Breast Cancer Research

RESEARCH

Open Access

Annotation-free deep learning algorithm trained on hematoxylin & eosin images predicts epithelial-to-mesenchymal transition phenotype and endocrine response in estrogen receptor-positive breast cancer

Kaimin Hu[1,2†], Yinan Wu[3†], Yajing Huang[4†], Meiqi Zhou[1,2], Yanyan Wang[1,2] and Xingru Huang[5*]

Abstract

Recent evidence indicates that endocrine resistance in estrogen receptor-positive (ER+) breast cancer is closely correlated with phenotypic characteristics of epithelial-to-mesenchymal transition (EMT). Nonetheless, identifying tumor tissues with a mesenchymal phenotype remains challenging in clinical practice. In this study, we validated the correlation between EMT status and resistance to endocrine therapy in ER+ breast cancer from a transcriptomic perspective. To confirm the presence of morphological discrepancies in tumor tissues of ER+ breast cancer classified as epithelial- and mesenchymal-phenotypes according to EMT-related transcriptional features, we trained deep learning algorithms based on EfficientNetV2 architecture to assign the phenotypic status for each patient utilizing hematoxylin & eosin (H&E)-stained slides from The Cancer Genome Atlas database. Our classifier model accurately identified the precise phenotypic status, achieving an area under the curve (AUC) of 0.886 at the tile-level and an AUC of 0.910 at the slide-level. Furthermore, we evaluated the efficacy of the classifier in predicting endocrine response using data from an independent ER+ breast cancer patient cohort. Our classifier achieved a predicting accuracy of 81.25%, and 88.7% slides labeled as endocrine resistant were predicted as the mesenchymal-phenotype, while 75.6% slides labeled as sensitive were predicted as the epithelial-phenotype. Our work introduces an H&E-based framework capable of accurately predicting EMT phenotype and endocrine response for ER+ breast cancer, demonstrating its potential for clinical application and benefit.

Keywords Deep learning, Estrogen receptor-positive breast cancer, Epithelial-to-mesenchymal transition, Digital pathology, Phenotypic classifier, Endocrine therapy response

†Kaimin Hu, Yinan Wu and Yajing Huang contributed equally to this work.

*Correspondence:
Xingru Huang
xingru.huang@qmul.ac.uk
[1]Department of Breast Surgery and Oncology, the Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China
[2]The Key Laboratory of Cancer Prevention and Intervention, China National Ministry of Education, the Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China
[3]Hangzhou Institute of Medicine (HIM), Zhejiang Cancer Hospital, Chinese Academy of Sciences, Hangzhou, China
[4]Department of Pathology, the Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China
[5]School of Electronic Engineering and Computer Science, Queen Mary University of London, London, UK

Introduction

The estrogen receptor-positive (ER+) subtype accounts for approximately 80% of all breast cancers [1]. The issue of endocrine resistance poses a major challenge in curing ER+ breast cancers, despite this particular subtype consistently having a better prognosis [2]. A high-mesenchymal cell state observed in cancer cell lines and human tumors has been associated with resistance to multiple treatment modalities across diverse cancer lineages [3]. Recent evidence has emerged that transition from an epithelial to a mesenchymal phenotype during cancer progression converts ER+ breast cancer into a state insensitive to ERα targeted agents [4–6]. Based on single-cell transcriptomics data, a subpopulation of "pre-adapted" treatment-naive cells was detected in luminal breast cancer [7]. These cells harbored features of epithelial-to-mesenchymal transition (EMT) and exhibited a survival advantage during short-term endocrine therapy [7]. Therefore, identifying cases with a mesenchymal phenotype among ER+ breast cancers could be a critical approach to discern patients with endocrine resistance. However, the existing EMT stratification systems are mostly based on complex molecular experiments or EMT-related gene signatures [8–11]. Currently, in clinical practice, there is still lack of a straightforward and efficient method to differentiate between the epithelial and mesenchymal phenotypes in tumor tissues.

In vitro data revealed that tamoxifen-resistant MCF7 cells exhibited higher expression of genes associated with EMT and invasiveness, compared to parental cells [4, 6]. Simultaneously, these cells displayed noticeable morphological changes [4]. Therefore, we hypothesized that endocrine-resistant ER+ breast cancers with a mesenchymal phenotype may exhibit distinct yet indiscernible morphological features in clinical hematoxylin & eosin (H&E) pathology images, which could be well distinguished via deep learning (DL) approaches. Deep neural networks, particularly convolutional neuronal networks (CNNs), are widely used algorithms for image classification [12]. In this study, we firstly established an EMT score to confirm the correlation between the EMT process and endocrine resistance in ER+ breast cancers. Based on the transcriptional classification, we trained a CNN-based deep learning network to accurately identify the phenotypic status of ER+ breast cancers and assist in predicting their endocrine response.

Methods

Cell culture and endocrine-resistant cell model

T47D cell line was obtained from ATCC and cultured according to ATCC's recommendations in Dulbecco's modified Eagle's medium (DMEM; Gibco) with 10% FBS (Gibco) and 1% penicillin/streptomycin (P/S). The fulvestrant-resistant T47D (T47D-FulvR) cell model was induced referring to previous research method [13], using a concentration of 0.5 μM of fulvestrant (Fulv, S1191, Selleck). All cells were maintained in a 5% CO_2 incubator at 37 ℃.

Gene expression datasets

Both T47D parental and T47D-FulvR cell samples were sequenced on an BGISEQ platform following the manufacturer's instructions. Publicly available gene-expression data and their corresponding clinical annotations for breast cancer patients were downloaded from the METABRIC [14] and The Cancer Genome Atlas (TCGA) databases [15]. The FPKM values were converted into transcripts per kilobase million. Besides, additional gene-expression profiles and clinical data from eight breast cancer cohorts, including GSE125738, GSE85536, GSE111563, GSE20181, GSE147271, GSE87411, GSE59515 and E-MTAB-9917 (details provided in Supplementary Table S1), were gathered from the Gene Expression Omnibus (GEO) and ArrayExpress databases [16, 17]. PAM50 subtypes were determined using the PAM50 classifier [18].

Generation of the EMT-score

Milena P. Mak and colleagues previously identified a pan-cancer EMT-related gene signature consisting of 77 genes [8]. From this signature, we selected the 75 genes that are most specific to breast cancer, including 52 'mesenchymal' genes and 23 'epithelial' genes (Supplementary Table S2), to form the EMT-related gene signature in this study. The EMT score for each breast cancer sample was calculated as the average mRNA expression level of 'mesenchymal' genes subtracted from that of 'epithelial' genes. The ER+ breast cancers, in both TCGA-BRCA and METABRIC cohorts, were then classified according to trisection of their EMT scores as epithelial (Epi-, defined by EMT score ≤ lowest 1/3), intermediate (defined by lowest 1/3 < EMT score < highest 1/3) or mesenchymal (Mes-, defined by EMT score ≥ highest 1/3) subtype [19].

Clinical logistic regression model

To investigate the predictive power of clinicopathologic characteristics in distinguishing between the Epi- and Mes-phenotypes, we developed a logistic regression model incorporating age, histological type, ER level, HER2 status, and PAM50 subtype. These features were extracted from the TCGA database and selected based on their significant differences between the two transcriptome phenotypes.

Gene set enrichment analysis (GSEA)

GSEA was performed to explore enriched pathways and to annotate RNA-seq data by utilizing predefined hallmark gene sets from the Molecular Signatures Database

version 2023.2.Hs, employing the "clusterProfiler" package for enrichment analyses of functional annotation estimates to each sample.

