

Expression levels of *SF3B3* correlate with prognosis and endocrine resistance in estrogen receptor-positive breast cancer

Yesim Gökmen-Polar¹, Yaseswini Neelamraju², Chirayu P Goswami³, Xiaoping Gu¹, Gouthami Nallamothu¹, Sarath Chandra Janga^{2,3,4} and Sunil Badve^{1,5,6}

¹Department of Pathology and Laboratory Medicine, Indianapolis, IN, USA; ²Department of Biohealth Informatics, School of Informatics and Computing, IUPUI, Indianapolis, IN, USA; ³Center for Computational Biology and Bioinformatics, Indiana University School of Medicine, Indianapolis, IN, USA; ⁴Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA; ⁵Department of Medicine, Indiana University School of Medicine, Indianapolis, IN, USA and ⁶Indiana University Melvin and Bren Simon Cancer Center, Indianapolis, IN, USA

De novo or acquired resistance to endocrine therapy limits its utility in a significant number of estrogen receptor-positive (ER-positive) breast cancers. It is crucial to identify novel targets for therapeutic intervention and improve the success of endocrine therapies. Splicing factor 3b, subunit 1 (*SF3B1*) mutations are described in luminal breast cancer albeit in low frequency. In this study, we evaluated the role of *SF3B1* and *SF3B3*, critical parts of the SF3b splicing complex, in ER-positive endocrine resistance. To ascertain the role of *SF3B1/SF3B3* in endocrine resistance, their expression levels were evaluated in ER-positive/endocrine-resistant cell lines (MCF-7/LCC2 and MCF-7/LCC9) using a real-time quantitative reverse transcription PCR (qRT-PCR). To further determine their clinical relevance, expression analysis was performed in a cohort of 60 paraffin-embedded ER-positive, node-negative breast carcinomas with low, intermediate, and high *Oncotype DX* recurrence scores. Expression levels of *SF3B1* and *SF3B3* and their prognostic value were validated in large cohorts using publicly available gene expression data sets including The Cancer Genome Atlas. *SF3B1* and *SF3B3* levels were significantly increased in ER α -positive cells with acquired tamoxifen (MCF-7/LCC2; both $P < 0.0002$) and fulvestrant/tamoxifen resistance (MCF-7/LCC9; $P = 0.008$ for *SF3B1* and $P = 0.0006$ for *SF3B3*). Expression levels of both MCF-7/LCC2 and MCF-7/LCC9 were not affected by additional treatments with E2 and/or tamoxifen. Furthermore, qRT-PCR analysis confirmed that *SF3B3* expression is significantly upregulated in *Oncotype DX* high-risk groups when compared with low risk ($P = 0.019$). Similarly, in publicly available breast cancer gene expression data sets, overexpression of *SF3B3*, but not *SF3B1*, was significantly correlated with overall survival. Furthermore, the correlation was significant in ER-positive, but not in ER-negative tumors. This is the first study to document the role of *SF3B3* in endocrine resistance and prognosis in ER-positive breast cancer. Potential strategies for therapeutic targeting of the splicing mechanism(s) need to be evaluated.

Modern Pathology (2015) 28, 677–685; doi:10.1038/modpathol.2014.146; published online 28 November 2014

The process of alternative splicing results in synthesis of multiple mRNA variants from a single gene. Earlier studies reported the importance of splice variants in a number of key genes regulating signal-

ing pathways such as apoptosis, metabolism, and angiogenesis. Alterations in expression of some splicing factors, genes that regulate splicing process of mature messenger RNA species from primary transcripts, have also been reported in various cancers including breast cancer.¹ The utilization of next-generation sequencing technologies have further revealed the presence of somatic mutations in splicing factors.² Among them, the *SF3B1* mutations occurred in several types of hematological malignancies as well as in solid cancers, such as breast, pancreatic cancers, and uveal melanomas.^{3–7} These mutations were shown to be

Correspondence: Dr Y Gökmen-Polar, PhD, Department of Pathology and Laboratory Medicine, Indiana University School of Medicine, 635 Barnhill Dr., MS 0038, Indianapolis, IN 46202, USA or S Badve, MD, FRCPath, Department of Pathology and Laboratory Medicine, Indiana University School of Medicine, 350 West 11th Street, IUHPL 4050, Indianapolis, IN 46202, USA. E-mail: ypolar@iu.edu or sbadve@iupui.edu

Received 27 May 2014; accepted 20 August 2014; published online 28 November 2014

associated with alternative splicing in lymphocytic leukemia and uveal melanomas.^{8,9} However, the impact of these factors in the clinical setting has not been well established in breast cancer, and their clinical utility is not clear.

SF3B1 is one of the 35 most frequently mutated genes using next-generation sequencing on 510 breast tumors.⁵ However, the frequency is low (2% of all tumors). Of the 15 nonsilent mutations, the majority were missense mutations. Patients with estrogen receptor (ER)-positive and HER2 subtypes harbored the majority of these mutations. The SF3B1 was also among the 18 significantly mutated genes in untreated ER-positive breast tumors from 77 patients accrued from two neoadjuvant aromatase inhibitor clinical trials.¹⁰ Taken together, these studies indicate the potential role of SF3B multi-protein complex in ER-positive breast cancers.

