



Genome Instability-Derived Genes Are Novel Prognostic Biomarkers for Triple-Negative Breast Cancer

Maoni Guo and San Ming Wang*

Cancer Centre and Institute of Translational Medicine, Faculty of Health Sciences, University of Macau, Macau, China

OPEN ACCESS

Edited by:

Ira Ida Skvortsova,
Innsbruck Medical University, Austria

Reviewed by:

Lili Zhang,
Beijing Hospital, China
Song Guo,
Shanghai Institute of Nutrition
and Health (CAS), China

*Correspondence:

San Ming Wang
sanmingwang@um.edu.mo

Specialty section:

This article was submitted to
Molecular and Cellular Oncology,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 27 April 2021

Accepted: 10 June 2021

Published: 12 July 2021

Citation:

Guo M and Wang SM (2021)
Genome Instability-Derived Genes Are
Novel Prognostic Biomarkers
for Triple-Negative Breast Cancer.
Front. Cell Dev. Biol. 9:701073.
doi: 10.3389/fcell.2021.701073

Background: Triple-negative breast cancer (TNBC) is an aggressive disease. Recent studies have identified genome instability-derived genes for patient outcomes. However, most of the studies mainly focused on only one or a few genome instability-related genes. Prognostic potential and clinical significance of genome instability-associated genes in TNBC have not been well explored.

Methods: In this study, we developed a computational approach to identify TNBC prognostic signature. It consisted of (1) using somatic mutations and copy number variations (CNVs) in TNBC to build a binary matrix and identifying the top and bottom 25% mutated samples, (2) comparing the gene expression between the top and bottom 25% samples to identify genome instability-related genes, and (3) performing univariate Cox proportional hazards regression analysis to identify survival-associated gene signature, and Kaplan–Meier, log-rank test, and multivariate Cox regression analyses to obtain overall survival (OS) information for TNBC outcome prediction.

Results: From the identified 111 genome instability-related genes, we extracted a genome instability-derived gene signature (GIGenSig) of 11 genes. Through survival analysis, we were able to classify TNBC cases into high- and low-risk groups by the signature in the training dataset (log-rank test $p = 2.66e-04$), validated its prognostic performance in the testing (log-rank test $p = 2.45e-02$) and Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) (log-rank test $p = 2.57e-05$) datasets, and further validated the predictive power of the signature in five independent datasets.

Conclusion: The identified novel signature provides a better understanding of genome instability in TNBC and can be applied as prognostic markers for clinical TNBC management.

Keywords: TNBC, mutation, copy number variation, genome instability, prognosis

INTRODUCTION

Triple-negative breast cancer (TNBC) accounts for ~15% of all breast cancer cases and is characterized by the absence of estrogen receptor (ER), progesterone receptor (PR), and epidermal growth factor receptor 2 (HER2) (Oakman et al., 2010). TNBC often occurs at a young age, highly aggressive, and metastatic with a poor prognosis (Haffty et al., 2006; Dent et al., 2007). A part of TNBC is related with the germline mutation in *BRCA1* and *BRCA2*, but the cause for most of the TNBC remains unclear (Song et al., 2014). With its unusual clinic outcome, a TNBC-specific prognostic signature will be highly valuable for clinical management of TNBC cases (Mersin et al., 2008; Tan and Swain, 2008).

Genomic instability is a major hallmark in cancer and an important prognostic factor associated with cancer progression and survival (Suzuki et al., 2003; Negrini et al., 2010). Genome instability-associated signatures have been identified in certain types of cancer. For example, a genomic instability-derived three-miRNA signature was used as a risk predictor for invasive breast cancer (Bao et al., 2021); a 12-genomic instability-derived gene expression signature was identified as a clinical outcome predictor for breast cancer (Habermann et al., 2009). Genome instability-related mutation and copy number variation (CNV) were also identified in TNBC (Schmitt et al., 2012). For instance, germline mutations in *BRCA1*, *BRCA2*, *ATM*, *PALB2*, *RAD51D*, and *RAD50* disrupted DNA damage repair pathways in TNBC (Wu et al., 2019); *FOSL1* had significantly higher CNV gains in TNBC than in other types of breast cancer (Serino et al., 2019); *PIK3CA* had high mutation frequency and copy number gains and highly ethnic-specific in TNBC (Jiang et al., 2019); somatic mutation and CNV-derived genomic metrics were significantly associated with immune prognostic category in TNBC (Karn et al., 2017). Furthermore, somatic mutations and CNVs were associated with dysregulation of multiple genes in TNBC. For example, mutations in *MYH9* and *HERC2* were both associated with lower lymphocyte-specific kinase (LCK) metagene expression in TNBC (Safonov et al., 2017); *JAK2* and *PD-L1* amplifications upregulated *PD-L1* expression by disturbing the JAK/STAT1 pathway in TNBC (Chen M. et al., 2018); high expression of *PD-L1* was associated with significant *CD274* gene copy number gain in TNBC (Guo et al., 2016). However, the clinical impact of these abnormalities as prognostic markers in TNBC remains largely unclear.

We hypothesized that there could be a genome instability-derived signature involved in the tumorigenesis and development of TNBC. We further reasoned that genomic instability-derived somatic mutation and CNV in TNBC could disturb gene expression in TNBC; therefore, expression difference in TNBC could be used as prognostic markers to predict clinical outcome of TNBC.

In this study, we first calculated the accumulative counts of somatic mutation and CNV in TNBC cases and selected the top and bottom 25% of the ranked cases. We then identified 111-genomic instability-derived genes to divide the cases into genomic unstable (GU) and genomic stable (GS) groups. Furthermore, we identified a genome instability-derived gene signature (GIGenSig) of 11 genes to classify TNBC cases

into high- and low-risk groups. We validated the results using multiple independent TNBC datasets. Our study provides a GIGenSig as a prognosis marker to predict the clinical outcome of TNBC.

MATERIALS AND METHODS

Datasets Used for the Study

We downloaded the METABRIC (Molecular Taxonomy of Breast Cancer International Consortium) breast cancer datasets (Pereira et al., 2016) from the cBioPortal database¹, including clinical information, gene expression, somatic mutation, and CNV data. The expression profile was processed as log intensity level of Illumina Human v3 microarray. We also downloaded the version 19 gene annotation file from the GENCODE database². Then, we randomly divided the TNBC cases into the training dataset and the testing dataset with the same size of living and deceased overall survival (OS) status in each dataset. The training dataset consisting of 150 TNBC samples was used to identify the prognostic signature and build the prognostic risk model; the testing dataset consisting of 149 TNBC patients was used to validate the prognostic model. Clinical characteristics for the training, testing, and METABRIC datasets are summarized in **Table 1**. The Cancer Genome Atlas (TCGA³), Shanghai TNBC data⁴ (Jiang et al., 2019; Goldman et al., 2020), and three additional independent datasets of GSE21653, GSE31448, and GSE25066 from the GEO database^{5,6,7} were used to validate the performance of the prognostic risk model (Hatzis et al., 2011; Sabatier et al., 2011a,b). For TCGA and Shanghai RNAseq datasets, we downloaded or processed the gene-level transcription estimates in $\log_2(x + 1)$ transformed RSEM normalized count. For the other three GEO microarray datasets, they were processed using the robust multichip average (RMA) algorithm for background adjustment (Irizarry et al., 2003a,b), and the Affymetrix GeneChip probe-level data were \log_2 transformed. The platform information for Affymetrix Human Genome U133 Plus 2.0 Array was downloaded from the Affymetrix website⁸. Gene expression data from the Affymetrix-based expression profiling were obtained by repurposing microarray probes based on the platform information and the gene annotation file from the GENCODE database (release 19, see text footnote 2).

Identification of Genome Instability-Related Genes in TNBC

To identify genome instability-related genes in TNBC, we first processed gene expression and genomic alteration profiles. For gene expression profile, we extracted the expression log intensity

¹<https://www.cbioportal.org/>

²<https://www.genecodegenes.org>

³<http://xena.ucsc.edu/>

⁴<https://www.ncbi.nlm.nih.gov/sra/?term=SRP157974>

⁵<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE21653>

⁶<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE31448>

⁷<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE25066>

⁸<http://www.affymetrix.com>

TABLE 1 | Clinical information for triple-negative breast cancer (TNBC) patient datasets used in this study.

