Figure S1

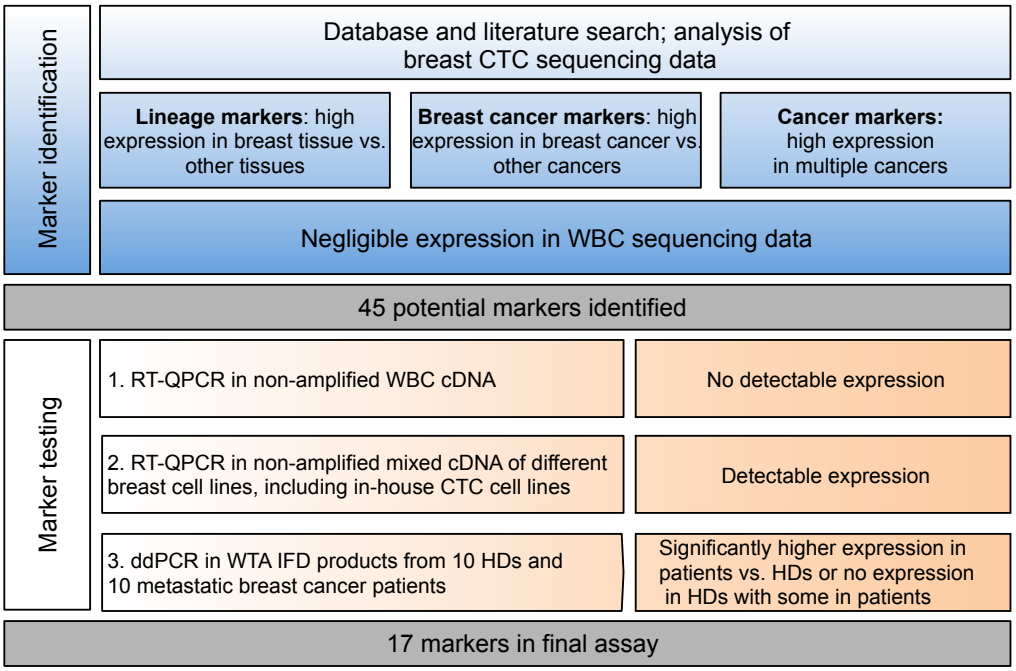


Figure S1. Schematic showing the process involved in the identification of the 17 breast CTC-specific transcripts. The queried databases include Oncomine (www.oncomine.org), GTEx (www.gtexportal.org) and TCGA (cancergenome.nih.gov).

Figure S2

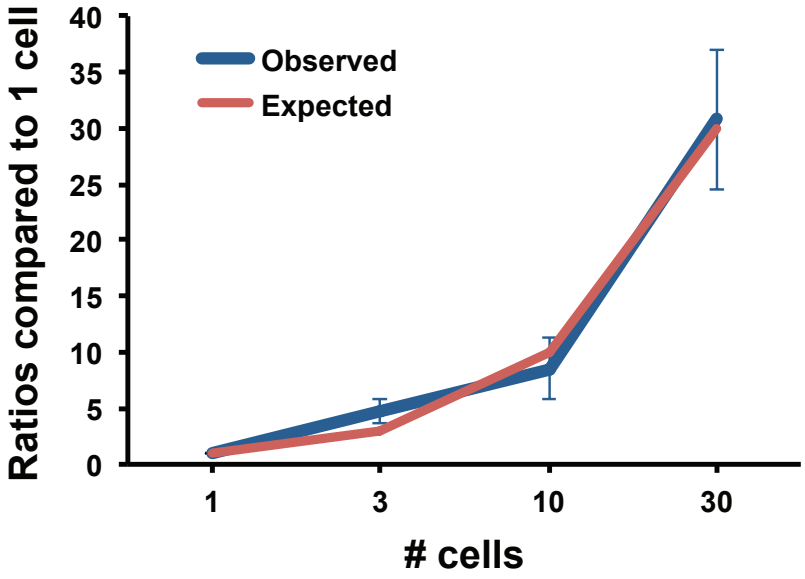


Figure S2. Expected and observed transcript signals from 1 to 30 CTCs (from a cultured CTC cell line BRX-142) individually manipulated into 4 ml of whole blood, followed by CTC-iChip enrichment and ddPCR analysis. Expected ratios are based on the number of cells added compared to the 1 cell sample; observed ratios are based on the average total CTC score for each sample compared to the 1 cell sample.

