Published in final edited form as:

Hum Pathol. 2010 December 1; 41(12): 1794–1801. doi:10.1016/j.humpath.2010.06.010.

Higher Levels of GATA3 Predict Better Survival in Women with Breast Cancer

Nam K. Yoon^{1,8}, Erin L. Maresh^{1,8}, Dejun Shen^{3,4,8}, Yahya Elshimali¹, Sophia Apple¹, Steve Horvath^{2,5,6}, Vei Mah¹, Shikha Bose^{1,7}, David Chia^{1,2}, Helena R. Chang^{2,3,4,9}, and Lee Goodglick^{1,2,9}

¹Department of Pathology and Laboratory Medicine, Los Angeles, California

²Jonsson Comprehensive Cancer Center, Los Angeles, California

³Gonda/UCLA Breast Cancer Research Laboratory, Los Angeles, California

⁴Revlon/UCLA Breast Cancer Center, Department of Surgery, Los Angeles, California

⁵Department of Biostatistics, Los Angeles, California

⁶Department of Human Genetics, David Geffen School of Medicine at UCLA, Los Angeles, California

⁷Department of Pathology and Laboratory Medicine, Cedars Sinai Medical Center, Los Angeles, California

Abstract

The GATA family members are zinc finger transcription factors involved in cell differentiation and proliferation. GATA3 in particular is necessary for mammary gland maturation, and its loss has been implicated in breast cancer development. Our goal was to validate the ability of GATA3 expression to predict survival in breast cancer patients. Protein expression of GATA3 was analyzed on a high density tissue microarray consisting of 242 cases of breast cancer. We associated GATA3 expression with patient outcomes and clinicopathological variables. Expression of GATA3 was significantly increased in breast cancer, *in situ* lesions, and hyperplastic tissue compared to normal breast tissue. GATA3 expression decreased with increasing tumor grade. Low GATA3 expression was a significant predictor of disease-related death in all patients, as well as in subgroups of estrogen receptor positive or low grade patients. Additionally, low GATA3 expression correlated with increased tumor size and estrogen and progesterone receptor negativity. GATA3 is an important predictor of disease outcome in breast cancer patients. This finding has been validated in a diverse set of populations. Thus, GATA3 expression has utility as a prognostic indicator in breast cancer.

Keywords

Tissue microarray; breast cancer; tumor marker; prognostic marker

© 2010 Elsevier Inc. All rights reserved.

To whom correspondence should be addressed: Lee Goodglick, Ph.D., Department of Pathology and Laboratory Medicine, David Geffen School of Medicine, University of California, Los Angeles, 10833 Le Conte Ave; Box 951732, CHS, Los Angeles, California, 90095-1747. Phone: (310) 825-9134; Fax: (310) 267-2104, lgoodglick@mednet.ucla.edu.

⁸Authors contributed equally

⁹Authors contributed equally

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

INTRODUCTION

Breast cancer is the most commonly diagnosed malignancy and the second leading cause of cancer death in women [1]. Encouragingly, the total numbers of deaths from breast cancer is decreasing, primarily due to more diligent surveillance and more refined therapeutic approaches. Nevertheless, breast cancer resulted in over 40,000 deaths in 2009 alone [1], underscoring the need for more effective means of detection, stratification of patient populations, and specific treatment of different sub-variants of this disease.

GATA binding protein 3 (GATA3) is one of six members of a family of zinc finger transcription factors that bind to the consensus DNA sequence (A/T)GATA(A/G). Binding of GATA members, in general, is thought to promote differentiation, development, and/or cell proliferation [2]. Notably, GATA3 is critical in T cell development and is required for Th2 differentiation [3–6]. GATA3 is functional in non-hematopoietic cells as well, playing a fundamental role in the development of the sympathetic nervous system, the kidney, adipose cells, cochlea, and the hair follicle in skin [7–11].