Characteristics		Training dataset (N = 150)	Testing dataset (N = 149)	Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) dataset (N = 299)	p-Value*
Age (%)	Age < 55	68(45.33)	75(50.34)	143(47.83)	0.687
	Age ≥ 55	82(54.67)	74(49.66)	156(52.17)	
Menopausal_status (%)	Pre	49(32.67)	58(38.93)	107(35.79)	0.529
	Post	101(67.33)	91(61.07)	192(64.21)	
Tumor stage (%)	I/II	100(66.67)	92(61.74)	192(64.21)	0.532
	III/IV	16(10.67)	9(6.04)	25(8.36)	
	Unknown	34(22.67)	48(32.21)	82(27.42)	
Grade (%)	G1/2	17(11.33)	22(14.77)	39(13.04)	0.664
	G3	132(88.00)	125(83.89)	257(85.95)	
	Unknown	1(0.67)	2(1.34)	3(1.01)	
OS_status (%)	Living	69(46.00)	69(46.31)	138(46.15)	0.999
	Deceased	81(54.00)	80(53.69)	161(53.85)	

*p-value: Chi square test.

levels (Illumina Human v3 microarray) for 16,331 protein-coding genes; for mutations, silent mutations were removed; for CNV profile, we only retained high-level amplification and homozygous deletion evaluated by GISTIC2 segment (Mermel et al., 2011). Then, we integrated TNBC gene expression and genomic alteration data as shown in **Supplementary Figure 1**: (1) extracted gene expression, somatic mutation, and CNV profiles in the 299 TNBC cases; (2) constructed a binary matrix by integrating somatic mutation and CNV profiles; (3) calculated the accumulated alterations for each case; (4) took the top 25% and the bottom 25% of alternated cases as GU group and GS group; (5) compared the gene expression between GU and GS groups by using the R package “limma”; and (6) identified genome instability-related genes with a Benjamini and Hochberg (BH) adjusted p -value < 0.05 and logFC (fold-change) > 1 or < -1 between GU group and GS group.

Functional Enrichment Analysis

We applied enrichGO and enrichKEGG functions in the Bioconductor package “clusterProfiler” to identify the functions and pathways of the genome instability-related genes (Yu et al., 2012). We also performed Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) functional enrichment annotation using the Database for Annotation, Visualization and Integrated Discovery (DAVID tool⁹, version 6.8) (Huang da et al., 2009). The Benjamini p -adjust < 0.05 was considered as statistically significant in “clusterProfiler” and p -value < 0.05 for DAVID analyses.

Statistical Analysis

To identify the genes predictive for TNBC OS, a univariate Cox proportional hazards regression analysis was performed to evaluate the relationship between the expression level of each gene and patient OS in the training dataset. Only the genes with a p -value < 0.05 were taken as statistically significant survival predictors. To construct a predictive model, a multivariate Cox

regression model was applied for these selected genes with OS in the training dataset. A risk score formula was built to evaluate the risk of each patient to develop TNBC as follows:

$$\text{Overall risk score (ORS)} = \sum_{i=1}^N (\text{Exp}_i \times \text{Coe}_i)$$

where, N is the number of prognostic genes, Exp is the gene expression value, and Coe is the estimated regression coefficient of the gene. A risk score for each patient was calculated by including the expression values of each selected gene, weighed by their estimated regression coefficients in the multivariate Cox regression analysis. The patients were divided into high- and low-risk groups using the median of the risk scores as the threshold. The receiver operating characteristic (ROC) curves were used to compute the sensitivity and specificity of overall prediction of the selected gene expression-based overall risk scores (ORSs) using the R package “survivalROC.” The area under curve (AUC) value was also calculated. The Kaplan–Meier method was applied to generate OS curves, and the log-rank test was used to assess the differences in OS between the high- and low-risk groups using the R package “survival.” Additionally, univariate and multivariate Cox proportional hazards regression, and data stratification analyses were performed to test whether the ORS was independent of other clinical features. Statistical significance was based on p -value < 0.05 and 95% confidence interval (CI) estimates.

To evaluate the performance of the risk model prediction, we randomly chose samples from the high- and low-risk group and trained a support vector machine (SVM) classifier based on the expression level of the selected genes in the risk model using the R package “e1071.” The 10-fold cross-validation method was used to evaluate the performance of the classifier. Plots of the ROC curve of the classifier and the calculation of the AUC were fulfilled using the R verification package. In addition, Chi-square test and Wilcoxon rank-sum test were also used in the study, and a p -value < 0.05 was considered as statistically significant. In differential expression analysis, the genes with the cutoff of

⁹<http://david.abcc.ncifcrf.gov/>

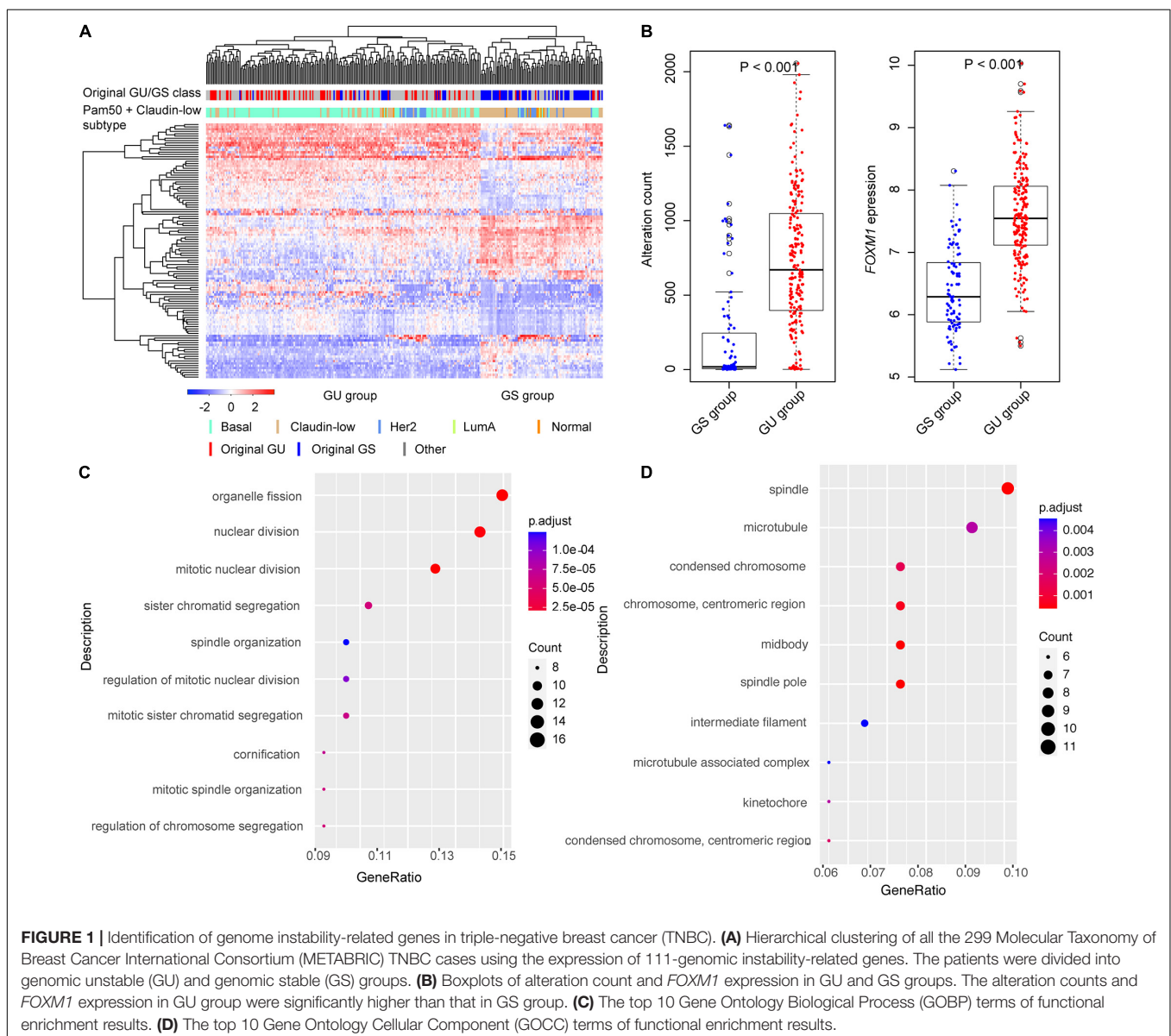
p -value < 0.05 and $\log_{2}FC > 1$ or < -1 between the two groups were regarded as statistically significant. All statistical analyses were performed using R version 3.6.3.

RESULTS

Identification of Genome Instability-Related Genes in TNBC

We identified 299 TNBC samples from the METABRIC breast cancer dataset and performed a systematic analysis (Supplementary Figure 1). To identify the genes associated with genomic instability, we calculated the cumulative count of alterations including somatic mutations and CNVs in each patient and sorted these counts in decreased order. The top 25% of TNBC patients ($n = 75$) were named as the GU group

and the bottom 25% ($n = 75$) as the GS group. We performed differential expression analysis between the GU and GS groups. We identified 111 differentially expressed genes between the two groups, 63 upregulated and 48 downregulated in the GU group (Supplementary Table 1). We performed unsupervised hierarchical clustering analysis for all 299 TNBC samples by the 111 genes and compared our GU/GS groups with the PAM50 and claudin-low subtype available from the METABRIC dataset. The results showed that 69.1% (143/207) of the samples in the GU group were classified as basal subtype and 65.2% (60/92) samples in the GS group as the claudin-low subtype (Figure 1A). We found that all patients were classified into either the GU group or the GS group, in which the cumulative alterations in the GU group were significantly higher than that of the GS group (Figure 1B, Wilcoxon test $p < 0.001$). We further compared the expression level of *FOXM1*, a genome



instability-related driver gene (Teh et al., 2010; Kim et al., 2013; Senfter et al., 2019), between the GU and GS groups. We found that the expression of *FOXM1* in the GU group was significantly higher than that in the GS group (Figure 1B, Wilcoxon test $p < 0.001$) and observed the expression of the classical proliferation gene *MKI67* was significantly higher in the GU group than that in the GS group (Supplementary Figure 2A, Wilcoxon test $p < 0.001$). To validate our defined GU/GS groups, we performed clustering analysis for the 111 genes in the 235 Shanghai TNBC samples and found that all samples were also significantly classified into the GU (191/235, 81.3%) and GS (44/235, 18.7%) groups (Supplementary Figure 3A). Besides, we compared the homologous recombination deficiency (HRD) levels between the GU/GS groups in the Shanghai dataset and observed that HRD level in the GU group was significantly higher than that in the GS group (Supplementary Figure 3B). We performed the same analysis using top 10% and bottom 10% as the cutoff for the GU/GS groups and received similar results from using the top 25% and bottom 25% as the cutoff (Supplementary Figure 2B).

