (hierarchical clustering and heat map)<sup>57,58</sup>, GSEA (enrichment of gene sets)<sup>63</sup> (Supplementary Fig. 8b and Supplementary Data 6).

### Flow cytometry (FACS)

For FACS/flow analyses, tumors were digested in sterile Epicult media (StemCell Technology, Canada), minced with sterile razor blades and incubated for 3 h in the presence of collagenase/hyaluronidase (1,000 Units/sample). Cells were washed with sterile filtered PBS supplemented with 1% BSA (PBS-BSA 1%) and filtered through a 40 mm nylon mesh (BD Biosciences, USA). Cells were then stained in a volume of 100 µL (PBS-BSA 1%) with CD44-APC antibody 100 ng/10<sup>6</sup>–10<sup>8</sup> cells (IM7, eBiosciences, USA, # 17-0441-82) on ice for 30 min and analyzed by flow cytometry at the MSKCC's flow core with BD FACS Aria I instrument (BD Biosciences, USA). Samples were analyzed for cell population distribution and sorted for viability (DAPI<sup>inc9</sup>) and CD44 expression. For flow plot analyses, samples were run using FlowJo 7.5 software (Tree Star, USA).

### Statistics

Data were expressed as mean ± 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 (FDR test), please refer to official website ([www.gsea-msigdb.org](http://www.gsea-msigdb.org)). Differences were considered statistically significant at *P* < 0.01 and *P* < 0.05.

### DATA AVAILABILITY

All data supporting the findings of this study are available with the article (including Supplementary information) or from the corresponding author upon reasonable request. The source data underlying Figs. 1a, b, d, 5a–c, 6a–e, g and Supplementary

Figs. 1a–c, f–h, 2b, f, g, i, 3a, f–k, 5a–c, 6a–e and 7b are provided as a Source Data file. The source data underlying Figs. 2a–h, 4a–f and 7a–f are provided as Supplementary Data 2, 4, 5, respectively. RNA-seq data have been deposited to ArrayExpress with accession numbers E-MTAB-10733 and E-MTAB-10738. MpBC RNA-seq data have been previously deposited to the Sequence Read Archive (SRP070780)<sup>33</sup>. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifiers PXD027087 and PXD027088.

Received: 19 December 2021; Accepted: 26 July 2022;

