

# NEW APPROACHES TO CLASSIFICATION AND DIAGNOSTIC PREDICTION OF BREAST CANCERS

EDITED BY: Aleix Prat and Mothaffar Rimawi

PUBLISHED IN: Frontiers in Oncology and Frontiers in Genetics





# frontiers

## Frontiers eBook Copyright Statement

The copyright in the text of individual articles in this eBook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this eBook is the property of Frontiers.

Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence.

The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or eBook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714

ISBN 978-2-88963-796-6

DOI 10.3389/978-2-88963-796-6

## About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

## Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

## Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: [researchtopics@frontiersin.org](mailto:researchtopics@frontiersin.org)

# NEW APPROACHES TO CLASSIFICATION AND DIAGNOSTIC PREDICTION OF BREAST CANCERS

Topic Editors:

**Aleix Prat**, Hospital Clínic de Barcelona, Spain

**Mothaffar Rimawi**, Baylor College of Medicine, United States

Despite many years of translational research in breast cancer, very few new biomarkers have been implemented for clinical use beyond estrogen receptor, progesterone receptor, and HER2. The main reason is that many promising biomarkers are clinically validated but lack analytical and clinical utility. One explanation is that proper validation of the predictive ability of the biomarker in independent datasets, and with a pre-planned statistical analysis, is not always performed. Thus, there is a need to identify new biomarkers or new ways to subclassify breast cancer patients that are reproducible and easy to implement in the clinical setting but, more importantly, that improve patient's outcomes.

**Citation:** Prat, A., Rimawi, M., eds. (2020). New Approaches to Classification and Diagnostic Prediction of Breast Cancers. Lausanne: Frontiers Media SA.  
doi: 10.3389/978-2-88963-796-6

# Table of Contents

- 06    *The C Allele of ATM rs11212617 Associates With Higher Pathological Complete Remission Rate in Breast Cancer Patients Treated With Neoadjuvant Metformin***  
Elisabet Cuyàs, Maria Buxó, Maria José Ferri Iglesias, Sara Verdura, Sonia Pernas, Joan Dorca, Isabel Álvarez, Susana Martínez, Jose Manuel Pérez-García, Norberto Batista-López, César A. Rodríguez-Sánchez, Kepa Amillano, Severina Domínguez, Maria Luque, Idoia Morilla, Agostina Stradella, Gemma Viñas, Javier Cortés, Jorge Joven, Joan Brunet, Eugeni López-Bonet, Margarita García, Samiha Saidani, Xavier Queralt Moles, Begoña Martín-Castillo and Javier A. Menendez
- 15    *Establishment and Verification of a Bagged-Trees-Based Model for Prediction of Sentinel Lymph Node Metastasis for Early Breast Cancer Patients***  
Chao Liu, Zeyin Zhao, Xi Gu, Lisha Sun, Guanglei Chen, Hao Zhang, Yanlin Jiang, Yixiao Zhang, Xiaoyu Cui and Caigang Liu
- 25    *Neoadjuvant Chemotherapy Alters Neuropilin-1, PlGF, and SNAI1 Expression Levels and Predicts Breast Cancer Patients Response***  
Noura Al-Zeheimi, Adviti Naik, Charles Saki Bakheit, Marwa Al Riyami, Adil Al Ajarrah, Suaad Al Badi, Khalid Al Baimani, Kamran Malik, Zamzam Al Habsi, Mansour S. Al Moundhri and Sirin A. Adham
- 37    *A Pathology-Based Combined Model to Identify PAM50 Non-luminal Intrinsic Disease in Hormone Receptor-Positive HER2-Negative Breast Cancer***  
Tomás Pascual, Miguel Martín, Aranzazu Fernández-Martínez, Laia Paré, Emilio Alba, Álvaro Rodríguez-Lescure, Giuseppe Perrone, Javier Cortés, Serafín Morales, Ana Lluch, Ander Urruticoechea, Blanca González-Farré, Patricia Galván, Pedro Jares, Adela Rodríguez, Nuria Chic, Daniela Righi, Juan Miguel Cejalvo, Giuseppe Tonini, Barbara Adamo, Maria Vidal, Patricia Villagrasa, Montserrat Muñoz and Aleix Prat
- 46    *Differentiation Between G1 and G2/G3 Phyllodes Tumors of Breast Using Mammography and Mammographic Texture Analysis***  
Wen Jing Cui, Cheng Wang, Ling Jia, Shuai Ren, Shao Feng Duan, Can Cui, Xiao Chen and Zhong Qiu Wang
- 55    *Androgen Receptor Expression and Association With Distant Disease-Free Survival in Triple Negative Breast Cancer: Analysis of 263 Patients Treated With Standard Therapy for Stage I-III Disease***  
Maria Vittoria Dieci, Vassilena Tsvetkova, Gaia Griguolo, Federica Miglietta, Mara Mantiero, Giulia Tasca, Enrico Cumerlato, Carlo Alberto Giorgi, Tommaso Giarratano, Giovanni Faggioni, Cristina Falci, Grazia Vernaci, Alice Menichetti, Eleonora Mioranza, Elisabetta Di Liso, Simona Frezzini, Tania Saibene, Enrico Orvieto and Valentina Guarneri
- 64    *TUFT1 Promotes Triple Negative Breast Cancer Metastasis, Stemness, and Chemoresistance by Up-Regulating the Rac1/ $\beta$ -Catenin Pathway***  
Weiguang Liu, Guanglei Chen, Lisha Sun, Yue Zhang, Jianjun Han, Yuna Dai, Jianchao He, Sufang Shi and Bo Chen



- 77 Mechanisms of Resistance to CDK4/6 Inhibitors: Potential Implications and Biomarkers for Clinical Practice**  
Amelia McCartney, Ilenia Migliaccio, Martina Bonechi, Chiara Biagioni, Dario Romagnoli, Francesca De Luca, Francesca Galardi, Emanuela Risi, Irene De Santo, Matteo Benelli, Luca Malorni and Angelo Di Leo
- 85 Comparison of GenesWell BCT Score With Oncotype DX Recurrence Score for Risk Classification in Asian Women With Hormone Receptor-Positive, HER2-Negative Early Breast Cancer**  
Mi Jeong Kwon, Jeong Eon Lee, Joon Jeong, Sang Uk Woo, Jinil Han, Byeong-il Kang, Jee-Eun Kim, Youngho Moon, Sae Byul Lee, Seonghoon Lee, Yoon-La Choi, Youngmi Kwon, Kyoung Song, Gyungyub Gong and Young Kee Shin
- 99 PAM50 Subtypes in Baseline and Residual Tumors Following Neoadjuvant Trastuzumab-Based Chemotherapy in HER2-Positive Breast Cancer: A Consecutive-Series From a Single Institution**  
Sonia Pernas, Anna Petit, Fina Climent, Laia Paré, J. Perez-Martin, Luz Ventura, Milana Bergamino, Patricia Galván, Catalina Faló, Idoia Morilla, Adela Fernandez-Ortega, Agostina Stradella, Montse Rey, Amparo Garcia-Tejedor, Miguel Gil-Gil and Aleix Prat
- 109 Corrigendum: PAM50 Subtypes in Baseline and Residual Tumors Following Neoadjuvant Trastuzumab-Based Chemotherapy in HER2-Positive Breast Cancer: A Consecutive-Series From a Single Institution**  
Sonia Pernas, Anna Petit, Fina Climent, Laia Paré, J. Perez-Martin, Luz Ventura, Milana Bergamino, Patricia Galván, Catalina Faló, Idoia Morilla, Adela Fernandez-Ortega, Agostina Stradella, Montse Rey, Amparo Garcia-Tejedor, Miguel Gil-Gil and Aleix Prat
- 111 Two Distinct Subtypes Revealed in Blood Transcriptome of Breast Cancer Patients With an Unsupervised Analysis**  
Wenlong Ming, Hui Xie, Zixi Hu, Yuanyuan Chen, Yanhui Zhu, Yunfei Bai, Hongde Liu, Xiao Sun, Yun Liu and Wanjun Gu
- 123 ANLN and KDR Are Jointly Prognostic of Breast Cancer Survival and Can Be Modulated for Triple Negative Breast Cancer Control**  
Xiaofeng Dai, Yi Mei, Xiao Chen and Dongyan Cai
- 134 Prognostic Value of Ki-67 in Patients With Resected Triple-Negative Breast Cancer: A Meta-Analysis**  
Qiang Wu, Guangzhi Ma, Yunfu Deng, Wuxia Luo, Yaqin Zhao, Wen Li and Qinghua Zhou
- 143 Deciphering HER2 Breast Cancer Disease: Biological and Clinical Implications**  
Ana Godoy-Ortiz, Alfonso Sanchez-Muñoz, Maria Rosario Chica Parrado, Martina Álvarez, Nuria Ribelles, Antonio Rueda Dominguez and Emilio Alba

- 159** *Different Pathological Complete Response Rates According to PAM50 Subtype in HER2+ Breast Cancer Patients Treated With Neoadjuvant Pertuzumab/Trastuzumab vs. Trastuzumab Plus Standard Chemotherapy: An Analysis of Real-World Data*  
Tamara Díaz-Redondo, Rocio Lavado-Valenzuela, Begoña Jimenez, Tomas Pascual, Fernando Gálvez, Alejandro Falcón, Maria del Carmen Alamo, Cristina Morales, Marta Amerigo, Javier Pascual, Alfonso Sanchez-Muñoz, Macarena González-Guerrero, Luis Vicioso, Aurora Laborda, Maria Victoria Ortega, Lidia Perez, Aranzazu Fernandez-Martinez, Nuria Chic, Jose Manuel Jerez, Martina Alvarez, Aleix Prat, Nuria Ribelles and Emilio Alba
- 166** *7-lncRNA Assessment Model for Monitoring and Prognosis of Breast Cancer Patients: Based on Cox Regression and Co-expression Analysis*  
Huayao Li, Chundi Gao, Lijuan Liu, Jing Zhuang, Jing Yang, Cun Liu, Chao Zhou, Fubin Feng and Changgang Sun
- 178** *OSbrca: A Web Server for Breast Cancer Prognostic Biomarker Investigation With Massive Data From Tens of Cohorts*  
Zhongyi Yan, Qiang Wang, Xiaoxiao Sun, Bingbing Ban, Zhendong Lu, Yifang Dang, Longxiang Xie, Lu Zhang, Yongqiang Li, Wan Zhu and Xiangqian Guo



## OPEN ACCESS

## Edited by:

Aleix Prat,  
Hospital Clínic de Barcelona, Spain

## Reviewed by:

Tarah Ballinger,  
Indiana University, Purdue University  
Indianapolis, United States  
Marcelo Rocha Cruz,  
Hospital Sírio-Libanês, Brazil

## \*Correspondence:

Begoña Martín-Castillo  
bmartin@iconcologia.net  
Javier A. Menendez  
jmenendez@iconcologia.net;  
jmenendez@idibgi.org

†These authors have contributed  
equally to this work

## Specialty section:

This article was submitted to  
Women's Cancer,  
a section of the journal  
Frontiers in Oncology

Received: 18 December 2018

Accepted: 06 March 2019

Published: 28 March 2019

## Citation:

Cuyàs E, Buxó M, Ferri Iglesias MJ,  
Verdura S, Pernas S, Dorca J,  
Álvarez I, Martínez S,  
Pérez-García JM, Batista-López N,  
Rodríguez-Sánchez CA, Amillano K,  
Domínguez S, Luque M, Morilla I,  
Stradella A, Viñas G, Cortés J,  
Joven J, Brunet J, López-Bonet E,  
García M, Saidani S, Queralt Moles X,  
Martín-Castillo B and Menéndez JA  
(2019) The C Allele of ATM  
rs11212617 Associates With Higher  
Pathological Complete Remission  
Rate in Breast Cancer Patients  
Treated With Neoadjuvant Metformin.  
Front. Oncol. 9:193.  
doi: 10.3389/fonc.2019.00193

# The C Allele of ATM rs11212617 Associates With Higher Pathological Complete Remission Rate in Breast Cancer Patients Treated With Neoadjuvant Metformin

Elisabet Cuyàs<sup>1,2†</sup>, Maria Buxó<sup>2†</sup>, Maria José Ferri Iglesias<sup>3</sup>, Sara Verdura<sup>1,2</sup>, Sonia Pernas<sup>4</sup>, Joan Dorca<sup>5</sup>, Isabel Álvarez<sup>6,7</sup>, Susana Martínez<sup>8</sup>, Jose Manuel Pérez-García<sup>9</sup>, Norberto Batista-López<sup>10</sup>, César A. Rodríguez-Sánchez<sup>11,12</sup>, Kepa Amillano<sup>13</sup>, Severina Domínguez<sup>14</sup>, Maria Luque<sup>15</sup>, Idoia Morilla<sup>4</sup>, Agostina Stradella<sup>4</sup>, Gemma Viñas<sup>5</sup>, Javier Cortés<sup>16</sup>, Jorge Joven<sup>17</sup>, Joan Brunet<sup>5,18,19</sup>, Eugeni López-Bonet<sup>20</sup>, Margarita García<sup>21</sup>, Samiha Saidani<sup>22</sup>, Xavier Queralt Moles<sup>3</sup>, Begoña Martín-Castillo<sup>22\*</sup> and Javier A. Menéndez<sup>1,2\*</sup>

<sup>1</sup> Program Against Cancer Therapeutic Resistance (ProCURE), Metabolism and Cancer Group, Catalan Institute of Oncology, Girona, Spain, <sup>2</sup> Girona Biomedical Research Institute (IDIBGI), Girona, Spain, <sup>3</sup> Laboratori Clínic Territorial, Parque Hospitalario Martí i Julià, Salt, Spain, <sup>4</sup> Breast Unit, Department of Medical Oncology, Catalan Institute of Oncology-Hospital Universitari de Bellvitge-Bellvitge Research Institute (IDIBELL), L'Hospitalet de Llobregat, Barcelona, Spain, <sup>5</sup> Medical Oncology, Catalan Institute of Oncology, Girona, Spain, <sup>6</sup> Medical Oncology Service, Hospital Universitario Donostia, Donostia-San Sebastián, Spain, <sup>7</sup> Biodonostia Health Research Institute, Donostia-San Sebastián, Spain, <sup>8</sup> Medical Oncology Department, Hospital de Mataró, Mataró, Barcelona, Spain, <sup>9</sup> Hospital Quirón, IOB Institute of Oncology, Barcelona, Spain, <sup>10</sup> Medical Oncology Service, Hospital Universitario de Canarias, San Cristóbal de La Laguna, Spain, <sup>11</sup> Medical Oncology Service, Hospital Universitario de Salamanca, Salamanca, Spain, <sup>12</sup> Instituto de Investigación Biomédica de Salamanca (IBSAL), Salamanca, Spain, <sup>13</sup> Medical Oncology, Hospital Universitari Sant Joan, Reus, Spain, <sup>14</sup> Medical Oncology Service, Hospital Universitario Araba, Vitoria-Gasteiz, Spain, <sup>15</sup> Department of Medical Oncology, Hospital Universitario Central de Asturias, Oviedo, Spain, <sup>16</sup> Department of Medical Oncology, Ramón y Cajal University Hospital, Madrid, Spain, <sup>17</sup> Unitat de Recerca Biomèdica, Hospital Universitari de Sant Joan, IISPV, Rovira i Virgili University, Reus, Spain, <sup>18</sup> Hereditary Cancer Programme, Catalan Institute of Oncology (ICO), Bellvitge Institute for Biomedical Research (IDIBELL), L'Hospitalet de Llobregat, Barcelona, Spain, <sup>19</sup> Hereditary Cancer Programme, Catalan Institute of Oncology (ICO), Girona Biomedical Research Institute (IDIBGI), Girona, Spain, <sup>20</sup> Department of Anatomical Pathology, Dr. Josep Trueta Hospital of Girona, Girona, Spain, <sup>21</sup> Clinical Research Unit, Catalan Institute of Oncology, L'Hospitalet de Llobregat, Barcelona, Spain, <sup>22</sup> Unit of Clinical Research, Catalan Institute of Oncology, Girona, Spain

**Background:** The minor allele (C) of the single-nucleotide polymorphism (SNP) rs11212617, located near the *ataxia telangiectasia mutated* (ATM) gene, has been associated with an increased likelihood of treatment success with metformin in type 2 diabetes. We herein investigated whether the same SNP would predict clinical response to neoadjuvant metformin in women with early breast cancer (BC).

**Methods:** DNA was collected from 79 patients included in the intention-to-treat population of the METTEN study, a phase 2 clinical trial of HER2-positive BC patients randomized to receive either metformin combined with anthracycline/taxane-based chemotherapy and trastuzumab or equivalent regimen without metformin, before surgery. SNP rs11212617 genotyping was assessed using allelic discrimination by quantitative polymerase chain reaction.

**Results:** Logistic regression analyses revealed a significant relationship between the *rs11212617* genotype and the ability of treatment arms to achieve a pathological complete response (pCR) in patients (odds ratio  $[OR]_{\text{genotype} \times \text{arm}} = 10.33$ , 95% confidence interval  $[CI]: 1.29\text{--}82.89$ ,  $p = 0.028$ ). In the metformin-containing arm, patients bearing the *rs11212617* C allele had a significantly higher probability of pCR ( $OR_{A/C,C/C} = 7.94$ , 95%CI:  $1.60\text{--}39.42$ ,  $p = 0.011$ ). Conversely, no association was found between *rs11212617* and clinical response in the reference arm ( $OR_{A/C,C/C} = 0.77$ , 95%CI:  $0.20\text{--}2.92$ ,  $p = 0.700$ ). After controlling for tumor size and hormone receptor status, the *rs11212617* C allele remained a significant predictor of pCR solely in the metformin-containing arm.

**Conclusions:** If reproducible, the *rs11212617* C allele might warrant consideration as a predictive clinical biomarker to inform the personalized use of metformin in BC patients.

**Trial Registration:** EU Clinical Trials Register, EudraCT number 2011-000490-30. Registered 28 February 2011, <https://www.clinicaltrialsregister.eu/ctr-search/trial/2011-000490-30/ES>.

**Keywords:** metformin, breast cancer, neoadjuvancy, HER2, ATM, *rs11212617*

## INTRODUCTION

The minor allele C of the noncoding single nucleotide polymorphism (SNP) *rs11212617*, which is located near the *ataxia telangiectasia mutated* (ATM) gene, was found to be associated with the metabolic response to the biguanide metformin in the first genome-wide association study (GWAS) carried out in 3,912 Europeans with type 2 diabetes (T2D) (1). Although lack of replication occurred in some studies aiming to verify the association between *rs11212617* and the effect of metformin in multiple ethnic groups (2), a meta-analysis in smaller cohorts suggested that the *rs11212617* C allele might be considered as the first robustly replicated common susceptibility locus associated with metformin treatment success in patients with T2D (3). Moreover, *rs11212617* remained a top signal with no other genome-significant hits in a more recent GWAS of 13,123 participants of different ancestries, but failed to associate with glycemic response to metformin in a systematic three-stage replication study (4). However, *rs11212617* has recently been shown to significantly affect not only the response to metformin in terms of insulin Z score, but also metformin plasma concentration (5). Mechanistic studies have shown that *rs11212617* increases enhancer activity and could lead to elevated expression of several target genes including ATM itself (6). Yet, almost nothing is known about the impact of the *rs11212617* C allele on the clinical efficacy of metformin in several ongoing clinical trials aiming to evaluate its potential benefits in a cancer setting (7).

A potential anti-cancer effect of metformin has gained considerable epidemiological and pre-clinical support over the last decade (7–10). First, a large number of population-based observational and cohort studies have suggested a cancer-preventive advantage associated with metformin usage

among T2D patients (11). Second, diabetic patients with breast cancer receiving metformin during neoadjuvant chemotherapy were reported to benefit from a 3-fold greater pathological complete response (pCR) when compared with those who did not receive metformin (12). Third, an ever-growing number of pre-clinical studies have proposed numerous cell-autonomous (e.g., AMPK/mTOR-related) and non-cell-autonomous (e.g., insulin/IGF-1-related) molecular mechanisms that have enthusiastically endorsed the clinical development of metformin as a novel anti-cancer drug (13–15). However, one should acknowledge that a metformin-driven cancer-preventive advantage does not necessarily imply an effective therapeutic efficacy in non-diabetic patients with established cancers, and it remains unclear whether the adjuvant use of metformin in combination with standard cancer therapy could translate into better clinical outcomes (16–19). Indeed, recent randomized studies reporting the use of metformin in cancer treatment have yielded mixed results in patients with advanced disease (20, 21). Although the results of much larger randomized studies, such as NCIC CTG MA.32, the most advanced adjuvant trial investigating the effects of metformin vs. placebo on invasive disease-free survival and other outcomes on early breast cancer in 3,649 women (22), will be of great interest to confirm or reject the causal nature of the suggested correlation between metformin use and survival benefit in cancer patients, it is also true that companion biomarker studies are urgently needed to refine tumor and patient selection when using metformin as an adjuvant to established cancer therapeutics.

We herein investigated whether the presence of the *rs11212617* C allele could predict the pathological complete response (pCR) in the METTEN study (23, 24), a randomized, open-label, multicenter, phase 2 trial of neoadjuvant metformin

in combination with trastuzumab and chemotherapy in women with early HER2-positive breast cancer.

## MATERIALS AND METHODS

### Subjects

The METTEN study was registered with the EU Clinical Trials Register and is available online (<https://www.clinicaltrialsregister.eu/ctr-search/trial/2011-000490-30/ES>).

Patients were randomly assigned to receive daily metformin (850 mg twice-daily) for 24 weeks concurrently with 12 cycles of weekly paclitaxel (80 mg/m<sup>2</sup>) plus trastuzumab (4 mg/kg loading dose followed by 2 mg/kg) followed by four cycles of 3 weekly fluorouracil (600 mg/m<sup>2</sup>), epirubicin (75 mg/m<sup>2</sup>), cyclophosphamide (600 mg/m<sup>2</sup>) with concomitant trastuzumab (6 mg/kg) (arm A), or equivalent sequential chemotherapy plus trastuzumab without metformin (arm B), followed by surgery. Patients had surgery within 4–5 weeks of the last cycle of neoadjuvant treatment (24). Post-surgery, patients received thrice-weekly trastuzumab to complete 1 year of neoadjuvant-adjuvant therapy. Genotyping of SNP *rs11212617* was carried out in the intention-to-treat (ITT) population ( $n = 79$ ), which included all randomly assigned patients who received at least one dose of study medication.

### Assessment of Pathological Complete Response (pCR)

pCR was defined as absence of invasive tumor cells on hematoxylin and eosin evaluation of the complete resected breast specimen (and all sample regional lymph nodes if lymphadenectomy was performed) following the completion of neoadjuvant systemic therapy. Residual ductal carcinoma *in situ* (DCIS) only was included in the definition of pCR (ypT0/is, ypN0) (24).

### Analytical Methods

Blood was drawn after an overnight fast. Serum glucose was measured in duplicate using the glucose oxidase method and serum insulin was measured in duplicate using the Human Insulin ELISA (Cat. # EZHI-14K, Merck Millipore, Billerica, MA). The lowest level of insulin that can be detected by this assay is 2 µU/mL when using a 20 µL sample size. Intra- and inter-assay coefficients of variation were below 6 and 11%, respectively. Fasting insulin resistance was calculated using the homeostasis model assessment (HOMA) using the following formula:  $\text{HOMA-IR} = \text{fasting glucose (mmol/L)} \times \text{fasting insulin (mU/L)} / 22.5$ .

### Genotyping of SNP *rs11212617*

The ATM *rs11212617* SNP variants were determined using the 5' exonuclease TaqMan-based allelic discrimination method (Applied Biosystems, assay ID C\_134213\_10).

### Statistical Analysis

Descriptive data were summarized using percentages, medians or means with their respective 25 and 75 percentiles or standard deviations as appropriate. Clinical baseline characteristics

between groups (non-pCR and pCR) were assessed using Chi-square or Fisher's exact test for categorical variables, student *t*-test for continuous variables with normal distribution, or Mann-Whitney *U* test for non-normal distributions. The assumption of normality was evaluated with the Shapiro-Wilk test. Changes in glucose, insulin, and HOMA-IR between pre and post treatment were compared using the Wilcoxon test. The R package Hardy-Weinberg (<http://www.jstatsoft.org/v64/i03/>) was employed to check whether the Hardy-Weinberg equilibrium holds among study population. Binary logistic regression was used to assess the prognostic effect of baseline *rs11212617* genotype on pCR. Unadjusted and adjusted odds ratios (ORs) with their relative 95% confidence intervals (CIs) were reported as a measure of association. All tests were 2-sided and  $P \leq 0.05$  was set as statistically significant. Statistical analyses were carried out using SPSS (IBM Corp. released 2017. IBM SPSS Statistics for Windows, Version 25.0; Armonk, NY) and STATA (StataCorp. 2013. Stata Statistical Software: Release 13; StataCorp LP, College Station, TX).

## RESULTS

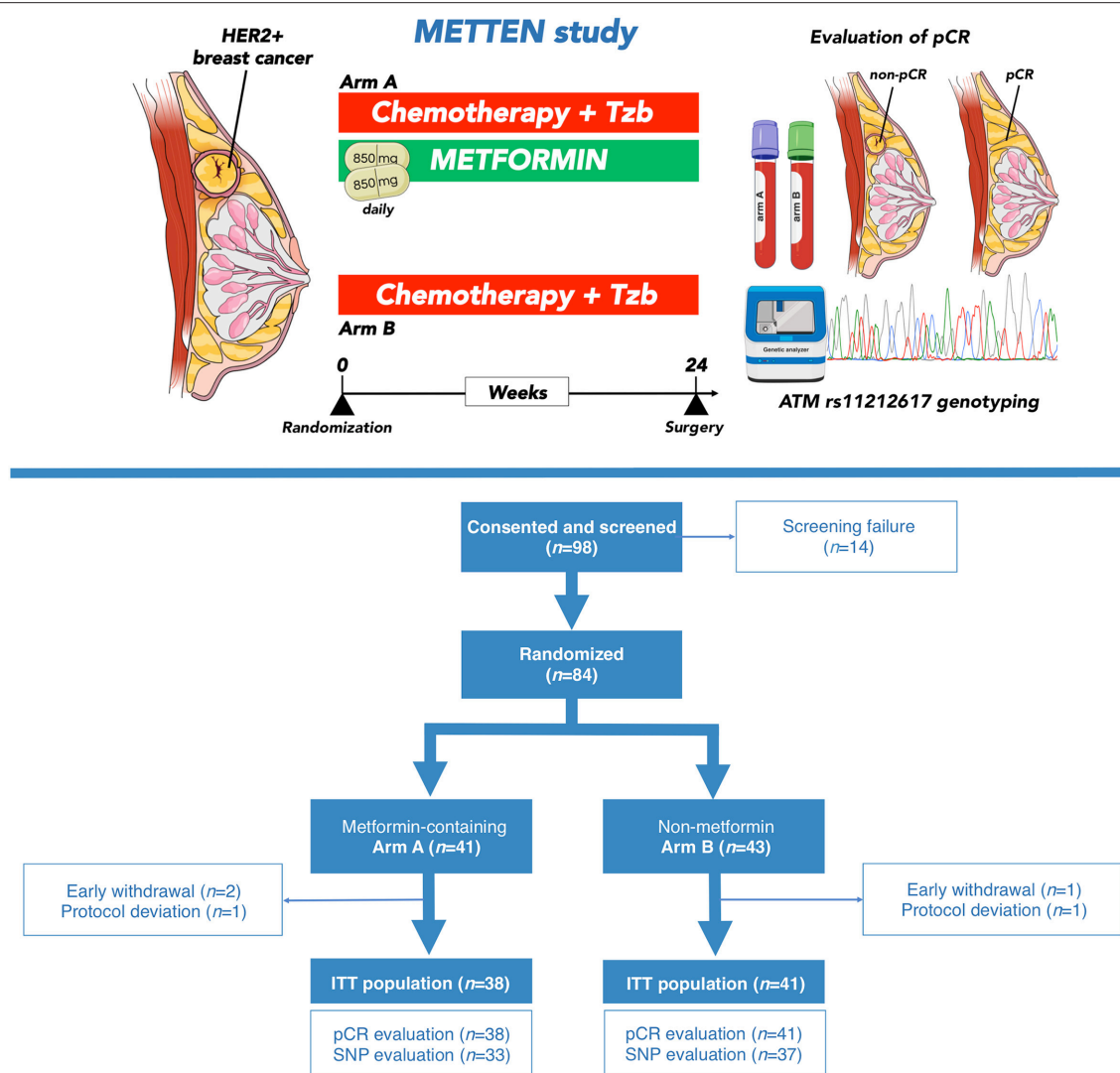
### Study Participants

This study was designed to evaluate the clinical relevance of the SNP *rs11212617* C allele with respect to its potential to predict a pCR in breast cancer patients with HER2 overexpression treated with metformin-containing neoadjuvant systemic therapy (Figure 1). We conducted the study with patients belonging to the ITT population of the METTEN trial, which included all randomly assigned patients who received at least one dose of study medication ( $n = 79$ ) (24). A flowchart describing the formation of each cohort in the study is shown in Figure 1. The baseline characteristics of those ITT patients who achieved pCR after neoadjuvant therapy and those who did not are shown in Table 1. The comparison of clinical-pathological variables at diagnosis between patients of each non-pCR/pCR cohort revealed no significant differences, except for hormone receptor status. The non-pCR group tended to have more estrogen receptor-negative and/or progesterone-positive tumors ( $p = 0.056$ ).

### Allele Frequencies of *rs11212617*

The *rs11212617* polymorphism was evaluable in most of the patient samples, and 70 of 79 patients (89%) were genotyped (Figure 1, Table 2). The A and C allelic frequencies of *rs11212617* in our patients were 69 and 31%, respectively. The frequencies of three genotypes in all the patients were 14.3% (C/C), 32.9% (A/C), and 52.9% (A/A). These genotype frequencies were very similar to those predicted by the Ensembl genome database for a Tuscany, in Italy (TSI) population, and slightly different to those observed in Europeans and the Iberian population in Spain (Table 2). Despite the small population size, there was no significant deviation in *rs11212617* genotype frequencies in our population from the Hardy-Weinberg expectation [HWE; Sum Equally Likely or More Extreme [SELOME]  $p = 0.0879$ ]. No significant differences were observed in the genotype frequencies





**FIGURE 1 |** The METTEN study. **(Top)** The open-label, multicenter, phase II randomized METTEN study was designed to evaluate the clinical activity, tolerability, and safety of adding metformin to neoadjuvant chemotherapy plus trastuzumab in operable, locally advanced, or inflammatory HER2-positive BC (23, 24). Women with primary, non-metastatic HER2-positive BC were randomly assigned (1:1) to receive daily metformin (850 mg twice-daily) for 24 weeks concurrently with 12 cycles of weekly paclitaxel plus trastuzumab followed by four cycles of 3 weekly fluorouracil, epirubicin, cyclophosphamide plus trastuzumab (arm A) or equivalent sequential chemotherapy plus trastuzumab without metformin (arm B), followed by surgery. The primary end point was pCR, defined as absence of invasive tumor cells on hematoxylin and eosin evaluation of the complete resected breast specimen (and all sample regional lymph nodes if lymphadenectomy was performed) following the completion of neoadjuvant systemic therapy. Residual ductal carcinoma *in situ* (DCIS) only was included in the definition of pCR (ypT0/is, ypN0). Between June 1, 2012 and March 17, 2016, 98 patients at 10 centers in Spain were recruited into the METTEN study. DNA sample collection was not included in the original study design and was added as addendum #3 in April 2012 to re-consent patients for an additional blood draw for germ line DNA extraction. DNA samples from 70 patients (89% of the full ITT cohort) were subsequently collected and genotyped for SNP *rs11212617*. **(Bottom)** Modified CONSORT diagram showing the 70 cases of HER2-positive BC patients used for the analysis of clinical response analysis to neoadjuvant metformin by the minor allele C of the SNP *rs11212617*.

of SNP *rs11212617* between the non-pCR and pCR cohorts in the ITT population (Table 1).

## Association Between *rs11212617* and Clinical Response

Frequency distributions of SNP *rs11212617* were similar between treatment arms (Table S1). Of the patients in the metformin-containing arm A, 81.2% of homo or heterozygous patients

for the *rs11212617* C allele achieved a pCR, whereas 64.7% of non-carrier patients did not achieve a pCR (Figure 2, top panels). Of the patients in the reference arm B, 58.8% of homo or heterozygous patients for the *rs11212617* C allele and 65% of non-carrier patients achieved a pCR, respectively (Figure 2, top panels). We employed logistic binary regression analyses to investigate the association between arm, ATM *rs11212617* genotype, and pCR. In bivariate analysis, we failed

**TABLE 1 |** Clinical characteristics of patients at baseline according to pathological complete response (pCR) status.

	Non-pCR (n = 31)	pCR (n = 48)	p-value
<b>Arm</b>			0.335
A	17 (54.8%)	21 (43.8%)	
B	14 (45.2%)	27 (56.3%)	
<b>SNP rs11212617<sup>a</sup></b>			0.214* <sup>1</sup>
A/A	18 (64.3%)	19 (45.2%)	
A/C	6 (21.4%)	17 (40.5%)	
C/C	4 (14.3%)	6 (14.3%)	
			0.118
A/A	18 (64.3%)	19 (45.2%)	
A/C, C/C	10 (35.7%)	23 (54.8%)	
<b>Age</b>			0.465
<50	20 (64.5%)	27 (56.3%)	
≥50	11 (35.5%)	21 (43.8%)	
Mean ± SD (min;max)	47.1 ± 11.9 (30;75)	48.0 ± 10.6 (23;71)	0.741
<b>Premenopausal status</b>			0.583
Post	11 (35.5%)	20 (41.7%)	
Pre+Peri	20 (64.5%)	28 (58.3%)	
<b>Body weight (kg)</b>			0.592
Mean ± SD (min;max)	64.3 ± 6.9 (48;78)	65.3 ± 10.2 (45.3;89.0)	
<b>Body-mass index</b>			0.179
<25	19 (61.3%)	22 (45.8%)	
≥25 (overweight)	12 (38.7%)	26 (54.2%)	
<b>Clinical tumor status</b>			0.077* <sup>1</sup>
cT2	18 (58.1%)	33 (68.8%)	
cT3	12 (38.7%)	10 (20.8%)	
cT4a	1 (3.2%)	0 (0.0%)	
cT4b	0 (0.0%)	4 (8.3%)	
cT4d	0 (0.0%)	1 (2.1%)	
<b>Clinical nodal stage</b>			0.581* <sup>1</sup>
cN0	6 (19.4%)	16 (33.3%)	
cN1	20 (64.5%)	24 (50.0%)	
cN2	1 (3.2%)	2 (4.2%)	
cN2a	1 (3.2%)	0 (0.0%)	
cN2b	0 (0.0%)	1 (2.1%)	
cN3	3 (9.7%)	4 (8.3%)	
cN3c	0 (0.0%)	1 (2.1%)	
<b>Hormone receptor status</b>			0.056
ER and/or PgR positive	21 (67.7%)	22 (45.8%)	
ER and PgR negative	10 (32.3%)	26 (54.2%)	
<b>Tumor grade<sup>b</sup></b>			1.000* <sup>1</sup>
G1	1 (4.0%)	1 (2.8%)	
G2	12 (48.0%)	18 (50.0%)	
G3	12 (48.0%)	17 (47.2%)	
<b>Baseline glucose (mmol/L)</b>			0.511
Mean ± SD (min;max)	5.2 ± 0.4 (4.4;6.0)	5.2 ± 0.5 (3.9;6.5)	
<b>Baseline insulin (mU/mL)</b>			0.834
Mean ± SD (min;max)	8.7 ± 12.2 (2.1;62.9)	8.2 ± 5.2 (3.0;21.6)	
<b>Baseline HOMA</b>			0.964
Mean ± SD (min;max)	1.9 ± 2.6 (0.5;13.1)	1.9 ± 1.2 (0.6;5.1)	

\*<sup>1</sup> Fisher exact test.<sup>a</sup>Data available for 70 of 79 patients.<sup>b</sup>Data available for 61 of 79 patients.**TABLE 2 |** Expected and observed SNP rs11212617 prevalence (%).

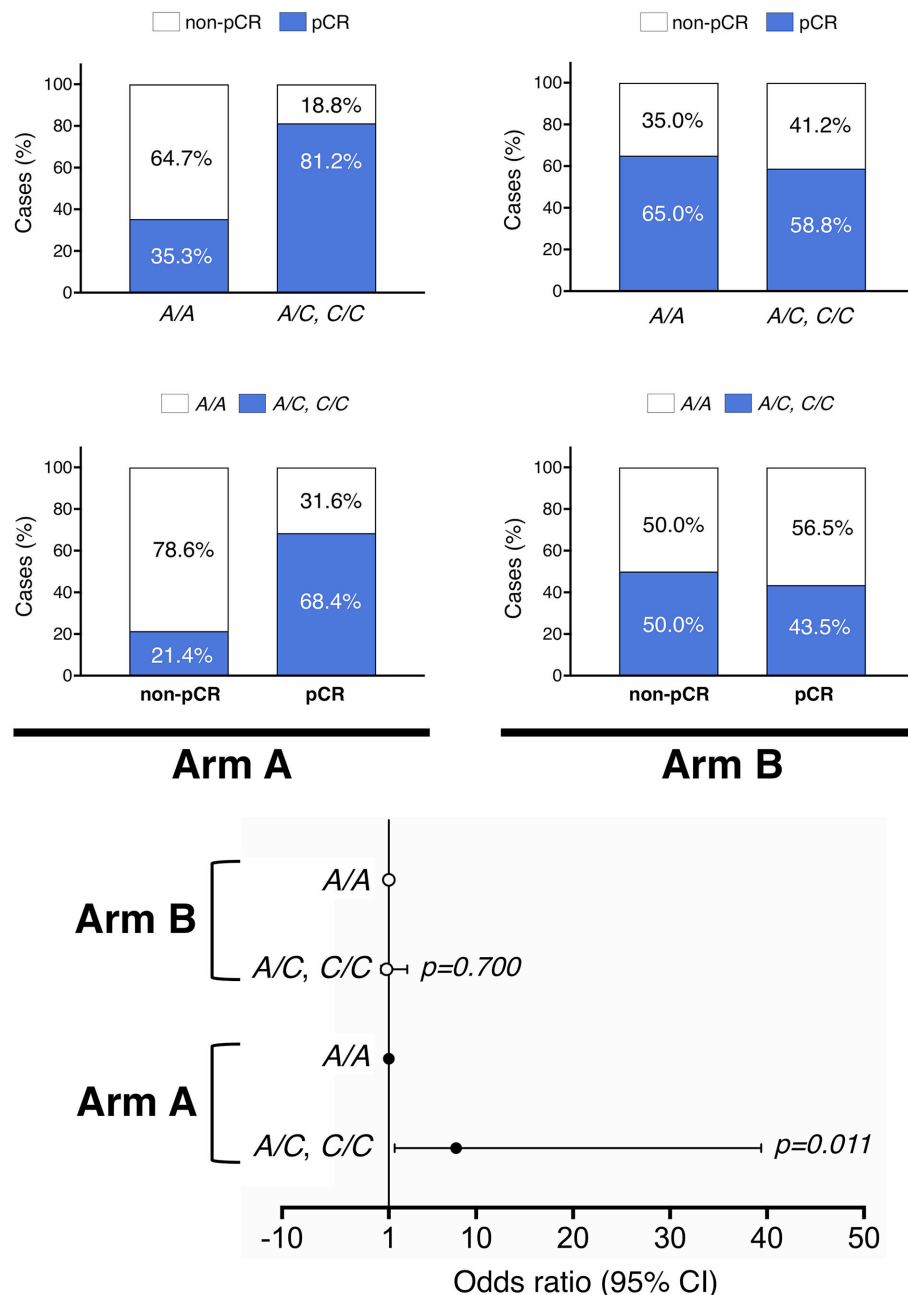
	Expected <sup>a</sup>			Observed	
	IBS <sup>b</sup>	EUR <sup>c</sup>	TSI <sup>d</sup>	METTEN trial (n = 70)	
ATM rs11212617	%	%	%	n	%
A/A	41.1	38.2	51.4	37	52.9
A/C	49.5	47.1	35.5	23	32.9
C/C	9.3	14.7	13.1	10	14.3

<sup>a</sup>[http://www.ensembl.org/Homo\\_sapiens/Variation/Population?db=core;r=\\$11:108411934-108412934;v=\\$rs11212617;vdb=\\$variation;vf=\\$6530681#373524\\_tablePanel](http://www.ensembl.org/Homo_sapiens/Variation/Population?db=core;r=$11:108411934-108412934;v=$rs11212617;vdb=$variation;vf=$6530681#373524_tablePanel).<sup>b</sup>IBS, Iberian Population in Spain.<sup>c</sup>EUR, European.<sup>d</sup>TSI, Tuscany in Italy.

to show predictive capacity of either the arm treatment or rs11212617 genotype with the probability of achieving pCR (Table S2). However, we observed a significant relationship between rs11212617 genotype and the ability of treatment arms to achieve pCR ( $OR_{\text{genotype} \times \text{arm}} = 10.33$ , 95%CI: 1.29–82.89,  $p = 0.028$ ; Table 3). This finding suggested that the direction and/or intensity of the relationship between rs11212617 genotype and pCR significantly varied in each treatment arm. Accordingly, the patients bearing the rs11212617 C allele in the metformin-containing arm had a significantly higher probability of pCR ( $OR_{A/C,C/C} = 7.94$ , 95%CI: 1.60–39.42,  $p = 0.011$ ; Figure 2, bottom panel). Conversely, no association was found between the presence of the rs11212617 C allele and clinical response in the (non-metformin) reference arm ( $OR_{A/C,C/C} = 0.77$ , 95%CI: 0.20–2.92,  $p = 0.700$ ; Figure 2, bottom panel). After additional adjusting for potential confounding tumor characteristics such as tumor size and hormone receptor (HR) status, a relationship between the rs11212617 genotype and the ability of treatment arms to achieve a pCR in patients remained significant (adjusted  $OR_{\text{genotype} \times \text{arm}} = 20.53$ , 95%CI: 1.97–213.79,  $p = 0.011$ ; Table S3). In the metformin-containing arm, the positive association between the presence of the rs11212617 C allele and pCR remained significant after accounting for tumor size and HR status (adjusted  $OR_{A/C,C/C} = 28.88$ , 95%CI: 2.20–378.73,  $p = 0.010$ ; Table S4). The lack of association between the rs11212617 C allele and pCR in the (non-metformin) reference arm was not altered after adjusting for these factors (Table S5).

## Association Between ATM rs11212617 and Metabolic Response

A Wilcoxon test was conducted to evaluate whether there was a significant relationship between the rs11212617 C allele and the metabolic response to each arm. In the reference arm, no significant relationship between rs11212617 C allele and reductions in glucose, insulin, or HOMA-IR index was evident (Table 4). In the metformin arm, however, there was a near-significant trend between the rs11212617 C allele and the metabolic response to metformin in terms of insulin reduction ( $p = 0.069$ ; Table 4).



**FIGURE 2 |** Association of ATM rs11212617 genotype and pCR by treatment arm. **(Top)** Rates of pCR stratified by the (C) rs11212617 genotype in patients randomized to receive either metformin combined with anthracycline/taxane-based chemotherapy and trastuzumab (arm A) or equivalent regimen without metformin (arm B). **(Bottom)** Relationship between the (C) rs11212617 genotype and the ability of treatment arms to achieve pCR.

## DISCUSSION

A significant number of neoadjuvant, adjuvant, and advanced disease trials are currently ongoing or have been proposed to elucidate whether metformin, when used at doses established for diabetes control, has the potential to be used in preventive and treatment settings as an adjuvant to established cancer therapeutics. In this

scenario, companion biomarker studies are urgently needed to define metformin efficacy and refine the tumor types and/or patient populations that are most likely to benefit from metformin-containing interventions.

To our knowledge, this is the first prospective study evaluating the relationship between the ATM SNP rs11212617 C allele, which has been associated with an increased likelihood of metformin treatment success in T2D (1, 3, 5), and



the clinical benefit of adding metformin to well-established neoadjuvant treatment regimens in breast cancer patients. Logistic regression analyses revealed a significant relationship between the *rs11212617* genotype and the ability of treatment arms to achieve a pCR. In the metformin-containing arm, patients bearing the *rs11212617* C allele had a significantly higher probability of pCR. Conversely, no association was found between *rs11212617* and clinical response in the reference arm. Because greater benefits from HER2-targeted neoadjuvant treatment in breast cancer are achieved in patients with small HR-negative tumors compared with patients with large HR-positive tumors (25), it is noteworthy that the capacity of the ATM *rs11212617* C allele to predict a higher chance of achieving a pCR in patients treated with neoadjuvant metformin was not altered after accounting for factors like tumor size and HR status.

A previous report by Reni et al. (21) failed to observe any association between the C allele of *rs11212617* and the clinical response to metformin in pancreatic cancer, but a significant relationship between the highest reduction of fasting plasma glucose and the CC genotype was observed. Our study suggests that the presence of the minor C allele of *rs11212617* might associate with a significant improvement in insulin sensitivity in HER2-positive breast cancer patients subjected to neoadjuvant metformin in combination with trastuzumab and chemotherapy. This was evidenced by a near significant reduction of circulating insulin levels and HOMA-IR index—which fairly correlates with the insulin sensitivity index calculated using the minimal model

approach (26), solely in those patients bearing SNP *rs11212617* C allele in the metformin-containing arm despite maintenance of blood glucose levels.

Limitations of this study are inherent in the design; in particular, the open-label nature of the study, and a relatively modest sample size. Further, because a concurrent analysis of well-characterized breast cancer biomarkers relevant for the putative mechanism of metformin was not achievable, it might be argued that the outcome predicted by the “favorable” C allele could be partially biased. Cancer cells expressing constitutively active phosphatidylinositol-3 kinase (PI3K) are proliferative regardless of the absence of insulin, and they can form dietary restriction (DR)-resistant tumors *in vivo* (27). Accordingly, because the binding of insulin to its receptors activates the PI3K/AKT/mammalian target of rapamycin (mTOR) signaling cascade, activating mutations in the *PIK3CA* oncogene might be expected to determine tumor response to DR-like pharmacological strategies targeting the insulin and mTOR pathways (27, 28). In our hands, however, breast cancer xenografts harboring the insulin-unresponsive, DR-resistant, *PIK3CA*-activating mutation H1047R remained largely sensitive to the anti-tumoral effects of metformin (29). Given that new groundbreaking research has shown how dietary approaches such as carb-restricted ketogenic diets can prevent the systemic glucose-insulin feedback that impairs the efficacy of PI3K inhibitors (30), our current findings, together with the ability of metformin to significantly augment the circulating levels of the ketone body beta-hydroxybutyrate in the metformin-containing arm of the METTEN study (*manuscript in preparation*), might have a significant impact on the design of future trials evaluating the potential of combining metformin with targeted therapy.

In summary, we have genotyped a subset of patients included in a neoadjuvant breast cancer trial to explore the effect of *rs11212617* variants on the clinical endpoint pCR, a powerful predictor of long-term outcome of patients with HER2-positive disease treated with neoadjuvant therapy with or without HER2-targeted agents (31–33). The present findings, although limited by the small effect size, suggest that further analyses using a larger

**TABLE 3 |** Association of the interaction between ATM *rs11212617* genotype and pCR by treatment arm.

		Odds ratio (95% CI)	p-value
ATM <i>rs11212617</i>	A/A	1	
	A/C, C/C	0.77 (0.20–2.92)	0.700
Arm	B	1	
	A	0.29 (0.08–1.14)	0.076
Genotype × Arm		10.33 (1.29–82.99)	0.028

**TABLE 4 |** Association of ATM *rs11212617* genotype with changes in glucose, insulin, and HOMA-IR pre- and post-treatment.

		Pre-treatment	Post-treatment	p-value <sup>a</sup>			Pre-treatment	Post-treatment	p-value <sup>a</sup>
Arm A	n	MD <sup>b</sup> (p25,p75)	MD (p25,p75)		Arm B	n	MD (p25,p75)	MD (p25,p75)	
<b>Glucose</b>					<b>Glucose</b>				
A/A	17	5.27 (4.99,5.54)	5.38 (5.13,5.66)	0.410	A/A	20	5.14 (4.82,5.30)	5.13 (4.60,5.77)	0.588
A/C,C/C	16	5.26 (4.84,5.43)	5.03 (4.79,5.36)	0.109	A/C,C/C	17	5.30 (4.84,5.50)	5.30 (4.92,5.64)	0.344
<b>Insulin</b>					<b>Insulin</b>				
A/A	14	4.78 (3.76,6.39)	7.22 (2.73,8.95)	0.245	A/A	9	5.55 (3.59,11.68)	7.85 (4.93,17.60)	0.441
A/C,C/C	8	6.38 (3.56,9.54)	4.39 (2.26,6.75)	0.069	A/C,C/C	10	6.46 (3.82,10.43)	4.27 (3.74,7.21)	0.169
<b>HOMA-IR<sup>c</sup></b>					<b>HOMA-IR</b>				
A/A	14	1.13 (0.97,1.55)	1.51 (0.64,2.10)	0.397	A/A	9	1.10 (0.80,2.79)	1.96 (0.94,5.36)	0.260
A/C,C/C	8	1.54 (0.82,2.25)	0.95 (0.46,1.65)	0.093	A/C,C/C	10	1.45 (0.86,2.61)	1.09 (0.80,1.80)	0.241

<sup>a</sup>Wilcoxon test.

<sup>b</sup>MD, Median.

<sup>c</sup>Homeostasis Model Assessment of Insulin Resistance.

number of breast cancer patients treated with metformin should verify whether a pharmacogenomic profile including the analysis of ATM SNP rs11212617 genotype might deserve consideration as a predictive clinical biomarker to inform the personalized use of metformin in a cancer setting.

## CONCLUSIONS

Association with a significantly augmented pCR rate was found in metformin-treated breast cancer patients that have a “favorable” C allele-containing ATM SNP rs11212617 genotype. Because achievement of pCR is an appropriate surrogate for significantly improved long-term clinical outcomes in high-risk breast cancer subtypes (34), future studies validating this association of favorable ATM rs11212617 genotype with improvements in relapse-free survival after surgery in the METTEN study (and retrospective outcome analyses for other clinical trials) should definitely determine whether the rs11212617 C allele may lead to actionable modifications for prospective clinical planning in metformin-based anti-breast cancer approaches.

## DATA AVAILABILITY

The datasets generated and analyzed during the current study are available from the corresponding authors on reasonable request.

## ETHICS STATEMENT

The hospital (Dr. Josep Trueta Hospital, Girona, Spain) ethics committee (Clinical Investigation Ethic Committee, CIEC) and independent institutional review boards at each site participating in the METTEN study approved the protocol and any amendments. All procedures were in accordance with the ethical standards of the institutional research committees and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study. The authors declared that they have no competing interests.

## REFERENCES

1. Zhou K, Bellenguez C, Spencer CC, Bennett AJ, Coleman RL, Tavendale R, et al. Common variants near ATM are associated with glycemic response to metformin in type 2 diabetes. *Nat Genet.* (2011) 43:117–20. doi: 10.1038/ng.735
2. Florez JC, Jablonski KA, Taylor A, Mather K, Horton E, White NH, et al. The C allele of ATM rs11212617 does not associate with metformin response in the Diabetes Prevention Program. *Diabetes Care.* (2012) 35:1864–70. doi: 10.2337/dc11-2301
3. van Leeuwen N, Nijpels G, Becker ML, Deshmukh H, Zhou K, Stricker BH, et al. A gene variant near ATM is significantly associated with metformin treatment response in type 2 diabetes: a replication and meta-analysis of five cohorts. *Diabetologia.* (2012) 55:1971–7. doi: 10.1007/s00125-012-2537-x

## AUTHOR CONTRIBUTIONS

BM-C and JM: conceptualization, supervision, and funding acquisition; BM-C, MB, and JM: methodology; MB, JM, and EC: formal analysis and visualization; EC, SV, MF, SP, JD, IA, SM, JP-G, NB-L, CR-S, KA, SD, ML, AS, IM, GV, JC, and JJ: investigation; JB, EL-B, MG, SS, and XQ: resources; EC, SS, and MB: data curation; JM: writing-original draft preparation; JM, JP-G, EC, and BM-C: writing-review and editing; BM-C: project administration.

## FUNDING

This work was supported by grants from the Ministerio de Sanidad, Servicios Sociales e Igualdad (EC10-125, Ayudas para el Fomento de la Investigación Clínica Independiente to BM-C). Work in the Menendez laboratory is supported by the Ministerio de Ciencia e Innovación [Grant SAF2016-80639-P, Plan Nacional de I+D+I, funded by the European Regional Development Fund (EU FEDER), Spain] and by an unrestricted research grant from the Fundació Oncolliga Girona (Lliga catalana d'ajuda al malalt de càncer, Girona).

## ACKNOWLEDGMENTS

The METTEN study was conceived and designed by BM-C and JM, and was sponsored by the Consortium for the Support of Biomedical Research Network (CAIBER) and the Catalan Institute of Oncology (ICO). The Unit of Clinical Research at the ICO in Girona and the Unit for Statistical and Methodological Assessment at the Girona Biomedical Research Institute (IDIBGI) were responsible for central data gathering and analysis. All authors had responsibility for the decision to submit for publication. The authors would like to thank Dr. Kenneth McCreath for editorial support.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2019.00193/full#supplementary-material>

4. Zhou K, Yee SW, Seiser EL, van Leeuwen N, Tavendale R, Bennett AJ, et al. Variation in the glucose transporter gene SLC2A2 is associated with glycemic response to metformin. *Nat Genet.* (2016) 48:1055–9. doi: 10.1038/ng.3632
5. Out M, Becker ML, van Schaik RH, Lehert P, Stehouwer CD, Kooy A. A gene variant near ATM affects the response to metformin and metformin plasma levels: a *post hoc* analysis of an RCT. *Pharmacogenomics.* (2018) 19:715–26. doi: 10.2217/pgs-2018-0010
6. Luizon MR, Eckalbar WL, Wang Y, Jones SL, Smith RP, Laurance M, et al. Genomic characterization of metformin hepatic response. *PLoS Genet.* (2016) 12:e1006449. doi: 10.1371/journal.pgen.1006449
7. Chae YK, Arya A, Malecek MK, Shin DS, Carneiro B, Chandra S, et al. Repurposing metformin for cancer treatment: current clinical studies. *Oncotarget.* (2016) 7:40767–80. doi: 10.18632/oncotarget.8194

8. Gonzalez-Angulo AM, Meric-Bernstam F. Metformin: a therapeutic opportunity in breast cancer. *Clin Cancer Res.* (2010) 16:1695–700. doi: 10.1158/1078-0432.CCR-09-1805
9. Thompson AM. Molecular pathways: preclinical models and clinical trials with metformin in breast cancer. *Clin Cancer Res.* (2014) 20:2508–15. doi: 10.1158/1078-0432.CCR-13-0354
10. Heckman-Stoddard BM, DeCensi A, Sahasrabudhe VV, Ford LG. Repurposing metformin for the prevention of cancer and cancer recurrence. *Diabetologia.* (2017) 60:1639–47. doi: 10.1007/s00125-017-4372-6
11. Xu H, Chen K, Jia X, Tian Y, Dai Y, Li D, et al. Metformin use is associated with better survival of breast cancer patients with diabetes: a meta-analysis. *Oncologist.* (2015) 20:1236–44. doi: 10.1634/theoncologist.2015-0096
12. Jiralerspong S, Palla SL, Giordano SH, Meric-Bernstam F, Liedtke C, Barnett CM, et al. Metformin and pathologic complete responses to neoadjuvant chemotherapy in diabetic patients with breast cancer. *J Clin Oncol.* (2009) 27:3297–302. doi: 10.1200/JCO.2009.19.6410
13. Goodwin PJ, Stambolic V, Lemieux J, Chen BE, Parulekar WR, Gelmon KA, et al. Evaluation of metformin in early breast cancer: a modification of the traditional paradigm for clinical testing of anti-cancer agents. *Breast Cancer Res Treat.* (2011) 126:215–20. doi: 10.1007/s10549-010-1224-1
14. Del Barco S, Vazquez-Martin A, Cufi S, Oliveras-Ferraro C, Bosch-Barrera J, Joven J, et al. Metformin: multi-faceted protection against cancer. *Oncotarget.* (2011) 2:896–917. doi: 10.18632/oncotarget.387
15. Foretz M, Guigas B, Bertrand L, Pollak M, Viollet B. Metformin: from mechanisms of action to therapies. *Cell Metab.* (2014) 20:953–66. doi: 10.1016/j.cmet.2014.09.018
16. Dowling RJ, Goodwin PJ, Stambolic V. Understanding the benefit of metformin use in cancer treatment. *BMC Med.* (2011) 9:33. doi: 10.1186/1741-7015-9-33
17. Dowling RJ, Niraula S, Stambolic V, Goodwin PJ. Metformin in cancer: translational challenges. *J Mol Endocrinol.* (2012) 48:R31–43. doi: 10.1530/JME-12-0007
18. Chandel NS, Avizonis D, Reczek CR, Weinberg SE, Menz S, Neuhaus R, et al. Are metformin doses used in murine cancer models clinically relevant? *Cell Metab.* (2016) 23:569–70. doi: 10.1016/j.cmet.2016.03.010
19. Menendez JA, Martin-Castillo B, Joven J. Metformin and cancer: Quo vadis et cui bono? *Oncotarget.* (2016) 7:54096–101. doi: 10.18632/oncotarget.10262
20. Kordes S, Pollak MN, Zwinderman AH, Mathôt RA, Weterman MJ, Beeker A, et al. Metformin in patients with advanced pancreatic cancer: a double-blind, randomised, placebo-controlled phase 2 trial. *Lancet Oncol.* (2015) 16:839–47. doi: 10.1016/S1470-2045(15)00027-3
21. Reni M, Dugnani E, Cereda S, Belli C, Balzano G, Nicoletti R, et al. (Ir)relevance of metformin treatment in patients with metastatic pancreatic cancer: an open-label, randomized phase II trial. *Clin Cancer Res.* (2016) 22:1076–85. doi: 10.1158/1078-0432.CCR-15-1722
22. Goodwin PJ, Parulekar WR, Gelmon KA, Shepherd LE, Ligibel JA, Hershman DL, et al. Effect of metformin vs placebo on and metabolic factors in NCIC CTG MA.32. *J Natl Cancer Inst.* (2015) 107:djv006. doi: 10.1093/jnci/djv006
23. Martin-Castillo B, Dorca J, Vazquez-Martin A, Oliveras-Ferraro C, Lopez-Bonet E, Garcia M, et al. Incorporating the antidiabetic drug metformin in HER2-positive breast cancer treated with neo-adjuvant chemotherapy and trastuzumab: an ongoing clinical-translational research experience at the Catalan Institute of Oncology. *Ann Oncol.* (2010) 21:187–9. doi: 10.1093/annonc/mdp494
24. Martin-Castillo B, Pernas S, Dorca J, Álvarez I, Martínez S, Pérez-García JM, et al. A phase 2 trial of neoadjuvant metformin in combination with trastuzumab and chemotherapy in women with early HER2-positive breast cancer: the METTEN Study. *Oncotarget.* (2018) 9:35687–704. doi: 10.18632/oncotarget.26286
25. Zhao B, Zhao H. Impact of clinicopathological characteristics on the efficacy of neoadjuvant therapy in patients with human epidermal growth factor receptor-2-positive breast cancer. *Int J Cancer.* (2018) 142:844–53. doi: 10.1002/ijc.31097
26. Fernandez-Real JM, Pugeat M, Grasa M, Broch M, Vendrell J, Brun J, et al. Serum corticosteroid-binding globulin concentration and insulin resistance syndrome: a population study. *J Clin Endocrinol Metab.* (2002) 87:4686–90. doi: 10.1210/jc.2001-011843
27. Kalaany NY, Sabatini DM. Tumours with PI3K activation are resistant to dietary restriction. *Nature.* (2009) 458:725–31. doi: 10.1038/nature07782
28. Brunet A. Cancer: when restriction is good. *Nature.* (2009) 458:713–14. doi: 10.1038/458713a
29. Cufi S, Corominas-Faja B, Lopez-Bonet E, Bonavia R, Pernas S, López IÁ, et al. Dietary restriction-resistant human tumors harboring the PIK3CA-activating mutation H1047R are sensitive to metformin. *Oncotarget.* (2013) 4:1484–95. doi: 10.18632/oncotarget.1234
30. Hopkins BD, Pauli C, Du X, Wang DG, Li X, Wu D, et al. Suppression of insulin feedback enhances the efficacy of PI3K inhibitors. *Nature.* (2018) 560:499–503. doi: 10.1038/s41586-018-0343-4
31. Cortazar P, Zhang L, Untch M, Mehta K, Costantino JP, Wolmark N, et al. Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. *Lancet.* (2014) 384:164–72. doi: 10.1016/S0140-6736(13)62422-8
32. Broglio KR, Quintana M, Foster M, Olinger M, McGlothlin A, Berry SM, et al. Association of pathologic complete response to neoadjuvant therapy in HER2-positive breast cancer with long-term outcomes: a meta-analysis. *JAMA Oncol.* (2016) 2:751–60. doi: 10.1001/jamaoncol.2015.6113
33. Untch M, von Minckwitz G, Gerber B, Schem C, Rezai M, Fasching PA, et al. Survival analysis after neoadjuvant chemotherapy with trastuzumab or lapatinib in patients with human epidermal growth factor receptor 2-positive breast cancer in the GeparQuinto (G5) Study (GBG 44). *J Clin Oncol.* (2018) 36:1308–16. doi: 10.1200/JCO.2017.75.9175
34. Bartsch R, de Azambuja E. I-SPY 2: optimising cancer drug development in the 21st century. *ESMO Open.* (2016) 1:e000113. doi: 10.1136/esmoopen-2016-000113

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Cuyàs, Buxó, Ferri Iglesias, Verdura, Pernas, Dorca, Álvarez, Martínez, Pérez-García, Batista-López, Rodríguez-Sánchez, Amillano, Domínguez, Luque, Morilla, Stradella, Viñas, Cortés, Joven, Brunet, López-Bonet, García, Saidani, Queralt Moles, Martin-Castillo and Menendez. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Establishment and Verification of a Bagged-Trees-Based Model for Prediction of Sentinel Lymph Node Metastasis for Early Breast Cancer Patients

Chao Liu<sup>1†</sup>, Zeyin Zhao<sup>2†</sup>, Xi Gu<sup>1†</sup>, Lisha Sun<sup>1</sup>, Guanglei Chen<sup>1</sup>, Hao Zhang<sup>1</sup>, Yanlin Jiang<sup>1</sup>, Yixiao Zhang<sup>3\*</sup>, Xiaoyu Cui<sup>2\*</sup> and Caigang Liu<sup>1\*</sup>

## OPEN ACCESS

### Edited by:

Mothaffar Rimawi,  
Baylor College of Medicine,  
United States

### Reviewed by:

Yongliang Yang,  
Dalian University of Technology (DUT),  
China  
Ramin Sadeghi,  
Mashhad University of Medical  
Sciences, Iran

### \*Correspondence:

Yixiao Zhang  
zhangyx201@hotmail.com  
Xiaoyu Cui  
cuixy@bme.neu.edu.cn  
Caigang Liu  
angel-s205@163.com

<sup>†</sup>Co-first authors

### Specialty section:

This article was submitted to  
Women's Cancer,  
a section of the journal  
Frontiers in Oncology

**Received:** 19 January 2019

**Accepted:** 27 March 2019

**Published:** 16 April 2019

### Citation:

Liu C, Zhao Z, Gu X, Sun L, Chen G,  
Zhang H, Jiang Y, Zhang Y, Cui X and  
Liu C (2019) Establishment and  
Verification of a Bagged-Trees-Based  
Model for Prediction of Sentinel Lymph  
Node Metastasis for Early Breast  
Cancer Patients. *Front. Oncol.* 9:282.  
doi: 10.3389/fonc.2019.00282

<sup>1</sup> Department of Breast Surgery, Shengjing Hospital of China Medical University, Shenyang, China, <sup>2</sup> Sino-Dutch Biomedical and Information Engineering School, Northeastern University, Shenyang, China, <sup>3</sup> Department of Urology Surgery, Shengjing Hospital of China Medical University, Shenyang, China

**Purpose:** Lymph node metastasis is a multifactorial event. Several scholars have developed nomograph models to predict the sentinel lymph nodes (SLN) metastasis before operation. According to the clinical and pathological characteristics of breast cancer patients, we use the new method to establish a more comprehensive model and add some new factors which have never been analyzed in the world and explored the prospect of its clinical application.

**Materials and methods:** The clinicopathological data of 633 patients with breast cancer who underwent SLN examination from January 2011 to December 2014 were retrospectively analyzed. Because of the imbalance in data, we used smote algorithm to oversample the data to increase the balanced amount of data. Our study for the first time included the shape of the tumor and breast gland content. The location of the tumor was analyzed by the vector combining quadrant method, at the same time we use the method of simply using quadrant or vector for comparing. We also compared the predictive ability of building models through logistic regression and Bagged-Tree algorithm. The Bagged-Tree algorithm was used to categorize samples. The SMOTE-Bagged Tree algorithm and 5-fold cross-validation was used to established the prediction model. The clinical application value of the model in early breast cancer patients was evaluated by confusion matrix and the area under receiver operating characteristic (ROC) curve (AUC).

**Results:** Our predictive model included 12 variables as follows: age, body mass index (BMI), quadrant, clock direction, the distance of tumor from the nipple, morphology of tumor molybdenum target, glandular content, tumor size, ER, PR, HER2, and Ki-67. Finally, our model obtained the AUC value of 0.801 and the accuracy of 70.3%. We used logistic regression to established the model, in the modeling and validation groups, the area under the curve (AUC) were 0.660 and 0.580. We used the vector combining quadrant method to analyze the original location of the tumor, which is more precise than



simply using vector or quadrant (AUC 0.801 vs. 0.791 vs. 0.701, Accuracy 70.3 vs. 70.3 vs. 63.6%).

**Conclusions:** Our model is more reliable and stable to assist doctors predict the SLN metastasis in breast cancer patients before operation.

**Keywords:** breast cancer, sentinel lymph nodes, metastasis prediction, model, bagged-trees

## INTRODUCTION

The incidence of breast cancer is the first in female malignant tumors, in which the highest incidence of breast cancer has been reported in Europe and the United States, however, in recent years, the incidence of breast cancer in China has annually increased (1, 2). Based on surgery as an important step in the treatment of breast cancer, in recent years, different individuals have never stopped exploration of a novel and optimum approach. Besides, NSABP-04, ASCOG-Z0011, and other tests have shown that for breast cancer patients with T1-T2 stage and clinical negative lymph node (cN0) during breast preservation surgery and total breast radiotherapy, axillary lymph node dissection (ALND) does not contain great benefits to the long-term survival of patients. As a result, sentinel lymph node biopsy (SLNB) has been gradually replaced with conventional ALND as a routine surgical method for early breast cancer patients (3–7). However, SLNB, as an invasive operation, leads to some postoperative complications. Although the corresponding incidence rate is lower than ALND, however, those complications should not be ignored. Moreover, SLNB has a high degree of professional requirement for physicians, and the richness of physician's experience directly affects the evaluation of the pathological status of sentinel lymph nodes (SLN).

In recent years, the concepts of precision medicine and individualized therapy have rapidly developed. We often ask the following questions: “Can SLNB be omitted for patients with lower probability of sentinel lymph node metastasis?” “OR For patients with higher probability of sentinel lymph node metastasis, can the SLNB be skipped and the armpit treatment to be directly conducted?” OR “For patients with neoadjuvant chemotherapy, should we have an SLNB before neoadjuvant chemotherapy, or after it?” The state of axillary lymph nodes is not only a key factor in determining the mode of surgery, but also an important prognostic factor, and before surgery, patients often would like to know whether there is a transfer of SLN. With the idea of micro-non-invasive operation, several scholars have developed and used mathematical models to predict the pathological state of SLN before operation, in which the most important predictive model was designed

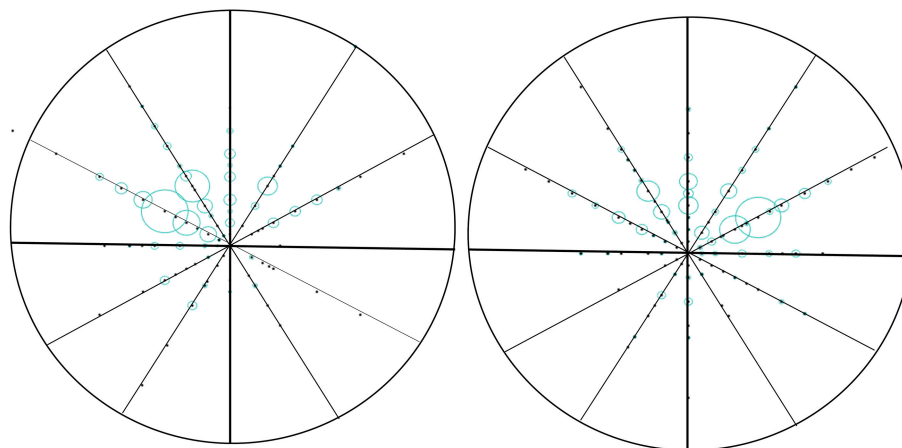
in 2007 at Memorial Sloan-Caitlin Cancer Center (MSKCC; NY, USA). It has been shown that with a receiver operating characteristic (ROC) curve of 0.75, a proper level of prediction and discrimination can be achieved (8–24). However, there are differences in the sources of patients (ethnic, regional, cultural, economic conditions, disease awareness, etc.), surgical methods, pathological evaluation methods, and other factors. Hence, it is difficult to have a predictive model that can be universally used. The clinical and pathological parameters for the application of different predictive models are not the same. Hence, the purpose of our study was to analyze the clinical and pathological data of early breast cancer patients in a more comprehensive way, and establish a predictive model for sentinel lymph node pathology. Technically, nomogram which is now used worldwide use multiple logistic regression (MLR) to predict a binary outcome based on a combination of risk factors. This well-established method has a limitation in that it incorporates only a few independent variables so that the model can accurately predict risk in independent datasets, by avoiding over-fitting to the given datasets. Such prediction models should also tolerate missing values, which are common in clinical datasets (15). Thus, we use SMOTE-Bagged Tree as a core algorithm to cope with a greater number of variables and that provide accurate prediction and robustness against missing values. In addition to the variables analyzed by other scholars, we added some specific variables, such as breast glandular content, molybdenum target tumor morphology, and primary location of the tumor (clock direction and distance from the nipple). Our study for the first time presents a model to analyze these factors as well.

## MATERIALS AND METHODS

### Patients

In this study, 633 patients with a clear state through sentinel lymph node examination (including lymph node biopsy and surgical treatment with ALND) were included. Analysis of clinical data involves the following variables: age, body mass index (BMI), tumor size, tumor location (quadrant, clock direction, distance from the nipple), clinical staging, pathological type, pathological classification, immunohistochemistry (IHC) [Estrogen receptor (ER), Progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2), (Ki-67)], grading of tumor tissue, menopausal state, molybdenum target glandular content, and morphology and sentinel lymph node metastasis of molybdenum. The patient's information was derived from the Shengjing Hospital of China Medical University (Shenyang, China) during January 2011 to December 2014. The patients with early-stage breast cancer who met the following criteria

**Abbreviations:** SLN, Sentinel lymph node; SLNB, Sentinel lymph node biopsy; SLN+, Sentinel lymph node positive; SLN-, Sentinel lymph node negative; ALN, Axillary lymph node; ALND, Axillary lymph node dissection; MLR, Multiple logistic regression; SMOTE, Synthetic minority oversampling; UIQ, Upper inner quadrant; UOQ, Upper outer quadrant; LOQ, Lower outer quadrant; LIQ, Lower inner quadrant; Central, Central quadrant; ROC, Receiver operating characteristics curve; AUC, Area under the receiver operating characteristics curve; BMI, Body mass index; IHC, Immunohistochemistry; ER, Estrogen receptor; PR, Progesterone receptor; HER2, Human epidermal growth factor receptor 2.



**FIGURE 1 |** Vector diagram of the location of the tumor in breast. We counted the location of the tumor in the direction of the clock and the distance from the nipple (the factor of the clock in the left and right sides of the breast has been taken into account). The number of cases transferred at the same location is proportional to the radius, and then the nipple is the center of the circle, in which the distance is the radius of the mapping.

were selected for treatment with SLNB: (1) diagnosis of breast cancer; (2) without receiving neoadjuvant chemotherapy; (3) the result for preoperative axillary lymph node to be negative, according to clinical and imaging examinations; (4) Primary diameter of tumor ranges at 0–5 cm; (5) Complete clinicopathological information; and (6) without pregnancy. Patients with incomplete data, metastasis of 3 and above axillary lymph nodes, distant metastasis, preoperative neoadjuvant chemotherapy, and radiotherapy were excluded. All patients involved in the present study signed a written informed consent form prior to their inclusion in the study. The study was approved by the Ethics Committee of Shengjing Hospital of China Medical University.

## Surgery Procedure

A standard breast cancer surgery was conducted on the basis of guidelines for the treatment of breast cancer patients in China. Surgery included primary tumor resection and lymph node biopsy (or ALND). The number and pathological status of SLN were detected after operation as well.

## Pathologic Evaluation

The Chinese breast cancer guidelines were used to evaluate surgical specimens. Tumors with > 10% nuclear-stained cells were considered positive for ER and PR. Ki67 expression > 20% was also considered positive. The HER-2 positivity was defined as a score of 3+ on IHC or amplification on FISH. If a pathologist scored the IHC 2+, the status of HER-2 was further investigated by FISH. In addition, the grade of breast cancer was determined by the Nottingham Histologic Scoring system. Tumor staging refers to the TNM staging method jointly conducted and published in form of the 8th edition by the International Anticancer Alliance (UICC) and the American Oncology Federation (AJCC) in 2018. The SLNs were step sectioned, stained with hematoxylin and eosin (H&E), and diagnosed by trained pathologists. Lymph nodes obtained after

ALND were evaluated using a single H&E stained section from each node. Metastases were defined as the presence of a tumor deposit >0.2 mm in diameter in at least one lymph node.

## Location of the Tumor

We use the polar coordinates to paint. We counted the location of the tumor in the direction of the clock and the distance from the nipple (the factor of the clock in the left and right sides of the breast has been taken into account). The number of cases transferred at the same location is proportional to the radius, and then the nipple is the center of the circle, in which the distance is the radius of the mapping (**Figure 1**).

## Statistical Analysis

In this study, MATLAB2018a was used for data processing, and statistical analysis was undertaken by using SPSS 25.0 statistical software (SPSS Inc., Chicago, IL, USA) and R software. The statistical significance level of the report was double-sided, and it was set to 0.05.

## Smote Algorithm Generates Data

There are a number of methods available to oversample a dataset used in a typical classification problem (using a classification algorithm to classify a set of images, given a labeled training set of images). The most common technique is known as SMOTE: Synthetic Minority Over-sampling Technique (25).

According to the natural occurrence of disease, breast cancer grows more in the outer upper quadrant. However, in the process of regular grouping, such case characteristic data is unbalanced (the proportion of the outer upper quadrant tumor is larger), which will affect the broad applicability of the model, so it is not appropriate to use the classifier to distinguish directly. Therefore, this study used the most appropriate smote algorithm to upsample the data to increase the balanced amount of data, reconstruct the training set, and obtain a relatively balanced training set.

The basic idea of smote algorithm is to analyze a few kinds of samples and add them to the dataset according to a few samples of synthetic samples, and the algorithm flow is as follows. (1) For each sample  $x$  in a few classes, the Euclidean distance is used as the criterion for calculating the distance of all samples in a few sample sets, and its  $k$  neighbor is obtained. (2) Set a sampling scale according to the unbalanced proportion of the sample to determine the sampling magnification  $n$ , for each minority sample  $X$ , from its  $K$  near neighbor randomly selected a number of samples, assuming that the selected neighbor is  $x_n$ . (3) For each randomly selected near neighbor  $X_n$ , a new sample is constructed with the original sample according to the following formula.

$$x_{new} = x + rand(0, 1) \times (\tilde{x} - x)$$

## Confusion Matrix

Confusion matrix is an important tool to evaluate the performance of classification model. A variety of evaluation indexes, such as true positive rate, false positive rate, true negative rate, false negative rate and accuracy, can be calculated by the confusion matrix. In particular, the confusion matrix distinguishes between false positives and false negatives of two different properties of miscalculation, which can be used to estimate the expected loss caused by miscalculation of the classification model. When the classification model returns the probability or score of each record belonging to the positive category, a confusion matrix can be obtained by specifying the threshold and making a positive judgment on all the probabilities or records that are rated above the threshold. By continuously changing the threshold value, different confusion matrices can be obtained, so that the ROC curve, and the performance of the classification model is evaluated and compared more comprehensively.

## Establishment of Predictive Models

First, we use the logistic regression method which was commonly used in other research centers to build prediction models, and then use the Bagged Tree algorithm to build prediction models. In the end we compared the results obtained by the two methods. The amplified data were classified using the Bagged Tree algorithm. It was used to analyze the following clinical candidate predictors: age, BMI, quadrant, clock direction, the distance of tumor from the nipple, morphology of tumor molybdenum target, glandular content, tumor size, ER, PR, HER2, and Ki-67. The trees in Bagged Trees were built on their own sampled datasets, and the training process was independent. The Bagged Tree algorithm extracts multiple random datasets to fit multiple decision tree models in order to improve model's performance. Each decision tree differs because of the subset data, and the final prediction results are determined based on the prediction of all trees (26). Accordingly, the versatility of predictions increases. The ROC curve was plotted, and the area under the curve (AUC) was here used to estimate the prediction accuracy of the model.

## RESULTS

### Patient Characteristics

Our study included 633 patients who underwent sentinel lymph node examination, in which 35.8% of whom had sentinel lymph node metastasis. The descriptive characteristics of the model population are presented in **Table 1**. In fact, the characteristic data of lymph node metastasis in breast cancer were not balanced, and utilizing a classifier to directly distinguish was not appropriate. Therefore, the SMOTE was used to sample the breast cancer dataset and reduce the imbalance of the training set. An oversampling algorithm was also used to add new information to the unbalanced data. In our study, 633 cases of raw data were sampled by SMOTE, in which 169 new data were generated according to the characteristics of the original data. We analyzed the newly generated data and raw data (a total of 802 cases) together.

### Predictors of SLNM

In the multi-factor analysis, we included the following variables associated with breast cancer SLN metastasis: age, BMI, quadrant, clock direction, the distance of tumor from the nipple, morphology of tumor molybdenum target, glandular content, tumor size, ER, PR, HER2, and Ki-67 (**Figure 2**).

### Construction and Validation of the Model by Logistic Regression

We are grateful for your advice. Our study included 633 patients randomized into a modeling set ( $n = 500$ ) and a validation set ( $n = 133$ ). The clinicopathological characteristics of the patients did not differ significantly between the two groups ( $P > 0.05$ ) in our study population (**Table 2**). The internal ROC curves in the modeling set and external ROC in the validation set were used to evaluate the model. In the modeling and validation groups, the AUC were 0.660 and 0.580 (**Figures 3, 4**).

### Construction and Validation of the Model by SMOTE-Bagged Tree Algorithm

A bagged tree was used to categorize samples. In order to obtain a reliable and stable model, a 5-fold cross-validation was used for verification. A Sentinel lymph node prediction program was established by SMOTE-bagged tree algorithm. Since the original location of the tumor was analyzed by the combination of vector and quadrant for the first time, we used the method of simply using quadrant or vector for comparison. Finally, our model obtained the AUC value of 0.801 (**Figure 5**), while the vector group is 0.791 (**Figure 6**) and the quadrant group is 0.701 (**Figure 7**). We use the confusion matrix to evaluate the accuracy of the model. The accuracy of our model is 70.3% (**Figure 8A**), compared to 70.3% for the vector group (**Figure 8B**) and 63.6% for the quadrant group (**Figure 8C**). The mentioned method provides an accurate and credible multi-variable prediction model.

The inclusion of imaging and pathological detection factors in an easy-to-use machine learning model facilitates the prediction of lymph node metastasis in patients before surgery. Because lymph node metastasis is similar in training set and validation

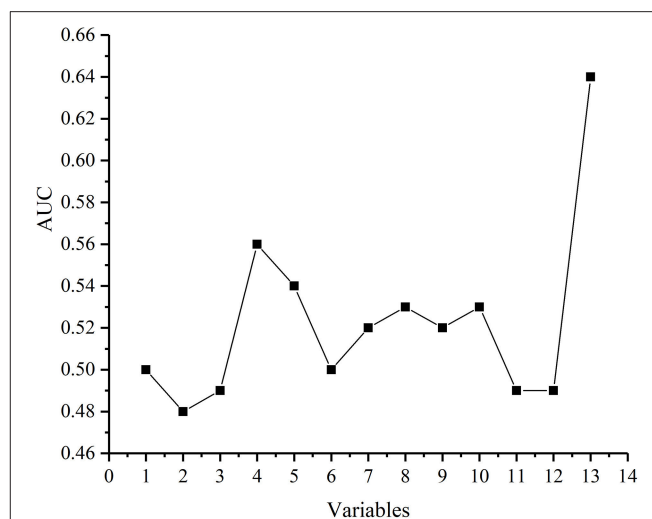
**TABLE 1 |** Comparison of descriptive characteristics of the SLN+ group and SLN- group for the model.

Variable	N	SLN+		SLN-		P-value
		Population	%	Population	%	
No. of cases	633	227	35.86	406	64.14	
Age						0.536
<40	52	17	32.69	35	67.31	
40–70	547	195	35.65	352	64.35	
>70	34	15	44.12	19	55.88	
AJCC T stage						0.149
T1a	1	0	0.00	1	100.00	
T1b	7	0	0.00	7	100.00	
T1c	349	118	33.81	231	66.19	
T2 < 3.0 cm	262	103	39.31	159	60.69	
T2 > 3.0 cm	14	6	42.86	8	57.14	
Left/Right breast						0.511
left	318	127	39.94	191	60.06	
right	315	100	31.75	215	68.25	
Neuroinvasion						0.102
Present	4	3	75.00	1	25.00	
Absent	629	224	35.61	405	64.39	
Lymphovascular invasion						0.232
Present	16	8	50.00	8	50.00	
Absent	617	219	35.49	398	64.51	
Tumor location						0.441
UOQ	313	122	38.98	191	61.02	
LOQ	47	15	31.91	32	68.09	
UIQ	127	38	29.92	89	70.08	
LIQ	42	16	38.10	26	61.90	
Central	104	36	34.62	68	65.38	
Tumor type						0.000
Ductal	609	211	34.65	398	65.35	
Lobular	10	4	40.00	6	60.00	
Mixed	10	10	100.00	0	0.00	
Mucinous	3	2	66.67	1	33.33	
Medullary	1	0	0.00	1	100.00	
Clinical stage						0.000
I	191	3	1.57	188	98.43	
II	440	223	50.68	217	49.32	
III	2	1	50.00	1	50.00	
Pathological grade						0.063
1	17	4	23.53	13	76.47	
2	582	205	35.22	377	64.78	
3	34	18	52.94	16	47.06	
ER						0.099
Present	481	181	37.63	300	62.37	
Absent	152	46	30.26	106	69.74	
PR						0.164
Present	427	161	37.70	266	62.30	

(Continued)

**TABLE 1 |** Continued

Variable	N	SLN+		SLN-		P-value
		Population	%	Population	%	
Absent	206	66	32.04	140	67.96	0.534
HER2						
Present	88	31	35.23	57	64.77	
Absent	545	196	35.96	349	64.04	0.960
Ki67						
Present	392	148	37.76	244	62.24	
Absent	241	79	32.78	162	67.22	0.214
Breast gland content						
<25%	99	31	31.31	68	68.69	
26–50%	184	42	22.83	142	77.17	
51–75%	305	113	37.05	192	62.95	
>75%	45	11	24.44	34	75.56	0.366
Tumor morphology						
Lump	378	130	34.39	248	65.61	
Mass with calcification	96	41	42.71	55	57.29	
No mass	159	56	35.22	103	64.78	0.726
BMI						
Too light	21	7	33.33	14	66.67	
Normal	313	108	34.50	205	65.50	
Overweight	299	112	37.46	187	62.54	

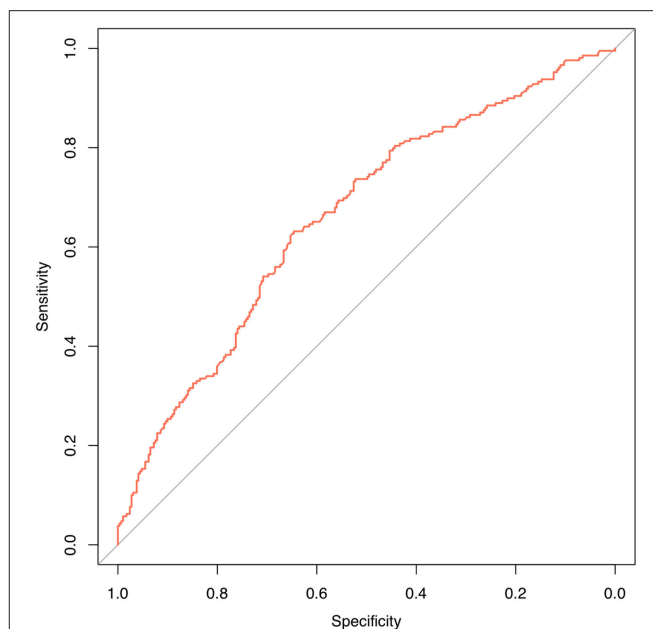
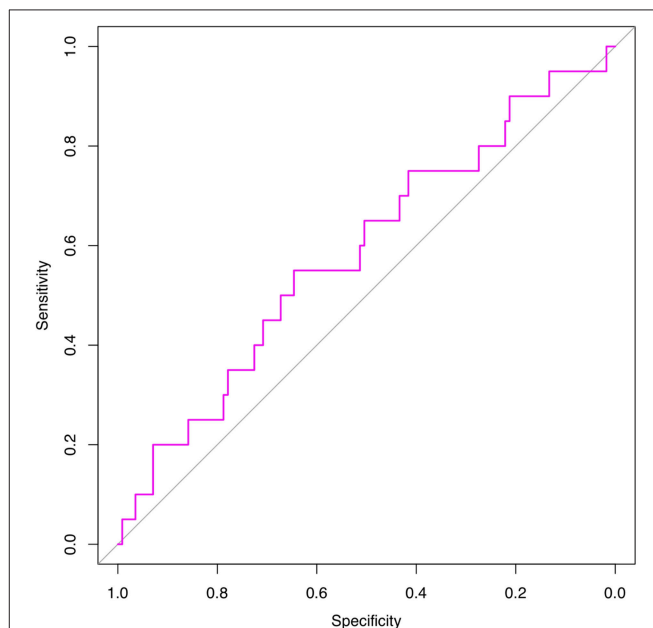
**FIGURE 2 |** The ability of each variable to predict breast cancer SLN metastasis. Each point represents in turn: Age, body mass index(BMI), Quadrant, Clock direction, The distance of tumor from the nipple, Morphology of tumor molybdenum target, Glandular content, Tumor size, ER, PR, HER2,Ki-67, Vector (Clock direction and The distance of tumor from the nipple).

set, a predictive model for breast cancer lymph node metastasis based on Bagged Tree algorithm is of great importance, and can be directly applied to verify dataset. The experimental results



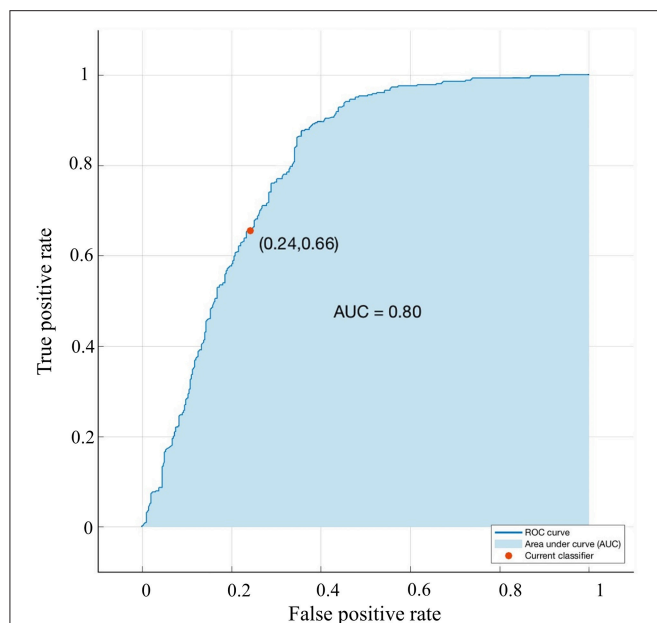
**TABLE 2 |** Comparison between modeling group and validation group by clinicopathological characteristics.

Variable	N	Modeling		Validation		P-value
		population	%	population	%	
No. of cases	633	500		133		
Age						0.892
<40	52	42	80.80	10	18.20	
40–70	547	432	79.00	115	21.00	
>70	34	26	76.50	8	5.40	
Tumor location						0.178
UOQ	313	241	77.00	72	23.00	
LOQ	47	42	89.40	5	10.60	
UIQ	127	96	75.60	31	24.40	
LIQ	42	36	85.70	6	4.30	
Central	104	85	81.70	19	18.30	
Tumor type						0.439
Ductal	609	483	79.3	126	20.70	
Lobular	10	6	60.00	4	40.00	
Mixed	10	7	70.00	3	30.00	
Mucinous	3	3	100.00	0	0.00	
Medullary	1	1	100.00	0	0.00	
Clinical stage						0.270
I	191	157	82.20	34	17.80	
II	440	342	77.70	98	22.30	
III	2	1	50.00	1	50.00	
Pathological grade						0.899
1	17	13	76.50	4	23.50	
2	582	461	79.20	161	20.80	
3	34	26	76.50	8	23.50	
ER						0.353
Present	481	384	79.80	97	20.20	
Absent	152	116	76.30	36	23.70	
PR						0.234
Present	427	343	80.30	84	19.70	
Absent	206	157	76.20	49	23.80	
HER2						0.802
Present	88	71	80.70	17	19.30	
Absent	545	427	78.80	115	21.20	
Ki67						0.960
Present	392	148	37.76	244	62.24	
Absent	241	79	32.78	162	67.22	
Breast gland content						0.769
<25%	99	82	82.80	17	17.20	
26–50%	184	144	78.30	40	21.70	
51–75%	305	238	78.00	67	22.00	
>75%	45	36	80.00	9	20.00	
Tumor morphology						0.056
Lump	378	311	82.33	67	17.70	
Mass with calcification	96	70	72.90	26	27.10	
No mass	159	119	74.80	40	25.20	
SLN						0.073
Present	227	170	74.90	57	24.10	
Absent	406	330	81.30	76	18.70	

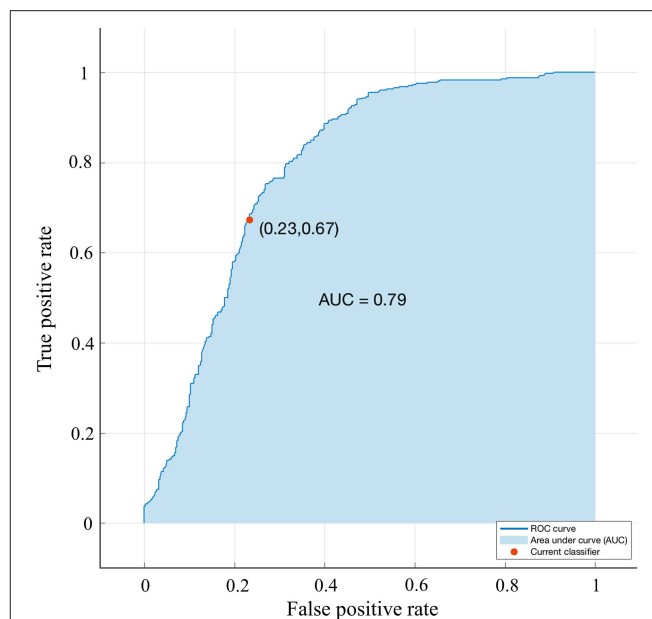
**FIGURE 3 |** Internal validation using a ROC curve. Established model by logistic regression. The AUC value is 0.660.**FIGURE 4 |** External validation using a ROC curve. Established model by logistic regression. The AUC value is 0.580.

showed that the prediction model for lymph node metastasis on the basis of SMOTE-Bagged Tee algorithm can effectively improve the rate of data utilization and assist doctors predict sentinel lymph node metastasis in breast cancer patients.

In order to make it easier to use in clinical applications, we have created an app that patients can use it easily by



**FIGURE 5 |** Validation using a ROC curve. Established model by Smote-Bagged-tree and used Vector combining quadrant method to analyze. The location of the tumor. The AUC value is 0.801.



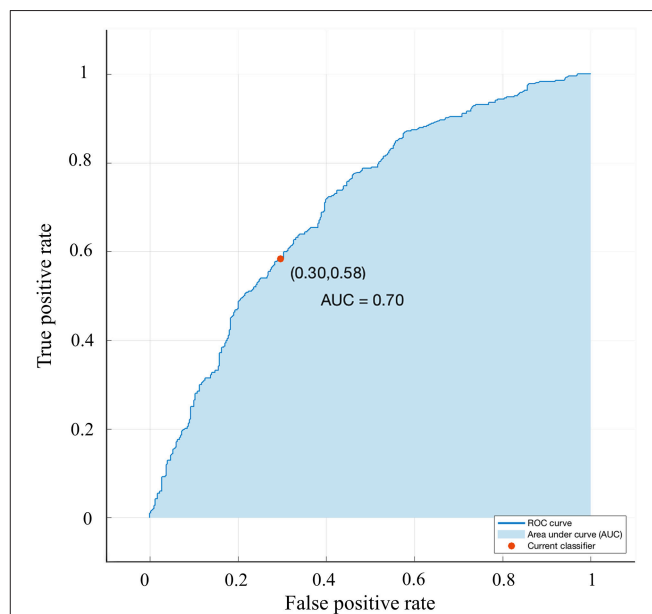
**FIGURE 6 |** Validation using a ROC curve. Established by Smote-Bagged-tree and used Simply using vector method to analyze the location of the tumor. The AUC value is 0.791.

a personal computer, laptop, or smart phone. The operation method is simple, only the user should input the patient's clinicopathological information into the app.

## DISCUSSION

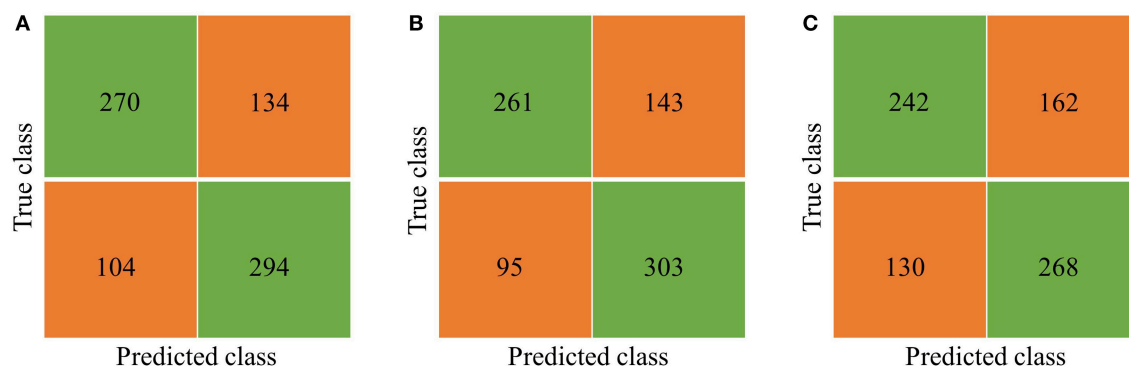
The prediction results obtained with the help of a predictive model are more credible than simple clinical guesses. The development of each predictive model is accomplished through the clinicopathological data of different populations. In our model, in terms of tumor size, we chose T1-T2 staging patients, because for the study of exemption from SLNB, the need for early clinical staging of the patient group, thus the size of tumor needs to be strictly limited. For type of tumor, we focused on invasive ductal carcinoma and mucous cancer. As invasive ductal carcinoma accounts for the vast majority, in order to avoid the formation of bias, the type of tumor was not taken into account in the inclusion of variables. Because chemotherapy can affect lymphatic vessels, SLNB has a high failure rate and a high false negative rate, hence, we ruled out patients with preoperative neoadjuvant chemotherapy. There were two variables for nerve infiltration and vascular thrombosis, and because the selected cases were early breast cancer patients, the number of cases in these two states was very limited. Moreover, these two variables could only be learned by breast cancer surgery, and the establishment of a preoperative prediction model was not practical, thus the variables of the model were not included.

In several studies, tumor size is the main predictor of sentinel lymph node metastasis (9–14). According to the results of statistical analysis, we also confirmed that the rate of positive



**FIGURE 7 |** Validation using a ROC curve. Established by Smote-Bagged-tree and used Simply using quadrant method to analyze the location of the tumor. The AUC value is 0.701.

SLN was lower in the group of patients with smaller tumors. This is also consistent with the results of NSABP-04, ASCOG-Z0011, IBCSG 23-01, and other tests. Axillary lymph node dissection (ALND) is not necessary for early breast cancer patients with 1 to 2 positive SLNs after undergoing lumpectomy, radiotherapy (RT), and systemic treatment (3, 4, 27–29).



**FIGURE 8 |** Confusion matrix. **(A)** Vector combining quadrant method, the Accuracy is 70.3%. **(B)** Simply using vector method, the Accuracy is 70.3%. **(C)** Simply using quadrant method, the Accuracy is 63.6%.

We compare and analyze the predictive ability of smote-bagged tree algorithm modeling and Logistic regression method modeling, and the results suggest that the model established by smote-bagged tree algorithm is more predictive. The main reasons are the following: (a) Because of the imbalance of data, using the Bagged-Tree algorithm can reduce the impact of data imbalance. It performs well on the category imbalance data and improves the prediction accuracy. (b) Feature normalization or standardization is not required. Especially when the scale of the feature is completely different or when the binary feature and the continuous feature exist simultaneously, the effect is very good.

To our knowledge, this is the first model analyzing and studying the original location of the tumor by both vector and quadrant. We analyzed the previously established axillary lymph node prediction models (8–15, 17–20), and found that there were different views on the effects of the primary location of the tumor on sentinel lymph node metastasis. The MSKCC model concluded that the risk of axillary lymph node metastasis in the upper inner quadrant was less, while there was no statistically different chance of axillary lymph node metastasis between other quadrants (11). In 2012, the SCH model proposed by Chinese scholars mentioned that in terms of the location of the tumor, the order of axillary lymph node metastasis from high to low should be the central quadrant, the lower inner quadrant, the outer upper quadrant, the lower outer quadrant, and the upper inner quadrant (20). This study takes into account the breast as a three-dimensional (3D) structure, thus we first proposed the primary location of the tumor through both vector and quadrant location analysis. By comparing the three methods, the results prove that: We used the vector combining quadrant method to analyze the original location of the tumor, which is more precise than simply using vector or quadrant. The combination of the two methods, the primary location of the tumor will be more precisely positioned.

BMI was used as an important factor in previous researches (11, 20, 30). To our knowledge, breast composition includes fat and glands, and BMI does not indicate the size of breast glands in the overall proportion of the breast, thus our study for the first

time presents the breast gland content as an independent factor for analysis.

Breast malignant tumor has its special form; for example, ultrasound is typically classified through the morphology, boundaries, activity, and other conditions of the tumor; molybdenum target will be classified according to the shape of the tumor and calcification (31). In retrospect, the morphology of the tumor was not analyzed as a factor in previous studies. Because the judgment of ultrasound is subjective, we use the tumor morphology under molybdenum target as a factor for statistical analysis. After deep learning of the image by computer, the picture of the breast molybdenum target image is automatically identified, and the shape of the tumor is scientifically grouped, and that is used as an important factor in the production of predictive software.

Looking back at the predictive models of other researches, the variables included in each model are different, and we believe that a predictive model should be simplified on the basis of ensuring accuracy, rather than only simply pursuing variables. Our predictive model contains 12 variables: age, BMI, quadrant, clock direction, the distance of tumor from the nipple, morphology of tumor molybdenum target, glandular content, tumor size, ER, PR, HER2, and Ki-67. Generally speaking, when the AUC value of a model is at the range of 0.7–0.8, the prediction ability of the model is superior. When the AUC value is at the range of 0.8–0.9, the prediction ability of the model is very good. Through the SMOTE-Bagged Tree algorithm, the AUC value obtained by our model is 0.80, which proves that our model has a proper prediction ability, and can be used for early breast cancer patients.

Our SLN prediction model appropriately predicts the risk of sentinel metastasis in patients. For patients with low risk of SLN metastasis, especially those who cannot tolerate SLNB surgery, or patients with high retention requirements for postoperative limb sensory function, no SLNB can be considered clinically to improve the patients' quality of life. For patients with high risk of SLN metastasis, axillary treatment can be directly performed, such as ALND, axillary radiotherapy, etc., especially for elderly or patients with poor foundation status, that can greatly shorten the

duration of operation time. For patients undergoing neoadjuvant chemotherapy, this model can be used to predict its SLN state before the new auxiliary chemotherapy, and our model does not include the patients after the new adjuvant radiotherapy, and does not apply to the prediction after the new adjuvant chemotherapy, limiting the prediction ability of the model.

In clinical work, more and more patients are eager to understand the pathological state of SLN before surgery. Predictive results obtained using objective predictive models are more believable than pure clinical guesses. According to the sentinel lymph node prediction model established by this research data, the overall prediction ability is very high, the result of the ROC Curve area is 0.801, which suggests that our prediction model has good predictive ability and strong stability, so we believe that the model can be generally applied to other groups of people. Compared with other models, we used the factors for the first time as follows: the location of the tumor was analyzed by vector combining quadrant method, the content of the breast glands, and the shape of the tumor, which caused that our model to be more sophisticated. Since the variables required by the model can be obtained by ultrasound, molybdenum target, hollow core needle biopsy, or open biopsy, the patient is able to learn the risk of SLN metastasis before the operation to predict the prognosis of the disease.

However, our predictive model contains some limitations. Firstly, in the next study, for the variant of tumor morphology under molybdenum target, we should incorporate more data, and continue to improve the depth of machine learning, so that it can be used to more detailed grouping of tumor morphology, and strive to further identify calcification patterns. Secondly, the breast is such a 3D structure, in which for finding the location of the tumor, we will then improve its grouping, and strive to complete the 3D positioning. In addition, this is a retrospective and single-center study, and our model has a proper diagnostic

ability, however, it needs to be further validated in other regions and populations.

## CONCLUSIONS

In summary, we have established an accurate, reliable, and user-friendly multi-variable predictive model. By adding several variables that have never been used in previous models, our model can be used to predict the risk of sentinel lymph node metastasis before breast cancer surgery, and it provide a reliable basis for the treatment of axillary lymph nodes.

## ETHICS STATEMENT

The study was granted ethical approval by the Ethical Committee of China Medical University and the Shengjing Hospital of China Medical University. All the patients provided written informed consent.

## AUTHOR CONTRIBUTIONS

ChL, ZZ, XG, YZ, XC, and CaL contributed conception and design of the study. XG, LS, GC, HZ, and YJ organized the database. ChL, ZZ, and LS performed the statistical analysis. ChL and ZZ wrote the first draft of the manuscript. ChL, ZZ, XG, YZ, XC, and CaL wrote sections of the manuscript. All authors contributed to manuscript revision, read and approved the submitted version.

## FUNDING

This study was supported by the National Natural Science Foundation of China (Grant 81572609 and 31601142).

## REFERENCES

- DeSantis CE, Ma J, Goding Sauer A, Newman LA, Jemal A. Breast cancer statistics, 2017, racial disparity in mortality by state. *Cancer J Clin.* (2017) 67:439448. doi: 10.3322/caac.21412
- Hortobagyi GN, de la Garza Salazar J, Pritchard K, Amadori D, Haidinger R, Hudis CA, et al. The global breast cancer burden: variations in epidemiology and survival. *Clin. Breast Cancer.* (2005) 6:391–401. doi: 10.3816/CBC.2005.n.043
- Krag DN, Anderson SJ, Julian TB, Brown AM, Harlow SP, Ashikaga T, et al. Technical outcomes of sentinel-lymph-node resection and conventional axillary-lymph-node dissection in patients with clinically node-negative breast cancer: results from the NSABP B-32 randomised phase III trial. *Lancet Oncol.* (2007) 8:881–8. doi: 10.1016/S1470-2045(07)70278-4
- Fisher B, Bauer M, Wickerham DL, Redmond CK, Fisher ER, Cruz AB, et al. Relation of number of positive axillary nodes to the prognosis of patients with primary breast cancer. An NSABP update. *Cancer.* (1983) 52:1551–7. doi: 10.1002/1097-0142(19831101)52:9<1551::AID-CNCR2820520902>3.0.CO;2-3
- Mansel RE, Fallowfield L, Kissin M, Goyal A, Newcombe RG, Dixon JM, et al. Randomized multicenter trial of sentinel node biopsy versus standard axillary treatment in operable breast cancer: the ALMANAC Trial. *J. Natl. Cancer Inst.* (2006) 98:599–609. doi: 10.1093/jnci/djj158
- Veronesi U, Paganelli G, Viale G, Luini A, Zurrida S, Galimberti V, et al. Sentinel-lymph-node biopsy as a staging procedure in breast cancer: update of a randomised controlled study. *Lancet Oncol.* (2006) 7:983–90. doi: 10.1016/S1470-2045(06)70947-0
- Purushotham AD, Upsoni S, Klevesath MB, Bobrow L, Millar K, Myles JP, et al. Morbidity after sentinel lymph node biopsy in primary breast cancer: results from a randomized controlled trial. *J Clin Oncol.* (2005) 23:4312–21. doi: 10.1200/JCO.2005.03.228
- Perhavec A, Perme MP, Hocevar M, Besic N, Zgajnar J. Ljubljana nomograms for predicting the likelihood of non-sentinel lymph node metastases in breast cancer patients with a positive sentinel lymph node. *Breast Cancer Res. Treat.* (2010) 119:357–66. doi: 10.1007/s10549-009-0561-4
- Beriwal S, Soran A, Kocer B, Wilson JW, Ahrendt GM, Johnson R. Factors that predict the burden of axillary disease in breast cancer patients with a positive sentinel node. *Am J Clin Oncol.* (2008) 31:34–8. doi: 10.1097/COC.0b013e318068419b
- Canavese G, Bruzzi P, Catturich A, Vecchio C, Tomei D, Del Mastro L, et al. A risk score model predictive of the presence of additional disease in the axilla in early-breast cancer patients with one or two metastatic sentinel lymph nodes. *Eur J Surg Oncol.* (2014) 40:835–42. doi: 10.1016/j.ejso.2014.03.005
- Bevilacqua JL, Kattan MW, Fey JV, Cody HS, III, Borgen PI, Van Zee KJ. Doctor, what are my chances of having a positive sentinel node? A validated nomogram for risk estimation. *J Clin Oncol.* (2007) 25:3670–9. doi: 10.1200/JCO.2006.08.8013
- Reynders A, Brouckaert O, Smeets A, Laenen A, Yoshihara E, Persyn E, et al. Prediction of non-sentinel lymph node involvement in breast cancer

- patients with a positive sentinel lymph node, *Breast*. (2014) 23:453–9. doi: 10.1016/j.breast.2014.03.009
13. Barron AU, Hoskin TL, Boughey JC. Predicting non-sentinel lymph node metastases in patients with a positive sentinel lymph node after neoadjuvant chemotherapy. *Ann Surg Oncol*. (2018) 25:2867–74. doi: 10.1245/s10434-018-6578-3
  14. Wu P, Zhao K, Liang Y, Ye W, Liu Z, Liang C. Validation of breast cancer models for predicting the nonsentinel lymph node metastasis after a positive sentinel lymph node biopsy in a Chinese population. *Technol Cancer Res Treat*. (2018) 17:1533033818785032. doi: 10.1177/1533033818785032
  15. Takada M, Sugimoto M, Naito Y, Moon HG, Han W, Noh DY, et al. Prediction of axillary lymph node metastasis in primary breast cancer patients using a decision tree-based model. *BMC Med Inform Decis Mak*. (2012) 12:54. doi: 10.1186/1472-6947-12-54
  16. Klar M, Jochmann A, Foeldi M, Stumpf M, Gitsch G, Stickeler E, et al. The MSKCC nomogram for prediction the likelihood of non-sentinel node involvement in a German breast cancer population. *Breast Cancer Res. Treat*. (2008) 112:523–31. doi: 10.1007/s10549-007-9884-1
  17. Chen JY, Chen JJ, Xue JY, Chen Y, Liu GY, Han QX, et al. Predicting Non-sentinel lymph node metastasis in a chinese breast cancer population with 1-2 positive sentinel nodes: development and assessment of a new predictive nomogram. *World J Surg*. (2015) 39:2919–27. doi: 10.1007/s00268-015-3189-z
  18. Liu C, Jiang Y, Gu X, Xu Z, Ai L, Zhang H, et al. Predicting level 2 axillary lymph node metastasis in a Chinese breast cancer population post-neoadjuvant chemotherapy: development and assessment of a new predictive nomogram. *Oncotarget*. (2017) 8:79147–56. doi: 10.18632/oncotarget.16131
  19. Zheng J, Cai S, Song H, Wang Y, Han X, Wu H, et al. Positive non-sentinel axillary lymph nodes in breast cancer with 1-2 sentinel lymph node metastases. *Medicine*. (2018) 97:e13015. doi: 10.1097/MD.00000000000013015
  20. Chen JY, Chen JJ, Yang BL, Liu ZB, Huang XY, Liu GY, et al. Predicting sentinel lymph node metastasis in a Chinese breast cancer population: assessment of an existing nomogram and a new predictive nomogram. *Breast Cancer Res Treat*. (2012) 135:839–48. doi: 10.1007/s10549-012-2219-x
  21. Koca B, Kuru B, Ozen N, Yoruker S, Bek Y. A breast cancer nomogram for prediction of non-sentinel node metastasis - validation of fourteen existing models. *Asian Pac J Cancer Prev*. (2014) 15:1481–8. doi: 10.7314/APJCP.2014.15.3.1481
  22. Qiu SQ, Aarnink M, van Maaren MC, Dorrius MD, Bhattacharya A, Veltman J, et al. Validation and update of a lymph node metastasis prediction model for breast cancer. *Eur J Surg Oncol*. (2018) 44:700–7. doi: 10.1016/j.ejso.2017.12.008
  23. van den Hoven I, Kuijt G, Roumen R, Voogd A, Steyerberg EW, Vergouwe Y. A head to head comparison of nine tools predicting non-sentinel lymph node status in sentinel node positive breast cancer women. *J Surg Oncol*. (2015) 112:133–8. doi: 10.1002/jso.23992
  24. Yildiz R, Urkan M, Hancerliogullari O, Kilbas Z, Ozturk E, Menten MO, et al. Comparison of five different popular scoring systems to predict nonsentinel lymph node status in patients with metastatic sentinel lymph nodes: a tertiary care center experience. *Springerplus*. (2015) 4:651. doi: 10.1186/s40064-015-1442-4
  25. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. *J Artif Intell Res*. (2002) 16:321–57. doi: 10.1613/jair.953
  26. Breiman L. Bagging predictors. *Mach Learn*. (1996) 24:123–40. doi: 10.1007/BF00058655
  27. Giuliano AE, Hunt KK, Ballman KV, Beitsch PD, Whitworth PW, Blumencranz W, et al. Axillary dissection vs no axillary dissection in women with invasive breast cancer and sentinel node metastasis: a randomized clinical trial. *JAMA*. (2011) 305:569–75. doi: 10.1001/jama.2011.90
  28. Sakorafas GH, Peros G, Cataliotti L, Vlastos G. Lymphedema following axillary lymph node dissection for breast cancer. *Surg Oncol*. (2006) 15:153–65. doi: 10.1016/j.suronc.2006.11.003
  29. della Rovere GQ, Bonomi R, Ashley S, Benson JR. Axillary staging in women with small invasive breast tumours. *Eur J Surg Oncol*. (2006) 32:733–7. doi: 10.1016/j.ejso.2006.04.016
  30. Otsuka M, Harkness E, Chen X. Local breast density at lesion sites in diagnostic and previous screening mammograms. *Breast Cancer Res*. (2014) 16(Suppl. 1):04. doi: 10.1186/bcr3699
  31. Li H, Meng X, Wang T, Tang Y, Yin Y. Breast masses in mammography classification with local contour features. *Biomed Eng Online*. (2017) 16:44. doi: 10.1186/s12938-017-0332-0

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Liu, Zhao, Gu, Sun, Chen, Zhang, Jiang, Zhang, Cui and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Neoadjuvant Chemotherapy Alters Neuropilin-1, PlGF, and *SNAI1* Expression Levels and Predicts Breast Cancer Patients Response

Noura Al-Zeheimi<sup>1</sup>, Adviti Naik<sup>2</sup>, Charles Saki Bakheit<sup>3</sup>, Marwa Al Riyami<sup>4</sup>, Adil Al Ajarrah<sup>5</sup>, Suaad Al Badi<sup>4</sup>, Khalid Al Baimani<sup>6</sup>, Kamran Malik<sup>7</sup>, Zamzam Al Habsi<sup>5</sup>, Mansour S. Al Moundhri<sup>6</sup> and Sirin A. Adham<sup>1\*</sup>

<sup>1</sup> Department of Biology, College of Science, Sultan Qaboos University, Muscat, Oman, <sup>2</sup> Qatar Biomedical Research Institute, Hamad Bin Khalifa University, Doha, Qatar, <sup>3</sup> Department of Mathematics and Statistics, Sultan Qaboos University, Muscat, Oman, <sup>4</sup> Department of Pathology, College of Medicine, Sultan Qaboos University, Muscat, Oman, <sup>5</sup> Department of Surgery, Sultan Qaboos University Hospital, Muscat, Oman, <sup>6</sup> Medical Oncology Unit, Department of Medicine, College of Medicine, Sultan Qaboos University Hospital, Muscat, Oman, <sup>7</sup> Department of Surgery, Wrexham Maelor Hospital, Wrexham, United Kingdom

## OPEN ACCESS

### Edited by:

Giorgio Seano,  
Institut Curie, France

### Reviewed by:

Luca Tamagnone,  
Institute for Cancer Research and  
Treatment (IRCC), Italy  
Aleix Prat,  
Hospital Clinic de Barcelona, Spain

### \*Correspondence:

Sirin A. Adham  
sadham@squ.edu.om;  
sirinadham@yahoo.com

### Specialty section:

This article was submitted to  
Cancer Molecular Targets and  
Therapeutics,  
a section of the journal  
Frontiers in Oncology

**Received:** 23 November 2018

**Accepted:** 10 April 2019

**Published:** 25 April 2019

### Citation:

Al-Zeheimi N, Naik A, Bakheit CS,  
Al Riyami M, Al Ajarrah A, Al Badi S, Al  
Baimani K, Malik K, Al Habsi Z, Al  
Moundhri MS and Adham SA (2019)  
Neoadjuvant Chemotherapy Alters  
Neuropilin-1, PlGF, and *SNAI1*  
Expression Levels and Predicts Breast  
Cancer Patients Response.  
Front. Oncol. 9:323.  
doi: 10.3389/fonc.2019.00323

Circulating proteins hold a potential benefit as biomarkers for precision medicine. Previously, we showed that systemic levels of neuropilin-1 (NRP-1) and its associated molecules correlated with poor-prognosis breast cancer. To further identify the role of NRP-1 and its interacting molecules in correspondence with patients' response to neoadjuvant chemotherapy (NAC), we conducted a comparative study on blood and tissue samples collected from a cohort of locally advanced breast cancer patients, before and after neoadjuvant chemotherapy (NAC). From a panel of tested proteins and genes, we found that the levels of plasma NRP-1, placenta growth factor (PlGF) and immune cell expression of the transcription factor *SNAI1* before and after NAC were significantly different. Paired *t*-test analysis of 22 locally advanced breast cancer patients showed that plasma NRP-1 levels were increased significantly ( $p = 0.018$ ) post-NAC in patients with pathological partial response (pPR). Kaplan–Meier analysis indicated that patients who received NAC cycles and their excised tumors remained with high levels of NRP-1 had a lower overall survival compared with patients whose tissue NRP-1 decreased post-NAC (log-rank  $p = 0.049$ ). *In vitro* validation of the former result showed an increase in the secreted and cellular NRP-1 levels in resistant MDA-MB-231 cells to the most common NAC regimen Adriamycin/cyclophosphamide+Paclitaxel (AC+PAC). In addition, NRP-1 knockdown in MDA-MB-231 cells sensitized the cells to AC and more profoundly to PAC treatment and the cells sensitivity was proportional to the expressed levels of NRP-1. Unlike NRP-1, circulating PlGF was significantly increased ( $p = 0.014$ ) in patients with a pathological complete response (pCR). *SNAI1* expression in immune cells showed a significant increase ( $p = 0.018$ ) in patients with pCR, consistent with its posited protective role. We conclude that increased plasma and tissue NRP-1 post-NAC correlate with pPR and shorter overall survival, respectively. These observations support the need to consider anti-NRP-1 as a potential targeted therapy for breast cancer patients who are

identified with high NRP-1 levels. Meanwhile, the increase in both PlGF and *SNAI1* in pCR patients potentially suggests their antitumorigenic role in breast cancer that paves the way for further mechanistic investigation to validate their role as potential predictive markers for pCR in breast cancer.

**Keywords:** neuropilin-1, biomarker, breast, blood, response, neoadjuvant, *SNAI1*, PlGF

## INTRODUCTION

Breast cancer patients with locally advanced disease are treated with preoperative cycles of neoadjuvant chemotherapy (NAC) regardless of a patient's molecular subtypes (1). Patients respond differently to the preoperative chemotherapy, either completely or partially, or do not respond at all (1, 2). Factors that determine the degree of a patient's response are not yet fully understood. Therefore, molecular biomarkers and precision medicine might help to answer this question. Peripheral blood sampling is a rapid, convenient, and non-invasive method to determine an individual's pathological and physiological states. Circulating growth factors and cytokines provide a snapshot of the systemic changes in response to cancer. For instance, the circulating levels of vascular endothelial growth factor (VEGF) were shown to drive tumor survival through angiogenesis, prompting the development of the anti-angiogenic neutralizing antibody, bevacizumab (3). This has led to an improvement in patients' overall survival when combined with chemotherapy in the treatment of metastatic colon cancer (4). However, the poor efficacy of bevacizumab for the treatment of advanced breast cancer led to the withdrawal of its approval for breast cancer in 2011 by the FDA (5). A receptor closely related to VEGF is Neuropilin-1 (NRP-1), which has been associated with the progression of different types of cancer including breast cancer (6–10) and direct NRP-1 targeting via miR-376a suppressed the progression of breast cancer cells (11). Therefore, current research suggests that targeting NRP-1 might be a new strategy for cancer treatment (12). NRP-1 is a non-signaling molecule with multiple functions depending on the ligand that binds to its extracellular domain. Genentech produced an antibody that targets NRP-1, which was combined with anti-VEGF in an experimental model to show their additive antitumor activity (13). Although there are many studies confirming that NRP-1 is involved in driving tumorigenicity, clinical investigations into NRP-1 levels in patients have not been well-explored to date. Similar to NRP-1, placental growth factor (PlGF) is a member of the VEGF family and known to mediate angiogenesis, with circulating PlGF shown to be a prognostic marker for cancer. A higher plasma PlGF level was associated with progression and recurrence in colorectal cancer and oral squamous cell carcinoma (14–16). Recently, we reported the expression of NRP-1 and other associated molecules in the plasma, immune cells, and tumor tissue of a breast cancer patients' cohort and confirmed the role of NRP-1, PlGF, and *SNAI1* in breast cancer progression (17). It is well-established that NRP-1 is associated with the worst breast cancer outcomes (18). To our knowledge, there are no previous reports investigating the levels of NRP-1 in locally advanced breast cancer patients

who receive NAC. Therefore, in this study, we explored the effect of NAC on the levels of plasma and tissue NRP-1 and PlGF, as well as validating their use as predictive/pharmacodynamic breast cancer biomarkers.

In this report, we are adding an extra finding, in which we show that the levels of circulating NRP-1 were significantly increased in patients who received NAC and had a partial response. Assessing patients who underwent NAC indicated that the levels of NRP-1 measured post-NAC were significantly higher in younger patients and patients with either a low or a medium body mass index (BMI), as well as in patients who remained with larger tumor size and partial response. Previously, we showed that *SNAI1* expression in the immune cells collected from the peripheral blood of breast cancer patients was significantly higher in patients with stage I disease compared with higher stages (17). In this report, we found that *SNAI1* expression in peripheral blood mononuclear cells (PBMCs) of patients who received NAC was significantly increased, especially in patients who showed a complete pathological response to the treatment, but did not increase in those who had a partial response, which indicates that *SNAI1* might be a good candidate to be used as a predictive marker for a complete pathological response. *In vitro* experiments on breast cancer MDA-MB-231 cells was done to validate the clinical observations. The knockdown of NRP-1 in the cells sensitized them to the common chemotherapy regimen (Adriamycin/cyclophosphamide + Taxane) used in the neoadjuvant setting, which can be translated in that, patients with low levels of NRP-1 might respond better to NAC and vice versa.

## MATERIALS AND METHODS

### Patients Characteristics

In a prospective setting, a cohort of 22 patients, diagnosed clinically and pathologically with locally advanced breast cancer at Sultan Qaboos University Hospital, was recruited. All 22 patients underwent NAC prior to surgery. Blood samples were collected from all 22 patients before and after the completion of NAC cycles and from 50 healthy controls. Tissue samples before (initial biopsy) and after treatment (excised tumor) were collected from all the 22 patients however, 12 out of the 22 tissue samples were only available for biomarker staining and the remaining 10 patients' tissue was not enough for research use therefore, we retrieved another 17 tissue samples from the pathology archive and were added in a retrospective setting to match the number of blood samples (22 prospective blood samples) + (12 prospective tissue + 17 retrospective tissue = 29 tissue samples). The study was approved by the ethical committee at the College of Medicine and Health Sciences, Sultan

Qaboos University (License #SQU.EU/162/14, MREC#1018). Informed signed consent was obtained from all participants. All experiments were performed in accordance with institutional and national guidelines.

