Am J Transl Res 2018;10(6):1677-1689
www.ajtr.org /ISSN:1943-8141/AJTR0072909

Original Article

Interleukin enhancer binding factor 2 is a prognostic biomarker for breast cancer that also predicts neoadjuvant chemotherapy responses

Zining Jin[1], Lu Xu[2], Lei Zhang[1], Min Zhao[1], Dongbao Li[1], Lijun Ye[1], Ying Ma[1], Siyu Ren[1], Hailan Yu[1], Danyu Wang[1], Chunyan Liang[1], Bo Chen[1]

Departments of [1]Breast Surgery, [2]Medical Oncology, The First Hospital of China Medical University, Shenyang 110001, Liaoning Province, China

Received January 17, 2018; Accepted June 1, 2018; Epub June 15, 2018; Published June 30, 2018

Abstract: Interleukin enhancer binding factor 2 (ILF2) participates in several aspects of DNA and RNA metabolism and regulates gene expression at multiple levels; however, its role in breast cancer remains undefined. The variant statuses of *ILF2* in human breast cancer were evaluated using the COSMIC database. Altered *ILF2* expression in normal breast tissue relative to cancer tissue and in breast cancer patients with different clinicopathological characteristics, molecular subtypes, clinical outcomes and chemotherapy responses were examined using the Oncomine, GOBO, Kaplan-Meier plotter and GEO datasets. To explore possible biological networks connected to ILF2 in breast cancer, we performed ingenuity pathway analysis on ILF2-related differentially expressed genes. We found that many breast cancers had increased *ILF2* copy number variations and increased *ILF2* expression. We also observed that elevated *ILF2* expression was correlated with aggressive features, such as high histological grade, BRCA1 mutations, and the triple-negative/basal-like subtype, which resulted in shorter survival in these cases. Moreover, *ILF2* expression predicted responses to anthracycline/taxane-based treatment. Ingenuity pathway analysis revealed that ILF2-related biological functions included promoting cell survival, viability, and proliferation, as well as cell cycle progression and DNA repair. Certain well-known oncogenes (MYC and HGF), cytokines (CSF2, IFNG and IL5) and microRNAs (miR-21, miR-155-5p and let-7) may participate in the ILF2 expression network in breast cancer. In summary, ILF2 is involved in the development and progression of breast cancer and may be a predictive biomarker for better responses to anthracycline/taxane-based treatments.

Keywords: ILF2, breast cancer, prognostic biomarker, predictive biomarker, neoadjuvant chemotherapy, pathway analysis

Introduction

Breast cancer is the most frequently diagnosed cancer among women and was the cause of an estimated 522,000 deaths worldwide in 2012 [1]. However, because of the biological complexity of breast cancer, it is often difficult to determine therapeutic responses, metastatic patterns, and clinical outcomes using only clinical parameters or classic pathological markers (oestrogen receptor [ER], progesterone receptor [PR], human epidermal growth factor receptor 2 [HER2], and Ki-67) [2]. Identifying new and effective biomarkers might offer an enhanced perspective on the clinical behaviour of breast cancer and/or provide new therapeutic targets.

Interleukin enhancer binding factor 2 (ILF2, also known as NF45) was initially identified as a member of the transcriptional activator complex that is crucial for interleukin 2 expression in T cells [3]. ILF2 was later found to be ubiquitously expressed in human tissues and, together with its binding partner NF90 (ILF3), to participate in multiple aspects of DNA and RNA metabolism, including DNA repair and replication [4, 5], transcription [6], translation [7, 8], mRNA splicing [9], micro-RNA biogenesis [10-12], and viral replication [13, 14]. Silencing *ILF2* leads to mitotic defects that can be partially explained by inhibiting DNA synthesis [15] and downregulating the translation of X-linked inhibitor of apoptosis (*XIAP*) and cellular inhibition of apoptosis protein 1 (*cIAP1*) mRNAs by reducing

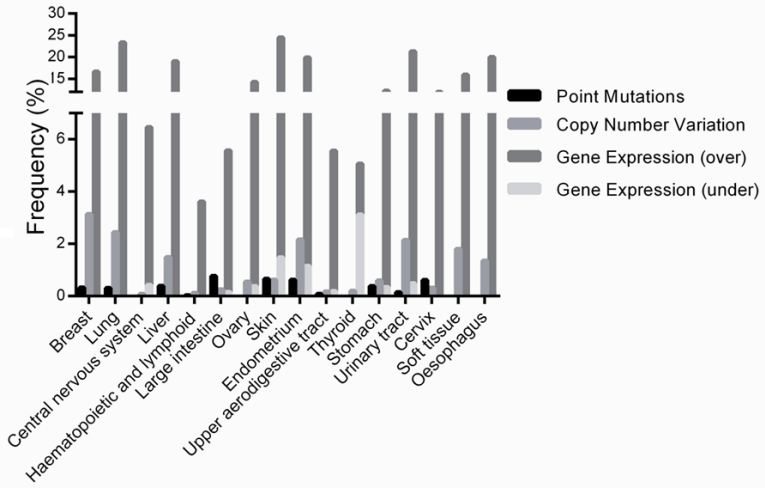


Figure 1. Alteration of ILF2 gene status across different human cancer types. Data on COSMIC database demonstrating 0.33% breast cancer patients (7/2,137 cases) exhibited ILF2 gene point mutations, 3.14% breast cancer patients (55/1,749 cases) exhibited ILF2 gene copy number gains, 16.67% breast cancer patients (184/1,104 cases) exhibited ILF2 mRNA up-regulation.

ILF2 IRES trans-acting protein factor activity [7, 8]. Recently, several studies have investigated the role of ILF2 in cancer and found that ILF2 is involved in the progression of various cancer types, such as pancreatic carcinoma [16, 17], hepatocellular carcinoma [12, 18], non-small cell lung cancer [19], oesophageal squamous cell carcinoma [20], and gastric cancer [21]. Although previous studies have demonstrated that certain ILF2 functions are cell type-dependent [14, 22], the precise role of ILF2 in breast cancer remains undefined. NF90, a crucial regulator and binding partner of ILF2, was recently reported to promote breast cancer tumourigenicity by activating urokinase-type plasminogen activator (*uPA*) transcription and inhibiting the processing of *uPA* mRNA-targeted microRNAs [23]. Based on these data, we hypothesized that ILF2 may also play a role in breast cancer progression.

This study used publicly available clinical datasets to identify correlations between *ILF2* expression and breast cancer prognoses. Our findings are the first to demonstrate that elevated *ILF2* expression is associated with breast cancer development and poor clinical outcomes. Higher *ILF2* expression was observed in triple-negative and basal-like breast cancers and predicted responses to anthracycline/taxane-based neoadjuvant chemotherapy. Finally, ingenuity pathway analysis (IPA) was used to analyse transcriptional changes downstream of increased *ILF2* expression to further explore the molecular mechanisms underlying the association between increased *ILF2* expression and aggressive behaviours in breast cancer.

Materials and methods

Catalogue of somatic mutations in cancer (COSMIC) database

The COSMIC database is a high-resolution resource for exploring the impact of somatic mutations in human cancer (http://cancer.sanger.ac.uk) [24]. The latest release (COSMIC v81) was used to find the variant ILF2 statuses in human breast cancer.

Oncomine

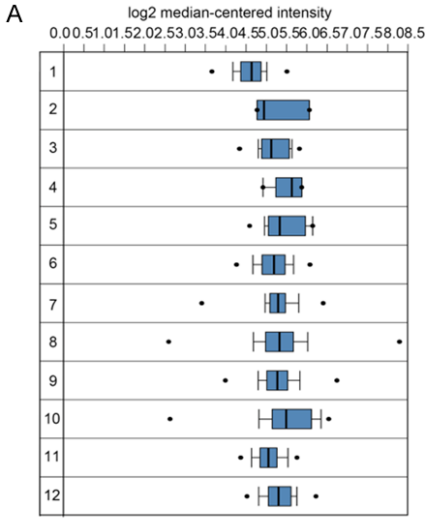
The Oncomine platform (https://www.oncomine.org/; version 4.5) is a cancer microarray database that includes the gene expression profile data of 91,866 samples from 729 datasets [25]. Alterations in *ILF2* mRNA expression were evaluated for normal breast tissues and breast cancers of different clinicopathological characteristics. Differences between groups were assessed by *P*-values complemented by fold change. To study the relative expression of ILF2 and cell cycle regulators, microarray data of 593 breast cancer cases from The Cancer Genome Atlas (TCGA) breast dataset were downloaded from Oncomine and analysed by linear regression.