To test whether these 111 differentially expressed genes were involved in important biological processes and pathways associated with genome instability, we performed functional enrichment analysis using R package “clusterProfiler” with BH adjustment for multiple testing. We identified multiple pathways associated with genomic instability, such as mitotic nuclear division, organelle fission, nuclear division, sister chromatid segregation, regulation of chromosome segregation, etc. (Figures 1C,D). We also performed DAVID analysis using the 111 genes. The results also showed that these genes were largely involved in cell cycle process, cell proliferation, and immune response (Supplementary Figure 4A), and some genes (such as *SFRP4*, *PRKCB*, *FZD9*, and *RAC2*) were known in involving TNBC development (Supplementary Figure 4B). The results highlight that the 111 differentially expressed genes were involved in tumorigenesis and development process of TNBC.

Identification of GIGenSig for Prognostic Prediction

To explore the potential prognostic value of the above genome instability-related genes, we divided all the METABRIC TNBC cases into two subsets, the training dataset ($n = 150$) and the testing dataset ($n = 149$). To identify prognostic-associated genes, we conducted univariate Cox regression analysis to calculate the relationship between the 111 gene expression and OS in the training dataset. The result showed that 11 of the genes were associated with OS in TNBC (Table 2, $p < 0.05$). We named these 11 genes as GIGenSig. To evaluate the prognostic potential of GIGenSig, we constructed a prognostic risk model for OS based on the expression of GIGenSig and coefficients of multivariate Cox analysis: $ORS = (0.137 \times PRKCB \text{ expression}) + (0.037 \times TFF3 \text{ expression}) + (-0.008 \times ART3 \text{ expression}) + (-0.071 \times CD52 \text{ expression}) + (-0.030 \times CD79A \text{ expression}) + (-0.155 \times FZD9 \text{ expression}) + (-0.010 \times GABRP \text{ expression}) + (-0.145 \times IRF8 \text{ expression}) + (-0.187 \times ITM2A \text{ expression}) + (-0.038 \times SOX10 \text{ expression}) + (-0.072 \times VGLL1 \text{ expression})$. Among the

GIGenSig, the coefficients of *PRKCB* and *TFF3* were positive, suggesting that they were risk factors for TNBC, and their high expression was associated with poor survival of TNBC. In contrast, the coefficients for the other nine genes (*ART3*, *CD52*, *CD79A*, *FZD9*, *GABRP*, *IRF8*, *ITM2A*, *SOX10*, and *VGLL1*) were negative, suggesting that these were protective factors, and their higher expression was associated with better survival.

Based on ORS values, TNBC patients in the training datasets were classified into two groups by their median ORS (-4.602), named as high-risk group and low-risk group. Survival plot showed that the OS survival in the low-risk group was significantly better than those of the high-risk group (Figure 2A, log-rank test $p = 2.66e-04$; HR = 2.718, 95% CI: 1.699–4.350), the 5-year survival rate in the low-risk group (73%) was higher than that in high-risk group (48%), and the 5-year ROC curve analysis provided an AUC of 0.648 (Figure 2B). In addition, SVM and 10-fold cross-validation showed that our risk model was robust to classify TNBC patients into high- and low-risk groups (Supplementary Figure 5A, AUC = 0.987, $p = 7.62e-33$).

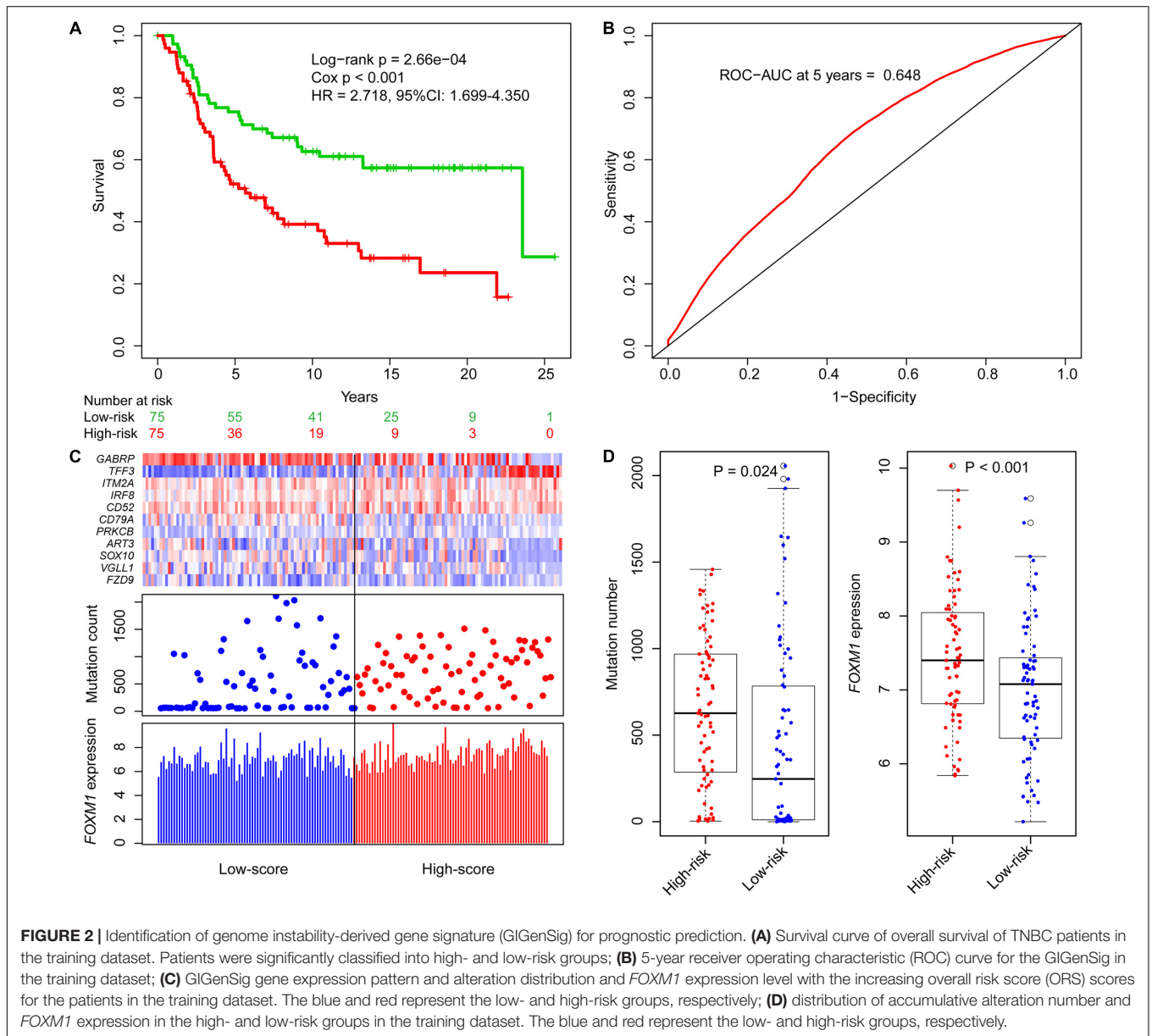
We ranked the ORS for patients in the training dataset to explore the differences of GIGenSig expression, alteration count, and *FOXM1* expression between low score and high score groups (Figure 2C). Clustering analysis showed that *PRKCB* and *TFF3* were upregulated in the high score group, whereas the other nine genes were upregulated in the low score group (Figure 2C). The differences of alteration count and *FOXM1* expression were both significant between the high-risk and low-risk groups (Figure 2D). The count of alterations in the high-risk group was significantly higher than that in the low-risk group (Figure 2D, Wilcoxon test $p = 0.024$). Additionally, *FOXM1* had significantly higher expression in the high-risk group than in the low-risk group (Figure 2D, Wilcoxon test $p < 0.001$).

