SpliceArray Profiling of Breast Cancer Reveals a Novel Variant of *NCOR2/SMRT* That Is Associated with Tamoxifen Resistance and Control of *ER*α Transcriptional Activity

Luduo Zhang[1], Chun Gong[1], Samantha L.Y. Lau[1], Nan Yang[1], Oscar G.W. Wong[1], Annie N.Y. Cheung[1], Janice W.H. Tsang[2], Kelvin Y.K. Chan[1,3], and Ui-Soon Khoo[1]

Abstract

Gene expression profiling aimed at classifying and prognosing breast cancer has yielded signatures with little, if any, concordance. However, expression arrays used in these studies do not discriminate alternate RNA splice isoforms that vary widely in cancer and may resolve this problem. In this study, we profiled splice isoforms in a panel of tamoxifen-sensitive and -resistant cell lines, defining a novel variant (BQ323636.1) of the nuclear receptor corepressor 2 (*NCOR2*) that was associated with tamoxifen resistance. Overexpression of this variant in a tamoxifen-sensitive cell line induced its resistance to tamoxifen. We confirmed our initial findings from cell lines in 77 breast tumors from a Chinese cohort, where BQ323636.1 expression was higher in tamoxifen-resistant patients than tamoxifen-sensitive patients. For patients who were estrogen receptor (ER)-positive and had received tamoxifen treatment, higher BQ323636.1 expression level correlated with distant metastasis. High expression level of BQ323636.1 was found to be associated with poorer overall and disease-free survival for patients who had received tamoxifen treatment. Notably, higher BQ323636.1 versus NCOR2 wild-type ratio was also associated with negative ER and progesterone receptor (PR) status, and triple-negative status (ER−/PR−/HER2− receptor status). Mechanistic investigations showed that under conditions of tamoxifen exposure, BQ323636.1 suppressed the transcriptional activity of *ER*α, exhibiting promoter-regulating functions. Our findings highlight a novel splice variant of the ERα corepressor NCOR2 as a candidate biomarker in breast cancer that not only predicts tamoxifen response but may be targeted to overcome tamoxifen resistance. *Cancer Res; 73(1); 246–55. ©2012 AACR.*

Introduction

Breast cancer is the most prevalent female cancer worldwide. About 70% of breast cancers express estrogen receptor (ER), a member of the nuclear hormone family of intracellular receptors. Estrogen binds to the ligand-binding domain of the ER and induces conformational change in protein structure for dimerization that is favorable for subsequent interaction with coactivator molecules (1). The activated ER then transactivates estrogen-responsive genes and induces subsequent cell response.

Tamoxifen, a selective ER modulator, is commonly used as first-line adjuvant treatment to prevent cancer recurrence for ER-positive (ER+) breast cancers. It reduces the annual breast cancer mortality by one-third and disease recurrence by almost half (2). In breast tissues, tamoxifen acts as an antagonist of ER, competes with estrogen for binding to ER, blocks the transcription activation activity of ER, and arrests cancer cells in the cell cycle. However, almost half of the patients with prolonged tamoxifen treatment will fail to respond to tamoxifen over time and develop disease recurrence eventually, that is, acquired tamoxifen resistance (3) for reasons as yet unclear. Postulated mechanisms are loss of ER expression, aberrant changes in coregulators, dysregulation of receptor tyrosine kinase signaling or cell survival signaling, and apoptosis (4).

Numerous gene expression profiling studies (5–10) have been done by different research groups with the aim of improving prognostication and classification of breast cancers, including tamoxifen responsiveness, to assist the selection of patients for the most appropriate adjuvant systemic therapies. However, each of the studies provided gene lists showing very little concordance in gene identity.

Alternative splicing is a key posttranscriptional mechanism for generating multiple protein products from a single gene, occurring in perhaps 40% to 60% of human genes. Alternatively spliced isoforms of a given protein can display

Authors' Affiliations: Departments of [1]Pathology and [2]Clinical Oncology, Li Ka Shing Faculty of Medicine, The University of Hong Kong; and [3]Department of Obstetrics and Gynecology, Tsan Yuk Hospital, Hong Kong, China

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

L. Zhang and C. Gong contributed equally to this work and are joint first authors.

Corresponding Author: Ui-Soon Khoo, Rm 323, 3/F, University Pathology Building, Queen Mary Hospital, Pokfulam Road, Hong Kong SAR, China. Phone: 852-2255-4410; Fax: 852-2218-5205; E-mail: uskhoo@pathology.hku.hk

American Association for *Cancer Research*

different and even antagonistic biologic functions. Microarray chips detect a small fragment of each gene transcript and unless specifically designed, are unlikely to discriminate the expression of all possible isoforms for most of the genes studied. Over or reduced expression of some gene isoforms that have different functions from their wild-type might account for differences in gene identity between the different gene expression profile models. In recent years, various studies have revealed a number of alternatively spliced variants that are associated with cancers (11, 12). Splicing microarray profiling in prostate cancer found that 30% splice variants of 222 prostate cancer–related genes had significant changes in expression between tumors and normal samples (13). A number of alternative spliced variants that distinguished tumors from normal tissues have also been identified in ovarian and breast cancers (14, 15).

We postulated that alternatively spliced variants may play a role in tamoxifen resistance in breast cancer and conducted custom designed splice variant profiling on a panel of breast cancer cell lines with differential response to tamoxifen treatment. A novel splice variant of the gene *NCOR2* was identified that was differentially expressed between parental tamoxifen-sensitive and -derived resistant breast cancer cell lines. The variant was further validated to significantly correlate with tamoxifen resistance and poor survival in breast cancers. Subsequent functional studies were carried out to elucidate the underlying role of this splice variant in tamoxifen resistance.

Materials and Methods

Cell culture

Human breast cancer cell lines MCF7 and ZR-75-1 were purchased from American Type Culture Collection and were authenticated by short tandem repeat profiling. MCF7 was cultured in Dulbecco's Modified Eagle's Medium supplemented with 10% FBS and 10 ng/mL human recombinant insulin and ZR-75-1 was cultured in Improved Minimum Essential Medium (IMEM) supplemented with 5% FBS. LCC2 and AK-47 were kindly provided by Dr. Robert Clarke (Georgetown University Medical School, Washington, D.C.; ref. 16). LCC2, a tamoxifen-resistant cell line derived from MCF7/LCC1 by stepwise *in vitro* selection of prolonged tamoxifen treatment was cultured in IMEM supplemented with 5% charcoal-stripped FBS and AK47, derived from ZR75-1, was cultured in IMEM supplemented with 5% charcoal-stripped FBS.

Human breast cancer samples

Primary tumors were collected from 77 patients diagnosed with breast cancer between the years 1993 and 2001 at Queen Mary Hospital (Hong Kong, China). Five- to 10-year clinical follow-up data and clinicopathologic information were available. The protocol was approved by the Institutional Review Board (UW 06-379 T/1404). Histologic sections of all cases were reviewed by a pathologist. Sections containing at least 70% of invasive carcinoma were used for total RNA extraction.

Total RNA extraction

TRIzol reagent (Invitrogen) was used for total RNA extraction. Breast cancer frozen tissues were cut in 20-μm thick cryostat sections for total RNA extraction.