Pathological cohorts

TCGA-BRCA image cohort A total of 536 ER+ breast cancer patients in the TCGA-BRCA cohort were classified as either Mes- or Epi-phenotype based on their transcriptional EMT scores. We collected a total of 1,076 eligible H&E-stained whole slide images (WSIs) corresponding to the aforementioned 536 cases from The Cancer Image Archive (TCIA) tissue slide dataset [20], with 534 labeled as Epi-phenotype and 542 as Mes-phenotype. The WSIs were scanned at 40× or 20× magnification.

ZEYY endocrine response cohort Patients with ER+ breast cancer who underwent core biopsy or surgery at the Second Affiliated Hospital of Zhejiang University between 2015 and 2021 were retrospectively identified from the database of the Department of Pathology. In this cohort, patients with more than 10% of invasive cancer cells positive for ER expression were designated as ER+ and received endocrine therapy with selective estrogen receptor modulators or aromatase inhibitors ± ovarian function suppression, according to the recommendation of the clinical physicians. Resistance to endocrine therapy encompassed both intrinsic and acquired resistance [21, 22]. Therefore, eligibility criteria for tumor tissues defined as endocrine resistance were as follows: (1) early breast cancers with a disease-free interval (DFI) of less than 24 months; (2) recurrent lesions that developed during endocrine therapy; (3) advanced or metastatic diseases that progressed after endocrine therapy. ER+ Cases with pathologically confirmed lymph node metastasis but without distant metastasis, who had achieved a DFI of more than 12 months after completing more than 5 years of adjuvant endocrine therapy, were classified as the endocrine-sensitive subgroup. After thoroughly reviewing the initial pathology reports, slide quality, treatment modalities, and prognostic information of eligible patients, a total of 63 cases with 144 H&E-stained WSIs were retrospectively collected for this study, including 25 resistant cases with 62 slides and 38 sensitive cases with 82 slides. The CONSORT diagram is shown in Supplementary Fig. 1. WSIs were scanned using digital slide scanner KF-PRO-400 (Jiangfeng Bio-Information Technology Co, Ningbo, China) at 40× magnification. Images with low quality, owing to extreme fading, low-resolution, or the absence of invasive tumor regions, were excluded.

Pathological image preprocessing and patch generation

Otsu's thresholding was utilized to separate the background from the foreground tissue using high-resolution (40×) H&E digital pathology images. Subsequently, all slides were tessellated into non-overlapping patches at 1024 × 1024 pixels and then downsampled by a factor of 2, resulting in patches of 512 × 512 pixels with a depth of 3 channels, while preserving the integrity of the pathological information.

Tumor tissue filtration

To train a filter for non-tumor tissues, we manually curated a subset of patches from 60 WSIs, purposefully selecting those that predominantly contained stroma, necrosis, lymphocytes, and blood vessels, which are typically considered non-tumorous for our analysis. Then we employed an EfficientNetV2B2 architecture as the backbone for the classifier capable of sifting through all patches to exclude these non-target elements. The entire datasets were processed through the filter to retain patches containing more than 50% of tumorous lesions.

Mes- and Epi-phenotype classification

For Mes- and Epi-phenotype classification, we developed another model based on EfficientNetV2B2 using patches detected above. Each patch was assigned either the Mes or Epi label of its corresponding slide based on the patient's EMT signature score. After classifying all tumor patches from the same slide, the mean prediction value of these patches determined the slide's final prediction result. Subsequently, all slides were classified using binary cross-entropy loss. We assessed the classifier's performance using an internal test set from the TCGA-BRCA cohort and validated it on an independent ZEYY endocrine response cohort. Specifically, for the TCGA-BRCA dataset, we split the data into training and internal test sets in a 9:1 ratio, with 90% of the samples used for training and the remaining 10% serving as the internal test set. Threshold optimization was not performed within the ZEYY cohort to avoid potential data leakage, which could overestimate model performance. Instead, pre-established thresholds from the TCGA-BRCA training set were used, ensuring an unbiased and robust evaluation of the model's generalizability to the independent ZEYY cohort. Performance metrics included area under the curve (AUC), accuracy, positive predictive value (PPV), and negative predictive value (NPV).

Model architecture and training details

The non-tumor tissue filtration and EMT phenotype classification models were both based on EfficientNetV2, trained using a transfer learning approach by fine-tuning from pre-trained weights on the ImageNet dataset to leverage learned features and accelerate convergence. To enhance model robustness and prevent overfitting, we implemented various data augmentation techniques, including random rotations, flips, and color jittering, to simulate the variability in real-world pathology images.

A cosine decay learning rate schedule with warm restarts was employed, allowing the learning rate to decrease in a cosine curve manner, resetting periodically to avoid local minima and ensure thorough exploration of the loss landscape. Weight decay and dropout were applied as regularization strategies to prevent overfitting, and a batch size of 8 was selected after empirical evaluation for optimal memory usage and convergence speed. The initial learning rate was set at $1e^{-5}$, carefully chosen to balance convergence rate and training stability. The cross-entropy loss function was utilized for training the classifier.

Heatmap generation for WSIs

To provide a comprehensive tissue assessment, we utilized the trained phenotypic status classifier on WSIs. The classifier analyzed the images in strides of 128 pixels, producing detailed heatmaps that visually depicted the model's activation levels, effectively emphasizing phenotypic statuses throughout the tissue.

Statistical analysis

For statistical analysis, the Mann-Whitney U test was used for numeric data, Pearson's chi-squared test and Fisher's exact test were used for categorical data. Box-and-whisker plots indicate the median, 25th and 75th percentiles, with whiskers representing the minima and maxima of the distributions through the ggplot2 R package. A two-sided p value of <0.05 was considered statistically significant. Analyses were performed using R software version 4.1.0 and SPSS version 20 (SPSS Inc., Chicago, IL).

Results

The correlation between endocrine resistance and EMT status in ER+ breast cancer

In vitro data from three independent datasets (GSE125738, GSE85536 and our data datasets) of T47D, a representative ER+ breast cancer cell line, showed that the transcriptional EMT scores were notably higher in both tamoxifen- and fulvestrant-resistant T47D cells, compared to parental cells (Fig. 1A). As was reported, there are morphological differences between the parental and tamoxifen-resistant MCF7 cells [4], a transition from tightly packed parental T47D cells to independently distributed cells was similarly observed in fulvestrant-resistant cells (Fig. 1B). Clinical data from four ER+ breast cancer cohorts (GSE111563, GSE20181, GSE147271 and GSE87411 datasets) also revealed increased EMT scores for residual lesions following neoadjuvant endocrine therapy, as compared to the initial samples (Fig. 1C). As