Gene expression-based outcome for breast cancer online (GOBO)

GOBO (http://co.bmc.lu.se/gobo) is a web-based analysis tool that uses Affymetrix U133A microarray gene expression data from 1,881 breast cancer patients and a 51-sample breast cancer cell line set [26]. We used the Gene Set Analysis (GSA) tumour to explore *ILF2* expression levels in different molecular subtypes and histological grades.

Kaplan-Meier plotter

The prognostic value of *ILF2* was determined by Kaplan-Meier analysis using the Kaplan-



Legend	Cancer and normal type	Number	*p* value	Fold change
1	Breast	144	/	/
2	Benign Breast Neoplasm	3	0.124	1.571
3	Breast Carcinoma	14	5.94E-05	1.478
4	Breast Phyllodes Tumor	5	0.004	1.857
5	Ductal Breast Carcinoma in Situ	10	3.59E-04	1.727
6	Invasive Breast Carcinoma	21	1.61E-06	1.487
7	Invasive Ductal and Invasive Lobular Breast Carcinoma	90	5.47E-30	1.608
8	Invasive Ductal Breast Carcinoma	1,556	2.71E-61	1.652
9	Invasive Lobular Breast Carcinoma	148	3.07E-36	1.586
10	Medullary Breast Carcinoma	32	1.14E-07	1.898
11	Mucinous Breast Carcinoma	46	2.87E-12	1.373
12	Tubular Breast Carcinoma	67	1.49E-27	1.628

Figure 2. ILF2 expression in human breast tumors. A. ILF2 expression in 144 normal breast tissues and 1,992 different histological types of breast tumor tissues using Curtis breast dataset from Oncomine. Fold changes and *P*-values indicate the fold differences and statistical significances of ILF2 over-expression in breast tumor tissues compared with normal breast tissues. B. ILF2 expression in breast cancer tissues of different histological grades using GOBO analysis (P ≤ 0.00001). The box reflects the interquartile range (25%-75%), line reflects the median and the dots indicate the extreme data values. C. ILF2 expression in breast cancer tissues of different histological grades from Hatzis breast dataset from Oncomine (P = 1.91E-13). D. ILF2 expression in breast cancer tissues of different histological grades using Curtis breast dataset from Oncomine (P = 8.06E-18).

Meier-plotter online software (http://kmplot.com/analysis/) with 5,143 breast cancer patients with a mean follow-up of 69 months [27]. To analyse relapse-free survival (RFS), overall survival (OS) and distant metastasis-free survival (DMFS), the patient samples were first split into two groups according to the median level of *ILF2* expression. If no obvious effect was found, the software identified the best cut-off for the separation of patients into high- and low-expressing groups and calculated the hazard ratio with 95% confidence intervals and the log-rank *P*-value.

Clinical datasets, gene expression profiling and functional annotation

Raw data from three publicly available datasets focusing on anthracycline/taxane-based neoadjuvant chemotherapy responses were downloaded from Gene Expression Omnibus (GEO) with the accession numbers GSE25055 [28], GSE25065 [28] and GSE20194 [29]. The classification performance of *ILF2* expression level was assessed by constructing a receiver operating characteristic (ROC) curve in SPSS version 18 (IBM, Armonk, NY, USA), and the predictive power was evaluated using the area under the ROC curve. Two-sided *P*-values < 0.05 were considered statistically significant.

To explore the possible molecular functions and biological networks of ILF2, the samples in GSE25055 and GSE25065 were split into two groups according to high (top 25%) and low (bottom 25%) *ILF2* expression. A total of 592 genes that had a log fold change ≥ 1 and a

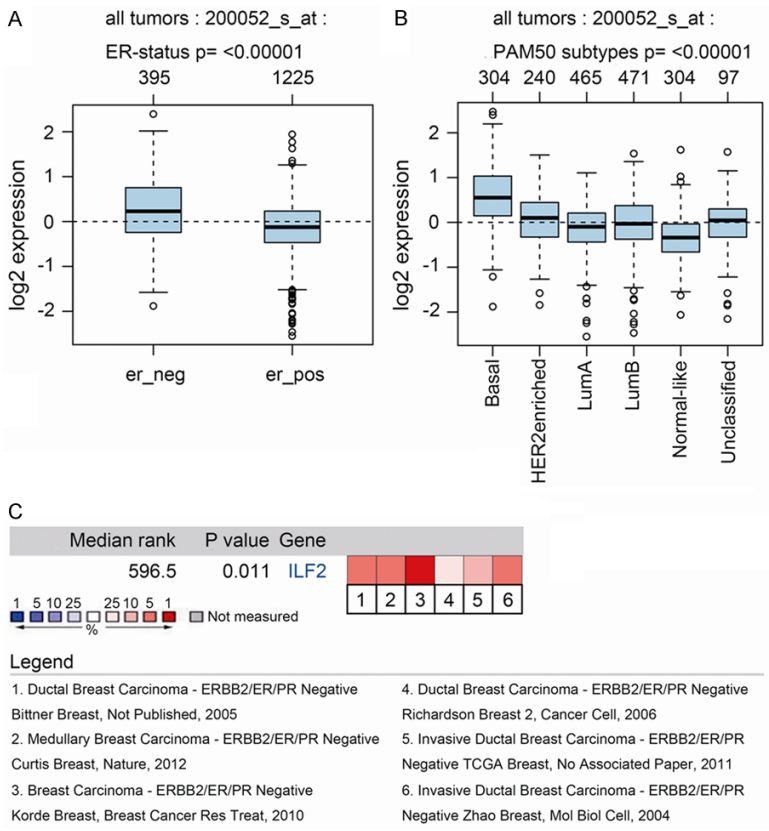


Figure 3. ILF2 expression in different molecular subtypes of breast cancer. A. ILF2 expression in 395 ER-negative and 1,225 ER-positive breast cancer tissues using GOBO analysis ($P \leq 0.00001$). B. ILF2 expression in 304 basal-like, 240 HER2-enriched, 465 luminal A, 471 luminal B, 304 normal-like and 97 unclassified breast cancer using GOBO analysis ($P \leq 0.00001$). C. ILF2 expression in TNBC. Meta-analysis of gene expression profiling for ILF2 gene in TNBC using Oncomine, with $P < 0.05$ and fold change > 1.5. The colored squares indicate the median rank for ILF2 across each analysis comparing TNBC with other subtypes.

P-value < 0.05 were considered significantly differentially expressed and were imported into IPA [30] (http://www.ingenuity.com), such that biological functions and diseases or canonical pathways of which ILF2 may participate could be inferred. Moreover, "Upstream analysis" and "Causal network analysis" were performed to identify molecules that were upstream regulators of *ILF2* and could possibly explain the observed expression changes.

Results

ILF2 variants in human cancers

The variant statuses of *ILF2* were examined in 29,379 cancer patients over 39 different types of cancer. Although only 16 cancer types possessed complete information for "Point Mutations", "Copy Number Variations" and "Gene Expression", the frequencies of *ILF2* variants in these patients are shown in **Figure 1**. The frequency of point mutations was not high, i.e., < 1%, in any cancer type. A gain in *ILF2* copy number was observed in several cancer types, such as tumours of the breast, lung, endometrium and urinary tract. The highest copy number variation frequency was 3.14%, which was found in breast cancer (among 1,749 patients). An overall trend towards increased *ILF2* expression in tumours was noted compared with the population norm. Taken together, the alterations in the copy number and gene expression suggested that ILF2 might participate in breast cancer development.

