Title

Gene expression associated with endocrine therapy resistance in estrogen receptor-positive breast cancer.

Journal

Scientific Reports, 15(1)

Authors

Jones, Veronica

Yin, Hongwei

Yuan, Yate-Ching

et al.

Publication Date

2025-02-28

DOI

10.1038/s41598-025-89274-9

Peer reviewed

scientific reports

Check for updates

OPEN

Gene expression associated with endocrine therapy resistance in estrogen receptor-positive breast cancer

Veronica Jones[1]✉, Hongwei Holly Yin[2], Yate-Ching Yuan[3], Yongzhe Wang[1], Sierra Min Li[4], Dana Aljaber[5], Angelica Sanchez[6], Christine Quinones[1], Dan Schmolze[7], Yuan Yuan[8], Joanne Mortimer[9], Lisa Yee[1], Laura Kruper[10], Tijana Jovanovic-Talisman[11], Jerneja Tomsic[6], Nancy Sanchez[6], Tanya Chavez[12], Ruth M. O'Regan[13], Qamar J. Khan[14], Melissa Davis[15], Kevin Kalinsky[16], Jane Meisel[16], Rick Kittles[17], Lorna Rodriguez-Rodriguez[1] & Victoria Seewaldt[6]

Despite endocrine therapy (ET), approximately 20–40% of Stage I–III estrogen receptor-positive breast cancer (ER + BC) patients experience recurrence. Recurrence while on ET is indicative of ET resistance. This study aimed to identify differentially expressed genes (DEGs) associated with recurrence during ET (ET resistance) and to explore gene expression differences across PAM50 molecular subtypes. Eighty tumor specimens from 79 patients treated at the City of Hope Comprehensive Cancer Center (2012–2016) were analyzed using NanoString technology. Fourteen patients (17.7%) experienced recurrence over a median follow-up of 68 months (range 35–104 months). Key upregulated DEGs in the recurrence group included *EZH2* (log2 fold change[log2FC]: 0.67, $p = 0.0017$), *WNT11* (log2FC: 1.08, $p = 0.0088$), *ITGB6* (log2FC: 0.80, $p = 0.0312$), and *TOP2A* (log2FC: 0.79, $p = 0.0381$). Downregulated DEGs included *SNAI2* (log2FC: − 0.63, $p = 0.0055$), *ITPR1* (log2FC: − 0.75, $p = 0.0083$), *CD10* (log2FC: − 0.70, $p = 0.0092$), *PTEN* (log2FC: − 0.29, $p = 0.0163$), *VRD* (log2FC: − 0.46, $p = 0.0184$), and *WNT5A* (log2FC: − 0.76, $p = 0.0272$). *EZH2* and *TOP2A* were positively correlated with proliferation scores, while *WNT11* and *ITGB6* emerged as potential biomarkers independently associated with recurrence. These findings suggest novel biomarker candidates that could help overcome ET resistance, reduce recurrence, and improve outcomes in ER + BC.

Precision medicine has changed breast cancer treatment. However, the heterogeneity of breast cancers, particularly estrogen receptor-positive breast cancer (ER + BC), has made the development of targeted agents challenging. Because of this, approximately 20–40% of Stage I-III patients with ER + BC will develop metastatic disease[1]. Once ER + BC recurs, patients have a threefold increased incidence of death, with as many as 65% dying within ten years after locoregional recurrence (LRR) and 80% within four years after distant recurrence (DR)[2–4].

[1]Department of Surgery, City of Hope, 1500 E Duarte Rd, Duarte, CA 91010, USA. [2]Department of High Throughput Screening, City of Hope, 1500 E Duarte Rd, Duarte, CA 91010, USA. [3]Department of Computational Quantitative Medicine, Beckman Research Institute, City of Hope, 1500 E Duarte Rd, Duarte, CA 91010, USA. [4]Department of Biostatistics, City of Hope, 1500 E Duarte Rd, Duarte, CA 91010, USA. [5]University of California, Davis School of Medicine, Davis, CA, USA. [6]Department of Population Sciences, City of Hope, 1500 E Duarte Rd, Duarte, CA 91010, USA. [7]Department of Pathology, City of Hope, 1500 E Duarte Rd, Duarte, CA 91010, USA. [8]Department of Medical Oncology, Cedars-Sinai Medical Center, Los Angeles, CA, USA. [9]Department of Medical Oncology and Therapeutics Research, City of Hope, 1500 E Duarte Rd, Duarte, CA 91010, USA. [10]Department of Surgery, Moffitt Cancer Center, Tampa, FL, USA. [11]Department of Cancer Biology and Molecular Medicine, City of Hope, 1500 E Duarte Rd, Duarte, CA 91010, USA. [12]University of California, San Diego, CA, USA. [13]Department of Medicine, University of Rochester, 125 Red Creek Dr, Rochester, NY 14623, USA. [14]Department of Medicine, University of Kansas, 3901 Rainbow Boulevard, Kansas City, KS 66160, USA. [15]Institute of Translational Genomic , Morehouse School of Medicine, Medicine, Morehouse College720 Westview Drive SW, Atlanta, GA 30310, USA. [16]Department of Hematology and Medical Oncology, Emory University, 1365-C Clifton Road NE, Atlanta, GA 30322, USA. [17]Office of Research Affairs, Morehouse College, Morehouse School of Medicine, 720 Westview Drive SW, Atlanta, GA 30310, USA. ✉email: vjones@coh.org

Given that ER + BC comprises approximately 70% of all breast cancer diagnoses, there is a need to define the molecular signaling pathways that drive therapy resistance, recurrence, and subsequent worse survival; drugs targeting these biomarkers will cause a dramatic improvement in breast cancer mortality.

While there has been thorough examination of gene expression in ER + BC, there has been limited evolution of precision medicine in the management of ER + BC. ER + BC are highly heterogenous with varying prognoses based on molecular expression; yet this subtype is treated with relative uniformity. Current management of locally advanced, non-metastatic ER + BC relies on local management combined with endocrine therapy with or without chemotherapy. Gene expression assays such as Oncotype DX ® and Mammaprint® are often used to guide the decision regarding chemotherapy administration; these tests do not guide the administration of endocrine therapy[5–7]. Tests such as the Breast Cancer Index (BCI) have been utilized to direct the length of administration of endocrine therapy but do not predict the development of resistance[8]. To date, predictive biomarkers guiding the administration of endocrine therapy have not been clinically utilized. Endocrine therapy is given to all ER + patients despite de novo resistance occurring in 15–20% of patients[9]. Moreover, the side effects related to endocrine therapy are not insignificant, with patients subjected to an increased risk of cardiovascular events and other morbidities[10]. Thus, determining the benefit of endocrine therapy is paramount to spare patients' unnecessary co-morbid conditions. Patients who are resistant to endocrine therapy and recur should be identified early to guide therapy and be spared morbid side effects.

Endocrine therapy resistance however has been traditionally difficult to study. This is because of the timing of endocrine therapy administration in the adjuvant setting. Because endocrine therapy is often given after the disease has been resected, resistance is only identified at the time of recurrence. Recurrence, therefore, serves as a critical endpoint to use to measure resistance to endocrine therapy. Recurrence data is not often available in publicly available cohorts as it requires in-depth longitudinal tracking of patients. Therein lies the strength of institutional cohorts to complement large publicly available cohorts; institutional data with complete clinical information including treatment and outcomes data may offer critical insights into biomarkers of recurrence while on endocrine therapy. In this study, we interrogated the mRNA expression profile of early-stage ER + BCs treated with standard-of-care endocrine therapy to determine expression associated with recurrence. Here in our institutional cohort, we report the differentially expressed genes (DEGs) significantly associated with early recurrence despite optimal endocrine therapy, indicative of ET resistance. This study identifies potential biomarker candidates of ET resistance, which may guide future research regarding upfront treatment for these cancers to aid in therapeutic decision-making in preventing recurrence and improving survival.

