Molecular Characterization and Classification of HER2-Positive Breast Cancer Inform Tailored Therapeutic Strategies

Yu-Wei Li[1,2], Lei-Jie Dai[1,2], Xiang-Rong Wu[1,2], Shen Zhao[1,2], Yu-Zheng Xu[1,2], Xi Jin[1,2], Yi Xiao[1,2], Ying Wang[1,2], Cai-Jin Lin[1,2], Yi-Fan Zhou[1,2], Tong Fu[1,2], Wen-Tao Yang[3], Ming Li[3], Hong Lv[3], Siyuan Chen[4], Anita Grigoriadis[4], Yi-Zhou Jiang[1,2], Ding Ma[1,2], and Zhi-Ming Shao[1,2]

ABSTRACT

HER2-positive breast cancer is an aggressive subtype that accounts for 15% to 20% of all breast cancers. Recent studies have suggested that HER2-positive breast cancer is a group of heterogeneous diseases with different sensitivities to standard treatment regimens. Revealing the molecular heterogeneity of HER2-positive breast cancer could potentially enable more precise treatment strategies. In this study, we performed multiomics profiling on a HER2-positive breast cancer cohort and identified four transcriptome-based subtypes. The classical HER2 (HER2-CLA) subtype comprised 28.3% of the samples and displayed high ERBB2 activation and significant benefit from anti-HER2 therapy. The immunomodulatory (HER2-IM) subtype (20%) featured an immune-activated microenvironment, potentially suitable for de-escalated treatment and immunotherapy. The luminal-like (HER2-LUM) subtype (30.6%) possessed similar molecular features of hormone receptor–positive HER2-negative breast cancer, suggesting endocrine therapy and CDK4/6 inhibitors as a potential therapeutic strategy. Lastly, the basal/mesenchymal-like (HER2-BM) subtype (21.1%) had a poor response to current dual HER2-targeted therapy and could potentially benefit from tyrosine kinase inhibitors. The molecular characteristics and clinical features of the subtypes were further explored across multiple cohorts, and the feasibility of the proposed treatment strategies was validated in patient-derived organoid and patient-derived tumor fragment models. This study elucidates the molecular heterogeneity of HER2-positive breast cancer and paves the way for a more tailored treatment.

Significance: Illumination of the inherent heterogeneity within HER2-positive breast cancers through the delineation of distinct molecular subtypes lays the groundwork for developing more personalized treatment strategies based on specific patient characteristics.

Introduction

HER2-positive breast cancer represents a distinct aggressive subtype, accounting for 15% to 20% of all breast cancers, and features the overexpression and amplification of HER2 (also referred to as *ERBB2*; refs. 1, 2). As novel anti-HER2-targeted therapies and treatment effectiveness have increased over the past few decades, patients have experienced a notable improvement in prognosis (3–5).

Recent studies suggested that HER2-positive breast cancer is not a uniform disease but a group of heterogeneous diseases with different sensitivities to standard treatment regimens (6–15). Therefore, more personalized and precise treatment might be needed to further enhance the response while avoiding unnecessary toxicity and cost (16, 17). Several studies have used both clinical and molecular parameters to guide the escalation or de-escalation of treatment in different risk groups of patients (17–22). However, as HER2-positive breast cancer is a molecularly heterogeneous disease, different regimens, rather than simply escalated or de-escalated treatment, may further improve the treatment options for different subgroups of patients with specific molecular features and targets. Meanwhile, practical methods to discriminate the heterogeneity were absent in clinical usage, restricting the generalization of precision management for HER2-positive breast cancers.

To achieve these goals, a more comprehensive understanding of the interpatient heterogeneity of HER2-positive breast cancer is required. Some pioneering studies have attempted to investigate the molecular heterogeneity of HER2-positive breast cancer, but neither the molecular profile nor the clinical significance has been fully elucidated (9). In this study, we conducted a multiomics profiling study of 180 patients with HER2-positive breast cancer with well-documented clinical records. We examined the heterogeneity of HER2-positive breast cancer by defining four precise subtypes with distinct characteristics and potential treatments, including a novel basal/mesenchymal-like HER2 subgroup that showed a unique EGFR-driving feature. Furthermore, we proposed feasible methods

[1]Department of Breast Surgery, Key Laboratory of Breast Cancer in Shanghai, Fudan University Shanghai Cancer Center, Shanghai, China. [2]Department of Oncology, Shanghai Medical College, Fudan University, Shanghai, China. [3]Department of Pathology, Fudan University Shanghai Cancer Center, Shanghai, China. [4]School of Cancer & Pharmaceutical Sciences, Faculty of Life Sciences and Medicine, King's College London, London, United Kingdom.

Y.-W. Li, L.-J. Dai, and X.-R. Wu contributed equally to this article.

Corresponding Authors: Yi-Zhou Jiang, Fudan University Shanghai Cancer Center, 270 Dong-An Road, Shanghai 200032, China. E-mail: yizhoujiang@fudan.edu.cn; Ding Ma, Fudan University Shanghai Cancer Center, 270 Dong-An Road, Shanghai 200032, China. E-mail: dma09@fudan.edu.cn; and Zhi-Ming Shao, Fudan University Shanghai Cancer Center, 270 Dong-An Road, Shanghai 200032, China. E-mail: zhimingshao@fudan.edu.cn.

Cancer Res 2024;84:3669–83

doi: 10.1158/0008-5472.CAN-23-4066

©2024 American Association for Cancer Research

to identify these subtypes in other cohorts by transcriptional data or immunohistochemistry (IHC) staining and validated our hypotheses in multiple clinical cohorts and patient-derived *in vitro* models.

Materials and Methods

Experimental model and study participant details

We retrospectively selected 191 consecutive patients with HER2-positive breast cancer who underwent surgery at the Department of Breast Surgery, Fudan University Shanghai Cancer Center (FUSCC), from January 1, 2013, to December 31, 2014. HER2 positivity was defined as evidence of HER2 protein overexpression measured by IHC or fluorescence in situ hybridization (FISH) according to the American Society of Clinical Oncology/College of American Pathologists guidelines (23). HER2-positive breast cancer samples were used for multiomics profiling. After quality control of tumor purity and other indicators, 180 patients with HER2-positive breast cancer were selected; all the enrolled patients were Chinese. In our cohort, 156 patients were tested HER2 IHC3+, and 33 patients underwent FISH testing were all positive.

To validate the differences in treatment efficacy among different subtypes under the existing standard therapy, we collected the clinical efficacy information and IHC staining results of 140 patients with HER2-positive breast cancer who underwent neoadjuvant treatment at FUSCC between June 1, 2021, and December 31, 2023. Detailed information on cohort collection and the process of sample preparation and multiomics data generation is provided in the Supplementary File. All tissue samples included in the present study were obtained after approval of the research by the FUSCC Ethics Committee, and each patient provided written informed consent.

Expression-based unsupervised clustering

We performed nonnegative matrix factorization (NMF) clustering (the "NMF" package v0.24.0 in R, nrun = 500) and consensus clustering (the "ConsensusClusterPlus" package v1.58.0 in R) to determine the optimal number of HER2-positive breast cancer subtypes in the mRNA data (24, 25). The optimal number of clusters was determined from the cophenetic correlation coefficient and the dispersion score of consensus matrixes. To determine the optimal number of genes in NMF clustering, we tested the clustering results by choosing genes based on the SD from the top 5% to the top 30% and finally chose the SD top 15% coding genes (2,976) to perform k-means clustering. Silhouette analysis was used to further assess the clustering quality.

In addition, consensus clustering was used to validate the optimal number. The optimal number of clusters was determined from the cumulative density function, which plots the corresponding empirical cumulative distribution, defined over a range between 0 and 1, and from a calculation of the proportion increase in the area under the cumulative density function curve.

Single-sample gene set and Kyoto Encyclopedia of Genes and Genomes enrichment analyses

Single-sample gene set enrichment analysis (ssGSEA) was used to calculate the characteristic pathway score corresponding to each subtype; this involved the "GSVA" function in the R package "GSVA" v1.42.0 (26). The analysis was also performed to validate the extrapolation in multiple cohorts. Furthermore, we used ssGSEA to evaluate the abundance of microenvironment cell subsets in each sample from the expression data (27). Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis was performed using the "enrichKEGG" function in the R package "clusterProfiler" v4.2.2 (28).

Extrapolating the clustering results to an extra HER2-positive tumor cohort

The extrapolation of mRNA-based subtypes to HER2-positive tumors in different cohorts was performed using the R package "randomForest" v4.7-1 (29). The random forest model was trained and tested using differentially expressed coding genes from four subtypes in the FUSCC HER2-positive tumor cohort. Before using the "randomForest" and "predict" functions, batch effects were removed using the "ComBat" function in the "sva" package v3.42.0 in R. Meanwhile, within the neoadjuvant therapy cohort, we meticulously selected three clinically accessible IHC markers: ESR1, FOXC1, and CD8A. These markers were subjected to pathologic staining and rigorous evaluation by expert pathologists.

Correlation network of metabolites

First, 669 polar metabolites detected in our database were included for further analysis (30). Then, the polar metabolite network was generated by calculating Spearman correlation. Given the potential confounding effect, negative correlations were not used. For a more concise representation and robust clustering effect, we retained metabolite interactions with correlation > 0.4 and FDR < 0.05 and removed the outlier nodes (KCore < 3) from the metabolite network. The ForceAtlas algorithm was used to lay out the network structures for both networks (31), and correlation networks were visualized in Gephi v0.9.3. Metabolites in distinct pathways were selected for visualization.

Patient-derived organoid culture and drug response test

We developed a biobank for human breast cancer organoid storage as previously described (32, 33). The organoids were resuscitated from the biobank and solidified by 30-minute incubation at 37°C and 5% CO_2 with 1 mL of breast cancer organoid medium. Organoid cell viability was evaluated with CellTiter-Glo 3D Cell Viability Assay (Promega, #G9683) after treatment with the indicated drugs. After RNA extraction from patient-derived organoid (PDO) tissues, the expression levels of specific genes were detected by qPCR (#RC112, #q311, Vazyme Biotech Co., Ltd.).

Patient-derived tumor fragment cultures and flow cytometry analysis

Tissue materials collected and qualified for patient-derived tumor fragment (PDTF) cultures were processed by cutting into small tumor fragments of 1 to 2 mm^3 size on ice as previously reported (34). A single PDTF per well was placed on the top of a 40 μL presolidified artificial extracellular matrix (ECM), followed by a second layer of the 40 μL matrix. Then plates were placed in an incubator at 37°C for 20 to 30 minutes for solidification. Next, 140 μL of tumor medium was added to the top, supplemented with either nivolumab (20 μg/mL final concentration) or human anti–β-Gal-hIgG4 as a control. PDTF cultures were kept at 37°C for 48 hours before flow cytometry analysis. T-cell activation was detected using antibodies from BioLegend as previously reported (34). PDTFs for flow cytometry analysis were manually retrieved from the matrix and processed into single-cell suspensions under each experimental condition. Cells were washed, resuspended in 50 μL of staining buffer (BioLegend) containing the indicate antibodies, and incubated for 20 minutes at 4°C. After washing twice, the cells were taken up in 100 μL of cell staining buffer and subjected to flow cytometry analysis. The supernatant from the PDTF cultures was used for ELISA detection (#EH0302, #EH0164, Wuhan Fine Biotech Co., Ltd.).

