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

Yesim Gökmen-Polar[1*], Yaseswini Neelamraju[3], Chirayu P. Goswami[4], Xiaoping Gu[1], Gouthami Nallamothu[1], Sarath Chandra Janga[3,4,5], and Sunil Badve[1,2,6*]

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

Grant Support
This work was supported by the Susan G Komen for the Cure® and 100 Voices of Hope to Sunil Badve.

Disclosure of Potential Conflicts of Interest
The authors disclose no potential conflicts of interest.

*Address of corresponding:

Yesim Gökmen-Polar, PhD
Department of Pathology and Laboratory Medicine
Indiana University School of Medicine
635 Barnhill Dr., MS 0038
Indianapolis, IN 46202
ypolar@iu.edu

Sunil Badve, MD, FRCPath
Department of Pathology and Laboratory Medicine
Indiana University School of Medicine
350 West 11th Street, IUHPL 4050
Indianapolis, IN 46202, USA
sbadve@iupui.edu

Running head: *SF3B1* and *SF3B3* expression in endocrine resistance

Keywords: Breast cancer, *SF3B1*, *SF3B3*, endocrine resistance, splicing factors

Abstract

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 Onco*type* DX recurrence scores. Expression levels of *SF3B1* and *SF3B3* and their prognostic value were validated in large cohorts using publicly available gene expression datasets including The Cancer Genome Atlas.

SF3B1 and *SF3B3* levels were significantly increased in ERα-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 Onco*type* DX high risk groups when compared with low risk ($P=0.019$). Similarly, in publicly-available breast cancer gene expression datasets, 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 splicing factor *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.

INTRODUCTION

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 signaling 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 that the presence of somatic mutations in splicing factors.[2] Among them, the splicing factor *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 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 ER-positive and HER2 subtypes harbored the majority of these mutations. The *SF3B1* was also amongst 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 multiprotein 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 premessenger RNA to form mature mRNA. SF3B1 interacts with other SF3B subunits including SF3B3.[11-14] Based

on the preclinical 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 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 Onco*type* DX cases using qRT-PCR. The findings were further confirmed in publicly available breast cancer datasets, including The Cancer Genome Atlas, 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.

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^{o}C in 5% CO_2:95% 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 (CCS) and 100 mg/mL penicillin as described previously.[18, 19]

Onco*type* 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 (greater than 1% expression as per ASCO-CAP guidelines) node-negative breast carcinomas at the Indiana University Simon Cancer Center based on their Onco*type* 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 (Suppl Table S1).

RNA isolation and qRT-PCR

RNA was extracted from 10μm-thick sections of archival paraffin blocks using RecoverAll™ total nucleic acid isolation kit (Life Technologies, Grand Island, NY). For breast cancer cell lines, RNAs were isolated (RNeasy isolation kit, Qiagen, VA) and treated with Turbo DNase® (Ambion, Foster City, CA) to remove contaminating DNA. The quality of RNA was assessed using the Nanodrop® ND-1000 spectrophotometer (ThermoScientific, Wilmington, DE). Total RNAs were reverse-transcribed using the high capacity cDNA reverse transcription kit (Life Technologies) according 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) 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 ΔΔCt 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 SD.

Analysis using The Cancer Genome Atlas (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 (BRCA) study with subtype classification (available at http://tcgadata.nci.nih.gov/tcga/tcgaHome2.jsp). Patients with breast cancer were categorized as having either of the four different subtypes (luminal A, luminal B, HER2-enriched, or basal) based on the PAM50 signature. The expression of *SF3B1* and *SF3B3* in all the 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 Wilcox 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 deaths (patients whose vital status is termed dead) and months to last followup (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 The Cancer Genome Atlas dataset.

Analysis of publicly available datasets

SF3B1 and *SF3B3* expression levels were analyzed based on ER-status, molecular subtypes, and other clinicopathological parameters using the datasets 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 datasets were described previously.[20] Association of outcome was investigated for each patient cohort with relapse-free survival or overall survival as endpoints 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 their normal counterparts, (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 both *SF3B1* (fold change: 2.92; *P*=0.0008) and *SF3B3*

(fold change: 3.58; *P*=0.0006). Interestingly, both *SF3B1* and *SF3B3* levels did not change in response to E2 alone, tamoxifen alone, or E2 and tamoxifen treatment in combination indicating that their 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 Onco*type* DX high recurrence scores

The Onco*type* 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. Since 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 to 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 to 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.

205 **Validation of *SF3B3* expression using specimens from The Cancer Genome Atlas and**
206 **Affymetrix datasets of breast cancer**

207 To assess the correlation of *SF3B3* expression in larger cohorts of breast cancer, we
208 analyzed the data of breast tumors from The Cancer Genome Atlas (TCGA) that can categorize
209 the samples to PAM50 subtypes.[5] The expression of *SF3B3* was higher in luminal B (*n*=290;
210 *P*=7.59e-12 HER2 (*n*=102; *P*=7.586e-19, and basal (*n*=166; *P*=1.66e-51) subtypes when
211 compared with luminal A (*n*=366) (Figure 3A), suggesting that the expression of *SF3B3*
212 correlated positively with the aggressiveness of subtypes. Luminal A tumors exhibited the lowest
213 *SF3B3* expression, while the sequential order of increase was observed in luminal B, HER2, and
214 basal subtype being the highest (Figure 3A).

215 Expression of *SF3B3* was also assessed using data from GOBO for all breast cancer
216 tumors (*n*=1881).[20] *SF3B3* expression was significantly lower in ER-positive tumors (*n*=1225)
217 compared to ER-negative (*n*=395; *P*<0.00001) (Figure 3B). Using the PAM50 subtypes (Figure
218 3C), the expression levels were higher in the order of luminal B (*n*=471), HER2 (*n*=240), and
219 basal subtype (*n*=304) being the highest, consistent with TCGA expression data. *SF3B3*
220 expression also correlated positively with higher grade having positive association with grade 3
221 cases (*n*=239 for grade 1, *n*=677 for grade 2, and *n*=495 for grade 3) (Suppl Figure S1).

