The gene encoding the estrogen receptor, ESR1, is differentially expressed in the primary tumors and metastases of patients with breast cancer based on survival outcomes and its expression correlates with survival.

Shahan Mamoor¹

¹Thomas Jefferson School of Law

San Diego, CA 92101

mamoorsk@tjssl.edu

Breast cancer is a leading cause of cancer death for women in the United States (1). Metastatic breast cancer is a major public health problem with a 5-year survival rate of 27% (1). We mined published microarray datasets to identify genes whose expression was most significantly different when comparing primary and metastatic tumors based on survival outcomes at 24 months (2, 3). In the primary tumors of patients with breast cancer, we identified the estrogen receptor alpha, ESR1 (also known as ER) as one of the most significantly differentially expressed genes when comparing tumor transcriptomes based on disease-free survival at 24 months. When comparing the transcriptomes of metastases from patients with stage IV metastatic breast cancer, we again found that ESR1 was among the genes whose expression was most significantly different based on overall survival at 24 months. In metastatic tissues, expression of ESR1 positively correlated with overall patient survival. Stratifying patients based on ESR1 expression revealed that higher ESR1 expression in both primary and metastatic tissues was generally associated with enhanced patient survival. Though established immunohistochemical classification of breast cancers groups tumors in a binary fashion based on ER expression at the protein level as ER+ and ER-, these data suggest that graded interpretation of estrogen receptor expression at the mRNA level, in primary tumors and in metastatic tissues, may be useful for prognostic purposes.

Keywords: breast cancer, metastatic breast cancer, tumor transcriptome analysis, ESR1, estrogen receptor alpha, ER, systems biology of breast cancer.

1 While the 5-year survival rate for patients with localized, stage I breast cancer is 99%
2 the 5-year survival rate for patients with metastatic breast cancer, cancer that has spread
3 beyond the breast to distant organ sites like the lungs, liver, brain and bones is a dismal 27%
4 (1). Metastasis is the major reason patients will die from solid tumors (4). Novel therapeutic
5 approaches to prevent, halt, and/or reverse metastasis are required. Understanding the
6 transcriptional landscape of primary and metastatic tumor tissues, and how this transcriptional
7 landscape differs most significantly between the primary and metastatic tumor tissues of
8 patients who survive the disease longer than patients who succumb to the disease is an
9 important step in the development of targeted therapeutics for metastatic breast cancer.
10

11 In this study we analyzed tumor transcriptome data from two independent published
12 microarray datasets generated using the primary and metastatic tumor tissues of women with
13 breast cancer for which there was associated survival data (2, 3). We found that the gene
14 encoding the estrogen receptor alpha, ESR1 (also known as ER) was among the genes whose
15 expression was most significantly different between the primary tumors of patients with breast
16 cancer when comparing tumor transcriptomes based on survival at 24 months. In patients
17 with stage IV metastatic breast cancer, ESR1 was again identified as a differentially expressed
18 gene when comparing the transcriptomes of metastatic tumor tissues based on patient survival
19 at 24 months. Further analyses revealed that ESR1 expression correlated with patient survival
20 in a linear fashion. Higher expression of ESR1, both in primary tumors and in metastatic
21 tissues correlated with enhanced patient survival. ESR1 mRNA levels in patient tumors may be
22 useful for prognostic purposes in identifying a group of patients with low ESR1 expression that
23 might benefit from more rigorous therapeutic regimens.
24
25
26

27 **Methods**

28 We used GSE48391 (2) and GSE124647 (3) for this differential gene expression analysis
in conjunction with GEO2R.

1 GSE48391 (2) was generated using Affymetrix Human Genome U133 Plus 2.0 Array
2 technology with n=70 for primary tumors of the breast from women with disease-free survival
3 greater than 24 months and an n=11 for primary tumors of the breast from women with
4 disease-free survival less than 24 months.

5 GSE124647 (3) was generated using Affymetrix Human Genome U133 Array from
6 biopsies from patients with stage IV metastatic breast cancer with an n=70 for metastatic
7 tumor tissues from women surviving greater than 24 months and an n=70 for metastatic tumor
8 tissues from women surviving less than 24 months.