Quantification and statistical analysis

Frequency tabulation and summary statistics were used to characterize the data distribution. The Student *t* test, ANOVA, Mann–Whitney–Wilcoxon test, and Kruskal–Wallis test were utilized to compare continuous variables and ordered categorical variables, whereas Pearson χ^2 test and Fisher exact test were used to compare unordered categorical variables. Survival curves were constructed using the Kaplan–Meier product limit method and compared with the log-rank test. The *P* values were adjusted to the FDR using the Benjamini–Hochberg procedure in multiple comparisons. All analyses were performed using R package v4.1.1 (https://cran.r-project.org/).

Data availability

Data generated in this study, the whole-exome sequencing data, copy-number alteration (CNA) data, RNA sequencing (RNA-seq) data, and metabolome data, have been deposited in the Genome Sequence Archive (GSA) database (https://ngdc.cncb.ac.cn/gsa/) and can be accessed by searching for BioProject PRJCA017539. All raw data used in this study are stored in its subproject Chinese Breast Cancer Genome Atlas and are available upon request from the corresponding author. The I-SPY2-990 Data Resource is publicly available in the NCBI Gene Expression Omnibus (GSE196096; ref. 15). TCGA and METABRIC datasets were collected from cBioPortal (www.cbioportal.org; refs. 35, 36). The NCT02326974 clinical trial data are publicly available in the NCBI Gene Expression Omnibus (GSE243375; ref. 37). The codes used in this study can be obtained from GitHub (https://github.com/Victorli02/FUSCC_HER2). All other raw data generated in this study are available upon request from the corresponding author.

Results

Multiomics profiling of 180 patients with HER2-positive breast cancer

To better uncover the molecular features of HER2-positive breast cancer, we assembled a multiomics HER2-positive breast cancer cohort (**Fig. 1A**; Supplementary Table S1). We obtained breast cancer samples from primary patients treated at FUSCC between 2013 and 2014. HER2-positive status was defined as evidence of HER2 protein overexpression measured by IHC or FISH (23). A total of 191 HER2-positive breast cancer samples were used to conduct multiomics profiling. After quality control, 180 patients with HER2-positive breast cancer were included in this study (**Fig. 1A**). Among these patients, 180 had RNA-seq data, 161 had CNA data, 151 had whole-exome sequencing data on primary tumor tissue and paired blood samples, and 124 had metabolomics data on primary tumor tissue. These patients had a median duration of follow-up of 84.3 months. In addition, 155 patients had detailed anti–HER2-targeted therapy records. Among them, 77.4% ($n = 120$) received adjuvant targeted therapy. This study also provided a multiomics dataset of HER2-positive breast cancer for further analysis.

HER2-positive breast cancer subtypes based on transcriptome data

We applied an unsupervised NMF clustering algorithm to the mRNA data (Supplementary Fig. S1A). Using the top 2,976 differentially expressed coding genes (SD top 15%), we classified 180 HER2-positive breast tumors into four separate subtypes according to the cophenetic correlation coefficient and the dispersion scores (Supplementary Fig. S1A–S1D; Supplementary Table S2A). Consensus clustering also suggested that four clusters were the optimal number of classifications (Supplementary Fig. S1E and S1F).

Through KEGG enrichment analysis and ssGSEA, we explored the population characteristics and enriched differential pathways for each subtype and entitled them accordingly (**Fig. 1B** and **C**; Supplementary Table S2B and S2C). The first subtype enriched carbon metabolism pathways and ERBB2-related pathways and was named the classical (HER2-CLA) subtype (28.3%); the second subtype was characterized by enrichment of multiple immune-related pathways and was named the immunomodulatory (HER2-IM) subtype (20%); the third subtype showed significant enrichment of estrogen receptor (ER)–related pathways and was named the luminal-like (HER2-LUM) subtype (30.6%); and the final subtype enriched the ECM and multiple carcinogenic pathways, with the receptor tyrosine kinase (RTK) pathways being particularly prominent, and was named the basal/mesenchymal-like (HER2-BM) subtype (21.1%).

We also compared our classifications with the clinically used IHC-based receptor subtypes and the PAM50 subtypes (Supplementary Fig. S1G and S1H). We found that hormone receptor (HR)–positive and HER2-positive (HR+ HER2+) accounted for 29.4% in HER2-CLA, 47.2% in HER2-IM, 94.5% in HER2-LUM, and 39.5% in HER2-BM subtypes, respectively (Supplementary Fig. S1G). In the HER2-LUM subtype, the proportion of PAM50 luminal-A and luminal-B subtypes was 70.9%, and 76.5% of all PAM50 luminal-A and luminal-B subtypes being classified as HER2-LUM, respectively (Supplementary Fig. S1H). The HER2-CLA and HER2-IM subtypes were mostly composed of PAM50 HER2-enriched breast cancers (accounting for 94% and 72%, respectively), and 79% of all HER2-enriched breast cancers comprised these two subtypes. The HER2-BM subtype did not exhibit a specific correlation with the PAM50-intrinsic subtypes. However, our findings suggest that HER2-BM constitutes a cohesive and unified subgroup (Supplementary Fig. S1I).

Molecular characteristic landscape of the HER2-positive breast cancer subtypes

To further understand the molecular characteristics of each subtype, we conducted omics profiling and comparative analyses (**Fig. 2**; Supplementary Table S3). The tumor mutational burden of nonsynonymous mutations was significantly higher in the HER2-CLA subtype (1.85 mutations per megabase) than in any other subtypes (**Fig. 2A**; Supplementary Fig. S2A). Heterogeneous distributions of mutational signatures were also observed across different NMF subtypes, in which HER2-CLA and HER2-IM breast cancers harbored a relatively higher proportion of APOBEC-related signatures (SBS2/13; **Fig. 2B**; Supplementary Fig. S2B). The cancer-associated genes most frequently mutated in this cohort included *TP53* (60%), *PIK3CA* (36%), *NF1* (7%), *GATA3* (5%), and *ARID1A* (4%; **Fig. 2C**; Supplementary Table S3B; ref. 38). *PIK3CA* mutations were most prevalent in HER2-LUM (49%; $P = 0.031$), which has been previously reported to be more common in ER+ breast cancer (35, 39). Additionally, *ARID1A* mutations were found only in HER2-CLA (13%; $P = 0.005$). However, the frequency of pathogenic or likely pathogenic germline mutations is relatively low (**Fig. 2D**). For somatic CNAs, the Genomic Identification of Significant Targets in Cancer (GISTIC) peaks were examined (**Fig. 2E**; Supplementary Table S3C and S3D). Differential alterations across the four subtypes included the gain of 6q24.1, 1q32.1, 3q26.32, 6q21,

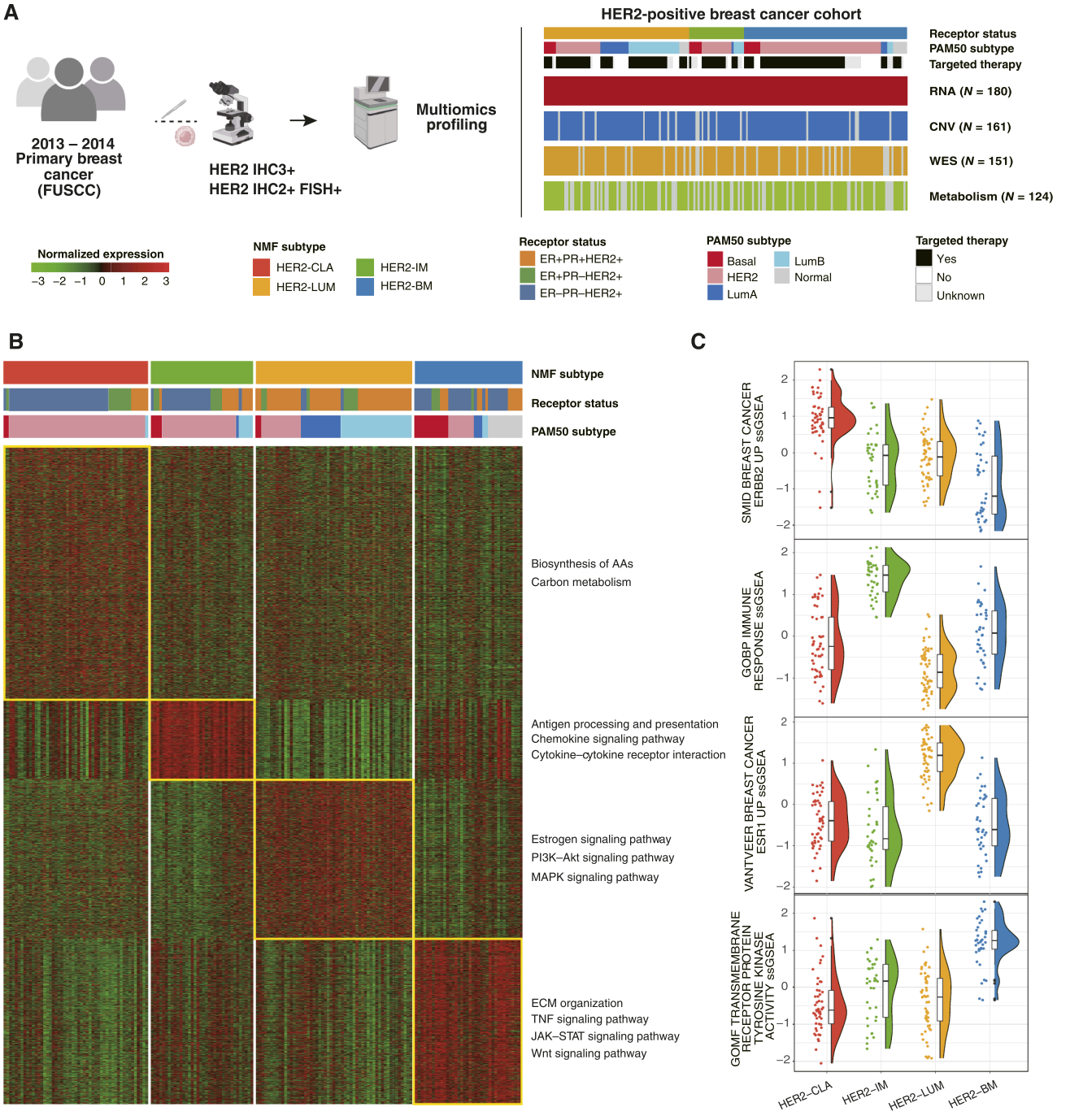


Figure 1.

Cohort information and the NMF subtypes of HER2-positive breast tumors. **A,** Flowchart of study cohort formation (left) and the clinical and omics information (right). **B,** Heatmap of the genes selected for NMF clustering. Annotated KEGG pathways of corresponding signature genes are on the right side. **C,** ssGSEA scores of characteristic pathways. CNV, copy-number variation; LumA, luminal A; LumB, luminal B; WES, whole-exome sequencing.