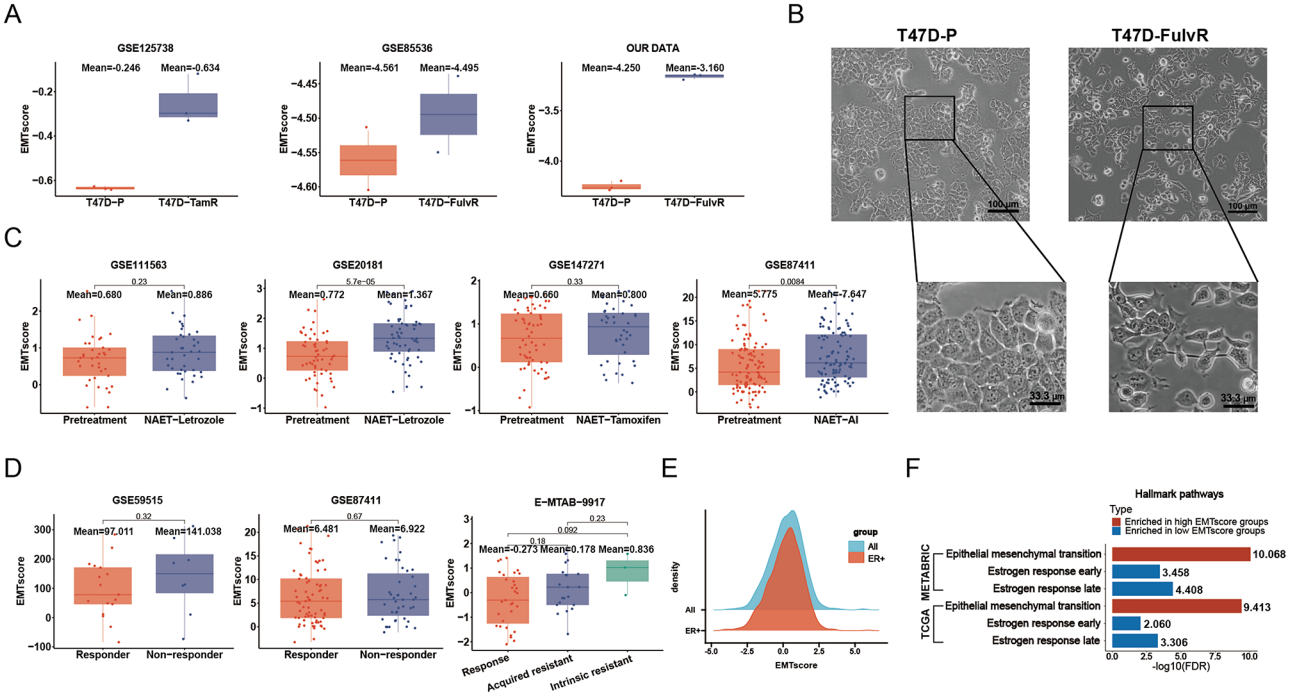


Fig. 1 The correlation between endocrine resistance and EMT status in ER+ breast cancer **A** Transcriptional EMT scores in parental and tamoxifen/fulvestrant-resistant T47D cells. **B** Morphological characteristics of the parental (T47D-P) and fulvestrant-resistant T47D cells (T47D-FulvR). **C** EMT scores in primary and residual lesions following neoadjuvant endocrine therapy from ER+ breast cancers. **D** EMT scores in ER+ breast cancer samples that exhibit response to endocrine therapy and those with intrinsic and acquired resistance. **E** The distribution of EMT scores in all and ER+ breast cancers from the TCGA-BRCA cohort. **F** Hallmark gene sets enriched in ER+ breast cancers with high- and low-EMT scores in both the TCGA-BRCA and METABRIC cohorts. For figures C and D, box-and-whisker plots indicate median; 25th and 75th percentiles; whiskers, minima and maxima of the distributions; the Wilcoxon signed rank test

Table 1 Comparison of clinicopathologic characteristics and molecular subtypes among estrogen receptor-positive breast cancers stratified into epithelial, intermediate, and mesenchymal phenotypes within the TCGA-BRCA cohort

Cases		Total (803)	Epithelial (268)	Intermediate (267)	Mesenchymal (268)	*P*-value
Age (years)	Mean ± SD	59 ± 13	61 ± 14	59 ± 13	57 ± 13	*0.002*
Menopausal status	Pre	163 (20.3%)	44 (16.4%)	62 (23.2%)	57 (21.3%)	*0.095*
	Post	526 (65.5%)	191 (71.3%)	170 (63.7%)	165 (61.6%)	
	Others	114 (14.2%)	33 (12.3%)	35 (13.1%)	46 (17.1%)	
Histological type	Ductal	538 (67.0%)	192 (71.6%)	183 (68.5%)	163 (60.8%)	*< 0.001*
	Lobular	190 (23.7%)	39 (14.6%)	68 (25.5%)	83 (31.0%)	
	Others*	75 (9.3%)	37 (13.8%)	16 (6%)	22 (8.2%)	
Clinical T stage	T1	212 (26.4%)	57 (21.3%)	72 (27%)	83 (31.0%)	*0.124*
	T2	452 (56.3%)	154 (57.5%)	156 (58.4%)	142 (53.0%)	
	T3	111 (13.8%)	44 (16.4%)	33 (12.4%)	34 (12.7%)	
	T4	26 (3.2%)	11 (4.1%)	6 (2.2%)	9 (3.3%)	
	unknown	2 (0.2%)	2 (0.7%)	0 (0%)	0 (0%)	
Clinical N stage	N0	354 (44.1%)	120 (44.8%)	112 (41.9%)	122 (45.5%)	*0.784*
	N1	284 (35.4%)	96 (35.8%)	96 (36%)	92 (34.3%)	
	N2	90 (11.2%)	27 (10.1%)	30 (11.2%)	33 (12.3%)	
	N3	59 (7.3%)	18 (6.7%)	22 (8.2%)	19 (7.1%)	
	unknown	16 (2.0%)	7 (2.6%)	7 (2.6%)	2 (0.8%)	
Clinical M stage	M0	651 (81.1%)	222 (82.8%)	212 (79.4%)	217 (81.0%)	*0.429*
	M1	16 (2.0%)	2 (0.8%)	8 (3%)	6 (2.2%)	
	unknown	136 (16.9%)	44 (16.4%)	47 (17.6%)	45 (16.8%)	
Clinical stage	I	140 (17.4%)	41 (15.3%)	50 (18.7%)	49 (18.3%)	*0.333*
	II	442 (55%)	157 (58.6%)	139 (52.1%)	146 (54.5%)	
	III	190 (23.7%)	64 (23.9%)	66 (24.7%)	60 (22.4%)	
	IV	15 (1.9%)	1 (0.4%)	8 (3%)	6 (2.2%)	
	unknown	16 (2.0%)	5 (1.8%)	4 (1.5%)	7(2.6%)	
ER level (%)	≥ 90	191 (23.8%)	71 (26.5%)	71 (26.6%)	49 (18.3%)	*0.020*
	≥ 50 and<90	83 (10.3%)	16 (6.0%)	28 (10.5%)	39 (14.6%)	
	≥ 10 and<50	42 (5.2%)	13 (4.9%)	10 (3.7%)	19 (7.1%)	
	<10	16 (2.0%)	5 (1.9%)	5 (1.9%)	6 (2.2%)	
	unknown	471 (58.7%)	163 (60.8%)	153 (57.3%)	155 (57.8%)	
PR expression	Positive	604 (75.2%)	189 (70.5%)	211 (79%)	204 (76.1%)	*0.191*
	Negative	112 (13.9%)	42 (15.7%)	34 (12.7%)	36 (13.4%)	
	unknown	87 (10.8%)	37 (13.8%)	22 (8.2%)	28 (10.4%)	
HER2 status	Positive	141 (17.6%)	36 (13.4%)	49 (18.4%)	56 (20.9%)	*0.073*
	Negative	507 (63.1%)	169 (63.1%)	174 (65.2%)	164 (61.2%)	
	unknown	155 (19.3%)	63 (23.5%)	44 (16.5%)	48 (17.9%)	
PAM50 subtype	LumA	393 (48.9%)	125 (46.6%)	132 (49.4%)	136 (50.7%)	*0.008*
	LumB	216 (26.9%)	92 (34.3%)	69 (25.8%)	55 (20.5%)	
	HER2	87 (10.8%)	20 (7.5%)	31 (11.6%)	36 (13.4%)	
	Basal	51 (6.4%)	20 (7.5%)	14 (5.2%)	17 (6.3%)	
	Normal	56 (7.0%)	11 (4.1%)	21 (7.9%)	24 (9.0%)	

*Other histological types: metaplastic, mucinous, mixed and other type carcinoma

ER, estrogen receptor; PR, progesterone receptor

illustrated in Fig. 1D, ER+ breast cancer samples with intrinsic and acquired resistance exhibit higher EMT scores compared to those that respond to endocrine therapy (GSE59515, GSE87411 and E-MTAB-9917 datasets).