SF3B1 encodes the largest subunit of the SF3B protein complex, a key component of the core spliceosome complex essential for the accurate excision of introns from pre-mRNA to form mature mRNA. SF3B1 interacts with other SF3B subunits including SF3B3.^{11–14} Based on the pre-clinical data, SF3B1 mutation (R1074H) confers resistance to spliceosome inhibitors, such as pladienolide, and impairs its binding to SF3B complex.¹⁵ Furthermore, SF3B1 and SF3B3 interactions are believed to be necessary to form a binding site for this class of inhibitors.^{15,16} Therefore, these genes may serve as potential new therapeutic targets in ER-positive breast cancer. However, there is a relative lack of data regarding the expression and prognostic value of these genes in ER-positive breast cancer and their clinical utility in endocrine therapy. Despite the presence of SF3B1 mutations, no SF3B3 mutations were detected in either The Cancer Genome Atlas (TCGA) or in neoadjuvant aromatase inhibitor clinical trials. Decreased SF3B3 expression has been reported in breast cancer, but attributed to loss of 16q, the location of the gene.¹⁷ In this study, the expression of SF3B1 and SF3B3 was analyzed in breast cancer cell lines with acquired resistance to endocrine therapy and in a cohort of Oncotype DX cases using real-time quantitative reverse transcription PCR (qRT-PCR). The findings were further confirmed in publicly available breast cancer data sets, including TCGA, and correlated their prognostic relevance and clinical utility for endocrine therapy. Here, we report the clinical relevance of SF3B1 and SF3B3 in development of innate and acquired endocrine resistance in ER-positive breast cancers.

Materials and methods

Breast Cancer Cell Lines

LCC2 (MCF-7/LCC2; resistant to tamoxifen), LCC9 (MCF-7/LCC9; resistant to fulvestrant (Faslodex; ICI

182,780) and cross-resistant to tamoxifen), and AZ (MCF-7/AZ control) cell lines were kind gifts from Dr R. Clarke (Georgetown University Medical School, Washington DC).^{18,19} Cell lines have been carefully maintained in a humidified tissue culture incubator at 37 °C in 5% CO₂:95% air atmosphere and stocks of the earliest passage cells have been stored. The cell lines were grown in phenol-red-free DMEM containing 5% charcoal-stripped fetal calf serum and 100 mg/ml penicillin as described previously.^{18,19}

Oncotype DX Samples

All protocols were reviewed and approved by the Institutional Review Board of Indiana University. Sixty archival formalin-fixed, paraffin-embedded tumor blocks were obtained from patients with ER-positive (> 1% expression as per ASCO-CAP guidelines) node-negative breast carcinomas at the Indiana University Simon Cancer Center based on their Oncotype DX recurrence score (19 low score, 21 intermediate score, and 20 high score). Four of the 60 cases had a lobular histology. Demographic and clinical characteristics of the patients were acquired from medical charts (Supplementary Table S1).

RNA Isolation and qRT-PCR

RNA was extracted from 10 μm-thick sections of archival paraffin blocks using RecoverAll total nucleic acid isolation kit (Life Technologies, Grand Island, NY, USA). For breast cancer cell lines, RNAs were isolated (RNeasy isolation kit, Qiagen, VA, USA) and treated with Turbo DNase (Ambion, Foster City, CA, USA) to remove contaminating DNA. The quality of RNA was assessed using the Nanodrop ND-1000 spectrophotometer (ThermoScientific, Wilmington, DE, USA). Total RNAs were reverse-transcribed using the high capacity cDNA reverse transcription kit (Life Technologies) according to the manufacturer's instructions. The mRNA levels of SF3B1 and SF3B3 were analyzed by qRT-PCR using TaqMan gene expression assays on an ABI Prism 7900 platform (Applied Biosystems, Foster City, CA, USA) with ACTB and GUSB as endogenous controls for normalization. All qRT-PCR reactions from tumor blocks and breast cancer cell lines were performed in duplicates and triplicates, respectively. The relative quantification of the gene expression changes (fold) was analyzed according to $\Delta\Delta C_t$ method using the Applied Biosystems DataAssist software v3.0. For qRT-PCR statistical analysis, Applied Biosystems DataAssist software v3.0 was performed and all graphs were generated using GraphPad Prism 5 software. The error bars were calculated and represented in terms of mean s.d.

Analysis Using TCGA Database

To validate the clinical relevance of *SF3B1* and *SF3B3* levels in larger cohorts, we obtained the normalized expression levels of *SF3B1* and *SF3B3* (level 3 data) in breast cancer patients enrolled in the TCGA breast invasive carcinoma study with subtype classification (available at <http://tcga-data.nci.nih.gov/tcga/tcgaHome2.jsp>). Patients with breast cancer were categorized as having one of four different subtypes (luminal A, luminal B, HER2-enriched, or basal) based on the PAM50 signature. The expression of *SF3B1* and *SF3B3* in all four subtypes was obtained and compared with luminal A breast tumors. The significance of change in expression of each subtype from that of the luminal A cohort was estimated using Wilcoxon test.

The clinical information for each patient was also obtained. To model survival, gene expression at or below 33rd percentile was considered low, at or above 67th percentile was considered high, and those falling within the 33rd and 67th percentile were considered medium. Overall survival was calculated from the date of initial diagnosis of breast cancer to disease-specific death (patients whose vital status is termed dead) and months to last follow up (for patients who are alive). Kaplan–Meier survival analysis was used to estimate association of the gene's expression with survival of patients. The 'survival' package in R (R Foundation for Statistical Computing) was used for statistical analyses. An analysis of relapse-free survival was not possible in TCGA data set.