SpliceArray profiling

SpliceArray profiling was conducted by ExonHit Therapeutics, Inc. Custom designed 4 × 44K microarray slides from Agilent Technologies were used. The probe sets were designed to detect the wild-type forms as recommended in National Center for Biotechnology Information RefSeq and all possible alternatively spliced forms. Each alternative splice event can be modeled with a short and a long form. The reference sequence of this gene can be either the short form or long form, depending on the alternative splice event type, and the alternatively spliced variant will be the other form.

A panel of breast cancer cell lines with differential response to tamoxifen was used in the SpliceArray profiling, as shown in Supplementary Table S1A. The cell lines were treated at tamoxifen concentrations 0 μmol/L, 0.1 μmol/L, and 1 μmol/L for 5 days before RNA extraction. Every sample was cultured and sampled in triplicate. All RNA samples went through quality control before cDNA synthesis and Cy3 dye labeling.

SpliceArray profiling analysis was conducted using Partek Genomics Suite. Analysis was conducted on each cell line cDNA sample individually. The probe intensity data were Log_2 transformed then quartile normalized to make the distribution of probe set intensities for each array similar within the set of arrays. Low or nonexpressed probe sets were filtered out if all samples' intensity values fell below the predetermined Log_2-based value.

Statistical analysis

For data obtained from the array profiling, a two-way ANOVA model was used to conduct statistical tests on the filtered probe set level intensities between each cell line sample group to generate P values and fold change values. "Top hit lists" were generated with $P \leq 0.05$ and fold change ≥ 1.8 as the cutoff, representing significant differences between the samples being compared. Clustering of the profiling data was conducted using Cluster 3.0 (17) and visualized using Java TreeView (18).

In vivo correlation analysis was based on the clinicopathologic characteristics of the 77 Chinese breast cancer samples collected from Queen Mary Hospital as described above. The expression levels of splice variant BQ323636.1 were analyzed against the tamoxifen responsiveness of the patients with breast cancer and other clinicopathologic features by independent sample t test. Survival analyses were done by Kaplan–Meier estimates and Cox regression model. P values of less than 0.05 were considered statistically significant.

Real-time quantitative PCR

Up to 5 μg of total RNAs were reverse transcribed into cDNA by SuperScript III reverse transcriptase (Invitrogen). Real-time PCR reaction was carried out with the ABI 7900HT Fast Real-time PCR system. For TaqMan real-time quantitative PCR

(qPCR), a standard curve was set up for quantification. The reaction volume was 15 μL in 96-well plate and 10 μL in 384-well plate. The reaction consisted of $1\times$ TaqMan Universal PCR Master Mix, No AmpErase UNG (Applied Biosystems), 0.1 μmol/L 6-carboxyfluorescein-labeled TaqMan MGB probe, 0.4 μmol/L forward and reverse primers, and 2 μL cDNA. The 2-step thermal cycling method consisted of $95°C$ for 10 minutes, followed by 40 cycles of $95°C$ for 15 seconds, and $60°C$ for 60 seconds. The primer and TaqMan probe sequences used were presented in Supplementary Table S2.

Western blot analysis

Cells were lysed and SDS-PAGE was conducted as previously described (19). Harvested cells were lysed in protein lysis buffer (Cell Signaling) containing $1\times$ Complete Protease Inhibitor Cocktail Tablet and 1 mmol/L phenylmethyl-sulfonylfluoride. Twenty micrograms of each protein sample were separated by SDS-PAGE using 8% gel. The proteins then were transferred to polyvinylidene fluoride membrane (Millipore) and probed with monoclonal antibody against NCOR2 (A01, Abnova), ERα (HC-20, Santa Cruz Biotech), 6xHis-tag (9F2, Wako), and β-actin (AC-74, Sigma). Secondary antibodies were: horseradish peroxidase–conjugated anti-rabbit antibody (P0448, Dako) and horseradish peroxidase-conjugated affiniPure anti-mouse antibody (Jackson ImmnunoResearch Laboratories).

MTT assay

Cells were seeded at a density of 1.5×10^4 in 96-well plates. On the day of conducting MTT assay, cells were incubated with culture medium containing 10% MTT (USB, Affymetrix) stock solution (5 mg/mL) at $37°C$ for 4 hours. After incubation, the medium was discarded and replaced with 100 μL isopropanol with 4 mmol/L HCl and 0.1% NP-40 to dissolve the purple precipitates. Extinction of the samples was measured by the Infinite 200 microtiter plate reader (Tecan) at 570 nm with a reference wavelength at 750 nm. Background MTT readings were subtracted from the samples.

Plasmids

pcDNA3.1-BQ323636.1 was constructed by inserting the human variant BQ323636.1 cDNA into pcDNA3.1(+) (Invitrogen). pcDNA3.1-6xHis-BQ323636.1 was constructed by Dr. Tori Chan, The University of Hong Kong, Hong Kong. pcDNA3.1-ERα was constructed by inserting full-length human ERα cDNA into pcDNA3.1(+). The luciferase reporter ERE-E1b-Luc was a gift from Dr. Carolyn Smith (Baylor College of Medicine, Houston, TX; ref. 20) and ERE-TK-Luc was a gift from Dr. Hung-Ying Kao (Case Western Reserve University, Cleveland, OH; ref. 21).

Transient transfection and luciferase assay

The ERE firefly reporter and pRL-SV40 were cotransfected with pcDNA3.1-ERα and pcDNA3.1-BQ323636.1 or its empty vector into HeLa cells with Lipofectamine 2000 transfection reagent (Invitrogen). Cells grown to 90% confluence were grown in Opti-MEM I-reduced serum medium without phenol red before transfection. The cells were transfected with desired amount of DNA or Lipofectamine 2000 reagent following the manufacturer's instruction. Six hours after transfection, the medium was changed to complete medium and incubated at $37°C$ with 5% CO_2.

After ligand treatment for 24 hours, the activities of firefly and *renilla* luciferases were measured sequentially using dual-luciferase reporter assay system (Promega) according to manufacturer's protocol. The intensities of firefly and *renilla* luminescence were recorded by Infinite F200 (Tecan). Background readings were subtracted from the samples.

Results

Eleven splice variants that were differentially expressed between tamoxifen sensitive and resistant breast cancer cell lines were selected for validation

To identify the alternatively spliced variants that correlate with tamoxifen resistance status in breast cancer, custom designed SpliceArray profiling was conducted. Breast cancer–related genes involved in cell signaling, cell proliferation, cell cycle, apoptosis and invasion, 417 in total, were chosen for profiling study. These 417 genes had a total of 4,088 evidence events investigated that were represented by 21,249 evidence probes. The evidence event refers to all the known and predicted alternative splice events identified in publicly available sequence databases of the gene.

