

1
2
3
4 **TFAP2C Expression in Breast Cancer - Correlation with Overall Survival Beyond**
5
6 **10 Years of Initial Diagnosis**
7
8
9

10
11 Susan M. Perkins¹, Casey Bales², Tudor Vladislav³, Sandra Althouse¹, Kathy D. Miller²,
12
13 George Sandusky³, Sunil Badve³, and Harikrishna Nakshatri^{4,5,6*}
14
15
16
17

18
19 Departments of Biostatistics¹, Medicine², Pathology and Laboratory Medicine³, Surgery⁴,
20
21 Biochemistry and Molecular Biology⁵, Indiana University School of Medicine,
22
23 Indianapolis, Indiana, USA 46202
24

25
26 Richard L. Roudebush VA Medical Center⁶, Indianapolis, Indiana, USA 46202
27
28
29
30

31 Running title: TFAP2C expression in breast cancer.
32
33
34
35

36 Keywords: TFAP2C, estrogen receptor, breast cancer, outcome
37
38
39
40

41 *Corresponding Author: Harikrishna Nakshatri, BVSc., PhD
42

43 C218C, 980 West Walnut St.
44

45 Indianapolis, IN, 46202
46

47 317 278 2238
48

49 317 274 0396 (fax)
50

51 hnakshat@iupui.edu
52
53
54
55
56

57
58 Authors have no conflict of interest to declare.
59
60

61 This is the author's manuscript of the article published in final edited form as:
62 Perkins, S. M., Bales, C., Vladislav, T., Althouse, S., Miller, K. D., Sandusky, G., ... Nakshatri, H. (2015).
63 TFAP2C expression in breast cancer: correlation with overall survival beyond 10 years of initial diagnosis.
64 Breast Cancer Research and Treatment, 152(3), 519–531. <http://doi.org/10.1007/s10549-015-3492-2>
65

1
2
3
4 **Abstract:**
5

6 **Purpose:** Recurrence and death in a significant number of patients with ER α -positive
7 breast cancer occurs 10-20 years after diagnosis. Prognostic markers for late events have
8 been more elusive. TFAP2C (AP2 γ) regulates the expression of ER α , the ER α pioneer
9 factors FOXA1 and GATA3, and controls ER α -dependent transcription. The purpose of
10 this investigation is to determine the long-term prognostic value of TFAP2C.
11
12
13
14
15
16
17
18

19 **Material and Methods:** A tissue microarray (TMA) consisting of breast tumors from
20 451 patients with median follow-up time of 10.3 years was created and tested for the
21 expression of TFAP2C by immunohistochemistry. Wilcoxon Rank-Sum and Kruskal-
22 Wallis tests were used to determine if TFAP2C H-scores correlate with other tumor
23 markers. Cox proportional hazards regression models were used to determine whether
24 TFAP2C H-scores and other tumor markers were related to overall and disease-free
25 survival in univariate and multivariable models.
26
27
28
29
30
31
32
33
34
35

36 **Results:** TFAP2C overexpression did not impact overall survival during the first 10 years
37 after diagnosis, but was associated with a shorter survival after 10 years (HR 3.40, 95%
38 CI: 1.58, 7.30; p-value=0.002). This late divergence persisted in ER-positive (HR 2.86,
39 95% CI: 1.29, 6.36; p-value=0.01) and endocrine therapy-positive subgroups (HR 4.19,
40 95% CI: 1.72, 10.23; p-value=0.002). For the ER+ and endocrine therapy subgroup, the
41 HR was 3.82 (95% CI: 1.53, 9.50; p-value=0.004). TFAP2C H-scores were not
42 correlated with other tumor markers or related to disease-free survival.
43
44
45
46
47
48
49
50
51
52

53 **Conclusion:** In this hypothesis-generating study, we show that higher TFAP2C scores
54 correlate with poor overall survival after 10 years of diagnosis in ER α -positive and
55 endocrine therapy-treated subgroups.
56
57
58
59
60
61
62
63
64
65

1
2
3
4 **Introduction:**

5 Breast cancers that express estrogen receptor alpha (ER α) are dependent on
6 estrogen (E2) for growth and are responsive to anti-estrogen treatment [1]. The
7
8 E2:ER α signaling network plays a role in breast cancer initiation and proliferation;
9
10 however, that same network potentially prevents metastatic progression [2, 3]. 30-40%
11
12 of patients develop resistance to anti-estrogen treatment and these resistant tumors
13
14 preferentially metastasize to bone [4, 5].
15
16
17
18
19

20 RNA-based intrinsic subtype classification and genome-aberration based
21
22 integrative cluster analysis have subclassified ER α -positive breast cancers into multiple
23
24 subtypes each with differing outcome [6, 7]. Thus, ER α -mediated signaling is not
25
26 uniform across all ER α -positive breast cancers and tumor-specific genomic aberrations
27
28 collude with ER α to confer diversity to ER α signaling and to overcome the effects of
29
30 anti-estrogens. Recent studies have demonstrated that ER α itself undergoes mutations in
31
32 metastatic lesions and these mutants are hypersensitive to E2 and/or resist the actions of
33
34 ER α antagonists [8-10]. Apart from mutation in ER α , which is observed in 15-20% of
35
36 metastatic breast cancers [11], very limited data are available on the other genomic
37
38 aberrations that collaborate with ER α to confer resistance to anti-estrogens.
39
40
41
42
43
44
45

46 Three distinct classes of signaling proteins control the function of ER α and
47
48 aberrations in these proteins may provide an additional avenue for tumors to escape the
49
50 actions of anti-estrogen. 1) Pioneer factors such as FOXA1, GATA3, TLE family
51
52 members, and PBX1 enable the recruitment of ER α to chromatin [12, 13].
53

54 Transcriptional signatures that are dependent on ER α and specific pioneer factors
55
56 identify tumors with differing response to endocrine therapy. For example, we, amongst
57
58
59
60
61
62
63
64
65

1
2
3
4 others, have shown that elevated FOXA1 expression correlates with better response to
5 anti-estrogen treatment [14-16]. However, there is some evidence for deleterious effects
6 of FOXA1 if expressed in metastatic lesions and/or in tumors that have acquired anti-
7 estrogen resistance [17, 18]. ER α -PBX1 dependent signature is associated with poor
8 outcome [19]. 2) Co-activators and co-repressors. Among them, overexpression of AIB1
9 is linked to anti-estrogen resistance and metastatic progression in a subset of breast
10 cancer [20]. 3) Kinases that phosphorylate ER α , which include AKT/PKB, ERK, SRC,
11 PLK, and PKA [21]. However, elevated activity of these kinases is not always associated
12 with poor outcome; elevated nuclear phospho-AKT(S473) correlates with better outcome
13 in ER α -positive breast cancer [22]. There are conflicting reports on the role of
14 ER α :PLK1 axis in anti-estrogen response and outcome [23, 24].

15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32 The focus of this study was on the pioneer factor TFAP2C, which regulates ER α
33 activity at multiple levels. First, TFAP2C is required for the expression of ER α and its
34 pioneer factors FOXA1 and GATA3 [25, 26]. Second, it is an E2-inducible gene [27].
35 Third, it is required for long-distance chromatin interaction and ER α -regulated gene
36 expression including the genes that characterize ER α -positive breast cancer [28].
37 Regulatory regions of genes that are dependent on chromatin looping for expression are
38 typically enriched for binding sites for ER α , FOXA1, and TFAP2C. Fourth, it controls
39 luminal cell-specific unliganded ER α cistrome [29]. Fifth, it governs luminal epithelial
40 phenotype of the breast by suppressing epithelial to mesenchymal transition [25]. ER α is
41 expressed predominantly in luminal cells. Thus, TFAP2C is an integral part of the
42 ER α :E2-mediated signaling in luminal breast epithelial cells.