## Clinical Assessment, and Definition of Patients' Response Post NAC

Patients' staging and response post NAC treatment (ypTN) were classified according to American Joint Committee on Cancer (AJCC). Patients were classified as pathological complete responders (pCR) when no invasive residual carcinoma (ypT0) was identified in either the breast or lymph nodes (ypN0). The presence of *in situ* carcinoma post NAC and the absence of residual invasive disease was also categorized as pCR. Therefore, these patients were staged as either ypT0 or ypTis, respectively. Pathological partial responders (pPR) were those cases in which residual invasive cancer was present with evidence of a response to treatment. These patients would therefore have ypT stage depending on residual tumor size. Changes indicative of response to chemotherapy included fibrosis, myxoid stroma, foamy macrophages, and chronic inflammation. Further response stratification was done to determine whether the molecules in test had any association with either tumor size (ypT) or lymph node status (ypN) according to the new definition of "Perfect Pathology Report" by national cancer institute NCI publication (19).

Regarding tumor grade after NAC, and based on cancer reporting guidelines by, the College of American Pathologists, USA (20), the American Joint Committee on Cancer (AJCC) (19), and the Royal College of Pathologists, UK (21) the grade post NAC was not considered for our comparative analysis since both guidelines clearly state that for most tumors the grade remains unchanged and the prognostic significance of a change in grade post NAC has not been determined. Hence the recommendation is to grade the tumor based on the pre-treatment core biopsy.

When the patient's outcome data during follow up was compared in respect to survival, the terms remission and relapse were used. Remission is defined as the absence of cancer in laboratory tests, physical examination, and radiological imaging after the completion of the prescribed treatment. Relapse is defined as the recurrence of cancer evidenced by the former tests. Patients' characteristics at diagnosis such as age, body mass index (BMI), hormone receptor status, breast cancer subtype, chemotherapy used, overall response, tumor size, and nodal status post NAC, disease status and survival are listed in Table 1.

## Serum Soluble Protein Detection by ELISA

Patient blood collected in EDTA-coated vacutainers was subjected to density gradient centrifugation with Histopaque (Sigma Aldrich, UK) at 400 g, with a break off for 30 min at room temperature. The separated plasma was frozen at  $-80^{\circ}\text{C}$  until further analysis. The concentration of soluble NRP-1 and PIGF was measured in the plasma samples or conditioned media using ELISA kits (R&D systems, USA) according to the manufacturer's instructions. ELISA validation test was done in our previous related published article (17).

**TABLE 1 |** Clinical information of breast cancer patients.

	Blood	Tissue
<b>PATIENTS CHARACTERISTICS</b>		
<b>Total</b>	22	29
<b>Age</b>		
≤50 years	86.4% (19/22)	68.9% (20/29)
>50 years	13.6% (3/22)	31.1% (9/29)
<b>BMI</b>		
18.5–24.9	22.7% (5/22)	24.2% (7/29)
25–29.9	27.3% (6/22)	13.8% (4/29)
≥30	40.9% (9/22)	58.6% (17/29)
Unknown	9.1% (2/22)	3.4% (1/29)
<b>Tumor size (Post NAC)</b>		
ypT0 (No tumor)	50% (11/22)	20.7% (6/29)
ypT1 (<2 cm)	18.2% (4/22)	31% (9/29)
ypT2 (2–5 cm)	22.7% (5/22)	20.7% (6/29)
ypT3 (>5 cm)	0% (0/22)	6.9% (2/29)
ypT4 (Tumors in the chest wall/skin or inflammatory breast cancer)	0% (0/22)	10.3% (3/29)
Not known	9.1% (2/22)	10.3% (3/29)
<b>Node status (Post NAC)</b>		
ypN0 (No nodes involved)	59.1% (13/22)	37.9% (11/29)
ypN1 (1–3 axillary lymph nodes)	22.7% (5/22)	37.9% (11/29)
ypN2 (4–9 axillary lymph nodes)	4.5% (1/22)	13.8% (4/29)
ypN3 (10 axillary lymph nodes or infraclavicular, supraclavicular or internal mammary nodes)	13.6% (3/22)	10.3% (3/29)
<b>Receptor status</b>		
ER positive	59% (13/22)	48.2% (14/29)
ER negative	41% (9/22)	51.8% (15/29)
PR positive	54.5% (12/22)	44.8% (13/29)
PR negative	45.5% (10/22)	55.2% (16/29)
HER2 positive	45.5% (10/22)	41.3% (12/29)
HER2 negative	54.5% (12/22)	58.7% (17/29)
<b>Subtypes</b>		
Luminal A	4.5% (1/22)	13.8% (4/29)
Luminal B	18.2% (4/22)	17.2% (5/29)
Luminal B like	41% (9/22)	17.2% (5/29)
HER2 type	31.8% (7/22)	24.2% (7/29)
Triple negative	4.5% (1/22)	27.6% (8/29)
<b>Chemotherapy*</b>		
6x DCH	9.1% (2/22)	6.9% (2/29)
4xAC+4xD	40.9% (9/22)	58.6% (17/29)
4xAC+4xD+ H	13.6% (3/22)	24.2% (7/29)
3xFEC+3xD	9.1% (2/22)	3.4% (1/29)
3xFEC+3xD+ H	4.5% (1/22)	3.4% (1/29)
4xAC+12xPaclitaxel+ H (weekly)	4.5% (1/22)	3.4% (1/29)
4xAC+4xDHP	13.6% (3/22)	0% (0/29)
6xDC	4.5% (1/22)	0% (0/29)

(Continued)



TABLE 1 | Continued

	Blood	Tissue
<b>Response</b>		
pCR	36.4% (8/22)	10.34% (3/29)
pPR	59.1% (13/22)	68.98% (20/29)
Stable disease	0.0% (0/22)	10.34% (3/29)
Progression	4.5% (1/22)	10.34% (3/29)
<b>Disease status</b>		
In remission	81.8% (18/22)	65.5% (19/29)
Relapsed	18.2% (4/22)	34.5% (10/29)
<b>Survival</b>		
Survived	100% (22/22)	37.9% (11/29)
Died	0% (0/22)	24.2% (7/29)
Did not complete 3years	100% (22/22)	37.9% (11/29)

\*D, Docetaxel; H, Herceptin; A, Adriamycin; C, Cyclophosphamide; F, Fluorouracil; E, Epirubicin; P, Pertuzumab.

## PBMC Isolation and RNA Extraction

Anti-coagulated blood was subjected to density gradient centrifugation with Histopaque (Sigma Aldrich, UK) for 30 min at 400 g (breakoff) at room temperature. The buffy coat layer containing PBMCs was isolated, washed twice with cold PBS and pelleted at 250 g at 4°C. RNA was extracted from the PBMCs using TRI reagent (Ambion, USA), phase separation with chloroform and overnight isopropanol precipitation. The RNA pellet was washed twice in 70% ethanol in DEPC water, then dried completely and resuspended in DEPC-treated water (Ambion, USA). RNA was quantified using the NanoDrop™ 2000c spectrophotometer (Thermo Scientific, USA) and RNA quality, as indicated by the 260/280, was determined to be in the range of 1.8–2.0. One microgram of extracted RNA was treated with DNase I (Ambion, Lithuania) for 15 min at room temperature and converted to cDNA using the high-capacity reverse transcription kit (Applied Biosystems, USA). Synthesized cDNA was diluted to a final concentration of 5 ng/ul in DEPC-treated water and stored at –80°C until further analysis.

## Quantitative Real-Time PCR

Real-time PCR was conducted using the SoAdvanced mastermix (Biorad, USA). Primers were designed using the Primer Express software (Applied Biosystems, USA) and are listed in Supplementary Tables 1, 2 in our previous report (17). Fifteen nanograms of cDNA was used per reaction. The CFX96 Real-time PCR Detection System (Biorad, USA) was used under the following conditions: enzyme activation at 95°C for 20 s, 40 cycles of denaturing at 95°C for 3 s, and annealing/extension at 63.4°C for 30 s. The specificity of PCR reactions was verified by the melt curve analysis of each amplified product. Each real-time PCR reaction was performed in duplicate. A no template control (NTC) was performed for each primer pair tested in all experimental runs. Commercially available reference cDNA (Clontech, USA) was utilized as an inter-plate calibrator to identify technical variations between experimental runs. The generated Ct results were analyzed using

the QBase data analysis software to generate relative expression values using the  $2^{-\Delta\Delta Ct}$  method of calculation. The *GUSB* gene was used for normalization in qRT-PCR since it was selected according to GeNorm analysis done in our previous study (17).

## Immunohistochemistry

Immunohistochemical analysis was conducted on 29 (12 prospective+17 retrospective) pathologically confirmed locally advanced breast cancer tumors in formalin-fixed paraffin-embedded tissue. Briefly, tissue sections (3 μm) were deparaffinized using xylene and rehydrated in graded ethanol and H<sub>2</sub>O. Antigens were retrieved using EDTA-pH 9 solution in 95°C water bath for 30 min. Endogenous peroxidases and residual blood were blocked/removed by 2% H<sub>2</sub>O<sub>2</sub> and the slides were washed in PBS followed by a wash in 0.05% Triton-x100 to permeabilize the cells. The tissues were blocked in 5% goat serum (Dako, USA) and incubated with primary antibody (Anti-Neuropilin 1 antibody (ab81321) or anti-PLGF antibody (ab196666) (Abcam, UK) at 4 °C overnight. Following, the slides were washed in PBS and incubated with the EnVision™ + Dual Link System-HRP (Dako, USA) labeled secondary antibody for 1 h at room temperature and followed by incubation with substrate chromogen solution (DAB) chromogen (Dako). The sections were counterstained using hematoxylin solution and dehydrated and mounted using DPX (Sigma, USA). Tissues were visualized using (NikonH600L) light microscope. Immuno Reactive Scoring (IRS) was performed for the stained slides using the following formula: IRS = SI (staining intensity) × PP (% of positive cells). Independent validation of the staining was done on normal human placenta tissue with and without primary antibody (antiNRP-1) **Supplementary Figure S1**.

## Cell Culture

MDA-MB-231 breast cancer cell line was purchased from Cell Lines Services CLS, Germany in 2015. The cells were maintained in monolayer cultures in 5% CO<sub>2</sub> incubator at 37°C. The MDA-MB-231 cells were sustained in DMEM (Sigma, USA) supplemented with 5 mM sodium pyruvate (Sigma, USA). The cells media was supplemented with 10% fetal bovine serum (Gibco®, USA), and 2 mg/L gentamicin (Gibco, USA).

## Establishment of Resistant MDA-MB-231 Cells

The treatment modalities of MDA-MB-231 cells was done to mimic the clinical NAC treatment of breast cancer patients. Briefly, MDA-MB-231 cells were treated *in vitro* with four cycles combination of 200 nM doxorubicin (Brand name Adriamycin, Pharmacia, Italy) and 600 nM cyclophosphamide (4xAC) (Brand name Cytosan, Baxter, Germany) followed by four cycles of 50 nM paclitaxel (4xAC+4xPAC) (Brand name Taxol, EBEWE Pharma, Austria). Each treatment cycle was 72 h long. After each cycle, the cells which remained attached were left to proliferate until confluency and the following cycle of treatment was initiated right after.

## NRP1 Knockdown Using CRISPR-Cas 9 System

CRISPR-Cas9 system was used to knockdown NRP-1 in MDA-MB-231 cells. Pre-designed NRP1 gRNA primers using the GeneArt™ CRISPR Search and Design tool (Thermo-Fisher) were used to synthesize gRNA for NRP1 knockdown (IVT-NRP1-gRNA-T2-F2: TAATACGACTCACTATAGACCAGGAG ATGTAAGG and IVT-NRP1-gRNA-T2-R2: TTCTAGCTCTAA AACGGTACCTTACATCTCCTGG. The gRNA was synthesized using GeneArt™ Precision gRNA Synthesis Kit (Thermo-Fisher). The gRNA Cleanup kit (Thermo-Fisher) was used for the purification of the generated gRNA before transfection. The concentration of the purified gRNA was measured using Nanodrop (Nanodrop 2000, USA) and the gRNA band was further checked in agarose gel (100 bp). A day before transfection,  $6 \times 10^5$  cells were seeded in 6-well plate. GeneArt Platinum Cas9 Nuclease kit (Thermo-fisher) was used for the transfection by mixing the Cas9 Nuclease and gRNA in addition to the transfection reagent lipofectamine™ CRISPRMAX in Opti-MEM media. The mixed complex was added to transfect the cells for 48 h in 5% CO<sub>2</sub> incubator at 37°C. Subsequently, the cells were diluted into 1:5, 1:10, and 1:50 to isolate the clones carrying the NRP1 knockdown. Using the colony disk isolation method, two NRP-1 knockdown clones # 15 and # 22 were selected according to the level of NRP-1 knockdown determined by western blot.

## Western Blot

NRP-1 protein expression was measured using western blotting technique. Briefly, the cells were washed with cold PBS, incubated with lysis buffer for 1–2 min (Cell Signaling technology, USA) in the presence of phenylmethylsulfonyl fluoride (PMSF) protease inhibitor (Sigma, Germany) (Sigma, Germany). Protein cell lysate was vortexed and centrifuged for 20 min at 4°C. The supernatants were collected, and protein quantification was done using Pierce bicinchoninic acid (BCA) protein assay kit (Thermo Fisher). Protein lysate samples were all adjusted to have 100 µg of total protein per sample and were electrophoretically separated using 7.5% sodium dodecyl sulfate polyacrylamide gel electrophoresis and then transferred onto polyvinylidene difluoride membranes (BioRad, USA). The immunoblots were then blocked with 5% non-fat milk and subsequently probed with rabbit primary monoclonal antibodies, NRP-1 (Ab Cam, UK Catalog #ab81321) or GAPDH as a normalizing internal control (cell signaling technologies catalog #2118) at a 1:1,000 dilution, incubated at 4°C overnight. The blots were then washed three times for 5 min with PBS and incubated with goat anti-rabbit IgG horseradish peroxidase conjugated secondary antibody at 1:5,000 dilutions (Abcam, UK) for 2 h at room temperature and developed using the clarity western ECL substrate (BioRad, USA). The densitometric analysis of the protein bands was performed using the Image lab software (BioRad, USA).

## Colony Formation Assay

The ability of the MDA-MB-231 cells parental or NRP-1 Knockdown variants to form colonies before and after treatment was tested. The cells were treated with combination of both 200 nM Adriamycin (doxorubicin) and 600 nM

cyclophosphamide or 50 nM paclitaxel [IC<sub>50</sub> (22)] for 72 h. Then, the cells were seeded at a density of 1,000 cells/well in 6-well plates (Corning, USA) and incubated for 14 days at 37°C in 5% CO<sub>2</sub> incubator and the media was changed twice during this period. After the 2 weeks of incubation, the colonies were washed with PBS and stained with 25% crystal violet/methanol solution for 15 min at room temperature. Crystal violet stain was removed, and the wells were washed with tap water.

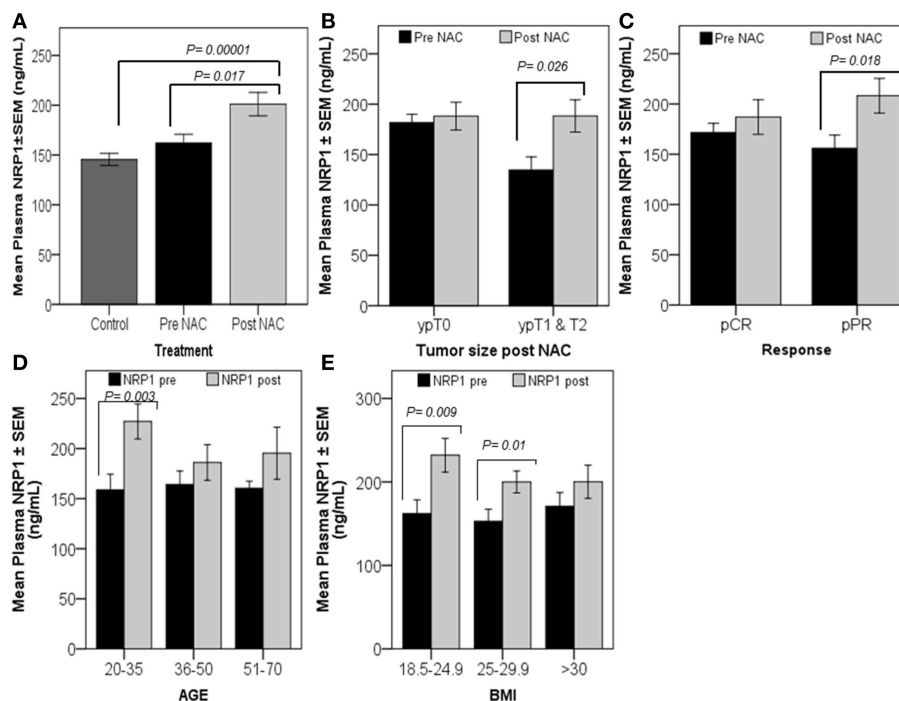
## Statistical Analysis

The paired *t*-test was used as a gold standard test for the differential expression of plasma or tissue proteins before and after NAC. The Shapiro–Wilk test for normality was conducted to determine the distribution of breast cancer patients in the cohort studied. While the distribution of the plasma protein dependent variables was confirmed to be normal, the PBMC gene expression data indicated non-normal distribution. Considering the heterogeneity in the population distribution and the unequal and limited number of cases among the subgroups studied, the PBMC gene expression data set was log<sub>10</sub>-transformed to attain normality in the population distribution. The results that were significant using the paired *t*-test were further tested by univariate analysis to compare the measured values in patient samples, before and after treatment, to the levels measured in the healthy controls used in our previous study, with Tukey used as a *post-hoc* test. The Kaplan–Meier curve was generated to calculate the overall survival of patients regarding the low or high immunoreactive score (IRS) for NRP-1 tissue expression and the log-rank test (Mantel-Cox) was used to indicate significance. Associations where  $p < 0.05$  were considered significant (95% confidence interval). Poorly represented subgroups ( $n < 3$ ) were excluded from the analysis to avoid interpretation errors. The IBM SPSS software (Version 22) was used for all statistical analyses and graph preparations.

## RESULTS

### Plasma NRP-1 Levels Are Induced by Chemotherapy and Correlate With a Poor Response in Patients

The univariate analysis comparing plasma NRP-1 levels in healthy controls ( $n = 50$ ), patients with locally advanced disease prior to the initiation of NAC (pre-NAC) ( $n = 22$ ), and post-treated patients (post-NAC) ( $n = 22$ ) indicated that there was a significant increase in plasma NRP-1 levels in post-treated patients, compared to their initial level ( $p = 0.017$ ) and the level of the healthy normal controls ( $p = 0.00001$ ) (Figure 1A). The paired *t*-test was used to determine the differential levels of plasma NRP-1 pre- and post-NAC. The analysis of data indicated that the levels of plasma NRP-1 were increased significantly ( $p = 0.026$ ) in patients who remained with a large tumor size ( $n = 9$ , ypT1&2) and partially responded to the treatment ( $p = 0.018$ ) ( $n = 13$ , pPR) (Figures 1B,C). However, no significant change was observed in the NRP1 level in patients who had no tumor post NAC ( $n = 11$ , ypT0) or in patients who showed complete response to NAC ( $n = 8$ , pCR). In



**FIGURE 1 |** Significant increase in plasma NRP-1, post-NAC, was correlated with poor response. Univariate analysis (A) showed a significant increase in plasma NRP-1, post-NAC, compared to the healthy controls and pre-NAC levels. The graphs (B–E) represent the paired *t*-test for the mean plasma NRP1 levels  $\pm$  SEM. The graphs reveal a significant increase in plasma NRP1, post-NAC, in patients who had tumors, ypT1&T2 (B), and in those with partial response (pPR) (C). NRP-1 levels were found to increase in younger patients and patients with either low or moderate BMI (D,E).  $p \leq 0.05$  is considered to indicate statistical significance.

addition, NRP-1 plasma levels, post-NAC, were significantly higher ( $p = 0.003$ ) in young patients ( $n = 6$ , 20–35 years) and in cases with a low or a medium BMI [ $(n = 5$ , BMI 18.5–24.9 ( $p = 0.009$ ) and  $n = 6$ , BMI 25–29.9 ( $p = 0.01$ )] (Figures 1D,E).

### Low Tissue NRP-1 Expression Post-NAC Is Correlated With Improved Survival

Univariate analysis showed that tissue NRP-1 expression was reduced in post-NAC specimens ( $p = 0.05$ ) of patients under remission ( $n = 19$ ), while there were no changes in relapsed patients ( $n = 10$ ) (Figure 2A). Similarly, tissue NRP-1 expression, post-NAC, was significantly decreased ( $p = 0.03$ ) in all surviving patients ( $n = 11$ ), but not in patients who died ( $n = 7$ ) (Figure 2B). A Kaplan–Meier graph indicated that patients who received NAC and remained with high tissue NRP-1 levels ( $n = 8$ ) had lower overall survival compared with patients whose tissue NRP-1 decreased post-NAC ( $n = 10$ ), (log-rank  $p = 0.049$ ) (Figure 2C). Representative immunohistochemistry images for tissue NRP-1 expression indicated a dramatic decrease in tissue NRP-1 levels, post-NAC, in the patients who survived, compared to those who died (Figure 2D).

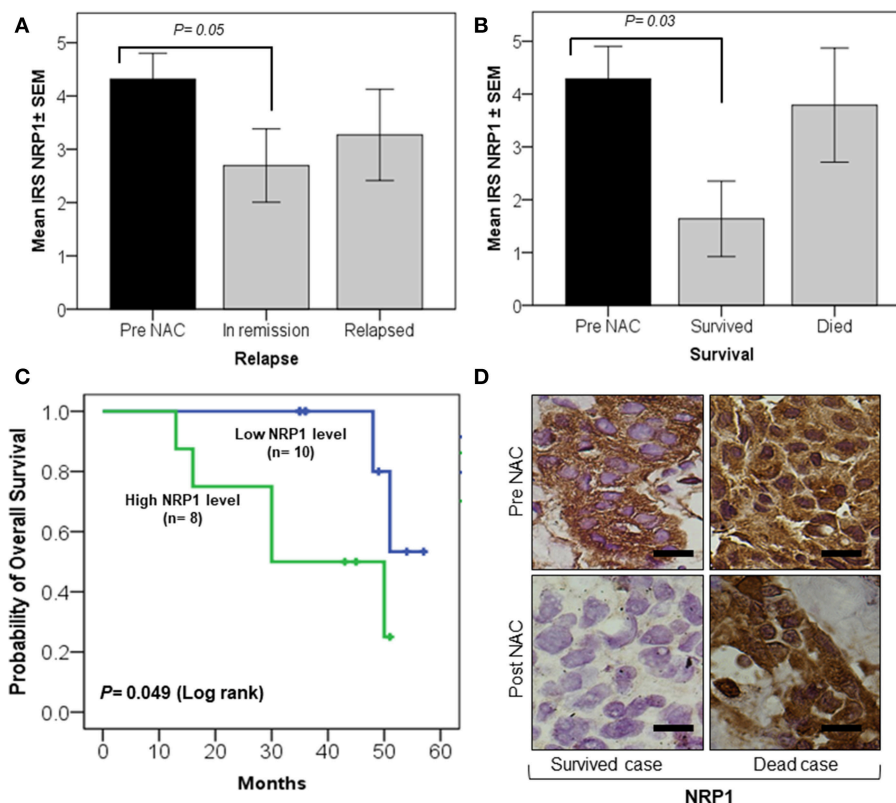
### Plasma Levels of PlGF Were Increased in Pathological Complete Responders

Univariate analysis indicated that the basal levels of plasma PlGF, pre-NAC ( $n = 22$ ), were significantly lower ( $p = 0.034$ )

than the levels found in healthy controls ( $n = 50$ ) (Figure 3A). The paired *t*-test indicated a relative increase in plasma levels of PlGF in patients who showed complete tumor regression ( $n = 11$ , ypT0) ( $p = 0.013$ ) and a pathological complete response ( $n = 8$ , pCR) ( $p = 0.014$ ) after NAC (Figures 3B,C). Increased plasma PlGF was observed in older and patients [ $n = 13$ , 36–50 years ( $p = 0.007$ ) and  $n = 3$ , 50–71 ( $p = 0.029$ )] and in patients with a high BMI [ $n = 9$ , BMI  $>30$  ( $p = 0.009$ )] (Figures 3D,E).

### SNAIL Expression in PBMCs Is Upregulated in Complete Responders

SNAIL levels measured in patients PBMCs indicated that this transcription factor is significantly increased in patients who had no residual tumor ( $n = 11$ , ypT0) ( $p = 0.025$ ) or diseased lymph nodes ( $n = 13$ , ypN0) after NAC and in pathological complete responders ( $n = 8$ , pCR) ( $p = 0.018$ ) (Figures 4A–C). Univariate analysis showed a significant decrease ( $p = 0.042$ ) in the initial expression (pre-NAC) of SNAIL in patients who showed a partial response ( $n = 13$ , pPR) to NAC, compared to the expression in the healthy controls ( $n = 50$ ) (Figure 4D). Additionally, a trend of increased SNAIL expression post-NAC was observed in patients with pCR ( $n = 8$ ) similar to the levels detected in the healthy controls (Figure 4E). SNAIL expression in PBMCs was significantly increased in young patients ( $n = 6$ , 20–35 years old) ( $p = 0.047$ ) (Figure 4F).



**FIGURE 2 |** Significant decrease in tissue NRP-1 expression, post-NAC, is linked with better overall patient survival. The graphs (A,B) represent the univariate analysis of NRP1  $\pm$  SEM IRS values in breast cancer tissues pre- and post-NAC. A significant decline in NRP-1 expression, post-NAC, was observed in the tumor tissue of patients under remission (A) and who survived (B). The Kaplan-Meier curve (C) shows a significant improvement in overall survival in patients whose tumors were represented with a low IRS tissue NRP-1 score after NAC. The representative images of immunohistochemistry staining (D) showed a decrease in tissue NRP-1 levels as a result of chemotherapy in surviving vs. non-surviving individuals, scale bar = 50  $\mu$ m.  $p \leq 0.05$  is considered to indicate statistical significance.

## Chemotherapy Treatment for MDA-MB-231 Cells Increased NRP-1 Expression Levels

The resistant MDA-MB-231 cells to chemotherapies 4xAC+4xPAC expressed higher levels of soluble NRP-1 in the conditioned media and exhibited increased levels of cellular NRP-1 as represented by the increase in the intensity of the protein band on western blot (Figures 5A,B).

## Neuropilin-1 Knockdown Decreased the Number of Colonies Formed by MDA-MB-231 Cells and Sensitized Them to Chemotherapy

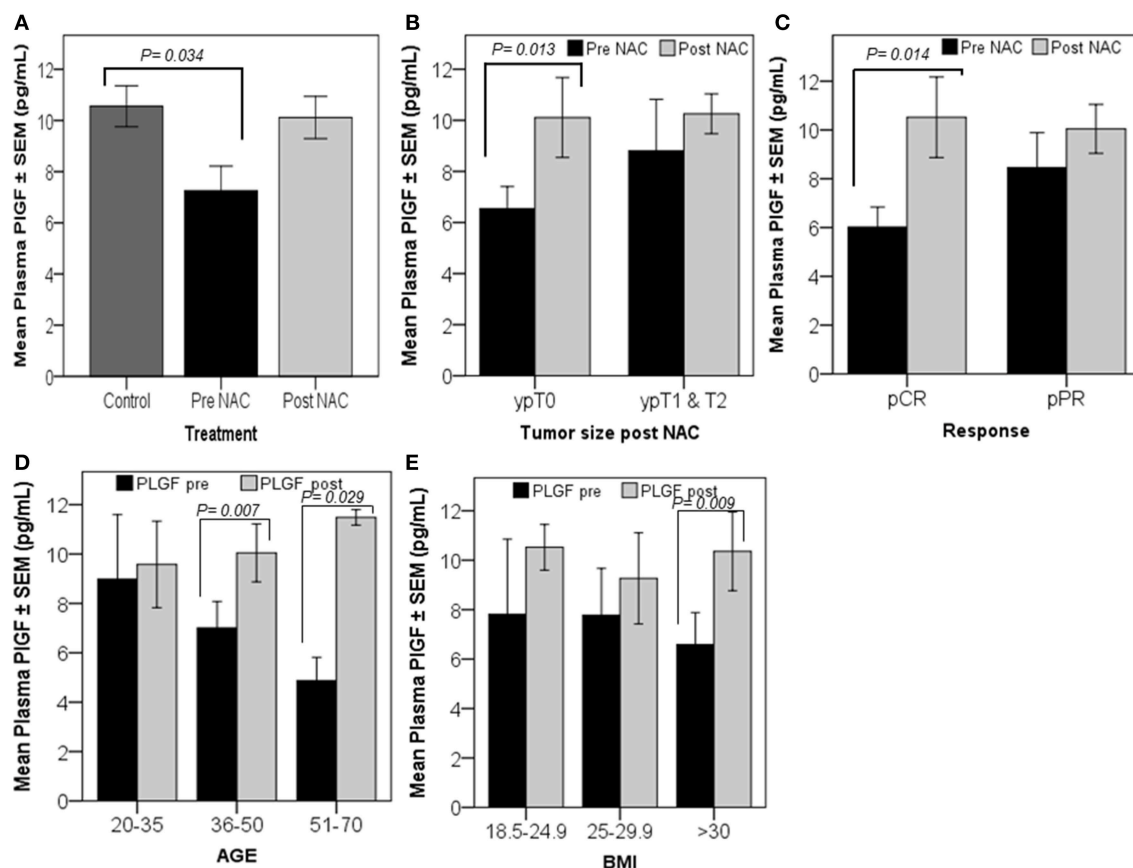
NRP-1 was efficiently knocked down in MDA-MB-231 cells using CRISPR Cas-9 as described in the materials and methods section. Two different clone variants; MDA- NRP-1 knockdown clone # 22 (58% knockdown) and MDA- NRP-1 knockdown clone # 15 (99% knockdown) were isolated (Figure 6). The quantification of NRP-1 expression was determined by measuring the density of the expressed bands as represented in the graph from three independent experiments (Figure 6). Clonogenic assay showed that NRP-1 knockdown caused a reduction in the ability of the cells to form colonies. Treating NRP-1 knockdown variants

with combination of Adriamycin and cyclophosphamide (AC) or paclitaxel differentially reduced the ability of the cells to form colonies, however. Treating control parental MDA-MB-231 cells with AC didn't not affect the formation of the colonies only paclitaxel reduced its clonogenic ability.

## DISCUSSION

A major obstacle of cancer management is drug resistance. Breast cancer patients with locally advanced breast cancer receive NAC to reduce tumor size, which makes it easier to be excised by surgery (12). The progression of the disease, post-NAC, is usually due to the presence of innate chemoresistant cells or acquired resistance throughout the cyclic treatment (23). In this study, we analyzed the plasma and PBMCs of breast cancer patients prior to the start of NAC and post-NAC. We also investigated the differential expression of proteins and genes that we have previously shown to be involved in poor-prognosis breast cancer cases (17); however, their role in predicting a response to chemotherapy has not been studied before. Plasma NRP-1, post-NAC, was upregulated in patients who were classified as partial responders to NAC. More importantly, the patients



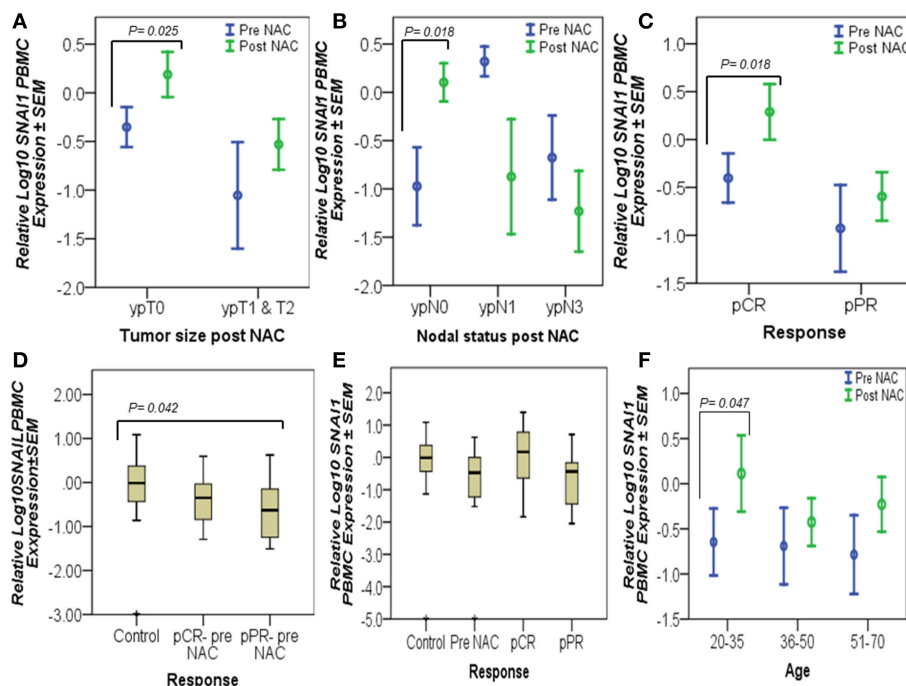


**FIGURE 3 |** Patients with high-circulating PlGF levels, post-NAC, showed a better response. Univariate analysis revealed a significant decrease in plasma PlGF levels in breast cancer patients measured prior to the initiation of chemotherapy (pre-NAC) compared to the levels measured in healthy controls (A). The graphs (B–E) represent the paired *t*-test results, which show the mean plasma PlGF ± SEM. A significant increase in the PlGF level was detected in patients who were tumor-free after finishing NAC cycles, ypT0 (B), and in patients who responded completely to the treatment, pCR (C). The graphs (D,E) show PlGF levels increased significantly post-NAC in patients aged between 36–50 and 51–70 years, and in patients who have a BMI >30.  $p \leq 0.05$  is considered to indicate statistical significance.

who died had high levels of tissue NRP-1, post-NAC, than surviving patients. This notion is interesting, since it is in concordance with previous findings in non-small lung cancer (NSCLC), where patients who had a high expression of NRP-1 after chemotherapy had shorter disease-free and overall survival (24). Another study indicated that overexpressing NRP-1 caused chemoresistance in pancreatic cancer cells through the MAPK signaling pathway (25). Similarly, we observed an increase in the expression levels of soluble and focal NRP-1 using ELISA and western blot, respectively, in an MDA-MB-231 breast cancer cell-chemoresistant model (4xAC+4xPAC) generated in our lab (Figures 5A,B). In addition, we recently reported that NRP-1 overexpression was induced by a combination of Adriamycin and cyclophosphamide-treated BT474 breast cancer cells (26). While the inhibition of NRP-1 increased chemosensitivity for different kinds of cancer cells (27). More recently, a study showed that the inhibition of NRP-1, using a small molecule antagonist, caused a combined reduction in angiogenic and tumorigenic ability (28). In line with the previous findings, the *in vitro* knockdown of

NRP-1 in MDA-MB-231 cells in this study supports the fact that NRP-1 high expression leads to more resistance to chemotherapy similar to a recent report which indicated the role of NRP-1 promoting resistance to oncogene targeted therapies (29). Therefore, our result confirms the usefulness of the strategy to target NRP-1 in combination with chemotherapy in patients with a partial response to NAC alone, which thus determines that NRP-1 is a pharmacodynamic biomarker in breast cancer.

Although we did not find any significance using Kaplan–Meier analysis between PlGF plasma or tissue expression, before and after chemotherapy, with patients' survival, the plasma levels were significantly high in complete responders (pCR). A similar increase in plasma PlGF was reported as a result of anti-angiogenic treatment (30); however, there are no reports on PlGF levels after NAC. Although we reported earlier that plasma PlGF levels were not significantly different from those in healthy controls (17), in this study, we showed that there was a significant decrease in pre-NAC plasma PlGF, when compared with healthy controls. This discrepancy arises from differential study design



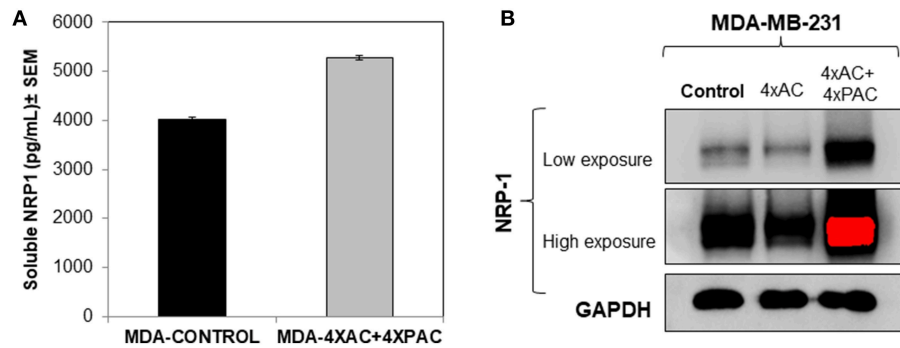
**FIGURE 4 |** Complete responders had significantly high *SNAI1* expression in their PBMCs. The graphs represent the relative Log<sub>10</sub> PBMC *SNAI1* ± SEM in controls and breast cancer patients, pre- and post-NAC, as measured by RT-qPCR. An upregulated level of *SNAI1*, post-NAC, was detected in patients who had no residual tumors (ypT0), no lymph node involvement (ypN0), and in patients with a complete response (pCR), as tested by the paired *t*-test (A–C). Similarly, univariate analysis (D,E) showed a significant decrease in the pre-NAC expression of *SNAI1* in pPR patients, compared to the expression in the healthy controls (D) and an increased trend in *SNAI1* expression, post-NAC, in pCR patients almost at the levels measured in the healthy controls (E). However, *SNAI1* expression declined in breast cancer patients when measured prior to the start of NAC (pre-NAC) and in those who did not respond to treatment (pPR) (E). In addition, a significant increase in *SNAI1* expression was detected in the PBMCs of young patients as a result of NAC using paired *t*-test (F).  $p \leq 0.05$  is considered to indicate statistical significance.

criteria, since our previous report was conducted on breast cancer patients, regardless of their disease stage or treatment plan; whereas, in this study, we only focused on those patients who presented with locally advanced disease and underwent NAC. A previous study showed the prognostic value of PIGF in patient tissue toward breast cancer progression (31), which is consistent with our previous findings that tissue PIGF is higher in metastatic breast cancer, compared with locally advanced breast cancer patients (17). In this study, we indicate that plasma PIGF increases significantly post-NAC in complete responders, thereby suggesting its potential use as a pharmacodynamic biomarker for breast cancer post NAC, similar to an earlier study in renal cancer, which described PIGF as a pharmacodynamic biomarker for anti-VEGF therapy (32).

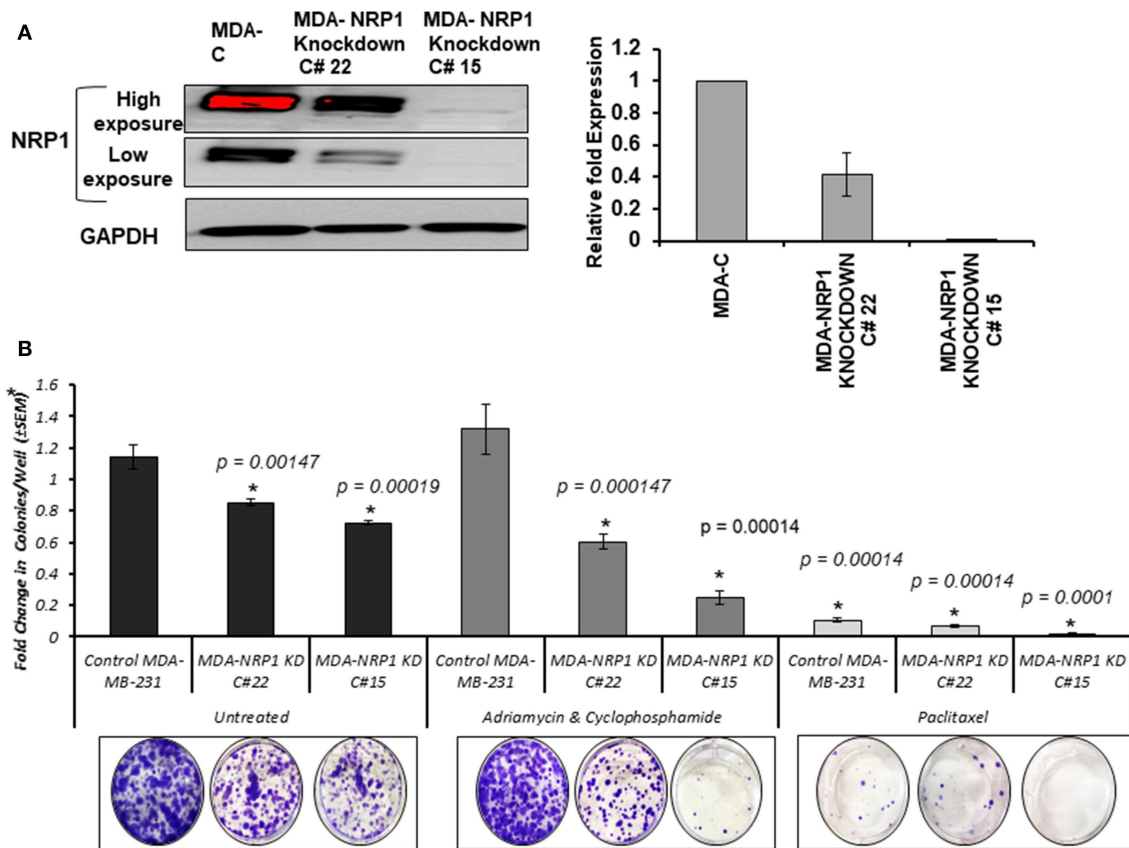
In addition to PIGF, the overexpression of *SNAI1* in PBMCs, post-NAC in complete responders (no residual tumor and no nodal disease), points to its protective role in breast cancer prognosis. The expression of *SNAI1* in complete responders (pCR) attained a similar level in the healthy controls. A previous report showed that the *SNAI1* protein product, Snail, was expressed at lower levels in breast tumor tissue compared with normal breast tissue (33). We reported a similar finding in PBMCs from breast cancer patients, who have significantly lower *SNAI1* expression compared with healthy controls (17).

## SUMMARY AND CONCLUSIONS

We conclude that NRP-1 expression in breast tumor tissue, post-NAC, is a potential predictive biomarker for breast cancer survival. Circulating plasma PIGF levels are lower in locally advanced breast cancer patients compared to healthy individuals, while they increased, post-NAC, in patients who responded completely to the treatment. *SNAI1* expression in immune cells are downregulated in breast cancer patients and increased, similar to healthy control levels in complete responders to NAC, indicating their potential protective role in breast cancer. All the studied molecules thus serve as good candidates for breast cancer prognosis and targeted treatment. Overall, the main aim of this study was to understand grossly the relationship between patients' response to NAC treatment regardless of the regimen used and potential molecular biomarkers but does not compare variables between the different chemotherapy treatment types, since treatment can't be always the exact same for each patient and depends on individual disease stage, subtype and sensitivity to drugs. And finally, the results of knocking down NRP-1 in MDA-MB-231 cells indicated that the cells became more sensitive to the treatment regardless of the drug used. Therefore, the *in vitro* results were consistent with the clinical observations in which patients with low levels of NRP-1, responded much better than those



**FIGURE 5 |** *In vitro* cellular analysis show that resistant MDA-MB-231 cells, treated with four cycles of Adriamycin and cyclophosphamide and four cycles of Paclitaxel (mimicking the clinical treatment setting of breast cancer), caused an increase in the levels of soluble (A) and cellular NRP-1 (B) detected by ELISA and western blot, respectively.



**FIGURE 6 |** NRP-1 knockdown reduced the colony formation ability of MDA-MB-231 cells and sensitized the cells to chemotherapy. (A) Western blot analysis showed the successful knockdown of NRP-1 to almost 58% in the isolated clone MDA- NRP-1 Knockdown Clone#22, and 99% in MDA-NRP-1 knockdown Clone#15. GAPDH was used for protein normalization. The graph represents the densitometry quantification of the mean ± standard error for the relative fold change in three independent experimental replicates and  $p < 0.05$  was considered the cut of value for significance. (B) The bar graph represents the average (±SEM) number of colonies from three independent replicates presented as the fold change in respect to the parental MDA-MB-231 untreated cells (MDA-C). The NRP-1 knockdown on its own reduced the ability of the MDA-MB-231 cells to form colonies. However, treatment with Adriamycin and Cyclophosphamide did not affect the control parental MDA-MB-231 cells, but it did decrease the number of colonies formed by the two NRP-1 knockdown cells. In the case of Paclitaxel treatment, the number of colonies were decreased in control parental MDA-MB-231 and the 58% NRP-1 knockdown cells (MDA- NRP1 knockdown C# 22) but totally inhibited the colony formation in the 99% NRP-1 knockdown cells (MDA- NRP1 knockdown C# 15). Asterisks indicate significantly different values from control untreated MDA-MD-231 cells ( $p < 0.05$ ).

who remained with high levels of NRP-1. The exploratory results obtained from such a small sample size study are still interesting for future validation on larger scale clinical research studies.

## AUTHOR CONTRIBUTIONS

NA-Z: methodology formal analysis, investigation, software, writing—review and editing. AN: data curation, methodology, software, writing—review and editing. CB: methodology formal analysis, software, writing—review and editing. MA: methodology formal analysis—review and editing. AA: data curation, investigation. SA and KA: methodology formal analysis. KM: methodology, writing—review and editing. ZA: methodology and data curation. MSA: conceptualization, methodology, investigation, writing—review and editing. SAA: conceptualization, supervision, investigation, writing—original draft, writing—review and editing, and visualization.

## REFERENCES

- Rouzier R, Perou CM, Symmans WF, Ibrahim N, Cristofanilli M, Anderson K, et al. Breast cancer molecular subtypes respond differently to preoperative chemotherapy. *Clin Cancer Res.* (2005) 11:5678–85. doi: 10.1158/1078-0432.CCR-04-2421
- Faneyte IF, Schrama JG, Peterse JL, Remijnse PL, Rodenhuis S, van De Vijver M. Breast cancer response to neoadjuvant chemotherapy: predictive markers and relation with outcome. *Br J Cancer.* (2003) 88:406. doi: 10.1038/sj.bjc.6600749
- Ferrara N, Hillan KJ, Novotny W. Bevacizumab (Avastin), a humanized anti-VEGF monoclonal antibody for cancer therapy. *Biochem Biophys Res Commun.* (2005) 333:328–35. doi: 10.1016/j.bbrc.2005.05.132
- Jang HJ, Kim BJ, Kim JH, Kim HS. The addition of bevacizumab in the first-line treatment for metastatic colorectal cancer: an updated meta-analysis of randomized trials. *Oncotarget.* (2017) 8:73009–16. doi: 10.18632/oncotarget.20314
- Tanne JH. FDA cancels approval for bevacizumab in advanced breast cancer. *BMJ.* (2011) 343:d7684. doi: 10.1136/bmj.d7684
- Ferrario C, Hostetter G, Bouchard A, Huneau MC, Mamo A, Basik M. Expression of neuropilin-1 and related proteins in breast cancer. *Proc Amer Assoc Cancer Res.* (2006) 66:1182.
- Zhang L, Wang H, Li C, Zhao Y, Wu L, Du X, et al. VEGF-A/Neuropilin 1 pathway confers cancer stemness via activating Wnt/ $\beta$ -catenin Axis in breast Cancer cells. *Cell Physiol Biochem.* (2017) 44:1251–62. doi: 10.1159/000485455
- Neufeld G, Kessler O. *The Neuropilins: Role and Function in Health and Disease.* Springer Nature (2017).
- Prud'homme GJ, Glinka Y. Neuropilins are multifunctional coreceptors involved in tumor initiation, growth, metastasis and immunity. *Oncotarget.* (2012) 3:921. doi: 10.18632/oncotarget.626
- Rizzolio S, Tamagnone L. Multifaceted role of neuropilins in cancer. *Curr Med Chem.* (2011) 18:3563–75. doi: 10.2174/092986711796642544
- Zhang L, Chen Y, Wang H, Zheng X, Li C, Han Z. mir-376a inhibits breast cancer cell progression by targeting neuropilin-1 nr. *Oncotargets Ther.* (2018) 11:5293. doi: 10.2147/OTT.S173416
- Chaudhary B, Khaled YS, Ammori BJ, Elkord E. Neuropilin 1: function and therapeutic potential in cancer. *Cancer Immunol Immunother.* (2014) 63:81–99. doi: 10.1007/s00262-013-1500-0
- Pan Q, Chanthery Y, Liang WC, Stawicki S, Mak J, Rathore N, et al. Blocking neuropilin-1 function has an additive effect with anti-VEGF to inhibit tumor growth. *Cancer Cell.* (2007) 11:53–67. doi: 10.1016/j.ccr.2006.10.018
- Bagley RG, Ren Y, Kurtzberg L, Weber W, Bangari D, Brondyk W, et al. Human choriocarcinomas: placental growth factor-dependent preclinical tumor models. *Int J Oncol.* (2012) 40:479–86. doi: 10.3892/ijo.2011.1257
- Cheng SJ, Lee JJ, Cheng SL, Chen HM, Chang HH, Wang YP, et al. Increased serum placenta growth factor level is significantly associated with progression, recurrence and poor prognosis of oral squamous cell carcinoma. *Oral Oncol.* (2012) 48:424–8. doi: 10.1016/j.oraloncology.2011.12.007
- Wei SC, Liang JT, Tsao PN, Hsieh FJ, Yu SC, Wong JM. Preoperative serum placenta growth factor level is a prognostic biomarker in colorectal cancer. *Dis Colon Rectum.* (2009) 52:1630–6. doi: 10.1007/DCR.0b013e3181afbda
- Naik A, Al-Zeheimi N, Bakheit CS, Al Riyami M, Al Jarrah A, Al Moundhri MS, et al. Neuropilin-1 associated molecules in the blood distinguish poor prognosis breast cancer: a cross-sectional study. *Sci Rep.* (2017) 7:3301. doi: 10.1038/s41598-017-03280-0
- Ghosh S, Sullivan CA, Zerkowski MP, Molinaro AM, Rimm DL, Camp RL, et al. High levels of vascular endothelial growth factor and its receptors (VEGFR-1, VEGFR-2, neuropilin-1) are associated with worse outcome in breast cancer. *Hum Pathol.* (2008) 39:1835–43. doi: 10.1016/j.humpath.2008.06.004
- Marchiò C, Maletta F, Annaratone L, Sapino A. The perfect pathology report after neoadjuvant therapy. *J Natl Cancer Inst Monogr.* (2015) 2015:47–50. doi: 10.1093/jncimonographs/lgv016
- Lester SC, Bose S, Chen YY, Connolly JL, de Baca ME, Fitzgibbons PL, et al. Protocol for the examination of specimens from patients with invasive carcinoma of the breast. *Arch Pathol Lab Med.* (2009) 133:1515–38. doi: 10.1043/1543-2165-133.10.1515
- Ellis I, Al-Sam S, Anderson N, Carder P, Deb R, Girling A, et al. *Pathology Reporting of Breast Disease in Surgical Excision Specimens Incorporating the Dataset for Histological Reporting of Breast Cancer.* London: The Royal College of Pathologists (2016).
- Aoudjit F, Vuori K. Integrin signaling inhibits paclitaxel-induced apoptosis in breast cancer cells. *Oncogene.* (2001) 20:4995. doi: 10.1038/sj.onc.1204554
- Kuczyński EA, Sargent DJ, Grothey A, Kerbel RS. Drug rechallenge and treatment beyond progression—implications for drug resistance. *Nat Rev Clin Oncol.* (2013) 10:571. doi: 10.1038/nrclinonc.2013.158
- Hong TM, Chen YL, Wu YY, Yuan A, Chao YC, Chung YC, et al. Targeting neuropilin 1 as an antitumor strategy in lung cancer. *Clin Cancer Res.* (2007) 13:4759–68. doi: 10.1158/1078-0432.CCR-07-0001
- Wey J, Gray M, Fan F, Belcheva A, McCarty M, Stoeltzing O, et al. Overexpression of neuropilin-1 promotes constitutive MAPK signalling and chemoresistance in pancreatic cancer cells. *Br J Cancer.* (2005) 93:233. doi: 10.1038/sj.bjc.6602663
- Naik A, Al-Yahyaee A, Abdullah N, Sam JE, Al-Zeheimi N, Yaish MW, et al. Neuropilin-1 promotes the oncogenic Tenascin-C/integrin  $\beta$ 3 pathway

## FUNDING

The Research Council of Oman (TRC) grant was given to SAA. TRC Grant # ORG/HSS/14/006 (TRC#137). The funding body did not contribute in any of the following: design of the study, data collection, data analysis, interpretation of data, or in writing the manuscript.

## ACKNOWLEDGMENTS

We would like to thank Sultan Qaboos University pharmacy for providing the chemotherapies used for cell culture experiment.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2019.00323/full#supplementary-material>



- and modulates chemoresistance in breast cancer cells. *BMC Cancer*. (2018) 18:533. doi: 10.1186/s12885-018-4446-y
27. Jia H, Cheng L, Tickner M, Bagherzadeh A, Selwood D, Zachary I. Neuropilin-1 antagonism in human carcinoma cells inhibits migration and enhances chemosensitivity. *Br J Cancer*. (2010) 102:541. doi: 10.1038/sj.bjc.6605539
  28. Powell J, Mota F, Steadman D, Soudy C, Miyauchi JT, Crosby S, et al. Small molecule neuropilin-1 antagonists combine antiangiogenic and antitumor activity with immune modulation through reduction of transforming growth factor beta (TGF $\beta$ ) production in regulatory T-cells. *J Med Chem*. (2018) 61:4135–54. doi: 10.1021/acs.jmedchem.8b00210
  29. Rizzolio S, Cagnoni G, Battistini C, Bonelli S, Isella C, Van Ginderachter JA, et al. Neuropilin-1 upregulation elicits adaptive resistance to oncogene-targeted therapies. *J Clin Invest*. (2018) 128:3976–90. doi: 10.1172/JCI99257
  30. Jain RK, Duda DG, Willett CG, Sahani DV, Zhu AX, Loeffler JS, et al. Biomarkers of response and resistance to antiangiogenic therapy. *Nat Rev Clin Oncol*. (2009) 6:327. doi: 10.1038/nrclinonc.2009.63
  31. Parr C, Watkins G, Boulton M, Cai J, Jiang WG. Placenta growth factor is over-expressed and has prognostic value in human breast cancer. *Eur J Cancer*. (2005) 41:2819–27. doi: 10.1016/j.ejca.2005.07.022
  32. Willett CG, Duda DG, di Tomaso E, Boucher Y, Ancukiewicz M, Sahani DV, et al. Efficacy, safety, and biomarkers of neoadjuvant bevacizumab, radiation therapy, and fluorouracil in rectal cancer: a multidisciplinary phase II study. *J Clin Oncol*. (2009) 27:3020–6. doi: 10.1200/JCO.2008.21.1771
  33. Chen S, Itoh T, Wu K, Zhou D, Yang C. Transcriptional regulation of aromatase expression in human breast tissue. *J Steroid Biochem Mol Biol*. (2002) 83:93–9. doi: 10.1016/S0960-0760(02)00276-5

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Al-Zeheimi, Naik, Bakheit, Al Riyami, Al Ajarrah, Al Badi, Al Baimani, Malik, Al Habsi, Al Moundhri and Adham. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# A Pathology-Based Combined Model to Identify PAM50 Non-luminal Intrinsic Disease in Hormone Receptor-Positive HER2-Negative Breast Cancer

## OPEN ACCESS

### Edited by:

Raquel Nunes,  
Johns Hopkins University,  
United States

### Reviewed by:

Tomas Reinert,  
Federal University of Rio Grande do  
Sul, Brazil  
Alessandro Igor Cavalcanti Leal,  
Johns Hopkins Medicine,  
United States

### \*Correspondence:

Aleix Prat  
alprat@clinic.cat

### Specialty section:

This article was submitted to  
Women's Cancer,  
a section of the journal  
Frontiers in Oncology

**Received:** 14 January 2019

**Accepted:** 02 April 2019

**Published:** 26 April 2019

### Citation:

Pascual T, Martin M, Fernández-Martínez A, Paré L, Alba E, Rodríguez-Lescure Á, Perrone G, Cortés J, Morales S, Lluch A, Urruticoechea A, González-Farré B, Galván P, Jares P, Rodríguez A, Chic N, Righi D, Cejalvo JM, Tonini G, Adamo B, Vidal M, Villagrasa P, Muñoz M and Prat A (2019) A Pathology-Based Combined Model to Identify PAM50 Non-luminal Intrinsic Disease in Hormone Receptor-Positive HER2-Negative Breast Cancer. *Front. Oncol.* 9:303. doi: 10.3389/fonc.2019.00303

Tomás Pascual<sup>1,2</sup>, Miguel Martín<sup>3,4,5</sup>, Aranzazu Fernández-Martínez<sup>6</sup>, Laia Paré<sup>1,2</sup>, Emilio Alba<sup>4,5,7</sup>, Álvaro Rodríguez-Lescure<sup>4,8</sup>, Giuseppe Perrone<sup>9</sup>, Javier Cortés<sup>10,11</sup>, Serafín Morales<sup>12</sup>, Ana Lluch<sup>4,5,13,14,15</sup>, Ander Urruticoechea<sup>16</sup>, Blanca González-Farré<sup>2,17</sup>, Patricia Galván<sup>1</sup>, Pedro Jares<sup>17</sup>, Adela Rodríguez<sup>1</sup>, Nuria Chic<sup>1</sup>, Daniela Righi<sup>9</sup>, Juan Miguel Cejalvo<sup>1</sup>, Giuseppe Tonini<sup>9</sup>, Barbara Adamo<sup>1</sup>, Maria Vidal<sup>1</sup>, Patricia Villagrasa<sup>2</sup>, Montserrat Muñoz<sup>1</sup> and Aleix Prat<sup>1,2\*</sup>

<sup>1</sup> Medical Oncology Department, Hospital Clínic de Barcelona, Barcelona, Spain, <sup>2</sup> SOLTI Breast Cancer Research Group, Barcelona, Spain, <sup>3</sup> Medical Oncology Department, Hospital Gregorio Marañón, Universidad Complutense, Madrid, Spain, <sup>4</sup> GEICAM (Spanish Breast Cancer Group), Madrid, Spain, <sup>5</sup> Centro de Investigación Biomédica en Red de Oncología, CIBERONC-ISCIII, Madrid, Spain, <sup>6</sup> Department of Genetics, University of North Carolina, Chapel Hill, NC, United States, <sup>7</sup> Medical Oncology Department, Hospital Universitario Virgen de la Victoria, IBIMA, Málaga, Spain, <sup>8</sup> Medical Oncology Department, Hospital Universitario de Elche, Elche, Spain, <sup>9</sup> Department of Medicine, Università Campus Bio-Medico di Roma, Rome, Italy, <sup>10</sup> IOB Institute of Oncology, Quironsalud Group, Madrid, Spain, <sup>11</sup> Vall d'Hebron Institute of Oncology (VHIO), Barcelona, Spain, <sup>12</sup> Medical Oncology Department, Hospital Arnau de Vilanova, Lleida, Spain, <sup>13</sup> Medical Oncology Department, Hospital Clínico Universitario, Valencia, Spain, <sup>14</sup> Biomedical Research Institute INCLIVA, Valencia, Spain, <sup>15</sup> Department of Medicine, Universitat de València, Valencia, Spain, <sup>16</sup> Department of Medical Oncology, Fundació Onkologikoa, Donostia, Spain, <sup>17</sup> Department of Pathology, Hospital Clínic de Barcelona, Barcelona, Spain

**Background:** In hormone receptor-positive (HR+)/HER2-negative breast cancer, the HER2-enriched and Basal-like intrinsic subtypes are associated with poor outcome, low response to anti-estrogen therapy and high response to chemotherapy. To date, no validated biomarker exists to identify both molecular entities other than gene expression.

**Methods:** PAM50 subtyping and immunohistochemical data were obtained from 8 independent studies of 1,416 HR+/HER2-negative early breast tumors. A non-luminal disease score (NOLUS) from 0 to 100, based on percentage of estrogen receptor (ER), progesterone receptor (PR) and Ki67 tumor cells, was derived in a combined cohort of 5 studies (training dataset) and tested in a combined cohort of 3 studies. The performance of NOLUS was estimated using Area Under the ROC Curve (AUC).

**Results:** In the training dataset ( $n = 903$ ) and compared to luminal disease, non-luminal disease had lower percentage of ER-positive cells (median 65.2 vs. 86.2%,  $p < 0.01$ ) and PR-positive cells (33.2 vs. 56.4%,  $p < 0.01$ ) and higher percentage of Ki67-positive cells (18.2 vs. 13.1%,  $p = 0.01$ ). A NOLUS formula was derived:  $-0.45 \times \text{ER} - 0.28 \times \text{PR} + 0.27 \times \text{Ki67} + 73.02$ . The proportion of non-luminal tumors in NOLUS-positive ( $\geq 51.38$ ) and NOLUS-negative ( $< 51.38$ ) groups was 52.6 and 8.7%, respectively. In the testing dataset ( $n = 514$ ), NOLUS was found significantly associated

with non-luminal disease ( $p < 0.01$ ) with an AUC 0.902. The proportion of non-luminal tumors in NOLUS-positive and NOLUS-negative groups was 76.9% (56.4–91.0%) and 2.6% (1.4–4.5%), respectively. The sensitivity and specificity of the pre-specified cutoff was 59.3 and 98.7%, respectively.

**Conclusions:** In the absence of gene expression data, NOLUS can help identify non-luminal disease within HR+/HER2-negative breast cancer.

**Keywords:** intrinsic subtype, non-luminal, PAM50, breast cancer, gene expression

## INTRODUCTION

Gene expression profiling has had a considerable impact on our understanding of hormone receptor-positive (HR+)/HER2-negative breast cancer biology (1, 2). During the last decade, two intrinsic molecular subtypes within HR+/HER2-negative disease (i.e., Luminal A and Luminal B) have been identified and intensively studied (3–5). These studies have led to well-validated prognostic gene expression-based tests such as Prosigna (6), OncotypeDX (7), MammaPrint (8), Breast Cancer Index (9), and EndoPredict (10). The implementation of these 4 platforms in the clinical practice has been essential in order to identify a subset of Luminal A tumors that can safely spare (neo)adjuvant chemotherapy treatments because of their good prognostic (11–13).

At the same time, cumulative evidence from recent studies suggests that 5–30% of HR+/HER2-negative tumors are not Luminal A or B by gene expression and fall into the HER2-enriched (HER2-E) and Basal-like categories (14). From a clinical perspective, these non-luminal tumors have been associated with low estrogen dependency (15–17), high chemo-sensitivity (18–20), potential lower activity of CDK4/6 inhibitors (21, 22) and poor outcome in both early and the advanced/metastatic breast cancer (22–24). Thus, clinical utility of the identification of the two non-luminal subtypes within HR+/HER2-negative disease is now being pursued.

In this study, we sought to validate a simple pathology-based model to help clinicians and researchers identify non-luminal disease within HR+/HER2-negative breast cancer in the absence of gene expression data.

## MATERIALS AND METHODS

### Study Design

PAM50 gene expression and pathology-based data from 1,416 HR+/HER2-negative early breast tumors were obtained from 8 independent studies that are summarized in Table 1 (20, 25–30). The GEICAM/9906 is a phase III adjuvant trial in women with lymph node-positive disease that compared treatment with fluorouracil, epirubicin, and cyclophosphamide (FEC) or with FEC followed by weekly paclitaxel (FEC-P) (25). A total of 531 HR+/HER2-negative tumor samples were analyzed (26). SOLTI-1007 NeoEribulin trial is a neoadjuvant trial within HER2-negative breast cancer, where patients were treated with eribulin monotherapy for 4 cycles (20). A total of 93 HR+/HER2-negative baseline tumor samples were

analyzed. Pre-operative endocrine treatment (PETx) cohort is a retrospective Spanish registry of 56 patients with HR+/HER2-negative disease treated with neoadjuvant endocrine therapy. From this study, baseline samples were analyzed (30). From GEICAM/2009-03\_CONVERTHER, a study that aimed to compared pathology and gene expression data between primary and metastatic tumor samples, we obtained 50 HR+/HER2-negative primary tumor samples (28, 31). GEICAM/2012-09 is a prospective study of the Spanish Breast Cancer Research Group to characterize the impact of Prosigna assay in adjuvant treatment decision of postmenopausal patients with HR+/HER2-negative breast cancer without nodal involvement (27). A total of 174 primary tumor samples were included. Hospital Clinic of Barcelona (HCB) cohort is a consecutive series of 194 tumor samples where Prosigna has been performed as routine clinical care (29). Università Campus Bio-Medico di Roma (CBM) cohort is a consecutive series of 145 tumor samples where Prosigna has been performed as routine clinical care (29). Instituto de Investigación Biomédica de Málaga (IBIMA) cohort includes 180 HR+/HER2-negative baseline tumors treated with neoadjuvant chemotherapy as routine clinical practice (18).

### Pathology-Based Data

The formalin-fixed paraffin-embedded tumor samples analyzed met the following criteria: (1) they were obtained from untreated primary tumors, (2) estrogen receptor (ER) and progesterone receptor (PR) positivity was defined as >1% positive tumor cells according to the ASCO/CAP guidelines (32), (3) HER2-negativity was defined according to the 2013 ASCO/CAP guidelines (33). Ki67 IHC was quantified according to the 2011 Guidelines developed by the International Ki67 in Breast Cancer working group (34).

### PAM50 Intrinsic Subtyping

A research-based PAM50 subtyping assay was performed using the nCounter as previously described (24, 35, 36), except in GEICAM/9906, where a research-based PAM50 qRT-PCR-based assay was used, and GEICAM/2012-09, HCB, IBIMA, and CBM datasets, which used the standardized and commercial version of the PAM50 assay (i.e., Prosigna®). Original subtype calls obtained from each study were used. From the research-based PAM50 version, we eliminated any tumor samples identified as normal-like.

**TABLE 1** | Main features of the cohorts analyzed in this study.

	GEICAM/ 9906	SOLTI- Neoeribulin	PETx	GEICAM/ 2009-03	GEICAM/ 2012-09	HCB	IBIMA	CBM
Dataset	Training	Training	Training	Training	Training	Testing	Testing	Testing
N	531	93	56	50	173	194	176	144
IHC	Centralized	Local	Local	Centralized	Centralized	Centralized	Centralized	Centralized
Platform	qRT-PCR	nCounter	nCounter	nCounter	nCounter	nCounter	nCounter	nCounter
PAM50 non-luminal disease (%)	77 (14.5)	12 (12.9)	3 (5.3)	7(14)	5 (2.9)	7 (3.6)	21 (11.9)	5 (3.5)
HER2-E (%)	71 (13.4)	1 (1.1)	3 (5.3)	6 (12)	4 (2.3)	4 (2.1)	7 (4.0)	3 (2.1)
Basal-like (%)	6 (1.3)	11 (11.8)	0	1 (2)	1 (0.6)	3 (1.5)	14 (7.9)	2 (1.4)

## Non-luminal Disease Score (NOLUS)

A combined score to identify non-luminal disease by PAM50 was derived from a combined dataset of 5 studies (i.e., training dataset) using ER, PR, and Ki67 levels (i.e., % of positive tumor cells). The optimal cutoff was defined as the point with the most significant (Fisher's exact test) split between Luminal and non-luminal disease. Once NOLUS was developed, the final model and cutoff were tested in 513 HR+/HER2-negative tumors (i.e., testing set) from 3 independent databases: HCB, IBIMA, and CBM studies.

## Statistical Analysis

Univariate and multivariable logistic regression analyses were done to investigate the association of each IHC biomarkers with non-luminal disease. Odds ratios (ORs) and 95% confidence intervals (CI) were calculated for each variable. The performance of NOLUS was estimated using Area Under the ROC Curve (AUC). 10-fold cross-validation was conducted (37). The significance level was set to a two-sided  $\alpha$  of 0.05. We used R version 3.3.1 for all the statistical analyses.

## RESULTS

### Proportion of Non-luminal Disease Within HR+/HER2-Negative Breast Cancer

A total of 903 HR+/HER2-negative tumor samples from 5 studies were used as the training dataset (Table 1). In this cohort, non-luminal subtypes represented 11.6% (105/903) of the cases, ranging from 2.9% in GEICAM/2012-09 to 14.5% in GEICAM/9906. As expected, a relationship between chemotherapy cohorts and higher proportion of non-luminal disease was found. The 3 chemotherapy cohorts had proportions of non-luminal disease >10%, whereas the 2 hormonotherapy cohorts, the Spanish neoadjuvant endocrine therapy registry (PETx) and the GEICAM/2012-09 prospective study, had 2.9 and 5.4% of non-luminal tumors, respectively.

### Expression of ER, PR, and Ki67 in Non-luminal Disease in the Training Dataset

ER, PR, and Ki67 were found differentially expressed ( $p < 0.001$ ) between PAM50 luminal ( $n = 798$ ) and non-luminal ( $n = 105$ )

disease. Non-luminal disease had lower percentage of ER-positive cells (median 65.2 vs. 86.2%,  $p < 0.01$ ) and PR-positive cells (33.2 vs. 56.4%,  $p < 0.01$ ) and higher percentage of Ki67-positive cells (18.2 vs. 13.1%,  $p = 0.01$ ) compared to luminal disease (Figure 1).

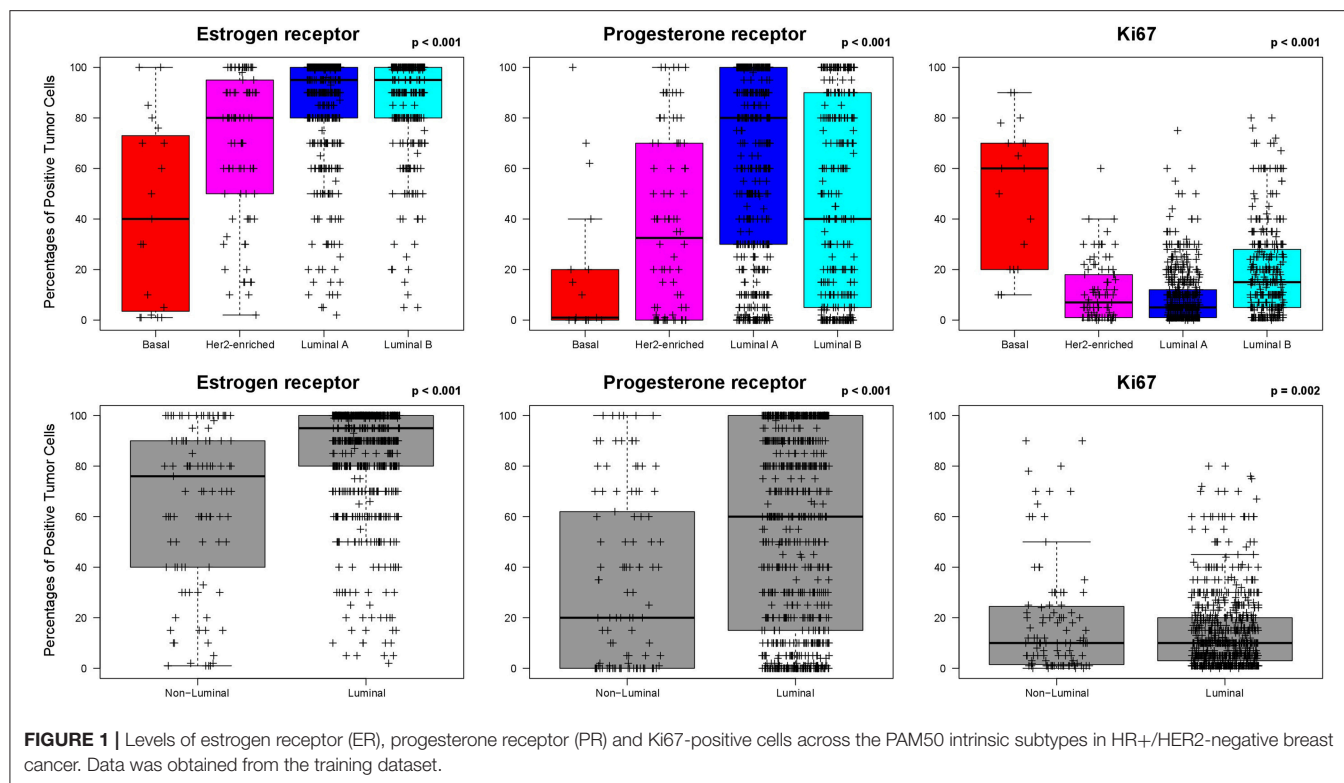
### Predicting Non-luminal Disease Using ER, PR, and Ki67

To evaluate if ER, PR, and Ki67 (measured as continuous variables) provide independent information from each other regarding the identification of non-luminal disease, a multivariable logistic regression model was applied (Table S1). Interestingly, the expression of the 3 biomarkers was found independently associated with non-luminal disease. Using this multivariable result, we developed a combined score, called non-luminal disease score (NOLUS), that weights the value of each biomarker to identify non-luminal disease. The estimated coefficient of each variable in the logistic model was used to derive NOLUS ( $0-100$ ) =  $-0.45 \cdot \text{ER\%} - 0.28 \cdot \text{PR\%} + 0.27 \cdot \text{Ki67\%} + 73$ , where ER, PR, and Ki67 are measured as continuous variables based on the percentage of positive tumor cells by immunohistochemistry.

Next, we identified a NOLUS cutoff to identify non-luminal disease based on the most significant split using a Fisher's exact test. Using this cutoff of 51.38, the proportion of NOLUS-positive ( $\geq 51.38$ ) tumors and NOLUS-negative ( $< 51.38$ ) tumors was 6.3 and 93.7%, respectively. In addition, the proportion of non-luminal tumors in NOLUS-positive and NOLUS-negative groups was 52.6% (95% CI 38.9–66.0) and 8.7% (95% CI 6.97–10.77), respectively ( $p < 0.001$ ) (Figure 2).

### Validation of NOLUS in the Testing Dataset

The testing dataset was composed of 514 HR+/HER2-negative tumor samples from 3 independent studies (HCB, IBIMA and CBM). The proportion of non-luminal disease here was 6.2% (33/514). NOLUS as a continuous variable was found significantly associated with non-luminal disease ( $p < 0.01$ ) with an AUC 0.902 (Figure 2). The proportion of non-luminal tumors in NOLUS-positive and NOLUS-negative groups was 76.9% (56.4–91.0) and 2.6% (1.4–4.5), respectively ( $p < 0.01$ ). The sensitivity was 59.3 and the specificity was 98.7%. To identify only HER2-E, the sensitivity was 42.8 and the specificity was 96.0%. To identify only Basal-like, the sensitivity was 53.9 and the specificity was 99.0%.



## NOLUS in All Datasets

We explored NOLUS in all datasets combined. The odds of being non-luminal subtype increase 6.8% for every point increase (OR = 1.068, 95% CI 1.06–1.08,  $p < 0.001$ ). The rates of non-luminal in NOLUS-negative and NOLUS-positive were 6.52 and 60.24%, respectively (Adjusted OR = 23.82, 95% CI 13.97–40.61,  $p < 0.001$ ) (Figure 3).

Finally, the model was validated using 10-fold cross validation. The data was separated into 10 sets, each set containing 10% of the data. For each validation round, 9 sets were used as training data, and the other set was used as testing data to validate the model using the linear discriminant analysis method. The accuracy of the model with 10-fold cross-validation was 0.97 (Cohen's kappa coefficient = 0.83).

## DISCUSSION

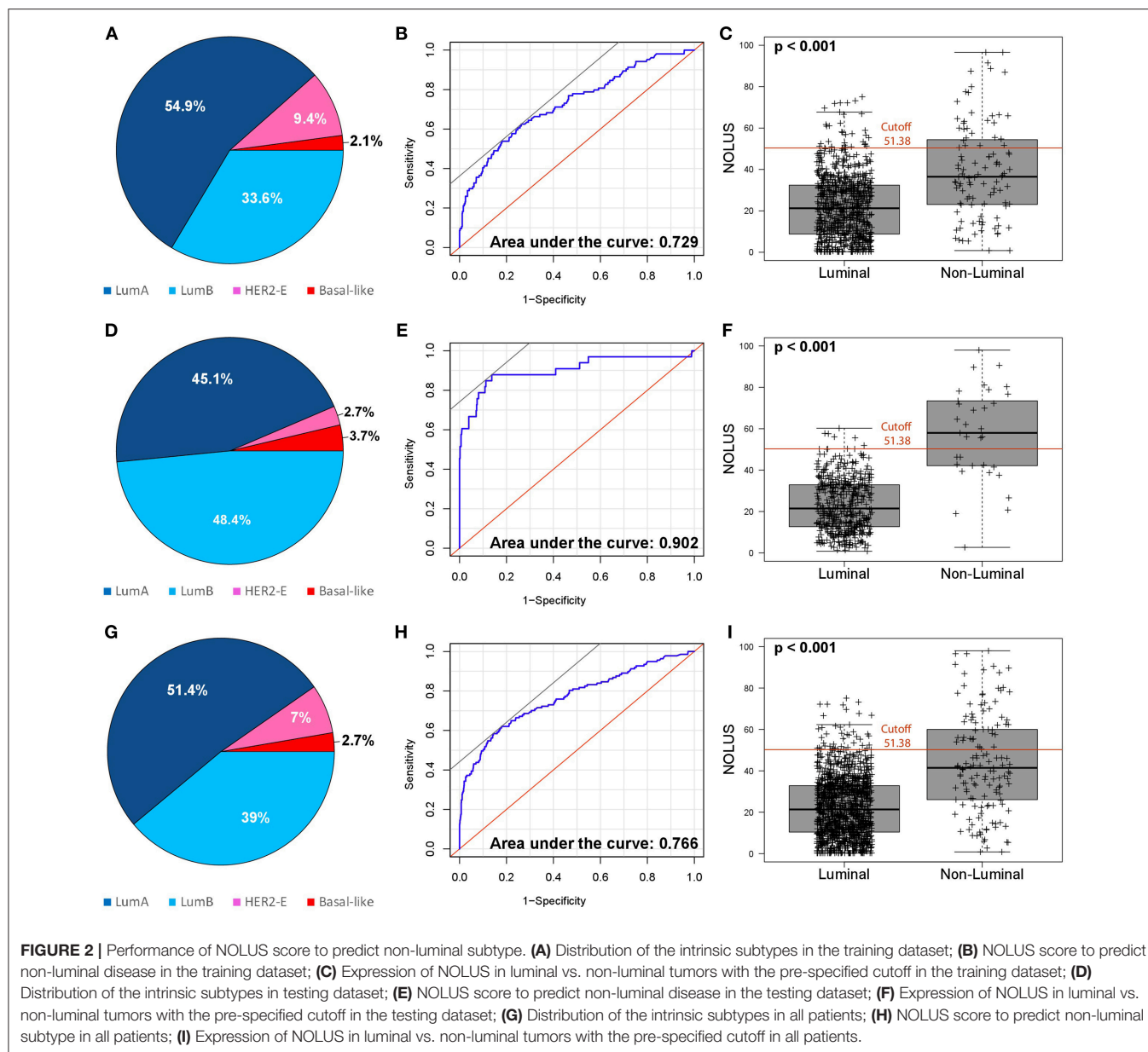
In this study, we aimed to identify a pathology-based model that is easy, fast and with the potential to be widely implemented to identify non-luminal disease within HR+/HER2-negative breast cancer when gene expression data is not available. The main reasons are that there is accumulating evidence that non-luminal disease within HR+/HER2-negative disease represents a distinct biological and clinical entity (14) that deserves substantial attention and that gene expression-based assays are not always readily available in daily clinical practice. To our knowledge, this is the first report to attempt to derive a pathology-based

predictive model to identify PAM50 non-luminal disease within HR+/HER2-negative disease.

The importance of intrinsic subtyping was highlighted in one of the most complete molecular characterization studies that has ever been performed in breast cancer (4). In this study, led by The Cancer Genome Atlas Project (TCGA), more than 500 primary breast cancer were extensively profiled at the DNA (i.e., methylation, chromosomal copy-number changes and somatic and germline mutations), RNA (i.e., miRNA and mRNA expression) and protein (i.e., protein and phosphor-protein expression) levels using the most recent technologies (4). In a particular analysis of over 300 primary tumors, 5 different data-types (i.e., all except DNA mutations) were combined together in a cluster of clusters in order to identify how many biological homogenous groups of tumors one can identify in breast cancer. The consensus clustering results showed the presence of 4 main entities of breast cancer but, more importantly, these 4 entities were found to be very-well recapitulated by the 4 main intrinsic subtypes (Luminal A, Luminal B, HER2-E, and Basal-like) as defined by mRNA expression only (3, 5, 6, 36, 38–40). Overall, these results suggest that intrinsic subtyping captures the vast majority of the biological diversity occurring in breast cancer.

Although the incidence of the Basal-like and HER2-E subtypes within HR+/HER2-negative tumors is below 10% in the primary disease setting (4), current evidence suggest that this frequency is much larger in the advanced/metastatic setting, specially following endocrine treatment (14). The increase proportion of the HER2-E subtype in the metastatic setting may be due to setting selection, a change in the biology of the tumor due to the

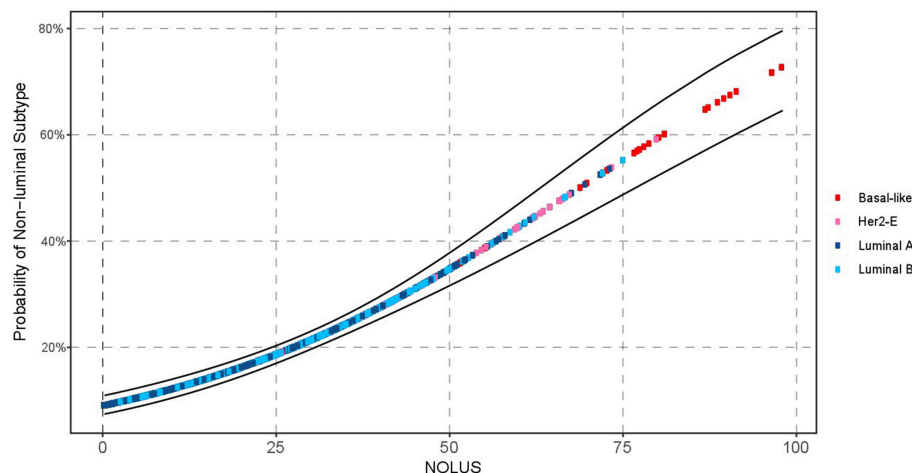




inherent evolution of the tumor or the effects of the treatment, or a combination of both. Current evidence supports this latter possibility. Patients with early HR+/HER2-negative/HER2-E breast cancer have a higher probability of relapse than luminal disease. Therefore, it is likely that a given population of patients with metastatic disease is more enriched for the HER2-E subtype compared to patients with early breast cancer. Moreover, using 123 pairs of primary vs. metastatic tumor samples with a high proportion of HR+/HER2-negative tumors, Cejalvo et al. (28) showed that the HER2-E signature and HER2-E subtype are enriched in the metastatic samples compared to primary tumors. For example, 13% of primary Luminal A and B tumors were identified as HER2-E in the relapsed tumor sample. Overall, the proportion of HER2-E tumors in primary vs. metastatic was 11.4

vs. 22%, respectively. Moreover, in a retrospective analysis of tumor samples from the BOLERO-2 study, where patients with HR+/HER2-negative advanced disease resistant to an aromatase inhibitor, the proportion of HER2-E in primary vs. metastatic tumors was 19 vs. 32% (41). Recently, gene expression data from the PALOMA-2 clinical trial have been presented (21, 22). In this retrospective analysis, which included 68% (445/666) of the tumors of both primary and metastatic tumors within the clinical trial population, the HER2-E population represented 19 and the Basal-like population represented 1%.

The prognostic value of the Basal-like and HER2-E intrinsic subtypes in HR+/HER2-negative breast cancer has been evaluated in several studies (22–24). For example, intrinsic subtyping performed in a cohort of 1,380 patients with ER+



**FIGURE 3 |** Probability of non-luminal disease as a function of NOLUS in all patients.

early breast cancer treated with 5 years of adjuvant tamoxifen-only (23) demonstrated the presence of a 7% of non-Luminal disease. These patients showed a statistically significant worse outcome compared to Luminal A subpopulation. The prognostic value of the HER2-E intrinsic subtype has been evaluated also in 3 retrospective studies involving HR+/HER2-negative metastatic patients (22, 24, 41). In the EGF30008 Phase III clinical trial, intrinsic subtyping was performed in a cohort of 821 patients with HR-positive disease (644 HER2-negative and 157 HER2+) treated in the first-line metastatic setting with either letrozole or letrozole plus lapatinib (24). Patients with HER2-E and Basal-like disease showed worse outcome in terms of progression free survival (PFS) and overall survival (OS) compared to Luminal A disease regardless of the HER2 status and treatment. Compared with the Luminal A subtype, the non-luminal subtypes showed a significantly decreased PFS independently of other clinical-pathological variables. Patients with HER2-E, and Basal-like subtypes had a 2.87, and 2.26 times higher risk of tumor progression, respectively. Median PFS differed across the intrinsic subtypes: Luminal A (16.9 months), Luminal B (11.0 months), HER2-E (4.7 months), and Basal-like (4.1 months). In the second study, PAM50 was performed in 261 tumor samples from the BOLERO-2 phase III trial (41). The subtype distribution was: 46.7% Luminal A, 21.5% HER2-negativeE, 15.7% Luminal B, 14.2% Normal-like and 1.9% Basal-like. Non-luminal disease was independently associated with poor PFS and OS compared to the luminal subtypes. In the third study, PAM50 was performed in 465 tumor samples from the PALOMA-2 phase III trial. Both non-luminal subtypes were associated with worse PFS compared to Luminal A subtype. These results support that non-luminal HR+/HER2-negative tumors are aggressive and require novel therapeutic approaches.

The ability of the Basal-like and HER2-E subtype to predict benefit from anti-estrogen therapy has been evaluated in the neoadjuvant setting. In the Z1031 neoadjuvant trial (16) within

ER+/HER2-negative disease, patients with HER2-E or Basal-like disease had persistently high surgical Ki67 levels (20%) after 4–6 months of treatment with an aromatase inhibitor, consistent with high-level estrogen-independent growth. In another retrospective study of 112 postmenopausal women with stages I–IIIB ER+ early breast cancer before and after 2-weeks' anastrozole treatment in a neoadjuvant trial, patients with HER2-E subtype ( $n = 9$  [8.0%]) or Basal-like subtype ( $n = 3$  [2.7%]) showed a poorer Ki67 response (mean Ki-67 change of  $-50.7$  and  $+15.3\%$ ) compared to Luminal A or B subtypes (mean Ki-67 change of  $-75\%$ ). Interestingly, this study also profiled post-treatment samples. As expected, the vast majority of Luminal A samples (31/32, 97%) continued being Luminal A. However, although the majority of Luminal B tumors became Luminal A (9/17, 53%), 12% (2/17) became HER2-E. Overall, this data, together with the poor PFS of the HER2-E subtype following endocrine therapy in EGF30008, BOLERO-2 and PALOMA 2 trials (22, 24, 41), suggest that both non-luminal subtypes within HR-positive disease might not benefit substantially from anti-estrogen therapy.

The ability of the Basal-like and HER2-E subtype to predict benefit from palbociclib has been recently evaluated in 465 samples of the PALOMA-2 study (22). The increase in median PFS in the HER2-E subtype was modest (2.8 months), compared to the increase in median PFS of 13.4 and 8.6 months in Luminal A and B subtypes, respectively. Regarding Basal-like, only 1 patient was identified and progressed at 6.4 months following letrozole plus palbociclib. This data suggest that non-luminal subtypes do not benefit much from CDK4/6 inhibition. In the neoadjuvant setting, Ma and colleagues conducted the NEOPALANA clinical trial with anastrozole and palbociclib. Two non-luminal tumors were identified by PAM50 (1 HER2-E and 1 Basal-like) and, interestingly, none of the 2 patients responded to the combined treatment (17).

The ability of the Basal-like and HER2-E subtype to predict chemotherapy sensitivity within HR+/HER2- disease has been

evaluated in the neoadjuvant setting. In one study, we evaluated the pathological complete response (pCR) rated in 451 patients with HR+/HER2-negative disease treated with standard multi-agent neoadjuvant chemotherapy (42). The pCR rates in the non-luminal subtype was 23.2% compared to 15% in Luminal B and 5% in Luminal A tumors. In another neoadjuvant study, Prat and colleagues evaluated the residual cancer burden (RCB) 0/1 rates of the intrinsic subtypes in 180 patients with HR+/HER2-negative disease treated with anthracycline/taxane-based chemotherapy (18). Concordant with the first study, the RCB0/1 rates were higher in the non-luminal subtypes (38.1%) compared to Luminal B (20.0%) and Luminal A (9.3%). Overall, this data suggests that within HR+/HER2-negative disease, non-luminal tumors are highly chemo-sensitive.

Our study has several limitations worth noting. For example, determination of ER, PR and Ki67 was not performed centrally in a single lab and, in 2 studies, IHC data was obtained from local pathology reports. In addition, each study used different pathology-based assays. Although this heterogeneity is a limitation, its effects must not be large since the proportion of non-luminal disease across studies was similar and the fact that NOLUS was able to predict non-luminal disease in both the training and testing sets with similar performance. Another limitation is that NOLUS is not a standardized assay; thus, analytical validity is lacking. However, the biomarkers that compose NOLUS (i.e., ER, PR, and Ki67) have not been standardized; thus, NOLUS will suffer from lack of standardization as well. Another aspect is that we did not aim to derive a model that could further distinguish Basal-like from HER2-E subtypes within non-luminal disease. The main reason is that at this point it is unclear what are the clinical implications of each of these entities both from a prognostic and predictive point of view. However, as more data is gathered, NOLUS could be updated in the future to further distinguish these 2 non-luminal subtypes. Finally, we do not provide clinical validation of the NOLUS predictor.

To conclude, NOLUS is a tool that, in the absence of gene expression-based assays, may help identify non-luminal

disease within HR+/HER2-negative breast cancer. Overall, the data clearly suggest that both non-luminal subtypes provide additional prognostic and predictive information beyond HR and HER2 status and may support more informed treatment decisions (1). For example, to identify patients who are not good candidates for endocrine therapy alone. Pivotal and large studies evaluating prognosis and treatment benefits can now apply NOLUS and further define the clinical validity and clinical utility of this biomarker.

## AUTHOR CONTRIBUTIONS

All authors participated in the design and/or interpretation of the reported results and participated in the acquisition and/or analysis of data. In addition, all authors participated in drafting and/or revising the manuscript and provided administrative, technical, or supervisory support.

## FUNDING

This work was supported by Instituto de Salud Carlos III - PI16/00904 (to AP), Banco Bilbao Vizcaya Argentaria Foundation (to AP), Pas a Pas (to AP), Save the Mama (to AP), Breast Cancer Research Foundation (to AP), Fundación Mutua Madrileña- Investigación en Salud 2018 (to AP), SEOM Translational Research Grant (to AF-M.), Banca d'Italia (to GP), and by T.C.I. telecomunicazioni (to GP).

## ACKNOWLEDGMENTS

We thank all the patients and their family members for participating in the studies.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2019.00303/full#supplementary-material>

## REFERENCES

- Prat A, Perou CM. Deconstructing the molecular portraits of breast cancer. *Mol Oncol.* (2011) 5:5–23. doi: 10.1016/j.molonc.2010.11.003
- Prat A, Pineda E, Adamo B, Galvan P, Fernandez A, Gaba L, et al. Clinical implications of the intrinsic molecular subtypes of breast cancer. *Breast.* (2015) 24 (Suppl. 2):S26–35. doi: 10.1016/j.breast.2015.07.008
- Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res.* (2010) 12:R68. doi: 10.1186/bcr2635
- Network CGA. Comprehensive molecular portraits of human breast tumours. *Nature.* (2012) 490:61. doi: 10.1038/nature11412
- Prat A, Parker JS, Fan C, Perou CM. PAM50 assay and the three-gene model for identifying the major and clinically relevant molecular subtypes of breast cancer. *Breast Cancer Res Treat.* (2012) 135:301–6. doi: 10.1007/s10549-012-2143-0
- 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) 27:1160–7. doi: 10.1200/JCO.2008.18.1370
- Sparano JA, Gray RJ, Makower DE, Pritchard KI, Albain KS, Hayes DE, et al. Prospective Validation of a 21-Gene Expression Assay in Breast Cancer. *N Engl J Med.* (2015) 373:2005–14. doi: 10.1056/NEJMoa1510764
- Cardoso F, Van't Veer LJ, Bogaerts J, Slaets L, Viale G, Delaloge S, et al. 70-Gene Signature as an Aid to treatment decisions in early-stage breast cancer. *N Engl J Med.* (2016) 375:717–29. doi: 10.1056/NEJMoa1602253
- Sgroi DC, Sestak I, Cuzick J, Zhang Y, Schnabel CA, Schroeder B, et al. Prediction of late distant recurrence in patients with oestrogen-receptor-positive breast cancer: a prospective comparison of the breast-cancer index (BCI) assay, 21-gene recurrence score, and IHC4 in the TransATAC study population. *Lancet Oncol.* (2013) 14:1067–76. doi: 10.1016/S1470-2045(13)70387-5
- Varga Z, Sinn P, Fritzsche F, Von Hochstetter A, Noske A, Schraml P, et al. Comparison of EndoPredict and Oncotype DX test results in

- hormone receptor positive invasive breast cancer. *PLoS ONE*. (2013) 8:e58483. doi: 10.1371/annotation/f715f38e-7aee-4d2b-8bbf-da0411dc6ef3
11. Dowsett M, Sestak I, Lopez-Knowles E, Sidhu K, Dunbier AK, Cowens JW, et al. Comparison of PAM50 risk of recurrence score with oncotype DX and IHC4 for predicting risk of distant recurrence after endocrine therapy. *J Clin Oncol*. (2013) 31:2783–90. doi: 10.1200/JCO.2012.46.1558
  12. Gnant M, Filipits M, Greil R, Stoeger H, Rudas M, Bago-Horvath Z, et al. Predicting distant recurrence in receptor-positive breast cancer patients with limited clinicopathological risk: using the PAM50 Risk of Recurrence score in 1478 postmenopausal patients of the ABCSG-8 trial treated with adjuvant endocrine therapy alone. *Ann Oncol*. (2014) 25:339–45. doi: 10.1093/annonc/mdt494
  13. Kwa M, Makris A, Esteva FJ. Clinical utility of gene-expression signatures in early stage breast cancer. *Nat Rev Clin Oncol*. (2017) 14:595–610. doi: 10.1038/nrclinonc.2017.74
  14. Cejalvo JM, Pascual T, Fernandez-Martinez A, Braso-Maristany F, Gomis RR, Perou CM, et al. Clinical implications of the non-luminal intrinsic subtypes in hormone receptor-positive breast cancer. *Cancer Treat Rev*. (2018) 67:63–70. doi: 10.1016/j.ctrv.2018.04.015
  15. Dunbier AK, Anderson H, Ghazoui Z, Salter J, Parker JS, Perou CM, et al. Association between breast cancer subtypes and response to neoadjuvant anastrozole. *Steroids*. (2011) 76:736–40. doi: 10.1016/j.steroids.2011.02.025
  16. Ellis MJ, Suman VJ, Hoog J, Lin L, Snider J, Prat A, et al. Randomized phase II neoadjuvant comparison between letrozole, anastrozole, and exemestane for postmenopausal women with estrogen receptor-rich stage 2 to 3 breast cancer: clinical and biomarker outcomes and predictive value of the baseline PAM50-based intrinsic subtype–ACOSOG Z1031. *J Clin Oncol*. (2011) 29:2342–9. doi: 10.1200/JCO.2010.31.6950
  17. Ma CX, Gao F, Luo J, Northfelt DW, Goetz M, Forero A, et al. NeoPalAna: Neoadjuvant Palbociclib, a Cyclin-Dependent Kinase 4/6 Inhibitor, and Anastrozole for Clinical Stage 2 or 3 Estrogen Receptor-Positive Breast Cancer. *Clin Cancer Res*. (2017) 23:4055–65. doi: 10.1158/1078-0432.CCR-16-3206
  18. Prat A, Galvan P, Jimenez B, Buckingham W, Jeiranian HA, Schaper C, et al. Prediction of Response to Neoadjuvant Chemotherapy Using Core Needle Biopsy Samples with the Prosigna Assay. *Clin Cancer Res*. (2016) 22:560–6. doi: 10.1158/1078-0432.CCR-15-0630
  19. Fujii T, Kogawa T, Dong W, Sahin AA, Moulder S, Litton JK, et al. Revisiting the definition of estrogen receptor positivity in HER2-negative primary breast cancer. *Ann Oncol*. (2017) 28:2420–8. doi: 10.1093/annonc/mdx397
  20. Prat A, Ortega V, Villagrasa P, Paré L, Galván P, Oliveira M, et al. Abstract P1-09-09: efficacy and gene expression results from SOLT11007 NEOERIBULIN phase II clinical trial in HER2-negative early breast cancer. *Cancer Res*. (2017) 77:P1-09-09. doi: 10.1158/1538-7445.SABCS16-P1-09-09
  21. Finn RS, Martin M, Rugo HS, Jones S, Im S-A, Gelmon K, et al. Palbociclib and letrozole in advanced breast cancer. *New Engl J Med*. (2016) 375:1925–36. doi: 10.1056/NEJMoa1607303
  22. Finn R, Liu Y, Martin M, Rugo H, Dieras V, Im S, et al. “Abstract P2-09-10: Comprehensive gene expression biomarker analysis of CDK 4/6 and endocrine pathways from the PALOMA-2 study”. *Cancer Res*. (2018) 78. doi: 10.1158/1538-7445.SABCS17-P2-09-10
  23. Prat A, Parker JS, Fan C, Cheang MC, Miller LD, Bergh J, et al. Concordance among gene expression-based predictors for ER-positive breast cancer treated with adjuvant tamoxifen. *Ann Oncol*. (2012) 23:2866–73. doi: 10.1093/annonc/mds080
  24. Prat A, Cheang MC, Galvan P, Nuciforo P, Pare L, Adamo B, et al. Prognostic value of intrinsic subtypes in hormone receptor-positive metastatic breast cancer treated with letrozole with or without lapatinib. *JAMA Oncol*. (2016) 2:1287–94. doi: 10.1001/jamaoncol.2016.0922
  25. Martín M, Rodríguez-Lescure Á, Ruiz A, Alba E, Calvo L, Ruiz-Borrego M, et al. Randomized phase 3 trial of fluorouracil, epirubicin, and cyclophosphamide alone or followed by paclitaxel for early breast cancer. *J Natl Cancer Inst*. (2008) 100:805–14. doi: 10.1093/jnci/djn151
  26. Martín M, Prat A, Rodríguez-Lescure Á, Caballero R, Ebbert MT, Munárriz B, et al. PAM50 proliferation score as a predictor of weekly paclitaxel benefit in breast cancer. *Breast Cancer Res Treat*. (2013) 138:457–66. doi: 10.1007/s10549-013-2416-2
  27. Martín M, González-Rivera M, Morales S, De La Haba-Rodríguez J, González-Cortijo L, Manso L, et al. Prospective study of the impact of the Prosigna assay on adjuvant clinical decision-making in unselected patients with estrogen receptor positive, human epidermal growth factor receptor negative, node negative early-stage breast cancer. *Curr Med Res Opin*. (2015) 31:1129–37. doi: 10.1185/03007995.2015.1037730
  28. Cejalvo JM, Martínez De Duenas E, Galvan P, García-Recio S, Burgues Gasion O, Pare L, et al. Intrinsic subtypes and gene expression profiles in primary and metastatic breast cancer. *Cancer Res*. (2017) 77:2213–21. doi: 10.1158/0008-5472.CAN-16-2717
  29. Fernandez-Martinez A, Pascual T, Perrone G, Morales S, De La Haba J, Gonzalez-Rivera M, et al. Limitations in predicting PAM50 intrinsic subtype and risk of relapse score with Ki67 in estrogen receptor-positive HER2-negative breast cancer. *Oncotarget*. (2017) 8:21930–7. doi: 10.18632/oncotarget.15748
  30. Larburu L, Pare L, Rezola R, Carrera M, Buch E, Gimenez J, et al. Abstract P4-04-05: primary endocrine therapy (PETx) induces PAM50 intrinsic subtype migration with prognostic implications. *Cancer Res*. (2018) 78:P4-04-05. doi: 10.1158/1538-7445.SABCS17-P4-04-05
  31. De Dueñas EM, Hernández AL, Zotano ÁG, Carrión RMP, López-Muñiz JJC, Novoa SA, et al. Prospective evaluation of the conversion rate in the receptor status between primary breast cancer and metastasis: results from the GEICAM 2009-03 ConvertHER study. *Breast Cancer Res Treat*. (2014) 143:507–15. doi: 10.1007/s10549-013-2825-2
  32. Hammond MEH, Hayes DF, Dowsett M, Allred DC, Hagerty KL, Badve S, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer (unabridged version). *Arch Pathol Lab Med*. (2010) 134:e48–72. doi: 10.1043/1543-2165-134.7.e48
  33. Wolff AC, Hammond MEH, Hicks DG, Dowsett M, Mcshane LM, Allison KH, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. *Arch Pathol Lab Med*. (2013) 138:241–56. doi: 10.1200/JCO.2013.50.9984
  34. Dowsett M, Nielsen TO, A'hern R, Bartlett J, Coombes RC, Cuzick J, et al. Assessment of Ki67 in breast cancer: recommendations from the International Ki67 in Breast Cancer working group. *J Natl Cancer Inst*. (2011) 103:1656–64. doi: 10.1093/jnci/djr393
  35. Vidal M, Peg V, Galvan P, Tres A, Cortes J, Ramon Y Cajal S, et al. Gene expression-based classifications of fibroadenomas and phyllodes tumours of the breast. *Mol Oncol*. (2015) 9:1081–90. doi: 10.1016/j.molonc.2015.01.003
  36. Llombart-Cussac A, Cortés J, Paré L, Galván P, Bermejo B, Martínez N, et al. HER2-enriched subtype as a predictor of pathological complete response following trastuzumab and lapatinib without chemotherapy in early-stage HER2-positive breast cancer (PAMELA): an open-label, single-group, multicentre, phase 2 trial. *Lancet Oncol*. (2017) 18:545–54. doi: 10.1016/S1470-2045(17)30021-9
  37. Peters A. Improved predictors (2018). Available online at: <http://www2.uaem.mx/r-mirror/web/packages/ipred/ipred.pdf>
  38. Cheang MC, Voduc KD, Tu D, Jiang S, Leung S, Chia SK, et al. Responsiveness of intrinsic subtypes to adjuvant anthracycline substitution in the NCIC.CTG MA.5 randomized trial. *Clin Cancer Res*. (2012) 18:2402–12. doi: 10.1158/1078-0432.CCR-11-2956

39. Prat A, Cheang MC, Martin M, Parker JS, Carrasco E, Caballero R, et al. Prognostic significance of progesterone receptor-positive tumor cells within immunohistochemically defined luminal A breast cancer. *J Clin Oncol.* (2013) 31:203–9. doi: 10.1200/JCO.2012.43.4134
40. Cheang MC, Martin M, Nielsen TO, Prat A, Voduc D, Rodriguez-Lescure A, et al. Defining breast cancer intrinsic subtypes by quantitative receptor expression. *Oncologist.* (2015) 20:474–82. doi: 10.1634/theoncologist.2014-0372
41. Prat A, Brase J, Cheng Y, Nuciforo P, Paré L, Pascual T, et al. PAM50 intrinsic subtype in hormone receptor-positive (HR+)/human epidermal growth factor receptor 2-negative (HER2-) advanced breast cancer (ABC) treated with exemestane (EXE) in combination with everolimus (EVE) or placebo (PBO): A correlative analysis of the phase III BOLERO-2 trial. *Eur J Cancer.* (2018) 92:S117–8. doi: 10.1016/S0959-8049(18)30574-4
42. Prat A, Fan C, Fernandez A, Hoadley KA, Martinello R, Vidal M, et al. Response and survival of breast cancer intrinsic subtypes following multi-agent neoadjuvant chemotherapy. *BMC Med.* (2015) 13:303. doi: 10.1186/s12916-015-0540-z

**Conflict of Interest Statement:** AP reports consulting and lecture fees from Nanostring Technologies outside the submitted work. ÁR-L reports Clinical Research from Amgen, Astra Zeneca, Boehringer-Ingelheim, GSK, Novartis, Pfizer, Roche/Genentech, Eisai, Celgene, and Pierre Fabre and Advisory Boards and Consulting from Novartis, Pfizer, Roche/Genentech, Eisai, and Celgene, outside the submitted work. GP reports lecture fees from Nanostring Technologies and Clinical Research funds from Astrazeneca, outside the submitted work.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Pascual, Martin, Fernández-Martínez, Paré, Alba, Rodríguez-Lescure, Perrone, Cortés, Morales, Lluch, Urruticoechea, González-Farré, Galván, Jares, Rodriguez, Chic, Righi, Cejalvo, Tonini, Adamo, Vidal, Villagrasa, Muñoz and Prat. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Differentiation Between G1 and G2/G3 Phyllodes Tumors of Breast Using Mammography and Mammographic Texture Analysis

Wen Jing Cui<sup>1†</sup>, Cheng Wang<sup>1,2†</sup>, Ling Jia<sup>3</sup>, Shuai Ren<sup>1</sup>, Shao Feng Duan<sup>4</sup>, Can Cui<sup>1</sup>, Xiao Chen<sup>1\*</sup> and Zhong Qiu Wang<sup>1\*</sup>

## OPEN ACCESS

### Edited by:

Aleix Prat,  
Hospital Clinic de Barcelona, Spain

### Reviewed by:

Yoichi Naito,  
National Cancer Center Hospital  
East, Japan  
Masahiko Tanabe,  
The University of Tokyo, Japan

### \*Correspondence:

Xiao Chen  
chxwin@163.com  
Zhong Qiu Wang  
zhq2001us@163.com

<sup>†</sup>These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
Women's Cancer,  
a section of the journal  
Frontiers in Oncology

**Received:** 11 March 2019

**Accepted:** 07 May 2019

**Published:** 29 May 2019

### Citation:

Cui WJ, Wang C, Jia L, Ren S,  
Duan SF, Cui C, Chen X and Wang ZQ  
(2019) Differentiation Between G1 and  
G2/G3 Phyllodes Tumors of Breast  
Using Mammography and  
Mammographic Texture Analysis.  
Front. Oncol. 9:433.  
doi: 10.3389/fonc.2019.00433

<sup>1</sup> Department of Radiology, Affiliated Hospital of Nanjing University of Chinese Medicine, Nanjing, China, <sup>2</sup> Department of Graduate, Bengbu Medical College, Bengbu, China, <sup>3</sup> Sir Run Run Hospital, Nanjing Medical University, Nanjing, China, <sup>4</sup> GE Healthcare China, Shanghai, China

**Purpose:** To determine the potential of mammography (MG) and mammographic texture analysis in differentiation between Grade 1 (G1) and Grade 2/ Grade 3 (G2/G3) phyllodes tumors (PTs) of breast.

**Materials and methods:** A total of 80 female patients with histologically proven PTs were included in this study. 45 subjects who underwent pretreatment MG from 2010 to 2017 were retrospectively analyzed, including 14 PTs G1 and 31 PTs G2/G3. Tumor size, shape, margin, density, homogeneity, presence of fat, or calcifications, a halo-sign as well as some indirect manifestations were evaluated. Texture analysis features were performed using commercial software. Receiver operating characteristic curve (ROC) was used to determine the sensitivity and specificity of prediction.

**Results:** G2/G3 PTs showed a larger size ( $>4.0$  cm) compared to PTs G1 (64.52 vs. 28.57%,  $p = 0.025$ ). A strong lobulation or multinodular confluent was more common in G2/G3 PTs compared to PTs G1 (64.52 vs. 14.29%,  $p = 0.004$ ). Significant differences were also observed in tumors' growth speed and clinical manifestations ( $p = 0.007$ , 0.022, respectively). Ten texture features showed significant differences between the two groups ( $p < 0.05$ ), Correlation\_AllDirection\_offset7\_SD and ClusterProminence\_AllDirection\_offset7\_SD were independent risk factors. The area under the curve (AUC) of imaging-based diagnosis, texture analysis-based diagnosis and the combination of the two approaches were 0.805, 0.730, and 0.843 (90.3% sensitivity and 85.7% specificity).

**Conclusions:** Texture analysis has great potential to improve the diagnostic efficacy of MG in differentiating PTs G1 from PTs G2/G3.

**Keywords:** phyllodes tumors, classification, mammography, artificial intelligence, machine learning

## INTRODUCTION

Phyllodes tumors (PTs) are rare breast fibroepithelial neoplasms that account for <1% (1, 2) of all breast tumors and 2–3% of all fibroepithelial breast lesions (3, 4). PTs were originally described in 1838 as “cystosarcoma phyllodes” because of their leaf-like pattern of growth and internal cystic degeneration. PTs usually showed benign biological manifestations. However, approximately 20–30% of resected PTs are malignant and approximately 25% of malignant ones show metastatic features (5). A prominent and widely accepted grading system has been reported by the World Health Organization (WHO) 3-tiered classification. PTs are classified as benign, borderline, and malignant based on the semi-quantitative evaluation of key histologic findings, which include stromal cellularity, stromal atypia, stromal mitosis, and stromal overgrowth (6).

PTs may occur in any age group from adolescents to the elderly but most commonly in women aged between 35 and 55 years (1, 4). Surgical resection is the fundamental treatment for PTs. However, surgical approaches are generally selected based on the histologic grade. Wide excision or mastectomy is usually performed in PTs Grade 2 (G2)/G3 (7–9). Therefore, the preoperative differentiation between PTs G1 and G2/G3 would be especially useful for surgery planning. Fine-needle biopsy is considered to be a highly accurate technique in PTs diagnosis. However, it is not proper to be used for PTs grading because of inadequate cytologic samples and the heterogeneous nature of the tissue composition in PTs (10, 11).

Various radiologic methods, including mammography (MG), ultrasound (US), and magnetic resonance imaging (MRI) have been used to preoperatively grade PTs (12). The MG and US showed limited potential in predicating PTs grades. MRI may be a useful imaging approach. However, some patients cannot undergo MRI examination because of biomedical metal stents or contraceptive ring implantations, which is very common among Chinese women. In addition, MRI examination is expensive and time consuming. Therefore, surgeons prefer direct operation after receiving US and MG examinations. It would be valuable to find a way to improve the diagnostic performance of MG or US.

Recently, artificial intelligent (AI) technology and radiomics, computer-aided texture analysis has been used for diagnosis, treatment response and prognosis evaluation in cancer patients. However, few studies have used the method of mammography combined with mammographic texture analysis to grade the PTs up to now. The purpose of this study was to determine the diagnostic performance of mammography and mammographic texture analysis in the differentiation between G1 and G2/G3 PTs.

**Abbreviations:** G1, Grade 1; G2/G3, Grade 2/Grade 3; PTs, Phyllodes tumors; MG, mammography; ROC, Receiver operating characteristic; US, ultrasound; AUC, Area under the curve; MRI, magnetic resonance imaging; CC, craniocaudal; MLO, mediolateral oblique; PACS, Picture Archiving and Communication Systems; ACR, American College of Radiology; ROIs, Region of interests; GLCM, Gray Level Co-occurrence Matrices; RLM, run-length matrix; AI, artificial intelligent; ADC, apparent diffusion coefficient.

## MATERIALS AND METHODS

The Declaration of Helsinki was adhered to throughout the entire study. The protocol was approved by the Institutional Review Board of the Affiliated Hospital of Nanjing University of Chinese Medicine. The need for informed consent was waived by the Institutional Review Board, due to the nature of this retrospective study.

### Patients

From February 2010 to October 2017, we obtained data from 80 female patients with surgically proven primary PTs, from our data warehouse. The patients' ages ranged from 25 to 70 years old (mean  $46.58 \pm 9.54$ ). The inclusion criteria were as follows: (1) patients with surgically proven primary PTs; (2) patients who did not undergo any treatment before surgery; (3) patients who underwent preoperative mammography; (4) with a visible lesion on the mammography images. Finally, 35 cases were excluded due to the absence of MG examination ( $n = 30$ ) or negative MG findings ( $n = 5$ ). A total of 45 patients were included in this study (Figure 1). According to the WHO 2012 classification for PTs, the PTs were divided into G1, G2, and G3 in this study. We obtained information about the tumors growth speed by tracking the patient's previous images (including mammography, ultrasound, and MRI) or by asking about the patients' feelings. A tumors diameter doubling within half a year is defined as a rapid growth tumor, while the remaining is defined as a slow growth tumor. Tactility was defined as hard like the forehead, medium like the nose and soft like the lips.

### Mammography Examinations and Images Analysis

Bilateral digital MG examinations were performed using the GIOTTOIMAGE 3D (IMS, Bologna, ITA), and choosing fully automatic exposure control mode, including the routine craniocaudal (CC), and mediolateral oblique (MLO) views. The DICOM images were obtained from the Picture Archiving and Communication Systems (PACS). Two radiologists (>8 years' experience in mammography), who were blinded to pathological findings, analyzed the images. The following imaging information was evaluated: tumor size, margin (well-defined or ill-defined border), shape (oval, weak lobulation, and strong lobulation /multinodular confluent), density (hypodensity, isodensity, or hyperdensity), homogeneity (yes or no), the presence of fat or calcifications, and the presence of a halo-sign (a low density fat ring caused by the tumor pushing against surrounding structures). In addition, some indirect manifestations, including breast composition categories of American College of Radiology (ACR), skin thickening, venectasia, and axillary lymphadenectasis (the short diameter >1 cm) were also evaluated. The size of the tumor was determined based on the maximum diameter either in a CC or MLO image. For quantitative data, we calculated the mean of two readers. For qualitative data, the final imaging features were confirmed when the two readers reached a consensus.

## Mammographic Texture Analysis

Region of interests (ROIs) were drawn manually to delineate the lesions using ITK-SNAP software. Since PTs have envelopes and the display rate of a halo-ring is as high as 91.11%(41/45) in this study, we outline ROIs of tumors with a halo-ring as the boundary. All the dicom images and ROIs were individually transferred to the texture analysis software package (Artificial Intelligent Kit-A.K., GE Healthcare). Subsequently, texture features were automatically calculated by the A.K. software package. The texture analysis was performed twice for each lesion, and mean values of texture features were calculated. The procedure is shown in **Figure 2**. Three categories of statistical methods including Histogram, Gray Level Cooccurrence Matrices (GLCM), and run-length matrix (RLM) were used. A total of 435 texture features were extracted from each image in our study.

## Statistical Analysis

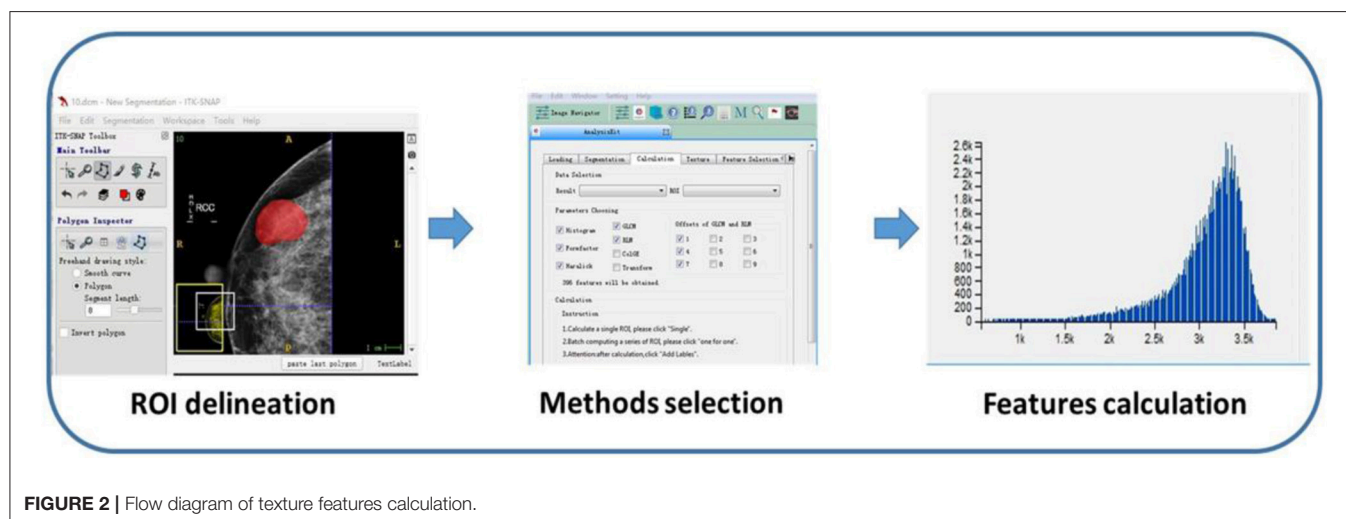
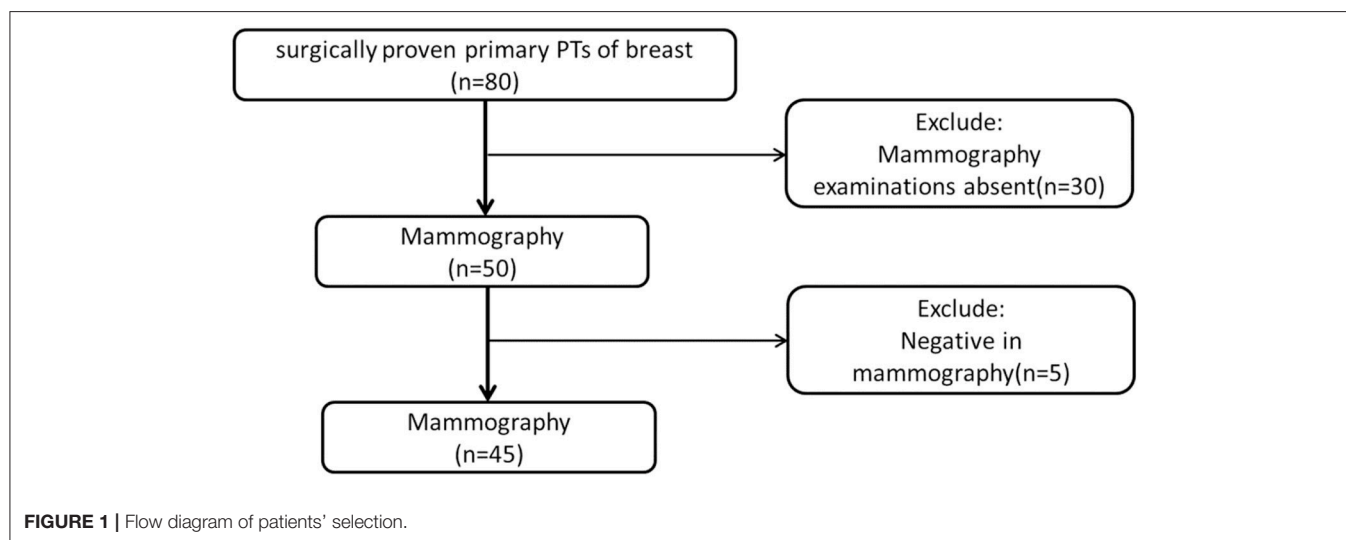
Statistical analyses were performed using IBM SPSS version 22.0 (IBM Corporation, New York). Quantitative data were

displayed as mean  $\pm$  SD. The Independent sample *t*-test and Mann-Whitney *U*-test was used for data with a normal and abnormal distribution, respectively. Categorical data were shown as a percentage and were analyzed using the Chi-square test or Fisher's exact test. Spearman correlation analysis and Logistic regression was used to show the relationship between texture features and tumor grade.  $P < 0.05$  were considered statistically significant. The Receiver operating characteristic (ROC) curve was adopted to determine the diagnostic sensitivity and specificity of the Mammography and Mammographic texture analysis.

## RESULTS

### Patients' Clinical Characteristics

The clinical characteristics of the 80 patients are summarized in **Table 1**. Each patient has only one lesion in the unilateral breast. All patients underwent surgery. There were 21 benign (26.25%), 38 borderline (47.50%), and 21 malignant tumors (26.25%). Fifteen of them underwent local excision, 52 underwent wide



**TABLE 1** | Clinical data of patients.

Clinical data	Total <i>n</i> = 80	PTs G1 <i>n</i> = 21	PTs G2/G3 <i>n</i> = 59	<i>p</i> -value
<b>Age (years)</b>	46.58 ± 9.54	45.33 ± 7.91	47.02 ± 10.16	
<b>Growth speed</b>				0.007
Slowly increased	49	18	31	
Rapidly increased	31	3	28	
<b>Clinical manifestations</b>				0.022
Mass	48	17	31	
Mass with pain	19	4	15	
Mass with pain and skin change	13	0	13	
<b>Stiffness</b>				0.353
Hard	45	10	35	
Medium	35	11	24	
Soft	0	0	0	
<b>Mobility</b>				0.178
Well	39	11	28	
Not good enough	27	9	18	
Poor	14	1	13	
<b>Location</b>				0.088
Upper inner quadrant	11	1	10	
Lower inner quadrant	5	1	4	
Upper outer quadrant	37	11	26	
Lower outer quadrant	8	5	3	

excision and 13 underwent mastectomy. Many PTs G2/G3 rapidly increased (diameter doubling) within half a year compared with PTs G1 (47.46 vs. 14.28%,  $p = 0.007$ ). PTs G2/G3 were more likely to cause pain and skin changes compared to PTs G1 ( $p = 0.022$ ). No significant differences were found in stiffness and mobility. Except for 19 lesions growing in the center or occupying the entire breast it was difficult to judge the origins, the location had no significance between these two groups.