High ILF2 expression is correlated with aggressive breast cancer phenotypes

To confirm that *ILF2* expression was elevated in breast cancer, another dataset with a larger sample size was selected. Analysis of the Curtis breast dataset, which includes 2,136 clinical samples from the Oncomine database, showed significantly higher *ILF2* expression in different histological breast cancer subtypes compared with normal breast tissue (**Figure 2A**). Moreover, elevated *ILF2* expression was observed in high histological grade breast cancer using GOBO analysis ($P \leq 0.00001$; **Figure 2B**). Several datasets in the Oncomine database presented similar results ($P < 0.05$). The results of two clinical cohorts with the largest sample sizes are shown in **Figure 2C** and **2D**.

In addition to clinical parameters, DNA microarray technologies facilitated the identification of five major breast cancer subtypes (luminal A, luminal B, HER2-positive, basal-like, and normal-like), which are associated with different prognoses and treatment decisions [31]. GOBO

Table 1. ILF2 expression is upregulated in TNBC

Legend	Clinical cohort	No. of TNBC[#]	No. of other subtypes	Fold change	*P*-value	Gene rank
1	Bittner breast	39	129	1.885	4.7E-07	214 (in top 2%)
2	Curtis breast	20	12	1.596	0.019	755 (in top 4%)
3	Korde Breast	21	39	1.767	3.1E-05	94 (in top 1%)
4	Richardson breast 2	18	19	1.681	0.015	3401 (in top 18%)
5	TCGA breast	46	250	1.775	4.1E-10	1077 (in top 6%)
6	Zhao breast	5	29	1.904	0.003	438 (in top 4%)

[#]triple negative breast cancer.

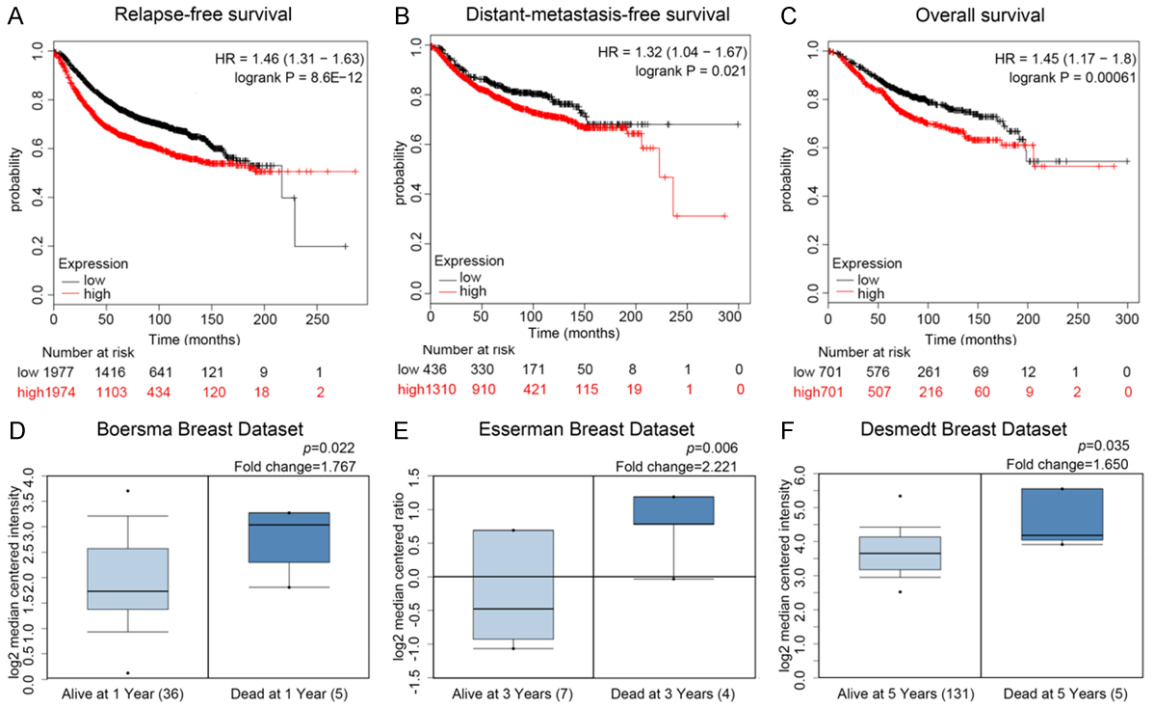


Figure 4. Elevated expression of ILF2 indicates worse clinical outcome in breast cancer patients. A-C. Kaplan-Meier relapse-free survival, distant metastasis-free survival and overall survival curves of breast cancer patients with high or low ILF2 expression, using Kaplan-Meier plotter analysis. D-F. ILF2 expression in breast cancer patients with different survival by using Boersma breast dataset, Esserman breast dataset, and Desmedt breast dataset from Oncomine.

analysis showed that *ILF2* expression was significantly higher in ER-negative than ER-positive breast cancers ($P \leq 0.00001$; **Figure 3A**). When investigating PAM50 intrinsic subtyping, the highest *ILF2* expression was observed in basal-like tissues (**Figure 3B**). To further confirm the distribution of *ILF2* expression in different breast cancer molecular subtypes, biomarker analysis was performed in the Oncomine database. Using six clinical cohorts, this analysis found that increased *ILF2* expression was significantly associated with triple negative breast cancer (TNBC) compared with the other molecular subtypes (fold change > 1.5, $P < 0.05$). The pooled results of the six clinical cohorts showed a significant increase in *ILF2* expression in TNBC ($P = 0.011$; **Figure 3C**); detailed information for each comparison is presented in **Table 1**. In summary, these findings indicated that elevated *ILF2* expression was correlated with more aggressive breast cancer subtypes, such as TNBC and basal-like breast cancer.

High ILF2 expression was correlated with poor prognoses

We further examined the influence of *ILF2* on the long-term survival of breast cancer patients

