

Y-box binding protein YBX1 and its correlated genes as biomarkers for poor outcomes in patients with breast cancer

Tomohiro Shibata¹, Eriko Tokunaga², Satoshi Hattori³, Kosuke Watari¹, Yuichi Murakami^{1,4}, Nami Yamashita⁵, Eiji Oki⁵, Junji Itou⁶, Masakazu Toi⁶, Yoshihiko Maehara⁷, Michihiko Kuwano⁴ and Mayumi Ono¹

¹Department of Pharmaceutical Oncology, Graduate School of Pharmaceutical Sciences, Kyushu University, Fukuoka, Japan

²National Hospital Organization Kyushu Cancer Center, Fukuoka, Japan

³Department of Integrated Medicine, Biomedical Statistics, Osaka University Graduate School of Medicine, Osaka, Japan

⁴Cancer Translational Research Center, St. Mary's Institute of Health Sciences, Kurume, Japan

⁵Department of Surgery and Science, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan

⁶Department of Breast Surgery, Graduate School of Medicine, Kyoto University, Kyoto, Japan

⁷Kyushu Central Hospital of the Mutual Aid Association of Public School Teachers, Fukuoka, Japan

Correspondence to: Mayumi Ono, **email:** mono@phar.kyushu-u.ac.jp

Keywords: YBX1; ER; breast cancer; predictive biomarker; antiestrogen resistance

Received: November 08, 2018

Accepted: December 05, 2018

Published: December 14, 2018

Copyright: Shibata et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

The enhanced expression of the Y-box binding protein YBX1 is consistently correlated with poor outcomes or reduced survival of breast cancer patients. However, the mechanism underlying the association between increased YBX1 expression and poor outcomes has yet to be revealed. We searched a database for the top 500 genes that are positively or negatively correlated with YBX1 and with ESR1 in breast cancer patients. We further examined the association between YBX1-correlated genes and breast cancer outcomes in patients at Kyushu University Hospital. More than 60% of genes that are positively correlated with YBX1 are also negatively correlated with ESR1. The enhanced expression levels of the top 20 positively correlated genes mostly predict negative outcomes, while the enhanced expression levels of the top 20 negatively correlated genes mostly predict positive outcomes. Furthermore, in breast cancer patients at Kyushu University Hospital, the expression levels of YBX1 and YBX1-positively correlated genes were significantly higher and the expression levels of genes negatively correlated with YBX1 were significantly lower in patients who relapsed after their primary surgery than in those who did not relapse. The expression of YBX1 together with the expression of its positively or negatively correlated genes may help to predict outcomes as well as resistance to endocrine therapies in breast cancer patients. Determining the expression of YBX1 and its closely correlated genes will contribute to the development of precision therapeutics for breast cancer.

INTRODUCTION

Growth in approximately 70% of breast cancers depends on estrogen receptor (ER) expression and its related signal transduction pathway, while growth in approximately 20% of breast cancers depends on the HER2/ERBB2 status [1, 2]. Endocrine therapies have provided significant benefits for patients with ER-positive breast cancers [3, 4], and HER2-targeted

therapeutics have similarly provided benefits for patients with HER2-positive breast cancers [2, 5]. However, the generation of refractory tumors that are resistant to these therapies is a serious obstacle to the further improvement of treatments for breast cancer patients [6, 7]. Identifying other genes or molecules involved in the progression of breast cancer is thus expected to contribute to the more precise prediction of the outcomes of and treatments for refractory tumors [8].

Y-box binding protein (YBX1), a DNA/RNA binding protein containing an evolutionarily conserved cold-shock domain, regulates transcription, translation, DNA damage repair, and other biological processes that occur in both the nucleus and cytoplasm [9–11]. Cytoplasmic YBX1 regulates mRNA stability and translation [12], and nuclear YBX1 plays a key role in transcriptional regulation via the Y-box binding site (inverted CCAAT box) [13, 14]. YBX1 plays a major role in the host's defense mechanisms against environmental cytotoxic stimuli as well as in the growth, survival, and drug resistance of tumor cells [10, 11, 15].

Because YBX1 was identified as an oncoprotein in breast cancer via transgenic knock-in animal models of various types of breast cancer [16], it is therefore expected that YBX1 may play a specific and essential role in the tumorigenesis and malignant progression of breast cancer. YBX1 converts human mammary epithelial cells into breast cancer cells capable of anchorage-independent growth [17]. It also promotes cell growth and increases the expression of HER2 as well as other genes that mediate the cell cycle, cell proliferation, and drug resistance in breast cancer cells [18–23]. YBX1 decreases the response to tamoxifen and fulvestrant in ER-positive breast cancer through the downregulation of the ER protein [24], and antiestrogen resistance is also mediated through YBX1 activation by FGFR2 [25]. We previously demonstrated that the resistance of breast cancer to antiestrogens is mediated through both increased proteasomal ER degradation and increased transcriptional activation of HER2 by YBX1 [26] and that YBX1 negatively and positively correlates with ER and HER2 expression, respectively, in clinical specimens from breast cancer patients as well as in breast cancer cells *in vitro* and *in vivo* [21, 26]. However, the mechanism underlying the predictive ability of YBX1 for a poor prognosis is still not well understood.

In the present study, we searched a database for the top 500 genes that are positively and negatively correlated with YBX1 in breast cancer. Most of the genes that were negatively correlated with YBX1 overlapped with genes that were positively associated with ESR1, and most of the genes that were positively correlated with YBX1 overlapped with genes that were negatively associated with ESR1. However, the clinical significance of YBX1 in breast cancer should be discussed in the context of genes facilitating ER-dependent or ER-independent growth and survival.

RESULTS

Genes that are positively or negatively correlated with YBX1 predict the outcomes of breast cancer patients

Previous studies from many laboratories have consistently shown that the enhanced expression of YBX1 protein or mRNA in mammary tumors is

significantly correlated with malignant progression or poor outcomes in patients with breast cancer [21, 23, 27, 28]. To identify the genes that are potentially regulated by YBX1, we evaluated RNA sequencing (RNA-seq) data from a cohort of 825 invasive breast cancer patients [29] obtained from The Cancer Genome Atlas (TCGA) (<https://cancergenome.nih.gov/>) [30, 31]. We first identified the top 500 genes that are positively or negatively correlated with YBX1 (Supplementary Tables 1–4). The top 20 genes that are significantly ($P < 0.001$) and positively correlated with YBX1 are listed in Figure 1A. Of these 20 genes, the enhanced expression of 8 (UQCRH, CTPS1, FMNL2, CDC20, B3GNT5, MTHFD1L, LRP8, and CDCA8) in tumors was found to be significantly associated with a poor prognosis for the patients (Figure 1A and Supplementary Figure 1A). The expression levels of all 8 of these genes were positively ($P < 0.001$) correlated with YBX1 expression levels (Supplementary Figure 1B). Furthermore, the top 20 genes that were found to be negatively correlated with YBX1 are listed in Figure 1B. ESR1 was found to be significantly and negatively correlated with YBX1, while the ER-coactivator genes FOXA1 and GATA3 were negatively correlated with YBX1 expression (Figure 1B). FOXA1 and GATA3 are involved in tumorigenesis and malignant progression in breast cancer cells in close correlation with the hormonal status [32, 33]. Of the top 20 negatively correlated genes, the enhanced expression of 14 (ESR1, CHCHD5, ATP5G2, GATA3, PCP2, THSD4, LRRC46, PEX11G, RUNDC1, ARSG, SIRT3, TBC1D9, ABAT, and EVL) in tumors was associated with a positive prognosis in patients (Figure 1B and Supplementary Figure 2A). The expression levels of all 14 of these genes were inversely ($P < 0.001$) correlated with YBX1 expression (Supplementary Figure 2B).

