

SUPPLEMENTARY MATERIAL De Martino et al “MALAT1 is required for estradiol to act as a mitogen in hormone receptor-positive breast cancer”

Supplementary Methods

The functional form of the nonlinear growth model was selected according to the observed proliferation pattern of each cell line. An exponential growth function was used for the LNCaP and T47D cell lines, whose longitudinal proliferation trajectories were adequately described by an approximately exponential increase over the observation period. A Hill growth function was used for the MCF7 and 22RV1 cell lines, whose proliferation profiles showed a more pronounced saturating pattern.

For all cell lines, observed IncuCyte proliferation values were normalized relative to the value measured at time zero, before treatment. For cell lines analyzed using the exponential-growth model, the normalized response was expressed as a percentage of the baseline value, such that the baseline observation corresponded to 100%. Thus, values above 100% indicate an increase in proliferation relative to the pre-treatment baseline.

For cell lines analyzed using the Hill-growth model, the baseline-normalized response was subsequently rescaled by subtracting 100 percentage points. Consequently, the response at time zero was equal to 0, and subsequent values represented the percentage-point increase relative to baseline. This rescaling was adopted to ensure consistency with the Hill function, which is equal to zero at time zero and increases toward an asymptotic response.

For both model classes, experimental condition was defined by the combination of transfection/gapmer status and hormonal exposure. Condition effects were incorporated as fixed effects, with the relevant reference condition specified in the corresponding model description. Random effects were included at the experimental-unit level to account for between-experiment variability and, where specified, for condition-dependent deviations within the same experiment. The mathematical specifications of the exponential and Hill nonlinear mixed-effects models are described in the following sections.

Nonlinear mixed-effects Hill model

Longitudinal IncuCyte proliferation data from the immortalised MCF7 and 22RV1 cell lines were analysed using nonlinear mixed-effects Hill growth models. In both cell lines, the experimental

condition was defined by the combination of gapmer/transfection status and hormonal exposure, yielding six condition levels. The reference condition was LacZ gapmer Vehicle. Experimental units, indexed by e , corresponded to independent experiments performed on different days under the same experimental protocol. Conditions were indexed by $j = 1, \dots, 6$, and repeated measurements within each experiment-condition combination were indexed by q .

The outcome y_{eqj} was the proliferation response at time t_{eqj} , expressed as a percentage of the pre-treatment value and subsequently shifted so that the response started at zero. The response was modelled using the Hill function

$$y_{eqj} = \frac{K_{ej} t_{eqj}^h}{t_{50,ej}^h + t_{eqj}^h} + \varepsilon_{eqj},$$

where K_{ej} is the asymptotic response, $t_{50,ej}$ is the time required to reach one half of the asymptotic response, h is the Hill coefficient, and ε_{eqj} is the residual error. Residual errors were assumed to have mean zero and common variance,

$$\varepsilon_{eqj} \sim N(0, \sigma^2).$$

To ensure positivity, K , t_{50} , and h were estimated on the logarithmic scale. Specifically,

$$K_{ej} = \exp(\eta_{K,ej}), t_{50,ej} = \exp(\eta_{t50,ej}), h = \exp(\eta_h).$$

MCF7 model

For the MCF7 cell line, five independent experiments were available. The asymptotic response and half-response time were modelled using condition-specific fixed effects, whereas the Hill coefficient was assumed to be common across conditions and experiments.

The log-asymptote was modelled using condition-specific fixed effects, with LacZ gapmer Vehicle as the reference condition:

$$\eta_{K,e1} = \beta_{K0},$$

$$\eta_{K,ej} = \beta_{K0} + \beta_{Kj}, j = 2, \dots, 6.$$

Here, $\eta_{K,ej}$ denotes the log-asymptote for experimental unit e under condition j , β_{K0} is the population-level log-asymptote under the reference condition LacZ gapmer Vehicle, and β_{Kj} is the fixed difference in log-asymptote between condition j and LacZ gapmer Vehicle.

Because no random effect was included for this parameter, $\eta_{K,ej}$ does not actually vary across

experimental units after conditioning on j . More precisely, the notation can therefore be simplified to:

$$\eta_{K,1} = \beta_{K0},$$

$$\eta_{K,j} = \beta_{K0} + \beta_{Kj}, \quad j = 2, \dots, 6.$$

Thus, on the original scale,

$$K_1 = \exp(\beta_{K0}),$$

and, for $j = 2, \dots, 6$,

$$K_j = \exp(\beta_{K0} + \beta_{Kj}).$$

where β_{K0} is the population-level log-asymptote for the reference condition LacZ gapmer Vehicle, β_{Kj} is the fixed difference in log-asymptote between condition j and LacZ gapmer Vehicle.