Results

Patient characteristics

Eighty specimens, representing 79 patients (one patient had bilateral disease), met the study inclusion criteria. Most patients had primary surgery performed between 2015 and 2016. Two had surgeries in 2012–2013. Of the 79 patients included in the analysis, the majority were non-Hispanic White (n = 42, 53.2%), 17 (21.5%) were Hispanic White, 17 (21.5%) were Asian, and two were American Indian and Alaska Native; one patient declined to self-report race and ethnicity. Table 1 shows patient demographic characteristics. The mean age was 56 (range 20–87), and the most common histology was invasive ductal carcinoma (n = 60, 75.9%). The mean tumor size was 3.33 cm (range 0.6–12 cm). Of the 79 patients, 25 (31.6%) had clinical nodal involvement, and 51 (64.6%) had pathological node involvement. Sixty-six tumors had Ki-67 performed; of those, 28 (35.4%) had a Ki-67 $\geq$ 20%. The median time of follow-up was 68 months (range 35–104).

At the time of treatment, four patients had an initial positive surgical margin; three underwent re-excision with ultimate clearing of the margins. Twenty patients (25.3%) received neoadjuvant chemotherapy. Of the 79 patients, 54 (68.4%) underwent radiation therapy, 30 (38%) received bone-directed therapy, 9 (11.4%) had ovarian suppression therapy, and all (100%) received adjuvant endocrine therapy (ET). For this study, endocrine therapy included selective estrogen receptor modulators (SERMs) such as tamoxifen and raloxifene, selective estrogen receptor degraders (SERDs) such as fulvestrant, as well as aromatase inhibitors (AI). Ovarian suppression was not included as endocrine therapy but was recorded in our cohort as a separate study variable. Patients may have received multiple agents throughout the observational period, based on tolerability and medical oncology preference. All patients were documented to have received and taken endocrine therapy, though due to the variability of administration, the study was not powered to detect differences across treatment regimens. Sixty-three patients received one type of endocrine therapy treatment (79.7%) and 15 patients received 2 or more types of endocrine therapy (19%). Of the patients who received 2 or more types of endocrine therapy, most reported intolerance to their prescribed medication as the reason for receiving an alternative endocrine therapy (n = 11, 73.3%) while some reported nonadherence as the reason for receiving an alternative endocrine therapy (n = 2, 13.3%); two patients' endocrine therapy were switched by their physicians in response to life changes such as menopause and surgery (13.3%).

Fourteen patients (17.7%) had recurrence; seven patients had locoregional recurrence (LRR) and seven had distant recurrence (DR). Mean time to recurrence was 32 months (range 6–51 months). There were three breast cancer deaths and one due to other causes. There was no difference between the recurrence and non-recurrence groups with respect to receipt of chemotherapy ($p = 0.70$). There was a positive association between high proliferation index as captured by Ki-67 $\geq$ 20% and recurrence; 64.3% of the recurrent patients had high Ki-67 ($p = 0.03$). Because of the heterogeneity of ER + BC, recurrence based on molecular subtype rather than simply immunohistochemical (IHC) staining was also ascertained. Most specimens were luminal A (n = 49, 62%) or luminal B (n = 27, 34.2%). All specimens were estrogen receptor positive and HER2 negative by IHC staining; however, on molecular subtyping three patient specimens were actually HER2 enriched (3.8%). Though not statistically significant, patients with luminal B breast cancer were more likely to recur (29.6% vs. 13.7% for luminal A, $p = 0.053$).

Characteristics	Total (n = 79; %)	Recurrence status		*P* value
		No Recurrence (n = 65; %)	Recurrence (n = 14; %)	
Age at diagnosis [mean, (SD)]	56.35 (13.61)	56.91 (13.09)	53.79 (16.13)	0.34
BMI at Diagnosis [mean, (SD)]	27.74 (5.58)	27.65 (5.4)	28.16 (6.56)	0.36
KI67 [mean, (SD)]	20.55 (17.82)	19.67 (18.87)	24.5 (11.88)	0.09
Race and ethnicity				
White, non-Hispanic	42 (53.16%)	32 (49.23%)	10 (71.43%)	0.21
Hispanic	17 (21.52%)	16 (24.62%)	1 (7.14%)	
Asian, non-Hispanic	17 (21.52%)	15 (23.08%)	2 (14.29%)	
AIAN, non-Hispanic	2 (2.53%)	1 (1.54%)	1 (7.14%)	
Missing	1 (1.27%)	1 (1.54%)	0 (0%)	
Diabetes at diagnosis				
No	70 (88.61%)	57 (87.69%)	13 (92.86%)	0.93
Yes	9 (11.39%)	8 (12.31%)	1 (7.14%)	
Menstrual				
Premenopausal	29 (36.71%)	23 (35.38%)	6 (42.86%)	0.8
Postmenopausal	49 (62.03%)	41 (63.08%)	8 (57.14%)	
Missing	1 (1.27%)	1 (1.54%)	0 (0%)	
Cancer stage				
I	2 (2.53%)	2 (3.08%)	0 (0%)	0.69
II	56 (70.89%)	45 (69.23%)	11 (78.57%)	
III	21 (26.58%)	18 (27.69%)	3 (21.43%)	
Tumor grade				
I	4 (5.06%)	2 (3.08%)	2 (14.29%)	0.22
II	63 (79.75%)	53 (81.54%)	10 (71.43%)	
III	12 (15.19%)	10 (15.38%)	2 (14.29%)	
Subtype				
Luminal A	49 (62.03%)	43 (66.15%)	6 (42.86%)	0.12
Luminal B	27 (34.18%)	19 (29.23%)	8 (57.14%)	
Her2-enriched	3 (3.8%)	3 (4.62%)	0 (0%)	
Breast procedure				
Total mastectomy	36 (45.57%)	29 (44.62%)	7 (50%)	0.2
Partial mastectomy	21 (26.58%)	20 (30.77%)	1 (7.14%)	
Nipple-sparing Mastectomy	14 (17.72%)	11 (16.92%)	3 (21.43%)	
Skin-sparing Mastectomy	8 (10.13%)	5 (7.69%)	3 (21.43%)	
Lymph node procedure				
Sentinel lymph node biopsy	38 (48.1%)	32 (49.23%)	6 (42.86%)	0.17
Axillary lymph node dissection	26 (32.91%)	18 (27.69%)	8 (57.14%)	
Both	9 (11.39%)	9 (13.85%)	0 (0%)	
None	2 (2.53%)	2 (3.08%)	0 (0%)	
Unknown	4 (5.06%)	4 (6.15%)	0 (0%)	
Lymph node status				
Negative	28 (35.44%)	24 (36.92%)	4 (28.57%)	0.78
Positive	51 (64.56%)	41 (63.08%)	10 (71.43%)	
Neoadjuvant therapy				
No	59 (74.68%)	49 (75.38%)	10 (71.43%)	0.99
Yes	20 (25.32%)	16 (24.62%)	4 (28.57%)	
Radiation therapy				
No	24 (30.38%)	17 (26.15%)	7 (50%)	0.16
Yes	54 (68.35%)	47 (72.31%)	7 (50%)	
Unknown	1 (1.27%)	1 (1.54%)	0 (0%)	
Chemotherapy				
No	29 (36.71%)	25 (38.46%)	4 (28.57%)	0.7
Yes	50 (63.29%)	40 (61.54%)	10 (71.43%)	
Bone directed therapy				
No	49 (62.03%)	40 (61.54%)	9 (64.29%)	0.99
Continued				

Characteristics	Total (n = 79; %)	Recurrence status		*P* value
		No Recurrence (n = 65; %)	Recurrence (n = 14; %)	
Yes	30 (37.97%)	25 (38.46%)	5 (35.71%)	
Ovarian suppression therapy				
No	70 (88.61%)	58 (89.23%)	12 (85.71%)	0.99
Yes	9 (11.39%)	7 (10.77%)	2 (14.29%)	
Endocrine therapy				
Anastrozole	22 (27.85%)	19 (29.23%)	3 (21.43%)	0.48
Exemestane	1 (1.27%)	1 (1.54%)	0 (0%)	
Letrozole	22 (27.85%)	18 (27.69%)	4 (28.57%)	
Tamoxifen	18 (22.78%)	16 (24.62%)	2 (14.29%)	
Two or more	15 (18.99%)	10 (15.38%)	5 (35.71%)	
Unknown	1 (1.27%)	1 (1.54%)	0 (0%)	

Table 1. Patient demographics and clinical characteristics stratified by recurrence status.