Figure S3

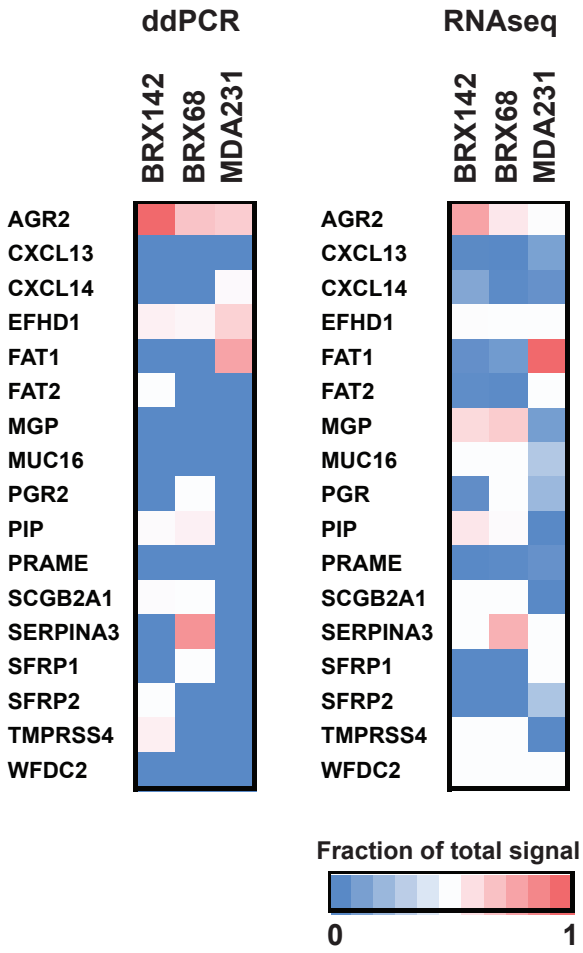


Figure S3. Heat map showing relative marker contribution to breast CTC assay signal as measured by ddPCR of 30 cells spiked into 4ml of HD blood (left) or bulk RNAseq (right) of BRx-142, BRx-68 and MDA-231 cells. Contribution of each marker was determined by dividing its individual signal (transcripts/ml for ddPCR or RPM for RNAseq) by the sum total of all 17 markes for each cell line.