The log half-response time was modelled using condition-specific fixed effects and experiment-level random effects. For the reference condition,

$$\eta_{t50,e1} = \beta_{t50,0} + b_{0e},$$

whereas, for $j = 2, \dots, 6$,

$$\eta_{t50,ej} = \beta_{t50,0} + \beta_{t50,j} + b_{0e} + b_{t50,ej}.$$

Here, $\eta_{t50,ej} = \log(t_{50,ej})$ is the log half-response time for experiment e under condition j , $\beta_{t50,0}$ is the population-level log half-response time under LacZ gapmer Vehicle, and $\beta_{t50,j}$ is the fixed difference in log half-response time between condition j and LacZ gapmer Vehicle. The random intercept b_{0e} represents the experiment-specific departure from the average log half-response time in the reference condition. For each non-reference condition, $b_{t50,ej}$ represents the additional experiment-specific deviation associated with the contrast between condition j and LacZ gapmer Vehicle.

The vector of experiment-level random effects was defined as

$$b_e = (b_{0e}, b_{t50,e2}, \dots, b_{t50,e6})^T,$$

and was assumed to follow

$$b_e \sim N_6(0, D_{t50}),$$

where D_{t50} is an unstructured 6×6 positive-definite covariance matrix. This covariance structure allows the random deviation in the reference condition and the five condition-specific random

contrast effects to be correlated within the same experiment. Random-effects vectors were assumed independent across experiments.

On the original scale, the half-response time was therefore:

$$t_{50,e1} = \exp(\beta_{t50,0} + b_{0e}),$$

for the reference condition, and:

$$t_{50,ej} = \exp(\beta_{t50,0} + \beta_{t50,j} + b_{0e} + b_{t50,ej}), j = 2, \dots, 6.$$

Finally, the Hill coefficient was modelled as a common fixed parameter:

$$\eta_h = \beta_h,$$

such that:

$$h = \exp(\beta_h).$$

Thus, the MCF7 model included condition-specific fixed effects for $\log(K)$ and $\log(t_{50})$, experiment-level random effects for $\log(t_{50})$, and a common fixed Hill coefficient. The six condition-specific curves were therefore allowed to differ in both their asymptotic response and their population-average half-response time, whereas between-experiment heterogeneity was modelled only for the half-response time.

22RV1 model

For the 22RV1 cell line, only two independent experiments were available. Therefore, a reduced version of the MCF7 Hill model was fitted to limit the number of population-level fixed-effect parameters to be estimated.

As in the MCF7 model, the log-asymptote was modelled using condition-specific fixed effects, with LacZ gapmer Vehicle as the reference condition:

$$\eta_{K,1} = \beta_{K0},$$

$$\eta_{K,j} = \beta_{K0} + \beta_{Kj}, j = 2, \dots, 6.$$

No random effect was included for the log-asymptote. The Hill coefficient was again assumed to be common across all conditions and experiments:

$$\eta_h = \beta_h, h = \exp(\beta_h).$$

Unlike the MCF7 model, the log half-response time did not include condition-specific fixed

effects. Instead, a common population-level fixed effect was assumed:

$$\eta_{t50,e1} = \beta_{t50,0} + b_{0e},$$
$$\eta_{t50,ej} = \beta_{t50,0} + b_{0e} + b_{t50,ej}, j = 2, \dots, 6.$$

Thus, $\beta_{t50,0}$ represents the common population-average log half-response time across conditions.

The random-effects vector and its distribution were specified as in the MCF7 model:

$$b_e = (b_{0e}, b_{t50,e2}, \dots, b_{t50,e6})^T,$$
$$b_e \sim N_6(0, D_{t50}).$$

Therefore, compared with the MCF7 model, the 22RV1 model retained condition-specific fixed effects for the log-asymptote and the same condition-dependent random-effects structure for the log half-response time, but omitted the five condition-specific fixed effects for the log half-response time.

Nonlinear mixed-effects exponential growth model

For cell lines LNCaP and T47D (three and five independent experiments, respectively) longitudinal IncuCyte proliferation was modelled using a nonlinear mixed-effects exponential growth model. For experimental unit e , condition j , and time point q , the response was modelled using the following parameterization:

$$y_{eqj} = \exp\{\eta_A + \lambda_{ej}t_{eqj}\} + \varepsilon_{eqj},$$

or, equivalently,

$$y_{eqj} = A \exp(\lambda_{ej}t_{eqj}) + \varepsilon_{eqj}, A = \exp(\eta_A).$$

Here, A is the common population-level initial response parameter, $\eta_A = \log(A)$ is its logarithmic-scale representation, t_{eqj} is the time post treatment, λ_{ej} is the growth-rate parameter for experimental unit e under condition j , and ε_{eqj} is the residual error.