In the TCGA-BRCA cohort, there was no difference in the distribution of EMT scores between all or ER+ breast cancers (Fig. 1E). Utilizing a trichotomy of the EMT scores, ER+ breast cancers from TCGA-BRCA were categorized into three phenotypes: mesenchymal (Mes, 268), intermediate (Int, 267) and epithelial (Epi, 268). The baseline clinical characteristics of patients across three phenotypes were detailed in Table 1. The mean age (± standard deviation) was 61 ± 14 years for patients with the Epi-phenotype, and 57 ± 13 years for those with the

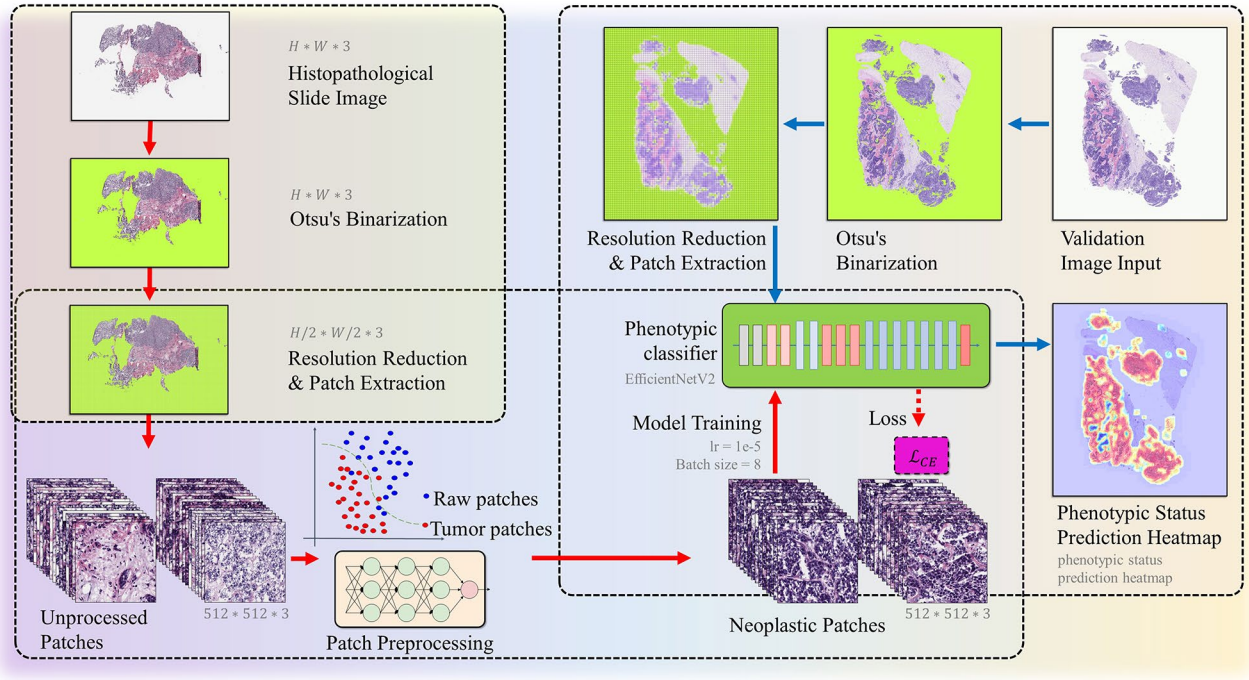


Fig. 2 Pipeline for preprocessing slides and training the model for ER+ breast cancer phenotype classification

Mes-phenotype, $P = 0.001$. A higher prevalence of invasive lobular breast cancer, decreased ER level and HER2-enriched subtype was observed in patients categorized as Mes-phenotype. There was no significant difference between patients classified as Mes- and Epi-phenotypes in menopausal status, tumor stage or PR expression. The clinical logistic regression model incorporating age, histological type, ER level, HER2 status, and PAM50 subtype demonstrated a limited capacity to differentiate between the Epi- and Mes-phenotypes, achieving an accuracy of 64.9% and an AUC of 0.690 (Supplementary Fig. 2). To investigate the biological functions of differentially expressed genes (DEGs) between Mes- and Epi-phenotypes, the hallmark modular of GSEA was performed. As expected, the DEGs were significantly enriched in the EMT, estrogen response early and late pathways (Fig. 1F). We further verified the above findings using data from the independent METABRIC cohort (Fig. 1F and Supplementary Table S3).

Taken together, these results indicated that both acquired and intrinsic endocrine resistance in breast cancer are correlated with a mesenchymal molecular phenotype. Nevertheless, clinical characteristics alone were not sufficient in distinguishing between the Epi- and Mes-phenotypes for ER+ breast cancer.

Deep learning framework distinguishes Mes- and Epi-phenotypes in ER+ breast cancer exclusively from H&E histopathology images

For clinical practice, it is too complicated and costly to uniformly predict the ER+ breast cancers to be Mes- or Epi-phenotype using transcriptional EMT-related gene signatures. Recently, deep learning has outperformed humans in certain tasks of medical data analysis, particularly in discerning morphological and genetic information for tumors using H&E-stained histopathology images efficiently and economically [23–27]. Next, we investigated whether deep learning could accurately recognize the Mes- and Epi-phenotypes categorized by our EMT score for ER+ breast cancers in TCGA-BRCA cohort. The overall workflow is shown in Fig. 2. WSIs were randomly divided into a training and a testing set, with a distribution of 90% for training and 10% for testing. The eligible 1076 WSIs were tessellated into a total of 917,934 non-overlapping patches of 512×512 pixel at a magnification of 20×. We first applied a non-tumor tissue filtration step, as described in the Methods section. Using a classifier based on the EfficientNetV2B2 architecture, patches with more than 50% of irrelevant features, such as necrosis, benign stroma, lymphocytes, and blood vessels, were excluded to ensure that the dataset primarily consisted of tumor-relevant regions. This process yielded 558,877 remaining patches, with an average of 519 patches per slide. With a total of 49,197 testing patches, our model attained a patch-level AUC of 0.886, along with a PPV of 0.807 and an NPV of 0.806