Validation of GIGenSig for Prognostic Prediction in Testing and Molecular Taxonomy of Breast Cancer International Consortium Datasets

To explore the prognostic performance of GIGenSig, we tested it using the testing dataset of 149 TNBC cases. Based on the ORS cutoff in the training dataset, the cases in the testing dataset were classified into high-risk group ($n = 73$) and low-risk group ($n = 76$). The OS of the low-risk group was significantly higher than that of the high-risk group (Figure 3A, log-rank test $p = 2.45e-02$; HR = 1.820, 95% CI: 1.099–3.023). Similarly, the 5-year survival rate in the low-risk group (72%) was also higher than that in the high-risk group (60%), and the 5-year ROC analysis yielded an AUC of 0.607 (Supplementary Figure 6A). Additionally, the SVM and 10-fold cross-validation showed that our risk model was robust to classify TNBC patients into high- and low-risk groups in the testing dataset (Supplementary Figure 5B, AUC = 0.980, $p = 8.64e-32$). We also displayed the clustering of GIGenSig, alteration count, and *FOXM1* expression level according to the increasing order of the ORS for each patient in the testing dataset (Figure 3C). The alteration count and *FOXM1* expression level were both significantly higher in the high-risk group than in the low-risk group (Figure 3E, Wilcoxon

TABLE 2 | Univariate Cox regression analysis for the 11 of 111-genome instability-related genes associated with overall survival in TNBC.

Gene	Genomic location	Coefficient	HR	95% CI	p-Value
<i>ART3</i>	chr4:76932337–77033955	-0.126	0.882	0.778–0.999	0.048
<i>CD52</i>	chr1:26644448–26647014	-0.171	0.843	0.715–0.993	0.042
<i>CD79A</i>	chr19:42381190–42385439	-0.150	0.860	0.751–0.986	0.031
<i>FZD9</i>	chr7:72848109–72850450	-0.187	0.829	0.689–0.999	0.049
<i>GABRP</i>	chr5:170190354–170241051	-0.115	0.891	0.817–0.971	0.009
<i>IRF8</i>	chr16:85932409–85956215	-0.228	0.796	0.646–0.982	0.033
<i>ITM2A</i>	chrX:78615881–78623164	-0.188	0.829	0.690–0.997	0.046
<i>PRKCB</i>	chr16:23847322–24231932	-0.212	0.809	0.657–0.996	0.045
<i>SOX10</i>	chr22:38366693–38383429	-0.154	0.857	0.755–0.973	0.017
<i>TFF3</i>	chr21:43731777–43735761	0.106	1.112	1.011–1.224	0.029
<i>VGLL1</i>	chrX:135614311–135638966	-0.160	0.852	0.739–0.983	0.028



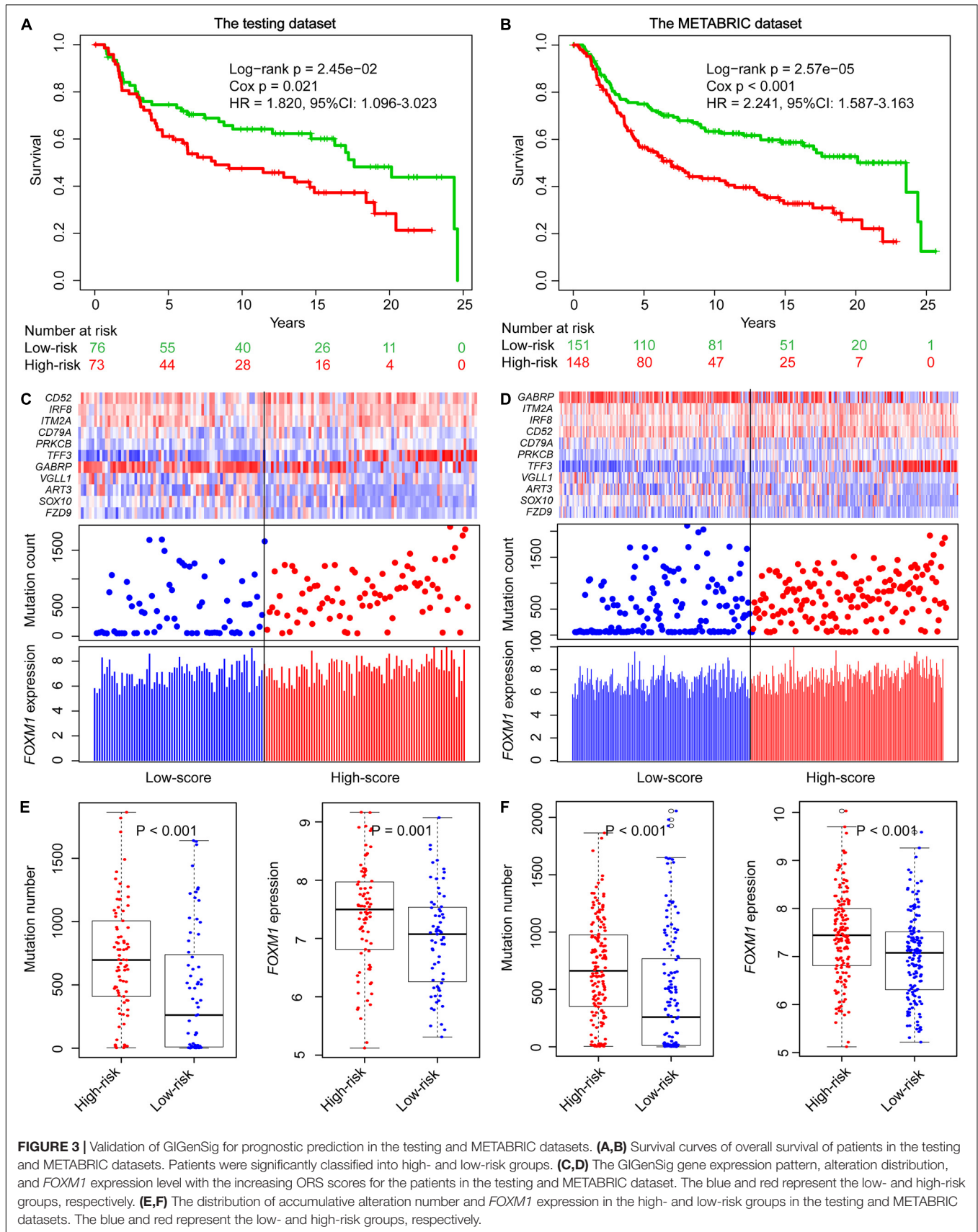
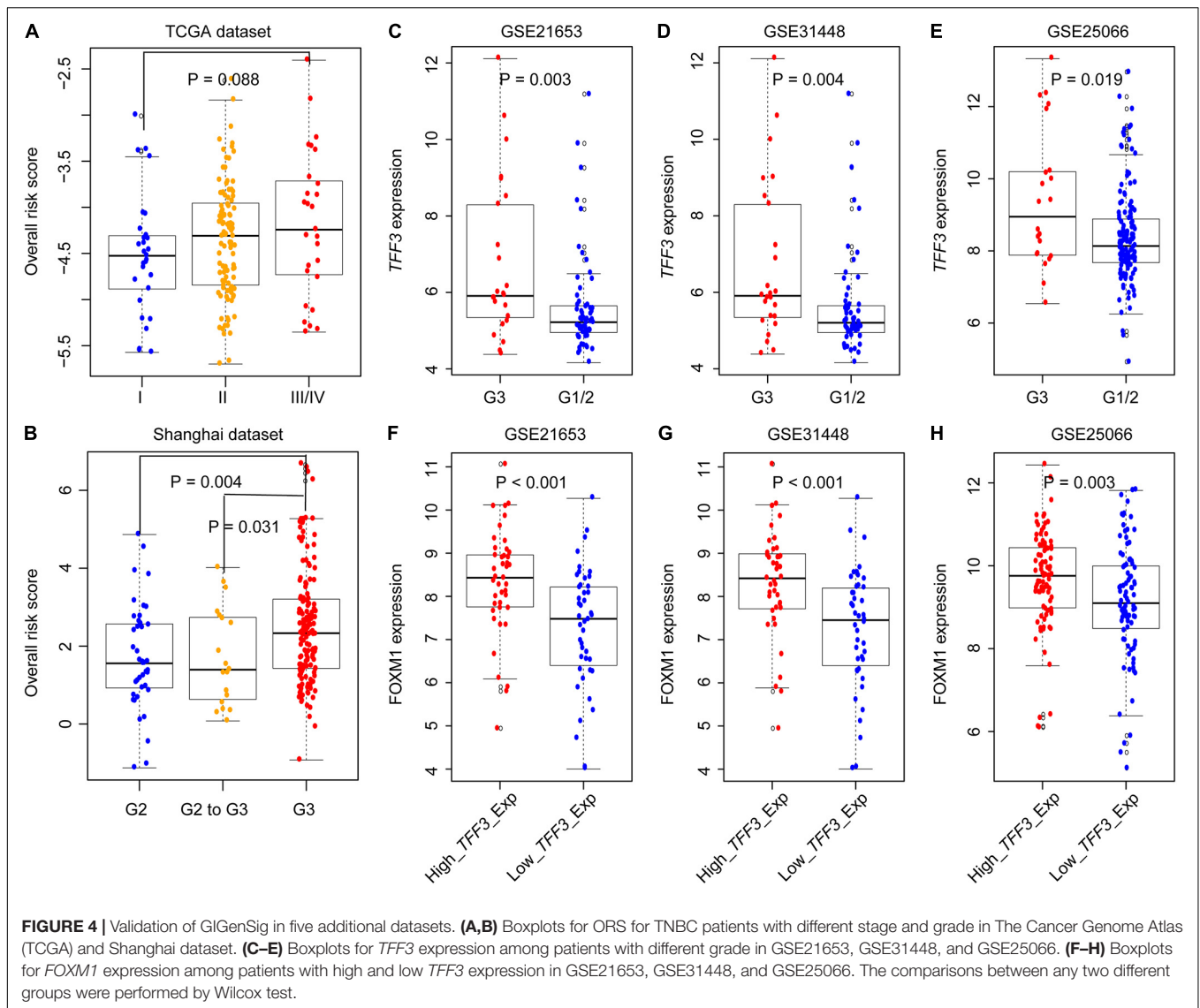


FIGURE 3 | Validation of GIGenSig for prognostic prediction in the testing and METABRIC datasets. **(A,B)** Survival curves of overall survival of patients in the testing and METABRIC datasets. Patients were significantly classified into high- and low-risk groups. **(C,D)** The GIGenSig gene expression pattern, alteration distribution, and FOXM1 expression level with the increasing ORS scores for the patients in the testing and METABRIC dataset. The blue and red represent the low- and high-risk groups, respectively. **(E,F)** The distribution of accumulative alteration number and FOXM1 expression in the high- and low-risk groups in the testing and METABRIC datasets. The blue and red represent the low- and high-risk groups, respectively.



test $p < 0.001$ for alteration count; Wilcoxon test $p = 0.001$ for *FOXM1* expression level).