222 We also analyzed the levels of *SF3B1* expression in TCGA and Affymetrix datasets. In
223 TCGA, *SF3B1* levels did not reach statistical significance in luminal B and HER2 in comparison
224 to luminal A tumors. *SF3B1* was significantly decreased only in basal subtype (*P*=0.04) when
225 compared with luminal A (Suppl Figure S2A). Consistent with this observation, *SF3B1* was
226 significantly lower in ER-negative tumors compared with ER-positive tumors in Affymetrix
227 datasets (*P*<0.00001) (Suppl Figure S2B), with highest expression in luminal A; whereas, its

expression was lower in luminal B, HER2, and basal subtypes (Suppl Figure S2C). On the other hand, high expression of *SF3B1* is associated with grade 1 (Suppl Figure S2D). These results suggest a differential role for *SF3B1* and *SF3B3* in breast cancer.

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 *SF3B1* and *SF3B3* in these subjects was categorized in three quantiles based on the low (*n*=324; blue line), medium (*n*=314; red line), and high (*n*=314; black line) expression (Figure 4A). The higher *SF3B3* expression was associated with shorter overall survival compared with lower expression in breast cancer patients (*P*=0.00461). In 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) (Suppl Figure S3).

We further assessed the correlation of *SF3B3* expression with overall survival and relapse-free survival in various categories including ER-status, and LN-status representative of 11 microarray datasets using the GOBO tool (Figures 4B–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 parameters (all tumors: node-negative; *P*=0.00551 and node-positive; *P*=0.00093) (Suppl 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 (*Suppl Figure S3C).

Analysis of *SF3B3* expression with clinical variables showed that low *SF3B3* levels and LN- 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 to LN+ tumors (Table 1 and Table 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 were 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 The Cancer Genome Atlas and Affymetrix datasets 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 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 the breast cancer mortality. The Onco*type* 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 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 (ATLAS) and the Adjuvant Tamoxifen Treatment Offer More (aTTom) 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 remains still a problem which might develop through multiple mechanisms including the deregulation of ER pathway, upregulation of growth factor signaling. This further switches the tumor pathways to hormone-independent and nonnuclear/nongenomic ER activities.[25]

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

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

311 To confirm the data observed, we performed analysis of large publically available gene
312 expression datasets. Analyses were performed using The Cancer Genome Atlas cohorts and
313 GOBO Affymetrix-based datasets. These analyses confirmed the association of high *SF3B3*
314 expression with high histological grade. Most importantly, larger datasets including The Cancer
315 Genome Atlas and microarray further confirmed the significant correlation of high *SF3B3* levels
316 with shorter overall survival in ER-positive but not in ER-negative patients. The association of
317 *SF3B3* expression with poor outcome was even more prominent in node-positive tumors than in
318 node-negative tumors. An analysis of relapse-free survival was not possible in The Cancer
319 Genome Atlas dataset due to the short followup period in this cohort. However, our analysis

using GOBO datasets documented that *SF3B3* expression was associated with shorter overall survival. In GOBO dataset, information about tamoxifen treated was available on 326 patients. In this subset analysis, high expression of *SF3B3* was associated with poor outcomes (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 datasets, although its mutations have been identified among the most significantly mutated genes in The Cancer Genome Atlas dataset.[5] This may suggest that *SF3B1* may not be a major driver in recurrence of these cancers. SF3B1 and SF3B3 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 analogues 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, another analogues 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 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.

Acknowledgement 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.

References

1. Grosso AR, Martins S, Carmo-Fonseca M. The emerging role of splicing factors in cancer. EMBO Rep 2008; 9:1087-93.
2. Yoshida K, Ogawa S. Splicing factor mutations and cancer. Wiley Interdiscip Rev RNA 2014.
3. Malcovati L, Papaemmanuil E, Bowen DT, et al. Clinical significance of SF3B1 mutations in myelodysplastic syndromes and myelodysplastic/myeloproliferative neoplasms. Blood 2011; 118:6239-46.
4. Biankin AV, Waddell N, Kassahn KS, et al. Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. Nature 2012; 491:399-405.
5. Cancer Genome Atlas N. Comprehensive molecular portraits of human breast tumours. Nature 2012; 490:61-70.
6. Harbour JW, Roberson ED, Anbunathan H, et al. Recurrent mutations at codon 625 of the splicing factor SF3B1 in uveal melanoma. Nat Genet 2013; 45:133-5.
7. Wang L, Lawrence MS, Wan Y, et al. SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. N Engl J Med 2011; 365:2497-506.
8. Furney SJ, Pedersen M, Gentien D, et al. SF3B1 mutations are associated with alternative splicing in uveal melanoma. Cancer Discov 2013; 3:1122-9.
9. Quesada V, Ramsay AJ, Lopez-Otin C. Chronic lymphocytic leukemia with SF3B1 mutation. N Engl J Med 2012; 366:2530.
10. Ellis MJ, Ding L, Shen D, et al. Whole-genome analysis informs breast cancer response to aromatase inhibition. Nature 2012; 486:353-60.

11. 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; 19:6796-802.
12. Kramer A, Gruter P, Groning K, Kastner B. Combined biochemical and electron microscopic analyses reveal the architecture of the mammalian U2 snRNP. J Cell Biol 1999; 145:1355-68.
13. 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; 20:4536-46.
14. 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; 21:4978-88.
15. Yokoi A, Kotake Y, Takahashi K, et al. Biological validation that SF3b is a target of the antitumor macrolide pladienolide. FEBS J 2011; 278:4870-80.
16. Kotake Y, Sagane K, Owa T, et al. Splicing factor SF3b as a target of the antitumor natural product pladienolide. Nat Chem Biol 2007; 3:570-5.
17. 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; 224:517-28.
18. 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; 53:3229-32.
19. 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; 57:3486-93.