Gene	log2 fold change	Standard error (log2 fold change)	Linear fold change	*P* value
EZH2	0.672888593	0.214944597	3.130521089	0.001744965
SNAI2	− 0.632297011	0.227948181	− 2.773862935	0.005539501
ITPR1	− 0.775259287	0.293925839	− 2.63760168	0.008349459
WNT11	1.078871165	0.411821509	2.619754288	0.008799314
CD10	− 0.699522153	0.268445313	− 2.605827402	0.009165263
PTEN	− 0.287858671	0.119846954	− 2.401885584	0.016310809
VRD	− 0.460172201	0.195164213	− 2.357871831	0.018380038
SYTL4	− 0.566694538	0.247211435	− 2.292347595	0.021885593
MDM2	− 0.334130962	0.147377142	− 2.267183074	0.023379044
WNT5A	− 0.755711835	0.342292419	− 2.207796004	0.027258497
IGF1R	− 0.527239042	0.244654459	− 2.155035493	0.031159057
ITGB6	0.799967093	0.372285529	2.148799863	0.031650268
TOP2A	0.701281843	0.338198902	2.073578118	0.038118509
MTOR	− 0.199119009	0.10038674	− 1.983519036	0.047309483
ANLN	0.557748652	0.285064381	1.956570829	0.050397947
WNT5B	− 0.431331553	0.223920978	− 1.926266833	0.054071065
WDR77	− 0.226717873	0.11857368	− 1.912042136	0.05587079
EGFR	− 0.608467987	0.32021188	− 1.900204288	0.057406316
MKI67	0.504882032	0.270587483	1.865873567	0.062059057
UBE2C	0.605437564	0.324599078	1.865185719	0.062155377

Table 2. Top 20 differentially expressed genes (log2 fold change) between recurrence and non-recurrence groups ranked by *P* value**.** The full list of results is provided in the supplementary material.

Gene expression analysis

We analyzed DEGs in the recurrent versus non-recurrent (NR) groups using a NanoString codeset. The DEGs significantly upregulated in the recurrent population included *EZH2* (log2fold change 0.67, $p = 0.0017$), *WNT11* (log2 fold change 1.08, $p = 0.0088$), *ITGB6* (log2fold change 0.80, $p = 0.0312$), and *TOP2A* (log2fold change 0.79, $p = 0.0381$). The significantly downregulated DEGs in the recurrent population included the following: *SNAI2* (log2 fold change -0.63, $p = 0.0055$), *ITPR1* (log2fold change − 0.75, $p = 0.0083$), *MME* (log2 fold change − 0.70, $p = 0.0092$), *PTEN* (log2fold change − 0.29, $p = 0.0163$), *VDR* (log2fold change − 0.46, $p = 0.0184$), *SYTL4* (log2 fold change − 0.57, $p = 0.0219$), *MDM2* (log2fold change − 0.33, $p = 0.0234$), *WNT5A* (log2fold change − 0.76, $p = 0.0272$), *IGF1R* (log2fold change − 0.53, $p = 0.0311$) and *MTOR* (log2 fold change − 0.20, $p = 0.047$). A list of DEGs is listed in Table 2 and a volcano plot is included as Fig. 1. *EZH2* expression was the most significant discriminator between the recurrent and non-recurrent groups with an area under the receiver operating characteristic (ROC) curve (AUC) 0.74, $p = 0.0042$ (Table 3). Given the association between the proliferative index Ki-67 recurrence, we next performed a correlation analysis with the proliferation score (created by NanoString based on proliferation genes) for each gene included in our codeset. The correlation analysis is included below in Table 4. Seventy-four genes were significantly correlated with the proliferation score. Of the genes significantly upregulated in the recurrence group, *EZH2* and *TOP2A* also showed positive significant correlation with proliferation.

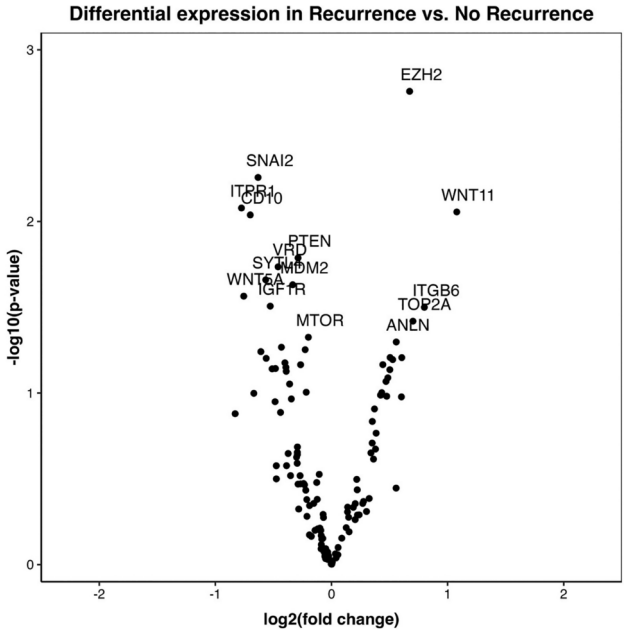

Fig. 1. Volcano plot for differentially expressed genes (log2 fold change) between recurrence and non-recurrence groups.

As expected, DEGs were identified across molecular subtypes as well (Table 5). DEGs with the most significant upregulation in luminal B specimens versus luminal A specimens included *PTTG1* (log2fold change 1.33, $p < 0.001$), *CCNB1* (log2 fold change 1.23, $p < 0.0001$), *MELK* (log2 fold change 1.39, $p < 0.0001$), *CDCA8* (log2 fold change 1.30, $p < 0.0001$), *KIF2C* (log2 fold change 1.36, $p < 0.0001$), *UBE2C* (log2 fold change 1.62, $p < 0.0001$), *TOP2A* (log2 fold change 1.69, $p < 0.0001$), *KNTC2* (log2 fold change 1.15, $p < 0.0001$), and *CDC20* (log2 fold change 1.20, $p < 0.0001$). Of those, *TOP2A* is not in the PAM50 molecular subtyping gene expression list. Downregulation in luminal B specimens compared to luminal A identified in *SFRP1* (log2 fold change − 2.60, $p < 0.0001$), *IGF1* (log2 fold change − 2.42, $p < 0.0001$), *FOXC1* (log2 fold change − 1.60, $p < 0.0001$), *EGFR* (log2 fold change − 1.46, $p < 0.0001$), and *ALDH1A1* (log2 fold change − 1.58 $p < 0.0001$). Of those, *IGF1* is not included in PAM50. This study identified gene expression differences associated with luminal B subtype, the more aggressive of the two subtypes.

We also performed a correlation analysis among all the genes (Fig. 2). The clustering patterns in the heatmap reveal significant co-expression among genes associated with cell cycle regulation (e.g., *CCNB1, UBE2C, CDCA8*), DNA replication and repair (e.g., *TOP2A, TYMS*), and epigenetic regulation (e.g., *EZH2*). These findings suggest that these genes are part of oncogenic networks driving breast cancer progression.

Discussion

This institutional study evaluated gene expression associated with ER + BC recurrence when treated optimally with endocrine therapy, indicative of endocrine therapy resistance. In doing so, we examined potential biomarkers of endocrine therapy resistance for drug development and clinical trial testing. There are currently no biomarkers of endocrine therapy resistance used in the upfront setting. Clinically, endocrine therapy resistance is only assessed in the recurrent or metastatic setting once endocrine therapy failure is suspected. Estrogen receptor gene ESR1 mutations are the most widely recognized proponent of endocrine therapy resistance[9,11]. ESR1 mutations however are developed over time and while very useful in the recurrent or metastatic setting, cannot be utilized in the upfront setting to guide therapy[12]. A comparable genetic aberration has not been identified in the upfront setting. This underscores the need for further understanding of drivers of endocrine therapy resistance prior to the development of recurrence to guide therapeutic decision-making.