Inverse association between genes correlated with YBX1 and those correlated with ESR1

The top 20 genes negatively correlated with YBX1 included many genes that are correlated with ESR1 and its effector genes, such as GATA3 and FOXA1 (Figure 1B). We next examined whether the genes correlated with YBX1 are associated with the genes correlated with ESR1. A Venn diagram revealed some overlap between the top 500 genes significantly correlated with YBX1 and those significantly correlated with ESR1 (Supplementary Tables 1–4). However, the top 500 genes positively correlated with ESR1 did not overlap with the top 500 genes positively correlated with YBX1 (Figure 1C and Supplementary Table 1). In contrast, 352 of the top 500 genes, including GATA3, FOXA1, SIRT3, MYB, and HDAC11, are common between the genes positively correlated with ESR1 and the genes negatively correlated with YBX1 (Figure 1D and Supplementary Table 2). Furthermore, 322 of the top

500 genes, including CDK6, UQCRH, and CTPS1, are common between the genes negatively correlated with ESR1 and the genes positively correlated with YBX1 (Figure 1E and Supplementary Table 3). However, there was no overlap between the top 500 genes that were negatively correlated with ESR1 and the top 500 genes that were negatively correlated with YBX1 (Figure 1F and Supplementary Table 4).

Genes positively or negatively correlated with YBX1 predict the overall survival of patients

We further assessed whether genes positively or negatively correlated with YBX1 could predict the outcomes of breast cancer patients ($n = 63$) at Kyushu University Hospital. Patients with high expression levels of YBX1 in tumors had worse outcomes than those with low expression levels (Figure 2A). The expression levels of 8 genes selected from among the genes positively correlated with YBX1 were mostly positively correlated with YBX1 mRNA expression levels (Figure 2B). Patients with higher expression levels of UQCRH, MTHFD1L, and CTPS1 showed significantly unfavorable outcomes compared to those with low expression levels. In contrast, patients with higher expression levels of CDC20, B3GNT5, LRP8, CDCA8, and FMNL2 showed no significant difference in overall survival compared to those with lower expression levels (Figure 2B). Furthermore, the expression levels of ESR1, GATA3, and SIRT3, which are genes negatively correlated with YBX1, showed an inverse association with the expression level of YBX1; patients with high expression levels of ESR1 and GATA3 in tumors had better outcomes than those with low expression levels (Figure 2C).

We next examined the association between YBX1 and the genes correlated with YBX1 in 13 breast cancer cell lines established from HR+/HER2-, HR-/HER2+, and HR-/HER2- tumors (Figure 3A). With the exception of UQCRH and LRP8, the expression levels of genes positively correlated with YBX1, such as CTPS1, FMNL2, CDC20, B3GNT5, MTHFD1L, and CDCA8, were significantly and positively correlated with the expression level of YBX1 in 13 breast cancer cell lines (Figure 3B). Furthermore, as shown in Figure 3C, the expression levels of ESR1 and GATA3 were inversely correlated with the expression level of YBX1, but there was no significant correlation between the expression levels of SIRT3 and YBX1.

We further examined the expression of 7 genes positively correlated with YBX1 and 2 genes negatively correlated with YBX1 when YBX1 expression was silenced in T-47D (HR+/HER2-) and Hs578T (HR-/HER2-) cells by siRNA (Figure 3D and 3E). The expression levels of 6 of the 7 genes that were positively correlated with the expression of YBX1 were reduced,

and the expression levels of the 2 genes that were negatively correlated with the expression of YBX1 were enhanced in YBX1-silenced T-47D cells (Figure 3D). The expression levels of 3 of the 7 genes that were positively correlated with the expression of YBX1 were also affected, and the expression levels of the two genes that were negatively correlated with the expression level of YBX1 were enhanced in YBX1-silenced Hs578T cells (Figure 3E).

The expression of YBX1 and its correlated genes predicts recurrence susceptibility

YBX1 promotes tumorigenesis and the malignant progression of breast cancer [21, 23, 27, 28]. Consistent with this notion, we explored whether YBX1 and the genes closely correlated with YBX1 were associated with breast cancer recurrence. The expression levels of YBX1 and its correlated genes were initially assessed when the tumors were surgically resected. The expression level of YBX1 in tumors was significantly ($P = 0.033$) higher in patients who relapsed after the initial surgical operation than in those who did not relapse (Figure 4A). Furthermore, the expression levels of 4 (UQCRH, CTPS1, B3GNT5, and CDCA8) of 8 genes positively correlated with YBX1 were significantly ($P < 0.05$) higher in recurrent tumors than in nonrecurrent tumors (Figure 4B). The expression levels of the other 4 genes (FMNL2, CDC20, MTHFD1L, and LRP8) were moderately higher in recurrent tumors than in nonrecurrent tumors. The expression levels of 2 of the 3 genes negatively correlated with YBX1 (ESR1 and GATA3) were significantly ($P < 0.05$) lower in patients who relapsed after their primary surgery than in those who did not relapse (Figure 4C), suggesting that YBX1 and its correlated genes could be used as predictive genes for recurrence and poor outcomes in breast cancer patients.

Associations of YBX1 and ESR1 with overall survival by a principal component analysis

The first 5 principal components constructed with the Kyushu University dataset (see Figure 2) had an 86% cumulative contribution ratio, so we restricted our attention to the first 5 principal components. Our previous studies demonstrated that HER2 is positively correlated with YBX1 in breast cancer [21, 26]. We thus added HER2 as a biomarker in addition to the 12 novel biomarkers in this analysis. Hereafter, the κ -th principal component is denoted by PRIN κ . The correlation coefficients between the 5 principal components and the 13 biomarkers are presented in Figure 5A. We applied principal component Cox regression, in which the first 5 principal components were included as explanatory variables. The first principal component, PRIN1, was positively associated with most biomarkers, including

A

Gene Symbol	Cytoband	Pearson Score	Spearman Score
● UQCRH	1p34.1	0.67	0.64
● CCDC82	11q21	0.61	0.64
● MSL3	Xp22.3	0.62	0.64
● CTPS1	1p34.1	0.65	0.63
● UTP11L	1p34.3	0.66	0.63
● FMNL2	2q23.3	0.63	0.63
● CDC20	1p34.1	0.65	0.62
● CDCA7	2q31	0.65	0.62
● B3GNT5	3q28	0.69	0.62
● MTHFD1L	6q25.1	0.63	0.62
● PRKY	Yp11.2	0.63	0.61
● UQCRHL	1p36.21	0.65	0.61
● PSAT1	9q21.2	0.65	0.61
● LRP8	1p34	0.61	0.6
● FAM49A	2p24.2	0.58	0.6
● MID1	Xp22	0.58	0.59
● ORC1	1p32	0.62	0.59
● CDCA8	1p34.3	0.61	0.59
● FAM171A1	10p13	0.65	0.59
● CLCN4	Xp22.3	0.58	0.58

B

Gene Symbol	Cytoband	Pearson Score	Spearman Score
○ ESR1	6q25.1	-0.7	-0.66
○ CHCHD5	2q13	-0.65	-0.66
○ GATM	19p13.3	-0.65	-0.64
○ FOXA1	14q21.1	-0.65	-0.64
○ ATP5G2	12q13.13	-0.65	-0.63
○ GATA3	10p15	-0.66	-0.63
○ MLPH	2q37.3	-0.69	-0.63
○ PCP2	19p13.2	-0.64	-0.63
○ THSD4	15q23	-0.67	-0.62
○ LRRC46	17q21.32	-0.62	-0.62
○ PEX11G	19p13.2	-0.61	-0.62
○ IGFALS	16p13.3	-0.6	-0.61
○ RUNDC1	17q21.31	-0.61	-0.61
○ ANXA9	1q21	-0.66	-0.61
○ ARSG	17q24.2	-0.62	-0.61
○ SIRT3	11p15.5	-0.6	-0.61
○ TBC1D9	4q31.21	-0.65	-0.61
○ ABAT	16p13.2	-0.61	-0.6
○ C9ORF116	9q34.3	-0.62	-0.6
○ EVL	14q32.2	-0.62	-0.6