Figure S4

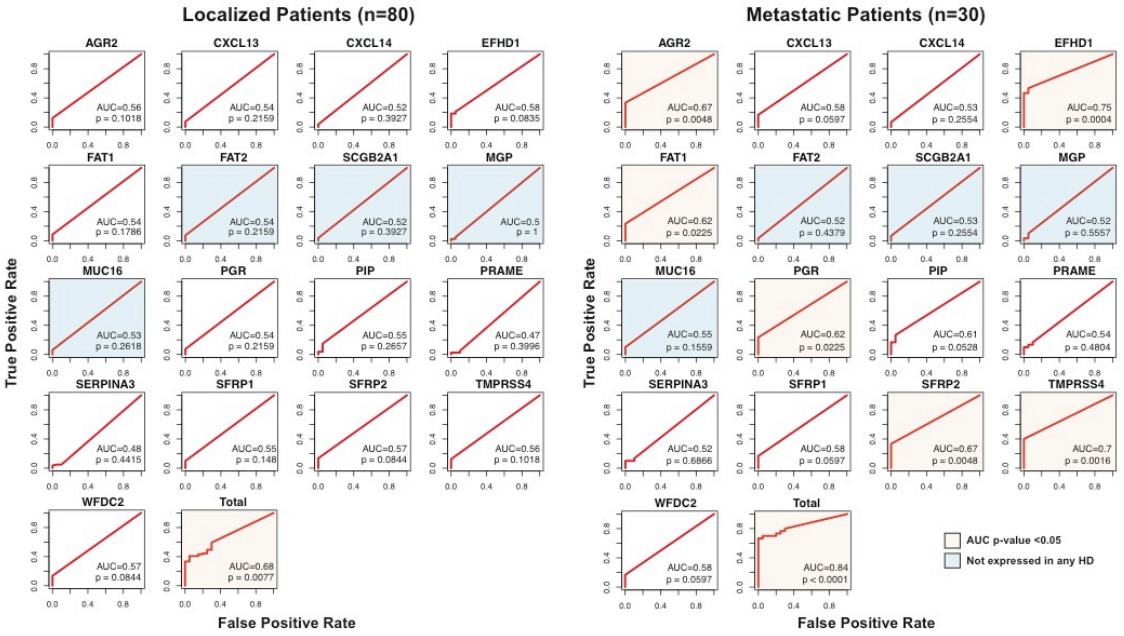


Figure S4. ROC analyses of individual markers and total CTC-Score in localized and metastatic patients. Comparisons were between 20 healthy female donors and pre-treatment samples from 80 localized patients (Stage I, II and III) or 30 mostly on-treatment metastatic patient samples. AUC values are shown; p-values are based on Wilcoxon rank sum test. Markers with p-value <0.05 are highlighted in orange; markers with 100% specificity are highlighted in blue.