## Mammography Findings

Subsequently, we evaluated the Mammographic findings of the 45 patients who met the study criteria. Their mean age was  $48.2 \pm 8.96$ . There were 14 benign (31.1%), 20 borderline (44.4%), and 11 malignant tumors (24.4%). Mammography findings of PTs are summarized in **Table 2**. Significant differences were found in tumor size, shape between G1 and G2/G3 PTs ( $p < 0.05$ ). Larger size ( $d > 4.0$  cm) were more common in G2/G3 PTs compared with PTs G1 (64.52 vs. 28.57%,  $p = 0.025$ ) (**Figure 3**). PTs G2/G3 showed strong lobulation or multinodular confluence compared to the PTs G1 [20/31 (64.52) vs. 2/14 (14.29%),  $p = 0.004$ ]. The lesions with strong lobulation or multinodular confluence showed a “multi-boundary sign” in MG because of the overlapped effect (**Figure 4**). Some low-grade PTs showed an ill-defined margin which was under the influence of the cover effect because of their small size and equal density to the surrounding gland (**Figure 5**). There were some limitations in the evaluation of PTs boundaries. There were no significant differences in density, homogeneity, the presence or absence of a halo ring,

**TABLE 2** | The mammography findings in phyllodes tumors (PTs) G1 and G2/G3.

Mammographic findings	Total <i>n</i> = 45	PTs G1 <i>n</i> = 14	PTs G2/G3 <i>n</i> = 31	<i>p</i> -value
<b>Size (cm)</b>	5.54 ± 3.67	4.11 ± 2.55	6.19 ± 3.93	0.077
≤ 4	21	10	11	0.025
> 4	24	4	20	
<b>Shape</b>				0.004
Oval	10	6	4	
Weak lobulation	13	6	7	
Strong lobulation or multinodular confluent	22	2	20	
<b>Mass margin</b>				0.147
Well- defined	33	8	25	
Ill- defined	12	6	6	
<b>Density</b>				1.000
Hypodensity	1	0	1	
Isodensity	25	8	17	
Hyperdensity	19	6	13	
<b>Homogeneity</b>				0.725
Yes	33	11	22	
No	12	3	9	
<b>Halo sign</b>				0.082
Presence	41	11	30	
Absence	4	3	1	
<b>Calcifications</b>				0.578
Presence	4	2	2	
Absence	41	12	29	
<b>Fat</b>				1.000
Presence	1	0	1	
Absence	44	14	30	

calcifications and fat between PTs G1 and PTs G2/G3. Similar results were observed for the indirect manifestations (**Table 3**).

ROC curve was adopted to determine the diagnostic sensitivity and specificity of Mammography findings. AUC was 0.805 with 64.5% of sensitivity and 85.7% of specificity (**Figure 6A**).

## Mammographic Texture Analysis

Total of 435 texture features were extracted from the mammographic images. Those texture features with significant differences between PTs G1 and PTs G2/G3 are shown in **Table 4**. Spearman correlation analysis also eliminated some parameters with strong a correlation (**Figure 7**). Finally, logistic regression showed that only two parameters were retained in our model. They were Correlation\_AllDirection\_offset7\_SD and ClusterProminence\_AllDirection\_offset7\_SD.

### Parameter 1:Correlation\_AllDirection\_offset7\_SD

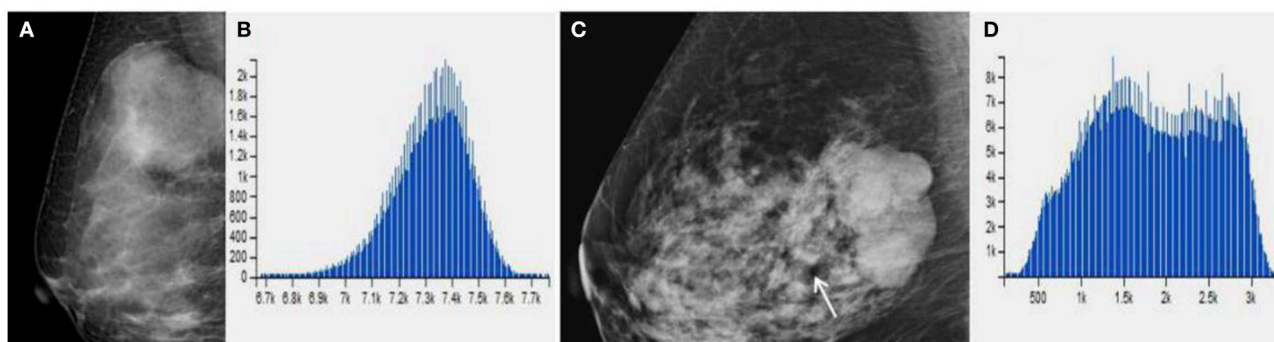
Correlation measures the similarity of the gray levels in neighboring pixels. Correlation\_AllDirection\_offset7\_SD is one of the 18 parameters related to the Correlation in AK Software.

$$\text{Formula: } - \sum_{i,j} \frac{(i - \mu)(j - \mu)g(i,j)}{\sigma^2}$$

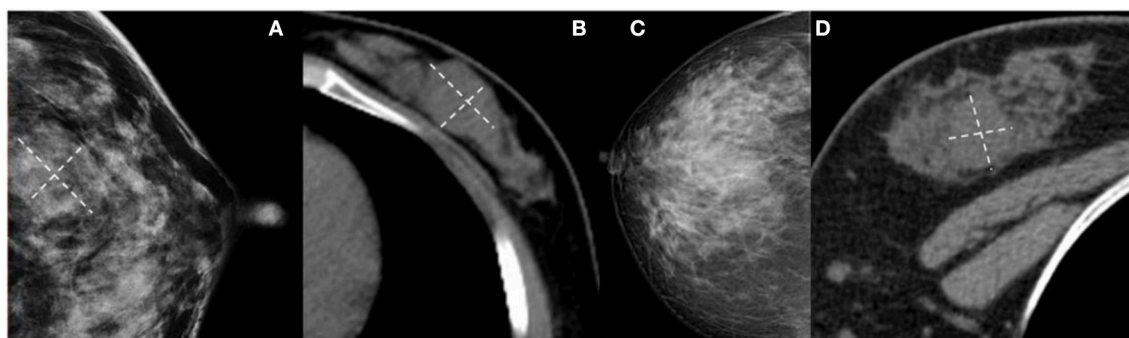




**FIGURE 3 |** (a) Malignant Phyllodes tumor (PT) of left breast in a 55-year-old woman. A mediolateral oblique mammogram shows a well-defined isodensity mass with a diameter of 13 cm. (b) Malignant PT of right breast in a 47-year-old woman. Mammogram shows a well-defined high-density mass with a diameter of 9 cm. The mass is partially surrounded by a lucent halo (arrows). (c,d) Malignant PT of left breast in a 38-year-old woman. CT can show cystic changes within the tumor. They are all well-defined masses with large size.



**FIGURE 4 |** (A,B) Benign Phyllodes tumor (PT) of right breast in 49-year-old woman, mammogram shows an ovoid mass with a diameter of 4.5 cm. (C) Borderline PT with a diameter of 4.5 cm in 63-year-old woman. Mammogram shows a mass formed by multiple nodules. (B,D) The histogram of the texture parameters of the two lesions also show a marked difference.



**FIGURE 5 |** (A,B) Benign Phyllodes tumor (PT) of left breast in 55-year-old woman. Mammogram shows an ill-defined isodensity mass. However, CT can show the boundary clearly. (C,D) Benign PT of right breast in 34-year-old woman. The lesion is not visible on mammogram, but clearly visible on CT. They are all affected by the cover effect of mammography.

## Parameter

### 2:ClusterProminence\_AllDirection\_offset7\_SD

Cluster Prominence is a measure of a symmetry of a given distribution. High values of this feature indicate that the symmetry of the image is low, in medical imaging low values of cluster prominence represent a smaller peak for the image gray

level value and usually the gray level difference between the forms is small.

$$\text{Formula: } \sum_{i,j} ((i - \mu) + (i + \mu)) 4g(i,j)$$



The texture features were associated with tumor grade (OR = 0.465, 95%CI:0.231–0.936; OR = 0.042, 95%CI:0.193–0.969, respectively).

**TABLE 3 |** The indirect manifestations on mammography in Phyllodes tumors (PTs) G1 and G2/G3.

Indirect mammography findings	Total <i>n</i> = 45	PTs G1 <i>n</i> = 14	PTs G2/G3 <i>n</i> = 31	<i>p</i> -value
<b>Breast composition categories of ACR</b>				0.889
a	2	1	1	
b	6	1	5	
c	30	10	20	
d	7	2	5	
<b>Skin thickening</b>				0.156
Presence	6	0	6	
Absence	39	14	25	
<b>Venectasia</b>				0.469
Presence	10	2	8	
Absence	35	12	23	
<b>Axillary lymphadenectasis</b>				0.530
Presence	2	1	1	
Absence	43	13	30	

ROC curve was adopted to determine the diagnostic sensitivity and specificity of Mammographic texture analysis. The AUC was 0.730. When the cut off value was 0.044, the sensitivity was 93.5%, and the specificity was 50% (**Figure 6B**).

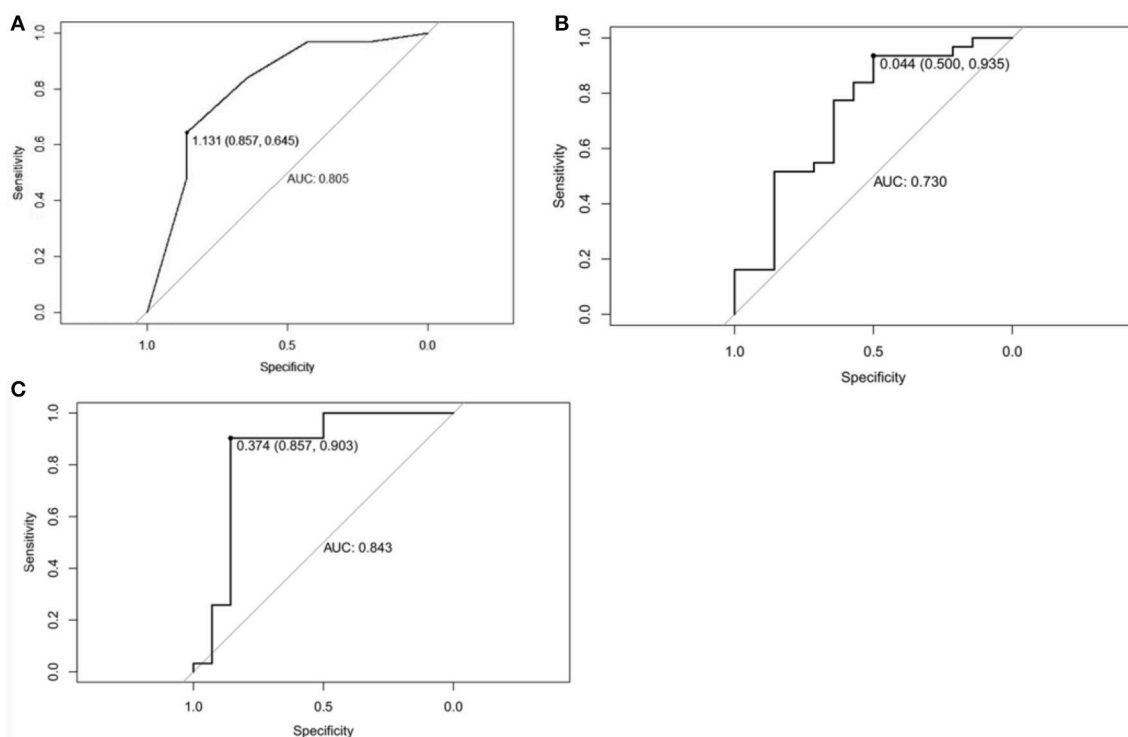
Subsequently, ROC curve was also adopted to determine the diagnostic sensitivity and specificity of Mammography findings + texture features. The AUC was 0.843 with 90.3% sensitivity and 85.7% specificity for predicting PTs G2/G3 tumors (**Figure 6C**).

Finally, we randomly selected 30 samples for internal validation, including nine benign (30%), 13 borderline (43.33%), and eight malignant tumors (26.67%). The AUC was 0.862 (85.7% sensitivity and 77.8% specificity). The verification results are similar to those of previous studies, which prove that the model is relatively stable.

## DISCUSSION

Previous studies have indicated that imaging approaches are useful in differentiating PTs G1 from PTs G2/G3. In the present study, we evaluated the role of texture features in grading PTs. Our data indicated that texture features are useful in grading PTs. Moreover, our data indicates that texture analysis can improve the diagnostic performance in differentiating PTs G1 and PTs G2/G3.

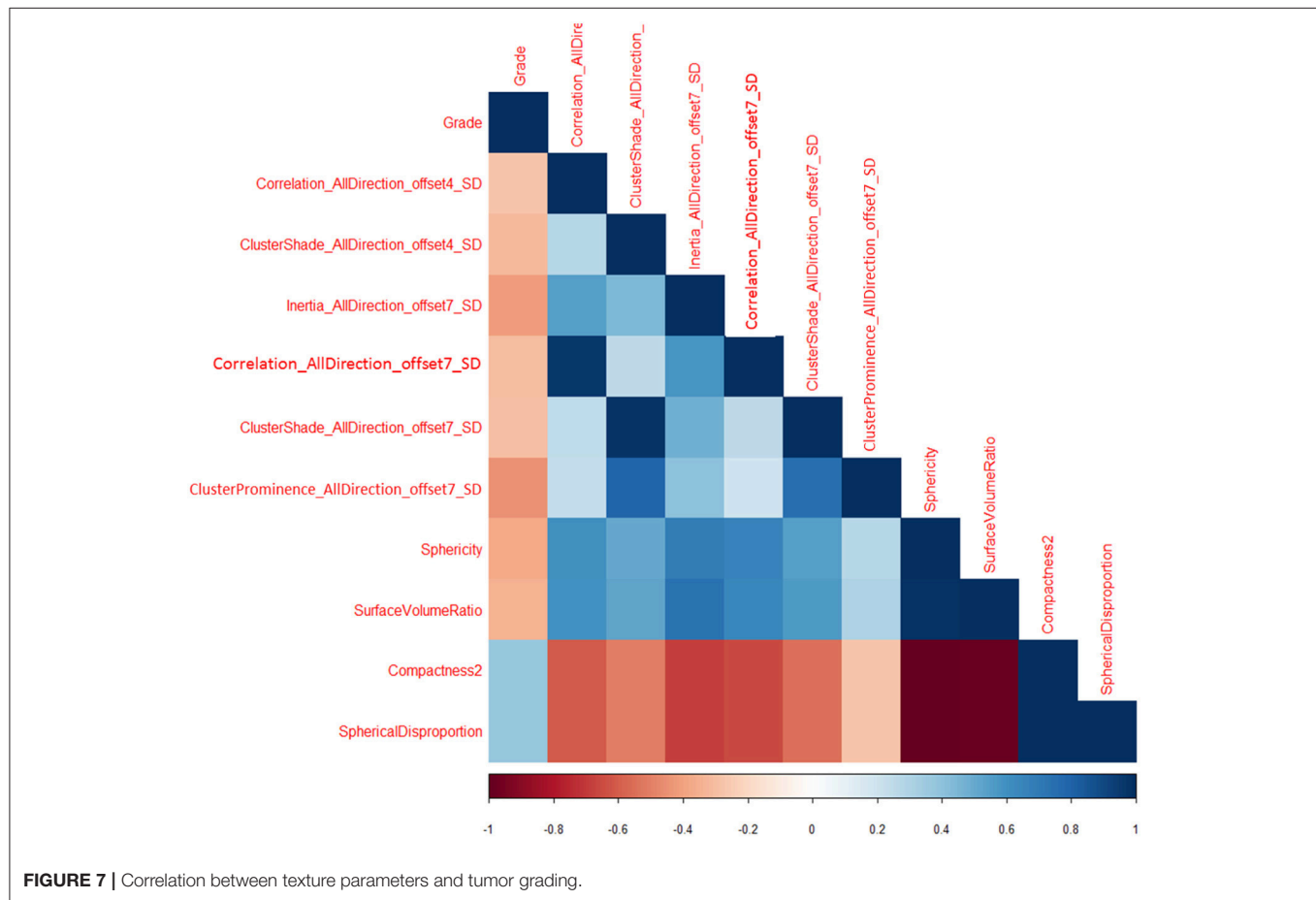
Surgical methods are associated with the grades of PTs. The preoperative differentiation would be especially useful for surgery planning. A fine-needle biopsy is an accurate method used



**FIGURE 6 |** (A) Receiver operating characteristic curve of Mammographic findings in predicting Phyllodes tumor (PT) G2/G3 tumors. The area under curve was 0.805. (B) Receiver operating characteristic curve of texture features in predicting Phyllodes tumor (PT) G2/G3 tumors. The area under curve was 0.730. (C) Receiver operating characteristic curve of Mammographic findings + texture features in predicting PTs G2/G3 tumors. The area under curve was 0.843.

**TABLE 4 |** Texture parameters in Phyllodes tumors (PTs) G1 and G2/G3.

Mammographic texture analysis	Total <i>n</i> = 45	PTs G1 <i>n</i> = 14	PTs G2/G3 <i>n</i> = 31	<i>p</i> -value
Sphericity	0.21 ± 0.06	0.24 ± 0.06	0.20 ± 0.06	0.021
Surface volume ratio	2146.45 ± 73.01	2174.47 ± 74.15	2133.79 ± 68.84	0.044
Compactness2	134.35 ± 74.93	105.36 ± 51.65	147.44 ± 79.96	0.042
Spherical disproportion	5.27 ± 1.87	4.51 ± 1.39	5.61 ± 1.95	0.035
Correlation_AllDirection_offset4_SD ( $\times 10^{-6}$ )	6.56 ± 5.45	8.63 ± 6.03	5.62 ± 4.99	0.043
Correlation_AllDirection_offset7_SD ( $\times 10^{-5}$ )	1.20 ± 1.08	1.77 ± 1.34	9.49 ± 8.43	0.008
ClusterShade_AllDirection_offset4_SD ( $\times 10^3$ )	8.86 ± 7.83	11.86 ± 10.25	7.50 ± 5.90	0.039
ClusterShade_AllDirection_offset7_SD ( $\times 10^3$ )	13.77 ± 11.84	18.93 ± 16.21	11.44 ± 8.58	0.024
ClusterProminence_AllDirection_offset7_SD ( $\times 10^6$ )	5.36 ± 4.31	7.03 ± 6.24	4.598217.58 ± 2.90	0.040
Inertia_AllDirection_offset7_SD	58.16 ± 38.71	72.87 ± 49.94	51.51 ± 31.16	0.043



in the diagnosis of PTs but cannot be used for classification, because of inadequate cytologic samples and the heterogeneous nature of the tissue composition(10, 11). It would be helpful to evaluate the PTs grades by using imaging approaches. However, the radiologic studies in PTs grading are very few because of the low incidence. Previous US, MG, and MRI studies indicated that a larger tumor size and irregular tumor shape are more common in higher grades of tumors than in lower grade tumors (9–15). Our data is consistent with those previous

findings, and we found that the multinodular confluent was characteristic imaging manifestation of PTs G2/G3. This is related to the degree of leaf-like growth in histology (2). An irregular cyst wall in an MRI, a tumor signal intensity lower than or equal to normal tissue on T2-weighted images and a low apparent diffusion coefficient (ADC) are all significantly correlated with the histologic grade. T1 weighted imaging signal in the G2/G3 PTs was higher than that in the PTs G1 (12).

Recently, texture analysis has been widely used to evaluate tumor heterogeneity. Texture parameters, such as entropy and kurtosis, show good performance in differentiating benign from malignant tumors (16, 17). Several studies also indicate that texture features are good predictors of tumor grades (18, 19). However, few studies have shown the role of texture features in PTs grading. We were the first one to use the method of mammographic texture analysis to grade the PTs up to now. Significant differences were found in 10 texture features and *Correlation\_AllDirection\_offset7\_SD* and *ClusterProminence\_AllDirection\_offset7\_SD* were the independent factors in identifying PTs G1 from PTs G2/G3. In addition, our data also indicated that Mammography can obtain good specificity but poor sensitivity, while texture analysis can obtain high sensitivity but poor specificity in differentiation. Interestingly, the combination of the two approaches can obtain both high sensitivity and specificity. Texture analysis can effectively improve the efficacy of mammography for PTs classification.

There are also several limitations in our study. First, since a mammography is a two-dimensional structural image, the recognition of functional, and three-dimensional structural images is absent, and texture analysis based on mammography may lose a lot of information. Second, as a retrospective study, selection bias cannot be avoided. Third, it is inevitable that the number of patients in this study is small for texture analysis study. There are two main reasons for the small number of cases: (1) The incidence of PTs is low, which only accounts for 1% of breast tumors. It is relatively difficult to collect cases for this. Second, texture analysis research requires a high consistency of Imaging equipment and parameters, in order to ensure the accuracy of texture analysis, some cases have to be excluded from the study. Because of the relatively small sample size, all cases were included for texture feature extraction. Then we performed internal validation to verify the results, aiming to improve the accuracy of the test set as much as possible under existing conditions. Finally, we compared the texture analysis results obtained in this study, with previous literature, and found that the two independent parameters we screened had been reported to have clear statistical significance in the benign and malignant differentiation of breast calcifications and evaluation of chemotherapy efficacy (20, 21), which further supported the credibility of the results of this study. In the future, we will expand the sample size to further improve the accuracy and repeatability of the study.

In conclusion, our data indicates that texture analysis based on Mammography has the potential to differentiate PTs G2/G3

from PTs G1. Combining Mammography and texture features can provide optimal predictions in the classification of PTs in mammography.

## DATA AVAILABILITY

All datasets generated for this study are included in the manuscript and/or the supplementary files.

## CONTRIBUTION TO THE FIELD STATEMENT

Phyllodes tumors (PTs) are rare breast fibroepithelial neoplasms that account for <1% of all breast tumors. PTs are classified as benign, borderline, and malignant. The preoperative differentiation between PTs G1 and G2/G3 would be especially useful for surgery planning. Wide excision or mastectomy is usually performed in PTs Grade2 (G2)/G3. A fine-needle biopsy should not be used for PTs grading because of the inadequate cytologic samples and the heterogeneous nature of the tissue composition. MRI may be a useful imaging approach but has so much contraindication as well as being costly. The MG and US showed limited potential in predicating PTs grades, however, our study first used the method of mammography combined with mammographic texture analysis to grade the PTs. The area under the curve (AUC) of imaging-based diagnosis, texture-based diagnosis and the combination of the two approaches were 0.805(64.5% sensitivity and 85.7% specificity), 0.730 (93.5% sensitivity and 50% specificity), and 0.843 (90.3% sensitivity and 85.7% specificity). Texture analysis can effectively improve the efficacy of mammography for PTs classification.

## AUTHOR CONTRIBUTIONS

WC and ZW: guarantor of integrity of the entire study. ZW and XC: study concepts and design. CW, WC, and CC: clinical studies. CW, SR, and SD: statistical analysis. WC, XC and LJ: manuscript editing.

## FUNDING

This study has received funding by the Key Research & Development Plan of Jiangsu Province (BE2017772). Talent peaks project of Jiangsu Provincial Hospital of TCM.

## ACKNOWLEDGMENTS

Thanks to the support and help of teachers and colleagues.

## REFERENCES

1. Lin CC, Chang HW, Lin CY, Chiu CF, Yeh SP. The clinical features and prognosis of phyllodes tumors: a single institution experience in Taiwan. *Int J Clin Oncol*. (2013) 18:614–20. doi: 10.1007/s10147-012-0442-4
2. Chng TW, Lee JY, Lee CS, Li H, Tan MH, Tan PH. Validation of the Singapore nomogram for outcome prediction in breast phyllodes tumours: an Australian cohort. *J Clin Pathol*. (2016) 169:1124–26. doi: 10.1136/jclinpath-2016-203951
3. Venter AC, Rosca E, Daina LG, Mutiu G, Pirte AN, D.Rahota. Phyllodes tumor: diagnostic imaging and histopathology findings. *Rom J Morphol Embryo*. (2015) 156:1397–402. doi: 10.3389/fonc.2019.00433

4. Plaza MJ, Swintelski C, Yaziji H, Torres-Salichs M, Esserman LE. Phyllodes tumor: review of key imaging characteristics. *Breast Dis.* (2015) 35:79–86. doi: 10.3233/BD-150399
5. Duman L, Gezer NS, Balci P, Altay C, Basara I, Durak MG, et al. Differentiation between phyllodes tumors and fibroadenomas based on mammographic sonographic and MRI features. *Breast Care.* (2016) 11:123–7. doi: 10.1159/000444377
6. Kim S, Kim JY, Kim DH, Jung WH, Koo JS. Analysis of phyllodes tumor recurrence according to the histologic grade. *Breast Cancer Res Treat.* (2013) 141:353–63. doi: 10.1007/s10549-013-2684-x
7. Onkendi EO, Jimenez RE, Spears GM, Harmsen WS, Ballman KV, Hieken TJ. Surgical treatment of borderline and malignant phyllodes tumors: the effect of the extent of resection and tumor characteristics on patient outcome. *Ann Surg Oncol.* (2014) 21:3304–9. doi: 10.1245/s10434-014-3909-x
8. Mitus J, Reinfuss M, Mitus JW, Jakubowicz J, Blecharz P, Wysocki WM, et al. Malignant phyllodes tumor of the breast: treatment and prognosis. *Breast J.* (2014) 20:639–44. doi: 10.1111/tbj.12333
9. Kawashima H, Miyati T, Ohno N, Ohno M, Inokuchi M, Ikeda H, et al. Differentiation between phyllodes tumours and fibroadenomas using intravoxel incoherent motion magnetic resonance imaging: comparison with conventional diffusion-weighted imaging. *Br J Radiol.* (2018) 91:20170687. doi: 10.1259/bjr.20170687
10. Franceschini G, Masetti R, Brescia A, Mule A, Belli P, Costantini M, et al. Phyllodes tumor of the breast: magnetic resonance imaging findings and surgical treatment. *Breast J.* (2005) 11:144–5. doi: 10.1111/j.1075-122X.2005.21586.x
11. Balaji R, Ramachandran KN. Magnetic resonance imaging of a benign phyllodes tumor of the breast. *Breast Care.* (2009):189–91. doi: 10.1159/000220604
12. Yabuuchi H, Soeda H, Matsuo Y, Okafuji T, Eguchi T, Sakai S, et al. Phyllodes tumor of the breast: correlation between MR findings and histologic grade. *Radiology.* (2006) 241:702–9. doi: 10.1148/radiol.2413051470
13. Farria DM, Gorczyca DP, Barsky SH, Barsky SH, Sinha S, Bassett LW, et al. Benign phyllodes tumor of the breast: MR imaging features. *AJR Am J Roentgenol.* (1996) 167:187–9. doi: 10.2214/ajr.167.1.8659370
14. Alhabshi SM, Rahmat K, Abu Hassan H, Westerhout CJ, Chandran PA. Advanced MRI applications and findings of malignant phyllodes tumour: review of two cases. *Jpn J Radiol.* (2013) 31:342–8. doi: 10.1007/s11604-013-0183-y
15. McCarthy E, Kavanagh J, O'Donoghue Y, McCormack E, D'Arcy C, O'Keeffe SA. Phyllodes tumours of the breast: radiological presentation, management and follow-up. *Br J Radiol.* (2014) 87:20140239. doi: 10.1259/bjr.20140239
16. Li Z, Yu L, Wang X, Yu H, Gao Y, Ren Y, et al. Diagnostic performance of mammographic texture analysis in the differential diagnosis of benign and malignant breast tumors. *Clin Breast Cancer.* (2018) 18:621–7. doi: 10.1016/j.clbc.2017.11.004
17. Kim JH, Ko ES, Lim Y, Lee KS, Han BK, Ko EY, et al. Breast cancer heterogeneity: mr imaging texture analysis and survival outcomes. *Radiology.* (2017) 282:665–75. doi: 10.1148/radiol.2016160261
18. Schieda N, Lim RS, Krishna S, McInnes MDF, Flood TAE, Thornhill R, et al. Diagnostic accuracy of unenhanced ct analysis to differentiate low-grade from high-grade chromophobe renal cell carcinoma. *AJR Am J Roentgenol.* (2018) 210:1079–87. doi: 10.2214/AJR.17.18874
19. Guo C, Zhuge X, Wang Z, Wang Q, Sun K, Feng Z, et al. Textural analysis on contrast-enhanced CT in pancreatic neuroendocrine neoplasms: association with WHO grade. *Abdom Radiol.* (2018) 44:576–85. doi: 10.1007/s00261-018-1763-1
20. Pereira RR Jr., Azevedo Marques PM, Honda MO, Kinoshita SK, Engelmann R, Muramatsu C, et al. Usefulness of texture analysis for computerized classification of breast lesions on mammograms. *J Dig Imaging.* (2007) 20:248–55. doi: 10.1007/s10278-006-9945-8
21. Ahmed A, Gibbs P, Pickles M, Turnbull L. Texture analysis in assessment and prediction of chemotherapy response in breast cancer. *J Magn Reson Imaging.* (2013) 38:89–101. doi: 10.1002/jmri.23971

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Cui, Wang, Jia, Ren, Duan, Cui, Chen and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Androgen Receptor Expression and Association With Distant Disease-Free Survival in Triple Negative Breast Cancer: Analysis of 263 Patients Treated With Standard Therapy for Stage I-III Disease

## OPEN ACCESS

### Edited by:

Mothaffar Rimawi,  
Baylor College of Medicine,  
United States

### Reviewed by:

Rachelle Johnson,  
Vanderbilt University Medical Center,  
United States

Alessandro Igor Cavalcanti Leal,  
Johns Hopkins Medicine,  
United States

### \*Correspondence:

Maria Vittoria Dieci  
mariavittoria.dieci@unipd.it

† These authors have contributed  
equally to this work and share first  
authorship

### Specialty section:

This article was submitted to  
Women's Cancer,  
a section of the journal  
Frontiers in Oncology

Received: 24 January 2019

Accepted: 13 May 2019

Published: 06 June 2019

### Citation:

Dieci MV, Tsvetkova V, Griguolo G,  
Miglietta F, Mantiero M, Tasca G,  
Cumerlato E, Giorgi CA, Giarratano T,  
Faggioni G, Falci C, Vernaci G,  
Menichetti A, Mioranza E, Di Liso E,  
Frezzini S, Saibene T, Orvieto E and  
Guarneri V (2019) Androgen Receptor  
Expression and Association With  
Distant Disease-Free Survival in Triple  
Negative Breast Cancer: Analysis of  
263 Patients Treated With Standard  
Therapy for Stage I-III Disease.  
Front. Oncol. 9:452.  
doi: 10.3389/fonc.2019.00452

Maria Vittoria Dieci<sup>1,2\*†</sup>, Vassilena Tsvetkova<sup>1,3†</sup>, Gaia Griguolo<sup>2</sup>, Federica Miglietta<sup>1</sup>, Mara Mantiero<sup>1</sup>, Giulia Tasca<sup>1,2</sup>, Enrico Cumerlato<sup>1</sup>, Carlo Alberto Giorgi<sup>2</sup>, Tommaso Giarratano<sup>2</sup>, Giovanni Faggioni<sup>2</sup>, Cristina Falci<sup>2</sup>, Grazia Vernaci<sup>1</sup>, Alice Menichetti<sup>1</sup>, Eleonora Mioranza<sup>2</sup>, Elisabetta Di Liso<sup>2</sup>, Simona Frezzini<sup>1</sup>, Tania Saibene<sup>4</sup>, Enrico Orvieto<sup>5</sup> and Valentina Guarneri<sup>1,2</sup>

<sup>1</sup> Department of Surgery, Oncology and Gastroenterology, University of Padova, Padova, Italy, <sup>2</sup> Medical Oncology 2, Istituto Oncologico Veneto IOV-IRCCS, Padova, Italy, <sup>3</sup> Anatomy and Histology Unit, Padova Hospital, Padova, Italy, <sup>4</sup> Breast Surgery, Istituto Oncologico Veneto IOV-IRCCS, Padova, Italy, <sup>5</sup> Pathology, Ulss 5 Polesana, Rovigo, Italy

**Background:** We evaluated immunohistochemical AR expression and correlation with prognosis in a large series of homogeneously treated patients with primary TNBC.

**Material and Methods:** Patients diagnosed with stage I-III TNBC between 2000 and 2015 at Istituto Oncologico Veneto who received treatment with surgery and neoadjuvant and/or adjuvant chemotherapy were included. Whole tissue slides were stained for AR. AR-positive expression was defined as >1% of positively stained tumor cells. Distant-disease-free survival (DDFS) was calculated from diagnosis to distant relapse or death. Late-DDFS was calculated from the landmark of 3 years after diagnosis until distant relapse or death.

**Results:** We included 263 primary TNBC patients. Mean AR expression was 14% (range 0–100%), and 29.7% ( $n = 78$ ) of patients were AR+. AR+ vs. AR- cases presented more frequently older age ( $p < 0.001$ ), non-ductal histology ( $p < 0.001$ ), G1-G2 ( $p = 0.003$ ), lower Ki67 ( $p < 0.001$ ) and lower TILs ( $p = 0.008$ ). At a median follow up of 81 months, 23.6% of patients experienced a DDFS event: 33.3% of AR+ and 19.5% of AR- patients ( $p = 0.015$ ). 5 years DDFS rates were 67.2% and 80.6% for AR+ and AR- patients (HR = 1.82 95%CI 1.10–3.02,  $p = 0.020$ ). AR maintained an independent prognostic role beyond stage, but when TILs were added to the model only stage and TILs were independent prognostic factors. AR was the only factor significantly associated with late-DDFS: 16.4% of AR+ and 3.4% of AR- patients experienced a DDFS after the landmark of 3 years after diagnosis ( $p = 0.001$ ). Late-DDFS rates at 5 years from the 3-year landmark were 75.8% for AR+ and 95.2% for AR- patients (log-rank  $p < 0.001$ ; HR = 5.67, 95%CI 1.90–16.94,  $p = 0.002$ ).



**Conclusions:** AR expression is associated with worse outcome for patients with TNBC. In particular, AR+ TNBC patients are at increased risk of late DDFS events. These results reinforce the rationale of AR targeting in AR+ TNBC.

**Keywords:** androgen receptor, triple negative, early breast cancer, androgen receptor, triple negative, early breast cancer, prognosis, late outcome

## INTRODUCTION

Triple negative breast cancer (TNBC) represents the most lethal breast cancer subtype, accounting for around 15% of all breast cancer diagnoses and being associated with an increased risk of relapse at distant sites, mostly occurring within the first 3 years from diagnosis (1). It is defined by the absence of expression of estrogen and progesterone receptors and lack of HER-2 overexpression/amplification. To date, chemotherapy remains the mainstay of systemic treatment for TNBC, since no relevant druggable targets have been identified (2).

In recent years, the application of genomic profiling techniques has allowed to dissect the heterogeneity of TNBC. At least four main TNBC subtypes have been defined (3, 4), including the luminal androgen receptor (LAR) class, which is enriched for hormonally regulated pathways and is dependent on AR signaling. The LAR subtype accounts for approximately 10–15% of TNBC and LAR cell lines have shown sensitivity to AR-antagonists (3, 4).

AR is found to be expressed by immunohistochemistry in 60–80% of breast cancers, less frequently in estrogen receptor-negative as compared to estrogen-receptor positive tumors (5). In TNBC series, the rate of AR-positive cases is generally 20–40% (5–8), with few studies showing rates up to 60% (9). Preclinical evidence shows that the AR effect depends on tumor subtype: in estrogen receptor-positive cancer cells AR activity is able to inhibit tumor growth (10), whereas in TNBC AR seems to retain an oncogenic effect (11, 12). With regards to the prognostic role of AR expression in patients cohorts, available evidence supports an association between AR expression and favorable prognosis for estrogen receptor-positive tumors (5, 13). In TNBC, data are more conflicting, with some studies showing a favorable prognosis associated with AR expression, some showing null results and others showing an association between AR expression and unfavorable outcome (5). Different methods of AR assessment and scoring, heterogeneity in patients cohorts and short follow up may have yielded to these contrasting results.

In this study, we evaluated AR expression by immunohistochemistry and its correlation with distant disease-free survival in a large cohort of patients with non-metastatic TNBC homogeneously treated with surgery and systemic chemotherapy.

## METHODS

### Patients Population

We included 263 patients with non-metastatic TNBC (estrogen receptor and progesterone receptor <10%, HER2 0/1+ by

immunohistochemistry and/or FISH non amplified) diagnosed from March 2000 to December 2015 at IRCCS Istituto Oncologico Veneto (Padova, Italy) who received treatment with surgery and neoadjuvant and/or adjuvant chemotherapy. Clinicopathological characteristics as well as treatment and follow up data were collected in a dedicated database. The study protocol was approved by the Ethical Committee of the Istituto Oncologico Veneto IRCCS (Padova, Italy). Written informed consent was obtained from patients.

### Pathology Assessments

AR expression was evaluated on the following FFPE primary tumor samples for main analyses: surgical sample for patients treated with primary surgery followed by adjuvant chemotherapy and diagnostic core-biopsy for patients treated with neoadjuvant chemotherapy followed by surgery.

In case of patients treated with neoadjuvant chemotherapy showing residual invasive breast cancer at the examination of the surgical sample, the FFPE surgical tumor block was also retrieved in order to conduct exploratory analysis of changes in AR expression from pre- to post-neoadjuvant chemotherapy.

AR nuclear staining was evaluated on whole sections by immunohistochemistry with the Dako AR441 antibody. AR was scored by a dedicated pathologist, blinded for clinical data, and was considered positive in case of staining in at least 1% of tumor cells, consistently with most recent studies (8, 9).

Tumor infiltrating lymphocytes (TILs) were evaluated according to consensus guidelines on hematoxylin and eosin-stained slides (14).

### Statistical Analysis

Statistical analysis was carried out using IBM SPSS (version 24) software.

Descriptive statistics were performed for patient demographics and clinical characteristics. For continuous variables, median and quartiles were computed. The  $\chi^2$  test or the Mann-Whitney non-parametric test were used to study association between variables, according to their nature (categorical or continuous). The Wilcoxon signed-rank test was used to study the changes in AR expression before and after neoadjuvant chemotherapy in the subset of patients who received this type of treatment and showing residual invasive disease on the surgical sample.

Distant-disease free survival (DDFS) was defined as the time from diagnosis to relapse at a distant site or death from any cause, whichever first. Late-DDFS analysis were performed from the landmark of 3 years after diagnosis until relapse at a distant site or death from any cause, whichever first. In late-DDFS analysis, patients with an event or censored before the

landmark point were excluded. The landmark for late-DDFS was defined based on the pattern of relapse for TNBC that shows a peak in the hazard rate of recurrence in the first 3 years after diagnosis (15). Overall survival (OS) was defined as the time from diagnosis to death from any cause. The Kaplan–Meier method was used to estimate survival curves, the log-rank test was used to test difference between groups. Univariate and multivariate Cox regression models were used to calculate HR and 95% CI. All reported *p*-values are two-sided, and significance level was set at  $p < 0.05$ .

## RESULTS

### Clinicopathological Characteristics and Association With AR

Mean AR expression level was 14% (range 0–100%). Of 263 TNBC patients, 29.7% ( $N = 78$ ) showed a positive AR expression. Images of representative slides are shown in **Figure 1**. Clinicopathological characteristics according to AR status are reported in **Table 1**. AR expression was significantly associated with older age ( $p = 0.002$ ), non-ductal histology ( $p < 0.001$ ), Grade 1–2 tumors ( $p = 0.003$ ), lower Ki67 ( $p < 0.001$ ), lower TILs ( $p = 0.008$ ). There was no difference in stage and treatment received according to AR. Considering neoadjuvant and adjuvant therapy combined, 73% of patients received both an anthracycline and a taxane as part of chemotherapy treatment.

### Survival Analyses

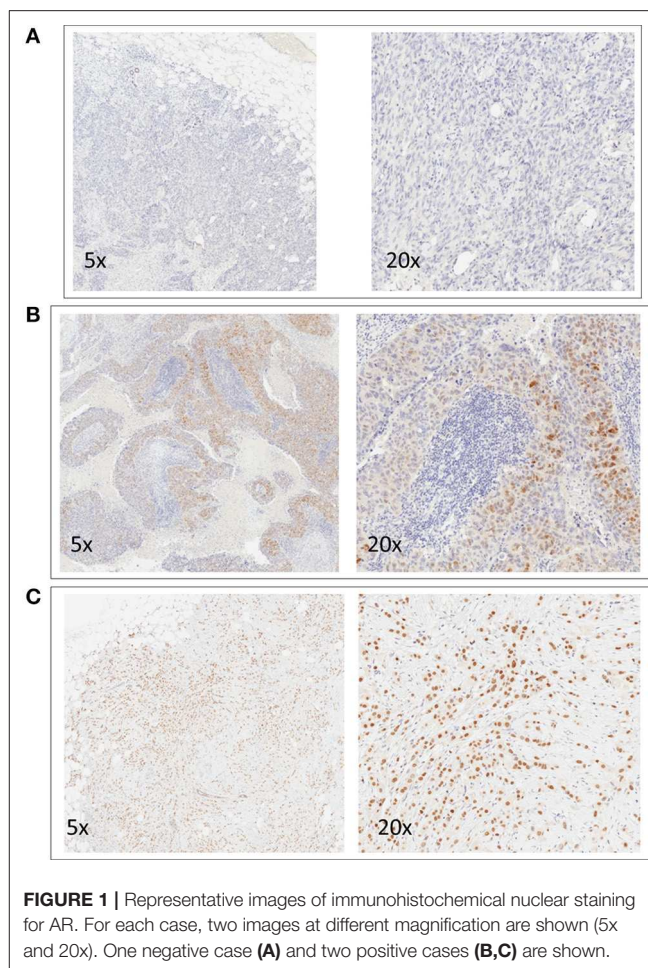
At a median follow up of 81 months (95% CI 74–87), 62 patients have experienced a DDFS event (23.6%). Type of DDFS event was: distant relapse in 56 patients (90%) and death in 6 patients (10%, two deaths occurred in patients with unresectable chest locoregional recurrence and 4 patients died without known breast cancer relapse). The rate of events was higher in AR+ as compared to AR- patients (33.3 and 19.5%, respectively).

As shown in **Figure 2A**, Patients with AR+ tumor showed worse DDFS as compared to AR- patients: 5 years DDFS rates were 67.2 and 80.6%, respectively (log-rank  $p = 0.018$ ). The HR for DDFS for the comparison of AR+ vs. AR- groups was 1.82 (95%CI 1.10–3.02,  $p = 0.020$ ).

**Figure 2B** shows OS Kaplan–Meier curves: 5 years OS rate was 79.9% for AR+ and 82.7% for AR- patients (log-rank  $p = 0.161$ ). The HR for OS for the comparison of AR+ vs. AR- patients was 1.48 (95% CI 0.85–2.58,  $p = 0.163$ ).

Univariate and multivariate cox models for DDFS are reported in **Table 2**.

In addition to AR, the other factors that were associated in univariate analysis with DDFS, were stage (Stage II–III vs. I,  $p = 0.024$ ) and TILs (considered as continuous variable for each 1% increment,  $p = 0.005$ ). In multivariate analysis including AR and Stage, both factors maintained an independent prognostic role (AR+ vs. AR-: HR = 1.74, 95%CI 1.05–2.88,  $p = 0.032$ ; Stage II–III vs. I: HR 3.05, 95%CI 1.83–5.08,  $p < 0.001$ ). When TILs were added to the multivariate model, only stage and TILs maintained an independent prognostic value. The HR for the association between AR status and DDFS in multivariate models including the three variables was 1.57 (95% CI 0.94–2.61,  $p = 0.084$ ).



**FIGURE 1** | Representative images of immunohistochemical nuclear staining for AR. For each case, two images at different magnification are shown (5x and 20x). One negative case (**A**) and two positive cases (**B,C**) are shown.

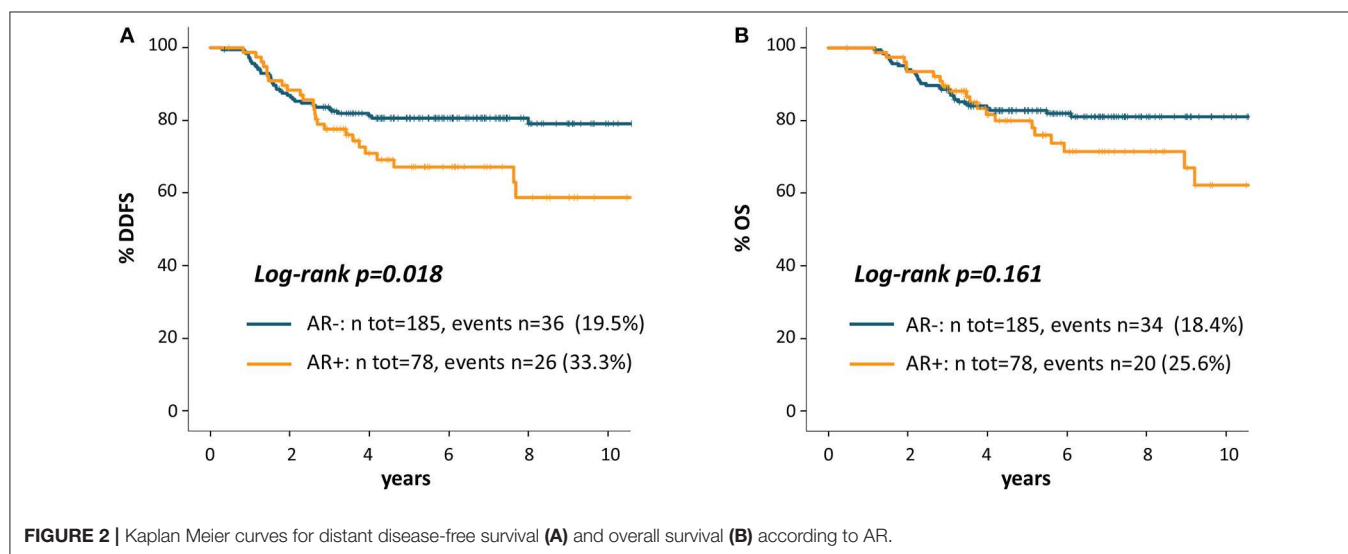
Since Kaplan Meier curves showed that the prognostic effect of AR on DDFS appeared driven by the occurrence of late recurrences in AR+ patients, we performed a landmark survival analysis for late-DDFS to study the association between AR and late outcome. This analysis included 203 patients who were DDFS-free at 3 years from initial diagnosis and were not censored before the landmark point:  $n = 55$  (27%) were AR+ and  $n = 148$  (73%) were AR-. At a median follow up of 47 months (95% CI 41–53)  $n = 14$  DDFS events have occurred. The rate of event was higher in AR+ (9/55, 16.4%) vs. AR- patients (5/148, 3.4%). Type of DDFS event included: 10 distant relapses and 4 deaths (1 in a patient with unresectable locoregional breast recurrence and 3 in patients without prior known breast cancer relapse). AR+ patients showed more frequently distant relapses ( $n = 8$  of 9 total events, 89%) as compared to AR- patients ( $n = 2$  of 5 total events, 40%).

Kaplan Meier curves in **Figure 3** shows that patients with AR+ tumor experienced a significantly worse late outcome as compared to AR- patients: late-DDFS rate at 5 years from the 3-years landmark were 75.8% for AR+ patients and 95.2% for AR- patients (log-rank  $p < 0.001$ ). Univariate late-DDFS cox

**TABLE 1** | Clinicopathological patients' characteristics by AR expression.

Patients' features		ALL (N = 263) N (%)	AR+ (N = 78) N (%)	AR- (N = 185) N (%)	p
Age years, median (Q1-Q3)		53 (44-66)	62 (47-70)	51 (42-62)	0.002
Hystotype	Ductal/NOS	237 (91.2)	66 (84.6)	171 (94.0)	<0.001
	Lobular	7 (2.7)	5 (6.4)	2 (1.1)	
	Apocrine	8 (3.1)	7 (9.0)	1 (0.5)	
	Metaplastic	5 (1.9)	0	5 (2.7)	
	Medullary	3 (1.2)	0	3 (1.6)	
AJCC Stage	I	83 (31.7)	24 (30.8)	59 (32.1)	0.886
	II	130 (49.6)	38 (48.7)	92 (50.0)	
	III	49 (18.7)	16 (20.5)	33 (17.9)	
Grade	G1-2	29 (11.9)	16 (21.1)	13 (7.7)	0.003
	G3	2515 (88.1)	60 (78.9)	155 (92.3)	
Ki67%, median (Q1-Q3)		55 (36-70)	40 (27-60)	60 (40-70)	<0.001
TILs%, median (Q1-Q3)		10 (5-30)	7 (2-20)	10 (5-30)	0.008
Neoadjuvant CT	Yes	108 (41.1)	27 (34.6)	81 (43.8)	0.167
	No	155 (58.9)	51 (65.4)	104 (56.2)	
Type of neoadjuvant CT	Anthra+tax	101 (93.5)	24 (88.9)	77 (95.1)	0.299
	Anthra	1 (0.9)	0	1 (1.2)	
	Tax	6 (5.6)	3 (11.1)	3 (3.7)	
Adjuvant CT	Yes	186 (71.3)	57 (73.1)	129 (70.5)	0.673
	No	75 (28.7)	21 (26.9)	54 (29.5)	
Type of adjuvant CT	Anthra+tax	96 (51.6)	28 (49.1)	68 (52.7)	0.523
	Anthra	37 (19.9)	14 (24.6)	23 (17.8)	
	Tax	8 (4.3)	1 (1.8)	7 (5.4)	
	Other	45 (24.2)	14 (24.6)	31 (24.0)	
Radiotherapy	Yes	165 (67.1)	48 (64.9)	117 (68.0)	0.629
	No	81 (32.9)	26 (35.1)	55 (32.0)	

N, number; AR, androgen receptor; p, p-value; Q1, first quartile; Q3, third quartile; NOS, not otherwise specified; AJCC, American Joint Committee on Cancer; TILs, tumor infiltrating lymphocytes; CT, chemotherapy; Anthra, anthracycline; Tax, taxane.



analysis for the comparison of AR+ vs. AR- patients showed HR = 5.67 (95% CI 1.90-16.94,  $p = 0.002$ ). No other factor showed a significant association with late-DDFS including: age (HR = 1.02,

95% CI 0.98-1.06,  $p = 0.377$ ), histologic Grade (Grade 3 vs. 1-2 HR = 1.96, 95% CI 0.25-15.56,  $p = 0.524$ ), stage (stage II-III vs. I, HR = 0.96, 95% CI 0.32-2.87,  $p = 0.943$ ) and TILs (HR =

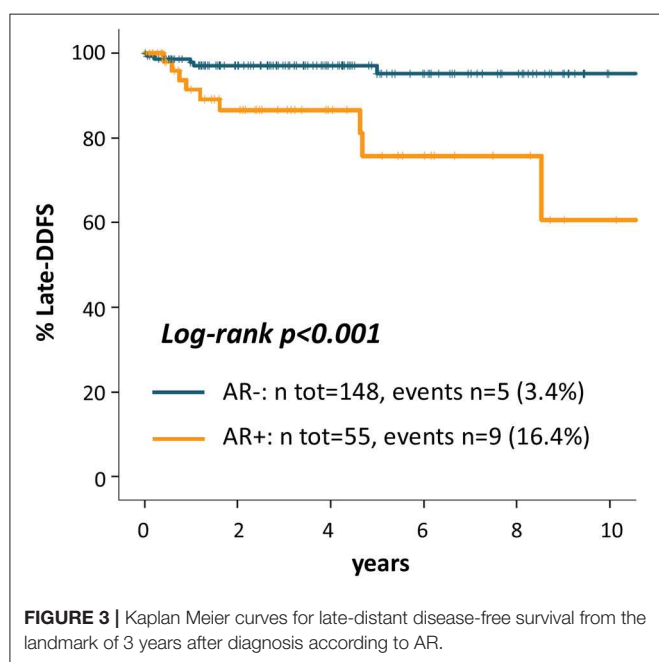
**TABLE 2 |** Univariate and multivariate DDFS cox models.

	Univariate			Multivariate model 1*			Multivariate model 2**		
	HR	95%CI	p	HR	95%CI	p	HR	95%CI	p
Age (continuous)	1.01	0.90–1.03	0.446	–	–	–	–	–	–
Grade 1-2	Ref			–	–	–	–	–	–
Grade 3	1.29	0.55–3.02	0.553						
Stage I	Ref			Ref			Ref		
Stage II-III	2.07	1.10–3.89	0.024	3.05	1.83–5.08	<0.001	2.34	1.26–4.47	0.008
TILs (1% increments)	0.98	0.96–0.99	0.005	–	–	–	0.98	0.96–0.99	0.004
AR –	Ref			Ref			Ref		
AR +	1.82	1.10–3.02	0.020	1.74	1.05–2.88	0.032	1.57	0.94–2.61	0.084

HR, hazard ratio; CI, confidence interval; p, p-value; TILs, tumor infiltrating lymphocytes, AR, androgen receptor.

\*Including stage and AR.

\*\*Including stage, AR and TILs.



0.98, 95% CI 0.95–1.01,  $p = 0.178$ ). However, number of events was low.

A list of cases with DDFS event and matched clinicopathological features is provided as **Table 3**. Moreover, exploratory additional survival analyses according to a cut-off of >10% of AR expression are reported in **Supplementary Figure 1**.

### Additional Analyses in Patients Treated With Neoadjuvant Chemotherapy

Of the 108 patients who received neoadjuvant chemotherapy, information on pathological response was available for 107 cases. A pathological complete response (pCR), defined as the absence of invasive cancer cells in the breast and axillary lymphnodes on the surgical specimen, was observed in 28% of cases ( $n = 30$ ). The rate of pCR was similar in

AR+ and AR- patients: 25.9 and 28.8%, respectively ( $p = 0.778$ ). Tumor tissue sample from the surgical specimen obtained after neoadjuvant chemotherapy was available for AR evaluation for  $n = 60$  patients without pCR (patients' flow diagram provided in **Supplementary Figure 2**). AR expression showed a non-significant decrease after neoadjuvant chemotherapy: mean 13% on the diagnostic core-biopsy and 10% on the paired surgical specimen (Wilcoxon signed-rank test  $p = 0.172$ ). All those cases that were AR- on the diagnostic core-biopsy were also AR- on the surgical specimen ( $n = 43$ ), whereas 41% of the 17 initially AR+ cases lost AR expression after neoadjuvant chemotherapy ( $\chi^2 p < 0.001$ ).

## DISCUSSION

In this study we showed that AR expression is associated with worse DDFS in TNBC patients treated with surgery and systemic chemotherapy. Although AR did not retain an independent prognostic value for DDFS in multivariate analysis in the total follow-up period, we found that AR expression was the only factor that resulted in a significant increase in the risk of late-DDFS event. Of note, the vast majority of events were distant relapses or deaths in patients with unresectable locoregional recurrences. Therefore, the potential confounding effect of deaths of unknown cause or not related to breast cancer (which may be relevant in studies with long-term follow up) is very limited.

We found that 30% of TNBC cases were classified as AR+, which is in line with a number of other studies (5–8). The correlation of AR+ status with other clinicopathologic characteristics such as older age, non-ductal histology, lower histologic grade, lower ki67 and lower TILs, is also consistent with other studies assessing AR by immunohistochemistry or evaluating the LAR molecular subtype (4, 9, 12, 16).

The available evidence on the prognostic role of AR for patients with early TNBC is conflicting. A recent metaanalysis reported that AR expression significantly predicts for a better



**TABLE 3 |** List of cases with a DDFS event and matched clinicopathological features entered in univariate and multivariable cox regression models.

Progressive number	Type of DDFS event	AR expression, %	Age, years	Grade	Stage	TILs, %
1	Death	0	58	G3	Stage III	20
2	Distant relapse	0	54	G3	Stage III	5
3	Distant relapse	0	36	G3	Stage I-II	1
4	Distant relapse	0	72	.	Stage I-II	30
5	Distant relapse	0	44	G3	Stage III	10
6	Distant relapse	0	47	G3	Stage III	5
7	Distant relapse	0	57	G3	Stage III	0
8	Distant relapse	0	44	.	Stage III	15
9	Distant relapse	0	48	G3	Stage I-II	10
10	Distant relapse	0	58	G3	Stage I-II	10
11	Distant relapse	0	41	G3	Stage III	60
12	Distant relapse	0	73	G3	Stage I-II	10
13	Distant relapse	0	57	G3	Stage I-II	32
14	Distant relapse	0	50	G3	Stage III	35
15	Death	0	77	G1-2	Stage I-II	2
16	Distant relapse	0	68	G3	Stage I-II	0
17	Distant relapse	0	55	G3	Stage I-II	22
18	Distant relapse	0	30	G1-2	Stage III	5
19	Distant relapse	0	77	G3	Stage I-II	1
20	Distant relapse	0	52	G3	Stage I-II	30
21	Distant relapse	0	39	G3	Stage I-II	7
22	Distant relapse	0	50	G3	Stage I-II	5
23	Death	0	46	G3	Stage I-II	15
24	Distant relapse	0	51	G3	Stage I-II	1
25	Distant relapse	0	42	G3	Stage III	1
26	Distant relapse	0	70	G3	Stage I-II	20
27	Distant relapse	0	46	G3	Stage I-II	5
28	Distant relapse	0	46	G1-2	Stage I-II	3
29	Distant relapse	0	80	G3	Stage I-II	3
30	Death	0	73	G3	Stage I-II	25
31	Distant relapse	0	70	G3	Stage I-II	35
32	Distant relapse	0	53	G3	Stage III	30
33	Distant relapse	0	41	G3	Stage III	5
34	Distant relapse	0	61	G3	Stage III	5
35	Distant relapse	0	70	G3	Stage I-II	10
36	Distant relapse	0	52	G3	Stage III	20
37	Distant relapse	1	47	.	Stage I-II	2
38	Distant relapse	1	45	G1-2	Stage I-II	10
39	Death	1	50	G3	Stage I-II	12
40	Distant relapse	1	37	G3	Stage I-II	7
41	Distant relapse	1	45	G3	Stage III	35
42	Distant relapse	2	64	G3	Stage III	30
43	Distant relapse	5	38	G3	Stage I-II	35
44	Distant relapse	5	30	G3	Stage I-II	7
45	Distant relapse	5	52	G3	Stage I-II	0
46	Distant relapse	5	52	G3	Stage I-II	2
47	Distant relapse	20	48	G3	Stage I-II	5
48	Distant relapse	30	64	G3	Stage I-II	2
49	Distant relapse	30	74	G3	Stage I-II	10
50	Death	40	82	G3	Stage III	5
51	Distant relapse	50	42	G3	Stage III	30

(Continued)



TABLE 3 | Continued

Progressive number	Type of DDFS event	AR expression, %	Age, years	Grade	Stage	TILs, %
52	Distant relapse	70	66	G3	Stage III	1
53	Distant relapse	75	74	G3	Stage III	5
54	Distant relapse	80	54	G3	Stage I-II	1
55	Distant relapse	85	65	G3	Stage III	1
56	Distant relapse	90	79	G3	Stage I-II	5
57	Distant relapse	90	84	G1-2	Stage I-II	7
58	Distant relapse	90	73	G3	Stage I-II	5
59	Distant relapse	95	41	G1-2	Stage III	5
60	Distant relapse	99	64	G3	Stage III	1
61	Distant relapse	99	55	G3	Stage III	10
62	Distant relapse	100	47	G3	Stage I-II	10

Events also considered in the late-DDFS analysis are highlighted in gray. DDFS, distant disease-free survival.

survival in TNBC (HR for DFS = 0.64, 95%CI 0.51-0.81 and HR for OS = 0.64, 95%CI 0.49-0.88) (13). Multivariate analysis was not available. It has to be noted that this was a study-level and not a patient-level meta-analysis, including studies that were heterogeneous for methods of AR scoring, clinical cohorts characteristics, treatment and length of follow up. At least two other retrospective studies were issued after the publication of this meta-analysis, reporting no association of AR with prognosis in TNBC (sample size of  $n = 130$  and  $n = 182$ , respectively) (17, 18). In addition, two other larger studies have recently demonstrated an unfavorable prognosis for AR+ TNBC patients (8, 9). In both these studies the Dako AR441 antibody was used and the definition of AR+ in immunohistochemistry was based on the >1% cut-off, consistently with the methods applied in our analysis. Data from the TNBC subset of the prospective Nurses' Health Studies cohorts ( $n = 581$ ) have reported, over a median follow up of 16.5 years, a significantly unfavorable breast cancer-specific survival in multivariable models for AR+ vs. AR- patients (8). In this study the prognostic impact of AR was evident in years 0–7 after diagnosis with an HR of 1.59 (95%CI 1.07–2.37) that maintained a similar value even >7 years after diagnosis, although not reaching statistical significance in this period (HR = 1.41, 95%CI 0.84–2.36). When looking at survival curves in this study, they result very similar to the ones reported in our analysis, with a separation of the curves for AR+ and AR- patients that starts around 3 years after diagnosis, supporting our findings of AR+ tumors being associated with an increased risk of late relapses. In another retrospective series of more than 300 TNBC (9), the significant association between AR+ and worse outcome was further refined by the combined evaluation of AR and forkhead-box A1 (FOXA1), a protein required for AR transcriptional activity (19). Indeed, patients with AR+/FOXA1+ TNBC showed a worse overall survival as compared to other patients in multivariable model (HR = 1.57, 95%CI 1.01-2.45) (9). Again, survival curves started to separate at around 3 years after diagnosis.

Although AR expression by itself can only be considered as a suboptimal surrogate of the molecular LAR TNBC

subtype (20), our results, together with the ones by Kensler et al. and Guiu et al. are consistent with findings suggesting the association of LAR subtype with poor prognosis in TNBC (16). Potential biological reasons for this association may include: the proposed oncogenic role of AR in TNBC (11, 12) and a distinct genomic landscape including an enrichment in somatic *PIK3CA* and *AKT1* mutations (9, 16). Moreover, AR+/LAR TNBC are associated with lower TILs (4), as also shown in our study. In particular, in our work, this correlation might explain the lack of independent prognostic role of AR for DDFS in the total follow-up period when both TILs and stage are added to the multivariate model.

Anti-androgen therapies are under investigation for breast cancer in different settings (12) and phase II studies in metastatic TNBC AR+ patients have already obtained encouraging results (21–23). If further validated by other studies, our results showing that TNBC AR+ patients are at increased risk of late DDFS event may be useful in planning the future development of antiandrogen adjuvant therapies in TNBC.

With regards to the subset of patients treated with neoadjuvant chemotherapy, we did not observe different rates of pCR according to AR, however sample size was limited. The majority of data indicate that TNBC with a positive AR expression or owing to the LAR subtype achieve lower rates of pCR as compared to other TNBC patients (4, 24–26), although other studies showed conflicting results (6). The achievement of pCR is associated with long-term outcome in TNBC. Whether and to which extent the less likelihood of pCR for AR+/LAR TNBC contributes to the long-term outcome of these patients is not clear at this time (25, 26). Moreover, interpretation of results from different studies is limited by the lack of concordance between the evaluation of AR expression by immunohistochemistry and the LAR classification by gene expression. The evaluation of combined chemotherapy and antiandrogen therapy is ongoing in the neoadjuvant setting (NCT02689427).

Our study has strengths, including: the large sample size, the homogeneous treatment received by patients which is consistent with contemporary standards (all patients treated with chemotherapy and surgery, the vast majority received both an anthracycline and a taxane), the methods for AR assessment in line with the most recent studies and the length of follow up (median 81 months), allowing to uncover the impact of AR on late outcome.

Limitations of our study include the retrospective nature and the low number of events in late-DDFS analysis that imposes caution in results interpretation and further validation in additional studies.

In conclusion, our results show that the evaluation of AR in TNBC is able to identify a subgroup of patients at worse prognosis, especially for the occurrence of late events. Further validation in other studies is warranted. These data support the

rationale for the ongoing evaluation of antiandrogen therapies in TNBC.

## AUTHOR CONTRIBUTIONS

MD planned and coordinated the manuscript. VT performed the pathology assessment. VG supervised the whole writing of the manuscript. Each author participated for appropriate portions of the content. All the authors conceived the review and approved of the final analysis and results.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2019.00452/full#supplementary-material>

## REFERENCES

- Dent R, Trudeau M, Pritchard KI, Hanna WM, Kahn HK, Sawka CA, et al. Triple-negative breast cancer: clinical features and patterns of recurrence. *Clin Cancer Res.* (2007) 13:4429–34. doi: 10.1158/1078-0432.CCR-06-3045
- Bianchini G, Balko JM, Mayer IA, Sanders ME, Gianni L. Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease. *Nat Rev Clin Oncol.* (2016) 13:674–90. doi: 10.1038/nrclinonc.2016.66
- Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest.* (2011) 121:2750–67. doi: 10.1172/JCI45014
- Lehmann BD, Jovanović B, Chen X, Estrada MV, Johnson KN, Shyr Y, et al. Refinement of triple-negative breast cancer molecular subtypes: implications for neoadjuvant chemotherapy selection. *PLoS ONE.* (2016) 11:e0157368. doi: 10.1371/journal.pone.0157368
- Ricciardelli C, Bianco-Miotto T, Jindal S, Butler LM, Leung S, McNeil CM, et al. The magnitude of androgen receptor positivity in breast cancer is critical for reliable prediction of disease outcome. *Clin Cancer Res.* (2018) 24:2328–41. doi: 10.1158/1078-0432.CCR-17-1199
- Loibl S, Müller BM, von Minckwitz G, Schwabe M, Roller M, Darb-Esfahani S, et al. Androgen receptor expression in primary breast cancer and its predictive and prognostic value in patients treated with neoadjuvant chemotherapy. *Breast Cancer Res Treat.* (2011) 130:477–87. doi: 10.1007/s10549-011-1715-8
- Wang C, Pan B, Zhu H, Zhou Y, Mao F, Lin Y, et al. Prognostic value of androgen receptor in triple negative breast cancer: a meta-analysis. *Oncotarget.* (2016) 7:46482–91. doi: 10.18632/oncotarget.10208
- Kensler KH, Poole ME, Heng YJ, Collins LC, Glass B, Beck AH, et al. Androgen receptor expression and breast cancer survival: results from the Nurses' Health Studies. *J Natl Cancer Inst.* (2018). doi: 10.1093/jnci/djy173. [Epub ahead of print].
- Guiu S, Mollevi C, Charon-Barra C, Boissière F, Crapez E, Chartron E, et al. Prognostic value of androgen receptor and FOXA1 co-expression in non-metastatic triple negative breast cancer and correlation with other biomarkers. *Br J Cancer.* (2018) 119:76–9. doi: 10.1038/s41416-018-0142-6
- Peters AA, Buchanan G, Ricciardelli C, Bianco-Miotto T, Centenera MM, Harris JM, et al. Androgen receptor inhibits estrogen receptor- $\alpha$  activity and is prognostic in breast cancer. *Cancer Res.* (2009) 69:6131–40. doi: 10.1158/0008-5472.CAN-09-0452
- Doane AS, Danso M, Lal P, Donaton M, Zhang L, Hudis C, et al. An estrogen receptor-negative breast cancer subset characterized by a hormonally regulated transcriptional program and response to androgen. *Oncogene.* (2006) 25:3994–4008. doi: 10.1038/sj.onc.1209415
- Gerratana L, Basile D, Buono G, De Placido S, Giuliano M, Minichillo S, et al. Androgen receptor in triple negative breast cancer: a potential target for the targetless subtype. *Cancer Treat Rev.* (2018) 68:102–10. doi: 10.1016/j.ctrv.2018.06.005
- Bozovic-Spasojevic I, Zardavas D, Brohée S, Ameye L, Fumagalli D, Ades F, et al. The prognostic role of androgen receptor in patients with early-stage breast cancer: a meta-analysis of clinical and gene expression data. *Clin Cancer Res.* (2017) 23:2702–12. doi: 10.1158/1078-0432.CCR-16-0979
- Salgado R, Denkert C, Demaria S, Sirtaine N, Klauschen F, Pruneri G, et al. The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: recommendations by an International TILs Working Group 2014. *Ann Oncol.* (2015) 26:259–71. doi: 10.1093/annonc/mdl450
- Cossetti RJ, Tyldesley SK, Speers CH, Zheng Y, Gelmon KA. Comparison of breast cancer recurrence and outcome patterns between patients treated from 1986 to 1992 and from 2004 to 2008. *J Clin Oncol.* (2015) 33:65–73. doi: 10.1200/JCO.2014.57.2461
- Bareche Y, Venet D, Ignatiadis M, Aftimos P, Piccart M, Rothe F, et al. Unravelling triple-negative breast cancer molecular heterogeneity using an integrative multiomic analysis. *Ann Oncol.* (2018) 29:895–902. doi: 10.1093/annonc/mdy024
- Kraby MR, Valla M, Opdahl S, Haugen OA, Sawicka JE, Engström MJ, et al. The prognostic value of androgen receptors in breast cancer subtypes. *Breast Cancer Res Treat.* (2018) 172:283–96. doi: 10.1007/s10549-018-4904-x
- Aleskandarany MA, Abduljabbar R, Ashankyty I, Elmouna A, Jerjees D, Ali S, et al. Prognostic significance of androgen receptor expression in invasive breast cancer: transcriptomic and protein expression analysis. *Breast Cancer Res Treat.* (2016) 159:215–27. doi: 10.1007/s10549-016-3934-5
- Robinson JL, Macarthur S, Ross-Innes CS, Tilley WD, Neal DE, Mills IG, et al. Androgen receptor driven transcription in molecular apocrine breast cancer is mediated by FoxA1. *EMBO J.* (2011) 30:3019–27. doi: 10.1038/emboj.2011.216
- Traina T, Miller K, Yardley D, O'Shaughnessy J, Cortes J, Awada A, et al. In: results from a phase 2 study of enzalutamide (ENZA), an androgen receptor (AR) inhibitor, in advanced AR+ triple-negative breast cancer (TNBC). *J Clin Oncol.* (2015) 33:1003. doi: 10.1200/jco.2015.33.15\_suppl.1003
- Gucalp A, Tolaney S, Isakoff SJ, Ingle JN, Liu MC, Carey LA, et al. Phase II trial of bicalutamide in patients with androgen receptor-positive, estrogen receptor-negative metastatic breast cancer. *Clin Cancer Res.* (2013) 19:5505–12. doi: 10.1158/1078-0432.CCR-12-3327
- Traina TA, Miller K, Yardley DA, Eakle J, Schwartzberg LS, O'Shaughnessy J, et al. Enzalutamide for the treatment of androgen

- receptor-expressing triple-negative breast cancer. *J Clin Oncol.* (2018) 36:884–90. doi: 10.1200/JCO.2016.71.3495
23. Bonnefoi H, Grellety T, Tredan O, Saghachian M, Dalenc F, Mailliez A, et al. A phase II trial of abiraterone acetate plus prednisone in patients with triple-negative androgen receptor positive locally advanced or metastatic breast cancer (UCBG 12–1). *Ann Oncol.* (2016) 27:812–8. doi: 10.1093/annonc/mdw067
  24. Echavarria I, López-Tarruella S, Picornell A, García-Saenz JÁ, Jerez Y, Hoadley K, et al. Pathological response in a triple-negative breast cancer cohort treated with neoadjuvant carboplatin and docetaxel according to Lehmann's refined classification. *Clin Cancer Res.* (2018) 24:1845–52. doi: 10.1158/1078-0432.CCR-17-1912
  25. Jovanović B, Mayer IA, Mayer EL, Abramson VG, Bardia A, Sanders ME, et al. A randomized phase II neoadjuvant study of cisplatin, paclitaxel with or without everolimus in patients with stage II/III triple-negative breast cancer (TNBC): responses and long-term outcome correlated with increased frequency of DNA damage response gene mutations, TNBC subtype, AR status, and Ki67. *Clin Cancer Res.* (2017) 23:4035–45. doi: 10.1158/1078-0432.CCR-16-3055
  26. Masuda H, Baggerly KA, Wang Y, Zhang Y, Gonzalez-Angulo AM, Meric-Bernstam F, et al. Differential response to neoadjuvant chemotherapy among

7 triple-negative breast cancer molecular subtypes. *Clin Cancer Res.* (2013) 19:5533–40. doi: 10.1158/1078-0432.CCR-13-0799

**Conflict of Interest Statement:** MD has received fees from EliLilly for consultancy role and participation on advisory boards; fees from Genomic Health for consultancy role; fees from Celgene for participation on advisory boards. VG has received honoraria from EliLilly and Roche for participation on advisory boards, and honoraria from AstraZeneca and Novartis.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Dieci, Tsvetkova, Griguolo, Miglietta, Mantiero, Tasca, Cumerlato, Giorgi, Giarratano, Faggioni, Falci, Vernaci, Menichetti, Mioranza, Di Liso, Frezzini, Saibene, Orvieto and Guarneri. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# TUFT1 Promotes Triple Negative Breast Cancer Metastasis, Stemness, and Chemoresistance by Up-Regulating the Rac1/ $\beta$ -Catenin Pathway

Weiguang Liu<sup>1\*</sup>, Guanglei Chen<sup>2</sup>, Lisha Sun<sup>2</sup>, Yue Zhang<sup>3</sup>, Jianjun Han<sup>1</sup>, Yuna Dai<sup>1</sup>, Jianchao He<sup>1</sup>, Sufang Shi<sup>1</sup> and Bo Chen<sup>4\*</sup>

<sup>1</sup> Department of Breast Surgery, Affiliated Hospital of Hebei University of Engineering, Handan, China, <sup>2</sup> Department of Breast Surgery, Shengjing Hospital of China Medical University, Shenyang, China, <sup>3</sup> Department of Physiology, Dalian Medical University, Dalian, China, <sup>4</sup> Department of Breast Surgery, The First Hospital of China Medical University, Shenyang, China

## OPEN ACCESS

### Edited by:

Mothaffar Rimawi,  
Baylor College of Medicine,  
United States

### Reviewed by:

Tomás Pascual Martínez,  
Hospital Clínic of Barcelona, Spain  
Masahiko Tanabe,  
The University of Tokyo, Japan

### \*Correspondence:

Weiguang Liu  
lwg1943@163.com  
Bo Chen  
chbyxl@163.com

### Specialty section:

This article was submitted to  
Women's Cancer,  
a section of the journal  
Frontiers in Oncology

**Received:** 28 March 2019

**Accepted:** 24 June 2019

**Published:** 09 July 2019

### Citation:

Liu W, Chen G, Sun L, Zhang Y,  
Han J, Dai Y, He J, Shi S and Chen B  
(2019) TUFT1 Promotes Triple  
Negative Breast Cancer Metastasis,  
Stemness, and Chemoresistance by  
Up-Regulating the Rac1/ $\beta$ -Catenin  
Pathway. *Front. Oncol.* 9:617.  
doi: 10.3389/fonc.2019.00617

**Objectives:** Triple negative breast cancer (TNBC) is a subtype of breast cancer with stronger invasion and metastasis, but its specific mechanism of action is still unclear. Tuft1 plays an important regulatory role in the survival of breast cancer cells; however, its role in regulating TNBC metastatic potential has not been well-characterized. Our aim was therefore to systematically study the mechanism of TUFT1 in the metastasis, stemness, and chemoresistance of TNBC and provide new predictors and targets for BC treatment.

**Methods:** We used western blotting and IHC to measure TUFT1 and Rac1-GTP expression levels in both human BC samples and cell lines. A combination of shRNA, migration/invasion assays, sphere formation assay, apoptosis assays, nude mouse xenograft tumor model, and GTP activity assays was used for further mechanistic studies.

**Results:** We demonstrated that silencing TUFT1 in TNBC cells significantly inhibited cell metastasis and stemness *in vitro*. A nude mouse xenograft tumor model revealed that TUFT1 knockdown greatly decreased spontaneous lung metastasis of TNBC tumors. Mechanism studies showed that TUFT1 promoted tumor cell metastasis and stemness by up-regulating the Rac1/ $\beta$ -catenin pathway. Moreover, mechanistic studies indicated that the lack of TUFT1 expression in TNBC cells conferred more sensitive to chemotherapy and increased cell apoptosis via down-regulating the Rac1/ $\beta$ -catenin signaling pathway. Further, TUFT1 expression positively correlated with Rac1-GTP in TNBC samples, and co-expression of TUFT1 and Rac1-GTP predicted poor prognosis in TNBC patients who treated with chemotherapy.

**Conclusion:** Our findings suggest that TUFT1/Rac1/ $\beta$ -catenin pathway may provide a potential target for more effective treatment of TNBC.

**Keywords:** triple negative breast cancer (TNBC), TUFT1, Rac1, metastasis, stemness, chemoresistance

## INTRODUCTION

Triple negative breast cancer (TNBC) is a subtype of BC that lacks estrogen or progesterone receptors and has no epidermal growth factor receptor 2 amplification, accounting for about 20% of the total breast cancer (1–3). TNBC is defined mainly based on its pathology. Its features overlap with those of basal-like BC, one of five subgroups based on microarray gene expression profiling (4, 5). TNBC usually presents with less favorable clinical features than other subtypes of breast cancer, for example, tumors proliferate faster, relapse earlier and metastasis more easily and is usually associated with poorer prognosis as a result (6–8). However, the mechanism by which TNBC's metastasis is less clear. In addition, there are currently no very effective targeted drugs available for TNBC, cytotoxic chemotherapy remains the main adjuvant therapy for this subtype of breast cancer (9). A more in-depth study of the mechanism of TNBC metastasis may be able to more efficiently find its target, and at the same time provide theoretical support for the exploration of new TNBC therapeutic drugs.

Tuftelin (TUFT1) is an acidic, hydrophilic, glycosylated, and phosphorylated protein. Sequence and characterization analysis has shown that TUFT1 is well conserved, with high homology across various species. The protein is considered to act on enamel mineralization and is involved in the interaction between mesenchymal ectoderm and autosomal enamel dysplasia during tooth development (10). Zhou et al. (11) demonstrated that the expression of TUFT1 protein in pancreatic cancer is higher than that in normal pancreatic tissue. Its expression is closely related to both the disease stage and local lymph node metastasis. Cell function experiments further confirmed that TUFT1 depletion reduced proliferation and metastasis of pancreatic cancer cells, and impaired various proteins expression related to epithelial-mesenchymal transition. The authors suggested that TUFT1 may affect HIF1 by influencing the expression of members of the Snail signaling pathway, which regulates epithelial mesenchymal transition. Our previous study found that inhibition of TUFT1 expression in breast cancer cells inhibited proliferation, affected the cell cycle, and induced apoptosis. In addition, we showed that suppression of TUFT1 affected the expression of the proteins RelA, Caspase 3, DUSP1, and Rac1 (12, 13). Kawasak et al. (14) found that TUFT1 activated the mTORC1 signaling pathway by regulating the Rab GTPase, and that the interaction between TUFT1 and RabGAP1 mediated intracellular lysosome localization and vesicle transport in tumor cells. However, the precise role of TUFT1 in breast cancer (BC), including the mechanics of TNBC's metastasis remain unclear.

Rac1 is a member of the Rho GTPases family, which is a subgroup of the Ras superfamily (15). Rac1 is activated by binding to GTP, while it is deactivated by binding to GDP, which makes it play an important role in many signaling pathways (16). Rac1 plays an important role in cancer progression (17), affecting cell adhesion, proliferation, migration, invasion, and cancer metastasis (18–20). The new study highlights the importance of Rac1 activation in cancer metastasis and acquired chemoresistance (21–24). One major mechanism by which Rac1 may provide resistance to chemotherapy is its role in apoptosis

regulation. Rac1-GTP can bind directly to the key apoptotic regulator Bcl-2 to elicit anti-apoptotic cell responses (25). Many studies have also proved that Rac1-GTP can affect the genes Nanog, Sox2, and Oct4, which play a central regulatory role in CSC (26–28). Rao et al. (29) showed that Rac1/ $\beta$ -catenin pathway participated in SEMA3F-mediated regulation of colorectal cancer cell stemness. In addition, Kawasak et al. (14) found that TUFT1 increased Rac1 levels through activation of the AKT/mTOR pathway. However, the functional mechanism of TUFT1 in metastasis, stemness, and chemoresistance of BC, especially in TNBC, has not been adequately characterized.

In this study, we showed that stable TUFT1 knockdown in TNBC cells drastically inhibited their migration, invasiveness, and CSC-like properties. Moreover, we found that the expression of TUFT1 increased significantly in TNBC samples. The co-expression of TUFT1 and Rac1-GTP suggested poor prognosis. Further functional studies showed that TUFT1 promoted TNBC cell metastasis, stemness, and chemoresistance by up-regulating the Rac1/ $\beta$ -catenin signaling pathway.

## MATERIALS AND METHODS

### Human Specimens

In our study, we recruited 60 pathologically confirmed TNBC patients at Affiliated Hospital of Hebei University of Engineering, between January 2014 and December 2014. All patients treated with anthracycline followed by taxanes chemotherapy after surgery. This study was carried out in accordance with the recommendations of ICMJE with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Ethics Committee of Affiliated Hospital of Hebei Engineering University.

### Human BC Cell Lines and Plasmids

HCC1937 cell line was obtained from the American Type Culture Collection (USA). MDA-MB-231 cell line was gained from the Chinese Academy of Sciences (China). Cells were cultured in RPMI-1,640 mixed with 10% FCS in an atmosphere containing 5% CO<sub>2</sub>. Recombinant retroviruses carrying PLNCX2-vector or PLNCX2-TUFT1 were synthesized based on relevant instructions (Clontech). MDA-MB-231 or HCC1937 cells with Polybrene [8  $\mu$ g/mL (Sigma-Aldrich)] were infected these retroviruses and then were selectively isolated with G418 [750  $\mu$ g/mL (Calbiochem)].

### RNA Interference

The recombinant adenoviruses encoding 2 different short-hairpin RNAs (shRNAs), respectively specific for human TUFT1 were designed and prepared from company (GeneChem, Shanghai, China). TUFT1-shRNA#1: AGAGAATTTAGAGATG CAT; TUFT1-shRNA#2: GGTGGAGTATTACGGTAAAC. Lentiviruses were transfected into cells based on the relevant instructions. The ability of TUFT1 knockdown was assessed by real-time quantitative PCR and western blot. Cell lines with over 80% efficacy were considered stable. More than 80% of the cell transfection efficiency was considered stable.



## IHC Analyses

TUFT1 (dilution 1:100, Abcam, USA), RAC1-GTP (dilution 1:800, NewEast Bioscience, USA), were purchased. The experimental method was carried out and the expression of TUFT1 and RAC1-GTP was evaluated semi-quantitatively according to the criteria described previously (12, 13). The analysis was performed by two independent pathologists.

## Real-Time PCR

Total RNA was extracted by Trizol (Invitrogen) for reverse transcription, according to manufacturer's instructions (Invitrogen). TUFT1 expression was examined by Real-time PCR according to the criteria described previously (12, 13).

## Western Blot

The rabbit antibodies used to detect TUFT1, Rac1,  $\beta$ -catenin, Nanog, SOX2, and OCT4 were obtained from AbCam (Cambridge, UK). Their protein levels were examined by western blot according to the criteria described previously (12, 13).

## Wound Healing Assay

Following the manufacturer's recommendations, marker pen was used on the back of the 6-well plate, horizontal lines were evenly drawn, about  $2 \times 10^5$  cells were added, and the next day, the gun head was scratched. Cells were washed with PBS for 3 times and serum-free medium was added. Incubate in a  $37^\circ\text{C}$ , 5%  $\text{CO}_2$  incubator. Sample at 0, 8, 24 h and take photos.

## Invasion Assay

The required number of chambers were placed in a new 24-well plate and 500  $\mu\text{L}$  serum-free medium was added to the upper and lower chambers, respectively. The preparation of serum-free cell suspension is usually  $5 \times 10^4$  cells/well (24-well plate). 500  $\mu\text{L}$  cell suspension was added to the upper chamber and 750  $\mu\text{L}$  30% FBS medium was added to the lower chamber. The incubator was incubated at  $37^\circ\text{C}$  for 24 h. Hematoxylin and eosin (H&E) stained cells to the lower surface of the membrane. Photographs are taken under a microscope.

## Transwell Assay

Serum-free cell suspension was prepared and counted, usually  $5 \times 10^4$  cells/well (24-well plate). Carefully remove the culture medium in the upper chamber and add 100  $\mu\text{L}$  cell suspension. Add 600  $\mu\text{L}$  30% FBS culture medium in the lower chamber. The incubator was incubated at  $37^\circ\text{C}$  for 24 h. The chamber was fixed in 4% paraformaldehyde for half an hour. 1–2 drops of staining solution were used to stain and transfer cells to the lower surface of the membrane for 1–3 min. Photographs are taken under a microscope.

## Sphere Formation Assay

Cell trypsin of each experimental group in the logarithmic growth phase was digested, serum-free medium was resuspended, cell suspensions were made, and counted. The cell suspension was inoculated in the ultra-low adhesion 6-well plate culture plate at a density of 10,000–20,000 cells/wells, and 2 mL serum-free medium DMEM/F12 was added to each well. Will the good cells in under the condition of  $37^\circ\text{C}$  and 5%  $\text{CO}_2$ , every 2–3 days

in liquid, extend the every 6–8 days. Observe cell balling and morphology under microscope at any time.

## Apoptosis Assay

After infection, supernatant was collected from cell culture in each experimental group in a 5 ml centrifuge tube. The cells were washed once by D-Hanks, the cells were digested by trypsin, and the culture supernatant was terminated. The cells were collected in the same 5 ml centrifuge tube. Centrifuge 1,500 rpm for 5 min and discard the supernatant. The cells were washed with PBS and precipitated once, centrifuged at 1,500 rpm for 5 min, and the cells were collected. The cells were washed with  $1 \times$  binding buffer for once, centrifuged at 1,500 rpm for 5 min, and the cells were collected. Cell suspension of 100  $\mu\text{L}$  ( $1 \times 10^5$ – $1 \times 10^6$  cells) was taken and stained with PI complex dyeing liquor (0.5 mL) for 10–15 min at room temperature. Flow cytometry was used for detection.

## Rac1-GTP Pull-Down Assay

Cells were splitting in buffer including 25 mM HEPES, 1% NP40, 10% glycerin, 5 mM  $\text{MgCl}_2$ , 1 mM DTT, 100 mM NaCl, and protease inhibitors. The pyrolysate was cultured on ice for 5 min and centrifuged for 1 min with  $10,000 \times g$ . Post-nuclear supernatant was tested for pull-down analysis of 30  $\mu\text{g}$  GST-RBD (Rac1) pre-coated GSH beads in each case. The beads and supernatant were cultured in a table at  $4^\circ\text{C}$  for 15 min. The beads were washed with a solution buffer containing 0.01% NP40, boiled with SDS PAGE, and separated. The beads were analyzed by Western blotting as shown above. NSC23766 (obtained from Tocris Bioscience) was used to inhibit Rac1 activation.

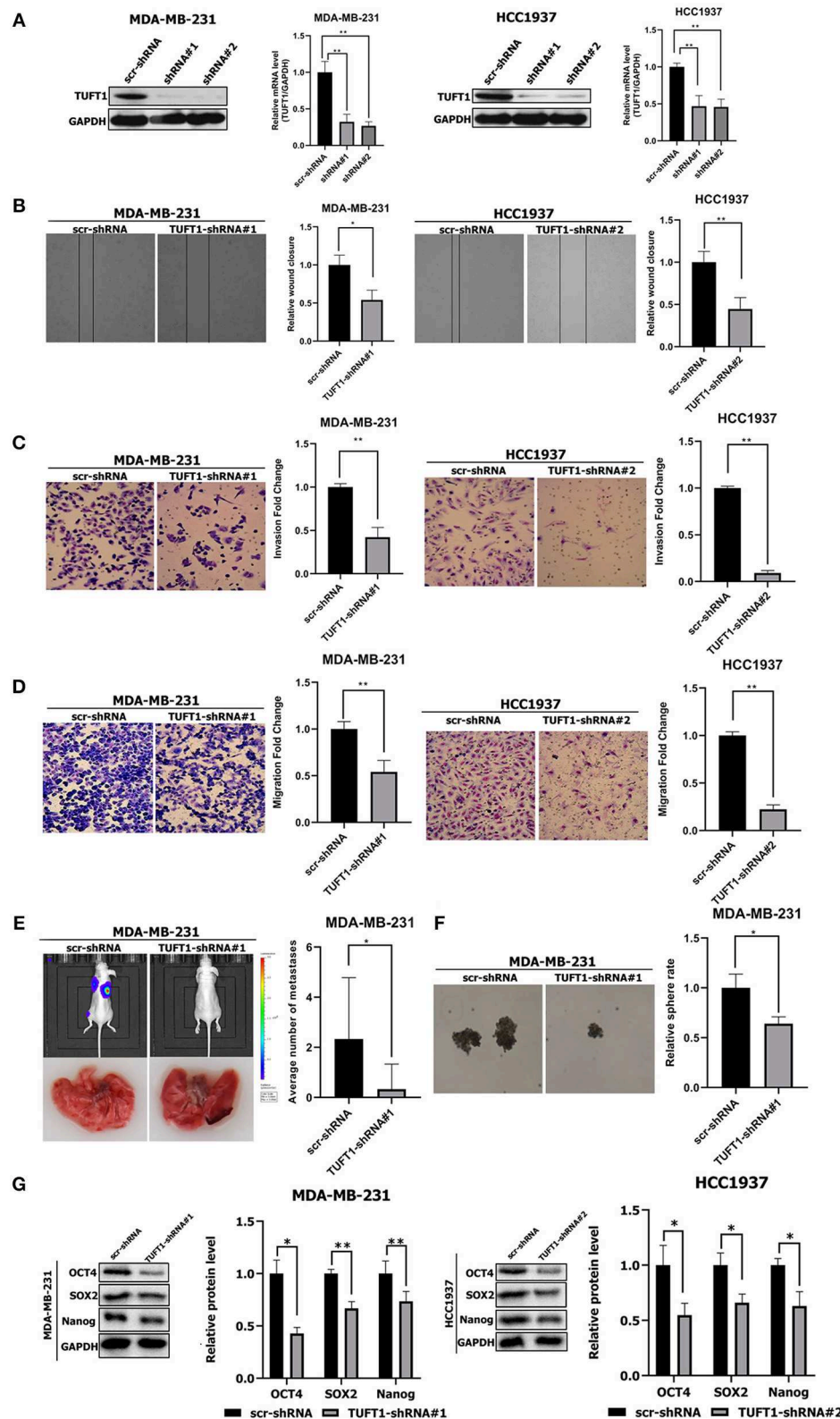
## Tumor Metastasis and Growth in Nude Mice

4–6-weeks female nude mice were obtained from the Shanghai Lingchang Biological Technology Ltd (Shanghai, China). The caudal vein was selectively injected into ShTUFT1—MDA-MB-231 cells. The nude mice were anesthetized by isoflurane gas using *in vivo* imaging instrument with gas anesthesia system. The mice were sacrificed at 10 weeks after treatment, and metastatic lung nodules were counted.

For the *in vivo* chemoresistance experiment, shTUFT1—MDA-MB-231 cells were injected into the flanks of nude mice (10 mice/group). Each group was divide randomly into two subgroups after 2 weeks that were either left untreated or received intraperitoneal injections of doxorubicin ( $4 \text{ mg kg}^{-1}$ ) every 5 days (three cycles), as previously described by Ghebeh et al. (30). Animal handling and research protocols were approved by the Ethics Committee of Affiliated Hospital of Hebei Engineering University.

## ONCOMINE Analysis

The mRNA levels of TUFT1 in BCs were determined through analysis of data from the ONCOMINE database ([www.oncomine.org](http://www.oncomine.org)). In our study, BC specimen data were compared with control datasets using student's *t*-test to examined the *p*-value. The fold change was defined as 2, and the *p*-value was set up at 0.01.



**FIGURE 1 |** Effects of TUFT1 on migration, invasiveness and stemness *in vivo* and *in vitro*. MDA-MB-231 and HCC1937 cells were infected with TUFT1-shRNA or scramble (scr)-shRNA. **(A)** TUFT1 protein and mRNA expression levels were reduced in MDA-MB-231 and HCC1937 cells with infection of adenovirus encoding (Continued)

**FIGURE 1 |** TUFT1-shRNA#1 and TUFT1-shRNA#2. Cell wound (B), invasiveness (C) and migration (D) were aberrant regulated after TUFT1 down-regulation in MDA-MB-231 or HCC1937 cells ( $n = 3$ ). (E) TUFT1 knockdown in MDA-MB-231 cells significantly reduced the number of lung metastatic nodules ( $n = 10$ ). (F) TUFT1 knockdown drastically reduced the number of mammary spheres formed by MDA-MB-231 cells ( $n = 3$ ). (G) The protein levels of Nanog, Sox2 and Oct4 were decreased by TUFT1 knockdown in both MDA-MB-231 and HCC1937 cells ( $n = 3$ ). Results are presented as means  $\pm$  SD. The statistical significance was assessed by student's *t*-test; \* $p < 0.05$ , \*\* $p < 0.01$ .

## Statistical Analysis

The SPSS 23.0 software was used for statistical analyses. Student's *t*-test and Pearson correlation test were used to compare the classified variables.  $p < 0.05$  was considered significant.

## RESULTS

### TUFT1 Regulates Metastasis and Stemness of TNBC Cells *in vitro* and *in vivo*

First, we performed TUFT1 knockdown in the HCC1937 and MDA-MB-231 TNBC cell lines using shRNA. Western blot and Real-time PCR revealed that TUFT1 protein and mRNA levels were prominently reduced in TUFT1-knockdown cells compared to control cells ( $p < 0.01$ , **Figure 1A**). Wound healing assays, invasion assays, and transwell assays were all used to examine the role of TUFT1 on the migration of TNBC cells. We found that TUFT1 down-regulation markedly reduced the migration of both TNBC cells compared to control cells, indicating that TUFT1 knockdown inhibits cell migratory ability ( $p < 0.05$ , **Figures 1B–D**).

To expand on our study *in vitro*, we next examined if TUFT1 could promote the metastasis in TNBC cells. ShTUFT1- MDA-MB-231 cells were injected into the caudal vein of nude mice. Then mice were sacrificed for quantitative analysis of lung metastatic nodules. Mice injected with ShTUFT1- MDA-MB-231 cells developed significantly fewer metastatic lung nodules than control mice ( $p < 0.05$ , **Figure 1E**). Taken together, results *in vitro* and *in vivo* reveal the metastatic potential of TUFT1 in TNBC cells.

CSCs play a key role in cancer metastasis (31, 32). We used a sphere formation assay to examine the role of TUFT1 on the stemness of TNBC cells. We found that TUFT1 knockdown drastically reduced the number of mammary spheres formed by MDA-MB-231 cells ( $p < 0.05$ , **Figure 1F**). Nanog, Sox2, and Oct4 play a central regulatory role in CSCs (26–28, 33, 34). We found that the Nanog, Sox2 and Oct4 levels were reduced by TUFT1 knockdown in both MDA-MB-231 and HCC1937 cells ( $p < 0.05$ , **Figure 1G**). These results reveal that TUFT1 is capable of significantly promoting CSC-like properties in TNBC cells.

### TUFT1 Promotes the Metastasis of TNBC Cells by Up-Regulating the Rac1/ $\beta$ -Catenin Pathway

To further investigate TUFT1-regulated metastasis in TNBC cells, we performed Rac1 activity assays following manipulation of TUFT1 expression levels. This revealed that knockdown of endogenous TUFT1 decreased Rac1-GTP levels in MDA-MB-231 cells ( $p < 0.01$ , **Figure 2A**), whereas TUFT1 overexpression increased Rac1-GTP levels in HCC1937 cells ( $p < 0.05$ ,

**Figure 2B**). These data indicate that TUFT1 promotes Rac1 activation in TNBC cells.

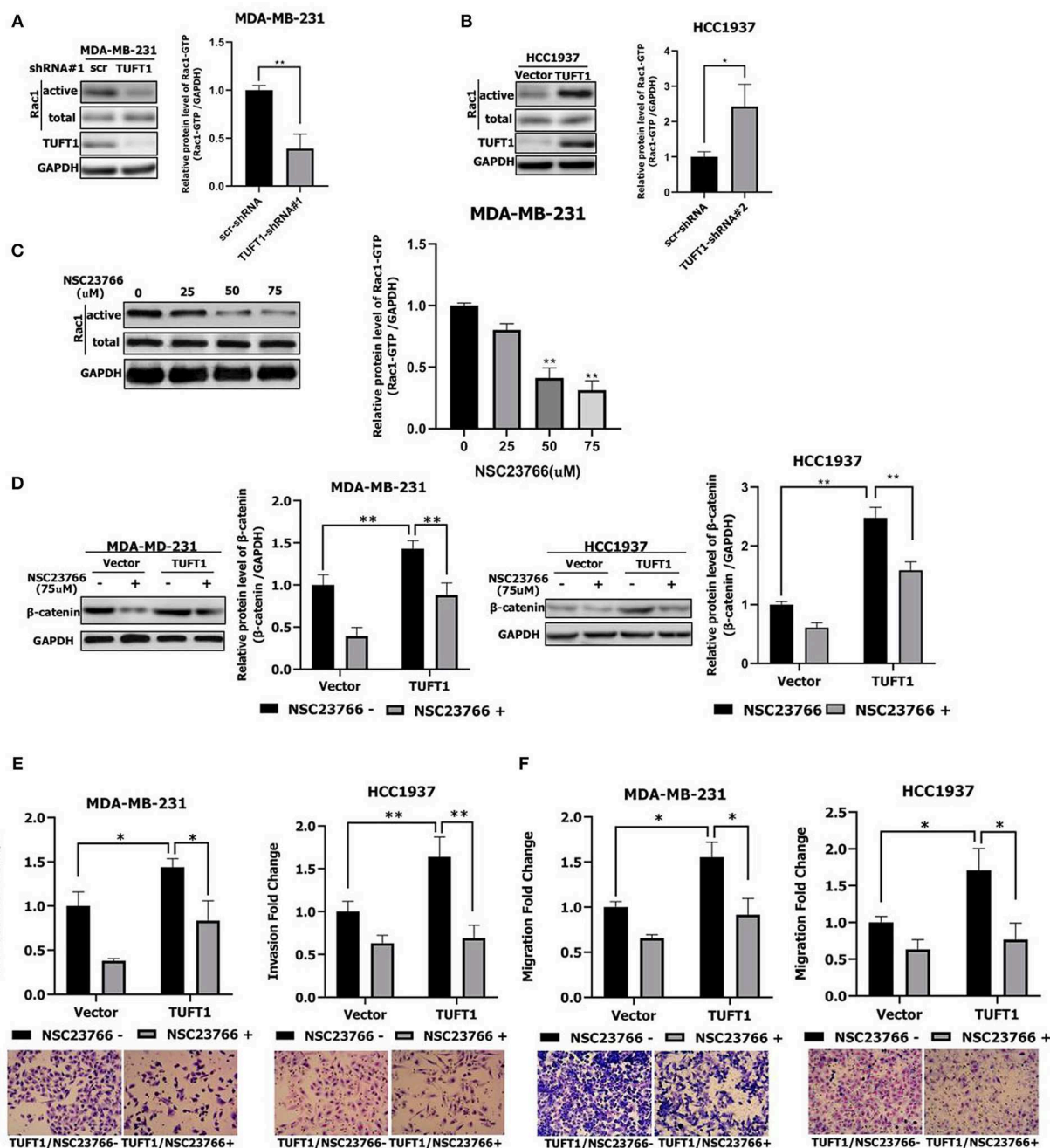
To investigate the potential role of Rac1 downstream of TUFT1, endogenous Rac1-GTP was inhibited using the Rac1 inhibitor NSC23766 (29) in TUFT1 overexpression TNBC cells. We confirmed that NSC23766-mediated inhibition of Rac1 was associated with a substantial reduction in its active form, Rac1-GTP (**Figure 2C**). The activation of Wnt/ $\beta$ -catenin pathway is related to the proliferation and metastasis of TNBC (35, 36). Interestingly, we found that  $\beta$ -catenin levels were significantly increased by TUFT1 overexpression in both TNBC cells ( $p < 0.01$ , **Figure 2D**). However, the increase in  $\beta$ -catenin induced by TUFT1 overexpression was significantly decreased by NSC23766 treatment in both TNBC cells, compared to the controls ( $p < 0.01$ , **Figure 2D**). Consistent with this, we observed that TUFT1-dependent TNBC cells metastasis was reversed in cells treated with NSC23766, as assessed by both invasion and transwell assays ( $p < 0.05$ , **Figures 2E,F**). In conclusion, these results suggest that Rac1 is necessary for TUFT1-dependent  $\beta$ -catenin activation and TNBC cells metastasis.

### TUFT1 Promotes the Stemness of TNBC Cells by Up-Regulating the Rac1 Signaling Pathway

To further investigate the regulation of TNBC cell stemness by TUFT1, we once again employed the Rac1 inhibitor NSC23766 (29) to inhibit endogenous Rac1-GTP in both TUFT1 overexpression TNBC cells. We found that Nanog, Sox2, and Oct4 levels were significantly increased by TUFT1 overexpression in both TNBC cells ( $p < 0.05$ , **Figure 3A**). However, the TUFT1-induced increase in Nanog, Sox2, and Oct4 was significantly decreased by NSC23766 treatment in both TNBC cells, compared to the corresponding controls ( $p < 0.05$ , **Figure 3A**). Consistent with this, we observed that NSC23766 treatment in MDA-MB-231 cells impaired TUFT1-dependent CSC-like properties, as assessed by the sphere formation assay ( $p < 0.01$ , **Figure 3B**).

### TUFT1 Inhibits Chemotherapy-Mediated Apoptosis in TNBC Cells by Targeting the Rac1/ $\beta$ -Catenin Signaling Pathway

ONCOMINE data showed that TUFT1 mRNA levels were significantly lower in epirubicin/docetaxel responder BC samples than epirubicin/docetaxel non-responder BC samples ( $p = 0.031$ , **Figure 4A**). To evaluate whether TUFT1 expression can directly contribute to resistance to chemotherapy in TNBC, we used MDA-MB-231-shTUFT1 cells (or control MDA-MB-231 cells) in a xenograft tumor model. IHC staining revealed that the tumors formed by the MDA-MB-231-TUFT1-shRNA

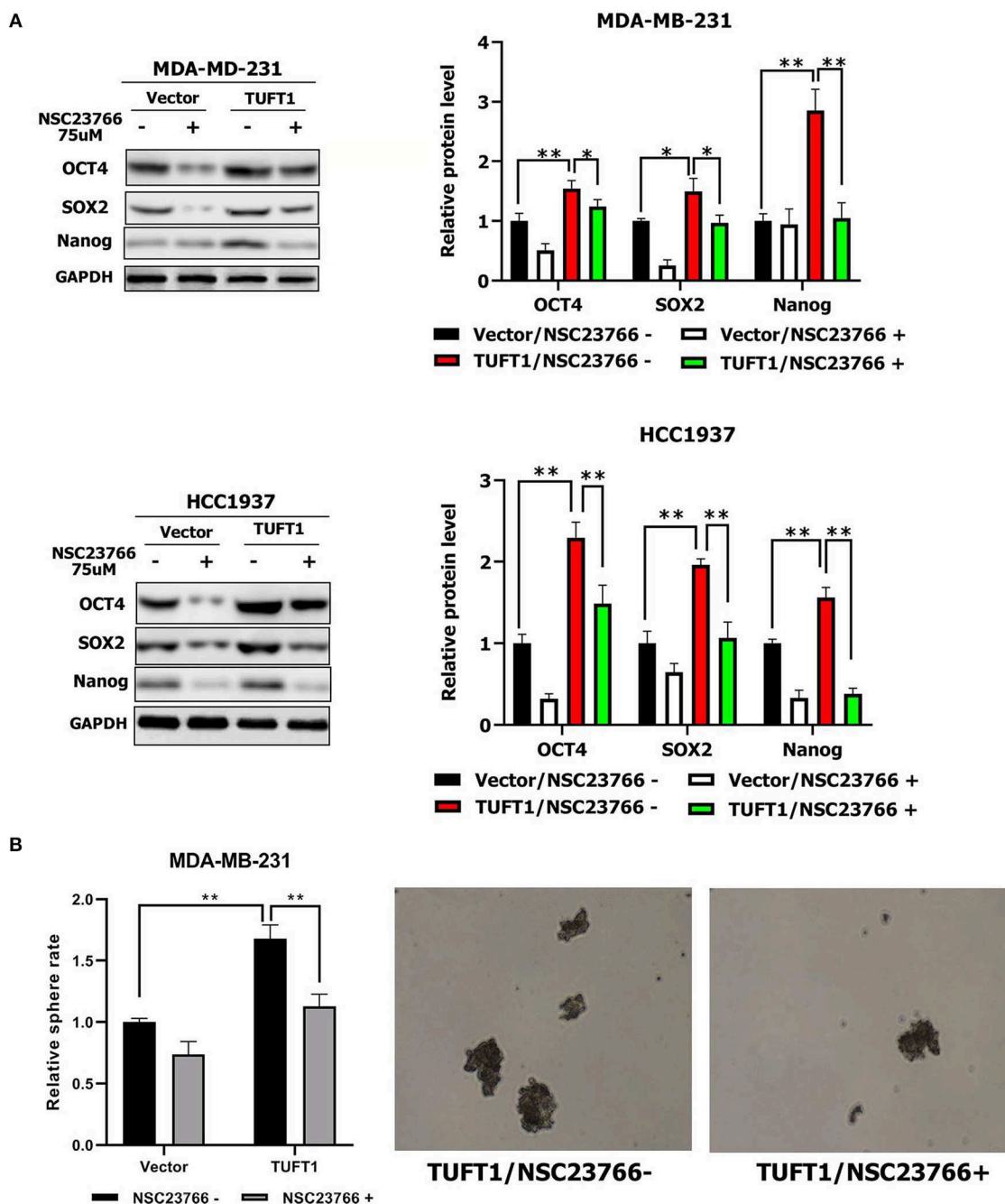


**FIGURE 2 |** Role of TUFT1 in TNBC migration and invasiveness by promoting the Rac1/β-catenin signaling pathway. **(A)** Down-regulation of TUFT1 by shRNA dramatically decreased the protein level of Rac1-GTP in MDA-MB-231 cells ( $n = 3$ ). **(B)** Overexpression of TUFT1 increased the protein level of Rac1-GTP in HCC1937 cells ( $n = 3$ ). **(C)** Western blot showed effect of Rac1 inhibitor NSC23766 on Rac1-GTP in MDA-MB-231 cells ( $n = 3$ ). **(D)** Down-regulation of Rac1-GTP treated with NSC23766 decreased expression of β-catenin in TUFT1—MDA-MB-231 and TUFT1—HCC1937 cells ( $n = 3$ ). Down-regulation of Rac1-GTP treated with NSC23766 decreased cell invasiveness **(E)** and migration in **(F)** in TUFT1—MDA-MB-231 and TUFT1—HCC1937 cells ( $n = 3$ ). Results are presented as means  $\pm$  SD. The statistical significance was assessed by student's  $t$ -test; \* $p < 0.05$ , \*\* $p < 0.01$ .

cells had lower TUFT1 expression than those formed by the control cells (**Figure 4B**). The size of tumors formed by TUFT1-positive cells was slightly reduced by doxorubicin treatment ( $p > 0.05$ , **Figure 4C**), whereas the size of the tumors formed by

TUFT1-negative cells was significantly reduced by doxorubicin treatment ( $p < 0.05$ , **Figure 4C**). These results show that the expression of TUFT1 is directly related to the increase of chemoresistance.



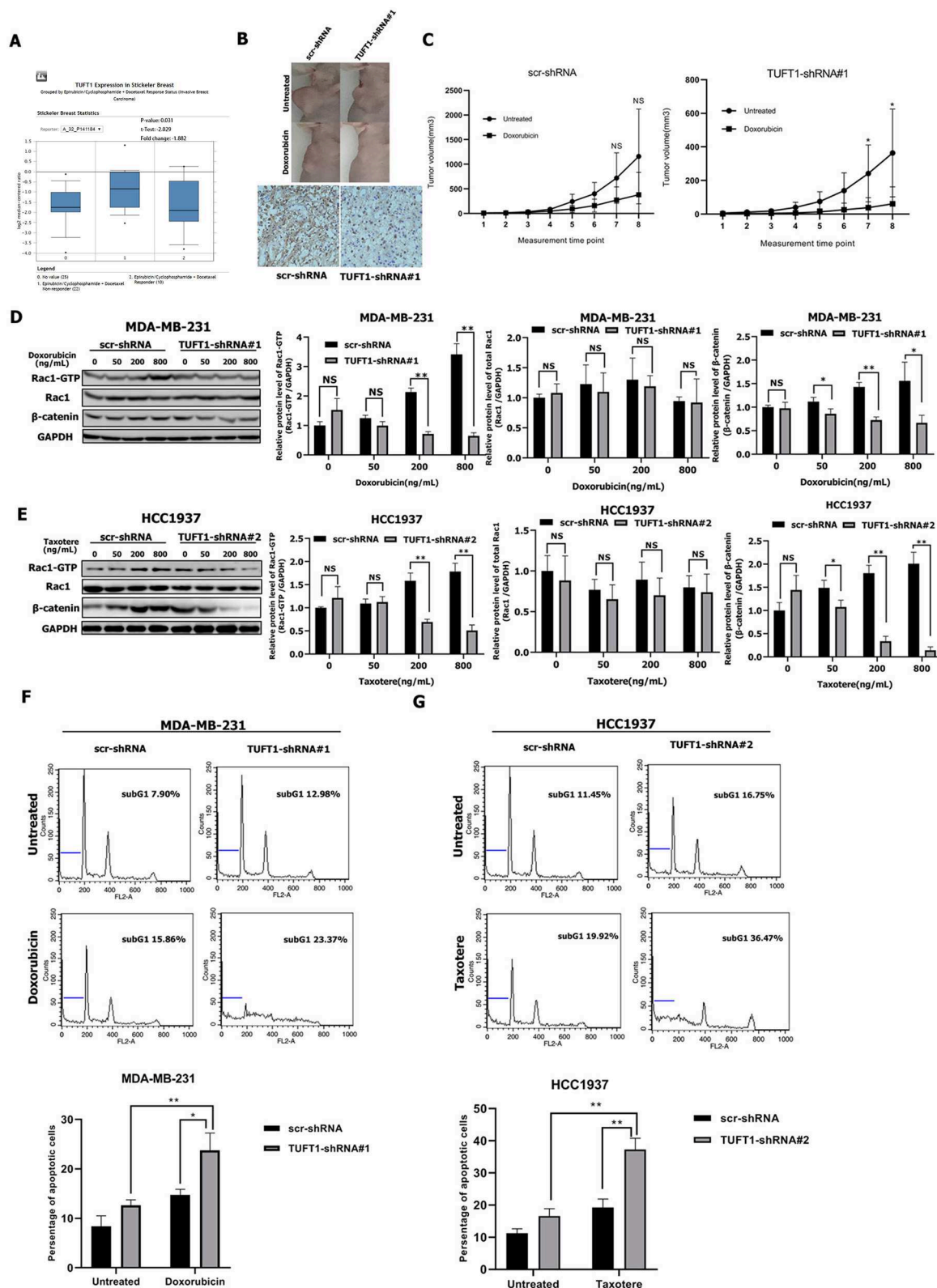


**FIGURE 3 |** Role of TUFT1 in TNBC stemness by promoting the Rac1 signaling pathway. **(A)** Down-regulation of Rac1-GTP treated with NSC23766 decreased expression of Nanog, Sox2, and Oct4 levels in TUFT1-MDA-MB-231 and TUFT1-HCC1937 cells ( $n = 3$ ). **(B)** Down-regulation of Rac1-GTP treated with NSC23766 decreased number of mammary spheres formed by TUFT1-MDA-MB-231 cells ( $n = 3$ ). Results are presented as means  $\pm$  SD. The statistical significance was assessed by student's  $t$ -test;  $p < 0.05$ ,  $**p < 0.01$ .

We next wondered whether TUFT1 confers resistance to chemotherapy in TNBC cells via the Rac1/ $\beta$ -catenin signaling pathway. Treatment of TUFT1-negative MDA-MB-231 cells with doxorubicin and HCC1937 cells with taxotere induced a decrease in both Rac1-GTP and  $\beta$ -catenin levels in a dose-dependent manner (Figures 4D,E). The protein levels of

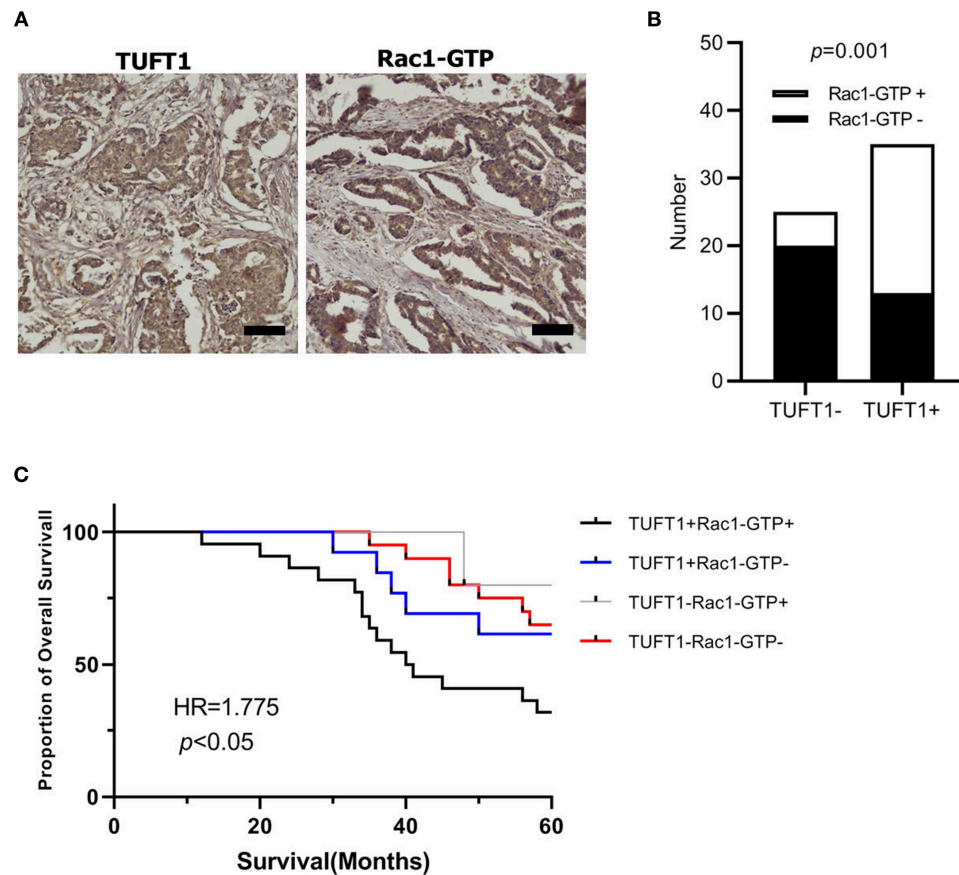
Rac1-GTP and  $\beta$ -catenin were significantly lower in TUFT1-negative cells than in TUFT1-positive cells following treatment with corresponding dose of doxorubicin and taxotere ( $p < 0.05$ , Figures 4D,E). However, the level of total Rac1 protein was unchanged (Figures 4D,E). Furthermore, we observed a significantly higher level of apoptosis in TUFT1-negative cells





**FIGURE 4 |** TUFT1-knockdown TNBC cells are more sensitive to doxorubicin and taxotere. **(A)** TUFT1 mRNA expression was lower in epirubicin/docetaxel responder BC samples by ONCOMINE analysis. **(B)** scr-shRNA- and TUFT1-shRNA-MDA-MB-231 cells were injected into nude mice as described in the Materials and Methods. (Continued)

**FIGURE 4 |** The tumor volumes were measured following treatment with or without doxorubicin ( $n = 5$ ). Representative images showing tumor formed in nude mice after injection with scr-shRNA- or TUFT1-shRNA cells and IHC staining of TUFT1 in tumor tissues. **(C)** Tumor volumes in four groups. **(D,E)** Western blot showing the expression levels of Rac1-GTP, Rac1 and  $\beta$ -catenin in scr-shRNA- and TUFT1-shRNA-MDA-MB-231 cells following treatment with various doses of doxorubicin or TUFT1-shRNA-HCC1937 cells following treatment with various doses of taxotere for 24 h ( $n = 3$ ). **(F,G)** Apoptotic cell death was detected by PI single staining method following treatment of scr-shRNA- and TUFT1-shRNA-MDA-MB-231 cells without or with 200 ng ml<sup>-1</sup> of doxorubicin or TUFT1-shRNA-HCC1937 cells without or with 200 ng ml<sup>-1</sup> of taxotere for 24 h ( $n = 3$ ). Numbers in the subG1 phase (blue bar) represent the percentage of apoptosis. Results are presented as means  $\pm$  SD. The statistical significance was assessed by student's t-test; \* $p < 0.05$ , \*\* $p < 0.01$ .



**FIGURE 5 |** The expression of TUFT1 and Rac1-GTP in 60 TNBC patients who had received anthracycline/taxanes chemotherapy after surgery. **(A)** Show the positive expression of TUFT1 and Rac1-GTP in serial sections. **(B)** Rac1-GTP positively correlated with TUFT1 expression in the TNBC samples. **(C)** Kaplan–Meier survival curves showing survival in 60 patients who received chemotherapy blotted in relation to TUFT1 and Rac1-GTP expression. Survival curves showing the poor overall survival in patients with tumors co-expressing TUFT1 and Rac1-GTP that received chemotherapy.

than in TUFT1-positive cells following treatment with 200 ng/mL doxorubicin or taxotere ( $p < 0.05$ , **Figures 4E,G**). These results indicate that TUFT1 may confer resistance to chemotherapy in TNBC cells by promoting cell apoptosis via the Rac1/ $\beta$ -catenin signaling pathway.

### TUFT1 and Rac1-GTP Expression Positively Correlate and Predict Poor Prognosis Following Treatment With Chemotherapy in TNBC

We next studied the clinical correlation of TUFT1 and Rac1-GTP using 60 TNBC specimens from patients who had received

anthracycline followed by taxanes chemotherapy after surgery. Examples of positive expression of TUFT1 and Rac1-GTP in serial sections are presented in **Figure 5A**. The level of TUFT1 protein was positively correlated with tumor size, histological grade and axillary lymph node metastasis ( $p = 0.010$ ,  $p = 0.005$ , and  $p = 0.010$ , respectively, **Table 1**). The level of Rac1-GTP protein positively correlated with TUFT1 expression in the TNBC samples ( $p = 0.001$ , **Table 2; Figure 5B**). We divided the patients into four groups according to the TUFT1 and Rac1-GTP expression in the TNBC samples. Our patient follow-up analysis showed that a total of 27 of 60 patients died, and the 5-years overall survival rate was 55.0%. Fourteen of the 22 patients with tumors co-expressing TUFT1 and Rac1-GTP were dead,

**TABLE 1** | The relationship between TUFT1 expression and the clinicopathological factors in TNBC patients who have received chemotherapy ( $n = 60$ ).

Variable	<i>n</i>	TUFT1 <sup>-</sup>	TUFT1 <sup>+</sup>	<i>p</i> -variable
Age				0.695
≥40	49	21	28	
<40	11	4	7	
Tumor size				0.010
T1	16	11	5	
T2–4	44	14	30	
Histological grades				0.005
I, II	28	17	11	
III	32	8	24	
Lymph node metastasis				0.010
–	18	12	6	
+	42	13	29	

“+,” positive; “–,” negative.

**TABLE 2** | Correlations between expression of TUFT1 and Rac1-GTP.

Variable	<i>n</i>	Rac1-GTP <sup>-</sup>	Rac1-GTP <sup>+</sup>	<i>p</i> -variable
TUFT1				0.001
–	25	20	5	
+	35	13	22	

“+,” positive; “–,” negative.

and this group displayed the lowest 5 years survival than other groups (log-rank test,  $p < 0.05$ , Hazard Ratio = 1.775, 95% CI of ratio = 0.986–3.195, **Figure 5C**). Therefore, TUFT1 and Rac1-GTP expression positively correlate and predict patient prognosis following treatment with chemotherapy in TNBC.

## DISCUSSION

To our knowledge, this is the first systematic study on the functional mechanism of TUFT1 mediated metastasis and stemness in TNBC. Zhou et al. (11) reported that TUFT1 overexpression promoted the metastasis of pancreatic cancer cells, and affected the expression of a number of epithelial-mesenchymal transformation-related proteins. They suggested that TUFT1 may affect HIF1 by influencing the expression of members of the Snail signaling pathway, which regulates epithelial-mesenchymal transition. Kawasak et al. (14) found that TUFT1 may be activated by the AKT/mTOR pathway to regulate tumor proliferation and metastasis. Compared to cells of other breast cancer subtypes, basal mesenchymal-like TNBC cells display increased migration, invasion, and metastatic potential (37). In this study, we found that TUFT1 promotes the metastasis of TNBC cells both *in vitro* and *in vivo*. CSCs have high tumorigenic capacity and are important features of new tumors (secondary and third foci) at locations other than those of the original tumor (38, 39). Here, we propose for the first time that TUFT1 can regulate the stemness of TNBC cells. TUFT1 knockdown in TNBC cells reduced the number of mammary

spheres and stemness-associated molecules. These results reveal that TUFT1 may promote the metastasis of TNBC cells by up-regulating their stem capacity.

Rac1, a member of the Rac subfamily of small GTPases, has its forms of active GTP-bound and inactive GDP-bound. Rac1 activity plays roles in the regulation of proliferation, differentiation, apoptosis, cell movement, and adhesion. Moreover, Rac1 has been shown to have an important role in tumor cell migration (40). Rac1-GTP interacts with different downstream effector molecules, thus affecting tumor invasion and metastasis (41).  $\beta$ -catenin, a target molecule of Rac1, is a key regulator of cell proliferation and metastasis (42, 43).  $\beta$ -catenin is a multi-gene nuclear transcription target. It can regulate the proliferation and metastasis of cancer cells (44, 45). Rac1 gene regulates  $\beta$ -catenin and locates its nucleus at the promoter TCF3/4 of target gene (46). Furthermore, active/inactive Rac1 state was shown to direct Rac1- $\beta$ -catenin complex to the nucleus in CRC cells (47). De et al. (36) demonstrated that Rac1 was activated by cascade of  $\beta$ -catenin-Tiam1/vav2 as downstream target of Wnt/ $\beta$ -catenin pathway activation during TNBC metastasis. However, our results show that TUFT1 can promote the metastasis of TNBC cells by activating Rac1 in the Rac1/ $\beta$ -catenin signaling pathway, suggesting that the TUFT1/Rac1/ $\beta$ -catenin axis may regulate metastasis in TNBC. NSC23766 reduces total  $\beta$ -catenin in CRC cells, thus demonstrating that Rac1 regulates stemness in CRC by activating Wnt/ $\beta$ -catenin signaling (29). Our study further implicates Rac1 and its downstream target  $\beta$ -catenin as critical molecules in the regulation of stemness in TNBC downstream of TUFT1. Our study identifies the TUFT1/Rac1/ $\beta$ -catenin axis as a novel regulator of metastasis and stemness in TNBC. However, how TUFT1 specifically regulates Rac1 expression, in a recent study, Kawasak et al. (14) found that TUFT1 activated the mTORC1 signaling pathway by regulating the Rab GTPase, and that the interaction of TUFT1 and RabGAP1 mediated intracellular lysosome localization and vesicle transport in BC cells, while Rac1 is the substrate of mTOR. In addition, through high-throughput differential gene screening, TUFT1 was found to be associated with Rab5 and Rac1 (13). Rab5 is responsible for regulating the early stage of vesicle transport. Once activated, Rab5 recruits a number of interacting proteins, such as Rac1 and Tiam1, which play an important role in tumor metastasis (48, 49). Díaz et al. (50) found that Rab5 activation could recruit Tiam1 around the endosome, thereby leading to the activation of Rac1. Based on this, we hypothesize that TUFT1 may initiate vesicle transport through activating Rab5, thereby affecting downstream Rac1 expression. So, regulatory processes may be complex, the relationship between TUFT1 and Rac1 needs further study.