Our study identified clinical characteristics associated with endocrine therapy resistance. The only clinical variable with statistically significant association with recurrence was the proliferation index Ki-67. Indeed, Ki-67 has been used as a biomarker of response in neoadjuvant endocrine therapy studies. In these studies, change in Ki-67 has been associated with response to endocrine therapy[13,14]. In the prospective randomized perioperative endocrine therapy in postmenopausal women with hormone-sensitive early breast cancer (POETIC) trial, Ki-67 was found to be associated with poor prognosis despite neoadjuvant and adjuvant endocrine therapy administration[15], further asserting that Ki-67 may be a biomarker candidate for endocrine therapy resistance and signaling the need for additional therapies. Therapeutic trials investigating the utility of endocrine therapy plus additional therapies such as the FELINE trial also used Ki-67 as a biomarker of therapy resistance[16]. In that study, significantly higher Ki-67 changes were seen across the treatment arms. The Ki-67 protein is present in all phases of the cell cycle; this may offer insights into which genes are upregulated in endocrine resistant tumors. Indeed, the gene *MIK67* is included among the genes comprising the proliferation score published by

Gene	AUC	*P* value
EZH2	0.74	0.0042
ANLN	0.73	0.0070
RRM2	0.69	0.0243
CDCA1	0.69	0.0263
CENPF	0.69	0.0251
CDCA8	0.69	0.0278
CCNB1	0.68	0.0318
CEP55	0.68	0.0340
SNAI2	0.68	0.0363
UBE2T	0.68	0.0363
MKI67	0.67	0.0439
MELK	0.67	0.0467
UBE2C	0.67	0.0497
CCNE1	0.66	0.0627
CD10	0.66	0.0682
GRB7	0.65	0.0763
TMEM45B	0.65	0.0772
KNTC2	0.65	0.0816
ITPR1	0.65	0.0863
TOP2A	0.65	0.0911

Table 3. Area under the receiver operating characteristic (ROC) curve (AUC) analysis of top 20 genes distinguishing recurrence versus non-recurrence status ranked by AUC**.** The full list of results is provided in the supplementary material.

NanoString. Genes significantly associated with recurrence and correlating with the proliferation score included *EZH2* and *TOP2A*.

Both *EZH2* and *TOP2A* are involved in the cell cycle. *EZH2* is postulated to be a regulator of the cell cycle[17]. Recent studies on *EZH2* have identified it as central to the epigenome through methylation of various transcription factors[18]. Interestingly, *EZH2* acts through the PTEN-MTOR pathway[17]; downregulation of these genes was associated with recurrence in our cohort. As demonstrated by Chen et.al[17], mechanisms by which high *EZH2* expression may propel ET resistance include modulation of the PI3K pathway, which has been implicated in ET resistance[19]. As suggested by Chen et.al., targeting both these may result in overcoming ET resistance. There are several *EZH2* inhibitors currently in Phase 1–2 clinical trials. *EZH2* also has effects on epithelial-mesenchymal transition and promotes the cell cycle. Indeed, heatmap analysis from our own work indicates increased co-expression of cell cycle pathway genes. Pre-clinical studies show that knockdown of *EZH2* also decreases cyclin D1 expression[17].

Also involved with transcription and DNA replication is the topoisomerase *TOP2A*. Because of this, its function is targeted by anthracycline based chemotherapy[20–22]. However, high expression has also been associated with endocrine therapy, specifically tamoxifen, resistance[23]. As identified by Chen et al., *TOP2A* expression is highly correlated with poor prognosis in breast cancer[24]. Similarly, as demonstrated by Ogino et al., *TOP2A* expression correlated with recurrence even in the absence of other clinicopathologic characteristics that would portend worse prognosis[25]. In our present study, *TOP2A* high expression was associated with worse prognosis. *TOP2A* expression has been demonstrated to be positively correlated with non-luminal A subtypes[26]. *TOP2A* expression is also independently associated with higher 21-gene recurrence scores in ER + patients and may be used to guide adjuvant therapy[27,28].

While targeting of *EZH2* and *TOP2A* is not currently clinically used outside of clinical trials, targeting of the cell cycle pathway has been the focus of the most recent developments in the treatment of ER + BC. Cyclin-dependent kinases 4/6 (CDK 4/6) inhibitors (e.g. ribociclib, abemaciclib, palbociclib) have shown benefit for locally advanced and metastatic ER + BC. For example, abemaciclib when combined with endocrine therapy is the first CDK 4/6 inhibitor to demonstrate a significant improvement in invasive disease-free survival (IDFS) in women with ER + , HER2 negative, node positive breast cancer at high risk of early recurrence[25]. However, it is currently unclear which patients will derive the most benefit. While CDK 4/6 inhibitors have promise, many women do not initially respond and, ultimately, most women develop resistance[12]. Several of the DEGs identified in our cohort, such as *EZH2*, interact directly with the cell cycle pathway and have potential to be used in combination with CDK 4/6 inhibitors to improve their efficacy.

Additional DEGs upregulated in the recurrent population included *WNT11* and *ITGB6*. *WNT* proteins contribute to both proliferation and metastasis as well as endocrine therapy resistance. *WNT11* participates in canonical and non-canonical signaling and is found downstream of estrogen receptor α signaling[29]. Our data demonstrates *WNT11* high expression associated with higher recurrence. Other studies have demonstrated worse outcomes associated with high expression of *WNT11*. As demonstrated by Menck et al., activation of *ROR2*, an alternative *WNT* receptor, was high in aggressive tumors displaying BRCA-like morphology and

Gene	Correlation	*P* value
CDCA8	0.779212413	6.64E−18
CCNB1	0.773921867	1.53E−17
TOP2A	0.770350826	2.65E−17
MKI67	0.765848065	5.24E−17
RRM2	0.764447002	6.45E−17
UBE2C	0.764353941	6.54E−17
PTTG1	0.747094474	7.63E−16
KIF2C	0.743722967	1.20E−15
TYMS	0.742170675	1.48E−15
MELK	0.73662885	3.08E−15
BIRC5	0.733229816	4.78E−15
CEP55	0.721798751	1.99E−14
CDC20	0.713438848	5.42E−14
KNTC2	0.707088872	1.13E−13
CDC6	0.702424934	1.92E−13
ANLN	0.687791216	9.47E−13
CENPF	0.672224212	4.68E−12
EXO1	0.66280183	1.18E−11
UBE2T	0.650349358	3.78E−11
ORC6L	0.632033222	1.91E−10
MYBL2	0.620359918	5.10E−10
CDCA1	0.60500048	1.74E−09
EZH2	0.601771708	2.24E−09
BRCA1	0.581745819	9.94E−09
CXXC5	0.503619875	1.41E−06
GSK3B	0.473946509	6.87E−06
CCNE1	0.472017394	7.58E−06
MMP11	0.471156781	7.91E−06
CCND1	0.470922033	8.01E−06
CDH1	0.439321094	3.64E−05
EpCAM	0.405567314	0.000156835
ESR1	0.399614254	0.000199749
FOXA1	0.376585369	0.000488323
BLVRA	0.34246372	0.001636368
CK.8.18	0.338300037	0.001879688
CLDN3 (claudin 3)	0.337345467	0.001939873
CLDN7 (claudin 7)	0.33015598	0.002451874
SLC39A6	0.304413806	0.005426111
GRB7	0.293206687	0.007509381
BCL2	0.277622495	0.011562105
WNT4	0.256995716	0.019765893
FGFR3	0.254964332	0.020794082
ACTR3B	0.23614595	0.032689911
MDM2	0.217814935	0.049322955
ERBB2	0.216011197	0.051281919
MAPT	0.215612781	0.051723127
CLDN4 (claudin 4)	0.200854461	0.070392094
WNT7B	0.199059245	0.072993078
SQLE	0.167393717	0.132794695
PIK3R3	0.165641069	0.136959075
TFF1	0.145974468	0.190675348
BAG1	0.141118311	0.206014049
SYTL4	0.125657771	0.260641452
LEF1	0.122468309	0.273032992
PSPHL	0.116799833	0.296017383
BCL2L1	0.109988994	0.325267551
Continued		