Figure S5

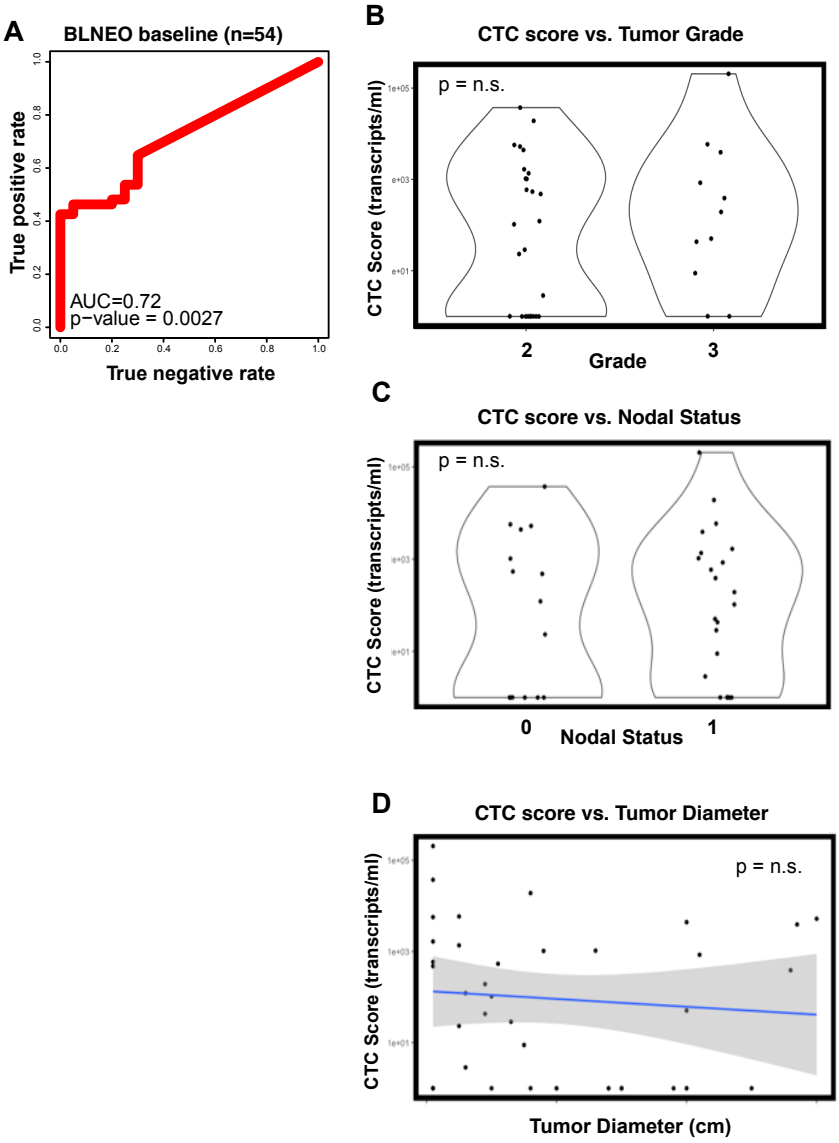


Figure S5. **A,** ROC analysis of baseline BLNEO samples. Comparison between 20 healthy female donors and pre-treatment samples from 54 BLNEO patients (Stages I, II and III). AUC value is shown; p-value was calculated using Wilcoxon rank sum test. **B,** Boxplot showing CTC scores in Grade 2 and Grade 3 tumors; p-value based on Wilcoxon Rank Sum test. **C,** Boxplot showing CTC scores in paients with (1) or without (0) nodal involvement at baseline; p-value based on Wilcoxon Rank Sum

test. D, Scatterplot of CTC score and tumor diameter at baseline. Best-fit line and confidence interval are shown, p-value based on linear regression statistics.

Figure S6

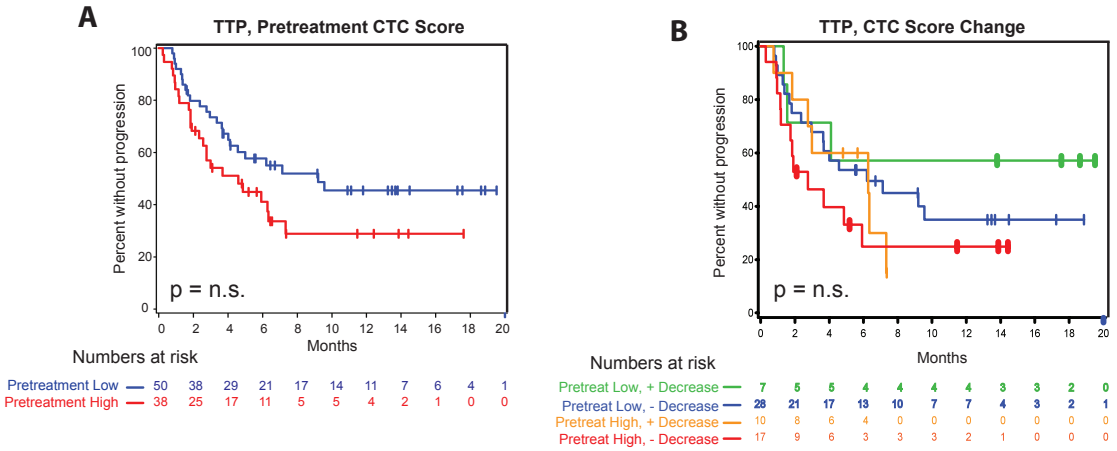


Figure S6. Pretreatment CTC-Score is not predictive of TTP **A**, Kaplan-Meier plot of TTP in patients in the TRACK cohort, based on CTC-Score at pre-treatment. Patients were divided into two groups at a cut-off of 3000 transcripts/ml (see Methods). p-values are based on multivariable Cox proportional hazards model (high CTC-Score (red) versus low CTC-Score (blue)). **B**, Kaplan-Meier plot of TTP in TRACK patients, based on changes in pre-treatment and 3-4 weeks on-treatment CTC-Scores. Groups are defined based on high or low signal at pretreatment (divided at 3000 transcripts/ml) and the magnitude of the change after 3-4 weeks of treatment (divided at 90% decrease in signal). Low pretreatment CTC-Score with >90% reduction in signal (green) versus low pretreatment CTC Score without >90% reduction in signal (blue) versus high pretreatment CTC-Score with >90% reduction in signal (orange) versus high CTC-Score without >90% reduction in signal (red).