As endocrine therapy or HER2 targeted therapy is ineffective for TNBC patients. Chemotherapy is the most effective treatment at present. In addition, more than 50% of TNBCs were resistant to adjuvant chemotherapy. Because of chemotherapeutic resistance, patients often have relapse and metastasis (51, 52). In 2015, experts at St. Gallen agreed to recommend anthracyclines and taxanes as the main adjuvant chemotherapeutic drugs for TNBC. However, the use of platinum antineoplastic drugs

is still controversial (53, 54). Here, we demonstrated that TUFT1 knockdown can reverse doxorubicin resistance in a TNBC xenograft tumor model. Meanwhile, TUFT1 suppression conferred sensitivity to chemotherapy and increased cell apoptosis via inhibition of Rac1/ $\beta$ -catenin signaling in TNBC cells. The mechanism of Rac1-mediated chemoresistance has been studied in several tumors (23, 55–57). We have found in previous studies that TUFT1 can inhibit the apoptosis of BC cells and the activation of Caspase 3 (13). Rac1 can regulate the DNA damage response, drug-induced apoptosis, and tumor metastasis by activating a number of stress-activated kinases, such as JNK and p38 kinase, which can regulate the activation of Caspase 3 (58, 59). In addition, dual specificity phosphatase-1 (DUSP1) can dephosphorylate all three family members of MAPK (ERK1/2, JNK1/2, p38 MAPK), which play a negative regulatory role in MAPK signaling pathway (60, 61). DUSP1 mediates breast cancer proliferation and chemotherapy resistance by inhibiting JNK pre-apoptotic signaling pathway (62, 63). TUFT1 can regulate DUSP1 expression in our previous studies (13), therefore, we consider whether there is a link between TUFT1/Rac1 pathway and DUSP1 to regulate downstream MAPK pathways, or whether TUFT1 directly mediates DUSP1 bypass signal to regulate apoptosis and chemoresistance of BC cells. This requires further study. CSCs as a target is a promising method for reversing chemoresistance, and activated Wnt/ $\beta$ -catenin pathway also can inhibit apoptosis of BC cells and confer the stemness of BC cells and lead to chemoresistance (64–66). Therefore, these results suggest that the TUFT1/Rac1/ $\beta$ -catenin axis can at least partially inhibit TNBC cells apoptosis and then promote doxorubicin/taxotere resistance in TNBC. Moreover, TUFT1 expression positively correlates with Rac1-GTP, and co-expression of TUFT1 and Rac1-GTP predicts poor patient prognosis in TNBC following adjuvant doxorubicin/taxotere treatment. Thus, TUFT1 may be a potential novel clinical therapy target for reversing chemoresistance in TNBC.

## CONCLUSIONS

In summary, we first systematic study on the functional mechanism of TUFT1 mediated metastasis, stemness and chemoresistance in TNBC. Our results find that TUFT1 can promotes the metastasis and stemness of TNBC cells via the RAC1/ $\beta$ -catenin pathway, meanwhile, TUFT1 could increase TNBC resistance to chemotherapy induced by RAC1/ $\beta$ -catenin

pathway. Therefore, our findings suggest that TUFT1 may provide a potential target for more effective treatment of TNBC. The mechanism of TUFT1 regulating Rac1 and the mechanism of TUFT1 mediating metastatic and apoptotic bypass signaling in TNBC cells need to be further explored.

## DATA AVAILABILITY

The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher.

## ETHICS STATEMENT

This study was carried out in accordance with the recommendations of ICMJE of guidelines, Ethics Committee of Affiliated Hospital of Hebei University of Engineering with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Ethics Committee of Affiliated Hospital of Hebei University of Engineering. This study was carried out in accordance with the recommendations of International Association of Veterinary Editors guidelines, Ethics Committee of Affiliated Hospital of Hebei University of Engineering. The protocol was approved by the Ethics Committee of Affiliated Hospital of Hebei University of Engineering.

## AUTHOR CONTRIBUTIONS

WL and BC conceived the study and provided the project direction. WL guided and performed the experiments, analyzed the data, and wrote the manuscript until the final submission version. WL, GC, LS, YZ, and JHa completed the cell experiments. WL, GC, YD, JHe, and SS assisted in performing the animal experiments.

## FUNDING

This research was supported in part by the Science and Technology Research and Development Project of Handan (Grant No. 1823208029ZC), the National Natural Science Foundation of China (Grant No. 31601142), and the Key Science and Technology Research Program of Hebei Provincial Department of Health (Grant No. 20190962).

## REFERENCES

- Harbeck N, Gnant M. Breast cancer. *Lancet*. (2017) 389:1134–50. doi: 10.1016/S0140-6736(16)31891-8
- Bianchini G, Balko JM, Mayer IA, Sanders ME, Gianni L. Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease. *Nat Rev Clin Oncol*. (2016) 13:674–90. doi: 10.1038/nrclinonc.2016.66
- Kumar P, Aggarwal R. An overview of triple-negative breast cancer. *Arch Gynecol Obstetr*. (2016) 293:247–69. doi: 10.1007/s00404-015-3859-y
- Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. *Nature*. (2000) 406:747–52. doi: 10.1038/35021093
- Thike AA, Cheok PY, Jara-Lazaro AR, Tan B, Tan P, Tan PH. Triple-negative breast cancer: clinicopathological characteristics and relationship with basal-like breast cancer. *Modern Pathol*. (2010) 23:123–33. doi: 10.1038/modpathol.2009.145
- Kast K, Link T, Friedrich K, Petzold A, Niedostatek A, Schoffer O, et al. Impact of breast cancer subtypes and patterns of metastasis on outcome. *Breast Cancer Res Treat*. (2015) 150:621–9. doi: 10.1007/s10549-015-3341-3



7. Kennecke H, Yerushalmi R, Woods R, Cheang MC, Voduc D, Speers CH, et al. Metastatic behavior of breast cancer subtypes. *J Clin Oncol*. (2010) 28:3271–7. doi: 10.1200/JCO.2009.25.9820
8. Dent R, Trudeau M, Pritchard KI, Hanna WM, Kahn HK, Sawka CA, et al. Triple-negative breast cancer: clinical features and patterns of recurrence. *Clin Cancer Res*. (2007) 13:4429–34. doi: 10.1158/1078-0432.CCR-06-3045
9. Bayraktar S, Gluck S. Molecularly targeted therapies for metastatic triple-negative breast cancer. *Breast Cancer Res Treat*. (2013) 138:21–35. doi: 10.1007/s10549-013-2421-5
10. Deutsch D, Palmon A, Fisher LW, Kolodny N, Termine JD, Young MF. Sequencing of bovine enamelin (“tuftelin”) a novel acidic enamel protein. *J Biol Chem*. (1991) 266:16021–8.
11. Zhou B, Zhan H, Tin L, Liu S, Xu J, Dong Y, et al. TUFT1 regulates metastasis of pancreatic cancer through HIF1-Snail pathway induced epithelial-mesenchymal transition. *Cancer Lett*. (2016) 382:11–20. doi: 10.1016/j.canlet.2016.08.017
12. Weiguang Liu JD, Min Zhao, Lei Zhang, Zining Jin, Bo Chen. Clinical implications of TUFT1 protein expression and correlation with RelA protein in breast cancer. *Int J Clin Exp Pathol*. (2017) 10:6544–51. doi: 10.1038/nrclinonc.2016.666544-6551
13. Liu W, Zhang L, Jin Z, Zhao M, Li Z, Chen G, et al. TUFT1 is expressed in breast cancer and involved in cancer cell proliferation and survival. *Oncotarget*. (2017) 8:74962–74. doi: 10.18632/oncotarget.20472
14. Kawasaki N, Isogaya K, Dan S, Yamori T, Takano H, Yao R, et al. TUFT1 interacts with RABGAP1 and regulates mTORC1 signaling. *Cell Disc*. (2018) 4:1. doi: 10.1038/s41421-017-0001-2
15. Ridley AJ. Rho GTPase signalling in cell migration. *Curr Opin Cell Biol*. (2015) 36:103–12. doi: 10.1016/j.ccb.2015.08.005
16. Hall A. Rho GTPases and the actin cytoskeleton. *Science*. (1998) 279:509–14. doi: 10.1126/science.279.5350.509
17. Orgaz JL, Herraiz C, Sanz-Moreno V. Rho GTPases modulate malignant transformation of tumor cells. *Small GTPases*. (2014) 5:e29019. doi: 10.4161/sgtp.29019
18. Burbelo P, Wellstein A, Pestell RG. Altered Rho GTPase signaling pathways in breast cancer cells. *Breast Cancer Res Treat*. (2004) 84:43–8. doi: 10.1023/B:BREA.0000018422.02237.f9
19. Baugher PJ, Krishnamoorthy L, Price JE, Dharmawardhane SF. Rac1 and Rac3 isoform activation is involved in the invasive and metastatic phenotype of human breast cancer cells. *Breast Cancer Res*. (2005) 7:R965–74. doi: 10.1186/bcr1329
20. Fritz G, Brachetti C, Bahlmann F, Schmidt M, Kaina B. Rho GTPases in human breast tumours: expression and mutation analyses and correlation with clinical parameters. *Br J Cancer*. (2002) 87:635–44. doi: 10.1038/sj.bjc.6600510
21. Zhang B, Zhang Y, Dagher MC, Shacter E. Rho GDP dissociation inhibitor protects cancer cells against drug-induced apoptosis. *Cancer Res*. (2005) 65:6054–62. doi: 10.1158/0008-5472.CAN-05-0175
22. Kregar BT, Johansen ER, Cerione RA, Antonyak MA. The enrichment of survivin in exosomes from breast cancer cells treated with paclitaxel promotes cell survival and chemoresistance. *Cancers*. (2016) 8:111. doi: 10.3390/cancers8120111
23. Hofbauer SW, Krenn PW, Ganghammer S, Asslaber D, Pichler U, Oberascher K, et al. Tiam1/Rac1 signals contribute to the proliferation and chemoresistance, but not motility, of chronic lymphocytic leukemia cells. *Blood*. (2014) 123:2181–8. doi: 10.1182/blood-2013-08-523563
24. Ikram M, Lim Y, Baek SY, Jin S, Jeong YH, Kwak JY, Yoon S. Co-targeting of Tiam1/Rac1 and Notch ameliorates chemoresistance against doxorubicin in a biomimetic 3D lymphoma model. *Oncotarget*. (2018) 9:2058–75. doi: 10.18632/oncotarget.23156
25. Velaithan R, Kang J, Hirpara JL, Loh T, Goh BC, Le Bras M, et al. The small GTPase Rac1 is a novel binding partner of Bcl-2 and stabilizes its antiapoptotic activity. *Blood*. (2011) 117:6214–26. doi: 10.1182/blood-2010-08-301283
26. Gangemi RM, Griffiro F, Marubbi D, Perera M, Capra MC, Malatesta P, et al. SOX2 silencing in glioblastoma tumor-initiating cells causes stop of proliferation and loss of tumorigenicity. *Stem Cells*. (2009) 27:40–8. doi: 10.1634/stemcells.2008-0493
27. Wen K, Fu Z, Wu X, Feng J, Chen W, Qian J. Oct-4 is required for an antiapoptotic behavior of chemoresistant colorectal cancer cells enriched for cancer stem cells: effects associated with STAT3/Survivin. *Cancer Lett*. (2013) 333:56–65. doi: 10.1016/j.canlet.2013.01.009
28. Zhang J, Espinoza LA, Kinders RJ, Lawrence SM, Pfister TD, Zhou M, et al. NANOG modulates stemness in human colorectal cancer. *Oncogene*. (2013) 32:4397–405. doi: 10.1038/ncr.2012.461
29. Rao J, Zhou ZH, Yang J, Shi Y, Xu SL, Wang B, et al. Semaphorin-3F suppresses the stemness of colorectal cancer cells by inactivating Rac1. *Cancer Lett*. (2015) 358:76–84. doi: 10.1016/j.canlet.2014.12.040
30. Ghebeh H, Al-Khaldi S, Olabi S, Al-Dhfyhan A, Al-Mohanna F, Barnawi R, et al. Fascin is involved in the chemotherapeutic resistance of breast cancer cells predominantly via the PI3K/Akt pathway. *Br J Cancer*. (2014) 111:1552–61. doi: 10.1038/bjc.2014.453
31. Nguyen LV, Vanner R, Dirks P, Eaves CJ. Cancer stem cells: an evolving concept. *Nat Rev Cancer*. (2012) 12:133–43. doi: 10.1038/nrc3184
32. Shukla S, Meeran SM. Epigenetics of cancer stem cells: pathways and therapeutics. *Biochim Biophys Acta*. (2014) 1840:3494–502. doi: 10.1016/j.bbagen.2014.09.017
33. Yang F, Zhang J, Yang H. OCT4, SOX2, NANOG positive expression correlates with poor differentiation, advanced disease stages, worse overall survival in HER2<sup>+</sup> breast cancer patients. *OncoTargets Ther*. (2018) 11:7873–81. doi: 10.2147/OTT.S173522
34. Tulsyan S, Agarwal G, Lal P, Mittal B. Significant association of combination of OCT4, NANOG, SOX2 gene polymorphisms in susceptibility and response to treatment in North Indian breast cancer patients. *Cancer Chemother Pharmacol*. (2014) 74:1065–78. doi: 10.1007/s00280-014-2588-4
35. Wei Z, Shaikh ZA. Cadmium stimulates metastasis-associated phenotype in triple-negative breast cancer cells through integrin and beta-catenin signaling. *Toxicol Appl Pharmacol*. (2017) 328:70–80. doi: 10.1016/j.taap.2017.05.017
36. De P, Carlson JH, Jepperson T, Willis S, Leyland-Jones B, Dey N. RAC1 GTPase signals Wnt-beta-catenin pathway mediated integrin-directed metastasis-associated tumor cell phenotypes in triple negative breast cancers. *Oncotarget*. (2017) 8:3072–103. doi: 10.18632/oncotarget.13618
37. Humphries B, Wang Z, Oom AL, Fisher T, Tan D, Cui Y, et al. MicroRNA-200b targets protein kinase Calpha and suppresses triple-negative breast cancer metastasis. *Carcinogenesis*. (2014) 35:2254–63. doi: 10.1093/carcin/bgu133
38. Magee JA, Piskounova E, Morrison SJ. Cancer stem cells: impact, heterogeneity, uncertainty. *Cancer Cell*. (2012) 21:283–96. doi: 10.1016/j.ccr.2012.03.003
39. Pang R, Law WL, Chu AC, Poon JT, Lam CS, Chow AK, et al. A subpopulation of CD26<sup>+</sup> cancer stem cells with metastatic capacity in human colorectal cancer. *Cell Stem Cell*. (2010) 6:603–15. doi: 10.1016/j.stem.2010.04.001
40. Lobos-Gonzalez L, Aguilar L, Diaz J, Diaz N, Urrea H, Torres VA, et al. E-cadherin determines Caveolin-1 tumor suppression or metastasis enhancing function in melanoma cells. *Pigment Cell Mel Res*. (2013) 26:555–70. doi: 10.1111/pcmr.12085
41. Palamidessi A, Frittoli E, Garre M, Faretta M, Mione M, Testa I, et al. Endocytic trafficking of Rac is required for the spatial restriction of signaling in cell migration. *Cell*. (2008) 134:135–47. doi: 10.1016/j.cell.2008.05.034
42. Hall A. The cytoskeleton and cancer. *Cancer Metast Rev*. (2009) 28:5–14. doi: 10.1007/s10555-008-9166-3
43. Hughes SC, Fehon RG. Understanding ERM proteins—the awesome power of genetics finally brought to bear. *Curr Opin Cell Biol*. (2007) 19:51–6. doi: 10.1016/j.ccb.2006.12.004
44. Goldstein B, Takeshita H, Mizumoto K, Sawa H. Wnt signals can function as positional cues in establishing cell polarity. *Dev Cell*. (2006) 10:391–6. doi: 10.1016/j.devcel.2005.12.016
45. Zhen T, Dai S, Li H, Yang Y, Kang L, Shi H, et al. MACC1 promotes carcinogenesis of colorectal cancer via beta-catenin signaling pathway. *Oncotarget*. (2014) 5:3756–69. doi: 10.18632/oncotarget.1993
46. Wu X, Tu X, Joeng KS, Hilton MJ, Williams DA, Long F. Rac1 activation controls nuclear localization of beta-catenin during canonical Wnt signaling. *Cell*. (2008) 133:340–53. doi: 10.1016/j.cell.2008.01.052
47. Buongiorno P, Pethe VV, Charames GS, Esufali S, Bapat B. Rac1 GTPase and the Rac1 exchange factor Tiam1 associate with Wnt-responsive promoters to enhance beta-catenin/TCF-dependent transcription in colorectal cancer cells. *Mol Cancer*. (2008) 7:73. doi: 10.1186/1476-4598-7-73
48. Teske C, Schweitzer C, Palamidessi A, Aust DE, Scita G, Weitz J, et al. Modulation of RAB5A early endosome trafficking in response to KRas



- mediated macropinocytic fluxes in pancreatic cancer cells. *Biochem Biophys Res Comm.* (2017) 493:528–33. doi: 10.1016/j.bbrc.2017.08.157
49. Kalin S, Hirschmann DT, Buser DP, Spiess M. Rabaptin5 is recruited to endosomes by Rab4 and Rabex5 to regulate endosome maturation. *J Cell Sci.* (2015) 128:4126–37. doi: 10.1242/jcs.174664
  50. Diaz J, Mendoza P, Ortiz R, Diaz N, Leyton L, Stupack D, et al. Rab5 is required in metastatic cancer cells for Caveolin-1-enhanced Rac1 activation, migration and invasion. *J Cell Sci.* (2014) 127:2401–6. doi: 10.1242/jcs.141689
  51. Longley DB, Johnston PG. Molecular mechanisms of drug resistance. *J Pathol.* (2005) 205:275–92. doi: 10.1002/path.1706
  52. Shen T, Zhang K, Siegal GP, Wei S. Prognostic value of E-cadherin and beta-catenin in triple-negative breast cancer. *Am J Clin Pathol.* (2016) 146:603–10. doi: 10.1093/ajcp/aqw183
  53. Hong J, Chen XS, Wu JY, Huang O, Zhu L, He JR, et al. [Analysis of the factors influencing adjuvant chemotherapy decisions for triple negative breast cancer]. *Zhonghua zhong liu za zhi.* (2017) 39:39–43. doi: 10.3760/cma.j.issn.0253-3766.2017.01.008
  54. Prat A, Pineda E, Adamo B, Galvan P, Fernandez A, Gaba L, et al. Clinical implications of the intrinsic molecular subtypes of breast cancer. *Breast.* (2015) 2 (24 Suppl):S26–35. doi: 10.1016/j.breast.2015.07.008
  55. Skvortsov S, Dudas J, Eichberger P, Witsch-Baumgartner M, Loeffler-Ragg J, Pritz C, et al. Rac1 as a potential therapeutic target for chemo-radioresistant head and neck squamous cell carcinomas (HNSCC). *Br J Cancer.* (2014) 110:2677–87. doi: 10.1038/bjc.2014.221
  56. Yoon C, Cho SJ, Chang KK, Park DJ, Ryeom SW, Yoon SS. Role of Rac1 pathway in epithelial-to-mesenchymal transition and cancer stem-like cell phenotypes in gastric adenocarcinoma. *Mol Cancer Res.* (2017) 15:1106–16. doi: 10.1158/1541-7786.MCR-17-0053
  57. Chen MJ, Wu DW, Wang YC, Chen CY, Lee H. PAK1 confers chemoresistance and poor outcome in non-small cell lung cancer via beta-catenin-mediated stemness. *Scient Rep.* (2016) 6:34933. doi: 10.1038/srep38463
  58. Fritz G, Kaina B. Rac1 GTPase, a multifunctional player in the regulation of genotoxic stress response. *Cell Cycle.* (2013) 12:2521–2. doi: 10.4161/cc.25807
  59. Fritz G, Henninger C. Rho GTPases: novel players in the regulation of the DNA damage response? *Biomolecules.* (2015) 5:2417–34. doi: 10.3390/biom5042417
  60. Wu GS. Role of mitogen-activated protein kinase phosphatases (MKPs) in cancer. *Cancer Metast Rev.* (2007) 26:579–85. doi: 10.1007/s10555-007-9079-6
  61. Hou MF, Chang CW, Chen FM, Wang SN, Yang SF, Chen PH, et al. Decreased total MKP-1 protein levels predict poor prognosis in breast cancer. *World J Surg.* (2012) 36:1922–32. doi: 10.1007/s00268-012-1608-y
  62. Wang HY, Cheng Z, Malbon CC. Overexpression of mitogen-activated protein kinase phosphatases MKP1, MKP2 in human breast cancer. *Cancer Lett.* (2003) 191:229–37. doi: 10.1016/S0304-3835(02)00612-2
  63. Rincon R, Zazo S, Chamizo C, Manso R, Gonzalez-Alonso P, Martin-Aparicio E, et al. c-Jun N-terminal kinase inactivation by mitogen-activated protein kinase phosphatase 1 determines resistance to taxanes and anthracyclines in breast cancer. *Mol Cancer Therap.* (2016) 15:2780–90. doi: 10.1158/1535-7163.MCT-15-0920
  64. Zhang L, Wang H, Li C, Zhao Y, Wu L, Du X, et al. VEGF-A/neuropilin 1 pathway confers cancer stemness via activating wnt/beta-catenin axis in breast cancer cells. *Cell Physiol Biochem.* (2017) 44:1251–62. doi: 10.1159/000485455
  65. Cheng S, Huang Y, Lou C, He Y, Zhang Y, Zhang Q. FSTL1 enhances chemoresistance and maintains stemness in breast cancer cells via integrin beta3/Wnt signaling under miR-137 regulation. *Cancer Biol Ther.* (2019) 20:328–37. doi: 10.1080/15384047.2018.1529101
  66. Mu J, Hui T, Shao B, Li L, Du Z, Lu L, et al. Dickkopf-related protein 2 induces G0/G1 arrest and apoptosis through suppressing Wnt/beta-catenin signaling and is frequently methylated in breast cancer. *Oncotarget.* (2017) 8:39443–59. doi: 10.18632/oncotarget.17055

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Liu, Chen, Sun, Zhang, Han, Dai, He, Shi and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Mechanisms of Resistance to CDK4/6 Inhibitors: Potential Implications and Biomarkers for Clinical Practice

Amelia McCartney<sup>1</sup>, Ilenia Migliaccio<sup>2</sup>, Martina Bonechi<sup>2</sup>, Chiara Biagioni<sup>3</sup>, Dario Romagnoli<sup>3</sup>, Francesca De Luca<sup>2</sup>, Francesca Galardi<sup>2</sup>, Emanuela Risi<sup>1</sup>, Irene De Santo<sup>1</sup>, Matteo Benelli<sup>3</sup>, Luca Malorni<sup>1,2</sup> and Angelo Di Leo<sup>1\*</sup>

<sup>1</sup> "Sandro Pitigliani" Department of Medical Oncology, Hospital of Prato, Prato, Italy, <sup>2</sup> "Sandro Pitigliani" Translational Research Unit, Hospital of Prato, Prato, Italy, <sup>3</sup> Bioinformatics Unit, Hospital of Prato, Prato, Italy

## OPEN ACCESS

### Edited by:

Mothaffar Rimawi,  
Baylor College of Medicine,  
United States

### Reviewed by:

Tomás Pascual Martínez,  
Hospital Clínic de Barcelona, Spain  
Nuria Chic,  
Hospital Clínic de Barcelona, Spain

### \*Correspondence:

Angelo Di Leo  
angelo.dileo@uslcentro.toscana.it

### Specialty section:

This article was submitted to  
Cancer Molecular Targets and  
Therapeutics,  
a section of the journal  
Frontiers in Oncology

**Received:** 10 May 2019

**Accepted:** 08 July 2019

**Published:** 23 July 2019

### Citation:

McCartney A, Migliaccio I, Bonechi M, Biagioni C, Romagnoli D, De Luca F, Galardi F, Risi E, De Santo I, Benelli M, Malorni L and Di Leo A (2019) Mechanisms of Resistance to CDK4/6 Inhibitors: Potential Implications and Biomarkers for Clinical Practice. *Front. Oncol.* 9:666. doi: 10.3389/fonc.2019.00666

The recent arrival of CDK4/6 inhibitor agents, with an approximate doubling of progression-free survival (PFS) associated with their use in hormone receptor-positive, HER2-negative advanced breast cancer (BC), has radically changed the approach to managing this disease. However, resistance to CDK4/6 inhibitors is considered a near-inevitability in most patients. Mechanisms of resistance to these agents are multifactorial, and research in this field is still evolving. Biomarkers with the ability to identify early resistance, or to predict the likelihood of successful treatment using CDK4/6 inhibitors are yet to be identified, and represent an area of unmet clinical need. Here we present selected mechanisms of resistance to CDK4/6 inhibitors, largely focussing on roles of Rb, cyclin E1, and the PIK3CA pathway, with discussion of associated biomarkers which have been investigated and applied in recent pre-clinical and clinical studies. These biological drivers may furthermore influence clinical treatment strategies adopted beyond CDK4/6 resistance.

**Keywords:** CDK4/6 inhibitors, biomarker, thymidine kinase-1, PIK3CA, resistance, palbociclib, ribociclib, abemaciclib

In normal cell signaling, mitogenic signaling via pathways including ER and PI3K/Akt/mTOR activates the cyclin D1-cyclin-dependent kinases 4 and 6 (CDK4/6) complex. Once activated, CDK4/6 phosphorylates the retinoblastoma protein (Rb), an event which causes Rb to lose the ability to bind to the E2F family of transcription factors. E2F is therefore released, activating gene transcription and thus initiating progression of the cell cycle from G1 to S phase (resulting in DNA synthesis) (1). Dysregulation of the CDK4/6 pathway, occasioning unchecked cell-cycle progression and proliferation, has been implicated in breast cancer (BC) via various mechanisms, including amplification of cyclin D1 (2), gain of CDK4, low expression of p18/high expression of RB1 (3) and inactivation of p16, the tumor suppressor protein and negative moderator of the cyclin D-CDK4/6 complex (4). The p16 protein, encoded by the *INK4a* gene, is a common target of inactivating mutations and deletions, operating upstream from RB (5).

The past decade has seen a radical shift in the management of advanced or metastatic hormone receptor-positive (HR+), human epidermal growth factor receptor-2 (HER2) negative BC. Endocrine therapy (ET) as monotherapy, previously thought to be the gold-standard in first and often subsequent lines of management, has been augmented and displaced in the treatment

hierarchy by the emergence of selective, small-molecule inhibitors of CDK4/6. Three such agents are currently in widespread clinical use, with all three, when given in combination with ET, resulting in an approximate doubling of progression-free survival (PFS) in patients with advanced HR+/HER2-negative BC, compared to ET plus placebo. These consistent positive PFS results have been demonstrated in large phase III trials in the upfront setting [PALOMA-2 (6) for palbociclib plus letrozole; MONALEESA-2 (7) for ribociclib plus letrozole and MONARCH-3 (8) for abemaciclib plus letrozole or anastrozole], as well as in populations previously treated with ET for advanced disease [PALOMA-3 for palbociclib plus fulvestrant (9); MONARCH-2 for abemaciclib and fulvestrant (10); and MONALEESA-3 for ribociclib plus fulvestrant (11)]. However, *de novo* or acquired resistance to CDK4/6 inhibitors is an almost ubiquitous inevitability, which has stimulated substantial interest in examining potential mechanisms of resistance, ways to overcome it, and methods of identifying it. Recently published reviews offer detailed insight into the myriad of likely mechanisms of resistance (12–14); conversely, we focus primarily on selected mechanisms and associated biomarkers that are of particular clinical interest, of whose potential has been validated in recent studies.

## MECHANISMS OF RESISTANCE—AVENUES FOR POSSIBLE MOLECULAR BIOMARKERS

### Loss of Retinoblastoma Susceptibility Gene Product (Rb) Function

Loss of Rb, the main target of CDK4/6, has been implicated by multiple preclinical studies in being a driver of resistance to CDK4/6 inhibitors (15–17). Without the inhibitory influence of Rb, transcription factors of the E2F family continue unchecked, thus facilitating unregulated cellular progression to S-phase entry. Conversely, higher levels of *RB1*, the gene encoding for Rb, and cyclin D1 (and lower levels of p16) are observed in human BC cell lines sensitive to palbociclib (18). Clinical evidence of acquired *RB1* mutations leading to CDK4/6 inhibitor resistance was recently reported in a case-series of three patients with metastatic BC who had genotyping performed in both tissue and blood samples before and after commencing a CDK4/6 inhibitor (19). Each somatic mutation was detected via ctDNA analyses performed at the point of disease progression, and was not present prior to initiation of CDK4/6 inhibition. Polyclonal *RB1* mutations were identified in patients assigned to the palbociclib arm of PALOMA-3, albeit at comparatively low frequency (4.7%) (20). RBSig, a gene expression signature of Rb loss-of-function, has been validated in identifying between palbociclib-sensitive and resistant BC cell lines (21), and has been associated with sensitivity to abemaciclib monotherapy in tumors derived from the neoMONARCH study (NCT02441948) (22). Similarly, a gene set containing E2F targets, the E2F regulon, was significantly associated with lack of PFS improvement from palbociclib combination in the PALOMA-3 trial (23). Of note, no interaction was found between treatment and *RB1* expression in the same

study, indicating that a wider analysis of RB pathway might be needed to identify resistant patients. Recent data emerging from genomic analysis of ER-positive BCs treated with CDK4/6 inhibitors revealed—not unsurprisingly—that loss of *RB1* was associated with treatment resistance. However, also implicated in resistance to CDK4/6 inhibition were loss-of-function mutations of *FAT1*, leading to cellular proliferation mediated via activation of the Hippo signaling pathway and elevations in CDK6, thus revealing an additional and intriguing potential mechanism of resistance (24).

### Cyclin E1 (CCNE1)

*CCNE1*, the gene that encodes cyclin E1, is upregulated in models with resistance to CDK4/6 inhibitors (25, 26). Data emerging from PALOMA-3 suggests that *CCNE1* expression is associated with benefit from palbociclib (23), in line with previous pre-clinical data which suggested *CCNE1* amplification is associated with acquired resistance to palbociclib (27), as well as exploratory data derived in the neoadjuvant NeoPalAna trial, associating high levels of *CCNE1* with palbociclib resistance (28). Tumor tissue procured from recurrent disease on trial was assessed via mRNA profiling, assessing a range of cell cycle-related genes. Although all biomarker groups derived benefit from palbociclib, those with low tumor *CCNE1* expression had a greater response (median PFS for those receiving palbociclib plus fulvestrant, 14.1 months; vs. 4.8 months in those receiving fulvestrant plus placebo) than those with high *CCNE1* expression (7.6 months vs. 4.0; palbociclib plus fulvestrant vs. fulvestrant plus placebo, respectively). The predictive power of *CCNE1* mRNA was stronger in metastatic biopsies (interaction  $p < 0.001$ ) than archived primary biopsy samples (interaction  $p = 0.09$ ). Investigators provided further validation in an independent cohort ( $N = 61$ ) drawn from the Preoperative Palbociclib (POP) Clinical Trial (NCT02008734), wherein high *CCNE1* mRNA expression correlated with a significantly lower anti-proliferative response to palbociclib. Contrastingly, no such association with *CCNE1* expression and PFS was found in the biomarker analyses of MONALEESA-2, with near-identical hazard ratios observed across expression groups (HR 0.54 95% CI 0.38–0.78 for *CCNE1* high expression; HR 0.53 95% CI 0.34–0.83 for low expression) (29). An earlier analysis of tumor specimens from PALOMA-2 also failed to demonstrate an association between *CCNE1* expression and benefit from palbociclib (30).

### Combining Rb and CCNE1—Possible Biomarker of CDK4/6 Sensitivity

A recent analysis of cell cycle-related markers found in a large panel of HR+ BC cell lines was recently reported, describing findings identified by gene-expression profiles and western blot (31). Both modalities identified that concurrent overexpression of *CCNE1* and down-regulation of Rb occurred at the time of palbociclib resistance. Subsequent *in silico* analyses, correlating the ratio between *CCNE1* and *RB1* expression levels (*CCNE1*/*RB1*) with palbociclib IC50 in a large dataset of cell lines, showed the ratio outperformed both *CCNE1* and/or *RB1* when they were utilized as sole markers. Furthermore, retrospective analyses showed *CCNE1*/*RB1* to be an adverse

prognostic factor, with the ability to differentiate between palbociclib sensitive and resistant patients enrolled in the neoadjuvant NeoPalAna trial (28).

### Initially Promising Targets but Negative Data: Cyclin D and p16

*CCND1*, the gene coding for cyclin D1 is amplified in approximately 15% of all BCs, and overexpression of cyclin D1 is observed in around 50% (32). Given the crucial role cyclin D1 plays in cell cycle mediation and its interplay with CDK4/6, it has been hypothesized that expression levels or dysregulation of cyclin D1 may relate to response to CDK4/6 inhibition. Similarly, intuitively, loss of p16<sup>INK4A</sup> and the consequent deficit of its usual inhibitory action on cyclin D1 would appear to be a reasonable premise of CDK4/6 inhibitor resistance. However, this is largely not been borne out in the clinical setting. In a phase II study of single-agent palbociclib, low p16 did not correlate with clinical outcome in Rb-positive, heavily pre-treated advanced BC (33). In the same study, amplification of cyclin D1 was also not associated with clinical benefit or PFS. PALOMA-1 failed to show any significant difference in PFS in patients whose tumors harbored evidence of a loss of p16 or *CCND1* amplification, compared to unselected patients (34). Expression levels of cyclin D1 was not associated with benefit from palbociclib in PALOMA-3 (23).

### PIK3CA—Activation of Growth Factor Signaling Pathways

Mutations in the phosphatidylinositol 3-kinase (PI3K) catalytic subunit (PIK3CA) are found in approximately 40% estrogen receptor-positive BC (3). The PI3K/mTOR pathway has been shown to be upregulated in response to chronic exposure to CDK4/6 inhibitors, which in turn upregulates cyclin D. In the absence of CDK4 and CDK6, activated cyclin D can activate CDK2, which subsequently drives cell cycle progression (27). Circulating tumor DNA sequencing was performed on 195 patients enrolled in the PALOMA-3 study, comparing baseline and end-of-treatment analyses (20), demonstrating the emergence of driver mutations in PIK3CA and *ESR1*. Patients with a history of greater drug exposure appeared more likely to develop driver gene mutations, perhaps underlining the role that drug pressure plays in clonal expansion. Contrastingly, PIK3CA mutations were detected in the circulating DNA of 129 patients enrolled in PALOMA-3, with no significant association observed with response to treatment (9). Similarly, biomarker analysis of MONALEESA-3 demonstrated consistent benefit from ribociclib plus fulvestrant, irrespective of PIK3CA alteration status, as detected in baseline circulating tumor DNA (35). Functioning downstream of PI3K is 3-phosphoinositide-dependent protein kinase 1 (PDK1), a vital requisite for the full activation of AKT (36). The PI3K-PDK1 signaling pathway has been implicated in mediating resistance to CDK4/6 inhibitors, with ribociclib-resistant BC cell lines demonstrating an increase in PDK1 levels following drug exposure, resulting in activation of the AKT pathway (37).

## PREDICTION OF SENSITIVITY OR EARLY RESPONSE TO CDK4/6 INHIBITION—A POSSIBLE ROLE FOR TK1

Thymidine kinase-1 (TK1) is an enzyme in the pyrimidine salvage pathway that plays a critical role in the synthesis of DNA and in cell proliferation (38). High TK1 levels and activity in primary BC tissue correlate with poor prognosis (39, 40). Malignant cells can secrete pathological levels of TK1 detectable in blood, whereas in disease-free controls, levels are low or undetectable (41), with similar patterns reported in membrane expression of TK1 (42). TK1 as a marker of cell proliferation has been known and studied for some decades, but until recently, widespread, reliable quantification of absolute levels and activity have been limited, with most historical tests being radioimmunoassay-based. DiviTum (Biovica International, Sweden) is a refined ELISA-based assay capable of estimating TK1 activity (TKa) in cell lines, plasma and serum. Previous studies have suggested baseline and repeated assessments of TKa during the course of treatment may provide prognostic information (43–45).

### TK1 as a Biomarker—Founding Data Within Endocrine Therapy Studies

Previous studies have validated the use of DiviTum, both as a prognostic marker, and as one of response to ET. A pilot study of 31 women with advanced HR+/HER2 negative BC commencing on a new line of palliative endocrine therapy showed that those with low baseline levels of plasma TKa had a median PFS of 25.9 months, vs. 5.9 months in those with high baseline levels ( $p = 0.012$ ) (46). Furthermore, patients whose TKa levels dropped after 1 month of ET demonstrated a significantly higher median PFS than those in whom TKa levels increased on treatment (14.5 months vs. 3.8, respectively;  $p = 0.0026$ ).

These findings were upheld by a second retrospective study of a larger, more heavily pre-treated population derived from the cohort of EFFECT, a landmark study which originally compared head-to-head palliative exemestane vs. fulvestrant (47). Again, baseline TKa levels proved prognostic: patients with low baseline readings had a median TTP (mTTP) of 5.03 vs. 2.57 in those with high baseline readings ( $p < 0.001$ ). Patients whose TKa increased from baseline after 3 months of treatment had a significantly shorter mTTP (3.39 months, 95% CI: 2.14–4.11) than those whose TKa did not increase (5.39 months, 95% CI: 4.01–6.68) ( $P = 0.0045$ ). After adjusting for major prognostic factors, TKa remained an independent marker (48).

### TK1 and CDK4/6 Inhibitors

Evidence of the prognostic role TK1 may play in HR+ BC patients has provided a proof-of-concept to justify moving investigation forward into the field of CDK4/6 inhibitors. Recent data suggests potential utility for TKa as a marker of CDK4/6 inhibition in patients receiving neoadjuvant ET plus palbociclib (49). DiviTum was employed in an analysis of serum samples derived from NeoPalAna, a neoadjuvant trial of 4 weeks of anastrozole monotherapy followed by four cycles of additional



palbociclib, followed by a subsequent palbociclib washout in all but eight patients (28, 49). TKa was shown to markedly reduce with the introduction of palbociclib, rising at washout (but remaining suppressed in patients who did not receive washout). There was high concordance between changes in TKa and tumor Ki67 in the same direction from baseline to C1D15 and from C1D15 to point of curative surgery. This led to the conjecture that TKa may be seen as a dynamic marker that signifies the presence or absence of palbociclib activity. However, there is some pre-clinical evidence suggesting that TKa may precede a significant reduction in cellular proliferation. In a panel of HR+ BC models—both with sensitivity to palbociclib, and with acquired resistance to the drug—exposure to escalating levels of palbociclib and its relation to cellular proliferation and TKa was examined (50). In palbociclib-sensitive models, TKa significantly reduced after 3 days of drug exposure compared to control ( $p < 0.05$ ). Concurrently, cellular proliferation (as assessed by methylene blue assay) was observed to drop significantly after a minimum of 6 days, suggesting TKa may be an early marker of proliferative inhibition in response to palbociclib. This phenomenon was not observed in models with acquired resistance to palbociclib. The prognostic ability of TKa has been clinically validated in planned translational studies of plasma derived from the TReND study (NCT02549430), a phase II trial which tested the activity and safety of single-agent palbociclib against palbociclib combined with the ET the patient had received (and progressed on) most recently before enrollment (51). Not unlike to previous findings in ET-based TK studies, TReND patients with a low baseline TKa had a significantly longer PFS compared to those whose levels were high at study commencement. Similarly, on treatment, those patients whose TKa levels rose had a shorter time to disease progression compared to those patients whose levels remained stable or dropped in response to treatment (52). The prognostic role of serum TK1 assessed at baseline and on treatment is being further explored in two ongoing clinical trials of luminal BC patients treated with CDK4/6 inhibitors and ET: BIOITALEE (NCT03439046), a Phase 3b biomarker study of ribociclib plus letrozole in the first-line setting and PYTHIA (NCT02536742), a phase 2 biomarker discovery trial of palbociclib and fulvestrant in patients with endocrine resistant disease.

## IMPLICATIONS ON THERAPEUTIC APPROACHES FOLLOWING PROGRESSION ON CDK4/6 INHIBITORS

### Primary Resistance to CDK4/6 Inhibitors

Approximately 10% of patients will have primary resistance to CDK4/6 inhibitors. Biomarkers may have future potential to identify such patients at baseline or soon after commencing treatment, thus facilitating an early switch to a more efficacious treatment. For instance, patients with evidence of functional Rb loss at baseline are not likely to benefit from CDK4/6 inhibition. Similarly, baseline evidence of increased cyclin E1 expression, or the CCNE1/RB ratio may also play a role in identifying these patients. Peripheral evidence of ongoing

neoplastic proliferation, as manifested by a rise in TK1 activity within a month of commencing therapy, may also provide a marker of early resistance.

### Secondary Resistance to CDK4/6 Inhibitors

An unanswered question regards the continuation of CDK4/6 inhibitors beyond progression on these agents. The premise that continuing a CDK4/6 inhibitor beyond progression may prove an effective strategy is being tested by several ongoing Phase 1 and 2 trials (MAINTAIN NCT02632045, PACE NCT03147287, NCT01857193, NCT 02871791, and TRINITI-1 NCT 02732119). Mutations in *RB1*, resulting in activation of other cell cycle factors, such as E2F and the Cyclin E-CDK2 axis, has been demonstrated in cases of acquired resistance (19, 53). This in turn results in independence from the CDK4/6 pathway for cell cycle progression from G1 to S phase. In such cases, in the setting of disease progression on a CDK4/6 inhibitor, concurrent biomarker evidence of a functional loss of Rb may support a switch to a new agent, rather than continuing CDK4/6 agents beyond progression.

### A Potential Role for PIK3CA Inhibitors?

PI3K-dependent activation of non-canonical cyclin D1-CDK2 and resultant recovery of Rb phosphorylation and S phase entry has been implicated in early resistance to CDK4/6 inhibition, with combined PI3K and CDK4/6 inhibition demonstrating the ability to overcome resistance to CDK4/6 inhibitors in BC cell lines (27). Hence, the role that PIK3CA inhibitors may play in overcoming resistance is of relevant interest. The SOLAR-1 trial randomly assigned 572 patients with pre-treated HR+/HER2-negative advanced BC to receive the oral PIK3CA inhibitor alpelisib plus fulvestrant or fulvestrant plus placebo (54). The primary endpoint was PFS in patients with PIK3CA mutations detectable in tumor tissue ( $n = 341$ ). After a median follow up of 20 months, the median PFS was almost double in mutation-positive patients receiving alpelisib compared to those receiving placebo (11.0 months vs. 5.7 months, respectively; HR 0.65 95% CI 0.50–1.25  $p = 0.00065$ ). Further data, reporting the efficacy of alpelisib according to mutational status evaluated by ctDNA, suggested an even greater clinical benefit than tissue analysis (55). In patients with a PIK3CA mutation detected via liquid biopsy ( $n = 186$ ), there was a 45% risk reduction in PFS (HR 0.55 95% CI 0.39–0.79). In the small number of patients who had previously received CDK4/6 inhibition ( $n = 20$ ), there was a 52% risk reduction in PFS in favor of alpelisib over placebo (HR 0.48 95% CI 0.17–1.36). Alpelisib is selective for the alpha isoform of PI3K, which has so far set it apart from pan-PI3K inhibitors, which have reported notably poor safety profiles (56, 57). Nevertheless, data from SOLAR-1 still reflect considerable toxicity. All-grade hyperglycaemia occurred in 64% of patients receiving alpelisib (37% occurring at grade 3/4), 58% reported diarrhea, 45% nausea and 36% developed rash (10% at grade 3/4). Five percent of patients discontinued from the alpelisib arm due to adverse events (54).

Whilst the number of patients with prior exposure to CDK4/6 inhibitors subsequently enrolled in SOLAR-1 was small, it is not unreasonable to consider—in patients harboring an actionable



mutation—PIK3CA inhibition following disease progression on a CDK4/6 agent. This is particularly relevant, given the likelihood of emergence of driver mutations of PIK3CA secondary to

previous ET and CDK4/6-targeted therapies (20). Whilst some pre-clinical data suggest that triplet therapy, combining ET plus CDK4/6 inhibitors with PIK3CA agents may be better in

**TABLE 1 |** Ongoing trials evaluating the combination of CDK4/6 inhibitors with PIK3CA agents in breast cancer [ClinicalTrials.gov June 2019].

ClinicalTrials.gov identifier	Phase	Setting/key eligibility	Intervention	Primary endpoint(s)
NCT02088684	Phase 1b/2* (*Phase 2 portion of study not opened due to sponsor decision)	ER+/HER2-negative metastatic or advanced disease	Ribociclib + fulvestrant + buparlisib or ribociclib + fulvestrant + alpelisib or ribociclib + fulvestrant	DLT (Phase 1b part only)
NCT02389842	Phase 1b (dose escalation and expansion phases)	Specific sub-group for patients with PIK3CA mutation	Palbociclib + taselisib + fulvestrant or palbociclib + taselisib or letrozole + palbociclib + taselisib	Recommended Phase 2 dose Safety and toxicity Anti-tumor response
NCT03939897	Phase 1/2	Endocrine resistant metastatic disease	Abemaciclib + fulvestrant + copanlisib vs. abemaciclib + fulvestrant	DLT PFS
NCT02684032	Phase 1b (dose escalation and expansion phases)	ER+/HER2 negative metastatic disease	Letrozole + palbociclib + gedatolisib or palbociclib + fulvestrant + gedatolisib	DLT rate ORR
NCT03128619	Phase 1/2	Locally advanced or metastatic disease not previously treated in advanced setting (Phase 1b portion); Stage I, II or III disease (Phase 2 (portion)	Copanlisib D1,8,15 + letrozole continuously or copanlisib D1,8,15 + palbociclib + letrozole continuously or palbociclib + letrozole for 14 days, followed by copanlisib D 1,8,15 + letrozole continuously	MTD Change in Ki67 on treatment DLT rate
NCT02154776	Phase 1 (dose escalation, open label)	ER+/HER2-negative advanced disease	Ribociclib + buparlisib + letrozole	DLT rate Safety and tolerability

DLT, dose limiting toxicity; ER+, endocrine receptor-positive; MTD, maximum tolerated dose; ORR, objective response rate; PFS, progression-free survival.

**TABLE 2 |** Currently enrolling trials recruiting patients with hormone receptor-positive metastatic breast cancer, evaluating the role of PIK3CA agents [ClinicalTrials.gov June 2019].

ClinicalTrials.gov identifier	Phase	Setting/key eligibility	Intervention	Primary endpoint(s)
NCT03006172	Phase 1 Dose escalation and expansion stages	PIK3CA mutation-positive	GDC-077 as single agent and also in combination with ET and CDK4/6i	% of pts with DLTs Recommended Phase 2 dose % of pts with AEs and SAEs
NCT03207529	Phase 1 Dose escalation and expansion stages	AR-positive and PTEN positive metastatic disease	Alpelisib + enzalutamide	MTD
NCT02705859	Phase 1b/2 Single arm	HER2-positive disease	Copanlisib and trastuzumab	MTD with trastuzumab CBR
NCT03767335	Phase 1b Dose escalation and expansion stages	HER2-positive disease (+/- HR-positivity)	MEN1611 + trastuzumab +/- fulvestrant (in HR-positive disease)	MTD Recommended Phase 2 dose
NCT03386162	Randomized Phase 2	PIK3CA mutation-positive Patients not progressing after 6–8 cycles of 1st or 2nd line chemotherapy	Fulvestrant + alpelisib vs. maintenance chemotherapy	PFS
NCT03056755	Phase 2 Non-comparative study	PIK3CA mutation-positive Post progression on CDK4/6i	Alpelisib + fulvestrant OR Alpelisib + letrozole	% of patients alive without progressive disease (at ~6 month mark)
NCT03337724	Phase 3 Randomized, double blind	PIK3CA/AKT1/PTEN-altered disease Accepting TNBC or HR+/HER2neg	Ipatasertib + paclitaxel vs. placebo+paclitaxel	PFS

AE, adverse event; AKT1, Alpha serine/threonine protein kinase; CBR, clinical benefit rate; CDK4/6i, Cyclin-dependent kinase 4/6 inhibitor; DLT, dose-limiting toxicity; HR, hormone receptor; ET, endocrine therapy; MTD, maximum tolerated dose; PIK3CA, phosphatidylinositol 3-kinase catalytic subunit; PFS, progression-free survival; PTEN, phosphate and tensin homolog tumor suppressor; SAE, serious adverse event; TNBC, triple negative breast cancer.

preventing acquired CDK4/6 resistance than doublet regimens (27), this approach may potentially come at the cost of increased toxicity (58). Clinical trials investigating the feasibility and utility of combining CDK4/6 and PI3K inhibition are ongoing (Table 1). Another alternative may be to expose patients to these agents sequentially rather than simultaneously, reserving PIK3CA inhibition for those harboring a druggable mutation following exposure to CDK4/6 inhibition. This tactic is being tested in a currently-recruiting phase II study, which will assess the efficacy and safety of combining alpelisib and ET in patients with PIK3CA mutations whose disease has progressed on or after receiving a CDK4/6 inhibitor plus ET (NCT03056755). A summary of clinical trials of PIK3CA agents, currently recruiting patients with endocrine receptor-positive advanced BC, is presented in Table 2.

## REFERENCES

- Burkhardt DL, Sage J. Cellular mechanisms of tumour suppression by the retinoblastoma gene. *Nat Rev Cancer*. (2008) 8:671–82. doi: 10.1038/nrc2c399
- Musgrove EA, Caldon CE, Barraclough J, Stone A, Sutherland RL. Cyclin D as a therapeutic target in cancer. *Nat Cancer Rev*. (2011) 11:558–72. doi: 10.1038/nrc3c090
- The Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature*. (2012) 490:61–70. doi: 10.1038/nature1e1412
- Shapiro GI. Cyclin-dependent kinase pathways as targets for cancer treatment. *J Clin Oncol*. (2006) 24:1770–83. doi: 10.1200/JCO.2005.03.7689
- Sherr CJ, Beach D, Shapiro GI. Targeting CDK4K and CDK6K: from discovery to therapy. *Cancer Discov*. (2016) 6:353–67. doi: 10.1158/2159-8290.CD-15-0894
- Finn RS, Martin M, Rugo HS, Jones S, Seock-Ah I, Gelmon K, et al. Palbociclib and letrozole in advanced breast cancer. *N Engl J Med*. (2016) 375:1925–36. doi: 10.1056/NEJMoa1607303
- Hortobagyi GN, Stemmer SM, Burris HA, Yap Y-S, Sonke GS, Paluch-Shimon S, et al. Ribociclib as first-line therapy for HR-positive, advanced breast cancer. *N Engl J Med*. (2016) 375:1738–48. doi: 10.1056/NEJMoa1609709
- Goetz MP, Toi M, Campone M, Sohn J, Paluch-Shimon S, Huober J, et al. MONARCH3H: Abemaciclib as initial therapy for advanced breast cancer. *J Clin Oncol*. (2017) 35:3638–46. doi: 10.1200/JCO.2017.75.6155
- Cristofanilli M, Turner NC, Bondarenko I, Ro J, Seock-Ah I, Masuda N, et al. Fulvestrant plus palbociclib versus fulvestrant plus placebo for treatment of hormone-receptor-positive, HER2R-negative metastatic breast cancer that progressed on previous endocrine therapy (PALOMA 3): final analysis of the multicentre, double-blind, phase 3 randomised controlled trial. *Lancet Oncol*. (2016) 17:425–39. doi: 10.1016/S1470-2045(15)00613-0
- Sledge GW Jr, Toi M, Neven P, Sohn J, Inoue K, Pivov X, et al. MONARCH 2: abemaciclib in combination with fulvestrant in women with HR+/HER2R-advanced breast cancer who had progressed while receiving endocrine therapy. *J Clin Oncol*. (2017) 35:2875–84. doi: 10.1200/JCO.2017.73.7585
- Slamon DJ, Neven P, Chia S, Fasching PA, De Laurentiis M, Seock-Ah I, et al. Phase III randomized study of ribociclib and fulvestrant in hormone receptor-positive, human epidermal growth factor receptor 2-negative advanced breast cancer: MONALEESA-3. *J Clin Oncol*. (2018) 36:2465–72. doi: 10.1200/JCO.2018.78.9909
- Guarducci C, Bonechi M, Boccalini G, Benelli M, Risi E, Di Leo A, et al. Mechanisms of resistance to CDK4K/6 inhibitors in breast cancer and potential biomarkers of response. *Breast Care*. (2017) 12:304–8. doi: 10.1159/000484167
- Pandey K, An H-J, Kim SK, Lee SA, Kim S, Lim SM, et al. Molecular mechanisms of resistance to CDK4K/6 inhibitors in breast cancer: a review. *Int J Cancer*. (2018) 145:1179–88. doi: 10.1002/ijc.32020

## AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

## FUNDING

This work was partially supported by the Breast Cancer Research Foundation [grant number BCRF 18-037].

## ACKNOWLEDGMENTS

The authors wish to acknowledge the Sandro Pitigliani Foundation, Prato, Italy, for their ongoing support.

- Portman N, Alexandrou S, Carson E, Wang S, Lim E, Caldon CE. Overcoming CDK4K/6 inhibitor resistance in ER-positive breast cancer. *Endocr Relat Cancer*. (2019) 26:R1R5–30. doi: 10.1530/ERC-18-0317
- Fry DW, Harvey PJ, Keller PR, Elliott WL, Meade M, Trachet E, et al. Specific inhibition of cyclin-dependent kinase 4/6 by PD0D332991 and associated antitumor activity in human tumor xenografts. *Mol Cancer Ther*. (2004) 3:1427–38.
- Dean JL, McClendon AK, Hickey TE, Butler LM, Tilley WD, Witkiewicz AK, et al. Therapeutic response to CK4K/6 inhibition to breast cancer defined by *ex vivo* analyses of human tumors. *Cell Cycle*. (2012) 11:2756–61. doi: 10.4161/cc.21195
- Witkiewicz AK, Knudsen ES. Retinoblastoma tumor suppressor pathway in breast cancer: prognosis, precision medicine and therapeutic interventions. *Breast Cancer Res*. (2014) 16:207. doi: 10.1186/bcr3r652
- Finn RS, Dering J, Conklin D, Kalous O, Cohen DJ, Desai AJ, et al. PD 0332991, 2991a selective cyclin D kinase 4/6 inhibitor, preferentially inhibits proliferation of luminal estrogen receptor-positive human breast cancer cell lines *in vitro*. *Breast Cancer Res*. (2009) 11:R77. doi: 10.1186/bcr2r419
- Condorelli R, Spring L, O'Shaughnessy J, Lacroix L, Bailleux C, Scott V, et al. Polyclonal RB1B mutations and acquired resistance to CDK4K/6 inhibitors in patients with metastatic breast cancer. *Ann Oncol*. (2018) 29:640–5. doi: 10.1093/annonc/mdx7x84
- O'Leary B, Cutts RJ, Liu Y, Hrebien S, Huang X, Fenwick K, et al. The genetic landscape and clonal evolution of breast cancer resistance to palbociclib plus fulvestrant in the PALOMA-3 trial. *Cancer Discov*. (2018) 8:1390–403. doi: 10.1158/2159-8290.CD-18-0264
- Malorni L, Piazza S, Ciani Y, Guarducci C, Bonechi M, Biagioni C, et al. A gene expression signature of retinoblastoma loss-of-function is a predictive biomarker of resistance to palbociclib in breast cancer cell lines and is prognostic in patients with ER positive early breast cancer. *Oncotarget*. (2016) 7:68012–22. doi: 10.18632/oncotarget.12010
- Hurvitz S, Martin M, Press M, Wijayawardana S, Brahmachary M, Ebert PJ, et al. Abstract PD2D-10: treatment with abemaciclib modulates the immune response in gene expression analysis of the neoMONARCH neoadjuvant study of abemaciclib in postmenopausal women with HR+, HER2R negative breast cancer. *Cancer Res*. (2019) 79(Suppl. 4):Abstract nr PD2-10. doi: 10.1158/1538-7445.SABCS18-PD2-10
- Turner NC, Liu Y, Zhu Z, Loi S, Colleoni M, Loibl S, et al. Cyclin E1E expression and palbociclib efficacy in previously treated hormone receptor-positive metastatic breast cancer. *J Clin Oncol*. (2019) 37:1169–78. doi: 10.1200/JCO.18.00925
- Li Z, Razavi P, Li Q, Toy W, Liu B, Ping C, et al. Loss of the FAT1 tumor suppressor promotes resistance to CDK4K/6 inhibitors via the hippo pathway. *Cancer Cell*. (2018) 34:893–905. doi: 10.1016/j.ccell.2018.11.006
- Taylor-Harding B, Aspuria PJ, Agadjanian H, Cheon DJ, Mizuno T, Greenberg D, et al. Cyclin E1E and RTK/RAS signalling drive CDK inhibitor

- resistance via activation of E2EF22F and ETS. *Oncotarget*. (2015) 6:696–714. doi: 10.18632/oncotarget.2673
26. Knudsen ES, Witkiewicz AK. The strange case of CDK4/6 inhibitors: mechanisms, resistance and combination strategies. *Trends Cancer*. (2017) 3:39–55. doi: 10.1016/j.trecan.2016.11.006
27. Herrera-Abreu MT, Palafox M, Asghar U, Rivas MA, Cutts RJ, Garcia-Murillas I, et al. Early adaption and acquired resistance to CDK4/6 inhibition in estrogen receptor-positive breast cancer. *Cancer Res*. (2016) 76:2301–13. doi: 10.1158/0008-5472.CAN-15-0728
28. Ma CX, Gao F, Luo J, Northfelt DW, Goetz M, Forero A, et al. NeoPalAna: neoadjuvant palbociclib, a cyclin-dependent kinase 4/6 inhibitor, and anastrozole for clinical stage 2 or 3 estrogen receptor-positive breast cancer. *Clin Cancer Res*. (2017) 23:4055–65. doi: 10.1158/1078-0432.CCR-16-3206
29. Hortobagyi GN, Paluch-Shimon S, Petrakova K, Villaneuva C, Chan A, Nusch A, et al. First-line ribociclib (RIB) + letrozole (LET) in hormone receptor-positive (HR+), HER2R-negative (HER2R-) advanced breast cancer (ABC): MONALEESA-2 biomarker analyses. *J Clin Oncol*. (2018) 36(Suppl. 15):1022. doi: 10.1200/JCO.2018.36.15\_suppl.1022
30. Finn RS, Yiu Y, Martin M, et al. Abstract P2P-09-10: comprehensive gene expression biomarker analysis of CDK4/6 and endocrine pathways from the PALOMA-2 study. *Cancer Res*. (2018) 78 (Suppl. 4):P2-09-10-P2-09-10. doi: 10.1158/1538-7445.SABCS1S7-P2P-09-10
31. Guarducci C, Bonechi M, Benelli M, Biagioni C, Boccalini G, Romagnoli D, et al. Cyclin E1E and Rb modulation as common events at time of resistance to palbociclib in hormone receptor-positive breast cancer. *NPJ Breast Cancer*. (2018) 4:38. doi: 10.1038/s41523-018-0092-4
32. Gillett C, Fantl V, Smith R, Risher C, Bartek J, Dickson C, et al. Amplification and overexpression of cyclin D1D in breast cancer detected by immunohistochemical staining. *Cancer Res*. (1994) 54:1812–17.
33. DeMichele A, Clare AS, Tan KS, Heitjan DF, Gramlich K, Gallagher M, et al. CDK4/6 inhibitor palbociclib (PD0332991) in Rb+ advanced breast cancer: phase II activity, safety, and predictive biomarker assessment. *Clin Cancer Res*. (2015) 21:995–1001. doi: 10.1158/1078-0432.CCR-14-2258
34. 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, HER2R-negative, advanced breast cancer (PALOMA-1/TRIO-18): a randomised phase 2 study. *Lancet Oncol*. (2015) 16:25–35. doi: 10.1016/S1547-0245(14)71159-3
35. Neven P, Petrakova K, Val Bianchi G. Abstract PD2D-05: biomarker analysis by baseline circulating tumor DNA alterations in the MONALEESA-3 study. *Cancer Res*. (2019) 79 (Suppl. 4):PD2-05-PD2-05. doi: 10.1158/1538-7445.SABCS1S8-PD2D-05
36. Alessi DR, James SR, Downes CP, Holmes AB, Gaffney PRJ, Reese CB, et al. Characterization of a 3-phosphoinositide-dependent protein kinase which phosphorylates and activates protein kinase B alpha. *Curr Biol*. (1997) 7:261–9. doi: 10.1016/S0960-9822(06)00122-9
37. Jansen VM, Bhola NE, Bauer JA, Formisano L, Lee K-M, Hutchinson KE, et al. Kinome-wide RNA interference screen reveals a role for PDK1K in acquired resistance to CDK4/6 inhibition in ER-positive breast cancer. *Cancer Res*. (2017) 77:2488–99. doi: 10.1158/0008-5472.CAN-16-2653
38. Topolcan O, Holubec L. The role of thymidine kinase in cancer diseases. *Expert Opin Med Diagn*. (2008) 2:129–41. doi: 10.1517/17530059.2.2.129
39. Romain S, Spyrtatos F, Descortes F, Daver A, Rostaing-Puissant B, Bougnoux P, et al. Prognostic role of DNA-synthesising enzyme activity (thymidine kinase and thymidylate synthase) in 908 T1T-T2T N0N-N1N, M0M breast cancers: a retrospective multicenter study. *Int J Cancer*. (2000) 87:860–8.
40. Broët P, Romain S, Daver A, Ricolleau G, Quillien V, Rallet A, et al. Thymidine kinase as a proliferative marker: clinical relevance in 1692 primary breast cancer patients. *J Clin Oncol*. (2001) 19:2778–87. doi: 10.1200/JCO.2001.19.11.2778
41. Robertson JE, O'Neill KL, Thomas MW, McKenna PG, Blamey RW. Thymidine kinase in breast cancer. *Br J Cancer*. (1990) 62:663–7. doi: 10.1038/bjc.1990.352
42. Weigel EG, Burrup W, Kovtun R, Velazquez EJ, Felsted AM, Townsend MH, et al. Membrane expression of thymidine kinase 1 and potential clinical relevance in lung, breast and colorectal malignancies. *Cancer Cell Int*. (2018) 10:135. doi: 10.1186/s12935-018-0633-9
43. Nisman B, Allweis T, Kadouri L, Mali B, Hamburger T, Baras M, et al. Comparison of diagnostic and prognostic performance of two assays measuring thymidine kinase 1 activity in serum of breast cancer patients. *Clin Chem Lab Med*. (2013) 51:439–47. doi: 10.1515/cclm-2012-0162
44. Bolayirli M, Papila C, Korkmaz GG, Papila B, Aydogan K, Karatas A, et al. Serum thymidine kinase 1 activity in solid tumour (breast and colorectal cancer) patients treated with adjuvant chemotherapy. *J Clin Lab Anal*. (2013) 27:220–6. doi: 10.1002/jcla.21587
45. Bjohle J, Bergqvist J, Gronowitz JS, Johansson H, Carlsson L, Einbeigi Z, et al. Serum thymidine kinase activity compared with CA 15-3 in locally advanced and metastatic breast cancer within a randomised trial. *Breast Cancer Res Treat*. (2013) 139:751–8. doi: 10.1007/s10549-013-2579-x
46. Bonechi M, Galardi F, Biagioni C, De Luca F, Bergqvist M, Neumüller M, et al. Plasma thymidine kinase-1 predicts outcome in patients with hormone receptor positive and HER2R negative metastatic breast cancer treated with endocrine therapy. *Oncotarget*. (2018) 27:16389–99. doi: 10.18632/oncotarget.24700
47. Chia S, Gradishar W, Bauriac L, Bines J, Amant F, Federico M, et al. Double-blind, randomized placebo-controlled trial of fulvestrant compared with exemestane after prior nonsteroidal aromatase inhibitor therapy in postmenopausal women with hormone receptor-positive, advanced breast cancer: Results from EFECT. *J Clin Oncol*. (2008) 26:1664–70. doi: 10.1200/JCO.2007.13.5822
48. McCartney A, Biagioni C, Schiavon G, Bergqvist M, Mattsson K, Miagliaccio I, et al. Prognostic role of serum thymidine kinase 1 activity in patients with hormone receptor-positive metastatic breast cancer: analysis of the randomised phase III Evaluation of Faslodex versus Exemestane Clinical Trial (EFECT). *Eur J Cancer*. (2019) 114:55–66. doi: 10.1016/j.ejca.2019.04.002
49. Bagegni N, Thomas S, Liu N, Luo J, Hoog J, Dorthfelt DW, et al. Serum thymidine kinase 1 activity as a pharmacodynamic marker of cyclin-dependent kinase 4/6 inhibition in patients with early-stage breast cancer receiving neoadjuvant palbociclib. *Breast Can Res*. (2017) 19:123. doi: 10.1186/s13058-017-0913-7
50. Bonechi M, Migliaccio I, Benelli M, Romagnoli D, Bergqvist M, Mattsson K, et al. Effects of palbociclib on thymidine kinase-1 (TK1K) in hormone receptor positive (HR+) breast cancer cell lines. *Cancer Res*. (2019) 79 (Suppl. 4):P6-09-02-P6-09-02. doi: 10.1158/1538-7445.SABCS1S8-P6-09-02
51. Malorni L, Curigliano G, Minisini AM, Cinieri S, Tondini CA, D'Hollander K, et al. Palbociclib as single agent or in combination with the endocrine therapy received before disease progression for estrogen receptor-positive, HER2R-negative metastatic breast cancer: TREnd trial. *Ann Oncol*. (2018) 29:1748–54. doi: 10.1093/annonc/mdy2y14
52. Malorni L, Biagioni C, De Luca F, Bonechi M, Curigliano G, Minisini AM, et al. Abstract 4416/10 plasma thymidine kinase activity in patients with luminal metastatic breast cancer treated with palbociclib within the phase II TREnd trial. In: *AACR Annual Meeting* (Atlanta, GA).
53. Cen L, Carlson BL, Schroeder MA, Ostrem JL, Kitange GJ, Mladek AC, et al. p1p6-CDK4K-Rb axis controls sensitivity to a cyclin-dependent kinase inhibitor PD0332991 in glioblastoma xenograft cells. *Neuro Oncol*. (2012) 14:870–81. doi: 10.1093/neuonc/nos1s14
54. André F, Ciruelos EM, Rubovszky G, Campone M, Loibl S, Rugo HS, et al. Alpelisib (ALP) + fulvestrant (FUL) for advanced breast cancer (ABC): results of the phase III SOLAR-1 trial. *Ann Oncol*. (2018) 29:abstract LBA3\_PR. doi: 10.1093/annonc/mdy4y24.010
55. Juric D, Ciruelos E, Rubovszky G, Campone M, Loibl S, Rugo HS, et al. Alpelisib + fulvestrant for advanced breast cancer: subgroup analyses from the phase III SOLAR-1 trial. *Cancer Res*. (2019) doi: 10.1158/1538-7445.SABCS1S8-GS3S-08. [Epub ahead of print].

56. Di Leo A, Johnston S, Lee KS, Ciruelos E, Lønning PE, Janni W, et al. Buparlisib plus fulvestrant in postmenopausal women with hormone-receptor-positive, HER2R-negative, advanced breast cancer progressing on or after mTOR inhibition (BELLE-3): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Oncol.* (2018) 19:P87–100. doi: 10.1016/S15470-2045(17)30688-5
57. Baselga J, Dent SF, Cortés J, Young-Hyuck I, Diéras V, Harbeck N, et al. Phase III study of taselisib (GDC-0032) + fulvestrant (FULV) v FULVE in patients (pts) with estrogen receptor (ER)-positive, PIK3KC33CA-mutant (MUT), locally advanced or metastatic breast cancer (MBC): primary analysis from SANDPIPER. *J Clin Oncol.* (2018) 36:18. doi: 10.1200/JCO.2018.36.18\_suppl.LBA1006
58. Cortés J, Im S-A, Holgado E, Perez-Garcia JM, Schmid P, Chavez-MacGregor M. The next era of treatment for hormone receptor-positive, HER2R-negative advanced breast cancer: triplet combination-based endocrine therapies. *Cancer Treat Rev.* (2017) 61:53–60. doi: 10.1016/j.ctrv.2017.09.011

**Conflict of Interest Statement:** AD has received honoraria from, and served as an advisory board member for, Pfizer, Lilly, AstraZeneca, and Novartis. AD has received research funding from Pfizer, AstraZeneca, and Novartis. LM has received consultation fees from Novartis and research support from Pfizer.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 McCartney, Migliaccio, Bonechi, Biagioni, Romagnoli, De Luca, Galardi, Risi, De Santo, Benelli, Malorni and Di Leo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Comparison of GenesWell BCT Score With Oncotype DX Recurrence Score for Risk Classification in Asian Women With Hormone Receptor-Positive, HER2-Negative Early Breast Cancer

Mi Jeong Kwon<sup>1,2†</sup>, Jeong Eon Lee<sup>3,4†</sup>, Joon Jeong<sup>5†</sup>, Sang Uk Woo<sup>6</sup>, Jinil Han<sup>7</sup>, Byeong-il Kang<sup>7</sup>, Jee-Eun Kim<sup>7</sup>, Youngho Moon<sup>7</sup>, Sae Byul Lee<sup>8</sup>, Seonghoon Lee<sup>6</sup>, Yoon-La Choi<sup>3,9,10</sup>, Youngmi Kwon<sup>11</sup>, Kyoung Song<sup>12</sup>, Gyungyub Gong<sup>13\*</sup> and Young Kee Shin<sup>14,15\*</sup>

## OPEN ACCESS

### Edited by:

Aleix Prat,  
Hospital Clinic de Barcelona, Spain

### Reviewed by:

Angel Luis Guerrero-Zotano,  
Instituto Valenciano de  
Oncología, Spain  
Tomás Pascual Martínez,  
Hospital Clinic de Barcelona, Spain

### \*Correspondence:

Gyungyub Gong  
gygong@amc.seoul.kr  
Young Kee Shin  
ykeeshin@snu.ac.kr

†These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
Women's Cancer,  
a section of the journal  
Frontiers in Oncology

**Received:** 13 March 2019

**Accepted:** 09 July 2019

**Published:** 24 July 2019

### Citation:

Kwon MJ, Lee JE, Jeong J, Woo SU,  
Han J, Kang B, Kim J-E, Moon Y, Lee  
SB, Lee S, Choi Y-L, Kwon Y, Song K,  
Gong G and Shin YK (2019)  
Comparison of GenesWell BCT Score  
With Oncotype DX Recurrence Score  
for Risk Classification in Asian Women  
With Hormone Receptor-Positive,  
HER2-Negative Early Breast Cancer.  
Front. Oncol. 9:667.  
doi: 10.3389/fonc.2019.00667

<sup>1</sup> Department of Pharmacy, College of Pharmacy, Kyungpook National University, Daegu, South Korea, <sup>2</sup> Research Institute of Pharmaceutical Sciences, Kyungpook National University, Daegu, South Korea, <sup>3</sup> Department of Health Sciences and Technology, Samsung Advanced Institute for Health Sciences and Technology (SAIHST), Sungkyunkwan University, Seoul, South Korea, <sup>4</sup> Department of Surgery, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, South Korea, <sup>5</sup> Department of Surgery, Gangnam Severance Hospital, Yonsei University College of Medicine, Seoul, South Korea, <sup>6</sup> Department of Surgery, Korea University Guro Hospital, Seoul, South Korea, <sup>7</sup> Gencurix, Inc., Seoul, South Korea, <sup>8</sup> Division of Breast Surgery, Department of Surgery, Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea, <sup>9</sup> Laboratory of Cancer Genomics and Molecular Pathology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, South Korea, <sup>10</sup> Department of Pathology and Translational Genomics, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, South Korea, <sup>11</sup> Center for Breast Cancer, National Cancer Center, Goyang-si, South Korea, <sup>12</sup> LOGONE Bio Convergence Research Foundation, Seoul, South Korea, <sup>13</sup> Department of Pathology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea, <sup>14</sup> Laboratory of Molecular Pathology and Cancer Genomics, College of Pharmacy, Seoul National University, Seoul, South Korea, <sup>15</sup> Department of Molecular Medicine and Biopharmaceutical Sciences, Graduate School of Convergence Science and Technology, Seoul National University, Seoul, South Korea

**Introduction:** The GenesWell Breast Cancer Test (BCT) is a recently developed multigene assay that predicts the risk of distant recurrence in patients with early breast cancer. Here, we analyzed the concordance of the BCT score with the Oncotype DX recurrence score (RS) for risk stratification in Asian patients with pN0-N1, hormone receptor-positive, human epidermal growth factor receptor 2 (HER2)-negative breast cancer.

**Methods:** Formalin-fixed, paraffin-embedded breast cancer tissues previously analyzed using the Oncotype DX test were assessed using the GenesWell BCT test. The risk stratification by the two tests was then compared.

**Results:** A total of 771 patients from five institutions in Korea were analyzed. According to the BCT score, 527 (68.4%) patients were classified as low risk, and 244 (31.6%) as high risk. Meanwhile, 134 (17.4%), 516 (66.9%), and 121 (15.7%) patients were categorized into the low-, intermediate-, and high-risk groups, respectively, according to the RS ranges used in the TAILORx. The BCT high-risk group was significantly associated with advanced lymph node status, whereas no association between RS risk groups and



nodal status was observed. The concordance between the two risk stratification methods in the overall population was 71.9% when the RS low-risk, and intermediate-risk groups were combined into one group. However, poor concordance was observed in patients aged  $\leq 50$  years and in those with lymph node-positive breast cancer.

**Conclusions:** The concordance between the BCT score and RS was low in women aged  $\leq 50$  years or with lymph node-positive breast cancer. Further studies are necessary to identify more accurate tests for predicting prognosis and chemotherapy benefit in this subpopulation.

**Keywords:** GenesWell BCT score, oncotype DX recurrence score, concordance, early breast cancer, risk classification, Asian population

## INTRODUCTION

Several multigene expression prognostic assays have been developed to overcome the limitations of clinical variables such as tumor size and nodal status for predicting prognosis in breast cancer (1). These assays are used to predict the risk of recurrence or distant metastasis after surgery and adjuvant hormone therapy in hormone receptor-positive early breast cancer to help treatment decisions regarding chemotherapy. MammaPrint (2) and Oncotype DX (3) are the first generation molecular prognostic assays; additional assays such as Prosigna (4–6) and EndoPredict (7) were developed later.

Oncotype DX (Genomic Health, Redwood City, CA, USA) is the most widely used multigene assay (3); it uses quantitative reverse transcription-polymerase chain reaction (qRT-PCR) to measure the expression of 21 genes in formalin-fixed, paraffin-embedded (FFPE) tissues. The Oncotype DX recurrence score (RS) also predicts the benefit of adding chemotherapy to hormone therapy in estrogen receptor (ER)-positive breast cancer (8, 9). Moreover, RS results are currently included in clinical guidelines for treatment decisions in early breast cancer (10–12). The American Joint Committee on Cancer eighth edition cancer staging system was recently revised to include this score for prognosis in breast cancer (13).

However, recent studies showed that other prognostic scores such as PAM50-based Prosigna risk of recurrence (ROR) score (6) and EPclin by EndoPredict (14) are more accurate than Oncotype DX RS for predicting the risk of distant recurrence in endocrine-treated postmenopausal patients with ER-positive breast cancer. Comparison of the prognostic value of six multigene signatures, including Clinical Treatment Score, four immunohistochemical markers (IHC4), RS, ROR, Breast Cancer Index (BCI), and EPclin in 774 postmenopausal women with ER-positive, human epidermal growth factor receptor 2 (HER2)-negative breast cancer also demonstrated that ROR, BCI, and EPclin are more prognostic for overall and late distant recurrence than RS in patients with lymph node-negative breast cancer (15). However, studies comparing Oncotype DX and other assays were performed in Western populations, and the results in Asian patients with breast cancer remain unclear.

Asian breast cancer differs from Western breast cancer in terms of age-specific incidence rates (16–18). Approximately

half of breast cancer patients (peak age: 45–50 years) are premenopausal in Asian countries, whereas 15–30% of Western breast cancer (peak age: 55–60 years) are premenopausal (19–21). In addition, distinct biological features of Asian breast cancer include higher prevalence of luminal B subtype, more frequent *TP53* mutation, and more active immune microenvironment, suggesting the needs for inclusion of more Asian women in clinical trials to unravel the ethnic difference of breast cancer (21, 22). However, most genomic algorithms for use in breast cancer tests are based on postmenopausal women in Western countries, which raises concerns regarding their prognostic or predictive value in Asian, or young breast cancer patients. Notably, recent data from the Trial Assigning Individualized Options for Treatment (TAILORx) (23) showed that there is no chemotherapy benefit in patients aged  $> 50$  years with hormone receptor-positive, HER2-negative, lymph node-negative breast cancer with a RS of 11–25, while those aged  $\leq 50$  years with a RS of 16–25 may benefit from chemotherapy. The trial results suggested that the predictive value of the RS for chemotherapy benefit or “number needed to treat (NNT)” can be different in Asian breast cancer patients, as this population includes a greater number of patients aged  $\leq 50$  years. The absolute risk reduction (ARR) and NNT for a RS of 21–25 was 6.5 and 15.4, while it was 1.6 and 62.5 for a RS of 16–20 (23), respectively. Meanwhile, the ARR and NNT for a RS  $\geq 26$  was 25.0 and 4.0, respectively (24). A recent study showed that tailored therapy based on Oncotype DX results could result in a net cost increase in initial care of American breast cancer if women aged  $\leq 50$  years with tumors with RS of 16–25 all chose to receive chemotherapy (25).

The GenesWell Breast Cancer Test (BCT) (Gencurix, Inc., Seoul, Korea) is a molecular prognostic assay that predicts the risk of 10-year distant metastasis in patients with pathologic N0 or N1 status (pN0-N1), hormone receptor-positive, HER2-negative breast cancer (26). This test is a qRT-PCR-based assay that measures the relative expression of six prognostic genes and two clinical variables using FFPE tumor tissues similar to the Oncotype DX. The ability of this assay to predict the chemotherapy benefit was also recently demonstrated in Asian breast cancer patients (27). Here, we aimed to assess the agreement in risk classification between the BCT score and the RS in a large sample of Asian breast cancer patients from multiple institutions.

## MATERIALS AND METHODS

### Patients and Tissue Samples

FFPE tumor blocks were obtained from patients meeting the following criteria: with hormone receptor-positive early breast cancer, underwent curative resection of the primary tumor at any of the five institutions (Samsung Medical Center, Asan Medical Center, Korea University Guro Hospital, Gangnam Severance Hospital in Seoul, and National Cancer Institute in Gyeonggi-do) in Korea between 2010 and 2017, and with a reportable RS. FFPE tumor tissues not eligible for the GenesWell BCT test or cases without sufficient tumor or clinical information were excluded. Hormone receptors (ER or progesterone receptor [PR]) and HER2 status were determined at local laboratories. The staining of ER or PR by immunohistochemistry (IHC) was scored using the semi-quantitative Allred score (AS) with a maximum score of 8, and AS  $\geq 2$  was considered as positive as described previously (28, 29). HER2 status was measured using the IHC, fluorescence *in situ* hybridization (FISH), or silver-enhanced *in situ* hybridization (SISH). According to the American Society of Clinical Oncology/College of American Pathologists guidelines, HER2 positivity was defined as an intensity of 3+ by IHC or as gene amplification ratio of  $\geq 2.0$  or average HER2 copy number  $\geq 6$  by FISH or SISH (30).

### Oncotype DX and BCT Tests

Samples were delivered to Genomic Health for Oncotype DX testing prior to the study. Tissue samples were prepared following the pathology guidelines of Oncotype DX. The RS results were determined by Genomic Health, as previously described (3).

Samples previously analyzed using the Oncotype DX test were used for the GenesWell BCT test. RNA was extracted from FFPE tissues, and samples containing sufficient residual RNA were subjected to qRT-PCR as previously described (26). The BCT score was calculated using two clinical variables (tumor size and nodal status) in combination with the relative expression of the six prognostic genes (*UBE2C*, *TOP2A*, *RRM2*, *FOXMI*, *MKI67*, and *BTN3A2*) (26). The expression of *ESR1*, *PGR*, and *ERBB2* was also quantified relative to the three reference genes (*CTBP1*, *CUL1*, and *UBQLN1*).

### Categorization of Risk Groups

Patients were categorized into BCT high-risk and low-risk groups according to the BCT scoring criteria reported previously (26). Briefly, patients with a BCT score  $< 4$  were classified as low risk, and those with a BCT score  $\geq 4$  were classified as high risk. For the Oncotype DX, two different RS ranges were used to classify patients. First, patients were grouped into low-risk (RS  $< 18$ ), intermediate-risk (RS 18–30), and high-risk (RS  $\geq 31$ ) groups using the originally validated cut-off (called clinical cut-off) (3). Second, patients were classified according to the RS ranges used in the TAILORx (called TAILORx cut-off) as low-risk (RS  $< 11$ ), intermediate-risk (RS 11–25), and high-risk (RS  $\geq 26$ ) groups (24, 31). Clinical risk was determined using the modified version

of Adjuvant! Online as reported in the Microarray in Node-Negative Disease May Avoid Chemotherapy (MINDACT) trial as previously described (27).

### Statistical Analysis

The association between clinicopathological parameters and the BCT score or the RS was analyzed using the Chi-square test. Chi-square test was also used to compare the distribution of each score between the subgroups. The Jonckheere-Terpstra test was used to determine trends in the association between gene expression and risk scores (32, 33). Differences were considered statistically significant at  $P < 0.05$ . All statistical analyses were performed using R 3.2.0 (<http://r-project.org>).

## RESULTS

### Patient Characteristics

The GenesWell BCT test was used to analyze 795 FFPE tissue samples from patients with pN0-N1, hormone receptor-positive, HER2-negative breast cancer with available RS results, and the BCT score was calculated for 771 patients. Sample availability is described in **Supplementary Figure 1**. The clinical characteristics of the patients included in the study are summarized in **Table 1**. All patients were Asians. The median age was 47 years (range, 23–79 years). A total of 66.7% and 33.3% of the patients were aged  $\leq 50$  years and  $> 50$  years, respectively. Most of the tumors were ductal carcinoma (85.1%), pN0 (80.3%), histologic grade 2 or 3 (82.2%), and nuclear grade 2 or 3 (91.8%).

### BCT Score-Based Risk Classification

Regarding BCT score distribution, the most common was 3–4 (30.7%), followed by 4–5.5 (27.1%) and 2–3 (22.6%) (**Figure 1A**). The BCT score distribution differed significantly between lymph node-negative and node-positive subgroups (**Figures 1B,C**) ( $P < 0.001$ ). Within each nodal subgroup, the BCT score distribution was similar between patients aged  $\leq 50$  years and those aged  $> 50$  years ( $P = 0.785$  for the lymph node-negative subgroup and  $P = 0.694$  for the node-positive subgroup) (**Figure 2**).

In the classification of patients according to the BCT score, 68.4% ( $n = 527$ ) of patients were included in the BCT low-risk group, whereas 31.6% ( $n = 244$ ) were in the BCT high-risk group (**Table 1** and **Figure 1A**). The proportion of BCT high-risk patients was higher in the node-positive (53.9%) than that in the node-negative subgroup (26.1%) (**Figures 1B,C**). Patients classified into the BCT high-risk group had significantly larger tumors ( $P < 0.001$ ), more advanced pN status ( $P < 0.001$ ), more advanced histologic grade ( $P < 0.001$ ), and higher nuclear grade ( $P < 0.001$ ) than those in the BCT low-risk group. No significant differences in age, PR status and histological type were observed between the two risk groups (**Table 1**).

### RS-Based Risk Classification

Patients were re-classified as low risk, intermediate risk, and high risk according to the RS results. The most frequent RS range was 11–15 (27.9%), followed by 18–25 (27.1%) (**Figure 1A**). The RS

**TABLE 1** | Clinical characteristics of the risk groups according to the BCT score.

			BCT score				
Characteristics	All		Low risk (<4)		High risk (≥4)		P-value
n, %	771		527	68.4%	244	31.6%	–
Age (years)							0.940
≤40	135	17.5%	90	66.7%	45	33.3%	
40–50	379	49.2%	260	68.6%	119	31.4%	
50–60	175	22.7%	122	69.7%	53	30.3%	
>60	82	10.6%	55	67.1%	27	32.9%	
ER							–
Positive	771	100.0%	527	68.4%	244	31.6%	
PR							0.470
Negative	78	10.1%	50	64.1%	28	35.9%	
Positive	693	89.9%	477	68.8%	216	31.2%	
Tumor size (cm)							<0.001
≤2.0	504	65.4%	414	82.1%	90	17.9%	
>2.0	267	34.6%	113	42.3%	154	57.7%	
pN							<0.001
0	619	80.3%	457	73.8%	162	26.2%	
1	152	19.7%	70	46.1%	82	53.9%	
Histologic grade							<0.001
1	137	17.8%	120	87.6%	17	12.4%	
2	542	70.3%	364	67.2%	178	32.8%	
3	92	11.9%	43	46.7%	49	53.3%	
Nuclear grade							<0.001
1	63	8.2%	55	87.3%	8	12.7%	
2	589	76.4%	406	68.9%	183	31.1%	
3	119	15.4%	66	55.5%	53	44.5%	
Histology							0.606
Ductal	656	85.1%	454	69.2%	202	30.8%	
Lobular	67	8.7%	44	65.7%	23	34.3%	
Mucinous	18	2.3%	10	55.6%	8	44.4%	
Others*	27	3.5%	18	66.7%	9	33.3%	
Unknown	3	0.4%	1	33.3%	2	66.7%	

\*Cribiform, ductal carcinoma with mucinous, tubular, mixed ductal and lobular, papillary, micropapillary, and metaplastic.

BCT, breast cancer test; ER, estrogen receptor; pN, pathologic nodal status; PR, progesterone receptor.

ER and PR status was assessed by immunohistochemistry.  $P < 0.05$  are marked in bold.

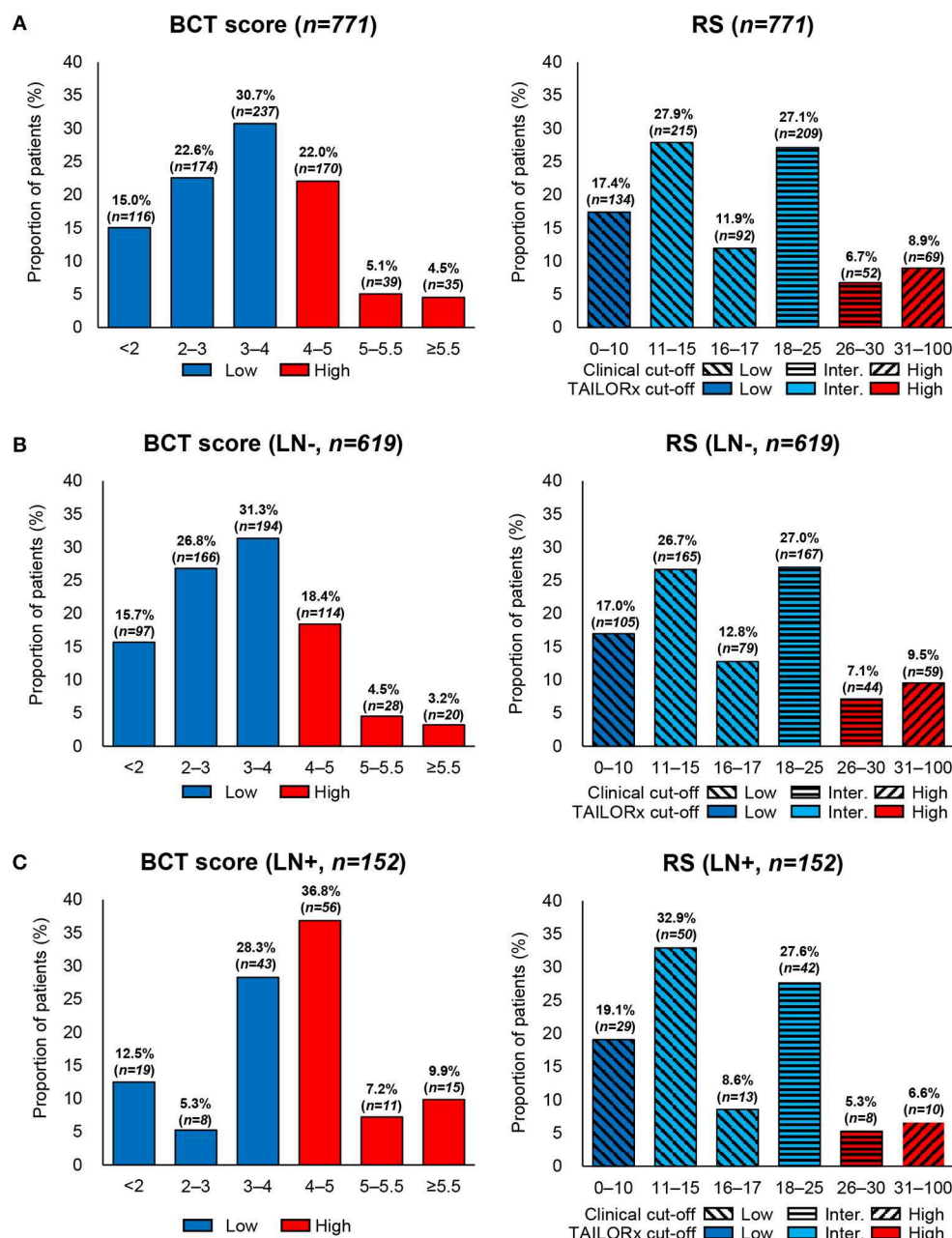
distribution was similar between the lymph node-negative and node-positive subgroups ( $P = 0.341$ ) (Figures 1B,C). However, a significant difference in the RS distribution according to age was observed in each nodal subgroup ( $P = 0.020$  for the lymph node-negative and  $P = 0.035$  for the node-positive subgroup) (Figure 2).

Using the original clinical cut-off, 441 (57.2%), 261 (33.9%), and 69 (8.9%) patients were classified as low risk, intermediate risk, and high risk, respectively (Supplementary Table 1). Meanwhile, based on the RS ranges used in TAILORx, 134 (17.4%), 516 (66.9%), and 121 (15.7%) patients were categorized as low risk, intermediate risk, and high risk, respectively (Supplementary Table 1). Compared with the risk classification using the original clinical cut-off, the TAILORx cut-off categorized more patients as intermediate risk and fewer as low risk.

The proportion of patients classified into the high-risk group according to the RS (8.9% using the clinical cut-off and 15.7% using the TAILORx cut-off) was lower than that of patients classified according to the BCT score (31.6%). In contrast to the BCT high-risk group, the RS high-risk group was not significantly associated with advanced pN status. Negative PR status was significantly correlated with a high RS ( $P < 0.001$ ) (Supplementary Table 1).

## Concordance Between the BCT Score and the RS

The concordance in risk stratification between the BCT score and the RS was analyzed using the RS ranges of TAILORx. The overall concordance between the two risk classifications was 71.9% when the RS low-risk and intermediate-risk groups were combined into one group (non-high-risk group, RS 0–25) (Table 2). Of



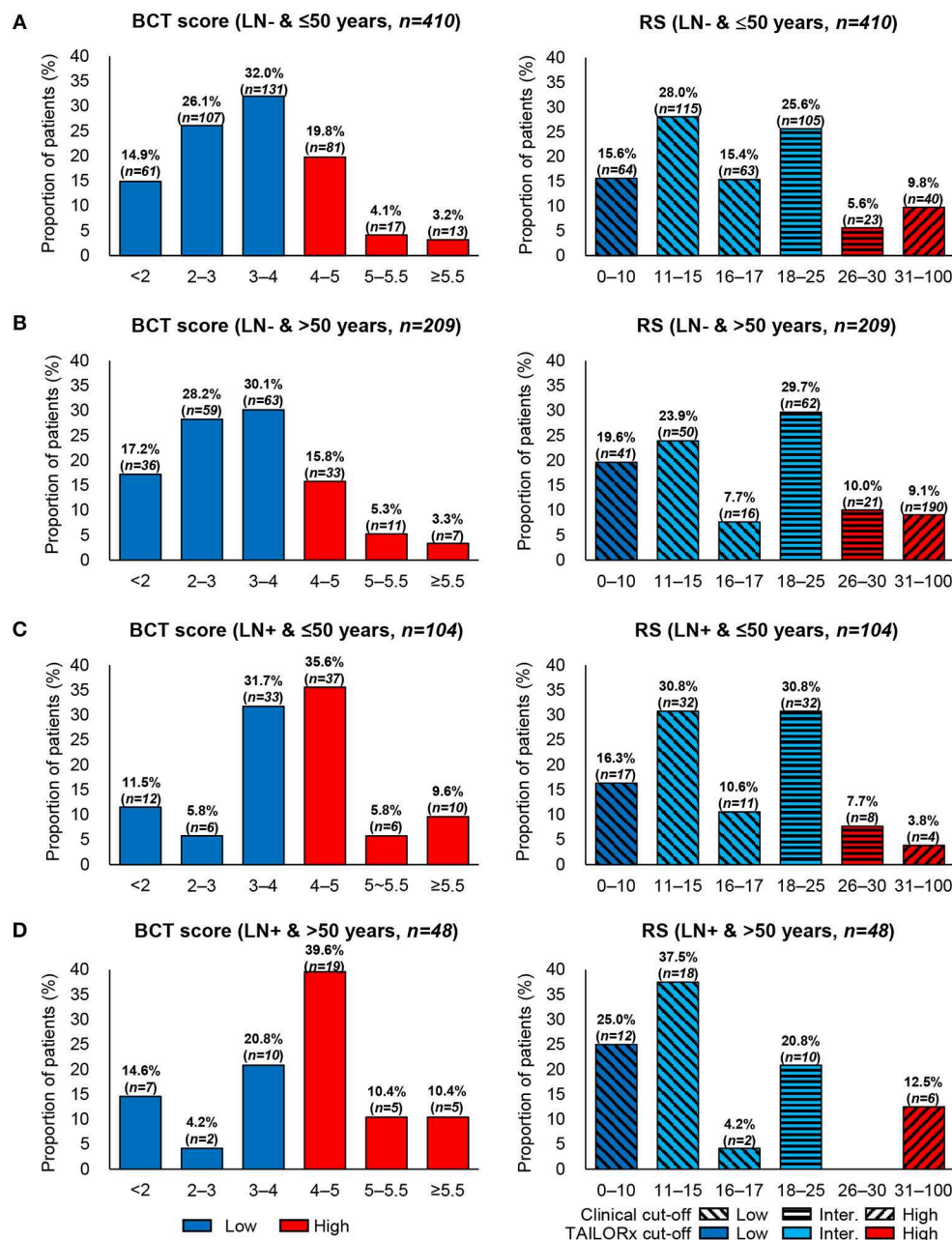
**FIGURE 1 |** Distribution of the BCT score and Oncotype DX RS by nodal status. Proportion of patients within each risk score range or risk group in **(A)** all patients ( $n = 771$ ), **(B)** lymph node-negative (LN-) patients ( $n = 619$ ), and **(C)** lymph node-positive (LN+) patients ( $n = 152$ ).

527 patients in the BCT low-risk group, 480 (91.9%) were classified as non-high risk according to the RS. Subgroup analysis according to nodal status showed that the concordance between the two scores was different in the lymph node-negative and node-positive subgroups. The overall concordance was higher in the lymph node-negative subgroup (76.6%) than that in the node-positive subgroup (52.6%) (Table 2).

We also assessed the concordance between the two scores according to age:  $\leq 50$  years and  $> 50$  years. Based on recent

findings on the benefits of chemotherapy for patients with a RS midrange score (11–25) from TAILORx (23), patients were categorized into chemobenefit and non-chemobenefit groups using different RS ranges for each age subgroup. In patients aged  $\leq 50$  years, those with RS 0–15 and RS  $\geq 16$  were categorized into non-chemobenefit and chemobenefit groups, respectively, whereas in patients aged  $> 50$  years, the RS ranges used for the classification into non-chemobenefit and chemobenefit groups were RS 0–25 and RS  $\geq 26$ , respectively.





**FIGURE 2 |** Distribution of the BCT score and Oncotype DX RS by age within each nodal subgroup. Proportion of patients within each risk score range according to age and nodal status. **(A)** Patients aged ≤50 years with lymph node-negative (LN-) breast cancer ( $n = 410$ ). **(B)** Patients aged >50 years with lymph node-negative (LN-) breast cancer ( $n = 209$ ). **(C)** Patients aged ≤50 years with lymph node-positive (LN+) breast cancer ( $n = 104$ ). **(D)** Patients aged >50 years with lymph node-positive (LN+) breast cancer ( $n = 48$ ).

The overall concordance was higher in women aged >50 years (72.8%) than in those aged ≤50 years (52.9%) (Table 2). However, in each nodal subgroup, the concordance results differed between patients aged ≤50 years and those aged >50 years. In patients with lymph node-negative breast cancer, the concordance was higher in those aged >50 years (77.5%)

than in those ≤50 years (53.2%) (Table 2). By contrast, in the lymph node-positive subgroup, the concordance was similar between patients aged >50 years (52.1%), and ≤50 years (51.9%) (Table 2). The highest concordance between the two scores was observed in patients aged >50 years with lymph node-negative breast cancer.



**TABLE 2 |** Concordance in risk stratification between the BCT score and Oncotype DX RS according to nodal status and age.

		All ( <i>n</i> = 771)			Lymph node-negative ( <i>n</i> = 619)			Lymph node-positive ( <i>n</i> = 152)		
		Oncotype DX RS (TAILORx cut-off)			Oncotype DX RS (TAILORx cut-off)			Oncotype DX RS (TAILORx cut-off)		
<i>n</i> (%)		Non-high risk (0–25)	High risk (≥26)	Total	Non-high risk (0–25)	High risk (≥26)	Total	Non-high risk (0–25)	High risk (≥26)	Total
BCT score	Low risk (<4)	480 (62.3%)	47 (6.1%)	527 (68.4%)	414 (66.9%)	43 (6.9%)	457 (73.8%)	66 (43.4%)	4 (2.6%)	70 (46.1%)
	High risk (≥4)	170 (22.0%)	74 (9.6%)	244 (31.6%)	102 (16.5%)	60 (9.7%)	162 (26.2%)	68 (44.7%)	14 (9.2%)	82 (53.9%)
	Total	650 (84.3%)	121 (15.7%)	771 (100.0%)	516 (83.4%)	103 (16.6%)	619 (100.0%)	134 (88.2%)	18 (11.8%)	152 (100.0%)
<b>≤50 YEARS</b>										
		Oncotype DX RS (TAILORx cut-off)			Oncotype DX RS (TAILORx cut-off)			Oncotype DX RS (TAILORx cut-off)		
<i>n</i> (%)		Non- chemobenefit (0–15)	Chemobenefit (≥16)	Total	Non- chemobenefit (0–15)	Chemobenefit (≥16)	Total	Non- chemobenefit (0–15)	Chemobenefit (≥16)	Total
BCT score	Low risk (<4)	168 (32.7%)	182 (35.4%)	350 (68.1%)	143 (34.9%)	156 (38.0%)	299 (72.9%)	25 (24.0%)	26 (25.0%)	51 (49.0%)
	High risk (≥4)	60 (11.7%)	104 (20.2%)	164 (31.9%)	36 (8.8%)	75 (18.3%)	111 (27.1%)	24 (23.1%)	29 (27.9%)	53 (51.0%)
	Total	228 (44.4%)	286 (55.6%)	514 (100.0%)	179 (43.7%)	231 (56.3%)	410 (100.0%)	49 (47.1%)	55 (52.9%)	104 (100.0%)
<b>&gt;50 YEARS</b>										
		Oncotype DX RS (TAILORx cut-off)			Oncotype DX RS (TAILORx cut-off)			Oncotype DX RS (TAILORx cut-off)		
<i>n</i> (%)		Non- chemobenefit (0–25)	Chemobenefit (≥26)	Total	Non- chemobenefit (0–25)	Chemobenefit (≥26)	Total	Non- chemobenefit (0–25)	Chemobenefit (≥26)	Total
BCT score	Low risk (<4)	159 (61.9%)	18 (7.0%)	177 (68.9%)	140 (67.0%)	18 (8.6%)	158 (75.6%)	19 (39.6%)	0 (0.0%)	19 (39.6%)
	High risk (≥4)	52 (20.2%)	28 (10.9%)	80 (31.1%)	29 (13.9%)	22 (10.5%)	51 (24.4%)	23 (47.9%)	6 (12.5%)	29 (60.4%)
	Total	211 (82.1%)	46 (17.9%)	257 (100.0%)	169 (80.9%)	40 (19.1%)	209 (100.0%)	42 (87.5%)	6 (12.5%)	48 (100.0%)

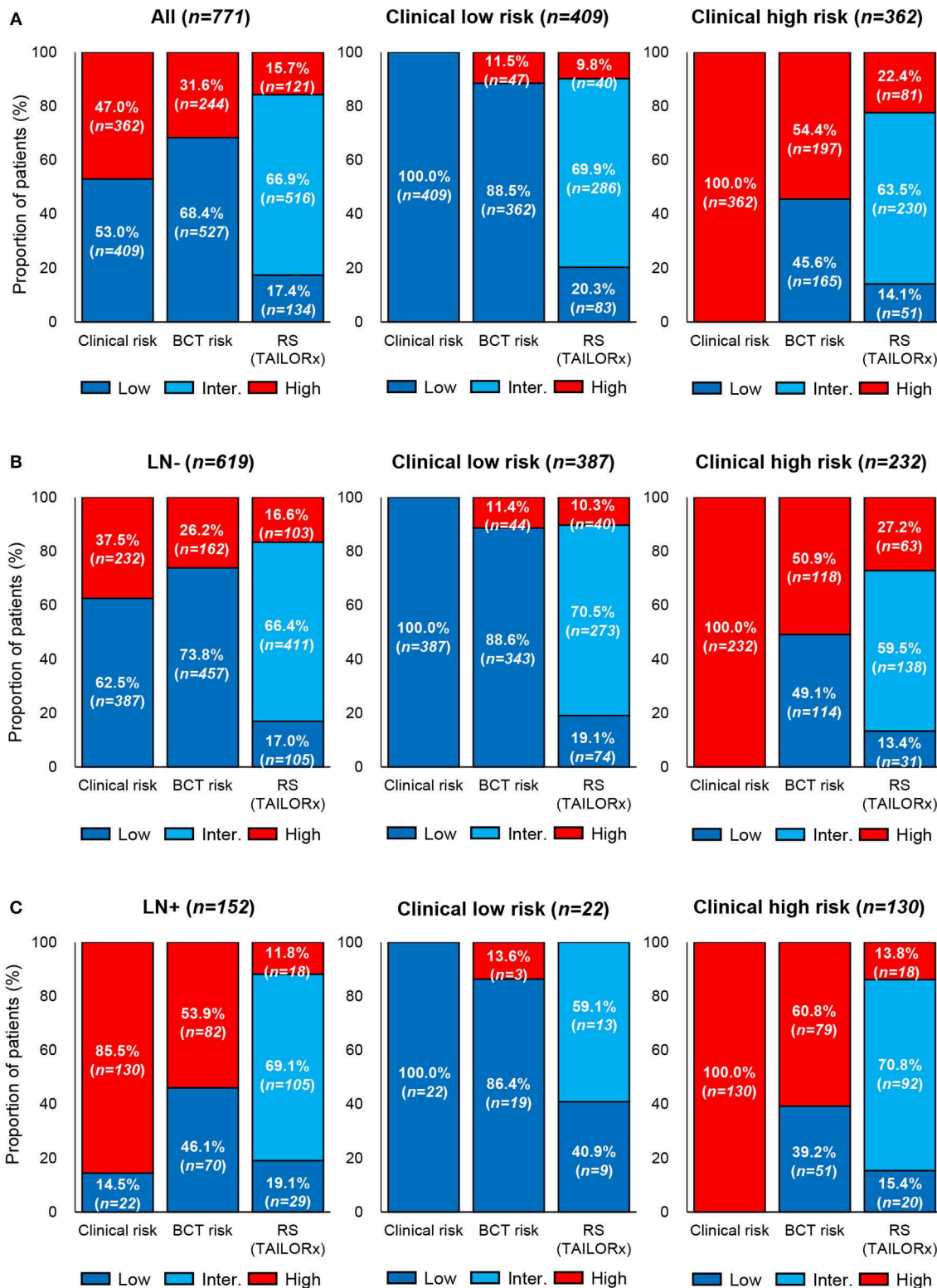
BCT, Breast Cancer Test; RS, recurrence score; TAILORx, Trial Assigning Individualized Options for Treatment.

## Comparison of Clinical Risk by Modified Adjuvant! Online With the BCT Score and the RS

The clinical risk of patients was examined using the modified Adjuvant! Online, and the clinical risk classification was compared with that obtained using the BCT score or the RS. Overall, 409 (53.0%), and 362 (47.0%) patients were categorized as clinical low risk and high risk, respectively (**Figure 3A**). Among patients in the clinical low-risk group, 11.5 and 9.8% were categorized as BCT high risk and RS high risk (≥26), respectively. Among patients in the clinical high-risk group, 45.6% and 77.6% were classified as BCT low risk and RS

non-high risk (0–25), respectively. The clinical risk classification according to nodal status was different. The proportion of patients categorized as clinical high risk was higher in the lymph node-positive subgroup (85.5%) than that in the node-negative subgroup (37.5%) (**Figures 3B,C**). The difference between the clinical risk and the risk stratification using the two tests was greater in the lymph node-positive subgroup than that in the node-negative subgroup.

Of note, a recent secondary analysis of TAILORx trial on the integration of clinical risk to RS showed that the RS ranges predicting chemotherapy benefit are different in young women aged ≤50 years according to clinical risk (34). Clinical low-risk



**FIGURE 3 |** Comparison of clinical risk with the risk classification by the BCT score or Oncotype DX RS. Proportion of patients within each risk group according to clinical risk assessment, BCT score, or RS in **(A)** all patients ( $n = 771$ ), **(B)** lymph node-negative (LN-) patients ( $n = 619$ ), and **(C)** lymph node-positive (LN+) patients ( $n = 152$ ). Clinical risk was determined using the modified Adjuvant! Online, as reported in the MINDACT trial. Risk classification by the RS was based on the recurrence score ranges used in the TAILORx.

patients with RS 0–20 and RS  $\geq 21$  were categorized into non-chemobenefit and chemobenefit groups, whereas in clinical high-risk group, the RS ranges used for the classification into non-chemobenefit and chemobenefit groups were RS 0–15 and RS  $\geq 16$ , respectively. Based on these findings, we further assessed the concordance between the BCT score and the RS in young patients aged  $\leq 50$  years. The overall concordance between the two risk classifications was 66.3% (341/514) and a higher concordance was observed in lymph node-negative subgroup (69.3% [284/410]) than node-positive subgroup (54.8% [57/104]) (Table 3).

Figure 4 shows the discordant results between the clinical risk and the risk classification using the two tests according to age within each nodal subgroup. In both nodal subgroups, the proportion of patients with discordant results between the clinical risk and risk by BCT score (i.e., either clinical low risk and BCT high risk or clinical high risk and BCT low risk) according to age was similar. By contrast, there was a difference in the proportion of patients with discordant results between the clinical risk and RS risk (i.e., either clinical low risk and RS chemobenefit or clinical high risk and RS non-chemobenefit) according to age. The RS categorized a higher proportion of patients into the chemobenefit group among clinical low-risk patients aged  $\leq 50$  years (21.2% [55/259] in the lymph node-negative subgroup and 12.5% [2/16] in the node-positive subgroup) than among those aged  $> 50$  years (10.2% [13/128] in the lymph node-negative subgroup and 0% [0/6] in the node-positive subgroup). Meanwhile, the proportion of RS non-chemobenefit patients among clinical high-risk patients was higher in women aged  $> 50$  years (66.7% [54/81] in the lymph node-negative subgroup and 85.7% [36/42] in the node-positive subgroup) than in those aged  $\leq 50$  years (37.7% [57/151] in the lymph node-negative subgroup and 43.2% [38/88] in the node-positive subgroup).

The risk stratification using the two tests in clinical high- or low-risk patients was different in specific subpopulations. In patients aged  $\leq 50$  years within the lymph node-negative subgroup ( $n = 259$ ), 21.2% of clinical low-risk patients were categorized into the chemobenefit group according to the RS, whereas 12.7% of patients were categorized as BCT high risk (Figure 4A). Among clinical high-risk patients aged  $> 50$  years in the lymph node-positive subgroup ( $n = 42$ ), 33.3 and 85.7% were classified as BCT low risk and non-chemobenefit, respectively, according to the RS (Figure 4D).

The prognostic value of the two scores was difficult to compare because of the short follow-up period. However, seven patients developed distant metastasis after surgery during the follow-up period in the present study. Both the BCT score and the RS categorized four of these patients as high risk (Supplementary Table 2).

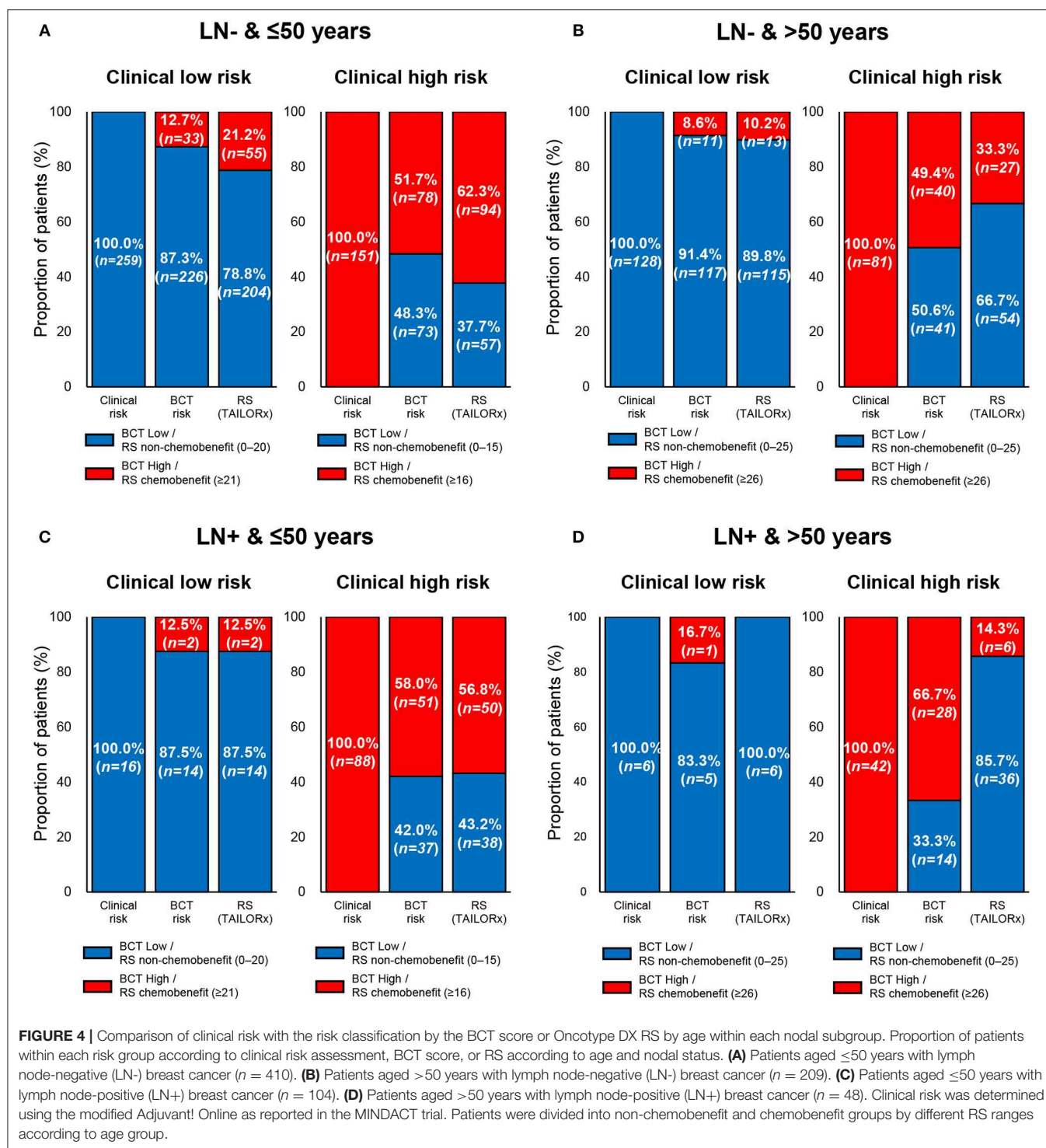
## Correlation of ER/PR/HER2 Expression With the BCT Score

The association of the two scores with the gene expression of *ESR1*, *PGR*, and *ERBB2* was assessed. Consistent with the RS algorithm including *ESR1* and *PGR* expression, there was a statistically significant trend toward lower *ESR1* and *PGR*

**TABLE 3 |** Concordance in risk stratification between the BCT score and Oncotype DX RS in patients aged  $\leq 50$  years according to clinical risk.

n (%)	All (n = 514)			Lymph node-negative (n = 410)			Lymph node-positive (n = 104)		
	Oncotype DX RS			Oncotype DX RS			Oncotype DX RS		
	(TAILORx cut-off according to clinical risk)			(TAILORx cut-off according to clinical risk)			(TAILORx cut-off according to clinical risk)		
	Non-chemobenefit (0–15 for clinical high risk or 0–20 for clinical low risk)	Chemobenefit ( $\geq 21$ for clinical low risk or $\geq 16$ for clinical high risk)	Total	Non-chemobenefit (0–15 for clinical high risk or 0–20 for clinical low risk)	Chemobenefit ( $\geq 21$ for clinical low risk or $\geq 16$ for clinical high risk)	Total	Non-chemobenefit (0–15 for clinical high risk or 0–20 for clinical low risk)	Chemobenefit ( $\geq 21$ for clinical low risk or $\geq 16$ for clinical high risk)	Total
BCT score									
Low risk (<4)	245 (47.7%)	105 (20.4%)	350 (68.1%)	217 (52.9%)	82 (20.0%)	299 (72.9%)	28 (26.9%)	23 (22.1%)	51 (49.0%)
High risk ( $\geq 4$ )	68 (13.2%)	96 (18.7%)	164 (31.9%)	44 (10.7%)	67 (16.3%)	111 (27.1%)	24 (23.1%)	29 (27.9%)	53 (51.0%)
Total	313 (60.9%)	201 (39.1%)	514 (100.0%)	261 (63.7%)	149 (36.3%)	410 (100.0%)	52 (50.0%)	52 (50.0%)	104 (100.0%)

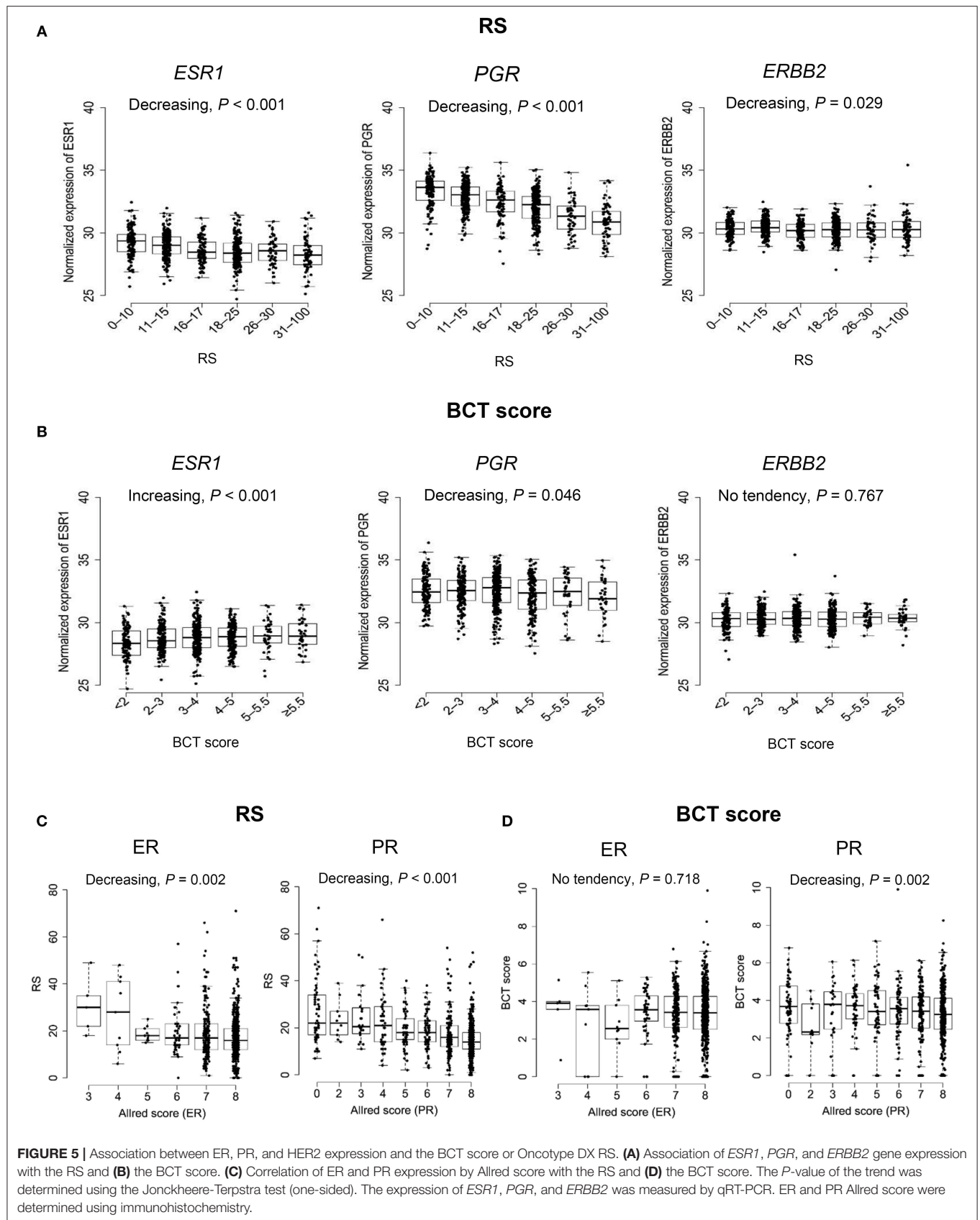
BCT, Breast Cancer Test; RS, recurrence score; TAILORx, Trial Assigning Individualized Options for Treatment. Clinical risk was determined using the modified Adjuvant! Online as reported in the MINDACT trial.



expression among patients with a higher RS (Jonckheere-Terpstra test,  $P < 0.001$ ) (Figure 5A). Similarly, *PGR* expression showed a decreasing trend in correlation with the BCT score ( $P = 0.046$ ) (Figure 5B). However, *ESR1* expression increased as the BCT score increased ( $P < 0.001$ ). *ERBB2* expression showed a decreasing trend as the RS increased ( $P = 0.029$ ), whereas

no significant association between *ERBB2* expression and the BCT score was observed. We also evaluated the correlation of the two scores with ER and PR expression by IHC. Negative correlation of ER ( $P = 0.002$ ), and PR expression ( $P < 0.001$ ) with the RS was observed (Figure 5C). There was no significant association between ER expression and the BCT score, whereas







BCT score showed a negative correlation with PR expression ( $P = 0.002$ ) (Figure 5D).

## Correlation of the RS With BCT Prognostic Genes

The correlation between the expression of six prognostic genes included in the BCT score and the RS was also examined. There was a statistically significant trend toward a higher expression of five proliferation-related genes (*UBE2C*, *TOP2A*, *RRM2*, *FOXM1*, and *MK167*) among patients with a higher RS (Jonckheere-Terpstra test,  $P < 0.001$ ) (Supplementary Figure 2). Although the expression of the immune response-related gene *BTN3A2* was negatively associated with the BCT score, it showed an increasing trend in correlation with the RS ( $P = 0.027$ ).

## DISCUSSION

The present study is the first to compare the BCT score and the RS for the risk classification of Asian patients with pN0-N1, hormone receptor-positive, HER2-negative breast cancer. The study is notable because of the inclusion of a large population of Asian patients from several institutions.

The present results showed a moderate concordance of 71.9% between the two scores for risk stratification using the RS ranges reported in TAILORx. The discrepancy in the risk classification between the BCT score and RS may be attributable to the different gene sets and algorithms used to calculate the score. Moreover, the BCT score algorithm includes clinical factors (tumor size and nodal status), which are not included in the RS. When compared the RS risk group distribution in this study with previous studies, similar distribution was observed. In the present study, 105 (17.0%), 411 (66.4%), and 103 (16.6%) patients were classified as low risk, intermediate risk, and high risk in lymph node-negative subgroup using TAILORx cut-off (Supplementary Table 1), which are similar to results from a TAILORx trial (low risk, 16.7%; intermediate risk, 69.0%, and high risk, 14.3%) (23). RS pooled risk group distribution from several studies was: low risk, 52.6%; intermediate risk, 35.9%, and high risk, 11.5%, respectively, when RS risk groups were defined using the original clinical cut-off (35). These results are also similar to our findings.