A panel of breast cancer cell lines with differential response to tamoxifen was used in the SpliceArray profiling (Supplementary Table S1A). AK 47 (ER-negative) can still respond to tamoxifen treatment; however, it is more resistant to tamoxifen compared with its parental cell line ZR-75-1 (ER+) as its IC_{50} was much higher (Supplementary Fig. S1). Cell line RNA samples were prepared in independent triplicates. The expression level of each alternatively spliced variant is expressed as the probe intensities of long form versus short form ratio. A large number of splice variants that had expression levels correlating with ER status or tamoxifen responsiveness status in the panel of breast cancer cell lines screened were observed (data not shown). To identify the significant and meaningful splice variants that are truly involved in tamoxifen resistance, splice variants that fulfilled at least one of the criteria below were selected:

(i) At basal level, they were differentially expressed (fold change >1.8, $P < 0.05$) between tamoxifen-sensitive (TamS) and tamoxifen-resistant (TamR) cell lines in both 2 pairs, that is, MCF7 versus LCC2; ZR75-1 versus AK47;

(ii) Under tamoxifen treatment, they had opposite trend of change (fold change >1.8, $P < 0.05$) between TamS and TamR cell lines in both pairs;

(iii) At basal level, they were differentially expressed (fold change >1.8, $P < 0.05$) between ER+ and ER− cell lines.

The significant fold change cutoff was set at 1.8, as this is the minimum difference that may be detected by TaqMan real-time quantitative PCR. Using these criteria, 11 splice variants of
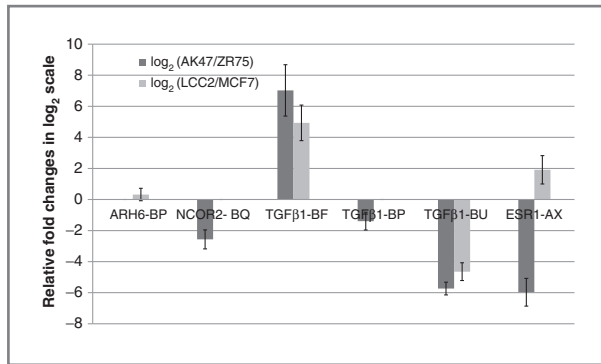
Figure 1. Fold changes of the validated splice variants' expression levels by qPCR. Fold changes of the validated splice variants in MCF7/LCC2 (gray bars) and AK47/ZR75-1 (black bars) cell line pairs were presented in log$_2$ scale. Negative fold changes represent upregulation of the variant, whereas positive fold changes represent downregulation of the variant. Data are expressed as mean ± SD of 3 independent experiments.

5 genes were selected (Supplementary Fig. S2) and subjected to validation by qPCR. The validation results were summarized in Supplementary Table S1B and the fold changes of their long versus short form ratio within the cell line pairs are illustrated in Fig. 1. At basal level (without tamoxifen treatment), 4 splice variants were validated in MCF7/LCC2 cell line pair and 5 splice variants were validated in ZR75-1/AK47 cell line pair (Fig. 1).

One of the splice variants of TGF-β1, BF756452.1, with a novel alternative splice acceptor site of 41 bps at the 5′ end upstream of exon 10 and a novel stop codon that leads to early termination of the protein before exon 10 was downregulated the most in both cell line pairs. The other splice variant BU149200.1, with a novel intron of 139 bps within the last exon, was the most upregulated variant in both cell line pairs. The splice variants of TGF-β1 validated from our study are being further investigated, but are outside the context of this article.

Splice variant of ERα, AX066402.1, with a deletion of the first 151 amino acids at the N-terminal of the protein due to exon 1 skipping had the relative large fold change in AK47/ZR75-1 cell line pair. However, because both the variant and the ERα wild-type had very low or no expression in ER− cell lines, its functional significance in ER− cells was limited, thus it was not chosen for functional studies.

Besides the 3 variants introduced above, fold change of BQ323636.1, a novel variant of nuclear receptor corepressor 2 (NCOR2), was significantly upregulated in AK47 compared with ZR75-1 (2.5-fold in log$_2$ scale). Exon 11 skipping in this splice variant leads to early termination of the protein product. Considering the known biologic functions of the genes validated in this study, NCOR2 was the most likely to bear functional relationship with tamoxifen resistance. It interacts with ERα when ERα is complexed with tamoxifen and represses the transcriptional activity of ERα. It has been shown to be directly involved in the antagonist activity of tamoxifen on ERα. Therefore, its splice variant, BQ323636.1, may play an important role in tamoxifen resistance and was selected for further analysis.

Differential expression of a novel splice variant of NCOR2 between tamoxifen-sensitive and -resistant breast cancer cell lines

By SpliceArray profiling, we had identified a novel alternatively spliced variant BQ323636.1 that was significantly differentially expressed in both MCF7/LCC2 and ZR75-1/AK47 cell line pairs both at basal level and under tamoxifen treatment (Supplementary Table S1B). BQ323636.1 has exon 11 skipped in mature mRNA, which results in early termination of its protein product, retaining only the N-terminal of the NCOR2 wild-type protein (Fig. 2A). To validate the differential expression of splice variant BQ323636.1 between TamS and TamR cell lines at the mRNA level, real-time TaqMan qPCR was conducted on a new batch of cell line mRNA samples in triplicates. Primers were designed to specifically target either this splice variant or the *NCOR2* wild-type form (NM_006312; Fig. 2B).

In the ZR75-1/AK47 cell line pair, qPCR results expressed as long versus short form (Fig. 2C, i), showed consistent results as in the SpliceArray profiling. To represent the expression level of BQ323636.1 in relation to that of its wild-type BQ323636.1 versus wild-type ratio (BQ/WT ratio), that is, the short versus long form, is shown in Fig. 2C, ii. Consistent with SpliceArray profiling data, the BQ/WT ratio was significantly higher in AK47 cell line (derived TamR) compared with its ZR75-1 cell line (parental TamS) by about 2-fold, both at basal level and under tamoxifen treatment. However, qPCR validation was unable to show significant differential expression of BQ323636.1 variant between the MCF7 and LCC2 cell line pair (data not shown). This may imply that the biologic role of BQ323636.1 in tamoxifen resistance involves loss of ER expression.

In addition, we also examined the expression level of the variant alone (Fig. 2C, iii) or wild-type alone (Fig. 2C, iv). Differential expression of BQ323636.1 was only observed under tamoxifen treatment, but not at basal level. This difference was however not observed for the *NCOR2* wild-type. The differential expression of BQ323636.1 only under tamoxifen treatment for an ER− TamR cell line suggests its potential functional role of this variant in presence of tamoxifen.

Expression of splice variant BQ323636.1 correlated with tamoxifen response status in cell lines

The protein product of the variant BQ323636.1 was predicted to be a truncated protein product of 362 amino acids with molecular mass of 42.6 kDa. Its amino acid sequence corresponds to the N-terminus of NCOR2 wild-type protein, with only the last 10 amino acids before the stop codon translated from exon 12 out of frame (Fig. 2A).