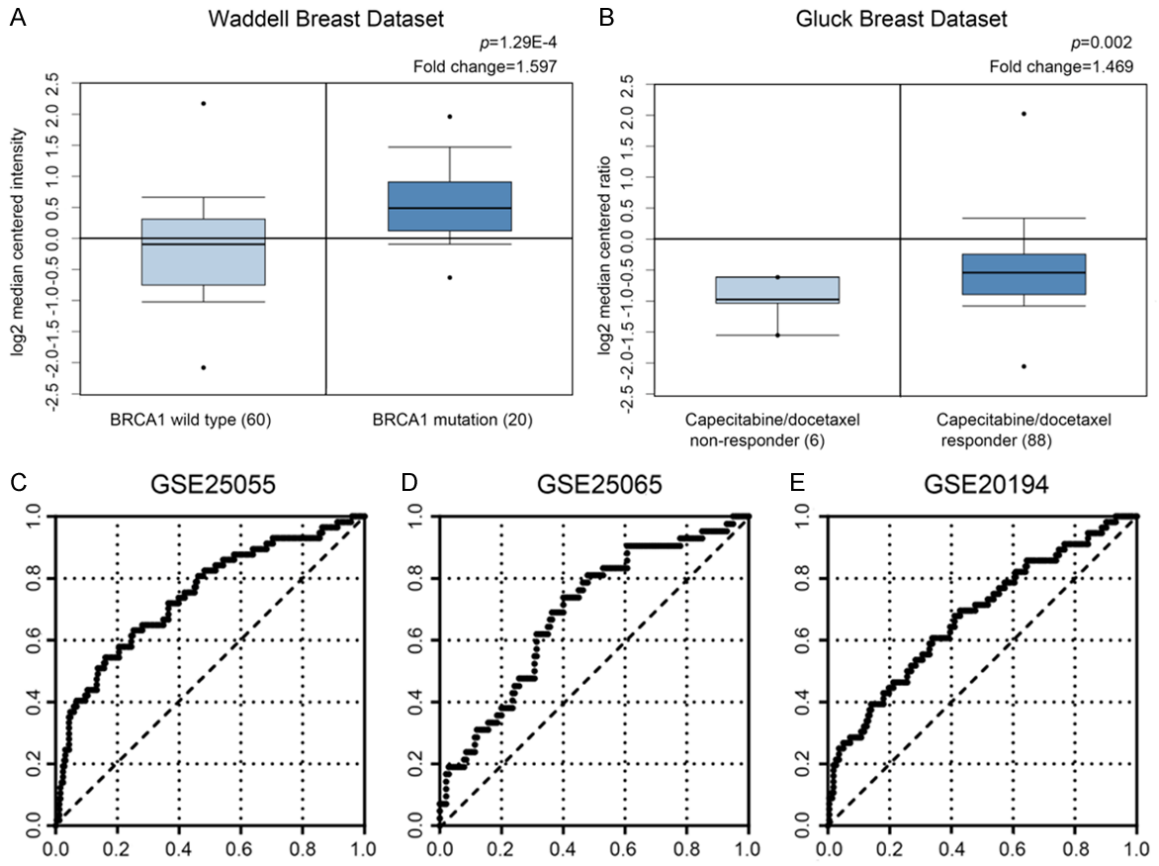


Figure 5. Elevated ILF2 expression correlated with BRCA1 mutation and response to chemotherapy in breast cancer. A. ILF2 expression in 60 BRCA1 wildtype patients and 20 BRCA1 mutation carriers using Waddell breast dataset from Oncomine (P = 1.29E-4). B. ILF2 expression in 6 capecitabine/docetaxel non-responders and 88 capecitabine/docetaxel responders using Gluck breast dataset from Oncomine (P = 0.002). C-E. ROC curves for ILF2 levels to discriminate the breast cancer patients who achieve pCR from non-pCR after anthracycline-taxane based neoadjuvant chemotherapy using data from GEO datasets GSE25055, GSE25065 and GSE20194.

using Kaplan-Meier plotter, which is a large publicly available clinical breast cancer microarray database. High *ILF2* mRNA expression was found to be an indicator of increased risk for relapse (P = 8.6E-12), distant metastasis (P = 0.021), and death (P = 0.00061; **Figure 4A-C**). This observation was also confirmed using three clinical cohorts from the Oncomine database (fold change > 1.5, P < 0.05), which showed that breast patients with elevated *ILF2* mRNA levels had relatively low 1-, 3- and 5-year OS rates (**Figure 4D-F**).

High ILF2 expression predicted anthracycline/taxane-based neoadjuvant chemotherapyresponses

BRCA proteins participate in DNA double-strand break (DSB) repair, and thus BRCA deficiencies increase genomic instability and promote the development of more aggressive tumour phenotypes [32, 33]. Oncomine analysis hinted that *ILF2* expression was higher in BRCA1 mutation carriers compared with BRCA1 wildtype patients (fold change = 1.597, P = 1.29E-4; **Figure 5A**). Moreover, we also found that the *ILF2* mRNA expression levels in pretreatment breast cancer biopsy specimens from chemosensitive patients were higher than in non-responders (fold change = 1.469, P = 0.002; **Figure 5B**).

BRCA1 mutation carriers show better responses to cytotoxic agents, which might be due to their deficiency in repairing the DNA DSBs induced by chemotherapy [34]. Additionally, based on the Oncomine results, *ILF2* expression appears to be correlated with chemoresponses. Therefore, we further investigated correlations between *ILF2* expression and che-

Table 2. AUC-ROC analysis of ILF2 expression in pretreatment biopsy specimens for predicting pCR

GEO datasets	All patients			AUC*	*P*-value	95% Confidence internal	
	Total	No. of pCR#	No. of non-pCR#			Lower bound	Upper bound
GSE25055	306	57	249	0.745	< 0.001	0.671	0.819
GSE25065	182	42	140	0.687	< 0.001	0.598	0.777
GSE20194	278	56	222	0.677	< 0.001	0.596	0.758

#pathological complete response. *area under the ROC curve.

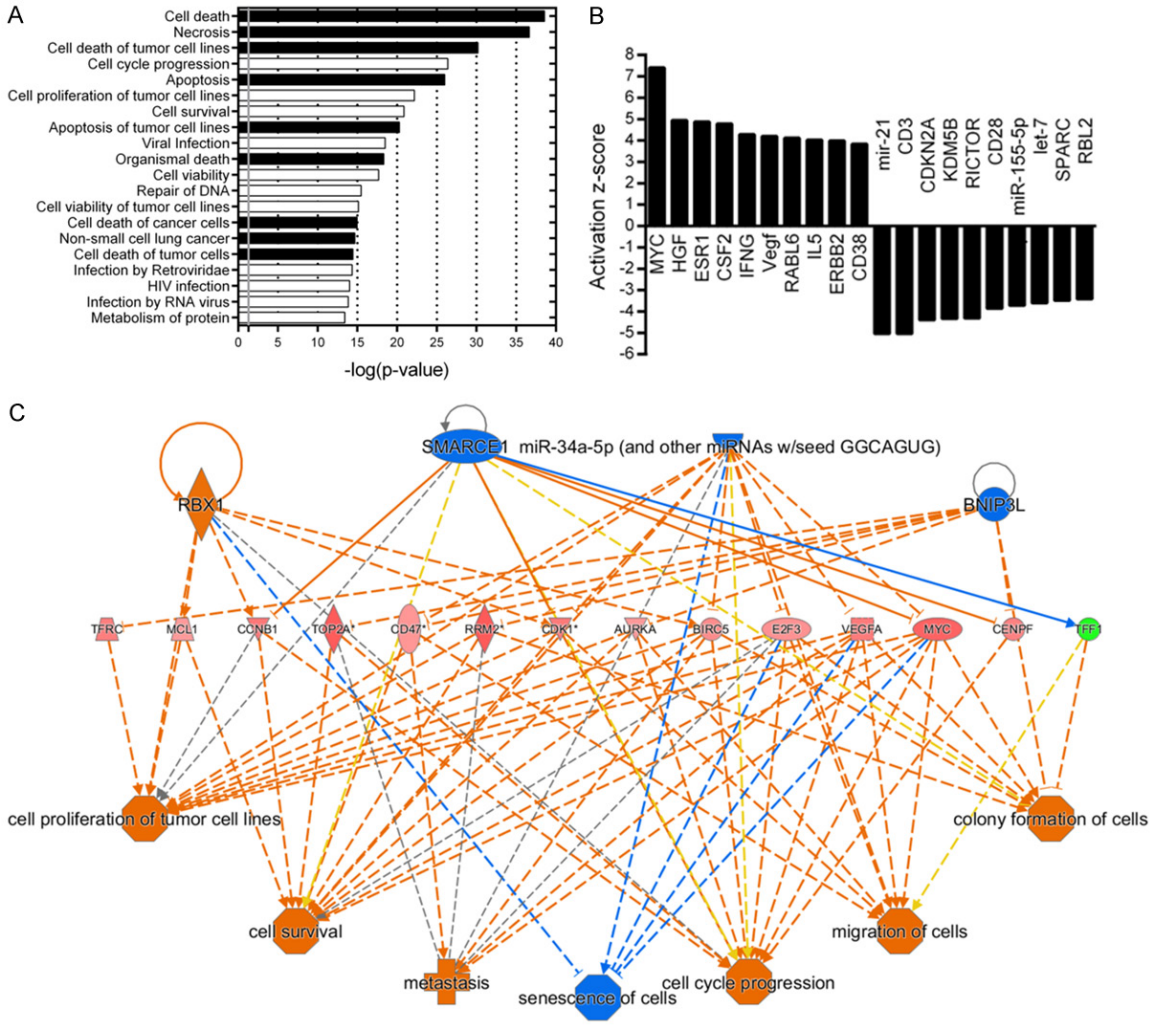


Figure 6. Biological functions and upstream regulators associated with ILF2-related DEGs in breast cancer, by using IPA. A. The top 20 diseases and functions. Activated (white) and inhibited (black) ILF2-related biological functions are presented as -log (*P*-value) for the probability that the specific function is affected. The gray vertical line represent least significant differences (-log (*P*-value) = 1.301/*P*-value = 0.05). B. The predicted upstream regulators. C. The predicted regulatory network. IPA's regulator effects algorithm connected the ILF2-related DEGs and predicted upstream regulators to downstream functions to generate regulator effects hypotheses with a consistency score. The diagram displays a result with high consistency score (15.50). Activated and inhibited regulators or functions are marked in orange and blue, respectively.

mo-responses in three larger GEO cohorts (GSE25055, GSE25065 and GSE20194) that received the commonly used anthracycline/taxane-based neoadjuvant chemotherapy regimen. Breast cancer patients with higher levels of *ILF2* expression in pre-treatment biopsy specimens were more likely to achieve pathological complete response (pCR) after anthra-