We also validated the prognostic power of GIGenSig in the METABRIC TNBC dataset. The patients were classified into two groups. The median survival time of the low-risk group was significantly higher than that of the high-risk group (Figure 3B, median: 23.6 vs. 7 years; log-rank test $p = 2.57e-05$; HR = 2.241, 95% CI: 1.587–3.63). The 5-year survival rate in the low-risk group was longer (72%) than that in the high-risk group (54%), and the 5-year ROC gave an AUC of 0.627 (Supplementary Figure 6B). In addition, the SVM and 10-fold cross-validation showed that our risk model was robust to classify TNBC patients into high- and low-risk groups in the METABRIC dataset (Supplementary Figure 5C, AUC = 0.980, $p = 5.30e-62$). A similar pattern was observed in the METABRIC dataset as in the training and testing datasets for the clustering of GIGenSig, alteration count, and *FOXM1* expression (Figure 3D). Additionally, both significant differences

were present for alteration count and *FOXM1* expression between the high-risk and low-risk groups (Figure 3F, Wilcoxon test $p < 0.001$ for alteration count; Wilcoxon test $p < 0.001$ for *FOXM1* expression level).

Validation of GIGenSig in Five Additional Datasets

We compared GIGenSig using two independent datasets, the TCGA and Shanghai TNBC data, to test if clinical stage and grade could have an impact on the prognosis of TNBC. As shown in Figure 4A, there was a close relationship between the stage of TNBC and ORS but not reaching a significant level in the TCGA dataset (Figure 4A, Wilcoxon test $p = 0.088$). ORS was also associated with TNBC grade in Shanghai dataset that ORS of grade 3 was significantly higher than that of grade 2 and grades 2–3 (Figure 4B, $p = 0.004$ for comparing with grade 2; $p = 0.031$ for comparing with grades 2–3; Wilcoxon test). We

TABLE 3 | Univariate and multivariate Cox regression analyses of the genome instability-derived gene signature (GIGenSig) and overall survival in different patient datasets.

Characteristics	Univariable analysis			Multivariable analysis		
	HR	95% CI	p-value	HR	95% CI	p-Value
Training dataset						
GIGenSig	2.718	1.699–4.350	<0.001	2.184	1.218–3.915	0.009
Age	1.022	1.004–1.041	0.019	1.042	1.007–1.079	0.018
Menopausal status	0.887	0.549–1.433	0.625	2.656	1.024–6.886	0.045
Tumor stage	1.358	0.691–2.672	0.375	0.963	0.458–2.024	0.921
Grade	1.187	0.593–2.375	0.629	1.870	0.822–4.255	0.135
Testing dataset						
GIGenSig	1.820	1.096–3.023	0.021	1.354	0.683–2.682	0.038
Age	1.024	1.008–1.040	0.004	1.035	1.000–1.072	0.052
Menopausal status	0.640	0.395–1.037	0.070	2.215	0.754–6.508	0.148
Tumor stage	4.682	2.255–9.721	<0.001	4.976	2.266–10.926	<0.001
Grade	0.899	0.504–1.605	0.719	1.485	0.689–3.203	0.313
METABRIC dataset						
GIGenSig	2.241	1.587–3.163	<0.001	1.647	1.080–2.512	0.020
Age	1.024	1.012–1.036	0.000	1.037	1.013–1.062	0.002
Menopausal status	0.736	0.525–1.031	0.075	2.270	1.137–4.531	0.020
Tumor stage	2.180	1.332–3.569	0.002	1.851	1.093–3.134	0.022
Grade	1.011	0.649–1.573	0.962	1.456	0.842–2.515	0.178

further validated GIGenSig in another three independent breast cancer datasets generated by the microarray platform (GSE21653, GSE31448, and GSE25066). We reannotated the microarray data to obtain the gene expression data and extracted the common clinical characteristics from the three datasets. We then examined the association of GIGenSig with TNBC genomic instability information in these three independent datasets. Among all the 111 genes in GIGenSig, we found that *TFF3* presented significantly higher expression level in grade 3 than in grades 1 and 2 in all three datasets (Figures 4C–E, $p = 0.003$ for GSE21653; $p = 0.004$ for GSE31448; and $p = 0.019$ for GSE25066; Wilcoxon test). Furthermore, we tested the relationship between *TFF3* expression and *FOXMI* expression in the three datasets. We observed that *FOXMI* expression in patients with high *TFF3* expression was significantly higher than that with low *TFF3* expression in all three datasets (Figures 4F–H, $p < 0.001$ for GSE21653; $p < 0.001$ for GSE31448; and $p = 0.003$ for GSE25066; Wilcoxon test).

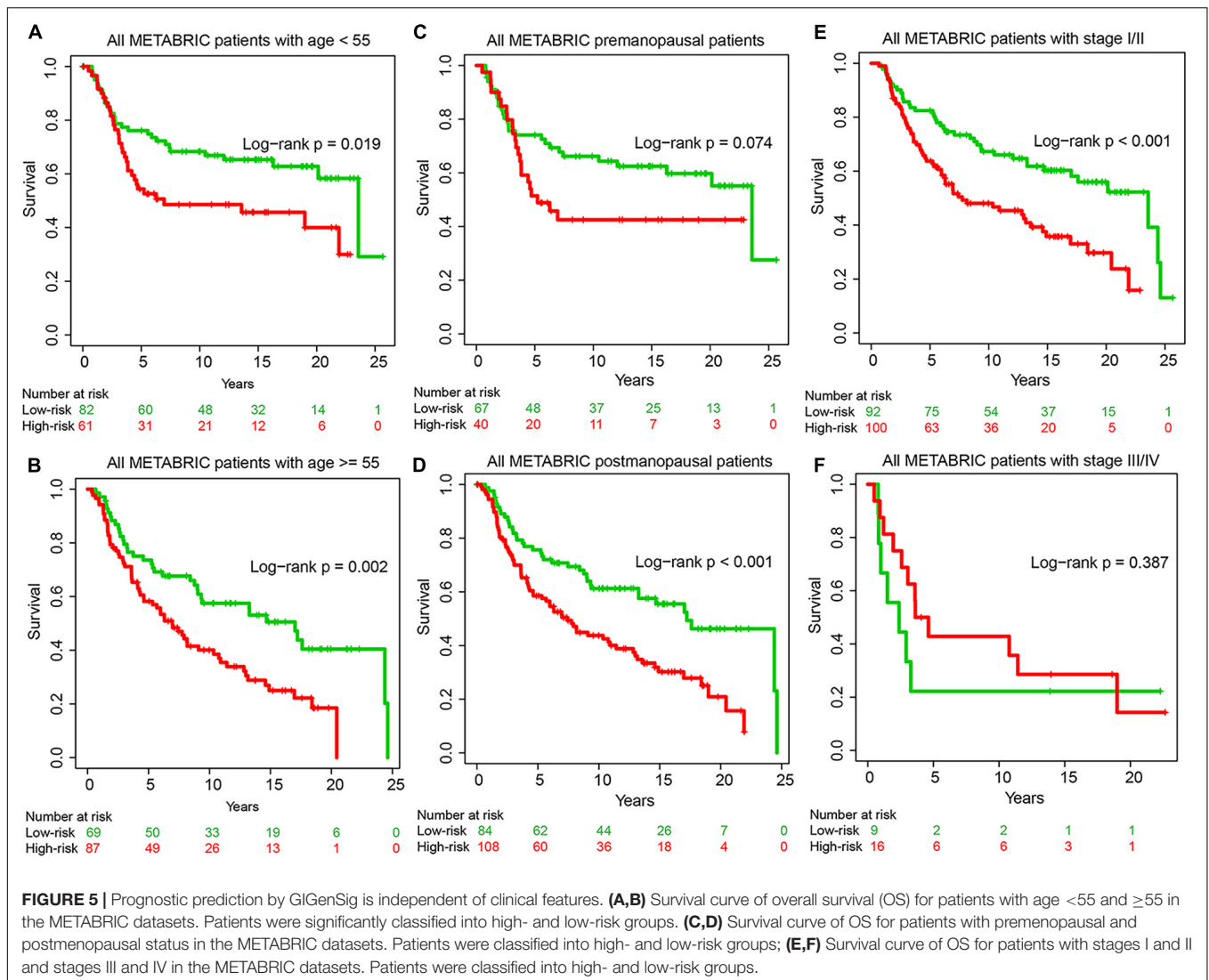
Prognostic Prediction by GIGenSig Is Independent of Clinical Features