1
2
3
4 Prognostic value of TFAP2C in breast cancer remains controversial. By analyzing
5
6
7 75 breast tumor tissues, Gee et al reported poor outcome and anti-estrogen resistance in
8
9 patients with TFAP2C overexpressing tumors [30]. Similar analysis of 54 tamoxifen-
10
11 sensitive and 38 tamoxifen-resistant tumors revealed a role of elevated TFAP2C in
12
13 tamoxifen resistance [31]. Friedrichs et al reported no prognostic value for TFAP2C by
14
15 analyzing 70 cases [32]. In animal models, TFAP2C has been shown to reduce ERBB2-
16
17 mediated mammary tumor incidence but promote tumor progression [33]. Given the
18
19 conflicting data, we sought to explore the prognostic value of TFAP2C in a large and
20
21 well-annotated tissue microarray with long-term follow-up. We were particularly
22
23 interested in the impact of TFAP2C on outcome 10 or more years after diagnosis given
24
25 the paucity of biomarkers that can be used for predicting late recurrence and death [34,
26
27
28
29
30
31 35].
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 **Materials and Methods:** Remark guidelines for tumor biomarker reporting studies were
5
6 followed in this study [36].
7

8
9 **Patients, Specimen Characteristics and TMA construction:** The TMA is composed of
10 duplicate cores from 460 cases of breast cancer in 453 patients. Patients with duplicate
11 cases, seven instances total, include two women with multiple primaries, four with
12 lumpectomy and mastectomy tissue, and one with bilateral/synchronous disease. These
13 samples were obtained from archival cases at Indiana University Health (University and
14 Methodist Hospitals) between 1989 and 2003. Patient medical records were reviewed for
15 demographics, pathology, treatment, and recurrence/follow-up information and are
16 described in Table 1. For this study, we excluded males (n=2). The Indiana University
17 IRB approved this project.
18
19
20
21
22
23
24
25
26
27
28
29
30

31 **Assay Methods/Immunostaining of TMA:** Antibody against TFAP2C was purchased
32 from Epitomics-Abcam (Cat#EP2692Y, Eugene, USA). The slides were deparaffinized
33 in xylene and rehydrated through graded alcohols to water. Antigen retrieval was
34 performed by immersing the slides in DAKO's EnVision FLEX low pH target retrieval
35 solution for 20 min. @ 100 degrees C., cooling to 85 degrees and placed in Dako wash
36 buffer. A 1:250 dilution of the primary antibody was added to the slides and incubated
37 20 minutes at room temperature. Following washing in TBST, visual detection was
38 performed with DAKO Envison Plus Polymer Link and DAB chromogen (DAKO). The
39 slides were washed and cover slipped. Control sections were treated with an isotype
40 control using the same concentration as primary antibody to verify the staining
41 specificity.
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 **Whole slide Digital Imaging:** The Aperio whole slide digital imaging system was used
5
6 for imaging. The Aperio Scan Scope CS system was used (360 Park Center Drive, Vista,
7
8 CA 92081). The system imaged all slides at 20x. The scan time ranged from 1½ minute
9
10 to a maximum time of 2.25 minutes. The whole images were housed and stored in their
11
12 Spectrum software system and images were shot from the whole slides.
13
14

15
16 **Automatic Image Quantitation:** The Positive Pixel Count algorithm was used to
17
18 quantify the amount of a specific stain present in a scanned slide image. A range of color
19
20 (range of hues and saturation) and three intensity ranges (weak, positive, and strong)
21
22 were masked and evaluated. The algorithm counted the number and intensity-sum in
23
24 each intensity range, along with three additional quantities: average intensity, ratio of
25
26 strong/total number, and average intensity of weak positive pixels.
27
28

29
30 The algorithm had a set of default input parameters when first selected—these inputs
31
32 have been pre-configured for Brown color quantification in the three intensity ranges
33
34 (220-175, 175-100, and 100-0). Pixels, which were stained, but did not fall into the
35
36 positive-color specification, were considered negative stained pixels—these pixels were
37
38 counted as well, so that the fraction of positive to total stained pixels was determined.
39
40

41
42 The algorithm was applied to an image by using ImageScope TMA Lab. This program
43
44 allowed us to cover each tissue core in the array, specify the input parameters, run the
45
46 algorithm, and view the results. Using the Aperio Positive Pixel algorithm, only staining
47
48 classified as “positive” or “strong positive” was used to calculate positivity; regions
49
50 classified as “weak positive” were mostly cytoplasmic and background staining, and were
51
52 not counted.
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 **Statistical Methods:** REMARK guidelines were followed for reporting study results
5
6
7 [37]. For the 7 subjects with multiple tumor samples available, we included only the
8
9 sample with the highest TFAP2C H-score. T-tests, Chi-square tests, Fisher's Exact tests,
10
11 or log-rank tests were used to compare patient and tumor variables between those with
12
13 TFAP2C H-scores vs. those without. Wilcoxon Rank Sum and Kruskal-Wallis tests were
14
15 used to determine if TFAP2C H-scores were correlated with other tumor markers. Cox
16
17 proportional hazards regression models were used to determine whether TFAP2C H-
18
19 scores and other variables were related to overall and disease-free survival either
20
21 univariately or in multivariable models. In these analyses, TFAP2C H-scores were
22
23 divided into low and high categories at the score of 15.9858 for overall survival (time
24
25 from surgery to death or censoring) and a score of 12.952 for disease-free survival (time
26
27 from surgery to first recurrence or censoring, excluding patients with M1 stage at
28
29 surgery). These cutoff values were determined by using the maximum chi-square value
30
31 for all score values between the 25th and 75th percentile [38]. TFAP2C high/low was
32
33 included in all multivariable models. Other tumor markers were included in multivariable
34
35 models only if significant using a α level of 0.05 in the univariate models and to preserve
36
37 the overall sample size we included missing as a separate category for PR Status, Tumor
38
39 Grade, and Tumor Stage. To assess non-proportional hazards, the AIC criterion was used
40
41 to compare non-nested models. Using this criterion, the cutoff for choosing Model 1 over
42
43 Model 2 is that the AIC for Model 1 is lower than that of Model 2 by at least 2.0. Overall
44
45 and disease-free survival curves were generated for *a priori* specified subgroups (by ER
46
47 status and endocrine therapy use) using the Kaplan-Meier Method and compared using
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 log-rank tests. Analyses were conducted using SAS Version 9.4. A α level of 5% was
5
6 used to determine statistical significance.
7
8

9 To further extend observations made at protein levels to mRNA levels, we
10
11 determined prognostic relevance of TFAP2C on overall and metastatic-free survival in
12
13 various publicly available gene expression array datasets using our recently developed
14
15 prognostic database PROGgene V2 [39]. The cBioportal database was used to determine
16
17 genomic aberration in TFAP2C in various cancers [40].
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 **Results:** Of the 451 subjects in the study, 385 (85%) had TFAP2C values available.
5
6 Clinical parameters of the subjects included in the study are summarized in Table 1 with
7
8 subgrouping of TFAP2C-positive and TFAP2C-negative cases. There were no
9
10 differences in the characteristics of the patients or their tumors based on whether or not
11
12 TFAP2C values were available. Median (range) follow-up from time of surgery was 10.3
13
14 years (0.04-21.8) for those with TFAP2C values and 10.3 years (1.1-17.5) for those
15
16 without (p = 0.2421).
17
18
19
20

21 In the TFAP2C-positive group, there were 126 recurrences (33%) and 163 deaths
22
23 (42%) during follow-up. The large number of HER-2/neu Unknown/Not Done was due
24
25 to HER-2/neu testing not being fully integrated into institutional standard of care until
26
27 April of 2001. There were no Unknown/Not Done values on or after that date. Also of
28
29 note, 62/284 (22%) of TFAP2C-positive ER positive subjects did not receive endocrine
30
31 therapy. The majority of these women (n= 46, 74%) were diagnosed prior to 2000 when
32
33 ER positive disease wasn't well understood and aromatase inhibitors were unproven [41]
34
35 Tamoxifen was the only available treatment and often not prescribed to elderly women,
36
37 those with small T1 tumors, or with other contraindication. Additionally, for
38
39 premenopausal women diagnosed prior to 1998, tamoxifen was incorrectly thought not to
40
41 be efficacious [42]. Specific reasons for lack of endocrine therapy in ER positive group
42
43 in this study include: patient refusal (n=3, 5%); comorbid condition (n=4, 6%); T1
44
45 tumor (n=17; 28%) age 70+ with or without T1 tumor or comorbidity (n=23, 37%);
46
47 weakly staining tumor (n=4, 6%); premenopausal (n=4, 6%); and unknown (n=7,
48
49 11%). Endocrine therapy was primarily adjuvant treatment; only 4% had known
50
51 metastatic disease.
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 **TFAP2C staining pattern in breast cancer:**
5

6 Representative TFAP2C immunostaining of breast tumors is shown in Figure 1. The
7 expression ranged from no expression to high expression and staining was restricted to
8 nucleus. Interestingly, expression in normal breast was mosaic and staining intensity was
9 similar to that of high expressing tumors. Thus, it is likely that tumors that express higher
10 levels of TFAP2C have emerged from cells that naturally express higher levels of
11 TFAP2C.
12
13
14
15
16
17
18
19
20