Table 3. The top 20 ILF2-related pathways from the IPA, based on most-significant *P*-values

Ingenuity Canonical Pathways	*P*-value	Ratio	Activation z-score
Cell cycle: G2/M DNA damage checkpoint regulation	5.89E-10	28.60%	-1.265
Protein ubiquitination pathway	5.89E-10	10.90%	NA
Sumoylation pathway	4.57E-06	14.60%	-1.265
RAN signaling	4.57E-06	41.20%	NA
Oxidative phosphorylation	1.41E-05	12.80%	NA
Mitochondrial dysfunction	0.00013	9.36%	NA
Estrogen-mediated S-phase entry	0.00069	25.00%	0.816
Mismatch repair in eukaryotes	0.001	31.20%	NA
Remodeling of epithelial adherens junctions	0.00107	13.00%	2.236
Cell cycle control of chromosomal replication	0.00112	17.50%	NA
Hypoxia signaling in the cardiovascular system	0.00155	12.20%	NA
Adipogenesis pathway	0.00182	8.96%	NA
Role of BRCA1 in DNA damage response	0.002	11.50%	-0.447
ATM signaling	0.00229	11.20%	0.378
Hereditary breast cancer signaling	0.00288	8.33%	NA
Mitotic roles of polo-like kinase	0.00288	12.10%	2
TCA cycle II (eukaryotic)	0.00302	21.70%	NA
Cholesterol biosynthesis I	0.00302	30.80%	NA
Cholesterol biosynthesis II (via 24, 25-dihydrolanosterol)	0.00302	30.80%	NA
Cholesterol biosynthesis III (via desmosterol)	0.00302	30.80%	NA

cycline/taxane-based neoadjuvant treatment in all three datasets (GSE25055, GSE25065 and GSE20194), with area under the ROC curve (AUC) values of 0.745 ($P < 0.001$), 0.687 ($P < 0.001$) and 0.677 ($P < 0.001$), respectively (**Figure 5C-E**; **Table 2**). These results indicated that ILF2 expression maybe a predictive biomarker for better responses to anthracycline/taxane-based treatment in breast cancer patients.

IPA reveals potential ILF2 regulatory mechanisms

We used IPA software to analyse the 592 ILF2-related differentially expressed genes (DEGs) (29 downregulated and 563 upregulated) and found multiple statistically significant related biological functions ($P < 0.05$), the top 20 of which are presented in **Figure 6A**. The most affected biological functions in high *ILF2*-expressing breast cancer tissues were involved in promoting cell survival, viability and proliferation. Additionally, the IPA software also predicted that *ILF2* overexpression was correlated with "cell cycle progression", "viral infection", "repair of DNA" and "metabolism of protein".

To more deeply explore the relevant mechanisms underlying ILF2-related biological functions, we identified 44 canonical pathways ($P < 0.05$), and the top 20 most significant pathways are listed in **Table 3**. Consistent with the observations above, these DEGs were highly enriched for pathways related to cell cycle progression such as "Cell Cycle: G2/M DNA Damage Checkpoint Regulation", "Estrogen-mediated S-phase Entry", "Cell Cycle Control of Chromosomal Replication" and "Mitotic Roles of Polo-Like Kinase". To further elucidate the role of ILF2 in G2/M transition, we explored correlations between *ILF2* expression and several cell cycle regulators using TCGA breast dataset from Oncomine. Activated Cyclin A/B-CDK1 complex is essential for G2/M phase progression [35]. Polo-like kinase 1 (PLK1) and Aurora A are well-established regulators of mitotic progression, and their hyperactivity can override DNA damage-induced G2 checkpoint arrest [36-38]. We found that *ILF2* mRNA levels in breast cancer tissues are positively correlated with those of *CCNA2* (Cyclin A2), *CCNB1* (Cyclin B1), *CDK1*, *AURKA* (Aurora A), *AURKB* (Aurora B) and *PLK1* (Supplementary Figure 1). DEGs were also enriched for pathways related to DNA repair, such as "Mismatch Repair in Eukaryotes", "Role of BRCA1 in DNA Damage Response" and "ATM Signalling". Additionally, we observed that pathways involved in energy

metabolism and biosynthesis were changed, including “Oxidative Phosphorylation”, ”Mitochondrial Dysfunction”, ”Adipogenesis pathway”, ”TCA Cycle” and “Cholesterol Biosynthesis”. Overactivation of “Remodelling of Epithelial Adherens Junctions” implied that breast cancers with high *ILF2* expression might possess strong invasive ability.

IPA upstream analysis was used to predict which upstream regulators might be activated or inhibited to explain the gene expression changes caused by *ILF2* expression. Several transcriptional regulators were predicted (z-score value ≥ 2 for activation and ≤ -2.0 for inhibition; $P < 0.05$), and the top 10 activated or inhibited upstream regulators are shown in **Figure 6B**. Regarding activated upstream regulators, except for certain well-known oncogenes such as MYC and HGF, activation of two of the most well-studied molecules involved in breast cancer, i.e., ESR1 and ERBB2, were also predicted, which implied that ILF2 might participate in the progression of luminal and HER2-positive breast cancer by activating the oestrogen pathway and ERBB2 pathway, respectively. Moreover, we noticed that several highly significant predicted upstream regulators were cytokines (such as CSF2, IFNG and IL5). Considering the role that ILF2 plays in the transcription of IL2 and IL13 in T cells [6, 39], we hypothesize that additional communication occurs between breast cancer cells and tumour-infiltrating lymphocytes in breast cancer tissues with high *ILF2* expression. With respect to inhibited upstream regulators, the upstream analysis revealed several microRNAs, such as miR-21, miR-155-5p and let-7. ILF2 has been reported to inhibit the maturation of several microRNAs, including let-7, miR-21 [10], miR-133a [11] and miR-7 [12], by restraining primary miRNA processing into precursor miRNAs. miR-21 (activation z-score = -5.041), let-7 (activation z-score = -3.604), and miR-133a-3p (activation z-score = -2.938) were also identified as upregulated regulators in the IPA upstream analysis, which adds confidence to the reliability of this method. Other predicted downregulated microRNAs such as miR-155-5p (activation z-score = -3.719), miR-1-3p (activation z-score = -3.225) and miR-124-3p (activation z-score = -3.141), which have been reported to influence breast cancer progression [40-42], might be novel ILF2 substrates.

Finally, causal networks were constructed to connect the DEGs and the predicted upstream regulators to downstream functions. **Figure 6C** illustrates one of the top mechanistic hypotheses that explains the expression changes observed in datasets grouped by *ILF2* levels.

Discussion

Due to both inter- and intra-tumoural heterogeneity in breast cancer, patients with the same tumour stage or grade can have different therapeutic responses, metastatic patterns, and/or clinical outcomes. Fortunately, the development of molecular oncology has identified several biomarkers that can provide additional information about the intrinsic characteristics of individual breast cancer patients. Some of them (e.g., ER, HER2) have been widely used in the clinic to predict patient outcomes and serve as therapeutic targets. ILF2 forms a heterodimeric core complex with NF90, which binds to nucleic acids, altering gene expression at multiple levels. In this work using publicly available clinical datasets, we found that ILF2 was involved in the development and progression of breast cancer and may be a predictive biomarker for better anthracycline/taxane-based treatment responses.

ILF2 is located on chromosome 1q21 [39], a region frequently amplified in breast cancer [43, 44]. We found that *ILF2* copy number gain was present in 3.14% of breast cancer patients and that there was higher *ILF2* mRNA expression in breast cancer tissues compared with normal breast tissues. Using publicly available clinical datasets with large sample sizes, we further showed that elevated *ILF2* expression was correlated with some aggressive features, such as high histological grade, BRCA1 mutations, the triple-negative/basal-like subtype, and shorter RFS, DMSF and OS. IPA indicated that ILF2-related DEGs were enriched in biological functions involved in promoting cell survival, viability, proliferation and inhibiting cell death, necrosis and apoptosis. Further analysis showed that several pathways correlated with cell cycle progression and DNA repair were also altered. These findings are consistent with observations in other cancer types. Depletion of the ILF2/NF90 complex sensitizes cervical carcinoma cells to genotoxic drug-induced apoptosis by restoring p53 expression

via reducing HPV-E6 RNA expression [45]. Zheng et al. [18] also reported that ILF2 blocked apoptosis in liver cancer cells by upregulating the anti-apoptotic proteins Bcl-2 and cIAP1, and downregulating the pro-apoptotic proteins Bax, Bak, and Bok. However, ILF2 also stimulates cell proliferation by promoting cell cycle progression in non-small cell lung cancer [19], pancreatic carcinoma [16], oesophageal squamous cell carcinoma [20], and glioma [46] cells.