20q13.2, and 11q13.3 and the loss of 4p14, 9p21.3, 12q24.23, 9p24.1, 8p23.2, and 5p11.2. Notably, 89% of all HER2-positive breast cancer cases exhibited 17q12 gain or amplification, which involves ERBB2, as determined by GISTIC. The amplification frequencies of 20q13.2 (*ZNF217*) and 11q13.3 (*CCND1*) were higher in HER2-LUM (43% and 27%, respectively); however, both of these genes were frequently amplified in ER+ breast cancers, as reported (40, 41).

In terms of metabolomics, we conducted a network analysis to explore subtype-specific polar metabolites (**Fig. 2F**; Supplementary Table S3E). HER2-CLA tumors had a higher abundance of

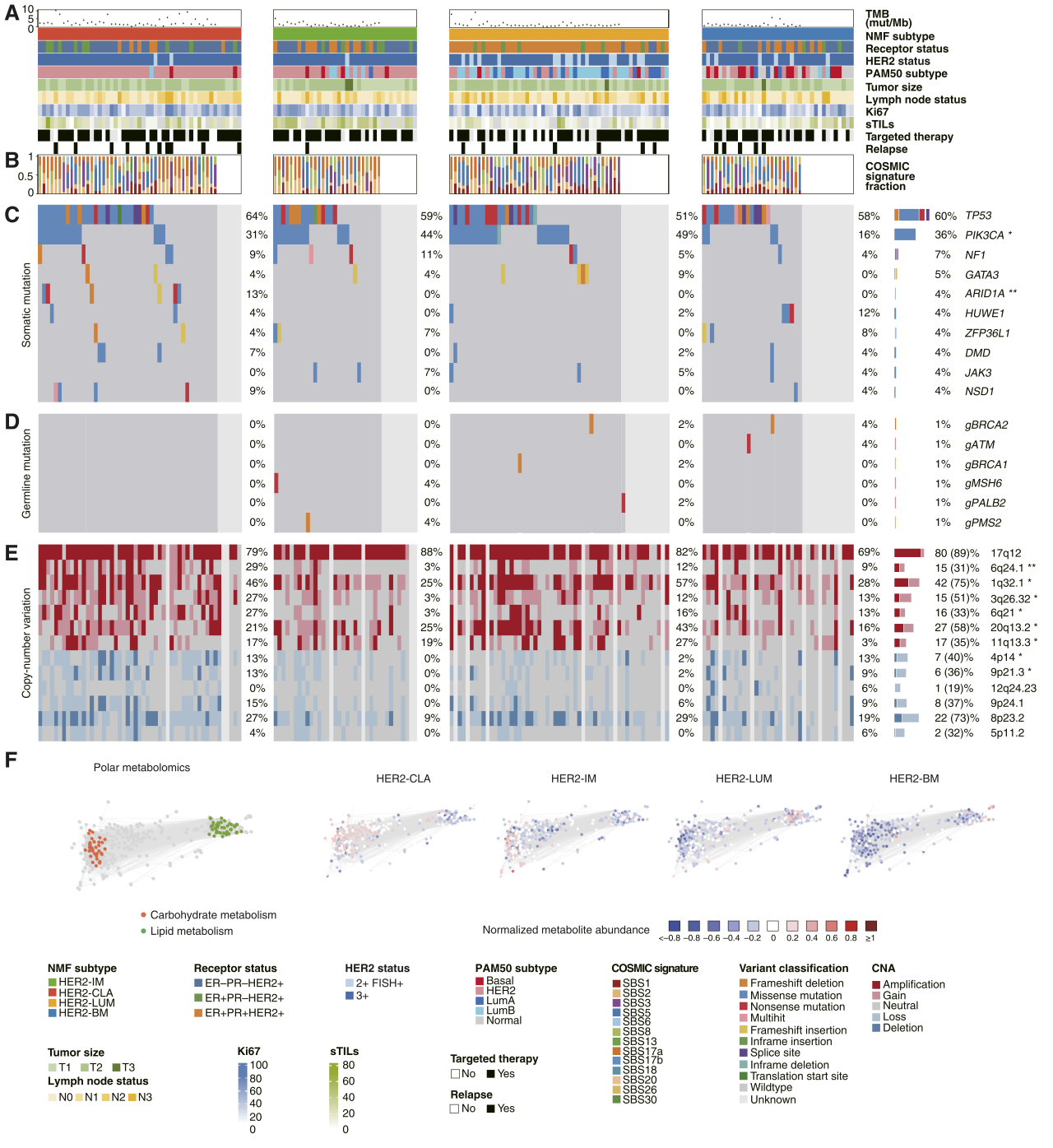


Figure 2.

Multiomics landscape of the NMF subtypes. **A,** Annotated tumor mutational burden, clinical data, and molecular features of HER2-positive breast cancer subtypes. NMF subtypes were used to order these samples. **A–D** show the same order of samples. **B,** Proportion of mutation signatures in each sample of the COSMIC (Catalogue of Somatic Mutations in Cancer) database. **C,** Somatic mutations of the top mutated genes. Genes were ordered by the total mutation frequencies of the cohort. *P* values were calculated from the Fisher exact test among the four subtypes (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$). **D,** Pathogenic or likely pathogenic germline mutations of the top mutated genes. Genes were ordered by the total mutation frequencies of the cohort. **E,** CNAs of significant differentially expressed GISTIC peaks (amplification, gain, neutrality, loss, and deletion were defined as GISTIC +2, +1, 0, −1, and −2, respectively). Differentially expressed GISTIC peaks were calculated by the amplification frequencies among the four subtypes using the Fisher exact test. Percentages represent the amplification frequency. Total frequencies of amplification and gain or loss and deletion peaks are in parentheses on the right. **F,** Differentially expressed metabolite correlation network based on 669 polar metabolites using a > 0.4 Spearman correlation and FDR < 0.05. Correlation networks were partitioned by a graph-clustering algorithm, and the average quantification of NMF subtypes in the correlation networks is presented. LumA, luminal A; LumB, luminal B; mut/Mb, mutations per megabase; sTIL, stromal tumor-infiltrating lymphocytes; TMB, tumor mutational burden.

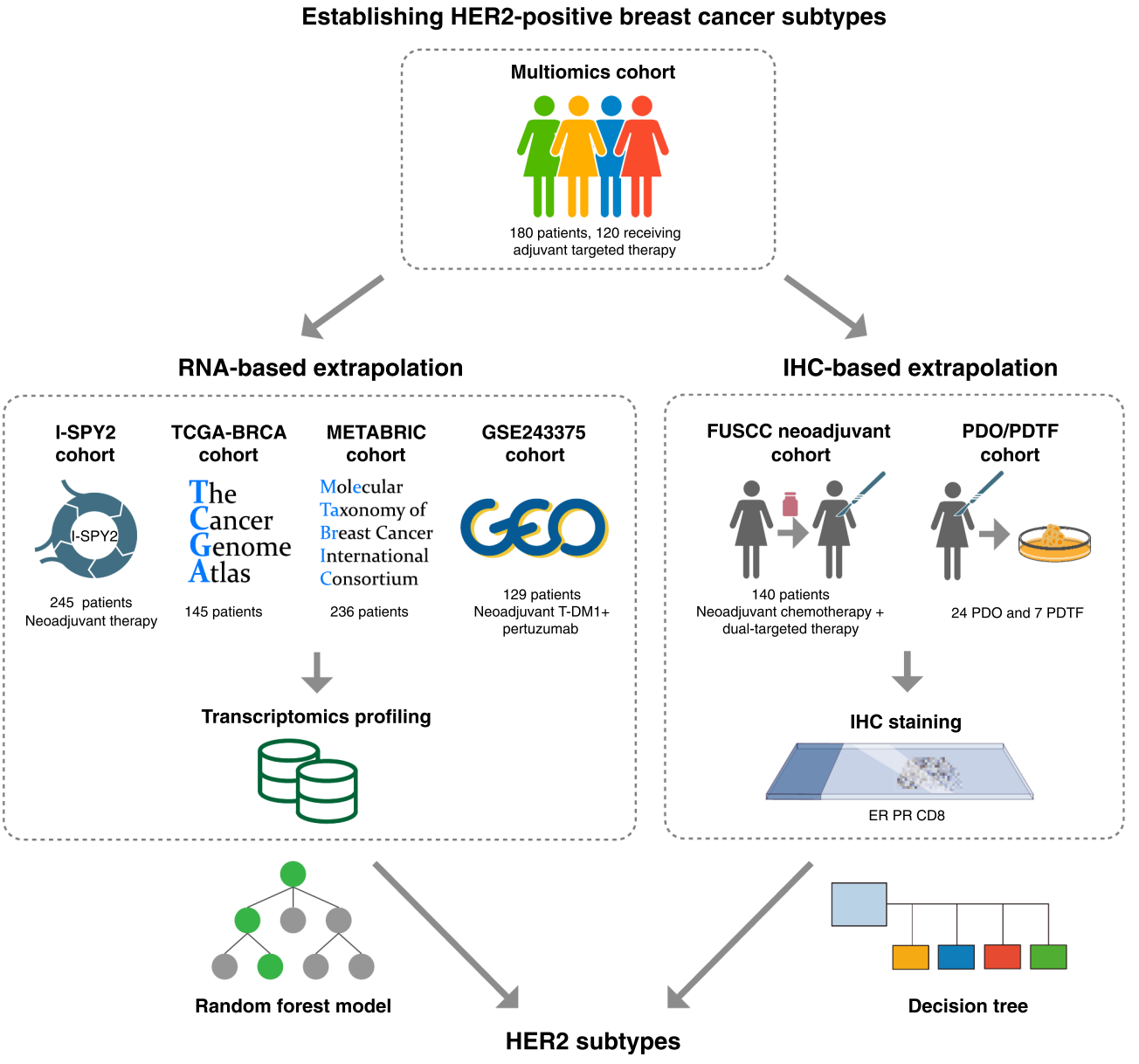


Figure 3.
Overview of the simplified extrapolation method and the use of different cohorts in this study.

carbohydrate metabolism metabolites. Lipid metabolism was relatively activated in the HER2-LUM subtype, which was reported to be associated with *ER* and *PI3K* signaling (42, 43).

Validation of HER2-positive breast cancer subtyping across additional cohorts

To extrapolate our findings to additional HER2-positive breast cancer cohorts, we developed two extrapolation methods based on RNA expression data and IHC staining (**Fig. 3**; Supplementary Fig. S3).

First, we established a random forest model using the subtype-specific differentially expressed genes obtained from the RNA differential analysis of the FUSCC cohort (Supplementary Fig. S3A; ref. 44). With this model, we predicted the NMF subtypes of other HER2-positive breast cancer cohorts that have undergone batch processing of RNA data (Supplementary Fig. S3A). We obtained corresponding samples for the four subtypes in the HER2-positive breast cancer cohorts from the I-SPY2 (15), TCGA (35), METABRIC (36), and GSE243375 (37) databases. The ssGSEA results indicated that the key molecular features of each subtype were validated, further demonstrating the robustness and reliability of NMF subtyping (Supplementary Fig. S3B).

For the convenience of the clinical application of our HER2-positive breast cancer subtypes, we also tried to establish an IHC-based subtyping method. First, we conducted an initial screening of markers for IHC staining by intersecting the differentially expressed RNAs from both the FUSCC and I-SPY2 cohorts (Supplementary Fig. S3C).