The population-level growth rate was modelled using an intercept and condition-specific fixed effects:

$$\lambda_{ej} = \beta_{\lambda 0} + x_j^T \beta_{\lambda} + z_j^T b_e^*,$$

where x_j is the design vector for the fixed condition effects and z_j is the corresponding design vector for the random effects.

With treatment coding and six condition levels, the random-effects vector for experimental unit e was

$$b_e^* = (b_{0e}, b_{\text{MALAT1_NT},e}, b_{\text{LACZ_DHT},e}, b_{\text{MALAT1_DHT},e}, b_{\text{LACZ_E2},e}, b_{\text{MALAT1_E2},e})^T,$$

where b_{0e} is the random deviation in the growth-rate intercept and the remaining components are random deviations associated with the condition contrasts.

The random-effects vector was assumed to follow

$$b_e^* \sim N_6(0, D_\lambda^*),$$

where D_λ^* is a general 6×6 positive-definite covariance matrix. Thus, the six random effects are allowed to be correlated within the same experimental unit. Random effects from different experimental units were assumed to be independent.

The residual errors were assumed to have mean zero and common standard deviation:

$$\varepsilon_{eqj} \sim N(0, \sigma^2).$$

Supplementary Figures

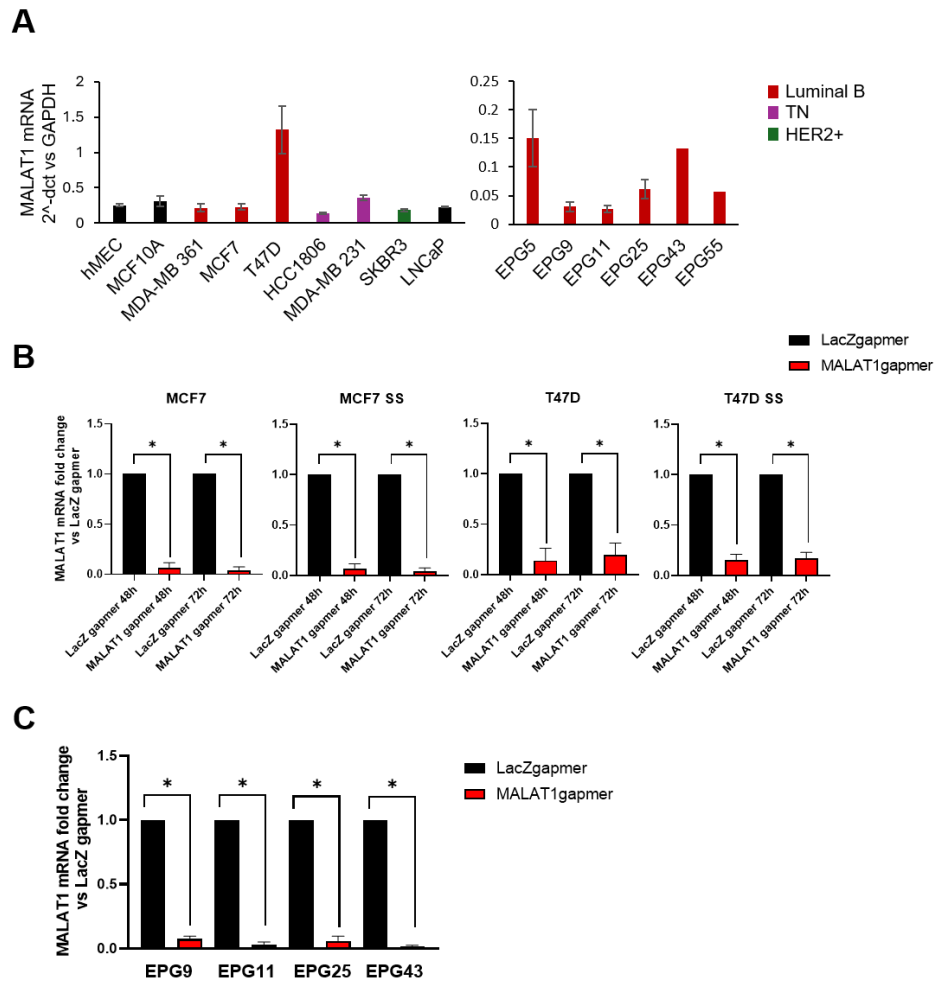
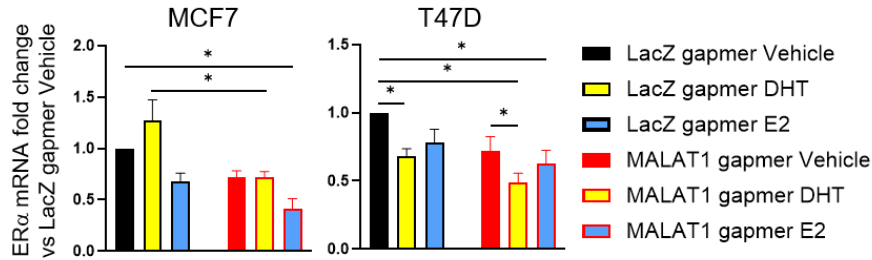
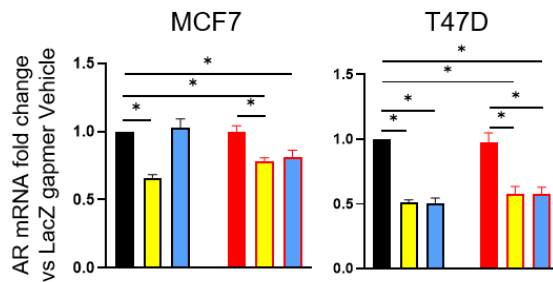