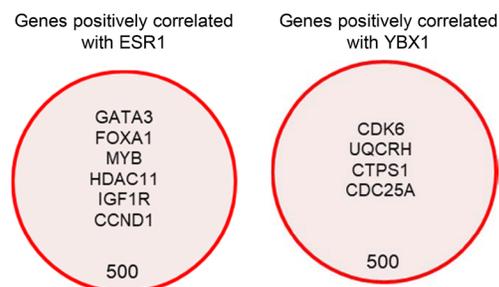
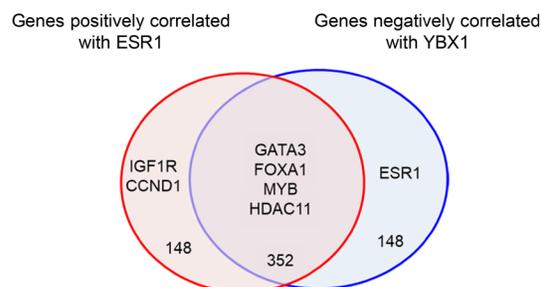
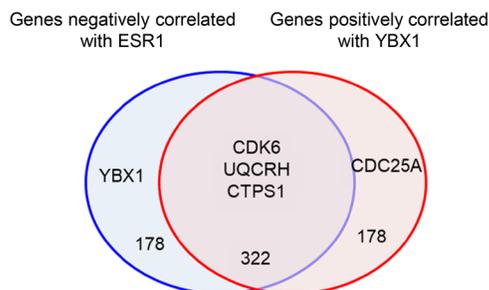
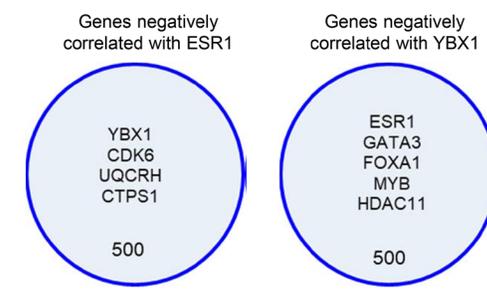
C**D****E****F**

Figure 1: Genes correlated with YBX1 are associated with prognostic factors. (A, B) Lists of the top 20 genes positively (A) or negatively (B) correlated with YBX1. RNA-seq expression data from human invasive breast cancer samples ($n = 825$) were analyzed for their association with YBX1. The genes are listed in order of Spearman's rank positive (A) or negative (B) correlation coefficients for the associations between YBX1 and the other genes. Black dots indicate poorly prognostic genes (see also Supplementary Figure 1). White dots indicate good prognostic genes (see also Supplementary Figure 2). (C) A Venn diagram showing the overlap of the top 500 genes that are positively correlated with ESR1 and the top 500 genes that are positively correlated with YBX1. The top 500 genes that are positively correlated with ESR1 showed no overlap with the top 500 genes that are positively correlated with YBX1. (D) A Venn diagram showing the overlap of the top 500 genes that are positively correlated with ESR1 and the top 500 genes that are negatively correlated with YBX1. Approximately 70% of the top 500 genes that are positively correlated with ESR1 overlap with the top 500 genes that are negatively correlated with YBX1. (E) A Venn diagram showing the overlap of the top 500 genes that are negatively correlated with ESR1 and the top 500 genes that are positively correlated with YBX1. Approximately 64% of the top 500 genes that are negatively correlated with ESR1 overlap with the top 500 genes that are positively correlated with YBX1. (F) A Venn diagram showing the overlap of the top 500 genes that are negatively correlated with ESR1 and the top 500 genes that are negatively correlated with YBX1. The top 500 genes that are negatively correlated with ESR1 showed no overlap with the top 500 genes that are negatively correlated with YBX1.