21 **Correlation of TFAP2C H-score with other tumor markers**
22

23 We generated TFAP2C H-score as described in the materials and methods and compared
24 expression with ER, PR, HER-2/neu, Nodal stage or Grade. TFAP2C levels were not
25 correlated with ER, PR, HER-2/neu, Nodal Stage, or Grade (Table 2).
26
27
28
29
30

31 **Overall Survival Analysis**
32

33 *Univariate:* In univariate analyses, variables significantly related to overall survival in the
34 Cox proportional hazards regression models were PR Status, Tumor Grade, Tumor Stage,
35 and Nodal Stage (Table 3). PR-, Higher Tumor Grade, Higher Tumor Stage, and Nodal
36 Stage-positive were correlated with lower survival. TFAP2C H-score was not related to
37 overall survival (Figure 2A, log rank test p-value 0.107); however examination of the
38 survival curve revealed that the survival curves for high and low TFAP2C were not
39 parallel, indicating non-proportional hazards. The curves diverged at the 10-year mark.
40 Cox models were compared that allowed for different hazard ratios for the time periods
41 before or at 10 years and after 10 years verses proportional hazards using the AIC
42 criterion, where it was found non-proportional hazard ratio provided a better fit. The
43 results from these models (both univariate and multivariable adjusting for significant
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 other tumor markers) indicated that the hazard ratio prior to or at 10 years was non-
5
6 significant, but that the hazard ratio greater than 10 years was significantly different
7
8 between the TFAP2C H-score categories. See Table 4 for multivariable model results.
9
10 After 10 years, the hazard ratio was 3.40 (95% CI 1.58, 7.30) when comparing High vs.
11
12 Low TFAP2C H-score categories.
13
14

15
16 *Subgroup analysis:* We conducted subgroup analyses on overall survival using the ER-
17
18 positive subgroup, endocrine therapy group, ER-positive on endocrine therapy, ER-
19
20 negative, and ER+/PR+/HER2- (Figure 2B-2F). For ER-positive, endocrine therapy, and
21
22 ER-positive on endocrine therapy, the results were similar to the overall results. Whether
23
24 adjusted for covariates or not, hazard ratios were not significantly different from one
25
26 prior to or at 10 years and were greater than one after 10 years. For ER-positive,
27
28 HR=2.86 (95% CI 1.29, 6.36). For endocrine therapy HR=4.19 (95% CI 1.72, 10.23), and
29
30 for ER-positive on endocrine therapy, HR=3.82 (95% CI 1.53, 9.50). Results for ER-
31
32 negative and ER+/PR+/HER2- were not significant.
33
34
35
36
37

38 **Disease Free Survival (DFS) Analysis**

39
40 *Univariate:* In univariate analyses, variables significantly related to disease free survival
41
42 in the Cox proportional hazards regression models were Tumor Grade, Tumor Stage, and
43
44 Nodal Stage (Table S1). Higher Tumor Grade, Higher Tumor Stage, and Nodal Stage-
45
46 positive were correlated with lower survival. TFAP2C H-score was not related to
47
48 disease-free survival (log rank test p-value 0.2481, Figure S1A).
49
50

51
52 *Multivariable:* In the multivariable analysis, Tumor Grade, Tumor Stage and Nodal
53
54 Status were independently related to disease free survival. Higher Tumor Grade and
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 higher or unknown Tumor Stage were correlated with lower disease free survival (Table
5
6
7 S2).

8
9 *Subgroup analysis:* We conducted subgroup analyses on disease free survival using the
10
11 ER-positive subgroup, endocrine therapy group, ER-positive on endocrine therapy, ER-
12
13 negative, and ER+/PR+/HER2- (Figure S1B to S1F) but the results were not statistically
14
15 significant (log rank test p-value 0.1097 for ER+, 0.1111 for those on endocrine therapy,
16
17 0.0707 for ER-positive on endocrine therapy, 0.8465 for ER-negative, and 0.1063 for
18
19 ER+/PR+/HER2-).
20
21
22

23 **Prognostic value of TFAP2C at mRNA level:**

24
25
26 Results of four mRNA analyses from public data are presented in Figure 3A-D. In the
27
28 gene expression array dataset described by Clarke et al., elevated expression of TFAP2C
29
30 in grade 3 tumors was associated with poor overall survival [43]. Similarly, in Miller et al
31
32 dataset, elevated expression of TFAP2C in PR-negative tumors is associated with poor
33
34 overall survival [44]. In the node-negative breast cancer dataset by Schmidt et al.,
35
36 TFAP2C overexpression was associated with poor metastasis-free survival and in the Loi
37
38 et al dataset, TFAP2C overexpression was associated with poor metastasis-free survival
39
40
41 [45, 46].
42
43
44

45
46 cBioportal database analysis revealed TFAP2C amplification in a variety of
47
48 cancers including 8% of breast cancer [40]. Curiously, 28% of primary breast cancers
49
50 that were adapted grow as xenografts showed amplification (Figure 3E). Thus, it appears
51
52 that TFAP2C-overexpressing tumor cells are selected during tumor
53
54 establishment/progression, at least under immunosuppressed condition.
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 **Discussion:**
5

6 Although ER α -positive breast cancer is associated with better outcome compared with
7 ER α -negative breast cancer, fear of recurrence is always in the minds of many patients.
8
9
10 Since recurrence can occur even after 15 years of initial diagnosis, development of
11 predictive biomarkers has proven difficult because of challenges in assimilating follow-
12 up information for a considerably longer time. Having overcome this limitation, we
13
14 present the role of TFAP2C in predicting recurrence and overall survival post-10 years of
15 initial diagnosis. We saw no relationship between TFAP2C and disease recurrence. For
16 overall survival, we saw no relationship between TFAP2C prior to or at 10 years
17 (essentially negative results). After 10 years, the data suggests that higher TFAP2C was
18 predictive of greater hazard of death. We do note that number of events is small in both
19 High (n=16 deaths) and Low (n=12 deaths) TFAP2C H-Score groups after 10 years of
20 diagnosis. In addition, because of lower number of events after 10-years, we were unable
21 to divide our TMA into training and validation sets. For a simple log-rank test, 21 deaths
22 total would be needed to have 80% power to detect a hazard ratio of 3.4 using a two-
23 sided test and 0.05 level of significance. Using a standard 50/50 split for the training and
24 validation sets, we would only have 14 events per set. Therefore, results presented above
25 should be considered “hypothesis generating” and need to be verified in a large TMA
26 dataset with extensive follow up as we have with our TMA. Few such TMAs have been
27 published, which we hope to analyze in future [14, 47]. Inclusion of proliferation markers
28 such as ki67 may allow us to determine whether TFAP2C expression is correlated with
29 proliferation, and potentially luminal B phenotype, which is planned in future. However,
30 the above limitation of the study is not due to quality of our TMA because in univariate
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 and/or multivariate analysis, markers such as Tumor Grade, Tumor stage and Nodal
5
6 metastasis status were associated with overall survival and/or disease-free survival
7
8 similar to many other published TMAs [16, 48, 49] (Tables 3, 4, S1 and S2).
9

10
11 Amongst ER α pioneer factors examined for prognostic utility, PBX1 mRNA
12
13 overexpression has been previously linked to poor outcome [19]. We had previously
14
15 demonstrated better outcome in all and ER α -positive breast cancer patients with tumors
16
17 that express higher levels of FOXA1 protein [16, 48]. Unlike the results of TFAP2C,
18
19 differences in outcome between high and low FOXA1 expressers were evident as early as
20
21 four-years post diagnosis [16]. However, similar to TFAP2C, FOXA1 had no prognostic
22
23 value in ER α -negative breast cancers. Thus, it is likely that prognostic utility of both
24
25 FOXA1 and TFAP2C is linked to their pioneer factor activity on ER α . This is not
26
27 unexpected based on intricate link in the expression/activity of ER α , TFAP2C, and
28
29 FOXA1 with TFAP2C being the regulator of both ER α and FOXA1 [25]. However,
30
31 downstream of this link, signaling by ER α :FOXA1 and ER α :TFAP2C may differ, which
32
33 contributes to differences in tumor cell behavior. In this respect, TFAP2A, a closely
34
35 related member of TFAP family, opposes the action of TFAP2C [50]. Balance between
36
37 these two family members may ultimately determine the ability of TFAP2C in regulating
38
39 ER α activity. An alternative possibility exists where TFAP2C and FOXA1 cooperate with
40
41 each other in anti-estrogen resistant breast cancers to afford distinct growth signaling
42
43 pathway that does not require estrogen. Recent studies have demonstrated that the
44
45 genomic activity of FOXA1 differs between tamoxifen-sensitive and tamoxifen-resistant
46
47 breast cancer cells suggesting that tamoxifen-resistance mechanism involves switch in the
48
49 activity rather than loss of FOXA1 expression [18, 51]. Since majority of TFAP2C:ER α
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 regulated genes are also enriched for FOXA1 binding sites [28], it will be interesting to
5
6
7 compare TFAP2C:FOXA1:ER α cistrome and transcriptome in anti-estrogen-sensitive
8
9 and anti-estrogen-resistant cells.