Figure S1. Basal *MALAT1* expression across breast cancer models and targeting efficiency. (A) Basal *MALAT1* expression measured by qRT-PCR as $2^{-\Delta C_t}$ relative to *GAPDH* in established mammary epithelial and breast cancer cell lines, the hormone-responsive prostate cancer cell line LNCaP, and patient-derived breast cancer cultures. Established models include hMEC, MCF10A, MDA-MB-361, MCF7, T47D, HCC1806, MDA-MB-231, SK-BR3, and LNCaP; primary cultures include ePG5, ePG9, ePG11, ePG25, ePG43, and ePG55. Black bars indicate the reference models hMEC, MCF10A, and LNCaP, while breast cancer models are grouped by the molecular categories shown in the color key. hMEC represents normal human mammary epithelial cells, MCF10A a non-tumorigenic mammary epithelial model, and LNCaP a hormone-responsive prostate cancer reference model. (B) *MALAT1* targeting efficiency in MCF7 and T47D 48 h and 72 h after transfection, under complete medium and hormone-deprived (SS) conditions. Expression is reported as fold change relative to LacZ-gapmer and shown as mean $\pm$ SEM from three independent biological experiments. * $p \leq 0.05$. (C) *MALAT1*-targeting efficiency 72 h after transfection in the patient-derived cultures ePG9, ePG11, ePG25, and ePG43. Expression is reported as fold change relative to LacZ-gapmer and shown as mean $\pm$ SEM from three independent biological experiments. $P \leq 0.05$.

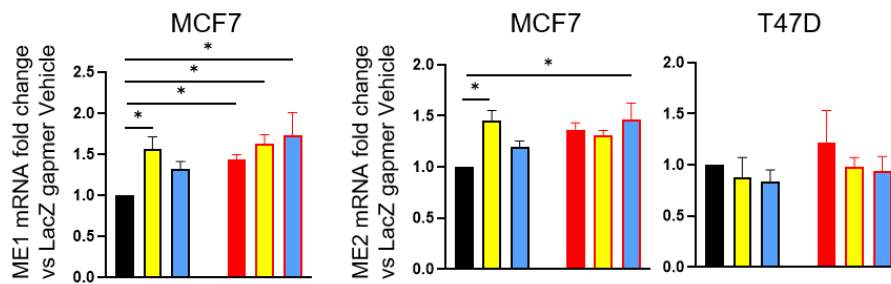
A



B



C



D

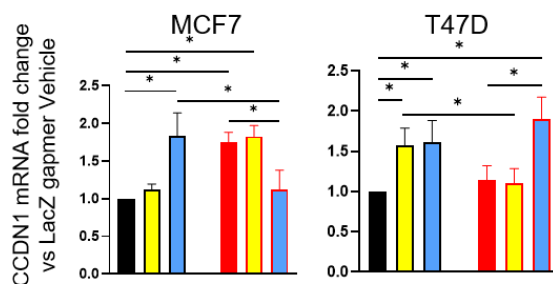


Figure S2. Effects of MALAT1 targeting and hormone treatment on receptor- and metabolic-gene transcripts. qRT-PCR analysis of hormone-receptor and metabolic-gene expression, *ESR1/ERα* (A), *AR* (B), *ME1* and *ME2* (C), and *CCND1* (D), in MCF7 and T47D cells following LacZ control or MALAT1-targeting gapmer transfection (72 h) and exposure to vehicle, DHT (10 nM, 48 h), or E2 (10 nM, 48 h). Expression is reported as fold change relative to LacZ-gapmer vehicle and shown as mean $\pm$ SEM from four independent biological experiments, as in Figure 2. $P \leq 0.05$.