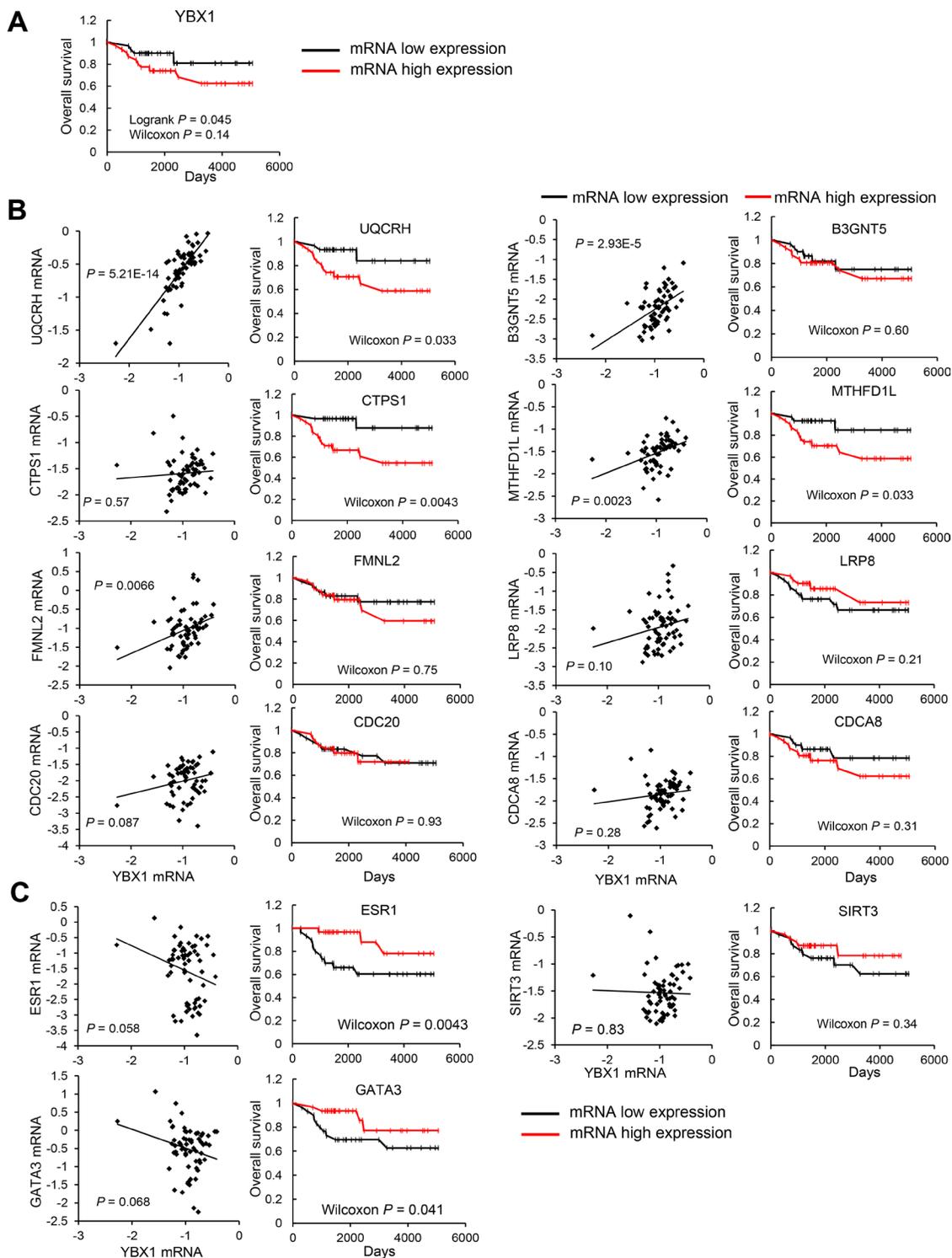


Figure 2: Genes correlated with YBX1 are closely associated with overall survival. (A) The Kaplan-Meier overall survival analysis according to YBX1 mRNA expression in patients with breast cancer. High expression levels of YBX1 are associated with a poor prognosis in breast cancer patients. (B) Left: Correlation data for the YBX1 mRNA expression level versus the mRNA expression levels of genes positively correlated with YBX1. The statistical significance of the correlations was determined using the χ^2 test. The linear regression curve is shown as a black line for significant correlations. Right: The Kaplan-Meier overall survival analysis according to the mRNA expression levels of genes positively correlated with the expression level of YBX1 in patients with breast cancer. (C) Left: Correlation data for the YBX1 mRNA expression level versus the mRNA expression levels of genes negatively correlated with YBX1. The statistical significance of the correlations was determined using the χ^2 test. The linear regression curve is shown as a black line for significant correlations. Right: The Kaplan-Meier overall survival analysis according to the mRNA expression levels of genes negatively correlated with YBX1 in patients with breast cancer.

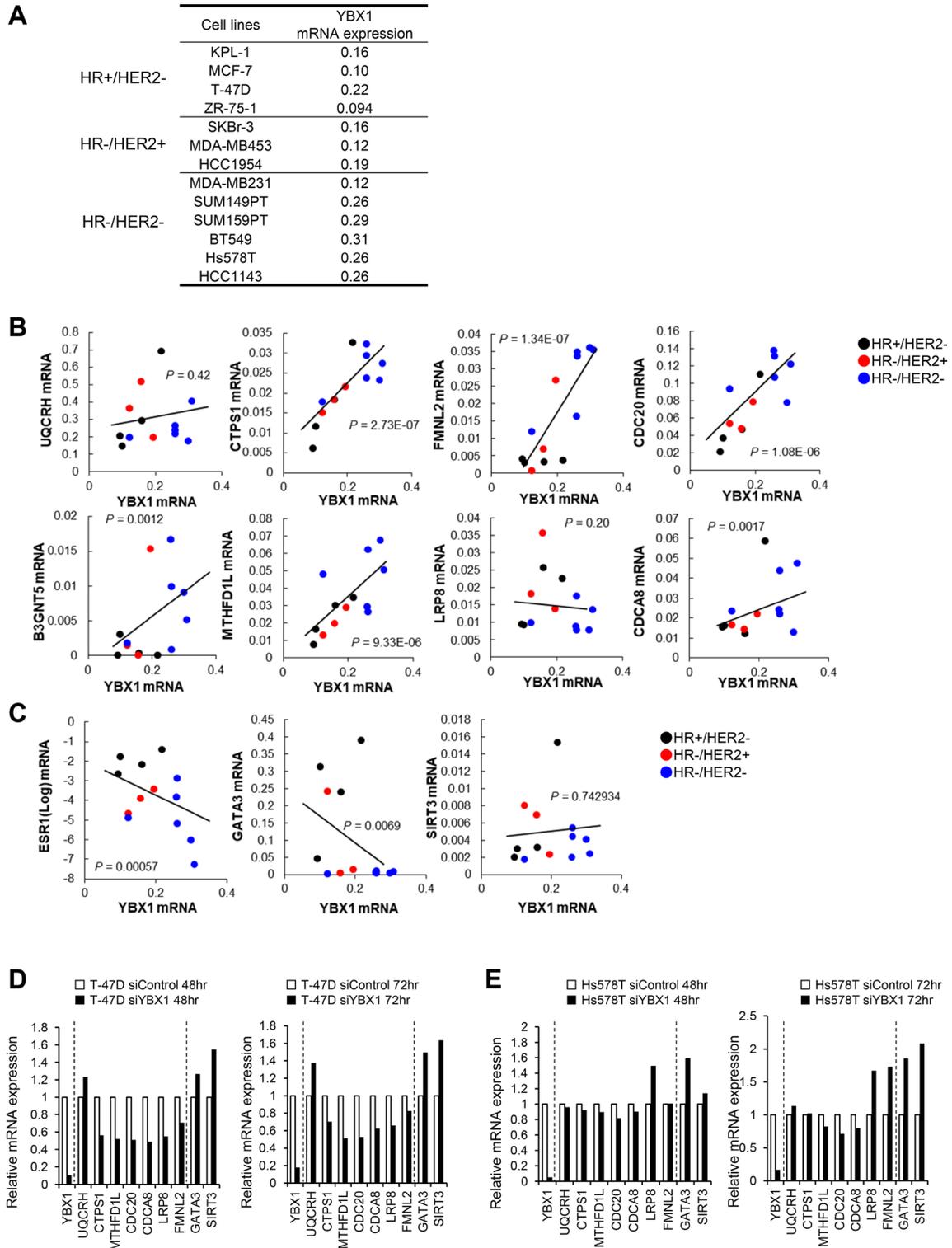


Figure 3: The close association of YBX1 and genes correlated with YBX1 in breast cancer cell lines. (A) List of breast cancer cell lines and YBX1 mRNA levels in three subtypes of breast cancer cell lines. (B) Correlation data for the YBX1 mRNA expression level versus the mRNA expression levels of genes positively correlated with YBX1 in 13 breast cancer cell lines. The statistical significance of the correlations was determined using the χ^2 test. The linear regression curve is shown as a black line for significant correlations. (C) Correlation data for the YBX1 mRNA expression level versus the mRNA expression levels of genes negatively correlated with YBX1 in 13 breast cancer cell lines. The statistical significance of the correlations was determined using the χ^2 test. The linear regression curve is shown as a black line for significant correlations. (D, E) Quantitative RT-PCR showing relative mRNA expression levels of genes positively or negatively correlated with the expression level of YBX1 in T-47D (D) or Hs578T (E) cells treated with YBX1 siRNA (200 nM) for 48 h (left) or 72 h (right).