10
11 Although chromosome 20q13.2 where TFAP2C is located is amplified in breast
12
13 cancer [52], TFAP2C protein does not appear to be overexpressed in majority of breast
14
15 cancer as we observed mosaic expression in normal breast and the expression levels in
16
17 normal breast appear to be similar to that in cancers with highest TFAP2C H-score. Thus,
18
19 it is likely that most of the cancers with elevated TFAP2C have originated from normal
20
21 cells that naturally express higher levels of TFAP2C. Indeed, Oncomine database
22
23 analysis did not reveal cancer-specific upregulation of TFAP2C except one study
24
25 showing 1.2-fold increase in ductal carcinoma (data not shown). Different outcome
26
27 observed in patients with tumors with higher and lower TFAP2C H-score or TFAP2C
28
29 mRNA levels could be related to how ER α activity differs between TFAP2C-high and
30
31 TFAP2C-low cells. To get insight into this possibility, we did coexpression analysis
32
33 using Oncomine and cBioportal. Both databases revealed a network comprising TFAP2C,
34
35 CITED1, CITED2, CITED4, and UBE2I. Among these genes, CITED1 is a co-activator
36
37 of ER α activity and is required for mammary gland development [53, 54]. Thus,
38
39 ER α :CITED1 transcriptional axis could differ in cells expressing higher levels of
40
41 TFAP2C compared to cells with lower TFAP2C.
42
43
44
45
46
47
48
49
50

51 In vitro studies in cell lines have revealed a role for TFAP2C in inducing the
52
53 expression of proto-oncogene RET independent of ER α [55]. TFAP2C has also been
54
55 shown repress p21 in cooperation with cMyc oncogene [56]. Thus, tumors originating
56
57 from cells with higher TFAP2C may be intrinsically more proliferative than tumors with
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

lower TFAP2C. Tumors with higher proliferative rate usually respond initially to chemo or hormonal therapy but eventually recur. Future studies focusing on association between TFAP2C and proliferation markers such as ki67 may help to establish a link between TFAP2C expression, proliferation, recurrence, and outcome.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Acknowledgements: We thank Janet Harlan and Jeanette Krohne for chart review, Lee Ann Baldrige for immunohistochemistry of TMA. IUPUI Signature center, IUSCC Breast Cancer Program, IUSCC Tissue Bank, and a grant from the Susan G. Komen for the Cure (SAC110025 to HN) supported this study.

Author’s contributions: SMP and SA did statistical analysis and wrote the statistical methods and results sections, CB did clinical annotation of the TMA, TV, GS ad SB did the pathology review, KDM did data analysis, and HN did public database search, and supervised the study. All authors contributed to manuscript writing. All authors edited and approved the manuscript.

Ethical Standards: The study was conducted in compliance with current laws of United States of America.

Conflict of Interest: The authors declare that they have no conflict of interest.

1
2
3
4 **References:**
5
6

- 7 1. Musgrove EA, Sutherland RL: Biological determinants of endocrine
8 resistance in breast cancer. *Nat Rev Cancer* 2009, 9:631-643.
- 9 2. Torres-Arzayus MI, Zhao J, Bronson R, Brown M: Estrogen-dependent
10 and estrogen-independent mechanisms contribute to AIB1-mediated
11 tumor formation. *Cancer Res* 2010, 70:4102-4111.
- 12 3. Hurtado A, Holmes KA, Geistlinger TR, Hutcheson IR, Nicholson RI,
13 Brown M, Jiang J, Howat WJ, Ali S, Carroll JS: Regulation of ERBB2 by
14 oestrogen receptor-PAX2 determines response to tamoxifen. *Nature*
15 2008, 456:663-666.
- 16 4. Wang J, Jarrett J, Huang CC, Satcher RL, Jr., Levenson AS: Identification of
17 estrogen-responsive genes involved in breast cancer metastases to the
18 bone. *Clin Exp Metastasis* 2007, 24:411-422.
- 19 5. Wei B, Wang J, Bourne P, Yang Q, Hicks D, Bu H, Tang P: Bone metastasis
20 is strongly associated with estrogen receptor-positive/progesterone
21 receptor-negative breast carcinomas. *Hum Pathol* 2008, 39:1809-1815.
- 22 6. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T,
23 Eisen MB, van de Rijn M, Jeffrey SS, et al: Gene expression patterns of
24 breast carcinomas distinguish tumor subclasses with clinical
25 implications. *Proc Natl Acad Sci U S A* 2001, 98:10869-10874.
- 26 7. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, Speed D,
27 Lynch AG, Samarajiwa S, Yuan Y, et al: The genomic and transcriptomic
28 architecture of 2,000 breast tumours reveals novel subgroups. *Nature*
29 2012, 486:346-352.
- 30 8. Jeselsohn R, Yelensky R, Buchwalter G, Frampton G, Meric-Bernstam F,
31 Gonzalez-Angulo AM, Ferrer-Lozano J, Perez-Fidalgo JA, Cristofanilli M,
32 Gomez H, et al: Emergence of constitutively active estrogen receptor-
33 alpha mutations in pretreated advanced estrogen receptor-positive
34 breast cancer. *Clin Cancer Res* 2014, 20:1757-1767.
- 35 9. Toy W, Shen Y, Won H, Green B, Sakr RA, Will M, Li Z, Gala K, Fanning S,
36 King TA, et al: ESR1 ligand-binding domain mutations in hormone-
37 resistant breast cancer. *Nat Genet* 2013, 45:1439-1445.
- 38 10. Robinson DR, Wu YM, Vats P, Su F, Lonigro RJ, Cao X, Kalyana-Sundaram
39 S, Wang R, Ning Y, Hodges L, et al: Activating ESR1 mutations in
40 hormone-resistant metastatic breast cancer. *Nat Genet* 2013, 45:1446-
41 1451.
- 42 11. Fuqua SA, Gu G, Rechoum Y: Estrogen receptor (ER) alpha mutations in
43 breast cancer: hidden in plain sight. *Breast Cancer Res Treat* 2014,
44 144:11-19.
- 45 12. Nakshatri H, Badve S: FOXA1 in breast cancer. *Expert Rev Mol Med* 2009,
46 11:e8.
- 47 13. Carroll JS, Liu XS, Brodsky AS, Li W, Meyer CA, Szary AJ, Eeckhoutte J,
48 Shao W, Hestermann EV, Geistlinger TR, et al: Chromosome-wide
49 mapping of estrogen receptor binding reveals long-range regulation
50 requiring the forkhead protein FoxA1. *Cell* 2005, 122:33-43.