20. Ringner M, Fredlund E, Hakkinen J, Borg A, Staaf J. GOBO: gene expression-based outcome for breast cancer online. PLoS One 2011; 6:e17911.

21. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. Nature 2000; 406:747-52.

22. 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; 351:2817-26.

23. 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; 381:805-16.

24. 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.

25. Osborne CK, Schiff R. Mechanisms of endocrine resistance in breast cancer. Annu Rev Med 2011; 62:233-47.

26. Rymond B. Targeting the spliceosome. Nat Chem Biol 2007; 3:533-5.

27. Bonnal S, Vigevani L, Valcarcel J. The spliceosome as a target of novel antitumour drugs. Nat Rev Drug Discov 2012; 11:847-59.

28. Fan L, Lagisetti C, Edwards CC, Webb TR, Potter PM. Sudemycins, novel small molecule analogues of FR901464, induce alternative gene splicing. ACS Chem Biol 2011; 6:582-9.

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; 8:2308-18.
30. Lagisetti C, Pourpak A, Goronga T, et al. Synthetic mRNA splicing modulator compounds with in vivo antitumor activity. J Med Chem 2009; 52:6979-90.
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) splicesome inhibitor administered intravenously on days 1 and 8 every 21 days to patients with solid tumors. Invest New Drugs 2013.
32. Dehm SM. Test-firing ammunition for spliceosome inhibition in cancer. Clin Cancer Res 2013; 19:6064-6.

Figure Legends

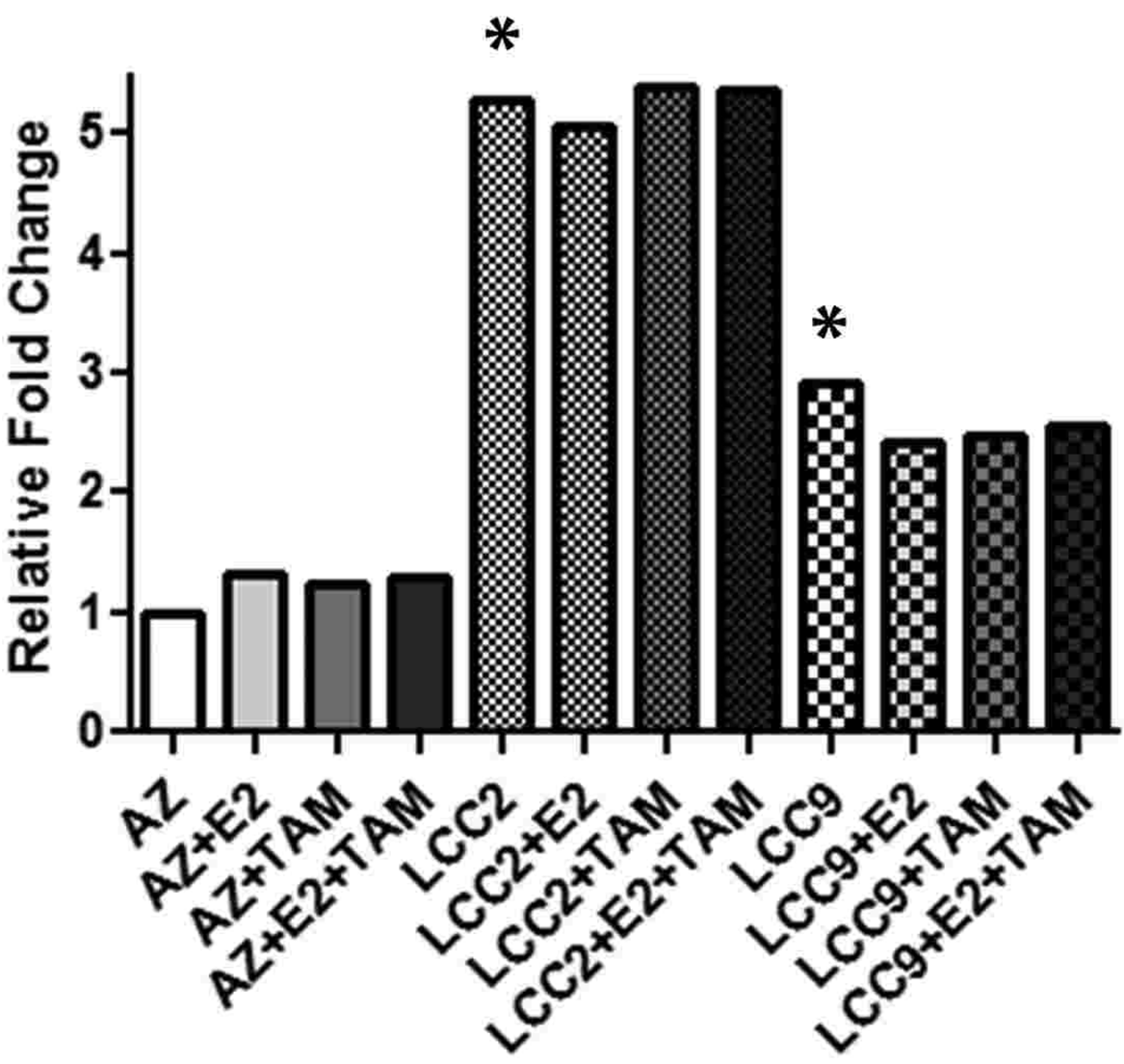
Figure 1. Expression levels of *SF3B1* and *SF3B3* in tamoxifen-resistant LCC2 (MCF7/LCC2), fulvestrant- and tamoxifen cross-resistant LCC9 (MCF7/LCC2), and control sensitive cell line (MCF-7/AZ) using qRT-PCR assay; Vehicle (ethanol), E2 (β-estradiol:10^{-10} M), Tamoxifen (4-OH-Tam:10^{-6} M), or in combination with E2 and Tamoxifen. **P*<0.0005; statistically significant.