GATA3 is a well-known factor in breast glandular cell development. In particular, GATA3 is necessary for embryonic mammary development and is actively involved in maintaining the differentiated state of luminal epithelial cells of the mammary gland in adults [12–13]. Not surprisingly, loss of GATA3 expression has been associated with breast cancer pathogenesis, with lower expression levels generally associating with estrogen and progesterone receptor negativity, Her2/neu over expression, and poor prognosis [14–27]. GATA3 overexpression is thought to contribute to aberrant aromatase expression in breast tumors [27–28].

Previously, we had also observed by gene expression array analysis and real time quantitative RT-PCR that GATA3 was significantly increased in breast cancer cases compared to normal breast epithelium [14]. Here we assessed GATA3 protein expression in breast cancer on a population of patients at the UCLA Medical Center using high-density tissue microarray (TMA) technology. Consistent with previous results, we found that lower levels of GATA3 expression predicted a poorer disease outcome in all patients as well as in the subgroups of ER + and low grade patients. Significantly, this study serves as an independent validation of results reported by others and thus emphasizes the importance of GATA3 as a prognostic tool and a potential target for therapeutic intervention.

MATERIALS AND METHODS

Breast tissue microarray

A high-density breast TMA was constructed using cores from formalin-fixed, paraffin embedded breast tissue donor blocks, consisting of 242 breast surgical cases of 210 patients who underwent surgery at the UCLA Medical Center between 1995 and 2000 [30–31]. At least three cores of each available histologic type were arrayed from the donor blocks. Of the 242 surgical cases, 179 cases were of invasive breast cancers of various histologic types. For outcome analysis, we removed surgical cases of patients who had received neoadjuvant therapy, resulting in 86 primary surgical cases of patients with invasive cancer who had disease-specific survival outcome and were informative for GATA3 protein expression.

Immunohistochemistry

Immunohistochemical staining of the breast TMA was performed using a standard two-step indirect avidin-biotin complex method (Vector Laboratories, Burlingame, CA) as previously described [31], using a mouse monoclonal anti-human GATA3 antibody (Santa Cruz Biotechnology, #SC-268, Lot H142). Briefly, 4 µm sections were deparaffinized, treated with

0.3% hydrogen peroxide in methanol, blocked with 5% goat serum, incubated with the primary antibody for 60 minutes and secondary for 60 minutes as described [31]. Diaminobenzidine was used for color detection [31].

A number of tests were performed to confirm the specificity of staining results for GATA3. First, expression results were confirmed using a primary antibody from another manufacturer, Abcam (Cambridge, MA). For these experiments we used rabbit anti-human GATA3 polyclonal antibody (catalogue # ab32858) at a final concentration of 5 µg/ml. Identical results were observed for GATA3 expression levels in breast cancer using the antibodies from either Abcam or Santa Cruz. Second, concentration-matched isotype control mouse IgG1 was used for negative controls. Under these conditions, no staining was observed. Finally, GATA3 peptide competition successfully inhibited nuclear staining. Anti-GATA3 polyclonal antibody (Abcam) was incubated with a 0, 1, 5 or 10 fold molar excess of inhibitory synthetic peptide (Abcam, catalogue # ab32857) for 2 hours on ice. Following incubation with peptide, the immunohistochemical protocol proceeded as described above.

In some experiments, breast sections were stained for the proliferation marker Ki67 (Dako, Carpinteria, CA; #M7240) at a final concentration of 1.33 µg/ml, or the apoptosis indicator, cleaved caspase 3 (Cell Signaling Technology, Danvers, MA; #9661S) at a final concentration of 1 µg/ml. The protocol used was as previously described [31].