If the differential expression of BQ323636.1 observed between ZR75-1 and AK47 has a functional role in biologic processes, such differential expression is also expected at the protein level. A NCOR2 polyclonal antibody (A01, Abnova) targets the first 1–111 amino acids of wild-type NCOR2 and was used to detect the protein product of BQ323636.1 in Western blot analysis. This antibody can detect the wild-type NCOR2 at size 250 kDa and the variant at size 42.6 kDa, respectively. Between protein size 40 and 50 kDa, a protein band was observed in both ZR75-1 and AK47, identified as the
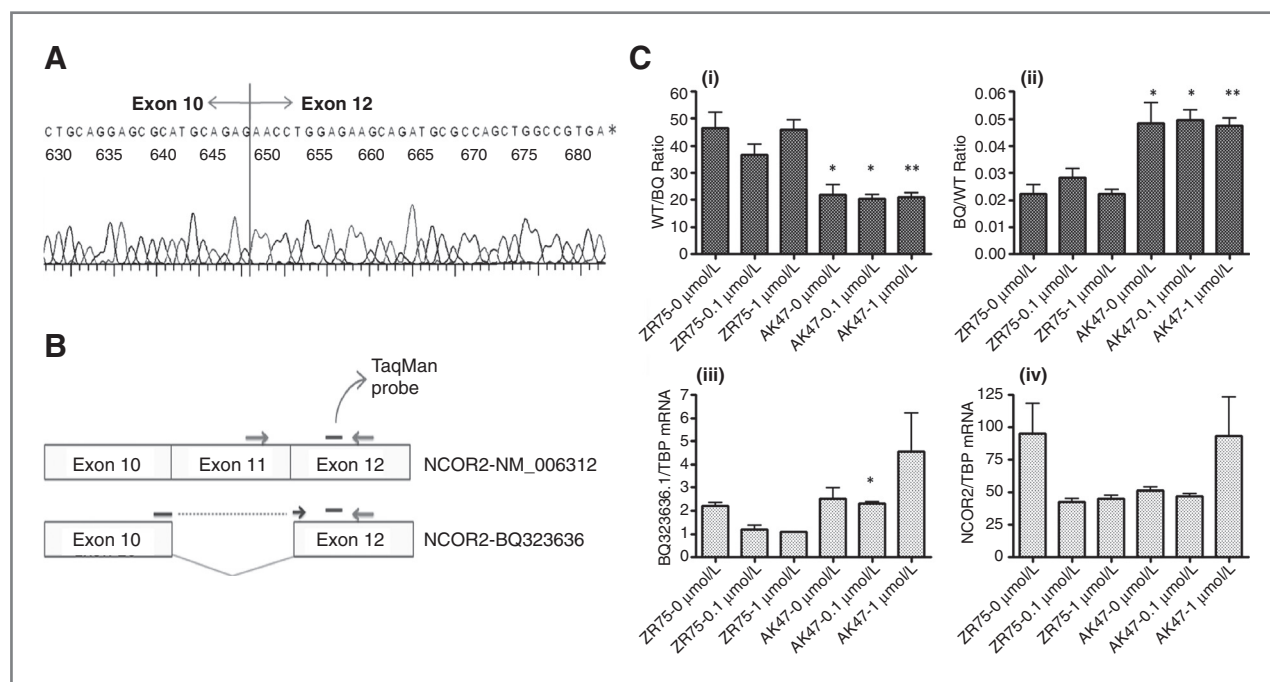
Figure 2. A, schematic sequence of the exon 10 and exon 12 junction on cDNA of BQ323636.1. *, the protein termination position of BQ323636.1 protein product with the stop codon TGA. B, schematic presentation of the primers and TaqMan probe. Schematic representation of the primers (arrows) and TaqMan probe that exclusively detect the *NCOR2* wild-type or BQ323636.1 variant by TaqMan qPCR. The forward primer of the variant of exon 10 and exon 12, targeting 16 bps at the 3′ end of exon 10 and 4 bps at the 5′ end of exon 12. A common reverse primer was used for both the wild-type and the variant. C, qPCR results on RNA prepared from ZR75-1 and AK47 under tamoxifen treatment. TaqMan real-time qPCR results from a new batch of RNA samples prepared from ZR75-1 and AK47 under tamoxifen treatment at specified concentrations for 5 days. i, *NCOR2*-WT versus BQ323636.1 (long vs. short form) ratio; ii, BQ323636.1 versus *NCOR2*-WT ratio (short vs. long form ratio); iii, BQ323636.1 mRNA expression level normalized by TATA box binding protein (TBP); iv, *NCOR2* wild-type mRNA expression level normalized by TBP. Expression level was expressed as the amount of copy numbers quantified by standard curve. *, $P < 0.05$; **, $P = 0.001$ in comparison with ZR75-1 treated with the same tamoxifen concentration. Data are expressed as mean ± SD of 3 independent experiments.

variant BQ323636.1 (Fig. 3A). Size of the variant was confirmed by overexpression of BQ323636.1 in both HeLa and 293T cell line with plasmid pcDNA3.1-His-BQ323636.1 (Supplementary Fig. S3). Consistent with the qPCR results, under tamoxifen treatment, the protein expression level of BQ323636.1 in AK47 was significantly higher than in ZR75-1 compared with that of wild-type. At basal level, such difference was also observed, but not as significant as under tamoxifen treatment.

To investigate whether the upregulation of BQ323636.1 was the cause of tamoxifen resistance, BQ323636.1 was transiently transfected into ER+, tamoxifen-sensitive cell line ZR-75-1, and MTT assays were conducted to assess the proliferation of cells in response to tamoxifen treatment. The decrease in cell proliferation observed under tamoxifen treatment was slower in BQ323636.1-overexpressing cells (Fig. 3B) than that of the vector control over a range of drug concentrations. This difference was statistically significant ($P < 0.05$) under 15 μmol/L tamoxifen treatment (Fig. 3C), indicating that the variant can mediate tamoxifen resistance in ER+ cells as well.

BQ323636.1 expression correlated with tamoxifen resistance, poor prognosis, and poor survival in Chinese patients with breast cancers

To reveal the clinical significance of this variant *in vivo*, we investigated for possible correlation of BQ323636.1 expression with clinicopathologic parameters in 77 breast cancer samples. BQ323636.1 mRNA expression was examined by TaqMan real-time qPCR and analyzed between clinical parameter groups using independent sample *t* test. Tamoxifen resistance was defined as those patients treated with tamoxifen only in the adjuvant setting who subsequently developed relapse or distant metastasis. Consistent with the observation in cell lines, the BQ323636.1 expression level was significantly higher in tamoxifen resistant (TamR) patients than those who were sensitive (TamS) to tamoxifen treatment ($P = 0.006$; Fig. 4A). The BQ/WT ratio showed similar trend, though not statistically significant (Supplementary Table S3).

Consistently, in 37 patients who were ER+ and had received tamoxifen treatment, higher BQ323636.1 expression level was significantly associated with development of distant metastasis ($P = 0.048$; Fig. 4B). However, BQ323636.1 expression level was not significantly correlated with chest wall/breast relapse ($P = 0.453$; Supplementary Fig. S4). This might be due to limited sample size, as among the 37 patients, there were only 27 patients whose relapse statuses were available.