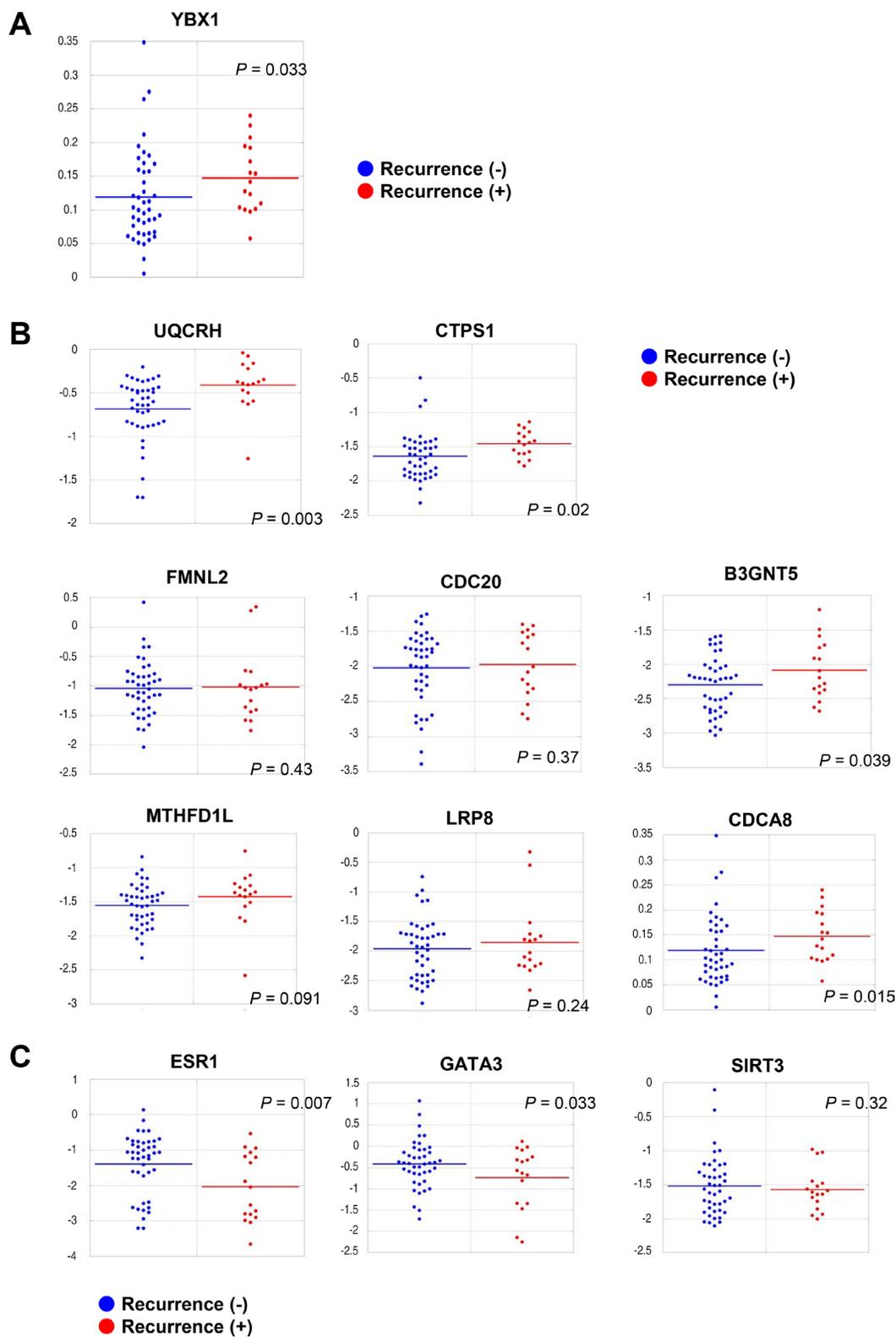


Figure 4: YBX1 is a predictive marker for recurrence and susceptibility to endocrine/chemotherapeutics. (A) Dot plots showing the mRNA expression levels of YBX1 in the primary tumors of patients who relapsed after the initial surgical operation and in the primary tumors of patients who did not relapse. P value, two-sided Student's t -test. (B) Dot plots showing the mRNA expression levels of genes positively correlated with the expression of YBX1 in the primary tumors of patients who relapsed after the initial surgical operation and in the primary tumors of patients who did not relapse. P value, two-sided Student's t -test. (C) Dot plots showing the mRNA expression level of genes negatively correlated with the expression of YBX1 in the primary tumors of patients who relapsed after the initial surgical operation and in the primary tumors of patients who did not relapse. P value, two-sided Student's t -test.

YBX1 and HER2. PRIN1 may represent the general association among the 13 biomarkers. However, PRIN1 was found to be nonessential for determining the prognosis by principal component Cox regression, while PRIN2 and PRIN5 were significantly associated with overall survival ($P = 0.0069$ for PRIN2 and $P = 0.0522$ for PRIN5) (Figure 5B). PRIN2 was positively associated with ESR1, GATA3, and SIRT3 and negatively associated with YBX1 and UQCRH. PRIN5 was positively associated with HER2 and negatively associated with CTPS1. These results suggest that the two mechanisms induced by PRIN2 and PRIN5 play critical roles in the prognosis of breast cancer patients. To demonstrate the idea more simply, we represented PRIN2 by YBX1 or ESR1 and PRIN5 by CTPS1 and examined their influence on overall survival using a simple Kaplan-Meier analysis (Figure 5C and 5D). YBX1 and CTPS1 as well as ESR1 and CTPS1 were found to be independently associated with the prognosis, supporting the idea that PRIN2 and PRIN5 play important roles in the prognosis of breast cancer patients.

DISCUSSION

The nuclear expression level of YBX1 and an increased total YBX1 expression level are predictive markers for poor prognosis in patients with breast cancer [19, 21, 23, 27, 28], ovarian cancer [34], prostate cancer [35], and other human malignancies [36]. Therefore, clarifying the precise mechanisms underlying the significant effect of YBX1 on prognosis will aid in the development of therapies for such cancers. In the present study, enhanced YBX1 expression was mostly correlated with enhanced expression of genes involved in cell growth, cell cycle, survival and DNA synthesis, as well as decreased expression of ESR1 and various ESR1 effector genes in breast cancer patients (Figures 1 and 2) and breast cancer cell lines (Figure 3). The increased or decreased expression levels of these genes correlated with the expression of YBX1 may thus limit the malignant progression of breast cancer (Figure 5E).

ER-targeted antiestrogens and other endocrine therapeutic drugs have greatly contributed to the improvement of breast cancer therapies [3, 4]. However, one serious problem is the emergence of tumors resistant to antiestrogen drugs during the treatment of ER-positive breast cancer patients. Our previous studies demonstrated a significant and reciprocal correlation between YBX1 and ER in breast cancers [21, 26], and showed that the enhanced expression of YBX1 markedly downregulates the expression of ESR1 and induces acquired resistance to antiestrogen therapeutics [26]. In the present study, more than 60% of the genes that were positively correlated with YBX1 were also negatively correlated with ESR1 (Figure 1E). Therefore, the enhanced expression of YBX1 is often reciprocally

accompanied by the reduced expression of ER-dependent genes, which likely promotes breast cancer progression by driving ER-independent cell growth and survival (Figure 5E). Furthermore, we found that patients with higher expression of YBX1 showed poor outcome in ER-positive patients treated with anti-estrogen therapy (Supplementary Figure 3).

Oncotype DX and MammaPrint are frequently used in clinical practice for deciding the suitability of adjuvant chemotherapy for patients with ER-positive and node-negative breast cancer [37, 38]. Oncotype DX and MammaPrint involve reverse transcription polymerase chain reactions for the expression of 21 and 70 genes, respectively, using RNA derived from breast tumor tissues. ER-positive early-stage breast cancer patients can be classified with statistical significance into low- and high-risk groups by these assays. However, in those assays, more than 20 genes are analyzed, and the outcomes of ER-negative breast cancer cannot be predicted. In the present study, we demonstrated that patients with tumors expressing higher YBX1 mRNA levels experienced recurrence more often than those with lower mRNA levels (Figure 4A). Genes that were positively correlated with YBX1 showed a higher expression in recurrent tumors than in nonrecurrent tumors (Figure 4B). In contrast, two genes that were negatively correlated with YBX1 showed lower expression levels in recurrent tumors than in nonrecurrent tumors (Figure 4C). Furthermore, the enhanced expression levels of YBX1 and its correlated gene CTPS1 were able to predict poor outcomes of breast cancer (Figure 5C). Determining the expression levels of YBX1 and its correlated genes can therefore aid in the identification of patients with a high risk of recurrence, and YBX1 and its correlated genes may be useful biomarkers for the precise prediction of breast cancer patients with a high risk of recurrence.

Our and other research groups previously reported that phosphorylation of YBX1 by various kinases including AKT, S6K and RSK through receptor tyrosine kinase and integrin linked kinase induces nuclear translocation of YBX1 in close context with transcriptional activation of various genes including drug resistance and tumor growth related genes [39–42]. Development of therapeutic drugs by targeting YBX1 activation process could contribute to overcome breast cancer through inhibiting expression of predictive genes in present study.