9
10 The Benjamin and Hochberg method of p-value adjustment was used for ranking of
11 differential expression but raw p-values were used for assessment of statistical significance of
12 global differential expression. Log-transformation of data was auto-detected, and the NCBI
13 generated category of platform annotation was used.

14
15 A statistical test was performed to evaluate the significance of difference between mRNA
16 expression levels of ESR1 in the primary tumors of women surviving less than 24 months as
17 compared to the primary tumors of women surviving greater than 24 months, as well as in the
18 metastatic tissues of women surviving less than 24 months as compared to the metastatic
19 tissue of women surviving greater than 24 months using a two-tailed, unpaired t-test with
20 Welch's correction. A statistical test was performed to compare time of disease-free survival of
21 patients when grouping their primary tumors by intervals of ESR1 expression into 4 "bins" of 20
22 or 21 and to compare time of patient overall survival when grouping their metastatic tumor
23 tissues by intervals of ESR1 expression into 4 "bins" of 35 using an ordinary one-way ANOVA
24 with Dunnett's multiple comparisons test. Only p-values less than 0.05 were considered
25 statistically significant. A simple linear regression analysis was performed along with Pearson
26 correlation, comparing the mRNA expression level of ESR1 in each patient's metastatic tumor
27 tissues and the amount of time they survived, and we evaluated the r and r squared value for
28

1 goodness of fit and the p-value to assess the significance of the non-zero, linear nature of the
2 relationship. A statistical test was performed to compare the percentage of stage IV metastatic
3 breast cancer patients with overall survival less than 20% based on ESR1 expression from
4 highest to lowest in four “bins” using an ordinary one-way ANOVA with Dunnett’s multiple
5 comparisons test. We used PRISM for all statistical analyses (Version 8.4.0)(455).
6

7 **Results**

8
9 We compared the transcriptomes of primary tumors and metastatic tissues from
10 women with breast cancer using two independently published microarray datasets (2, 3) to
11 understand how gene expression in tumor tissues most significantly differed between breast
12 cancer patients based on survival outcomes at 24 months.
13

14 ESR1 is differentially expressed in breast cancer when comparing primary tumors based on 15 disease-free survival of the patient at 24 months.

16 We identified ESR1 as among the most differentially expressed genes when comparing
17 primary tumors from 81 women with breast cancer who possessed disease-free survival
18 greater than or less than 24 months using a published dataset (2). When sorting each of the
19 microarray-detected mRNAs expressed in primary tumors based on change in expression
20 between patients with disease-free survival greater or less than 24 months, ESR1 ranked 136
21 out of 54674 total transcripts. Differential expression of ESR1 in primary tumors of the breast
22 was statistically significant (Table 1; $p=0.00136651$).
23

24 25 ESR1 is differentially expressed in metastatic breast cancer when comparing metastases 26 based on overall survival of the patient at 24 months.

27 We analyzed a second published dataset (3), this containing transcriptome data from
28 the metastases of 140 women with stage IV breast cancer. Again, we identified ESR1 as
among the most differentially expressed genes when comparing metastatic tumor tissues from

women with metastatic breast cancer possessing overall survival greater than or less than 24 months. In this dataset, four separate mRNAs encoded by the ESR1 gene were identified as among the transcripts whose expression was most quantitatively different when comparing metastatic tissues based on overall survival at 24 months. When sorting each of the microarray-detected mRNAs expressed in metastases based on change in expression between patients with overall survival greater or less than 24 months, the four ESR1 transcripts ranked 3, 39, 139, and 152 out of 22283 total transcripts. Differential expression of each of these ESR1 transcripts in metastases from patients with stage IV breast cancer was statistically significant (Table 2; $p= 6.26E-07$, $p= 6.15E-05$, $p= 5.24E-04$ and $p= 5.77E-04$, respectively).

ESR1 is expressed at significantly higher levels in the primary tumors of patients with disease-free survival greater than 24 months.