Subsequently, potential markers for each subtype were determined through a literature review and integration of characteristic molecular features specific to each subtype (Supplementary Fig. S3C). Finally, we attempted to identify the HER2-LUM subtype by utilizing the crucial gene *ESR1* (ER) in the ER pathway, identified the HER2-BM subtype through the classical basal-like marker *FOXC1* (45, 46), and recognized the HER2-IM subtype by using the classical immune-related indicator *CD8A* (47).

We further collected paraffin-embedded sections of our HER2-positive breast cancer samples for IHC staining and performed receiver operating characteristic (ROC) curve analysis to evaluate the identification efficacy. A total of 100 paraffin-embedded tissue sections were collected for staining and randomly divided into a training set ($n = 66$) and a validation set ($n = 34$; Supplementary Fig. S3D). Based on the protein expression level of ER, the HER2-LUM subtype was effectively identified. The area under the curve (AUC) values were 0.892 in the training cohort and 0.871 in the validation cohort. HER2-BM was identified by using FOXC1 (AUC in the training cohort: 0.818; AUC in the validation cohort: 0.952). As a classical immunomarker, the protein expression of CD8A distinguished this group precisely (AUC in the training cohort: 0.850; AUC in the validation cohort: 0.867). As HER2-CLA was a classic *ERBB2*-driven subtype, we considered using an exclusive method to identify this subgroup. According to the Youden index, we determined the IHC staining cutoff values for ER, FOXC1, and CD8A to be 40%, 5%, and 15%, respectively.

Finally, HER2-positive breast cancers were classified into four subtypes based on the IHC staining scores (Supplementary Fig. S3E). The overall Cohen κ coefficient of the model was 0.616. The percentage of samples classified as the same subtype by both mRNA- and IHC-based classification methods was 72%. The receptor subtype, PAM50 subtypes and molecular features of both mRNA- and IHC-based subtyping methods were generally consistent (Supplementary Fig. S3E). The IHC-based subtyping method was subsequently utilized to expand the sample selection in the validation cohort and PDO and PDTF models.

The *ERBB2*-driven HER2-CLA subtype represents a prioritized population for intensifying anti–HER2-targeted therapies

As the classical driver gene of HER2-positive breast cancer, we first compared the activation of *ERBB2* between four subtypes. At the global level, we initially identified significant enrichment of multiple *ERBB2* signaling pathways in the HER2-CLA subtype through ssGSEA (**Fig. 4A**). Further analysis revealed a substantial increase in both the estimated copy number and mRNA expression of *ERBB2* within the HER2-CLA subtype compared with the other subtypes (**Fig. 4B**). Furthermore, the RNA expression levels of key genes in the downstream RAS–MAPK and Akt–mTOR pathways driven by *ERBB2* in the HER2-CLA subtype were also higher (**Fig. 4B**).

Furthermore, when compared with other HER2-enriched subtype in the PAM50 classification, the expression levels of *ERBB2* and its downstream key genes in the HER2-CLA subtype were significantly higher (Supplementary Fig. S4A). This implies that the HER2-CLA subtype represents a subgroup within HER2-positive breast cancer primarily influenced by *ERBB2*.

Next, we investigated whether the HER2-CLA subtype significantly benefits more from HER2-targeted therapies. Several *ERBB2*-related signatures that predicted the response to HER2-targeted therapy were dramatically upregulated in HER2-CLA patients, shedding light on the potential clinical benefit of anti–HER2-targeted therapies (**Fig. 4C**; ref. 15). We also observed that the HER2-CLA subtype significantly benefited from targeted therapy in our cohort, with improved recurrence-free survival (RFS; HR, 0.12; 95% confidence interval, 0.03–0.48; $P = 0.003$) and overall survival (HR, 0.04; 95% confidence interval, 0.004–0.34; $P = 0.003$; **Fig. 4D**).

As a growing number of HER2-targeted therapies have emerged, we were eager to investigate whether the HER2-CLA subtype can benefit from escalating HER2-targeted treatment strategies. In the I-SPY2 cohort, we found higher pathologic complete response (pCR) rates for the HER2-CLA subtype with dual HER2-targeted therapy compared with single HER2-targeted therapy (**Fig. 4E**). Notably, the pCR rate for the paclitaxel + trastuzumab + pertuzumab arm reached 93%, whereas the pCR rate for the paclitaxel + trastuzumab arm was only 40% (**Fig. 4E**). However, this phenomenon was not observed in the non–HER2-CLA HER2-enriched subgroup (Supplementary Fig. S4B). Collectively, our findings further highlight that *ERBB2* signaling–activated HER2-CLA populations are sensitive to intensifying anti–HER2-targeted regimens.

The immune microenvironment–activated HER2-IM subtype represents a potential population for de-escalation therapy

The biological features of HER2-IM patients were initially explored through gene set enrichment analysis, which revealed significant enrichment of immune activation and immune response pathways (**Fig. 5A**). Additionally, we noted markedly upregulated expression profiles of immunostimulators within HER2-IM cases compared with the other subtypes (**Fig. 5B**). We then conducted ssGSEA to estimate the relative abundance of 24 microenvironment cell types among the patients (27). The HER2-IM subtype exhibited higher abundances of both adaptive and active innate immune cells, which suggested a strong antitumor immune response (**Fig. 5C**).

Furthermore, we evaluated the immune status of the HER2-IM subtype at the pathologic level. A notably increased abundance of estimated stromal tumor-infiltrating lymphocytes was found in the HER2-IM subtype (**Fig. 5D**), and the IHC staining scores of CD8 further clearly verified these findings (**Fig. 5E** and **F**). These results indicated that the immune microenvironment of HER2-IM was in an activated state. We further investigated the therapeutic response of HER2-IM.

We investigated the clinical efficacy of the HER2-IM subtype across different cohorts (**Fig. 3**). In the extra FUSCC HER2-positive breast cancer neoadjuvant therapy cohort, the HER2-IM subtype had the highest pCR rate from standard treatment (chemotherapy and dual-targeted therapy) compared with the other subtypes (**Fig. 5G**; Supplementary Table S4). In the I-SPY2 cohort, our results indicated that the HER2-IM subtype had a relatively higher pCR rate for different treatments (**Fig. 5H**), and the treatment efficacy for HER2-IM was also higher than that for the other subtypes in overall effectiveness (Supplementary Fig. S5A). In the NCT02326974 clinical trial (GSE243375) cohort, the HER2-IM subtype also exhibited higher efficacy in neoadjuvant treatment containing T-DM1 compared with the other subtypes (Supplementary Fig. S5E). Additionally, in TCGA cohort, we observed that HER2-IM patients had better overall survival than other patients with HER2-positive breast cancer (Supplementary Fig. S5F). In our cohort, the HER2-IM subtype demonstrated a favorable prognosis with various treatment strategies (Supplementary Fig. S5G and S5H), suggesting the potential for de-escalation approaches, such as reducing chemotherapy usage or shortening the duration of targeted treatment.

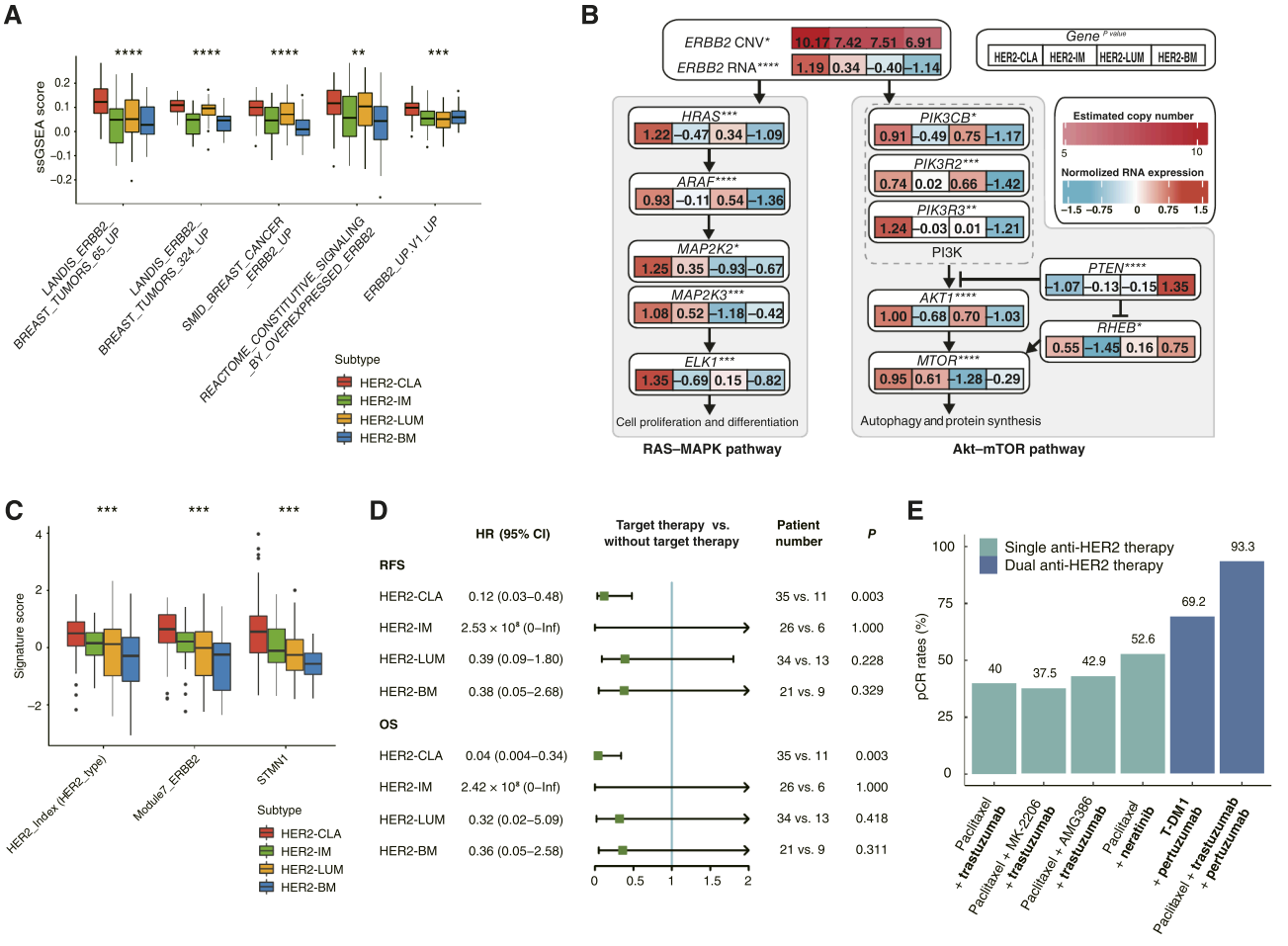


Figure 4.

Biological characteristics and treatment responses of the HER2-CLA subtype. **A–C,** Comparison of ERBB2-related characteristics, incorporating the ssGSEA scores (**A**), the copy-number variation and expression profile among ERBB2 subtypes and critical nodes in the downstream RAS–MAPK and Akt–mTOR pathways (**B**), and ERBB2-related efficacy signatures (**C**) among HER2 subtypes. The Kruskal–Wallis test was conducted to compare the significant differences among different subtypes. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$. **D,** Forest plot comparing the benefits of targeted therapy in different subtypes vs. without targeted therapy. **E,** Comparison of pCR rates in HER2-CLA patients from the I-SPY2 cohort with different therapeutic strategies. CNV, copy-number variation.