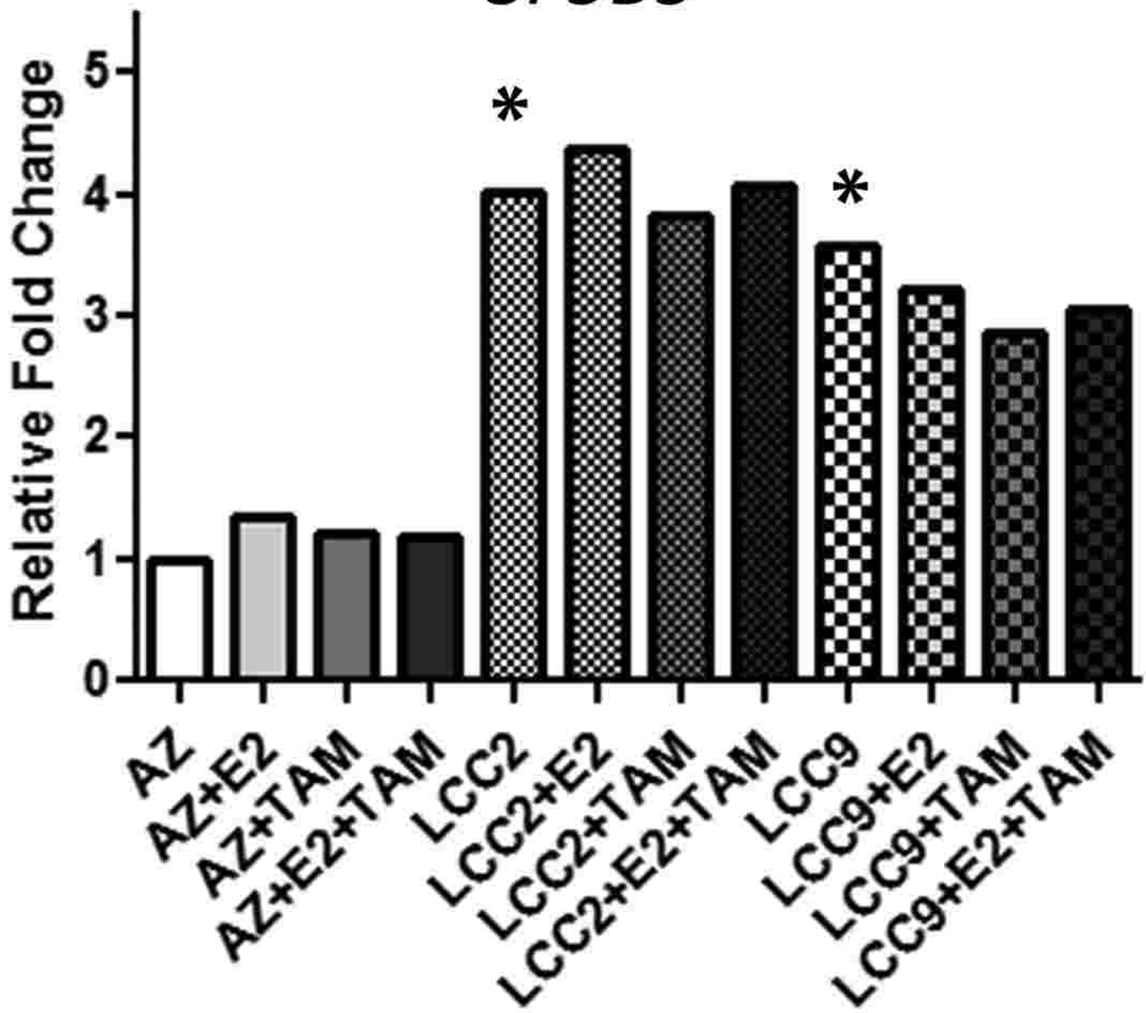
Figure 2. Expression of *SF3B1* and *SF3B3* in patients with ER-positive tumors using qRT-PCR assay. The expression of *SF3B1* and *SF3B3* is higher in patients with high recurrence scores; LS=low score, IS=intermediate score and HS=high score in Onco*type* DX. **P*=0.019; statistically significant.