14. Habashy HO, Powe DG, Rakha EA, Ball G, Paish C, Gee J, Nicholson RI, Ellis IO: Forkhead-box A1 (FOXA1) expression in breast cancer and its prognostic significance. *Eur J Cancer* 2008, 44:1541-1551.
15. Thorat MA, Marchio C, Morimiya A, Savage K, Nakshatri H, Reis-Filho JS, Badve S: FOXA1 expression in breast cancer is associated with Luminal subtype and good prognosis. *J Clin Pathol* 2007.
16. Badve S, Turbin D, Thorat MA, Morimiya A, Nielsen TO, Perou CM, Dunn S, Huntsman DG, Nakshatri H: FOXA1 Expression in Breast Cancer Correlation with Luminal Subtype A and Survival. *Clin Cancer Res* 2007, 13:4415-4421.
17. Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ, Brown GD, Gojis O, Ellis IO, Green AR, et al: Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* 2012, 481:389-393.
18. Hurtado A, Holmes KA, Ross-Innes CS, Schmidt D, Carroll JS: FOXA1 is a key determinant of estrogen receptor function and endocrine response. *Nat Genet* 2011, 43:27-33.
19. Magnani L, Ballantyne EB, Zhang X, Lupien M: PBX1 genomic pioneer function drives ERalpha signaling underlying progression in breast cancer. *PLoS Genet* 2011, 7:e1002368.
20. Osborne CK, Bardou V, Hopp TA, Chamness GC, Hilsenbeck SG, Fuqua SA, Wong J, Allred DC, Clark GM, Schiff R: Role of the estrogen receptor coactivator AIB1 (SRC-3) and HER-2/neu in tamoxifen resistance in breast cancer. *J Natl Cancer Inst* 2003, 95:353-361.
21. Anbalagan M, Rowan BG: Estrogen receptor alpha phosphorylation and its functional impact in human breast cancer. *Mol Cell Endocrinol* 2015 doi: 10.1016/j.mce.2015.01.016.
22. Badve S, Collins NR, Bhat-Nakshatri P, Turbin D, Leung S, Thorat M, Dunn SE, Geistlinger TR, Carroll JS, Brown M, et al: Subcellular localization of activated AKT in estrogen receptor- and progesterone receptor-expressing breast cancers: potential clinical implications. *Am J Pathol* 2010, 176:2139-2149.
23. Wierer M, Verde G, Pisano P, Molina H, Font-Mateu J, Di Croce L, Beato M: PLK1 Signaling in Breast Cancer Cells Cooperates with Estrogen Receptor-Dependent Gene Transcription. *Cell Rep* 2013, 3:2021-2032.
24. Bhola NE, Jansen VM, Bafna S, Giltane JM, Balko JM, Estrada MV, Meszoely I, Mayer I, Abramson V, Ye F, et al: Kinome-wide functional screen identifies role of PLK1 in hormone-independent, ER-positive breast cancer. *Cancer Res* 2015, 75:405-414.
25. Cyr AR, Kulak MV, Park JM, Bogachek MV, Spanheimer PM, Woodfield GW, White-Baer LS, O'Malley YQ, Sugg SL, Olivier AK, et al: TFAP2C governs the luminal epithelial phenotype in mammary development and carcinogenesis. *Oncogene* 2015, 34:436-444.
26. Woodfield GW, Chen Y, Bair TB, Domann FE, Weigel RJ: Identification of primary gene targets of TFAP2C in hormone responsive breast carcinoma cells. *Genes Chromosomes Cancer* 2010, 49:948-962.

- 1
 - 2
 - 3
 - 4
 - 5
 - 6
 - 7
 - 8
 - 9
 - 10
 - 11
 - 12
 - 13
 - 14
 - 15
 - 16
 - 17
 - 18
 - 19
 - 20
 - 21
 - 22
 - 23
 - 24
 - 25
 - 26
 - 27
 - 28
 - 29
 - 30
 - 31
 - 32
 - 33
 - 34
 - 35
 - 36
 - 37
 - 38
 - 39
 - 40
 - 41
 - 42
 - 43
 - 44
 - 45
 - 46
 - 47
 - 48
 - 49
 - 50
 - 51
 - 52
 - 53
 - 54
 - 55
 - 56
 - 57
 - 58
 - 59
 - 60
 - 61
 - 62
 - 63
 - 64
 - 65
27. Orso F, Cottone E, Hasleton MD, Ibbitt JC, Sismondi P, Hurst HC, De Bortoli M: Activator protein-2gamma (AP-2gamma) expression is specifically induced by oestrogens through binding of the oestrogen receptor to a canonical element within the 5'-untranslated region. *Biochem J* 2004, 377:429-438.
28. Tan SK, Lin ZH, Chang CW, Varang V, Chng KR, Pan YF, Yong EL, Sung WK, Cheung E: AP-2gamma regulates oestrogen receptor-mediated long-range chromatin interaction and gene transcription. *Embo J* 2011, 30:2569-2581.
29. Caizzi L, Ferrero G, Cutrupi S, Cordero F, Ballare C, Miano V, Reineri S, Ricci L, Friard O, Testori A, et al: Genome-wide activity of unliganded estrogen receptor-alpha in breast cancer cells. *Proc Natl Acad Sci U S A* 2014, 111:4892-7.
30. Gee JM, Eloranta JJ, Ibbitt JC, Robertson JF, Ellis IO, Williams T, Nicholson RI, Hurst HC: Overexpression of TFAP2C in invasive breast cancer correlates with a poorer response to anti-hormone therapy and reduced patient survival. *J Pathol* 2009, 217:32-41.
31. Guler G, Iliopoulos D, Guler N, Himmetoglu C, Hayran M, Huebner K: Wwox and Ap2gamma expression levels predict tamoxifen response. *Clin Cancer Res* 2007, 13:6115-6121.
32. Friedrichs N, Jager R, Paggen E, Rudlowski C, Merkelbach-Bruse S, Schorle H, Buettner R: Distinct spatial expression patterns of AP-2alpha and AP-2gamma in non-neoplastic human breast and breast cancer. *Mod Pathol* 2005, 18:431-438.
33. Jager R, Friedrichs N, Heim I, Buttner R, Schorle H: Dual role of AP-2gamma in ErbB-2-induced mammary tumorigenesis. *Breast Cancer Res Treat* 2005, 90:273-280.
34. Sgroi DC, Carney E, Zarrella E, Steffel L, Binns SN, Finkelstein DM, Szymonifka J, Bhan AK, Shepherd LE, Zhang Y, et al: Prediction of late disease recurrence and extended adjuvant letrozole benefit by the HOXB13/IL17BR biomarker. *J Natl Cancer Inst* 2013, 105:1036-1042.
35. Bianchini G, Pusztai L, Karn T, Iwamoto T, Rody A, Kelly C, Muller V, Schmidt S, Qi Y, Holtrich U, et al: Proliferation and estrogen signaling can distinguish patients at risk for early versus late relapse among estrogen receptor positive breast cancers. *Breast Cancer Res* 2013, 15:R86.
36. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM: REporting recommendations for tumor MARKer prognostic studies (REMARK). *Breast Cancer Res Treat* 2006, 100:229-235.
37. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM: REporting recommendations for tumor MARKer prognostic studies (REMARK). *Nat Clin Pract Urol* 2005, 2:416-422.
38. Zhou QaD, B: A SAS® Macro for biomarker analysis using maximally selected chi-square statistics with application in oncology. *PharmaSUG* 2012:SP12.

- 1
 - 2
 - 3
 - 4
 - 5
 - 6
 - 7
 - 8
 - 9
 - 10
 - 11
 - 12
 - 13
 - 14
 - 15
 - 16
 - 17
 - 18
 - 19
 - 20
 - 21
 - 22
 - 23
 - 24
 - 25
 - 26
 - 27
 - 28
 - 29
 - 30
 - 31
 - 32
 - 33
 - 34
 - 35
 - 36
 - 37
 - 38
 - 39
 - 40
 - 41
 - 42
 - 43
 - 44
 - 45
 - 46
 - 47
 - 48
 - 49
 - 50
 - 51
 - 52
 - 53
 - 54
 - 55
 - 56
 - 57
 - 58
 - 59
 - 60
 - 61
 - 62
 - 63
 - 64
 - 65
39. Goswami CP, Nakshatri H: PROGgeneV2: enhancements on the existing database. *BMC Cancer* 2014, 14:970.
40. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, et al: The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* 2012, 2:401-404.
41. Bonneterre J, Thurlimann B, Robertson JF, Krzakowski M, Mauriac L, Koralewski P, Vergote I, Webster A, Steinberg M, von Euler M: Anastrozole versus tamoxifen as first-line therapy for advanced breast cancer in 668 postmenopausal women: results of the Tamoxifen or Arimidex Randomized Group Efficacy and Tolerability study. *J Clin Oncol* 2000, 18:3748-3757.
42. Tamoxifen for early breast cancer: an overview of the randomised trials. Early Breast Cancer Trialists' Collaborative Group. *Lancet* 1998, 351:1451-1467.
43. Clarke C, Madden SF, Doolan P, Aherne ST, Joyce H, O'Driscoll L, Gallagher WM, Hennessy BT, Moriarty M, Crown J, et al: Correlating transcriptional networks to breast cancer survival: a large-scale coexpression analysis. *Carcinogenesis* 2013, 34:2300-2308.
44. Miller LD, Smeds J, George J, Vega VB, Vergara L, Ploner A, Pawitan Y, Hall P, Klaar S, Liu ET, Bergh J: An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. *Proc Natl Acad Sci U S A* 2005, 102:13550-13555.
45. Schmidt M, Bohm D, von Torne C, Steiner E, Puhl A, Pilch H, Lehr HA, Hengstler JG, Kolbl H, Gehrman M: The humoral immune system has a key prognostic impact in node-negative breast cancer. *Cancer Res* 2008, 68:5405-5413.
46. Loi S, Haibe-Kains B, Desmedt C, Lallemand F, Tutt AM, Gillet C, Ellis P, Harris A, Bergh J, Foekens JA, et al: Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade. *J Clin Oncol* 2007, 25:1239-1246.
47. Sarwar N, Kim JS, Jiang J, Peston D, Sinnett HD, Madden P, Gee JM, Nicholson RI, Lykkesfeldt AE, Shousha S, et al: Phosphorylation of ERalpha at serine 118 in primary breast cancer and in tamoxifen-resistant tumours is indicative of a complex role for ERalpha phosphorylation in breast cancer progression. *Endocr Relat Cancer* 2006, 13:851-861.
48. Mehta RJ, Jain RK, Leung S, Choo J, Nielsen T, Huntsman D, Nakshatri H, Badve S: FOXA1 is an independent prognostic marker for ER-positive breast cancer. *Breast Cancer Res Treat* 2012, 131:881-890.
49. Habashy HO, Rakha EA, Aleskandarany M, Ahmed MA, Green AR, Ellis IO, Powe DG: FOXO3a nuclear localisation is associated with good prognosis in luminal-like breast cancer. *Breast Cancer Res Treat* 2011, 129:11-21.
50. Bogachek MV, Chen Y, Kulak MV, Woodfield GW, Cyr AR, Park JM, Spanheimer PM, Li Y, Li T, Weigel RJ: Sumoylation pathway is required