We extracted exact mRNA expression values for ESR1 from each primary tumor to directly compare ESR1 levels between the two groups of patients: those with disease-free survival greater than 24 months, and those with disease-free survival less than 24 months. ESR1 was expressed at significantly higher levels in the primary tumors of patients with disease-free survival greater than 24 months when compared to those with disease-free survival less than 24 months (Figure 1; $p=0.0074$).

ESR1 is expressed at significantly higher levels in the metastatic tumors of patients with overall survival greater than 24 months.

We also extracted exact mRNA expression values for ESR1 from each metastatic tumor tissue to directly compare ESR1 levels between patients with overall survival greater than 24 months, and patients with overall survival less than 24 months. Again, ESR1 was expressed at higher levels in patients with enhanced survival metrics (Figure 2). ESR1 was expressed at significantly higher levels in the metastases of patients with overall survival greater than 24

months when compared to the metastases of patients with overall survival less than 24 months (Figure 2; $p < 0.0001$).

Linear regression analysis demonstrates ESR1 expression in metastases to multiple organ sites generally correlates with overall patient survival in metastatic breast cancer.

Since ESR1 was expressed at significantly higher levels in the metastatic tumor tissues of patients with overall survival greater than 24 months, we asked whether ESR1 expression levels in the metastatic tissues of patients with stage IV breast cancer generally correlated with overall survival. Simple linear regression analysis revealed that ESR1 expression in metastatic tissues was significantly correlated with overall survival in patients with stage IV breast cancer (Figure 3; $p = 0.0031$). In general, as the expression of ESR1 in metastatic tissues increased, so did the overall survival of that patient; this was not nearly a perfect correlation, however (Figure 3; $r = 0.3198$; $r^2 = 0.06168$). We concluded that ESR1 expression in metastases correlated with overall survival of patients with stage IV breast cancer.

Binning primary tumors based on ESR1 expression demonstrates mRNA level ESR1 expression can help predict enhanced survival outcomes in patients with breast cancer.

We asked whether ESR1 expression in primary tumors, broadly speaking, would be able to predict disease-free survival. In other words, did patients with the highest levels of ESR1 expression in their primary tumors possess greater disease-free survival than patients with lowest levels of ESR1 expression? We grouped the 140 metastatic tumor tissues isolated from patients with metastatic breast cancer into 4 “bins” of 35 tumors (Figure 4). While patients in Bin 2 and Bin 3 possessed mean disease-free survival greater than that of patients in Bin 1, this was not statistically significant (Figure 4; $p = 0.8226$ and $p = 0.1059$, respectively). However, patients in Bin 4, patients whose tumors expressed ESR1 mRNA at the highest levels, displayed significantly higher disease-free survival than patients in Bin 1 (Figure 4; $p = 0.0096$). In this dataset, patients whose primary tumors expressed the highest levels of

ESR1 possessed significantly increased disease-free survival than patients whose primary tumors expressed the lowest levels of ESR1.

Binning metastatic tissues based on ESR1 expression demonstrates mRNA level ESR1 expression in metastases can help predict enhanced survival outcomes in patients with stage IV breast cancer.

We also performed a similar analysis with the four separate ESR1 transcripts we identified as being among the mRNAs whose expression changed most significantly between those with overall survival greater or less than 24 months, asking whether ESR1 expression in metastases, broadly speaking, would be able to predict overall survival in patients with stage IV metastatic breast cancer. In other words, did patients with the highest levels of ESR1 expression in their metastases possess greater overall survival than patients with lowest levels of ESR1 expression? We grouped the 140 metastases isolated from patients with stage IV breast cancer into 4 “bins” of 35 tumors each (Figure 5). For the two ESR1 transcripts whose expression changed most significantly between patients based on overall survival at 24 months (IDs 211234_x_at and 211233_x_at), again, patients in Bin 2 and Bin 3 possessed mean overall survival greater than that of patients in Bin 1, but this was not statistically significant (Figure 5; $p=0.6449$ and $p=0.1923$ for Bins 2 and 3 [211234_x_at], and $p=0.9589$ and $p=0.1130$ for Bins 2 and 3 [211233_x_at], respectively). However, patients in Bin 4, patients whose metastatic tissues expressed ESR1 mRNA at the highest levels, displayed significantly higher mean overall survival than patients in Bin 1, and this was the case for both of the most differentially expressed ESR1 mRNAs (Figure 5; $p=0.0002$ and $p=0.0053$). While mean overall survival was higher in Bin 4 patients than in Bin 1 patients for the remaining two differentially expressed ESR1 mRNAs, this was not statistically significant (Figure 5; $p=0.1597$ and $p=0.1071$, respectively). Thus, in this dataset, patients whose metastases expressed the highest levels of ESR1 possessed significantly increased overall survival than patients whose metastases expressed the lowest levels of ESR1. We concluded that ESR1 expression, not only in primary