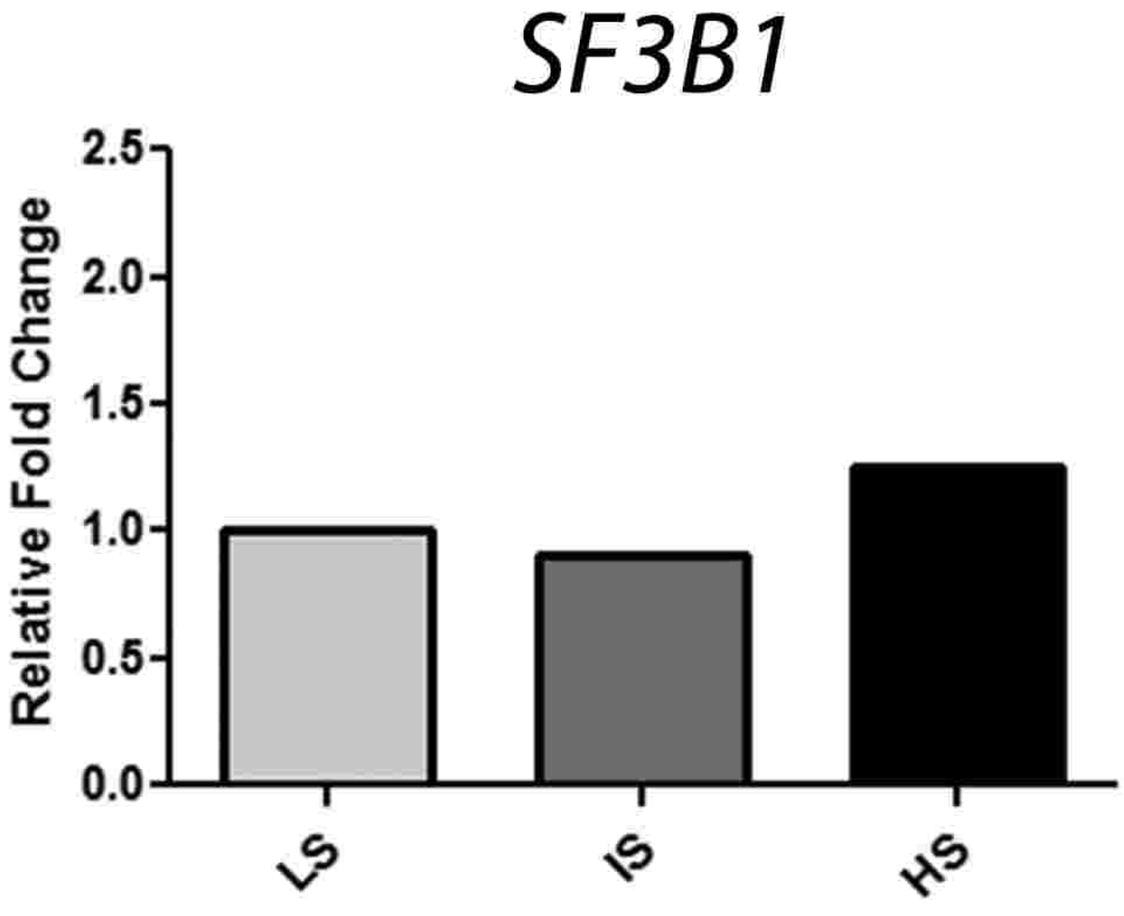
Figure 3. *SF3B3* expression levels using The Cancer Genome Atlas dataset and Affymetrix datasets using GOBO online analysis. A) PAM50 subtypes in The Cancer Genome Atlas, B) ER status (GOBO), and C) PAM50 subtypes (GOBO).

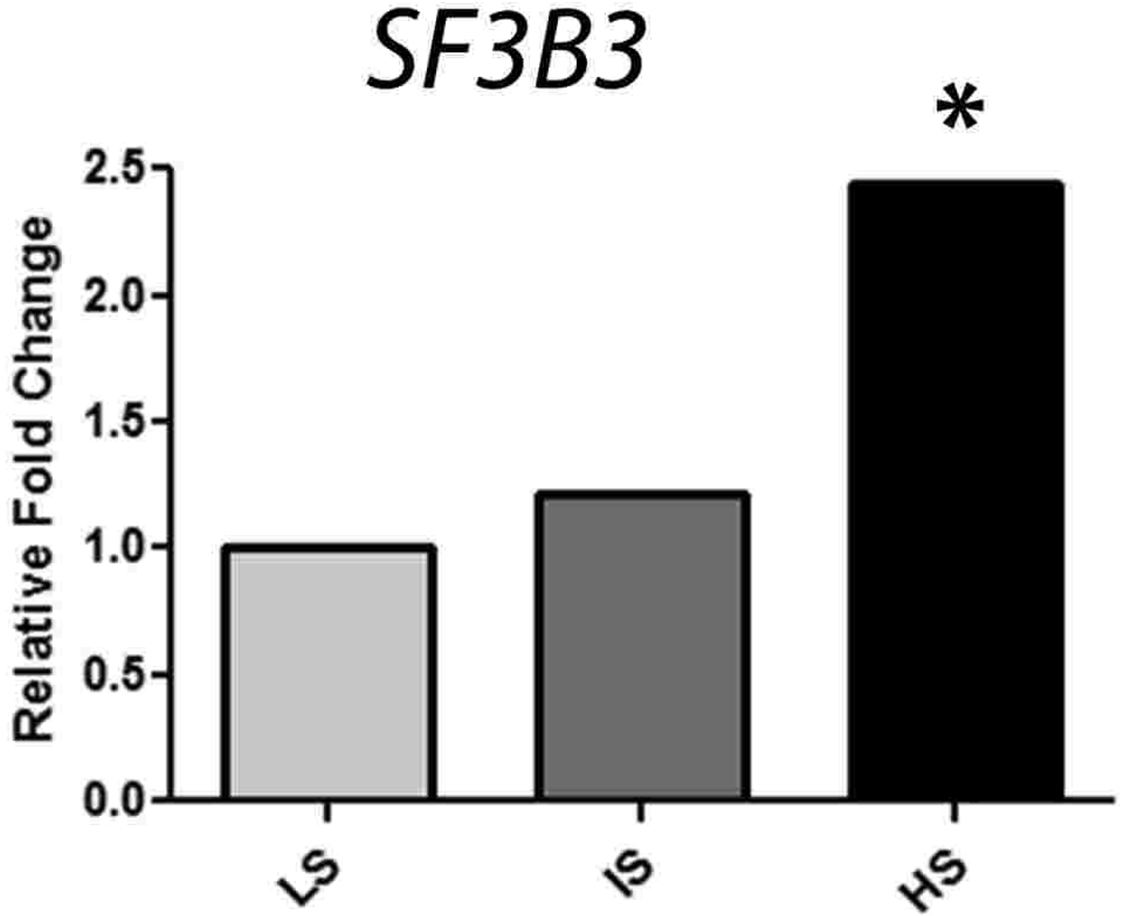
Figure 4. Kaplan-Meier analysis using The Cancer Genome Atlas dataset and Affymetrix datasets using GOBO online analysis. A) Overall survival in The Cancer Genome Atlas (months); B) 10-year overall survival (overall survival) as endpoint for all tumors, ER-positive tumors, and ER-negative tumors; and C) 10-year recurrence-free survival as endpoint for all tumors, ER-positive tumors, ER-negative tumors and ER-positive tamoxifen (Tam)-treated tumors. The expression analysis was stratified into three quantiles based on *SF3B3* expression; low expression (black), medium expression (light grey), and high expression (dark grey) expression.