The results showed that the agreement between the BCT score and the RS differed according to nodal status and age. Better concordance was found in the lymph node-negative subgroup than in the node-positive subgroup and in patients aged  $>50$  years than in those  $\leq 50$  years. Accordingly, the highest concordance between the two scores for risk classification was observed in patients aged  $>50$  years with lymph node-negative breast cancer. This was related to the differences in risk assignment by the BCT score or the RS according to nodal status or age. The poor concordance in the lymph node-positive subgroup may be associated with the different risk assignment by the BCT score between the two subgroups. The proportion of patients classified as high risk according to the BCT was higher in lymph node-positive than that in node-negative patients, whereas

the RS yielded a similar pattern of risk assignment between the two subgroups. Given that advanced nodal status is a strong unfavorable prognostic factor (36, 37), it is not surprising that the proportion of patients categorized as BCT high risk was higher in the lymph node-positive subgroup than that in the node-negative subgroup. By contrast, the distribution of RS ranges differed between the two age subgroups, whereas the BCT score distribution was similar in each age subgroup. This may explain the large difference in risk stratification by the two risk scores in women aged  $\leq 50$  years.

Following the previous TAILORx results, a recent secondary analysis of TAILORx trial further found that clinical risk stratification provided additional prognostic information to hormone receptor-positive, HER2-negative, lymph node-negative breast cancer patients aged  $\leq 50$  years with RS 16–25 (34). Importantly, the study showed that there was no benefit from chemotherapy for women aged  $\leq 50$  years with RS 16–20 and at clinical low risk, whereas patients with RS 16–25 and at clinical high risk do benefit from chemotherapy. Based on these results, we categorized patients aged  $\leq 50$  years into non-chemobenefit and chemobenefit groups using different RS ranges according to clinical risk. Patients with RS 0–20 and RS  $\geq 21$  were categorized into non-chemobenefit and chemobenefit groups in clinical low-risk group, whereas in the clinical high-risk group, the RS ranges used for the classification of non-chemobenefit and chemobenefit groups were RS 0–15 and RS  $\geq 16$ , respectively and we assessed the concordance in risk stratification between the two tests. Similar to the agreement between the two risk classifications not considering clinical risk, the concordance in patients aged  $\leq 50$  years was lower than that in patients aged  $>50$  years. The agreement between clinical risk and risk stratification using the two tests varied depending on age. In the subgroup analysis by age in each nodal subgroup, the proportion of patients with discordant results between clinical risk and RS risk was different between patients aged  $\leq 50$  years and those  $>50$  years. The risk stratification using the two tests in clinical high- or low-risk patients was different in specific subpopulations including patients aged  $\leq 50$  years with lymph node-negative breast cancer and patients aged  $>50$  years with lymph node-positive breast cancer. These results raised a question regarding which risk stratification is more appropriate in these subpopulations. Moreover, these results suggest the need for further studies to identify more accurate risk score for predicting the risk of recurrence or chemotherapy benefit in Asian breast cancer patients aged  $\leq 50$  years.

Because the clinical data was based on a short follow-up period, a direct comparison of the prognostic and predictive values of the BCT score with the RS was not possible in this study. Therefore, the results are not sufficient to determine which test is more accurate for predicting the risk of recurrence or chemotherapy benefit in hormone receptor-positive, HER2-negative early breast cancer. However, the BCT high-risk group was significantly associated with larger tumor size and advanced nodal status, whereas the RS showed no significant relationship with nodal status. Moreover, in a recent study that compared the prognostic value of six multigene signatures in postmenopausal patients with ER-positive, HER2-negative breast

cancer, combined genomic and clinical models such as ROR and EPclin were more prognostic for late distant recurrence than other molecular signatures in lymph node-positive patients (15). These findings suggest that the BCT score based on combined gene expression and clinical variables, is likely to have a better prognostic value than RS in lymph node-positive patients.

## CONCLUSIONS

The present results showed a moderate accordance in risk assignment between the two scores, whereas the concordance was lower in patients aged  $\leq 50$  years or those with lymph node-positive disease. Further studies are necessary to directly compare the prognostic and predictive values of the two tests in Asian breast cancer patients aged  $\leq 50$  years.

## DATA AVAILABILITY

All datasets generated for this study are included in the manuscript and/or the **Supplementary Files**.

## ETHICS STATEMENT

The study was approved by the review board of five institutions (Samsung Medical Center, Asan Medical Center, Korea University Guro Hospital, Gangnam Severance Hospital in Seoul, and National Cancer Institute in Gyeonggi-do) in Korea

and was performed in accordance with the Declaration of Helsinki. Because the study was retrospective in nature, the requirement for informed consent was waived.

## AUTHOR CONTRIBUTIONS

YKS and GG conceived the study and participated in its design. JEL, JJ, SUW, SBL, SL, Y-LC, and YK were involved in data acquisition. MJK and JH drafted the manuscript. MJK, JEL, JJ, SUW, JH, GG, and YKS analyzed and interpreted the data. JH performed statistical analyses. BK, J-EK, YM, and KS provided administrative, technical, or material support. JEL, JJ, SUW, GG, and YKS participated in critical revision of the manuscript with respect to important intellectual content. YKS supervised the study. All authors read and approved the final manuscript.

## FUNDING

We are very grateful for the financial support of the Research Institute of Pharmaceutical Sciences, Seoul National University College of Pharmacy.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2019.00667/full#supplementary-material>

## REFERENCES

- Gyorffy B, Hatzis C, Sanft T, Hofstatter E, Aktas B, Pusztai L. Multigene prognostic tests in breast cancer: past, present, future. *Breast Cancer Res.* (2015) 17:11. doi: 10.1186/s13058-015-0514-2
- van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med.* (2002) 347:1999–2009. doi: 10.1056/NEJMoa021967
- Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med.* (2004) 351:2817–26. doi: 10.1056/NEJMoa041588
- 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) 27:1160–7. doi: 10.1200/JCO.2008.18.1370
- Nielsen TO, Parker JS, Leung S, Voduc D, Ebbert M, Vickery T, et al. A comparison of PAM50 intrinsic subtyping with immunohistochemistry and clinical prognostic factors in tamoxifen-treated estrogen receptor-positive breast cancer. *Clin Cancer Res.* (2010) 16:5222–32. doi: 10.1158/1078-0432.CCR-10-1282
- Dowsett M, Sestak I, Lopez-Knowles E, Sidhu K, Dunbier AK, Cowens JW, et al. Comparison of PAM50 risk of recurrence score with oncotype DX and IHC4 for predicting risk of distant recurrence after endocrine therapy. *J Clin Oncol.* (2013) 31:2783–90. doi: 10.1200/JCO.2012.46.1558
- Filipits M, Rudas M, Jakesz R, Dubsky P, Fitzal F, Singer CF, et al. A new molecular predictor of distant recurrence in ER-positive, HER2-negative breast cancer adds independent information to conventional clinical risk factors. *Clin Cancer Res.* (2011) 17:6012–20. doi: 10.1158/1078-0432.CCR-11-0926
- Paik S, Tang G, Shak S, Kim C, Baker J, Kim W, et al. Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *J Clin Oncol.* (2006) 24:3726–34. doi: 10.1200/JCO.2005.04.7985
- Albain KS, Barlow WE, Shak S, Hortobagyi GN, Livingston RB, Yeh IT, et al. Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial. *Lancet Oncol.* (2010) 11:55–65. doi: 10.1016/S1470-2045(09)70314-6
- Harris LN, Ismaila N, McShane LM, Andre F, Collyar DE, Gonzalez-Angulo AM, et al. Use of Biomarkers to guide decisions on adjuvant systemic therapy for women with early-stage invasive breast cancer: American society of clinical oncology clinical practice guideline. *J Clin Oncol.* (2016) 34:1134–50. doi: 10.1200/JCO.2015.65.2289
- Coates AS, Winer EP, Goldhirsch A, Gelber RD, Gnant M, Piccart-Gebhart M, et al. Tailoring therapies—improving the management of early breast cancer: St Gallen international expert consensus on the primary therapy of early breast cancer 2015. *Ann Oncol.* (2015) 26:1533–46. doi: 10.1093/annonc/mdv221
- National Comprehensive Cancer Network. *NCCN Clinical Practice Guidelines in Oncology: Breast Cancer Version 2.2016*. (2017). Available online at: [http://www.nccn.org/professionals/physician\\_gls/pdf/breast.pdf](http://www.nccn.org/professionals/physician_gls/pdf/breast.pdf)
- Giuliano AE, Connolly JL, Edge SB, Mittendorf EA, Rugo HS, Solin LJ, et al. Breast cancer—major changes in the American joint committee on cancer eighth edition cancer staging manual. *CA Cancer J Clin.* (2017) 67:290–303. doi: 10.3322/caac.21393
- Buus R, Sestak I, Kronenwett R, Denkert C, Dubsky P, Krappmann K, et al. Comparison of endopredict and epclin with oncotype DX recurrence score for prediction of risk of distant recurrence after endocrine therapy. *J Natl Cancer Inst.* (2016) 108:djw149. doi: 10.1093/jnci/djw149
- Sestak I, Buus R, Cuzick J, Dubsky P, Kronenwett R, Denkert C, et al. Comparison of the performance of 6 prognostic signatures for estrogen receptor-positive breast cancer: a secondary analysis of a randomized clinical trial. *JAMA Oncol.* (2018) 4:545–53. doi: 10.1001/jamaoncol.2017.5524
- Althuis MD, Dozier JM, Anderson WF, Devesa SS, Brinton LA. Global trends in breast cancer incidence and mortality 1973–1997. *Int J Epidemiol.* (2005) 34:405–12. doi: 10.1093/ije/dyh414

17. Shin HR, Joubert C, Boniol M, Hery C, Ahn SH, Won YJ, et al. Recent trends and patterns in breast cancer incidence among Eastern and Southeastern Asian women. *Cancer Causes Control*. (2010) 21:1777–85. doi: 10.1007/s10552-010-9604-8
18. Liu L, Zhang J, Wu AH, Pike MC, Deapen D. Invasive breast cancer incidence trends by detailed race/ethnicity and age. *Int J Cancer*. (2012) 130:395–404. doi: 10.1002/ijc.26004
19. Leong SP, Shen ZZ, Liu TJ, Agarwal G, Tajima T, Paik NS, et al. Is breast cancer the same disease in Asian and Western countries? *World J Surg*. (2010) 34:2308–24. doi: 10.1007/s00268-010-0683-1
20. Min SY, Kim Z, Hur MH, Yoon CS, Park EH, Jung KW, et al. The basic facts of Korean breast cancer in 2013: results of a nationwide survey and breast cancer registry database. *J Breast Cancer*. (2016) 19:1–7. doi: 10.4048/jbc.2016.19.1.1
21. Kan Z, Ding Y, Kim J, Jung HH, Chung W, Lal S, et al. Multi-omics profiling of younger Asian breast cancers reveals distinctive molecular signatures. *Nat Commun*. (2018) 9:1725. doi: 10.1038/s41467-018-04129-4
22. Yap YS, Lu YS, Tamura K, Lee JE, Ko EY, Park YH, et al. Insights into breast cancer in the east vs. the west: a review. *JAMA Oncol*. (2019). doi: 10.1001/jamaoncol.2019.0620. [Epub ahead of print].
23. Sparano JA, Gray RJ, Makower DF, Pritchard KI, Albain KS, Hayes DE, et al. Adjuvant chemotherapy guided by a 21-gene expression assay in breast cancer. *N Engl J Med*. (2018) 379:111–21. doi: 10.1056/NEJMoa1804710
24. Sparano JA, Paik S. Development of the 21-gene assay and its application in clinical practice and clinical trials. *J Clin Oncol*. (2008) 26:721–8. doi: 10.1200/JCO.2007.15.1068
25. Mariotto A, Jayasekera J, Petkov V, Schechter CB, Enewold L, Helzlsouer KJ, et al. Expected Monetary impact of oncotype DX score-concordant systemic breast cancer therapy based on the TAILORx Trial. *J Natl Cancer Inst*. (2019) 112. doi: 10.1093/jnci/djz068. [Epub ahead of print].
26. Gong G, Kwon MJ, Han J, Lee HJ, Lee SK, Lee JE, et al. A new molecular prognostic score for predicting the risk of distant metastasis in patients with HR+/HER2- early breast cancer. *Sci Rep*. (2017) 7:45554. doi: 10.1038/srep45554
27. Kwon MJ, Lee SB, Han J, Lee JE, Lee JW, Gong G, et al. BCT score predicts chemotherapy benefit in Asian patients with hormone receptor-positive, HER2-negative, lymph node-negative breast cancer. *PLoS ONE*. (2018) 13:e0207155. doi: 10.1371/journal.pone.0207155
28. Harvey JM, Clark GM, Osborne CK, Allred DC. Estrogen receptor status by immunohistochemistry is superior to the ligand-binding assay for predicting response to adjuvant endocrine therapy in breast cancer. *J Clin Oncol*. (1999) 17:1474–81. doi: 10.1200/JCO.1999.17.5.1474
29. Choi YL, Oh E, Park S, Kim Y, Park YH, Song K, et al. Triple-negative, basal-like, and quintuple-negative breast cancers: better prediction model for survival. *BMC Cancer*. (2010) 10:507. doi: 10.1186/1471-2407-10-507
30. Wolff AC, Hammond ME, Hicks DG, Dowsett M, McShane LM, Allison KH, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American society of clinical oncology/college of American pathologists clinical practice guideline update. *Arch Pathol Lab Med*. (2014) 138:241–56. doi: 10.5858/arpa.2013-0953-SA
31. Sparano JA, Gray RJ, Makower DF, Pritchard KI, Albain KS, Hayes DE, et al. Prospective Validation of a 21-gene expression assay in breast cancer. *N Engl J Med*. (2015) 373:2005–14. doi: 10.1056/NEJMoa1510764
32. Jonckheere R. A distribution-free k-sample test against ordered alternatives. *Biometrika*. (1954) 41:133–45. doi: 10.1093/biomet/41.1-2.133
33. Terpstra TJ. The asymptotic normality and consistency of Kendall's test against trend, when ties are present in one ranking. *Indagationes Mathematicae*. (1952) 14:327–33. doi: 10.1016/S1385-7258(52)50043-X
34. Sparano JA, Gray RJ, Ravdin PM, Makower DF, Pritchard KI, Albain KS, et al. Clinical and genomic risk to guide the use of adjuvant therapy for breast cancer. *N Engl J Med*. (2019). 380:2395–405. doi: 10.1056/NEJMoa1904819
35. Varga Z, Sinn P, Fritzsche F, von Hochstetter A, Noske A, Schraml P, et al. Comparison of EndoPredict and Oncotype DX test results in hormone receptor positive invasive breast cancer. *PLoS ONE*. (2013) 8:e58483. doi: 10.1371/journal.pone.0058483
36. Ciarrocca M, Goldstein LJ. Prognostic and predictive factors in early-stage breast cancer. *Oncologist*. (2004) 9:606–16. doi: 10.1634/theoncologist.9-6-606
37. Donegan WL. Tumor-related prognostic factors for breast cancer. *CA Cancer J Clin*. (1997) 47:28–51. doi: 10.3322/canjclin.47.1.28

**Conflict of Interest Statement:** JH, BK, J-EK, and YM are salaried employees of Gencurix. YKS holds a patent application related to the content of this article.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Kwon, Lee, Jeong, Woo, Han, Kang, Kim, Moon, Lee, Lee, Choi, Kwon, Song, Gong and Shin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# PAM50 Subtypes in Baseline and Residual Tumors Following Neoadjuvant Trastuzumab-Based Chemotherapy in HER2-Positive Breast Cancer: A Consecutive-Series From a Single Institution

Sonia Pernas<sup>1\*</sup>, Anna Petit<sup>2</sup>, Fina Climent<sup>2</sup>, Laia Paré<sup>3</sup>, J. Perez-Martin<sup>4</sup>, Luz Ventura<sup>1</sup>, Milana Bergamino<sup>1</sup>, Patricia Galván<sup>3</sup>, Catalina Falo<sup>1</sup>, Idoia Morilla<sup>1</sup>, Adela Fernandez-Ortega<sup>1</sup>, Agostina Stradella<sup>1</sup>, Montse Rey<sup>5</sup>, Amparo Garcia-Tejedor<sup>6</sup>, Miguel Gil-Gil<sup>1</sup> and Aleix Prat<sup>3\*</sup>

## OPEN ACCESS

### Edited by:

Michael Gnant,  
Medical University of Vienna, Austria

### Reviewed by:

Luis Schwarz,  
Auna Oncosalud, Peru  
Ariella Hanker,  
UT Southwestern Medical Center,  
United States

### \*Correspondence:

Sonia Pernas  
spernas@iconcologia.net  
Aleix Prat  
alprat@clinic.cat

### Specialty section:

This article was submitted to  
Women's Cancer,  
a section of the journal  
Frontiers in Oncology

**Received:** 26 April 2019

**Accepted:** 16 July 2019

**Published:** 06 August 2019

### Citation:

Pernas S, Petit A, Climent F, Paré L, Perez-Martin J, Ventura L, Bergamino M, Galván P, Falo C, Morilla I, Fernandez-Ortega A, Stradella A, Rey M, Garcia-Tejedor A, Gil-Gil M and Prat A (2019) PAM50 Subtypes in Baseline and Residual Tumors Following Neoadjuvant Trastuzumab-Based Chemotherapy in HER2-Positive Breast Cancer: A Consecutive-Series From a Single Institution. *Front. Oncol.* 9:707. doi: 10.3389/fonc.2019.00707

<sup>1</sup> Department of Medical Oncology-Breast Cancer Unit, Institut Català d'Oncologia (ICO)-H.U.Bellvitge-Institut d'Investigació Biomèdica de Bellvitge (IDIBELL), Universitat de Barcelona, Barcelona, Spain, <sup>2</sup> Department of Pathology-Breast Cancer Unit, Institut Català d'Oncologia (ICO)-H.U.Bellvitge-Institut d'Investigació Biomèdica de Bellvitge (IDIBELL), Universitat de Barcelona, Barcelona, Spain, <sup>3</sup> Department of Medical Oncology, Hospital Clínic de Barcelona, Universitat de Barcelona, Barcelona, Spain, <sup>4</sup> Clinical Research Unit, Institut Català d'Oncologia (ICO)-L'Hospitalet, Barcelona, Spain, <sup>5</sup> Department of Pharmacy, Institut Català d'Oncologia (ICO)-L'Hospitalet, Barcelona, Spain, <sup>6</sup> Department of Gynecology-Breast Cancer Unit, Institut Català d'Oncologia (ICO)-H.U.Bellvitge-Institut d'Investigació Biomèdica de Bellvitge (IDIBELL), Universitat de Barcelona, Barcelona, Spain

**Introduction:** HER2-enriched subtype has been associated with higher response to neoadjuvant anti-HER2-based therapy across various clinical trials. However, limited data exist in real-world practice and regarding residual disease. Here, we evaluate the association of HER2-enriched with pathological response (pCR) and gene expression changes in pre- and post-treatment paired samples in HER2-positive breast cancer patients treated outside of a clinical trial.

**Methods:** We evaluated clinical-pathological data from a consecutive series of 150 patients with stage II-IIIC HER2-positive breast cancer treated from August 2004 to December 2012 with trastuzumab-based neoadjuvant chemotherapy. Expression of 105 breast cancer-related genes, including the PAM50 genes, was determined in available pre- and post-treatment formalin-fixed paraffin-embedded tumor samples using the nCounter platform. Intrinsic molecular subtypes were determined using the research-based PAM50 predictor. Association of genomic variables with total pCR was performed.

**Results:** The pCR rate was 53.3%, with higher pCR among hormonal receptor (HR)-negative tumors (70 vs. 39%;  $P < 0.001$ ). A total of 89 baseline and 28 residual tumors were profiled, including pre- and post-treatment paired samples from 26 patients not achieving a pCR. HER2-enriched was the predominant baseline subtype not only in the overall and HR-negative cohorts (64 and 75%, respectively), but also in the HR-positive cohort (55%). HER2-enriched was associated with higher pCR rates compared to non-HER2-enriched subtypes (65 vs. 31%; OR = 4.07, 95% CI 1.65–10.61,  $P < 0.002$ ) and this



association was independent of HR status. In pre- and post-treatment paired samples from patients not achieving a pCR, a lower proportion of HER2-enriched and twice the number of luminal tumors were observed at baseline, and luminal A was the most frequent subtype in residual tumors. Interestingly, most (81.8%) HER2-enriched tumors changed to non-HER2-enriched, whereas most luminal A samples maintained the same subtype in residual tumors.

**Conclusions:** Outside of a clinical trial, PAM50 HER2-enriched subtype predicts pCR beyond HR status following trastuzumab-based chemotherapy in HER2-positive disease. The clinical value of intrinsic molecular subtype in residual disease warrants further investigation.

**Keywords:** breast cancer, HER2, pathological complete response, gene expression, molecular intrinsic subtype, residual disease, paired samples

## INTRODUCTION

Significant advances have occurred in the treatment of HER2-positive breast cancer that have dramatically improved survival and changed its natural history (1–6). In the neoadjuvant setting, the introduction of HER2-targeted agents to chemotherapy has considerably enhanced the achievement of a pathological complete response (pCR) (7–10). This has translated into important gains in survival in early HER2-positive disease (11–13). Despite these improvements, HER2-positive breast cancer remains a clinically and biologically heterogeneous disease with different treatment sensitivities and survival outcomes (14–16). Thus, identification of these distinct groups of patients using molecular-based biomarkers is needed.

Among different molecular biomarkers evaluated to date in HER2-positive disease, intrinsic molecular subtypes (i.e., luminal A, luminal B, HER2-enriched, and basal-like) identified by gene expression analysis have now shown consistent data across several clinical trials. Specifically, the HER2-enriched subtype has been associated with a higher likelihood of achieving a pCR following neoadjuvant anti-HER2-based chemotherapy compared to non-HER2-enriched disease (15, 17–21). However, limited data exist to date (1) outside a clinical trial setting and (2) regarding residual disease and gene expression changes in paired samples.

Based on this prior evidence, the primary aim of this study was to test the association of the HER2-enriched subtype with pCR in a consecutive series of HER2-positive breast cancer patients homogeneously treated with trastuzumab-based neoadjuvant chemotherapy at a single comprehensive cancer center. As a secondary aim, we explored biological changes between baseline and surgery specimens in patients with residual disease after neoadjuvant treatment. Initial clinical results of this series were previously published (22).

## METHODS

### Clinical-Pathological Data

Clinicopathological data were evaluated in a series of 150 women with stages II to IIIC (T4d included) HER2-positive

breast cancer consecutively treated at Institut Català d'Oncologia (ICO)-Hospitalet (Barcelona, Spain) between August 2004 and December 2012. Treatment schema consisted of weekly paclitaxel 80 mg/m<sup>2</sup> for 12 weeks followed by 4 cycles of 5-Fluoracil, Epirubicin, and Cyclophosphamide (600/75/60 mg/m<sup>2</sup>) every 21 days. During the 24 weeks of neoadjuvant systemic treatment, concomitant trastuzumab 2 mg/kg (after a 4 mg/kg loading dose) was administered. Surgery was performed 3–4 weeks after the last dose of chemotherapy. Left ventricular ejection fraction was monitored every 12 weeks during treatment and in the follow-up period every 6 months for the first 2 years and then annually. Adjuvant hormonal therapy and radiotherapy were administered per institutional guidelines. Additional 6 months of adjuvant trastuzumab were also recommended since 2006. This study was approved by the Institutional Review Board of H.U. Bellvitge, L'Hospitalet (Barcelona), and all patients signed informed consent forms to allow molecular analyses to be performed on their tissue samples.

Estrogen receptor (ER) and progesterone receptor (PR) status were determined by immunohistochemistry (IHC) at baseline core biopsies and in post-treatment surgical specimens with residual disease and considered positive if >1% of tumor cells were stained. HER2 positivity was determined by IHC and fluorescence *in situ* hybridization according to 2007 ASCO/CAP guidelines (23). pCR was defined as the absence of invasive cancer both in the breast and lymph nodes, regardless of the presence of *in situ* carcinoma (ypT0/isypN0).

### Gene Expression Analysis and Intrinsic Subtyping

Hematoxylin and eosin-stained slides from formalin-fixed paraffin-embedded (FFPE) baseline core biopsies and post-neoadjuvant surgical specimens of patients with residual disease were examined to confirm the presence of invasive tumor cells and to determine the minimum surface area. For RNA purification, 1–5 10 µm FFPE slides were used for each tumor specimen. A minimum of 100 ng of total RNA was used to measure the expression of 105 breast cancer-related genes, including the PAM50 genes, 5 housekeeping genes, and 50 additional genes (related to proliferation, cell



**TABLE 1** | Baseline patient characteristics of the entire cohort and of patients with genomic data.

	All patients <i>n</i> = 150	Patients with genomic data <i>n</i> = 91
Age, median (range)	50 (27–79)	51 (27–76)
<b>HISTOLOGICAL GRADE</b>		
Grade 2	47 (31.3%)	33 (36.3%)
Grade 3	101 (67.3%)	57 (62.6%)
Not evaluable	2 (1.3%)	1 (1.1%)
<b>TUMOR STAGE</b>		
T1	9 (6.0%)	7 (7.8%)
T2	97 (64.7%)	55 (61.1%)
T3	15 (10.0%)	8 (8.9%)
T4b	16 (10.6%)	10 (11.1%)
T4d	13 (8.7%)	10 (11.1%)
<b>NODAL STATUS</b>		
N0	34 (22.5%)	18 (20%)
N1	90 (59.6%)	57 (63.3%)
N2	15 (9.9%)	10 (11.1%)
N3	11 (7.3%)	5 (5.6%)
<b>HORMONAL RECEPTOR STATUS</b>		
ER+ PR+	56 (37.4%)	36 (39.6%)
ER+ PR–	22 (14.4%)	11 (12.1%)
ER– PR+	2 (1.2%)	2 (2.2%)
ER– PR–	70 (47.0%)	42 (46.1%)
<b>Ki67</b>		
<20	27 (18.0%)	18 (19.8%)
≥20	119 (79.3%)	73 (80.2%)
Not evaluable	4 (2.7%)	

ER, estrogen receptor; PR, progesterone receptor.

cycle, and angiogenesis/hypoxia). Gene expression analyses and comparison of pre- and post-treatment samples were performed at Vall d'Hebron Institute of Oncology (VHIO) using the nCounter platform (Nanostring Technologies, Seattle, WA, USA). Data were log base 2 transformed and normalized using housekeeping genes selected.

Intrinsic subtyping (luminal A, luminal B, HER2-enriched, basal-like, and normal-like) was performed using the research-based PAM50 intrinsic subtype predictor as previously described (24, 25).

## Statistical Analysis

Association between two variables was evaluated using Student's *t*-test, Pearson's  $\chi^2$  test or Fisher's exact test. Univariate and multivariate logistic regression analyses were done to investigate the association of each variable with pCR. Odds ratios (OR) and 95% confidence intervals (CI) were calculated for each variable. The significance level was set to a two-sided  $\alpha$  of 0.05. To identify genes whose expression was significantly different between paired pre- and post-treated samples, we used a paired two-class significance analysis of microarrays (SAM) with a false discovery rate (FDR) <5%. All statistical tests were two sided, and the

statistical significance level was set to <0.05. We used R version 3.2.2 for all the statistical analyses (<http://cran.r-project.org>).

## RESULTS

Baseline clinicopathologic characteristics of the overall cohort of patients (*n* = 150) and from those with tissue samples available for gene expression (*n* = 91) are listed in **Table 1**. A flow diagram of the study population is shown in **Figure S1**. The baseline median tumor size was 30 mm and 35% of patients had locally advanced breast cancer. All 150 patients underwent surgery; therefore, all were evaluable for pathological response. Lumpectomy was performed in 87 patients (58%). Overall, 80 of 150 patients (53.3%, 95% CI 0.45–0.61) achieved a pCR in the breast and lymph nodes. Interestingly, 10 patients out of 13 (77%) with inflammatory breast cancer experienced a pCR. HR-negative disease was significantly associated with higher pCR rates (69.6% [48/69] vs. 39.5% [32/81] in HR-positive; *p* < 0.001). Age, tumor size, histological differentiation grade, or Ki67 were not associated with pCR.

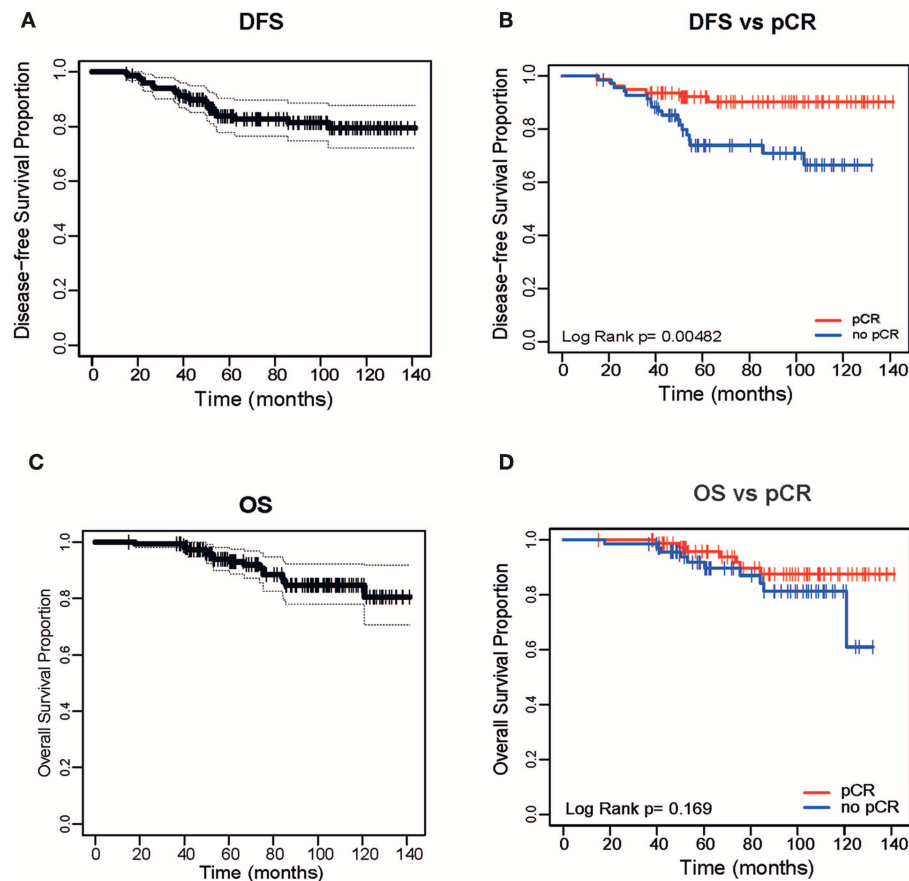
With a median of follow-up of 79 months (range 15–141 months), median disease-free survival (DFS) was not reached (**Figure 1A**); DFS was 83% (95% CI 72.1–87.6%). There were 25 relapses (16.7%): 16 patients had initially HR-positive tumors and 9 HR-negative tumors. Median time to progression was 32 months (range 8–96 months). This time differed significantly per HR status: 19.8 months in HR-negative tumors and double (39.5 months) in HR-positive ones (*p* = 0.023). Achieving a pCR was significantly associated with an improved DFS in the overall cohort (**Figure 1B**) and by HR status (**Figure S2**). There were 7 relapses (8.7%) in the pCR group vs. 18 (25.7%) in the group of residual disease (*p* = 0.005, OR 3.28, 95% CI 1.37–7.86). Median overall survival (OS) was not reached (**Figure 1C**). OS was 88.7% (95% CI 70.6–91.8%). There were 17 deaths, the majority due to disease progression and 3 due to other causes (none of these related to treatment). In contrast to DFS, achieving a pCR was not significantly associated with an improved OS (**Figure 1D**).

## Baseline Subtype Distribution

Of the 89 available baseline samples for gene expression analyses, 40 were HR-negative and 49 were HR-positive. At baseline, most tumors were classified as HER2-enriched subtype by PAM50 (64%), followed by luminal A (11.2%), normal-like (9%) basal-like (7.9%), and luminal B (7.9%). Subtype distribution differed significantly between HR-status. Basal-like subtype was identified only in HR-negative disease, whereas luminal A and B were identified only in HR-positive samples (**Figure 2**) HER2-enriched was the predominant one in both HR-negative tumors (75%) and HR-positive tumors (55%).

## Association of Intrinsic Subtypes and Gene Expression With pCR

Higher rates of pCR were observed in HER2-enriched tumors compared to non-HER2-enriched subtypes (64.9 vs. 31.2%, OR = 4.07, 95% CI 1.65–10.61, *p* < 0.002) regardless of HR status (**Figure 3**). None of the luminal A samples achieved a pCR and only two samples with luminal B disease (28.6%) achieved a pCR.



**FIGURE 1 |** Disease Free Survival of the entire cohort (A) and based on pathological complete response (pCR) (B). Overall Survival (OS) of the entire cohort (C) and based on pCR (D).

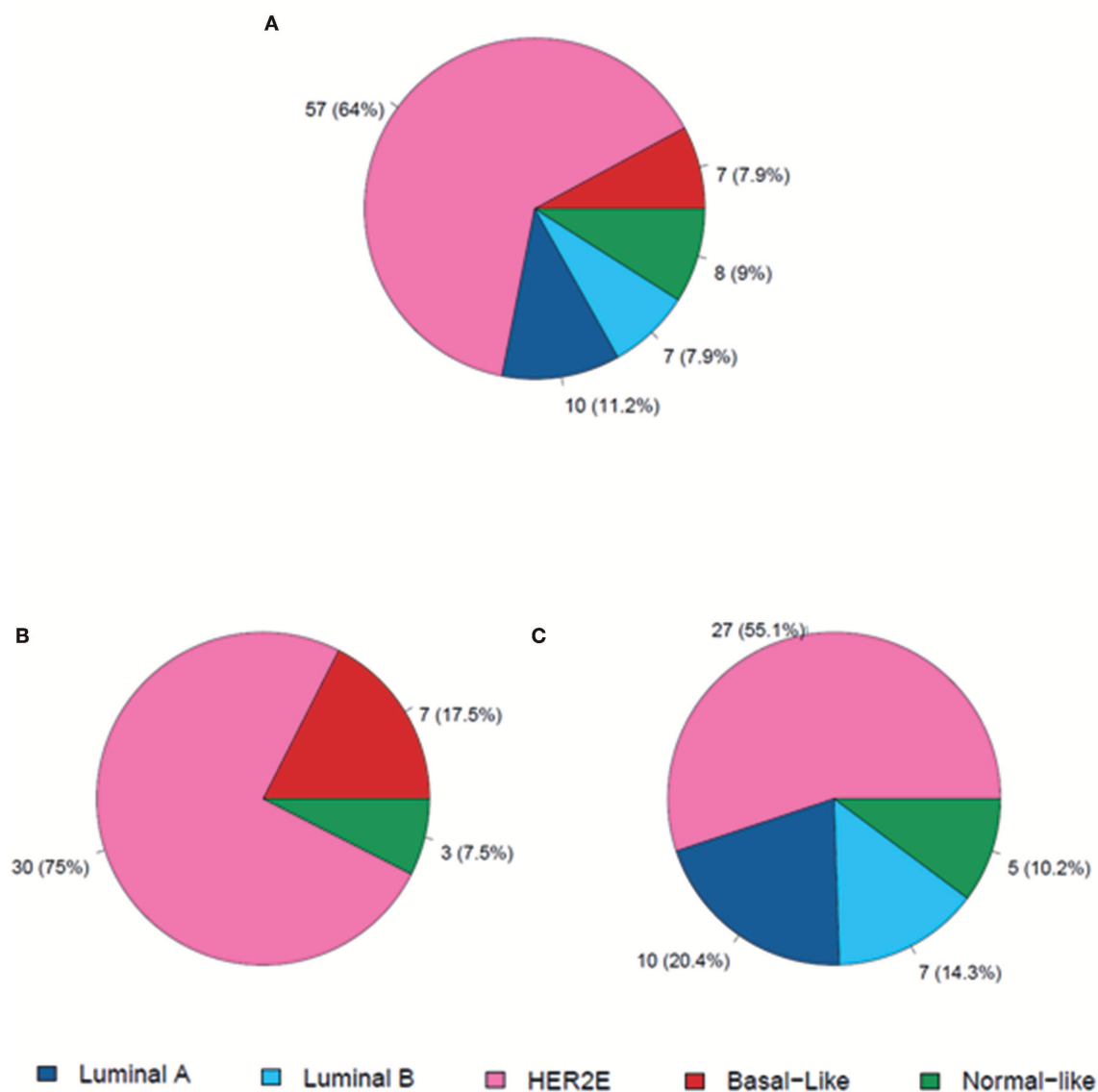
We evaluated the association between PAM50 signatures, HR status (by IHC), and ki67 (by IHC and by gene expression) with pCR. HR-negative status and five of the eight PAM50 signatures (HER2-enriched, ROR-S based on subtype contents, ROR-P based on subtype contents and proliferation index, Basal-like, and Proliferation score) were significantly associated with pCR, whereas luminal A was associated with non-pCR ( $p < 0.001$ ). HR-negative status, HER2-enriched and luminal A signatures demonstrated the strongest association in predicting pathological response (Figure 4A). After adjusting for HR status, HER2-enriched, ROR-S and ROR-P were significantly associated with pCR and luminal A with non-pCR (Figure 4B).

We then assessed the association between individual expression of 105 genes and pCR. The expression of 14 genes was significantly associated with pCR, including ERBB2, CCNE1, genes involved in cell survival and migration (like FGFR4 and GRB7), and genes related with DNA repair and replication pathway (EXO1, ORC6L, and RRM2). On the contrary, the expression of 21 genes was significantly associated with non-pCR, including BCL2, ESR1, GATA3, KRT19, MYC, PGR, PIK3CA, and SLC39A6 (Supplemental Data).

## Residual Disease and Paired Samples From Patients Not Achieving a pCR

Out of the 66 patients with residual disease at surgery, gene expression was successfully performed in 28 surgical specimens (42.4%). Residual subtype distribution was as follows: normal-like (50.0%), luminal A (32.1%), HER2-enriched (14.3%), and luminal B (3.5%). Of these 28 surgical specimens with residual disease, 26 had pre- and post-treatment paired samples. As expected, the baseline distribution of the intrinsic subtypes in this cohort of patients that did not achieved a pCR, differed from the overall cohort (Figure 5), with less proportion of HER2-enriched subtype (42.3 vs. 64%) and nearly double the proportion of luminal samples (42.3 vs. 19.1%). Regarding changes in intrinsic subtypes in pre- and post-treatment paired samples with residual disease, most of HER2-enriched tumors (81.8%) converted to non-HER2-enriched, whereas 66.7% of luminal A samples maintained the same subtype. Interestingly, in this cohort of paired samples there were 7 conversions to HER2-negative in residual disease (10 cases in the overall cohort).

Next, we analyzed changes in the expression of the 8 PAM50 signatures in those 52 pre- and post-treatment paired

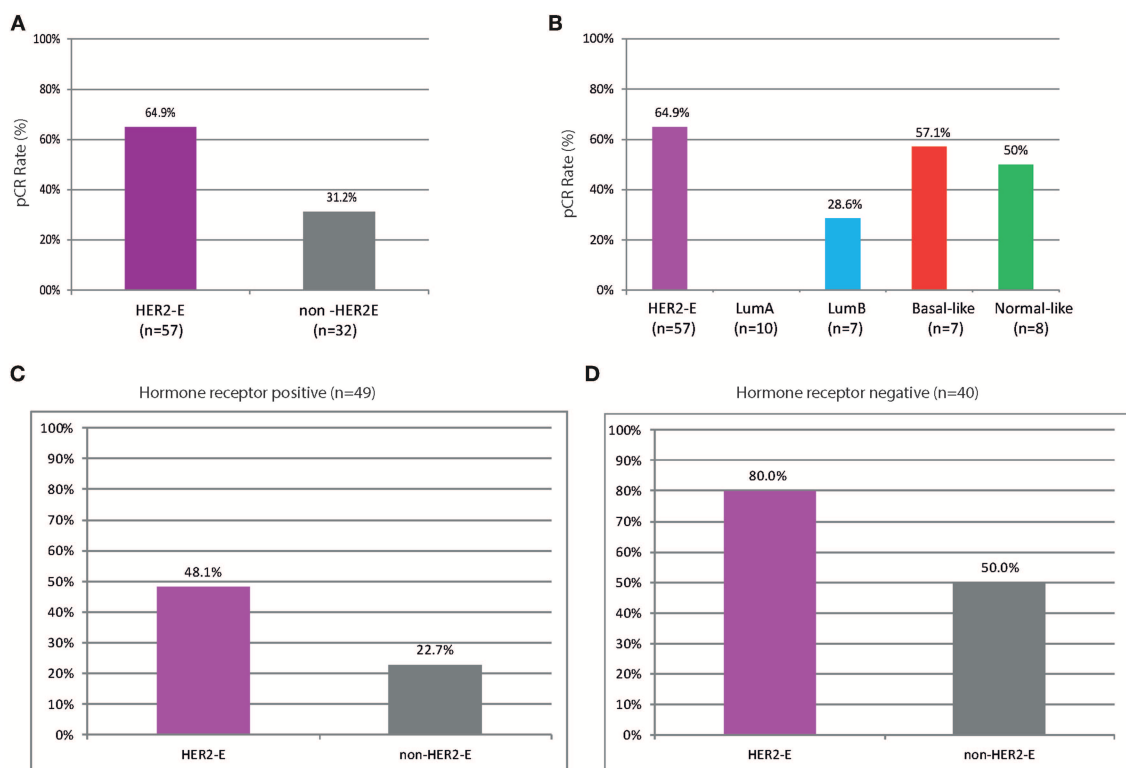


**FIGURE 2 |** Distribution of molecular subtypes at baseline. **(A)** In all patients ( $n = 89$ ); **(B)** Patients with HR-negative disease ( $n = 40$ ); **(C)** Patients with HR-positive disease ( $n = 49$ ).

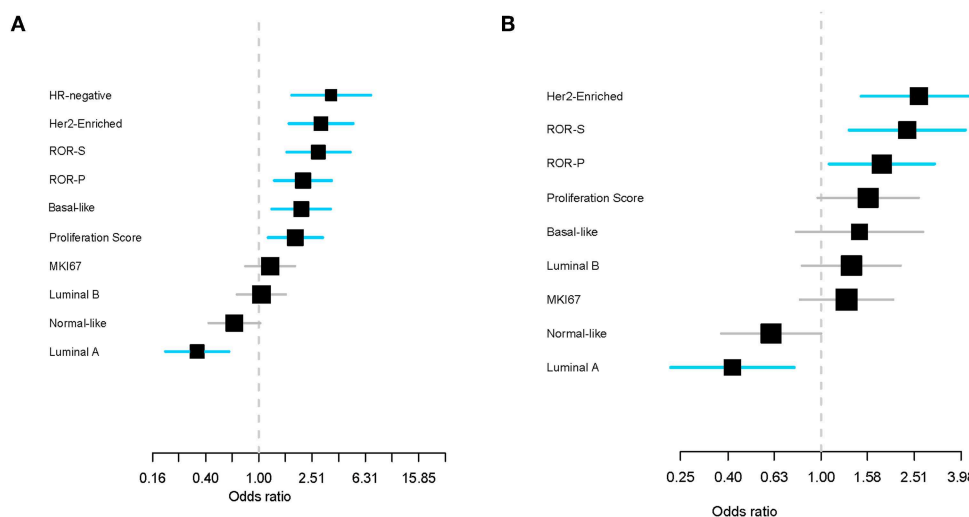
samples. Most of them underwent significant changes: a decrease in expression of HER2-enriched and luminal B signatures, proliferation score, ROR-S, and ROR-P PAM50 signatures were observed in most samples, as well as an important increase in luminal A and normal-like signatures. On the contrary, basal-like signatures showed no changes (**Figure S3**). Regarding single genes, 90 changed significantly, with a false discovery rate of  $<5\%$ . Thirty-five genes, mostly related to stroma (CAV1, VIM, MET, MMP) were overexpressed in post-treatment samples compared to baseline, whereas 55 genes decreased in expression (**Supplemental Data**). Most of the downregulated genes in post-treatment samples are involved in functions such as cell cycle and proliferation (EXO1, CENPF, MKI67).

## DISCUSSION

HER2-positive breast cancer is indeed a clinically and biologically heterogeneous disease not fully recapitulated by HR status. In this consecutive series of HER2-positive breast cancer patients treated with trastuzumab-based primary chemotherapy, all the main intrinsic molecular subtypes were identified by gene expression analyses. Intrinsic subtype distribution differed significantly between HR-negative and HR-positive tumors. Importantly, HER2-enriched was the predominant subtype, not only in the overall and HR-negative cohorts (64 and 75%, respectively) but also in the HR-positive subgroup (55%). Tumor heterogeneity within this series of HER2-positive breast cancer modulated response to neoadjuvant treatment. The highest pCR rate was



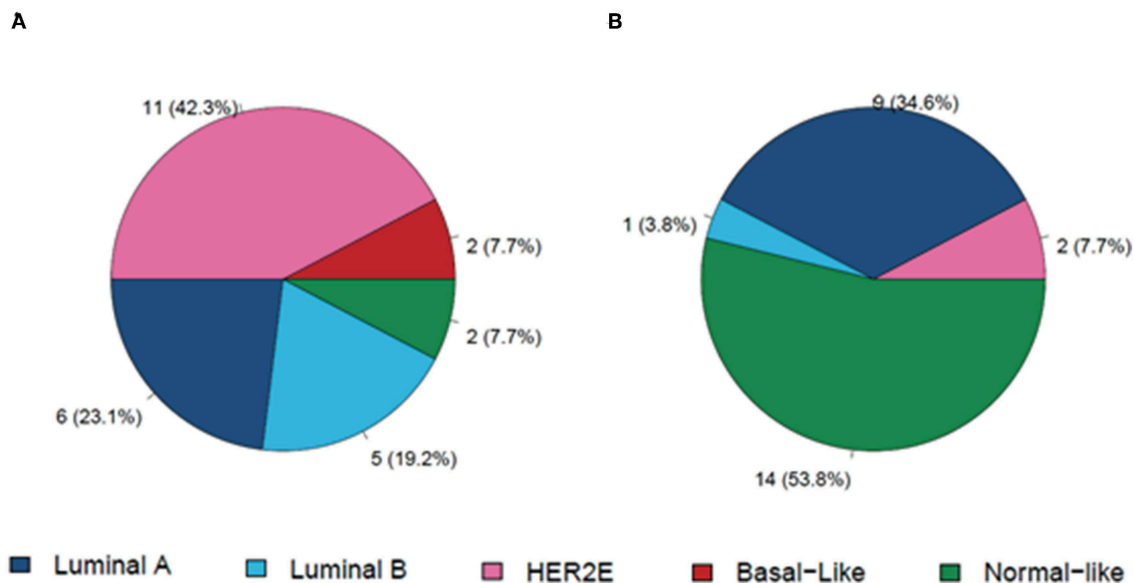
**FIGURE 3 |** Pathological complete response (pCR) in breast and axilla across the intrinsic subtypes of breast cancer in **(A,B)** the overall cohort; **(C)** Patients with HR-positive disease ( $n = 49$ ); **(D)** Patients with HR-negative disease ( $n = 40$ ). HER2-E, HER2-enriched; non-HER2-E, non-HER2-enriched.



**FIGURE 4 |** Effect of PAM50 signatures (as continuous variables) on pathological complete response (pCR) in the univariate analysis **(A)** and after adjusting for hormone receptor variables **(B)**. Each signature has been standardized to have a mean of 0 and a standard deviation of 1. The size of the square is inversely proportional to the standard error. Horizontal bars represent the 95% CIs of ORs. Statistically significant variables are shown in blue. Each gene signature has been evaluated individually and ranked ordered based on the estimated OR. ROR-S, risk of recurrence score based on subtype; ROR-P risk of recurrence score based on subtype and proliferation.

among patients with HER2-enriched tumors, which was more than double the pCR rate of patients with non-HER2-enriched tumors (65 vs. 31%), even in patients with HR-positive tumors (48 vs. 23%).

HER2-enriched subtype has consistently been associated with achieving the highest rate of pCR among HER2-positive tumors (15, 17–21), even in the absence of chemotherapy, with just dual HER2-blockade (21). In the clinical trials that have evaluated



**FIGURE 5 |** Distribution of molecular subtypes in a cohort of patients with residual disease and paired baseline (A) and surgical specimens (B). HER2E, HER2-enriched.

efficacy of HER2-targeted agents (e.g., trastuzumab, pertuzumab, and lapatinib) in combination with neoadjuvant chemotherapy or dual blockade alone, the pCR observed among the HER2-enriched subtype varies between 41 and 70%, with the highest rate being achieved with dual-HER2 blockade and chemotherapy. In our study, the pCR rate of the HER2-enriched subgroup was 65%, similar to that achieved in the CALGB study (70%), one of the highest ever described in HER2-positive breast cancer (15), regardless of treatment arm or HR status. It is important to note that, in our study, treatment consisted of single-trastuzumab given concomitantly with anthracycline-and-taxanes-based neoadjuvant chemotherapy. The overall pCR rate of 53% in our series is similar to that achieved in the ACOSOG Z1041 trial (20), a trial designed to compare the pCR rate either in a sequential or concurrent regimen of an anthracycline-and-taxanes-based chemotherapy and trastuzumab (like the one used in our study), which ultimately found no difference between both arms. In this trial, cases classified as HER2-enriched subtype by RNA-seq analysis were also more likely to achieve a pCR compared to non-HER2-enriched tumors.

To our knowledge, our study is the first one to demonstrate the association of five out of eight PAM50 signatures (HER2-enriched, ROR-S, ROR-P, Basal-like, and Proliferation score) with pCR, whereas luminal A signature was associated with non-pCR. Moreover, HR-negative status and HER2-enriched subtype (and signature) demonstrated the strongest association in predicting pCR and luminal A signature with non-pCR. Importantly, intrinsic subtype was an independent, additional predictive factor of pCR to HR status.

Regarding baseline distribution of molecular subtypes, our results are in accordance with previous reports such as the PAMELA trial (21) and the APT trial (26, 27) where the largest subset of baseline samples was classified as HER2-enriched (66.9

and 65%, respectively). In contrast, in the CALGB40601 (15) study and the Cher-LOB trial (18), the proportion of HER2-enriched subtype at baseline was similar to that of luminal A and luminal B (31 and 27%, respectively) and luminal subtypes predominated among HR-positive tumors. This fact could explain why the overall pCR rate in the control arm in the Cher-LOB trial (with the same treatment schema as in our series) was surprisingly low (25%) (18). Another explanation is that the PAMELA trial, the APT trial, and our study all used the nCounter platform, whereas the CALGB40601 study used RNAseq and Cher-LOB used microarrays.

Limited data exist regarding the distribution of molecular subtypes in residual disease after neoadjuvant therapy and in pre- and post-treatment paired samples. In the present study, we examined changes in gene expression and molecular subtype in paired samples of patients with residual disease. A lower proportion of HER2-enriched subtypes and almost twice the number of luminal tumors than in the overall cohort were found at baseline. The most frequent subtype in not eradicated post-treated tumors, excluding normal-like, was the luminal A subtype, as occurred in the CALGB40601 study (15). In the paired samples, most HER2-enriched tumors changed to non-HER2-enriched, whereas most luminal A samples maintained the same subtype. The observed changes in molecular subtype could be attributed to reduced proliferation and/or changes in tumor and stroma cellularity. Residual tumors also showed a substantial modulation of genes, with downregulation of genes involved in proliferation and cell cycle function and upregulation of those related mostly to stroma. However, gene expression analyses cannot distinguish between intra-tumor heterogeneity, stromal alterations or a true treatment effect and may be a mixture of all three. The down regulation of the HER2-enriched, luminal B and proliferation PAM50 signatures (proliferation score, ROR-S, and



ROR-P) and the overexpression of the luminal A and normal-like signatures, seen in those paired samples from our study could be explained by peritumoral stromal contamination. We note that these analyses should be interpreted with caution, due to the exploratory nature and small sample size of the study.

This study has several strengths and limitations. It was done in a real-world setting, at a single institution, and it has a long-term follow-up. Patients were homogeneously treated with trastuzumab-based therapy and the study also included evaluation of pre- and post-treatment paired samples. Nevertheless, gene expression did not include immune signatures, which other studies have found to be an independent predictor of response to HER2 targeting beyond PAM50 intrinsic subtypes (15, 18–20), and mutational status (such as PIK3CA) was not analyzed either. Additionally, the current standard neoadjuvant therapy for HER2-positive breast cancer include dual-HER2-blockade with trastuzumab and pertuzumab.

Biologic heterogeneity within HER2-positive breast cancer can determine response to treatment and prognosis as shown in clinical trials, and in everyday clinical practice as shown in our study. Yet, not all HER2-positive breast cancer patients may need to be treated in the same manner. The combination of HER2-targeted therapy alone (dual HER2 blockade with or without endocrine therapy) has shown activity in a substantial percentage of patients, eradicating HER2-positive tumors without chemotherapy and with a favorable toxicity profile (21, 28, 29). However, we need to be able to identify which patients can benefit from this de-escalation strategy and if there is a survival benefit in achieving a pCR with just dual blockade and no chemotherapy. Interestingly, findings from additional exploratory subgroup analyses in the NOAH study (30) showed that the prognostic effect of pCR for event-free survival and overall survival was statistically significant only in patients treated with chemotherapy and trastuzumab and not in patients treated with chemotherapy alone. What does seem clear is that HER2 expression as a single biomarker of treatment response is not enough to develop rational individualized therapeutic regimens. There is an urgent need to find robust predictive biomarkers of response or resistance to the anti-HER2 approach, other than HER2-positivity, in order to individualize treatment and identify different populations of patients who need more treatment or others who may avoid unnecessary treatments and their related toxicities. Serial changes in gene expression, tumor cells or immune cells, as was done in the PAMELA trial (21, 31), may identify early predictive markers of response or resistance than just baseline or residual intrinsic subtypes alone.

## CONCLUSION

Our data show that, outside of a clinical trial, PAM50 HER2-enriched intrinsic subtype predicts pCR beyond HR status following trastuzumab-based chemotherapy in HER2-positive disease. The clinical value of intrinsic molecular subtype in residual disease warrants further investigation.

## DATA AVAILABILITY

All datasets generated for this study are included in the manuscript and/or the **Supplementary Files**.

## ETHICS STATEMENT

This study was approved by the Institutional Review Board of H.U. Bellvitge, L'Hospitalet (Barcelona), and all patients signed informed consent forms to allow molecular analyses to be performed on their tissue samples.

## AUTHOR CONTRIBUTIONS

SP and AP: conception and design. SP, APe, FC, CF, IM, AF-O, AS, MR, AG-T, MG-G, and AP: provision of study materials or patients. SP, APe, FC, LP, JP-M, LV, MB, PG, CF, IM, AF-O, AS, MR, MG-G, and AP: collection and assembly of data. SP, APe, FC, LP, JP-M, and AP: data analysis and interpretation. SP, LP, and AP: manuscript writing. All authors: final approval of manuscript and accountable for all aspects of the work.

## FUNDING

This work was partially supported by a research grant Roche (to SP), Pas a Pas (to AP); Save the Mama (to AP), Instituto de Salud Carlos III—PI16/00904 (to AP); Career Catalyst Grant CCR13261208 from the Susan Komen Foundation (to AP).

## ACKNOWLEDGMENTS

We would like to acknowledge all the patients and their families for participating in this study and also all the members of the Breast Cancer Unit of the Institut Català d'Oncologia (ICO) H.U.Bellvitge. We thank Kaitlyn T. Bifolck, BA, for her editorial support. We thank CERCA Programme/Generalitat de Catalunya for institutional support.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2019.00707/full#supplementary-material>

**Figure S1** | Flow diagram detailing the study population.

**Figure S2** | Survival outcomes based on pathological complete response (pCR). Disease-free survival (DFS) by hormone receptor (HR) status **(A)**; Overall survival (OS) by HR status **(B)**; DFS by HER2-Enriched and non-HER2-Enriched **(C)**; OS by HER2-Enriched and non-HER2-Enriched **(D)**.

**Figure S3** | Changes in the PAM50 signatures in paired samples (i.e., baseline vs. post-treatment) of patients not achieving a pCR.

**Table S1** | Changes in the expression of the 105 genes between pre- and post-treatment paired samples in patients not achieving a pathologic complete response (pCR).

## REFERENCES

- Slamon D, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, et al. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N Engl J Med.* (2001) 344:783–92. doi: 10.1056/NEJM200103153441101
- Swain SM, Baselga J, Kim SB, Ro J, Semiglazov V, Campone M, et al. CLEOPATRA Study Group. Pertuzumab, trastuzumab, and docetaxel in HER2-positive metastatic breast cancer. *N Engl J Med.* (2015) 372:724–34. doi: 10.1056/NEJMoa1413513
- Piccart-Gebhart MJ, Procter M, Leyland-Jones B, Goldhirsch A, Untch M, Smith I, et al. Herceptin adjuvant (HERA) trial study team. Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. *N Engl J Med.* (2005) 353:1659–72. doi: 10.1056/nejmoa052306
- Romond EH, Perez EA, Bryant J, Suman VJ, Geyer CE Jr, Davidson NE, et al. Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. *N Engl J Med.* (2005) 353:1673–84. doi: 10.1056/NEJMoa052122
- Slamon D, Eiermann W, Robert N, Pienkowski T, Martin M, Press M, et al. Breast cancer international research group. Adjuvant trastuzumab in HER2-positive breast cancer. *N Engl J Med.* (2011) 365:1273–83. doi: 10.1056/NEJMoa0910383
- Moja L, Tagliabue L, Balduzzi S, Parmelli E, Pistotti V, Guarneri V, et al. Trastuzumab containing regimens for early breast cancer. *Cochrane Database Syst Rev.* (2012) 18:CD006243. doi: 10.1002/14651858.CD006243.pub2
- Buzdar AU, Valero V, Ibrahim NK, Francis D, Broglio KR, Theriault RL, et al. Neoadjuvant therapy with paclitaxel followed by 5-fluorouracil, epirubicin, and cyclophosphamide chemotherapy and concurrent trastuzumab in human epidermal growth factor receptor 2-positive operable breast cancer: an update of the initial randomized study population and data of additional patients treated with the same regimen. *Clin Cancer Res.* (2007) 13:228–33. doi: 10.1158/1078-0432.CCR-06-1345
- Gianni L, Eiermann W, Semiglazov V, Manikhas A, Lluch A, Tjulandin S, et al. Neoadjuvant chemotherapy with trastuzumab followed by adjuvant trastuzumab versus neoadjuvant chemotherapy alone, in patients with HER2-positive locally advanced breast cancer (the NOAH trial): a randomised controlled superiority trial with a parallel HER2-negative cohort. *Lancet.* (2010) 375:377–84. doi: 10.1016/S0140-6736(09)61964-4
- Untch M, Rezaei M, Lobl S, Fasching PA, Huober J, Tesch H, et al. Neoadjuvant treatment with trastuzumab in HER2-positive breast cancer: results from the GeparQuattro study. *J Clin Oncol.* (2010) 28:2024–31. doi: 10.1200/JCO.2009.23.8451
- Pernas S, Barroso-Sousa, Tolane SM. Optimal treatment of early stage HER2-positive breast cancer. *Cancer.* (2018) 124:4455–66. doi: 10.1002/cncr.31657
- Untch M, Fasching PA, Konecny GE, Hasmuller S, Lebeau A, Kreienberg R, et al. Pathologic complete response after neoadjuvant chemotherapy plus trastuzumab predicts favorable survival in human epidermal growth factor receptor 2-overexpressing breast cancer: results from the TECHNO. *J Clin Oncol.* (2011) 29:3351–7. doi: 10.1200/JCO.2010.31.4930
- von Minckwitz G, Untch M, Blohmer J-U, Costa SD, Eidtmann H, Fasching PA, et al. Definition and impact of pathologic complete response on prognosis after neoadjuvant chemotherapy in various intrinsic breast cancer subtypes. *J Clin Oncol.* (2012) 30:1796–804. doi: 10.1200/JCO.2011.38.8595
- Cortazar P, Zhang L, Untch M, Mehta K, Costantino JP, Wolmark N, et al. Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. *Lancet.* (2014) 384:164–72. doi: 10.1016/S0140-6736(13)62422-8
- Prat A, Carey LA, Adamo B, Vidal M, Tabernero J, Cortes J, et al. Molecular features and survival outcomes of the intrinsic subtypes within HER2-positive breast cancer. *J Natl Cancer Inst.* (2014) 106:152. doi: 10.1093/jnci/dju152
- Carey LA, Berry DA, Cirincione 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. doi: 10.1200/JCO.2015.62.1268
- Prat A, Perou CM. Deconstructing the molecular portraits of breast cancer. *Mol Oncol.* (2011) 5:5–23. doi: 10.1016/j.molonc.2010.11.003
- Prat A, Bianchini G, Thomas M, Belousov A, Cheang MCU, Koehler A, et al. Research-Based PAM50 subtype predictor identifies higher responses and improved survival outcomes in HER2-Positive breast cancer in the NOAH Study. *Clin Cancer Res.* (2014) 20:511–21. doi: 10.1158/1078-0432.CCR-13-0239
- Dieci MV, Prat A, Tagliafico E, Paré L, Ficarra G, Bisagni G, et al. Integrated evaluation of PAM50 subtypes and immune modulation of pCR in HER2-positive breast cancer patients treated with chemotherapy and HER2-targeted agents in the CherLOB trial. *Ann Oncol.* (2016) 27:1867–73. doi: 10.1093/annonc/mdw262
- Fumagalli D, Venet D, Ignatiadis M, Azim HA, Maetens M, Rothe F, et al. RNA sequencing to predict response to neoadjuvant anti-HER2 therapy. A second analysis of the NeoALTTO randomized clinical trial. *JAMA Oncol.* (2017) 3:227–34. doi: 10.1001/jamaoncol.2016.3824
- 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. doi: 10.1093/annonc/mdx048
- Llombart-Cussac A, Cortes J, Pare L, Galvan P, Bermejo B, Martinez N, et al. HER2-enriched subtype as a predictor of pathological complete response following trastuzumab and lapatinib without chemotherapy in early-stage HER2-positive breast cancer (PAMELA): an open-label, single-group, multicenter, phase 2 trial. *Lancet Oncol.* (2017) 18:545–54. doi: 10.1016/S1470-2045(17)30021-9
- Pernas S, Gil-Gil M, de Olza MO, Gumà A, Climent F, Petit A, et al. Efficacy and safety of concurrent trastuzumab plus weekly paclitaxel-FEC as primary therapy for HER2-positive breast cancer in everyday clinical practice. *Breast Cancer Res Treat.* (2012) 134:1161–8. doi: 10.1007/s10549-012-2149-7
- Wolff AC, Hammond ME, Schwartz JN, Hagerty KL, Allred DC, Cote RJ, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *J Clin Oncol.* (2007) 25:118e45. doi: 10.1043/1543-2165(2007)13118:ASOCCO2.0.CO;2
- Prat A, Cheang MU, Galván P, Nuciforo P, Pare L, Adamo B, et al. Prognostic value of intrinsic subtypes in hormone receptor-positive metastatic breast cancer treated with letrozole with or without lapatinib. *JAMA Oncol.* (2016) 2:1287–94. doi: 10.1001/jamaoncol.2016.0922
- Vidal M, Peg V, Galván P, Tres A, Cortes J, Ramon y Cajal S, et al. Gene expression-based classifications of fibroadenomas and phyllodes tumours of the breast. *Mol Oncol.* (2015) 9:1081–90. doi: 10.1016/j.molonc.2015.01.003
- Tolane 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. doi: 10.1056/NEJMoa1406281
- Tolane SM, Guo H, Pernas S, Barry WT, Dillon DA, Ritterhouse L, et al. Seven-year follow-up analysis of adjuvant paclitaxel and trastuzumab trial for node-negative, human epidermal growth factor

- receptor 2-positive breast cancer. *J Clin Oncol.* (2019) 2:JCO1900066. doi: 10.1200/JCO.19.00066
28. Gianni L, Pienkowski T, Im Y-H, Tseng L-M, Liu M-C, Lluch A, et al. Efficacy and safety of neoadjuvant pertuzumab and trastuzumab in women with locally advanced, inflammatory, or early HER2-positive breast cancer (NeoSphere): a randomised, multicentre, open-label, phase 2 trial. *Lancet Oncol.* (2012) 13:25–32. doi: 10.1016/S1470-2045(11)70336-9
  29. Rimawi MF, Mayer IA, Forero A, Nanda R, Goetz MP, Rodriguez AA, et al. Multicenter phase II study of neoadjuvant lapatinib and trastuzumab with hormonal therapy and without chemotherapy in patients with human epidermal growth factor receptor 2-overexpressing breast cancer: TBCRC 006. *J Clin Oncol.* (2013) 31:1726–31. doi: 10.1200/JCO.2012.44.8027
  30. Gianni L, Eiermann W, Semiglazov V, Lluch A, Tjulandin S, Zambetti M, et al. Neoadjuvant and adjuvant trastuzumab in patients with HER2-positive locally advanced breast cancer (NOAH): follow-up of a randomised controlled superiority trial with a parallel HER2-negative cohort. *Lancet Oncol.* (2014) 15:640–7. doi: 10.1016/S1470-2045(14)70080-4
  31. Nuciforo P, Pascual T, Cortés J, Llombart-Cussac A, Fasani R, Paré L, et al. A predictive model of pathologic response based on tumor cellularity and tumor-infiltrating lymphocytes (CelTIL) in HER2-positive breast cancer treated with chemo-free dual HER2 blockade. *Ann Oncol.* (2018) 29:170–7. doi: 10.1093/annonc/mdx647

**Conflict of Interest Statement:** SP has received honoraria for talks and travel grants from Roche, outside of the submitted work and has served as an advisor/consultant to Polyphor. CF has received travel grants from Celgene, outside of the submitted work. AS has received honoraria for talks and travel grants from Roche and Eisai, outside of the submitted work. MG-G has received honoraria for talks from Roche, Pfizer, Novartis, and Pierre Fabre and travel grants from Roche and Daiichi-Sankyo all outside of the submitted work. MG-G has served as an advisor/consultant to Pfizer, Novartis, and Daiichi-Sankyo. Advisory role of APr for Nanostring Technologies.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Pernas, Petit, Climent, Paré, Perez-Martin, Ventura, Bergamino, Galván, Falo, Morilla, Fernandez-Ortega, Stradella, Rey, Garcia-Tejedor, Gil-Gil and Prat. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Corrigendum: PAM50 Subtypes in Baseline and Residual Tumors Following Neoadjuvant Trastuzumab-Based Chemotherapy in HER2-Positive Breast Cancer: A Consecutive-Series From a Single Institution

Sonia Pernas<sup>1\*</sup>, Anna Petit<sup>2</sup>, Fina Climent<sup>2</sup>, Laia Paré<sup>3</sup>, J. Perez-Martin<sup>4</sup>, Luz Ventura<sup>1</sup>, Milana Bergamino<sup>1</sup>, Patricia Galván<sup>3</sup>, Catalina Falo<sup>1</sup>, Idoia Morilla<sup>1</sup>, Adela Fernandez-Ortega<sup>1</sup>, Agostina Stradella<sup>1</sup>, Montse Rey<sup>5</sup>, Amparo Garcia-Tejedor<sup>6</sup>, Miguel Gil-Gil<sup>1</sup> and Aleix Prat<sup>3\*</sup>

## OPEN ACCESS

### Edited and reviewed by:

Michael Grant,  
Medical University of Vienna, Austria

### \*Correspondence:

Sonia Pernas  
spernas@iconcologia.net  
Aleix Prat  
alprat@clinic.cat

### Specialty section:

This article was submitted to  
Women's Cancer,  
a section of the journal  
Frontiers in Oncology

Received: 21 August 2019

Accepted: 11 September 2019

Published: 25 September 2019

### Citation:

Pernas S, Petit A, Climent F, Paré L,  
Perez-Martin J, Ventura L,  
Bergamino M, Galván P, Falo C,  
Morilla I, Fernandez-Ortega A,  
Stradella A, Rey M, Garcia-Tejedor A,  
Gil-Gil M and Prat A (2019)  
Corrigendum: PAM50 Subtypes in  
Baseline and Residual Tumors  
Following Neoadjuvant  
Trastuzumab-Based Chemotherapy in  
HER2-Positive Breast Cancer: A  
Consecutive-Series From a Single  
Institution. *Front. Oncol.* 9:967.  
doi: 10.3389/fonc.2019.00967

<sup>1</sup> Department of Medical Oncology-Breast Cancer Unit, Institut Català d'Oncologia (ICO)-H.U.Bellvitge-Institut d'Investigació Biomèdica de Bellvitge (IDIBELL), Universitat de Barcelona, Barcelona, Spain, <sup>2</sup> Department of Pathology-Breast Cancer Unit, Institut Català d'Oncologia (ICO)-H.U.Bellvitge-Institut d'Investigació Biomèdica de Bellvitge (IDIBELL), Universitat de Barcelona, Barcelona, Spain, <sup>3</sup> Department of Medical Oncology, Hospital Clínic de Barcelona, Universitat de Barcelona, Barcelona, Spain, <sup>4</sup> Clinical Research Unit, Institut Català d'Oncologia (ICO)-L'Hospitalet, Barcelona, Spain, <sup>5</sup> Department of Pharmacy, Institut Català d'Oncologia (ICO)-L'Hospitalet, Barcelona, Spain, <sup>6</sup> Department of Gynecology-Breast Cancer Unit, Institut Català d'Oncologia (ICO)-H.U.Bellvitge-Institut d'Investigació Biomèdica de Bellvitge (IDIBELL), Universitat de Barcelona, Barcelona, Spain

**Keywords:** breast cancer, HER2, pathological complete response, gene expression, molecular intrinsic subtype, residual disease, paired samples

## A Corrigendum on

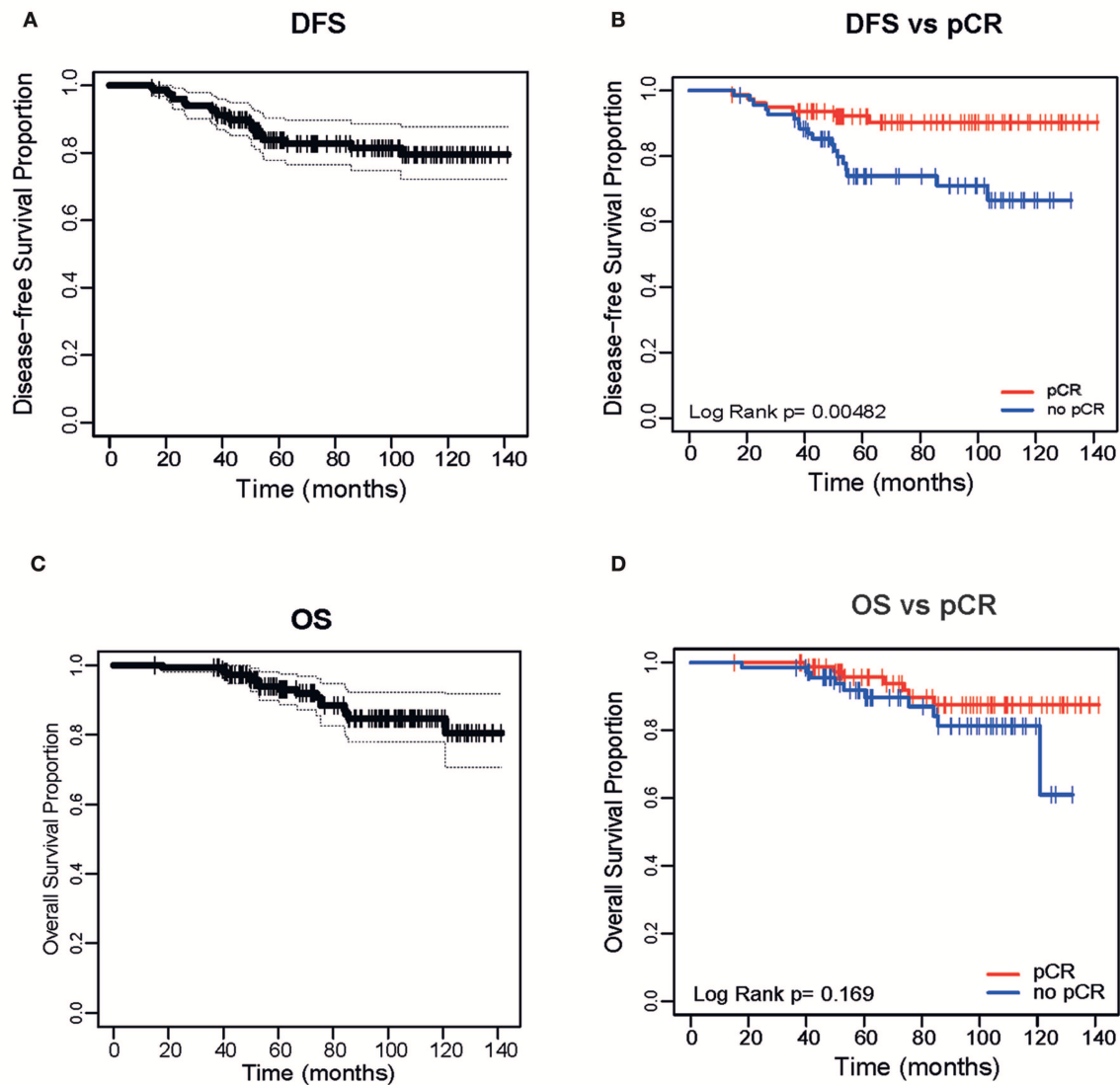
### PAM50 Subtypes in Baseline and Residual Tumors Following Neoadjuvant Trastuzumab-Based Chemotherapy in HER2-Positive Breast Cancer: A Consecutive-Series From a Single Institution

by Pernas, S., Petit, A., Climent, F., Paré, L., Perez-Martin, J., Ventura, L., et al. (2019). *Front. Oncol.* 9:707. doi: 10.3389/fonc.2019.00707

In the original article, there was a mistake in **Figure 1** as published. The colors of the labels used for **Figures 1B,D** were incorrect. pCR should be in red and non-pCR should be in blue. The corrected **Figure 1** appears below.

The authors apologize for this error and state that this does not change the scientific conclusions of the article in any way. The original article has been updated.

Copyright © 2019 Pernas, Petit, Climent, Paré, Perez-Martin, Ventura, Bergamino, Galván, Falo, Morilla, Fernandez-Ortega, Stradella, Rey, Garcia-Tejedor, Gil-Gil and Prat. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



**FIGURE 1 |** Disease Free Survival of the entire cohort **(A)** and based on pathological complete response (pCR) **(B)**. Overall Survival (OS) of the entire cohort **(C)** and based on pCR **(D)**.





# Two Distinct Subtypes Revealed in Blood Transcriptome of Breast Cancer Patients With an Unsupervised Analysis

Wenlong Ming<sup>1†</sup>, Hui Xie<sup>2†</sup>, Zixi Hu<sup>1</sup>, Yuanyuan Chen<sup>1</sup>, Yanhui Zhu<sup>2</sup>, Yunfei Bai<sup>1</sup>, Hongde Liu<sup>1</sup>, Xiao Sun<sup>1</sup>, Yun Liu<sup>2\*</sup> and Wanjun Gu<sup>1\*</sup>

<sup>1</sup> State Key Laboratory of Bioelectronics, School of Biological Science and Medical Engineering, Southeast University, Nanjing, China, <sup>2</sup> The First Affiliated Hospital of Nanjing Medical University, Nanjing, China

## OPEN ACCESS

### Edited by:

Aleix Prat,  
Hospital Clínic de Barcelona, Spain

### Reviewed by:

Nuria Chic,  
Hospital Clínic de Barcelona, Spain  
Masahiko Tanabe,  
The University of Tokyo, Japan

### \*Correspondence:

Yun Liu  
liuyun@njmu.edu.cn  
Wanjun Gu  
wanjingu@gmail.com

<sup>†</sup>These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
Women's Cancer,  
a section of the journal  
Frontiers in Oncology

**Received:** 03 July 2019

**Accepted:** 16 September 2019

**Published:** 01 October 2019

### Citation:

Ming W, Xie H, Hu Z, Chen Y, Zhu Y,  
Bai Y, Liu H, Sun X, Liu Y and Gu W  
(2019) Two Distinct Subtypes  
Revealed in Blood Transcriptome of  
Breast Cancer Patients With an  
Unsupervised Analysis.  
Front. Oncol. 9:985.  
doi: 10.3389/fonc.2019.00985

**Background:** Breast cancer (BC) is a highly heterogeneous cancer. The interaction between immune system and BC is complex, widespread yet unclear. In this study, we aimed to reveal the heterogeneity of host systemic immune response to BC and understand the possible mechanisms that may drive the heterogeneity using transcriptomic data from peripheral blood mononuclear cells (PBMCs).

**Methods:** Transcriptome-wide gene expressions of PBMCs in 33 BC patients were generated by RNA sequencing. An unsupervised clustering algorithm was employed to discover PBMC transcriptome subtypes among BC patients. Association analysis between PBMC subtypes and age, clinical stage, abundance of immune cells, and other clinical factors was performed to understand the underlying biological processes that may drive this heterogeneity. Immune gene signature identification and *in silico* survival analysis were performed to investigate the potential clinical implications of these PBMC subtypes. The findings were validated using the whole blood transcriptomes of an independent cohort.

**Results:** We observed that established BC subtypes were not associated with PBMC gene expression profiles. Instead, we discovered and validated two new BC subtypes using PBMC transcriptome, which have distinct immune cell proportions, especially for lymphocytes ( $P = 5.22 \times 10^{-12}$ ) and neutrophils ( $P = 1.13 \times 10^{-14}$ ). Enrichment analysis of differentially expressed genes revealed that these two subtypes had distinct patterns of immune responses, including osteoclast differentiation and interleukin-10 signaling pathway. We developed two immune gene signatures that can differentiate these two BC PBMC subtypes. Further analysis suggested they had the ability to predict the clinical outcome of BC patients.

**Conclusions:** PBMC transcriptome profiles can classify BC patients into two distinct subtypes. These two subtypes are mainly shaped by different immune cell abundance, which may have implications on clinical outcomes.

**Keywords:** peripheral blood mononuclear cells, immune gene signature, unsupervised analysis, breast cancer subtype, breast cancer survival

## INTRODUCTION

Breast cancer (BC) is now the most frequently diagnosed cancer and the sixth leading cause of cancer-related death among Chinese women (1). To gain better outcomes, the early diagnosis, prognosis and treatment monitoring are critically important (1). However, BC is well-known as a highly heterogeneous malignant tumor, both molecularly and histologically. At present, BC has been classified into five intrinsic molecular subtypes, including luminal-A, luminal-B, HER2-enriched, basal-like, and normal-like (2–5). Each subtype has distinct gene expression profiles, which is associated with cancer prognosis, disease progression, cancer metastasis, and therapeutic resistance (2–5). Based on several clinical and pathological factors, such as estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) status, BC is routinely divided into several subtypes in clinical implementation (6, 7). These clinical classifications are frequently used to guide the treatment of BC patients (6, 7).

Although genetic and epigenetic changes are the key causes of BC, both the innate and adaptive immune system may play substantial roles in BC progression and metastasis as well (8). The presence of cancer cells can activate different immune cells to undergo various phenotypic and functional changes, and eventually kill cancer cells or promote the proliferation of cancer cells (9, 10). Several studies have attempted to detect the presence of cancers by profiling the gene expression in peripheral blood mononuclear cells (PBMCs) from BC patients (11–14) and some other malignant tumors (15, 16). They have proposed several PBMC gene expression signatures that can significantly differentiate cancer patients from healthy controls (12, 13, 15, 16). Furthermore, expression profiles of several immune-related genes in PMBCs from BC patients can predict the relapse of triple negative BC (11, 14). These findings indicated that transcriptomic analysis of peripheral blood immune cells (PBMCs) might be a practical way to evaluate the host systemic immune responses against cancer cells. Notably, this is especially valuable, since the collection of blood samples is non-invasive and convenient as compared to the sampling of tumor tissues (11). However, the human immune system is substantially variable (17). A wide range of factors, such as age, sex, genetic background, and some environmental influences, can perturb and shape the blood transcriptome (17). The relationship between immune system and BC is intricate, and many unanswered questions remain (8, 18). Among them, one of the most important issues is to explore the heterogeneity of blood transcriptome of BC patients and the clinical relevance of this heterogeneity.

In this study, we aimed to reveal the heterogeneity of host systemic immune response to BC and understand the possible mechanisms that drive the heterogeneity. First, we measured the transcriptome-wide gene expressions in PBMC samples from 33 BC patients using RNA sequencing (RNA-seq), and correlated the gene expression profiles with known clinical classifications. Next, we performed an unsupervised cluster analysis on PBMC expressions to reveal the heterogeneity among BC patients and *de novo* classified BC patients with distinct host response patterns.

Then, we validated the PBMC subtypes in an independent BC dataset. Furthermore, we investigated possible clinical factors that may be related to the PBMC subtypes of BC patients, including age, clinical stages and the abundance of immune cells. Finally, we explored the potential of using PBMC gene signatures to predict the clinical outcome of BC patients.

## MATERIALS AND METHODS

### Overview of Patient Cohorts

In this study, we recruited 33 BC patients from the First Affiliated Hospital of Nanjing Medical University, between July and September 2017, as a discovery cohort. All patients participated anonymously in consideration of privacy and security concerns. The detailed baseline demographic information of the discovery cohort is listed in **Table 1**. In IHC subtyping, ER positive, HER2 negative, high PR expression (more than 20%) and low Ki-67 expression (<14%) patients were defined as luminal-A subtype. ER positive, HER2 negative, low PR expression (<20%) or high Ki-67 expression (more than 14%) patients were defined as luminal-B subtype. Additionally, ER positive and HER2 positive patients were defined as luminal-B subtype as well (19). Upon recruitment, fresh peripheral blood samples were collected before clinical treatment. To validate the unsupervised classification of PBMC transcriptome in BC patients, we also downloaded the whole blood gene expression data and the clinical features of another BC cohort from European Genome-phenome Archive (accession number: EGAD00010001063) (20). This validation cohort includes 173 BC patients in the Norwegian Women and Cancer Study (21). The whole blood transcriptome was quantified by Illumina Human AWG-6 and HT12, including microarray expression data for 16,782 genes (21). The baseline characteristics of BC patients in the validation cohort are shown in **Additional File 1**. To estimate the proportion of tumor infiltrated lymphocytes (TILs) in BC, we also downloaded the transcriptome level gene expression data of 173 tumor tissue samples for all patients in the validation cohort from European Genome-phenome Archive (accession number: EGAD00010001064) (21).

### Isolation of Total RNA From PBMC and RNA-Seq

PBMC samples of 33 BC patients in the discovery cohort were isolated from whole blood applying Ficoll-Paque Premium (GE Healthcare, IL, USA) according to the manufacturer's instructions. Total RNA was extracted from PBMC using TRIzol reagent (Invitrogen, CA, USA) and purified with the mirVana RNA Isolation Kit (Ambion, Massachusetts, USA) in accordance with the manufacturer's protocol. The purity and concentration of RNA were determined from OD260/280 readings using NanoDrop ND-1000. RNA integrity was determined by 1% formaldehyde denaturing gel electrophoresis. Only RNA extracts with RNA integrity number values >6 were used for further experiments. The isolated RNAs were immediately frozen in liquid nitrogen, and stored at  $-80^{\circ}\text{C}$ . RNA-seq libraries were constructed by *Ovation* human FFPE RNA-seq library systems

**TABLE 1** | Demographics of BC patients in the discovery cohort.

Characteristic	All patients (n = 33)
<b>Age (y)*</b>	51.3 (24–77)
<b>Menopausal status</b>	
Premenopausal	0
Postmenopausal	0
Not available	33
<b>Histological type</b>	
Invasive ductal carcinoma	30
Invasive lobular carcinoma	1
Intraductal carcinoma	2
<b>ER status</b>	
Positive	22
Negative	11
<b>PR status</b>	
Positive	19
Negative	14
<b>HER2 status</b>	
Positive	19
Negative	14
<b>Ki-67 status</b>	
Less than 20%	9
More than 20%	24
<b>IHC-based subtypes</b>	
Luminal-A	16
Luminal-B	6
HER2-positive	3
Triple negative	8
<b>Pathological stage</b>	
Stage I	16
Stage II	11
Stage III	3
Not available	3

Unless otherwise indicated, data are number of patients. \*Data for continuous variables are means, with ranges in parentheses.

(NuGEN Technologies, CA, USA) and sequenced on Illumina HiSeq X Ten platform (Illumina, CA, USA) using paired-end 150 bp runs.

## RNA-Seq Data Analysis

RNA-seq reads were aligned to human genome 19 by HISAT2 (22), quantified by *featureCounts* (23) and assembled by *StringTie* (24). The expression level of genes was quantified in forms of both counts data and normalized FPKM (fragments per kilobase of exon per million reads mapped). In total, expression values of 57,773 unique genes in PBMC samples of BC patients in the discovery cohort were measured. Considering the different types of gene expression profiles in the discovery and validation cohorts, GLM in *DESeq2* (25) was used to perform the differential gene expression analysis for RNA-seq data, while linear models in *limma* (26) was used for microarray data. Genes with a fold change in expression level of <0.25 or >4.0 and FDR-corrected

$P < 0.01$  were identified as significant differentially expressed genes (DEGs). The annotation and enrichment visualization of DEGs were accomplished using *Metascape* (<http://metascape.org>) (27) and *Reactome* pathway database (<https://reactome.org/>) (28). The Gene Ontology (GO) terms, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways and *Reactome* pathways with a  $P < 1 \times 10^{-5}$  in the enrichment analysis were retained.

## Discovery and Validation of the PBMC Subtypes

We used unsupervised consensus clustering (29) to discover intrinsic PBMC subtypes in the discovery and validation cohorts, respectively. The consensus clustering is a resampling-based method to represent the consensus across multiple runs of a clustering algorithm and to assess the stability of the discovered clusters (29). The method, which is robust and insensitive to the initial conditions, has been widely used to identify biologically meaningful clusters (29). In detail, we first selected the top 5,000 variable genes measured by median absolute deviation as the most informative genes for class detection. Then, we performed a bootstrap procedure with 80% item resampling and 80% gene resampling on the PBMC gene expression profiles 10,000 times using the agglomerative hierarchical clustering algorithm with the *Spearman* distance metric. We selected the optimal number of clusters that corresponds to the most stable consensus matrices and the most unambiguous cluster assignments across permuted clustering runs by varying the number of clusters from 2 to 10 (29). This process determined the optimal number of intrinsic unsupervised clusters defined by PBMC transcriptome in the discovery cohort. To validate the result, we implemented the same procedure on the validation cohort. In addition, we used in-group proportion (IGP) statistical analysis (30) to demonstrate the existence of the clusters in the validation cohort and evaluate the reproducibility of the clusters derived from consensus clustering in the two independent cohorts. IGP provides a quantitative approach to measure the similarity between the clusters. IGP will be 100% if the clusters are identical between two datasets and will be 0% conversely. Due to the different types of expression values in the two datasets, we normalized the expression data by Z-score prior to the IGP statistical analysis. The consensus clustering and IGP analysis were performed in R (<https://www.r-project.org/>) (31).

## Estimation of the Abundance of Major Immune Cells Using Gene Expression Profiles

We used *CIBERSORT* (<https://cibersort.stanford.edu/>) (32) with the LM22 signature gene matrix (32) to characterize the proportion of immune cells in the PBMC sample of each BC patient in both discovery and validation cohorts. *CIBERSORT* is able to accurately estimate cell composition of complex tissues from their gene expression profiles, including the immune cells in human PBMC samples (32). We obtained the proportion of seven major immune cell types, including lymphocytes (consisting of all types of B cells, T cells, and NK cells), monocytes, macrophages (consisting of M0, M1, M2 macrophages), dendritic cells

(consisting of resting and activated dendritic cells), mast cells (consisting of resting and activated mast cells), eosinophils and neutrophils. All subsequent analysis of immune cell proportions in this study was based on the estimation of these seven major cell types.

## Survival Analysis

We identified the immune-related gene signatures using their expression in PBMC samples. To explore the implication of the immune-related gene signatures on the patient's survival, we used *Kaplan-Meier-plotter* (<http://www.kmplot.com/>) (33) to perform *in silico* survival analysis. *Kaplan-Meier-plotter* is able to assess the effect of 54,000 genes on cancer survival in 21 cancer types, including BC, using their expression profiles in the tumor tissue (33).

## Statistical Analysis

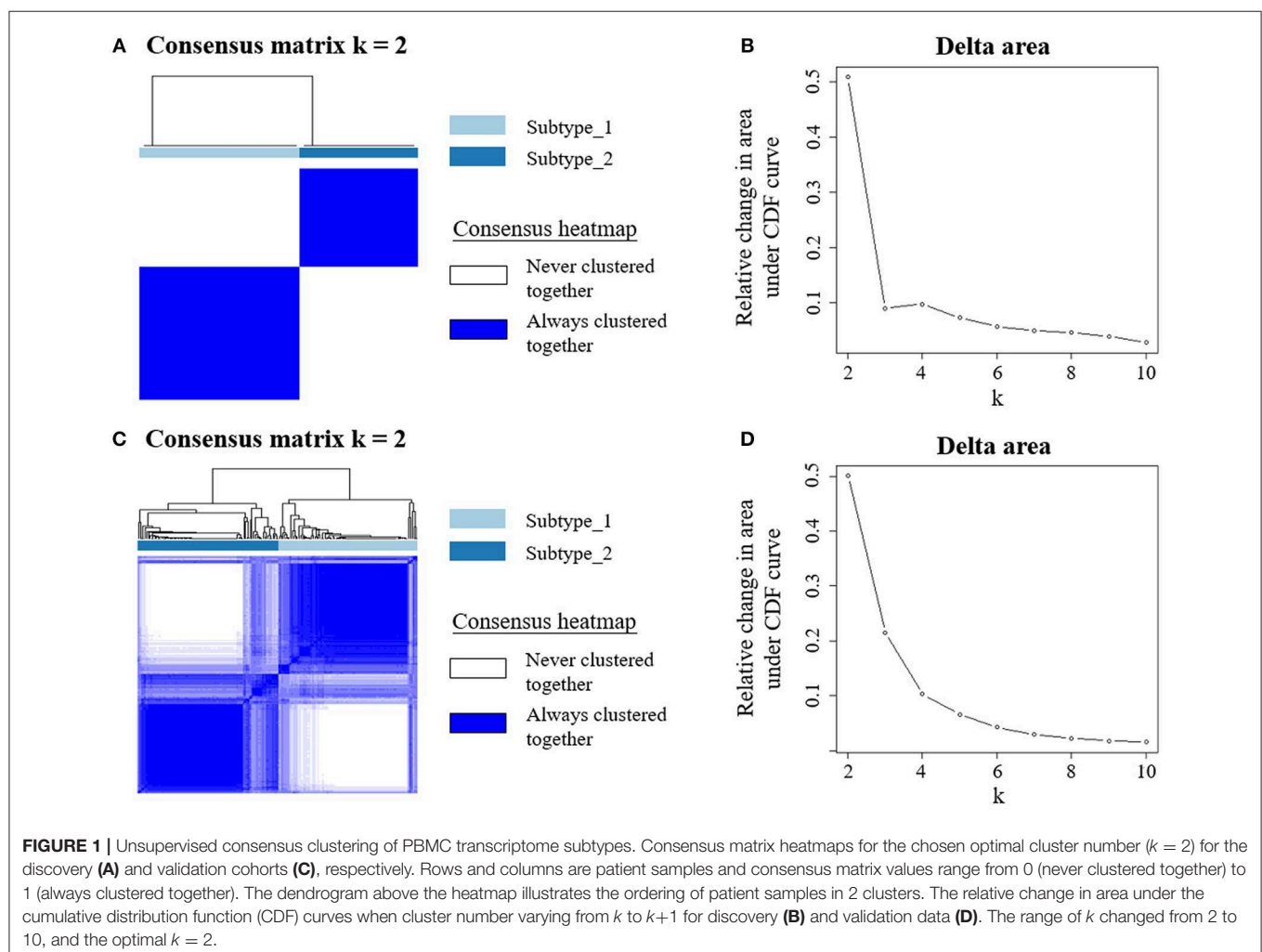
To compare the clinical characteristics, cell proportions and established subtypes between clusters in both cohorts, we performed the *Fisher's* exact test or *Pearson's* chi-squared test for categorical variables and the *Student's t*-test for continuous

variables. All statistical analysis were performed in R (<https://www.r-project.org/>) (31).

## RESULTS

### Established Clinical Classifications Cannot Explain PBMC Expression Heterogeneity Among BC Patients

First, we explored the heterogeneity of PBMC transcriptome among the BC patients. We observed that a substantial number of genes varied significantly in expression in PBMC samples of the BC patients in both cohorts (**Additional File 2**). To explain this variation, we projected the PBMC transcriptome differences among BC patient groups onto known clinical classification. In the discovery cohort, the status of three immunohistochemistry (IHC) markers was available for each patient. We classified BC patients using all three IHC markers' status and compared the gene expression of BC patients with different ER, PR, and HER2 status. No significant difference was found between BC patients with different IHC markers' status (**Additional File 3**).





In the validation cohort, only the status of ER and HER2 was available. We tested the expression differences in patients with ER and HER2 status. Again, we found no significant difference (**Additional File 4**). In addition, gene expression profile of the matched tumor tissue is available for each patient in the validation cohort. With the expression data, we further classified the patients in the validation cohort into PAM50 subtypes (2) and investigated the PBMC transcriptome variations among these patient groups. The result indicated that PBMC gene expression in the BC patients with different PAM50 subtypes are statistically similar (**Additional File 4**). All these results suggested that the established known subtypes based on IHC marker and PAM50 were not associated with PBMC gene expression in BC patients.

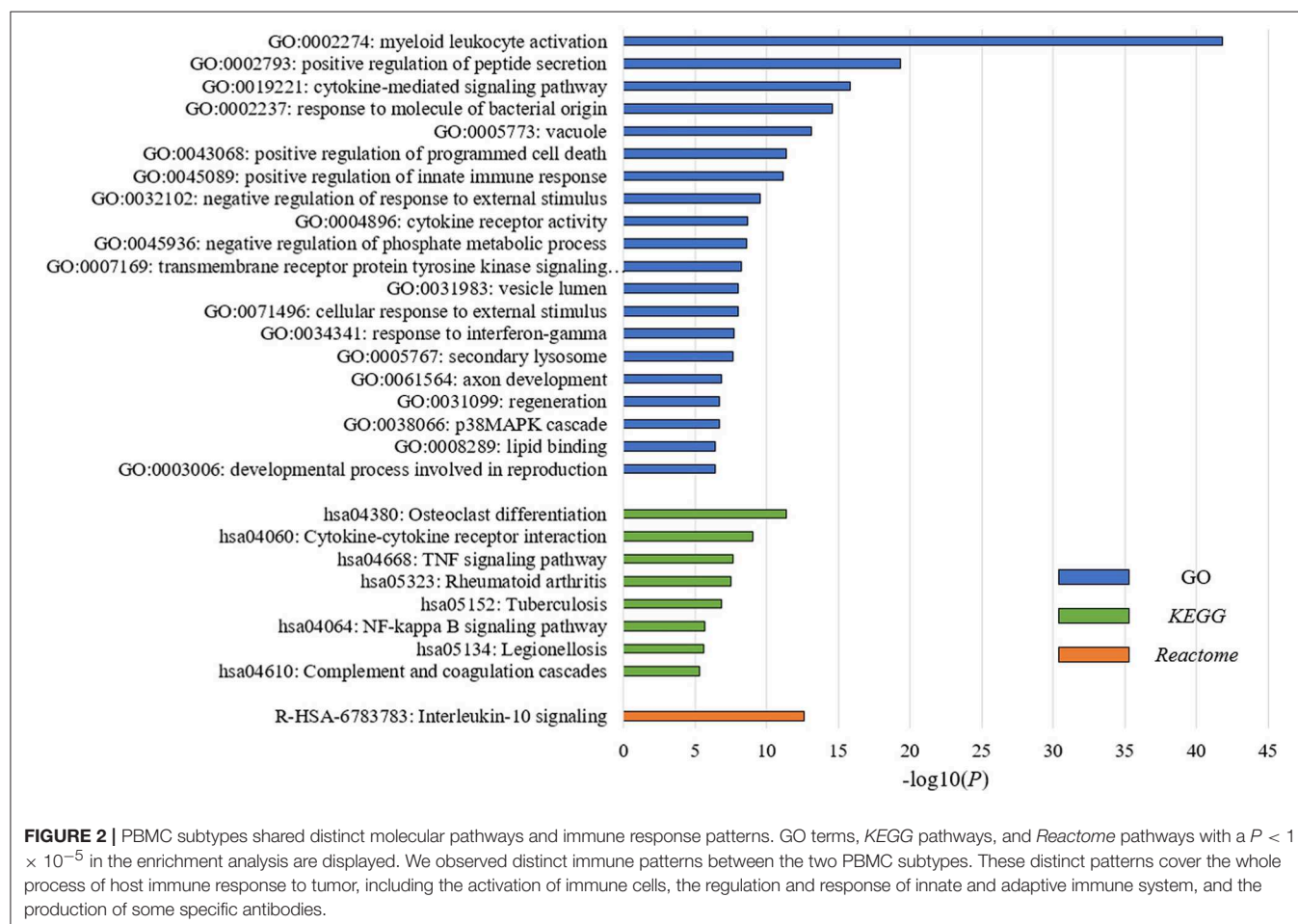
## Identification and Validation for PBMC Transcriptome-Based Subtypes for BC Patients

Next, we employed an unsupervised clustering algorithm to classify the BC patients into *de novo* groups based on their heterogeneity of systemic immune response to BC. We selected the top 5,000 genes with the highest median absolute deviation of expression values in the discovery cohort, and classified BC patients into two clusters, subtype\_1 and subtype\_2 (**Figure 1A**),

using the consensus clustering algorithm (29). The 2-cluster solution corresponded to the largest cluster number that induced the least incremental change in the area under the cumulative distribution function (CDF) curves while keeping the maximal consensus within clusters and the minimal rate of ambiguity in cluster assignments (**Figure 1B**). Finally, subtype\_1 includes 19 patients (58%), while subtype\_2 includes 14 patients (42%).

To confirm this *de novo* classification, we independently applied the same analysis procedure (29) on the validation dataset, which is whole blood transcriptome data. Interestingly, we observed that the samples in the validation cohort were also clustered into two optimal clusters, which is very similar to that identified in the discovery dataset (**Figures 1C,D**). We evaluated the reproducibility of the two PBMC subtypes across the discovery and validation cohorts using in-group proportion (IGP) statistic (30). The IGP values are 89.8 and 75.3% for subtype\_1 and subtype\_2, respectively, indicating that both subtypes had high consistency between the two cohorts. This suggested that these two PBMC transcriptome subtypes are robust across different BC cohorts.

To understand the underlying biological mechanisms that differ in these two PBMC subtypes, we performed differential gene expression analysis using *DESeq2* (25). We observed 1,988 DEGs between these two subtypes in the discovery cohort.





**TABLE 2 |** Differences of established BC subtypes and clinical characteristics in PBMC subtypes in the discovery cohort.

Clinical factors	Subtype_1 (n = 19)	Subtype_2 (n = 14)	P-value
<b>ER status</b>			1 <sup>a</sup>
ER+ subtype	13	9	
ER- subtype	6	5	
<b>PR status</b>			1 <sup>a</sup>
PR+ subtype	11	8	
PR- subtype	8	6	
<b>HER2 status</b>			0.723 <sup>a</sup>
HER2+ subtype	10	9	
HER2- subtype	9	5	
<b>IHC-based subtype</b>			0.309 <sup>a</sup>
Luminal-A	10	6	
Luminal-B	3	3	
HER2-positive	3	0	
Triple negative	3	5	
<b>Histological type</b>			0.496 <sup>a</sup>
Invasive ductal carcinoma	16	14	
Invasive lobular carcinoma	1	0	
Intraductal carcinoma	2	0	
<b>Pathological stage</b>			0.169 <sup>a</sup>
Stage I	10	6	
Stage II	4	7	
Stage III	3	0	
<b>Age*(y)</b>	47.9 (24–73)	55.9 (41–77)	0.052 <sup>b</sup>

Unless otherwise indicated, data are number of patients or the P-value of statistical test.

<sup>a</sup>P-value for the Fisher's exact test.

<sup>b</sup>P-value for the Student's t-test.

\*Data for the continuous variables are means with ranges in parentheses.

In enrichment analysis for the DEGs, the top 20 significantly enriched GO terms are related to immune regulation (**Figure 2**). Among them, myeloid leukocyte activation was the most significant GO term. Similarly, the enriched *KEGG* pathways and *Reactome* pathways (**Figure 2**) include osteoclast differentiation and interleukin-10 signaling, which associate to host immune response. The results suggested that the major differences between these two subtypes may be explained by their different immune responses to BC.

## PBMC Transcriptome Subtypes Are Distinct in Terms of Immune Cell Abundance

Then, we investigated possible clinical factors that relate to the two subtypes in the BC patients, including age, clinical stage, established BC subtype, blood immune cell abundance, and TILs. In the discovery cohort, there was no statistical difference between the two subtypes in terms of age, histological type or clinical stage (**Table 2**), or age, menopausal status or weight in the validation cohort (**Table 3**). Moreover, we found that the known established BC subtypes, including IHC marker status, IHC-based subtypes, and PAM50 intrinsic molecular subtypes,

**TABLE 3 |** Differences of established BC subtypes and clinical characteristics in PBMC subtypes in the validation cohort.

Clinical factors	Subtype_1 (n = 88)	Subtype_2 (n = 85)	P-value
<b>ER status</b>			0.301 <sup>a</sup>
ER+ subtype	68	71	
ER- subtype	20	14	
<b>HER2 status</b>			0.6973 <sup>a</sup>
HER2+ subtype	18	22	
HER2- subtype	70	63	
<b>PAM50 molecular subtype</b>			0.656 <sup>a</sup>
Luminal-A	24	24	
Luminal-B	21	22	
HER2-enriched	11	15	
Basal-like	15	14	
Normal-like	17	10	
<b>Menopausal status</b>			0.429 <sup>a</sup>
Premenopausal	6	7	
Postmenopausal	74	64	
<b>Weight*(kg)</b>	72.9 (50–120)	70.2 (50–150)	0.219 <sup>b</sup>
<b>Age*(y)</b>	56.7 (43–104)	55.9 (44–66)	0.445 <sup>b</sup>

Unless otherwise indicated, data are number of patients or the P-value of statistical test.

<sup>a</sup>P-value for Pearson chi-square test.

<sup>b</sup>P-value for Student's t-test.

\*Data for continuous variables are means, with ranges in parentheses.

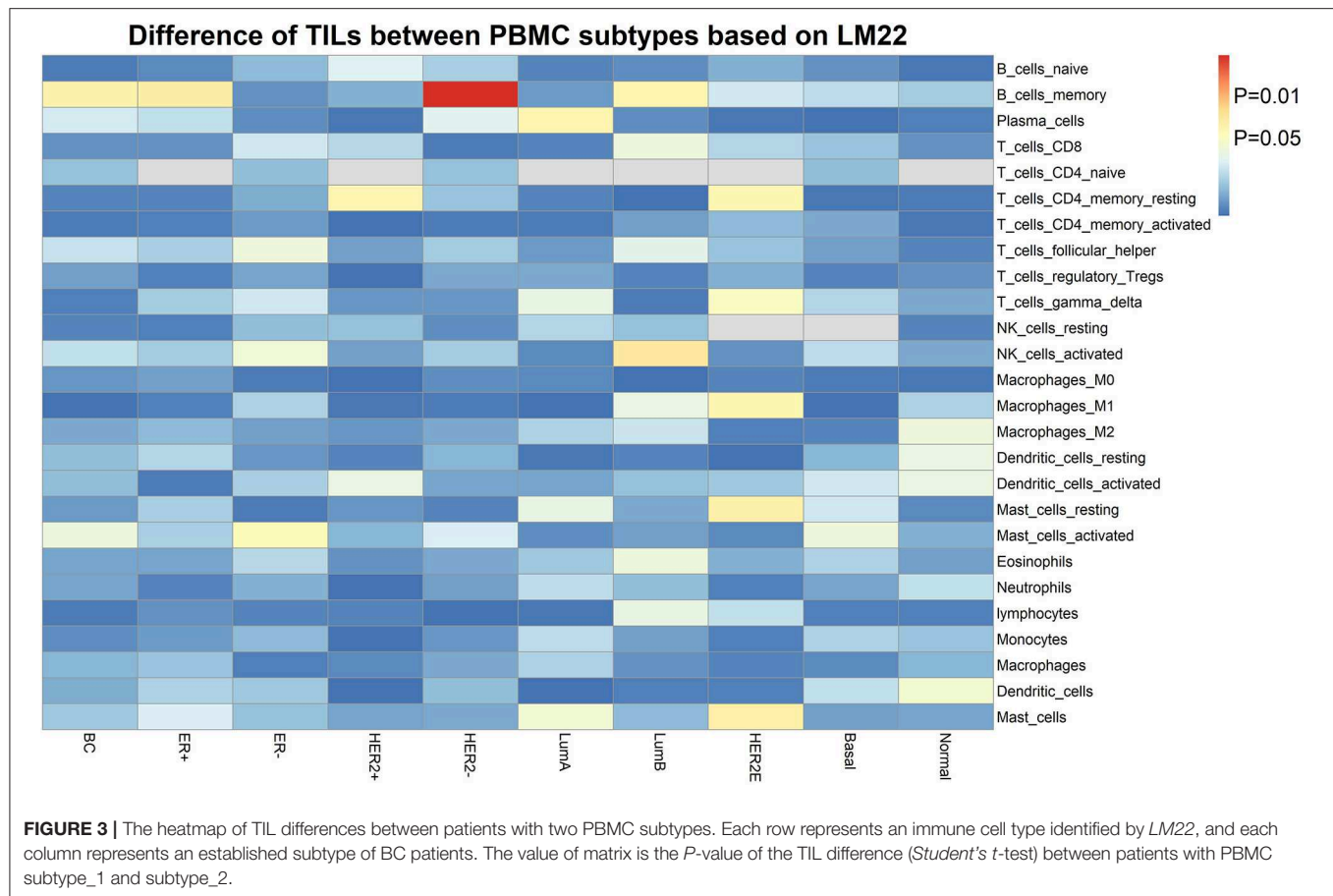
**TABLE 4 |** Differences of immune cell components in PBMC subtypes in the discovery and validation cohorts.

Immune cell types	Discovery cohort (n = 33)	Validation cohort (n = 173)
Lymphocytes	$5.22 \times 10^{-12}^*$	$5.80 \times 10^{-18}^*$
Monocytes	$5.29 \times 10^{-5}^*$	0.509
Macrophages	0.579	0.00026*
Dendritic cells	0.001*	0.252
Mast cells	0.022	0.00076*
Eosinophils	0.399	0.166
Neutrophils	$1.13 \times 10^{-14}^*$	$1.86 \times 10^{-24}^*$

Unless otherwise indicated, data are the P-value of Student's t-test. \* $P < 0.01$ .

cannot account for the differences between PBMC transcriptome subtypes (**Tables 2, 3**), because both PBMC subtypes contained the BC patients with IHC marker status and PAM50 subtypes.

Interestingly, we observed significant differences in proportion of lymphocytes (in the discovery cohort:  $P = 5.22 \times 10^{-12}$ ; in the validation cohort:  $P = 5.80 \times 10^{-18}$ ) and proportion of neutrophils (in the discovery cohort:  $P = 1.13 \times 10^{-14}$ ; in the validation cohort:  $P = 1.86 \times 10^{-24}$ ) between the two PBMC transcriptome-based subtypes (**Table 4**). Furthermore, we calculated the neutrophil-to-lymphocyte ratio (NLR), a common and stable hematological indicator that can reflect the inflammatory state of the body (34, 35). In comparing the NLR values between the two subtypes, we also observed a



significant difference (in the discovery cohort:  $P = 6.60 \times 10^{-6}$ ; in the validation cohort:  $P = 9.08 \times 10^{-21}$ ). Other immune cells, such as the monocytes in the discovery cohort and the macrophages in the validation cohort, do not show a significant difference (Table 4).

Furthermore, we assessed the TIL differences in tumor tissue samples of patients with different PBMC transcriptome subtypes. We estimated the proportion of immune cells in the tumor tissue sample of each BC patient in the validation cohort, using CIBERSORT with the LM22 signature (32). We found the tumor infiltration of memory B cells is statistically different in BC patients with two PBMC transcriptome subtypes (Figure 3), including all BC patients ( $P = 0.032$ ), ER+ patients ( $P = 0.027$ ), Luminal-B patients ( $P = 0.036$ ) and HER2- patients ( $P = 0.0022$ ). Additionally, memory resting CD4+ T cells is differentially infiltrated in cancer tissues of patients with different PBMC subtypes in HER2+ patients ( $P = 0.034$ ) and HER2-enriched patients ( $P = 0.037$ ).

These results suggested that the composition of immune cells in PBMCs and TILs in tumor tissues, rather than age, clinical stage, and known BC subtypes, are related to the heterogeneity of PBMC transcriptome in BC patients.

## PBMC Transcriptome Subtypes May Be Related to BC Survival

Finally, we tried to explore the implications of the PBMC transcriptome heterogeneity on BC management. In the previous results, we found no difference in several available clinical characteristics between the two subtypes (Table 3). However, NLR, which is an indicator of the inflammation level, differed between the two subtypes. The inflammation level has important potential in predicting the clinical outcome of BC (36). We investigated if patients with different PBMC subtypes have different survival rate. Twenty-eight immune-related genes were identified in the pathway of osteoclast differentiation, which is the most enriched KEGG pathway (Table 5). Expression values of all the 28 genes were significantly higher in subtype\_2 than in subtype\_1 (Figure 4). Using Kaplan-Meier-plotter (33), we observed that the tissue expression values of the 28-gene signature had the ability to predict the clinical outcomes of all subtypes of BC patients (Figure 5A), as well as ER positive patients (Figure 5B), basal-like patients (Figure 5C) and clinical stage III patients (Figure 5D). The high expression of these 28 genes in tumor tissue, including IFNGR1, IFNGR2, IL1A, IL1B, TLR2, TLR4, FOSL1, and CSF1, associates with a lower

**TABLE 5 |** Gene symbols of the 28-gene signature.

	Gene symbol
28-gene signature	TYROBP, IFNGR1, GAB2, TNFRSF1A, PTGS2, NFKB2, NFKBIA, SIRPB1, NFKBIB, RELB, IL1A, IL1R1, IL1B, TLR4, TLR2, FCGR2A, IFNGR2, FCGR3B, JUNB, FOSL1, JUN, SOCS3, SIRPA, CR1, LILRB3, LILRA2, LILRA6, CSF1

risk of cancer recurrence and better survival rate in BC patients (**Figure 5**).

Furthermore, we repeated the analysis above and identified 16 immune-related genes in the most enriched *Reactome* pathway (**Additional File 5**). Similarly, 16 genes including IL1R2, CXCL1, CXCL8, PTGS2, IL1A, IL1RN, and CSF1 were highly expressed in the subtype\_2 BC patients (**Additional File 6**). High expression of these genes in tumor tissue, were related to a low risk of recurrence and better survival rate in all subtypes of BC patients, ER positive patients, luminal-A patients, luminal-B patients and clinical stage III patients (**Additional File 7**). However, both gene signatures had no statistical power in differentiating the clinical outcomes of PR positive patients, HER2 positive patients, HER2-enriched patients, or other clinical stages BC patients (detailed in **Additional Files 7, 8**).

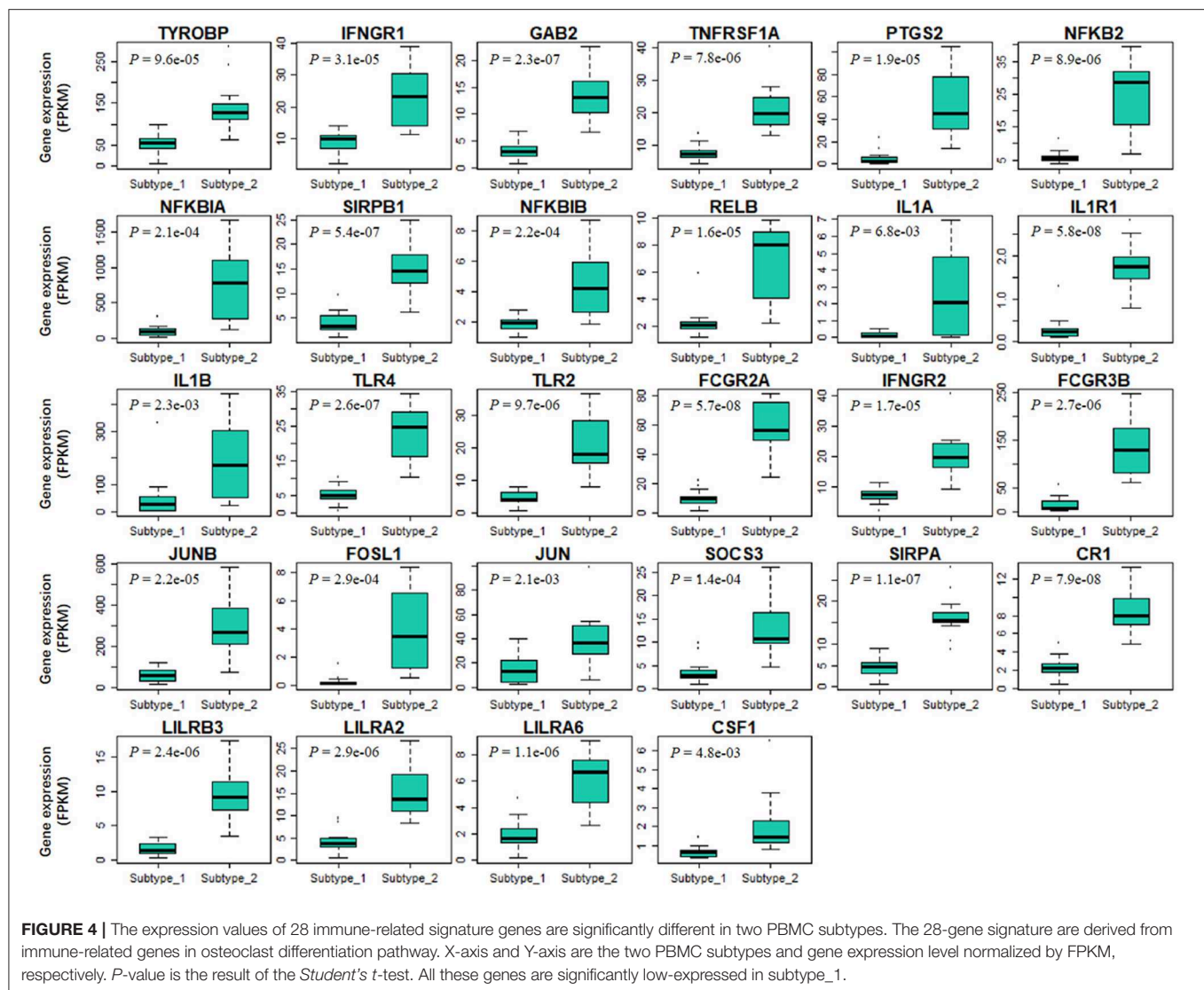
## DISCUSSION

In this study, we revealed substantial heterogeneity of PBMC transcriptome in BC patients (**Additional File 2**) and identified two subtypes based on the PBMC gene expression profiles (**Figure 1**). Our results indicated that these two subtypes had distinct molecular pathways in host immune response and regulation (**Figure 2**). We observed that the PBMC-transcriptome based subtyping was a novel and independent classification for BC patients. The essential molecular basis of the subtyping reflects the interaction between host immune system and BC. We found that the proportion of immune cells in peripheral blood, especially lymphocytes and neutrophils, shaped the significant differences between the two subtypes (**Table 4**). Furthermore, two gene signatures that discriminates these two PBMC subtypes are able to predict the clinical outcomes of BC patients (**Figure 5** and **Additional File 7**). Importantly, such subtyping is general and robust, since they were independently observed in both the discovery dataset and validation dataset. In the discovery dataset, we quantified PBMC transcriptome using RNA-seq technology, while the transcriptome data in the validation dataset was gene expression array (12). Although the quantification platform and source samples are different in these two datasets, the findings are consistent (**Figure 1**). However, a future study using a large prospective cohort will be highly helpful to validate these two PBMC subtypes in BC, since the sample size in the discovery cohort is relatively small.

Current clinical classifications did not reflect the heterogeneity of interactions between BC and host immune system (**Tables 2, 3**). This is consistent with several previous findings, suggesting that transcriptional fingerprint of BC

subtypes is not the predominant signal in the patient's systemic immune response (14, 21). Thus, it was difficult to classify BC patients into classical BC subtypes using the PBMC expression profiles. The classification of the established BC subtypes was based on the expression of several important makers in tumor tissue, including ER, PR, and HER2 (6, 7). In contrast, PBMCs contains the major inflammatory or supportive cells, which are composed of the main stromal components of tumor microenvironment and govern the systemic inflammatory responses in human malignancies, including BC (37). Therefore, it was reasonable that PBMC transcriptome cannot mirror the different expression profiles in tissue samples among BC patients of different clinical subtypes. Instead, PBMC gene expression profiles might be useful for early diagnosis of human cancers, such as BC and colorectal cancer (11–13, 38).

In order to explore the heterogeneity of host systemic immune response to BC, we employed an unsupervised clustering algorithm to cluster BC patients using PBMC gene expression data, and revealed two distinct subtypes (**Figure 1**). Functional annotation and enrichment analysis displayed distinguishing immune patterns between the two subtypes (**Figure 2**). These distinct patterns covered the whole process of host immune response to tumor, including the activation of immune cells, the regulation and response of innate and adaptive immune system, and the production of some specific antibodies. Considering *KEGG* categorizes genes into meaningful biological pathways, which makes the interpretation more straightforward (39), we focused on the enriched *KEGG* pathways below. In our results, osteoclast differentiation, cytokine-cytokine receptor interaction and TNF signaling pathway were the top three *KEGG* pathways that had distinct expression patterns between the two subtypes. Osteoclasts are multinucleated cells of monocyte/macrophage origin that degrade bone matrix. The differentiation of osteoclasts is dependent on a tumor necrosis factor (TNF) family cytokine, receptor activator of nuclear factor (NF)- $\kappa$ B ligand (RANKL), as well as macrophage colony-stimulating factor (M-CSF) (40). BC frequently metastasizes to the skeleton, interfering with the normal bone remodeling process and inducing bone degradation (41, 42). Cytokines are highly inducible, secretory proteins that mediate intercellular communication in the immune system. Cytokine and cytokine receptor interaction are regarded as crucial aspects of inflammation and tumor immunology (43). Although the exact initiation process of BC is unknown, inflammation has been proposed as an important factor in tumor initiation, promotion, angiogenesis, and metastasis, in which cytokines are prominent players (44, 45). Moreover, many studies suggested that cytokines play an important role in the regulation of both induction and protection in BC (46, 47). TNF is a proinflammatory cytokine that plays a critical role in diverse cellular events, including cell proliferation, differentiation and apoptosis (48). TNF- $\alpha$  is an important inflammatory factor that acts as a master switch in establishing an intricate link between inflammation and cancer (48). A wide variety of evidence has pointed to a pivotal role of TNF- $\alpha$  in tumor proliferation, migration, invasion and angiogenesis, including BC (49, 50). These enriched pathways hinted that the different status of inflammation may partly explain the differences between PBMC



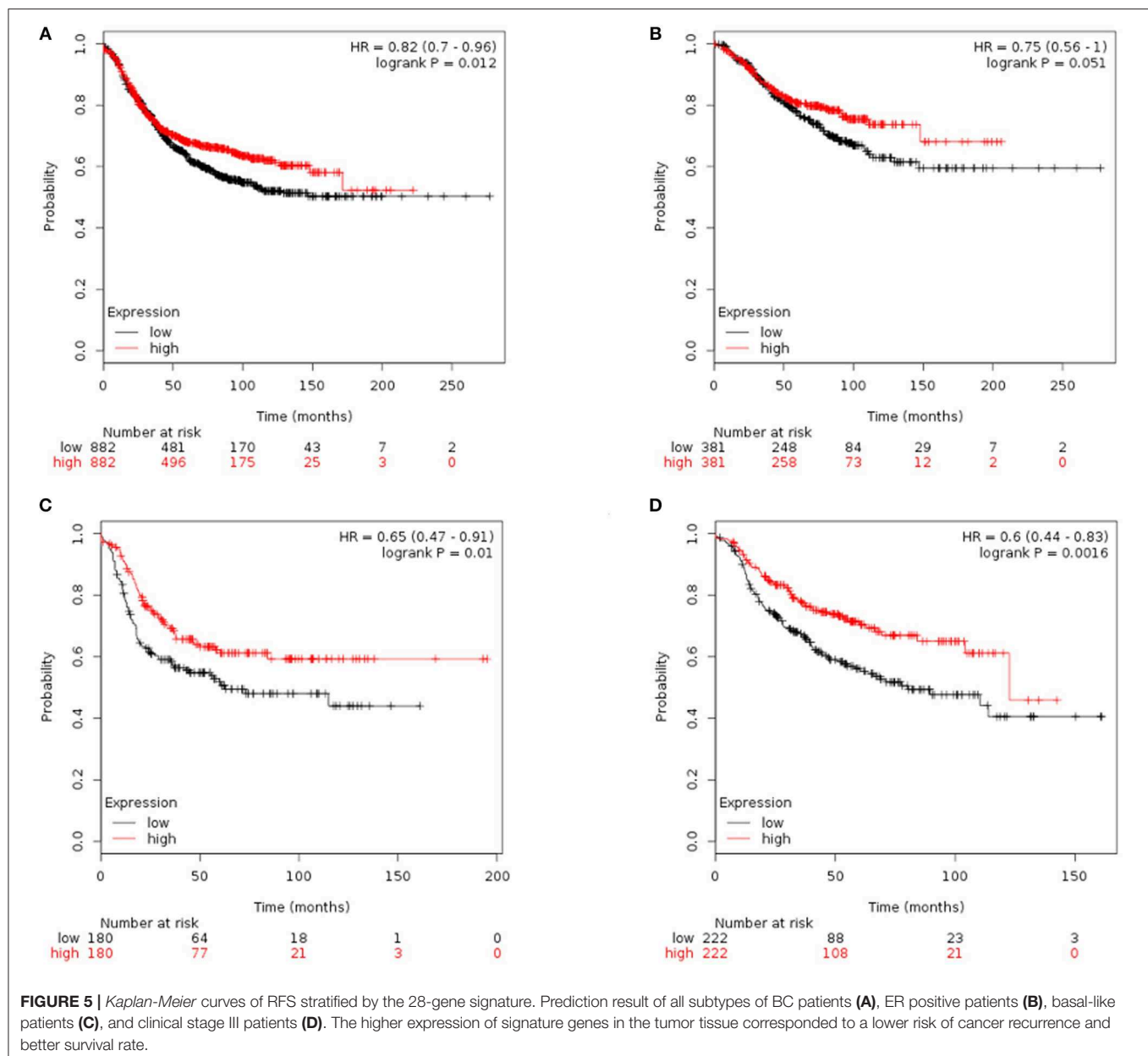
transcriptome subtypes of BC patients, which may be related to BC metastasis.