1 tumors, but also in metastases could help predict overall survival in many patients with stage
2 IV breast cancer.

3
4 Binning metastatic tissues based on ESR1 expression demonstrates mRNA level ESR1
5 expression in metastases can help predict enhanced survival outcomes in patients with stage
6 IV breast cancer.

7
8 Finally, we asked in patients with stage IV metastatic breast cancer whether we could
9 detect differences in the percentage of patients with poor overall survival - less than 20 months
10 - based on the level of ESR1 expressed in patient metastatic tumor tissue. Again, we grouped
11 the metastases of these 140 patients into 4 “bins” of 35 each, and compared ESR1 expression
12 levels for each the four differentially expressed ESR1 transcripts identified to overall survival
13 metrics for each patient. While the majority of patients with the highest levels of ESR1
14 expression in metastatic tissues possessed overall survival greater than 20 months (Table 3;
15 22.14% of patients with overall survival less than 20 months in Bin 4), the majority of patients
16 with the lowest ESR1 expression levels in metastatic tissues possessed overall survival less
17 than 20 months (Table 3; 65.71% patients with overall survival less than 20 months in Bin 1).

18
19
20 Thus, using two independently published microarray datasets (2, 3), we identified ESR1
21 as among the genes whose expression was most significantly different between the primary
22 tumors and metastases of patients with breast cancer when comparing tumor tissues based on
23 disease-free survival and overall survival, respectively. ESR1 expression in patient metastases
24 generally correlated, in a linear fashion, with overall survival in patients with stage IV metastatic
25 breast cancer. In both primary tumors and in metastases, respectively relative disease-free
26 survival and overall survival could be predicted using ESR1 tumor expression. A majority of
27 stage IV metastatic breast cancer patients whose metastases expressed the highest ESR1
28 mRNA levels possessed overall survival greater than 20 months, while a majority of stage IV
metastatic breast cancer patients whose metastases expressed the lowest ESR1 mRNA levels

possessed overall survival less than 20 months. In general, increasing ESR1 mRNA levels in primary and metastatic tumor tissues translated into enhanced survival outcomes for patients with breast cancer.

Discussion

Expression of the estrogen receptor at the protein level is one major characteristic by which patients with breast cancers are distinguished (5). The difference between ER+ and ER- cancers has more importance than a mere label because this informs patient treatment; patients with ER+ cancers will receive endocrine therapy such as tamoxifen, letrozole, anastrozole, and exemestane (6). We sought to understand how gene expression most significantly differs in the primary and metastatic tumor tissues of patients with breast cancer when comparing patients based on disease-free survival (2) or overall survival (3), through differential gene expression analysis using published microarray datasets (2, 3). We found in one dataset (2) that the gene encoding the estrogen receptor alpha, ESR1 (also known as ER) was among the genes whose expression was most quantitatively different between the primary tumors of patients with disease-free survival greater or less than 24 months. In a separate dataset (3), we found that the ESR1 was among the genes whose expression was most quantitatively different between the metastatic tissues of patients with stage IV metastatic breast cancer with overall survival greater or less than 24 months. ESR1 expression in patient metastases generally correlated in a linear fashion with overall survival of the patient in one group of patients with stage IV breast cancer. ESR1 expression in primary tumors could predict relative disease-free survival outcomes, as patients whose tumors expressed the highest levels of ESR1 in their primary tumors possessed significantly higher disease-free survival than patients whose primary tumors expressed the lowest levels of ESR1. ESR1 expression in metastases was also relatively predictive of overall survival for patients with stage IV metastatic breast cancer: patients whose metastases expressed the highest levels of ESR1

1 possessed significantly increased overall survival than patients whose metastases expressed
2 the lowest levels of ESR1.