Published online: 23 August 2022

## REFERENCES

- Bernstein, L. & Lacey, J. V. Jr. Receptors, associations, and risk factor differences by breast cancer subtypes: positive or negative? *J. Natl Cancer Inst.* **103**, 451–453 (2011).
- Hennighausen, L. & Robinson, G. W. Signaling pathways in mammary gland development. *Dev. Cell* **1**, 467–475 (2001).
- Siersbaek, R., Kumar, S. & Carroll, J. S. Signaling pathways and steroid receptors modulating estrogen receptor alpha function in breast cancer. *Genes Dev.* **32**, 1141–1154 (2018).
- Pedram, A., Razandi, M., Wallace, D. C. & Levin, E. R. Functional estrogen receptors in the mitochondria of breast cancer cells. *Mol. Biol. Cell* **17**, 2125–2137 (2006).
- Hamilton, D. J. et al. Estrogen receptor alpha activation enhances mitochondrial function and systemic metabolism in high-fat-fed ovariectomized mice. *Physiol. Rep.* **4**, e12913 (2016).
- Osborne, C. K. Aromatase inhibitors in relation to other forms of endocrine therapy for breast cancer. *Endocr. Relat. Cancer* **6**, 271–276 (1999).
- Cole, M. P., Jones, C. T. & Todd, I. D. A new anti-oestrogenic agent in late breast cancer. An early clinical appraisal of ICI46474. *Br. J. Cancer* **25**, 270–275 (1971).
- Johnston, S. J. & Cheung, K. L. Fulvestrant - a novel endocrine therapy for breast cancer. *Curr. Med. Chem.* **17**, 902–914 (2010).
- Toy, W. et al. ESR1 ligand-binding domain mutations in hormone-resistant breast cancer. *Nat. Genet.* **45**, 1439–1445 (2013).
- Jeselsohn, R. et al. Emergence of constitutively active estrogen receptor-alpha mutations in pretreated advanced estrogen receptor-positive breast cancer. *Clin. Cancer Res.* **20**, 1757–1767 (2014).
- Robinson, D. R. et al. Activating ESR1 mutations in hormone-resistant metastatic breast cancer. *Nat. Genet.* **45**, 1446–1451 (2013).
- Pejerrey, S. M. et al. The Impact of ESR1 Mutations on the Treatment of Metastatic Breast. *Cancer Horm. Cancer* **9**, 215–228 (2018).
- Razavi, P. et al. The genomic landscape of endocrine-resistant advanced breast cancers. *Cancer Cell* **34**, 427–438.e426 (2018).
- Finlay-Schultz, J. & Sartorius, C. A. Steroid hormones, steroid receptors, and breast cancer stem cells. *J. Mammary Gland Biol. Neoplasia* **20**, 39–50 (2015).
- Rodriguez, D. et al. The central contributions of breast cancer stem cells in developing resistance to endocrine therapy in estrogen receptor (ER)-positive breast cancer. *Cancers* **11**, 1028 (2019).
- Sansone, P. et al. Self-renewal of CD133(hi) cells by IL6/Notch3 signalling regulates endocrine resistance in metastatic breast cancer. *Nat. Commun.* **7**, 10442 (2016).
- Sansone, P. et al. Evolution of cancer stem-like cells in endocrine-resistant metastatic breast cancers is mediated by stromal microvesicles. *Cancer Res.* **77**, 1927–1941 (2017).
- García-Martínez, L., Zhang, Y., Nakata, Y., Chan, H. L. & Morey, L. Epigenetic mechanisms in breast cancer therapy and resistance. *Nat. Commun.* **12**, 1786 (2021).
- Foulkes, W. D., Smith, I. E. & Reis-Filho, J. S. Triple-negative breast cancer. *N. Engl. J. Med.* **363**, 1938–1948 (2010).
- Shen, M. et al. A review of current progress in triple-negative breast cancer therapy. *Open Med. (Wars.)* **15**, 1143–1149 (2020).
- Treeck, O., Schuler-Toprak, S. & Ortmann, O. Estrogen actions in triple-negative breast cancer. *Cells* **9**, 2358 (2020).
- Howell, A., Osborne, C. K., Morris, C. & Wakeling, A. E. ICI 182,780 (Faslodex): development of a novel, “pure” antiestrogen. *Cancer* **89**, 817–825 (2000).
- Robertson, J. F. ICI 182,780 (Fulvestrant)—the first oestrogen receptor down-regulator—current clinical data. *Br. J. Cancer* **85**(Suppl 2), 11–14 (2001).
- Toy, W. et al. Activating ESR1 mutations differentially affect the efficacy of ER antagonists. *Cancer Discov.* **7**, 277–287 (2017).
- Sabbah, M., Radanyi, C., Redeuilh, G. & Baulieu, E. E. The 90 kDa heat-shock protein (hsp90) modulates the binding of the oestrogen receptor to its cognate DNA. *Biochem. J.* **314**, 205–213 (1996).
- Chandarlapaty, S. et al. SNX2112, a synthetic heat shock protein 90 inhibitor, has potent antitumor activity against HER kinase-dependent cancers. *Clin. Cancer Res.* **14**, 240–248 (2008).
- Lizio, M. et al. Gateways to the FANTOM5 promoter level mammalian expression atlas. *Genome Biol.* **16**, 22 (2015).
- Hon, C. C. et al. An atlas of human long non-coding RNAs with accurate 5' ends. *Nature* **543**, 199–204 (2017).
- Mayr, C. & Bartel, D. P. Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* **138**, 673–684 (2009).
- Wang, Z. et al. Identification, cloning, and expression of human estrogen receptor-alpha36, a novel variant of human estrogen receptor-alpha66. *Biochem. Biophys. Res. Commun.* **336**, 1023–1027 (2005).
- Wang, Q. et al. Tamoxifen enhances stemness and promotes metastasis of ERalpha36(+) breast cancer by upregulating ALDH1A1 in cancer cells. *Cell Res.* **28**, 336–358 (2018).
- Varley, K. E. et al. Recurrent read-through fusion transcripts in breast cancer. *Breast Cancer Res. Treat.* **146**, 287–297 (2014).
- Piscuoglio, S. et al. Genomic and transcriptomic heterogeneity in metaplastic carcinomas of the breast. *NPJ Breast Cancer* **3**, 48 (2017).
- Khan, S. A., Rogers, M. A., Khurana, K. K., Meguid, M. M. & Numann, P. J. Estrogen receptor expression in benign breast epithelium and breast cancer risk. *J. Natl Cancer Inst.* **90**, 37–42 (1998).
- Lin, C. Y. et al. Discovery of estrogen receptor alpha target genes and response elements in breast tumor cells. *Genome Biol.* **5**, R66 (2004).
- Hebbard, L. et al. CD44 expression and regulation during mammary gland development and function. *J. Cell Sci.* **113**, 2619–2630 (2000).
- Bellerby, R. et al. Overexpression of specific CD44 isoforms is associated with aggressive cell features in acquired endocrine resistance. *Front. Oncol.* **6**, 145 (2016).
- Domenici, G. et al. A Sox2-Sox9 signalling axis maintains human breast luminal progenitor and breast cancer stem cells. *Oncogene* **38**, 3151–3169 (2019).
- Christin, J. R. et al. Stem cell determinant SOX9 promotes lineage plasticity and progression in basal-like breast cancer. *Cell Rep.* **31**, 107742 (2020).
- Kaminska, K. 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.* **23**, 26 (2021).
- Giessrigl, B. 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* **109**, 2751–2762 (2013).
- Kingston, B. et al. Genomic profile of advanced breast cancer in circulating tumour DNA. *Nat. Commun.* **12**, 2423 (2021).
- Tealco-Cruz, A. C. & Ramirez-Jarquín, J. O. Polyubiquitination inhibition of estrogen receptor alpha and its implications in breast cancer. *World J. Clin. Oncol.* **9**, 60–70 (2018).
- Horigome, T. et al. Purification and characterization of mouse uterine estrogen receptor under conditions of varying hormonal status. *Endocrinology* **121**, 2099–2111 (1987).
- Boonyaratankornkit, V. et al. Extranuclear signaling by sex steroid receptors and clinical implications in breast cancer. *Mol. Cell Endocrinol.* **466**, 51–72 (2018).
- Zhou, Z., Zhou, J. & Du, Y. Estrogen receptor alpha interacts with mitochondrial protein HADHB and affects beta-oxidation activity. *Mol. Cell Proteom.* **11**, M111011056 (2012).
- O'Brien, C. S., Farnie, G., Howell, S. J. & Clarke, R. B. Breast cancer stem cells and their role in resistance to endocrine therapy. *Horm. Cancer* **2**, 91–103 (2011).
- Leung, E. Y. et al. Endocrine therapy of estrogen receptor-positive breast cancer cells: early differential effects on stem cell markers. *Front. Oncol.* **7**, 184 (2017).
- Szostakowska, M., Trebinska-Stryjewska, A., Grzybowska, E. A. & Fabisiwicz, A. Resistance to endocrine therapy in breast cancer: molecular mechanisms and future goals. *Breast Cancer Res. Treat.* **173**, 489–497 (2019).
- Yan, Y., Yu, L., Castro, L. & Dixon, D. ERalpha36, a variant of estrogen receptor alpha, is predominantly localized in mitochondria of human uterine smooth muscle and leiomyoma cells. *PLoS ONE* **12**, e0186078 (2017).
- Thiebaut, C. et al. The role of ERalpha36 in development and tumor malignancy. *Int. J. Mol. Sci.* **21**, 4116 (2020).
- Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V. & Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **1**, 2856–2860 (2006).
- Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).
- Strillacci, A. et al. MiR-101 downregulation is involved in cyclooxygenase-2 overexpression in human colon cancer cells. *Exp. Cell Res.* **315**, 1439–1447 (2009).
- Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