Survival analysis was conducted to investigate the correlation between the variant's expression level and the clinical outcomes of the patients with breast cancer. Among the 77 patients, 37 patients, who had received tamoxifen treatment, were dichotomized by the BQ323636.1 expression level (cutoff
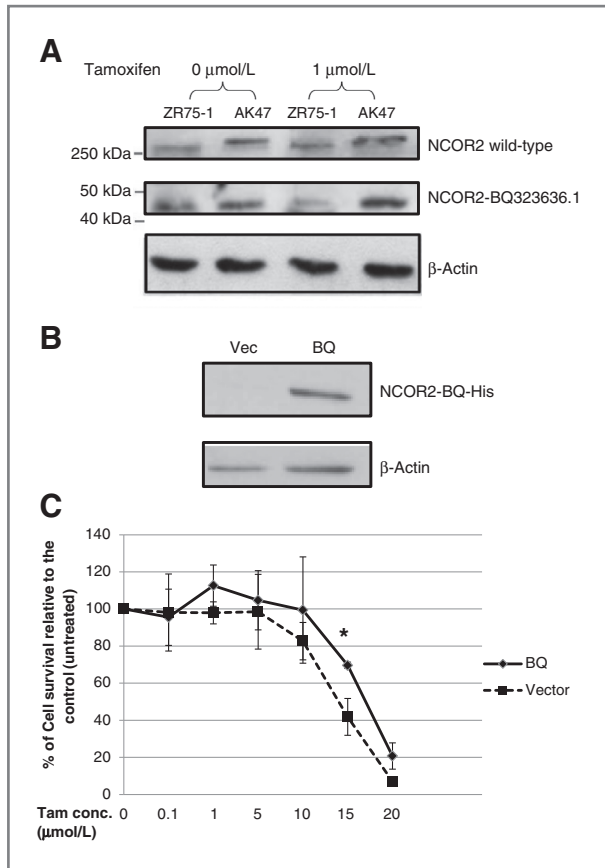
Figure 3. A, protein level of BQ323636.1 in ZR75-1 and AK47 with or without tamoxifen treatment. Protein level of BQ323636.1 in cell lines ZR75-1 and AK47 with vehicle (EtOH) treatment or 1 μmol/L tamoxifen treatment for 5 days. BQ323636.1 is differentially expressed between ZR75-1 and AK47 under 1 μmol/L tamoxifen treatment. The Western blots were repeated 3 times on 3 separate batches of protein lysates. The representative blot is shown here. B and C, overexpression of BQ323636.1 in tamoxifen-sensitive ZR-75-1 mediates its resistance to tamoxifen. ZR-75-1 was transiently transfected with BQ323636.1 or an empty vector and treated with tamoxifen for 24 hours. Cell proliferation was measured by MTT assay. When ZR-75-1 cells were overexpressing BQ323636.1 or a vector control (B), the decrease in cell proliferation observed under tamoxifen treatment was slower over a range of drug concentrations in BQ323636.1-overexpressing cells compared with the control and this difference was statistically significant ($P < 0.05$) when treated with 15 μmol/L tamoxifen (C). Data are expressed as mean ± SD of 3 independent experiments.

value is the median). It was found that higher BQ323636.1 expression level was associated with both poorer overall survival ($P = 0.041$; Fig. 4C) and poorer disease-free survival ($P = 0.008$; Fig. 4D). This association remained significant after being adjusted by tumor size, tumor grade, HER2 and PR status of the patients (Table 1), and analyzed by multivariate Cox regression model (HR = 11.8989; $P = 0.030$ for overall survival; HR = 10.764; $P = 0.038$ for disease-free survival; Table 2). Though there were only 37 patients included, these significant findings would support the relevance of this variant in acquiring resistance among patients who have received tamoxifen treatment.

Moreover, it was found that both BQ323636.1 expression ($P = 0.005$) and the BQ/WT ratio ($P = 0.008$; Supplementary

Fig. 5A and Supplementary Table S3) correlated to the *de novo* ER status of the patients at diagnosis, being significantly higher in ER− patients. Progesterone receptor (PR) is the downstream target gene of ERα. Following the pattern for ER status, the BQ323636.1 expression level was significantly higher in PR− patients ($P = 0.035$; Supplementary Fig. 5B). The BQ/WT ratio also showed a similar trend but not statistically significant. While there was no correlation observed for HER-2 status, the BQ/WT ratio significantly correlated with triple-negative (negative for ER, PR, and HER2 overexpression) cancers ($P = 0.018$; Supplementary Fig. 5C). There was no significant correlation between BQ323636.1 expression level and other clinicopathologic parameters including tumor size, tumor grade, and tumor type (Supplementary Table S3).

BQ323636.1 modulates ERα transcriptional activation activity

As tamoxifen directly targets ERα in breast cancer cells, we next investigated whether BQ323636.1 modulates ERα transcription activation activity when ERα is bound by estrogen (activated ERα) or tamoxifen (inactivated ERα).

Studies elucidating the role of NCOR2 on ERα transcription activity have been conflicting depending on the luciferase reporter used. Using the ERE-TK-Luc reporter, NCOR2 acted as a corepressor for ERα-mediated transcription (21, 22), whereas on the ERE-E1b-Luc reporter, it was shown that NCOR2 was required for full ERα transcriptional activity (20). We therefore used both luciferase reporters in turn to test the modulation of BQ323636.1 on ERα-mediated transcription activation (Fig. 5).

When ERα was unliganded (treated with vehicle EtOH), BQ323636.1 overexpression significantly suppressed the transcription on both ERE-E1b and ERE-TK promoters. On E2 treatment, BQ323636.1 overexpression exhibited opposing effects on luciferase activity between ERE-E1b and ERE-TK promoters. On ERE-E1b promoter, overexpression of BQ323636.1 induced significant decrease in transcription activity (Fig. 5A); whereas on ERE-TK promoter, significant increase in the transcription activity was seen (Fig. 5B). On the basis of previous reports of NCOR2 functions on ERE transcription, that is, as a coactivator on ERE-E1b promoter and a corepressor ERE-TK promoter, BQ323636.1 seems to exert the opposite effect in modulating ERα transcriptional activation activity compared with wild-type NCOR2.

In the presence of tamoxifen, regardless of whether E2 was present, BQ323636.1 consistently exhibited suppression effect on ERE transcription for both EREs. Overexpression of BQ323636.1 decreased luciferase activity by 70% for ERE-E1b promoter and by 50% for ERE-TK promoter. In conclusion, BQ323636.1 was shown to be able to modulate the transcriptional activation activity of ERα, indicating its functional importance in breast cancer.