MicroRNAs are short non-coding RNAs that regulate many crucial cancer-related pathways by post-transcriptionally silencing mRNAs [47]. The intimate and complicated relationship between ILF2 and microRNAs was also revealed in cancer cells. Sakamoto et al. [12] found that that NF90/ILF2 complex promoted the proliferation of hepatocellular carcinoma cells by restraining miR-7 biogenesis. ILF2 also inhibited the maturation of let-7, miR-21, and miR-133a [10, 11], which are known to affect cancer progression [48-50]. However, *ILF2* expression is also regulated by microRNAs. Recently, a study reported that miR-7 reduced *ILF2* expression and regulated epithelial-mesenchymal transition-related genes in pancreatic carcinoma [17]. Similarly, we observed that several microRNAs, including miR-155-5p, miR-1-3p, miR-124-3p, and some of the previously reported ILF2-regulated microRNAs (miR-21, let-7, and miR-133a-3p) were predicted to be inhibited upstream regulators of the ILF2-related expression network in breast cancer. This result indicated that ILF2-induced breast cancer progression may partially rely on interactions with microRNAs.

In this study, we also found that breast cancer patients with elevated *ILF2* expression were more likely to achieve pCR after anthracycline/taxane-based neoadjuvant treatment. Unlike our study, Colla et al. [9] found that depleting *ILF2* increased the sensitivity of multiple myeloma cells to DNA damaging agents by affecting YB-1 nuclear localization and thereby decreasing homologous recombination repair. The IPA results offer certain clues to explain the contradictory effects of ILF2 on multiple myeloma and breast cancer. After DNA damage caused by chemotherapy or radiotherapy, cell cycle arrest at the G2/M checkpoint allows DNA repair. Interestingly, previous research has shown that ILF2 plays roles in both cell cycle progression and DNA repair. According to our IPA results, although selected pathways involved in DNA repair were disrupted, a clearer trend towards promoting cell cycle progression can be observed in breast cancer. We found that upregulated *ILF2* expression was correlated with downregulation of the G2/M DNA damage checkpoint and PLK pathway activation. Additionally, we observed that elevated *ILF2* expression in breast cancer tissues was positively correlated with the expression of *CCNA2*, *CCNB1*, *CDK1*, *AURKA*, *AURKB* and PLK1, which are mitotic inducers and checkpoint regulators. Taken together, these data suggest that *ILF2* overexpression may lead to defective checkpoint control, which causes cell cycle progression and the loss of genomic integrity in breast cancer. As G2/M checkpoint defects have been confirmed to contribute to the sensitivity to DNA damaging agents [51], the higher pCR rates in breast cancer patients with elevated *ILF2* expression was consistent with previous studies. However, breast cancer patients with higher expression of proliferation genes are more likely to achieve pCR [52], thus ILF2-induced cellular proliferation could also lead to increased pCR rates. Although more breast cancers with elevated *ILF2* expression respond to chemotherapy initially, early relapses occur because of genomic instability caused by overriding the checkpoint.

Finally, we are the first group to investigate the roles of ILF2 in breast cancer development and progression. This study benefited from a large sample size facilitated by publicly available online datasets, which offer opportunities to derive more reliable and precise results. Gene expression data used in IPA were sourced from pre-treatment tumour biopsies from clinical breast cancer cases, and hence interactions between various cell types in the tumour microenvironment could be explored. This approach might be more suitable because ILF2 has been found to regulate cytokine synthesis in lymphocytes [6, 39]. Nevertheless, certain limitations in our study should be considered when interpreting the results. First, all results are based on *ILF2* expression at the mRNA level; future protein-level investigations are warranted. Moreover, ILF2-related biological functions, canonical pathways and upstream regulators were calculated based on prior biological knowledge in the IPA ingenuity knowl-

edge database. These findings must be further confirmed by mechanistic studies performed in breast cancer cells.

In conclusion, we showed that *ILF2* may be involved in breast cancer carcinogenesis and that elevated *ILF2* expression was correlated with certain aggressive features in breast cancer, such as high histological grade, BRCA1 mutations and the triple-negative/basal-like subtype. More importantly, *ILF2* can help select breast cancer patients who are likely to benefit from anthracycline/taxane-based neoadjuvant chemotherapy.

Acknowledgements

This work was supported by Subject Development Project of China Medical University (No. 3110117053) and National Natural Science Foundation Project (No. 81372811).

Disclosure of conflict of interest

None.

Address correspondence to: Bo Chen, Department of Breast Surgery, First Hospital of China Medical University, 155 North Nanjing Street, Heping District, Shenyang 110001, Liaoning Province, China. E-mail: chbyxl@163.com

References

[1] Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J and Jemal A. Global cancer statistics, 2012. CA Cancer J Clin 2015; 65: 87-108.

[2] Prat A and Perou CM. Deconstructing the molecular portraits of breast cancer. Mol Oncol 2011; 5: 5-23.

[3] Kao PN, Chen L, Brock G, Ng J, Kenny J, Smith AJ and Corthesy B. Cloning and expression of cyclosporin A- and FK506-sensitive nuclear factor of activated T-cells: NF45 and NF90. J Biol Chem 1994; 269: 20691-20699.

[4] Shamanna RA, Hoque M, Lewis-Antes A, Azzam EI, Lagunoff D, Pe'ery T and Mathews MB. The NF90/NF45 complex participates in DNA break repair via nonhomologous end joining. Mol Cell Biol 2011; 31: 4832-4843.

[5] Karmakar S, Mahajan MC, Schulz V, Boyapaty G and Weissman SM. A multiprotein complex necessary for both transcription and DNA replication at the beta-globin locus. EMBO J 2010; 29: 3260-3271.

[6] Kiesler P, Haynes PA, Shi L, Kao PN, Wysocki VH and Vercelli D. NF45 and NF90 regulate HS4-dependent interleukin-13 transcription in T cells. J Biol Chem 2010; 285: 8256-8267.

[7] Graber TE, Baird SD, Kao PN, Mathews MB and Holcik M. NF45 functions as an IRES transacting factor that is required for translation of cIAP1 during the unfolded protein response. Cell Death Differ 2010; 17: 719-729.

[8] Faye MD, Graber TE, Liu P, Thakor N, Baird SD, Durie D and Holcik M. Nucleotide composition of cellular internal ribosome entry sites defines dependence on NF45 and predicts a posttranscriptional mitotic regulon. Mol Cell Biol 2013; 33: 307-318.

[9] Marchesini M, Ogoti Y, Fiorini E, Aktas Samur A, Nezi L, D'Anca M, Storti P, Samur MK, Ganan-Gomez I, Fulciniti MT, Mistry N, Jiang S, Bao N, Marchica V, Neri A, Bueso-Ramos C, Wu CJ, Zhang L, Liang H, Peng X, Giuliani N, Draetta G, Clise-Dwyer K, Kantarjian H, Munshi N, Orlowski R, Garcia-Manero G, DePinho RA and Colla S. ILF2 is a regulator of RNA splicing and DNA damage response in 1q21-Amplified multiple myeloma. Cancer Cell 2017; 32: 88-100, e106.

[10] Sakamoto S, Aoki K, Higuchi T, Todaka H, Morisawa K, Tamaki N, Hatano E, Fukushima A, Taniguchi T and Agata Y. The NF90-NF45 complex functions as a negative regulator in the microRNA processing pathway. Mol Cell Biol 2009; 29: 3754-3769.