Gene	Correlation	*P* value
GPR160	0.103921339	0.352827026
STAT1	0.100620527	0.368410265
IFT140	0.093500694	0.403425521
TMEM45B	0.084839463	0.448554184
CD24	0.055832505	0.618334598
MLPH	0.054744129	0.625206102
MAPK10	0.035368253	0.752416168
PIK3R2	0.033390968	0.765848792
TP53	0.026793101	0.811152099
VRD	0.017273614	0.877586589
PHGDH	0.013281968	0.905725288
FGFR4	0.011273749	0.919928283
CRYBB2	− 0.004153213	0.970459711
PLCB1	− 0.011508003	0.918270158
NAT1	− 0.011781239	0.916336548
MTOR	− 0.017314964	0.877295829
IGF1R	− 0.034290807	0.75972693
PIK3CA	− 0.037377838	0.738839095
GPC4	− 0.042647094	0.703626004
ITGB6	− 0.045614611	0.684060457
PGR	− 0.054114736	0.629194477
CDH3	− 0.057018358	0.610884788
EGF	− 0.061705162	0.581831722
CREBBP	− 0.084074502	0.452670601
APOE	− 0.093838936	0.40171909
CD44	− 0.095898028	0.391422963
WDR77	− 0.099605893	0.373283552
MIA	− 0.099682834	0.372912639
ITPR1	− 0.104627535	0.349546977
BORCS7	− 0.113655973	0.309297385
SP110	− 0.121748758	0.275882314
WNT5A	− 0.12324229	0.269990252
HMGA2	− 0.13240452	0.235702707
EGFR	− 0.147662287	0.185542411
MUC1	− 0.153306567	0.169106151
TP.63	− 0.156074734	0.1614484
KRT5	− 0.167528491	0.132478492
KRT14	− 0.177529524	0.110573404
WNT2	− 0.193487627	0.081559804
SNAI1	− 0.194864019	0.079372497
KRT17	− 0.196106589	0.077438285
WNT11	− 0.213896798	0.053659172
CD27	− 0.222101647	0.044913553
WNT10A	− 0.234914541	0.033635869
CD29	− 0.24125079	0.029003355
PIK3R1	− 0.252996515	0.021833369
WNT6	− 0.283350732	0.009892677
PIK3CG	− 0.287048039	0.008930632
WNT5B	− 0.288309215	0.008621764
CD8A	− 0.326122427	0.002789486
PTEN	− 0.345150151	0.00149486
MYC	− 0.37216891	0.000575397
CD49	− 0.372674579	0.000564754
PIK3CD	− 0.403042148	0.000173876
THY1	− 0.404109345	0.000166476
PIK3R5	− 0.407359329	0.000145692
Continued		

Gene	Correlation	*P* value
SFRP1	− 0.437049905	4.04E−05
TWIST1	− 0.536793832	2.01E−07
AKT3	− 0.55870238	4.91E−08
FOXC1	− 0.561970455	3.94E−08
IGF1	− 0.573443683	1.79E−08
CD10	− 0.580079178	1.12E−08
TWIST2	− 0.586473834	7.06E−09
HGF	− 0.59162025	4.83E−09
SNAI2	− 0.606832734	1.51E−09
SFRP2	− 0.607831005	1.40E−09
PRKCA	− 0.630304239	2.22E−10
ZEB2	− 0.632133332	1.90E−10
VIM	− 0.639380628	1.01E−10
ZEB1	− 0.659206617	1.66E−11
CCND2	− 0.661548668	1.33E−11
ALDH1A1	− 0.662491142	1.21E−11

Table 4. Pearson's correlation between each gene and proliferation score. Seventy-four genes were significantly correlated with the proliferation score. Of the genes significantly upregulated in the recurrence group, *EZH2* and *TOP2A* also showed positive significant correlation with proliferation. *WNT11* is highlighted as the most upregulated DEG in the recurrence group with no significant correlation seen with the proliferation score, indicative of an independent process.

Gene	log2 fold change	Standard error (log2 fold change)	Linear fold change	*P* value
SFRP1	− 2.602804016	0.272356995	− 9.55658958	1.22E−21
IGF1	− 2.41858579	0.266520174	− 9.074681858	1.14E−19
FOXC1	− 1.598321763	0.177217831	− 9.01896697	1.90E−19
PTTG1	1.332699618	0.153700815	8.670738801	4.29E−18
EGFR	− 1.461561517	0.173687307	− 8.414901153	3.93E−17
CCNB1	1.226817148	0.14817301	8.279626286	1.24E−16
MELK	1.393930667	0.171318428	8.136489936	4.07E−16
CDCA8	1.297511265	0.160352822	8.091602329	5.89E−16
KIF2C	1.3637245	0.175398265	7.775017055	7.54E−15
UBE2C	1.619548221	0.210037004	7.71077566	1.25E−14
ALDH1A1	− 1.575617077	0.205863689	− 7.653691055	1.95E−14
TOP2A	1.694887784	0.222322791	7.623544921	2.47E−14
KNTC2	1.145905063	0.151080407	7.584736422	3.33E−14
CDC20	1.197871711	0.158620538	7.551807158	4.29E−14
CEP55	1.233521387	0.163818958	7.5297841	5.08E−14
RRM2	1.405638309	0.187541493	7.495079008	6.63E−14
BIRC5	1.6079831	0.21722356	7.402434149	1.34E−13
TWIST2	− 1.457440043	0.197184995	− 7.391232003	1.45E−13
MKI67	1.28204435	0.175421019	7.30838502	2.70E−13
MYBL2	1.557696393	0.222340722	7.005897885	2.45E−12

Table 5. Top 20 differentially expressed genes (log2 fold change) between molecular subtype luminal B versus luminal A groups ranked by *P* value. The full list of results is provided in the supplementary material. Gene expression in the luminal A versus luminal B population. Statistically significant genes are highlighted in yellow. Upregulated genes in the luminal B population include *PTTG1, CCNB1, MELK, CDCA8, KIF2C, UBE2C, TOP2A*, *KNTC2*, and *CDC20*. Downregulated genes include *SFRP1, IGF1, FOXC1, EGFR*, and *ALDH1A1*. In addition to genes included in the PAM50 molecular assay, *TOP2A* and *IGF1* were identified as genes associated with molecular classification.

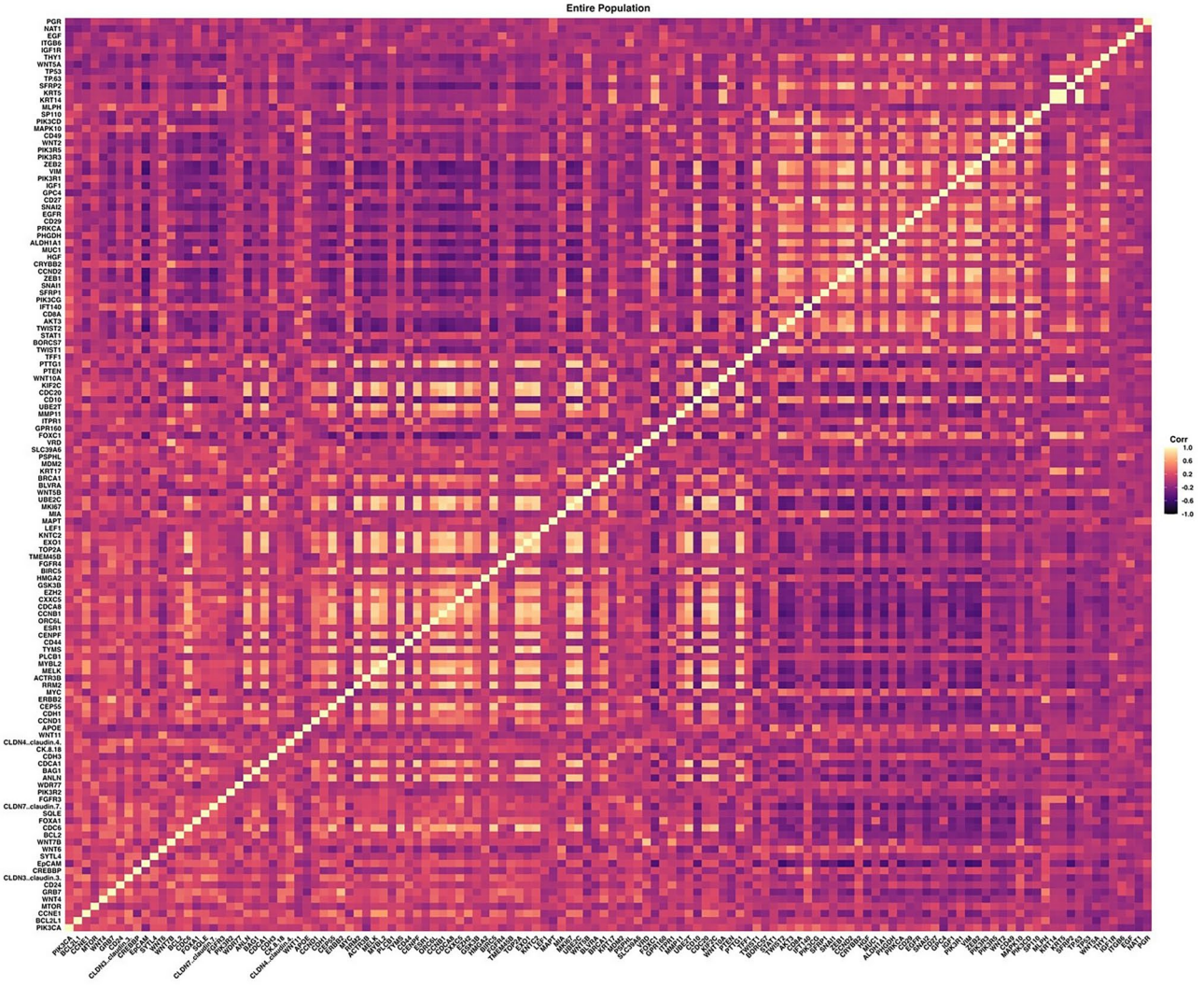

Fig. 2. Heatmap of Pearson's correlation among 128 genes for all patients.