To explore whether the prognostic ability of GIGenSig was independent of age, menopausal status, and tumor stage, we performed univariate and multivariate Cox regression analysis. The results showed that the GIGenSig was significantly associated with TNBC OS in the three datasets when adjusted by age, menopausal status, tumor stage, and grade (Table 3). The METABRIC patients were divided into two groups according to age with <55 and ≥ 55 , pre- and post-status of menopause, and tumor stage with I/II and III/IV. The patients were classified into high- and low-risk groups according to the median risk

scores in the training dataset. The results revealed that the patients were significantly classified into two groups by age (Figures 5A,B; log-rank test $p = 0.019$ for age < 55 group; log-rank test $p = 0.002$ for age ≥ 55 group), menopausal (Figures 5C,D; log-rank test $p = 0.074$ for premenopausal group; log-rank test $p < 0.001$ for postmenopausal group), and stage subset (Figures 5E,F; log-rank test $p < 0.001$ for stage I/II group; log-rank test $p = 0.387$ for stage III/IV group). The classification for patients in the stage III/IV group was not significant probably due to the smaller sample size ($n = 25$) in this group. These results indicated that GIGenSig served as a prognostic signature for TNBC independent of age, menopausal status, and tumor stage.

Genome Instability-Derived Gene Signature Performs Better Than Other Prognostic Signatures

To further explore the prognostic performance of the GIGenSig, we compared GIGenSig with other TNBC prognostic signatures including the two-gene signature (Alsaleem et al., 2020), the five-gene signature (Wang et al., 2018), the eight-gene signature (Kim et al., 2019), and the 19-gene signature (Qian et al., 2017) using the METABRIC dataset. The result showed that the 5-year AUC (0.627) of OS for GIGenSig was significantly higher than that of the two-gene signature (AUC = 0.534), the five-gene signature (AUC = 0.571), the eight-gene signature (AUC = 0.546), and the 19-gene signature (AUC = 0.615) (Figure 6). The results demonstrated that the GIGenSig provided better prognostic prediction for TNBC than the other four signatures.

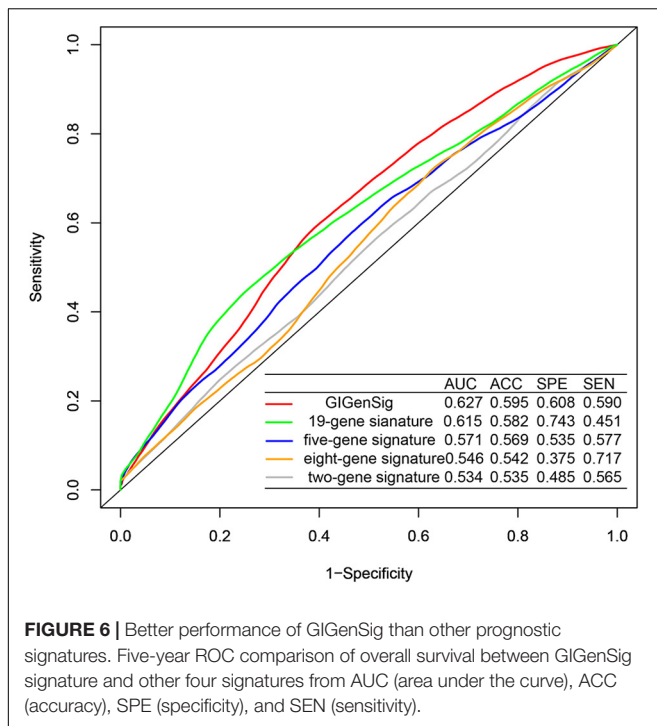


DISCUSSION

Compared with TNBC treatment, limited progress has been made in TNBC prognosis (Sulaiman et al., 2018a,b; Mou and Wang, 2019; Zhang et al., 2019). Prognostic study usually evaluates clinical features, such as tumor size, stage, grade, etc., which provide limited mechanistic information to understand the relationship between prognosis and the disease (Echavarría et al., 2018; Johansson et al., 2018; Park et al., 2019). Genome instability and abnormal gene expression are common features in cancer (Telli et al., 2016; Tutt et al., 2018; Huang et al., 2019). While the relationship between genome instability and dysregulation of gene expression in cancer has been studied, genome-wide characterization for its prognostic value in TNBC has not been systematically analyzed (Grady and Carethers, 2008; Rao et al., 2017; Kalimutho et al., 2019).

As shown from our current study, molecular evidence from genome instability and abnormal gene expression is a rich

resource to identify prognostic signatures as the prognostic marker for TNBC. In our study, we identified 111 genome instability-related genes by integrating mutation, CNV, and gene expression from TNBC. Functional analysis showed that these 111 genomic instability-related genes were enriched in the pathways associated with mitotic process. Dysregulation of mitotic processes can impact DNA replication involving mitotic nuclear division, nuclear division, and organelle fission, contributing to genome instability and OS of TNBC (Chen L. et al., 2018; Si et al., 2020; Suo et al., 2020). For example, the feedback loop between Drp1-mediated mitochondrial fission and Notch signaling pathway can promote TNBC cell survival via increasing survivin expression (Chen L. et al., 2018), and silibinin-induced mitochondrial fission can cause mitophagy preventing silibinin-induced apoptosis in TNBC (Si et al., 2020). Functional annotation with DAVID tool revealed that these 111-genomic instability-related genes play important roles in carcinogenetic pathways, such as affecting cell cycle (Kastan and Bartek, 2004), uncontrolled cell proliferation



(Evan and Vousden, 2001), and abnormal immune response (Desrichard et al., 2016).

From the 111 genes, we further identified a GIGenSig with 11 genes (*ART3*, *CD52*, *CD79A*, *FZD9*, *GABRP*, *IRF8*, *ITM2A*, *PRKCB*, *SOX10*, *TFE3*, and *VGLL1*). Our study demonstrated that the GIGenSig effectively divided TNBC into high- and low-risk groups in the training dataset and its prognostic value was validated independently in multiple testing datasets. Besides, GIGenSig was significantly associated with genomic alteration pattern and *FOXO1* expression, which are important predictors of genome instability. Certain genes in GIGenSig are known to be closely related to tumorigenesis and development of TNBC. For example, *ART3* overexpression regulated TNBC cell functions by activating AKT and ERK pathways (Tan et al., 2016); *GABRP* and *VGLL1* were present in basal-like/triple-negative phenotype, and their expression levels were associated with OS of TNBC (Castilla et al., 2014; Sizemore et al., 2014); and *ITM2A* and *SOX10* were prognostic biomarkers for TNBC and potential therapeutic targets (Harbhajanka et al., 2018; Abuderman et al., 2020).

Furthermore, our identified GIGenSig can have clinical significance in TNBC treatment. It has been reported that Ki67 encoded by *MKI67* plays an important role in the prognosis and treatment of breast cancer (Yerushalmi et al., 2010; Li et al., 2015; Jurikova et al., 2016). Using the METABRIC dataset, we also observed that older TNBC patients had significantly lower *MKI67* expression than younger TNBC patients (Supplementary Figure 7A, Wilcoxon test $p = 2.59 \times 10^{-3}$). However, Ki67 did not show prognostic value for patients used in our study (Supplementary Figure 7B). This implies that a combination of multiple

genes provides better prognostic power than does a single gene. Our study also found that TNBC patients with chemotherapy treatment had low risk (84 vs. 73 with chemotherapy in the low-risk and high-risk groups, Supplementary Figure 7C), demonstrating the effectiveness of chemotherapy as the main therapeutic strategy in TNBC treatment (Denkert et al., 2017; Chaudhary et al., 2018; Lyons, 2019).

In conclusion, our study provides a genome instability-based TNBC prognostic signature to predict the clinical outcome of TNBC. Further tests with more datasets and clinical information will validate its value for clinical TNBC applications.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. These are included in the article/Supplementary Material.

AUTHOR CONTRIBUTIONS

MG and SW conceptualized the study and reviewed and edited the manuscript. MG formulated the methodology, performed the investigation and visualization, and wrote the original draft. SW supervised and secured the funding. Both authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by grants from the Macau Science and Technology Development Fund (085/2017/A2 and 0077/2019/AMJ), the University of Macau (SRG2017-00097-FHS and MYRG2019-00018-FHS), and the Faculty of Health Sciences, the University of Macau (FHSIG/SW/0007/2020P, Startup fund) (SW).