Discussion

Our SpliceArray profiling identified a novel splice variant BQ323636.1 of *NCOR2* that had significantly higher expression
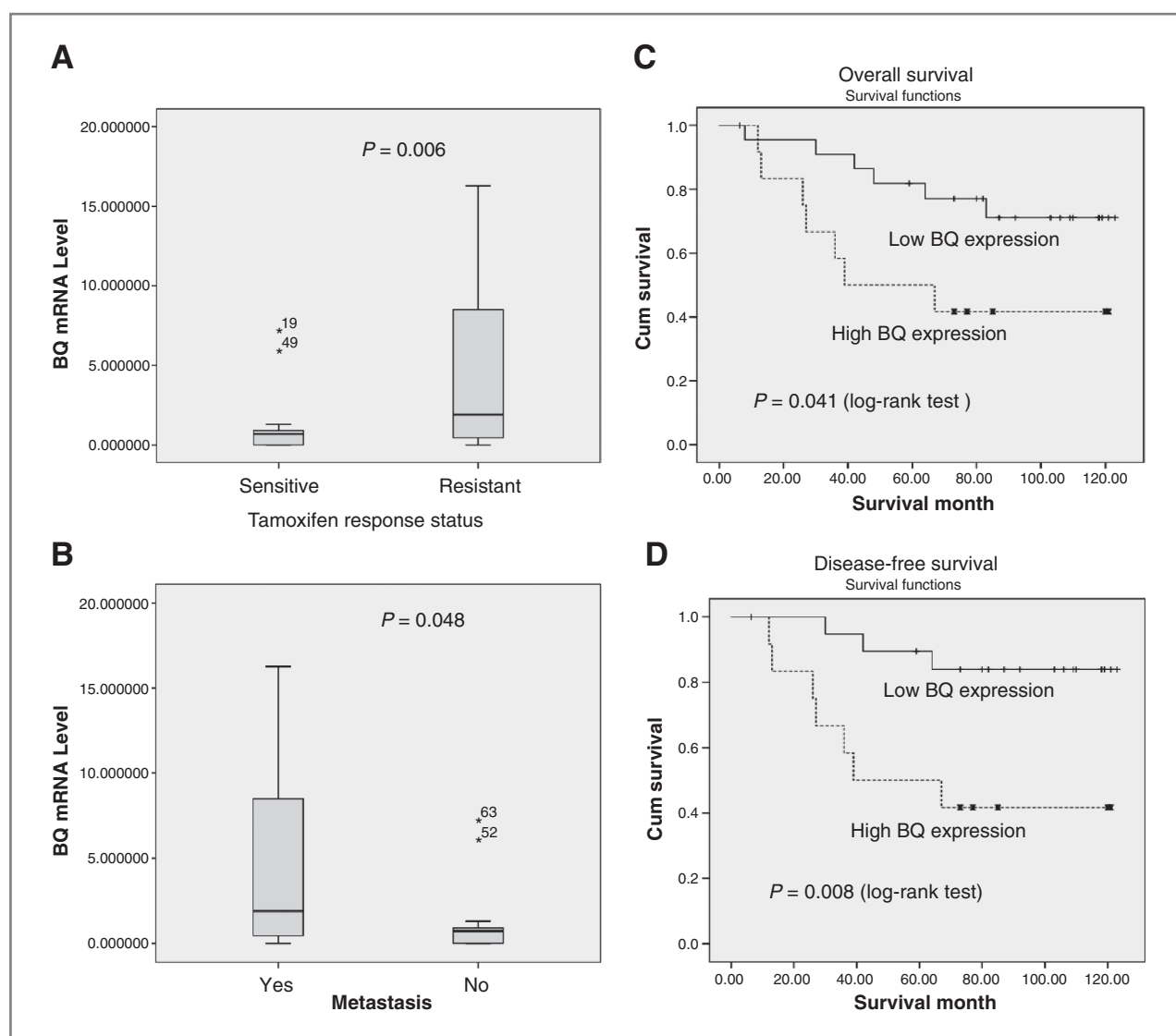
Figure 4. A, BQ mRNA expression level correlates with tamoxifen response status in patients with ER+ breast cancer. The horizontal line represents the median of the group. B, BQ mRNA expression level correlates with metastasis in patients with ER+ breast cancer. The horizontal line represents the median of the group. C, BQ323636.1 expression level significantly correlates with overall survival. Thirty-seven patients, who had received tamoxifen treatment, were dichotomized by the BQ323636.1 expression level (cutoff value is the median). Overexpression of BQ323636.1 was associated with poorer overall survival ($P = 0.041$, log-rank test) among these patients. D, BQ323636.1 expression level significantly correlates with disease-free survival. Overexpression of BQ323636.1 was associated with poorer disease-free survival ($P = 0.008$, log-rank test) in patients who have received tamoxifen treatment.

in derived tamoxifen-resistant cell line AK47 than its parental tamoxifen-sensitive cell line ZR75-1, both at mRNA level and protein level. *In vitro*, overexpression of the variant in tamoxifen-sensitive cell line ZR-75-1 induced its resistance to tamoxifen. *In vivo*, it consistently correlated with tamoxifen resistance in breast cancer samples. Higher expression of BQ323636.1 was associated with tamoxifen resistance and, consequently, poorer overall and disease-free survival. Higher expression of BQ323636.1 was associated with negative ER and PR status and metastasis; higher BQ323636.1 to wild-type ratio was associated with *de novo* negative ER status, and triple-negative status — all features representing poor response to

tamoxifen, poor prognosis, and more aggressive phenotypes in breast cancer.

NCOR2, also commonly known as SMRT (the silencing mediator of retinoic acid and thyroid hormone receptor), was initially discovered as the suppressor for retinoic acid and thyroid hormone receptors (23). As ERα is a critical ligand-activated transcription factor in breast cells controlling multiple important cell signaling pathways, its transcription activation on its target genes is strictly regulated by coactivators and corepressors. There are more than 23 corepressors of ERα identified that bind to different functional domains of ERα and suppress the activity of ERα

through various mechanisms (24). NCOR2, together with its homologous gene NCOR, are the best characterized corepressors for ERα and interact with ERα in the presence of antagonist, that is, tamoxifen (25–27). When tamoxifen acts as an antagonist in the cell context, it binds to ERα and induces ERα to undergo conformational change that is more favorable for the binding of corepressors rather than coactivators (28, 29), so as to suppress the transcriptional activation activity of ERα.

We therefore investigated whether BQ323636.1 also regulated ERα transcription activation activity by luciferase assay. On E2 treatment, BQ323636.1 was observed to exhibit an opposite effect in modulating ERα transcriptional activation activity compared with wild-type NCOR2. For ERE-E1b promoter, on which NCOR2 was reported to act as a coactivator when E2 was present (25, 30), overexpression of BQ323636.1 significantly suppressed ERE-E1b transcription. For ERE-TK promoter, on which NCOR2 was reported to be a corepressor (22), overexpression of BQ323636.1 significantly promoted ERE-TK transcription. However, under tamoxifen treatment, regardless of whether E2 was present, BQ323636.1 consistently suppressed the transcription on both ERE-E1b and ERE-TK promoters. Therefore, when ERα is liganded with E2, the effect of NCOR2 and BQ323636.1 on its transcriptional activation activity is promoter-specific. Regardless of whether NCOR2 wild-type acts as a coactivator or corepressor on the ERE promoter, BQ323636.1 exerts an effect opposite to that of its wild-type. However, once tamoxifen is applied, BQ323636.1

Table 1. Thirty-seven patients with breast cancer who had been treated with tamoxifen were included in survival analyses

Clinical parameter	Status	Sample size n (%)
Patients	Received tamoxifen treatment	37
Tumor size	>3 cm	29 (78)
	≥3 cm	8 (22)
Tumor grade	1 and 2	14 (38)
	3	17 (40)
	Missing	6 (16)
PR status	Positive	2 (5)
	Negative	27 (73)
	Missing	8 (22)
HER2 status	Positive	9 (24)
	Negative	19 (52)
	Missing	9 (24)
Overall survival	Alive	22 (59)
	Died	14 (38)
	Missing	1 (3)
Disease-free survival	Alive without disease	22 (59)
	Died of cancer	11 (30)
	Missing	4 (11)

Table 2. Survival analyses for tamoxifen-treated patients using multivariate Cox regression model