increased *WNT11*, which in turn increased tumor invasion and brain metastases[30]. Similarly, *WNT11* was a key player in estradiol related signaling in ER + BCs with its expression contributing to apoptosis-resistance[31]. The paucity of additional *WNT* expression in the DEG list is likely due to *WNT* signaling predominating in triple negative subtypes, which were not included in this investigation. The role of *ITGB6* in ER + BC has yet to be fully elucidated. There is indication that it mediates aggressive biology and therapy resistance in other subtypes such as HER2-enriched tumors[32]. Its effect may be mediated through modulation of the Rho-Rac pathway, which mediates actin assembly. While more investigation is needed to elucidate the role of *ITGB6* in breast cancer progression, *WNT11* may represent a targetable biomarker associated with endocrine therapy resistance.

Conversely, *WNT5A* downregulation was associated with recurrence in our cohort. This has been identified in other investigations of *WNT5A* demonstrating *WNT5A* as a tumor suppressor gene[33]. Additional down regulated genes in the recurrence group included *SNAI2*, *ITPR1*, *MME*, *PTEN*, *VDR*, *SYTL4*, *MDM2*, *IGF1*, and *MTOR*. In literature, these genes have variable impact on breast cancer prognosis. Several of them interact in known pathways associated with aggressive biology, such as *MTOR* and *IGF1* in the *PI3K*/*AKT*/*MTOR* pathway[34]. Downregulation of *PTEN* is associated with increased *PI3K* activity and loss of *PTEN* associated with aggressive biology[35]. *VDR* expression is protective in other cancer types though its role in breast cancer has yet to be fully elucidated[36]. *IGF1* allows breast cancer cells to proliferate and is linked to estrogen signaling[37]. In one study evaluating expression of *IGF1* in relation to tamoxifen resistance, higher *IGF1* expression was associated with increased time to development of endocrine resistance[38]. One possible mechanism by which this may occur is via *IGF1* activation of PR and ER expression[39]. *IGF1* increases ERα expression and increases ERα activity[40]. While uninhibited expression of *IGF1R* can lead to proliferation via the ER pathway, it also induces greater receptivity to estrogen blocking agents. In high ER and IGF1 expressors, inhibition of both may lead to prolonged endocrine sensitivity and decreased relapse. The role of these genes in tumorigenesis is complex and likely involves multiple pathways.

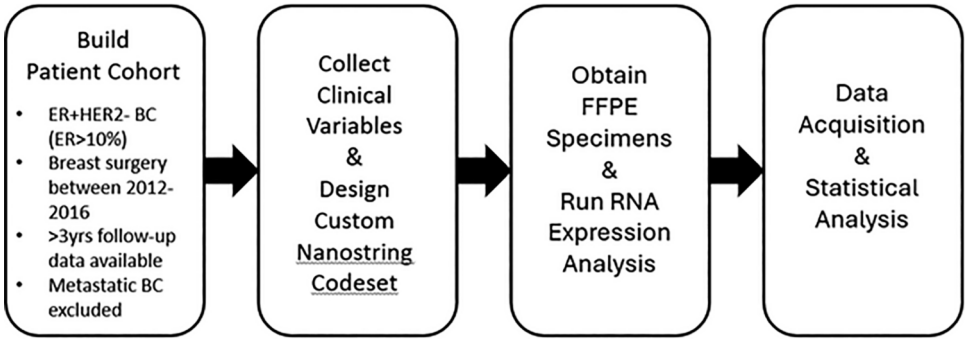

Fig. 3. Study design flowchart.

The strength of this study lies in the robust details on therapy administration and recurrence, clinical data not widely available in public cohorts. This allows elucidation of gene expression associated with therapy resistance. The retrospective analysis allowed examination of long-term outcomes, such as recurrence and metastasis, in the setting of standard of care. This is the gold standard, especially with recent data showing that pathologic response may be a less reliable outcome measure to predict efficacy of treatment and long-term outcomes in ER + BC[41]. However, a limitation of this study is its retrospective nature and smaller sample size. Another limitation is the low representation of Black/African American patients in the study. ER + BC portends a two-fold higher mortality in Black/African American women compared to other races[42,43]. Larger studies must include representation of these patients to fully understand the landscape of aggressive luminal breast cancer. More research is needed to validate gene lists appropriate for certain populations.

Prevention of breast cancer recurrence is essential to improve overall outcomes since metastatic disease is almost always fatal. Given the de novo resistance to endocrine therapy, there is an urgent need to identify targetable biomarkers in the upfront setting to prevent relapse and improve survival. Our study demonstrates key DEGs associated with endocrine therapy resistance and worse survival. Further studies are necessary to elucidate gene–gene correlation and pathway regulation, as well as to identify the clinical utility of targeting these genes to prevent recurrence and improve outcomes.

Methods

Study design and sample selection

Surgical pathology archives were queried for HER2- ER + BC formalin-fixed paraffin-embedded (FFPE) specimens resected during the study period. All breast cancer patients treated with surgery at City of Hope Comprehensive Cancer Center between the years 2012–2016 with documented written informed consent for use of specimens for research [(COH Institutional Review Board (IRB) 07047] were queried for eligibility. These were serially requested, beginning with year 2016. Because tumors with less than 10% ER positivity behave like triple-negative diseases, a cutoff of 10% was chosen[12]. Eligibility criteria included clinical tumor size of at least 2 cm or clinical tumor size of less than 2 cm with biopsy-proven axillary nodal metastasis. Eligibility criteria also included documented follow up of at least 3 years. Patients with metastatic breast cancer were excluded. Specimens were selected and analyzed based on eligibility criteria outlined in Fig. 3. All patients included were treated with endocrine therapy, including selective estrogen receptor modulators (SERMs), selective estrogen receptor degraders (SERDs), and aromatase inhibitors (AI) following resection of the primary tumor. Recurrence was only considered if occurring at least three months after initial treatment; this cutoff was chosen to exclude persistent and potentially de novo metastatic disease[13]. All recurrences, including locoregional and distant were included. Ethical approval was obtained for this study from the COH Institutional Review Board (IRB 18423) in accordance with the Declaration of Helsinki.

Clinical, tumor, and patient characteristics

Clinical variables were collected including patient demographics (race/ethnicity, age, BMI), clinical staging, type of breast surgery, type of axillary surgery, surgical pathology (size of tumor, number of positive lymph nodes, total number of lymph nodes retrieved), type and timing of adjuvant therapies including systemic chemotherapy, endocrine therapy, and radiation treatment, timing and location of recurrence, and survival.

NanoString codeset design

A custom NanoString nCounter panel (NanoString Technologies, WA) including 145 unique genes (128 endogenous and 17 housekeeping genes) was designed to include known ER + BC pathways as well as genes associated with race-related outcomes including *PSPHL*, *SQLE*, *CRYBB2*, and *MUC1*[44–47] (Table S4). Pathways included apoptosis, basal, cell cycle, epithelial to mesenchymal transition (EMT), ER, immune response, metabolism, PI3K, proliferation, *WNT*, and stem cell pathways[48–51]. PAM50 genes were included for molecular subtype determination, and subtype analysis was performed by NanoString and provided[52]. Risk of recurrence scores were not generated.