The correlation of PBMC heterogeneity to BC metastasis is also confirmed by the differential analysis of immune cell proportions. Our results showed significant differences of the proportions of lymphocytes and neutrophils in the peripheral blood and the neutrophil-to-lymphocyte ratio (NLR) in the two subtypes (Table 4). The proportion of lymphocytes in subtype<sub>1</sub> was higher than that in subtype<sub>2</sub>, whereas neutrophils were merely the major component of PBMCs in subtype 2. Several previous studies suggested that peripheral blood lymphocytes expressed abundant information about the interactions between the tumors and the host immune system, which are useful biomarkers for predicting the risk of cancer occurrence and recurrence (51, 52). Neutrophils, altering the local microenvironment by releasing inflammatory signals and promoting the formation of metastases, were considered as the main driving force of pulmonary metastatic colonization

of BC cells (36, 53, 54). Neutrophils were also observed to be useful biomarkers for clinical BC diagnosis and prognosis assessment (36, 53, 54). Additionally, the pre-treatment NLR was a prognostic factor for BC (34, 35, 55, 56). A higher NLR was associated with poorer recurrence-free survival in BC patients (34, 35, 55, 56). In addition to immune cells that are circulating in the peripheral blood, BC patients with different PBMC transcriptome subtypes showed distinct TILs in tumor tissues (Figure 3). Although the precise role of tumor-infiltrating lymphocytes in cancer development and metastasis is not well-understood and remains controversial, accumulating evidences suggest that the adaptive immunity mediated by T and B lymphocytes provides a critical foundation for effective and sustained antitumor responses (57).

Above evidences hinted that patients with different PBMC transcriptome subtypes may have different clinical outcomes. However, due to the limitation of small sample size and insufficient clinical data, the direct association between the





PBMC subtypes and disease recurrence or cancer survival remains unexplored in our analysis. To partially overcome this, we identified two immune-related gene signatures in PBMCs and examined their power of predicting clinical outcomes using *in silico* prognostic analysis on their expressions in BC tissue samples. Both gene signatures showed the ability to predict the survival of BC patients (Figure 5 and Additional File 7), which is similar to the findings observed by Foulds et al. (14). In their study, they measured PBMC expression values of 800 immune-related genes and investigated their implications on clinical outcomes. They reported that the expression of CD163, CXCR4, and THBS1 in PBMCs could predict the relapse-free survival for triple negative BC patients (14). In

our results, the higher expression of signature genes in tumor tissue corresponded to a lower risk of cancer recurrence and better survival rate. Interestingly, the BC patients with subtype\_1 might had smaller metastasis probability and better prognosis, because they had higher proportion of lymphocytes, smaller proportion of neutrophils and lower NLR. However, the expression values of the two sets of signature genes were down-regulated in subtype\_1. Therefore, we proposed that the up-regulation expression of immune-related genes in peripheral blood is probably related to a down-regulated expression in tumor tissue. This is very similar to the findings in literatures that the regulation of immune-related gene expression is opposite in blood and tissue (58).



## CONCLUSIONS

In conclusion, we identified two new subtypes of BC based on their PBMC expression profiles. The two PBMC transcriptome subtypes had distinct immune patterns, which was associated with different immune cell abundances. *In silico* prognostic analysis suggested that BC patients of the two subtypes may have different clinical outcomes. Although this classification is probably useful for personalized BC management, further investigation in a large prospective setting is required to ascertain their clinical values.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher.

## ETHICS STATEMENT

The study was approved by the ethical committee of the First Affiliated Hospital of Nanjing Medical University. All samples were used according to the ethical guidelines of the 1975 Declaration of Helsinki and obtained with the patients' understanding that it might be published. The written informed consent was obtained from the participants of this study.

## REFERENCES

- Fan L, Strasser-Weippl K, Li JJ, St. Louis J, Finkelstein DM, Yu KD, et al. Breast cancer in China. *Lancet Oncol.* (2014) 15:e279–89. doi: 10.1016/S1470-2045(13)70567-9
- Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. *Nature.* (2000) 406:747–52. doi: 10.1038/35021093
- Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA.* (2001) 98:10869–74. doi: 10.1073/pnas.191367098
- Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res.* (2010) 12:R68. doi: 10.1186/bcr2635
- Voduc KD, Cheang MCU, Tyldesley S, Gelmon K, Nielsen TO, Kennecke H. Breast cancer subtypes and the risk of local and regional relapse. *J Clin Oncol.* (2010) 28:1684–91. doi: 10.1200/JCO.2009.24.9284
- Hammond MEH, Hayes DF, Dowsett M, Allred DC, Hagerty KL, Badve S, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer (unabridged version). *Arch Pathol Lab Med.* (2010) 134:e48–72. doi: 10.1043/1543-2165-134.7.e48
- Harris LN, Ismaila N, McShane LM, Andre F, Collyar DE, Gonzalez-Angulo AM, et al. Use of biomarkers to guide decisions on adjuvant systemic therapy for women with early-stage invasive breast cancer: American Society of Clinical Oncology Clinical Practice Guideline. *J Clin Oncol.* (2016) 34:1134–50. doi: 10.1200/JCO.2015.65.2289
- Nagarajan D, McArdle SEB. Immune landscape of breast cancers. *Biomedicine.* (2018) 6:20. doi: 10.3390/biomedicine6010020
- Noy R, Jeffrey Pollard W. Tumor-associated macrophages: from mechanisms to therapy. *Immunity.* (2014) 41:49–61. doi: 10.1016/j.immuni.2014.06.010
- Talmadge JE, Gabrilovich DI. History of myeloid-derived suppressor cells. *Nat Rev Cancer.* (2013) 13:739. doi: 10.1038/nrc3581
- Suzuki E, Sugimoto M, Kawaguchi K, Pu F, Uozumi R, Yamaguchi A, et al. Gene expression profile of peripheral blood mononuclear cells may contribute to the identification and immunological classification of breast cancer patients. *Breast Cancer.* (2018) 26:282–9. doi: 10.1007/s12282-018-0920-2
- Dumeaux V, Ursini-Siegel J, Flatberg A, Fjosne HE, Frantzen JO, Holmen MM, et al. Peripheral blood cells inform on the presence of breast cancer: a population-based case-control study. *Int J Cancer.* (2015) 136:656–67. doi: 10.1002/ijc.29030
- Hensler M, Vančurová I, Becht E, Palata O, Strnad P, Tesarová P, et al. Gene expression profiling of circulating tumor cells and peripheral blood mononuclear cells from breast cancer patients. *OncoImmunology.* (2016) 5:e1102827. doi: 10.1080/2162402X.2015.1102827
- Foulds GA, Vadakekolathu J, Abdel-Fatah TMA, Nagarajan D, Reeder S, Johnson C, et al. Immune-phenotyping and transcriptomic profiling of peripheral blood mononuclear cells from patients with breast cancer: identification of a 3 gene signature which predicts relapse of triple negative breast cancer. *Front Immunol.* (2018) 9:2028. doi: 10.3389/fimmu.2018.02028
- Baine MJ, Chakraborty S, Smith LM, Mallya K, Sasson AR, Brand RE, et al. Transcriptional profiling of peripheral blood mononuclear cells in pancreatic cancer patients identifies novel genes with potential diagnostic utility. *PLoS ONE.* (2011) 6:e17014. doi: 10.1371/journal.pone.0017014
- Ciarloni L, Hosseinian Ehrensberger S, Imaizumi N, Monnier-Benoit S, Nichita C, Myung SJ, Kim JS, et al. Development and clinical validation of a blood test based on 29-gene expression for early detection of colorectal cancer. *Clin Cancer Res.* 22:4604–11. (2016) doi: 10.1158/1078-0432.CCR-15-2057
- Brodin P, Davis MM. Human immune system variation. *Nat Rev Immunol.* (2017) 17:21–9. doi: 10.1038/nri.2016.125
- Standish LJ, Sweet ES, Novack J, Wenner CA, Bridge C, Nelson A, et al. Breast cancer and the immune system. *J Soc Integr Oncol.* (2008) 6:158–68. doi: 10.2310/7200.2008.0027

## AUTHOR CONTRIBUTIONS

This study was conceptualized and designed by WM, XS, HL, YL, and WG. Samples were collected and provided by HX and YZ. RNA sequencing was completed by YB. ZH and YC contributed to the analysis of RNA-seq data and the estimation of the abundance of immune cells. WM completed the discovery and validation of new subtypes, and performed the survival analysis and statistical analysis. The draft manuscript was developed by WM and WG. All authors reviewed the draft and provided comments, contributing to the final version of the manuscript, read and approved the submission and publication.

## FUNDING

This work was funded by grants from National Key R&D Program of China (2018YFC1314900, 2018YFC1314902), Key Research & Development Program of Jiangsu Province (BE2016002-3), National Natural Science Foundation of China (61571109), and the Fundamental Research Funds for the Central Universities (2242017K3DN04).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2019.00985/full#supplementary-material>

19. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thürlimann B, Senn HJ. Panel, strategies for subtypes—dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. *Ann Oncol.* (2011) 22:1736–47. doi: 10.1093/annonc/mdr304
20. Lappalainen I, Almeida-King J, Kumanduri V, Senf A, Spalding JD, Ur-Rehman S, et al. The European Genome-phenome Archive of human data consented for biomedical research. *Nat Genet.* (2015) 47:692–5. doi: 10.1038/ng.3312
21. Dumeaux V, Fjukstad B, Fjosne HE, Frantzen JO, Holmen MM, Rodegerdts E, et al. Interactions between the tumor and the blood systemic response of breast cancer patients. *PLoS Comput Biol.* (2017) 13:e1005680. doi: 10.1371/journal.pcbi.1005680
22. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods.* (2015) 12:357–60. doi: 10.1038/nmeth.3317
23. Smyth GK, Shi W, Liao Y. FeatureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics.* (2013) 30:923–30. doi: 10.1093/bioinformatics/btt656
24. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol.* (2015) 33:290–5. doi: 10.1038/nbt.3122
25. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* (2014) 15:550. doi: 10.1186/s13059-014-0550-8
26. Ritchie ME, Smyth GK, Phipson B, Wu D, Hu Y, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* (2015) 43:e47. doi: 10.1093/nar/gkv007
27. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun.* (2019) 10:1523. doi: 10.1038/s41467-019-09234-6
28. Fabregat A, Jupe S, Matthews L, Sidiropoulos K, Gillespie M, Garapati P, et al. The reactome pathway knowledgebase. *Nucleic Acids Res.* (2018) 46:D649–55. doi: 10.1093/nar/gkx1132
29. Monti S, Tamayo P, Mesirov J, Golub T. Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. *Mach Learn.* (2003) 52:91–118. doi: 10.1023/a:1023949509487
30. Kapp AV, Tibshirani R. Are clusters found in one dataset present in another dataset? *Biostatistics.* (2006) 8:9–31. doi: 10.1093/biostatistics/kxj029
31. R Core Team. *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing (2017).
32. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods.* (2015) 12:453–7. doi: 10.1038/nmeth.3337
33. Györfy B, Lanczky A, Eklund AC, Denkert C, Budczies J, Li Q, et al. An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. *Breast Cancer Res Treat.* (2010) 123:725–31. doi: 10.1007/s10549-009-0674-9
34. Koh CH, Bhoo-Pathy N, Ng KL, Jabir RS, Tan GH, See MH, et al. Utility of pre-treatment neutrophil-lymphocyte ratio and platelet-lymphocyte ratio as prognostic factors in breast cancer. *Br J Cancer.* (2015) 113:150–8. doi: 10.1038/bjc.2015.183
35. Ethier JL, Desautels D, Templeton A, Shah PS, Amir E. Prognostic role of neutrophil-to-lymphocyte ratio in breast cancer: a systematic review and meta-analysis. *Breast Cancer Res.* (2017) 19:2. doi: 10.1186/s13058-016-0794-1
36. Whyte MKB, Walmsley SR, Grecian R. The role of neutrophils in cancer. *Br Med Bull.* (2018) 128:5–14. doi: 10.1093/bmb/ldy029
37. Mishra S, Srivastava AK, Suman S, Kumar V, Shukla Y. Circulating miRNAs revealed as surrogate molecular signatures for the early detection of breast cancer. *Cancer Lett.* (2015) 369:67–75. doi: 10.1016/j.canlet.2015.07.045
38. Aaroe J, Lindahl T, Dumeaux V, Saebo S, Tobin D, Hagen N, et al. Gene expression profiling of peripheral blood cells for early detection of breast cancer. *Breast Cancer Res.* (2010) 12:R7. doi: 10.1186/bcr2472
39. Baranzini SE, Galwey NW, Wang J, Khankhanian P, Lindberg R, Pelletier D, et al. Pathway and network-based analysis of genome-wide association studies in multiple sclerosis. *Hum Mol Genet.* (2009) 18:2078–90. doi: 10.1093/hmg/ddp120
40. Asagiri M, Takayanagi H. The molecular understanding of osteoclast differentiation. *Bone.* (2007) 40:251–64. doi: 10.1016/j.bone.2006.09.023
41. Le Pape F, Vargas G, Clézardin P. The role of osteoclasts in breast cancer bone metastasis. *J Bone Oncol.* (2016) 5:93–5. doi: 10.1016/j.jbo.2016.02.008
42. David Waning L, Khalid Mohammad S, Theresa Guise A. Cancer-associated osteoclast differentiation takes a good look in the miR(NA)ror. *Cancer Cell.* (2013) 24:407–9. doi: 10.1016/j.ccr.2013.10.001
43. Lee M, Rhee I. Cytokine signaling in tumor progression. *Immune Netw.* (2017). 17:214–27. doi: 10.4110/in.2017.17.4.214
44. Grivnenkov SI, Greten FR, Karin M. Immunity, inflammation, and cancer. *Cell.* (2010) 140:883–99. doi: 10.1016/j.cell.2010.01.025
45. Esquivel-Velázquez M, Ostoa-Saloma P, Palacios-Arreola MI, Nava-Castro KE, Castro JI, Morales-Montor J. The role of cytokines in breast cancer development and progression. *J Interferon Cytokine Res.* (2015) 35:1–16. doi: 10.1089/jir.2014.0026
46. Sasser AK, Sullivan NJ, Studebaker AW, Hendey LF, Axel AE, Hall BM. Interleukin-6 is a potent growth factor for ER-alpha-positive human breast cancer. *FASEB J.* (2007) 21:3763–70. doi: 10.1096/fj.07-8832com
47. Zu X, Zhang Q, Cao R, Liu J, Zhong J, Wen G, Cao D. Transforming growth factor- $\beta$  signaling in tumor initiation, progression and therapy in breast cancer: an update. *Cell Tissue Res.* (2012) 347:73–84. doi: 10.1007/s00441-011-1225-3
48. Liu ZG. Molecular mechanism of TNF signaling and beyond. *Cell Res.* (2005) 15:24–7. doi: 10.1038/sj.cr.7290259
49. Wu Y, Zhou BP. TNF- $\alpha$ /NF- $\kappa$ B/Snail pathway in cancer cell migration and invasion. *Br J Cancer.* (2010) 102:639–44. doi: 10.1038/sj.bjc.6605530
50. Wolczyk D, Zaremba-Czogalla M, Hryniewicz-Jankowska A, Tabola R, Grabowski K, Sikorski AF et al. TNF- $\alpha$  promotes breast cancer cell migration and enhances the concentration of membrane-associated proteases in lipid rafts. *Cell Oncol.* (2016) 39:353–63. doi: 10.1007/s13402-016-0280-x
51. Lando C, Ceppi M, Bonassi S, Bigatti MP, Cebulski-Wasilewska A, Fabianova E, et al. An increased micronucleus frequency in peripheral blood lymphocytes predicts the risk of cancer in humans. *Carcinogenesis.* (2007) 28:625–31. doi: 10.1093/carcin/bgl177
52. Ali HR, Provenzano E, Liu B, Dawson SJ, Caldas C, Rakha E, et al. Association between CD8+ T-cell infiltration and breast cancer survival in 12 439 patients. *Ann Oncol.* (2014) 25:1536–43. doi: 10.1093/annonc/mdl191
53. Wculek SK, Malanchi I. Neutrophils support lung colonization of metastasis-initiating breast cancer cells. *Nature.* (2015) 528:413–7. doi: 10.1038/nature16140
54. Ocana A, Nieto-Jiménez C, Pandiella A, Templeton AJ. Neutrophils in cancer: prognostic role and therapeutic strategies. *Mol Cancer.* (2017) 16:137. doi: 10.1186/s12943-017-0707-7
55. Azab B, Bhatt VR, Phookan J, Murukutla S, Kohn N, Terjanian T, et al. Usefulness of the neutrophil-to-lymphocyte ratio in predicting short- and long-term mortality in breast cancer patients. *Ann Surg Oncol.* (2012) 19:217–24. doi: 10.1245/s10434-011-1814-0
56. Jia W, Wu J, Jia H, Yang Y, Zhang X, Chen K, et al. The peripheral blood neutrophil-to-lymphocyte ratio is superior to the lymphocyte-to-monocyte ratio for predicting the long-term survival of triple-negative breast cancer patients. *PLoS ONE.* (2015) 10:e0143061. doi: 10.1371/journal.pone.0143061
57. Salgado R, Denkert C, Demaria S, Sirtaine N, Klauschen F, Pruneri G, et al. The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: recommendations by an International TILs Working Group 2014. *Ann Oncol.* (2014) 26:259–71. doi: 10.1093/annonc/mdl450
58. Zhang XC, Wang J, Shao GG, Wang Q, Qu X, Wang B, et al. Comprehensive genomic and immunological characterization of Chinese non-small cell lung cancer patients. *Nat Commun.* (2019) 10:1772. doi: 10.1038/s41467-019-09762-1

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Ming, Xie, Hu, Chen, Zhu, Bai, Liu, Sun, Liu and Gu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# ANLN and KDR Are Jointly Prognostic of Breast Cancer Survival and Can Be Modulated for Triple Negative Breast Cancer Control

Xiaofeng Dai<sup>1\*</sup>, Yi Mei<sup>2</sup>, Xiao Chen<sup>2</sup> and Dongyan Cai<sup>3</sup>

<sup>1</sup> Wuxi School of Medicine, Jiangnan University, Wuxi, China, <sup>2</sup> School of Biotechnology, Jiangnan University, Wuxi, China, <sup>3</sup> Department of Oncology, Affiliated Hospital of Jiangnan University, Wuxi, China

## OPEN ACCESS

### Edited by:

Aleix Prat,  
Hospital Clínic de Barcelona,  
Spain

### Reviewed by:

Jie Tan,  
Johns Hopkins Medicine,  
United States  
Eva Hernando,  
New York University,  
United States  
Tomás Pascual Martínez,  
Hospital Clínic de Barcelona,  
Spain

### \*Correspondence:

Xiaofeng Dai  
xiaofeng.dai@jiangnan.edu.cn

### Specialty section:

This article was submitted to  
Cancer Genetics,  
a section of the journal  
Frontiers in Genetics

**Received:** 21 March 2019

**Accepted:** 26 July 2019

**Published:** 04 October 2019

### Citation:

Dai X, Mei Y, Chen X and Cai D  
(2019) ANLN and KDR Are Jointly  
Prognostic of Breast Cancer Survival  
and Can Be  
Modulated for Triple Negative  
Breast Cancer Control.  
Front. Genet. 10:790.  
doi: 10.3389/fgene.2019.00790

**Purpose:** Kinase insert domain receptor (KDR) is the primary vascular endothelial growth factor receptor mediating survival, growth, and migration of endothelial cells and is expressed also in various tumor cells through autocrine production. The PI3K/Pten pathway is one of the downstream signalings affected by KDR activation and most commonly altered in breast cancer. Here, we investigate whether KDR expression is associated with members in PI3K/Pten signaling on the prognosis of breast cancer patients.

**Methods:** PI3K/Pten pathway components were defined by mapping The Cancer Genome Atlas (TCGA) protein data to the KEGG database complemented by literature searching, accounting for 36 proteins subject to the interaction analysis with KDR on breast cancer patient survival. The identified interaction gene pair was subjected to *in vitro* validation following functional analysis.

**Results:** Anillin (ANLN) was found to interact with KDR at translational and transcriptional levels using the public TCGA protein expression data and five gene expression datasets. Favorable prognosis corresponds to high protein but low gene expression of ANLN when KDR is highly expressed. Externally modulating cells toward low ANLN and high KDR gene expression was shown to transit triple negative cells toward a luminal-like state with increased level of ER and elevated sensitivity to Tamoxifen.

**Conclusion:** Our study proposes a two-gene panel prognostic of breast cancer survival and a novel therapeutic strategy for triple negative breast cancer control via transiting cancer cells towards a luminal-like state sensitive to established targeted therapy.

**Keywords:** ANLN, KDR, interaction, state transition, subtype, survival

## INTRODUCTION

Vascular endothelial growth factor receptors (VEGFRs) are receptor tyrosine kinases mediating the survival, growth, and migration of endothelial cells through paracrine signaling (Deng et al., 2018). The downstream effects of VEGFR activation are mediated by a number of signaling cascades such as the mitogen-activated protein kinase and the PI3K/Pten pathways, where PI3K/Pten is frequently altered in breast cancers (Li et al., 2017). The intimate connections and regulatory relationships between VEGFR and PI3K/Pten signaling in tumors motivate us to investigate the joint prognostic

value of VEGFR and components involved in the PI3K/Pten pathway on breast cancer clinical outcome. We conducted pair-wise interaction survival analysis between kinase insert domain receptor (KDR) [also named VEGFR2 and is the primary VEGFR (Takahashi and Shibuya, 2005)] and PI3K/Pten players at both transcriptional and translational levels using data retrieved from The Cancer Genome Atlas (TCGA), European Genome-Phenome Archive (METABRIC) (Curtis et al., 2012), and Gene Expression Omnibus database (Edgar et al., 2002), followed by a series of experimental validations. We demonstrate that low *ANLN* and high *KDR* gene expression is associated with favorable breast cancer outcome; externally forcing cancer cells to exhibit such a profile could transit cells from the triple negative to luminal-like phenotype and sensitize cells to Tamoxifen (Kumar et al., 2018) treatment due to possibly upregulated ER expression. Our results contribute in identifying a two-gene panel prognostic of breast cancer clinical outcome and propose a combined therapeutic strategy for triple negative breast cancer control.

## MATERIALS AND METHODS

### Data

Data used in this study are summarized in **Supplementary Table 1**.

### Protein Expression Data

The level 2 primary breast tumor reverse-phase protein microarrays data were retrieved from TCGA (<http://cancergenome.nih.gov>), which contains 385 samples. Super curve log2 values were linearized, median centered by the median across all samples, and normalized by the median across the entire panel of antibodies following the protocol (<https://www.mdanderson.org/research/research-resources/core-facilities/functional-proteomics-rppa-core/faq.html>).

### Gene Expression Data

The level 3 primary breast tumor mRNA expression data were retrieved from TCGA, which includes 514 samples and 65 breast cancer death events. The mRNA data were produced using Agilent 244K Custom Gene Expression G4502A-07-3 platform, locally weighted scatterplot smoothing normalized followed by log2 transformation of the ratio between two channels.

The mRNA expression data from METABRIC (Curtis et al., 2012) were retrieved with permission, which include 1,293 samples and 295 breast cancer death events. The mRNA data were produced using Affymetrix SNP 6.0 and normalized using the quantile-based approach.

Three public datasets from GEO (Edgar et al., 2002), i.e., GSE6532 (Loi et al., 2007) and GSE22220 (Buffa et al., 2011), and GSE24450 (Muranen et al., 2011) were retrieved. GSE6532, including 87 samples (with 28 relapsed cases), was produced using Affymetrix Human Genome U133 Plus 2.0 Array and quantile normalized in robust multiarray analysis (Bolstad et al., 2003). GSE22220 was composed of 216 samples (including 82 distant relapsed events), produced using Illumina HumanRefSeq-8\_V1 expression BeadChips, and normalized using the quantile-based approach. GSE24450 contains 183 primary breast tumors (39 breast

cancers died of breast cancer or having distant metastasis), produced using Illumina HumanHT-12\_V3 Expression BeadChips, and quantile normalized.

### Histopathological Data

The histopathological data were retrieved from TCGA, which contains information on ER, PR, HER2, tumor size, nodal metastasis, and the tumor, node, and metastasis (TNM) stage (**Table 1**).

## Computational Methods

### Expression Interaction Survival Analysis

The primary players of the PI3K/Pten pathway were defined using Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa and Goto, 2000) supplemented by relevant literatures (Suzuki et al., 2005; Brouxhon et al., 2013; Quann et al., 2013; Thuma and Zoller, 2013). We first conducted survival analysis on pair-wise interactions at the translational level. In total, there were 142 antibodies available in TCGA, representing 114 unique proteins, among which 31 were involved in the PI3K/Pten pathway. These 31 genes plus 5 reported players of the PI3K/Pten pathway (Suzuki et al., 2005; Brouxhon et al., 2013; Quann et al., 2013) constitute the gene panel used in the interaction analysis (**Supplementary Table 2, Supplementary Figure 1**). Significant interactions at the translational level were selected for analysis at the transcriptional level following the same analytical procedure.

While TCGA data were used at the translational level, five datasets (TCGA, METABRIC, GSE6532, GSE22220, and GSE24450) were used at the transcriptional level. Anillin (ANLN) and KDR expressions were split into high and low levels at the splitting point optimized by grid searching (Barto, 1985). Binarized data were fitted into a Cox regression model, which include both the effect of each component and the interaction. In addition, a model without the interaction term was built for each pair. The *p* value from the chi-square test of the likelihood ratio between the model including the interaction term and the one without was used to assess the significance of the interaction. Kaplan–Meir plots were drawn to visualize the interactive effect.

Meta-analysis was applied in the analysis at the transcriptional level using the “metagen” function from the “meta” R package to assess the combined effect of the five datasets. The meta *p* value from the Fisher method (Fisher, 1932) was used to assess the significance of the interaction term. Stratified analysis, i.e., the survival was analyzed for one gene as stratified by the expression of the other, was conducted at both the protein and gene expression levels using the same statistical assessment methods.

Different death events were available in different datasets, i.e., 15-year breast cancer specific death in METABRIC, 10-year overall survival in TCGA data, 15-year relapse free survival in GSE6532, 10-year relapse free survival in GSE22220 data, and 10-year breast cancer specific death in GSE24450.

### Histopathological Association Analysis

Samples were binarized into high and low expression of ANLN and KDR. The associations between tumors with different protein expressions of ANLN and KDR, and histopathological markers including ER, PR, HER2, T, N, TNM stage, and subtype



**TABLE 1 |** Associations of the interaction between ANLN and KDR with histopathological parameters. The expression level, “high” or “low,” refers to that of ANLN and KDR, respectively, in the represented order. “ER,” “PR,” and “HER2” are cell receptors canonically used for breast cancer subtyping, “T” represents the size of the original tumor and whether it has invaded nearby tissue, “N” describes the nearby lymph nodes involved, “TNM stage” is an international standard for classifying the extent of spread of cancer based on “T,” “N,” and “M” (“M” describes distant metastasis). “Subtype” refers to PAM50 molecular subtyping, and ER-PR-HER2 histochemistry staining system was used to assess the subtyping status if PAM50 subtyping was not available; “LumAorB” means that PAM50 is “NA,” ER or PR is positive, HER2 is negative; “TNG” is short for triple negative group. Patients were analyzed by ANLN and KDR protein expression, with the number and percentage of patients in each category being summarized as “No.” and “(%).” Chi-squared test and 1,000 permutations of Monte Carlo simulations were conducted to assess the significance of associations of the two-gene interaction with each histopathological parameter.

ANLN:KDR	All		High:High		High:Low		Low:High		Low:Low		Chi-square	Monte Carlo
	No.	(%)	No.	(%)	No.	(%)	No.	(%)	No.	(%)	p	p
<b>ER</b>												
–	95	25.13%	28	28.57%	41	46.07%	7	6.09%	19	25.00%	1.91E–09	1.00E–04
+	283	74.87%	70	71.43%	48	53.93%	108	93.91%	57	75.00%		
<b>PR</b>												
–	148	39.05%	44	44.90%	48	53.93%	21	18.26%	35	45.45%	5.49E–07	1.00E–04
+	231	60.95%	54	55.10%	41	46.07%	94	81.74%	42	54.55%		
<b>HER2</b>												
–	203	75.46%	55	82.09%	44	72.13%	71	83.53%	33	58.93%	4.20E–03	5.20E–03
+	66	24.54%	12	17.91%	17	27.87%	14	16.47%	23	41.07%		
<b>T</b>												
1	90	23.50%	21	21.21%	16	18.18%	34	29.31%	19	23.75%	0.2776657	0.288471153
2+	293	76.50%	78	78.79%	72	81.82%	82	70.69%	61	76.25%		
<b>N</b>												
0	178	47.21%	53	54.08%	39	44.83%	53	47.32%	33	41.25%	0.3624281	0.360863914
1+	199	52.79%	45	45.92%	48	55.17%	59	52.68%	47	58.75%		
<b>TNM Stage</b>												
1	77	20.48%	17	17.35%	12	13.95%	32	28.07%	16	20.51%	0.0365787	0.0359964
2	213	56.65%	65	66.33%	46	53.49%	58	50.88%	44	56.41%		
3	86	22.87%	16	16.33%	28	32.56%	24	21.05%	18	23.08%		
<b>Subtype</b>												
Basal	28	7.37%	11	11.11%	10	11.36%	2	1.74%	5	6.41%	2.63E–07	1.00E–04
Her2	23	6.05%	4	4.04%	11	12.50%	1	0.87%	7	8.97%		
LumA	195	51.32%	50	50.51%	36	40.91%	77	66.96%	32	41.03%		
LumAorB	22	5.79%	10	10.10%	2	2.27%	6	5.22%	4	5.13%		
LumB	76	20.00%	14	14.14%	12	13.64%	27	23.48%	23	29.49%		
TNG	36	9.47%	10	10.10%	17	19.32%	2	1.74%	7	8.97%		

classification were analyzed separately. The statistical significance was assessed by chi-square test and Monte Carlo simulation on 10,000 permutations in R.

## Experimental Materials

### Cell Culture

One human normal mammary epithelial cell line (MCF10A), one luminal cell line (MCF7), and two triple negative breast cancer cell lines (MDAMB231 and SUM159PT) were included in the experiment. Cells were bought from the American Type Culture Collection, with mycoplasma tested and verified by sequencing.

MCF10A cells were cultured in Dulbecco’s modified eagle medium (DMEM)/F12 (Gibco) supplemented with 5% charcoal-stripped horse serum (Gibco), 10 µg/ml insulin (PeproTech), 20 ng/ml epithelia growth factor (PeproTech), and  $1.4 \times 10^{-6}$  mol/l hydrocortisone (PeproTech). MCF7 and MDAMB231 cells were cultured in DMEM supplemented with 10% fetal bovine serum (Gibco). SUM159PT cells were cultured in F12 (Gibco) supplemented with 5% fetal bovine serum (Gibco), 20 µg/ml insulin (PeproTech), 1% HEPES (PeproTech),  $2.8 \times 10^{-6}$  mol/l

hydrocortisone (PeproTech). Assay ready cells were prepared by culturing cells in a large batch and aliquoting them into ampules that were kept in liquid nitrogen in solution containing 90% fetal bovine serum and 10% dimethyl sulfoxide. Immediately prior to transfection, cells were thawed and washed with culture medium, and cell number was counted using a hemocytometer (Thermo).

## Experimental Protocols

### Cell Transfection

$1 \times 10^6$  cells per well were added in 2 ml of culture medium and transferred to black clear bottom tissue-culture treated six-well plates (Nalgene #167018). Cells were incubated overnight and achieved 70–80% confluence before transfection. Medium was replaced by 2 ml serum-free medium before transfection. One hundred microliter Optimem medium (Gibco) containing 1 µg sgRNA plasmids (sgRNAs were listed in **Supplementary Table 3**) and 1 µg dCas9-synergistic activation mediator (SAM) plasmids were added to 100 µl Optimem medium containing 6 µl lipo2000 transfection reagent per well and mixed for 15–20 min prior to transfection. The mixture was transferred to a six-well plate and



incubated at 37°C for 5–8 h in the presence of 5% CO<sub>2</sub> (HERA Cell 150i, Thermo Scientific). Serum-free medium was replaced by 2 ml medium containing 10% serum. Cells were incubated at 37°C for 24 h and then subjected to stable clone selection under 4- $\mu$ l 200 mg/ml G418 and 5- $\mu$ l 0.1 mg/ml puromycin pressure for 2 months.

### qPCR Assay

After transfection, cells were collected and extracted for total RNA using TRIzol reagent (TianGen) at 3 days after transfection. The cDNA was synthesized using PrimeScript RT reverse transcriptase (Takara). Primers for quantitative reverse transcription PCR (qRT-PCR) are listed in **Supplementary Table 4**. The absorbance value was recorded at the extension stage. The relative expression level was calculated using the  $2^{-\Delta\Delta C_t}$  methods. All qRT-PCR experiments were performed using ABI Step one plus Real-Time PCR System (ABI) following Takara protocol.

### Proliferation Assay

Eight thousand cells per well were added in 100  $\mu$ l of culture medium and transferred to black clear bottom tissue-culture treated 96-well plates (Nalgene #167008). Cells were incubated overnight and achieved 70–80% confluence before transfection, cells transfection as described above. For cell proliferation measurement, 10  $\mu$ l per well of CKK-8 (Dojindo) was added, and absorbance was detected using EZ Read 800 microplate Reader (Biochrom) after cell incubation at 37°C for 2 h.

### Invasion Assay

After transfection, cells were incubated until they form confluent monolayers. Wounds were made using a pipette tip, and photographs were taken immediately (0 h), 12, 24, and 36 h after wounding. Distance change between the two edges of wounded area due to cell migration was measured and computed at each time point. Results were presented as the migration rate.

Student's *t* test was computed using R to evaluate the statistical significance on cell migration, and *p* values were computed as the two-tailed probability at 95% confidence from a standard normal distribution.

### Flow Cytometry Assay

The proportion of cancer stem cell was assessed by FACSCalibur flow cytometer (BD). Cultured cells were washed twice with phosphate-buffered saline (PBS) and then harvested using trypsin. Detached cells were washed once in PBS and stained using ALDEFLUOR™ kit (STEMCELL Technologies) at the room temperature (RT) in the darkness for 30 min. Labeled cells were washed and fixed in PBS and analyzed using flow cytometer.

### Western Blot Assay

Cultured cells were washed twice using ice-cold PBS and lysed in radioimmunoprecipitation assay lysis buffer supplemented with protease inhibitors for 5 min on ice and centrifuged at 12,000g for 10 min before supernatants collection. The protein concentration was estimated using the BCA Protein Assay Kit (Tiangen). Proteins (50  $\mu$ g) per lane were resolved by sodium dodecyl sulfate polyacrylamide gel electrophoresis and transferred to polyvinylidene fluoride membrane. After blocking with 5%

nonfat dried milk powder in Tris-buffered saline plus Tween-20 buffer, the membrane was incubated using the appropriate primary Abs (Proteintech) at 4°C overnight followed by secondary Abs (Proteintech) for 2 h at RT. Ab binding was visualized by developing the blot using enhanced chemiluminescence reagent. The bands were visualized using OmegaLumG (UVP) followed by analysis using the Image J software. Western blot was performed after 72 h of construct transfection.

### Drug Response Assay

MCF10A, MCF7, MDAMB231, and AdKu (ANLN downregulation and KDR upregulation) cells were used in the experiment. Eight Tamoxifen concentrations (1, 10, 25, 100, 250, 1,000, 2,500, and 10,000 nM) with six replicates were designed. Also included in each plate were the negative control and drug-free negative control at each drug concentration with six replicates. Tamoxifen (Sigma) was added to cells after they form confluent monolayers. Ten microliters per well of CKK-8 was added 48 h after adding Tamoxifen, and absorbance was detected using an EZ Read 800 microplate reader after cell incubation at 37°C for 2 h. The dose–response curve of Tamoxifen treatment and IC50 values were obtained for each siRNA in each cell line using the “drc” package in R, where a four-parameter log logistic model (LL.4) was used for data fitting. Statistical significance on IC50 alteration was evaluated by Student's *t* test using R.

### Mouse *In Vivo* Study

$1 \times 10^6$  MDAMB231 and AdKu-231 cells suspended in PBS were injected subcutaneously to six female BALB/c mice aged 4–6 weeks with the average weight of  $20 \pm 5$  g, respectively. Mice were divided into two groups, i.e., MDAMB231 group, AdKu-231 group, depending on the tumor cells subcutaneously injected, and each group included four mice by design. Tumor volume was calculated using Equation (1)

$$V = \frac{\pi \times L \times W^2}{6} \quad (1)$$

where “V,” “L,” and “W” each represents volume, the largest diameter, and smallest diameter of the tumor, respectively.

Tumor growth measuring started when tumor lesion appeared and recorded every 3 days. Mice were killed at the 24th day after the initial appearance of tumor lesions.

## RESULTS

### Opposite Interactions Between ANLN and KDR at Translational and Transcriptional Levels

Among the 36 proteins being analyzed (**Supplementary Table 2**, **Supplementary Figure 1**), anillin (encoded by ANLN) was found to interact with KDR (also named VEGFR2), and such an interaction affected breast cancer survival with statistical significance at both the translational (**Supplementary Figure 2**, 51 and 44% were optimized for ANLN and KDR binarization, respectively) and transcriptional (**Supplementary Figure 3**, 51 and 32%

were optimized for *ANLN* and *KDR* binarization, respectively) levels. Interactions between *ANLN* and *KDR* were confirmed by fitting the Cox regression model, where the fitness significantly improved when the interaction term was included at both translational ( $p = 0.006$ ) and transcriptional (meta-analysis from five public datasets  $p = 0.024$ ) levels (Table 2). No significant univariate clinical association was observed at the translational level for neither protein (Supplementary Figure 2). At the transcriptional level, *ANLN* had an independent main effect that was exemplified by *KDR* overexpression (Supplementary Figure 3), i.e., Fisher meta-analysis  $p$  value for *ANLN* was  $8.07\text{e-}11$  and became  $3.59\text{e-}11$  when *KDR* expression was high in the stratified analysis.

Interestingly, concomitant low *ANLN* and high *KDR* protein expression was associated with poor clinical outcome ( $\text{HR} = 3.16$ ) but conveyed protective effect ( $\text{HR} < 1$  for four out of five datasets) at the transcriptional level (Table 2, Figure 1, Supplementary Figures 2 and 3). In other words, low *ANLN* and high *KDR* gene expression shared the same clinical association with concomitant overexpression of both proteins, which was associated with favorable clinical outcome; concomitant high levels of both *ANLN* and *KDR* expression shared the same clinical outcome with patients having low *ANLN* and high *KDR* expression, which was associated with poor clinical outcome (Figure 1, Table 2).

We constructed two cell lines, namely, AdKu-231 and AdKu-159, with low *ANLN* and high *KDR* gene expression (Figure 2A). *ANLN* expression was significantly reduced ( $p = 0.008$  for AdKu-231,  $p = 0.002$  for AdKu-159) and that of *KDR* was significantly upregulated ( $p = 0.004$  for AdKu-231,  $p = 0.005$  for AdKu-159) in AdKu cells (Figure 2A). Western blotting showed concomitant overexpression of both proteins in both AdKu cells (Figure 2C). These results suggest that the observed opposite clinical associations at the translational and

transcriptional levels lie in the reverse expression of *ANLN* at both gene and protein expression levels.

## Low *ANLN* and High *KDR* Gene Expression Is Associated With Less Malignant Breast Cancer Cell Features

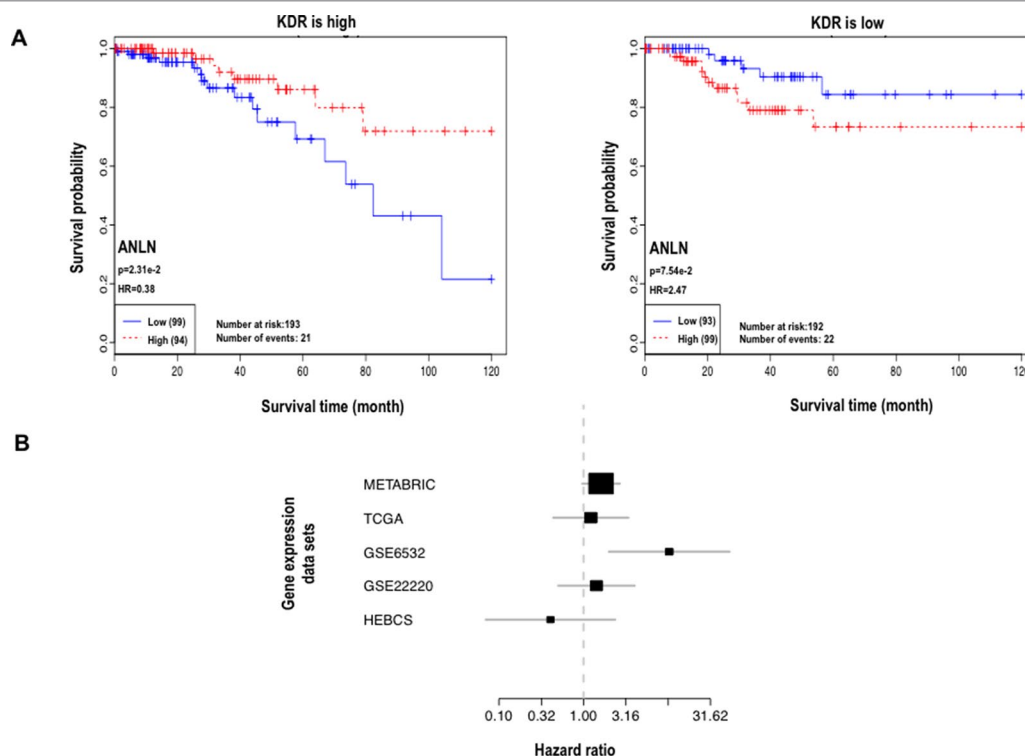
*KDR* and *ANLN* were positively correlated at the transcriptional level when *ANLN* gene expression was perturbed in triple negative breast cancer cell lines SUM159PT and MDAMB231 (Figures 3A, B). In brief, *KDR* gene expression was significantly reduced ( $p = 5.54\text{e-}4$  in SUM159PT,  $p = 0.010$  in MDAMB231) once *ANLN* was effectively downregulated ( $p$  values were 0.001 and  $3.53\text{e-}4$ , respectively, in SUM159PT and MDAMB231). When *ANLN* was sufficiently overexpressed ( $p$  values for upregulating *ANLN* were  $2.00\text{e-}4$  and  $3.81\text{e-}4$  in SUM159PT and MDAMB231, respectively), *KDR* expression increased with statistical significance ( $p = 5.12\text{e-}4$  in SUM159PT,  $p = 0.002$  in MDAMB231). Similarly, the expression of both genes was positively correlated when *KDR* was modulated in triple negative breast cancer cells (Figures 3C, D). That is, *ANLN* expression was significantly altered in the consistent direction with *KDR* ( $p = 0.010$  for downregulation in SUM159PT,  $p = 2.41\text{e-}4$  for downregulation in MDAMB231,  $p = 4.36\text{e-}5$  for upregulation in SUM159PT,  $p = 4.72\text{e-}5$  for upregulation in MDAMB231) when *KDR* expression was effectively down- and upregulated ( $p = 0.004$  and  $p = 7.60\text{e-}4$  for downregulation in SUM159PT and MDAMB231, respectively;  $p = 9.41\text{e-}4$  and  $p = 9.14\text{e-}5$  for upregulation in SUM159PT and MDAMB231, respectively).

We did not observe any significant alteration on *KDR* gene expression when modulating that of *ANLN* in the luminal breast cancer cell line MCF7 and normal breast cell

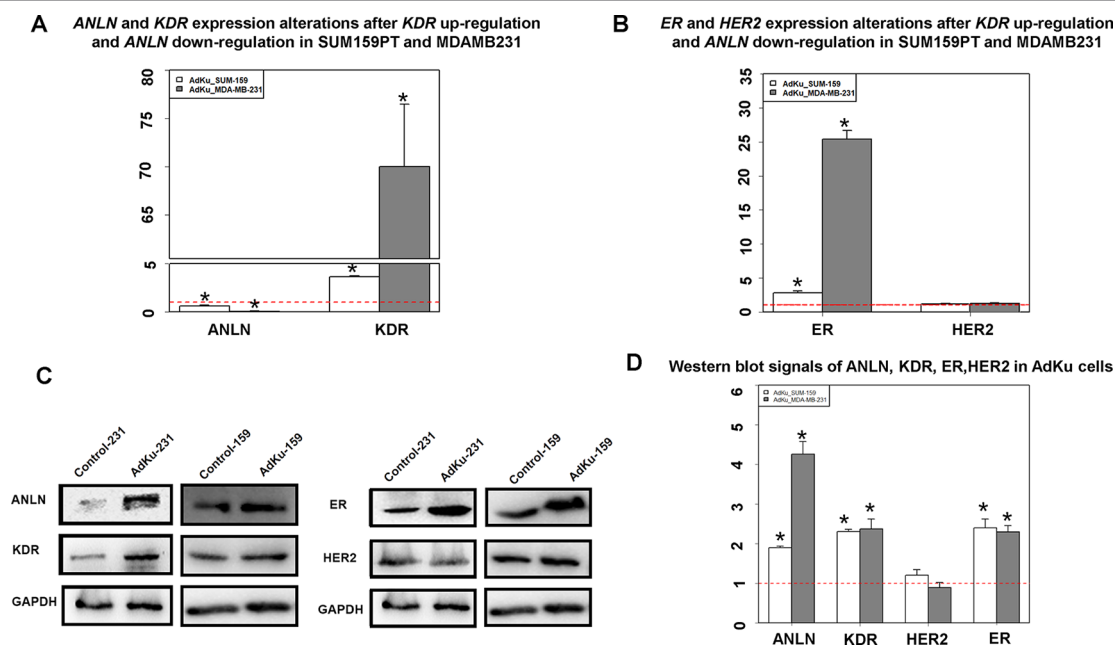
**TABLE 2 |** Statistics of the model including the interactions between *ANLN* and *KDR* at the expression levels. “GEX” and “PEX” each represents the gene expression and protein expression, respectively. The expression level, “high” or “low,” each refers to that of *ANLN* and *KDR*, respectively, in the presented order. The 51 and 44% were (optimized using TCGA PEX data) used as the splitting point for binarizing *ANLN* and *KDR* PEX data, respectively; and 51 and 32% (optimized using METABRIC GEX data) were used as the splitting points for GEX data binarization, accordingly. “HR” and “95%CI” are the hazard ratio and 95% confidence interval ([low, high]) for each pair, respectively. The  $p$  value for the interaction term ( $p_{\text{inter}}$ ) comes from the chi-square test, which shows the significance of the improvement of the model including the interaction term as compared with the model without interactions. “Meta-analysis” is conducted for GEX data, the meta-analysis  $p$  value (fixed-effects model given that no heterogeneity was detected) for each genotype combination is obtained using “metagen” from R package “meta,” and the meta-analysis for the interaction term is obtained using the Fisher’s method from the  $p$  values ( $p_{\text{inter}}$ ).

Analysis	Data	ANLN:KDR	High:Low	Low:High	High:High	$p_{\text{inter}}$
PEX	TCGA	HR	2.5	<b>3.16</b>	0.16	0.0061
		95%CI	[0.88,7.10]	<b>[1.16,8.64]</b>	[0.04,0.63]	
GEX	METABRIC	HR	1.33	0.88	1.61	0.0692
		95%CI	[0.87,2.04]	[0.60,1.31]	[0.96,2.68]	
	TCGA	HR	0.87	0.92	1.42	0.5105
		95%CI	[0.38,2.02]	[0.43,1.97]	[0.50,4.02]	
	GSE6532	HR	0.58	0.21	10.29	0.0044
		95%CI	[0.18,1.91]	[0.06,0.75]	[1.99,53.18]	
	GSE22220	HR	2.55	0.92	1.22	0.7063
		95%CI	[1.08,5.99]	[0.39,2.18]	[0.44,3.39]	
	GSE24450	HR	3.39	3.19	0.41	0.2951
		95%CI	[0.68,16.80]	[0.72,14.16]	[0.07,2.36]	
Meta	GEX datasets	$p$	0.0779	0.365	<b>0.0244</b>	0.0235
		Method	Metagen	<b>Metagen</b>	Metagen	

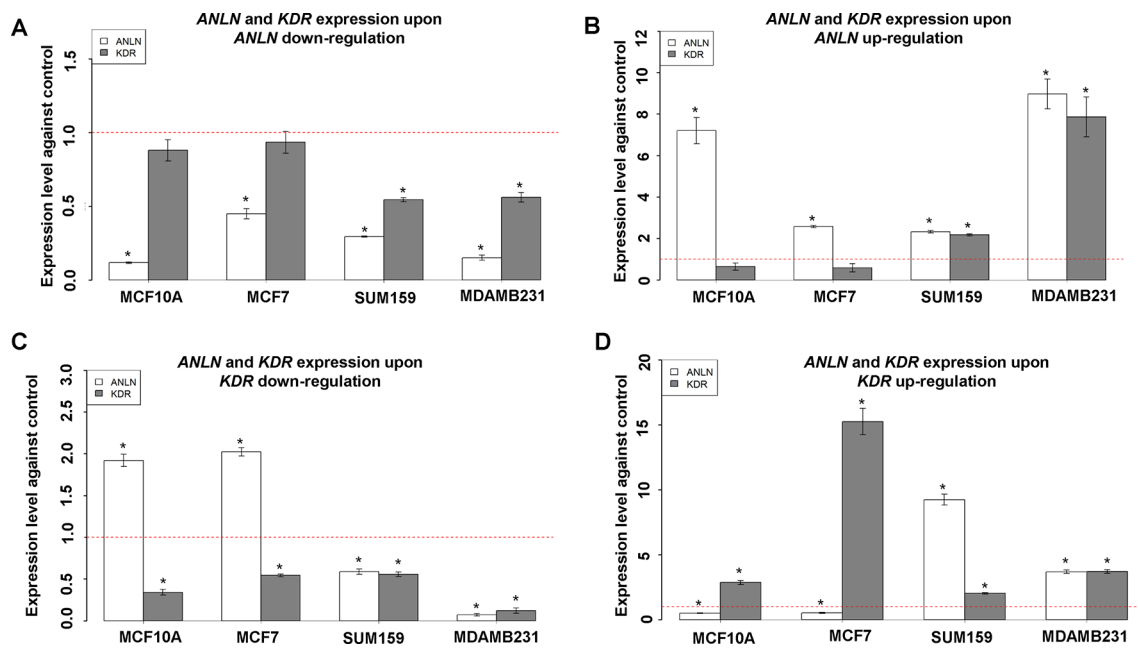
Significance of Bold Values that conveys risky effect is Low:High at the PEX level and High:High at the GEX level.



**FIGURE 1 |** Analysis of interactions between ANLN and KDR. **(A)** Kaplan–Meier plot at the protein expression level. **(B)** Forest plot at the gene expression level. In the forest plot, each line represents the confidence interval of a study, where a longer line represents a smaller dataset; each black box represents a point estimation, where a larger area represents a higher weight the dataset contributes to the meta-analysis.



**FIGURE 2 |** Expression of KDR, ANLN, ER, and HER2 in AdKu cells derived from triple negative breast cancer cells. **(A)** Expression of *KDR* and *ANLN* at the transcriptional level. **(B)** Expression of *ER* and *HER2* at the transcriptional level. **(C)** Expression of ANLN, KDR, ER, and HER2 at the translational level. **(D)** Western blot signaling intensities normalized by that of GAPDH for ANLN, KDR, ER, and HER2 in AdKu cells. \* represents statistical significance ( $p < 0.05$ ) Student's *t* test. The red dotted line represents the expression level where no external modulation was done. MDAMB231 and SUM159PT cells were used to derive AdKu cells. The red dotted line represents the expression level where no external modulation was done.



**FIGURE 3 |** Interactions between *ANLN* and *KDR* in different breast cancer cells. **(A)** *KDR* gene expression after downregulating *ANLN* in each cell line. **(B)** *KDR* gene expression after upregulating *ANLN* in each cell line. **(C)** *ANLN* gene expression after downregulating *KDR* in each cell line. **(D)** *ANLN* gene expression after upregulating *ANLN* in each cell line. Bars represent mean  $\pm$  SD from at least three independent experiments. \* represents statistical significance ( $p < 0.05$ ) from Student's  $t$  test. The red dotted line represents the expression level where no external modulation was done. SUM159 and MDAMB231 are triple negative breast cancer cells, MCF7 is a luminal breast cancer cell line, and MCF10A represents normal breast epithelial cells.

line MCF10A (Figures 3A, B). *ANLN* gene expression was significantly modulated both up- and downwards ( $p$  values for downregulation were  $2.49 \times 10^{-4}$  and  $1.81 \times 10^{-4}$  in MCF7 and MCF10A, for upregulation were  $1.33 \times 10^{-5}$  and  $6.14 \times 10^{-4}$  in MCF7 and MCF10A, respectively), and no significant alteration was observed for *KDR* gene expression. However, we observed significant mutual suppression between *ANLN* and *KDR* gene expression in the luminal cell line MCF7 and normal breast cells MCF10A (Figures 3C, D). That is, *ANLN* was significantly downregulated ( $p = 8.41 \times 10^{-4}$  in MCF7 and  $p = 0.002$  in MCF10A) when *KDR* was upward modulated ( $p = 1.53 \times 10^{-4}$  in MCF7,  $p = 3.97 \times 10^{-4}$  in MCF10A), and significantly upregulated ( $p = 5.82 \times 10^{-5}$  in MCF7 and  $p = 7.36 \times 10^{-4}$  in MCF10A) when *KDR* was downward modulated ( $p = 0.008$  in MCF7,  $p = 7.72 \times 10^{-4}$  in MCF10A).

### Modulated Cells With Low ANLN and High KDR Gene Expression Exhibit Less Malignant Cancer Features

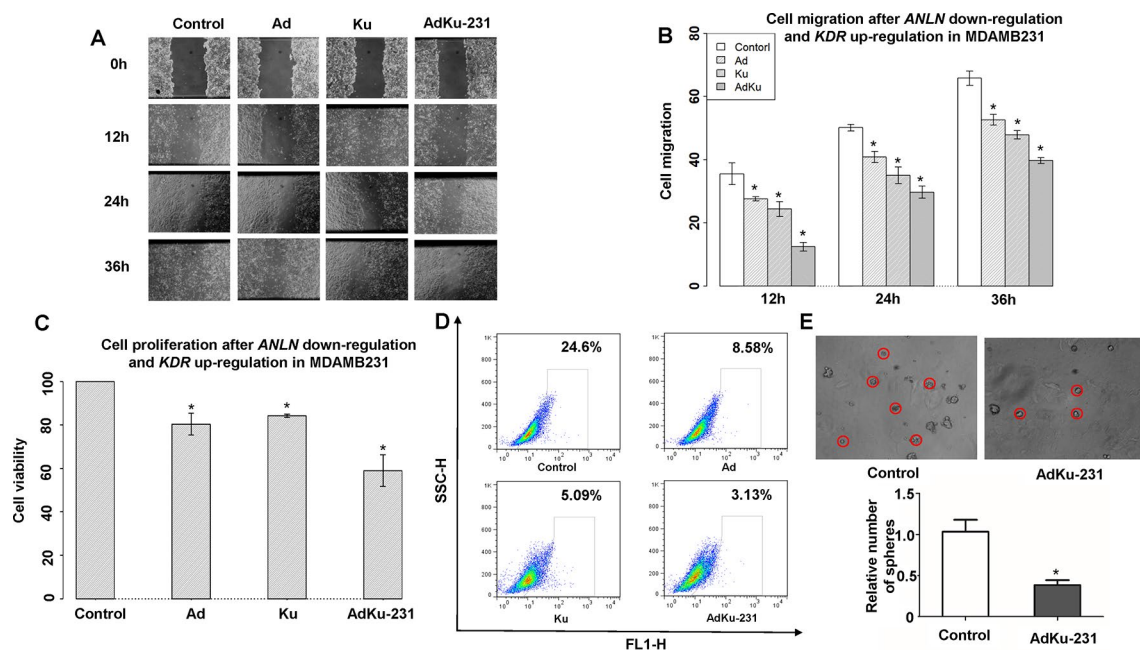
We constructed a stable cell line, AdKu-231, with reduced *ANLN* and increased *KDR* gene expression from the triple negative breast cancer cell line MDAMB231 using the Crispr technique (sgRNAs are listed in Supplementary Table 3). *ANLN* and *KDR* were effectively modulated ( $p = 0.008$  for knocking down *ANLN* and  $p = 0.004$  for upregulating *KDR*, Figure 2A). The migration of AdKu-231 cells was significantly recessed as measured at 12 ( $p = 4.28 \times 10^{-4}$ ,  $6.40 \times 10^{-5}$ ,  $0.0017$  as compared with MDAMB231, Ad, Ku), 24 ( $p = 8.71 \times 10^{-5}$ ,  $0.002$ ,

$0.046$  as compared with MDAMB231, Ad, Ku), and 36 ( $p = 5.13 \times 10^{-5}$ ,  $3.36 \times 10^{-4}$ ,  $0.001$  as compared with MDAMB231, Ad, Ku) hours (Figures 4A, B). The growth of AdKu-231 cells was significantly reduced as compared with MDAMB231 ( $p = 1.91 \times 10^{-5}$ ), Ad ( $p = 8.99 \times 10^{-5}$ ), and Ku ( $p = 2.80 \times 10^{-4}$ ) cells (Figure 4C). The percentage of cancer stem cells was considerably reduced from 24.6% in MDAMB231 to 8.58% in Ad cells, to 5.09% in Ku cells, and to 3.13% in AdKu-231 cells (Figure 4D), and the relative number of spheres was reduced to 38% in AdKu-231 cells as compared with the control ( $p = 0.009$ , Figure 4E).

ER expression was significantly elevated in AdKu-231 cells, with  $p = 0.006$  and  $p = 0.007$ , respectively, at the transcriptional and translational levels as compared with MDAMB231 cells (Figures 2B, D). Similar expression profiles were observed in AdKu-159 cells (Figures 2B, D). Histopathological association analysis revealed that ER status was significantly affected by the protein expression of *ANLN* and *KDR*, with the  $p$  value from chi-square test being  $1.91 \times 10^{-9}$  and the  $p$  value from 1,000 permutations of Monte Carlo simulation being  $1 \times 10^{-4}$ . All three primary cell surface receptors used for breast cancer subtyping (ER, PR, and HER2) were significantly associated with *ANLN* and *KDR* expression (Table 1), suggesting that the synergistic effect of *ANLN* and *KDR* can affect cells' transition from the triple negative to the luminal-like phenotype.

AdKu-231 cells show increased sensitivity to Tamoxifen, a commercialized drug-targeting ER-positive tumors. IC<sub>50</sub> of AdKu-231 cells ( $29.75 \mu\text{M}$ ) dropped to two-thirds of that of MDAMB231 ( $48.19 \mu\text{M}$ ) and was close to that of MCF7





**FIGURE 4 |** Cell morphological alterations in the stable AdKu cells derived from MDAMB231. Comparisons on the (A) images and (B) measured areas of cell migration, (C) cell proliferation, (D) stem cell percentage, and (E) cell self-renew ability. “Ad” represents the stable cell line with reduced ANLN gene expression, “Ku” represents the stable cell line with increased KDR gene expression, and “AdKu-231” means both are regulated. \* represents statistical significance ( $p < 0.05$ ) from Student's  $t$  test.

**TABLE 3 |** IC50 of each cell line in response to Tamoxifen, Doxorubicin, or their combination. “IC50-STD” represents the standard deviation of IC50. “AdKu” represents the stable cell line we established with reduced ANLN and increased KDR gene expression.

Cell line	Drug	IC50	IC50-STD
MCF10A	Tamoxifen	16.245	0.3269
MCF7	Tamoxifen	25.4251	1.2181
AdKu	Tamoxifen	29.7514	1.4571
MDAMB231	Tamoxifen	48.1855	2.393
MCF10A	Tamoxifen + 10 nm Doxorubicin	12.9009	0.3451
MCF7	Tamoxifen + 10 nm Doxorubicin	36.3579	2.9832
AdKu	Tamoxifen + 10 nm Doxorubicin	39.0639	2.8264
MDAMB231	Tamoxifen + 10 nm Doxorubicin	42.4501	3.1901
MCF10A	Tamoxifen + 150 nm Doxorubicin	0.3773	0.5591
MCF7	Tamoxifen + 150 nm Doxorubicin	0.6219	1.0033
AdKu	Tamoxifen + 150 nm Doxorubicin	3.3216	2.9751
MDAMB231	Tamoxifen + 150 nm Doxorubicin	39.9377	3.2996
MCF10A	Doxorubicin	154.4694	10.7738
MCF7	Doxorubicin	184.7564	13.1385
AdKu	Doxorubicin	141.9265	20.4547
MDAMB231	Doxorubicin	242.127	20.6849

(25.43  $\mu$ M) (Table 3, Figure 5). We also tested the sensitivity of AdKu-231 cells in response to the synergistic effect of Tamoxifen and Doxorubicin as compared with MDAMB231, MCF7, and MCF10A (Figure 5). Combined use of Tamoxifen and Doxorubicin largely increased cells' sensitivities. While cancer cells share similar Tamoxifen IC50s which are distinctive from that of normal cells when Tamoxifen was combinatorially used with 10 nm Doxorubicin (lowest tested dose, Figure 5), AdKu-231 shares a similar Tamoxifen response

curve with MCF7 and MCF10A, which is distinct from that of MDAMB231 under IC50 dose of Doxorubicin (Figure 5).

*In vivo* study showed slower growth of AdKu-231 cells than MDAMB231 cells ( $p = 0.004$ , Figure 6), which is consistent from what we observed from *in vitro* experiments.

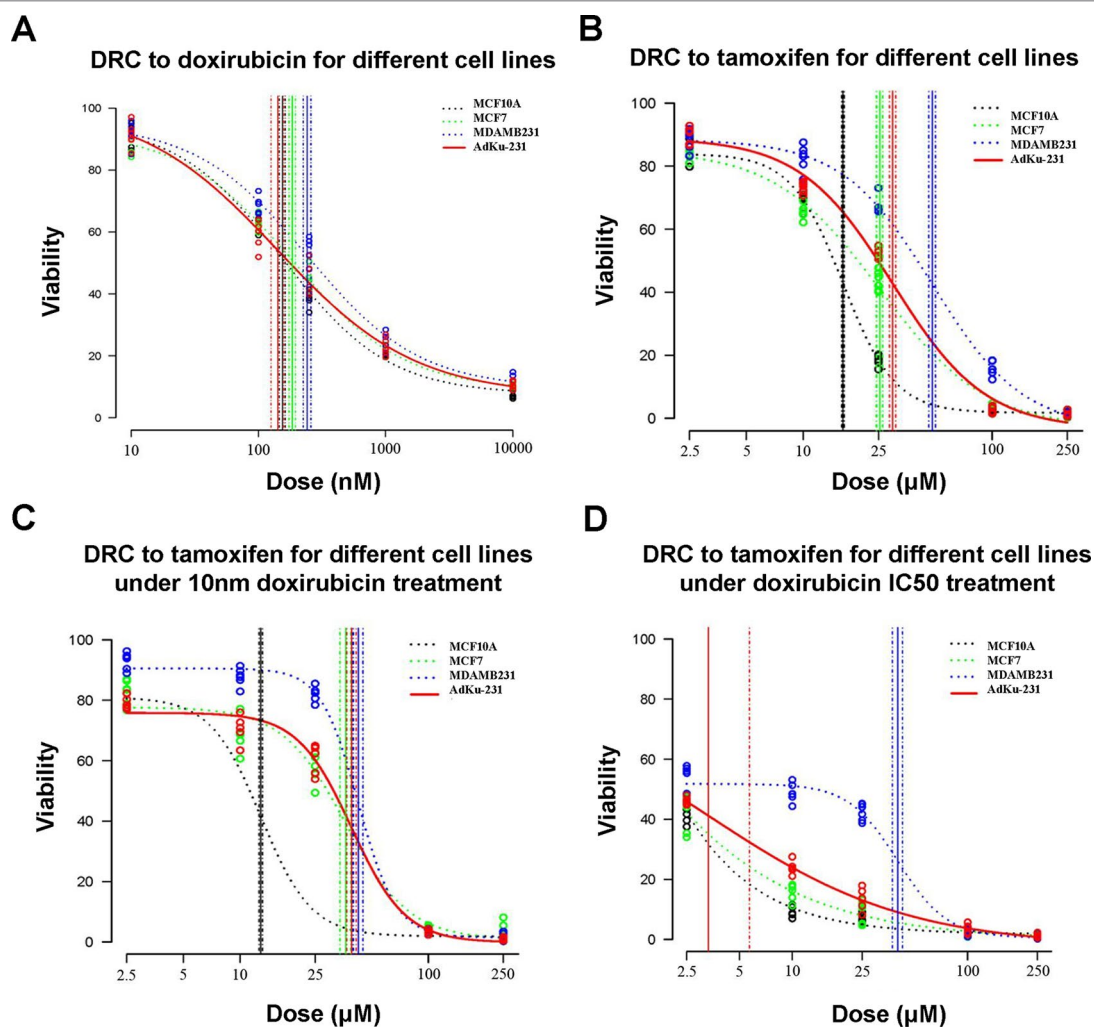
## DISCUSSION

Anillin (encoded by *ANLN*), a relatively poorly understood actin-binding protein involved in cytokinesis and the PI3K/Pten pathway (Suzuki et al., 2005), was found to interact with KDR at both transcriptional and translational levels with opposite clinical implications (Figure 1). That is, patients with low *ANLN* and high *KDR* gene expression shared similar favorable clinical outcomes with patients having concomitant high levels of both proteins. Such findings were validated by qPCR and Western blot (Figure 2).

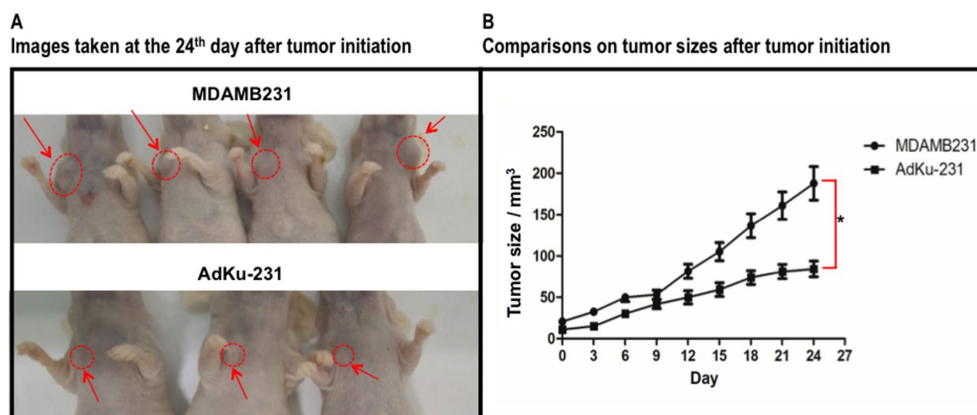
These inconsistent clinical associations were driven by *ANLN*, i.e., low *ANLN* expression at the transcriptional level corresponded to *ANLN* high expression at the translational level under *KDR* abundance (Figure 2). The  $p$  value and HR were  $1.72e-7$  and 0.54 for patients with *ANLN* low expression, which dropped to  $7.09e-8$  and 0.47, respectively, once *KDR* was upregulated in addition. This implicates that *ANLN* drove the main effect of this interaction and *KDR* has an amplification effect on *ANLN* functionalities in breast cancer.

*ANLN* mRNA abundance was associated with increased hazard of breast cancer death (Supplementary Figure 3).





**FIGURE 5 |** Comparison on cell viabilities in response to Tamoxifen, Doxorubicin, and combined use of Tamoxifen and Doxorubicin among different cell lines. Drug response curves under the treatment of (A) Tamoxifen, (B) Doxorubicin, (C) combined used of Tamoxifen and 10 nm Doxorubicin, and (D) combined use of Tamoxifen and IC50 Doxorubicin. AdKu-231 was used in this figure.



**FIGURE 6 |** Growth comparison between mouse tumors injected with MDAMB231 and AdKu-231 cells. (A) Images taken at the 24th day after tumor initiation. (B) Comparisons on tumor sizes after tumor initiation. One mouse injected with AdKu-231 died during the analysis and was dropped out from this study.

*ANLN* mRNA expression during tumor progression was measured in a diverse spectrum of tumors including breast cancers as well as normal tissues, which showed an increasing trend from the normal to the metastatic state (Wang et al., 2016). Knocking down *ANLN* could significantly decrease the invasiveness and growth of tumor cells (Calvo et al., 2013; Zhang et al., 2018). *ANLN* was recently proposed as a prognostic biomarker independent of KI-67 (known proliferation marker) and being essential for cell cycle progression in primary breast cancers (Magnusson et al., 2016). These converge to the favorable prognostic value of low *ANLN* mRNA expression among patients and are suggestive of the driving role of *ANLN* in the identified joint prognostic value.

The differential regulatory relationships between *ANLN* and *KDR* in different breast cancer cell lines and normal breast cells (Figure 3) suggest a potential network rewiring between more and less malignant states in breast cancer cells, which warrants validation at the transcriptional level. Low *ANLN* and high *KDR* gene expression is associated with a favorable clinical outcome, and low *ANLN* is naturally accompanied by decreased *KDR* in malignant tumor cells (Figure 3); by externally upregulating *KDR* and downregulating *ANLN* in triple negative cells MDAMB231, we established a cell line sharing similar phenotypical features with luminal breast cancer cells. Cell proliferation, migration, and cancer stem cell assays all suggest that AdKu cells are less malignant than MDAMB231. AdKu cells exhibit similar drug response curve with MCF7 cells under Tamoxifen (Kumar et al., 2018) treatment, suggesting that triple negative cells may be treated using the same strategy as luminal cells if *ANLN* was suppressed and *KDR* was upregulated at the transcriptional level. Indeed, ER, the target of Tamoxifen, was overexpressed on AdKu cells, explaining the demonstrated sensitivity of AdKu cells to Tamoxifen. Triple negative breast cancers are more malignant than the other subtypes and lack effective targeted therapeutic modalities. Triple negative cancers are conventionally treated by chemotherapy or radiotherapy, which are not selective on cancer cells and can considerably reduce the life quality of patients. Poly-ADP ribose polymerase inhibitors target BRCA1-deficient breast cancer cells which cannot represent triple negative breast cancers in general. Our results suggest a novel strategy for triple negative breast cancer control by concomitantly modulating *ANLN* and *KDR* gene expression while administering Tamoxifen to triple negative patients. That is, by transiting triple negative cancer cells to a less malignant state via concomitantly modulating *ANLN* and *KDR* gene expression, we could obtain desired clinical results using the same strategy as that for luminal cancers. Efforts devoted to cancer state transition, though few, do exist. It was reported that knocking down either *ERN1* or *ALPK1* could push bipotential breast tumor-initiating cells towards the luminal fate (Strietz et al., 2016). Different than that, we focus on the synergistic effects of two pathways (as represented by the identified two genes) on breast cancer state transition, both computationally and experimentally. Importantly, we show direct evidence of combined therapeutic efficacy of

the proposed approach, which suggests an emerging cancer therapeutic modality and has profound clinical implications.

## CONCLUSION

We report that concomitant low *ANLN* and high *KDR* gene expression is associated with favorable breast cancer survival. Externally modulating breast cancer cells towards low *ANLN* and high *KDR* gene expression can transit cells from the triple negative to luminal-like phenotype and sensitize cells to Tamoxifen treatment. This implicates a novel joint therapeutic approach combating against triple negative breast cancers.

## DATA AVAILABILITY

All datasets analyzed for this study are included in the manuscript and the **Supplementary files**.

## ETHICS STATEMENT

All animal experiments were performed in accordance with the laboratory animal guidelines and with the approval of the Animal Experimentations Ethics Committee, Jiangnan University.

## AUTHOR CONTRIBUTIONS

XFD designed, supervised and financed the project, and drafted the paper. XFD and XC conducted computational analysis. YM and DYK conducted the experimental validations. YM and XC prepared the figures. All authors have read and proved the content of the manuscript.

## FUNDING

This study was supported by the National Science and Technology Major Project of China (grant number: 2018ZX10302205-004-002), Natural Science Foundation of Jiangsu Province (grant number: BK20161130), the Six Talent Peaks Project in Jiangsu Province (grant number: SWYY-128), Postgraduate Education Reform Project of Jiangsu Province, and Research Funds for the Medical School of Jiangnan University ESI special cultivation project (grant number: 1286010241170320). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2019.00790/full#supplementary-material>

## REFERENCES

- Barto, A. G. (1985). Learning by statistical cooperation of self-interested neuron-like computing elements. *Hum. Neurobiol.* 4 (4), 229–256.
- Bolstad, B. M., Irizarry, R. A., Astrand, M., and Speed, T. P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 19, 185–193. doi: 10.1093/bioinformatics/19.2.185
- Brouxhon, S. M., Kyrkanides, S., Teng, X., Raja, V., O'banion, M. K., Clarke, R., et al. (2013). Monoclonal antibody against the ectodomain of E-cadherin (DECMA-1) suppresses breast carcinogenesis: involvement of the HER/PI3K/Akt/mTOR and IAP pathways. *Clin. Cancer Res.* 19 (12), 3234–3246. doi: 10.1158/1078-0432.CCR-12-2747
- Buffa, F. M., Camps, C., Winchester, L., Snell, C. E., Gee, H. E., Sheldon, H., et al. (2011). MicroRNA associated progression pathways and potential therapeutic targets identified by integrated mRNA and microRNA expression profiling in breast cancer. *Cancer Res.* 71, 5635–5645. doi: 10.1158/0008-5472.CAN-11-0489
- Calvo, F., Ege, N., Grande-Garcia, A., Hooper, S., Jenkins, R. P., Chaudhry, S. I., et al. (2013). Mechanotransduction and YAP-dependent matrix remodelling is required for the generation and maintenance of cancer-associated fibroblasts. *Nat. Cell Biol.* 15 (6), 637–646. doi: 10.1038/ncb2756
- Curtis, C., Shah, S. P., Chin, S. F., Turashvili, G., Rueda, O. M., Dunning, M. J., et al. (2012). The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* 486, 346–352. doi: 10.1038/nature10983
- Deng, Y., Yang, Y., Yao, B., Ma, L., Wu, Q., Yang, Z., et al. (2018). Paracrine signaling by VEGF-C promotes non-small cell lung cancer cell metastasis via recruitment of tumor-associated macrophages. *Exp. Cell Res.* 364, 208–216. doi: 10.1016/j.yexcr.2018.02.005
- Edgar, R., Domrachev, M., and Lash, A. E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 30, 207–210. doi: 10.1093/nar/30.1.207
- Fisher, R. A. (1932). *Statistical methods for research workers*. Edinburgh: Oliver and Boyd.
- Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30. doi: 10.1093/nar/28.1.27
- Kumar, B. N. P., Puvvada, N., Rajput, S., Sarkar, S., Mahto, M. K., Yallapu, M. M., et al. (2018). Targeting of EGFR, VEGFR2, and Akt by engineered dual drug encapsulated mesoporous silica-gold nanoclusters sensitizes Tamoxifen-resistant breast cancer. *Mol. Pharm.* 15, 2698–2713. doi: 10.1021/acs.molpharmaceut.8b00218
- Li, X., Zhang, R., Liu, Z., Li, S., and Xu, H. (2017). The genetic variants in the PTEN/PI3K/AKT pathway predict susceptibility and CE(A)F chemotherapy response to breast cancer and clinical outcomes. *Oncotarget* 8, 20252–20265. doi: 10.18632/oncotarget.15690
- Loi, S., Haibe-Kains, B., Desmedt, C., Lallemand, F., Tutt, A. M., Gillet, C., et al. (2007). Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade. *J. Clin. Oncol.* 25, 1239–1246. doi: 10.1200/JCO.2006.07.1522
- Magnusson, K., Gremel, G., Rydén, L., Pontén, V., Uhlén, M., Dimberg, A., et al. (2016). ANLN is a prognostic biomarker independent of Ki-67 and essential for cell cycle progression in primary breast cancer. *BMC Cancer* 16 (1), 904. doi: 10.1186/s12885-016-2923-8
- Muranen, T. A., Greco, D., Fagerholm, R., Kilpivaara, O., Kampjarvi, K., Aittomaki, K., et al. (2011). Breast tumors from CHEK2 1100delC-mutation carriers: genomic landscape and clinical implications. *Breast Cancer Res.* 13. doi: 10.1186/bcr3015
- Quann, K., Gonzales, D. M., Mercier, I., Wang, C., Sotgia, F., Pestell, R. G., et al. (2013). Caveolin-1 is a negative regulator of tumor growth in glioblastoma and modulates chemosensitivity to temozolomide. *Cell Cycle* 12, 1510–1520. doi: 10.4161/cc.24497
- Strietz, J., Stepputtis, S. S., Preca, B. T., Vannier, C., Kim, M. M., Castro, D. J., et al. (2016). ERN1 and ALPK1 inhibit differentiation of bi-potential tumor-initiating cells in human breast cancer. *Oncotarget* 7 (50), 83278–83293. doi: 10.18632/oncotarget.13086
- Suzuki, C., Daigo, Y., Ishikawa, N., Kato, T., Hayama, S., Ito, T., et al. (2005). ANLN plays a critical role in human lung carcinogenesis through the activation of RHOA and by involvement in the phosphoinositide 3-kinase/AKT pathway. *Cancer Res.* 65, 11314–11325. doi: 10.1158/0008-5472.CAN-05-1507
- Takahashi, H., and Shibuya, M. (2005). The vascular endothelial growth factor (VEGF)/VEGF receptor system and its role under physiological and pathological conditions. *Clin. Sci.* 109, 227–241. doi: 10.1042/CS20040370
- Thuma, F., and Zoller, M. (2013). EpCAM-associated claudin-7 supports lymphatic spread and drug resistance in rat pancreatic cancer. *Int. J. Cancer* 133 (4), 855–866. doi: 10.1002/ijc.28085
- Wang, G., Shen, W., Cui, L., Chen, W., Hu, X., and Fu, J. (2016). Overexpression of Anillin (ANLN) is correlated with colorectal cancer progression and poor prognosis. *Cancer Biomarkers* 16, 459–465. doi: 10.3233/CBM-160585
- Zhang, S., Nguyen, L. H., Zhou, K., Tu, H. C., Sehgal, A., Nassour, I., et al. (2018). Knockdown of anillin actin binding protein blocks cytokinesis in hepatocytes and reduces liver tumor development in mice without affecting regeneration. *Gastroenterology* 154, 1421–1434. doi: 10.1053/j.gastro.2017.12.013.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Dai, Mei, Chen and Cai. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Prognostic Value of Ki-67 in Patients With Resected Triple-Negative Breast Cancer: A Meta-Analysis

Qiang Wu<sup>1†</sup>, Guangzhi Ma<sup>1†</sup>, Yunfu Deng<sup>1†</sup>, Wuxia Luo<sup>2</sup>, Yaqin Zhao<sup>3</sup>, Wen Li<sup>1\*</sup> and Qinghua Zhou<sup>1\*</sup>

<sup>1</sup> Lung Cancer Center & Institute, West China Hospital, Sichuan University, Chengdu, China, <sup>2</sup> Department of Oncology, Chengdu First People's Hospital, Chengdu, China, <sup>3</sup> Cancer Center, West China Hospital, Sichuan University, Chengdu, China

## OPEN ACCESS

### Edited by:

Mothaffar Rimawi,  
Baylor College of Medicine,  
United States

### Reviewed by:

Yoichi Naito,  
National Cancer Center Hospital  
East, Japan  
Xiaosong Chen,  
Shanghai Jiao Tong University, China

### \*Correspondence:

Wen Li  
liwensc@163.com  
Qinghua Zhou  
zhouqh135@163.com

<sup>†</sup>These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
Women's Cancer,  
a section of the journal  
Frontiers in Oncology

Received: 09 May 2019

Accepted: 30 September 2019

Published: 17 October 2019

### Citation:

Wu Q, Ma G, Deng Y, Luo W, Zhao Y,  
Li W and Zhou Q (2019) Prognostic  
Value of Ki-67 in Patients With  
Resected Triple-Negative Breast  
Cancer: A Meta-Analysis.  
Front. Oncol. 9:1068.  
doi: 10.3389/fonc.2019.01068

**Background:** Ki-67 is a widely used marker of tumor proliferation, but the prognostic value of ki-67 in triple-negative breast cancer (TNBC) has not been comprehensively reviewed. This meta-analysis was conducted to evaluate the association between ki-67 expression and survival of patients with resected TNBC.

**Materials and Methods:** Relevant studies, evaluating the prognostic impact of pretreatment ki-67 in resected TNBC patients, were identified from PubMed, Embase, Web of Science, China National Knowledge Infrastructure, and Cochrane Library until March 14, 2019. Hazard ratios (HRs) with 95% confidence intervals (CI) were calculated as effect values for disease-free survival (DFS) and overall survival (OS).

**Results:** In present meta-analysis, 35 studies with 7,716 enrolled patients were eligible for inclusion. Pooled results showed that a high ki-67 expression was significantly associated with poor DFS (HR = 1.73, 95% CI: 1.45–2.07,  $p < 0.001$ ) and poor OS (HR = 1.65, 95% CI: 1.27–2.14,  $p < 0.001$ ) in resected TNBC. In the subgroup analysis, when a cutoff of Ki-67 staining  $\geq 40\%$  was applied, the pooled HR for DFS and OS was 2.30 (95% CI 1.54–3.44,  $p < 0.001$ ) and 2.95 (95% CI 1.67–5.19,  $p < 0.001$ ), respectively.

**Conclusion:** A high Ki-67 expression is a poor prognostic factor of resected TNBC. The cut-off of ki-67  $\geq 40\%$  is associated with a greater risk of recurrence and death compared with lower expression rates, despite the Ki-67 threshold with the greatest prognostic significance is as yet unknown.

**Keywords:** Ki-67, triple-negative breast cancer, TNBC, prognosis, meta-analysis

## INTRODUCTION

Breast cancer is one of the most frequently diagnosed cancers and the leading cause of cancer morbidity in women worldwide. It affected more than 1.6 million individuals in 2012 and constituted  $\sim 15\%$  of all cancer-related deaths among females (1). Triple-negative breast cancer (TNBC) is a subtype of breast cancer and accounts for about 12 to 17% of all breast cancers (2). Due to lacking the expression of estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor type 2 (HER2) in tumor cells, patients with TNBC are neither sensitive to endocrine therapy nor therapies targeted to HER2 (3). TNBC is usually a high-grade invasive ductal carcinoma without a special pathological type, and it is also a heterogeneous



disease because some of these patients are obviously sensitive to chemotherapy with likelihood to achieve a favorable prognosis (4, 5). Thus, sufficient and valid prognostic factors of TNBC should be identified.

Ki-67, a non-histone nuclear protein, is present in the cell nucleus during all of the active phases of the cell cycle ( $G_1$ , S,  $G_2$ , and mitosis) but absent in quiescent cells ( $G_0$ ), which makes it a widely used biomarker of tumor proliferation and a crucial element of pathological assessment (6, 7). The prognostic significance of Ki-67 has been extensively evaluated in various malignancies, including breast cancer. Ki-67 is established as a vital factor in the distinction between luminal A and luminal B breast cancer subtypes by the 2011 and 2013 St. Gallen International Breast Cancer Conference (8, 9). Unlike its role in luminal diseases whose low Ki-67 expression achieves an enhanced prognosis after standard systematic treatments, the prognostic value of Ki-67 in TNBC is still unclear and no consensus has been reached (10). Therefore, this study focused on the assessment of the prognostic value of Ki-67 in resected TNBC patients.

## METHODS

Our meta-analysis was conducted in line with the “Preferred Reporting Items for Systematic Reviews and Meta-Analyses” (PRISMA) statement (11).

### Search Strategy

A comprehensive electronic search of PubMed, Embase, Web of Science, China National Knowledge Infrastructure, and Cochrane Library was conducted without language restriction to identify all relevant full-length studies on the prognostic role of Ki-67 in patients with TNBC. To retrieve data as much as possible, we expand the search scope by using the keywords as follow: (“Ki-67” or “mib-1” or “proliferative marker”) and (“breast cancer” or “breast tumor” or “breast carcinoma” or “breast neoplasm”). The beginning date was not limited, and the search was up to March 14, 2019. References cited in eligible studies were also searched manually to obtain additional pertinent articles.

### Study Selection

The inclusion criteria were as follows: (1) studies or subsets in studies investigating the association between Ki-67 and prognosis in resected TNBC who has received neo-adjuvant or adjuvant treatment; (2) studies have adequate data for calculation including the hazard ratio (HR) and its corresponding 95% confidence interval (CI), and (3) the threshold value of Ki-67 was determined by pretreatment biopsy specimen.

The exclusion criteria were as follows: (1) non-original research articles with limited data, such as reviews, letters, comments, conference abstracts, or case reports; (2) studies without adequate survival or recurrence data for further calculation; (3) studies involving metastatic diseases; (4) overlapping or duplicate data; and (5) studies with a sample size of <30 analyzable cases.

## Data Extraction and Quality Assessment

The following data was extracted: first author's name, year of publication, country, study design and sample size, demographic characteristics (e.g., age, gender, and geographical background), cut-off value of Ki-67 expression, percentage of positive lymph nodes, treatment, and the HR with 95% CI of disease-free survival (DFS) and overall survival (OS). Multivariate outcomes were preferred when multivariate and univariate analyses performed simultaneously.

Newcastle–Ottawa Scale (NOS) was used to examine the qualities of the included studies (12). This evaluation tool covered the selection, comparability, and clinical outcomes, and studies were considered to be of high quality when they scored 6 or more.

## Statistical Analysis

Prognostic outcomes, including DFS and OS, were the primary endpoints of this study. DFS was defined as the interval period from the date of operation to the first observation of recurrence or the last follow-up without evidence of recurrence. OS was defined as the time from the first diagnosis of primary breast cancer to the time of death from any cause. HRs with 95% CIs for prognostic outcomes were extracted for further calculation. For those that were indirectly given in publications, published data and figures from original papers were extracted to calculate the corresponding HRs by utilizing the methods described by Tierney et al. (13).

Cochrane's  $Q$  ( $P < 0.1$  was considered significant) and Higgins's  $I^2$  ( $I^2 > 50\%$  was considered substantially heterogeneous) statistic tests were used to evaluate the heterogeneity among the eligible studies (14). A fixed-effect model would be preferred in the analyses to acquire precise results if the heterogeneities were insignificant. Otherwise, a random-effect model should be utilized (15). Subgroup analyses were also conducted to investigate the role of Ki-67 in specific populations and the potential source of heterogeneity. Publication bias was assessed with a funnel plot via Egger's and Begg's tests, and results were considered insignificant when  $P > 0.1$  (16). Sensitivity analysis was performed to explore the influence of individual studies on the summarized results.

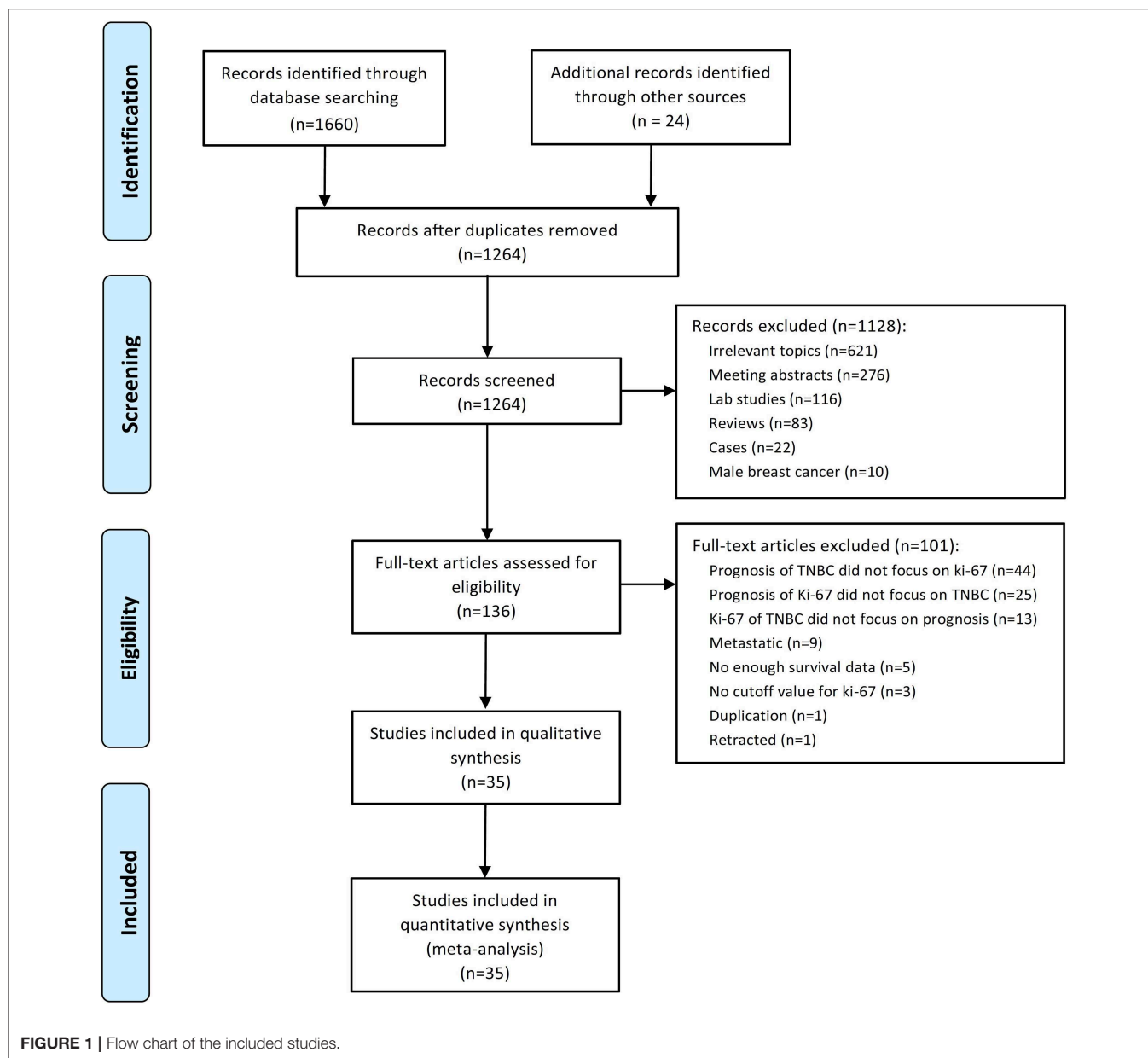
Kaplan–Meier curves were recognized by Engauge Digitizer 4.1 (free software downloaded for <http://getdata-graph-digitizer.com/>). All tests were two sided, and  $P < 0.05$  indicated statistical significance. Data analyses were performed with Stata 12.0 (StataCorp LP, TX, USA).

## RESULTS

### Selection of Studies

A total of 1,684 potential studies were identified by the search algorithm. After duplicates were removed, abstracts of the 1,264 remaining studies were reviewed. Of these studies, 1,128 were excluded, and 136 potentially relevant studies were selected for further examination. A total of 101 studies were excluded because the prognosis of TNBC did not focus on Ki-67 ( $n = 44$ ); the prognosis of Ki-67 did not highlight TNBC ( $n = 25$ ); and Ki-67 of TNBC did not cover prognosis ( $n = 13$ ), metastatic disease ( $n = 9$ ), insufficient survival data ( $n = 5$ ), no cutoff for Ki-67 ( $n = 3$ ),





duplication ( $n = 1$ ), and retracted study ( $n = 1$ ). Finally, 35 studies regarding the prognostic role of Ki-67 in TNBC subjected to neo-adjuvant or adjuvant chemotherapy were eligible for this meta-analysis (17–51). The flow diagram of studies selection was summarized in **Figure 1**.

## Study Characteristics

A total of 7,716 patients with TNBC were enrolled in the 35 included studies for analyses. The patients' median age ranged from 50 to 60 years, and the median follow-up varied from 11 to 112 months. The cutoff of Ki-67 was 10%–50%. The article quality assessed by NOS was 6–9, and 80% of the included studies had a quality of 7–9. None of these studies included patients who underwent surgery alone without neoadjuvant or adjuvant

treatment. **Table 1** summarizes the main characteristics of the included studies.

## Relationship Between Ki-67 Expression and Prognosis

In **Figure 2**, 29 studies reported the association between Ki-67 and DFS, whereas 24 determined the OS. The pooled HR of DFS comparing the high Ki-67 expression level to the low was 1.73 (95% CI: 1.45–2.07;  $p < 0.001$ ; **Figure 2A**). No significant heterogeneity ( $I^2 = 43.7\%$ ) was found, and the fixed effect model was used. The pooled HR of OS was 1.65 (95% CI: 1.27–2.14;  $p < 0.001$ ; **Figure 2B**), and moderate heterogeneity ( $I^2 = 62.6\%$ ) existed among these studies.

**TABLE 1** | Characteristics of the included studies in this meta-analysis.

References	Analysis of survival data	Neoadjuvant or adjuvant	Number of patients	Chemotherapy regimen	Region	Test tissue	Ki-67 cut off value (%)	ER/PR cut off value (%)
Toyama et al. (17)	UVA	Adjuvant	71	5-FU	Asian	Operative specimen	10	10
Trere et al. (18)	Survival Curve	Adjuvant	24	CMF	European	Operative specimen	20	10
Lee et al. (19)	MVA	Adjuvant	1550	NR	Asian	Operative specimen	20	10
Nishimura et al. (20)	MVA	Adjuvant	356	CMF, CE(F), Taxane	Asian	Operative specimen	20	10
Wang et al. (21)	UVA	Adjuvant	42	Paclitaxel-based	Asian	Operative specimen	30	10
Keam et al. (23)	MVA	Neoadjuvant	105	Taxel/anthracyclin-based	Asian	Pre-NAC	10	10
Li et al. (24)	UVA	Adjuvant	125	NR	Asian	Operative specimen	NR	10
Masuda et al. (25)	MVA	Neoadjuvant	33	Taxel/anthracyclin-based	Asian	Pre-NAC	50	10
Miyashita et al. (26)	MVA	Adjuvant	102	NR	Asian	Operative specimen	40	1
Kashiwagi et al. (22)	Survival Curve	Adjuvant	190	5-FU/anthracyclin-based	Asian	Operative specimen	30	1
Munzone et al. (27)	Survival Curve	Adjuvant	496	Anthracyclin-based, CMF	European	Operative specimen	35	1
Ryu et al. (28)	UVA	Adjuvant	94	Anthracyclin-based	Asian	Operative specimen	10	10
Xue et al. (29)	MVA	Adjuvant	913	Taxel/anthracyclin-based, CMF	Asian	Operative specimen	14	5
Huang et al. (30)	MVA	Adjuvant	185	Taxel/anthracyclin-based, CMF	Asian	Operative specimen	20	1
Milde-Langosch et al. (31)	MVA	Adjuvant	95	CMF/EC	European	Operative specimen	NR	NR
Xu et al. (32)	MVA	Adjuvant	122	CMF	Asian	Operative specimen	10	10
Yamashita et al. (33)	MVA	Adjuvant	82	NR	Asian	Operative specimen	30	1
Zhang et al. (34)	MVA	Adjuvant	428	Taxel/anthracyclin-based	Asian	Operative specimen	14	10
Zhou et al. (35)	UVA	Adjuvant	31	Taxel/anthracyclin-based, CMF	Asian	Operative specimen	10	10
Schmidt et al. (38)	Survival Curve	Adjuvant/neoadjuvant	103	Taxel-based	European	NR	14	1
Park et al. (36)	UVA	Adjuvant	1551	NR	Asian	Operative specimen	50	1
Pistelli et al. (37)	UVA	Adjuvant	81	Anthracyclin-based, CMF	European	Operative specimen	30	10
Hao et al. (40)	MVA	Adjuvant	571	NR	Asian	Operative specimen	35	1
Khalifa et al. (41)	UVA	Adjuvant/neoadjuvant	106	Taxel/anthracyclin-based	European	Pre-NAC	20	10
Liu et al. (42)	UVA	Adjuvant	154	NR	Asian	Operative specimen	30	1
Asano et al. (39)	UVA	Neoadjuvant	61	Taxel/anthracyclin-based	Asian	Pre-NAC	14	1
Wang et al. (43)	MVA	Adjuvant	363	Taxel/anthracyclin-based	Asian	Operative specimen	40	1
Yang and Han (50)	UVA	Adjuvant	199	NR	Asian	Operative specimen	14	NR
Yue et al. (44)	MVA	Adjuvant	192	NR	Others	Operative specimen	50	1

(Continued)

TABLE 1 | Continued

References	Analysis of survival data	Neoadjuvant or adjuvant	Number of patients	Chemotherapy regimen	Region	Test tissue	Ki-67 cut off value (%)	ER/PR cut off value (%)
Zakaria et al. (45)	MVA	Adjuvant	77	Taxel/anthracyclin-based	Others	Operative specimen	10	10
Zhong et al. (46)	UVA	Adjuvant	90	NR	Asian	Operative specimen	14	1
Ieni et al. (47)	MVA	Adjuvant	65	NR	European	Operative specimen	20	1
Kwon et al. (48)	UVA	Adjuvant/neoadjuvant	230	Taxel/anthracyclin-based	Asian	Pre-NAC	20	1
Najafi et al. (49)	Survival Curve	Adjuvant	119	NR	Asian	Operative specimen	20	10
Wang and Liu (51)	MVA	Adjuvant	110	NR	Asian	Operative specimen	20	NR

UVA, univariate analysis; MVA, multivariate analysis; NR, not reported; NAC, neoadjuvant chemotherapy; CMF, cyclophosphamide, methotrexate, and 5-fluorouracil; CEF, cyclophosphamide, epirubicin, and 5-FU; EC, epirubicin/cyclophosphamide.

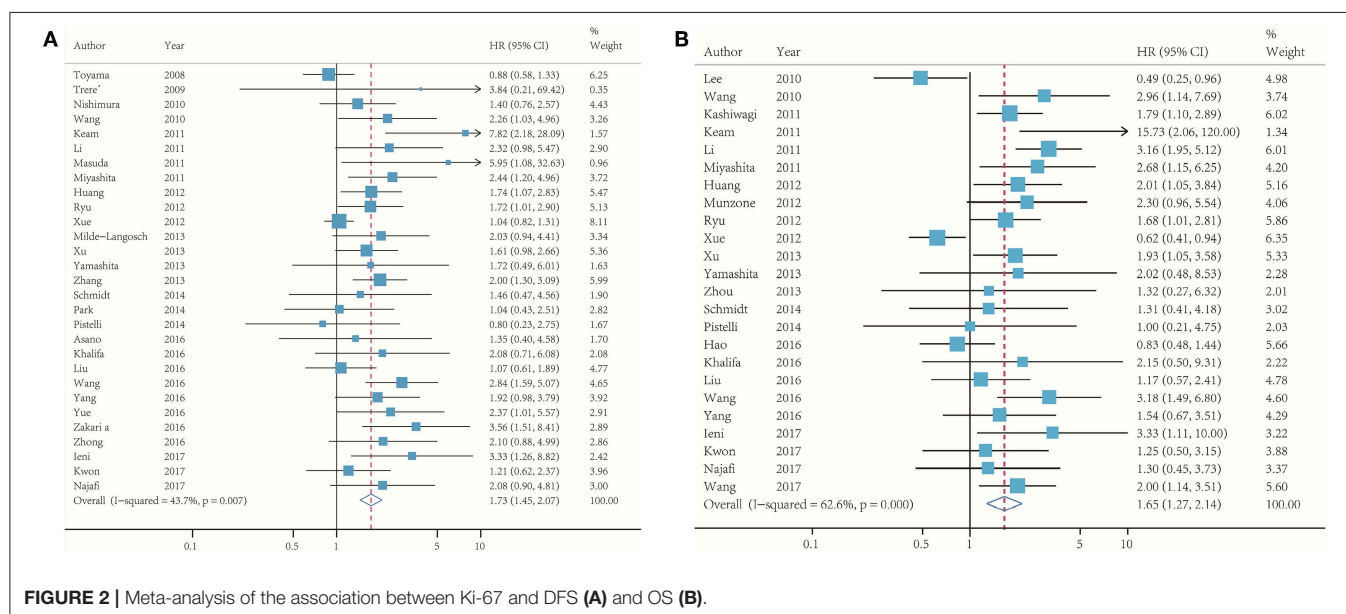


FIGURE 2 | Meta-analysis of the association between Ki-67 and DFS (A) and OS (B).

## Subgroup Analyses

Subgroup analyses were conducted in accordance with Ki-67 cutoffs, positive ER/PR expression thresholds (1% or 10%), treatment strategies (neo-adjuvant or adjuvant), and geographic regions (Europe, Asian, or other regions). Despite the limited number of studies in some subgroups, the results of DFS (Figure 3A) and OS (Figure 3B) stratified by these factors were consistent. Noticeably, the pooled HR for DFS and OS was 2.30 (95% CI 1.54–3.44,  $p < 0.001$ ) and 2.95 (95% CI 1.67–5.19,  $p < 0.001$ ), respectively, under the circumstance of a cutoff of Ki-67 staining  $\geq 40\%$ .

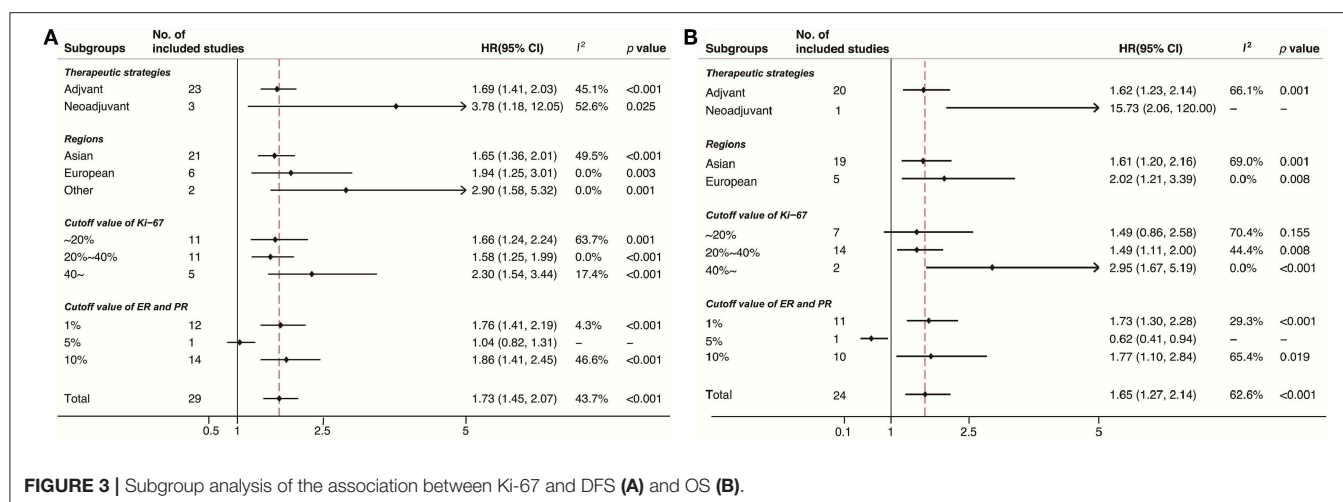
## Publication Bias

In Begg's plots of publication bias,  $p$ -value was 0.209 (Figure 4), implying that publication bias did not exist in the present meta-analysis.

## DISCUSSION

TNBC has a worse prognosis than other phenotypes of breast cancer because of its aggressive biology and insensitivity to targeted therapy (52). Biomarkers useful in the selection of appropriate treatment strategies and the prediction of prognosis should be identified.

Previous studies demonstrated the prognostic role of Ki-67, as a critical biomarker of cell proliferation, in various malignancies that originate from organs and tissues, such as prostate, stomach, esophagus, cervix, and breast. A high expression level of Ki-67 protein was accompanied with poor prognostic outcomes (53). Several meta-analyses have shown that a high Ki-67 expression level is associated with the likelihood of achieving a pathological complete response (pCR) after patients with TNBC receive neo-adjuvant chemotherapy (NAC), and these patients may

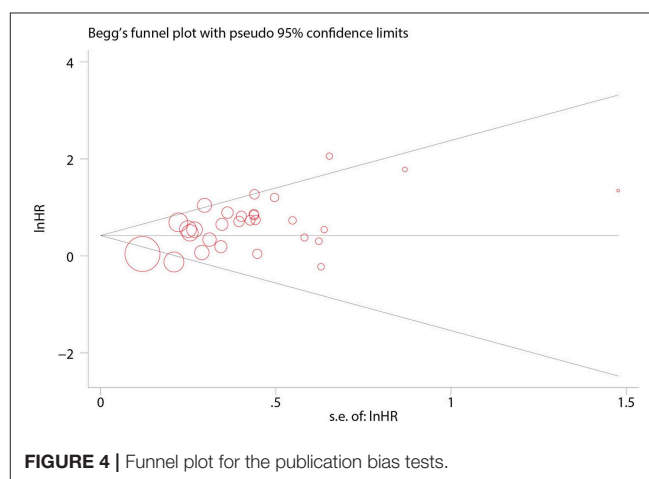


have favorable outcomes. Nevertheless, most of these studies included small sample sizes and contained diverse cut-offs of Ki-67 (54, 55).

In this meta-analysis, data were pooled to assess the prognostic value of Ki-67 in patients who suffered from resected TNBC and received neo-adjuvant or adjuvant chemotherapy. The results showed that patients with a high Ki-67 expression substantially had worse DFS and OS than their counterparts regardless of treatment strategies, study regions, Ki-67 cutoffs, or ER/PR thresholds.

Despite the consistency obtained in our study, the optimized cutoff of Ki-67 is still under deliberation (56). Some investigators suggested that Ki-67 should be used as a continuous marker to fully reflect the biological behavior of tumor proliferation and simultaneously resolve the cutoff issue; however, confronting diverse therapeutic strategies is impractical for clinical decision making (7). A previous meta-analysis indicated that a 25% cutoff of Ki-67 is adequate to distinguish patients with breast cancer at different risks of death (57). The cutoff selection of Ki-67 may be apparent if this parameter is considered within each subtype, and a 14% cutoff for the classification of luminal A and luminal B cancers was proposed in the 2011 St. Gallen Consensus (9). Considering that the baseline Ki-67 values of TNBC are usually higher than those of luminal diseases, Leskandarany et al. reported that the optimized Ki-67 cutoff within a TNBC subgroup population is 70% as determined by X-tile (58). Different Ki-67 values were selected as a cut-point in our included studies, and the threshold of Ki-67 varied between 10 and 50%. The subgroup analysis based on the Ki-67 cutoff indicated that the prediction was significant in all of the subgroups except one subgroup (Ki-67 < 20%). This finding might indicate that further prospective studies should be performed to optimize the cutoff of Ki-67 in TNBC.

Baseline Ki-67 confirms the high chemosensitivity of highly proliferating TNBC after patients receive NAC, TNBC with a high Ki-67 expression likely has a high rate of pCR, which predicts favorable outcomes (59). However, studies have shown that TNBC with a high Ki-67 expression is associated with



a poor prognosis because of rapid recurrence within 3 years despite a high pCR rate. A Korean study has demonstrated that a high Ki-67 expression ( $\geq 10\%$ ) is significantly associated with poor relapse-free survival and OS in preoperative TNBC despite a high pCR rate (26). Our subgroup analyses showed that a high Ki-67 expression is an adverse prognostic factor of DFS and OS both in the two groups of patients treated with adjuvant or neo-adjuvant therapy. Keam et al. reported that patients who suffer from TNBC and receive neo-adjuvant therapy with a high Ki-67 expression have a pattern of early recurrence. By contrast, the low-Ki-67-expressing subgroup did not have any pattern, indicating that a high Ki-67 expression, which indicated a high proliferation potential, might result in early recurrence. This phenomenon might partly explain why a high Ki-67 expression remained an adverse prognostic factor in the neo-adjuvant subgroup (23).

The American Society of Clinical Oncology and the College of American Pathologists Guideline Recommendations indicated that the cutoff for positive ER or PR should be  $\geq 1\%$  of immunoreactive tumor cell nuclei in 2010, and the previous threshold was  $>10\%$ . Hence, a subgroup analysis classified by ER cut-off was performed. The results showed that a high



Ki-67 expression was an adverse prognostic factor of all the subgroups, indicating that Ki-67 might be a prognostic factor of patients whose ER expression ranged from 2 to 10. Another study showed that defining triple-negative breast cancer as HER2-negative breast cancer with <10% rather than <1% of ER and progesterone receptor expression because HER2-negative primary breast cancer with ER < 10% clinically behaves like TNBC in terms of survival outcomes (60). This phenomenon might partly explain why Ki-67 was a poor prognostic factor of this patient subgroup.

Subgroup analyses on regions where these studies were conducted yielded the following classifications: Europe, Asia, and others. The results showed that a high Ki-67 expression was consistently an adverse prognostic factor of DFS and OS in these three subgroups. Moreover, the pooled data showed that TNBC was more likely to recur in Europe than in Asia. However, only eight studies were from Europe, while 27 studies were from Asia. Therefore, these findings should be carefully considered, and further studies should be performed to verify these results.

Notably, our study has a few limitations. First, due to linguistic constraints, we included studies written in English and Chinese only, hence publications in other languages could have been omitted. Second, we failed to perform subgroup analyses on other parameters, such as age or tumor stage, because of insufficient background information and thus might cause heterogeneity in the pooled results. Other clinical heterogeneities among studies, such as different NAC and adjuvant regimens, were not analyzed.

In conclusion, this study demonstrated that higher Ki-67 expression is a poorer prognostic factor of resected TNBC. The cut-off of ki-67  $\geq 40\%$  is associated with a greater risk of recurrence and death compared with lower expression rates, despite the Ki-67 threshold with the greatest prognostic significance is as yet unknown.

## DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the manuscript/supplementary files.

## AUTHOR CONTRIBUTIONS

QW, WLi, and QZ contributed to the conception and design of this research. QW, GM, YD, WLu, YZ, WLi, and QZ contributed to the drafting of the article and final approval of the submitted version, and contributed to data analyses and the interpretation and completion of the figures and tables. All authors read and approved the final manuscript.

## FUNDING

This study was supported by grants from the National Natural Science Foundation of China (81572288). The sponsor had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

## REFERENCES

- Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin.* (2015) 65:87–108. doi: 10.3322/caac.21262
- Foulkes WD, Smith IE, Reis-Filho JS. Triple-negative breast cancer. *N Engl J Med.* (2010) 363:1938–48. doi: 10.1056/NEJMra1001389
- Penault-Llorca F, Viale G. Pathological and molecular diagnosis of triple-negative breast cancer: a clinical perspective. *Ann Oncol.* (2012) 23:vi19–22. doi: 10.1093/annonc/mds190
- Weigelt B, Pusztai L, Ashworth A, Reis-Filho JS. Challenges translating breast cancer gene signatures into the clinic. *Nat Rev Clin Oncol.* (2011) 9:58–64. doi: 10.1038/nrclinonc.2011.125
- Chen X, Li J, Gray WH, Lehmann BD, Bauer JA, Shyr Y, et al. TNBCtype: a subtyping tool for triple-negative breast cancer. *Cancer Inform.* (2012) 11:147–56. doi: 10.4137/CIN.S9983
- Scholz T, Gerdes J. The Ki-67 protein: from the known and the unknown. *J Cell Physiol.* (2000) 182:311–22. doi: 10.1002/(SICI)1097-4652(200003)182:3<311::AID-JCP1>3.0.CO;2-9
- Denkert C, Budczies J, von Minckwitz G, Wienert S, Loibl S, Klauschen F. Strategies for developing Ki67 as a useful biomarker in breast cancer. *Breast.* (2015) 24:S67–72. doi: 10.1016/j.breast.2015.07.017
- Goldhirsch A, Winer EP, Coates AS, Gelber RD, Piccart-Gebhart M, Thürlimann B, et al. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. *Ann Oncol.* (2013) 24:2206–23. doi: 10.1093/annonc/mdt303
- Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thürlimann B, Senn HJ, et al. Strategies for subtypes—dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. *Ann Oncol.* (2011) 22:1736–47. doi: 10.1093/annonc/mdr304
- Simon RM, Paik S, Hayes DF. Use of archived specimens in evaluation of prognostic and predictive biomarkers. *J Natl Cancer Inst.* (2009) 101:1446–52. doi: 10.1093/jnci/djp335
- Knobloch K, Yoon U, Vogt PM. Preferred reporting items for systematic reviews and meta-analyses (PRISMA) statement and publication bias. *J Craniomaxfac Surg.* (2011) 39:91–2. doi: 10.1016/j.jcms.2010.11.001
- Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. *Eur J Epidemiol.* (2010) 25:603–5. doi: 10.1007/s10654-010-9491-z
- Tierney JF, Stewart LA, Ghersi D, Burdett S, Sydes MR. Practical methods for incorporating summary time-to-event data into meta-analysis. *Trials.* (2007) 8:16. doi: 10.1186/1745-6215-8-16
- Higgins J, Thompson S, Deeks J, Altman D. Statistical heterogeneity in systematic reviews of clinical trials: a critical appraisal of guidelines and practice. *J Health Serv Res Policy.* (2002) 7:51–61. doi: 10.1258/1355819021927674
- Kontopantelis E, Reeves D. Performance of statistical methods for meta-analysis when true study effects are non-normally distributed: a simulation study. *Stat Methods Med Res.* (2012) 21:409–26. doi: 10.1177/0962280210392008
- Song F, Gilbody S. Bias in meta-analysis detected by a simple, graphical test. Increase in studies of publication bias coincided with increasing use of meta-analysis. *BMJ (Clinical research ed.).* (1998) 316:471. doi: 10.1136/bmj.316.7129.469
- Toyama T, Yamashita H, Kondo N, Okuda K, Takahashi S, Sasaki H, et al. Frequently increased epidermal growth factor receptor (EGFR) copy numbers and decreased BRCA1 mRNA expression in Japanese triple-negative breast cancers. *BMC Cancer.* (2008) 8:309. doi: 10.1186/1471-2407-8-309
- Treré D, Brighenti E, Donati G, Ceccarelli C, Santini D, Taffurelli M, et al. High prevalence of retinoblastoma protein loss in triple-negative breast cancers and its association with a good prognosis in patients treated with

- adjuvant chemotherapy. *Ann Oncol.* (2009) 20:1818–23. doi: 10.1093/annonc/mdp209
19. Lee JA, Kim KI, Bae JW, Jung YH, An H, Lee ES, et al. Triple negative breast cancer in Korea-distinct biology with different impact of prognostic factors on survival. *Breast Cancer Res Treat.* (2010) 123:177–87. doi: 10.1007/s10549-010-0998-5
  20. Nishimura R, Osako T, Okumura Y, Hayashi M, Toyozumi Y, Arima N. Ki-67 as a prognostic marker according to breast cancer subtype and a predictor of recurrence time in primary breast cancer. *Exp Ther Med.* (2010) 1:747–54. doi: 10.3892/etm.2010.133
  21. Wang J, Liu Y, Ji R, Gu Q, Zhao X, Liu Y, et al. Prognostic value of the X-linked inhibitor of apoptosis protein for invasive ductal breast cancer with triple-negative phenotype. *Hum Pathol.* (2010) 41:1186–95. doi: 10.1016/j.humpath.2010.01.013
  22. Kashiwagi S, Yashiro M, Takashima T, Aomatsu N, Ikeda K, Ogawa Y, et al. Advantages of adjuvant chemotherapy for patients with triple-negative breast cancer at Stage II: usefulness of prognostic markers E-cadherin and Ki67. *Breast Cancer Res.* (2011) 13:R122. doi: 10.1186/bcr3068
  23. Keam B, Im SA, Lee KH, Han SW, Oh DY, Kim JH, et al. Ki-67 can be used for further classification of triple negative breast cancer into two subtypes with different response and prognosis. *Breast Cancer Res.* (2011) 13:R22. doi: 10.1186/bcr2834
  24. Li C, Li R, Song H, Wang D, Feng T, Yu X, et al. Significance of AEG-1 expression in correlation with VEGF, microvessel density and clinicopathological characteristics in triple-negative breast cancer. *J Surg Oncol.* (2011) 103:184–92. doi: 10.1002/jso.21788
  25. Masuda H, Masuda N, Kodama Y, Ogawa M, Karita M, Yamamura J, et al. Predictive factors for the effectiveness of neoadjuvant chemotherapy and prognosis in triple-negative breast cancer patients. *Cancer Chemother Pharmacol.* (2011) 67:911–7. doi: 10.1007/s00280-010-1371-4
  26. Miyashita M, Ishida T, Ishida K, Tamaki K, Amari M, Watanabe M, et al. Histopathological subclassification of triple negative breast cancer using prognostic scoring system: five variables as candidates. *Virchows Archiv Int J Pathol.* (2011) 458:65–72. doi: 10.1007/s00428-010-1009-2
  27. Munzone E, Botteri E, Sciandivasci A, Curigliano G, Nolè F, Mastropasqua M, et al. Prognostic value of Ki-67 labeling index in patients with node-negative, triple-negative breast cancer. *Breast Cancer Res Treat.* (2012) 134:277–82. doi: 10.1007/s10549-012-2040-6
  28. Ryu DW, Lee CH. Outcome of triple-negative breast cancer in patients with or without markers regulating cell cycle and cell death. *J Korean Surg Soc.* (2012) 83:187–95. doi: 10.4174/jkss.2012.83.4.187
  29. Xue C, Wang X, Peng R, Shi Y, Qin T, Liu D, et al. Distribution, clinicopathologic features and survival of breast cancer subtypes in Southern China. *Cancer Sci.* (2012) 103:1679–87. doi: 10.1111/j.1349-7006.2012.02339.x
  30. Huang L, Liu Z, Chen S, Liu Y, Shao Z. A prognostic model for triple-negative breast cancer patients based on node status, cathepsin-D and Ki-67 index. *PLoS ONE.* (2013) 8:e83081. doi: 10.1371/journal.pone.0083081
  31. Milde-Langosch K, Karn T, Müller V, Witzel I, Rody A, Schmidt M, et al. Validity of the proliferation markers Ki67, TOP2A, and RacGAP1 in molecular subgroups of breast cancer. *Breast Cancer Res Treat.* (2013) 137:57–67. doi: 10.1007/s10549-012-2296-x
  32. Xu J, Wu X, Zhou WH, Liu AW, Wu JB, Deng JY, et al. Aurora-A identifies early recurrence and poor prognosis and promises a potential therapeutic target in triple negative breast cancer. *PLoS ONE.* (2013) 8:e56919. doi: 10.1371/journal.pone.0056919
  33. Yamashita N, Tokunaga E, Kitao H, Hisamatsu Y, Taketani K, Akiyoshi S, et al. Vimentin as a poor prognostic factor for triple-negative breast cancer. *J Cancer Res Clin Oncol.* (2013) 139:739–46. doi: 10.1007/s00432-013-1376-6
  34. Zhang J, Wang Y, Yin Q, Zhang W, Zhang T, Niu Y. An associated classification of triple negative breast cancer: the risk of relapse and the response to chemotherapy. *Int J Clin Exp Pathol.* (2013) 6:1380–91.
  35. Zhou L, Li K, Luo Y, Tian L, Wang M, Li C, et al. Novel prognostic markers for patients with triple-negative breast cancer. *Hum Pathol.* (2013) 44:2180–7. doi: 10.1016/j.humpath.2013.03.021
  36. Park HS, Jang MH, Kim EJ, Kim HJ, Lee HJ, Kim YJ, et al. High EGFR gene copy number predicts poor outcome in triple-negative breast cancer. *Mod Pathol.* (2014) 27:1212–22. doi: 10.1038/modpathol.2013.251
  37. Pistelli M, Caramanti M, Biscotti T, Santinelli A, Pagliacci A, De Lisa M, et al. Androgen receptor expression in early triple-negative breast cancer: clinical significance and prognostic associations. *Cancers.* (2014) 6:1351–62. doi: 10.3390/cancers6031351
  38. Schmidt G, Meyberg-Solomayer G, Gerlinger C, Juhasz-Böss I, Herr D, Rody A, et al. Identification of prognostic different subgroups in triple negative breast cancer by Her2-neu protein expression. *Archiv Gynecol Obstet.* (2014) 290:1221–9. doi: 10.1007/s00404-014-3331-4
  39. Asano Y, Kashiwagi S, Goto W, Kurata K, Noda S, Takashima T, et al. Tumour-infiltrating CD8 to FOXP3 lymphocyte ratio in predicting treatment responses to neoadjuvant chemotherapy of aggressive breast cancer. *Br J Surg.* (2016) 103:845–54. doi: 10.1002/bjs.10127
  40. Hao S, He ZX, Yu KD, Yang WT, Shao ZM. New insights into the prognostic value of Ki-67 labeling index in patients with triple-negative breast cancer. *Oncotarget.* (2016) 7:24824–31. doi: 10.18632/oncotarget.8531
  41. Khalifa J, Duprez-Paumier R, Filleron T, Lacroix Triki M, Jouve E, Dalenc F, et al. Outcome of pN0 triple-negative breast cancer with or without lymph node irradiation: a single institution experience. *Breast J.* (2016) 22:510–9. doi: 10.1111/tbj.12626
  42. Liu YX, Wang KR, Xing H, Zhai XJ, Wang LP, Wang W. Attempt towards a novel classification of triple-negative breast cancer using immunohistochemical markers. *Oncol Lett.* (2016) 12:1240–56. doi: 10.3892/ol.2016.4778
  43. Wang W, Wu J, Zhang P, Fei X, Zong Y, Chen X, et al. Prognostic and predictive value of Ki-67 in triple-negative breast cancer. *Oncotarget.* (2016) 7:31079–87. doi: 10.18632/oncotarget.9075
  44. Yue Y, Astvatsaturyan K, Cui X, Zhang X, Fraass B, Bose S. Stratification of prognosis of triple-negative breast cancer patients using combinatorial biomarkers. *PLoS ONE.* (2016) 11:e0149661. doi: 10.1371/journal.pone.0149661
  45. Zakaria F, El-Mashad N, Mohamed D. Androgen receptor expression as a prognostic and predictive marker in triple-negative breast cancer patients. *Alexandria J Med.* (2016) 52:131–40. doi: 10.1016/j.ajme.2015.06.002
  46. Zhong Z, Shan M, Wang J, Liu T, Shi Q, Pang D. Decreased Wnt5a expression is a poor prognostic factor in triple-negative breast cancer. *Med Sci Monitor.* (2016) 22:1–7. doi: 10.12659/MSM.894821
  47. Ieni A, Barresi V, Licata L, Cardia R, Fazzari C, Nuciforo G, et al. Immunoexpression of lactoferrin in triple-negative breast cancer patients: a proposal to select a less aggressive subgroup. *Oncol Lett.* (2017) 13:3205–9. doi: 10.3892/ol.2017.5859
  48. Kwon J, Eom KY, Koo TR, Kim BH, Kang E, Kim SW, et al. A prognostic model for patients with triple-negative breast cancer: importance of the modified nottingham prognostic index and age. *J Breast Cancer.* (2017) 20:65–73. doi: 10.4048/jbc.2017.20.1.65
  49. Najafi S, Mozaffari HR, Sadeghi M. Clinicopathological features of non-metastatic triple negative breast cancer. *Iran J Blood Cancer.* (2017) 9:18–23. doi: 10.1007/978-3-319-69980-6\_2
  50. Yang M, Han X. Nomogram model based on biomarkers for predicting the prognosis of triple negative breast cancer. *Chin J Cancer Prev Treat.* (2016) 23:988–995.
  51. Wang C, Liu R. Expression of CD147 in triple-negative breast cancer and its association With the prognosis. *J Shanghai Jiao Tong Univ.* (2017) 37:55–59.
  52. Yao H, He G, Yan S, Chen C, Song L, Rosol TJ, et al. Triple-negative breast cancer: is there a treatment on the horizon? *Oncotarget.* (2017) 8:1913–24. doi: 10.18632/oncotarget.12284
  53. Yerushalmi R, Woods R, Ravdin PM, Hayes MM, Gelmon KA. Ki67 in breast cancer: prognostic and predictive potential. *Lancet Oncol.* (2010) 11:174–83. doi: 10.1016/S1470-2045(09)70262-1
  54. Chen X, He C, Han D, Zhou M, Wang Q, Tian J, et al. The predictive value of Ki-67 before neoadjuvant chemotherapy for breast cancer: a systematic review and meta-analysis. *Fut Oncol.* (2017) 13:843–57. doi: 10.2217/fon-2016-0420

55. Tian M, Zhong Y, Zhou F, Xie C, Zhou Y, Liao Z. Effect of neoadjuvant chemotherapy in patients with triple-negative breast cancer: a meta-analysis. *Oncol Lett.* (2015) 9:2825–32. doi: 10.3892/ol.2015.3072
56. Dowsett M, Nielsen TO, A'Hern R, Bartlett J, Coombes RC, Cuzick J, et al. Assessment of Ki67 in breast cancer: recommendations from the International Ki67 in Breast Cancer working group. *J Natl Cancer Inst.* (2011) 103:1656–64. doi: 10.1093/jnci/djr393
57. Petrelli F, Viale G, Cabiddu M, Barni S. Prognostic value of different cut-off levels of Ki-67 in breast cancer: a systematic review and meta-analysis of 64,196 patients. *Breast Cancer Res Treat.* (2015) 153:477–91. doi: 10.1007/s10549-015-3559-0
58. Aleskandarany MA, Green AR, Benhasouna AA, Barros FF, Neal K, Reis-Filho JS, et al. Prognostic value of proliferation assay in the luminal, HER2-positive, and triple-negative biologic classes of breast cancer. *Breast Cancer Res.* (2012) 14:R3. doi: 10.1186/bcr3084
59. Gamucci T, Pizzuti L, Sperduti I, Mentuccia L, Vaccaro A, Moscetti L, et al. Neoadjuvant chemotherapy in triple-negative breastcancer: a multicentric retrospective observational study in real-life setting. *J Cell Physiol.* (2018) 233:2313–23. doi: 10.1002/jcp.26103
60. Fujii T, Kogawa T, Dong W, Sahin AA, Moulder S, Litton JK, et al. Revisiting the definition of estrogen receptor positivity in HER2-negative primary breast cancer. *Ann Oncol.* (2017) 28:2420–8. doi: 10.1093/annonc/mdx397

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Wu, Ma, Deng, Luo, Zhao, Li and Zhou. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Deciphering HER2 Breast Cancer Disease: Biological and Clinical Implications

Ana Godoy-Ortiz<sup>1,2\*</sup>, Alfonso Sanchez-Muñoz<sup>1,2</sup>, Maria Rosario Chica Parrado<sup>2</sup>, Martina Álvarez<sup>2</sup>, Nuria Ribelles<sup>1,2</sup>, Antonio Rueda Dominguez<sup>1,2</sup> and Emilio Alba<sup>1,2,3</sup>

<sup>1</sup> Unidad de Gestión Clínica Intercentros de Oncología Médica, Hospitales Universitarios Regional y Virgen de la Victoria de Málaga, Málaga, Spain, <sup>2</sup> Laboratorio de Biología Molecular del Centro de Investigaciones Médico-Sanitarias de Málaga (CIMES), Instituto de Investigación Biomédica de Málaga (IBIMA), Universidad de Málaga (UMA), Málaga, Spain, <sup>3</sup> Centro de Investigación Biomédica en Red de Oncología, CIBERONC-ISCIII, Madrid, Spain

## OPEN ACCESS

### Edited by:

Mothaffar Rimawi,  
Baylor College of Medicine,  
United States

### Reviewed by:

Howard Donninger,  
University of Louisville, United States  
Parvin Mehdipour,  
Tehran University of Medical  
Sciences, Iran

### \*Correspondence:

Ana Godoy-Ortiz  
anagodort@gmail.com

### Specialty section:

This article was submitted to  
Cancer Genetics,  
a section of the journal  
Frontiers in Oncology

**Received:** 16 June 2019

**Accepted:** 09 October 2019

**Published:** 29 October 2019

### Citation:

Godoy-Ortiz A, Sanchez-Muñoz A, Chica Parrado MR, Álvarez M, Ribelles N, Rueda Dominguez A and Alba E (2019) Deciphering HER2 Breast Cancer Disease: Biological and Clinical Implications. *Front. Oncol.* 9:1124. doi: 10.3389/fonc.2019.01124

The main obstacle for designing effective treatment approaches in breast cancer is the extensive and the characteristic heterogeneity of this tumor. The vast majority of critical genomic changes occurs during breast cancer progression, creating a significant variability within primary tumors as well as between the primary breast cancer and their metastases, a hypothesis have already demonstrated in retrospective studies (1). A clear example of this is the HER2-positive breast cancer. In these tumors, we can find all of the transcriptional subtypes of breast cancer, even the basal like or luminal A subtypes. Although the HER2-enriched is the most representative transcriptional subtype in the HER2-positive breast cancer, we can find it too in breast cancers with HER2-negative status. This intrinsic subtype shows a high expression of the HER2 and is associated with proliferation-related genes clusters, among other features. Therefore, two hypotheses can be suggested. First, the HER2 amplification can be a well-defined driver event present in all of the intrinsic subtypes, and not a subtype marker isolated. Secondly, HER2-enriched subtype can have a distinctive transcriptional landscape independent of HER2 amplification. In this review, we present an extensive revision about the last highlights and advances in clinical and genomic settings of the HER2-positive breast cancer and the HER2-enriched subtype, in an attempt to improving the knowledge of the underlying biology of both entities and to explaining the intrinsic heterogeneity of HER2-positive breast cancers.