	Sig.	HR (95% CI)
Multivariate Cox regression analysis of overall survival for tamoxifen-treated patients with breast cancer		
BQ Tamtreated[a]	0.030[b]	11.889 (1.275–110.865)
Tumor size	0.377	2.370 (0.350–16.060)
PR status	0.030[b]	0.057 (0.004–0.759)
HER2 status	0.901	1.122 (0.183–6.894)
Tumor grade	0.366	2.030 (0.437–9.425)
Multivariate Cox regression analysis of disease free survival for tamoxifen-treated patients with breast cancer		
BQ Tamtreated[a]	0.038[b]	10.764 (1.143–101.381)
Tumor size	0.490	1.947 (0.294–12.908)
PR status	0.129	0.101 (0.005–1.943)
HER2 status	0.915	1.104 (0.178–6.840)
Tumor grade	0.374	2.285 (0.369–14.151)

NOTE: Variables used in the Cox regression model were dichotomized as shown in Table 1. Top, higher BQ323636.1 expression level significantly associated with higher risk of death after being adjusted by various clinical parameters including tumor size, tumor grade, PR status, and HER2 status (HR = 11.889; $P = 0.030$). Bottom, higher BQ323636.1 expression level significantly associated with higher risk of dying from cancer (disease-free survival) after being adjusted by various clinical parameters including tumor size, tumor grade, PR status, and HER2 status (HR = 10.764; $P = 0.038$).
[a]Cutoff =1.05, which is the median of tamoxifen-treated patients' BQ mRNA level.
[b]$P < 0.05$.

uniformly exhibits suppressive effect on the transcription of both ERE promoters.

Splice variant BQ323636.1 skips exon 11 but retains the N-terminal 362 amino acids (42.6 kDa) of the NCOR2 wild-type protein, corresponding to the first repression domain RD-1 (Supplementary Fig. 6A). NCOR2 is a large protein (250 kDa) with its N-terminus interacting with other transcriptional corepressors to regulate transcription and its C-terminus binding to nuclear receptors, that is, ERα (Supplementary Fig. 6B). At the N-terminus of NCOR2, amino acids 1–316 interact with other transcription corepressors such as GPS2 (G protein pathway suppressor 2; ref. 21) and TBLR1 (transducin β-like 1 X-linked receptor 1; ref. 31; Supplementary Fig. 6B). The 2 SANT motifs, corresponding to amino acids 432–489 and amino acids 611–669, are critical components of a deacetylase activation domain (DAD) that bind and activate HDAC3. The repression domain RD-3 interacts with HDAC 4 and HDAC5, and the repression domain RD-4 interacts with HDAC3, GPS2, and TBLR1. As the SANT motifs important for NCOR2-HDAC3 interaction are missing in BQ323636.1 protein product, the modulation of BQ323636.1 on ERα transcriptional activity may not be through histone deacetylases (HDAC) but through the
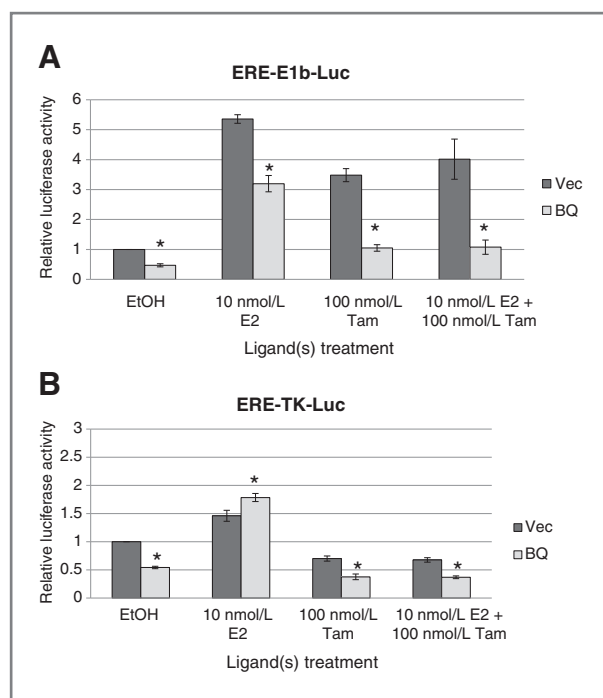
A

ERE-E1b-Luc

B

ERE-TK-Luc

Figure 5. BQ323636.1 suppresses ERα transcriptional activity in a promoter-specific manner. HeLa cells were transfected with 50 ng pcDNA3.1-ERα; 20 ng pRL-SV40; either 1 µg ERE-E1b-Luc reporter (A) or 500 ng ERE-TK-Luc reporter (B); and either pcDNA3.1-BQ323636.1 (BQ) or pcDNA3.1 empty vector (Vec) to balance the total amount of DNA transfected. Cells were treated with 0.1% vehicle (EtOH), 10 nmol/L E2, 100 nmol/L tamoxifen (Tam) or both for 24 hours before luciferase measurements. Data represent the mean ± SD of 4 independent experiments. *, $P < 0.05$ compared with the empty vector control group.

interaction with other corepressors, possibly GPS2 or TBLR1. The variant may compete with NCOR2 wild-type for binding to these protein partners and thus inhibit the function of its wild-type to facilitate the development of tamoxifen resistance (Supplementary Fig. S7). More comprehensive functional studies will be needed to elucidate the role of this variant in tamoxifen resistance.

If BQ323636.1 is functionally important, its corresponding amino acid sequence should be highly conserved among different species. Indeed, the protein product of BQ323636.1 (highlighted by the frame in Supplementary Fig. 6C), was the most conserved region in NCOR2 with above 90% similarity across all 7 species. This implies that BQ323636.1 contains the most important functional domain in NCOR2 protein.

Very low mRNA expression level of BQ323636.1 compared with the NCOR2 wild-type was observed in previous qPCR results; however, in Western blot analysis, the protein level of BQ323636.1 was rather high, similar, or even higher than its wild-type form. Therefore, there may be some posttranscriptional mechanisms promoting the translation of BQ323636.1 or protecting it from degradation. A phosphorylation site at Ser2410 on NCOR2 has been reported to be the recognition

motif to export NCOR2 out of the nucleus for proteasome-mediated degradation (32). This phosphorylation site is deleted in BQ323636.1, which may protect it from protein degradation, whereas the NCOR2 wild-type that contains the phosphorylation site is recognized to undergo protein degradation. This may be a possible explanation for the relatively high protein expression of BQ323636.1 observed in breast cancer cell lines.