Gene expression date preprocessing

Once eligible specimens were identified by the pathologist on study (D. Schmolze), tumor blocks were sectioned. RNA was extracted using miRNeasy FFPE kit (Qiagen). RNA concentration was assessed with the Nanodrop spectrophotometer ND-1000 and Qubit 3.0 Fluorometer (Thermo Scientific, CA). RNA fragmentation and quality control were determined by 2100 Bioanalyzer (Agilent, CA). Gene expression was analyzed by the NanoString nCounter platform (NanoString Technologies, WA). RNA was hybridized with codeset at 65 °C for 16 h. Samples were processed on the NanoString Prep Station and the target-probe complex was immobilized onto the analysis cartridge. Cartridges were scanned by the nCounter Digital Analyzer for digital counting of molecular barcodes corresponding to each target at 280 fields of view.

Gene expression data were processed using nSolver Analysis Software v3.0 (NanoString Technologies, Seattle, WA) with the Advanced Analysis Module, which facilitates quality control, normalization, Pathview plotting, and immune cell profiling. Data normalization involved subtracting the mean plus one standard deviation of eight negative controls, while technical variation was addressed using six internal positive controls. Input volume corrections were performed using eight internal housekeeping genes through the geNorm algorithm. Additionally, bioinformatics analyses, including pathway enrichment, were performed using the Ingenuity Pathway Analysis (IPA) tool to provide deeper biological insights (version 73620684, http://www.ingenuity.com).

Statistical analysis

Patient demographic and clinical characteristics, stratified by recurrence and non-recurrence status, were summarized in Table 1. Group comparisons were conducted using t-tests for continuous variables and Chi-squared tests for categorical variables to assess baseline differences. The primary objective of this study was to identify DEGs associated with recurrence in ER + breast cancer patients treated with optimal endocrine therapy. Differential expression (DE) analyses were performed using negative-binomial regression models with a log2 link function, yielding fold-change estimates in gene expression between recurrence and non-recurrence groups. To address the secondary objective—evaluating gene expression differences across PAM50 molecular subtypes (luminal A vs. luminal B)—we applied the same DE analysis framework. Results of DE analyses were visualized with volcano plots, and the complete list of DEGs, ranked by statistical significance, is provided in supplementary tables. ROC curve analyses were performed to evaluate the discriminatory power of individual genes in distinguishing recurrence from non-recurrence, with AUC values and significance assessed via Wilcoxon rank-sum tests. Furthermore, Pearson's correlation analyses were utilized to investigate associations between gene expression and proliferation scores, as well as pairwise correlations among genes. All analyses relied on normalized gene expression counts derived from the NanoString platform and nSolver Analysis Software v3.0, ensuring robust and reproducible results. Statistical significance was set at p value < 0.05, and all analyses were conducted using R software (version 4.4.1), adhering to rigorous scientific standards.

Ethical approval and consent to participate

Ethical approval was obtained for this study from the COH Institutional Review Board (IRB 18423) in accordance with the Declaration of Helsinki; written informed consent was obtained from participants.

Data availability

The nucleic acid sequencing datasets generated or analyzed during the current study in the institutional cohort are publicly available on the National Center for Biotechnology Information Gene Expression Omnibus (NCBI GEO) biorepository.

Received: 31 October 2023; Accepted: 4 February 2025
Published online: 28 February 2025

References

1. Zhang, X. H., Giuliano, M., Trivedi, M. V., Schiff, R. & Osborne, C. K. Metastasis dormancy in estrogen receptor-positive breast cancer. *Clin. Cancer Res.* **19**(23), 6389–6397 (2013).
2. Dent, R. et al. Factors associated with breast cancer mortality after local recurrence. *Curr. Oncol.* **21**(3), e418–e425 (2014).
3. Anderson, S. J. et al. Prognosis after ipsilateral breast tumor recurrence and locoregional recurrences in patients treated by breast-conserving therapy in five National Surgical Adjuvant Breast and Bowel Project protocols of node-negative breast cancer. *J. Clin. Oncol.* **27**(15), 2466–2473 (2009).
4. Sopik, V., Sun, P. & Narod, S. A. Predictors of time to death after distant recurrence in breast cancer patients. *Breast Cancer Res. Treat.* **173**(2), 465–474 (2019).
5. Sparano, J. A. et al. Adjuvant chemotherapy guided by a 21-gene expression assay in breast cancer. *N. Engl. J. Med.* **379**(2), 111–121 (2018).
6. Kalinsky, K. et al. 21-gene assay to inform chemotherapy benefit in node-positive breast cancer. *N. Engl. J. Med.* **385**(25), 2336–2347 (2021).
7. Soliman, H. et al. MammaPrint guides treatment decisions in breast cancer: Results of the IMPACt trial. *BMC Cancer* **20**(1), 81 (2020).
8. Noordhoek, I. et al. Breast cancer index predicts extended endocrine benefit to individualize selection of patients with HR(+) early-stage breast cancer for 10 years of endocrine therapy. *Clin. Cancer Res.* **27**(1), 311–319 (2021).
9. Lei, J. T., Anurag, M., Haricharan, S., Gou, X. & Ellis, M. J. Endocrine therapy resistance: New insights. *Breast* **48**(Suppl 1), S26-s30 (2019).
10. Krauss, K. & Stickeler, E. Endocrine therapy in early breast cancer. *Breast Care* **15**(4), 337–346 (2020).
11. Hartkopf, A. D., Grischke, E. M. & Brucker, S. Y. Endocrine-resistant breast cancer: Mechanisms and treatment. *Breast Care* **15**(4), 347–354 (2020).
12. Reinert, T. et al. ESR1 mutations are not a common mechanism of endocrine resistance in patients with estrogen receptor-positive breast cancer treated with neoadjuvant aromatase inhibitor therapy. *Front. Oncol.* **10**, 342 (2020).