We identified an immune-activated HER2-positive breast cancer subgroup that was not distinguishable by the conventional PAM50 classification. Hence, we hypothesized that patients with this subtype could also potentially benefit from immunotherapy. To test this hypothesis, we used PDTFs to analyze the immunologic response of human tumor tissue to *ex vivo* PD1 blockade (Supplementary Table S5; ref. 34). The results of flow cytometry showed that the CD3+ total T-cell activation markers (ICOS and CD25) of the HER2-IM subtype increased more significantly from baseline after immunotherapy than those of the other subtypes ($P = 0.04$ for both; **Fig. 5I** and **J**). Additionally, cytokines (IFNγ and TNFα) were significantly increased in the HER2-IM subtype following PD1 blockade (Supplementary Fig. S5I). In summary, our results revealed the potential application of de-escalation strategies for the immune microenvironment–activated HER2-IM subtype; additionally, this subtype could benefit from immunotherapy.

The HER2-LUM subtype resembles HR+HER2− tumors in molecular features and responds to endocrine therapy and CDK4/6 inhibitors

As the most prevalent subtype (30.6%), HER2-LUM was predominantly composed of HR+HER2+ tumors (Supplementary Fig. S1G). However, compared with non–HER2-LUM HR+HER2+ tumors, the IHC staining scores and RNA expression levels of ER and progesterone receptor (PR) were significantly higher in the HER2-LUM subtype (**Fig. 6A** and **B**). These results indicated that HER2-LUM may represent a subgroup of HR+HER2+ tumors driven by the HR pathway. Furthermore, principal component analysis revealed that the molecular profile of HER2-LUM was more closely related to that of HR+HER2− tumors (**Fig. 6C**).

To determine whether the HER2-LUM subtype could benefit from the treatment strategies used for HR+HER2− tumors, we calculated and compared the sensitivity to endocrine therapy (SET) index (48–50), a verified transcriptomic algorithm for predicting

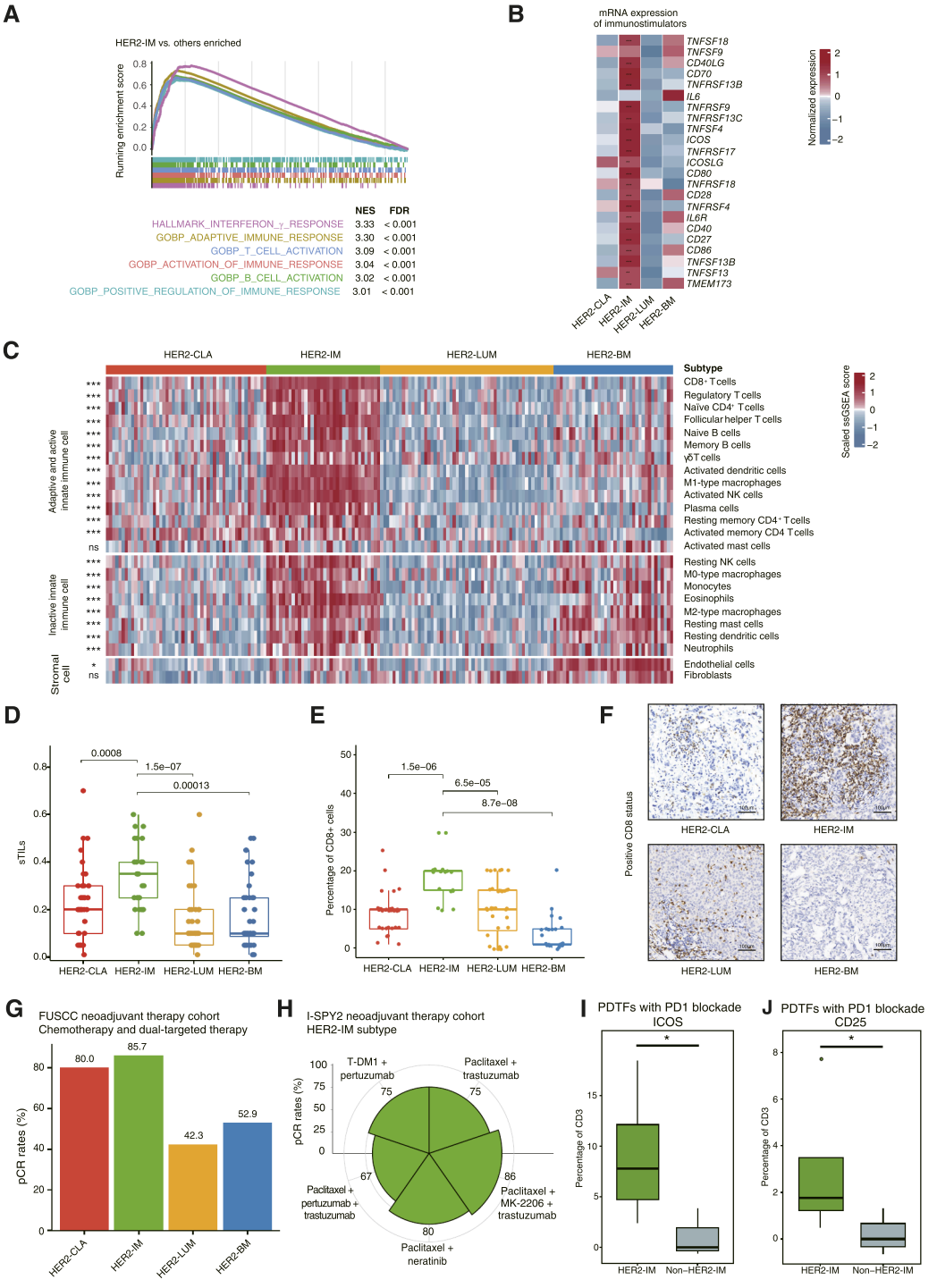


Figure 5.

Biological characteristics and treatment responses of the HER2-IM subtype. **A,** GSEA comparing the HER2-IM subtype with other subtypes. **B** and **C,** Comparison of several immune features utilizing the Kruskal–Wallis test, including the expression profiles of several immunostimulators (**B**) and the estimated abundance of 24 microenvironment cell types (**C**). *P* values were calculated from the Mann–Whitney–Wilcoxon test. *, FDR < 0.05; **, FDR < 0.01; ***, FDR < 0.001. **D,** Estimated abundance of stromal tumor-infiltrating lymphocytes among different subtypes. **E,** Comparison of IHC staining scores of CD8 among different subtypes. **F,** Representative IHC staining results of CD8 on different subtypes. **G,** Comparison of treatment efficacy in different therapeutic regimens for HER2-IM patients in the FUSCC HER2-positive breast cancer neoadjuvant therapy cohort. **H,** Comparison of pCR rates in HER2-IM patients from the I-SPY2 cohort with different therapeutic strategies. **I** and **J,** Δ values of the T-cell activation markers ICOS (**I**) and CD25 (**J**) between the anti-PD1-treated and untreated conditions in tumors with the HER2-IM subtype and the other subtypes. *P* values were calculated by the Mann–Whitney–Wilcoxon test. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$. sTIL, stromal tumor-infiltrating lymphocytes.

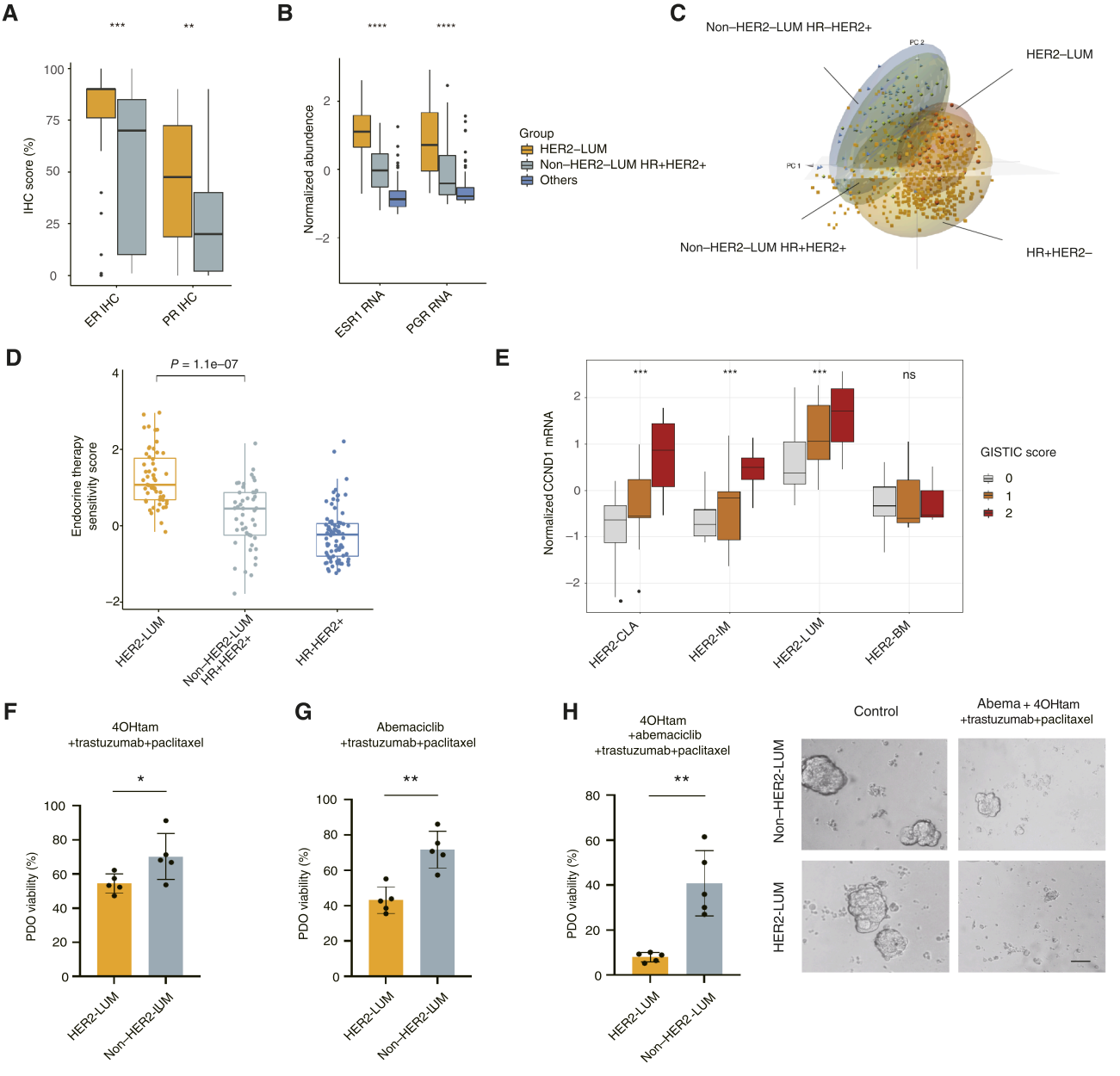


Figure 6.