Analysis of Publicly Available Data Sets

SF3B1 and *SF3B3* expression levels were analyzed based on ER status, molecular subtypes, and other clinicopathological parameters using the data sets from the gene expression-based outcome for breast cancer online algorithm (GOBO).²⁰ GOBO is a web-based analysis tool that utilizes 11 publicly available Affymetrix U133A gene expression data curated from 1881 breast cancer patients with associated stage, grade, nodal status, and intrinsic molecular classification.²¹ Of all 1881 tumors, the groups were distributed as follows: (1) ER-positive tumors ($n = 1225$), (2) ER-negative tumors ($n = 395$), (3) systemically untreated patients ($n = 927$), and (4) patients treated with tamoxifen alone ($n = 326$). Clinical characteristics of individual data sets were described previously.²⁰ Association of outcome was investigated for each patient cohort with relapse-free survival or overall survival as end points and 10-year censoring in the above groups. The Kaplan–Meier survival analysis was calculated using Cox proportional hazard model, and the score test of the proportional hazard model is equivalent to the log-rank test.

Results

SF3B1 and SF3B3 Are Upregulated in Acquired Endocrine Resistance Models

To further determine the relevance of the identified genes in endocrine resistance, we analyzed the relative expression levels of these genes in a panel of endocrine-resistant, ER-positive breast cancer cell lines using qRT-PCR. *SF3B1* and *SF3B3* expression levels were significantly upregulated in LCC2 cells compared to control cells (AZ). (fold change: 5.30; $P = 0.0002$ and fold change: 4.012; $P = 0.0002$, respectively) (Figure 1). LCC9 cells also exhibited significant upregulation for *SF3B1* (fold change: 2.92; $P = 0.0008$) and *SF3B3* (fold change: 3.58; $P = 0.0006$). Interestingly, *SF3B1* and *SF3B3* levels did not change in response to E2 alone, tamoxifen alone, or E2 and tamoxifen treatment in combination, indicating that their

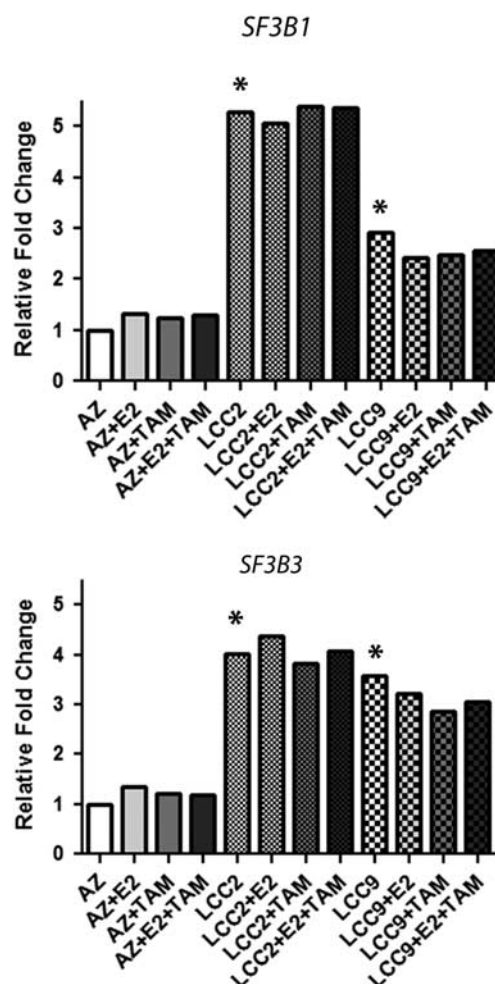


Figure 1 Expression levels of *SF3B1* and *SF3B3* in tamoxifen-resistant LCC2 (MCF-7/LCC2), fulvestrant and tamoxifen-cross-resistant LCC9 (MCF-7/LCC2), and control sensitive cell line (MCF-7/AZ) using real-time quantitative reverse transcription PCR assay; vehicle (ethanol), E2 (β -estradiol: 10^{-10} M), tamoxifen (4-OH-Tam: 10^{-6} M), or in combination with E2 and tamoxifen. * $P < 0.05$; statistically significant.

upregulation is independent of E2 stimulation. These findings suggest the importance of elevated expression of *SF3B1* and *SF3B3* in endocrine resistance and the likelihood of recurrence of ER-positive breast cancer.

SF3B3 Is Significantly Upregulated in Cases with Oncotype DX High-Recurrence Scores

The Oncotype DX recurrence score in current practice predicts the likelihood of distant recurrence in tamoxifen-treated patients with node-negative, ER-positive breast cancer.²² Furthermore, analysis of the National Surgical Adjuvant Breast and Bowel Project B14 clinical trial has shown that patients with high-recurrence score have innate resistance to tamoxifen. However, the assay does not provide the mechanistic basis for endocrine resistance. As tamoxifen resistance might contribute to early recurrence, we first decided to determine the expression levels of *SF3B1* and *SF3B3* in a cohort of 60 paraffin-embedded ER-positive, node-negative breast carcinomas with low, intermediate, and high (19, 21, and 20 cases, respectively) recurrence scores. qRT-PCR analysis revealed that expression of *SF3B3* was upregulated in cases with high-recurrence scores compared with the low-recurrence score cases (fold change: 2.44-fold; $P=0.019$) (Figure 2). However, expression of *SF3B1* was slightly increased and did not reach statistical significance (fold change: 1.26-fold). Expression of *SF3B1* and *SF3B3* remained nonsignificant in cases with intermediate score compared with cases with low scores being 0.90-fold and 1.21-fold, respectively. These results suggest that high expression of *SF3B3* may indicate the likelihood of high recurrence and might contribute to endocrine resistance in ER-positive cancers.