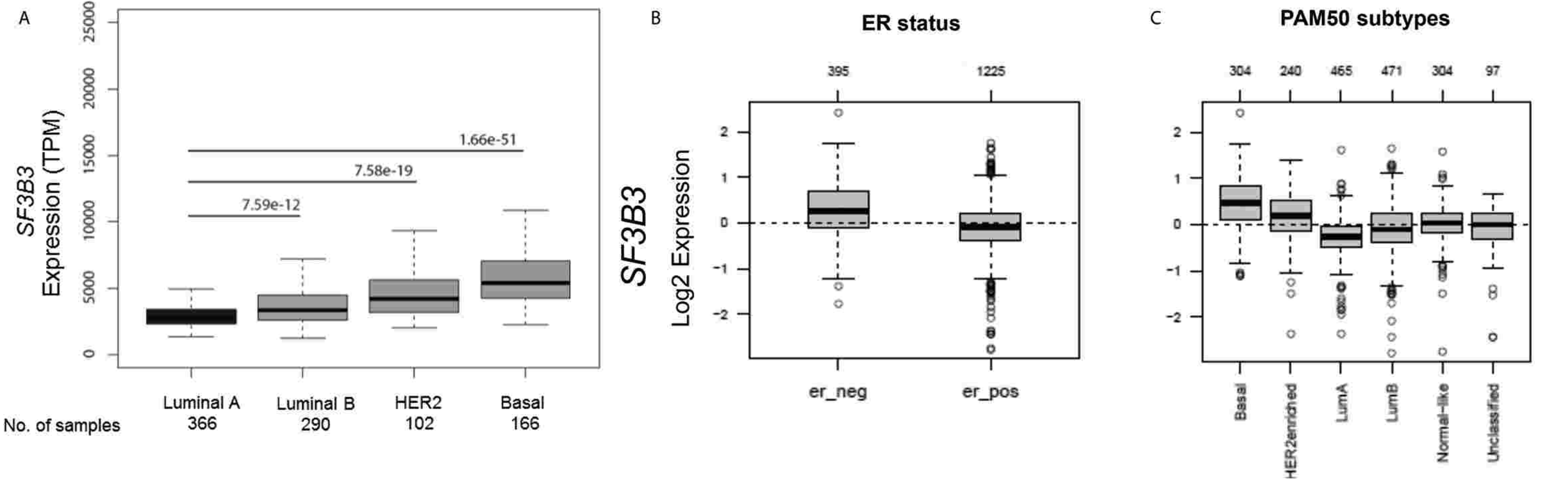
458
459


SF3B1
Relative Fold Change
0
1
2
3
4
5
*
*
AZ
AZ+E2
AZ+TAM
AZ+E2+TAM
LCC2
LCC2+E2
LCC2+TAM
LCC2+E2+TAM
LCC9
LCC9+E2
LCC9+TAM
LCC9+E2+TAM