3
4 One study demonstrated that low levels of ESR1 expression predicted tamoxifen
5 resistance in ER+ breast cancers (7). This is biologically intuitive as tamoxifen is an endocrine
6 therapy and relevant because patients with low levels of ESR1, despite this predicted
7 resistance, are still classified as having ER+ breast cancer and receive endocrine therapy.
8

9
10 A 2007 published analysis of survival data from the National Cancer Institute's
11 Surveillance Epidemiology and End Results (SEER) program demonstrated that the peak
12 incidence of deaths from breast cancer occurs 2 to 4 years from diagnosis, with annual hazard
13 of breast cancer death for women with ER- breast cancers spiking at approximately 2 years
14 and is strikingly higher at this time point for women with ER- breast cancers as compared to
15 women with ER+ breast cancers (8). The results from our analyses, though assessing different
16 survival outcomes, are generally consistent with SEER survival data demonstrating that
17 patients whose tumors express less ESR (or, no ESR) possess worse survival outcomes than
18 patients whose tumors expression more ESR (or, some level of ESR).
19

20
21 Borderline estrogen receptor positive cancers are defined by positivity of ER expression
22 by immunohistochemistry from 1 to 10% (9). A study found that most ER-positive tumors
23 (defined by IHC) tumors with expressed deemed from 1% to 9% display similarity at the
24 molecular level to ER-negative basal-like tumors (9). Only a minority of 1% to 9% ER-positive
25 tumors displayed similarity at the molecular level to ER+, endocrine-sensitive tumors. This
26 same study concluded that a second assessment of ER expression using RNA levels, in
27 addition to protein levels by IHC, could help identify patients who had been diagnosed with
28 ER+ cancers but would likely not benefit from endocrine therapy (9).

1 Here, we analyzed survival outcomes based on mRNA expression of ESR1; differential
2 expression of ESR1 is just one way this gene is important in breast cancer. Alterations of the
3 ESR1 gene at the genetic level have been identified by multiple groups in patients with breast
4 cancer. This includes gene fusions by rearrangement, for example to the CCDC170 gene (10);
5 copy number variation by gene amplification (11), mutation of the ESR1 gene (12-14), non-
6 coding single-nucleotide variants that alter expression of ESR1 (15) as well as non-coding
7 RNAs that can influence ESR1 expression (16).
8

9 At least one of these studies linked changes at the genetic level with changes in ESR1
10 expression levels (11). Amplification of the ESR1 gene was identified in greater than 20% of
11 breast cancers in one study of patients treated with tamoxifen, and this copy number variation
12 correlated with increases in ESR1 protein expression (11). Patients whose primary breast
13 tumors contained ESR1 amplification possessed significantly increased survival when
14 compared to patients with ER-positive tumors with no ESR1 amplification, and when compared
15 to ER-negative cancers. Patients whose tumor DNA had more than 10 copies of ESR1 survived
16 longer periods of time than patients whose tumor DNA had either 2, 3-4, or 5-10 copies of the
17 ESR1 gene. Allred scoring is a method of scoring ER positivity based on immunohistochemistry
18 (17). While ER+ breast cancer patients with tumor Allred scores 7-8 survived longer than
19 patients with tumor Allred scores of 0-2 and 3-6, patients whose tumors contained ESR1
20 amplification possessed even greater survival compared to patients whose tumors with Allred
21 scores of 7-8. ESR1 copy number variation through gene amplification correlated with ESR1
22 expression and increased ESR1 expression was statistically linked with increased patient
23 survival.
24
25