Validation of SF3B3 Expression Using Specimens from TCGA and Affymetrix Data Sets of Breast Cancer

To assess the correlation of *SF3B3* expression in larger cohorts of breast cancer, we analyzed the data

of breast tumors from TCGA that can categorize the samples to PAM50 subtypes.⁵ The expression of *SF3B3* was higher in luminal B ($n=290$; $P=7.59e-12$), HER2 ($n=102$; $P=7.586e-19$), and basal ($n=166$; $P=1.66e-51$) subtypes when compared with luminal A ($n=366$) (Figure 3a), suggesting that the expression of *SF3B3* correlated positively with the aggressiveness of subtypes. Luminal A tumors exhibited the lowest *SF3B3* expression, while the sequential order of increase was observed in luminal B, HER2, and basal subtype being the highest (Figure 3a).

Expression of *SF3B3* was also assessed using data from GOBO for all breast cancer tumors ($n=1881$).²⁰ *SF3B3* expression was significantly lower in ER-positive tumors ($n=1225$) compared with ER-negative ($n=395$; $P<0.00001$) (Figure 3b). Using the PAM50 subtypes (Figure 3c), the expression levels were higher in the order of luminal B ($n=471$), HER2 ($n=240$), and basal subtype ($n=304$) being the highest, consistent with TCGA expression data. *SF3B3* expression also correlated positively with higher grade having positive association with grade 3 cases ($n=239$ for grade 1, $n=677$ for grade 2, and $n=495$ for grade 3) (Supplementary Figure S1).

We also analyzed the levels of *SF3B1* expression in TCGA and Affymetrix data sets. In TCGA, *SF3B1* levels did not reach statistical significance in luminal B and HER2 in comparison with luminal A tumors. *SF3B1* was significantly decreased only in basal subtype ($P=0.04$) when compared with luminal A (Supplementary Figure S2A). Consistent with this observation, *SF3B1* was significantly lower in ER-negative tumors compared with ER-positive tumors in Affymetrix data sets ($P<0.00001$) (Supplementary Figure S2B), with highest expression in luminal A; whereas, its expression was lower in luminal B, HER2, and basal subtypes (Supplementary Figure S2C). On the other hand, high expression of *SF3B1* is associated with grade 1 (Supplementary Figure S2D). These results suggest a differential role for *SF3B1* and *SF3B3* in breast cancer.

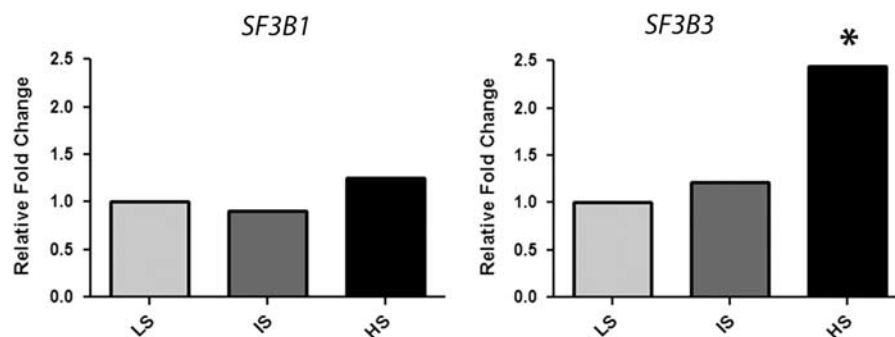


Figure 2 Expression of *SF3B1* and *SF3B3* in patients with estrogen receptor-positive tumors using real-time quantitative reverse transcription PCR assay. The expression of *SF3B1* and *SF3B3* is higher in patients with high-recurrence scores; LS, low score; IS, intermediate score; HS, high score in Oncotype DX. * $P=0.019$; statistically significant.

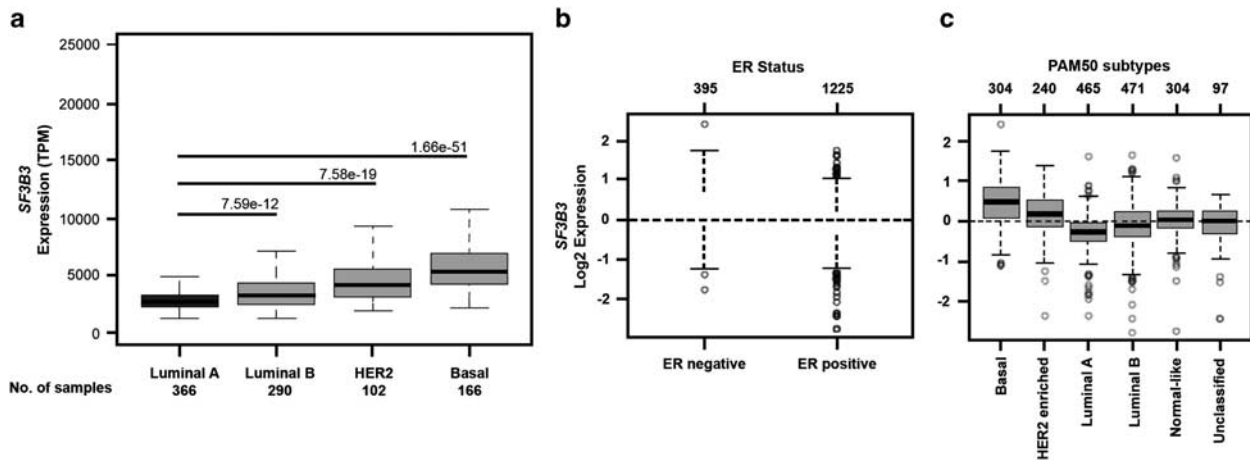


Figure 3 *SF3B3* expression levels using The Cancer Genome Atlas (TCGA) data set and Affymetrix data sets using GOBO online analysis. (a) PAM50 subtypes in TCGA, (b) estrogen receptor (ER) status (GOBO), and (c) PAM50 subtypes (GOBO).