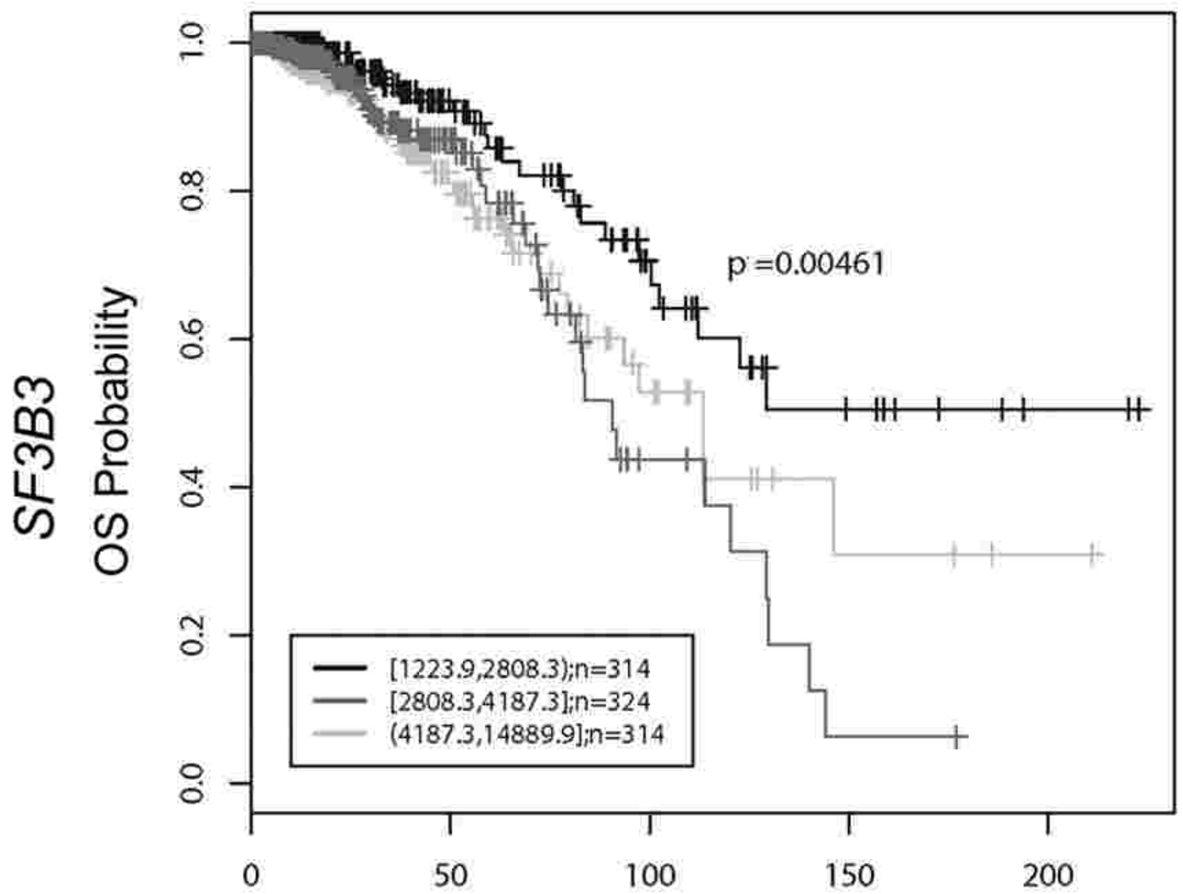

SF3B3
Relative Fold Change
0
1
2
3
4
5
*
*
AZ
AZ+E2
AZ+TAM
AZ+E2+TAM
LCC2
LCC2+E2
LCC2+TAM
LCC2+E2+TAM
LCC9
LCC9+E2
LCC9+TAM
LCC9+E2+TAM


SF3B1
Relative Fold Change
2.5
2.0
1.5
1.0
0.5
0.0
LS
IS
HS

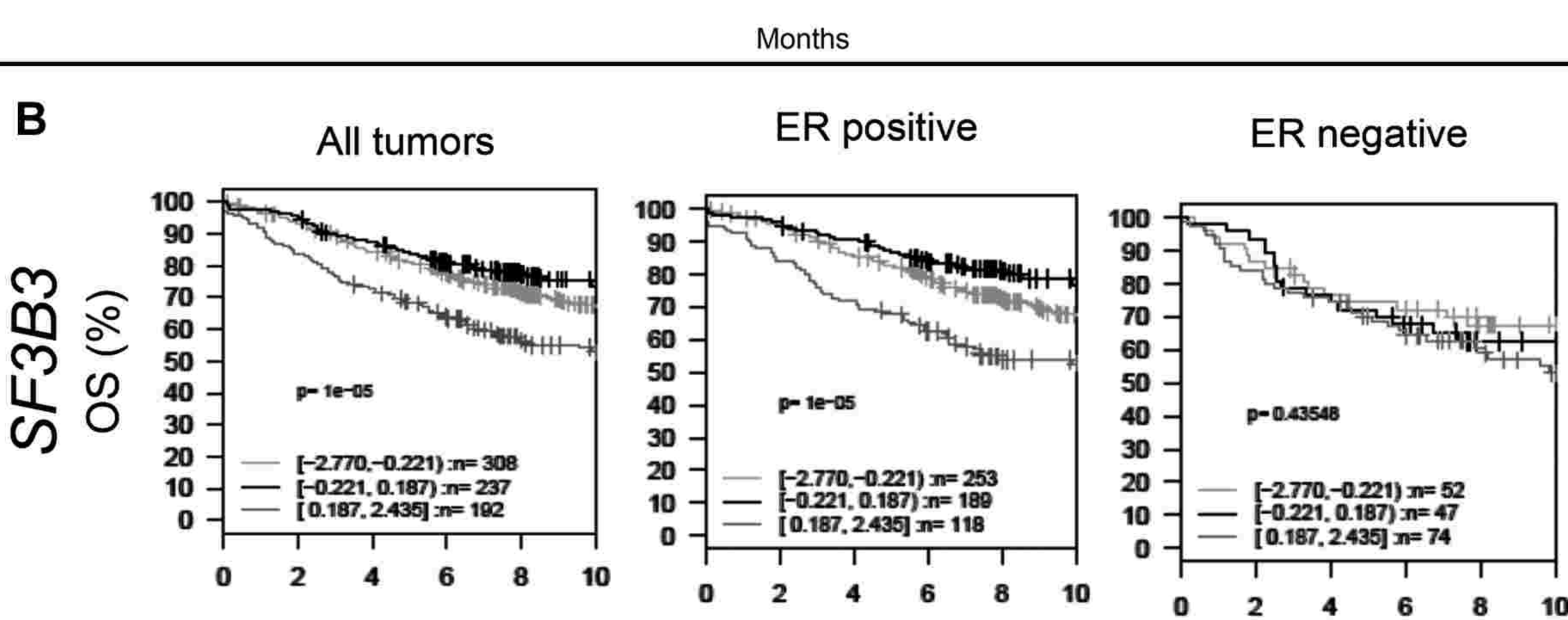

SF3B3
*
Relative Fold Change
2.5
2.0
1.5
1.0
0.5
0.0
LS
IS
HS


A
SF3B3
Expression (TPM)
0
5000
10000
15000
20000
25000
1.66e-51
7.58e-19
7.59e-12
Luminal A
Luminal B
HER2
Basal
No. of samples
366
290
102
166
B
ER status
SF3B3
Log2 Expression
395
1225
er_neg
er_pos
C
PAM50 subtypes
304
240
465
471
304
97
Basal
HER2enriched
LumA
LumB
Normal-like
Unclassified

A


SF3B3
OS Probability
p =0.00461
[1223.9,2808.3);n=314
[2808.3,4187.3];n=324
(4187.3,14889.9];n=314
0.0
0.2
0.4
0.6
0.8
1.0
0
50
100
150
200
Months