26 We found that ESR1 was among the most differentially expressed genes in the primary
27 tumors of patients with breast cancer and in the metastatic tissues of patients with stage IV
28 metastatic breast cancer when comparing patients tumor transcriptomes based on survival
outcomes at 24 months by analyzing published microarray data from two separate groups of

1 patients with breast cancer (2, 3). While these patient groups are relatively small, the results
2 were clear: the change in ESR1 expression between the primary tumors and metastatic tissues
3 of patients based on disease-free survival or overall survival at 24 months was greater than the
4 change in more than 99% of remainder of the tumor transcriptome. While a majority of stage IV
5 metastatic breast cancer patients whose metastases expressed the highest ESR1 mRNA levels
6 possessed overall survival greater than 20 months, a majority of stage IV metastatic breast
7 cancer patients whose metastases expressed the lowest ESR1 mRNA levels possessed overall
8 survival less than 20 months. In general, increasing ESR1 mRNA levels in primary and
9 metastatic tumor tissues translated into better survival outcomes for patients with breast cancer.
10
11 It may be beneficial to interpret ESR1 expression levels in a graded fashion, to identify patients
12 who might require more rigorous treatment protocols based on prediction of poor survival
13 outcomes and when deciding which patients might not benefit from endocrine therapy despite
14 being labeled as having ER+ breast cancers (8, 9).

References

1. ACS Cancer Facts & Figures 2019, <https://www.cancer.net/cancer-types/breast-cancer-metastatic/statistics>.
2. Huang, C.C., Tu, S.H., Lien, H.H., Jeng, J.Y., Huang, C.S., Huang, C.J., Lai, L.C. and Chuang, E.Y., 2013. Concurrent gene signatures for han chinese breast cancers. *PLoS One*, 8(10).
3. Sinn, B.V., Fu, C., Lau, R., Litton, J., Tsai, T.H., Murthy, R., Tam, A., Andreopoulou, E., Gong, Y., Murthy, R. and Gould, R., 2019. SET ER/PR: a robust 18-gene predictor for sensitivity to endocrine therapy for metastatic breast cancer. *NPJ breast cancer*, 5(1), pp.1-8.
4. Gupta, G.P. and Massagué, J., 2006. Cancer metastasis: building a framework. *Cell*, 127(4), pp.679-695.
5. Onitilo, A.A., Engel, J.M., Greenlee, R.T. and Mukesh, B.N., 2009. Breast cancer subtypes based on ER/PR and Her2 expression: comparison of clinicopathologic features and survival. *Clinical medicine & research*, 7(1-2), pp.4-13.
6. Waks, A.G. and Winer, E.P., 2019. Breast cancer treatment: a review. *Jama*, 321(3), pp. 288-300.
7. Jatoi, I., Chen, B.E., Anderson, W.F. and Rosenberg, P.S., 2007. Breast cancer mortality trends in the United States according to estrogen receptor status and age at diagnosis. *Journal of Clinical Oncology*, 25(13), pp.1683-1690.
8. Kim, C., Tang, G., Pogue-Geile, K.L., Costantino, J.P., Baehner, F.L., Baker, J., Cronin, M.T., Watson, D., Shak, S., Bohn, O.L. and Fumagalli, D., 2011. Estrogen receptor (ESR1) mRNA expression and benefit from tamoxifen in the treatment and prevention of estrogen receptor-positive breast cancer. *Journal of clinical oncology*, 29(31), p.4160.
9. Iwamoto, T., Booser, D., Valero, V., Murray, J.L., Koenig, K., Esteva, F.J., Ueno, N.T., Zhang, J., Shi, W., Qi, Y. and Matsuoka, J., 2012. Estrogen receptor (ER) mRNA and ER-related gene expression in breast cancers that are 1% to 10% ER-positive by immunohistochemistry. *J Clin Oncol*, 30(7), pp.729-734.
10. Veeraraghavan, J., Tan, Y., Cao, X.X., Kim, J.A., Wang, X., Chamness, G.C., Maiti, S.N., Cooper, L.J., Edwards, D.P., Contreras, A. and Hilsenbeck, S.G., 2014. Recurrent ESR1-CCDC170 rearrangements in an aggressive subset of oestrogen receptor-positive breast cancers. *Nature communications*, 5(1), pp.1-12.