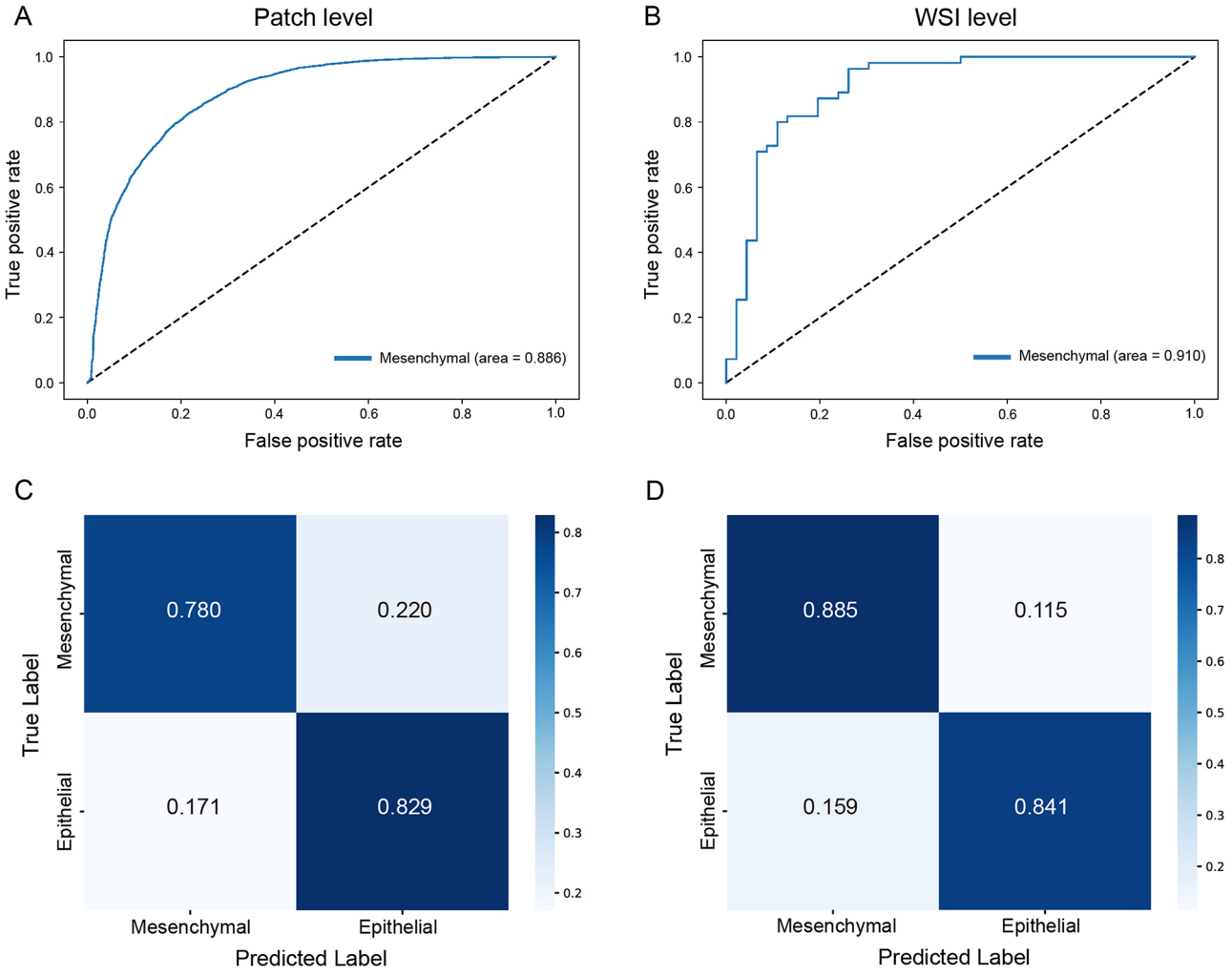


Fig. 3 The DL-based classification model distinguishes the Mes- and Epi-phenotypes in ER+ breast cancer based on H&E histopathology images. **A** Patch-level ROC curve and **C** confusion matrix of the phenotype classifier. **B** Slide-level ROC curve and **D** confusion matrix of the phenotype classifier. The Mes-phenotype was defined as the positive class in the ROC curves

for the identification of the Mes-phenotype (Fig. 3A, C). After removing images with low quality or deficient tissue information, 96 WSIs were deployed in the testing set at the slide-level, and our model achieved an AUC of 0.910, a PPV of 0.868 and an NPV of 0.860 for the identification of the Mes-phenotype, indicating that the model, by aggregating information from multiple patches, substantially achieved superior performance (Fig. 3B, D). To address the limitations of manual exclusion, we conducted additional evaluations on an augmented test set, which included 52,271 testing patches from 102 WSIs without manually excluding low-quality or irrelevant images. On this unfiltered dataset, the model achieved a patch-level AUC of 0.864 and a slide-level AUC of 0.889 for identifying the Mes-phenotype. These results highlight the importance of evaluating the model on datasets with varying levels of manual intervention to assess its robustness and potential for clinical application. Then we generated heatmaps using patch-predicted probabilities to visualize the prediction made by our model. The H&E slides and their corresponding predicted heatmaps from four representative cases were shown in Fig. 4.

Our classification model accurately predicts endocrine therapy response in ER+ breast cancer

Next, we evaluated whether our Epi- and Mes-phenotype classification model could directly predict the outcome of endocrine therapy for patients with ER+ breast cancer. In this test, the H&E stains with endocrine therapy response data from the ZEYY patient cohort were used as input to the phenotype-classifier, and the predictions (Epi- or Mes-phenotype) were tabulated against the response labels. The clinicopathological variables for this cohort were provided in Table 2. As depicted in the confusion matrix in Fig. 5A, our classification model achieved a predicting accuracy of 81.25% (117/144), and 55/62 (88.7%)

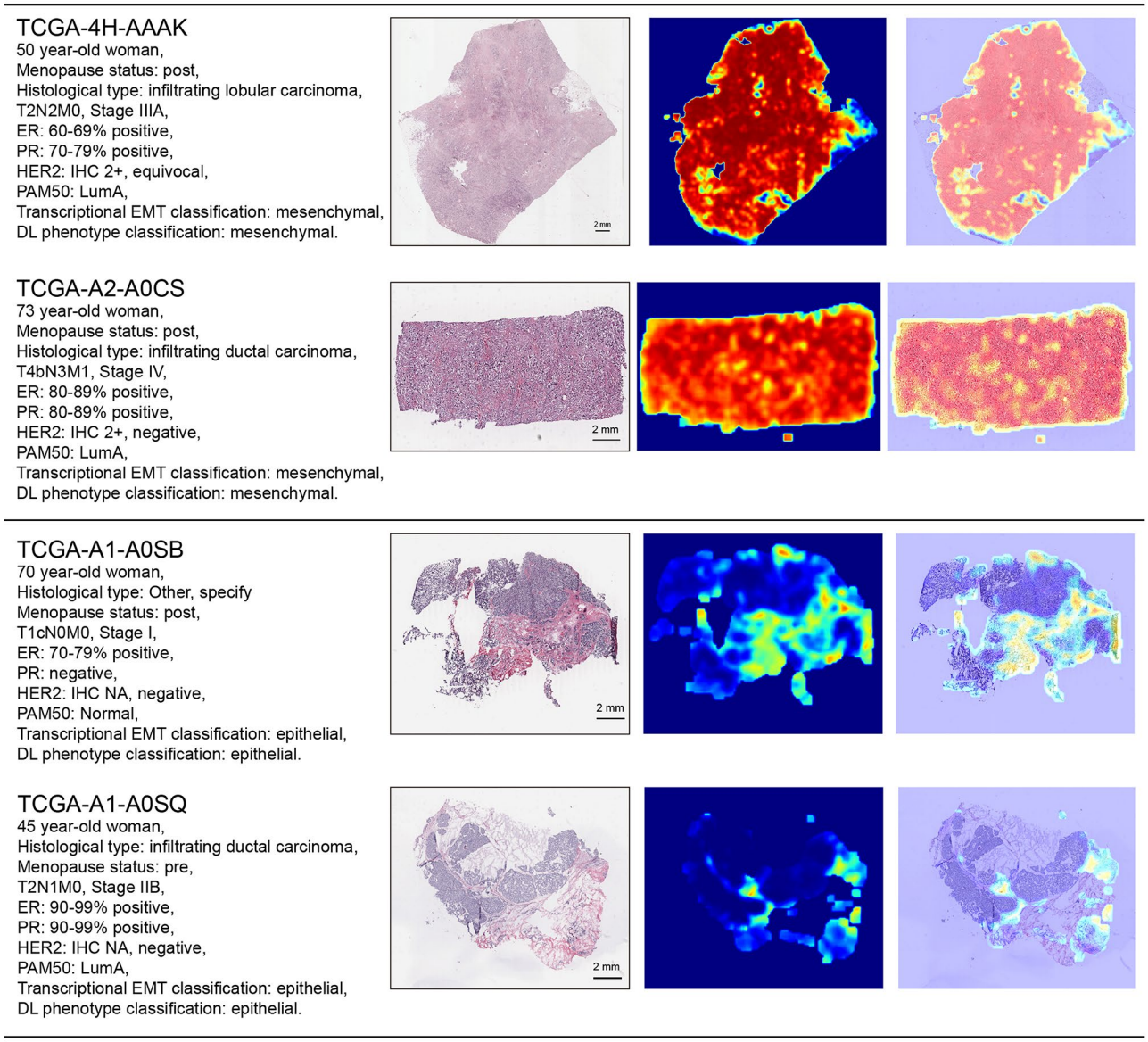


Fig. 4 Four representative H&E slides from TCGA test set and their predicted heatmaps

slides labeled as endocrine resistant were predicted as the Mes-phenotype, while 62/82 (75.6%) slides labeled as endocrine sensitive were predicted as the Epi-phenotype. The H&E slides and their corresponding predicted heatmaps from two representative cases were displayed in Fig. 5B. These results illustrated that the phenotype distinguished by our deep learning method can effectively predict the benefit of endocrine therapy for ER+ breast cancer patients.