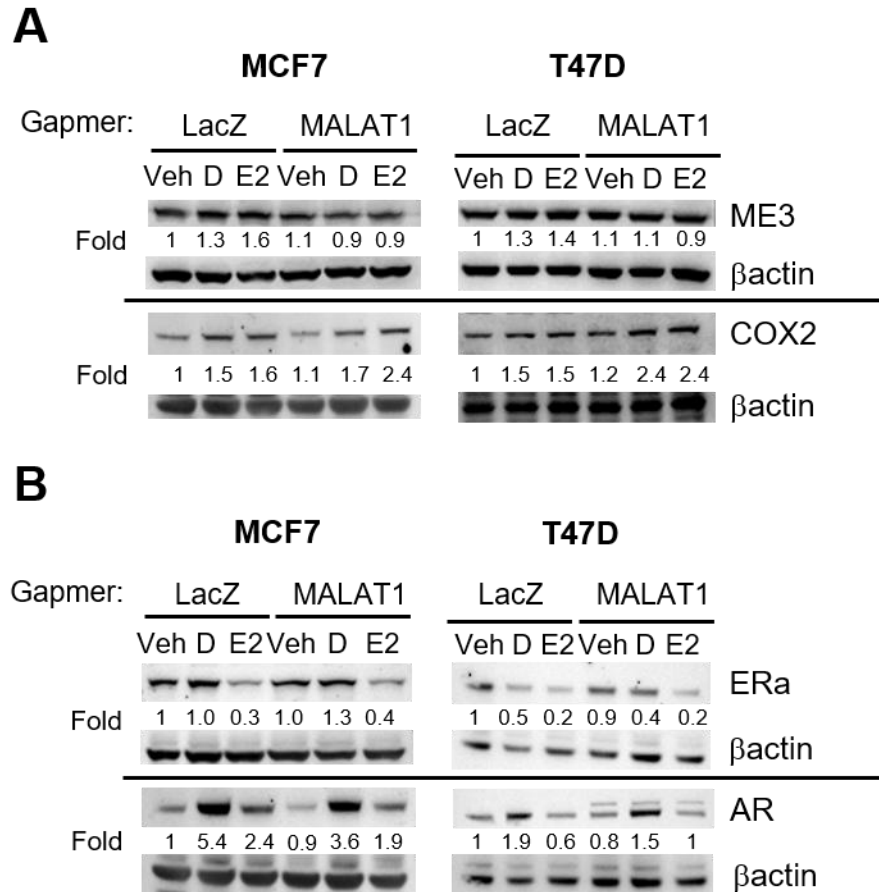
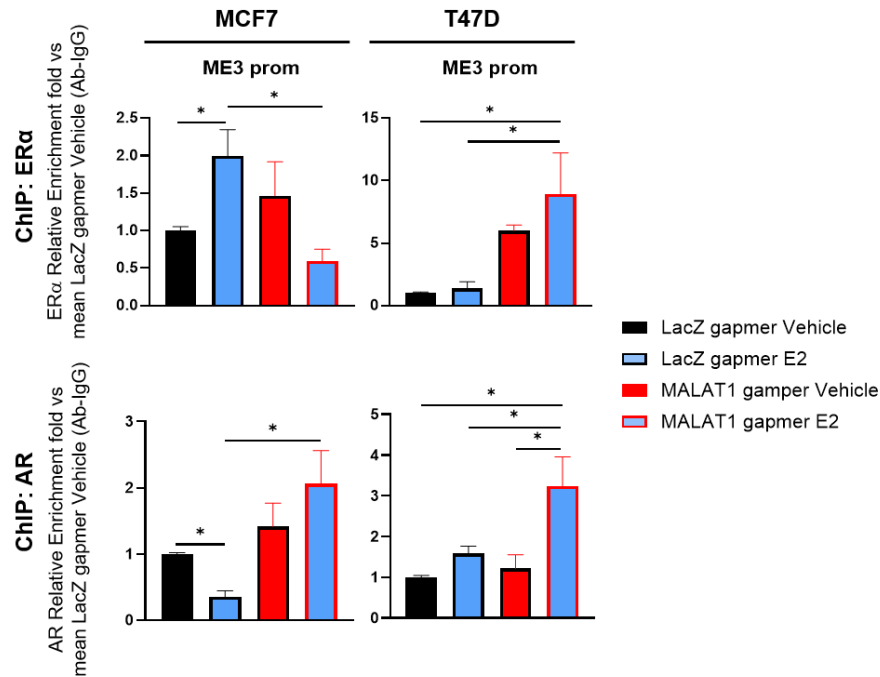