- 1 11. Holst, F., Stahl, P.R., Ruiz, C., Hellwinkel, O., Jehan, Z., Wendland, M., Lebeau, A.,
2 Terracciano, L., Al-Kuraya, K., Jänicke, F. and Sauter, G., 2007. Estrogen receptor alpha
3 (ESR1) gene amplification is frequent in breast cancer. *Nature genetics*, 39(5), pp.655-660.
- 4 12. Chandarlapaty, S., Chen, D., He, W., Sung, P., Samoila, A., You, D., Bhatt, T., Patel, P., Voi,
5 M., Gnant, M. and Hortobagyi, G., 2016. Prevalence of ESR1 mutations in cell-free DNA and
6 outcomes in metastatic breast cancer: a secondary analysis of the BOLERO-2 clinical
7 trial. *JAMA oncology*, 2(10), pp.1310-1315.
- 8 13. Robinson, D.R., Wu, Y.M., Vats, P., Su, F., Lonigro, R.J., Cao, X., Kalyana-Sundaram, S.,
9 Wang, R., Ning, Y., Hodges, L. and Gursky, A., 2013. Activating ESR1 mutations in
10 hormone-resistant metastatic breast cancer. *Nature genetics*, 45(12), p.1446.
- 11 14. Toy, W., Shen, Y., Won, H., Green, B., Sakr, R.A., Will, M., Li, Z., Gala, K., Fanning, S., King,
12 T.A. and Hudis, C., 2013. ESR1 ligand-binding domain mutations in hormone-resistant
13 breast cancer. *Nature genetics*, 45(12), p.1439.
- 14 15. Bailey, S.D., Desai, K., Kron, K.J., Mazrooei, P., Sinnott-Armstrong, N.A., Treloar, A.E.,
15 Dowar, M., Thu, K.L., Cescon, D.W., Silvester, J. and Yang, S.C., 2016. Noncoding somatic
16 and inherited single-nucleotide variants converge to promote ESR1 expression in breast
17 cancer. *Nature genetics*, 48(10), p.1260.
- 18 16. Tomita, S., Abdalla, M.O.A., Fujiwara, S., Matsumori, H., Maehara, K., Ohkawa, Y., Iwase,
19 H., Saitoh, N. and Nakao, M., 2015. A cluster of noncoding RNAs activates the ESR1 locus
20 during breast cancer adaptation. *Nature communications*, 6(1), pp.1-15.
- 21 17. Harvey, J.M., Clark, G.M., Osborne, C.K. and Allred, D.C., 1999. Estrogen receptor status by
22 immunohistochemistry is superior to the ligand-binding assay for predicting response to
23 adjuvant endocrine therapy in breast cancer. *Journal of clinical oncology*, 17(5), pp.
24 1474-1481.
- 25
- 26
- 27
- 28

Rank	ID	p-value	t	B	Fold change	Gene	Gene name
136	205225_at	0.00136651	3.31	-1.37597	1.279 ± 0.2618	ESR1	estrogen receptor 1

Table 1: ESR1 is differentially expressed in breast cancer when comparing primary tumors based on disease-free survival of the patient at 24 months.

The rank of differential expression globally (out of 54674 total transcripts), the probe ID, the p-value of global differential expression, t, a moderated t-statistic, B, the log-odds of differential expression between the two groups compared, the gene and gene name are listed in this chart.

Rank	ID	p-value	t	B	Fold change	Gene	Gene name
3	211234_x_at	6.26E-07	5.21955147	5.59237	1.1718 ± 0.1931	ESR1	estrogen receptor 1
39	211233_x_at	6.15E-05	4.13079083	1.56819	1.1259 ± 0.1864	ESR1	estrogen receptor 1
139	217190_x_at	5.24E-04	3.54962541	-0.28596	1.1586 ± 0.2809	ESR1	estrogen receptor 1
152	205225_at	5.77E-04	3.52211026	-0.36811	1.0777 ± 0.1248	ESR1	estrogen receptor 1

Table 2: Transcripts encoded by the ESR1 gene represent 4 out of the 250 most differentially expressed RNA messages in metastatic breast cancer when comparing metastases based on overall survival of the patient at 24 months.