Figure S7

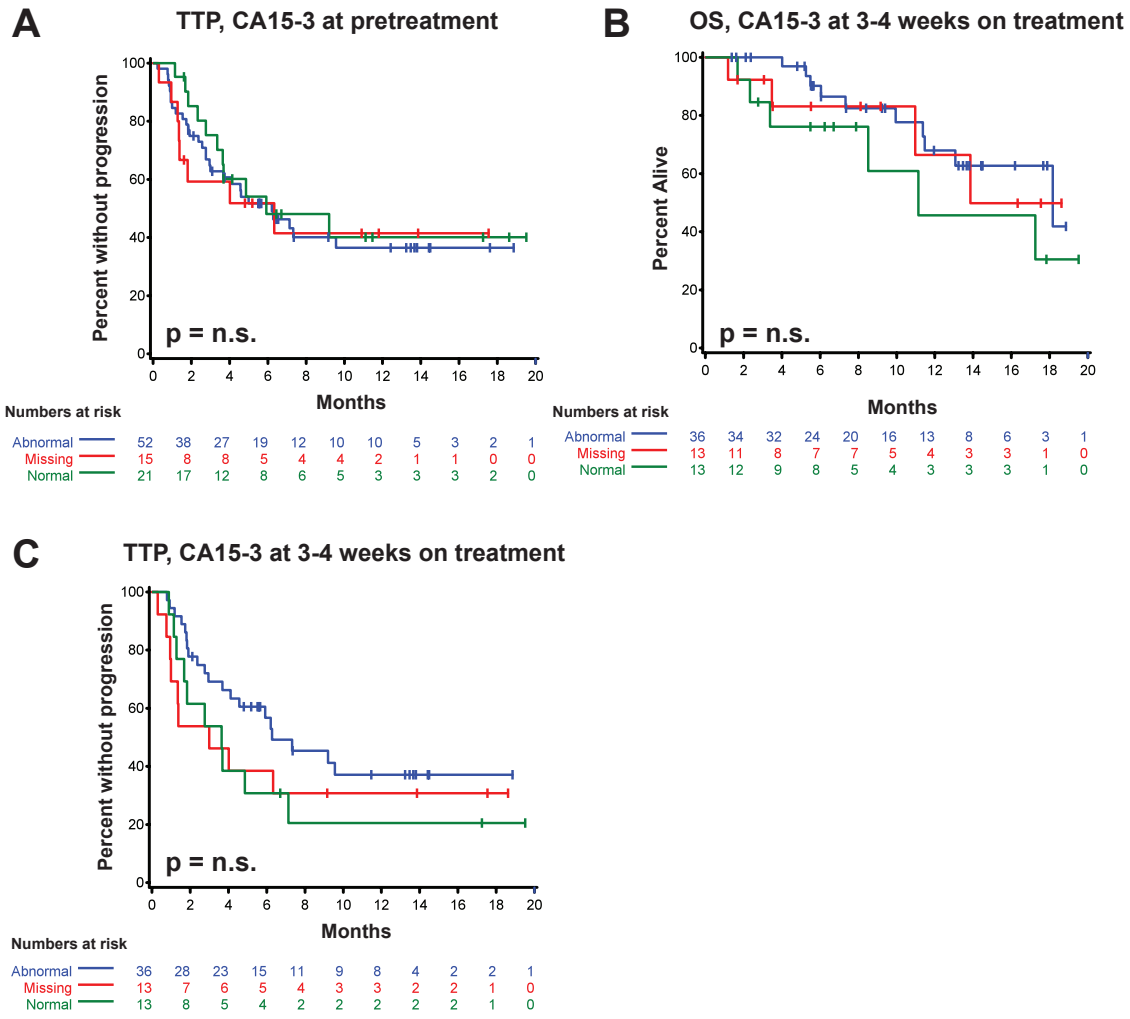


Figure S7. CA15-3 tumor marker levels at baseline or 3-4 weeks on treatment are not predictive of OS or TTP. **A**, Kaplan-Meier plot of TTP in TRACK patients, based on CA15-3 levels at pre-treatment. Groups are defined based on normal CA15-3 levels (<=30, green), abnormal CA15-3 levels (>30, blue), and missing CA15-3 levels (NA, red). P-value was calculated using log-rank test. **B**, Kaplan-Meier plot of OS in TRACK patients, based on CA15-3 levels at 3-4 weeks on treatment. Groups are

defined based on normal CA15-3 levels (<=30, green), abnormal CA15-3 levels (>30, blue), and missing CA15-3 levels (NA, red). P-values are based on log-rank