13. Dowsett, M. et al. Short-term changes in Ki-67 during neoadjuvant treatment of primary breast cancer with anastrozole or tamoxifen alone or combined correlate with recurrence-free survival. *Clin. Cancer Res.* **11**(2 Pt 2), 951s-s958 (2005).
14. Mosly, D., Turnbull, A., Sims, A., Ward, C. & Langdon, S. Predictive markers of endocrine response in breast cancer. *World J. Exp. Med.* **8**(1), 1–7 (2018).
15. Smith, I. et al. Long-term outcome and prognostic value of Ki67 after perioperative endocrine therapy in postmenopausal women with hormone-sensitive early breast cancer (POETIC): An open-label, multicentre, parallel-group, randomised, phase 3 trial. *Lancet Oncol.* **21**(11), 1443–1454 (2020).
16. Khan, Q. J. et al. Letrozole + ribociclib versus letrozole + placebo as neoadjuvant therapy for ER+ breast cancer (FELINE trial). *J. Clin. Oncol.* **38**(15_suppl), 505–505 (2020).
17. Chen, Y., Zhu, H., Luo, Y., Tong, S. & Liu, Y. EZH2: The roles in targeted therapy and mechanisms of resistance in breast cancer. *Biomed. Pharmacother.* **175**, 116624 (2024).
18. Liu, Y. & Yang, Q. The roles of EZH2 in cancer and its inhibitors. *Med. Oncol.* **40**(6), 167 (2023).
19. Nunnery, S. E. & Mayer, I. A. Targeting the PI3K/AKT/mTOR pathway in hormone-positive breast cancer. *Drugs* **80**(16), 1685–1697 (2020).
20. Gerashchenko, T. S. et al. Different morphological structures of breast tumors demonstrate individual drug resistance gene expression profiles. *Exp. Oncol.* **40**(3), 228–234 (2018).
21. Litviakov, N. et al. Expression of M2 macrophage markers YKL-39 and CCL18 in breast cancer is associated with the effect of neoadjuvant chemotherapy. *Cancer Chemother. Pharmacol.* **82**(1), 99–109 (2018).
22. Zhu, J., Muskhelishvili, L., Tong, W., Borlak, J. & Chen, M. Cancer genomics predicts disease relapse and therapeutic response to neoadjuvant chemotherapy of hormone sensitive breast cancers. *Sci. Rep.* **10**(1), 8188 (2020).
23. Guan, Q. et al. Identification of tamoxifen-resistant breast cancer cell lines and drug response signature. *Front. Mol. Biosci.* **7**, 564005 (2020).
24. Chen, G. et al. Identification of candidate biomarkers correlated with poor prognosis of breast cancer based on bioinformatics analysis. *Bioengineered* **12**(1), 5149–5161 (2021).
25. Ogino, M. et al. Implications of topoisomerase (TOP1 and TOP2α) expression in patients with breast cancer. *In Vivo* **34**(6), 3483–3487 (2020).
26. Shigematsu, H. et al. Overexpression of topoisomerase II alpha protein is a factor for poor prognosis in patients with luminal B breast cancer. *Oncotarget* **9**(42), 26701–26710 (2018).
27. Zhu, L. et al. Clinical analysis of 21-gene recurrence score test in hormone receptor-positive early-stage breast cancer. *Oncol Lett* **17**(6), 5469–5480 (2019).
28. Sparano, J. A., Goldstein, L. J., Davidson, N. E., Sledge, G. W. Jr. & Gray, R. TOP2A RNA expression and recurrence in estrogen receptor-positive breast cancer. *Breast Cancer Res. Treat.* **134**(2), 751–757 (2012).
29. Xu, X., Zhang, M., Xu, F. & Jiang, S. Wnt signaling in breast cancer: biological mechanisms, challenges and opportunities. *Mol Cancer* **19**(1), 165 (2020).
30. Menck, K. et al. WNT11/ROR2 signaling is associated with tumor invasion and poor survival in breast cancer. *J. Exp. Clin. Cancer Res.* **40**(1), 395 (2021).
31. Lin, Z., Reierstad, S., Huang, C. C. & Bulun, S. E. Novel estrogen receptor-alpha binding sites and estradiol target genes identified by chromatin immunoprecipitation cloning in breast cancer. *Cancer Res.* **67**(10), 5017–5024 (2007).
32. Sabapathi, N. et al. Prognostic significance of FOXC1 in various cancers: A systematic review and meta-analysis. *Mol. Diagn. Ther.* **23**(6), 695–706 (2019).
33. Jönsson, M., Dejmek, J., Bendahl, P. O. & Andersson, T. Loss of Wnt-5a protein is associated with early relapse in invasive ductal breast carcinomas. *Cancer Res.* **62**(2), 409–416 (2002).
34. Chalhoub, N. & Baker, S. J. PTEN and the PI3-kinase pathway in cancer. *Annu. Rev. Pathol.* **4**, 127–150 (2009).
35. Hollestelle, A., Elstrodt, F., Nagel, J. H., Kallemeijn, W. W. & Schutte, M. Phosphatidylinositol-3-OH kinase or RAS pathway mutations in human breast cancer cell lines. *Mol. Cancer Res.* **5**(2), 195–201 (2007).
36. Feldman, D., Krishnan, A. V., Swami, S., Giovannucci, E. & Feldman, B. J. The role of vitamin D in reducing cancer risk and progression. *Nat. Rev. Cancer* **14**(5), 342–357 (2014).
37. Fagan, D. H. & Yee, D. Crosstalk between IGF1R and estrogen receptor signaling in breast cancer. *J. Mammary Gland Biol. Neoplasia* **13**(4), 423–429 (2008).
38. Chong, K., Subramanian, A., Sharma, A. & Mokbel, K. Measuring IGF-1, ER-α and EGFR expression can predict tamoxifen-resistance in ER-positive breast cancer. *Anticancer Res.* **31**(1), 23–32 (2011).
39. Katzenellenbogen, B. S. Estrogen receptors: Bioactivities and interactions with cell signaling pathways. *Biol. Reprod.* **54**(2), 287–293 (1996).
40. Bartella, V., De Marco, P., Malaguarnera, R., Belfiore, A. & Maggiolini, M. New advances on the functional cross-talk between insulin-like growth factor-I and estrogen signaling in cancer. *Cell. Signal* **24**(8), 1515–1521 (2012).
41. Conforti, F. et al. Evaluation of pathological complete response as surrogate endpoint in neoadjuvant randomised clinical trials of early stage breast cancer: Systematic review and meta-analysis. *BMJ* **375**, e066381 (2021).
42. Rauscher, G. H. et al. Racial disparity in survival from estrogen and progesterone receptor-positive breast cancer: Implications for reducing breast cancer mortality disparities. *Breast Cancer Res. Treat.* **163**(2), 321–330 (2017).
43. Vidal, G., Bursac, Z., Miranda-Carboni, G., White-Means, S. & Starlard-Davenport, A. Racial disparities in survival outcomes by breast tumor subtype among African American women in Memphis, Tennessee. *Cancer Med.* **6**(7), 1776–1786 (2017).
44. Costantino, N. S., Freeman, B., Shriver, C. D. & Ellsworth, R. E. Outcome disparities in African American compared with European American women with ER+HER2-tumors treated within an equal-access health care system. *Ethn. Dis.* **26**(3), 407–416 (2016).
45. Rummel, S., Penatzer, C. E., Shriver, C. D. & Ellsworth, R. E. PSPHL and breast cancer in African American women: Causative gene or population stratification?. *BMC Genet.* **15**, 38 (2014).
46. Parada, H. Jr. et al. Race-associated biological differences among luminal A and basal-like breast cancers in the Carolina Breast Cancer Study. *Breast Cancer Res.* **19**(1), 131 (2017).
47. Stewart, P. A., Luks, J., Roycik, M. D., Sang, Q. X. & Zhang, J. Differentially expressed transcripts and dysregulated signaling pathways and networks in African American breast cancer. *PLoS One* **8**(12), e82460 (2013).
48. Wu, J. et al. Elevated HMGA2 expression is associated with cancer aggressiveness and predicts poor outcome in breast cancer. *Cancer Lett.* **376**(2), 284–292 (2016).
49. Ding, Y. C. et al. Molecular subtypes of triple-negative breast cancer in women of different race and ethnicity. *Oncotarget* **10**(2), 198–208 (2019).
50. Dai, X., Xiang, L., Li, T. & Bai, Z. Cancer hallmarks, biomarkers and breast cancer molecular subtypes. *J Cancer* **7**(10), 1281–1294 (2016).
51. Adeola, F. Normalization of gene expression by quantitative RT-PCR in human cell line: Comparison of 12 endogenous reference genes. *Ethiop. J. Health Sci.* **28**(6), 741–748 (2018).
52. Parker, J. S. et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J. Clin. Oncol.* **27**(8), 1160–1167 (2009).

Acknowledgements

We are grateful to the City of Hope Shared Resources: Molecular Pathology, Anatomic Pathology, and Mul-

ti-Scale Translational Research. All these cores at City of Hope are supported by the National Cancer Institute of the National Institutes of Health under award number P30CA033572. We would also like to thank the COH Lee Graff Medical and Scientific Library for assistance with the literature review. Importantly, we would like to sincerely thank the City of Hope patients for the donation of their tissue for research.

Author contributions

VJ-conceptualization, methodology, investigation, formal analysis, resources, data curation, writing-original draft, writing-review and editing, project administration, and funding acquisition HY,YCY, YW-formal analysis, investigation, data curation, writing-original draft, writing-review and editing, and visualization SML-methodology, formal analysis, and writing review and editing DA-formal analysis, data curation, and visualization AS, CQ, JT, NS, TC-investigation, data curation, and project administration YY, JM, LY, LK, TJT, RO, QK, MD, KK, JM-writing-review, and editing RK, LR -writing-review and editing and supervision VS-conceptualization, methodology, resources, writing reviews and editing, supervision, and funding acquisition.

Funding

Philanthropy, City of Hope (COH) Shared Resources Pilot Award, American Cancer Society-Institutional Research Grant, K12 Paul Calabresi Career Development Award for Clinical Oncology.

Declarations

Competing interests

Yuan Yuan declares Grant/Research Support from Imugene, Genentech, Pfizer, Merck, Edenbridge, Agenus; Consultant for Gilead Sciences and Novartis; On the Speakers Bureau for AstraZeneca, Daiichi-Sankyo, Merck, Pfizer and Gilead Sciences. Kevin Kalinsky declares Employment/Stock: Spouse—EQRX, Grail (Prior Employee); Advisory/Consulting: Eli-Lilly, Pfizer, Novartis, AstraZeneca, Daiichi Sankyo, Puma, 4D Pharma, Oncosec, Immunomedics, Puma, Merck, Seattle Genetics, Mersana, and Cyclocel; Institutional Research Funding: Genentech/Roche, Novartis, Eli-Lilly, AstraZeneca, Daiichi Sankyo, Ascentage. The remaining authors have no conflicts of interest.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-89274-9.

Correspondence and requests for materials should be addressed to V.J.