1
2
3
4 to maintain the basal breast cancer subtype. *Cancer Cell* 2014, 25:748-
5 761.

- 6
7 51. Wright TM, Wardell SE, Jasper JS, Stice JP, Safi R, Nelson ER, McDonnell
8 DP: Delineation of a FOXA1/ERalpha/AGR2 regulatory loop that is
9 dysregulated in endocrine therapy-resistant breast cancer. *Mol Cancer*
10 *Res* 2014, 12:1829-1839.
11
12 52. Nikolsky Y, Sviridov E, Yao J, Dosymbekov D, Ustyansky V, Kaznacheev
13 V, Dezsó Z, Mulvey L, Macconnaill LE, Winckler W, et al: Genome-wide
14 functional synergy between amplified and mutated genes in human
15 breast cancer. *Cancer Res* 2008, 68:9532-9540.
16
17 53. Yahata T, Shao W, Endoh H, Hur J, Coser KR, Sun H, Ueda Y, Kato S,
18 Isselbacher KJ, Brown M, Shioda T: Selective coactivation of estrogen-
19 dependent transcription by CITED1 CBP/p300-binding protein. *Genes*
20 *Dev* 2001, 15:2598-2612.
21
22 54. Howlin J, McBryan J, Napoletano S, Lambe T, McArdle E, Shioda T,
23 Martin F: CITED1 homozygous null mice display aberrant pubertal
24 mammary ductal morphogenesis. *Oncogene* 2006, 25:1532-1542.
25
26 55. Spanheimer PM, Woodfield GW, Cyr AR, Kulak MV, White-Baer LS, Bair
27 TB, Weigel RJ: Expression of the RET proto-oncogene is regulated by
28 TFAP2C in breast cancer independent of the estrogen receptor. *Ann*
29 *Surg Oncol* 2013, 20:2204-2212.
30
31 56. Wong PP, Miranda F, Chan KV, Berlato C, Hurst HC, Scibetta AG: Histone
32 demethylase KDM5B collaborates with TFAP2C and Myc to repress the
33 cell cycle inhibitor p21(cip) (CDKN1A). *Mol Cell Biol* 2012, 32:1633-
34 1644.
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 **Figure Legends:**
5

6
7 Figure 1: Immunostaining pattern of TFAP2C in breast cancers and normal breast.
8

9
10
11 Figure 2: Relationship between TFAP2C-H scores and overall survival of all patients or
12 patients subdivided based on tumor characteristics and/or treatment. A blue line
13 represents high TFAP2C H-score, whereas a red line represents Low TFAP2C H-score.
14
15
16
17

18
19 Figure 3: Prognostic significance of TFAP2C mRNA in breast cancer. Public databases
20 were used to classify breast tumors into TFAP2C-high and TFAP2C-low (divided at
21 median) subgroups and then analyzed based on different tumor characteristics as
22 indicated (A-D). E) TFAP2C genome aberration in various cancers was determined using
23 the cBioportal.
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Table 1. Description of the patients and characteristics of their tumors (n=451)

Characteristic	Have TFAP2C (N=385)	Missing TFAP2C (N=66)	p-value*
Age at Diagnosis, y			0.0891
Mean (SD)	58.60 (14.86)	61.9 (13.53)	
Median	58.00	64.00	
Range	27.00-94.00	34.00-90.00	
Race, <i>n</i> (%)			0.8800
White	306 (79.48%)	55 (83.33%)	
African American	74 (19.22%)	11 (16.67%)	
Asian	4 (1.04%)	0	
Other	1 (0.26%)	0	
PR Status, <i>n</i> (%)			0.0637
Negative	120 (31.17%)	14 (21.21%)	
Positive	227 (58.96%)	48 (72.73%)	
Unknown	38 (9.87%)	4 (6.06%)	
ER Status, <i>n</i> (%)			0.0995
Negative	79 (20.52%)	8 (12.12%)	
Positive	284 (73.77%)	55 (83.33%)	
Not Done	22 (5.71%)	3 (4.55%)	
HER-2/neu, <i>n</i> (%)			0.1165
Negative	160 (41.56%)	30 (45.45%)	
Positive	50 (12.99%)	4 (6.06%)	
Unknown/Not Done	175 (45.45%)	32 (48.48%)	
Tumor Grade, <i>n</i> (%)			0.0778
I	100 (25.97%)	24 (36.36%)	
II	159 (41.30%)	27 (40.91%)	
III	94 (24.42%)	9 (13.64%)	
Unknown	32 (8.31%)	6 (9.09%)	
T Stage, <i>n</i> (%)			0.0684
T0	2 (0.52%)	1 (1.52%)	

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Characteristic	Have TFAP2C (N=385)	Missing TFAP2C (N=66)	p-value*
T1	212 (55.06%)	46 (69.70%)	
T2	131 (34.03%)	14 (21.21%)	
T3	29 (7.53%)	2 (3.03%)	
T4	10 (2.60%)	2 (3.03%)	
TX/Unknown	1 (0.26%)	1 (1.52%)	
N Stage, n(%)			0.3587
N0	233 (60.52%)	38 (57.58%)	
N1	101 (26.23%)	21 (31.82%)	
N2	25 (6.49%)	1 (1.52%)	
N3	9 (2.34%)	1 (1.52%)	
NX/Unknown	17 (4.42%)	5 (7.58%)	
M Stage, n(%)			0.1429
M0	291 (75.58%)	48 (72.73%)	
M1	16 (4.16%)	0	
MX/Unknown	78 (20.26%)	18 (27.27%)	
Endocrine Therapy, n(%)			
Yes	243 (63.12%)	36 (54.55%)	0.2308
No	138 (35.84%)	29 (43.94%)	
Unknown	4 (1.04%)	1 (1.52%)	
Follow-up, y, median (95% CI)	10.3 (0.04, 21.8)	10.3 (1.1, 17.5)	0.2421
Recurrence, n(%)			0.1027
Yes	126 (32.73%)	15 (22.73%)	
No	258 (67.01%)	51 (77.27%)	
Unknown	1 (0.26%)	0	
Death, n(%)			0.1081
Yes	163 (42.34%)	21 (31.82%)	
No	222 (57.66%)	45 (68.18%)	
Overall survival, y, median (95% CI)	11.8 (11.2, 12.5)	NE (10.7, NE)	0.2117

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Characteristic	Have TFAP2C (N=385)	Missing TFAP2C (N=66)	p-value*
Disease-free survival, y, median (95% CI)***	10.8 (10.2, 11.9)	11.4 (10.2, NE)	0.2311

*t-test for age; chi-square or fisher's exact test, where appropriate, for categorical; log-rank for time to event; unknown values were not included in tests

**All unknown/not done values were prior to April 2001, when testing became fully integrated as standard of care at our institution.