High Expression of *SF3B3* Correlates with Poor Prognosis in Patients with ER-Positive Breast Cancer

We next analyzed the correlation of *SF3B1* and *SF3B3* expression with overall survival using TCGA data of breast cancer subjects. The expression of *SF3B3* in these subjects was categorized in three quantiles based on low ($n = 324$; black line), medium ($n = 314$; light gray line), and high ($n = 314$; dark gray line) expressions (Figure 4a). The higher *SF3B3* expression was associated with shorter overall survival compared with lower expression in breast cancer patients ($P = 0.00461$). In a cohort with high *SF3B3* levels, the overall survival probability was 79.6% and 47.0% at 50 and 100 months, respectively, and those with low *SF3B3* levels showed a survival probability of 90.4% and 67.4% at 50 and 100 months, respectively. On the other hand, *SF3B1* levels were not associated with survival (high *SF3B1*, 82.8% and 46.9% at 50 and 100 months versus low *SF3B1*, 87.9% and 58.6% for 50 and 100 months, respectively) (Supplementary Figure S3a).

We further assessed the correlation of *SF3B3* expression with overall survival and relapse-free survival in various categories including ER status, and tamoxifen-treated cases representative of 11 microarray data sets using the GOBO tool (Figures 4b and c). High expression of *SF3B3* was associated with shorter overall survival in all tumors ($P = 1e - 05$) and in ER-positive tumors ($P = 1e - 05$), but not in ER-negative tumors (Figure 4b). Higher expression of *SF3B3* was also associated with relapse-free survival in all tumors ($P = 0.00001$) and in ER-positive tumors ($P = 2e - 05$) (Figure 4c). Significant association of poor relapse-free survival and high *SF3B3* was also observed in ER-positive tamoxifen-treated population ($P = 0.02573$). Consistent with overall survival, the expression levels were not associated with relapse-free survival in ER-negative tumors.

Correlation of high *SF3B3* expression with shorter overall survival stayed significant for other para-

eters (all tumors: node-negative; $P = 0.00551$ and node-positive; $P = 0.00093$) (Supplementary Figure S3b). *SF3B3* was also significantly correlated with relapse-free survival for these parameters (all tumors: node-negative; $P = 0.01679$ and node-positive; $P = 8e - 05$). In ER-positive tumors, node-negative tumors were more significantly associated with high *SF3B3* (Supplementary Figure S3c).

Analysis of *SF3B3* expression with clinical variables showed that low *SF3B3* levels and node-negative tumors were significantly associated with better overall survival (hazard ratio (HR) = 0.46; 95% CI = 0.33–0.64; $P = 3.838e - 06$) and relapse-free survival (HR = 0.72; 95% CI = 0.54–0.96; $P = 0.0266$) compared with node-positive tumors (Tables 1 and 2). On the other hand, HR for larger tumors (> 20 mm) was significantly high for overall survival (HR = 1.94; 95% CI = 1.38–2.72; $P = 0.0001$) and for relapse-free survival (HR = 1.92; 95% CI = 1.45–2.54; $P = 5.388e - 06$), being a significant factor for survival and relapse. Age (> 50) stayed significant for overall survival (HR = 1.58; 95% CI = 1.11–2.24; $P = 0.0117$), but not relapse-free survival. Grade was not significantly associated with either overall survival or relapse-free survival in multivariable analysis. No significant correlation was observed for *SF3B1* expression with overall survival and/or relapse-free survival or in any categories in both TCGA and Affymetrix data sets in GOBO analysis (data not shown). These results further validate the involvement of high expression of *SF3B3* with poor prognosis using large cohorts of breast cancer and its potential contribution to endocrine resistance in ER-positive cancers.

Discussion

ER-positive subtype constitutes 65–70% of all breast cancers. ER expression is a strong predictive factor for efficacy of endocrine therapy. Although