Our present study indicates that the expression of YBX1 and its correlated genes could be used to predict not only poor outcomes but also resistance to endocrine therapeutics and chemotherapeutics in patients with breast cancer. Further follow-up studies in a larger number of patients would be required to confirm our present findings. Targeting YBX1 is expected to help further improve the utility of precision medicine for breast cancer.

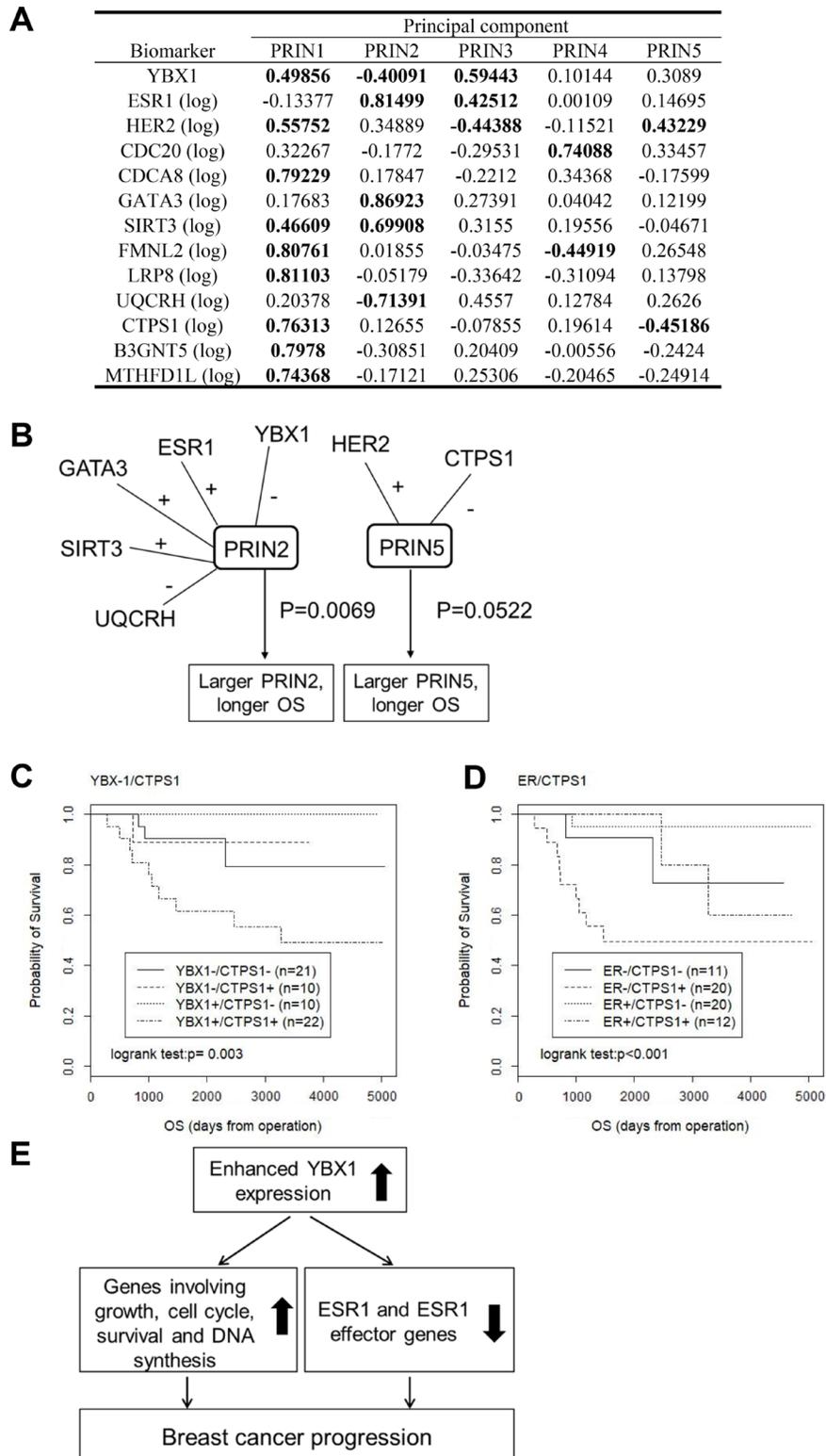


Figure 5: YBX1 and its correlated genes are predictive markers for a poor prognosis. (A) Correlation coefficient between the principal components and biomarkers. (B) Association between the principal components and the overall survival; biomarkers with correlation coefficients $> |0.4|$ with principal components were linked with the sign attached. (C) Kaplan-Meier plots for the four groups defined by the expression of YBX1 and CTPS1, where positivity (+) is defined as a biomarker value higher than its median. (D) Kaplan-Meier plots for the four groups defined by the expression of ESR1 and CTPS1, where positivity (+) is defined as a biomarker value higher than its median. (E) Our hypothetical model showing that YBX1 promotes the malignant progression of breast cancer in close correlation with the enhanced expression of its effector genes involved in cell growth, cell cycle, survival and DNA synthesis, as well as the decreased expression of ESR1 and various ESR1 effector genes.

MATERIALS AND METHODS

Cell culture

The human breast cancer cell lines MCF-7, T-47D, SKBr-3, MDA-MB231, MDA-MB453, ZR-75-1, HCC1954, BT549, Hs578T, and HCC1143 were purchased from the American Type Culture Collection (Manassas, VA, USA). SUM159PT and SUM149PT cells were purchased from Asterand (Detroit, MI, USA). KPL-1 was purchased from Health Science Research Resources Bank (Osaka, Japan). KPL-1, MCF-7, T-47D, SKBr-3, MDA-MB231, MDA-MB453, and Hs578T cells were cultured at 37°C in DMEM supplemented with 10% fetal bovine serum (FBS) in a humidified atmosphere containing 5% CO₂. SUM159PT and SUM149PT cells were maintained with Ham's F-12 nutrient mixture containing 5% FBS, 5 µg/mL insulin, 1 µg/mL hydrocortisone, and 10 mmol/L HEPES. ZR-75-1, HCC1954, and HCC1143 cells were cultured with RPMI containing 10% FBS. BT549 cells were cultured with RPMI containing 10% FBS and 0.8 µg/mL insulin. All cell lines were passaged for ≤ 6 months and were not further tested or authenticated by the authors.

Transfection of small interfering RNA

The siRNA corresponding to the nucleotide sequence of YBX1 (siYBX1; 5'-GGUCCCCACCUUACUACAU-3') was purchased from QIAGEN Inc. (Valencia, CA, USA). Cells were transfected with siRNA duplexes using Lipofectamine RNAiMAX and Opti-MEM (Invitrogen) according to the manufacturer's recommendations.

Patient information

Breast cancer tissue specimens were obtained from 64 Japanese patients who underwent surgery without neoadjuvant systemic therapy at the Department of Surgery and Science, Kyushu University Hospital, between 2004 and 2014. The study was approved by the institutional review board of the university (30–40). Immediately after surgery, the specimens for extraction of the total RNA were placed in liquid nitrogen and stored at –80°C. The clinical data were obtained from the patients' medical records. Endocrine therapy consisted of an aromatase inhibitor, either tamoxifen or toremifene for postmenopausal women, and an LH-RH agonist for premenopausal women. Chemotherapy regimens, including epirubicin and cyclophosphamide, 5-fluorouracil (5-FU), cyclophosphamide, methotrexate and 5-FU and taxanes, were administered based on the clinicopathological findings. The median follow-up was 2678 days. An overview of the clinical information regarding the patients' age, sex, pathological diagnosis of the primary tumor, status of recurrence, stage, adjuvant therapy, and immunohistochemistry scores of ER, PGR,

and HER2 are presented in Supplementary Table 5. To avoid confusion with the TCGA dataset, we have referred to these data as the Kyusyu University dataset in this paper.