Figure S3. Representative protein changes after MALAT1 targeting and hormone treatment. Representative Western blot analyses from two independent biological experiments performed in MCF7 and T47D cells following LacZ control or MALAT1-targeting gapmer treatment (72 h) and exposure to vehicle (Veh), DHT (D; 10 nM, 48 h), or E2 (10 nM, 48 h). (A) ME3 and COX-2 protein levels in MCF7 and T47D cells. (B) ER α and AR protein levels in the same cell lines. β -actin was used as a loading control. Numbers represent densitometric analysis of protein levels normalized to β -actin and expressed as fold change vs LacZ gapmer vehicle.

A



B

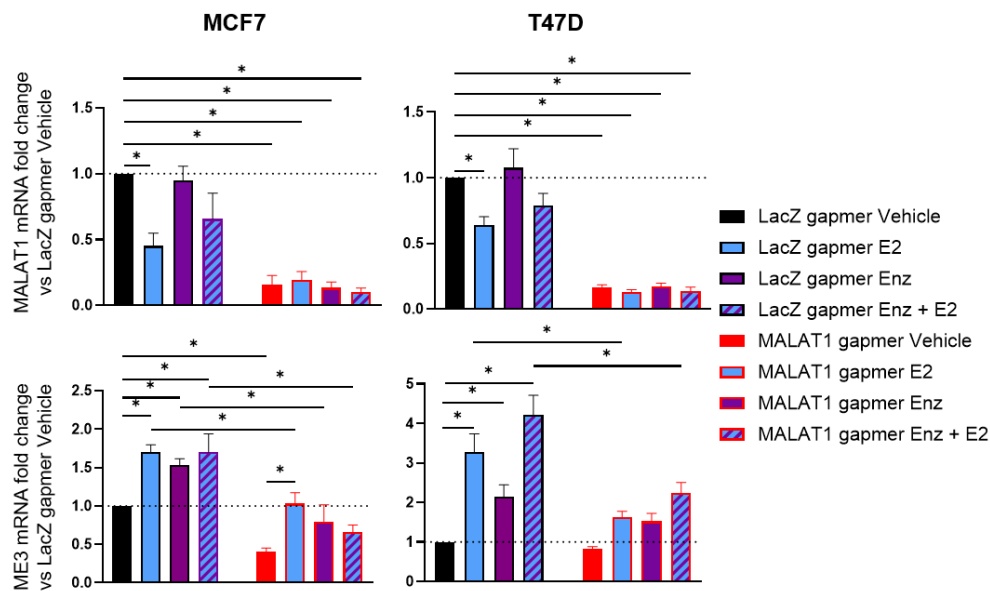


Figure S4. Locus-specific analysis of ERα and AR recruitment and ME3 expression. (A) ChIP-qPCR analysis of ERα and AR occupancy at the *ME3* promoter in MCF7 and T47D cells following LacZ control or MALAT1-targeting gapmer transfection (48 h) and treatment with vehicle or E2 (100 nM, 45 min). ERα occupancy is shown in the upper panels and AR occupancy in the lower panels. Enrichment was normalized to the corresponding IgG control and expressed as fold enrichment relative to the mean LacZ gapmer vehicle condition $\pm$ SEM from three independent biological experiments. (B) qRT-PCR analysis of *MALAT1* and *ME3* mRNA levels in MCF7 and T47D cells after LacZ or MALAT1 gapmer transfection (72 h) and treatment with vehicle, E2 (10 nM, 48 h), ENZ (10 μ M, 48 h), or ENZ + E2. Expression is reported as fold change relative to LacZ-gapmer vehicle and shown as mean $\pm$ SEM from three independent biological experiments as in Figure 4. $P \leq 0.05$.