56. Anders, S., Reyes, A. & Huber, W. Detecting differential usage of exons from RNA-seq data. *Genome Res.* **22**, 2008–2017 (2012).
57. de Hoon, M. J., Imoto, S., Nolan, J. & Miyano, S. Open source clustering software. *Bioinformatics* **20**, 1453–1454 (2004).
58. Saldanha, A. J. Java Treeview—extensible visualization of microarray data. *Bioinformatics* **20**, 3246–3248 (2004).
59. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).
60. Kuleshov, M. V. et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**, W90–W97 (2016).
61. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
62. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
63. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).

## ACKNOWLEDGEMENTS

This work was supported by grants from the National Institute of Health to MSKCC's Cancer Center (P30CA008748) and to S.C. (R01CA204999), and a grant from the 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.

## AUTHOR CONTRIBUTIONS

Study concepts: A.S., P.S., D.L., S.C., and J.B. Study design: A.S., S.C., and J.B. In vivo and ex vivo experiments: A.S., P.S., V.K.R., Q.C., E.d.S., and R.S.R. Microscopy experiments: A.S., M.T., V.B., and I.D.P. RNA capture-sequencing experiments: A.S., F.M., B.H.-L., D.B., and M.F.B. Proteomics experiments: A.S., R.C.H., and F.B. Flow cytometry experiments: A.S., V.K.R., and R.S.R. RNA-seq data processing: A.S., A.P., and B.L. MpBC study: F.P. and J.S.R.-F. Metabolic assays: A.S. and M.B. Cell culture: A.S., C.M. V.K.R., Y.C., and I.D.P. Data analysis/interpretation: A.S., S.C., and J.B. Statistical analysis: A.S. Manuscript preparation: A.S., S.C., and J.B. Financial Support: E.S., S.C., and J.B.

## COMPETING INTERESTS

The authors declare no non-financial competing interests but declare the following competing financial interests for S.C.: consulting honorarium from Novartis, Lilly, Inivata, Ultivue, AstraZeneca, and Sanofi; grant support from Paige.AI, Ambryx, AstraZeneca and Daiichi-Sankyo.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41523-022-00470-6>.

**Correspondence** and requests for materials should be addressed to Sarat Chandralapaty or Jacqueline Bromberg.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022