**Keywords:** breast cancer, HER2-positive, intrinsic subtype, heterogeneity, HER2-enriched, molecular

## INTRODUCTION

Breast cancer (BC) is the most common malignant tumor in women and one of the principal causes of cancer mortality in this sex, despite significant improvements obtained in the last decades. Conversely, male breast cancer is a rare disease with an incidence of <1% and mainly classified by immunohistochemistry as a luminal disease (2). BC is modeled by a group of heterogeneous diseases, at both an inter- and intra-tumoral level. All of them share a substantial morphological and molecular heterogeneity, what affect to his clinical behavior and therapeutic response. A crucial objective in the treatment of any cancer disease is to perform clinical decisions through a comprehensive insight of the molecular profile of the tumor to predict the probable clinical outcome of the disease individually. By the expansion of high-throughput



molecular technologies, we can analyze changes in the genetic, epigenetic and proteomics contexts, so that allows improving in the comprehension of the complexity of BC biology.

One biomarker with reported heterogeneity in BC is the *Human Epidermal Growth Factor Receptor 2* (HER2), a component of the *EGF* receptor (EGFR) family. The overexpression of this biomarker defined the HER2-positive disease. Traditionally, HER2-positive breast cancer (HER2+ BC) has been associated with a worse prognosis and inferior outcomes in survival. However, over the last years, several therapeutic advances have been improved the clinical treatment of HER2+ disease, and thus, its prognosis. After the discovery of the intrinsic subtypes through gene expression analysis, and later transcriptomic and genomic studies, there is sufficient evidence that HER2+ BC is an entity with a large heterogeneity at multiple levels (3), including cell-to-cell. There has been discrepancy about the determination of the clinical status of HER2+ over the last years, with several guidelines and updates in order to find a formal and universal consensus. In clinical practice, HER2+ tumors are categorized by immunohistochemistry (IHC) and/or by *in situ hybridization* (ISH) in order to tailor the different therapeutic approaches (4).

The gene expression profiling has had a large-scale impact in the progress about the knowledge of the biological heterogeneity of this tumor (5). However, in this ambit, there is a considerable variability as well, what makes it even more difficult to categorize the basis of pathological diagnosis and therapeutic approach. The principal molecular subtypes of BC have widely characterized, and within HER2+ BC the most representative intrinsic subtype is the HER2-enriched (HER2-E). However, we can find HER2+ BC with luminal A, luminal B, or even the basal-like subtype (6). The intrinsic subtype HER2-E is defined generally by a higher expression of HER2 at the RNA and protein level than other subtypes, in addition the increased expression of the tumor proliferation-related genes (6, 7). Recent studies confirm that this subtype obtains the best clinical and therapeutic results by anti-HER2 therapies, with or without chemotherapy, in both adjuvant and neoadjuvant scenarios, and regardless of the clinical status of HER2 (3). Nonetheless, no more than 50% of clinically HER2+ tumors are HER2-E, and what is more exciting, we can also find this subtype in clinically HER2-negative BC, which do not receive HER2-therapies since these drugs are not approved for the treatment of clinically HER2-negative breast tumors. Therefore, we consider it is highly important to perform an extensive revision about the latest highlights and advances in clinical outcomes and genomic features within HER2+ BC and its most representative intrinsic subtype, HER2-E, with a previous extensive revision from the state of science in which these advances are based.

## CURRENT CLASSIFICATION OF BREAST CANCER

Intertumoral heterogeneity of BC is initially illustrated with a clinical staging of the disease. The TNM staging system by the *American Joint Committee on Cancer and Union for International*

*Cancer Control* (AJCC/UICC) adds information about tumor features such as size, regional lymph-node involvement or the presence of distant metastases (8). After the clinical diagnosis, the first step is the assessment of histological criteria on the primary tumor obtained by surgery and/or a core biopsy, encompassing morphology-base and immunohistochemical (IHC) analyses for testing the biomarker profile. This is a classical and non-molecular classification of BC, and sets the standard in the usual clinical practice. Classic pathological criteria, such as histological type, tumor size, grade and axillary lymph node status, are relevant for the initial prognostic evaluation (9). The expression of hormone receptors [estrogen (ER) and progesterone receptors (PR)] by IHC and the overexpression and/or amplification of HER2 by IHC and/or ISH gives additional predictive value, being elementary for guiding algorithms of treatment (9, 10), as will be discussed in the following two sections.

## Histopathological Subtypes: Morphologic Heterogeneity

The histopathological classification of BC is set by the 2012 *World Health Organization* (WHO) (11). Most of the breast cancers are adenocarcinomas, with around 70–80% defined as invasive ductal carcinomas not otherwise specified (IDC-NOS) (11). The rest, around 25–30%, are characterized by “*histological special types*” such as papilar, metaplastic, cribriform, apocrine, or mucinous carcinomas, among others (11). The majority of special types is rare and differ strongly about prognosis and response to the treatments (12). The tumor grade is the other important intrinsic characteristic of tumoral heterogeneity (13, 14).

## Immunohistochemistry: ER, PR, and HER2

Via the characterization of ER, PR, and HER2 status, we can divide BC in three phenotypes or entities. Hormone receptor-positive breast cancers are defined as positive by expression of ER and/or PR receptor equal to 1% or higher of invasive cancer cells (15). ER and PR receptors are expressed around 80 and 65% of breast cancers, respectively (16). Although estrogen receptor-positive tumors co-express PR in the majority of breast cancers, some cases are ER+/PR– and less frequently, ER–/PR+. The response to hormonal therapy seems to be major in breast tumors with positivity for ER and PR, with lower rates in ER+/PR– and ER–/PR+ tumors (11).

Approximately 15–20% of BC has HER2 overexpression and/or amplification, and over 50% of these co-expressing hormone receptors (13, 17). These tumors are called HER2+ BC. The remaining, with negativity for hormonal receptors and HER2, are denominated triple-negative breast cancers. A fourth protein marker, the androgen receptor (AR), is immunoexpressed in 60–80% of breast cancers, with similar proportions to prostate tumors, and specially expressed in HER2+ and triple-negative breast tumors. However, its determination is still not justified in clinical practice as there is no targeted treatment approved for this marker. Other biomarkers with heterogeneous expression include the *epidermal growth factor receptor* (EGFR), p53, c-myc, and proliferation markers such as Ki-67 (14, 18). Ki-67 is a nuclear protein, expressed in all phases of the cell cycle except G0, and a cellular

marker of proliferation with prognostic and predictive value (16, 19).

Even so, this current and basic classification of human breast tumors presents a number of important limitations. The main one is the variability in therapeutic response and clinical outcomes, even for tumors with similar clinical and pathological features. Secondly, this classification provides limited knowledge into the biology and the molecular pathways that divide the BC in distinct subtypes and stages, stepping away from the personalized treatment paradigm.

## Molecular and Genomic Classification of Breast Cancer

Expression analysis has provided an opportunity to explore comprehensive molecular profiling of BC. Differences in gene expressions patterns display basic alterations in the tumor cell biology and are associated with significant variation in terms of clinical behavior, survival (17, 20–22), and treatment outcomes (23–37). The identification of several molecular subtypes was the first insight into the molecular heterogeneity of the BC (20). Five main intrinsic subtypes have been identified based solely on gene expression patterns using DNA microarrays (20, 22): *luminal A*, *luminal B*, *HER2 overexpressing* or *HER2-enriched* (HER2-E) and *basal like*, with another less characterized group named *normal breast-like*. They are called as “intrinsic subtypes of breast cancer” and they have exposed crucial differences in several aspects. The tumor heterogeneity within hormone receptor-positive breast cancers are encompassed by the *luminal A* and *luminal B* subtypes, with better survival outcomes with respect to the non-luminal intrinsic subtypes. The *luminal B* breast tumor expresses hormonal receptors same as the luminal A subtype, but generally having low PR, high proliferation, high grade and worse response to hormonal therapy. At the molecular level, this subtype seems to be dramatically distinct from luminal A, at levels of gene expression, gene copy, or somatic aberrations. All of these features, confers it worse prognosis than the other luminal intrinsic subtype (5).

In 2009, Parker et al. (25) introduced a gene expression-based test named PAM50, which identifies the intrinsic molecular subtypes in four well-established transcriptional subtypes, through the expression of 50 genes in formalin-fixed paraffin embedded (FFPE) tumor tissues: *luminal A*, *luminal B*, *basal-like*, and *HER2-enriched* (25, 28, 31). The intrinsic subtypes overlap with staining of ER, PR and HER2 protein expression by IHC and complemented with ISH for testing HER2 gene amplification. However, several studies have assessed and compared the classification of breast tumors based on the PAM50 gene expression with the classification based on pathological criteria, and a low concordance rate was found in the majority of these studies (31, 34–42). For example, in a combined analysis of data from several studies including a total of 5,994 independent tumor samples, the discordance rate was found to be present in 30.72% across all patients (43). The majority of these studies performed central assessment of pathology-based biomarkers, which normally shows less discrepancies than local

determination (15). Therefore, the two methods should never be considered the same to identify intrinsic biology of BC.

Nonetheless, the diverse genomic landscape of BC is not completely captured through histopathological or transcriptomic analysis. Changes in gene expression patterns are influenced by the underlying genomic structure, and we have evidence that some features of the intrinsic subtypes can be defined by copy number profiling (5, 29, 44). The development of next-generation sequencing technologies has allowed for the characterization of the mutational landscape of this disease, with the identification of novel cancer genes that found it to be recurrently mutated in BC (6, 36, 45, 46). The relevant of integration of the intrinsic subtype with genomic analysis are highlighted in one of the most complete and important molecular characterization studies that have ever been performed in BC (5). In this study, led by *The Cancer Genome Atlas Project* (TCGA), more than 600 primary tumors were extensively profiling at the DNA (methylation, copy-number alterations, somatic and germline mutations), RNA (i.e., miRNA sequencing and mRNA expression) and protein levels (5) (**Table 1**; **Figure 1**). After the analysis of more than 300 primary tumors, five different data-types were mixed together in a cluster of 10 clusters. The consensus clustering analysis identified four major groups of BC, which were found to be very-well summarize by the four molecular intrinsic subtypes defined by mRNA expression only (47) (**Figure 2**).

Thus, all breast cancers show significant genetic diversity. Inherited variants, represented by the single-nucleotide polymorphisms (SNPs) and copy number variants (CNVs), can have an impact in a germline genetic landscape of the individual and inducing the cancer development. The single-nucleotide variants (mutations) and copy number aberrations (CNAs) are genomic changes at somatic level, thus variations acquired that contribute to the initiation and the dissemination of sporadic breast tumors (48). In a recent study, the authors integrated analysis of both, genomic and transcriptomic data, in 2,000 breast tumors as part of the METABRIC consortium (36) dataset, and proposed an alternative molecular classification (48) (**Table 2**). Germline variants and somatic alterations were found to be linked with changes in gene expressions, and the CNAs reported the greatest variability. Clustering analysis of joint copy number and gene expression data from the cis-associated gene reported 10 new molecular subgroups or *integrative clusters* with the capacity of dividing the main intrinsic subtypes into independent groups. Each integrative clusters are characterized by distinct CNAs, gene expression changes, clinical characteristics and different survival outcomes (48). This extensive heterogeneity, as a result of different cell-of-origins and molecular variations, makes that the response of patients to treatments remained variable and difficult to predict.

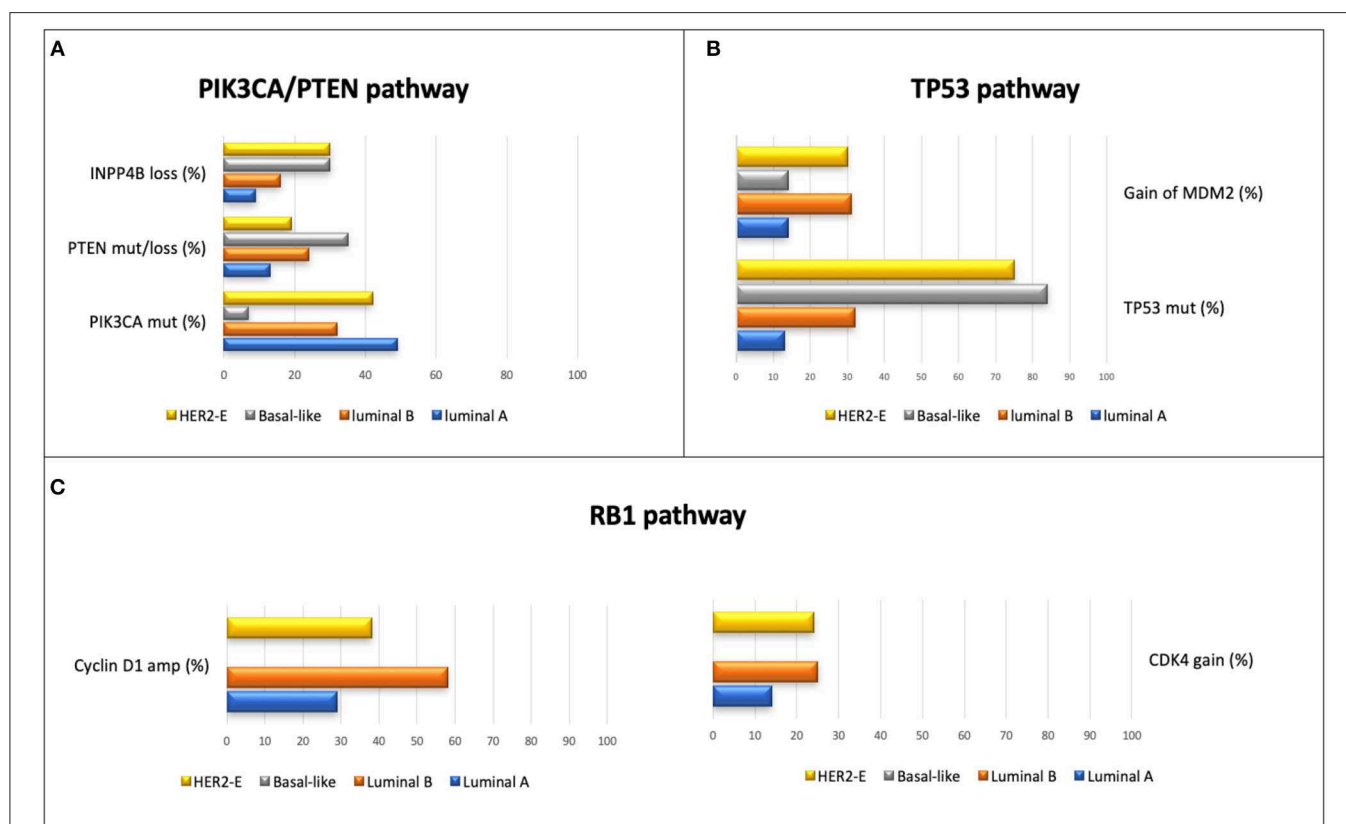
## HER2-POSITIVE BREAST CANCER AND HER2-ENRICHED SUBTYPE

A clear example of complex heterogeneity, inter- and intratumoral, is the HER2+ BC. ERBB2/HER2 is an oncogene coding for a tyrosine kinase receptor that activates oncogenic

**TABLE 1** | Main data about mRNA expression, copy number, DNA mutations and protein expression in the breast cancer tissue samples analyzed in the TCGA project (5).

Subtype/ cluster	Luminal A	Luminal B	Basal-like	HER2-E
mRNA expression	High ER cluster; low proliferation signature	Lower ER cluster; high proliferation signature	Basal-signature; high proliferation	HER2 amplicon signature; high proliferation
Copy number	Most diploid; many with quiet genomes; 1q, 8q, 8q11 gain; 8o, 16q loss, 11q13.3 amp (24%)	Most aneuploidy; many with focal amp; 1q, 8q, 8p11 gain; 8p, 16q loss, 11q13.3 amp (51%); 8p11.23 amp (28%)	Most aneuploidy; high genomic instability; 1q, 10p gain; 8p, 5q loss; MYC focal gain (40%)	Most aneuploidy; high genomic instability; 1q, 8q gain; 8o lossM 17q12 focal ERBB2 amp (71%)
DNA mutations	PIK3CA (49%); TP53 (12%), GATA3 (14%), MAP3K1 (14%)	TP53 (32%); PIK3CA (32%); MAP3K1 (5%)	TP53 (84%); PIK3CA (7%)	TP53 (75%); PIK3CA (42%); PIK3R1 (8%)
Protein expression	High estrogen signaling; high MYB; RPPA reactive subtypes	Less estrogen signaling; high POXM1 and MYC; RPPA reactive subtypes	High expression of DNA repair proteins, PTEN and INPP4B loss signature (pAKT)	High protein and phosphoprotein expression of EGFR and HER2

Amp, amplification; mut, mutation. Percentages are based on 466 tumor samples (463 patients).

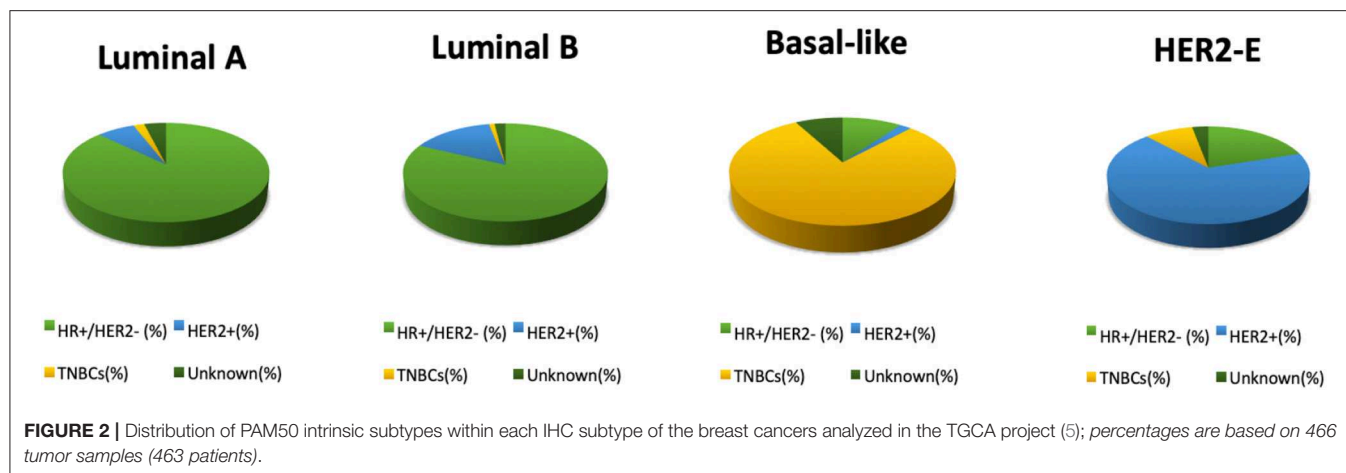


**FIGURE 1** | Principal alterations in most representative pathways in BC, according the intrinsic profiling, analyzed in the TCGA project (5); percentages are based on 466 tumor samples (463 patients). **(A)** Principal alterations in TP53 pathway. **(B)** Principal alterations in PIK3CA/PTEN pathway. **(C)** principal alterations in RB1 pathway. Within the basal-like intrinsic subtype, the main alterations found in this pathway were RB1 mut/loss (20%) and amplification of Cyclin E1 (9%). The expression degree of CDKN2C and RB1 was low and high, respectively, in the luminal A subtype, unlike what was reported in tumors with basal-like subtype.

pathways related with increase proliferation, angiogenesis and invasiveness, resulting in an highly aggressive neoplasm with poor outcomes that others BC (49, 50). The ERBB2/HER2 gene is located in chromosomal region 17q12-21 and its amplification occurs in around 15–20% of breast cancers (10). Overexpression of the protein kinase receptor enables patients with HER2+ BC to benefit from antibody-based and anti-kinase based

therapies that target this receptor, either with a combination of these targeted therapies and chemotherapy, or through dual anti-HER2 therapy without chemotherapy (51–69). This therapeutic approach, has completely changed the prognosis of HER2+ tumors.

So far, the HER2+ BC has been considered as a simple entity. Although the HER2 receptor itself has a dominant



role, and the efficacy of the anti-HER2 agents support it, it is increasing the evidence that HER2 is a phenotype with one of the most extensive and specific heterogeneity (4–6, 70). HER2+ breast cancers vary clearly in their genome variations, gene expression programs, cell-of-origin and cell plasticity, what impact in their microenvironment, prognosis and therapeutic outcomes.

## Immunohistochemistry Criteria: Past, Present, and Future

The HER2 status assessment was established by The *American Society of Clinical Oncology* and the *College of American Pathologists* (ASCO/CAP), with the publication of guidelines with recommendations for testing the level of HER2 protein overexpression by IHC and the HER2 gene amplification determined by ISH, both on FFPE breast tumor tissues. The first ASCO/CAP guideline was published in 2007 (71), and updated in 2013 (72, 73) and 2018 (4) (Table 3). In the last update, the experts refined some controversial criteria of the older guidelines and tried to systematize the testing algorithm for the unusual categories of HER2 ISH results (4) (Table 4). The results of these tests are graded semi-quantitatively as either 0 (negative), 1+ (negative), 2+ (equivocal) or 3+ (positive) by IHC, and classify as amplification (positive, Group 5), equivocal (Group 2,3,4) or negative (Group 1) by ISH. In all of these guidelines, when the HER2 status is negative by IHC and/or ISH, is not indicated the confirmation by an alternate assay. In contrast, the HER2 equivocal cases, by either HER2 IHC or HER2 ISH assays, must be analyzed with an secondary HER2 testing method, or on different tissue blocks with the same testing approach (4, 72). The answer about which of the two methods (IHC or ISH) is better for evaluating the HER2 status, continues to be unknown. Also, with the two latest updates, an important problem was added respecting the 2007 ASCO/CAP guidelines: more HER2 equivocal cases are diagnosed which an increase in reflex HER2 testing (74).

The concordance between HER2 gene status and HER2 protein expression is generally high, even though discordance between IHC and ISH assay is not uncommon. Both methods

detect biological different targets, HER2 protein and HER2 gene expression, respectively, and each assay has its own advantages and disadvantages. The main discordant results are caused by tumor heterogeneity (4, 75–79) focusing mainly in HER2-equivocal cases (4, 73, 76), being a critical factor in the accurate HER2 status evaluation. The ASCO/CAP 2013 guidelines defined heterogeneity as findings of between 5 and 50% of total cells with HER2/CEN17 ratio >2.0 or >6 Her2 signals/cells (72), and the ASCO/CAP 2018 update such as the presence of any aggregated population of amplified cells comprising >10% of the tumor cells on the slide (4). In the low-grade HER2 amplification cases (defined as HER2/CEN17 ratio between 2 and 4) a significant HER2 genetic heterogeneity is detected more frequently than breast cancers with a high-grade HER2 amplification (defined as HER2/CEN17 ratio  $\geq 4.0$ ) and HER2 protein overexpression (defined by IHC 3+) (48, 80). Thus, the evaluation of HER2 through IHC staining and gene amplification, can be remarkably heterogeneous and this could affect the selection of patients, the therapeutic response and the disease-free survival (DFS) rates (76, 81). With an incidence among the studies of 5–40% of HER2 intratumoral heterogeneity (ITH), it cannot be ignored.

HER2 IHC and HER2 ISH tests are employed to select patients for HER2-targeted therapy, and each assay have their advantages and weakness. With the object of improving the assessment of the individual HER2 ITH in tumor samples, Nitta and colleagues, elaborated and validated a protocol in FFPE xenograft tumor tissue sections and in FFPE BC tissue-microarray (TMA) slides, that allows simultaneous brightfield-microscopy detection of HER2 protein and HER2 gene expression, called first *tricolor HER2 gene-protein assay* (GPA) (82). This test exposed the heterogeneity of HER2 protein expression in different BC cells populations (82). A recent study with this assay reported relevant and clinical implications of this intra-heterogeneity (83). Through the combined assessment of HER2 gene amplification and HER2 protein status, five patterns were established. Three of them (type 3 to 5) were defined as a heterogeneous HER2 status and if the tumor case presented any of these types, it related to have ITH. Type 1 (homogeneous HER2 gene amplification and HER2 protein overexpression in all tumor



**TABLE 2 |** Main features of the integrative clusters (48) in 2,000 breast tumors samples.

IntClust	Frequency (n)	Expression (n, %)	Molecular features	PAM50 subtype (n, %)	Prognosis (HR 5, 10 year DSS)
1	139	ER+: 123 (88.5%) PR+: 60 (43%) HER2+: 20 (14.4%)	17q23 amplification High genomic instability	Basal: 9 (6.5%) HER2-E: 21 (15 %) LumA: 11 (7.9%) LumB: 90 (64.8%) Normal: 8 (5.8%)	Intermediate 0.80, 0.69
2	72	ER+: 69 (95.8%) PR+: 51 (70.8%) HER2+: 3 (4.2%)	11q13/14 amplification High genomic instability	Basal: 2 (2.8%) HER2-E: 6 (8.3 %) LumA: 25 (34.7%) LumB: 36 (50%) Normal: 3 (4.2%)	Poor 0.78, 0.51
3	290	ER+: 278 (95.9%) PR+: 211 (72.8%) HER2+: 1 (0.3%)	Paucity of copy number changes Low genomic instability	Basal: 4 (1.4%) HER2-E: 9 (3.1 %) LumA: 195 (67.9%) LumB: 43 (15%) Normal: 36 (12.5%)	Good 0.93, 0.88
4	343	ER+: 238 (69.4%) PR+: 155 (45.2%) HER2+: 20 (5.8%)	CNA devoid Low genomic instability	Basal: 64 (18.7%) HER2-E: 34 (10 %) LumA: 106 (31%) LumB: 29 (8.5%) Normal: 109 (31.9%)	Good 0.89, 0.76
5	190	ER+: 79 (41.6%) PR+: 40 (21.1%) HER2+: 181 (14.4%)	ERBB2 amplification Intermediate genomic instability	Basal: 21 (11%) HER2-E: 108 (56.8 %) LumA: 18 (9.5%) LumB: 33 (17.4%) Normal: 10 (5.3%)	Poor 0.62, 0.45
6	85	ER+: 123 (88.5%) PR+: 60 (43%) HER2+: 20 (14.4%)	8p12 amplification High genomic instability	Basal: 3 (3.5%) HER2-E: 10 (11.8%) LumA: 23 (27.1%) LumB: 43 (50.6%) Normal: 6 (7.1%)	Intermediate 0.83, 0.59
7	190	ER+: 187 (98.4%) PR+: 150 (79%) HER2+: 2 (1.1%)	16p gain, 16q loss, 8q amplification Intermediate genomic instability	Basal: 3 (1.6%) HER2-E: 9 (4.8 %) LumA: 123 (65.1%) LumB: 41 (21.7%) Normal: 13 (6.9%)	Good 0.94, 0.81
8	299	ER+: 297 (99%) PR+: 236 (78.9%) HER2+: 1 (0.3%)	1q gain, 16q loss Intermediate genomic instability	Basal: 1 (0.3%) HER2-E: 9 (3%) LumA: 192 (64.2%) LumB: 89 (29.8%) Normal: 8 (2.7%)	Good 0.88, 0.78
9	146	ER+: 125 (85.6%) PR+: 79 (54.1%) HER2+: 10 (6.9%)	8q gain, 20q amplification High genomic instability	Basal: 20 (13.8%) HER2-E: 26 (18%) LumA: 24 (16.6%) LumB: 70 (48.3%) Normal: 5 (3.5%)	Intermediate 0.78, 0.62
10	226	ER+: 25 (11.1%) PR+: 19 (8.4%) HER2+: 6 (2.7%)	5q loss, 8q gain, 10p gain, 12 p gain Intermediate genomic instability	Basal: 202 (89.4%) HER2-E: 8 (3.5%) LumA: 1 (0.4%) LumB: 14 (6.2%) Normal: 1 (0.4%)	Poor 0.71, 0.68

IntClust, integrative cluster; DSS, disease-specific survival; ER+, estrogen receptor; PR+, progesterone receptor.

cells) and type 2 (homogeneously amplified HER2 gene tumor cells, but without HER2 protein overexpression) were defined as homogenous HER2 status. The type 1 and type 2 were previously reported as “micro-heterogeneity” (42, 84, 85), what can only be detected by GPA. In the final analyses, the HER2 ITH was an independent factor associated with incomplete

pathological response to anti-HER2 neoadjuvant chemotherapy in a cohort of 64 patients (83). Thus, a histopathological-level, a test that allows the recognition of discordance between HER2 gene amplification and protein expression simultaneously, could improve the clinical selection of patients for anti-HER2 therapies, due to a better accuracy of the HER2 IHT in the HER2+ BC.

**TABLE 3 |** 2018 ASCO/CAP summary recommendations [original recommendations and focused update recommendations (4)].

2013 ASCO/CAP recommendations		2018 ASCO/CAP recommendations
<b>HER2 IHC CRITERIA</b>		
Specimens to be tested	All newly diagnosed patients with breast cancer must have a HER2 test performed. Patients who then develop metastatic disease must have a HER2 test performed in a metastatic site, if tissue sample is available.	No change
HER2 score 0 (negative)	No staining is observed or membrane staining that is incomplete and is faint/barely perceptible and within $\leq 10\%$ of tumor cells.	No change
HER2 score 1+ (negative)	Incomplete membrane staining that is faint/barely perceptible and within $> 10\%$ of tumor cells.	No change
HER2 score 2+ (equivocal)	<ul style="list-style-type: none"> <li>Circumferential membrane staining that is incomplete and/or weak/moderate and within <math>&gt; 10\%</math> of tumor cells, or</li> <li>Complete and circumferential membrane staining that is intense and within <math>\leq 10\%</math> of the invasive tumor cells.</li> </ul>	<ul style="list-style-type: none"> <li>Weak to moderate complete membrane staining observed in <math>&gt; 10\%</math> of tumor cells.</li> <li>Basolateral staining for HER2 in a rare subtype of breast cancer with micropapillary histology and circumferential staining that is intense but <math>&lt; 10\%</math> of the tumor cells.</li> </ul>
HER2 score 3+ (positive)	Circumferential membrane staining that is complete, intense, and with $> 10\%$ of tumor cells that must show homogeneous, dark circumferential (chicken wire).	No change
<b>HER2 ISH CRITERIA</b>		
Amplification	Dual-probe Her2/CEP17 ratio $\geq 2.0$ ; with an average Her2 gene copy number $\geq 4.0$ signals/cell (Group 1) Dual-probe Her2/CEP17 ratio $< 2.0$ with an average Her2 gene copy number $\geq 6.0$ signals/cell (Group 3) Dual-probe Her2/CEP17 ratio $\geq 2.0$ with an average Her2 gene copy number $< 4.0$ signals/cell (Group 2)	Dual-probe Her2/CEP17 ratio $\geq 2.0$ ; with an average Her2 gene copy number $\geq 4.0$ signals/cell (Group 1) Dual-probe Her2/CEP17 ratio $< 2.0$ with an average Her2 gene copy number $\geq 6.0$ signals/cell (Group 3) <sup>†</sup> Dual-probe Her2/CEP17 ratio $\geq 2.0$ with an average Her2 gene copy number $< 4.0$ signals/cell (Group 2) <sup>†</sup> If a case has is Group 3 and 2, a definitive diagnosis will be rendered based on additional work-up. If not already assessed by the institution or laboratory performing the ISH test, IHC testing for HER2 should be performed using sections from the same tissue sample used for ISH, and the slides from both ISH and IHC should be reviewed together to guide the selection of areas to score by ISH.
Equivocal	Single-probe average Her2 gene copy $\geq 4.0$ and $\leq 6.0$ signals/cell Dual-probe Her2/CEP17 signal ratio of $< 2.0$ with an average Her2 gene copy number $\geq 4.0$ and $\leq 6.0$ signals/cell (Group 4).	Dual-probe Her2/CEP17 signal ratio of $< 2.0$ with an average Her2 gene copy number $\geq 4.0$ and $\leq 6.0$ signals/cell (Group 4) If a case has an Her2 gene copy $\geq 4.0$ and $< 6.0$ signals/cell ((Group 4) <sup>†</sup> , formerly diagnosed as ISH positive for HER2, a definitive diagnosis will be rendered based on additional work-up. If not already assessed by the institution or laboratory performing the ISH test, IHC testing for HER2 should be performed using sections from the same tissue sample used for ISH, and the slides from both ISH and IHC should be reviewed together to guide the selection of areas to score by ISH.
Non-amplification	Single-probe average Her2 gene copy $< 4.0$ signals/cell Dual-probe Her2/CEP17 signal ratio of $< 2.0$ with an average Her2 gene copy number of $< 4$ signals/cell (Group 5)	No change
<b>Acceptable (IHC and ISH) tests</b>	Should preferentially use an FDA-approved IHC, brightfield ISH, or FISH assay	No change

CAP, College of American Pathologists; CEP17, chromosome enumeration probe 17; ER, estrogen receptor; FDA, US Food and Drug Administration; FISH, fluorescent in situ hybridization; HER2, human epidermal growth factor receptor 2; IHC, immunohistochemistry; ISH, in situ hybridization.

<sup>†</sup> In the 2013 Guideline Update, the work-up of cases in the less common dual-probe ISH categories (groups 2 to 4) include only ISH as additional work-up on diagnosis.

## Molecular Portraits

HER2+ BC has been historically divided in two distinct diseases based on the expression of hormonal receptors, while the gene expression analyses have proved that HER2+ BC is constituted of all the main intrinsic subtypes. In the HR+/HER2+ BC, two intrinsic subtypes are predominantly isolated: *Luminal B* and *HER2-E* (43). Within HR-/HER2+ tumors, around 50–88% have the *HER2-E* subtype, followed by other poor prognostic subtypes such as the *luminal B* or the *basal-like* subtype (41).

The *HER2-E* subtype is defined by high expression of HER2-related and proliferation-related genes of the 17q amplicon (e.g., ERBB2/HER2 and GRB7), an average expression of luminal-related genes (e.g., ESR1, FGFR4, FOXA1, and PGR) and proteins, and by low or missing expression of basal-related genes and proteins (e.g., cytokeratins 5 and 6, OFXC19) (1, 5). At the DNA level, these tumors are characterized by the greatest number of mutations across the genome. About 70–75% and 40% of HER2-E tumors are TP53 and PIK3CA mutated, respectively

**TABLE 4 |** Summary of test result scenarios and recommended final HER2 status (4).

Group	Biology	HER2/CEP17 ratio	HER2 copy number	2018 ASCO/CAP recommendation
1	Classic HER2 amplified cancer <sup>†</sup>	≥2.0	≥4.0	Positive
2	Monosomy 17 <sup>†</sup>	≥2.0	<4.0	Negative, unless concurrent IHC 3+
3	Co-amplification, previously polysomy 17 <sup>†</sup>	<2.0	≥6.0	Negative, unless concurrent IHC 2+ or 3+
4	Borderline/equivocal <sup>†</sup>	<2.0	≥4.0 and <6.0	Negative, unless concurrent IHC 3+
5	Classic HER2 non-amplified cancer <sup>†</sup>	<2.0	<4.0	Negative

<sup>†</sup>Around 95% of breast tumors tested for HER2 by dual-probe ISH correspond to group 1 (HER2 positive) and group 5 (HER2 negative).

<sup>†</sup>The overall prevalence of subgroups 2, 3, and 4 among all breast cancers undergoing HER2 testing is estimated to be about 5%, but within and individual laboratory, the frequency ISH results can be increased.

(5, 6, 44) (**Figure 2**). Thus, any HER2+ BC can be included in the *HER2-E*, *basal-like*, or *luminal* molecular subtypes, and this affect significantly to their biological behavior and therapeutic outcomes. Conversely, the *HER2-E* subtype seems to capture some, but not all clinically HER2+ tumors, while *HER2-E* tumors can be identified within HER2-negative breast tumors, both in hormone receptor-positive or negative profiling (5, 6, 37, 44).

The concept of intrinsic subtypes has provided large insights into the heterogeneity of HER2+ disease. Prat et al. performed an analysis with data of TCGA (5) and METABRIC studies (36) with the purpose to evaluated how molecular subtypes and clinical HER2 status (defined by 2007 ASCO/CAP guidelines and/or DNA copy-number data) overlapped (44). HER2+ BC had a higher frequency of *HER2-E* subtype (47 vs. 7.1% in HER2-negative tumors), with a lower frequency of *luminal A* (10.7 vs. 39%) and *basal-like* subtypes (14.1 vs. 23.4%). Conversely, the ratio of HER2+ BC was 64.6% in *HER2-E* vs. 20, 14.4, and 7.3% in *luminal B*, *basal-like* and *luminal A* subtypes, respectively (44). Among HER2+ and HER2-negative BC, <5% genes were found to be expressed differently within each molecular subtype, and respect to the subtype, the genes significant up-regulated in HER2+ breast cancers, were found enriched for genes located in the 17q12 and 17q21 DNA amplicons. The HER2 gene expression and the expression of other 17q12 amplicon genes, were significantly upper in HER2+ tumors with *HER2-E* and *basal-like* intrinsic subtypes. Finally, after a clustering analysis of a METABRIC dataset of the most variable genes across the four subtypes, the results revealed that overall profile of them is largely maintained regardless of the clinical HER2 status, except for the *HER2-E* subtype (44). Thus, it seems that of gene expression the HER2+ BC of a given subtype is practically indistinguishable from a HER2-negative tumor with the identical subtype, except for the higher expression of genes in or close to the HER2 amplicon on 17q in the HER2+ tumors.

In the study about the ten integrative clusters previously described (48), ERBB2 amplified cancers joined in the integrative Cluster 5 (IntClust), unlike the classification of the intrinsic subtypes of Perou et al. (20), or with the analyses of Prat et al. (44). Several publications, have been compared the prognostic value of the 10 integrative clusters classification in front of the intrinsic subtypes, and the authors concluded that they do not confer supplementary information apart from the provided by the intrinsic subtype (44).

The TCGA dataset study also offers the opportunity to examine additional characteristics of the intrinsic subtype based on HER2 status (5). Through the analysis of protein expression, miRNA, DNA methylation and gene expression, slight molecular differences between HER2+ and HER2-negative tumors within each subtype were detected. The vast majority of proteins up-regulated in HER2+ BC derived again from genes located in the 17qDNA region. After the publication of the TCGA study, the last and distinctive study with a similar approach was published in July 2016 (6). The complex molecular heterogeneity within HER2+ disease was highlighted and explained for the first time by whole-sequencing genome (WGS) and transcriptome sequencing data from HER2+ BC samples (6). The authors selected a total of 289 HER2+ breast cancers with FFPE tissues identified within the French PHRE/SIGNAL programs (86, 87). An overall of 99 selected tumors were analyzed for genome-wide expression portraits, out of which 64 tumors and matched normal DNA were subjected to WGS. On the basis of gene-expression data in an unsupervised hierarchical cluster analysis, four groups were defined with specific genomic alterations (somatic mutations, copy-number changes, and structural alterations). Groups A and B encompassed most HR-positive tumors, and groups C and D mostly contained HR-negative tumors. Using the PAM50 assay to identify the intrinsic subtypes, the tumors were mainly *luminal B* (A and B groups) and *HER2E* (in C and D groups), with only a marginal number of *luminal A* and *basal* tumors (6). These groups displayed specific genomic alterations too. All samples in group D and none in group A displayed mutations in TP53, while only one sample in group D harbored a mutation in PIK3CA, with equal distribution of such mutations in the other groups. A similar gradient, was also observed in terms of genomic and cell of origin transcriptomic signatures (6, 88). Group D showed more genomic instability and a progenitor *luminal* signature. In contrast, group A was more stable and showed a typical mature *luminal* signature (88). These observations are concordant with the cell-of-origin scheme (88–91), in which the intra-tumoral heterogeneity reflects the developmental stage of the epithelial mammary cells. Thus, multiple phenotypes can emerge from one cell-of-origin depending on the initiating genetic event (91).

Thanks to WGS data the authors obtained information about the amplification process itself and about how and maybe when it is arising. The process was consistent with a *breakage-fusion-bridge* (BFB) folding mechanism, supported by the sequence of copy numbers and the orientation of clipped reads (88, 92). However, the present of long distance and inter-chromosomal rearrangements supported that the amplification is a complex

phenomenon, probably comprising multiple amplicons on the same or different chromosomes and several interlaced mechanisms (88). All of this suggests that HER2 amplification, although probably strongly selected, is an embedded event that is superimposed on the standard time course of the breast carcinogenesis (88).

Another relevant article recently published, with genomic and transcriptome analysis too, concluded in a similar theory: HER2 could be defined as a pan-cancer phenomenon (93). The authors explored genomics data (RNA sequencing, expression and copy number changes) across three cohorts of patients [TCGA (5), METABRIC (36) consortium and the USO1062 phase III trial population (94)], with more than 3,000 breast tumors samples analyzed. PAM50 was employed for classifying the intrinsic subtypes. Their results were similar to the previously described: (i) the concordance between HER2 amplification and *HER2-E* subtype was really poorly (only 47% of HER2 amplify tumors presented this intrinsic subtype); (ii) it was find no evidence for cooperating copy number drivers with HER2 outside chromosome 17, and finally (iii) after the transcriptional profiling of the *HER2-E* subtype, the authors reported that HER2+ tumors are hormonally driven, either by ER in hormone receptor-positive and *HER2-E* BC, or by AR in hormone receptor-negative and *HER2-E* BC (93).

## CLINICAL IMPLICATIONS

Trastuzumab was approved in 2001 for metastatic BC patients after the results reported by Slamon et al. (51), in a randomized clinical trial. In adjuvant setting, data from five randomized trials showed a significant improvement in DFS in women with early HER2+ BC after adjuvant treatment with an anti-HER2 antibody called trastuzumab. Latest updates confirmed a benefit sustained over time, resulting finally in a significant improvement in overall survival (OS). In the same way, the treatment with anti-HER2 therapy plus chemotherapy, improved the outcomes in OS in patients with metastatic disease, with numerous randomized clinical trials of anti-HER2 therapies published. To date, the level of expression of HR and HER2 status continue guiding the algorithm of treatment for the HER2+ BC in the clinical practice. Other pathological variables (tumor size, nodal status) provided independent prognostic information. However, if we take into consideration the intrinsic subtype that characterized the tumor, the impact of the clinical and pathological features its decreases considerably.

After the first clinical trial of a HER2-targeted therapy for BC (51), improving strategies to select patients candidate for these therapies has become a critical element to the successful development of anti-HER2 drugs. To date, this selection remaining based on the degree of HER2 positivity in the tumor, by IHC and/or ISH scores (50–53). None biomarker beyond HER2 itself has demonstrated clinical utility across the majority of randomized clinical trials published. Further, although patients with HER2+ disease obtain the greatest benefit from anti-HER2

treatment, the response is greatly heterogeneous, and a substantial proportion of patients present primary or secondary resistance.

The relationship between the grade of HER2 amplification or protein overexpression and the measure of benefit from the different anti-HER2 therapies, has been largely assessed in both early and metastatic disease studies. Available evidence supports a higher probability of success to these therapies in tumors with an increased HER2 protein expression or greater HER2 mRNA levels, although lower HER2 expression or mRNA levels have been associated with clinical benefit too (95–101). Several studies in the neoadjuvant context, have showed an association between rates of pathological complete response (pCR) and a higher HER2 amplification, increased HER2 mRNA levels or HER2 protein overexpression (99–101). In adjuvant studies, such association not impacted either DFS or OS. What's more, centralized laboratory analysis of HER2 testing in the NSABP-B31 (102) and NCCTG N9831 (103) adjuvant trastuzumab trials found a treatment benefit in women with HER2-negative tumors.

Respect to the expression of HR, in the neoadjuvant setting different trials has exhibited heterogeneous response rates after neoadjuvant with chemotherapy and anti-HER2 therapy between hormone receptor-positive and receptor-negative tumors, that is not limited to trastuzumab (10). Achieving a pCR seems to have a significant impact in patient outcomes, with the strongest correlation found in HER2+ BC without expression of hormonal receptors. However, the greatest benefit from anti-HER2 drugs in hormone receptor-negative breast cancers, has not been found in the 3-large adjuvant clinical trials evaluating 1-year of trastuzumab vs. placebo, and both groups seem to obtain similar benefits (103).

Another example that confirm the clinical impact of the HER2 heterogeneity is a phase II study led by the *Danna-Farber Cancer Institute* and recently presented at ASCO 2019 (104). In this clinical trial, the patients received neoadjuvant treatment with 6 cycles of T-DM1 plus pertuzumab. The authors assessment the heterogeneity in basal time (by baseline ultrasound-guide core biopsies from two distinct areas of each tumor), and this entity was defined as at least one of the six areas with either (1) HER2 positivity by ISH in more than 5% and <50% of tumor cells, or (2) a tumoral area with negative result for HER2. Among the 164 patients included, the heterogeneity in HER2 was identify in 10% of evaluable cases without any pCR among cases classified as heterogeneous, being the Residual Cancer Burden (RCB) III the pathological response more frequent in these patients. Secondary analysis also demonstrated a significant relation between pCR (or RCB-0) and HER2 3+ vs. HER2 2+ by IHC. The association between heterogeneity and pCR remained significant when adjusted by hormone receptor status and HER2 IHC measurement (104). These findings, as well as those previously described by Nitta et al. (83), confirm that the ITH is a distinct entity, more diverse than we could expect with the classic pathological evaluation. The heterogeneity in HER2+ BC exist and the treatment of these patients only with anti-HER2 therapies can be insufficient. This entity may need treated with chemotherapy plus anti-HER2 drugs and with novel treatment approaches.



## Sensitivity to Anti-HER2 Based Chemotherapy

The impact of the intrinsic subtyping has been researched retrospectively, either trials evaluating anti-HER2 based chemotherapy in the neoadjuvant [i.e., NeoALTTO (105), CALGB-40601 (84), NOAH (42), CHER-LOB (85) and BERENICE (106)] and adjuvant [i.e., NSABP-B31 (107) and N9831 (108)] settings. Again, in all of these analyses the impact of the HR status, HER2 amplification or the HER2 expression at the protein or mRNA levels, fall into a second or third level such as predictive biomarkers with respect to the intrinsic subtype. In the neoadjuvant setting, when HER2+ BC were classified by PAM50 molecular assay, *HER2-E* subtype was associated with a higher pCR rate (exceeding 50% in all trials) and DFS rates compared to *non-HER2-E* subtypes, following either trastuzumab plus chemotherapy treatment (42, 84, 85) or with dual HER2 blockade without chemotherapy.

## Sensitivity to Dual HER2 Blockade-Only

Nowadays, an area with great interest for the oncologist community is to identify what patients might be treated with a regimen based on dual HER2 blockade without chemotherapy. It has been presented results of several neoadjuvant studies, which submit that a subgroup of patients with HER2+ BC are especially sensitive to the dual HER2 blockade, achieves pCR rates around 70%, so that could potentially be treated without chemotherapy (109).

The *HER2-E* breast tumors are driven by HER2/EGFR signaling, such as it showed, through a silico and omic analyses, in the TCGA breast cancer project (5). So, this intrinsic subtype should benefit the most from anti-HER2 dual-blockade. The benefit achieved in HER-negative BC with *HER2-E* intrinsic subtype can be explained because these tumors preserve the higher expression of EGFR, with independence of expression degree of hormonal receptors (7). However, the greater response rate in the *HER2-E* subtype in previous studies could not distinguish anti-HER2 sensitivity vs. cytotoxic therapy-sensitivity. *HER2-E* subtype could be a predictor itself of anti-HER2 therapy benefit, and this theory should be validated in future randomized trials. If this happened, this intrinsic subtype could help to select a group of patients with HER2+ BC that might be cured with anti-HER2 drugs without chemotherapy, or patients with metastatic disease that can be treated with less intensive treatment, such as dual HER2 blockade-only.

## Immune Infiltration

The *tumor-infiltrating lymphocytes (TILs)* are white bloodstream cells that migrate toward the tumor. In this heterogeneous group of cells, we have found several types of white cells, including *T cells*, *B cells*, and even *Natural-Killer (NK)* cells, although the *T cells* are the most representative. Overall, TILs comprising the majority of mononuclear immune infiltrates from the innate and adaptive immune response, with rates that depending of tumor type and stage. An important feature of these cells is that their functions changes dynamically, throughout tumor progression and in response to oncology treatments, being able to acquire dramatically opposite functions. The TILs represent

pre-existing anti-tumor immunity, with prognostic relevance and predictive value in BC, especially for HER2+ and triple negative breast cancers (110, 111), although the BC has not classically been considered as an immunogenic neoplasm. In contrast to mucosal tissues, normal breast tissue contain limited aggregates of immune cells (112).

In HER2+ BC patients, the TILs are linked to favorable long-term prognosis and survival outcomes, both on early (110, 113–115) and metastatic disease (116). Within HER2+ BC, non-luminal subtypes have the highest levels of TILs, especially the *HER2-E* intrinsic subtype (7, 117). This has been associated with higher rates of pCR and better survival outcomes following chemotherapy and anti-HER2 neoadjuvant treatment (118), and with response to immunotherapy, as suggest the results from the PANACEA phase IB/II trial (119).

However, in multivariable models adjusted for PAM50 subtypes, TILs seems lost their significant association with better outcomes, due that the intrinsic subtype profiling appears encompasses the information provided by TILs (7). Thus, if immunotherapy aspires to obtain relevance in the treatment of BC, future trials should explore these new therapies according to the intrinsic subtype, especially in the HER2+ BC.

## Therapeutic Resistance

Different resistance mechanisms to anti-HER2 therapy have been described, which mostly favoring the reactivation of the HER2 pathway or its downstream signaling (109, 120). Most of the therapeutic failures in the treatment of HER2+ BC come from acquired resistance by sub-clones of cells that are highly selected by the therapeutic pressure. The real prevalence and clinical impact of these mechanisms remain largely unclear, majority of them involve genetic or epigenetic aberrations, and have been mainly described in relation to single HER2 blockade (120). Therefore, these mechanisms should clearly be reviewed, because antiHER2 combinations could select different alterations respect to single HER2 blockade.

Among the main mechanisms described we have (1) an incomplete blockade of the HER2 receptor with the activation of compensatory mechanisms by the HER2 receptors family; (2) the activation of alternative receptor tyrosine kinases (RTKs) or other membrane receptors outside of the HER2 family [such as *insulin-like growth factor 1 receptor (IGF-1R)*, *AXL Receptor Tyrosine Kinase (AXL)* or *MET* (121)] and (3) the alterations in downstream signaling pathways, especially in the PI3K/AKT/mTOR axis. The hyperactivation of PI3K/AKT/mTOR pathway is the best characterized and seems to be the alteration most important to initiate and perpetuate the resistance to anti-HER2 therapies in HER2+ tumors with any degree of the hormone receptor expression (120). Activating mutations in PIK3CA (122) or reduced levels of tumor suppressor genes (mutations or loss of PTEN, and loss of INPP4-B, among others) are the main molecular alterations than maintain this hyperactivation. The role of targeting these pathways has been evaluated in numerous randomized clinical trials. Among these trials, we have the BOLERO-1 and BOLERO-3, both evaluating the role of everolimus, an mTOR inhibitor, in combination with trastuzumab plus paclitaxel as



first-line treatment (BOLERO-1) (63) or in combination with trastuzumab and vinorelbine in trastuzumab-resistant advanced HER2-positive BC (BOLERO 3) (123). The results of them were disappointing and the increase in toxicity very significant. The most relevant data of both studies comes from the combined biomarker analyses that reported an improvement in PFS for patients that harboring PIK3CA mutations or PTEN loss and were treated with everolimus (124). Current efforts have focused on evaluating the activity, in combination with anti-HER2 treatment, of pan-PI3K and alpha-specific PI3K inhibitors. Until now the alpha-specific PI3K inhibitors are the drugs with the most promising therapeutic results and less incidence of serious toxicities respects the pan-PI3K inhibitors (125). These drugs, such as alpelisib (BYL719) (126), target the PI3K- $\alpha$  protein, the most frequently altered PI3K isoform in solid tumors and breast cancers, encoded by the PIK3CA gene and with a prominent role in PI3K signaling.

Another relevant mechanism recently proposed is related to the activity of the cyclinD1-cyclin-dependent kinase 4/6 (CDK 4/6) axis. Their enhanced activation can be driven by cyclin D1/CDK4 overexpression or CD4 mutations, causing resistance to hormonal treatment in hormone receptor-positive breast cancers (127). We already have preclinical evidence (128, 129) and from controlled trials (130) about the role of Cyclin D1-CDK 4/6 axis in the anti-HER2 resistance. Using transgenic mouse models, Goel et al. (128) showed that the suppression of CDK4 activity reduces TSC2 (tuberin) phosphorylation, with a partial suppression of mTORC1 and, hence p70-S6K activity, which relieve feedback inhibition of EGFR family kinases rendering cells more sensitive to the effects of EGFR/HER2 inhibitors and overcome acquired resistance to anti-HER2 treatment. The chemotherapy-free trastuzumab-pertuzumab-palbociclib-fulvestrant combination tested in neoadjuvant setting, has recently exhibited promising activity in terms of reduction of ki67 and rate of pCR for breast tumors with positivity of HER2+ and hormonal receptors (130). So, the combination of CKD4/6 and HER2 inhibitors could be a valid option to chemotherapy-containing regimens, at least in a subgroup of patients.

So, hypothetically, the vast majority of resistance mechanisms described could be targeted by drugs that are already available, such as inhibitors of ER, cyclins, mTOR or FGFR1 (109, 120). However, the potential therapeutic advantage of combining these drugs with standard anti-HER2 therapy should be weighed against the potential risk of serious toxicities. Moreover, as a result of intra- and inter-tumoral heterogeneity, different mechanisms can co-exist in a same patient, keeping the potential possibility to contemporaneously target all resistant tumor clones. The *HER2-E* is the second intrinsic subtype, after the *luminal A*, with greater percentage of aberrations in PI3K/AKT/mTOR axis (by PI3KCA mutations/loss of PTEN) and alterations in RB1 pathway (by Cyclin D1 amplification and/or CDK4 gain). So, if both axis are implicated in the resistance of anti-HER2 treatment, this intrinsic subtype could be the most appropriated to design future clinical trials that testing the role of targeting all pathways simultaneously and to prevent of development of acquired resistances, independently of pathological evaluation of HER2, which does not seem to

adequately measure the ITH, an entity already established as other potential resistance mechanism.

## CONCLUSION

To date, amplification and/or overexpression of HER2 remains the only biomarker regarding treatment decisions with anti-HER2 drugs, but it is insufficient itself to clarify the heterogeneous therapeutic outcomes. The complex heterogeneity of the HER2+ BC is a critical aspect, as it has been described at multiple levels: intra-tumoral, at gene expression, transcriptomic and genomic levels. The HER2+ BC do not represent a subtype itself, but are instead dispersed along the whole breast cancer spectrum, from hormone receptor-positive luminal to hormone receptor-negative basal phenotype, with genome variations accordingly to these phenotypes and incidentally defined by a specific gene amplification. Perhaps, combining phenotypic (i.e., gene expression groups) and mechanistic (i.e., co-amplifications) characteristics, may improve the actual classification of HER2+ BC, with the identification of more homogeneous subgroups and improving the knowledge of the genetic mechanisms implicated in the heterogeneity of this disease. This could lead to rational therapeutic strategies, exploring additional pathways and genes co-amplified with ERBB2, especially relevant for patients who show an initial weak response or that exhibit treatment resistance, patients with a particularly poor prognosis.

Although HER2 amplification is traditionally associated with *HER2-E* transcriptional subtype, these are substantially distinct. HER2 amplification seems an oncogenic driver present in all subtypes in place of a biomarker itself of an intrinsic subtype, and its strong enrichment in the *HER2-E* subtype has masked the nature of this entity. Taking into consideration only the intrinsic subtype, any prognostic value attributable to clinical and pathological variables such as the degree, ER/PR or HER2 status by IHC and/or ISH, disappears, as happens with the amplification of HER2 isolated taken as a predictive factor itself.

We already have data of efficacy for anti-HER2 therapy in patients with HER2-negative tumors, with a considerable proportion of patients with HER2+ breast cancers not achieving such clinical benefit. Overall, the evidence so far suggests that all BC with *HER2-E* intrinsic subtype benefit from anti-HER2 treatment. Although much remains to be done, with the data available and presented in this review, it seems that the *HER2-E* intrinsic subtype would be a more appropriate biomarker to assess the real benefit of anti-HER2 treatment in all phenotypes of BC.

Respecting the actual HER2+ BC therapeutic setting, the most recent studies try to improve the results of patients adding new anti-HER2 drugs, still without a selection by molecular features, thus, achieving a discrete therapeutic benefit in the most of these trials, and increasing toxicity and costs. The intrinsic molecular subtyping of BC fairly has extended our knowledge about the behavior of this tumor and should have an established place in the clinical practice. After this revision, we would like to conclude that the *HER2-E* subtype should be established itself as the best predictor of prognosis and clinical outcomes of the BC with this intrinsic subtype, what would allow for the extension of the use of anti-HER2 drugs for HER2-negative tumors and to improve

the selection of patients with HER2+ BC for combination of anti-HER2 therapies.

## AUTHOR CONTRIBUTIONS

AG-O: authorship and complete writing of the manuscript. AS-M and AR: general review of the manuscript and

contribution of bibliography. MC: help with the interpretation of molecular data. MP: contribution and bibliographical guidance and has helped too with the interpretation of molecular data. NE: contribution and bibliographical guidance. EC: general review, partial correction of the manuscript, and contribution of bibliography. All authors read and approved the final manuscript.

## REFERENCES

- Cejalvo JM, de Dueñas EM, Galván P, García-Recio S, Gasión OB, Paré L, et al. Intrinsic subtypes and gene expression profiles in primary and metastatic breast cancer. *Cancer Res.* (2017) 77:2213–21. doi: 10.1158/0008-5472.CAN-16-2717
- Sánchez-Muñoz A, Vicioso L, Santonja A, Álvarez M, Plata-Fernández Y, Miramón J, et al. Male breast cancer: correlation between immunohistochemical subtyping and PAM50 intrinsic subtypes, and the subsequent clinical outcomes. *Mod Pathol.* (2018) 31:299–306. doi: 10.1038/modpathol.2017.129
- Harbeck N. Insights into biology of luminal HER2 vs. enriched HER2 subtypes: therapeutic implications. *Breast.* (2015) 24:S44–8. doi: 10.1016/j.breast.2015.07.011
- 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. doi: 10.1200/JCO.2018.77.8738
- Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature.* (2012) 490:61–70. doi: 10.1038/nature11412
- Ferrari A, Vincent-Salomon A, Pivot X, Sertier A-S, Thomas E, Tonon L, et al. A whole-genome sequence and transcriptome perspective on HER2-positive breast cancers. *Nat Commun.* (2016) 7:12222. doi: 10.1038/ncomms12222
- Cejalvo JM, Pascual T, Fernández-Martínez A, Brasó-Maristany F, Gomis RR, Perou CM, et al. Clinical implications of the non-luminal intrinsic subtypes in hormone receptor-positive breast cancer. *Cancer Treat Rev.* (2018) 67:63–70. doi: 10.1016/j.ctrv.2018.04.015
- Giuliano AE, Edge SB, Hortobagyi GN. Eighth edition of the AJCC cancer staging manual: breast cancer. *Ann Surg Oncol.* (2018) 25:1783–5. doi: 10.1245/s10434-018-6486-6
- Elston CW, Ellis IO, Pinder SE. Pathological prognostic factors in breast cancer. *Crit Rev Oncol Hematol.* (1999) 31:209–23. doi: 10.1016/S1040-8428(99)00034-7
- Loibl S, Gianni L. HER2-positive breast cancer. *Lancet.* (2017) 389:2415–29. doi: 10.1016/S0140-6736(16)32417-5
- Lakhani SR, Ellis IO, Schnitt SJ, Tan PH, van de Vijver MJ. *WHO Classification of Tumours of the Breast.* Available online at: <http://publications.iarc.fr/Book-And-Report-Series/Who-Iarc-Classification-Of-Tumours/WHO-Classification-Of-Tumours-Of-The-Breast-2012> (accessed March 14, 2019).
- Weigelt B, Geyer FC, Reis-Filho JS. Histological types of breast cancer: how special are they? *Mol Oncol.* (2010) 4:192–208. doi: 10.1016/j.molonc.2010.04.004
- Harris L, Fritzsche H, Mennel R, Norton L, Ravdin P, Taube S, et al. American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. *J Clin Oncol.* (2007) 25:5287–312. doi: 10.1200/JCO.2007.14.2364
- Rakha EA, Reis-Filho JS, Baehner F, Dabbs DJ, Decker T, Eusebi V, et al. Breast cancer prognostic classification in the molecular era: the role of histological grade. *Breast Cancer Res.* (2010) 12:207. doi: 10.1186/bcr2607
- Hammond MEH, Hayes DF, Dowsett M, Allred DC, Hagerty KL, Badve S, et al. American Society of Clinical Oncology/College of American Pathologists Guideline Recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. *J Clin Oncol.* (2010) 134:e48–72. doi: 10.1200/JOP.777003
- de Azambuja E, Cardoso F, de Castro G, Colozza M, Mano MS, Durbecq V, et al. Ki-67 as prognostic marker in early breast cancer: a meta-analysis of published studies involving 12,155 patients. *Br J Cancer.* (2007) 96:1504–13. doi: 10.1038/sj.bjc.6603756
- Wirapati P, Sotiriou C, Kunkel S, Farmer P, Pradervand S, Haibe-Kains B, et al. Meta-analysis of gene expression profiles in breast cancer: toward a unified understanding of breast cancer subtyping and prognosis signatures. *Breast Cancer Res.* (2008) 10:R65. doi: 10.1186/bcr2124
- Turashvili G, Brogi E. Tumor heterogeneity in breast cancer. *Front Med.* (2017) 4:227. doi: 10.3389/fmed.2017.00227
- Yerushalmi R, Woods R, Ravdin PM, Hayes MM, Gelmon KA. Ki67 in breast cancer: prognostic and predictive potential. *Lancet Oncol.* (2010) 11:174–83. doi: 10.1016/S1470-2045(09)70262-1
- Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. *Nature.* (2000) 406:747–52. doi: 10.1038/35021093
- Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA.* (2001) 98:10869–74. doi: 10.1073/pnas.191367098
- Sørlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci USA.* (2003) 100:8418–23. doi: 10.1073/pnas.0932692100
- Carey LA, Perou CM, Livasy CA, Dressler LG, Cowan D, Conway K, et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. *JAMA.* (2006) 295:2492–502. doi: 10.1001/jama.295.21.2492
- Rouzier R, Perou CM, Symmans WF, Ibrahim N, Cristofanilli M, Anderson K, et al. Breast cancer molecular subtypes respond differently to preoperative chemotherapy. *Clin Cancer Res.* (2005) 11:5678–85. doi: 10.1158/1078-0432.CCR-04-2421
- Parker JS, Mullins M, Cheang MCU, Leung S, Voduc D, Vickery T, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol.* (2009) 27:1160–7. doi: 10.1200/JCO.2008.18.1370
- Hugh J, Hanson J, Cheang MCU, Nielsen TO, Perou CM, Dumontet C, et al. Breast cancer subtypes and response to docetaxel in node-positive breast cancer: use of an immunohistochemical definition in the BCIRG 001 trial. *J Clin Oncol.* (2009) 27:1168–76. doi: 10.1200/JCO.2008.18.1024
- Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res.* (2010) 12:R68. doi: 10.1186/bcr2635
- Nielsen TO, Parker JS, Leung S, Voduc D, Ebbert M, Vickery T, et al. A comparison of PAM50 intrinsic subtyping with immunohistochemistry and clinical prognostic factors in tamoxifen-treated estrogen receptor-positive breast cancer. *Clin Cancer Res.* (2010) 16:5222–32. doi: 10.1158/1078-0432.CCR-10-1282
- Prat A, Ellis MJ, Perou CM. Practical implications of gene-expression-based assays for breast oncologists. *Nat Rev Clin Oncol.* (2011) 9:48–57. doi: 10.1038/nrclinonc.2011.178
- Cheang MCU, Voduc KD, Tu D, Jiang S, Leung S, K. Chia S, et al. Responsiveness of intrinsic subtypes to adjuvant anthracycline substitution in the NCIC.CTG MA.5 randomized trial. *Clin Cancer Res.* (2012) 18:2402–12. doi: 10.1158/1078-0432.CCR-11-2956

31. Chia SK, Bramwell VH, Tu D, Shepherd LE, Jiang S, Vickery T, et al. A 50-gene intrinsic subtype classifier for prognosis and prediction of benefit from adjuvant tamoxifen. *Clin Cancer Res.* (2012) 18:4465–72. doi: 10.1158/1078-0432.CCR-12-0286
32. Esserman LJ, Berry DA, Cheang MCU, Yau C, Perou CM, Carey L, et al. Chemotherapy response and recurrence-free survival in neoadjuvant breast cancer depends on biomarker profiles: results from the I-SPY 1 TRIAL (CALGB 150007/150012; ACRIN 6657). *Breast Cancer Res Treat.* (2012) 132:1049–62. doi: 10.1007/s10549-011-1895-2
33. Prat A, Cheang MCU, Martin M, Parker JS, Carrasco E, Caballero R, et al. Prognostic significance of progesterone receptor-positive tumor cells within immunohistochemically defined luminal A breast cancer. *J Clin Oncol.* (2013) 31:203–9. doi: 10.1200/JCO.2012.43.4134
34. Ellis MJ, Suman VJ, Hoog J, Lin L, Snider J, Prat A, et al. Randomized phase II neoadjuvant comparison between letrozole, anastrozole, and exemestane for postmenopausal women with estrogen receptor-rich stage 2 to 3 breast cancer: clinical and biomarker outcomes and predictive value of the baseline PAM50-based intrinsic subtype-ACOSOG Z1031. *J Clin Oncol.* (2011) 29:2342–9. doi: 10.1200/JCO.2010.31.6950
35. Dowsett M, Sestak I, Lopez-Knowles E, Sidhu K, Dunbier AK, Cowens JW, et al. Comparison of PAM50 risk of recurrence score with oncotype DX and IHC4 for predicting risk of distant recurrence after endocrine therapy. *J Clin Oncol.* (2013) 31:2783–90. doi: 10.1200/JCO.2012.46.1558
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. doi: 10.1038/nature10983
37. 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. doi: 10.1093/annonc/mdx048
38. Bastien RRL, Rodríguez-Lescure Á, Ebbert MTW, Prat A, Munárriz B, Rowe L, et al. PAM50 breast cancer subtyping by RT-qPCR and concordance with standard clinical molecular markers. *BMC Med Genomics.* (2012) 5:44. doi: 10.1186/1755-8794-5-44
39. Carey LA, Berry DA, Ollila D, Harris L, Krop IE, Weckstein D, et al. Clinical and translational results of CALGB 40601: a neoadjuvant phase III trial of weekly paclitaxel and trastuzumab with or without lapatinib for HER2-positive breast cancer. *JCO.* (2013) 31:500. doi: 10.1200/jco.2013.31.15\_suppl.500
40. Gnant M, Filipits M, Greil R, Stoeger H, Rudas M, Bago-Horvath Z, et al. Predicting distant recurrence in receptor-positive breast cancer patients with limited clinicopathological risk: using the PAM50 Risk of Recurrence score in 1478 postmenopausal patients of the ABCSG-8 trial treated with adjuvant endocrine therapy alone. *Ann Oncol.* (2014) 25:339–45. doi: 10.1093/annonc/mdt494
41. Prat A, Perou CM. Deconstructing the molecular portraits of breast cancer. *Mol Oncol.* (2011) 5:5–23. doi: 10.1016/j.molonc.2010.11.003
42. Prat A, Bianchini G, Thomas M, Belousov A, Cheang MCU, Koehler A, et al. Research-based PAM50 subtype predictor identifies higher responses and improved survival outcomes in HER2-positive breast cancer in the NOAH study. *Clin Cancer Res.* (2014) 20:511–21. doi: 10.1158/1078-0432.CCR-13-0239
43. Prat A, Pineda E, Adamo B, Galván P, Fernández A, Gaba L, et al. Clinical implications of the intrinsic molecular subtypes of breast cancer. *Breast.* (2015) 24:S26–35. doi: 10.1016/j.breast.2015.07.008
44. Prat A, Carey LA, Adamo B, Vidal M, Tabernero J, Cortés J, et al. Molecular features and survival outcomes of the intrinsic subtypes within HER2-positive breast cancer. *J Natl Cancer Inst.* (2014) 106: dju152. doi: 10.1093/jnci/dju152
45. Banerji S, Cibulskis K, Rangel-Escareno C, Brown KK, Carter SL, Frederick AM, et al. Sequence analysis of mutations and translocations across breast cancer subtypes. *Nature.* (2012) 486:405–9. doi: 10.1038/nature11154
46. Stephens PJ, Tarpey PS, Davies H, Van Loo P, Greenman C, Wedge DC, et al. The landscape of cancer genes and mutational processes in breast cancer. *Nature.* (2012) 486:400–4. doi: 10.1038/nature11017
47. Pascual T, Martin M, Fernández-Martínez A, Paré L, Alba E, Rodríguez-Lescure Á, et al. A pathology-based combined model to identify PAM50 non-luminal intrinsic disease in hormone receptor-positive HER2-negative breast cancer. *Front Oncol.* (2019) 9:303. doi: 10.3389/fonc.2019.00303
48. Dawson S-J, Rueda OM, Aparicio S, Caldas C. A new genome-driven integrated classification of breast cancer and its implications. *EMBO J.* (2013) 32:617–28. doi: 10.1038/emboj.2013.19
49. Baselga J. Why the epidermal growth factor receptor? The rationale for cancer therapy. *Oncologist.* (2002) 7(Suppl. 4):2–8. doi: 10.1634/theoncologist.7-suppl\_4-2
50. 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. doi: 10.1126/science.3798106
51. Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, et al. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N Engl J Med.* (2001) 344:783–92. doi: 10.1056/NEJM200103153441101
52. Paik S, Bryant J, Tan-Chiu E, Romond E, Hiller W, Park K, et al. Real-world performance of HER2 testing—National Surgical Adjuvant Breast and Bowel Project experience. *J Natl Cancer Inst.* (2002) 94:852–4. doi: 10.1093/jnci/94.11.852
53. Piccart-Gebhart MJ, Procter M, Leyland-Jones B, Goldhirsch A, Untch M, Smith I, et al. Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. *N Engl J Med.* (2005) 353:1659–72. doi: 10.1056/NEJMoa052306
54. Burstein HJ, Harris LN, Marcom PK, Lambert-Falls R, Havlin K, Overmoyer B, et al. Trastuzumab and vinorelbine as first-line therapy for HER2-overexpressing metastatic breast cancer: multicenter phase II trial with clinical outcomes, analysis of serum tumor markers as predictive factors, and cardiac surveillance algorithm. *J Clin Oncol.* (2003) 21:2889–95. doi: 10.1200/JCO.2003.02.018
55. Joensuu H, Kellokumpu-Lehtinen P-L, Bono P, Alanko T, Kataja V, Asola R, et al. Adjuvant docetaxel or vinorelbine with or without trastuzumab for breast cancer. *N Engl J Med.* (2006) 354:809–20. doi: 10.1056/NEJMoa053028
56. Dennis S, Wolfgang E, Nicholas R, Tadeusz P, Miguel M, Michael P, et al. Adjuvant trastuzumab in HER2-positive breast cancer. *N Engl J Med.* (2011) 365:1273–83. doi: 10.1056/NEJMoa0910383
57. Baselga J, Cortés J, Kim S-B, Im S-A, Hegg R, Im Y-H, et al. Pertuzumab plus trastuzumab plus docetaxel for metastatic breast cancer. *N Engl J Med.* (2012) 366:109–19. doi: 10.1056/NEJMoa113216
58. Cameron D, Casey M, Oliva C, Newstat B, Imwalle B, Geyer CE. Lapatinib plus capecitabine in women with HER-2-positive advanced breast cancer: final survival analysis of a phase III randomized trial. *Oncologist.* (2010) 15:924–34. doi: 10.1634/theoncologist.2009-0181
59. Andersson M, Lidbrink E, Bjerre K, Wist E, Enevoldsen K, Jensen AB, et al. Phase III randomized study comparing docetaxel plus trastuzumab with vinorelbine plus trastuzumab as first-line therapy of metastatic or locally advanced human epidermal growth factor receptor 2-positive breast cancer: the HERNATA study. *J Clin Oncol.* (2011) 29:264–71. doi: 10.1200/JCO.2010.30.8213
60. Gianni L, Pienkowski T, Im Y-H, Tseng L-M, Liu M-C, Lluch A, et al. 5-year analysis of neoadjuvant pertuzumab and trastuzumab in patients with locally advanced, inflammatory, or early-stage HER2-positive breast cancer (NeoSphere): a multicentre, open-label, phase 2 randomised trial. *Lancet Oncol.* (2016) 17:791–800. doi: 10.1016/S1470-2045(16)00163-7
61. Rimawi MF, Mayer IA, Forero A, Nanda R, Goetz MP, Rodríguez AA, et al. Multicenter phase II study of neoadjuvant lapatinib and trastuzumab with hormonal therapy and without chemotherapy in patients with human epidermal growth factor receptor 2-overexpressing breast cancer: TBCRC 006. *J Clin Oncol.* (2013) 31:1726–31. doi: 10.1200/JCO.2012.44.8027
62. Perez EA, Romond EH, Suman VJ, Jeong J-H, Sledge G, Geyer CE, et al. Trastuzumab plus adjuvant chemotherapy for human epidermal growth factor receptor 2-positive breast cancer: planned joint analysis of overall survival from NSABP B-31 and NCCTG N9831. *J Clin Oncol.* (2014) 32:3744–52. doi: 10.1200/JCO.2014.55.5730
63. Hurvitz SA, Andre F, Jiang Z, Shao Z, Mano MS, Neciosup SP, et al. Combination of everolimus with trastuzumab plus paclitaxel as first-line treatment for patients with HER2-positive advanced breast cancer



- (BOLERO-1): a phase 3, randomised, double-blind, multicentre trial. *Lancet Oncol.* (2015) 16:816–29. doi: 10.1016/S1470-2045(15)00051-0
64. Gelmon KA, Boyle FM, Kaufman B, Huntsman DG, Manikhas A, Di Leo A, et al. Lapatinib or trastuzumab plus taxane therapy for human epidermal growth factor receptor 2-positive advanced breast cancer: final results of NCIC CTG MA.31. *J Clin Oncol.* (2015) 33:1574–83. doi: 10.1200/JCO.2014.56.9590
  65. Verma S, Miles D, Gianni L, Krop IE, Welslau M, Baselga J, et al. Trastuzumab emtansine for HER2-positive advanced breast cancer. *N Engl J Med.* (2012) 367:1783–91. doi: 10.1056/NEJMoa1209124
  66. Perez EA, Suman VJ, Davidson NE, Gralow JR, Kaufman PA, Visscher DW, et al. Sequential versus concurrent trastuzumab in adjuvant chemotherapy for breast cancer. *J Clin Oncol.* (2011) 29:4491–7. doi: 10.1200/JCO.2011.36.7045
  67. Tolane 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. doi: 10.1056/NEJMoa1406281
  68. Prat A, Baselga J. Dual human epidermal growth factor receptor 2 (HER2) blockade and hormonal therapy for the treatment of primary HER2-positive breast cancer: one more step toward chemotherapy-free therapy. *J Clin Oncol.* (2013) 31:1703–6. doi: 10.1200/JCO.2012.48.4998
  69. Prat A, Baselga J. The role of hormonal therapy in the management of hormonal-receptor-positive breast cancer with co-expression of HER2. *Nat Clin Pract Oncol.* (2008) 5:531–42. doi: 10.1038/ncponc1179
  70. Montemurro F, Di Cosimo S, Arpino G. Human epidermal growth factor receptor 2 (HER2)-positive and hormone receptor-positive breast cancer: new insights into molecular interactions and clinical implications. *Ann Oncol.* (2013) 24:2715–24. doi: 10.1093/annonc/mdt287
  71. Wolff AC, Hammond MEH, Schwartz JN, Hagerty KL, Allred DC, Cote RJ, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *J Clin Oncol.* (2006) 25:118–45. doi: 10.1200/JCO.2006.09.2775
  72. Wolff AC, Hammond MEH, Hicks DG, Dowsett M, McShane LM, Allison KH, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. *J Clin Oncol.* (2013) 31:3997–4013. doi: 10.1200/JCO.2013.50.9984
  73. Bethune GC, Veldhuijzen van Zanten D, MacIntosh RF, Rayson D, Younis T, Thompson K, et al. Impact of the 2013 American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 (HER2) testing of invasive breast carcinoma: a focus on tumours assessed as “equivocal” for HER2 gene amplification by fluorescence in-situ hybridization. *Histopathology.* (2015) 67:880–7. doi: 10.1111/his.12723
  74. Varga Z, Noske A. Impact of modified 2013 ASCO/CAP guidelines on HER2 testing in breast cancer. One year experience. *PLoS ONE.* (2015) 10:e0140652. doi: 10.1371/journal.pone.0140652
  75. Nitta H, Kelly BD, Allred C, Jewell S, Banks P, Dennis E, et al. The assessment of HER2 status in breast cancer: the past, the present, and the future. *Pathol Int.* (2016) 66:313–24. doi: 10.1111/pin.12407
  76. Seol H, Lee HJ, Choi Y, Lee HE, Kim YJ, Kim JH, et al. Intratumoral heterogeneity of HER2 gene amplification in breast cancer: its clinicopathological significance. *Mod Pathol.* (2012) 25:938–48. doi: 10.1038/modpathol.2012.36
  77. Hou Y, Nitta H, Li Z. HER2 gene protein assay is useful to determine HER2 status and evaluate HER2 heterogeneity in HER2 equivocal breast cancer. *Am J Clin Pathol.* (2017) 147:89–95. doi: 10.1093/ajcp/aqw211
  78. Hanna WM, Rüschhoff J, Bilous M, Coudry RA, Dowsett M, Osamura RY, et al. HER2 *in situ* hybridization in breast cancer: clinical implications of polysomy 17 and genetic heterogeneity. *Mod Pathol.* (2014) 27:4–18. doi: 10.1038/modpathol.2013.103
  79. Yang Y-L, Fan Y, Lang R-G, Gu F, Ren M-J, Zhang X-M, et al. Genetic heterogeneity of HER2 in breast cancer: impact on HER2 testing and its clinicopathologic significance. *Breast Cancer Res Treat.* (2012) 134:1095–102. doi: 10.1007/s10549-012-2046-0
  80. Brunelli M, Manfrin E, Martignoni G, Miller K, Remo A, Reghellin D, et al. Genotypic intratumoral heterogeneity in breast carcinoma with HER2/neu amplification: evaluation according to ASCO/CAP criteria. *Am J Clin Pathol.* (2009) 131:678–82. doi: 10.1309/AJCP09VUTZWZXB MJ
  81. Bartlett AI, Starczyński J, Robson T, Maclellan A, Campbell FM, van de Velde CJH, et al. Heterogeneous HER2 gene amplification: impact on patient outcome and a clinically relevant definition. *Am J Clin Pathol.* (2011) 136:266–74. doi: 10.1309/AJCP09EN6AQMWETZZ
  82. Nitta H, Kelly BD, Padilla M, Wick N, Brunhoeber P, Bai I, et al. A gene-protein assay for human epidermal growth factor receptor 2 (HER2): brightfield tricolor visualization of HER2 protein, the HER2 gene, and chromosome 17 centromere (CEN17) in formalin-fixed, paraffin-embedded breast cancer tissue sections. *Diagn Pathol.* (2012) 7:60. doi: 10.1186/1746-1596-7-60
  83. Hou Y, Nitta H, Wei L, Banks PM, Portier B, Parwani AV, et al. HER2 intratumoral heterogeneity is independently associated with incomplete response to anti-HER2 neoadjuvant chemotherapy in HER2-positive breast carcinoma. *Breast Cancer Res Treat.* (2017) 166:447–57. doi: 10.1007/s10549-017-4453-8
  84. Carey LA, Berry DA, Cirincione 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. doi: 10.1200/JCO.2015.62.1268
  85. Dieci MV, Prat A, Tagliafico E, Paré L, Ficarra G, Bisagni G, et al. Integrated evaluation of PAM50 subtypes and immune modulation of pCR in HER2-positive breast cancer patients treated with chemotherapy and HER2-targeted agents in the CherLOB trial. *Ann Oncol.* (2016) 27:1867–73. doi: 10.1093/annonc/mdw262
  86. Pivot X, Romieu G, Debled M, Pierga J-Y, Kerbrat P, Bachelot T, et al. 6 months versus 12 months of adjuvant trastuzumab for patients with HER2-positive early breast cancer (PHARE): a randomised phase 3 trial. *Lancet Oncol.* (2013) 14:741–8. doi: 10.1016/S1470-2045(13)70225-0
  87. Kramar A, Bachelot T, Madrange N, Pierga J-Y, Kerbrat P, Espié M, et al. Trastuzumab duration effects within patient prognostic subgroups in the PHARE trial. *Ann Oncol.* (2014) 25:1563–70. doi: 10.1093/annonc/mdu177
  88. Ferrari A, Sertier A-S, Vincent-Salomon A, Pivot X, Pauporté I, Saintigny P, et al. A phenotypic and mechanistic perspective on heterogeneity of HER2-positive breast cancers. *Mol Cell Oncol.* (2016) 3:e1232186. doi: 10.1080/23723556.2016.1232186
  89. Visvader JE, Stingl J. Mammary stem cells and the differentiation hierarchy: current status and perspectives. *Genes Dev.* (2014) 28:1143–58. doi: 10.1101/gad.242511.114
  90. Koren S, Bentires-Alj M. Breast tumor heterogeneity: source of fitness, hurdle for therapy. *Mol Cell.* (2015) 60:537–46. doi: 10.1016/j.molcel.2015.10.031
  91. Skibinski A, Kuperwasser C. The origin of breast tumor heterogeneity. *Oncogene.* (2015) 34:5309–16. doi: 10.1038/ncr.2014.475
  92. Greenman CD, Pleasance ED, Newman S, Yang F, Fu B, Nik-Zainal S, et al. Estimation of rearrangement phylogeny for cancer genomes. *Genome Res.* (2012) 22:346–61. doi: 10.1101/gr.118414.110
  93. Daemen A, Manning G. HER2 is not a cancer subtype but rather a pan-cancer event and is highly enriched in AR-driven breast tumors. *Breast Cancer Res.* (2018) 20:8. doi: 10.1186/s13058-018-0933-y
  94. Wilson TR, Yu J, Lu X, Spoerke JM, Xiao Y, O'Brien C, et al. The molecular landscape of high-risk early breast cancer: comprehensive biomarker analysis of a phase III adjuvant population. *NPJ Breast Cancer.* (2016) 2:16022. doi: 10.1038/npjbcancer.2016.22
  95. Baselga J, Cortés J, Im S-A, Clark E, Ross G, Kiermaier A, et al. Biomarker analyses in CLEOPATRA: a phase III, placebo-controlled study of pertuzumab in human epidermal growth factor receptor 2-positive, first-line metastatic breast cancer. *J Clin Oncol.* (2014) 32:3753–61. doi: 10.1200/JCO.2013.54.5384
  96. Baselga J, Lewis Phillips GD, Verma S, Ro J, Huober J, Guardino AE, et al. Relationship between tumor biomarkers and efficacy in EMILIA, a phase III study of trastuzumab emtansine in HER2-positive metastatic breast cancer. *Clin Cancer Res.* (2016) 22:3755–63. doi: 10.1158/1078-0432.CCR-15-2499
  97. Kim S-B, Wildiers H, Krop IE, Smitt M, Yu R, Lysbet de Haas S, et al. Relationship between tumor biomarkers and efficacy in TH3RESA, a phase

- III study of trastuzumab emtansine (T-DM1) vs. treatment of physician's choice in previously treated HER2-positive advanced breast cancer. *Int J Cancer*. (2016) 139:2336–42. doi: 10.1002/ijc.30276
98. Schneeweiss A, Chia S, Hegg R, Tausch C, Deb R, Ratnayake J, et al. Evaluating the predictive value of biomarkers for efficacy outcomes in response to pertuzumab- and trastuzumab-based therapy: an exploratory analysis of the TRYPHAEENA study. *Breast Cancer Res.* (2014) 16:R73. doi: 10.1186/bcr3690
99. Denkert C, Huober J, Loibl S, Prinzler J, Kronenwett R, Darb-Esfahani S, et al. HER2 and ESR1 mRNA expression levels and response to neoadjuvant trastuzumab plus chemotherapy in patients with primary breast cancer. *Breast Cancer Res.* (2013) 15:R11. doi: 10.1186/bcr3384
100. Scaltriti M, Nuciforo P, Bradbury I, Sperinde J, Agbor-Tarh D, Campbell C, et al. High HER2 expression correlates with response to the combination of lapatinib and trastuzumab. *Clin Cancer Res.* (2015) 21:569–76. doi: 10.1158/1078-0432.CCR-14-1824
101. Arnould L, Arveux P, Couturier J, Gelly-Marty M, Loustalot C, Ettore F, et al. Pathologic complete response to trastuzumab-based neoadjuvant therapy is related to the level of HER-2 amplification. *Clin Cancer Res.* (2007) 13:6404–9. doi: 10.1158/1078-0432.CCR-06-3022
102. Paik S, Kim C, Wolmark N. HER2 status and benefit from adjuvant trastuzumab in breast cancer. *N Engl J Med.* (2008) 358:1409–11. doi: 10.1056/NEJMc0801440
103. Perez EA, Press MF, Dueck AC, Jenkins RB, Kim C, Chen B, et al. Immunohistochemistry and fluorescence *in situ* hybridization assessment of HER2 in clinical trials of adjuvant therapy for breast cancer (NCCTG N9831, BCIRG 006, and BCIRG 005). *Breast Cancer Res Treat.* (2013) 138:99–108. doi: 10.1007/s10549-013-2444-y
104. Meeting Library. *HER2 Heterogeneity as a Predictor of Response to Neoadjuvant T-DM1 Plus Pertuzumab: Results From a Prospective Clinical Trial*. Available online at: <https://meetinglibrary.asco.org/record/174997/abstract> (accessed September 8, 2019).
105. de Azambuja E, Holmes AP, Piccart-Gebhart M, Holmes E, Di Cosimo S, Swaby RF, et al. Lapatinib with trastuzumab for HER2-positive early breast cancer (NeoALTTO): survival outcomes of a randomised, open-label, multicentre, phase 3 trial and their association with pathological complete response. *Lancet Oncol.* (2014) 15:1137–46. doi: 10.1016/S1470-2045(14)70320-1
106. Swain SM, Ewer MS, Viale G, Delaloge S, Ferrero J-M, Verrill M, et al. Pertuzumab, trastuzumab, and standard anthracycline- and taxane-based chemotherapy for the neoadjuvant treatment of patients with HER2-positive localized breast cancer (BERENICE): a phase II, open-label, multicenter, multinational cardiac safety study. *Ann Oncol.* (2018) 29:646–53. doi: 10.1093/annonc/mdx773
107. Pogue-Geile KL, Song N, Jeong J-H, Gavin PG, Kim S-R, Blackmon NL, et al. Intrinsic subtypes, PIK3CA mutation, and the degree of benefit from adjuvant trastuzumab in the NSABP B-31 trial. *J Clin Oncol.* (2015) 33:1340–7. doi: 10.1200/JCO.2014.56.2439
108. Perez EA, Ballman KV, Mashadi-Hossein A, Tenner KS, Kachergus JM, Norton N, et al. Intrinsic subtype and therapeutic response among HER2-positive breast tumors from the NCCTG (Alliance) N9831 trial. *J Natl Cancer Inst.* (2017) 109:djw207. doi: 10.1093/jnci/djw207
109. Rimawi MF, Schiff R, Osborne CK. Targeting HER2 for the treatment of breast cancer. *Annu Rev Med.* (2015) 66:111–28. doi: 10.1146/annurev-med-042513-015127
110. Xu T, He B-S, Liu X-X, Hu X-X, Lin K, Pan Y-Q, et al. The predictive and prognostic role of stromal tumor-infiltrating lymphocytes in HER2-positive breast cancer with trastuzumab-based treatment: a meta-analysis and systematic review. *J Cancer.* (2017) 8:3838–48. doi: 10.7150/jca.21051
111. Ochi T, Bianchini G, Ando M, Nozaki F, Kobayashi D, Criscitiello C, et al. Predictive and prognostic value of stromal tumor-infiltrating lymphocytes before and after neoadjuvant therapy in triple negative and HER2-positive breast cancer. *Eur J Cancer.* (2019) 118:41–8. doi: 10.1016/j.ejca.2019.05.014
112. Degnim AC, Brahmbhatt RD, Radisky DC, Hoskin TL, Stallings-Mann M, Laudenschlager M, et al. Immune cell quantitation in normal breast tissue lobules with and without lobulitis. *Breast Cancer Res Treat.* (2014) 144:539–49. doi: 10.1007/s10549-014-2896-8
113. Salgado R, Denkert C, Campbell C, Savas P, Nuciforo P, Nuciforo P, et al. Tumor-infiltrating lymphocytes and associations with pathological complete response and event-free survival in HER2-positive early-stage breast cancer treated with lapatinib and trastuzumab: a secondary analysis of the NeoALTTO trial. *JAMA Oncol.* (2015) 1:448–54. doi: 10.1001/jamaoncol.2015.0830
114. Bianchini G, Pusztai L, Pienkowski T, Im Y-H, Bianchi GV, Tseng L-M, et al. Immune modulation of pathologic complete response after neoadjuvant HER2-directed therapies in the NeoSphere trial. *Ann Oncol.* (2015) 26:2429–36. doi: 10.1093/annonc/mdv395
115. 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. doi: 10.1016/j.ctrv.2017.04.005
116. Luen SJ, Salgado R, Fox S, Savas P, Eng-Wong J, Clark E, et al. Tumor-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. doi: 10.1016/S1470-2045(16)30631-3
117. Nuciforo P, Pascual T, Cortés J, Llombart-Cussac A, Fasani R, Paré L, et al. A predictive model of pathologic response based on tumor cellularity and tumor-infiltrating lymphocytes (CETIL) in HER2-positive breast cancer treated with chemo-free dual HER2 blockade. *Ann Oncol.* (2018) 29:170–7. doi: 10.1093/annonc/mdx647
118. Cortazar P, Zhang L, Untch M, Mehta K, Costantino JP, Wolmark N, et al. Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. *Lancet.* (2014) 384:164–72. doi: 10.1016/S0140-6736(13)62422-8
119. Loi S, Giobbie-Hurder A, Gombos A, Bachelot T, Hui R, Curigliano G, et al. Pembrolizumab plus trastuzumab in trastuzumab-resistant, advanced, HER2-positive breast cancer (PANACEA): a single-arm, multicentre, phase 1b–2 trial. *Lancet Oncol.* (2019) 20:371–82. doi: 10.1016/S1470-2045(18)30812-X
120. Vernieri C, Milano M, Brambilla M, Mennitto A, Maggi C, Cona MS, et al. Resistance mechanisms to anti-HER2 therapies in HER2-positive breast cancer: current knowledge, new research directions and therapeutic perspectives. *Crit Rev Oncol Hematol.* (2019) 139:53–66. doi: 10.1016/j.critrevonc.2019.05.001
121. Kennedy SP, Hastings JF, Han JZR, Croucher DR. The under-appreciated promiscuity of the epidermal growth factor receptor family. *Front Cell Dev Biol.* (2016) 4:88. doi: 10.3389/fcell.2016.00088
122. Bahrami A, Khazaei M, Shahidsales S, Hassanian SM, Hasanzadeh M, Maftouh M, et al. The therapeutic potential of PI3K/Akt/mTOR inhibitors in breast cancer: rational and progress. *J Cell Biochem.* (2018) 119:213–22. doi: 10.1002/jcb.26136
123. André F, O'Regan R, Ozguroglu M, Toi M, Xu B, Jerusalem G, et al. Everolimus for women with trastuzumab-resistant, HER2-positive, advanced breast cancer (BOLERO-3): a randomised, double-blind, placebo-controlled phase 3 trial. *Lancet Oncol.* (2014) 15:580–91. doi: 10.1016/S1470-2045(14)70138-X
124. André F, Hurvitz S, Fasolo A, Tseng L-M, Jerusalem G, Wilks S, et al. Molecular alterations and everolimus efficacy in human epidermal growth factor receptor 2-overexpressing metastatic breast cancers: combined exploratory biomarker analysis from BOLERO-1 and BOLERO-3. *JCO.* (2016) 34:2115–24. doi: 10.1200/JCO.2015.63.9161
125. Yang J, Nie J, Ma X, Wei Y, Peng Y, Wei X. Targeting PI3K in cancer: mechanisms and advances in clinical trials. *Mol Cancer.* (2019) 18:26. doi: 10.1186/s12943-019-0954-x
126. *Phase I Study of Alpelisib (BYL-719) and Trastuzumab Emtansine (T-DM1) in HER2-Positive Metastatic Breast Cancer (MBC) After Trastuzumab and Taxane Therapy*. Springer. Available online at: <https://link.springer.com/article/10.1007%2Fs10549-018-4792-0> (accessed September 8, 2019).
127. Overcoming CDK4/6 inhibitor resistance. In: *ER-Positive Breast Cancer in: Endocrine-Related Cancer*, Vol. 26 (2019). Available online at: <https://erc.bioscientifica.com/view/journals/erc/26/1/ERC-18-0317.xml> (accessed September 6, 2019).



128. Goel S, Wang Q, Watt AC, Tolaney SM, Dillon DA, Li W, et al. Overcoming therapeutic resistance in HER2-positive breast cancers with CDK4/6 inhibitors. *Cancer Cell*. (2016) 29:255–69. doi: 10.1016/j.ccell.2016.02.006
129. Finn RS, Dering J, Conklin D, Kalous O, Cohen DJ, Desai AJ, et al. PD 0332991, a selective cyclin D kinase 4/6 inhibitor, preferentially inhibits proliferation of luminal estrogen receptor-positive human breast cancer cell lines *in vitro*. *Breast Cancer Res*. (2009) 11:R77. doi: 10.1186/bcr2419
130. Gianni L, Bisagni G, Colleoni M, Del Mastro L, Zamagni C, Mansutti M, et al. Neoadjuvant treatment with trastuzumab and pertuzumab plus palbociclib and fulvestrant in HER2-positive, ER-positive breast cancer (NA-PHER2): an exploratory, open-label, phase 2 study. *Lancet Oncol*. (2018) 19:249–56. doi: 10.1016/S1470-2045(18)30001-9

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Godoy-Ortiz, Sanchez-Muñoz, Chica Parrado, Álvarez, Ribelles, Rueda Dominguez and Alba. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



OPEN ACCESS

**Edited by:**

Takayuki Ueno,  
Cancer Institute Hospital of Japanese  
Foundation for Cancer  
Research, Japan

**Reviewed by:**

Alessandro Igor Cavalcanti Leal,  
Johns Hopkins Medicine,  
United States  
Shigehira Saji,  
Fukushima Medical University, Japan

**\*Correspondence:**

Nuria Ribelles  
nuriaribelles@uma.es

**Specialty section:**

This article was submitted to  
Women's Cancer,  
a section of the journal  
Frontiers in Oncology

**Received:** 14 May 2019

**Accepted:** 18 October 2019

**Published:** 05 November 2019

**Citation:**

Díaz-Redondo T,  
Lavado-Valenzuela R, Jimenez B,  
Pascual T, Gálvez F, Falcón A,  
Alamo MdC, Morales C, Amerigo M,  
Pascual J, Sanchez-Muñoz A,  
González-Guerrero M, Vicioso L,  
Laborda A, Ortega MV, Perez L,  
Fernandez-Martinez A, Chic N,  
Jerez JM, Alvarez M, Prat A,  
Ribelles N and Alba E (2019) Different  
Pathological Complete Response  
Rates According to PAM50 Subtype in  
HER2+ Breast Cancer Patients  
Treated With Neoadjuvant  
Pertuzumab/Trastuzumab vs.  
Trastuzumab Plus Standard  
Chemotherapy: An Analysis of  
Real-World Data.  
Front. Oncol. 9:1178.  
doi: 10.3389/fonc.2019.01178

# Different Pathological Complete Response Rates According to PAM50 Subtype in HER2+ Breast Cancer Patients Treated With Neoadjuvant Pertuzumab/Trastuzumab vs. Trastuzumab Plus Standard Chemotherapy: An Analysis of Real-World Data

**Tamara Díaz-Redondo**<sup>1,2</sup>, **Rocio Lavado-Valenzuela**<sup>1,3</sup>, **Begoña Jimenez**<sup>1,2</sup>, **Tomas Pascual**<sup>4</sup>, **Fernando Gálvez**<sup>5</sup>, **Alejandro Falcón**<sup>6</sup>, **Maria del Carmen Alamo**<sup>7</sup>, **Cristina Morales**<sup>8</sup>, **Marta Amerigo**<sup>9</sup>, **Javier Pascual**<sup>10</sup>, **Alfonso Sanchez-Muñoz**<sup>1,2</sup>, **Macarena González-Guerrero**<sup>11</sup>, **Luis Vicioso**<sup>1,12</sup>, **Aurora Laborda**<sup>3</sup>, **Maria Victoria Ortega**<sup>1,12</sup>, **Lidia Perez**<sup>1,12</sup>, **Aranzazu Fernandez-Martinez**<sup>13</sup>, **Nuria Chic**<sup>4</sup>, **Jose Manuel Jerez**<sup>1,14</sup>, **Martina Alvarez**<sup>1,3,12</sup>, **Aleix Prat**<sup>4</sup>, **Nuria Ribelles**<sup>1,2\*</sup> and **Emilio Alba**<sup>1,2,3</sup>

<sup>1</sup> Instituto de Investigación Biomédica de Málaga (IBIMA), Hospitales Universitarios Regional y Virgen de la Victoria, Málaga, Spain, <sup>2</sup> Unidad de Gestión Clínica Oncología Intercentros, Hospitales Universitarios Regional y Virgen de la Victoria, Málaga, Spain, <sup>3</sup> Laboratorio de Biología Molecular del Cáncer, Centro de Investigaciones Médico-Sanitarias (CIMES), Universidad de Málaga, Málaga, Spain, <sup>4</sup> Translational Genomics and Targeted Therapeutics in Solid Tumors Lab (IDIBAPS), Hospital Clinic de Barcelona, Barcelona, Spain, <sup>5</sup> Hospital Universitario Ciudad de Jaén, Jaén, Spain, <sup>6</sup> Hospital Universitario Virgen del Rocío, Sevilla, Spain, <sup>7</sup> Hospital Universitario Virgen Macarena, Sevilla, Spain, <sup>8</sup> Hospital Reina Sofía, Córdoba, Spain, <sup>9</sup> Hospital Juan Ramón Jiménez, Huelva, Spain, <sup>10</sup> Hospital Costa del Sol, Marbella, Spain, <sup>11</sup> Hospital Universitario Puerta del Mar, Cádiz, Spain, <sup>12</sup> Unidad de Gestión Clínica Anatomía Patológica Intercentros, Hospitales Universitarios Regional y Virgen de la Victoria, Málaga, Spain, <sup>13</sup> Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States, <sup>14</sup> Departamento de Lenguajes y Ciencias de la Computación, Universidad de Málaga, Málaga, Spain

**Background:** Double blockade with pertuzumab and trastuzumab combined with chemotherapy is the standard neoadjuvant treatment for HER2-positive early breast cancer. Data derived from clinical trials indicates that the response rates differ among intrinsic subtypes of breast cancer. The aim of this study is to determine if these results are valid in real-world patients.

**Methods:** A total of 259 patients treated in eight Spanish hospitals were included and divided into two cohorts: Cohort A (132 patients) received trastuzumab plus standard neoadjuvant chemotherapy (NAC), and Cohort B received pertuzumab and trastuzumab plus NAC (122 patients). Pathological complete response (pCR) was defined as the complete disappearance of invasive tumor cells. Assignment of the intrinsic subtype was realized using the research-based PAM50 signature.

**Results:** There were more HER2-enriched tumors in Cohort A (70 vs. 56%) and more basal-like tumors in Cohort B (12 vs. 2%), with similar luminal cases in both cohorts

(luminal A 12 vs. 14%; luminal B 14 vs. 18%). The overall pCR rate was 39% in Cohort A and 61% in Cohort B. Better pCR rates with pertuzumab plus trastuzumab than with trastuzumab alone were also observed in all intrinsic subtypes (luminal PAM50 41 vs. 11.4% and HER2-enriched subtype 73.5 vs. 50%) but not in basal-like tumors (53.3 vs. 50%). In multivariate analysis the only significant variables related to pCR in both luminal PAM50 and HER2-enriched subtypes were treatment with pertuzumab plus trastuzumab (Cohort B) and histological grade 3.

**Conclusions:** With data obtained from patients treated in clinical practice, it has been possible to verify that the addition of pertuzumab to trastuzumab and neoadjuvant chemotherapy substantially increases the rate of pCR, especially in the HER2-enriched subtype but also in luminal subtypes, with no apparent benefit in basal-like tumors.

**Keywords:** breast cancer, real-world data, neoadjuvant, pertuzumab, trastuzumab

## INTRODUCTION

The contribution of anti-HER2 therapies to the management of HER2-positive breast cancer patients is undeniable both in metastatic and adjuvant settings (1–6). In the same way, significant benefit was observed with neoadjuvant trastuzumab treatment, obtaining pathological complete response (pCR) rates from 25 to 46% (7–12). These results were improved with the use of pertuzumab combined with trastuzumab, reaching pCR rates between 49 and 69% (8, 13–17).

Several authors have shown that all four intrinsic subtypes can be found in clinically HER2+ tumors. Although the majority of cases are HER2 enriched (40–72%), luminal A (10–27%), luminal B (10–28%), and basal-like tumors (7–14%) are also represented (13, 17–24). This distribution may vary depending on the hormone receptor status. In the hormone-receptor-negative subset, the main intrinsic subtype was HER2 enriched (51–85%), with fewer cases of luminal (luminal A 0.7–24%; luminal B 3–11%), and basal-like subtypes (9–28%) (12, 19, 20, 22, 24, 25). In contrast, in hormone-receptor-positive tumors, luminal subtypes were more frequent (luminal A 28–44%; luminal B 24–48%) than HER2-enriched (8–32%) or basal-like (0.5–2.5%) ones (12, 20, 22, 25, 26).

This heterogeneity is also reflected in the magnitude of the benefit of neoadjuvant anti-HER2 therapies. In the NOAH trial, pCR rate obtained in the trastuzumab arm was higher in the HER2-enriched subtype and in tumors with high ROR scores (18). This association was also observed in patients treated with double HER2 blockade (i.e., trastuzumab plus lapatinib or trastuzumab plus pertuzumab) plus chemotherapy. Neoadjuvant trastuzumab with or without lapatinib shows pCR rates of 50–70% in HER2-enriched tumors, 9–34% in luminal A, 17–36% in luminal B, and 25–38% in basal-like cases (12, 25, 27). Similarly, pCR rates by intrinsic subtype in patients treated with neoadjuvant pertuzumab plus trastuzumab were 70–83% in the HER2-enriched subtype, 16–45% in luminal A, 16–52% in luminal B, and 20–85% in basal-like tumors (13, 17, 22, 23). In a series of patients with Blueprint-defined subtypes, the pCR rate was 76% in the HER2+ type, 31% in the luminal type, and 43% in the basal type (28).

The aim of this work was to evaluate whether the effect of neoadjuvant pertuzumab combined with trastuzumab in comparison with trastuzumab alone varies as a function of PAM50-defined intrinsic subtypes in a real-world cohort of patients with HER2-positive early breast cancer.

## MATERIALS AND METHODS

### Patients

A total of 254 patients with HER2+ early breast cancer consecutively treated with standard neoadjuvant chemotherapy (NAC) in eight Spanish hospitals were included in the study. The whole population was divided in two cohorts: Cohort A received trastuzumab plus NAC, and Cohort B was treated with pertuzumab and trastuzumab plus NAC. Standard NAC included taxanes with or without anthracyclines. Adjuvant radiotherapy was performed according local practice. Adjuvant endocrine therapy was administered in all hormone-receptor-positive patients.

Patient data were derived from the patients' clinical records and original pathology reports. Although the analysis was retrospective, the data were collected prospectively.

The study was approved by local ethics committees. Written informed consent was obtained from each participant.

### Definition of pCR, Hormone Receptor, HER2 Status, Immunohistochemical Phenotype, and Intrinsic Subtype

pCR was defined as the complete disappearance of invasive tumor cells (ypT0 or ypTis and ypN0). All pathological determinations were performed on diagnostic biopsies. Tumors were classified as estrogen-receptor and progesterone-receptor-positive if  $\geq 1\%$  of tumor cells were stained. HER2+ status was defined by an immunohistochemistry score of 3+ or a HER2 amplification ratio of 2.0 or more by FISH or SISH. Hormone-receptor-positive cases were classified as Luminal-HER2 (luminal immunophenotype) and those with negative hormone receptors such ones as HER2+ (HER2+ immunophenotype).

Assignment of the intrinsic subtype was realized using the research-based PAM50 signature as previously described (29) in order to categorize all cases as one of the following subtypes: luminal A, luminal B, HER2-enriched, basal-like, and normal-like.

## Statistical Analysis

Associations between variables and pCR were evaluated by the chi-squared test or Fisher's exact test. Multivariate logistic regression analyses were used to evaluate the association of each intrinsic subtype with pCR and included the variables that showed significant associations in univariate analyses. Cases with unknown data for any of the variables considered were excluded from multivariate analyses.

All the tests were two-sided, and a *P*-value of <0.05 was considered to indicate statistical significance. Analyses were carried out using the R system for statistical computing (version 3.5.2).

## RESULTS

### Patients' Characteristics

A total of 254 patients were included in the study: 132 patients were treated with NAC plus trastuzumab (Cohort A) and 122 patients with NAC plus pertuzumab and trastuzumab (Cohort B). The clinical characteristics are outlined in **Table 1**. There were more cases with a greater tumor burden in Cohort B than in Cohort A (T3 or T4 tumor size: 33 vs. 19%; stage I: 1 vs. 14%). Slightly more patients were treated with taxanes alone in Cohort B (14 vs. 7%).

The overall pCR rate was 39% in Cohort A and 61% in Cohort B. The immunohistochemical phenotype distribution was similar in both cohorts: luminal-HER2 69 vs. 61% and HER2+ cases 31 vs. 39%. Regarding PAM50-assigned subtypes, there were more HER2-enriched tumors in Cohort A (70 vs. 56%) and more basal-like tumors in Cohort B (12 vs. 2%), with similar luminal case distributions (luminal A 12 vs. 14%; luminal B 14 vs. 18%).

### Association Between Variables and pCR

In the whole population (**Table 2**), pCR was significantly related to the type of treatment (Cohort A 39.4% vs. Cohort B 60.6%; *P* = 0.0011), histological grade (grade 1 + 2 35.5% vs. grade 3 62.2%; *P* = 0.0007), Ki67 level (<20% 28.9% vs. 20–50% 60.8% vs. >50% 54.2%; *P* = 0.003), immunohistochemical phenotype (luminal HER2 38.7% vs. HER2+ 69.6%; *P* = 0.000005), and PAM50-based subtype (luminal A 21.2% vs. luminal B 31.7% vs. HER2-enriched 60% vs. basal like 52.9%; *P* = 0.0004). Similar results were observed in separate analyses of each cohort (**Table 2**).

The better results found in cohort B in the whole population were also observed in an evaluation of different subpopulations (**Table 3**). Thus, immunohistochemical luminal tumors showed greater pCR with pertuzumab and trastuzumab treatment (48.6 vs. 30.8%; *P* = 0.03) and also HER2+ patients (58.5 vs. 79.6%; *P* = 0.06). In addition, in the luminal PAM50-based subtype, a pCR rate of 11.4% was obtained with trastuzumab treatment vs. 41% with combination treatment (*P* = 0.008) and in the HER2-enriched subtype, these rates were 50 vs. 73.5% (*P* = 0.004).

**TABLE 1 |** Patient characteristics.

	Cohort A (T)		Cohort B (P+T)	
Total	132		122	
Menopausal status				
Premenopausal	57	43%	57	47%
Postmenopausal	50	38%	52	43%
NA	25	19%	13	11%
Age				
<50	63	48%	58	48%
≥50	69	52%	64	52%
Tumor size				
T1	26	20%	14	11%
T2	78	59%	67	55%
T3	15	11%	27	22%
T4	10	8%	13	11%
NA	3	2%	1	1%
Nodal status				
Negative	55	42%	44	36%
Positive	77	58%	77	63%
NA			1	1%
Stage				
I	18	14%	1	1%
II	90	68%	90	74%
III	21	16%	30	25%
NA	3	2%	1	1%
Grade				
1–2	62	47%	37	30%
3	45	34%	45	37%
ND	25	19%	40	33%
Hormone receptor				
Negative	41	31%	48	39%
Positive	91	69%	74	61%
Ki67				
<20	25	19%	13	11%
20–50	44	33%	76	62%
>50	29	22%	30	25%
ND	34	26%	3	2%
NAC				
Taxanes <sup>a</sup>	9	7%	17	14%
Taxanes + Anthracyclines <sup>b</sup>	123	93%	105	86%
pCR				
No	80	61%	48	39%
Yes	52	39%	74	61%
Immunohistochemical phenotype				
Luminal-HER2	91	69%	74	61%
HER2	41	31%	48	39%
PAM50-based subtype				
Luminal A	16	12%	17	14%
Luminal B	19	14%	22	18%

(Continued)



**TABLE 1 |** Continued

	Cohort A (T)		Cohort B (P+T)	
HER2-enriched	92	70%	68	56%
Basal-like	2	2%	15	12%
Normal	3	2%		

NA, Not available.

<sup>a</sup>Paclitaxel-Trastuzumab; Docetaxel-Trastuzumab; Paclitaxel-Trastuzumab-Pertuzumab; Docetaxel-Trastuzumab-Pertuzumab.

<sup>b</sup>Epirubicin-Cyclophosphamide followed by Docetaxel-Trastuzumab; Epirubicin-Cyclophosphamide followed by Paclitaxel-Trastuzumab; Adriamycin-Cyclophosphamide followed by Docetaxel-Trastuzumab; Adriamycin-Cyclophosphamide followed by Paclitaxel-Trastuzumab; Fluorouracil-Epirubicin-Cyclophosphamide followed by Docetaxel-Trastuzumab; Fluorouracil-Epirubicin-Cyclophosphamide followed by Paclitaxel-Trastuzumab; Epirubicin-Cyclophosphamide followed by Docetaxel-Trastuzumab-Pertuzumab; Epirubicin-Cyclophosphamide followed by Paclitaxel-Trastuzumab-Pertuzumab; Adriamycin-Cyclophosphamide followed by Docetaxel-Trastuzumab-Pertuzumab; Adriamycin-Cyclophosphamide followed by Paclitaxel-Trastuzumab-Pertuzumab; Fluorouracil-Epirubicin-Cyclophosphamide followed by Docetaxel-Trastuzumab-Pertuzumab; Fluorouracil-Epirubicin-Cyclophosphamide followed by Paclitaxel-Trastuzumab-Pertuzumab.

**TABLE 2 |** Association between variables and pCR.

	Whole population		Cohort A (T)		Cohort B (P+T)	
	pCR (%)	P	pCR (%)	P	pCR (%)	P
Cohort A	39.4	0.001				
Cohort B	60.6					
Grade 1–2	35.5	0.00007	27.4	0.038	40.5	0.002
Grade 3	62.2		48.9		75.5	
Ki67 <20%	28.9	0.002	20.0	0.006	46.1	0.4
Ki67 20–50%	60.8		56.8		63.1	
Ki67 >50%	54.2		41.3		66.0	
Luminal-HER2	38.7	0.000005	30.7	0.004	48.6	0.001
HER2	69.6		58.5		79.1	
Luminal A	21.2	0.00004	6.2	0.0004	35.2	0.007
Luminal B	31.7		15.7		45.4	
HER2-E	60.0		50.0		73.5	
Basal-like	52.9		50.0		53.3	
Normal	33.3		33.3		0.0	

## Multivariate Analyses

The variables that remained significantly associated with pCR in the whole population were treatment Cohort B [Odds Ratio (OR) 2.5; 95% CI 1.07–6;  $P = 0.036$ ], histological grade 3 (OR 3.41; 95% CI 1.48–8.09;  $P = 0.004$ ), immunophenotype HER2+ (OR 3.82; 95% CI 1.39–11.6;  $P = 0.01$ ), and PAM50-based HER2-enriched subtype (OR 2.98; 95% CI 1.39–11.6;  $P = 0.02$ ) (Table 4).

In the cohort of patients treated with trastuzumab alone, grade 3 (OR 5.1; 95% CI 1.5–20.7;  $P = 0.01$ ) and immunophenotype HER2+ (OR 9.8; 95% CI 2.0–75.3;  $P = 0.01$ ) were the only variables independently associated with a higher probability of pCR, and in the cohort of patients that received pertuzumab and trastuzumab, these variables were grade 3 (OR 3.4; 95% CI 1.1–10.8;  $P = 0.03$ ) and PAM50-based HER2-enriched subtype (OR 3.7; 95% CI 1.2–11;  $P = 0.02$ ) (Table 4).

**TABLE 3 |** Association between variables and pCR in specific subpopulations.

	pCR (%)	P
<b>Luminal-HER2 immunophenotype</b>		
Cohort A	30.8	0.03
Cohort B	48.6	
<b>HER2+ immunophenotype</b>		
Cohort A	58.5	0.06
Cohort B	79.6	
<b>Luminal PAM50 subtypes</b>		
Cohort A	11.4	0.008
Cohort B	41.0	
<b>HER2-enriched PAM50 subtypes</b>		
Cohort A	50.0	0.004
Cohort B	73.5	

In an analysis of luminal PAM50-based tumors, the variables that remained significantly associated with pCR were treatment Cohort B (OR 4.2; 95% CI 1.05–22.4;  $P = 0.05$ ), and grade 3 (OR 4.5; 95% CI 1.1–19.0;  $P = 0.03$ ); this was also true in the HER2-enriched subgroup (Cohort B OR 2.7; 95% CI 1.01–7.6;  $P = 0.05$ . Grade 3 OR 4.1; 95% CI 1.6–11.2;  $P = 0.003$ ) (Table 4).

## DISCUSSION

Our study provides valuable information from the real world about neoadjuvant anti-HER2 treatment in early breast cancer, showing that the rate of pCR obtained by double blockade with pertuzumab plus trastuzumab exceeds by 20% that obtained with trastuzumab alone. The pCR rate observed in our series with pertuzumab and trastuzumab treatment (60.6%) is in the range of responses observed in the published phase II-III trials (45.8–69.8%) (8, 13–15, 17, 22). Moreover, the pCR rate found in patients treated with trastuzumab alone (39.4%) is in agreement with previous data (31–46%) (7–12). Interestingly, the greater efficacy shown by the combination of pertuzumab and trastuzumab in our study was despite the fact that the patients in this cohort had worse prognostic characteristics than those who received trastuzumab alone, with a higher percentage of tumors larger than 5 cm or a greater number of cases with nodal involvement. Lower pCR rates were observed in patients with the luminal immunophenotype in both the cohort treated with pertuzumab and trastuzumab and in the one receiving trastuzumab alone. This finding is consistent with previously published data (8, 9, 12, 15, 17, 28).