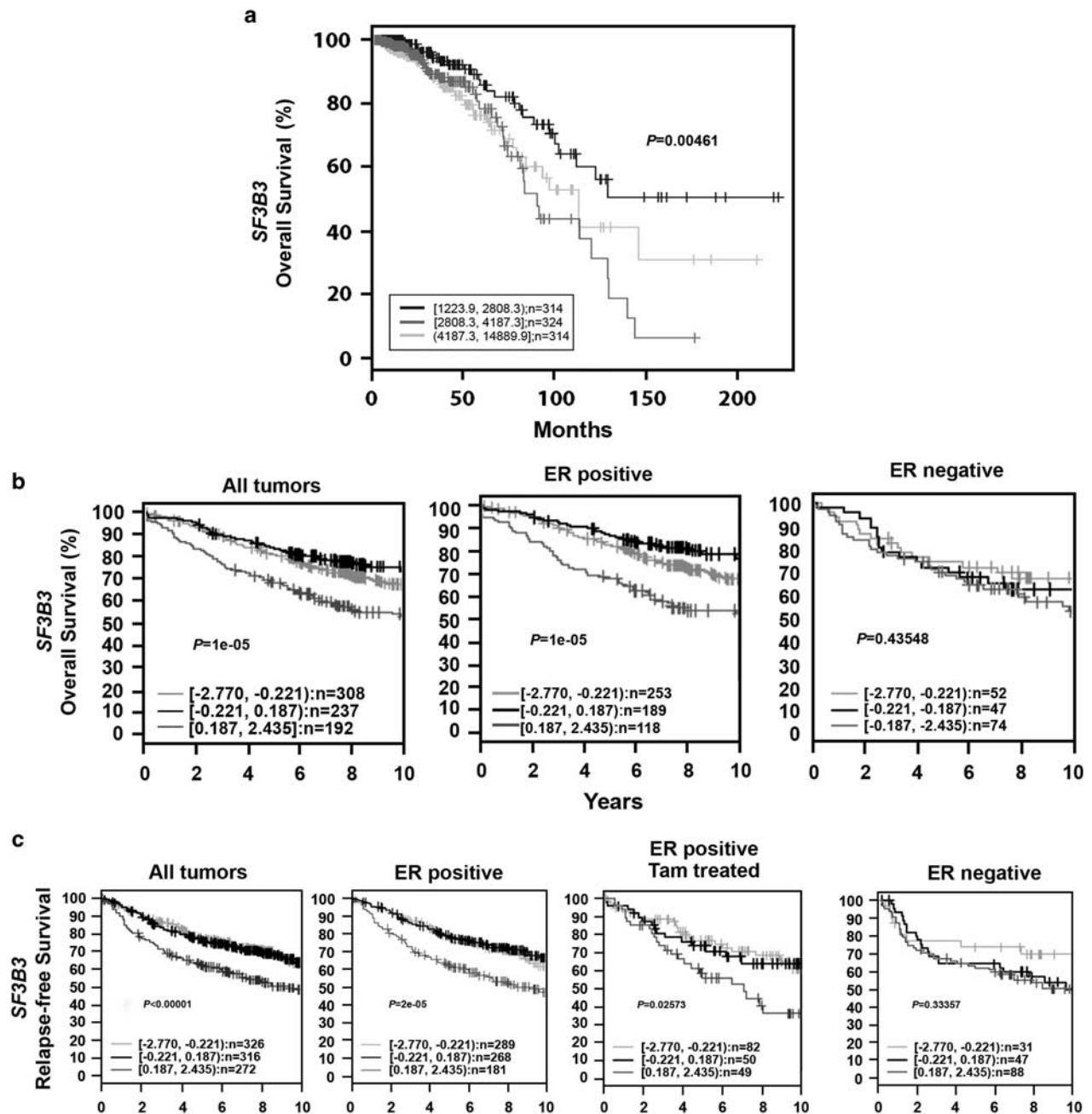


Figure 4 Kaplan–Meier analysis using The Cancer Genome Atlas (TCGA) data set and Affymetrix data sets using GOBO online analysis. (a) Overall survival in TCGA (months); (b) 10-year overall survival (OS) as end point for all tumors, estrogen receptor (ER)-positive tumors, and ER-negative tumors; and (c) 10-year relapse-free survival (RFS)-free survival as end point for all tumors, ER-positive tumors, ER-positive tamoxifen (Tam)-treated tumors, and ER-negative tumors. The expression analysis was stratified into three quantiles based on *SF3B3* expression; low expression (black), medium expression (light gray), and high expression (dark gray).

endocrine therapy is effective in the early stage of ER-positive breast cancer, recurrence and resistance to therapy is the principal cause of morbidity and mortality from breast cancer. The exact process by which this occurs is complex. Thus, it is crucial to identify targets with better prognostic and predictive ability to improve the success of endocrine therapies and prevent breast cancer mortality. The Oncotype DX recurrence score in current practice

predicts the likelihood of distant recurrence in tamoxifen-treated patients with node-negative, ER-positive breast cancer.²² However, like other signatures for ER-positive cancers, it is predominantly based on proliferation and does not provide predictive ability to overcome *de novo* or acquired resistance.

Choice of endocrine therapy varies based on the menopausal status of women. In premenopausal

Table 1 Multivariable analysis in ER-positive patients for overall survival (Affymetrix data sets-GOBO)

Variable	HR (95% CI)	P value*
Low <i>SF3B3</i> vs high <i>SF3B3</i>	0.6 (0.41–0.89)	0.011
Node negative vs node positive	0.46 (0.33–0.64)	3.838e – 06
Grade 3 vs grade 1 or grade 2	1.3 (0.89–1.88)	0.7196
Age ≥ 50 years vs < 50 years	1.58 (1.11–2.24)	0.0117
Tumor size ≥ 20 mm vs < 20 mm	1.94 (1.38–2.72)	0.0001

Abbreviations: CI, confidence interval; HR, hazard ratio.
* $P \leq 0.05$ is considered as statistically significant.

Table 2 Multivariable analysis in ER-positive patients for relapse-free survival (Affymetrix data sets-GOBO)

Variable	HR (95% CI)	P value*
Low <i>SF3B3</i> vs high <i>SF3B3</i>	0.62 (0.45–0.86)	0.0046
Node negative vs node positive	0.72 (0.54–0.96)	0.0266
Grade 3 vs grade 1 or grade 2	1.09 (0.79–1.5)	0.6127
Age ≥ 50 years vs < 50 years	0.81 (0.61–1.09)	0.1634
Tumor size ≥ 20 mm vs < 20 mm	1.92 (1.45–2.54)	5.388e – 06

Abbreviations: CI, confidence interval; HR, hazard ratio.
* $P \leq 0.05$ is considered as statistically significant.