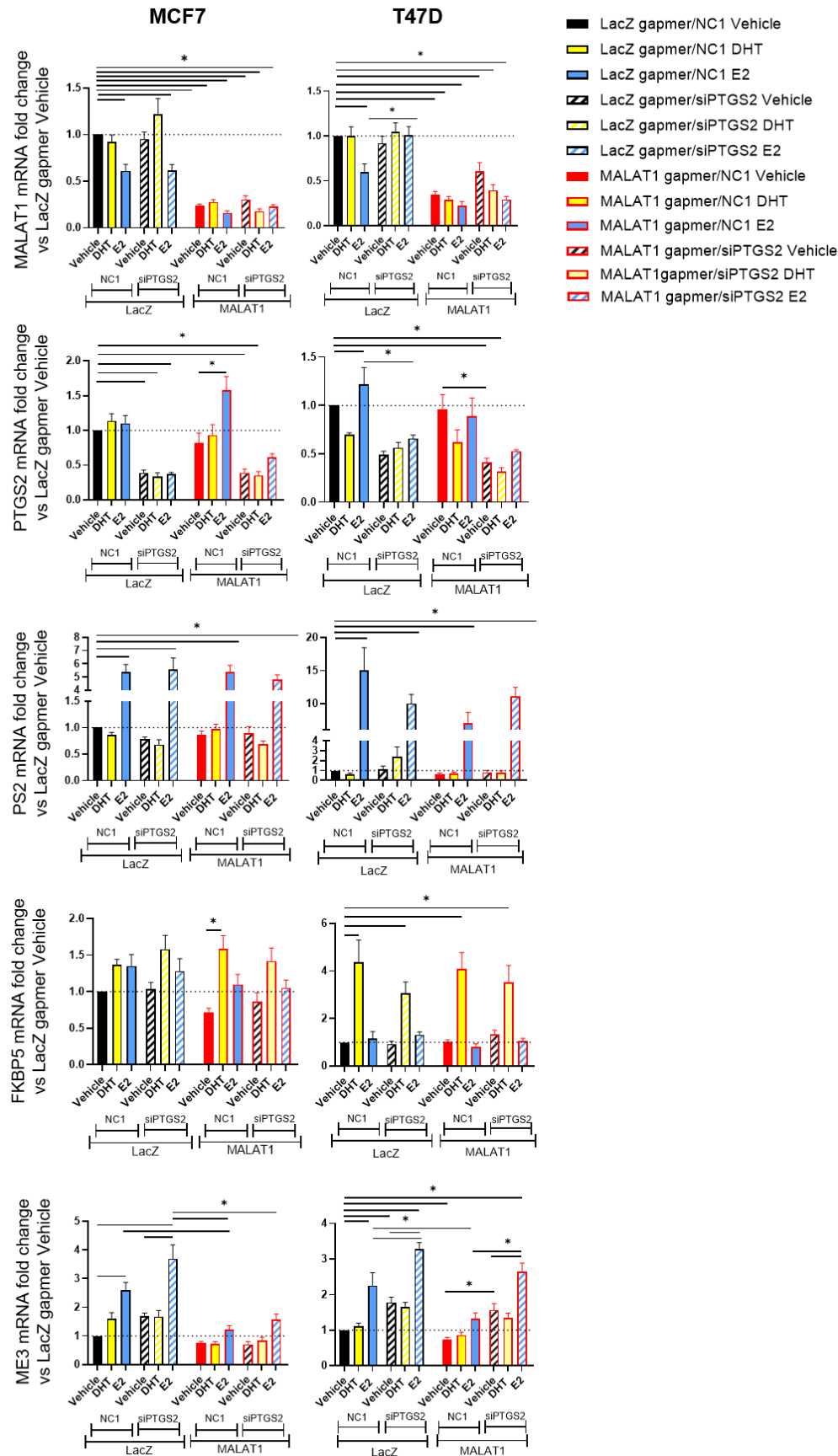


Figure S5. Molecular validation of MALAT1 and PTGS2 co-targeting experiments. qRT-PCR analysis of *MALAT1*, *PTGS2*, *PS2/TFF1*, *FKBP5*, and *ME3* mRNA levels in MCF7 and T47D cells co-transfected with LacZ control or MALAT1-targeting gapmer and either NC1 control siRNA or PTGS2-targeting siRNA. Cells were analyzed after 72 h of transfection under vehicle, DHT (10 nM, 48 h), or E2 (10 nM, 48 h) conditions. Solid bars indicate NC1-treated conditions and hatched bars indicate PTGS2-targeting siRNA conditions. Expression is reported relative to the LacZ-gapmer/NC1 vehicle reference and shown as mean $\pm$ SEM from three independent biological experiments, as in Figure 5. $P \leq 0.05$.