Quantitative reverse transcription polymerase chain reaction (qRT-PCR)

Total RNA was isolated from human breast tumor tissue using ISOGEN (Nippon Gene Co., Ltd., Tokyo, Japan) according to the manufacturer's instructions. The RNA concentration was assessed by spectrophotometry at 260 nm. qRT-PCR was performed using the Real-Time PCR system 7300 (Applied Biosystems, Foster City, CA, USA). In these analyses, the average expression levels of genes correlated with YBX1 were calculated, and the patients were then classified into high- and low-mRNA expression groups using the average as a cut-off point. Survival curves were plotted using the Kaplan-Meier method and compared using the log-rank test or Wilcoxon's test. Survival data were evaluated using a multivariate Cox proportional hazards model. Genes were evaluated as having either a poor prognostic ability ($P < 0.05$) or a good prognostic ability ($P < 0.05$).

Statistical analyses

In this manuscript, we used two clinical datasets. The first dataset was from a cohort of 825 invasive breast cancer patients obtained from TCGA (the cBioPortal for Cancer Genomics; <http://www.cbioportal.org>). Using Spearman's rank correlation coefficient, we confirmed the negative association between ESR1 and YBX1. In addition, by ranking genes in order of Spearman's rank correlation coefficients for the associations between YBX1 and those genes (Figure 1A and 1B), the importance of the association between ESR1 and YBX1 was investigated. The associations between the expression levels of YBX1 and its correlated genes were assessed using χ^2 tests. The second dataset was the Kyushu University dataset, the details of which are described in the patient information section above. Using these data, we examined the associations among the mRNA levels of 13 biomarkers and their influence on overall survival, which was defined as the duration from surgery to death due to any cause and could be right-censored at the date of last follow-up. To this end, we employed a principal component Cox regression analysis. To avoid biases due to influential observations, we assessed the skewness of the mRNA levels of each biomarker, and the mRNA levels of biomarkers other than YBX1 were log-transformed. Regarding the principal components as underlying independent mechanisms behind the 13 biomarkers (possibly pathways), we tried to speculate about the underlying mechanism in breast cancer patients. Using principal component Cox regression, the associations between the underlying mechanisms and prognosis were investigated. To simplify interpretation,

we focused on the principal components attaining an 80% cumulative contribution ratio, which was interpreted as 80% variation of the 13 biomarkers being explained by these principal components.

Author contributions

TS, ET, KW, YMurakami, MK and MO conceived the study design and analytical concept. SH performed statistical analyses and drafted the manuscript. ET, SH and MT contributed to the interpretation of results and the critical revision of the manuscript. TS, KW and JI designed and performed cell culture experiments. ET, NY, EO and YMaehara contributed to the acquisition of patient information and pathology data.

CONFLICTS OF INTEREST

The authors declare that no conflicts of interest exist.

FUNDING

This work is supported by JSPS KAKENHI grants number 17K15508 (T.S.), the Fukuoka Foundation for Sound Health Cancer Research Fund (T.S.), the Life Science Foundation of Japan (M.O.), and St. Mary's Institute of Health Sciences (K.W., M.K., M.O).

REFERENCES

1. Clarke CA, Keegan TH, Yang J, Press DJ, Kurian AW, Patel AH, Lacey JV Jr. Age-specific incidence of breast cancer subtypes: understanding the black-white crossover. *J Natl Cancer Inst.* 2012; 104:1094–101.
2. Kümler I, Tuxen MK, Nielsen DL. A systematic review of dual targeting in HER2-positive breast cancer. *Cancer Treat Rev.* 2014; 40:259–70.
3. Davies C, Pan H, Godwin J, Gray R, Arriagada R, Raina V, Abraham M, Medeiros Alencar VH, Badran A, Bonfill X, Bradbury J, Clarke M, Collins R, et al. Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of estrogen receptor-positive breast cancer: ATLAS, a randomised trial. *Lancet.* 2013; 381:805–16.
4. Graham J, Pitz M, Gordon V, Grenier D, Amir E, Niraula S. Clinical predictors of benefit from fulvestrant in advanced breast cancer: A Meta-analysis of randomized controlled trials. *Cancer Treat Rev.* 2016; 45:1–6.
5. Smith I, Procter M, Gelber RD, Guillaume S, Feyereislova A, Dowsett M, Goldhirsch A, Untch M, Mariani G, Baselga J, Kaufmann M, Cameron D, Bell R, et al. 2-year follow-up of trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer: a randomised controlled trial. *Lancet.* 2007; 369:29–36.
6. Ma CX, Reinert T, Chmielewska I, Ellis MJ. Mechanisms of aromatase inhibitor resistance. *Nat Rev Cancer.* 2015; 15:261–75.
7. Jeselsohn R, Buchwalter G, De Angelis C, Brown M, Schiff R. ESR1 mutations—a mechanism for acquired endocrine resistance in breast cancer. *Nat Rev Clin Oncol.* 2015; 12:573–83.
8. Sparano JA. TAILORx: trial assigning individualized options for treatment (Rx). *Clin Breast Cancer.* 2006; 7:347–50.
9. Matsumoto K, Wolffe AP. Gene regulation by Y-box proteins: coupling control of transcription and translation. *Trends Cell Biol.* 1998; 8:318–33.
10. Kohno K, Izumi H, Uchiumi T, Ashizuka M, Kuwano M. The pleiotropic functions of the Y-box-binding protein, YB-1. *Bioessays.* 2003; 25:691–8.
11. Kuwano M, Oda Y, Izumi H, Yang SJ, Uchiumi T, Iwamoto Y, Toi M, Fujii T, Yamana H, Kinoshita H, Kamura T, Tsuneyoshi M, Yasumoto K, et al. The role of nuclear Y-box binding protein 1 as a global marker in drug resistance. *Mol Cancer Ther.* 2004; 3:1485–92.
12. Fukuda T, Ashizuka M, Nakamura T, Shibahara K, Maeda K, Izumi H, Kohno K, Kuwano M, Uchiumi T. Characterization of the 5'-untranslated region of YB-1 mRNA and autoregulation of translation by YB-1 protein. *Nucleic Acids Res.* 2004; 32:611–22.
13. Ladomery M, Sommerville J. A role for Y-box proteins in cell proliferation. *Bioessays.* 1995; 17:9–11.
14. Ohga T, Koike K, Ono M, Makino Y, Itagaki Y, Tanimoto M, Kuwano M, Kohno K. Role of the human Y box-binding protein YB-1 in cellular sensitivity to the DNA-damaging agents cisplatin, mitomycin C, and ultraviolet light. *Cancer Res.* 1996; 56:4224–8.
15. Kuwano M, Uchiumi T, Hayakawa H, Ono M, Wada M, Izumi H, Kohno K. The basic and clinical implications of ABC transporters, Y-box-binding protein-1 (YB-1) and angiogenesis-related factors in human malignancies. *Cancer Sci.* 2003; 94:9–14.
16. Bergmann S, Royer-Pokora B, Fietze E, Jürchott K, Hildebrandt B, Trost D, Leenders F, Claude JC, Theuring F, Bargou R, Dietel M, Royer HD. YB-1 provokes breast cancer through the induction of chromosomal instability that emerges from mitotic failure and centrosome amplification. *Cancer Res.* 2005; 65:4078–87.
17. Davies AH, Reipas KM, Pambid MR, Berns R, Stratford AL, Fotovati A, Firmino N, Astanehe A, Hu K, Maxwell C, Mills GB, Dunn SE. YB-1 transforms human mammary epithelial cells through chromatin remodeling leading to the development of basal-like breast cancer. *Stem Cells.* 2014; 32:1437–50.
18. Bargou RC, Jürchott K, Wagener C, Bergmann S, Metzner S, Bommert K, Mapara MY, Winzer KJ, Dietel M, Dörken B, Royer HD. Nuclear localization and increased levels of transcription factor YB-1 in primary human breast cancers are associated with intrinsic MDR1 gene expression. *Nat Med.* 1997; 3:447–50.