test. **C**, Kaplan-Meier plot of TTP in TRACK patients, based on CA15-3 levels at 3-4 weeks on treatment. Groups are defined based on normal CA15-3 levels (<=30, green), abnormal CA15-3 levels (>30, blue), and missing CA15-3 levels (NA, red). P-value based on log-rank test.

Figure S8

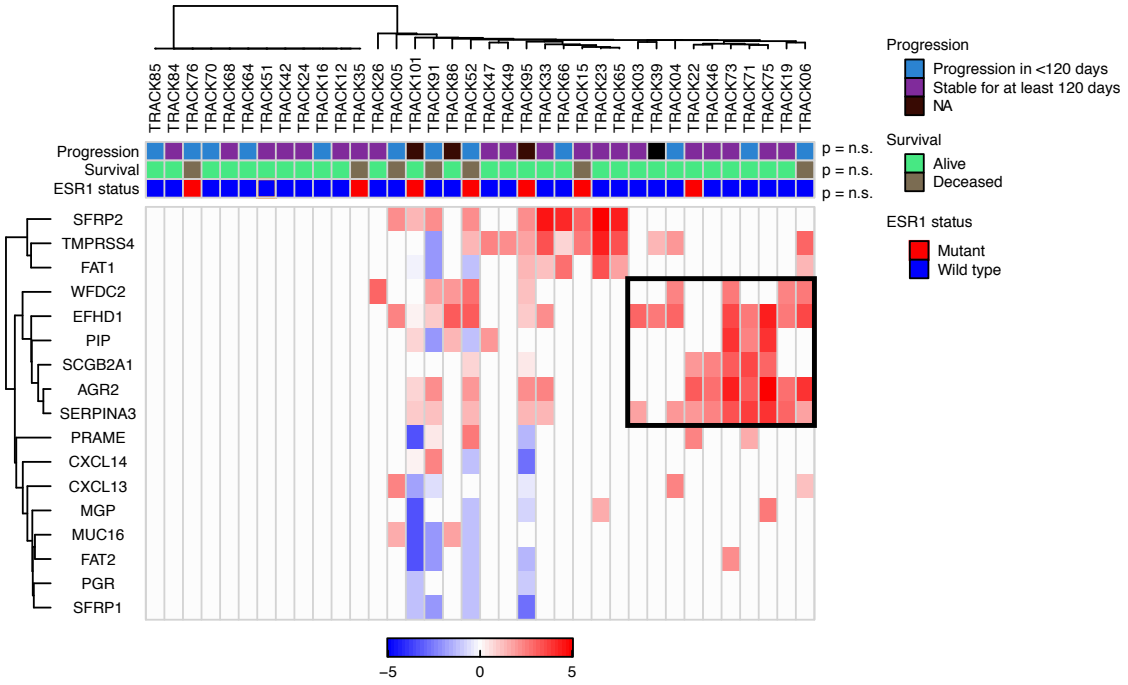


Figure S8. Unsupervised clustering of breast CTC marker expression in pre-treatment samples from patients with HR+ disease starting endocrine-based treatment. High levels of the 6 RS genes in pretreatment samples do not correlate with faster disease progression (<120 days), poor survival, or the presence of ESR1 mutations (Fisher's exact test). Clustering performed using single linkage.