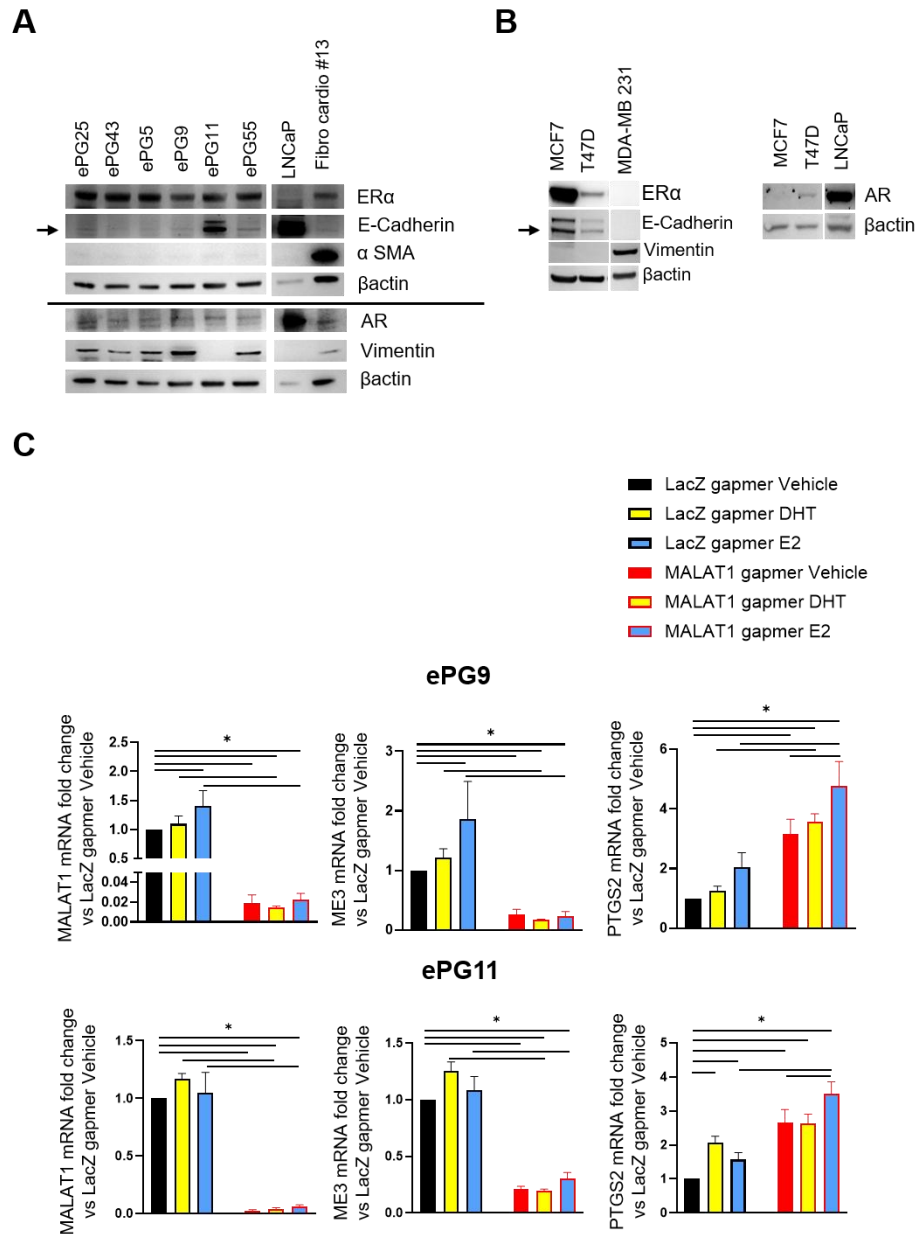


Figure S6. Molecular and phenotypic characterization of patient-derived primary breast cancer cultures.

(A) Western blot characterization of the patient-derived breast cancer cultures ePG25, ePG43, ePG5, ePG9, ePG11, and ePG55 for ERα, AR, E-cadherin, α-smooth-muscle actin (SMA), and vimentin, with β-actin as a loading control. LNCaP cells and cardiac fibroblasts were included as reference samples where indicated. (B) Western blot comparison of MCF7 and T47D cells with MDA-MB-231 or LNCaP as reference cells for epithelial, mesenchymal, estrogen-receptor, and androgen-receptor markers, with β-actin as a loading control. White vertical lines indicate non-contiguous lanes of the same gel. (C) qRT-PCR analysis of *MALAT1*, *ME3*, and *PTGS2* mRNA levels in the representative patient-derived cultures ePG9 and ePG11 after LacZ control or MALAT1-targeting gapmer transfection (72 h) and treatment with vehicle, DHT (10 nM, 48 h), or E2 (10 nM, 48 h). Expression is reported as fold change relative to LacZ-gapmer vehicle and shown as mean ± SEM from three independent biological experiments. $P \leq 0.05$.