Biological characteristics and treatment strategies of the HER2-LUM subtype. **A** and **B,** Comparison of IHC staining scores (**A**) and RNA expression levels (**B**) for ER and PR. *P* values were calculated from the Mann–Whitney–Wilcoxon test (**A**) and Kruskal–Wallis test (**B**). **C,** Principal component analysis 3D plots of the relationship between HER2-LUM and other HER2+ (including HR+ and HR−) and HR+HER2− tumors. **D,** Comparison of the endocrine therapy sensitivity scores among HER2-LUM and other tumors. *P* values were calculated from the Mann–Whitney–Wilcoxon test. **E,** mRNA levels of *CCND1* across different GISTIC scores. *P* values were calculated from the Kruskal–Wallis test. **F–H,** Results of the cell viability assay testing the efficacy of 10 μmol/L 4OHtam (**F**), 1 μmol/L abemaciclib (**G**), and their combination (**H**) based on trastuzumab and paclitaxel on PDOs from different subtypes. Scale bar, 100 μm. *P* values are from the Mann–Whitney–Wilcoxon test. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$.

sensitivity to endocrine therapy. We found that HER2-LUM exhibits greater sensitivity to endocrine therapy, suggesting that this subtype was more likely to benefit from endocrine treatment (**Fig. 6D**). Moreover, under different GISTIC score statuses, the RNA expression levels of *CCND1* in HER2-LUM were higher than those of the other subtypes, which implied a potential benefit from CDK4/6 inhibitor treatment (**Fig. 6E**; ref. 51).

In the neoadjuvant therapy cohort from FUSCC, we discovered that in comparison with HER2-CLA (pCR rate = 79.0%) and HER2-IM (pCR rate = 83.3%), HER2-LUM (pCR rate = 41.3%) exhibited unsatisfactory response to standard treatment (chemotherapy and dual-targeted therapy; **Fig. 5G**). The pCR rates of HER2-LUM in different therapeutic arms from the I-SPY2 and GSE243375 cohorts were limited (Supplementary Fig. S5C and S5E).

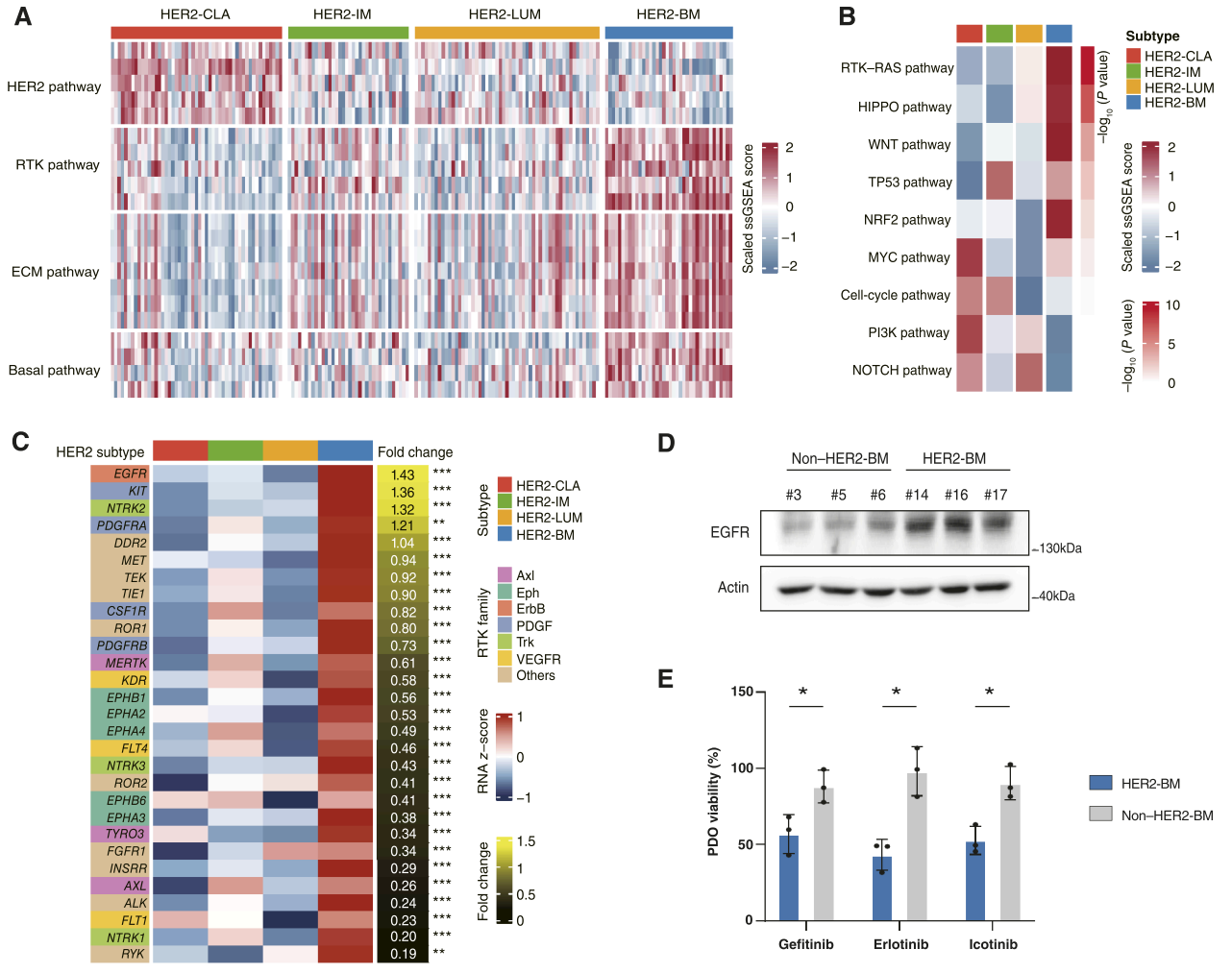


Figure 7.

Biological characteristics and treatment strategies of the HER2-BM subtype. **A,** ssGSEA scores of HER2-BM subtype characteristic pathways. **B,** ssGSEA scores of key oncogenic signaling pathways among the four subtypes. **C,** RNA expression levels of key RTK genes in the HER2-BM subtype. **D,** Western blot images showing the expression of *EGFR* in HER2-BM tumors or non–HER2-BM tumors. **E,** Results of the cell viability assay testing the efficacy of gefitinib (10 μmol/L), erlotinib (5 μmol/L), and icotinib (10 μmol/L) based on trastuzumab and paclitaxel on PDOs from different subtypes. *P* values are from the Mann–Whitney–Wilcoxon test. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

Therefore, further validation of intensified treatment strategies based on PDOs for HER2-LUM was necessary.

We compared the inhibitory effects of 4-hydroxytamoxifen (4OHtam), CDK4/6 inhibitors, and their combination in addition to chemotherapy and targeted therapy on HER2-LUM and non–HER2-LUM PDOs (**Fig. 6F–H**; Supplementary Table S5). We found that HER2-LUM subtype PDOs exhibited greater sensitivity to the combination of 4OHtam or CDK4/6 inhibitors with chemotherapy and targeted therapy (**Fig. 6F** and **G**). Furthermore, the combination of 4OHtam and CDK4/6 inhibitors further enhanced the inhibition of HER2-LUM subtype PDOs (**Fig. 6H**). After treatment with 4OHtam and CDK4/6 inhibitors, we observed a significant downregulation of RNA expression levels of genes related to endocrine therapy and the cell-cycle pathway in the HER2-LUM subtype (Supplementary Fig. S6). In conclusion, we identified a subgroup within the HER2-positive breast cancer population with molecular characteristics more akin to HR+HER2− breast cancer. Through analysis and experimental validation, we demonstrated that this subgroup exhibits enhanced sensitivity to additional endocrine and CDK4/6 inhibitor treatments.

The HER2-BM subtype represents a unique subgroup with activated RTK signaling in HER2-positive breast cancer

Although HER2-BM seems heterogeneous from previous receptor- and PAM50-based subtyping methods, our analysis results suggested that this subtype represents a cohesive and distinct group (Supplementary Fig. S1D and S1I). In the FUSCC neoadjuvant therapy cohort, the standard treatment of the HER2-BM subtype demonstrated unsatisfactory efficacy (pCR rate = 50.0%) in contrast to HER2-CLA (pCR rate = 79.0%) and HER2-IM (pCR rate = 83.3%; **Fig. 5G**). In addition, in the I-SPY2 and GSE243375 cohorts,

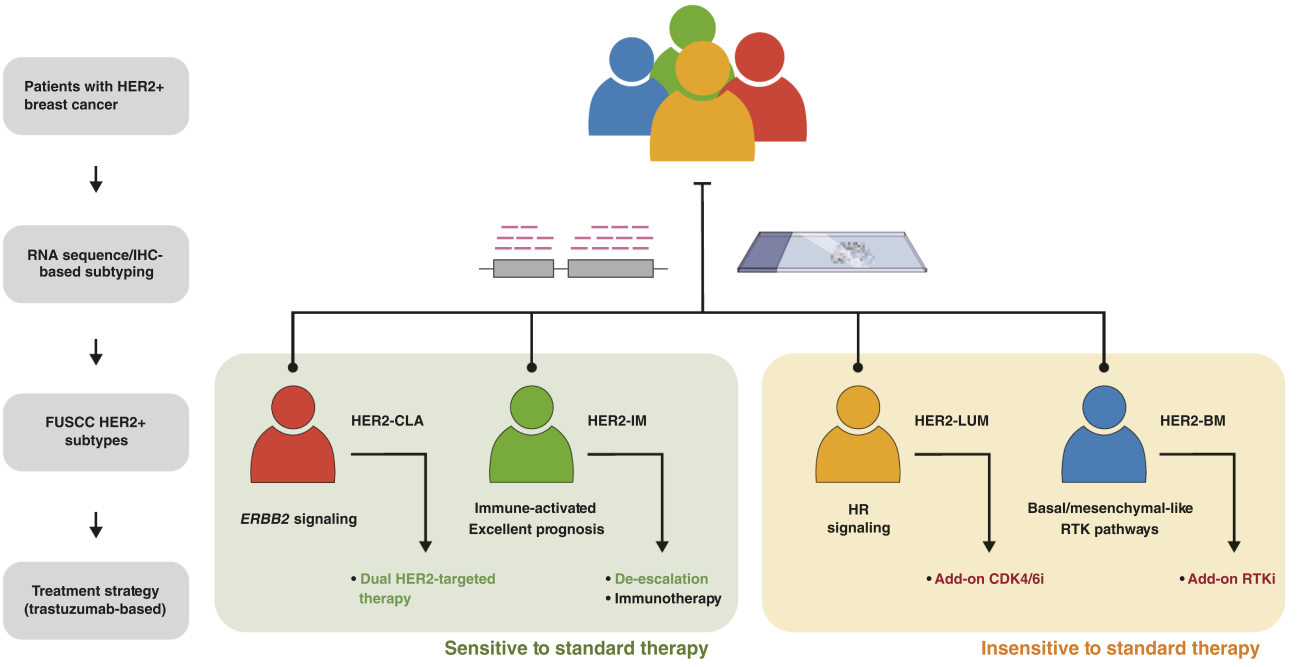


Figure 8.