19. Janz M, Harbeck N, Dettmar P, Berger U, Schmidt A, Jürchott K, Schmitt M, Royer HD. Y-box factor YB-1 predicts drug resistance and patient outcome in breast cancer independent of clinically relevant tumor biologic factors HER2, uPA and PAI-1. *Int J Cancer*. 2002; 97:278–82.
20. Wu J, Lee C, Yokom D, Jiang H, Cheang MC, Yorida E, Turbin D, Berquin IM, Mertens PR, Iftner T, Gilks CB, Dunn SE. Disruption of the Y-box binding protein-1 results in suppression of the epidermal growth factor receptor and HER-2. *Cancer Res*. 2006; 66:4872–9.
21. Fujii T, Kawahara A, Basaki Y, Hattori S, Nakashima K, Nakano K, Shirouzu K, Kohno K, Yanagawa T, Yamana H, Nishio K, Ono M, Kuwano M, et al. Expression of HER2 and estrogen receptor α depends upon nuclear localization of Y-box binding protein-1 in human breast cancers. *Cancer Res*. 2008; 68:1504–12.
22. Basaki Y, Taguchi K, Izumi H, Murakami Y, Kubo T, Hosoi F, Watari K, Nakano K, Kawaguchi H, Ohno S, Kohno K, Ono M, Kuwano M. Y-box binding protein-1 (YB-1) promotes cell cycle progression through CDC6-dependent pathway in human cancer cells. *Eur J Cancer*. 2010; 46:954–65.
23. Wu K, Chen K, Wang C, Jiao X, Wang L, Zhou J, Wang J, Li Z, Addya S, Sorensen PH, Lisanti MP, Quong A, Ertel A, et al. Cell fate factor DACH1 represses YB-1-mediated oncogenic transcription and translation. *Cancer Res*. 2014; 74:829–39.
24. Ito T, Kamijo S, Izumi H, Kohno K, Amano J, Ito K. Alteration of Y-box binding protein-1 expression modifies the response to endocrine therapy in estrogen receptor-positive breast cancer. *Breast Cancer Res Treat*. 2012; 133:145–59.
25. Campbell TM, Castro MAA, de Oliveira KG, Ponder BAJ, Meyer KB. ER α Binding by transcription factors NFIB and YBX1 enables FGFR2 signaling to modulate estrogen responsiveness in breast cancer. *Cancer Res*. 2018; 78:410–21.
26. Shibata T, Watari K, Izumi H, Kawahara A, Hattori S, Fukumitsu C, Murakami Y, Takahashi R, Toh U, Ito KI, Ohdo S, Tanaka M, Kage M, et al. Breast cancer resistance to antiestrogens is enhanced by increased ER degradation and ERBB2 expression. *Cancer Res*. 2017; 77:545–56.
27. Habibi G, Leung S, Law JH, Gelmon K, Masoudi H, Turbin D, Pollak M, Nielsen TO, Huntsman D, Dunn SE. Redefining prognostic factors for breast cancer: YB-1 is a stronger predictor of relapse and disease-specific survival than estrogen receptor or HER-2 across all tumor subtypes. *Breast Cancer Res*. 2008; 10:R86.
28. Lasham A, Samuel W, Cao H, Patel R, Mehta R, Stern JL, Reid G, Woolley AG, Miller LD, Black MA, Shelling AN, Print CG, Braithwaite AW. YB-1, the E2F pathway, and regulation of tumor cell growth. *J Natl Cancer Inst*. 2012; 104:133–46.
29. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature*. 2012; 490:61–70.
30. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov*. 2012; 2:401–4.
31. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E, Cerami E, Sander C, Schultz N. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal*. 2013; 6:p11.
32. Green KA, Carroll JS. Oestrogen-receptor-mediated transcription and the influence of co-factors and chromatin state. *Nat Rev Cancer*. 2007; 7:713–22.
33. Hisamatsu Y, Tokunaga E, Yamashita N, Akiyoshi S, Okada S, Nakashima Y, Taketani K, Aishima S, Oda Y, Morita M, Maehara Y. Impact of GATA-3 and FOXA1 expression in patients with hormone receptor-positive/HER2-negative breast cancer. *Breast Cancer*. 2015; 22:520–8.
34. Kamura T, Yahata H, Amada S, Ogawa S, Sonoda T, Kobayashi H, Mitsumoto M, Kohno K, Kuwano M, Nakano H. Is nuclear expression of Y box-binding protein-1 a new prognostic factor in ovarian serous adenocarcinoma? *Cancer*. 1999; 85:2450–4.
35. Giménez-Bonafé P, Fedoruk MN, Whitmore TG, Akbari M, Ralph JL, Ettinger S, Gleave ME, Nelson CC. YB-1 is upregulated during prostate cancer tumor progression and increases P-glycoprotein activity. *Prostate*. 2004; 59:337–49.
36. Kosnopfel C, Sinnberg T, Schitteck B. Y-box binding protein 1--a prognostic marker and target in tumour therapy. *Eur J Cell Biol*. 2014; 93:61–70.
37. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T, Hiller W, Fisher ER, Wickerham DL, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med*. 2004; 351:2817–26.
38. Cardoso F, van't Veer LJ, Bogaerts J, Slaets L, Viale G, Delaloge S, Pierga JY, Brain E, Causeret S, DeLorenzi M, Glas AM, Golfopoulos V, Goulioti T, et al. 70-Gene Signature as an Aid to Treatment Decisions in Early-Stage Breast Cancer. *N Engl J Med*. 2016; 375:717–29.
39. Sutherland BW, Kucab J, Wu J, Lee C, Cheang MC, Yorida E, Turbin D, Dedhar S, Nelson C, Pollak M, Leighton Grimes H, Miller K, Badve S, et al. Akt phosphorylates the Y-box binding protein 1 at Ser102 located in the cold shock domain and affects the anchorage-independent growth of breast cancer cells. *Oncogene*. 2005; 24:4281–92.
40. Evdokimova V, Ruzanov P, Anglesio MS, Sorokin AV, Ovchinnikov LP, Buckley J, Triche TJ, Sonenberg N, Sorensen PH. Akt-mediated YB-1 phosphorylation activates translation of silent mRNA species. *Mol Cell Biol*. 2006; 26:277–92.
41. Basaki Y, Hosoi F, Oda Y, Fotovati A, Maruyama Y, Oie S, Ono M, Izumi H, Kohno K, Sakai K, Shimoyama T,

Nishio K, Kuwano M. Akt-dependent nuclear localization of Y-box-binding protein 1 in acquisition of malignant characteristics by human ovarian cancer cells. *Oncogene*. 2007; 26:2736–46.

42. Stratford AL, Fry CJ, Desilets C, Davies AH, Cho YY, Li Y, Dong Z, Berquin IM, Roux PP, Dunn SE. Y-box

binding protein-1 serine 102 is a downstream target of p90 ribosomal S6 kinase in basal-like breast cancer cells. *Breast Cancer Res*. 2008; 10:R99.