***excludes M1

NE = not estimable

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Table 2. Bivariate analysis of TFAP2C H-score with other tumor markers

Variable	TFAP2C H-score Median (25 th percentile, 75 th percentile)						p-value*
	n	values	n	values	n	values	
	Negative			Positive			
ER	79	10.9 (0, 59.2)	284	12.6 (0, 40.9)			0.9097
PR	120	11.9 (0, 53.5)	227	13.3 (0, 38.6)			0.9958
HER-2/neu	160	10.8 (0, 48.2)	50	0 (0, 31.2)			0.5646
ER+/PR+/HER-	73	0 (0, 62.5)	134	0 (0, 42.0)			0.6318
Nodal Status	233	12.5 (0, 36.3)	135	12.2 (0, 46.6)			0.9539
	Grade 1		Grade 2		Grade 3		
Tumor Grade	100	10.8 (0, 24.8)	159	12.6 (0, 41.0)	94	16.3 (0, 65.8)	0.1135

*from Wilcoxon Rank Sum test for Hormone Receptor Status and Kruskal-Wallis test for Tumor Grade

Table 3. Univariate analysis of other tumor markers for Overall Survival

Variable	n	Comparison*	HR (95% CI)	p-value**
ER	426	ER- vs. ER+	1.18 (0.82, 1.71)	0.3874
PR	409	PR- vs. PR+	1.59 (1.16, 2.17)	0.0041
HER-2/neu	244	HER-2/neu- vs. HER-2/neu+	0.66 (0.41, 1.06)	0.0832
ER+/PR+/HER-	240	No vs. Yes	1.48 (0.94, 2.33)	0.0946
Tumor Grade	413	Grade 1 vs. Grade 2	0.78 (0.53, 1.15)	<0.0001
		Grade 1 vs. Grade 3	0.41 (0.27, 0.62)	
		Grade 2 vs. Grade 3	0.53 (0.37, 0.75)	
T Stage	449	T0/1 vs. T2	0.59 (0.43, 0.82)	<0.0001
		T0/1 vs. T3/4	0.31 (0.20, 0.47)	
		T2 vs. T3/4	0.52 (0.33, 0.80)	
N Stage	429	N+ vs. N0	1.62 (1.19, 2.19)	0.0020

*referent group listed last

**from Wald Chi-square test

Table 4. Multivariable analysis for Overall Survival (N=360)

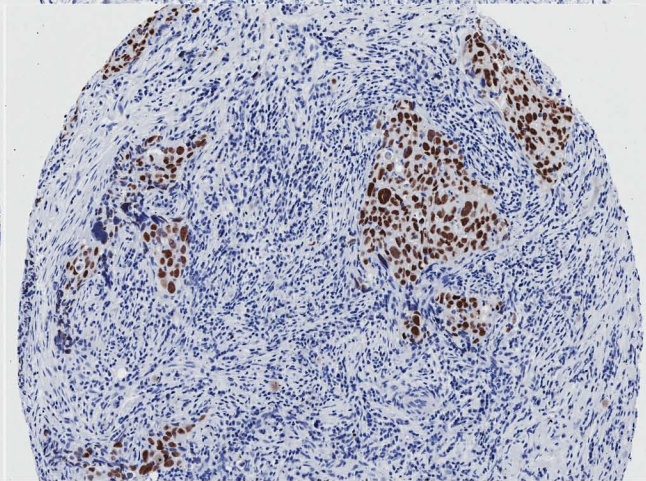
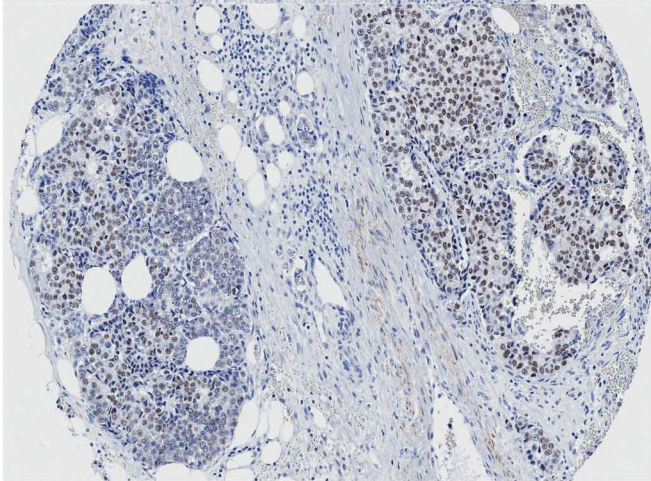
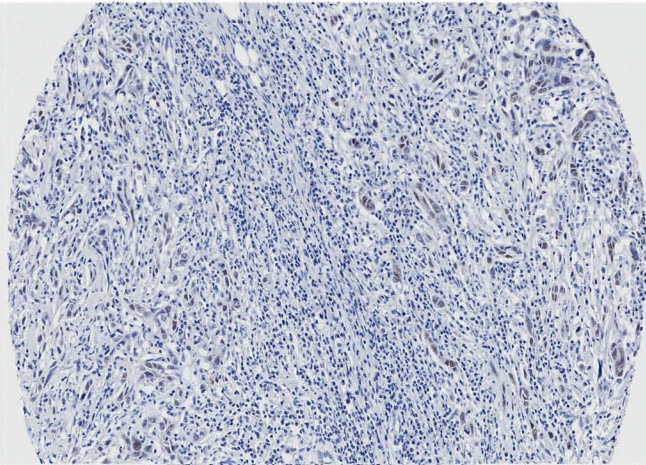
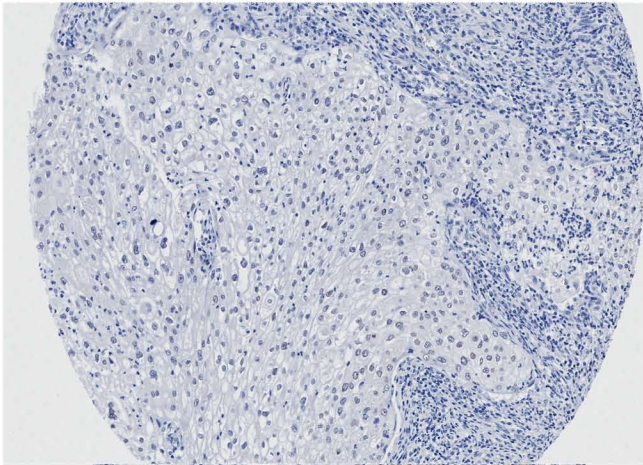
Variable	Comparison*	HR (95% CI)	p-value**
PR Status	Negative vs. Positive	1.25 (0.86, 1.84)	0.4312
	Not done vs. Positive	1.29 (0.70, 2.36)	
Tumor Grade	Grade 1 vs. Grade 2	0.75 (0.47, 1.20)	0.0463
	Grade 1 vs. Grade 3	0.48 (0.28, 0.81)	
	Grade 1 vs. Unknown	0.70 (0.35, 1.39)	
	Grade 2 vs. Grade 3	0.63 (0.42, 1.97)	
	Grade 2 vs. Unknown	0.93 (0.51, 1.72)	
	Grade 3 vs. Unknown	1.46 (0.78, 2.76)	
T Stage	T0/1 vs. T2	0.80 (0.54, 1.18)	0.0898
	T0/1 vs. T3/4	0.53 (0.31, 0.93)	
	T0/1 vs. TX	0.21 (0.03, 1.71)	
	T2 vs. T3/4	0.67 (0.40, 1.12)	
	T2 vs. TX	0.27 (0.03, 2.10)	
	T3/4 vs. TX	0.40 (0.05, 3.28)	
N Stage	N0 vs. N+	1.56 (1.07, 2.27)	0.0298
TFAP2C H-score Category	≤ 10 years High vs. Low	1.01 (0.69, 1.47)	0.9662
	> 10 years High vs. Low	3.40 (1.58, 7.30)	0.0018