Figure S9

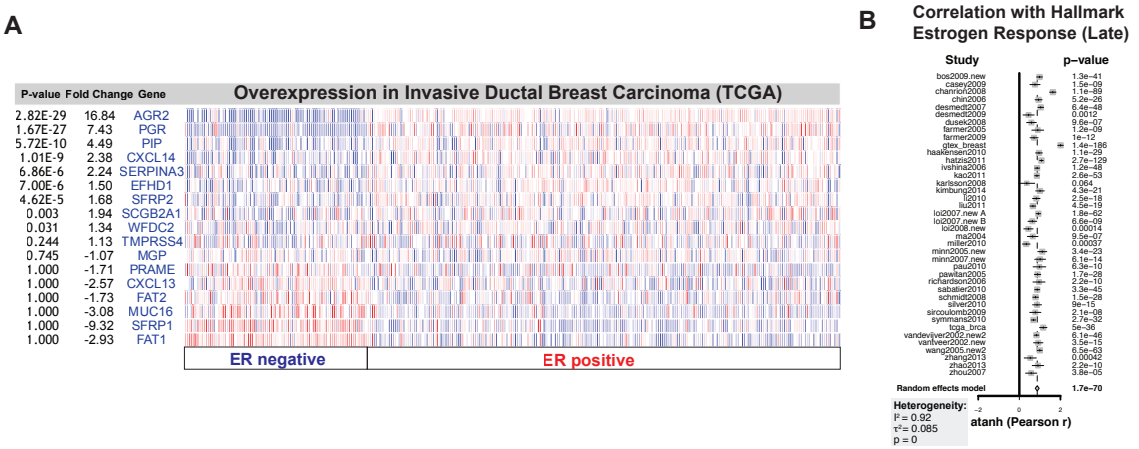
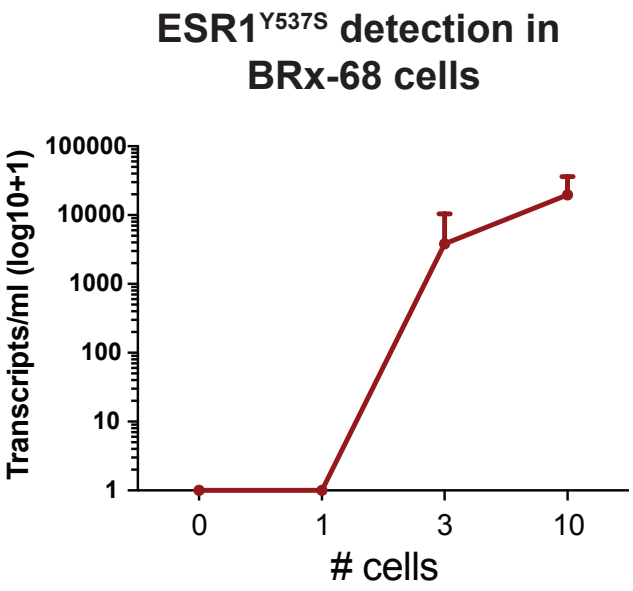


Figure S9. RS genes are associated with ESR1 signaling and endocrine resistance. **A**, Heatmap showing the expression (log2 median-centered ratios) of the 17 breast CTC markers in ER-negative and ER-positive invasive ductal breast carcinoma samples (TCGA). Markers are arranged from most enriched to the most depleted in ER+ samples. Fold changes are shown on the left; p-values are based on two-sample t-tests. **B,** Forest plot showing positive correlation (atanh Pearson's r) between the 6 RS-gene metascore and the Hallmark Estrogen Receptor Signaling (Late) signature across multiple publicly available gene expression data sets.