[11] Todaka H, Higuchi T, Yagyu K, Sugiyama Y, Yamaguchi F, Morisawa K, Ono M, Fukushima A, Tsuda M, Taniguchi T and Sakamoto S. Overexpression of NF90-NF45 represses myogenic MicroRNA biogenesis, resulting in development of skeletal muscle atrophy and centronuclear muscle fibers. Mol Cell Biol 2015; 35: 2295-2308.

[12] Higuchi T, Todaka H, Sugiyama Y, Ono M, Tamaki N, Hatano E, Takezaki Y, Hanazaki K, Miwa T, Lai S, Morisawa K, Tsuda M, Taniguchi T and Sakamoto S. Suppression of MicroRNA-7 (miR-7) biogenesis by nuclear factor 90-Nuclear factor 45 complex (NF90-NF45) controls cell proliferation in hepatocellular carcinoma. J Biol Chem 2016; 291: 21074-21084.

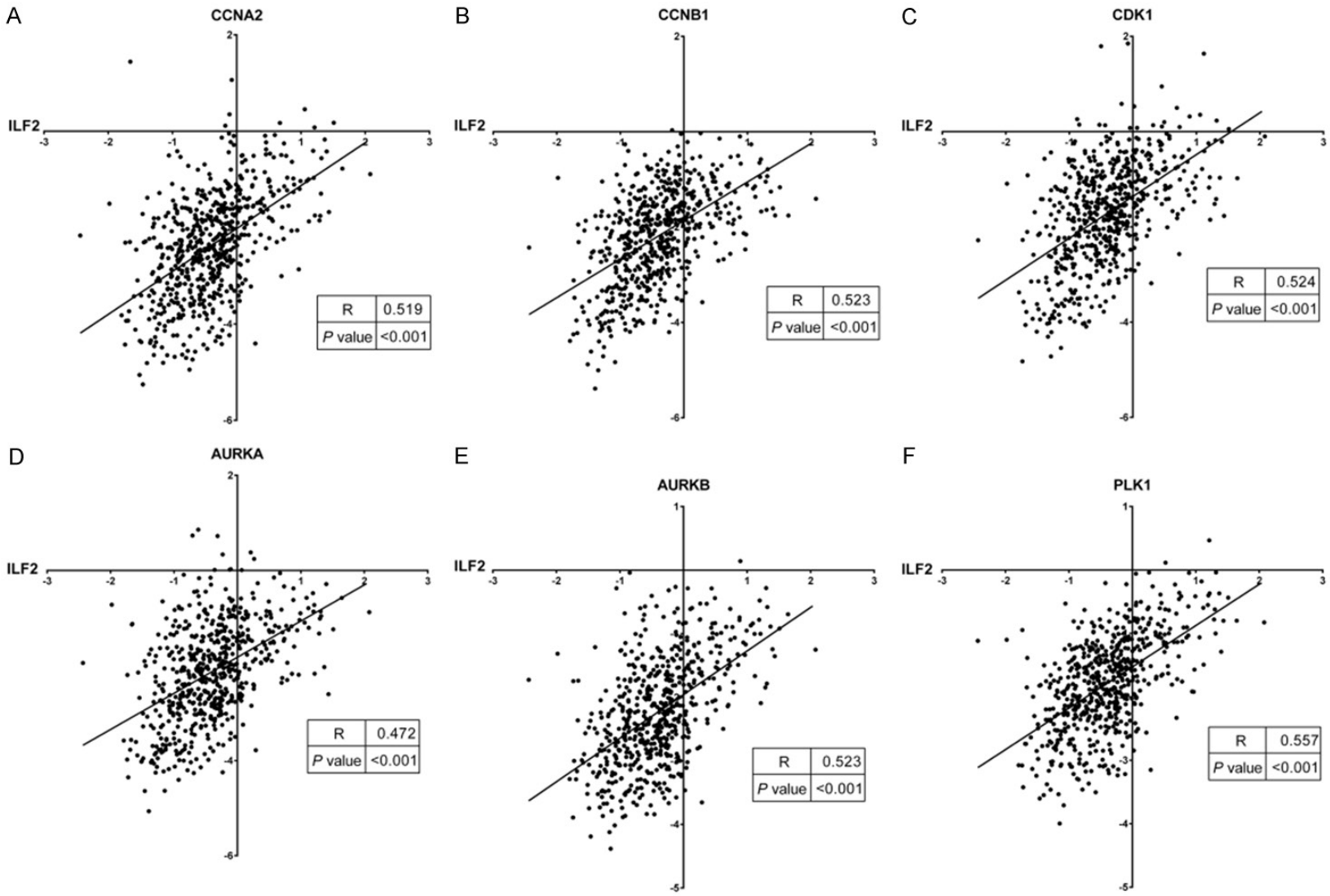
[13] Isken O, Baroth M, Grassmann CW, Weinlich S, Ostareck DH, Ostareck-Lederer A and Behrens SE. Nuclear factors are involved in hepatitis C virus RNA replication. RNA 2007; 13: 1675-1692.

[14] Merrill MK and Gromeier M. The double-stranded RNA binding protein 76:NF45 heterodimer inhibits translation initiation at the rhinovirus type 2 internal ribosome entry site. J Virol 2006; 80: 6936-6942.

[15] Guan D, Altan-Bonnet N, Parrott AM, Arrigo CJ, Li Q, Khaleduzzaman M, Li H, Lee CG, Pe'ery T and Mathews MB. Nuclear factor 45 (NF45) is a regulatory subunit of complexes with