women with hormone receptor-positive disease, tamoxifen, a selective estrogen receptor modulator, is considered the standard endocrine therapy of choice in combination with or without ovarian suppression/ablation. Recent results of the Adjuvant Tamoxifen: Longer Against Shorter and the Adjuvant Tamoxifen Treatment Offer More trials demonstrated the beneficial effect of extended tamoxifen therapy in ER-positive breast cancer and provided an important treatment approach in high-risk young patients.^{23,24} However, despite these successes, the resistance to tamoxifen still remains a problem which might develop through multiple mechanisms, including the deregulation of ER pathway and upregulation of growth factor signaling. This further switches the tumor pathways to hormone-independent and nonnuclear/nongenomic ER activities.²⁵

A growing body of evidence suggests that splicing factors may provide prognostic and predictive utility in several cancers including breast cancer.² In particular, the identification of mutations in splicing factor genes may present candidates with prognostic and predictive utility. They also may serve as excellent therapeutic targets as they regulate alternative splicing of multiple genes that contribute to cancer recurrence and resistance to therapies. Herein, we demonstrated that *SF3B3* was significantly upregulated in tamoxifen-resistant LCC2 and in fulvestrant-resistant and tamoxifen cross-resistant LCC9 cell lines, suggesting its association with resistance to endocrine therapy.

In this study, we also demonstrated that higher expression of *SF3B3* was seen in ER-negative cancers than in ER-positive cancers. In ER-positive cancers, consistent with the cell line studies, the expression of *SF3B3* was higher in luminal B tumors than in luminal A tumors, reiterating the possible role in aggressive types of ER-positive cancers. Consistent with the aggressive nature of tumors, we further showed that higher expression of *SF3B3*, but not *SF3B1*, was significantly associated with Oncotype DX high-score cases.

To confirm the data observed, we performed analysis of large publically available gene expression data sets. Analyses were performed using TCGA cohorts and GOBO Affymetrix-based data sets. These analyses confirmed the association of high *SF3B3* expression with high histological grade. Most importantly, larger data sets, including TCGA and microarray, further confirmed the significant correlation of high *SF3B3* levels with shorter overall survival in ER-positive, but not in ER-negative patients. The association of *SF3B3* expression with poor outcome was even more prominent in node-positive tumors than in node-negative tumors. An analysis of relapse-free survival was not possible in TCGA data set due to the short follow-up period in this cohort. However, our analysis using GOBO data set documented that *SF3B3* expression was associated with shorter overall survival. In GOBO data set, information about tamoxifen treated was available on 326 patients. In this subset analysis, high expression of *SF3B3* was associated with poor outcome (Figure 4). This supports the postulated association of high *SF3B3* expression in tamoxifen resistance.

SF3B1 was not correlated with relapse-free survival in either ER-positive or ER-negative breast cancers lacking any prognostic relevance using the breast cancer data sets, although its mutations have been identified among the most significantly mutated genes in TCGA data set.⁵ This suggests that *SF3B1* may not be a major driver in recurrence of these cancers. *SF3B1* and *SF3B3* proteins are known to form a complex that is integral for splicing the RNAs. Thus the lack of association of *SF3B1* with survival is surprising, and will be a subject of further studies. Similarly, the role of *SF3B3* and *SF3B1* in HER2 and basal subtypes is not clear, as they did not significantly correlate with relapse-free survival in these tumors. This may indicate differential regulation of *SF3B3* based on the ER status.

The combination of endocrine therapy with other drugs targeting key molecules involved in endocrine resistance is the most promising approach to prevent and/or overcome endocrine resistance and benefit these breast cancer patients. The spliceosome inhibitors targeting SF3B complex may serve as potential candidates. Several groups have developed inhibitors targeting spliceosomal pathway to reduce the cellular growth of cancer cells. Among them, synthetic analogs, such as spliceostatin A,

meayamycin, sudemycins, and E7107 are currently under study.^{16,26–28} Meayamycin has been found more potent and stable than its natural compound FR901464 in various tumor cell lines, including MCF-7 and MDA-MB231 breast cancer cell lines.²⁹ Sudemycins, analogs of parental compound FR901464, have also been developed to target SF3b and shown to modulate alternative splicing in human tumor xenografts.^{28,30} E7107 entered clinical trials as an anticancer agent. However, the incidence of a few cases of vision loss led to study discontinuation.^{31,32} This may be drug specific and other inhibitors need to be investigated.

In conclusion, our findings suggest a role for the elevated expression of SF3B3 in poor prognosis and tamoxifen resistance, leading to recurrence of ER-positive breast cancer. Confirmation of these findings could lead to new treatment strategies, including spliceosome inhibitors to reverse *de novo* and acquired tamoxifen resistance. Such therapies could prevent not only early, but also late recurrence in ER-positive subtype, which is a major cause of anxiety in breast cancer survivors.

Acknowledgments

Sunil Badve is a Komen Scholar and the studies were supported by funds from the Susan G Komen for the Cure and 100 Voices of Hope.

Disclosure/conflict of interest

The authors declare no conflict of interest.