Figure S10

A



B

Patient ID	ESR1 mutation by genotyping	Biopsy Site	Mutataion by CTC-ddPCR
BRX294	L536R Y537N	Metastatic	L536R Y537N

Figure S10. Detection of ESR1 mutations in-vitro and in an ESR1 mutant patient. **A**, Detection of $ESR1^{Y537S}$ by RNA-based ddPCR assay, following the addition of 1 to 10 individually manipulated cultured CTCs (BRx-68) into 4ml of healthy donor (HD) whole blood and CTC-iChip microfluidic enrichment (n=3, dots represent means, error bars show standard deviation). **B**, A table showing the detection of $ESR1^{Y537N}$ and $ESR1^{L536R}$ mutations by ddPCR of an on-treatment blood sample from a patient previously shown to harbor these mutations by genotyping of the metastatic biopsy.

Figure S11

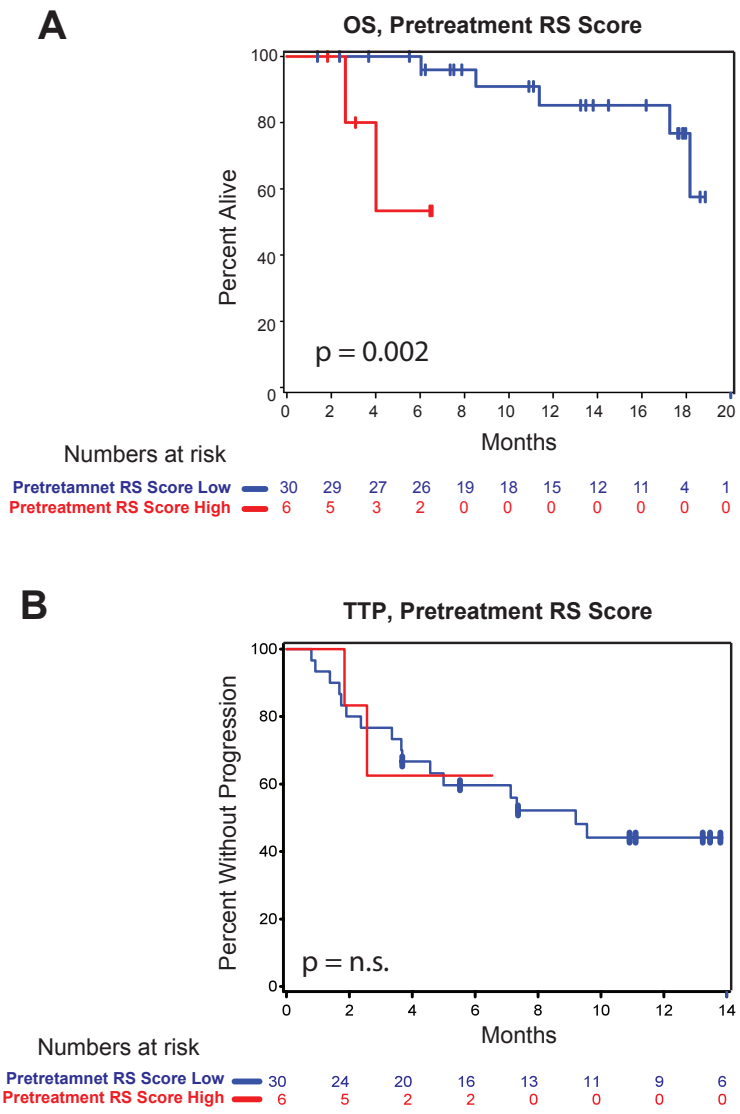


Figure S11. Kaplan Meier plots of OS and TTP in HR+ patients receiving endocrine therapy based on RS score at pretreatment. Patients were divided into two groups at a RS Score cut-off of 12300 transcripts/ml. Cases with high RS Score (red) are compared with those having a low RS Score (blue).