Discussion

In this study, we validated the correlation between EMT status and resistance to endocrine therapy in ER+ breast cancer from a transcriptomic perspective. Subsequently, we trained a DL model for tumor tissue filtering and phenotypic classification according to transcriptomic phenotypes, utilizing H&E-stained histologic images. Our method can automatically detect informative regions without the necessity for annotation of interest region-level. Leveraging data from an independent patient cohort that furnished details on endocrine responses, our classification model demonstrated outstanding performance in predicting the benefits of endocrine therapy in ER+ breast cancer.

In current clinical practice, breast cancer patients with ER+ tumors are expected to benefit from receiving endocrine therapies. However, ER status alone is not an adequate predictive indicator for the response of endocrine therapy. Previous studies reported that endocrine resistance in ER+ breast cancer is commonly driven by

Table 2 Clinicopathologic characteristics of 63 estrogen receptor-positive breast cancers from the ZEYY endocrine response cohort

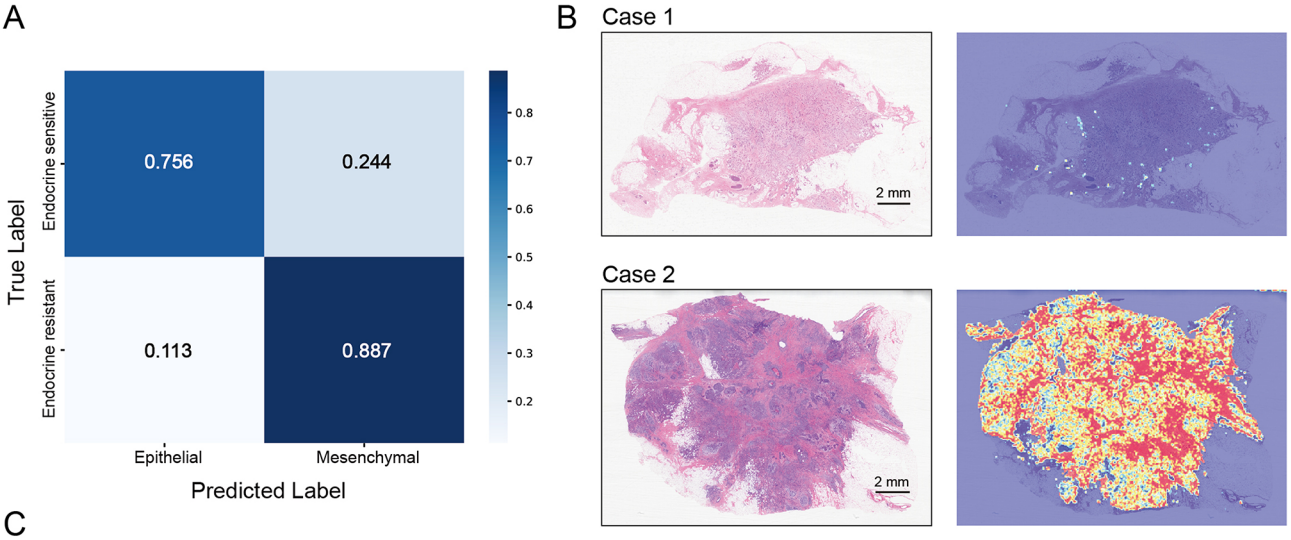
Cases	Sensitive (38)	Resistant (25)	*P*-value
Age (years, mean ± SD)	51 ± 9	50 ± 11	*0.658*
ER level (%)			
≥ 90	27 (71.1%)	17 (68.0%)	*1.000*
≥ 50 and<90	10 (26.3%)	7 (28.0%)	
≥ 10 and<50	1 (2.6%)	1 (4.0%)	
PR expression			
positive	36 (94.7%)	23 (92.0%)	*1.000*
negative	2 (5.3%)	2 (8.0%)	
HER2 status			
positive	4 (10.5%)	2 (8.0%)	*1.000*
negative	34 (89.5%)	23 (92.0%)	
Histological type			
Ductal	27 (71.1%)	22 (88.0%)	*0.034*
Lobular	3 (7.9%)	3 (12.0%)	
Others*	8 (21.0%)	0 (0.0%)	
Ki-67 expression			
>20%	15 (39.5%)	12 (48.0%)	*0.328*
≤ 20%	23 (60.5%)	13 (52.0%)	
Clinical T stage			
T1	15 (39.5%)	/	
T2	20 (52.6%)	/	
T3 and T4	3 (7.9%)	/	
Clinical N stage			
N1	17 (44.7%)	/	
N2	15 (39.5%)	/	
N3	6 (15.8%)	/	
Clinical stage			
I and II	17 (44.7%)	/	
III and IV	21 (55.3%)	/	
Resistance type(22)			
intrinsic	/	19 (76.0%)	
acquired	/	6 (24.0%)	

*Other histological types: neuroendocrine, mucinous, tubular, micropapillary, and mixed carcinoma

ER, estrogen receptor; PR, progesterone receptor

ligand-independent ER reaction [2]. Genetic aberrations in the ESR1 gene were identified as one important mechanism, but they did not alter the percentage or intensity of ER expression in either primary tumors or metastasis [28]. Only about 10% of endocrine-resistant breast cancers were reported to lose ER expression [29]. In vitro data demonstrated that endocrine-resistant ER+ breast cancer cells acquired an EMT-related transcriptional features [2]. In the present study, we established an EMT-related transcriptional gene score and confirmed that both intrinsic and acquired endocrine resistance in ER+ breast cancer is correlated with a mesenchymal phenotype. Nevertheless, clinical characteristics cannot well distinguish between the epithelial and mesenchymal phenotypes in ER+ breast cancer. Our results indicated that an EMT-related multi-gene-based assay could be effective in predicting response to endocrine therapy for ER+ breast cancers. However, such an assay is expensive and challenging to calibrate for application in broader patient populations. Accordingly, there is an urgent medical demand for a cost-effective and simplified method to differentiate ER+ breast cancers into epithelial or mesenchymal phenotypes.