*referent group listed last

**from Wald Chi-square test

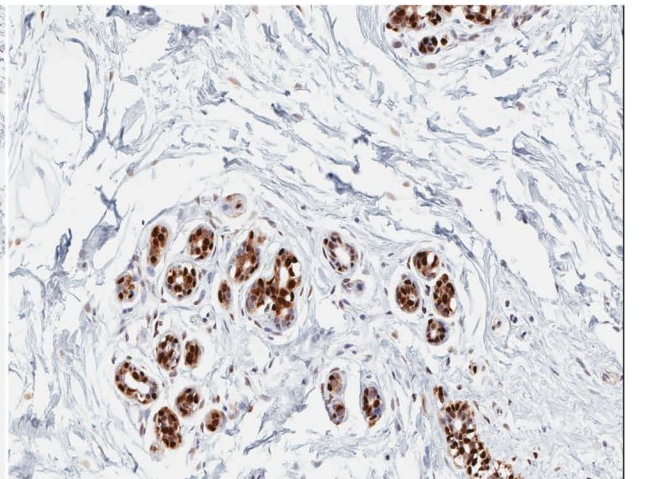
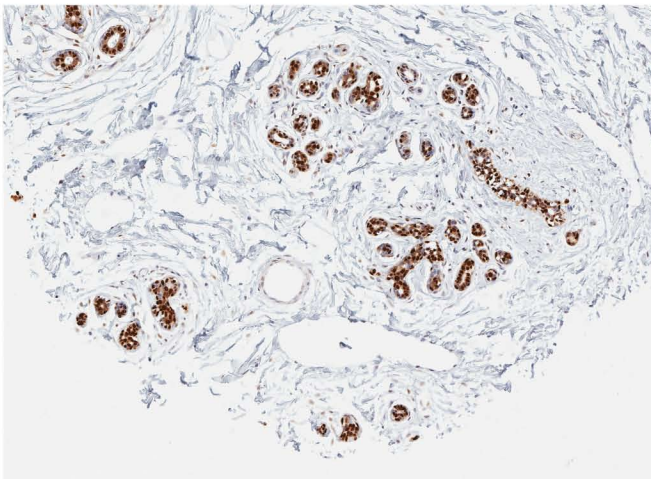
Negative

Low



Medium

High



Normal breast 10X

Normal breast 20X

Figure 1

Figure

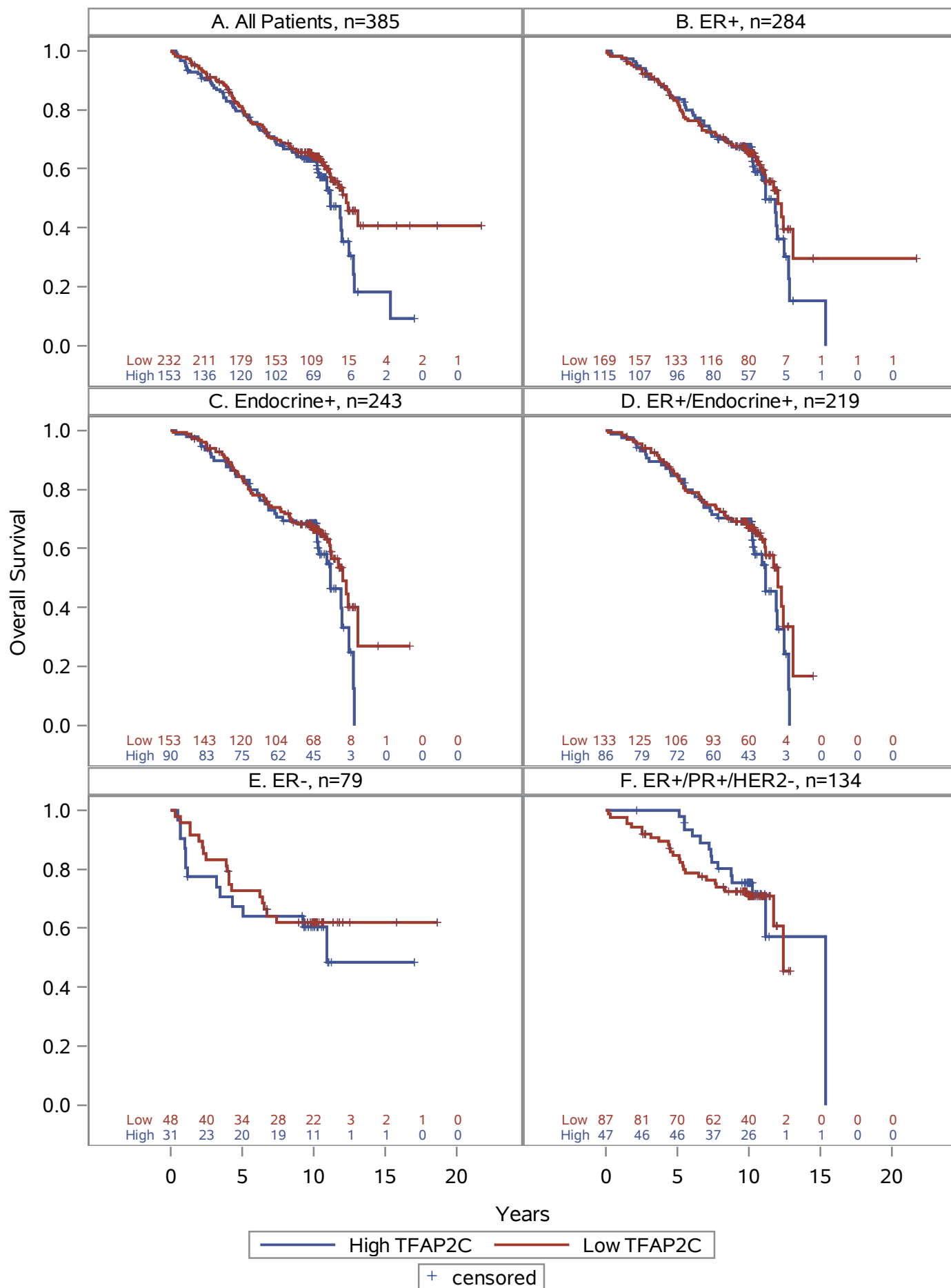
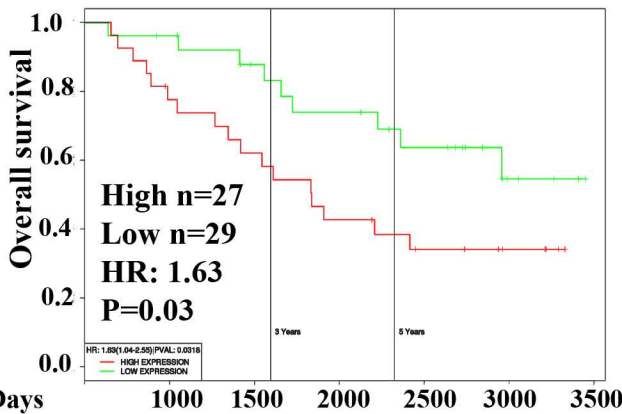
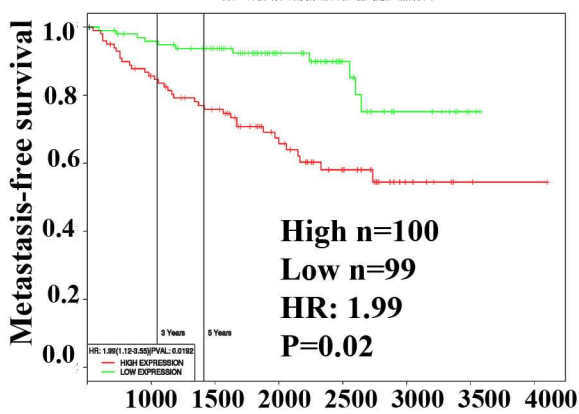


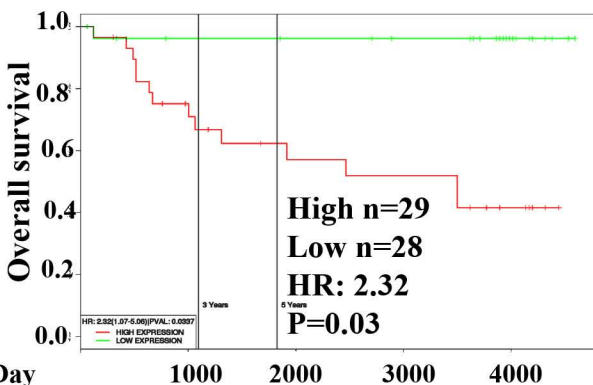
Figure **GSE42568- Grade 3 subgroup**



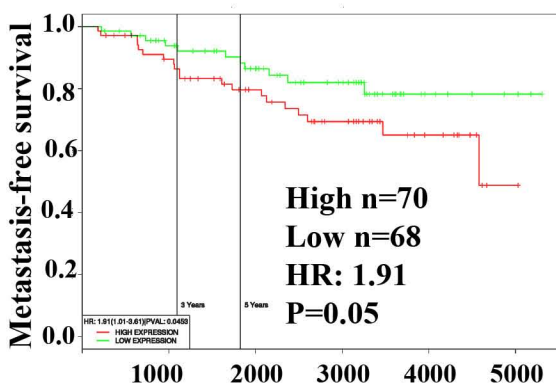
C **GSE11121 Node-negative subgroup**



B **GSE3494 PR-negative subgroup**



D **GSE6532 ER+ subgroup**



— High expression
— Low expression

E