The rank of differential expression globally (out of 22283 total transcripts), the probe ID, the p-value of global differential expression, t, a moderated t-statistic, B, the log-odds of differential expression between the two groups compared, the gene and gene name are listed in this chart.

Figure 1: ESR1 is expressed at significantly higher levels in the primary tumors of patients with disease-free survival greater than 24 months.

The expression of ESR1 at the mRNA level in primary tumors of the breast is graphically represented, compared between patients with disease-free survival less than (left) or greater than 24 months (right), with mean value marked and p-value noted.

Figure 2: ESR1 is expressed at significantly higher levels in the metastatic tumors of patients with overall survival greater than 24 months.

The expression of ESR1 at the mRNA level in metastases to multiple organ sites is graphically represented, compared between patients with overall survival less than (left) or greater than 24 months (right), with mean value marked and p-value noted.

Figure 3: Linear regression analysis demonstrates ESR1 expression in metastases to multiple organ sites generally correlates with overall patient survival in metastatic breast cancer.

The mRNA expression level of ESR1 in metastases, on the X-axis, is compared to the overall survival of each patient, on the Y-axis. A linear regression model was generated and the goodness of fit (r^2) as well as the statistical significance of the model's non-zero nature (p) is listed.

Disease-Free Survival binned by ESR1 expression
primary tumors

Figure 4: Binning primary tumors based on ESR1 expression demonstrates mRNA level ESR1 expression can help predict enhanced survival outcomes in patients with breast cancer.

Disease-free survival is graphically represented with mean values marked, comparing patients when grouped into 4 “bins” of 20 or 21 patients each based on ESR1 mRNA levels in primary tumors, from lowest to highest ESR1 mRNA expression levels, with p-value above evaluating significance of difference in disease-free survival compared to Bin 1, tumors with the lowest ESR1 expression levels.

Figure 5: Binning metastatic tissues based on ESR1 expression demonstrates mRNA level ESR1 expression in metastases can help predict enhanced survival outcomes in patients with stage IV breast cancer.

Overall survival is graphically represented with mean values marked, comparing patient overall survival when grouped into 4 “bins” of 35 patients each based on ESR1 mRNA levels in metastases, from lowest to highest ESR1 mRNA expression levels, with p-value above evaluating significance of difference in overall survival compared to Bin 1, metastatic tumor tissues with the lowest ESR1 expression levels. Analysis of each of the four differentially expressed ESR1 transcripts is presented here.

Figure 6: In patients with stage IV metastatic breast cancer, the majority of patients whose metastatic tissues express the lowest amount of ESR1 mRNA possess overall survival less than 20 months, while the majority of patients whose metastatic tissues express the highest amount of ESR1 mRNA possess overall survival greater than 20 months.

The mean percentage of patients with stage IV metastatic breast cancer possessing overall survival less than 20 months, averaging the results from analyses using each of the four differentially expressed ESR1 transcripts, is graphically represented here, with p-value above evaluating significance of difference in percentage of patients surviving less than 20 months compared to Bin 1, metastatic tumor tissues with the lowest ESR1 expression levels.

	<u>Bin 1 (lowest ESR1 expression)</u>	<u>Bin 2</u>	<u>Bin 3</u>	<u>Bin 4 (highest ESR1 expression)</u>
Mean % Patients with OS less than 20 months	65.71	52.14	38.57	22.14
Standard Deviation	4.666	3.595	4.949	6.335

Table 3: Average percentage of patients with overall survival less than 20 months when metastases are grouped based on ESR1 mRNA expression levels from lowest to highest.

The standard deviation is listed below the mean percentage of patients with overall survival less than 20 months, averaging the results from analyses using each of the four differentially expressed ESR1 transcripts.