Schematic diagram of FUSCC HER2+ subtypes and precision treatment for HER2+ breast cancer. A schematic diagram showing the overview of FUSCC HER2+ subtypes, including classified methods, molecular characteristics, and treatment strategies. The rectangles with green and orange background indicate the patient sensitivity and insensitivity to standard trastuzumab-based therapy, respectively.

the efficacy of different treatment arms was also limited (Supplementary Fig. S5D and S5E).

Hence, we proceeded to delve into the molecular characteristics and potential targets of this subtype. Using RNA-seq data, we performed GSEA and discovered that the RTK pathway was significantly enriched in HER2-BM patients (Supplementary Fig. S7A and S7B). Further analysis revealed that HER2-BM exhibited RTK pathway activation beyond *ERBB2* signaling, alongside the activation of the ECM and basal-like tumor-associated pathways (**Fig. 7A**). We also compared the ssGSEA scores of key oncogenic signaling pathways among the four subtypes (52, 53). The enrichment of the RTK–RAS pathway was most prominent in HER2-BM (**Fig. 7B**). Therefore, we further explored specific RTK activation in the HER2-BM subtype. By comparing the expression levels of key genes from different RTK subfamilies (Supplementary Fig. S7C), we identified the significantly upregulated key RTKs in the HER2-BM subtype (**Fig. 7C**). We found that *EGFR* was the most differentially expressed RTK gene in the HER2-BM subtype compared with the other subtypes (**Fig. 7C**). Through Western blotting, we validated that *EGFR* was highly expressed in HER2-BM tumor samples (**Fig. 7D**).

Additionally, in PDO model experiments, we found that compared with the other subtypes, the HER2-BM subtype demonstrated increased sensitivity to *EGFR*-targeting inhibitors gefitinib ($P = 0.03$), erlotinib ($P = 0.01$), and icotinib ($P = 0.01$; **Fig. 7E**; Supplementary Table S5). After treatment with the EGFR inhibitor, we observed a significant downregulation of EGFR-related pathway genes in the HER2-BM subtype (Supplementary Fig. S7D). In conclusion, our subtyping identified a distinct subgroup within HER2-positive breast cancer that exhibits a poor response to current standard treatment while displaying characteristic RTK pathways activation. This suggests the potential therapeutic approach of using tyrosine kinase inhibitors (TKI) that encompass *EGFR* targeting.

Discussion

In this study, we established a new HER2-positive breast cancer cohort with multiple omics dimensions. This dataset allowed us to distinguish four transcriptome-based molecular subtypes of HER2-positive breast cancer that reflect its molecular diversity. These four subtypes provide an important basis for us to tailor more precise treatments according to their biological features and treatment responses. Furthermore, we investigated the association between subtypes and optimal treatment strategies in multiple cohorts and PDO models (**Fig. 8**).

The anti-HER2 therapies have notably improved the outcomes of HER2-positive breast cancer, but molecular stratification is pivotal for further optimizing the existing treatments. Treatment strategies tailored to the intertumoral heterogeneity can enhance efficacy and reduce toxicity (54, 55). Some studies have tried to discover the subgroups of HER2-positive breast cancer, but the clinical guidance offered remained limited (9, 15). A more comprehensive understanding of the molecular heterogeneity of HER2-positive breast cancer is needed to optimize its precision treatment strategy. We established four subtypes to uncover the molecular diversity. The HER2-CLA subtype represents an identified group that is significantly characterized by high *ERBB2* activation and significantly benefits from HER2-targeted therapy. The HER2-IM subtype is an immune microenvironment–activated group and may be a candidate for de-escalation strategies (56, 57). The HER2-LUM subtype is characterized by similar molecular features of the HR+HER2− breast cancer, thereby suggesting

potential treatment strategies involving endocrine therapy and CDK4/6 inhibitors. The HER2-BM subtype represents a unique subgroup characterized by both basal-like tumor features and an enrichment of mesenchymal components. Furthermore, the activation of RTK pathways suggests the combination therapy with TKIs, especially with TKIs showing a stronger inclination for *EGFR* inhibition.

The PAM50 classification is an important tool for identifying breast cancer molecular subtypes to guide treatment decisions. However, there remains room for optimization in the application of the PAM50 classification method within HER2-positive breast cancer (58). Here, our classification is centered on HER2-positive breast cancer and could offer insights beyond what PAM50 classification can provide. For example, HER2-CLA and HER2-IM subtypes contribute to the refinement of the conventional PAM50 HER2-enriched subtype, exhibiting relative favorable responses to standard HER2-targeted therapies. However, these two subtypes differ in their molecular characteristics and clinical management strategies. The precision treatment of the HER2-CLA subtype relies upon more potent anti–HER2-targeted therapies. The previously unrecognized HER2-IM subtype demonstrates favorable therapeutic efficacy possibly linked to immune microenvironment activation (10, 56). Additionally, findings from the PDTF experiments suggest that the HER2-IM subtype might be suitable for treatment strategies incorporating immunotherapy in HER2-positive breast cancer (59, 60). The HER2-BM subtype is an undefined subgroup with poor treatment response to existing dual HER2-targeted therapies. We observed heightened sensitivity of this subtype to *EGFR* inhibitors, implying that clinical treatment strategies could favor the selection of TKI drugs with a stronger propensity for targeting *EGFR*. Simultaneously, the HER2-BM subtype can consider treatment regimens involving the combination of TKIs (61). Additionally, we found limited responses of the HER2-LUM and HER2-BM subtypes to treatment regimens containing novel antibody–drug conjugates. Collectively, our study could help better inform clinical management for individual patients with HER2-positive breast cancer, which might hopefully increase treatment efficacy.

Nevertheless, this study had some limitations. The existing conclusion about therapeutic benefits still requires further data for confirmation. The potential therapeutic strategies and specific drug combinations proposed in this study require further validation and refinement through more prospective clinical evidence, particularly in the form of randomized controlled trials that encompass corresponding precision treatment regimens for the four subtypes. Additionally, due to certain limitations in experimental models derived from patient samples, the specific biological mechanisms of certain subtypes remain to be explored in depth, which would enable the formulation of more precise treatment approaches based on these findings.

In conclusion, we performed a comprehensive multiomics analysis of HER2-positive breast cancer and revealed its molecular subtypes and biological diversity. We proposed more tailored treatment strategies for each subtype, which may improve efficacy and reduce overtreatment and extra costs for patients with specific subtypes of HER2-positive breast cancer.

Authors' Disclosures

No disclosures were reported.

Authors' Contributions

Y.-W. Li: Software, formal analysis, validation, writing–original draft. **L.-J. Dai:** Software, formal analysis, validation. **X.-R. Wu:** Data curation, software, formal analysis, validation. **S. Zhao:** Software, validation. **Y.-Z. Xu:** Resources, data curation, software, formal analysis. **X. Jin:** Resources, data curation. **Y. Xiao:** Resources, data curation, software. **Y. Wang:** Resources, software. **C.-J. Lin:** Resources, data curation, software, formal analysis. **Y.-F. Zhou:** Resources, data curation, software. **T. Fu:** Resources, data curation, software. **W.-T. Yang:** Resources, validation. **M. Li:** Resources, software. **H. Lv:** Resources, software. **S. Chen:** Software, methodology. **A. Grigoriadis:** Software, methodology. **Y.-Z. Jiang:** Resources, data curation, supervision, writing–review and editing. **D. Ma:** Resources, data curation, software, supervision, writing–review and editing. **Z.-M. Shao:** Resources, supervision.

Acknowledgments

This work was supported by grants from the National Key Research and Development Project of China (2020YFA0112304 and 2021YFF1201300), the National Natural Science Foundation of China (92159301, 82341003, 82272822, and 82272704), the Program of Shanghai Academic/Technology Research Leader (20XD1421100), the Shanghai Key Laboratory of Breast Cancer (12DZ2260100), the Natural Science Foundation of Shanghai (22ZR1479200), the Shanghai Rising-Star Program (23QA1401400), the Chen Guang Project supported by the Shanghai Municipal Education Commission and Shanghai Education Development Foundation (21CGA07), and the Medical Engineering Joint Fund of Fudan University. We also appreciate the constructive suggestions on experimental design provided by Li-Tong Yao, Fan Yang, Yun-Song Yang, Tian-Jian Yu, and Xiu-Zhi Zhu.

Note

Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

Received December 21, 2023; revised April 19, 2024; accepted August 8, 2024; published first August 26, 2024.

References

1. Loibl S, Gianni L. HER2-positive breast cancer. Lancet 2017;389:2415–29.
2. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science 1987;235:177–82.
3. Oh D-Y, Bang Y-J. HER2-targeted therapies—a role beyond breast cancer. Nat Rev Clin Oncol 2020;17:33–48.
4. Sharma P. Major strides in HER2 blockade for metastatic breast cancer. N Engl J Med 2020;382:669–71.
5. Pernas S, Tolaney SM. HER2-positive breast cancer: new therapeutic frontiers and overcoming resistance. Ther Adv Med Oncol 2019;11:1758835919833519.
6. Hayes DF. HER2 and breast cancer—a phenomenal success story. N Engl J Med 2019;381:1284–6.
7. Brasó-Maristany F, Griguolo G, Pascual T, Paré L, Nuciforo P, Llombart-Cussac A, et al. Phenotypic changes of HER2-positive breast cancer during and after dual HER2 blockade. Nat Commun 2020;11:385.
8. McNamara KL, Caswell-Jin JL, Joshi R, Ma Z, Kotler E, Bean GR, et al. Spatial proteomic characterization of HER2-positive breast tumors through neoadjuvant therapy predicts response. Nat Cancer 2021;2:400–13.
9. Ferrari A, Vincent-Salomon A, Pivot X, Sertier AS, Thomas E, Tonon L, et al. A whole-genome sequence and transcriptome perspective on HER2-positive breast cancers. Nat Commun 2016;7:12222.
10. Fumagalli D, Venet D, Ignatiadis M, Azim HA Jr, Maetens M, Rothé F, et al. RNA sequencing to predict response to neoadjuvant anti-HER2 therapy: a secondary analysis of the NeoALTTO randomized clinical trial. JAMA Oncol 2017;3:227–34.