ACKNOWLEDGMENTS

We are thankful to the Information and Communication Technology Office (ICTO), University of Macau, for providing the High-Performance Computing Cluster (HPCC) resources and facilities for the study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2021.701073/full#supplementary-material>

REFERENCES

- Abuderman, A. A., Harb, O. A., and Gertallah, L. M. (2020). Prognostic and clinicopathological values of tissue expression of MFAP5 and ITM2A in triple-negative breast cancer: an immunohistochemical study. *Contemp. Oncol. (Pozn)* 24, 87–95. doi: 10.5114/wo.2020.97520
- Alsalem, M. A., Ball, G., Toss, M. S., Raafat, S., Aleskandarany, M., Joseph, C., et al. (2020). A novel prognostic two-gene signature for triple negative breast cancer. *Mod. Pathol.* 33, 2208–2220. doi: 10.1038/s41379-020-0563-7
- Bao, S., Hu, T., Liu, J., Su, J., Sun, J., Ming, Y., et al. (2021). Genomic instability-derived plasma extracellular vesicle-microRNA signature as a minimally invasive predictor of risk and unfavorable prognosis in breast cancer. *J. Nanobiotechnol.* 19:22. doi: 10.1186/s12951-020-00767-3
- Castilla, M. A., Lopez-Garcia, M. A., Atienza, M. R., Rosa-Rosa, J. M., Diaz-Martin, J., Pecero, M. L., et al. (2014). VGLL1 expression is associated with a triple-negative basal-like phenotype in breast cancer. *Endocr. Relat. Cancer* 21, 587–599. doi: 10.1530/ERC-13-0485
- Chaudhary, L. N., Wilkinson, K. H., and Kong, A. (2018). Triple-negative breast cancer: who should receive neoadjuvant chemotherapy? *Surg. Oncol. Clin. N. Am.* 27, 141–153. doi: 10.1016/j.soc.2017.08.004
- Chen, L., Zhang, J., Lyu, Z., Chen, Y., Ji, X., Cao, H., et al. (2018). Positive feedback loop between mitochondrial fission and Notch signaling promotes survivin-mediated survival of TNBC cells. *Cell Death Dis.* 9:1050. doi: 10.1038/s41419-018-1083-y
- Chen, M., Pockaj, B., Andreozzi, M., Barrett, M. T., Krishna, S., Eaton, S., et al. (2018). JAK2 and PD-L1 amplification enhance the dynamic expression of PD-L1 in triple-negative breast cancer. *Clin. Breast Cancer* 18, e1205–e1215. doi: 10.1016/j.clbc.2018.05.006
- Denkert, C., Liedtke, C., Tutt, A., and von Minckwitz, G. (2017). Molecular alterations in triple-negative breast cancer—the road to new treatment strategies. *Lancet* 389, 2430–2442. doi: 10.1016/S0140-6736(16)32454-0
- Dent, R., Trudeau, M., Pritchard, K. I., Hanna, W. M., Kahn, H. K., Sawka, C. A., et al. (2007). Triple-negative breast cancer: clinical features and patterns of recurrence. *Cancer Res.* 13(15 Pt 1), 4429–4434. doi: 10.1158/1078-0432.CCR-06-3045
- Desrichard, A., Snyder, A., and Chan, T. A. (2016). Cancer Neoantigens and Applications for Immunotherapy. *Clin. Cancer Res.* 22, 807–812. doi: 10.1158/1078-0432.CCR-14-3175
- Echavarría, I., Lopez-Tarruella, S., Picornell, A., Garcia-Saenz, J. A., Jerez, Y., Hoadley, K., et al. (2018). Pathological response in a triple-negative breast cancer cohort treated with neoadjuvant carboplatin and docetaxel according to lehmans's refined classification. *Clin. Cancer Res.* 24, 1845–1852. doi: 10.1158/1078-0432.CCR-17-1912
- Evan, G. I., and Vousden, K. H. (2001). Proliferation, cell cycle and apoptosis in cancer. *Nature* 411, 342–348. doi: 10.1038/35077213
- Goldman, M. J., Craft, B., Hastie, M., Repecka, K., McDade, F., Kamath, A., et al. (2020). Visualizing and interpreting cancer genomics data via the Xena platform. *Nat. Biotechnol.* 38, 675–678. doi: 10.1038/s41587-020-0546-8
- Grady, W. M., and Carethers, J. M. (2008). Genomic and epigenetic instability in colorectal cancer pathogenesis. *Gastroenterology* 135, 1079–1099. doi: 10.1053/j.gastro.2008.07.076
- Guo, L., Li, W., Zhu, X., Ling, Y., Qiu, T., Dong, L., et al. (2016). PD-L1 expression and CD274 gene alteration in triple-negative breast cancer: implication for prognostic biomarker. *Springerplus* 5:805. doi: 10.1186/s40064-016-2513-x
- Habermann, J. K., Doering, J., Hautaniemi, S., Roblick, U. J., Bundgen, N. K., Nicorici, D., et al. (2009). The gene expression signature of genomic instability in breast cancer is an independent predictor of clinical outcome. *Int. J. Cancer* 124, 1552–1564. doi: 10.1002/ijc.24017
- Haffty, B. G., Yang, Q., Reiss, M., Kearney, T., Higgins, S. A., Weidhaas, J., et al. (2006). Locoregional relapse and distant metastasis in conservatively managed triple negative early-stage breast cancer. *J. Clin. Oncol.* 24, 5652–5657. doi: 10.1200/JCO.2006.06.5664
- Harbhajanka, A., Chahar, S., Miskimen, K., Silverman, P., Harris, L., Williams, N., et al. (2018). Clinicopathological, immunohistochemical and molecular correlation of neural crest transcription factor SOX10 expression in triple-negative breast carcinoma. *Hum. Pathol.* 80, 163–169. doi: 10.1016/j.humpath.2018.06.007
- Hatzis, C., Pusztai, L., Valero, V., Booser, D. J., Esserman, L., Lluch, A., et al. (2011). A genomic predictor of response and survival following taxane-anthracycline chemotherapy for invasive breast cancer. *JAMA* 305, 1873–1881. doi: 10.1001/jama.2011.593
- Huang, Y., Li, W., Yan, W., Wu, J., Chen, L., Yao, X., et al. (2019). Loss of PICH promotes chromosome instability and cell death in triple-negative breast cancer. *Cell Death Dis.* 10:428. doi: 10.1038/s41419-019-1662-6
- Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57. doi: 10.1038/nprot.2008.211
- Irizarry, R. A., Bolstad, B. M., Collin, F., Cope, L. M., Hobbs, B., and Speed, T. P. (2003a). Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res.* 31:e15. doi: 10.1093/nar/gng015
- Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U., et al. (2003b). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4, 249–264. doi: 10.1093/biostatistics/4.2.249
- Jiang, Y. Z., Ma, D., Suo, C., Shi, J., Xue, M., Hu, X., et al. (2019). Genomic and transcriptomic landscape of triple-negative breast cancers: subtypes and treatment strategies. *Cancer Cell* 35, 428–440.e425. doi: 10.1016/j.ccell.2019.02.001
- Johansson, A. L. V., Andersson, T. M., Hsieh, C. C., Jirstrom, K., Cnattingius, S., Fredriksson, I., et al. (2018). Tumor characteristics and prognosis in women with pregnancy-associated breast cancer. *Int. J. Cancer* 142, 1343–1354. doi: 10.1002/ijc.31174
- Jurikova, M., Danihel, L., Polak, S., and Varga, I. (2016). Ki67, PCNA, and MCM proteins: markers of proliferation in the diagnosis of breast cancer. *Acta Histochem.* 118, 544–552. doi: 10.1016/j.acthis.2016.05.002
- Kalimutho, M., Nones, K., Srihari, S., Duijff, P. H. G., Waddell, N., and Khanna, K. K. (2019). Patterns of genomic instability in breast cancer. *Trends Pharmacol. Sci.* 40, 198–211. doi: 10.1016/j.tips.2019.01.005
- Karn, T., Jiang, T., Hatzis, C., Sanger, N., El-Balat, A., Rody, A., et al. (2017). Association between genomic metrics and immune infiltration in triple-negative breast cancer. *JAMA Oncol.* 3, 1707–1711. doi: 10.1001/jamaoncol.2017.2140
- Kastan, M. B., and Bartek, J. (2004). Cell-cycle checkpoints and cancer. *Nature* 432, 316–323. doi: 10.1038/nature03097
- Kim, E. K., Park, A. K., Ko, E., Park, W. Y., Lee, K. M., Noh, D. Y., et al. (2019). Risk stratification of triple-negative breast cancer with core gene signatures associated with chemoresponse and prognosis. *Breast Cancer Res. Treat.* 178, 185–197. doi: 10.1007/s10549-019-05366-x
- Kim, Y. H., Choi, M. H., Kim, J. H., Lim, I. K., and Park, T. J. (2013). C-terminus-deleted FoxM1 is expressed in cancer cell lines and induces chromosome instability. *Carcinogenesis* 34, 1907–1917. doi: 10.1093/carcin/bgt134
- Li, J., Chen, Z., Su, K., and Zeng, J. (2015). Clinicopathological classification and traditional prognostic indicators of breast cancer. *Int. J. Clin. Exp. Pathol.* 8, 8500–8505.
- Lyons, T. G. (2019). Targeted therapies for triple-negative breast cancer. *Curr. Treat. Options Oncol.* 20:82. doi: 10.1007/s11864-019-0682-x
- Mermel, C. H., Schumacher, S. E., Hill, B., Meyerson, M. L., Beroukhi, R., and Getz, G. (2011). GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* 12:R41. doi: 10.1186/gb-2011-12-4-r41
- Mersin, H., Yildirim, E., Berberoglu, U., and Gulben, K. (2008). The prognostic importance of triple negative breast carcinoma. *Breast* 17, 341–346. doi: 10.1016/j.breast.2007.11.031
- Mou, E., and Wang, H. (2019). LncRNA LUCAT1 facilitates tumorigenesis and metastasis of triple-negative breast cancer through modulating miR-5702. *Biosci. Rep.* 39:BSR20190489. doi: 10.1042/BSR20190489
- Negrini, S., Gorgoulis, V. G., and Halazonetis, T. D. (2010). Genomic instability—an evolving hallmark of cancer. *Nat. Rev. Mol. Cell Biol.* 11, 220–228. doi: 10.1038/nrm2858
- Oakman, C., Viale, G., and Di Leo, A. (2010). Management of triple negative breast cancer. *Breast* 19, 312–321. doi: 10.1016/j.breast.2010.03.026
- Park, J. H., Jonas, S. F., Bataillon, G., Criscitiello, C., Salgado, R., Loi, S., et al. (2019). Prognostic value of tumor-infiltrating lymphocytes in patients with early-stage triple-negative breast cancers (TNBC) who did not receive adjuvant chemotherapy. *Ann. Oncol.* 30, 1941–1949. doi: 10.1093/annonc/mdz395