NF90/110 involved in mitotic control. Mol Cell Biol 2008; 28: 4629-4641.
[16] Wan C, Gong C, Ji L, Liu X, Wang Y, Wang L, Shao M, Yang L, Fan S, Xiao Y, Wang X, Li M, Zhou G and Zhang Y. NF45 overexpression is associated with poor prognosis and enhanced cell proliferation of pancreatic ductal adenocarcinoma. Mol Cell Biochem 2015; 410: 25-35.
[17] Bi Y, Shen W, Min M and Liu Y. MicroRNA-7 functions as a tumor-suppressor gene by regulating ILF2 in pancreatic carcinoma. Int J Mol Med 2017; 39: 900-906.
[18] Cheng S, Jiang X, Ding C, Du C, Owusu-Ansah KG, Weng X, Hu W, Peng C, Lv Z, Tong R, Xiao H, Xie H, Zhou L, Wu J and Zheng S. Expression and critical role of interleukin enhancer binding factor 2 in hepatocellular carcinoma. Int J Mol Sci 2016; 17.
[19] Ni T, Mao G, Xue Q, Liu Y, Chen B, Cui X, Lv L, Jia L, Wang Y and Ji L. Upregulated expression of ILF2 in non-small cell lung cancer is associated with tumor cell proliferation and poor prognosis. J Mol Histol 2015; 46: 325-335.
[20] Ni S, Zhu J, Zhang J, Zhang S, Li M, Ni R, Liu J, Qiu H, Chen W, Wang H and Guo W. Expression and clinical role of NF45 as a novel cell cycle protein in esophageal squamous cell carcinoma (ESCC). Tumour Biol 2015; 36: 747-756.
[21] Yin ZH, Jiang XW, Shi WB, Gui QL and Yu DF. Expression and clinical significance of ILF2 in gastric cancer. Dis Markers 2017; 2017: 4387081.
[22] Ye J, Jin H, Pankov A, Song JS and Blelloch R. NF45 and NF90/NF110 coordinately regulate ESC pluripotency and differentiation. RNA 2017; 23: 1270-1284.
[23] Hu Q, Lu YY, Noh H, Hong S, Dong Z, Ding HF, Su SB and Huang S. Interleukin enhancer-binding factor 3 promotes breast tumor progression by regulating sustained urokinase-type plasminogen activator expression. Oncogene 2013; 32: 3933-3943.
[24] Forbes SA, Beare D, Boutselakis H, Bamford S, Bindal N, Tate J, Cole CG, Ward S, Dawson E, Ponting L, Stefancsik R, Harsha B, Kok CY, Jia M, Jubb H, Sondka Z, Thompson S, De T and Campbell PJ. COSMIC: somatic cancer genetics at high-resolution. Nucleic Acids Res 2017; 45: D777-D783.
[25] Rhodes DR, Kalyana-Sundaram S, Mahavisno V, Varambally R, Yu J, Briggs BB, Barrette TR, Anstet MJ, Kincead-Beal C, Kulkarni P, Varambally S, Ghosh D and Chinnaiyan AM. Oncomine 3.0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles. Neoplasia 2007; 9: 166-180.
[26] Ringner M, Fredlund E, Hakkinen J, Borg A and Staaf J. GOBO: gene expression-based outcome for breast cancer online. PLoS One 2011; 6: e17911.
[27] Gyorffy B, Lanczky A, Eklund AC, Denkert C, Budczies J, Li Q and Szallasi Z. 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: 725-731.
[28] Hatzis C, Pusztai L, Valero V, Booser DJ, Esserman L, Lluch A, Vidaurre T, Holmes F, Souchon E, Wang H, Martin M, Cotrina J, Gomez H, Hubbard R, Chacon JI, Ferrer-Lozano J, Dyer R, Buxton M, Gong Y, Wu Y, Ibrahim N, Andreopoulou E, Ueno NT, Hunt K, Yang W, Nazario A, DeMichele A, O'Shaughnessy J, Hortobagyi GN and Symmans WF. A genomic predictor of response and survival following taxane-anthracycline chemotherapy for invasive breast cancer. JAMA 2011; 305: 1873-1881.
[29] Popovici V, Chen W, Gallas BG, Hatzis C, Shi W, Samuelson FW, Nikolsky Y, Tsyganova M, Ishkin A, Nikolskaya T, Hess KR, Valero V, Booser D, Delorenzi M, Hortobagyi GN, Shi L, Symmans WF and Pusztai L. Effect of training-sample size and classification difficulty on the accuracy of genomic predictors. Breast Cancer Res 2010; 12: R5.
[30] Kramer A, Green J, Pollard J Jr and Tugendreich S. Causal analysis approaches in ingenuity pathway analysis. Bioinformatics 2014; 30: 523-530.
[31] Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Lonning PE and Borresen-Dale AL. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A 2001; 98: 10869-10874.
[32] Tutt A and Ashworth A. The relationship between the roles of BRCA genes in DNA repair and cancer predisposition. Trends Mol Med 2002; 8: 571-576.
[33] Lakhani SR, Jacquemier J, Sloane JP, Gusterson BA, Anderson TJ, van de Vijver MJ, Farid LM, Venter D, Antoniou A, Storfer-Isser A, Smyth E, Steel CM, Haites N, Scott RJ, Goldgar D, Neuhausen S, Daly PA, Ormiston W, McManus R, Scherneck S, Ponder BA, Ford D, Peto J, Stoppa-Lyonnet D, Bignon YJ, Struewing JP, Spurr NK, Bishop DT, Klijn JG, Devilee P, Cornelisse CJ, Lasset C, Lenoir G, Barkardottir RB, Egilsson V, Hamann U, Chang-Claude J, Sobol H, Weber B, Stratton MR and Easton DF. Multifactorial analysis of differences between sporadic breast cancers and cancers involving BRCA1 and BRCA2 mutations. J Natl Cancer Inst 1998; 90: 1138-1145.
[34] Chappuis PO, Goffin J, Wong N, Perret C, Ghadirian P, Tonin PN and Foulkes WD. A significant response to neoadjuvant chemotherapy in BRCA1/2 related breast cancer. J Med Genet 2002; 39: 608-610.

[35] Otto T and Sicinski P. Cell cycle proteins as promising targets in cancer therapy. Nat Rev Cancer 2017; 17: 93-115.
[36] Seki A, Coppinger JA, Jang CY, Yates JR and Fang G. Bora and the kinase Aurora a cooperatively activate the kinase Plk1 and control mitotic entry. Science 2008; 320: 1655-1658.
[37] Macurek L, Lindqvist A, Lim D, Lampson MA, Klompmaker R, Freire R, Clouin C, Taylor SS, Yaffe MB and Medema RH. Polo-like kinase-1 is activated by aurora A to promote checkpoint recovery. Nature 2008; 455: 119-123.
[38] van Vugt MA, Bras A and Medema RH. Polo-like kinase-1 controls recovery from a G2 DNA damage-induced arrest in mammalian cells. Mol Cell 2004; 15: 799-811.
[39] Zhao G, Shi L, Qiu D, Hu H and Kao PN. NF45/ILF2 tissue expression, promoter analysis, and interleukin-2 transactivating function. Exp Cell Res 2005; 305: 312-323.
[40] Kong W, He L, Richards EJ, Challa S, Xu CX, Permuth-Wey J, Lancaster JM, Coppola D, Sellers TA, Djeu JY and Cheng JQ. Upregulation of miRNA-155 promotes tumour angiogenesis by targeting VHL and is associated with poor prognosis and triple-negative breast cancer. Oncogene 2014; 33: 679-689.
[41] Liu C, Zhang S, Wang Q and Zhang X. Tumor suppressor miR-1 inhibits tumor growth and metastasis by simultaneously targeting multiple genes. Oncotarget 2017; 8: 42043-42060.
[42] Wang Y, Chen L, Wu Z, Wang M, Jin F, Wang N, Hu X, Liu Z, Zhang CY, Zen K, Chen J, Liang H, Zhang Y and Chen X. miR-124-3p functions as a tumor suppressor in breast cancer by targeting CBL. BMC Cancer 2016; 16: 826.
[43] Guan XY, Meltzer PS, Dalton WS and Trent JM. Identification of cryptic sites of DNA sequence amplification in human breast cancer by chromosome microdissection. Nat Genet 1994; 8: 155-161.
[44] James LA, Mitchell EL, Menasce L and Varley JM. Comparative genomic hybridisation of ductal carcinoma in situ of the breast: identification of regions of DNA amplification and deletion in common with invasive breast carcinoma. Oncogene 1997; 14: 1059-1065.
[45] Shamanna RA, Hoque M, Pe'ery T and Mathews MB. Induction of p53, p21 and apoptosis by silencing the NF90/NF45 complex in human papilloma virus-transformed cervical carcinoma cells. Oncogene 2013; 32: 5176-5185.
[46] Huang Q, He X, Qiu X, Liu X, Sun G, Guo J, Ding Z, Yang L, Ban N, Tao T and Wang D. Expression of NF45 correlates with malignant grade in gliomas and plays a pivotal role in tumor growth. Tumour Biol 2014; 35: 10149-10157.
[47] Jansson MD and Lund AH. MicroRNA and cancer. Mol Oncol 2012; 6: 590-610.
[48] Johnson CD, Esquela-Kerscher A, Stefani G, Byrom M, Kelnar K, Ovcharenko D, Wilson M, Wang X, Shelton J, Shingara J, Chin L, Brown D and Slack FJ. The let-7 microRNA represses cell proliferation pathways in human cells. Cancer Res 2007; 67: 7713-7722.
[49] Pfeffer SR, Yang CH and Pfeffer LM. The role of miR-21 in cancer. Drug Dev Res 2015; 76: 270-277.
[50] Yang QS, Jiang LP, He CY, Tong YN and Liu YY. Up-regulation of MicroRNA-133a inhibits the MEK/ERK signaling pathway to promote cell apoptosis and enhance radio-sensitivity by targeting EGFR in esophageal cancer in vivo and in vitro. J Cell Biochem 2017; 118: 2625-2634.
[51] Rundle S, Bradbury A, Drew Y and Curtin NJ. Targeting the ATR-CHK1 Axis in cancer therapy. Cancers (Basel) 2017; 9.
[52] Gianni L, Zambetti M, Clark K, Baker J, Cronin M, Wu J, Mariani G, Rodriguez J, Carcangiu M, Watson D, Valagussa P, Rouzier R, Symmans WF, Ross JS, Hortobagyi GN, Pusztai L and Shak S. Gene expression profiles in paraffin-embedded core biopsy tissue predict response to chemotherapy in women with locally advanced breast cancer. J Clin Oncol 2005; 23: 7265-7277.



Supplementary Figure 1. Levels of ILF2 mRNA expression in breast cancer tissues are positively correlated with those of CCNA2 (A), CCNB1 (B), CDK1 (C), AURKA (D), AURKB (E) and PLK1 (F), based on TCGA breast dataset obtained from Oncomine.