11. Lesurf R, Griffith OL, Griffith M, Hundal J, Trani L, Watson MA, et al. Genomic characterization of HER2-positive breast cancer and response to neoadjuvant trastuzumab and chemotherapy-results from the ACOSOG Z1041 (Alliance) trial. Ann Oncol 2017;28:1070–7.
12. Prat A, Guarneri V, Paré L, Griguolo G, Pascual T, Dieci MV, et al. A multivariable prognostic score to guide systemic therapy in early-stage HER2-positive breast cancer: a retrospective study with an external evaluation. Lancet Oncol 2020;21:1455–64.
13. Villacampa G, Tung NM, Pernas S, Paré L, Bueno-Muiño C, Echavarría I, et al. Association of HER2DX with pathological complete response and survival outcomes in HER2-positive breast cancer. Ann Oncol 2023;34:783–95.
14. Tolaney SM, Tarantino P, Graham N, Tayob N, Parè L, Villacampa G, et al. Adjuvant paclitaxel and trastuzumab for node-negative, HER2-positive breast cancer: final 10-year analysis of the open-label, single-arm, phase 2 APT trial. Lancet Oncol 2023;24:273–85.
15. Wolf DM, Yau C, Wulfkuhle J, Brown-Swigart L, Gallagher RI, Lee PRE, et al. Redefining breast cancer subtypes to guide treatment prioritization and maximize response: predictive biomarkers across 10 cancer therapies. Cancer Cell 2022;40:609–23.e6.
16. Harbeck N. Neoadjuvant and adjuvant treatment of patients with HER2-positive early breast cancer. Breast 2022;62(Suppl 1):S12–6.
17. Nitz UA, Gluz O, Christgen M, Grischke EM, Augustin D, Kuemmel S, et al. De-escalation strategies in HER2-positive early breast cancer (EBC): final analysis of the WSG-ADAPT HER2+/HR− phase II trial: efficacy, safety, and predictive markers for 12 weeks of neoadjuvant dual blockade with trastuzumab and pertuzumab ± weekly paclitaxel. Ann Oncol 2017;28:2768–72.
18. Tolaney SM, Barry WT, Dang CT, Yardley DA, Moy B, Marcom PK, et al. Adjuvant paclitaxel and trastuzumab for node-negative, HER2-positive breast cancer. N Engl J Med 2015;372:134–41.
19. Pivot X, Romieu G, Debled M, Pierga J-Y, Kerbrat P, Bachelot T, et al. 6 months versus 12 months of adjuvant trastuzumab in early breast cancer (PHARE): final analysis of a multicentre, open-label, phase 3 randomised trial. Lancet 2019;393:2591–8.
20. von Minckwitz G, Procter M, de Azambuja E, Zardavas D, Benyunes M, Viale G, et al. Adjuvant pertuzumab and trastuzumab in early HER2-positive breast cancer. N Engl J Med 2017;377:122–31.
21. Martin M, Holmes FA, Ejlertsen B, Delaloge S, Moy B, Iwata H, et al. Neratinib after trastuzumab-based adjuvant therapy in HER2-positive breast cancer (ExteNET): 5-year analysis of a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet Oncol 2017;18:1688–700.
22. von Minckwitz G, Huang C-S, Mano MS, Loibl S, Mamounas EP, Untch M, et al. Trastuzumab emtansine for residual invasive HER2-positive breast cancer. N Engl J Med 2019;380:617–28.
23. Wolff AC, Hammond MEH, Allison KH, Harvey BE, Mangu PB, Bartlett JMS, et al. Human epidermal growth factor receptor 2 testing in breast cancer: American society of clinical oncology/college of American pathologists clinical practice guideline focused update. J Clin Oncol 2018;36:2105–22.
24. Sra S, Dhillon IS. Generalized nonnegative matrix approximations with Bregman divergences. In: Advances in neural information processing systems; 2005;18. pp. 283–290.
25. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics 2010;26:1572–3.
26. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics 2013;14:7.
27. Xiao Y, Ma D, Zhao S, Suo C, Shi J, Xue MZ, et al. Multi-omics profiling reveals distinct microenvironment characterization and suggests immune escape mechanisms of triple-negative breast cancer. Clin Cancer Res 2019;25:5002–14.
28. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 2012;16:284–7.
29. Liaw A, Wiener M. Classification and regression by randomForest. R News 2002;2:18–22.
30. Chen D, Zhang Y, Wang W, Chen H, Ling T, Yang R, et al. Identification and characterization of robust hepatocellular carcinoma prognostic subtypes based on an integrative metabolite-protein interaction network. Adv Sci (Weinh) 2021;8:e2100311.
31. Chen Y-J, Roumeliotis TI, Chang Y-H, Chen C-T, Han C-L, Lin M-H, et al. Proteogenomics of non-smoking lung cancer in east Asia delineates molecular signatures of pathogenesis and progression. Cell 2020;182:226–44.e17.
32. Gong Y, Ji P, Yang Y-S, Xie S, Yu T-J, Xiao Y, et al. Metabolic-pathway-based subtyping of triple-negative breast cancer reveals potential therapeutic targets. Cell Metab 2021;33:51–64.e9.
33. Yang F, Xiao Y, Ding J-H, Jin X, Ma D, Li D-Q, et al. Ferroptosis heterogeneity in triple-negative breast cancer reveals an innovative immunotherapy combination strategy. Cell Metab 2023;35:84–100.e8.
34. Voabil P, de Bruijn M, Roelofsen LM, Hendriks SH, Brokamp S, van den Braber M, et al. An ex vivo tumor fragment platform to dissect response to PD-1 blockade in cancer. Nat Med 2021;27:1250–61.
35. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature 2012;490:61–70.
36. Curtis C, Shah SP, Chin S-F, Turashvili G, Rueda OM, Dunning MJ, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature 2012;486:346–52.
37. Li Z, Metzger Filho O, Viale G, dell'Orto P, Russo L, Goyette M-A, et al. HER2 heterogeneity and treatment response-associated profiles in HER2-positive breast cancer in the NCT02326974 clinical trial. J Clin Invest 2024;134:e176454.
38. Gillette MA, Satpathy S, Cao S, Dhanasekaran SM, Vasaikar SV, Krug K, et al. Proteogenomic characterization reveals therapeutic vulnerabilities in lung adenocarcinoma. Cell 2020;182:200–25.e35.
39. Toska E, Osmanbeyoglu HU, Castel P, Chan C, Hendrickson RC, Elkabets M, et al. PI3K pathway regulates ER-dependent transcription in breast cancer through the epigenetic regulator KMT2D. Science 2017;355:1324–30.
40. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA Jr, Kinzler KW. Cancer genome landscapes. Science 2013;339:1546–58.
41. Hsu P-Y, Hsu H-K, Lan X, Juan L, Yan PS, Labanowska J, et al. Amplification of distant estrogen response elements deregulates target genes associated with tamoxifen resistance in breast cancer. Cancer Cell 2013;24:197–212.
42. Kulkoyluoglu-Cotul E, Arca A, Madak-Erdogan Z. Crosstalk between estrogen signaling and breast cancer metabolism. Trends Endocrinol Metab 2019;30:25–38.
43. Martínez-Reyes I, Chandel NS. Cancer metabolism: looking forward. Nat Rev Cancer 2021;21:669–80.
44. Breiman L. Random forests. Machine Learn 2001;45:5–32.
45. Ray PS, Wang J, Qu Y, Sim M-S, Shamonki J, Bagaria SP, et al. FOXC1 is a potential prognostic biomarker with functional significance in basal-like breast cancer. Cancer Res 2010;70:3870–6.
46. Han B, Bhowmick N, Qu Y, Chung S, Giuliano AE, Cui X. FOXC1: an emerging marker and therapeutic target for cancer. Oncogene 2017;36:3957–63.
47. Raskov H, Orhan A, Christensen JP, Gögenur I. Cytotoxic $CD8^+$ T cells in cancer and cancer immunotherapy. Br J Cancer 2021;124:359–67.
48. Du L, Yau C, Brown-Swigart L, Gould R, Krings G, Hirst GL, et al. Predicted sensitivity to endocrine therapy for stage II-III hormone receptor-positive and HER2-negative (HR+/HER2−) breast cancer before chemo-endocrine therapy. Ann Oncol 2021;32:642–51.
49. Speers CW, Symmans WF, Barlow WE, Trevarton A, The S, Du L, et al. Evaluation of the sensitivity to endocrine therapy index and 21-gene breast recurrence score in the SWOG S8814 trial. J Clin Oncol 2023;41:1841–8.
50. Sinn BV, Fu C, Lau R, Litton J, Tsai T-H, Murthy R, et al. $SET_{ER/PR}$: a robust 18-gene predictor for sensitivity to endocrine therapy for metastatic breast cancer. NPJ Breast Cancer 2019;5:16.
51. Finn RS, Crown JP, Lang I, Boer K, Bondarenko IM, Kulyk SO, et al. The cyclin-dependent kinase 4/6 inhibitor palbociclib in combination with letrozole versus letrozole alone as first-line treatment of oestrogen receptor-positive, HER2-negative, advanced breast cancer (PALOMA-1/TRIO-18): a randomised phase 2 study. Lancet Oncol 2015;16:25–35.
52. Sanchez-Vega F, Mina M, Armenia J, Chatila WK, Luna A, La KC, et al. Oncogenic signaling pathways in the cancer genome Atlas. Cell 2018;173:321–37.e10.
53. Ding L, Bailey MH, Porta-Pardo E, Thorsson V, Colaprico A, Bertrand D, et al. Perspective on oncogenic processes at the end of the beginning of cancer genomics. Cell 2018;173:305–20.e10.
54. Carey LA, Berry DA, Cirrincione CT, Barry WT, Pitcher BN, Harris LN, et al. Molecular heterogeneity and response to neoadjuvant human epidermal growth factor receptor 2 targeting in CALGB 40601, a randomized phase III trial of paclitaxel plus trastuzumab with or without lapatinib. J Clin Oncol 2016;34:542–9.
55. Luen SJ, Salgado R, Fox S, Savas P, Eng-Wong J, Clark E, et al. Tumour-infiltrating lymphocytes in advanced HER2-positive breast cancer treated with

pertuzumab or placebo in addition to trastuzumab and docetaxel: a retrospective analysis of the CLEOPATRA study. Lancet Oncol 2017;18:52–62.
56. Fernandez-Martinez A, Pascual T, Singh B, Nuciforo P, Rashid NU, Ballman KV, et al. Prognostic and predictive value of immune-related gene expression signatures vs tumor-infiltrating lymphocytes in early-stage ERBB2/HER2-positive breast cancer: a correlative analysis of the CALGB 40601 and PAMELA trials. JAMA Oncol 2023;9:490–9.
57. Solinas C, Ceppi M, Lambertini M, Scartozzi M, Buisseret L, Garaud S, et al. Tumor-infiltrating lymphocytes in patients with HER2-positive breast cancer treated with neoadjuvant chemotherapy plus trastuzumab, lapatinib or their combination: a meta-analysis of randomized controlled trials. Cancer Treat Rev 2017;57:8–15.
58. Krug K, Jaehnig EJ, Satpathy S, Blumenberg L, Karpova A, Anurag M, et al. Proteogenomic landscape of breast cancer tumorigenesis and targeted therapy. Cell 2020;183:1436–56.e31.
59. Huober J, Barrios CH, Niikura N, Jarząb M, Chang Y-C, Huggins-Puhalla S-L, et al. Atezolizumab with neoadjuvant anti-human epidermal growth factor receptor 2 therapy and chemotherapy in human epidermal growth factor receptor 2-positive early breast cancer: primary results of the randomized phase III IMpassion050 trial. J Clin Oncol 2022;40:2946–56.
60. Ahn HK, Sim SH, Suh KJ, Kim MH, Jeong JH, Kim J-Y, et al. Response rate and safety of a neoadjuvant pertuzumab, atezolizumab, docetaxel, and trastuzumab regimen for patients with ERBB2-positive stage II/III breast cancer: the Neo-PATH phase 2 nonrandomized clinical trial. JAMA Oncol 2022;8:1271–7.
61. Wu J, Jiang Z, Liu Z, Yang B, Yang H, Tang J, et al. Neoadjuvant pyrotinib, trastuzumab, and docetaxel for HER2-positive breast cancer (PHEDRA): a double-blind, randomized phase 3 trial. BMC Med 2022;20:498.