By revealing the clinical significance and the underlying function of BQ323636.1 in tamoxifen resistance and breast tumor progression, splice variant BQ323636.1 may be developed as a potential prognostic biomarker for tamoxifen treatment response and furthermore as a novel therapeutic target to overcome tamoxifen resistance in clinical application. Tamoxifen resistance is a complex outcome involving several mechanisms and pathways. It is unlikely that any single gene or mechanism confers tamoxifen resistance. In this study, we showed that alternatively spliced variants may also play a critical role in tamoxifen resistance, which is an important perspective that was neglected often in previous tamoxifen resistance studies. By integrating splice variant that is significant in causing tamoxifen resistance with other previous well-characterized targets, a combination set of more effective therapeutic targets may be developed as a novel therapy to overcome this clinical obstacle in breast cancer.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: K.Y.K. Chan, U.-S. Khoo
Development of methodology: L. Zhang, C. Gong, S.L.Y. Lau, K.Y.K. Chan
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): L. Zhang, C. Gong, S.L.Y. Lau, J.W.H. Tsang, U.-S. Khoo
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): L. Zhang, C. Gong, K.Y.K. Chan
Writing, review, and/or revision of the manuscript: L. Zhang, C. Gong, N. Yang, J.W.H. Tsang, K.Y.K. Chan, U.-S. Khoo
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): L. Zhang, A.N.Y. Cheung, J.W.H. Tsang, K.Y.K. Chan
Study supervision: K.Y.K. Chan, U.-S. Khoo
Verified the cell lines used in the study: O.G.W. Wong

References

1. McDonnell DP, Norris JD. Connections and regulation of the human estrogen receptor. Science 2002;296:1642–4.
2. Group EBCTC. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 2005;365:1687–717.
3. Clarke R, Leonessa F, Welch JN, Skaar TC. Cellular and molecular pharmacology of antiestrogen action and resistance. Pharmacol Rev 2001;53:25–71.
4. Musgrove EA, Sutherland RL. Biological determinants of endocrine resistance in breast cancer. Nat Rev Cancer 2009;9:631–43.
5. Chanrion M, Negre V, Fontaine H, Salvetat N, Bibeau F, Mac Grogan G, et al. A gene expression signature that can predict the recurrence of tamoxifen-treated primary breast cancer. Clin Cancer Res 2008;14: 1744–52.
6. Jansen MP, Foekens JA, van Staveren IL, Dirkzwager-Kiel MM, Ritstier K, Look MP, et al. Molecular classification of tamoxifen-resistant breast carcinomas by gene expression profiling. J Clin Oncol 2005;23: 732–40.
7. Loi S, Haibe-Kains B, Desmedt C, Wirapati P, Lallemand F, Tutt AM, et al. Predicting prognosis using molecular profiling in estrogen receptor-positive breast cancer treated with tamoxifen. BMC Genomics 2008;9:239.
8. Ma XJ, Wang Z, Ryan PD, Isakoff SJ, Barmettler A, Fuller A, et al. A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen. Cancer Cell 2004;5:607–16.
9. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med 2004;351:2817–26.
10. Vendrell JA, Robertson KE, Ravel P, Bray SE, Bajard A, Purdie CA, et al. A candidate molecular signature associated with tamoxifen failure in primary breast cancer. Breast Cancer Res 2008;10:R88.
11. Fackenthal JD, Godley LA. Aberrant RNA splicing and its functional consequences in cancer cells. Dis Model Mech 2008;1:37–42.
12. Brinkman BM. Splice variants as cancer biomarkers. Clin Biochem 2004;37:584–94.
13. Zhang C, Li HR, Fan JB, Wang-Rodriguez J, Downs T, Fu XD, et al. Profiling alternatively spliced mRNA isoforms for prostate cancer classification. BMC Bioinformatics 2006;7:202.
14. Venables JP, Klinck R, Bramard A, Inkel L, Dufresne-Martin G, Koh C, et al. Identification of alternative splicing markers for breast cancer. Cancer Res 2008;68:9525–31.
15. Klinck R, Bramard A, Inkel L, Dufresne-Martin G, Gervais-Bird J, Madden R, et al. Multiple alternative splicing markers for ovarian cancer. Cancer Res 2008;68:657–63.
16. Wong LJ, Dai P, Lu JF, Lou MA, Clarke R, Nazarov V. AIB1 gene amplification and the instability of polyQ encoding sequence in breast cancer cell lines. BMC cancer 2006;6:111.
17. Heard J, Kaufmann W, Guan X. A novel method for large tree visualization. Bioinformatics 2009;25:557–8.
18. Saldanha AJ. Java Treeview–extensible visualization of microarray data. Bioinformatics 2004;20:3246–8.
19. Chen J, Gomes AR, Monteiro LJ, Wong SY, Wu LH, Ng TT, et al. Constitutively nuclear FOXO3a localization predicts poor survival and promotes Akt phosphorylation in breast cancer. PLoS ONE 2010;5: e12293.
20. Peterson TJ, Karmakar S, Pace MC, Gao T, Smith CL. The silencing mediator of retinoic acid and thyroid hormone receptor (SMRT) corepressor is required for full estrogen receptor alpha transcriptional activity. Mol Cell Biol 2007;27:5933–48.
21. Cheng X, Kao HY. G protein pathway suppressor 2 (GPS2) is a transcriptional corepressor important for estrogen receptor alpha-mediated transcriptional regulation. J Biol Chem 2009;284:36395–404.
22. Romano A, Adriaens M, Kuenen S, Delvoux B, Dunselman G, Evelo C, et al. Identification of novel ER-alpha target genes in breast cancer cells: gene- and cell-selective co-regulator recruitment at target promoters determines the response to 17beta-estradiol and tamoxifen. Mol Cell Endocrinol 2010;314:90–100.
23. Chen JD, Evans RM. A transcriptional co-repressor that interacts with nuclear hormone receptors. Nature 1995;377:454–7.
24. Dobrzycka KM, Townson SM, Jiang S, Oesterreich S. Estrogen receptor corepressors – a role in human breast cancer? Endocr Relat Cancer 2003;10:517–36.
25. Smith CL, Nawaz Z, O'Malley BW. Coactivator and corepressor regulation of the agonist/antagonist activity of the mixed antiestrogen, 4-hydroxytamoxifen. Mol Endocrinol 1997;11:657–66.
26. Lavinsky RM, Jepsen K, Heinzel T, Torchia J, Mullen TM, Schiff R, et al. Diverse signaling pathways modulate nuclear receptor recruitment of N-CoR and SMRT complexes. Proc Natl Acad Sci U S A 1998;95: 2920–5.
27. Varlakhanova N, Snyder C, Jose S, Hahm JB, Privalsky ML. Estrogen receptors recruit SMRT and N-CoR corepressors through newly recognized contacts between the corepressor N terminus and the receptor DNA binding domain. Mol Cell Biol 2010;30:1434–45.
28. Shang Y, Hu X, DiRenzo J, Lazar MA, Brown M. Cofactor dynamics and sufficiency in estrogen receptor-regulated transcription. Cell 2000; 103:843–52.
29. Liu XF, Bagchi MK. Recruitment of distinct chromatin-modifying complexes by tamoxifen-complexed estrogen receptor at natural target gene promoters in vivo. J Biol Chem 2004;279:15050–8.
30. Kaasa SO, Peterson T, Morris EK, Thompson WC. Statistical inference and forensic evidence: evaluating a bullet lead match. Law Hum Behav 2007;31:433–47.
31. Zhang XM, Chang Q, Zeng L, Gu J, Brown S, Basch RS. TBLR1 regulates the expression of nuclear hormone receptor co-repressors. BMC Cell Biol 2006;7:31.
32. Hoberg JE, Popko AE, Ramsey CS, Mayo MW. IkappaB kinase alpha-mediated derepression of SMRT potentiates acetylation of RelA/p65 by p300. Mol Cell Biol 2006;26:457–71.