References

- Grosso AR, Martins S, Carmo-Fonseca M. The emerging role of splicing factors in cancer. *EMBO Rep* 2008;9:1087–1093.
- Yoshida K, Ogawa S. Splicing factor mutations and cancer. *Wiley Interdiscip Rev RNA* 2014;5:445–459.
- Malcovati L, Papaemmanuil E, Bowen DT, *et al*. Clinical significance of SF3B1 mutations in myelodysplastic syndromes and myelodysplastic/myeloproliferative neoplasms. *Blood* 2011;118:6239–6246.
- Biankin AV, Waddell N, Kassahn KS, *et al*. Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. *Nature* 2012;491:399–405.
- Cancer Genome Atlas N. Comprehensive molecular portraits of human breast tumours. *Nature* 2012;490:61–70.
- Harbour JW, Roberson ED, Anbunathan H, *et al*. Recurrent mutations at codon 625 of the splicing factor SF3B1 in uveal melanoma. *Nat Genet* 2013;45:133–135.
- Wang L, Lawrence MS, Wan Y, *et al*. SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. *N Engl J Med* 2011;365:2497–2506.
- Furney SJ, Pedersen M, Gentien D, *et al*. SF3B1 mutations are associated with alternative splicing in uveal melanoma. *Cancer Discov* 2013;3:1122–1129.
- Quesada V, Ramsay AJ, Lopez-Otin C. Chronic lymphocytic leukemia with SF3B1 mutation. *N Engl J Med* 2012;366:2530.
- Ellis MJ, Ding L, Shen D, *et al*. Whole-genome analysis informs breast cancer response to aromatase inhibition. *Nature* 2012;486:353–360.
- Das BK, Xia L, Palandjian L, *et al*. Characterization of a protein complex containing spliceosomal proteins SAPs 49, 130, 145, and 155. *Mol Cell Biol* 1999;19:6796–6802.
- Kramer A, Gruter P, Groning K, *et al*. Combined biochemical and electron microscopic analyses reveal the architecture of the mammalian U2 snRNP. *J Cell Biol* 1999;145:1355–1368.
- Will CL, Schneider C, MacMillan AM, *et al*. A novel U2 and U11/U12 snRNP protein that associates with the pre-mRNA branch site. *EMBO J* 2001;20:4536–4546.
- Will CL, Urlaub H, Achsel T, *et al*. Characterization of novel SF3b and 17S U2 snRNP proteins, including a human Prp5p homologue and an SF3b DEAD-box protein. *EMBO J* 2002;21:4978–4988.
- Yokoi A, Kotake Y, Takahashi K, *et al*. Biological validation that SF3b is a target of the antitumor macrolide pladienolide. *FEBS J* 2011;278:4870–4880.
- Kotake Y, Sagane K, Owa T, *et al*. Splicing factor SF3b as a target of the antitumor natural product pladienolide. *Nat Chem Biol* 2007;3:570–575.
- Hungermann D, Schmidt H, Natrajan R, *et al*. Influence of whole arm loss of chromosome 16q on gene expression patterns in oestrogen receptor-positive, invasive breast cancer. *J Pathol* 2011;224:517–528.
- Brunner N, Frandsen TL, Holst-Hansen C, *et al*. MCF7/LCC2: a 4-hydroxytamoxifen resistant human breast cancer variant that retains sensitivity to the steroidal antiestrogen ICI 182,780. *Cancer Res* 1993;53:3229–3232.
- Brunner N, Boysen B, Jirus S, *et al*. MCF7/LCC9: an antiestrogen-resistant MCF-7 variant in which acquired resistance to the steroidal antiestrogen ICI 182,780 confers an early cross-resistance to the nonsteroidal antiestrogen tamoxifen. *Cancer Res* 1997;57:3486–3493.
- Ringner M, Fredlund E, Hakkinen J, *et al*. GOBO: gene expression-based outcome for breast cancer online. *PLoS One* 2011;6:e17911.
- Perou CM, Sorlie T, Eisen MB, *et al*. Molecular portraits of human breast tumours. *Nature* 2000;406:747–752.
- Paik S, Shak S, Tang G, *et al*. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 2004;351:2817–2826.
- Davies C, Pan H, Godwin J, *et al*. Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial. *Lancet* 2013;381:805–816.
- Gray RG, Rea D, Handley K, *et al*. aTTom: Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years in 6,953 women with early breast cancer. *J Clin Oncol* 2013;31:5.
- Osborne CK, Schiff R. Mechanisms of endocrine resistance in breast cancer. *Annu Rev Med* 2011;62:233–247.
- Rymond B. Targeting the spliceosome. *Nat Chem Biol* 2007;3:533–535.
- Bonnal S, Vigevani L, Valcarcel J. The spliceosome as a target of novel antitumour drugs. *Nat Rev Drug Discov* 2012;11:847–859.
- Fan L, Lagisetti C, Edwards CC, *et al*. Sudemycins, novel small molecule analogues of FR901464, induce

- alternative gene splicing. *ACS Chem Biol* 2011;6: 582–589.
- 29 Albert BJ, McPherson PA, O'Brien K, *et al*. Meayamycin inhibits pre-messenger RNA splicing and exhibits picomolar activity against multidrug-resistant cells. *Mol Cancer Ther* 2009;8:2308–2318.
- 30 Lagisetti C, Pourpak A, Goronga T, *et al*. Synthetic mRNA splicing modulator compounds with in vivo antitumor activity. *J Med Chem* 2009;52:6979–6990.
- 31 Hong DS, Kurzrock R, Naing A, *et al*. A phase I, open-label, single-arm, dose-escalation study of E7107, a precursor messenger ribonucleic acid (pre-mRNA) splicing inhibitor administered intravenously on days 1 and 8 every 21 days to patients with solid tumors. *Invest New Drugs* 2013;32:436–444.
- 32 Dehm SM. Test-firing ammunition for spliceosome inhibition in cancer. *Clin Cancer Res* 2013;19: 6064–6066.

Supplementary Information accompanies the paper on Modern Pathology website (<http://www.nature.com/modpathol>)