Although most tumors positive for HER2 by immunohistochemistry or by *in situ* hybridization correspond to the intrinsic HER2-enriched subtype, it is possible to identify any of the remaining intrinsic subtypes in this type of tumor (19, 29, 30). Surprisingly, the percentage of cases by intrinsic subtype in our two patient cohorts differ to some extent, despite the fact that the processing of the tumor samples was performed in the same laboratory, albeit at different times. In the group of patients treated with pertuzumab and trastuzumab, 56% of the cases corresponded to the HER2-enriched, 14% to the luminal

**TABLE 4 |** Multivariate logistic regression of pCR.

	Whole population			Cohort A (T)			Cohort B (P+T)			Luminal PAM50			HER2-E PAM50		
	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P
Cohort B	2.5	1.07–6.00	0.03	–	–	–	–	–	–	4.2	1.05–22.4	0.05	2.7	1.01–7.6	0.05
Grade 3	3.41	1.48–8.09	0.004	5.1	1.5–20.7	0.01	3.4	1.1–10.8	0.03	4.5	1.1–19.0	0.03	4.1	1.6–11.2	0.003
HER2Ihc	3.82	1.39–11.6	0.01	9.8	2.0–75.3	0.01	–	–	–	–	–	–	–	–	–
HER2-E	2.98	1.19–7.7	0.02	–	–	–	3.7	1.2–11	0.02	–	–	–	–	–	–

OR, odds ratio; CI, Confidence Interval; HER2Ihc, immunophenotype HER2; HER2E, HER2 enriched.

A, 18% to the luminal B, and 12% to the basal-like subtype, and this distribution is in agreement with previously published data (13, 17, 18, 20, 22, 23). However, in the cohort of patients who received trastuzumab alone, there was a higher percentage of HER2-enriched cases (70%), a lower number of basal-like tumors (2%), and a similar amount of luminal tumors (luminal A 12%; luminal B 14%). Similar data were reported by Perez et al. from NCCTG N9831 Trial (21) and more recently by Tolane et al. from the APT Trial (24).

Anti-HER2 therapies are more beneficial in HER2-enriched tumors, but all intrinsic subtypes benefit from this type of treatment in both the adjuvant (20, 21) and neoadjuvant settings, and the HER2-enriched subtype benefits the most (13, 17, 18, 22, 23, 28). According to these data, our patients with HER2-enriched tumors obtained the highest pCR rate with both treatment schedules. Furthermore, the use of pertuzumab and trastuzumab was the only variable, together with the histological grade, that provided independent predictive information for pCR events in both HER2-enriched tumors (OR 2.7) and patients with luminal subtypes (OR 4.2). Although the number of patients was small, the basal-like subtype shows no benefit with the use of anti-HER2 therapy, achieving nearly the same pCR rate with pertuzumab and trastuzumab as with trastuzumab alone.

To our knowledge, there is no published series of real-world patients with early HER2+ breast cancer treated with NAC plus pertuzumab and trastuzumab or trastuzumab alone, in which the intrinsic subtypes have been established according to the PAM50 definition and their relationship with the pCR rate analyzed. Beitsch et al. (28) published data from patients included in a prospective registry, of whom 178 were treated with NAC plus trastuzumab and 119 with NAC plus pertuzumab and trastuzumab and in which the molecular subtype was defined by Blueprint platform. Their results agree with ours, showing a higher response with double HER2 blockade vs. treatment with trastuzumab alone in the HER2+ type (76% vs. 57%) and the luminal type (31% vs. 8%) and no differences in the basal type (43% vs. 45%). Recently, Fashing et al. (31) published their results from a series of patients included in an ongoing registry comparing two cohorts of patients that received

neoadjuvant treatment with chemotherapy plus trastuzumab or chemotherapy plus trastuzumab and pertuzumab. In agreement with our results, there was a greater number of pCR in patients treated with pertuzumab plus trastuzumab with an adjusted OR for double HER2 blockade vs. trastuzumab alone of 2.04 (95% CI 1.24–3.35).

Our results confirm the data obtained from clinical trials in patients treated in clinical practice, showing that the addition of pertuzumab to trastuzumab and neoadjuvant chemotherapy increase the pCR rate substantially, especially in the HER2-enriched subtype but also in luminal subtypes, with no apparent benefit in basal-like tumors.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Comité de Ética de la Investigación Provincial de Málaga, Servicio Andaluz de Salud, Consejería de Salud, Junta de Andalucía. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

EA, TD-R, and AP made substantial contributions to the conception of the work. TD-R, RL-V, BJ, TP, FG, AF, MCA, CM, MAL, JP, AS-M, MG-G, LV, AL, MO, LP, AF-M, NC, and MAM contributed to the acquisition of the data. EA, JJ, TD-R, and NR contributed to the analysis of data. NR and EA drafting the work. All authors approved the final version of the manuscript.

## ACKNOWLEDGMENTS

The authors acknowledge support through grant TIN2017-88728-C2-1-R from MICINN SPAIN.

## REFERENCES

- Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, et al. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N Engl J Med*. (2001) 344:783–92. doi: 10.1056/NEJM200103153441101
- Cameron D, Piccart-Gebhart MJ, Gelber RD, Procter M, Goldhirsch A, de Azambuja E, et al. 11 years' follow-up of trastuzumab after adjuvant chemotherapy in HER2-positive early breast cancer: final analysis of the HERceptin Adjuvant (HERA) trial. *Lancet*. (2017) 389:1195–205. doi: 10.1016/S0140-6736(16)32616-2
- Perez EA, Romond EH, Suman VJ, Jeong JH, Sledge G, Geyer CE Jr, et al. Trastuzumab plus adjuvant chemotherapy for human epidermal growth factor receptor 2-positive breast cancer: planned joint analysis of overall survival from NSABP B-31 and NCCTG N9831. *J Clin Oncol*. (2014) 32:3744–52. doi: 10.1200/JCO.2014.55.5730
- Slamon D, Eiermann W, Robert N, Pienkowski T, Martin M, Press M, et al. Ten year follow-up of the BCIRG-006 trial comparing doxorubicin plus cyclophosphamide followed by docetaxel (AC-T) with doxorubicin plus cyclophosphamide followed by docetaxel and trastuzumab (AC-TH) with docetaxel, carboplatin and trastuzumab (TCH) in HER2+ early breast cancer patients. *Cancer Res*. (2015) 76. doi: 10.1158/1538-7445.SABCS15-S5-04
- Swain SM, Kim SB, Cortes J, Ro J, Semiglazov V, Campone M, et al. Pertuzumab, trastuzumab, and docetaxel for HER2-positive metastatic breast cancer (CLEOPATRA study): overall survival results from a randomised, double-blind, placebo-controlled, phase 3 study. *Lancet Oncol*. (2013) 14:461–71. doi: 10.1016/S1470-2045(13)70130-X
- Bachelot T, Ciruelos E, Schneeweiss A, Puglisi F, Peretz-Yablonski T, Bondarenko I, et al. Preliminary safety and efficacy of first-line pertuzumab combined with trastuzumab and taxane therapy for HER2-positive locally recurrent or metastatic breast cancer (PERUSE). *Ann Oncol*. (2019) 30:766–73. doi: 10.1093/annonc/mdz061
- Gianni L, Eiermann W, Semiglazov V, Manikhas A, Lluch A, Tjulandini S, et al. Neoadjuvant chemotherapy with trastuzumab followed by adjuvant trastuzumab versus neoadjuvant chemotherapy alone, in patients with HER2-positive locally advanced breast cancer (the NOAH trial): a randomised controlled superiority trial with a parallel HER2-negative cohort. *Lancet*. (2010) 375:377–84. doi: 10.1016/S0140-6736(09)61964-4
- Gianni L, Pienkowski T, Im YH, Roman L, Tseng LM, Liu MC, et al. Efficacy and safety of neoadjuvant pertuzumab and trastuzumab in women with locally advanced, inflammatory, or early HER2-positive breast cancer (NeoSphere): a randomised multicentre, open-label, phase 2 trial. *Lancet Oncol*. (2012) 13:25–32. doi: 10.1016/S1470-2045(11)70336-9
- Baselga J, Bradbury I, Eidtmann H, Di Cosimo S, de Azambuja E, Aura C, et al. Lapatinib versus trastuzumab for HER2-positive early breast cancer (NeoALTTO): a randomised, open-label, multicentre, phase 3 trial. *Lancet*. (2012) 379:633–40. doi: 10.1016/S0140-6736(11)61847-3
- Guarneri V, Frassoldati A, Bottini A, Cagossi K, Bisagni G, Sarti S, et al. Preoperative chemotherapy plus trastuzumab, lapatinib, or both in human epidermal growth factor receptor 2-positive operable breast cancer: results of the randomized phase II CHER-LOB study. *J Clin Oncol*. (2012) 30:1989–95. doi: 10.1200/JCO.2011.39.0823
- Untch M, Loibl S, Bischoff J, Eidtmann H, Kaufmann M, Blohmer JU, et al. Lapatinib versus trastuzumab in combination with neoadjuvant anthracycline-taxane-based chemotherapy (GeparQuinto, GBG 44): a randomised phase 3 trial. *Lancet Oncol*. (2012) 13:135–44. doi: 10.1016/S1470-2045(11)70397-7
- Carey LA, Berry DA, Cirincione 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. doi: 10.1200/JCO.2015.62.1268
- Gavila J, Oliveira M, Pascual T, Pérez J, Canes J, González X, et al. Association of intrinsic subtype and immune genes with pathological complete response in the OPTIHER-HEART phase II clinical trial following neoadjuvant trastuzumab/pertuzumab-based chemotherapy in HER2-positive breast cancer. *Cancer Res*. (2018) 78. doi: 10.1158/1538-7445.SABCS17-P2-09-04
- Schneeweiss A, Chia S, Hickish T, Harvey V, Eniu A, Hegg R, et al. Pertuzumab plus trastuzumab in combination with standard neoadjuvant anthracycline-containing and anthracycline-free chemotherapy regimens in patients with HER2-positive early breast cancer: a randomized phase II cardiac safety study (TRYPHAENA). *Ann Oncol*. (2013) 24:2278–84. doi: 10.1093/annonc/mdt182
- Untch M, Jackisch C, Schneeweiss A, Conrad B, Aktas B, Denkert C, et al. Nab-paclitaxel versus solvent-based paclitaxel in neoadjuvant chemotherapy for early breast cancer (GeparSepto-GBG 69): a randomised, phase 3 trial. *Lancet Oncol*. (2016) 17:345–56. doi: 10.1016/S1470-2045(15)00542-2
- Hurvitz SA, Martin M, Symmans WF, Jung KH, Huang CS, Thompson AM, et al. Neoadjuvant trastuzumab, pertuzumab, and chemotherapy versus trastuzumab emtansine plus pertuzumab in patients with HER2-positive breast cancer (KRISTINE): a randomised, open-label, multicentre, phase 3 trial. *Lancet Oncol*. (2018) 19:115–26. doi: 10.1016/S1470-2045(17)30716-7
- Swain SM, Ewer MS, Viale G, Delaloge S, Ferrero JM, Verrill M, et al. Pertuzumab, trastuzumab, and standard anthracycline- and taxane-based chemotherapy for the neoadjuvant treatment of patients with HER2-positive localized breast cancer (BERENICE): a phase II, open-label, multicenter, multinational cardiac safety study. *Ann Oncol*. (2018) 29:646–53. doi: 10.1093/annonc/mdx773
- Prat A, Bianchini G, Thomas M, Belousov A, Cheang MC, Koehler A, et al. Research-based PAM50 subtype predictor identifies higher responses and improved survival outcomes in HER2-positive breast cancer in the NOAH study. *Clin Cancer Res*. (2014) 20:511–21. doi: 10.1158/1078-0432.CCR-13-0239
- Prat A, Carey LA, Adamo B, Vidal M, Tabernero J, Cortes J, et al. Molecular features and survival outcomes of the intrinsic subtypes within HER2-positive breast cancer. *J Natl Cancer Inst*. (2014) 106:dju152. doi: 10.1093/jnci/dju152
- Pogue-Geile KL, Song N, Jeong JH, Gavin PG, Kim SR, Blackmon NL, et al. Intrinsic subtypes, PIK3CA mutation, and the degree of benefit from adjuvant trastuzumab in the NSABP B-31 trial. *J Clin Oncol*. (2015) 33:1340–7. doi: 10.1200/JCO.2014.56.2439
- Perez EA, Ballman KV, Mashadi-Hosseini A, Tenner KS, Kachergus JM, Norton N, et al. Intrinsic subtype and therapeutic response among HER2-positive breast tumors from the NCCTG (Alliance) N9831 trial. *J Natl Cancer Inst*. (2017) 109:djw207. doi: 10.1093/jnci/djw207
- Prat A, Slamon D, Hurvitz S, Press MF, Lewis Phillips G, Valverde VL, et al. Association of intrinsic subtypes with pathological complete response (pCR) in the KRISTINE neoadjuvant phase 3 clinical trial in HER2-positive early breast cancer (EBC). Abstract PD3–06. *Cancer Res*. (2017) 78. doi: 10.1158/1538-7445.SABCS17-PD3-06
- Ignatiadis M, Van den Eynden G, Roberto S, Fornili M, Bareche Y, Desmedt C, et al. Tumor-infiltrating lymphocytes in patients receiving trastuzumab/pertuzumab-based chemotherapy: a TRYPHAENA substudy. *J Natl Cancer Inst*. (2019) 111:69–77. doi: 10.1093/jnci/djy076
- Tolaney SM, Guo H, Pernas S, Barry WT, Dillon DA, Ritterhouse L, et al. Seven-year follow-up analysis of adjuvant paclitaxel and trastuzumab trial for node-negative, human epidermal growth factor receptor 2-positive breast cancer. *J Clin Oncol*. (2019) 37:1868–75. doi: 10.1200/JCO.19.00066
- Dieci MV, Prat A, Tagliafico E, Pare L, Ficarra G, Bisagni G, et al. Integrated evaluation of PAM50 subtypes and immune modulation of pCR in HER2-positive breast cancer patients treated with chemotherapy and HER2-targeted agents in the CherLOB trial. *Ann Oncol*. (2016) 27:1867–73. doi: 10.1093/annonc/mdw262
- Cejalvo JM, Pascual T, Fernandez-Martinez A, Braso-Maristany F, Gomis RR, Perou CM, et al. Clinical implications of the non-luminal intrinsic subtypes in hormone receptor-positive breast cancer.

- Cancer Treat Rev.* (2018) 67:63–70. doi: 10.1016/j.ctrv.2018.04.015
27. Fumagalli D, Venet D, Ignatiadis M, Azim HA Jr, Maetens M, Rothe F, et al. RNA sequencing to predict response to neoadjuvant anti-HER2 therapy: a secondary analysis of the NeoALTTO randomized clinical trial. *JAMA Oncol.* (2016) 3:227–34. doi: 10.1001/jamaoncol.2016.3824
  28. Beitsch P, Whitworth P, Baron P, Rotkis MC, Mislowsky AM, Richards PD, et al. Pertuzumab/trastuzumab/CT versus trastuzumab/CT therapy for HER2+ breast cancer: results from the prospective neoadjuvant breast registry symphony trial (NBRST). *Ann Surg Oncol.* (2017) 24:2539–46. doi: 10.1245/s10434-017-5863-x
  29. 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) 27:1160–7. doi: 10.1200/JCO.2008.18.1370
  30. The Cancer Genome Atlas N, Koboldt DC, Fulton RS, McLellan MD, Schmidt H, Kalicki-Veizer J, et al. Comprehensive molecular portraits of human breast tumours. *Nature.* (2012) 490:61. doi: 10.1038/nature11412
  31. Fasching PA, Hartkopf AD, Gass P, Haberle L, Akpolat-Basci L, Hein A, et al. Efficacy of neoadjuvant pertuzumab in addition to chemotherapy and trastuzumab in routine clinical treatment of patients with primary breast cancer: a multicentric analysis. *Breast Cancer Res Treat.* (2019) 173:319–28. doi: 10.1007/s10549-018-5008-3

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Díaz-Redondo, Lavado-Valenzuela, Jimenez, Pascual, Gálvez, Falcón, Alamo, Morales, Amerigo, Pascual, Sanchez-Muñoz, González-Guerrero, Vicioso, Laborda, Ortega, Perez, Fernandez-Martinez, Chic, Jerez, Alvarez, Prat, Ribelles and Alba. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# 7-lncRNA Assessment Model for Monitoring and Prognosis of Breast Cancer Patients: Based on Cox Regression and Co-expression Analysis

Huayao Li<sup>1†</sup>, Chundi Gao<sup>2†</sup>, Lijuan Liu<sup>3,4</sup>, Jing Zhuang<sup>3,4</sup>, Jing Yang<sup>3</sup>, Cun Liu<sup>2</sup>, Chao Zhou<sup>3,4</sup>, Fubin Feng<sup>3,4</sup> and Changgang Sun<sup>5\*</sup>

<sup>1</sup> College of Chinese Medicine, Shandong University of Traditional Chinese Medicine, Jinan, China, <sup>2</sup> College of First Clinical Medicine, Shandong University of Traditional Chinese Medicine, Jinan, China, <sup>3</sup> Department of Oncology, Weifang Traditional Chinese Hospital, Weifang, China, <sup>4</sup> Department of Oncology, Affiliated Hospital of Weifang Medical University, Weifang, China, <sup>5</sup> Chinese Medicine Innovation Institute, Shandong University of Traditional Chinese Medicine, Jinan, China

## OPEN ACCESS

### Edited by:

Aleix Prat,  
Hospital Clínic de Barcelona, Spain

### Reviewed by:

Luigi Formisano,  
University of Naples Federico II, Italy  
Chuanxin Wang,  
Second Hospital, Shandong  
University, China

### \*Correspondence:

Changgang Sun  
scgdoctor@126.com

<sup>†</sup>These authors have contributed  
equally to this work and share first  
authorship

### Specialty section:

This article was submitted to  
Women's Cancer,  
a section of the journal  
Frontiers in Oncology

Received: 25 January 2019

Accepted: 15 November 2019

Published: 03 December 2019

### Citation:

Li H, Gao C, Liu L, Zhuang J, Yang J,  
Liu C, Zhou C, Feng F and Sun C  
(2019) 7-lncRNA Assessment Model  
for Monitoring and Prognosis of  
Breast Cancer Patients: Based on  
Cox Regression and Co-expression  
Analysis. *Front. Oncol.* 9:1348.  
doi: 10.3389/fonc.2019.01348

**Background:** Breast cancer is one of the deadliest malignant tumors worldwide. Due to its complex molecular and cellular heterogeneity, the efficacy of existing breast cancer risk prediction models is unsatisfactory. In this study, we developed a new lncRNA model to predict the prognosis of patients with BRCA.

**Methods:** BRCA-related differentially-expressed long non-coding RNA were screened from the Cancer Genome Atlas database. A novel lncRNA model was developed by univariate and multivariate analyses to predict the prognosis of patients with BRCA. The efficacy of the model was verified by TCGA-based breast cancer samples. Identified lncRNA-related mRNA based on the co-expression method.

**Results:** We constructed a 7-lncRNA breast cancer prediction model including LINC00377, LINC00536, LINC01224, LINC00668, LINC01234, LINC02037, and LINC01456. The breast cancer samples were divided into high-risk and low-risk groups based on the model, which verified the specificity and sensitivity of the model. The Area Under Curve (AUC) of the 3- and 5-year Receiver Operating Characteristic curve were 0.711 and 0.734, respectively, indicating that the model has good performance.

**Conclusion:** We constructed a 7-lncRNA model to predict the prognosis of patients with BRCA, and suggest that these lncRNAs may play a specific role in the carcinogenesis of BRCA.

**Keywords:** breast cancer, univariate and multivariate Cox analyses, bioinformatic analysis, 7-lncRNA model, co-expression analysis

## INTRODUCTION

Breast cancer (BRCA) is considered as the leading cause of death among gynecologic neoplasias. The treatment of BRCA has markedly improved due to advances in early screening and the development of anticancer strategies (1). However, breast cancer still exhibits a high recurrence rate (2). Studies have shown that the prognosis of breast cancer is affected by many factors like

age, tumor size, grade, lymph node involvement, lymphovascular invasion, histology, hormone-receptor status, c-erbB2 status, and positive margins (3). Due to the pathogenic complexity of breast cancer, although many breast cancer prognostic biomarkers have been discovered, prognosis remains a difficult problem (4, 5). There is a need to construct a new breast cancer risk prediction model to improve the treatment of breast cancer patients. Due to the gene signature is yet limited in coding genes and microRNAs, to prove the necessity to develop the lncRNA model for predicting BRCA survival.

In the post-genomic era, many genome sequencing techniques have emerged (6). These tools provide new ideas and insights for tumor diagnosis and prognosis prediction. These next-generation sequencing methods and the data can thereby help better identify clinical biomarkers of cancer. The discovery of long non-coding RNA (lncRNA) has dramatically altered our understanding of cancer. The expression and dysregulation of lncRNAs is more cancer-type specific than the protein-coding genes (7). The latest research shows that lncRNAs play key roles in gene regulation and carcinogenesis, including proliferation, adhesion, migration, and apoptosis (8). Given the heterogeneity of BRCA and the complexity of non-coding RNAs, a panel of lncRNA biomarkers may be more precise and stable for BRCA prognosis (9). Shi et al. (10), based on The Cancer Genome Atlas (TCGA) database, constructed a 31-lncRNA model, which might be able to predict Overall Survival (OS) in patients with lung adenocarcinoma with high accuracy. Long et al. (11), by integrating the high-throughput data from the TCGA database, screened four genes (*CENPA*, *SPPI*, *MAGEB6*, and *HOXD9*) using univariate, Lasso, and multivariate Cox-regression analyses to develop the hepatocellular carcinoma prognostic model.

In this study, we screened breast cancer-associated differentially-expressed lncRNAs from the TCGA database and developed a new lncRNA model to predict the prognosis of patients with BRCA. It is well-known that lncRNAs could affect the function of proteins and cells directly or indirectly due to their involvement in the regulation of mRNA (12). Therefore, we have further explored the function of lncRNA in the model by studying the function of lncRNA-related mRNA. In summary, the use of lncRNA features provides a deeper insight into the prognosis of BRCA, which may be helpful in guiding the treatment.

## MATERIALS AND METHODS

### Data Source

The lncRNA expression profiles and the corresponding clinical information from the patients with BRCA were obtained from The Cancer Genome Atlas (TCGA: <https://cancergenome.nih.gov/>) (13); a total of 1,208 samples, including 112 healthy and 1,096 BRCA samples. BRCA samples with incomplete prognostic

information were excluded, and the average expression level was used as the final expression data of the same patient mRNA and lncRNA. A total of 1,076 BRCA samples were selected for further construction of the prognostic risk model and co-expression analysis. As the information was retrieved from the TCGA database, a public database, further ethical approvals do not apply to our research. Data collection and processing are in line with TCGA data policies for protecting human subjects (<http://cancergenome.nih.gov/publications/publicationsguidelines>).

**TABLE 1 |** Specific baseline clinical characteristic of 1,076 breast cancer patients.

1,076 breast cancer patients	
<b>Age</b>	
<60 years	572
≥60 years	504
<b>Stage</b>	
I	180
II	610
III	244
IV	19
Unknown	23
<b>Pathologic T stage</b>	
T1-2	897
T3-4	176
Unknown	3
<b>Pathologic N stage</b>	
N0-1	862
N2-3	194
Unknown	20
<b>Pathologic M stage</b>	
M0	896
M1	21
Unknown	159
<b>Estrogen receptor</b>	
Positive	790
Negative	237
Unknown	49
<b>Progesterone receptor</b>	
Positive	683
Negative	341
Unknown	52
<b>HER2</b>	
Positive	161
Negative	554
Unknown	361
<b>Survival time</b>	
≤1 years	185
1 years <	482
≤3 years	
3 years <	167
≤5 years	
>5 years	242

**Abbreviations:** lncRNAs, long non-coding RNAs; BRCA, breast cancer; OS, overall survival; TCGA, the Cancer Genome Atlas; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; ROC, receiver operating characteristic; AUC, Area Under Curve; BP, biological process; CC, cellular component; MF, molecular function.

## Identification of Differentially-Expressed lncRNAs and mRNAs

To identify the lncRNAs and mRNAs differentially expressed between the BRCA and the healthy samples, the downloaded lncRNA and mRNA data were standardized and differential-expression analysis was performed using the edgeR software package in the R software. The lncRNAs and mRNAs were differentially expressed with an absolute  $|\log FC| > 2$  and  $p < 0.01$  were considered for subsequent analysis. The logFC indicates the fold change in the expression of each lncRNA and mRNA between BRCA and healthy breast tissue samples. Volcano plot of the differentially-expressed lncRNAs and mRNAs was obtained using the R software.

## Definition of the lncRNA-Related Prognostic Model

The lncRNA-related prognostic model was constructed based on the prognostic characteristics of lncRNA, and the correlation between overall survival (OS) and lncRNA expression levels was studied using univariate and multivariate Cox regression analysis. Differences were assessed by univariate Cox proportional hazards regression analysis using R survival kits. For the association between expressed lncRNA and the overall survival, the lncRNA was considered significant when the  $p$ -value was  $<0.01$  in the univariate Cox-regression analysis and was selected for multivariate Cox-regression analysis. Subsequently, multivariate Cox-regression analysis was performed to evaluate the contribution of genes as independent prognostic factors inpatient survival. A stepwise approach was used to further select the best model. A lncRNA-based prognostic risk score was calculated based on a linear combination of regression coefficients from the

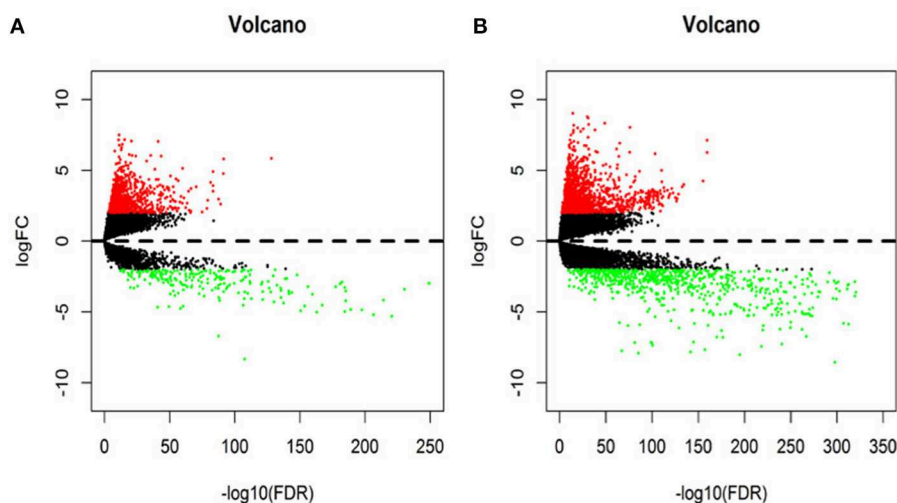
multivariate Cox-regression model ( $\beta$ ) and its expression levels (10, 11).

$$\text{Prognostic index} = \sum_{i=1}^N \text{Exp}_i \times \beta_i$$

The Rpackage was used to find the optimal median threshold. According to the optimal median threshold, the survival data of 1,076 patients with BRCA were divided into low-risk and high-risk groups. Kaplan-Meier (KM) survival curves were generated to assess OS in low-risk or high-risk cases and time-dependent receiver operating characteristic (ROC) curve analysis was performed to calculate area under the curve (AUC) values to assess the predictive power of the model (14). Subsequently, we applied the model to patients with

**TABLE 2 |** Thirteen prognosis-related lncRNAs obtained based on univariate Cox regression analysis ( $P < 0.01$ ).

Name	HR	z	p-value
LINC02037	1.243690056	4.120958497	3.77E-05
LINC01234	1.154170798	3.811036114	1.38E-04
LINC00668	1.105563899	3.700969179	2.15E-04
LINC01456	1.132314635	3.598894647	3.20E-04
LINC01592	1.238001411	3.465525511	5.29E-04
LINC02418	1.154200697	3.230285935	1.24E-03
LINC01854	1.221094553	2.881476012	3.96E-03
C6orf99	1.225252162	2.837096926	4.55E-03
LINC00536	1.117384364	2.763536856	5.72E-03
LINC01224	0.916544112	-2.70404596	6.85E-03
LINC02408	1.19605941	2.681361271	7.33E-03
LINC00377	0.748711948	-2.67297483	7.52E-03
LINC01574	1.145235673	2.589564334	9.61E-03



**FIGURE 1 |** The volcano diagram about differentially expresses lncRNAs (A) and mRNAs (B) between breast cancer tissue and normal tissue samples. Red dots represent up-regulated RNA and green dots represent down-regulated RNA.

stage I, II, III, and Her2 positive BRCA to test the sensitivity and effectiveness of the model for survival prediction. In addition, we compared the predictive performance of 7-lncRNA model with traditional clinical risk factors (including age, TNM, stage, ER, PR, and HER2 status) by univariate and multivariate Cox analysis. First of all, univariate Cox analysis found factors closely related to the prognosis of patients. Then, the effects of many factors on survival time were analyzed at the same time, and the independent prognostic factors could be used to evaluate the survival of patients.  $P < 0.05$  was used as the cutoff condition to verify the ability of the model to evaluate the prognosis and sensitivity of patients.

### Co-expression Method Predicts lncRNA-Related mRNAs

To better explore the function of the relevant lncRNAs in the risk assessment model, the related mRNAs were predicted by co-expression methods based on the Pearson correlation. The related mRNAs were screened for functional enrichment analysis according to  $|COR| > 0.25$ ,  $p < 0.05$ . In addition,

the lncRNA-mRNA co-expression network was visualized using Cytoscape.

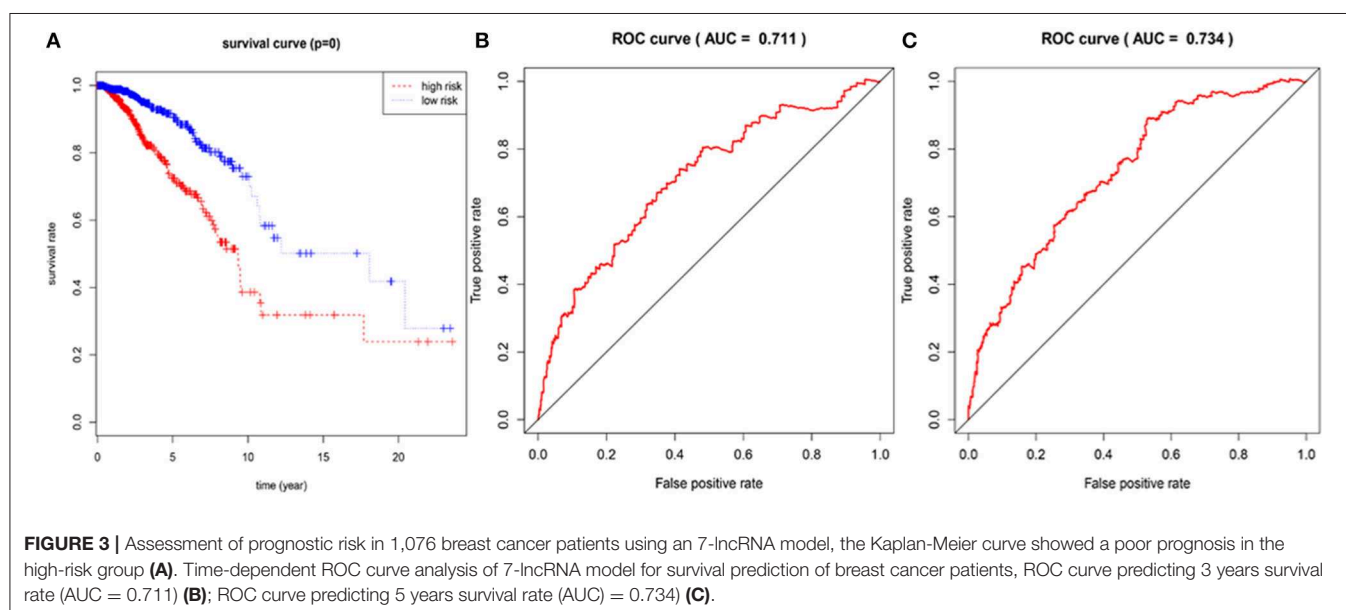
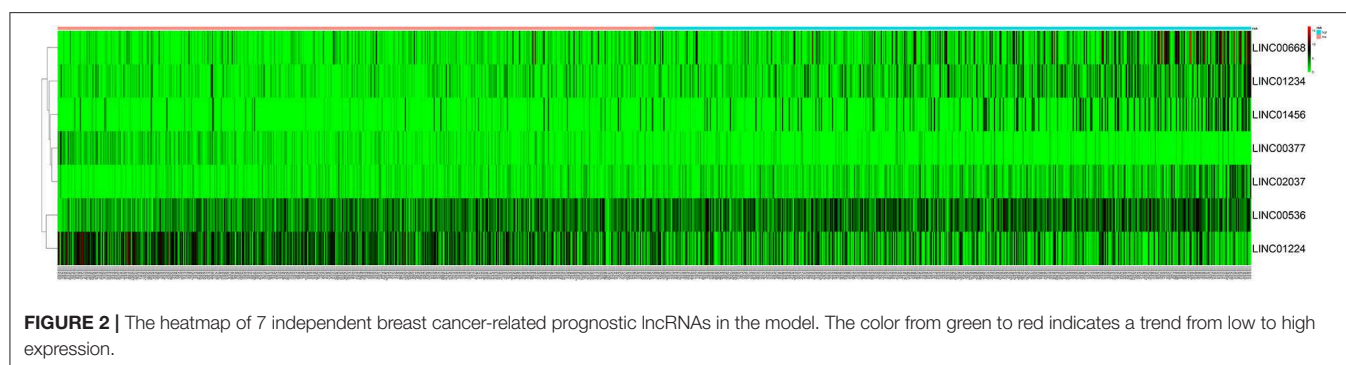
### GO and KEGG Analysis of lncRNA-Related mRNA

To understand the underlying biological pathways between lncRNA and the related mRNAs, the database for annotation, visualization, and integrated discovery (DAVID) (<http://david.abcc.ncifcrf.gov/>) was used to perform functional enrichment analysis (15). Subsequently, lncRNA-related mRNAs were analyzed using the gene ontology (GO) database (<http://www.geneontology.org>). Finally, significantly enriched GO terms were selected to analyze their biological function. The Kyoto Encyclopedia of Genes and Genomes (KEGG; <http://www.kegg.jp/>) was used to perform the pathway enrichment analysis.

## RESULTS

### Differentially Expressed lncRNAs and mRNAs in BRCA Patients

In this study, 1,208 samples were downloaded from the TCGA database and were used to identify differentially-expressed



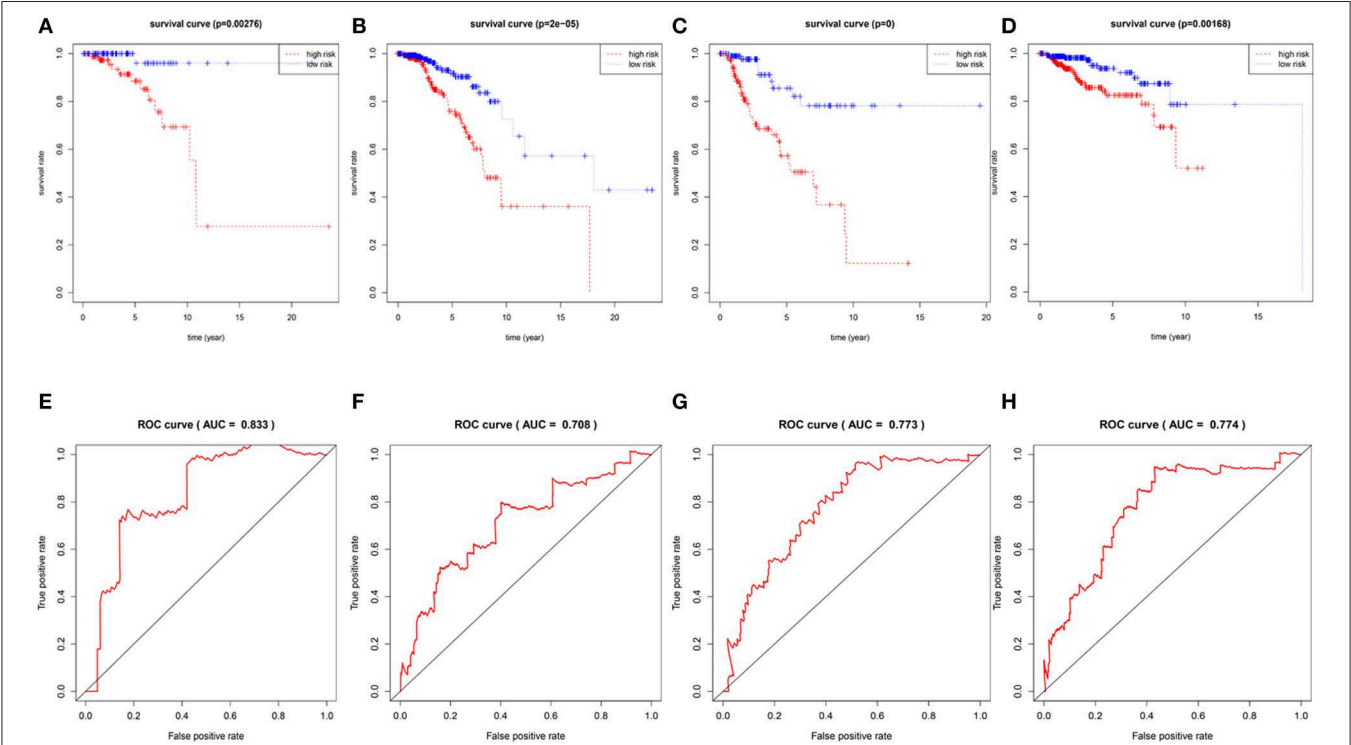


lncRNAs and mRNAs in BRCA patients, We analyzed the specific baseline clinical characteristic of 1,076 BRCA patients presented in **Table 1**. A total of 1,059 differentially expressed lncRNAs were obtained in accordance with  $|\log FC| > 2$  and  $p < 0.01$ . This included 842 upregulated lncRNAs and 217 downregulated lncRNAs (**Figure 1A**), and 2,138 differentially-expressed mRNAs

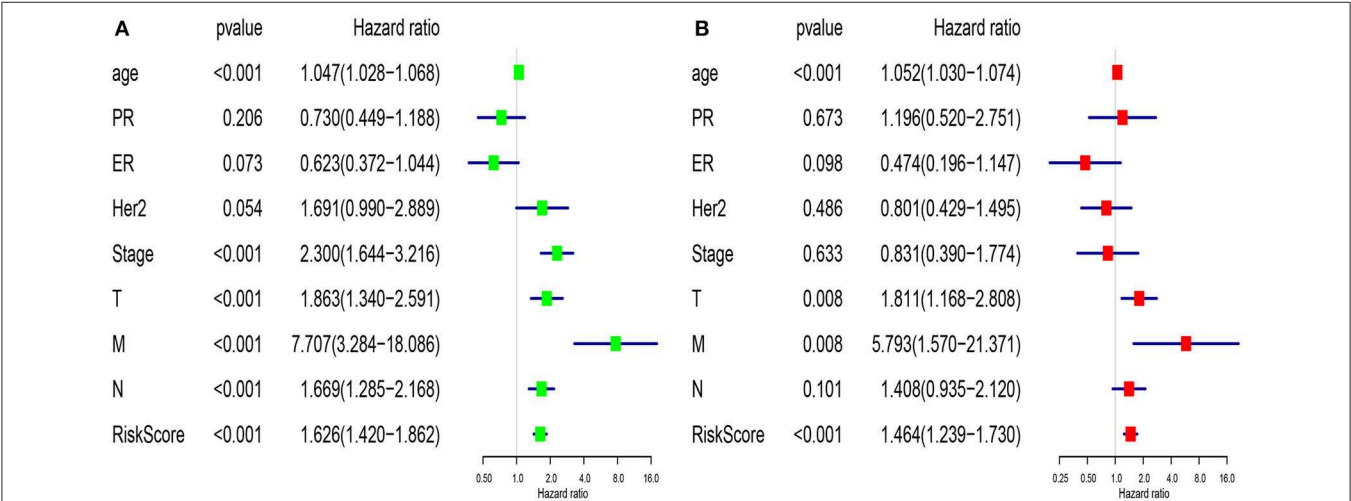
included 1,375 upregulated mRNAs and 763 downregulated mRNAs (**Figure 1B**).

Derivation of lncRNA Prognostic Model

After excluding lncRNA without specific names and lack of corresponding studies, a total of 282 differentially-expressed



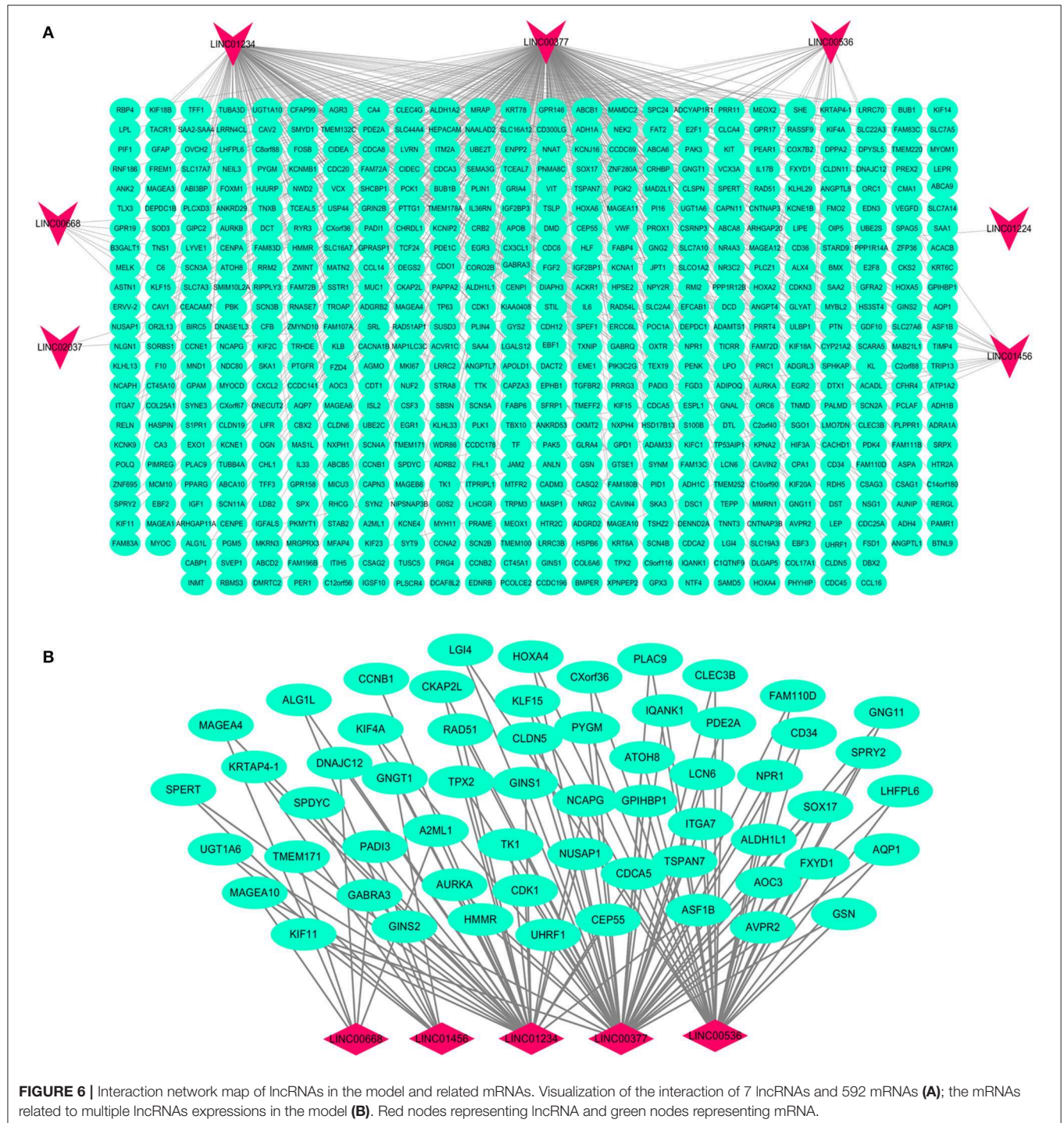
**FIGURE 4 |** Verification the specificity and sensitivity of the 7-lncRNA prognostic model. The Kaplan-Meier curve of patients with stage I, stage II, stage III, and Her2-positive BRCA (**A–D**); the ROC curve of the model at 3 years of OS with stage I, stage II, stage III, and Her2-positive BRCA, the AUC values were 0.883, 0.708, 0.773, 0.774 (**E–H**).



**FIGURE 5 |** Univariate (**A**) and multivariate (**B**) analysis of clinic pathologic factors for overall survival of breast cancer patients from TCGA.

lncRNAs remained for further study. Firstly, we performed a univariate Cox-regression analysis to study the correlation between differentially-expressed lncRNA and OS of BRCA patients. With a  $p < 0.01$  as an identification standard, a total of 13 lncRNAs were obtained, which were significantly associated with OS in BRCA patients (Table 2). Subsequently, based on the primary screening using univariate Cox-regression analysis, we obtained seven lncRNAs that were used to construct a predictive

model by performing stepwise multivariate Cox-regression analysis. They were LINC00377, LINC00536, LINC01224, LINC00668, LINC01234, LINC02037, and LINC01456 and the cluster dendrogram for these lncRNA is shown in Figure 2. The predictive model was characterized by the linear combination of the expression levels of the seven lncRNAs weighted by their relative coefficients from the multivariate Cox regression as follows:



Prognostic index (PI) =  $(-0.2611 \times \text{expression level of LINC00377}) + (0.0960 \times \text{expression level of LINC00536}) + (-0.0966 \times \text{expression level of LINC01224}) + (0.0738 \times \text{expression level of LINC00668}) + (0.1014 \times \text{expression level of LINC01234}) + (0.2020 \times \text{expression level of LINC02037}) + (0.0627 \times \text{expression level of LINC01456})$ .

Of these seven lncRNAs obtained by Cox-regression analysis, five (LINC00536, LINC00668, LINC01234, LINC02037, and LINC01456) showed positive coefficients, suggesting that these lncRNAs have a higher risk and their expression corresponds to the shorter OS in BRCA patients. In addition, the risk prediction correlation analysis between the seven lncRNAs is presented in **Supplementary Figure 1**. At the same time, the remaining two lncRNAs (LINC00377 and LINC01224) showed negative coefficients. Although the risk associated with these two lncRNAs is not higher, they are still important links in the prognosis model. These seven lncRNAs together constitute a prognostic model for patients with BRCA.

In the 1,076 BRCA patients, the median of the prognostic score was obtained as the grouping threshold by calculating the risk scores for the expression of the seven lncRNAs. With a median PI as the group threshold, 538 patients with a prognostic score above the PI threshold were classified as high risk, while 538 patients below the PI threshold were assigned to the low-risk group. We found that Kaplan-Meier survival curve analysis of the high-risk and low-risk groups based on the prognostic risk model constructed by the seven lncRNAs showed that the overall survival rate of the high-risk group was lower, and the difference between the two groups was statistically significant (**Figure 3A**). Subsequently, the prognostic ability of the 7-lncRNA prognostic model was evaluated by calculating the AUC of the time-dependent ROC curve. Based on earlier results of the RUC curve, the higher the AUC, the better is the prediction performance of the model. For 3- and 5-year survival times, the AUC of the 7-lncRNA BRCA patient prognostic model was 0.711 and 0.734, respectively, indicating that the predictive model is highly sensitive and specific (**Figures 3B,C**).

**TABLE 3 |** Functional enrichment analysis of lncRNA-related mRNAs.

Category	Term	Count	P-Value
Biological Processes	Cell division	45	1.13E-15
	Mitotic nuclear division	37	5.70E-15
	Positive regulation of cell proliferation	32	4.11E-05
	Cell proliferation	26	1.49E-04
	Cell adhesion	26	3.81E-03
	Response to drug	23	1.65E-04
	Sister chromatid cohesion	21	3.18E-11
	Cell surface receptor signaling pathway	19	1.90E-03
	DNA replication	17	2.07E-05
	G2/M transition of mitotic cell cycle	16	1.86E-05
Molecular Function	ATP binding	61	1.08E-02
	Calcium ion binding	35	5.20E-03
	Protein kinase binding	28	2.39E-05
	Chromatin binding	21	1.36E-02
	Microtubule binding	19	5.29E-05
	Transcriptional activator activity, RNA polymerase II core promoter proximal region sequence-specific binding	16	4.87E-03
	ATPase activity	15	1.19E-03
	Heparin binding	14	1.02E-03
	Transporter activity	14	7.65E-03
	Microtubule motor activity	12	2.41E-05
Cellular component	Plasma membrane	148	1.38E-02
	Cytosol	122	1.36E-02
	Extracellular region	79	2.15E-05
	Integral component of plasma membrane	67	2.85E-04
	Extracellular space	64	3.66E-04
	Centrosome	23	1.09E-02
	Microtubule	22	5.53E-04
	Apical plasma membrane	20	1.48E-03
	Midbody	17	1.83E-06
	Kinetochore	16	1.77E-08

Enrichment analysis of biological processes, molecular function, and cellular component ( $P < 0.05$ ).



To confirm the validity and sensitivity of the 7-lncRNA model for predicting survival, we applied the model to risk assessment in patients with stage I, stage II, stage III, and HER2 positive BRCA. Patients were divided into high-risk and low-risk groups using a median risk score (value = 0.965). The Kaplan-Meier curve results showed that the high-risk groups of patients with stage I, stage II, stage III, and Her2-positive BRCA were closely associated with poor prognosis (Figures 4A–D). In addition, the ROC curve indicated that the AUC values of the model were 0.883, 0.708, 0.773, 0.774 at 3 years of OS (Figures 4E–H), indicating that the 7-lncRNA model we constructed had certain specificity and sensitivity in evaluating the prognosis of patients with BRCA.

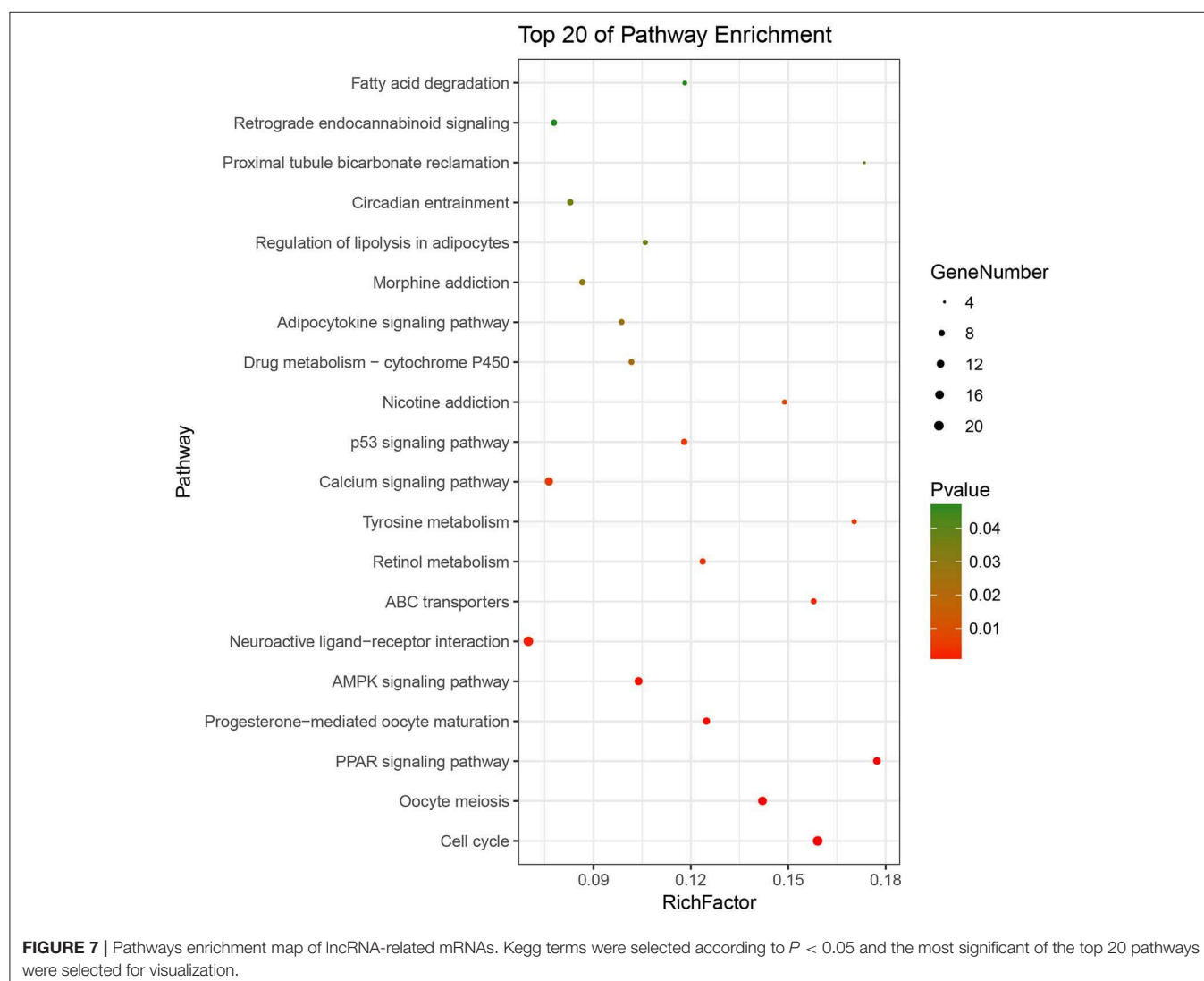
### Comprehensive Assessment of Model Predictive Performance and Routine Clinical Risk Factors

We compared the predictive performance of the 7-lncRNA model with conventional clinical risk factors, including age,

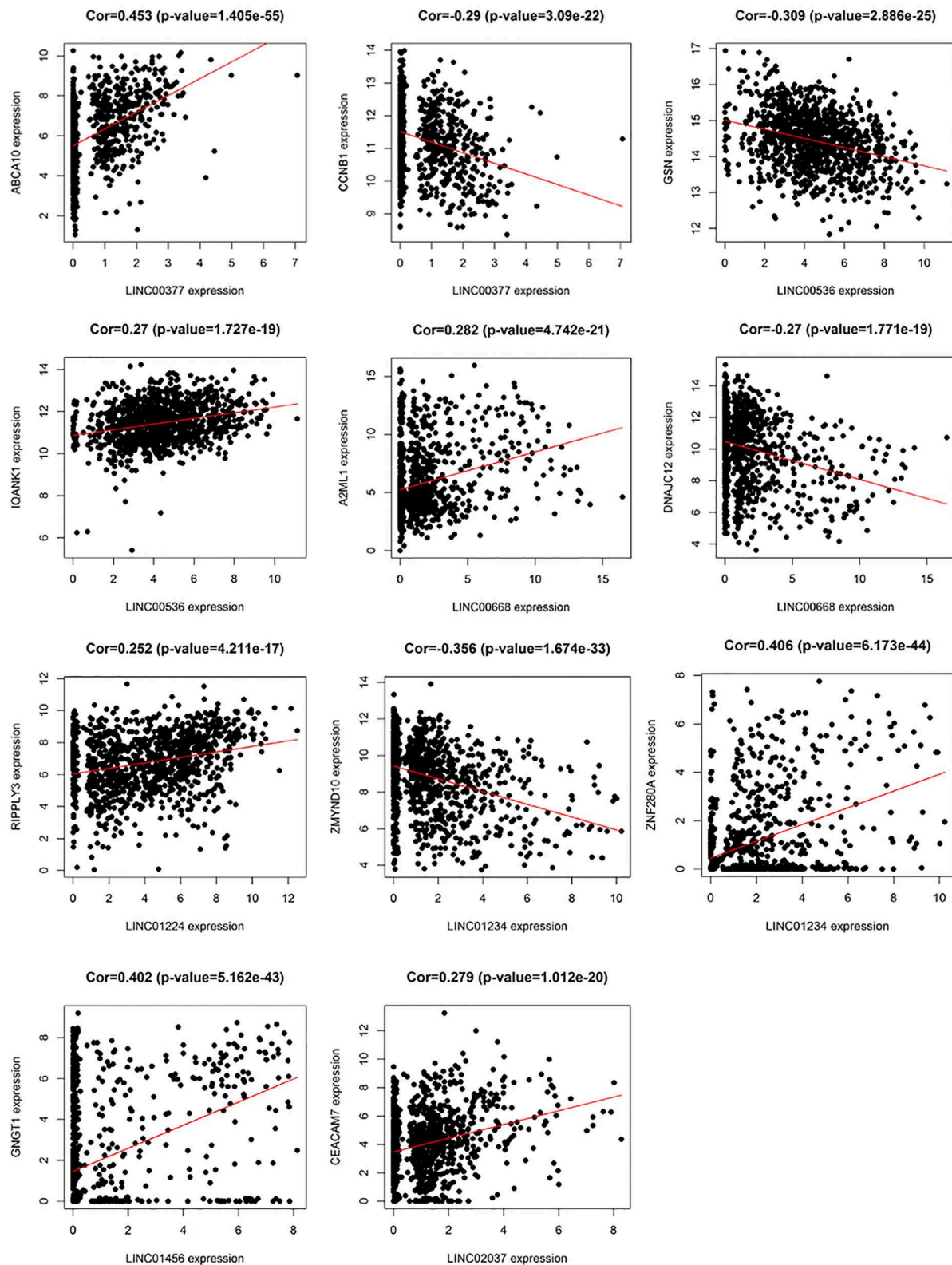
TNM, Stage, ER, PR, and HER2 status. Univariate analysis found that age, Stage, TNM stage, and predictive performance of the 7-lncRNA model were closely related to prognosis (Figure 5A). Further multivariate analysis found that predictive performance of age, T, M, and 7-lncRNA models could be used as independent prognostic factors to assess patient outcomes (Figure 5B).

### Functional Assessment of lncRNA-Related mRNA

Based on the BRCA-related lncRNA and mRNA expression data from the TCGA database, co-expression analysis was performed using the Pearson correlation with  $|COR| > 0.25$  and  $p < 0.05$  as the cutoff. A total of 592 mRNAs were found to be closely related to the 7 lncRNAs (Figure 6). The functions of the lncRNA-related mRNAs were determined using DAVID bioinformatics resources 6.8. The results of GO analysis mainly include Biological Process (BP), Molecular Function (MF), and Cellular Component (CC) (Table 3). We selected the most significant 10 enrichment results in the 3 parts for analysis. The process of enrichment in BP







**FIGURE 8 |** The relationship between 7 lncRNAs and related mRNAs (only the mRNAs that are most positively and negatively correlated with 7 lncRNAs are listed according to the correlation coefficient).

mainly includes cell division, cell proliferation, cell adhesion, and DNA replication, processes that are closely related to the growth and proliferation of tumor cells. The characteristics of enrichment in MF are mainly ATP binding, calcium-ion binding, chromatin binding, and protein-kinase binding, and those related to CC are plasma membrane, cytosol, integral component of plasma membrane, and the extracellular region. Five hundred ninety-two mRNAs were mainly enriched in 20 signaling pathways (Figure 7), including cell cycle, oocyte meiosis, and other cell division and proliferation pathways; and cancer-related signaling pathways, such as PPAR signaling pathway, neuroactive ligand-receptor interaction, and p53 signaling pathway.

In addition, we identified up-regulated and down-regulated mRNA with the highest correlation coefficient with 7 lncRNAs, and obtained a total of 11 mRNAs, including *ABCA10*, *CCNB1*, *GSN*, *IQANK1*, *A2ML1*, *DNAJC12*, *RIPPLY3*, *ZMYND10*, *ZNF280A*, *GNGT1*, and *CEACAM7* (Figure 8).

## DISCUSSION

BRCA is still one of the deadliest malignant tumors worldwide (16). Due to its complex molecular and cellular heterogeneity, the efficacy of existing breast cancer risk prediction models is unsatisfactory (17). High recurrence rate of breast cancer is one of the causes of high mortality. Therefore, in order to reduce mortality and improve the prognosis of BRCA, there is a need to construct a new breast cancer risk prediction model for clinical use. Clinicians should be able to develop individualized treatment plans for BRCA patients, establish strategies for prevention and early detection of BRCA recurrence, more frequently track high-risk populations, and perform regular clinical examinations for early diagnosis and recurrence of BRCA based on the predictions of the model.

In this study, BRCA-related differentially-expressed lncRNAs and mRNAs were obtained based on high-throughput RNA sequencing and clinical data of BRCA patients from the TCGA database. Subsequently, univariate and multivariate Cox analysis was performed to establish a risk model for predicting BRCA prognosis. Finally, BRCA prognostic risk prediction model was constructed using seven lncRNAs (LINC00377, LINC00536, LINC01224, LINC00668, LINC01234, LINC02037, and LINC01456). Applying the prognostic model to the TCGA BRCA dataset, breast cancer patients can be divided into high-risk and low-risk groups. The three- and 5-year AUC values for the time-dependent ROC curve were 0.771 and 0.734, respectively, indicating that the 7-lncRNA model has a good performance in survival prediction. By exploring the correlation between differentially-expressed lncRNAs and mRNAs, lncRNA-related mRNAs were identified to further study the function of the 7 lncRNAs and the molecular mechanisms involved in breast cancer progression.

In the current study, among these 7 lncRNAs, LINC00668, LINC01234, and LINC01456 have been shown to play a role in the pathogenesis and prognosis of cancer. Zhao et al. (18) showed that in laryngeal squamous cell carcinoma, the expression levels of LINC00668 were associated with age, pathological

differentiation degree, T stage, clinical stage, and cervical lymph node metastasis, and using a series of bioinformatics tools and *in vitro* experiments, proved that knockdown of LINC00668 can inhibit the proliferation, migration, and invasion ability of laryngeal squamous cell carcinoma cells. Zhang et al. (19) found that the expression of LINC00668 was negatively correlated with miR-297 expression in oral squamous cell carcinoma, and further found that LINC00668 promoted oral squamous cell carcinoma tumorigenesis via miR-297/VEGFA axis. In addition, Zhang et al. (20) found that knockdown of LINC00668 significantly inhibited the proliferation of gastric cancer cells *in vitro* and *in vivo*, and the significant increase in expression was associated with gastric cancer outcomes and prognosis. In our study, we found that the expression of LINC00668 is associated with *A2ML1* and *DNAJC12*; of which *A2ML1* has been shown to be closely related to the treatment of lung squamous cell carcinoma and can be used as a potential prognostic biomarker (21). Bubnov et al. (22) used genome-wide microarray Sentrix HumanWD-6V3 BeadChip (Illumina) to analyze gene expression pattern in 15 invasive adenocarcinoma samples and 15 healthy breast tissue samples, and found that *DNAJC12*, a member of the HSP40/DNAJ family, was significantly elevated. In addition, De Bessa et al. (23) found that *DNAJC12* is an estrogen target gene, its expression can be used as a marker of the ER activity, and that it may have a predictive value in response to hormonal therapy.

LINC01234 has been shown to be significantly associated with cancer treatment and prognosis in colon, gastric, and breast cancer (24–26). Chen et al. (27) found that LINC01234 expression was significantly upregulated in gastric cancer tissue and was associated with larger tumor size, advanced TNM stage, lymph node metastasis, and shorter survival. Furthermore, knockdown of LINC01234 induced apoptosis, arrested growth, and inhibited tumorigenesis in mouse xenografts. In our study, LINC01234 was found to be associated with *ZMYND10* and *ZNF280A*. *ZMYND10*, a candidate tumor suppressor gene, is frequently downregulated in nasopharyngeal carcinoma and many other tumors like gastric cancer, due to hypermethylation of the promoter (28). Functional evidence suggests that the *ZMYND10* gene inhibits tumor growth in animal experiments (29). According to reports, LINC01456 is a risk factor in ovarian cancer and is involved in the progression of ovarian cancer (30). In our study, we found a positive correlation between *GNGT1* and LINC01456 expression.

So far, no studies have reported any association between LINC00377, LINC00536, LINC01224, and LINC02037, and cancer. However, in our study, LINC00377 was found to be associated with expression of *ABCA10* and *CCNB1*. Ho et al. (31) found that *ABCA10* is involved in the pathogenesis of osteosarcoma, while Elsnerova et al. (32) found that the expression level of *ABCA10* was significantly associated with progression-free survival in ovarian cancer. *CCNB1* belongs to the highly conserved cyclin family and is significantly overexpressed in various cancer types. Ding et al. (33), showed that *CCNB1* had a significant predictive power in distant metastasis free survival, disease free survival, recurrence free survival, and overall survival of ER+ breast cancer patients. They also found that *CCNB1* was closely associated with hormone

therapy resistance. LINC00536 was found to be associated with expression of *GSN* and *IQANK1*, a ubiquitous actin filament-cleaving protein and a well-known downregulated target in breast tumors (34). *GSN* overexpression studies in MDA-MB231 and MCF-7 cells indicated that increased expression of *GSN* can result in changes in cell proliferation and cell-cycle progression (35). In addition, Chang et al. (36) showed that LINC01224 is associated with the expression of *RIPPLY3*, LINC02037 is associated with the expression of *CEACAM7*, and *CEACAM7* is found to be a potential prognostic biomarker for colorectal cancer.

The use of the TCGA database broadens the range of models for cancer survival prediction. Compared with the previously constructed breast cancer lncRNA prognosis model (37, 38), the patient's sample data in the TCGA database is large, and the clinical information is complete, and there is complete prognosis survival data of breast cancer patients. The ROC curve can be used to assess the specificity and sensitivity of the model (AUC >0.7 indicates that the model has good sensitivity). The 7-lncRNA prognostic model we developed has the potential to predict the prognosis of patients with BRCA and is specific and sensitive. In addition, whether univariate or multivariate Cox-regression analysis, the predictive performance of the 7-lncRNA model we constructed can be a good assessment of prognosis, further indicating the evaluation value of the model. In addition, as the lncRNAs used in the model have a predictive effect on the prognosis of patients with BRCA, further experimental studies can be conducted to investigate the role of these lncRNAs in the pathogenesis of BRCA in order to provide new ideas and insights for treatment. However, current research still has some limitations, we attempted to validate the predictive performance of the 7-lncRNA model in other large breast cancer data sets. Unfortunately, due to the limitations of the clinical mutation information of breast cancer and patient prognosis information, we did not find a data set that met the verification requirements. So it is

necessary to propose effective strategies such as including longer follow-up duration to validate the results and multiple regression modeling methods to improve the accuracy of the model.

## CONCLUSION

We constructed a 7-lncRNA prognostic model to reliably predict the prognosis of patients with BRCA, and these lncRNAs may play a role in the carcinogenesis of BRCA. Further functional studies are needed to elucidate the molecular mechanisms behind the roles of these lncRNAs in BRCA.

## DATA AVAILABILITY STATEMENT

This manuscript contains previously unpublished data. The name of the repository and accession number are not available.

## AUTHOR CONTRIBUTIONS

CS, HL, and CG conceived and designed the study. LL, JZ, and JY performed data analysis. CL, CZ, and FF contributed analysis tools. HL and CG wrote the paper.

## FUNDING

This work was supported by the grants from National Natural Science Foundation of China (81673799) and National Natural Science Foundation of China Youth Fund (81703915).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2019.01348/full#supplementary-material>

**Figure S1** | The risk prediction correlation analysis between the seven lncRNAs.

## REFERENCES

- Kim YA, Cho H, Lee N, Jung SY, Sim SH, Park IH, et al. Doxorubicin-induced heart failure in cancer patients: a cohort study based on the Korean National Health Insurance Database. *Cancer Med.* (2018) 7:6084–92. doi: 10.1002/cam4.1886
- Yang B, Chou J, Tao Y, Wu D, Wu X, Li X, et al. An assessment of prognostic immunity markers in breast cancer. *NPJ Breast Cancer.* (2018) 4:35. doi: 10.1038/s41523-018-0088-0
- Senkus E, Kyriakides S, Ohno S, Penault-Llorca F, Poortmans P, Rutgers E, et al. Primary breast cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol.* (2015) 26(Suppl. 5), v8–30. doi: 10.1093/annonc/mdv298
- Lou W, Liu J, Ding B, Xu L, Fan W. Identification of chemoresistance-associated miRNAs in breast cancer. *Cancer Manag Res.* (2018) 10:4747–57. doi: 10.2147/CMAR.S172722
- Tian T, Wang M, Zhu Y, Zhu W, Yang T, Li H, et al. Expression, clinical significance, and functional prediction of MNX1 in breast cancer. *Mol Ther Nucleic Acids.* (2018) 13:399–406. doi: 10.1016/j.omtn.2018.09.014
- Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet.* (2009) 10:57–63. doi: 10.1038/nrg2484
- Yan X, Hu Z, Feng Y, Hu X, Yuan J, Zhao SD, et al. Comprehensive genomic characterization of long non-coding RNAs across human cancers. *Cancer Cell.* (2015) 28:529–40. doi: 10.1016/j.ccell.2015.09.006
- Huarte M. The emerging role of lncRNAs in cancer. *Nat Med.* (2015) 21:1253–61. doi: 10.1038/nm.3981
- Mendell JT. Targeting a long noncoding RNA in breast cancer. *N Engl J Med.* (2016) 374:2287–9. doi: 10.1056/NEJMcibr1603785
- Shi X, Tan H, Le X, Xian H, Li X, Huang K, et al. An expression signature model to predict lung adenocarcinoma-specific survival. *Cancer Manag Res.* (2018) 10:3717–32. doi: 10.2147/CMAR.S159563
- Long J, Zhang L, Wan X, Lin J, Bai Y, Xu W, et al. A four-gene-based prognostic model predicts overall survival in patients with hepatocellular carcinoma. *J Cell Mol Med.* (2018) 22:5928–38. doi: 10.1111/jcmm.13863
- Liu SJ, Lim DA. Modulating the expression of long non-coding RNAs for functional studies. *EMBO Rep.* (2018) 19:e46955. doi: 10.15252/embr.201846955
- Tomczak K, Czerwinska P, Wiznerowicz M. The cancer genome atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol (Pozn).* (2015) 19:A68–77. doi: 10.5114/wo.2014.47136
- Yoneda K, Kuwata T, Chikaishi Y, Mori M, Kanayama M, Takenaka M, et al. Detection of circulating tumor cells with a novel microfluidic

- system in malignant pleural mesothelioma. *Cancer Sci.* (2018) 110:726–33. doi: 10.1111/cas.13895
15. Jiao X, Sherman BT, Huang da W, Stephens R, Baseler MW, Lane HC, et al. DAVID-WS: a stateful web service to facilitate gene/protein list analysis. *Bioinformatics.* (2012) 28:1805–6. doi: 10.1093/bioinformatics/bts251
  16. Desreux JAC. Breast cancer screening in young women. *Eur J Obstet Gynecol Reprod Biol.* (2018) 230:208–11. doi: 10.1016/j.ejogrb.2018.05.018
  17. Braden AM, Stankowski RV, Engel JM, Onitilo AA. Breast cancer biomarkers: risk assessment, diagnosis, prognosis, prediction of treatment efficacy and toxicity, and recurrence. *Curr Pharm Des.* (2014) 20:4879–98. doi: 10.2174/1381612819666131125145517
  18. Zhao L, Cao H, Chi W, Meng W, Cui W, Guo W, et al. Expression profile analysis identifies the long non-coding RNA landscape and the potential carcinogenic functions of LINC00668 in laryngeal squamous cell carcinoma. *Gene.* (2018) 687:47–55. doi: 10.1016/j.gene.2018.11.020
  19. Zhang CZ. Long intergenic non-coding RNA 668 regulates VEGFA signaling through inhibition of miR-297 in oral squamous cell carcinoma. *Biochem Biophys Res Commun.* (2017) 489:404–12. doi: 10.1016/j.bbrc.2017.05.155
  20. Zhang E, Yin D, Han L, He X, Si X, Chen W, et al. E2F1-induced upregulation of long noncoding RNA LINC00668 predicts a poor prognosis of gastric cancer and promotes cell proliferation through epigenetically silencing of CKIs. *Oncotarget.* (2016) 7:23212–26. doi: 10.18632/oncotarget.6745
  21. Zhang W, Cui Q, Qu W, Ding X, Jiang D, Liu H. TRIM58/cg26157385 methylation is associated with eight prognostic genes in lung squamous cell carcinoma. *Oncol Rep.* (2018) 40:206–16. doi: 10.3892/or.2018.6426
  22. Bubnov V, Moskalev E, Petrovskiy Y, Bauer A, Hoheisel J, Zaporozhan V. Hypermethylation of TUSC5 genes in breast cancer tissue. *Exp Oncol.* (2012) 34:370–2. Available online at: <http://dspace.nbuv.gov.ua/handle/123456789/139864>
  23. De Bessa SA, Salaorni S, Patrao DF, Neto MM, Brentani MM, Nagai MA. JDP1 (DNAJC12/Hsp40) expression in breast cancer and its association with estrogen receptor status. *Int J Mol Med.* (2006) 17:363–7. doi: 10.3892/ijmm.17.2.363
  24. Guo W, Wang Q, Zhan Y, Chen X, Yu Q, Zhang J, et al. Transcriptome sequencing uncovers a three-long noncoding RNA signature in predicting breast cancer survival. *Sci Rep.* (2016) 6:27931. doi: 10.1038/srep27931
  25. Gu J, Li Y, Fan L, Zhao Q, Tan B, Hua K, et al. Identification of aberrantly expressed long non-coding RNAs in stomach adenocarcinoma. *Oncotarget.* (2017) 8:49201–16. doi: 10.18632/oncotarget.17329
  26. Liu JX, Li W, Li JT, Liu F, Zhou L. Screening key long non-coding RNAs in early-stage colon adenocarcinoma by RNA-sequencing. *Epigenomics.* (2018) 10:1215–28. doi: 10.2217/epi-2017-0155
  27. Chen X, Chen Z, Yu S, Nie F, Yan S, Ma P, et al. Long Noncoding RNA LINC01234 functions as a competing endogenous RNA to regulate CBFβ expression by sponging miR-204-5p in gastric cancer. *Clin Cancer Res.* (2018) 24:2002–14. doi: 10.1158/1078-0432.CCR-17-2376
  28. Xiao K, Yu Z, Shi DT, Lei Z, Chen H, Cao J, et al. Inactivation of BLU is associated with methylation of Sp1-binding site of BLU promoter in gastric cancer. *Int J Oncol.* (2015) 47:621–31. doi: 10.3892/ijo.2015.3032
  29. Cheng Y, Ho RL, Chan KC, Kan R, Tung E, Lung HL, et al. Anti-angiogenic pathway associations of the 3p21.3 mapped BLU gene in nasopharyngeal carcinoma. *Oncogene.* (2015) 34:4219–28. doi: 10.1038/onc.2014.353
  30. Meng X, Jin-Cheng G, Jue Z, Quan-Fu M, Bin Y, Xu-Feng W. Protein-coding genes, long non-coding RNAs combined with microRNAs as a novel clinical multi-dimension transcriptome signature to predict prognosis in ovarian cancer. *Oncotarget.* (2017) 8:72847–59. doi: 10.18632/oncotarget.20457
  31. Ho XD, Phung P, Q Le V, H Nguyen V, Reimann E, Prans E, et al. Whole transcriptome analysis identifies differentially regulated networks between osteosarcoma and normal bone samples. *Exp Biol Med (Maywood).* (2017) 242:1802–11. doi: 10.1177/1535370217736512
  32. Elsnierova K, Mohelnikova-Duchonova B, Cerovska E, Ehrlichova M, Gut I, Rob L, et al. Gene expression of membrane transporters: importance for prognosis and progression of ovarian carcinoma. *Oncol Rep.* (2016) 35:2159–70. doi: 10.3892/or.2016.4599
  33. Ding K, Li W, Zou Z, Zou X, Wang C. CCNB1 is a prognostic biomarker for ER+ breast cancer. *Med Hypotheses.* (2014) 83:359–64. doi: 10.1016/j.mehy.2014.06.013
  34. Gur-Dedeoglu B, Konu O, Bozkurt B, Ergul G, Seckin S, Yulug IG. Identification of endogenous reference genes for qRT-PCR analysis in normal matched breast tumor tissues. *Oncol Res.* (2009) 17:353–65. doi: 10.3727/096504009788428460
  35. Chen ZY, Wang PW, Shieh DB, Chiu KY, Liou YM. Involvement of gelsolin in TGF-β1 induced epithelial to mesenchymal transition in breast cancer cells. *J Biomed Sci.* (2015) 22:90. doi: 10.1186/s12929-015-0197-0
  36. Chang J, Huang L, Cao Q, Liu F. Identification of colorectal cancer-restricted microRNAs and their target genes based on high-throughput sequencing data. *Onco Targets Ther.* (2016) 9:1787–94. doi: 10.2147/OTT.S93338
  37. Wang K, Li J, Xiong Y, Zeng Z, Zhang X, Li H. A potential prognostic long noncoding RNA signature to predict recurrence among ER-positive breast cancer patients treated with tamoxifen. *Sci Rep.* (2018) 8:3179. doi: 10.1038/s41598-018-21581-w
  38. Sun J, Chen X, Wang Z, Guo M, Shi H, Wang X, et al. A potential prognostic long non-coding RNA signature to predict metastasis-free survival of breast cancer patients. *Sci Rep.* (2015) 5:16553. doi: 10.1038/srep16553

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Li, Gao, Liu, Zhuang, Yang, Liu, Zhou, Feng and Sun. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# OSbrca: A Web Server for Breast Cancer Prognostic Biomarker Investigation With Massive Data From Tens of Cohorts

Zhongyi Yan<sup>1†</sup>, Qiang Wang<sup>1†</sup>, Xiaoxiao Sun<sup>1</sup>, Bingbing Ban<sup>1</sup>, Zhendong Lu<sup>1</sup>, Yifang Dang<sup>1</sup>, Longxiang Xie<sup>1</sup>, Lu Zhang<sup>1</sup>, Yongqiang Li<sup>1</sup>, Wan Zhu<sup>2</sup> and Xiangqian Guo<sup>1\*</sup>

<sup>1</sup> Cell Signal Transduction Laboratory, Department of Preventive Medicine, Bioinformatics Center, School of Basic Medical Sciences, School of Software, Institute of Biomedical Informatics, Henan University, Kaifeng, China, <sup>2</sup> Department of Anesthesia, Stanford University, Stanford, CA, United States

## OPEN ACCESS

### Edited by:

Aleix Prat,

Hospital Clínic de Barcelona, Spain

### Reviewed by:

Yoichi Naito,

National Cancer Center Hospital

East, Japan

Steven Narod,

University of Toronto, Canada

### \*Correspondence:

Xiangqian Guo

xqguo@henu.edu.cn

<sup>†</sup>These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to

Women's Cancer,

a section of the journal

Frontiers in Oncology

Received: 27 April 2019

Accepted: 15 November 2019

Published: 20 December 2019

### Citation:

Yan Z, Wang Q, Sun X, Ban B, Lu Z, Dang Y, Xie L, Zhang L, Li Y, Zhu W and Guo X (2019) OSbrca: A Web Server for Breast Cancer Prognostic Biomarker Investigation With Massive Data From Tens of Cohorts. *Front. Oncol.* 9:1349. doi: 10.3389/fonc.2019.01349

Potential prognostic mRNA biomarkers are exploited to assist in the clinical management and treatment of breast cancer, which is the first life-threatening tumor in women worldwide. However, it is technically challenging for untrained researchers to process high dimensional profiling data to screen and validate the potential prognostic values of genes of interests in multiple cohorts. Our aim is to develop an easy-to-use web server to facilitate the screening, developing, and evaluating of prognostic biomarkers in breast cancers. Herein, we collected more than 7,400 cases of breast cancer with gene expression profiles and clinical follow-up information from The Cancer Genome Atlas and Gene Expression Omnibus data, and built an Online consensus Survival analysis web server for Breast Cancers, abbreviated OSbrca, to generate the Kaplan–Meier survival plot with a hazard ratio and log rank *P*-value for given genes in an interactive way. To examine the performance of OSbrca, the prognostic potency of 128 previously published biomarkers of breast cancer was reassessed in OSbrca. In conclusion, it is highly valuable for biologists and clinicians to perform the preliminary assessment and validation of novel or putative prognostic biomarkers for breast cancers. OSbrca could be accessed at <http://bioinfo.henu.edu.cn/BRCA/BRCAList.jsp>.

**Keywords:** survival, breast cancer, prognosis, biomarker, OSbrca

## INTRODUCTION

Breast cancer is one of the leading cancers and the primary cause of mortality in women. The global burden of breast cancer is still increasing (1). It is predicted that by 2021, the incidence of breast cancer will increase to 85 per 100,000 women in China (2). Currently, clinicopathological risk factors are primarily used to estimate prognosis. These clinicopathological risks include stage, histological grade, tumor size, lymph node infiltrate, and so on (3). Molecular subtypes influence the survival of breast cancer. According to three protein expression statuses [estrogen receptor, progesterone receptor, and human epidermal growth factor receptor 2 (HER2)], breast cancer can be categorized into four classes: luminal A, luminal B, basal-like, and HER2+ (4). Because of the heterogeneity and survival difference of breast cancer, the utmost interests for researchers are how to validate the prognostic and predictive candidate genes

in appropriately powered breast cancer cohorts using the massive published expression levels of various genes profiles with clinical outcome.

So far, a number of poor clinical outcome associated genes have been identified. The most famous prognostic significance of breast cancer is the estrogen receptor gene, which is expressed in 50–70% of clinical tumor cases (5). Progesterone receptor and HER2 are two other important prognostic-related and predictive genes for breast cancer. In addition, a lot of new prognostic genes are exploited for diagnosing and curing breast cancer, such as breast cancer 1/2, TP53, cyclin D1, cyclin E, cathepsin D, cystatin E/M, and plexin B1 (6–8). Many studies showed that using multigenes as a panel of biomarkers may work more accurately to predict clinical outcome (9). Therefore, multivariate cohorts are needed to identify novel genes, and these genes need to be exploited to cure and evaluate prognosis of breast cancer.

By combining clinical follow-up data and high-throughput profiling data, we have reached a better understanding in the study of breast carcinoma. In this study, we collected the gene expression profiling data with follow-up information of breast cancers, which were mainly from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) database. Our aim is to provide a high powerful web server with massive data to generate survival plots to assess the relevance of the expression levels of interested genes on the clinical outcome for breast cancer patients. The Online consensus Survival analysis web server for Breast Cancers offers a web server to clinicians or non-bioinformatics researchers to appraise or exploit potential prognostic genes. Users can predict the prognostic potency of gene of interests using OSbrca.

## METHODS AND EXPERIMENT

### Data Collection

The gene expression profiling datasets for breast cancer were mainly composed of TCGA and GEO cohorts (Table 1) according to the following four criteria: (1) the cohort must have at least 50 breast cancer cases, (2) the cohort must contain individual clinical follow-up information, (3) the probe annotation should be completed or probe could be translated to gene symbol by ID conversion, such as DIVID, and (4) only platforms with more than 50 individual samples were selected if GEO cohorts having more than one platform.

### Development of OSbrca

The OSbrca server is deployed in a tomcat server as previously described with minor modification (10). In brief, front-end application was exploited in HTML and JSP to retrieve user inputs and display the output on the web page. Java and R were also used in the server application to control the analysis request and return the results. The gene expression profiles and clinical data were stored and managed by the SQL Server database. The R and SQL Server were linked by third middleware (The R packages, “RODBC” and “JDBC”). The R package “survminer” and “survival” generate Kaplan–Meier (KM) survival curves with log-rank *P*-value and calculate the hazard ratio (HR) with 95% confidence intervals (95%CI). The KM survival curves measure

the effect of genes on survival using breast cancer data (11). Log-rank test is the standard method of survival data comparison, which is widely used in survival analysis (12). HR and 95% confidence interval (95% CI) were calculated by univariate Cox regression analysis. OSbrca can be accessed in <http://bioinfo.henu.edu.cn/BRCA/BRCAList.jsp>.

## Collection and Authenticating Previously Reported Prognostic Biomarkers of Breast Cancer

To collect previously published biomarkers of breast cancer in the PubMed, three key words were used: breast cancer, prognostic, and biomarker. One hundred and twenty-eight previously identified prognostic biomarkers are listed in Table S1. To examine the performance of OSbrca, each reported prognostic biomarker was analyzed in OSbrca, by categorizing patients with “upper 25%” (the upper 25% expression vs. the bottom 75% expression). In addition, OSbrca is a web server for cross-validation of the potential prognostic biomarkers among tens of breast cancer cohorts. As a result, the methodology of validation in OSbrca includes two parts. First, we performed the validation of prognostic biomarkers between different breast cancer cohorts, and this independent validation between cohorts is of great importance for biomarker development; second, validation of previously reported prognostic biomarkers in OSbrca presented the reliability of OSbrca.

## RESULTS

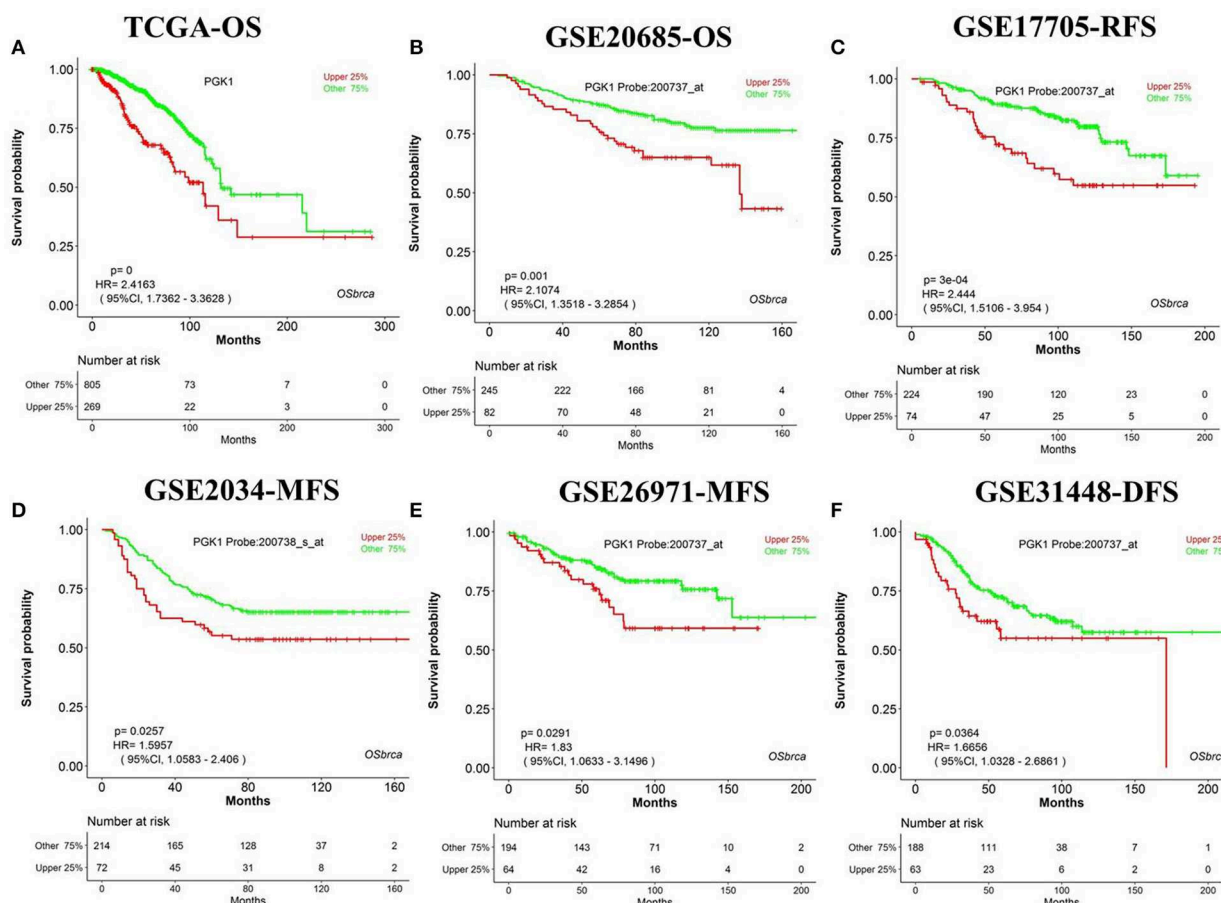
### Collection of Gene Expression Profiles With Clinical Follow-Up Information of Breast Cancer

Breast cancer is the leading mortality in women and is one of the most widely studied cancers. Thus, the urge for breast cancer patient is to exploit novel therapy target and prognostic biomarkers, which would offer the opportunities to assist the clinical management and treatment. However, it is technically challenging for untrained researchers to process the high dimensional profiling data to screen and validate the potential prognostic values of genes of interests in multiple cohorts. To build OSbrca, we have collected more than 7,400 samples of breast cancer expression profiles with clinical follow-up information, mainly obtained from TCGA (1,092 samples) and GEO cohorts (6,364 samples) (Table 1). OSbrca includes overall survival (OS, 3,786 patients from 23 cohorts), progression-free interval (1,096 patients only from TCGA cohort), progression-free survival (1,096 patients only from TCGA cohort), disease-specific survival (1,499 patients from three cohorts), disease-free interval (952 patients only from TCGA cohort), recurrence-free survival (RFS, 2,207 patients from 19 cohorts), disease-free survival (DFS, 1,632 patients from 11 cohorts), and metastasis-free survival (MFS, 2,508 patients from 16 cohorts). In other words, the OSbrca can predict those eight survival endpoints basing on breast cancer clinical information, such as RFS.

**TABLE 1 |** The basic information of The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) of breast cancer cohorts in Online consensus Survival analysis web server for Breast Cancers (OSbrca).

	Datasets	Cohort	Platform	Survival	No.#	References
1	NIH and NHGRI	TCGA	DCC	OS, PFI, PFS, DSS, DFI	1,083, 1,096, 1,096, 1,078,952	(19)
2	Chapel Hill	GSE10885	GPL1390	OS, RFS	94, 95	(20)
3	Chapel Hill	GSE10886	GPL1390	OS, RFS	178, 178	(21)
4	Chapel Hill	GSE10893	GPL1390	OS, RFS	155, 156	(22)
5	Leverkusen	GSE11121	GPL96	MFS	200	(23)
6	San Diego	GSE12093	GPL96	DFS	136	(24)
7	Rotterdam	GSE12276	GPL570	MFS	204	(25)
8	Carlsbad	GSE1379	GPL1223	DFS	60	(26)
9	Stockholm	GSE1456	GPL96	OS, MFS	159, 159	(27)
10	Woburn	GSE17705	GPL570	RFS	298	(28)
11	Chapel Hill	GSE18229	GPL887	OS, RFS	53, 53	(29)
12	Chapel Hill	GSE18229	GPL1390	OS, RFS	164, 165	(29)
13	San Diego	GSE2034	GPL96	MFS	286	(30)
14	Taipei	GSE20685	GPL570	OS, MFS	327, 327	(31)
15	Toronto	GSE20711	GPL570	OS, RFS	88, 88	(32)
16	Marseille	GSE21653	GPL570	DFS	248	(33, 34)
17	Helsinki	GSE24450	GPL6947	OS, DFS	183, 183	(35, 36)
18	New York	GSE2603	GPL96	MFS	82	(37)
19	Chapel Hill	GSE2607	GPL1390	OS, RFS	52, 52	(38)
20	Chapel Hill	GSE26338	GPL887	OS, RFS	56, 56	(39)
21	Chapel Hill	GSE26338	GPL1390	OS, RFS	173, 174	(39)
22	Köln	GSE26971	GPL96	MFS	258	(40)
23	Chapel Hill	GSE2741	GPL1390	OS, RFS	61, 61	(41)
24	Toronto	GSE2990	GPL96	RFS	109	(42)
25	Amsterdam	GSE31364	GPL14378	DFS	72	(43)
26	Durham	GSE3143	GPL8300	OS	158	(44)
27	Marseille	GSE31448	GPL570	DFS	251	(34)
28	Taipei	GSE33926	GPL7264	MFS	51	(45)
29	Singapore	GSE3494	GPL96	DSS	237	(46)
30	Chapel Hill	GSE3521	GPL1390	OS, RFS	84, 84	(47)
31	Chapel Hill	GSE35629	GPL1390	OS, RFS	53, 53	(48)
32	Milan	GSE37181	GPL6884	MFS	123	(49)
33	Bethesda	GSE37751	GPL6244	OS	61	(50)
34	Bethesda	GSE39004	GPL6244	OS	61	(50, 51)
35	Bangalore	GSE40206	GPL4133	MFS	61	(52)
36	Amsterdam	GSE41994	GPL16233	DFS	103	(53)
37	Dublin	GSE42568	GPL570	OS, RFS	104, 104	(54)
38	Winston-Salem	GSE45255	GPL96	DFS, MFS, DSS	94, 136, 134	(55)
39	Taipei	GSE48391	GPL570	DFS	81	(56)
40	Singapore	GSE4922	GPL96	DFS	249	(57)
41	Chicago	GSE5327	GPL96	MFS	58	(58)
42	Taipei	GSE53752	GPL7264	MFS	51	(45)
43	Chapel Hill	GSE6130	GPL1390	OS, RFS	86, 87	(59)
44	Toronto	GSE6532	GPL96	RFS, MFS	119, 239	(60–62)
45	Toronto	GSE7390	GPL96	OS, RFS, MFS	198, 198, 198	(63, 64)
46	Toronto	GSE9195	GPL570	RFS, MFS	77, 77	(61, 62)
47	Montpellier	GSE9893	GPL5049	OS, DFS	155, 155	(65)
Total##					7456	

# The number of samples only includes follow-up information; ## only the sum of the highest survival number. OS, overall survival; PFI, progression-free interval; PFS, progression-free survival; DSS, disease-specific survival; DFI, disease-free interval; RFS, recurrence-free survival; DFS, disease-free survival; MFS, metastasis-free survival.



**FIGURE 1 |** Validation previously reported gene *PGK1* in Online consensus Survival analysis web server for Breast Cancers (OSbrca). *PGK1* gene is high expressed in tumor tissue as a worse prognostic survival biomarker in breast cancer. (A) OS of TCGA, (B) OS of GSE20685, (C) RFS of GSE17705, (D) MFS of GSE2034, (E) MFS of GSE269721, (F) DFS of GSE31448. *PGK1*, phosphoglycerate kinase 1; OS, overall survival; DFS, disease-free survival; RFS, recurrence-free survival; MFS, metastasis-free survival.

## The Architecture of the OSbrca Web Server for Breast Cancer

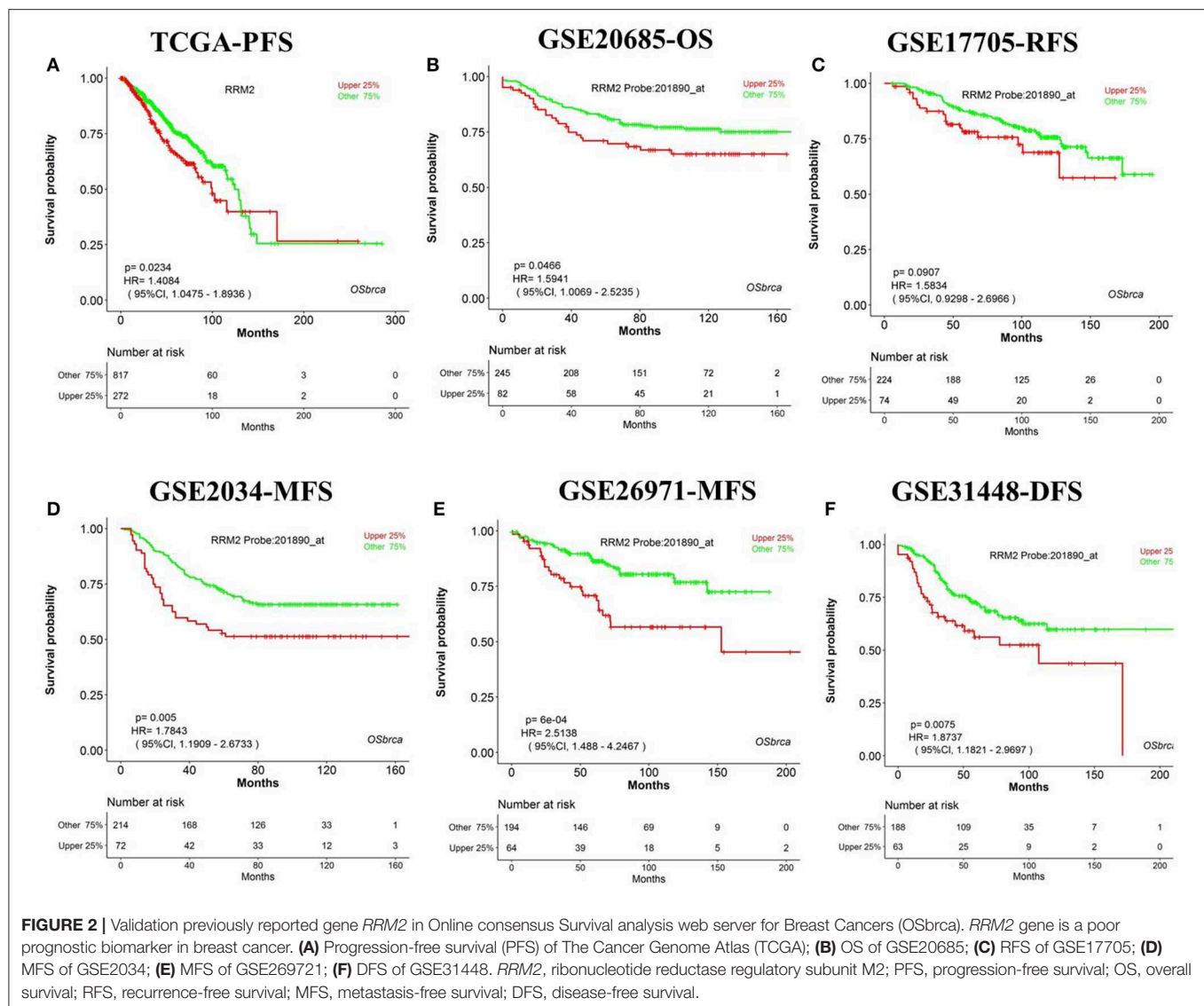
Based on the expression profiles and clinical outcome of breast cancers, OSbrca can determine the prognostic values of interested genes using KMPlot, HR, and log-rank *P*-value. OSbrca has implemented several optional clinical confounding factors, such as data source, age, stage, histological type, molecular subtype, survival, and ER/PgR/HER2 status. Users can select different cutoff, such as the upper 25%, for gene expression levels when categorizing the breast cancer population. The interface of the OSbrca is simple and friendly. Users could input the particular official gene symbol with all the default parameters and then click “Kaplan–Meier plot” button. The KMPlot with HR and log-rank *P*-value will be displayed on the output web page.

## Evaluation of the Previously Reported Prognostic Biomarkers of Breast Cancer in OSbrca

We have designed OSbrca to be a user-friendly and easy-to-use online web server to analyze and evaluate the prognostic values

of particular genes in 48 breast cancer cohorts using existing high-throughput profiling breast cancer data. To measure the performance and determine the reliability of OSbrca, we have collected previously published prognostic biomarkers of breast cancer (Table S1) and tested their prognostic potency in OSbrca. Fu et al. have demonstrated that *PGK1* was overexpressed in tumor tissue and was an indication of worse survival biomarker in breast cancer (13). Using OSbrca, we showed that *PGK1* gene was indeed a poor survival biomarker in breast cancer cohorts (top 6 samples): TCGA [OS, HR (95% CI) = 2.42 (1.74–3.36),  $P < 0.0001$ ], GSE20685 [OS, HR (95% CI) = 2.11 (1.35–3.39),  $P = 0.001$ ], GSE17705 [RFS, HR (95% CI) = 2.44 (1.51–3.95),  $P < 0.001$ ], GSE2034 [MFS, HR (95% CI) = 1.60 (1.06–2.41),  $P = 0.0257$ ], GSE269721 [MFS, HR (95% CI) = 1.83 (1.06–3.15),  $P = 0.0291$ ], and GSE31448 [DFS, HR (95% CI) = 1.67 (1.03–2.69),  $P = 0.0364$ ] (Figure 1). We also test another reported poor DFS biomarker *RRM2*. Figure 2 shows that *RRM2* gene was an indication of worse survival indicator in five out of six breast cancer cohorts (top 6 samples), except in the cohort of GSE17705 (Figure 2). One hundred and twenty-eight previous reported prognostic biomarkers were validated in OSbrca shown





in **Table S1**. Based on our studies using OSbrca, 62% analyzed biomarkers (79/128) showed consistent performance as reported in the literature, but some biomarkers showed contradictory outcomes to previous results. Taking the *AOCA1* gene as an another example, a previous study showed that the *AOCA1* gene could potentially predict a worse clinical prognosis in breast cancer (14). However, the analysis from OSbrca suggested that breast cancer patients with the overexpression of the *AOCA1* gene would potentially have a better clinical outcome (**Table S1**). In summary, all the results showed that the OSbrca web server is very reliable through validating previously reported biomarkers of breast cancer.

## DISCUSSION

Breast cancer is widely profiled by RNA-sequences and gene microarrays, such as TCGA. Thus, the core and focus issue

is how to excavate potential therapy targets and to develop prognostic biomarkers by possessing massive high-throughput profiles. Based on massive data of different cohorts, we integrated 48 cohorts of breast cancer datasets and established an online web server, named OSbrca. OSbrca implanted a selective set of clinical parameters, including tumor grade, age, status of ER/PgR/HER2, menopause status, and so on. The OSbrca could output the KMPlot with HR and log rank *P*-value for given genes in an interactive way. In addition, users can study genes in a particular country or race using OSbrca, such as Chinese breast cancer patients. Herein, we retrospectively validated the previously reported prognostic biomarkers of breast cancer. The results showed that most previous reported biomarkers could be identified by some different cohorts of OSbrca (**Figures 1, 2, and Table S1**). In addition, OSbrca is an across-validation web server used to exploit breast cancer biomarkers based on different independent cohorts of breast cancer. Cross-validation in OSbrca means that it is important to exploit prognostic

biomarkers among tens of breast cancer cohorts and also presents the reliability of OSbrca.

So far, there are some online prognostic websites for breast cancer, such as KM plotter (11), PROGgene (15), ITTACA (16), PrognoScan (17), OncoLnc, and GEPIA (18), but the size of datasets used in these tools is relatively small and limited compared to OSbrca. Specifically, OSbrca integrates 48 cohorts that contain more than 7,400 patients with RNA-sequencing and gene microarray data. It allows researchers to revisit previous protein biomarkers and exploit novel prognostic biomarkers. There are some limitations of this study, such as the loss of different platform integration, lacking noncoding gene information, which will be solved in the new-version of this tool. In addition, when new cohorts become available, we will update OSbrca in a timely manner.

In conclusion, the OSbrca web server integrates more than 7,400 follow-up breast samples and is highly valuable for researchers with a limited bioinformatics background to access and uncover prognostic-related biomarkers for breast cancer.

## DATA AVAILABILITY STATEMENT

The data for this manuscript can be accessed at OSbrca <http://bioinfo.henu.edu.cn/BRCA/BRCAList.jsp>. The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher.

## REFERENCES

1. Akram M, Iqbal M, Daniyal M, Khan AU. Awareness and current knowledge of breast cancer. *Biol Res.* (2017) 50:33. doi: 10.1186/s40659-017-0140-9
2. Ziegler RG, Anderson WF, Gail MH. Increasing breast cancer incidence in China: the numbers add up. *J Natl Cancer Inst.* (2008) 100:1339–41. doi: 10.1093/jnci/djn330
3. Tofigh A, Suderman M, Paquet ER, Livingstone J, Bertos N, Saleh SM, et al. The prognostic ease and difficulty of invasive breast carcinoma. *Cell Rep.* (2014) 9:129–42. doi: 10.1016/j.celrep.2014.08.073
4. Sana M, Malik HJ. Current and emerging breast cancer biomarkers. *J Cancer Res Therapeutics.* (2015) 11:508–13. doi: 10.4103/0973-1482.163698
5. Han SJ, Guo QQ, Wang T, Wang YX, Zhang YX, Liu F, et al. Prognostic significance of interactions between ER alpha and ER beta and lymph node status in breast cancer cases. *Asian Pacific J Cancer Prev.* (2013) 14:6081–4. doi: 10.7314/APJCP.2013.14.10.6081
6. Donepudi M, Kondapalli K, Amos S, Venkateshan P. Breast cancer statistics and markers. *J Cancer Res Therapeutics.* (2014) 10:506–11. doi: 10.4103/0973-1482.137927
7. Liu J, Lichtenberg T, Hoadley KA, Poisson LM, Lazar AJ, Cherniack AD, et al. An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics. *Cell.* (2018) 173:400–16.e11. doi: 10.1158/1538-7445.AM2018-3287
8. Rody A, Holtrich U, Gaetje R, Gehrmann M, Engels K, von Minckwitz G, et al. Poor outcome in estrogen receptor-positive breast cancers predicted by loss of *Plexin B1*. *Clin Cancer Res.* (2007) 13:1115–22. doi: 10.1158/1078-0432.CCR-06-2433
9. Alexe G, Dalgin GS, Scandfeld D, Tamayo P, Mesirov JP, DeLisi C, et al. High expression of lymphocyte-associated genes in node-negative HER2+ breast cancers correlates with lower recurrence rates. *Cancer Res.* (2007) 67:10669–76. doi: 10.1158/0008-5472.CAN-07-0539

## AUTHOR CONTRIBUTIONS

XG: research design. ZY, XS, LX, BB, ZL, YD, WZ, and LZ: collect and deal with data. ZY, LX, XS, YD, LZ, YL, and XG: draft of the manuscript. QW and XG: establish OSbrca Web Server. BB, XS, ZL, and YL: collect and validate previous reported biomarkers of breast cancer. ZY, QW, XS, LX, BB, YD, LZ, WZ, and XG: critical revision of the manuscript.

## FUNDING

This study was supported by the following funding: Kaifeng Science and Technology Major Project (18ZD008), National Natural Science Foundation of China (Nos. 81602362 and 81801569), Program for Science and Technology Development in Henan Province (Nos. 162102310391, 172102210187, and 192102310302), Program for Innovative Talents of Science and Technology in Henan Province (No. 18HASTIT048), Supporting grants of Henan University (Nos. 2015YBZR048 and B2015151), and Yellow River Scholar Program (No. H2016012).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2019.01349/full#supplementary-material>

10. Wang Q, Xie L, Dang Y, Sun X, Xie T, Guo J, et al. OSlms: a web server to evaluate the prognostic value of genes in leiomyosarcoma. *Front Oncol.* (2019) 9:190. doi: 10.3389/fonc.2019.00190
11. Györfy B, Lanczky A, Eklund AC, Denkert C, Budczies J, Li Q, et al. An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. *Breast Cancer Res Treat.* (2010) 123:725–31. doi: 10.1007/s10549-009-0674-9
12. Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. analysis and examples. *Br J Cancer.* (1977) 35:1–39. doi: 10.1038/bjc.1977.1
13. Fu D, He C, Wei J, Zhang Z, Luo Y, Tan H, et al. PGK1 is a potential survival biomarker and invasion promoter by regulating the HIF-1 $\alpha$ -mediated epithelial-mesenchymal transition process in breast cancer. *Cell Physiol Biochem.* (2018) 51:2434–44. doi: 10.1159/000495900
14. Lai Y-CC, Chang S-J, Kostoro J, Kwan A-L, Chai C-Y. Vascular adhesion protein-1 as indicator of breast cancer tumor aggressiveness and invasiveness. *APMIS.* (2018) 126:755–61. doi: 10.1111/apm.12885
15. Goswami CP, Nakshatri H. PROGgene: gene expression based survival analysis web application for multiple cancers. *J Clin Bioinformatics.* (2013) 3:22. doi: 10.1186/2043-9113-3-22
16. Elfilali A, Lair S, Verbeke C, La RP, Radvanyi F, Barillot E. ITTACA: a new database for integrated tumor transcriptome array and clinical data analysis. *Nucleic Acids Res.* (2006) 34:D613–6. doi: 10.1093/nar/gkj022
17. Mizuno H, Kitada K, Nakai K, Sarai A. PrognoScan: a new database for meta-analysis of the prognostic value of genes. *BMC Med Genomics.* (2009) 2:18. doi: 10.1186/1755-8794-2-18
18. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res.* (2017) 45:W98–102. doi: 10.1093/nar/gkx247

19. Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, Ellrott K, et al. The cancer genome atlas pan-cancer analysis project. *Nat Genet.* (2013) 45:1113–20. doi: 10.1038/ng.2764
20. Hennessy BT, Gonzalez-Angulo AM, Stemke-Hale K, Gilcrease MZ, Krishnamurthy S, Lee JS, et al. Characterization of a naturally occurring breast cancer subset enriched in epithelial-to-mesenchymal transition and stem cell characteristics. *Cancer Res.* (2009) 69:4116–24. doi: 10.1158/0008-5472.CAN-08-3441
21. 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) 27:1160–7. doi: 10.1200/JCO.2008.18.1370
22. Weigman VJ, Chao HH, Shabalín AA, He X, Parker JS, Nordgard SH, et al. Basal-like Breast cancer DNA copy number losses identify genes involved in genomic instability, response to therapy, and patient survival. *Breast Cancer Res Treat.* (2012) 133:865–80. doi: 10.1007/s10549-011-1846-y
23. Schmidt M, Böhm D, von Törne C, Steiner E, Puhl A, Pilch H, et al. The humoral immune system has a key prognostic impact in node-negative breast cancer. *Cancer Res.* (2008) 68:5405–13. doi: 10.1158/0008-5472.CAN-07-5206
24. Zhang Y, Sieuwerts AM, McGreevy M, Casey G, Cufer T, Paradiso A, et al. The 76-gene signature defines high-risk patients that benefit from adjuvant tamoxifen therapy. *Breast Cancer Res Treat.* (2009) 116:303–9. doi: 10.1007/s10549-008-0183-2
25. Bos PD, Zhang XH, Nadal C, Shu W, Gomis RR, Nguyen DX, et al. Genes that mediate breast cancer metastasis to the brain. *Nature.* (2009) 459:1005–9. doi: 10.1038/nature08021
26. Ma XJ, Wang Z, Ryan PD, Isakoff SJ, Barmettler A, Fuller A, et al. A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen. *Cancer Cell.* (2004) 5:607–16. doi: 10.1016/j.ccr.2004.05.015
27. Pawitan Y, Bjöhle J, Amler L, Borg AL, Eghazi S, Hall P, et al. Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts. *Breast Cancer Res.* (2005) 7:R953–64. doi: 10.1186/bcr1325
28. Symmans WF, Hatzis C, Sotiriou C, Andre F, Peintinger F, Regitnig P, et al. Genomic index of sensitivity to endocrine therapy for breast cancer. *J Clin Oncol.* (2010) 28:4111–9. doi: 10.1200/JCO.2010.28.4273
29. Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res.* (2010) 12:R68. doi: 10.1186/bcr2635
30. Wang Y, Klijn JG, Zhang Y, Sieuwerts AM, Look MP, Yang F, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary Breast Cancer. *Lancet.* (2005) 365:671–9. doi: 10.1016/S0140-6736(05)17947-1
31. Kao KJ, Chang KM, Hsu HC, Huang AT. Correlation of microarray-based breast cancer molecular subtypes and clinical outcomes: implications for treatment optimization. *BMC Cancer.* (2011) 11:143. doi: 10.1186/1471-2407-11-143
32. Dedeurwaerder S, Desmedt C, Calonne E, Singhal SK, Haibe-Kains B, Defrance M, et al. DNA methylation profiling reveals a predominant immune component in breast cancers. *EMBO Mol Med.* (2011) 3:726–41. doi: 10.1002/emmm.201100801
33. Sabatier R, Finetti P, Cervera N, Lambaudie E, Esterni B, Mamessier E, et al. A gene expression signature identifies two prognostic subgroups of basal breast cancer. *Breast Cancer Res Treat.* (2011) 126:407–20. doi: 10.1007/s10549-010-0897-9
34. Sabatier R, Finetti P, Adelaide J, Guille A, Borg JP, Chaffanet M, et al. Down-regulation of ECRG4, a candidate tumor suppressor gene, in human breast cancer. *PLoS ONE.* (2011) 6:e27656. doi: 10.1371/journal.pone.0027656
35. Muranen TA, Greco D, Fagerholm R, Kilpivaara O, Kampjarvi K, Aittomäki K, et al. Breast tumors from CHEK2 1100delC-mutation carriers: genomic landscape and clinical implications. *Breast Cancer Res.* (2011) 13:R90. doi: 10.1186/bcr3015
36. Heikkinen T, Greco D, Pelttari LM, Tommiska J, Vahteristo P, Heikkilä P, et al. Variants on the promoter region of PTEN affect breast cancer progression and patient survival. *Breast Cancer Res.* (2011) 13:R130. doi: 10.1186/bcr3076
37. Minn AJ, Gupta GP, Siegel PM, Bos PD, Shu W, Giri DD, et al. Genes that mediate breast cancer metastasis to lung. *Nature.* (2005) 436:518–24. doi: 10.1038/nature03799
38. Perreard L, Fan C, Quackenbush JF, Mullins M, Gauthier NP, Nelson E, et al. Classification and risk stratification of invasive breast carcinomas using a real-time quantitative RT-PCR assay. *Breast Cancer Res.* (2006) 8:R23. doi: 10.1186/bcr1399
39. Harrell JC, Prat A, Parker JS, Fan C, He X, Carey L, et al. Genomic analysis identifies unique signatures predictive of brain, lung, and liver relapse. *Breast Cancer Res Treat.* (2012) 132:523–35. doi: 10.1007/s10549-011-1619-7
40. Filipits M, Rudaš M, Jakesz R, Dubsy P, Fitzal F, Singer CF, et al. A new molecular predictor of distant recurrence in ER-positive, HER2-negative breast cancer adds independent information to conventional clinical risk factors. *Clin Cancer Res.* (2011) 17:6012–20. doi: 10.1158/1078-0432.CCR-11-0926
41. Weigelt B, Hu Z, He X, Livasy C, Carey LA, Ewend MG, et al. Molecular portraits and 70-gene prognosis signature are preserved throughout the metastatic process of breast cancer. *Cancer Res.* (2005) 65:9155–8. doi: 10.1158/0008-5472.CAN-05-2553
42. Sotiriou C, Wirapati P, Loi S, Harris A, Fox S, Smeds J, et al. Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. *J Natl Cancer Inst.* (2006) 98:262–72. doi: 10.1093/jnci/djj052
43. Molloy TJ, Roepman P, Naume B, van't Veer LJ. A prognostic gene expression profile that predicts circulating tumor cell presence in breast cancer patients. *PLoS ONE.* (2012) 7:e32426. doi: 10.1371/journal.pone.0032426
44. Bild AH, Yao G, Chang JT, Wang Q, Potti A, Chasse D, et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature.* (2006) 439:353–7. doi: 10.1038/nature04296
45. Kuo WH, Chang YY, Lai LC, Tsai MH, Hsiao CK, Chang KJ, et al. Molecular characteristics and metastasis predictor genes of triple-negative breast cancer: a clinical study of triple-negative breast carcinomas. *PLoS ONE.* (2012) 7:e45831. doi: 10.1371/journal.pone.0045831
46. Miller LD, Smeds J, George J, Vega VB, Vergara L, Ploner A, et al. An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. *Proc Natl Acad Sci USA.* (2005) 102:13550–5. doi: 10.1073/pnas.0506230102
47. Hu Z, Fan C, Livasy C, He X, Oh DS, Ewend MG, et al. A compact VEGF signature associated with distant metastases and poor outcomes. *BMC Med.* (2009) 7:9. doi: 10.1186/1741-7015-7-9
48. Ulirsch J, Fan C, Knafl G, Wu MJ, Coleman B, Perou CM, et al. Vimentin DNA methylation predicts survival in breast cancer. *Breast Cancer Res Treat.* (2013) 137:383–96. doi: 10.1007/s10549-012-2353-5
49. Callari M, Musella V, Di Buduo E, Sensi M, Miodini P, Dugo M, et al. Subtype-dependent prognostic relevance of an interferon-induced pathway metagene in node-negative. *Breast Cancer Mol Oncol.* (2014) 8:1278–89. doi: 10.1016/j.molonc.2014.04.010
50. Terunuma A, Putluri N, Mishra P, Mathe EA, Dorsey TH, Yi M, et al. MYC-driven accumulation of 2-hydroxyglutarate is associated with breast cancer prognosis. *J Clin Invest.* (2014) 124:398–412. doi: 10.1172/JCI71180
51. Tang W, Zhou M, Dorsey TH, Prieto DA, Wang XW, Ruppén E, et al. Integrated proteotranscriptomics of breast cancer reveals globally increased protein-mRNA concordance associated with subtypes and survival. *Genome Med.* (2018) 10:94. doi: 10.1186/s13073-018-0602-x
52. Bashir M, Damineni S, Mukherjee G, Kondaiah P. Activin-A signaling promotes epithelial-mesenchymal transition, invasion, and metastatic growth of breast cancer. *NPJ Breast Cancer.* (2015) 1:15007. doi: 10.1038/npjbcancer.2015.7
53. Jansen MP, Knijnenburg T, Reijm EA, Simon I, Kerkhoven R, Droog M, et al. Hallmarks of aromatase inhibitor drug resistance revealed by epigenetic profiling in breast cancer. *Cancer Res.* (2013) 73:6632–41. doi: 10.1158/0008-5472.CAN-13-0704
54. Clarke C, Madden SF, Doolan P, Aherne ST, Joyce H, O'Driscoll L, et al. Correlating transcriptional networks to breast cancer survival: a large-scale coexpression analysis. *Carcinogenesis.* (2013) 34:2300–8. doi: 10.1093/carcin/bgt208

55. Nagalla S, Chou JW, Willingham MC, Ruiz J, Vaughn JP, Dubey P, et al. Interactions between immunity, proliferation and molecular subtype in breast cancer prognosis. *Genome Biol.* (2013) 14:R34. doi: 10.1186/gb-2013-14-4-r34
56. Huang CC, Tu SH, Lien HH, Jeng JY, Huang CS, Huang CJ, et al. Concurrent gene signatures for han chinese breast cancers. *PLoS ONE.* (2013) 8:e76421. doi: 10.1371/journal.pone.0076421
57. Ivshina AV, George J, Senko O, Mow B, Putti TC, Smeds J, et al. Genetic reclassification of histologic grade delineates new clinical subtypes of breast cancer. *Cancer Res.* (2006) 66:10292–301. doi: 10.1158/0008-5472.CAN-05-4414
58. Minn AJ, Gupta GP, Padua D, Bos P, Nguyen DX, Nuyten D, et al. Lung metastasis genes couple breast tumor size and metastatic spread. *Proc Natl Acad Sci USA.* (2007) 104:6740–5. doi: 10.1073/pnas.0701138104
59. Mullins M, Perreard L, Quackenbush JF, Gauthier N, Bayer S, Ellis M, et al. Agreement in breast cancer classification between microarray and quantitative reverse transcription PCR from fresh-frozen and formalin-fixed, paraffin-embedded tissues. *Clin Chem.* (2007) 53:1273–9. doi: 10.1373/clinchem.2006.083725
60. Loi S, Haibe-Kains B, Desmedt C, Lallemand F, Tutt AM, Gillet C, et al. Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade. *J Clin Oncol.* (2007) 25:1239–46. doi: 10.1200/JCO.2006.07.1522
61. Loi S, Haibe-Kains B, Desmedt C, Wirapati P, Lallemand F, Tutt AM, et al. Predicting prognosis using molecular profiling in estrogen receptor-positive breast cancer treated with tamoxifen. *BMC Genomics.* (2008) 9:239. doi: 10.1186/1471-2164-9-239
62. Loi S, Haibe-Kains B, Majjaj S, Lallemand F, Durbecq V, Larsimont D, et al. PIK3CA mutations associated with gene signature of low mTORC1 signaling and better outcomes in estrogen receptor-positive breast cancer. *Proc Natl Acad Sci USA.* (2010) 107:10208–13. doi: 10.1073/pnas.0907011107
63. Desmedt C, Piette F, Loi S, Wang Y, Lallemand F, Haibe-Kains B, et al. Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series. *Clin Cancer Res.* (2007) 13:3207–14. doi: 10.1158/1078-0432.CCR-06-2765
64. Patil P, Bachant-Winner PO, Haibe-Kains B, Leek JT. Test set bias affects reproducibility of gene signatures. *Bioinformatics.* (2015) 31:2318–23. doi: 10.1093/bioinformatics/btv157
65. Chanrion M, Negre V, Fontaine H, Salvétat N, Bibeau F, Mac Grogan G, et al. A gene expression signature that can predict the recurrence of tamoxifen-treated primary breast cancer. *Clin Cancer Res.* (2008) 14:1744–52. doi: 10.1158/1078-0432.CCR-07-1833

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Yan, Wang, Sun, Ban, Lu, Dang, Xie, Zhang, Li, Zhu and Guo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Advantages of publishing in Frontiers



## OPEN ACCESS

Articles are free to read  
for greatest visibility  
and readership



## FAST PUBLICATION

Around 90 days  
from submission  
to decision



## HIGH QUALITY PEER-REVIEW

Rigorous, collaborative,  
and constructive  
peer-review



## TRANSPARENT PEER-REVIEW

Editors and reviewers  
acknowledged by name  
on published articles

## Frontiers

Avenue du Tribunal-Fédéral 34  
1005 Lausanne | Switzerland

Visit us: [www.frontiersin.org](http://www.frontiersin.org)

Contact us: [info@frontiersin.org](mailto:info@frontiersin.org) | +41 21 510 17 00



## REPRODUCIBILITY OF RESEARCH

Support open data  
and methods to enhance  
research reproducibility



## DIGITAL PUBLISHING

Articles designed  
for optimal readership  
across devices



## FOLLOW US

@frontiersin



## IMPACT METRICS

Advanced article metrics  
track visibility across  
digital media



## EXTENSIVE PROMOTION

Marketing  
and promotion  
of impactful research



## LOOP RESEARCH NETWORK

Our network  
increases your  
article's readership