B


SF3B3
OS (%)
All tumors
p= 1e-05
[−2.770,−0.221) :n= 308
[−0.221, 0.187) :n= 237
[0.187, 2.435] :n= 192
ER positive
p= 1e-05
[−2.770,−0.221) :n= 253
[−0.221, 0.187) :n= 189
[0.187, 2.435] :n= 118
ER negative
p= 0.43548
[−2.770,−0.221) :n= 52
[−0.221, 0.187) :n= 47
[0.187, 2.435] :n= 74
0
2
4
6
8
10
Years

C

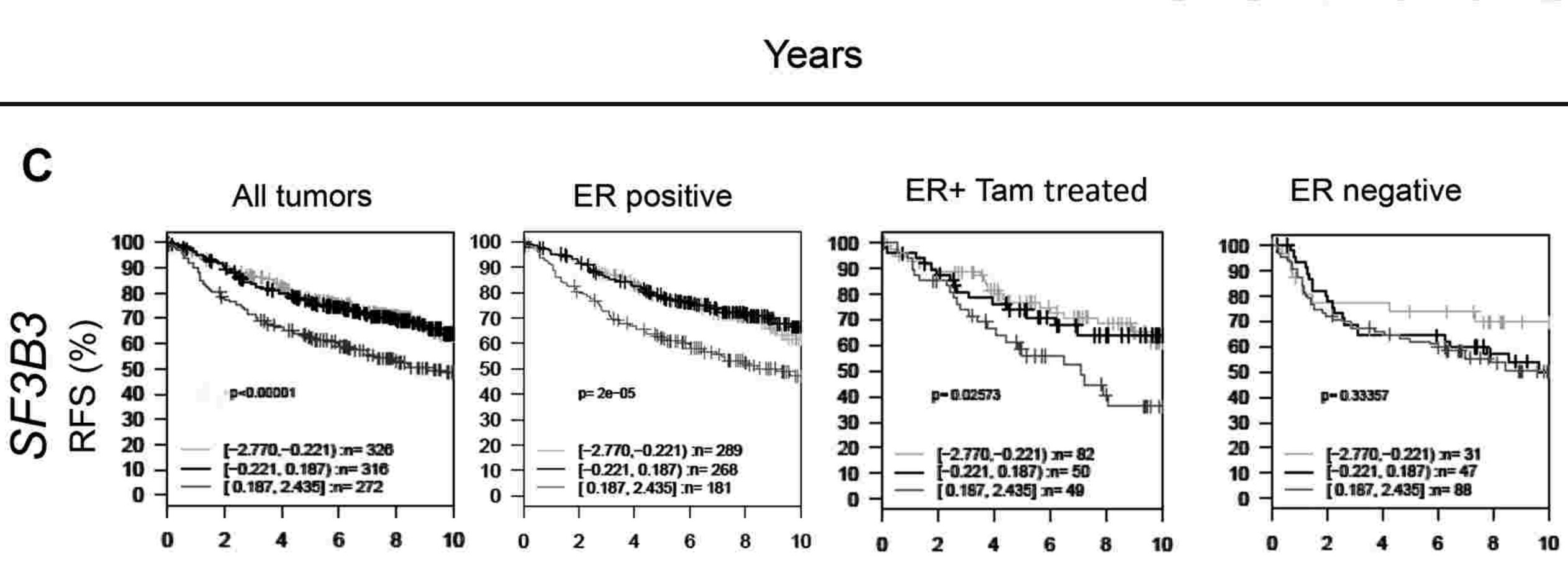

SF3B3
RFS (%)
All tumors
p<0.00001
[−2.770,−0.221) :n= 326
[−0.221, 0.187) :n= 316
[0.187, 2.435] :n= 272
ER positive
p= 2e-05
[−2.770,−0.221) :n= 289
[−0.221, 0.187) :n= 268
[0.187, 2.435] :n= 181
ER+ Tam treated
p= 0.02573
[−2.770,−0.221) :n= 82
[−0.221, 0.187) :n= 50
[0.187, 2.435] :n= 49
ER negative
p= 0.33357
[−2.770,−0.221) :n= 31
[−0.221, 0.187) :n= 47
[0.187, 2.435] :n= 88
0
2
4
6
8
10
Years

Table 1. Multivariable Analysis in ER+ patients for overall survival (Affymetrix datasets-GOBO)

Variable	**HR (95% CI)**	***P* value***
Low *SF3B3* versus high *SF3B3*	0.6 (0.41-0.89)	0.011
LN- versus LN+	0.46 (0.33-0.64)	3.838e-06
Grade 3 versus Grade1-Grade 2	1.3 (0.89-1.88)	0.7196
Age ≥50yrs versus <50yrs	1.58 (1.11-2.24)	0.0117
Tumor size ≥20mm versus <20mm	1.94 (1.38-2.72)	0.0001

HR, hazard ratio; CI, confidence interval.

**P≤0.05 is considered as statistically significant.*

Table 2. Multivariable Analysis in ER+ patients for relapse-free survival (Affymetrix datasets-GOBO)

Variable	**HR (95% CI)**	***P* value***
Low *SF3B3* versus high *SF3B3*	0.62 (0.45-0.86)	0.0046
LN- versus LN+	0.72 (0.54-0.96)	0.0266
Grade 3 versus Grade1-Grade 2	1.09 (0.79-1.5)	0.6127
Age ≥50yrs versus <50yrs	0.81 (0.61-1.09)	0.1634
Tumor size ≥20mm versus <20mm	1.92 (1.45-2.54)	5.388e-06

HR, hazard ratio; CI, confidence interval.

**P≤0.05 is considered as statistically significant.*