DL-based artificial intelligence has been developed and applied in tumor pathology, including for diagnosis, subtyping, prognostic prediction, as well as identification of specific pathological features [30]. Recent studies have addressed aspects of accurate determination of histologic grade, ER and HER2 status for breast cancer via DL approaches with affordable H&E stains [26, 27, 31–34]. Naik et al. [27], designed a multiple instance learning-based deep neural network called ReceptorNet, which achieved an AUC of 0.92 on the test set for the estimation of ER status. The DeepGrade model developed by Wang et al. [26] stratified the intermediate-risk Nottingham histological grade 2 cases into DG2-low



C

Patients	Histological type	Surgical procedure	Stage	IHC	Treatment protocol	Prognosis
Case 1: A 59-year-old postmenopausal woman	IDC mixed with focal micropapillary carcinoma	Modified radical mastectomy	pT2N2M0 Grade 2 tumor size: 2.5*1.3cm positive LN: 9/16	ER 100% (+) PR 90% (+) HER2 (1+) Ki-67 5% (+)	adjuvant EC*4-T*4→ radiotherapy→Anastrozole	No recurrence after 8 years of follow-up
Case 2: A 43-year-old premenopausal woman	IDC	Lumpectomy + SLNB	pT1cN0(sn)M0 Grade 2 tumor size: 1.7*1.2cm	ER 90% (+) PR 100% (+) HER2 (0) Ki-67 30% (+)	adjuvant TC*4→ radiotherapy→Tamoxifen	Liver metastasis within 2 years of endocrine therapy

Fig. 5 The DL-based classification model predicts endocrine therapy response in ER+ breast cancer based on H&E histopathology images. **A** Confusion matrix of the phenotype classifier used as endocrine therapy response predictor. **B** Two representative H&E slides and corresponding predicted heatmaps. **C** Clinical characteristics of the two representative ER+ breast cancer cases. IDC, invasive ductal carcinoma; SLNB, sentinel lymph node biopsy; LN, lymph node; IHC, immunohistochemistry; ER, estrogen receptor; PR, progesterone receptor; E, epirubicin; C, cyclophosphamide; T, docetaxel

and DG2-high subgroups with significant prognostic difference (HR = 2.94, 95% CI: 1.24–6.97, $P = 0.015$). Farahmand and colleagues [34] trained a CNN-based classifier for estimating HER2 status on H&E-stained WSIs that were manually annotated for tumor regions of interest (ROIs). Their classifier achieved an AUC of 0.90 at the slide-level and an AUC of 0.81 on an independent TCGA test set. In light of these above, certain biological traits influence the morphological features of breast cancer cells that are imperceptible to the naked eye, and DL methods can be an effective way to accurately recognize these subtle features.

In vitro data from Bi et al. [4] and our study showed that parental ER+ breast cancer cell lines displayed characteristic epithelial cell morphology, forming tightly packed cobblestone-like clusters. In contrast, endocrine-resistant cells started to spread as individual cells and acquired a phenotype resembling mesenchymal cells. To confirm the presence of morphological differences in tumor tissues of ER+ breast cancers classified as Epi- and Mes-phenotypes based on EMT-related transcriptional features, we developed an EfficientNetV2-based classifier model to assign the phenotypic status of ER+ breast cancer utilizing H&E-stained histologic images. Our classifier accurately identified the specific phenotypic status, achieving an AUC of 0.886 at the tile-level and an AUC of 0.910 at the slide-level. Furthermore, we evaluated the accuracy of the classifier model in predicting the response to endocrine therapy in ER+ breast cancer patients with data from an independent patient cohort that provided response information. The results indicated that 89.9% of samples classified as the Epi-phenotype by our classifier responded to endocrine treatment, while 73.3% of those predicted as the Mes-phenotype showed resistance to endocrine therapy. Considering that patients in this test set could have several pathological slides from distinct tumor lesions, we assessed the consistency of classification across slides from the same patient. We found a 92.0% agreement rate for endocrine-resistant patients and a 68.4% agreement rate for endocrine-sensitive cases. The inconsistencies in classification results among different slides from the same patient suggest a potential correlation with intratumoral heterogeneity. Indeed, our phenotype classifier for ER+ breast cancer, trained to detect morphological features within histological images, has proven to be a dependable predictor of endocrine therapy response, offering particularly valuable insights for patients who exhibit resistance to this treatment.

The EfficientNetV2 architecture we utilized is indeed a state-of-the-art CNN designed for scalability and

efficiency in image classification tasks. It incorporates advancements such as progressive learning, where the network begins with a smaller model and gradually increases both depth and width during training, enabling the network to effectively capture fine-grained features from pathology images. The EfficientNetV2 model employs a compound scaling method that uniformly scales network width, depth, and resolution using a set of fixed scaling coefficients. This methodology enhances the model's adaptability to various computational budgets and image resolutions. In the realm of digital pathology, EfficientNetV2's scaling capabilities empower it to handle large image sizes typical of WSIs. In this study, our tumor tissue filter and phenotype classifier model based upon the EfficientNetV2 algorithm can automatically identify informative tumor regions, recognize subtle morphological traits to distinguish between epithelial and mesenchymal phenotypes, and predict endocrine treatment response in ER + breast cancer. The heatmaps produced by our model showed a remarkable agreement with tumor ROIs annotated by pathologists.

This study also has several limitations. First, it was retrospective with convenience samples, and the extensibility of our classifier in predicting endocrine therapy response for ER+ breast cancer was validated using solely single-center retrospective data. Further validation and optimization of the model should be pursued *via* prospective data. Second, the quality of pathological images could potentially impact the robustness and accuracy of our model. Third, our classifier model was trained and tested with a binary classification system. Variability in cut-off values may yield disparate in classification outcomes, potentially rendering our model unsuitable for tumors exhibiting substantial heterogeneity or those in an intermediate EMT state.

In summary, our work demonstrated the potential to enhance theragnosis for ER+ breast cancer by harnessing biological morphological characteristics imperceptible to clinicians via deep learning. The phenotype predicted by our classifier model is a useful indicator to augment ER status for endocrine treatment recommendation. This paves new avenues for the translating transcriptional classification into a DL model based on H&E stains, which could soon be implemented in clinical practice. This approach has the potential to enhance the accessibility of molecular stratification, especially in low-resource settings.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13058-025-01959-1.

Supplementary Material 1: Fig. 1 CONSORT Diagram: Patient inclusion and exclusion criteria in the ZEYY endocrine response cohort. Fig. 2. The logistic regression model distinguishes the Mes- and Epi-phenotypes in ER+ breast cancer based on clinical characteristics. A Confusion matrix. B ROC curve. The Mes-phenotype was defined as the positive class in the ROC curve. Table S1. Details of the eight datasets included in our study. Table S2. The EMT related gene signature composed by 75 genes in our study. Table S3. Comparison of clinicopathologic features and molecular subtypes between epithelial and mesenchymal estrogen receptor-positive breast cancers in the METABRIC cohort.

Acknowledgements

This research was supported in part by National Natural Science Foundation of China (82203077). The funder played no role in study design, data collection, analysis and interpretation of data, or the writing of this manuscript.
The authors would like to thank Bailai Hu and Lifang Yao, Department of Pathology, the Second Affiliated Hospital, Zhejiang University School of Medicine for their assistance in acquiring the eligible whole slide images in our study.

Author contributions

KM.H and XR.H performed study concept and design; KM.H, XR.H and YN.W performed development of methodology and writing, review and revision of the paper; XR.H and YN.W provided acquisition, analysis and interpretation of data, and statistical analysis; YJ.H, MQ.Z and YY.W provided technical and material support. All authors reviewed and approved the final manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (82203077).

Data availability

Data is provided within the manuscript or supplementary information files. The raw data are available upon reasonable request.

Declarations

Ethics approval and consent to participate

The research ethics committee of the Second Affiliated Hospital of Zhejiang University approved this retrospective study (approval no. 20220731) and waived the requirement for informed consent since it solely utilized pre-existing medical data. This study was conducted in compliance with the ethical standards of the participating institution and adhered to the tenets outlined in the Declaration of Helsinki.

Competing interests

The authors declare no competing interests.

Conflict of interest

The authors declare no competing interests.