- Pereira, B., Chin, S. F., Rueda, O. M., Vollan, H. K., Provenzano, E., Bardwell, H. A., et al. (2016). The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. *Nat. Commun.* 7:11479. doi: 10.1038/ncomms11479
- Qian, J., Chen, H., Ji, X., Eisenberg, R., Chakravarthy, A. B., Mayer, I. A., et al. (2017). A 3q gene signature associated with triple negative breast cancer organ specific metastasis and response to neoadjuvant chemotherapy. *Sci. Rep.* 7:45828. doi: 10.1038/srep45828
- Rao, C. V., Asch, A. S., and Yamada, H. Y. (2017). Frequently mutated genes/pathways and genomic instability as prevention targets in liver cancer. *Carcinogenesis* 38, 2–11. doi: 10.1093/carcin/bgw118
- Sabatier, R., Finetti, P., Adelaide, J., Guille, A., Borg, J. P., Chaffanet, M., et al. (2011a). Down-regulation of ECRG4, a candidate tumor suppressor gene, in human breast cancer. *PLoS One* 6:e27656. doi: 10.1371/journal.pone.0027656
- Sabatier, R., Finetti, P., Cervera, N., Lambaudie, E., Esterni, B., Mamessier, E., et al. (2011b). A gene expression signature identifies two prognostic subgroups of basal breast cancer. *Breast Cancer Res. Treat.* 126, 407–420. doi: 10.1007/s10549-010-0897-9
- Safonov, A., Jiang, T., Bianchini, G., Gyorffy, B., Karn, T., Hatzis, C., et al. (2017). Immune gene expression is associated with genomic aberrations in breast cancer. *Cancer Res.* 77, 3317–3324. doi: 10.1158/0008-5472.CAN-16-3478
- Schmitt, M. W., Prindle, M. J., and Loeb, L. A. (2012). Implications of genetic heterogeneity in cancer. *Ann. N. Y. Acad. Sci.* 1267, 110–116. doi: 10.1111/j.1749-6632.2012.06590.x
- Senfter, D., Samadaei, M., Mader, R. M., Gojo, J., Peyrl, A., Krupitza, G., et al. (2019). High impact of miRNA-4521 on FOXM1 expression in medulloblastoma. *Cell Death Dis.* 10:696. doi: 10.1038/s41419-019-1926-1
- Serino, L. T. R., Jucoski, T. S., Morais, S. B., Fernandes, C. C. C., Lima, R. S., Urban, C. A., et al. (2019). Association of FOSL1 copy number alteration and triple negative breast tumors. *Genet. Mol. Biol.* 42, 26–31. doi: 10.1590/1678-4685-gmb-2017-0267
- Si, L., Fu, J., Liu, W., Hayashi, T., Mizuno, K., Hattori, S., et al. (2020). Silibinin-induced mitochondria fission leads to mitophagy, which attenuates silibinin-induced apoptosis in MCF-7 and MDA-MB-231 cells. *Arch. Biochem. Biophys.* 685:108284. doi: 10.1016/j.abb.2020.108284
- Sizemore, G. M., Sizemore, S. T., Seachrist, D. D., and Keri, R. A. (2014). GABA(A) receptor pi (GABRP) stimulates basal-like breast cancer cell migration through activation of extracellular-regulated kinase 1/2 (ERK1/2). *J. Biol. Chem.* 289, 24102–24113. doi: 10.1074/jbc.M114.593582
- Song, H., Cicek, M. S., Dicks, E., Harrington, P., Ramus, S. J., Cunningham, J. M., et al. (2014). The contribution of deleterious germline mutations in BRCA1, BRCA2 and the mismatch repair genes to ovarian cancer in the population. *Hum. Mol. Genet.* 23, 4703–4709. doi: 10.1093/hmg/ddu172
- Sulaiman, A., McGarry, S., Lam, K. M., El-Sahli, S., Chambers, J., Kaczmarek, S., et al. (2018a). Co-inhibition of mTORC1, HDAC and ESR1alpha retards the growth of triple-negative breast cancer and suppresses cancer stem cells. *Cell Death Dis.* 9:815. doi: 10.1038/s41419-018-0811-7
- Sulaiman, A., McGarry, S., Li, L., Jia, D., Ooi, S., Addison, C., et al. (2018b). Dual inhibition of Wnt and Yes-associated protein signaling retards the growth of triple-negative breast cancer in both mesenchymal and epithelial states. *Mol. Oncol.* 12, 423–440. doi: 10.1002/1878-0261.12167
- Suo, H. D., Tao, Z., Zhang, L., Jin, Z. N., Li, X. Y., Ma, W., et al. (2020). Coexpression network analysis of genes related to the characteristics of tumor stemness in triple-negative breast cancer. *Biomed. Res. Int.* 2020:7575862. doi: 10.1155/2020/7575862
- Suzuki, K., Ohnami, S., Tanabe, C., Sasaki, H., Yasuda, J., Katai, H., et al. (2003). The genomic damage estimated by arbitrarily primed PCR DNA fingerprinting is useful for the prognosis of gastric cancer. *Gastroenterology* 125, 1330–1340. doi: 10.1016/j.gastro.2003.07.006
- Tan, A. R., and Swain, S. M. (2008). Therapeutic strategies for triple-negative breast cancer. *Cancer J.* 14, 343–351. doi: 10.1097/PPO.0b013e31818d839b
- Tan, L., Song, X., Sun, X., Wang, N., Qu, Y., and Sun, Z. (2016). ART3 regulates triple-negative breast cancer cell function via activation of Akt and ERK pathways. *Oncotarget* 7, 46589–46602. doi: 10.18632/oncotarget.10306
- Teh, M. T., Gemenetzidis, E., Chaplin, T., Young, B. D., and Philpott, M. P. (2010). Upregulation of FOXM1 induces genomic instability in human epidermal keratinocytes. *Mol. Cancer* 9:45. doi: 10.1186/1476-4598-9-45
- Telli, M. L., Timms, K. M., Reid, J., Hennessy, B., Mills, G. B., Jensen, K. C., et al. (2016). Homologous recombination deficiency (HRD) score predicts response to platinum-containing neoadjuvant chemotherapy in patients with triple-negative breast cancer. *Clin. Cancer Res.* 22, 3764–3773. doi: 10.1158/1078-0432.CCR-15-2477
- Tutt, A., Tovey, H., Cheang, M. C. U., Kernaghan, S., Kilburn, L., Gazinska, P., et al. (2018). Carboplatin in BRCA1/2-mutated and triple-negative breast cancer BRCAness subgroups: the TNT trial. *Nat. Med.* 24, 628–637. doi: 10.1038/s41591-018-0009-7
- Wang, S., Beeghly-Padiel, A., Cai, Q., Cai, H., Guo, X., Shi, L., et al. (2018). Gene expression in triple-negative breast cancer in relation to survival. *Breast Cancer Res. Treat.* 171, 199–207. doi: 10.1007/s10549-018-4816-9
- Wu, J., Mamidi, T. K. K., Zhang, L., and Hicks, C. (2019). Integrating germline and somatic mutation information for the discovery of biomarkers in triple-negative breast cancer. *Int. J. Environ. Res. Public Health* 16:1055. doi: 10.3390/ijerph16061055
- Yerushalmi, R., Woods, R., Ravdin, P. M., Hayes, M. M., and Gelmon, K. A. (2010). Ki67 in breast cancer: prognostic and predictive potential. *Lancet Oncol.* 11, 174–183. doi: 10.1016/S1470-2045(09)70262-1
- Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287. doi: 10.1089/omi.2011.0118
- Zhang, B., Shetti, D., Fan, C., and Wei, K. (2019). miR-29b-3p promotes progression of MDA-MB-231 triple-negative breast cancer cells through downregulating TRAF3. *Biol. Res.* 52:38. doi: 10.1186/s40659-019-0245-4

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 © 2021 Guo 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.