Received: 10 October 2024 / Accepted: 1 January 2025
Published online: 12 January 2025

References

1. Giaquinto AN, Sung H, Miller KD, Kramer JL, Newman LA, Minihan A, et al. Breast Cancer Stat 2022 CA Cancer J Clin. 2022;72(6):524–41.
2. Hanker AB, Sudhan DR, Arteaga CL. Overcoming endocrine resistance in breast Cancer. Cancer Cell. 2020;37(4):496–513.
3. Viswanathan VS, Ryan MJ, Dhruv HD, Gill S, Eichhoff OM, Seashore-Ludlow B, et al. Dependency of a therapy-resistant state of cancer cells on a lipid peroxidase pathway. Nature. 2017;547(7664):453–7.
4. Bi M, Zhang Z, Jiang YZ, Xue P, Wang H, Lai Z, et al. Enhancer reprogramming driven by high-order assemblies of transcription factors promotes phenotypic plasticity and breast cancer endocrine resistance. Nat Cell Biol. 2020;22(6):701–15.
5. Wu DP, Zhou Y, Hou LX, Zhu XX, Yi W, Yang SM, et al. Cx43 deficiency confers EMT-mediated tamoxifen resistance to breast cancer via c-Src/PI3K/Akt pathway. Int J Biol Sci. 2021;17(10):2380–98.
6. Lombardo Y, Faronato M, Filipovic A, Vircillo V, Magnani L, Coombes RC. Nicastrin and Notch4 drive endocrine therapy resistance and epithelial to

mesenchymal transition in MCF7 breast cancer cells. Breast Cancer Res. 2014;16(3):R62.
7. Hong SP, Chan TE, Lombardo Y, Corleone G, Rotmensz N, Bravaccini S, et al. Single-cell transcriptomics reveals multi-step adaptations to endocrine therapy. Nat Commun. 2019;10(1):3840.
8. Mak MP, Tong P, Diao L, Cardnell RJ, Gibbons DL, William WN, et al. A Patient-Derived, Pan-cancer EMT signature identifies global molecular alterations and Immune Target Enrichment following epithelial-to-mesenchymal transition. Clin Cancer Res. 2016;22(3):609–20.
9. Byers LA, Diao L, Wang J, Saintigny P, Girard L, Peyton M, et al. An epithelial-mesenchymal transition gene signature predicts resistance to EGFR and PI3K inhibitors and identifies Axl as a therapeutic target for overcoming EGFR inhibitor resistance. Clin Cancer Res. 2013;19(1):279–90.
10. Lamouille S, Xu J, Derynck R. Molecular mechanisms of epithelial-mesenchymal transition. Nat Rev Mol Cell Biol. 2014;15(3):178–96.
11. Jeong H, Ryu YJ, An J, Lee Y, Kim A. Epithelial-mesenchymal transition in breast cancer correlates with high histological grade and triple-negative phenotype. Histopathology. 2012;60(6B):E87–95.
12. Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Duan Y, Al-Shamma O, et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J Big Data. 2021;8(1):53.
13. Li Y, Zhao J, Gutgesell LM, Shen Z, Ratia K, Dye K, et al. Novel pyrrolopyridone bromodomain and Extra-terminal Motif (BET) inhibitors effective in endocrine-resistant ER + breast Cancer with Acquired Resistance to Fulvestrant and Palbociclib. J Med Chem. 2020;63(13):7186–210.
14. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature. 2012;486(7403):346–52.
15. Cancer Genome Atlas N. Comprehensive molecular portraits of human breast tumours. Nature. 2012;490(7418):61–70.
16. Edgar R, Domrachev M, Lash AE. Gene expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 2002;30(1):207–10.
17. Parkinson H, Kapushesky M, Shojatalab M, Abeygunawardena N, Coulson R, Farne A, et al. ArrayExpress–a public database of microarray experiments and gene expression profiles. Nucleic Acids Res. 2007;35(Database issue):D747–50.
18. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol. 2009;27(8):1160–7.
19. Zhang SY, Ren XY, Wang CY, Chen XJ, Cao RY, Liu Q, et al. Comprehensive characterization of Immune Landscape based on Epithelial-Mesenchymal Transition Signature in OSCC: implication for prognosis and immunotherapy. Front Oncol. 2021;11:587862.
20. Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, et al. The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. J Digit Imaging. 2013;26(6):1045–57.
21. Sledge GW Jr., Toi M, Neven P, Sohn J, Inoue K, Pivot X, et al. MONARCH 2: Abemaciclib in Combination with Fulvestrant in Women with HR+/HER2- advanced breast Cancer who had progressed while receiving endocrine therapy. J Clin Oncol. 2017;35(25):2875–84.
22. Cardoso F, Paluch-Shimon S, Senkus E, Curigliano G, Aapro MS, Andre F, et al. 5th ESO-ESMO international consensus guidelines for advanced breast cancer (ABC 5). Ann Oncol. 2020;31(12):1623–49.
23. Kather JN, Pearson AT, Halama N, Jager D, Krause J, Loosen SH, et al. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. Nat Med. 2019;25(7):1054–6.
24. Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyo D, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. Nat Med. 2018;24(10):1559–67.
25. Echle A, Rindtorff NT, Brinker TJ, Luedde T, Pearson AT, Kather JN. Deep learning in cancer pathology: a new generation of clinical biomarkers. Br J Cancer. 2021;124(4):686–96.
26. Wang Y, Acs B, Robertson S, Liu B, Solorzano L, Wahlby C, et al. Improved breast cancer histological grading using deep learning. Ann Oncol. 2022;33(1):89–98.
27. Naik N, Madani A, Esteva A, Keskar NS, Press MF, Ruderman D, et al. Deep learning-enabled breast cancer hormonal receptor status determination from base-level H&E stains. Nat Commun. 2020;11(1):5727.
28. Moreira-Dinzey J, Zhan H, Rozenblit M, Krishnamurti U, Harigopal M, Zhong M, et al. The correlation of ESR1 genetic aberrations with estrogen receptor and progesterone receptor status in metastatic and primary estrogen receptor-positive breast carcinomas. Hum Pathol. 2023;137:56–62.
29. Shiino S, Kinoshita T, Yoshida M, Jimbo K, Asaga S, Takayama S, et al. Prognostic impact of discordance in hormone receptor status between primary and recurrent sites in patients with recurrent breast Cancer. Clin Breast Cancer. 2016;16(4):e133–40.
30. Jiang Y, Yang M, Wang S, Li X, Sun Y. Emerging role of deep learning-based artificial intelligence in tumor pathology. Cancer Commun (Lond). 2020;40(4):154–66.
31. Couture HD, Williams LA, Geradts J, Nyante SJ, Butler EN, Marron JS, et al. Image analysis with deep learning to predict breast cancer grade, ER status, histologic subtype, and intrinsic subtype. NPJ Breast Cancer. 2018;4:30.
32. Jaroensri R, Wulczyn E, Hegde N, Brown T, Flament-Auvigne I, Tan F, et al. Deep learning models for histologic grading of breast cancer and association with disease prognosis. NPJ Breast Cancer. 2022;8(1):113.
33. Shamai G, Binenbaum Y, Slossberg R, Duek I, Gil Z, Kimmel R. Artificial Intelligence Algorithms to assess Hormonal Status from tissue microarrays in patients with breast Cancer. JAMA Netw Open. 2019;2(7):e197700.
34. Farahmand S, Fernandez AI, Ahmed FS, Rimm DL, Chuang JH, Reisenbichler E, et al. Deep learning trained on hematoxylin and eosin tumor region of interest predicts HER2 status and trastuzumab treatment response in HER2 + breast cancer. Mod Pathol. 2022;35(1):44–51.

Publisher's note