The level of protein expression was semiquantitatively assessed by a pathologist (YE) blinded to all clinico-pathological variables, who noted both the intensity and frequency of nuclear staining in glandular epithelial cells of each spot. A second pathologist (VM) spotchecked 20% of the cores with an inter-observer variation of less than 5%. A weighted (integrated) score for which combined relative protein expression levels with the frequency of epithelial cells staining at each intensity was calculated using the following formula: $(3(\%a) + 2(\%b) + 1(\%c)) / 100$, where a, b, and c represent the percentage of cells staining at strong, moderate, and weak staining, respectively. For case-level and outcomes analyses, the median expression value was calculated and used for each case similar to methods previously described [32–33].

Statistical analyses

Statistical analyses were performed using StatView Version 5.0 (SAS Institute, Cary, NC) or with the R software package (<http://www.r-project.org>) as previously described [32–33]. Spotlevel GATA3 expression was compared against spot grade and spot histology using Mann-Whitney and Kruskal-Wallis tests for two-group and multi-group comparisons, respectively. Case-level GATA3 expression was evaluated as a continuous variable and was also dichotomized as low expression or high expression using an optimized cut-point identified as previously described [32,34]. As a continuous variable, GATA3 expression was compared against known clinico-pathological variables using Mann-Whitney and Kruskal Wallis tests for group comparisons and Spearman's two-tailed correlation test for nonparametric correlative analyses. As a dichotomized variable, comparisons were done using Fisher's Exact for 2×2 comparisons and the Chi-Square test for $2 \times n$ comparisons. The difference in the cumulative disease-specific survival of patients split by their GATA3 expression was visualized using Kaplan-Meier curves and the statistical difference calculated using the log-rank test. Univariate Cox proportional hazards regression models were used to calculate the prognostic significance of GATA3 expression and other clinico-pathological variables. A multivariate Cox model was used to determine the prognostic significance of GATA3 after correcting for conventional prognostic variables.

RESULTS

GATA3 protein expression patterns in breast tissues

Expression of the transcription factor GATA3 has been shown to be important for normal breast glandular cell development as well for maintaining the differentiated state of luminal epithelial cells [12–13]. Here we examined the expression level and frequency of GATA3 in tissue from breast cancer patients from the UCLA Medical Center using a high-density tissue microarray (TMA) platform consisting of 242 patients with a total of 2,040 spots. Representative images of GATA3 expression in different histologies are shown in Figure 1. We observed clear, predominantly nuclear GATA3 protein expression in luminal epithelial cells and some faint, diffuse staining in the cytoplasm of strongly nuclear-positive cells (Figure 1). Some degree of cytoplasmic expression had been previously described [35]. We did not observe staining in the myoepithelial cells or in the surrounding stroma. Normal ducts and lobules were focally positive, with generally increased expression in malignant cells. Specificity of the immunoassay was determined by the lack of staining with a non-immune primary antibody, through a complete blocking of nuclear staining using 5 and 10 fold molar excess of GATA3 competitive peptide, and through an identical pattern of staining using anti-GATA3 primary antibodies from 2 separate manufacturers (see Materials and Methods).

We next examined GATA3 expression in each histologic type as described in Materials and Methods. As shown in Figure 2A, there were significantly higher levels of GATA3 expression in ductal hyperplasia lesions (DH, $P < 0.0001$), ductal carcinoma *in situ* (DCIS, $P = 0.0006$) and invasive ductal carcinoma (IDC, $P < 0.0001$) compared to normal ductal epithelium. Normal epithelium was represented by either normal tissue adjacent to a neoplasm or tissue from elective breast reduction surgery. GATA3 was significantly associated with spot grade (Figure 2B) ($P < 0.0001$). Low-grade tumors had the highest GATA3 expression, with a steady decrease in GATA3 expression with increasing grade.

Patients with relatively low levels of GATA3 have a poorer disease prognosis

We next determined whether GATA3 expression levels had predictive value for disease-specific survival. First, as a continuous variable, GATA3 approached significance as a predictor of survival in a univariate Cox model with lower levels of GATA3 associating with increased risk of death ($P = 0.055$, Table 2). However, as a dichotomized variable, GATA3 was a highly significant predictor of survival. Figure 3 shows the Kaplan-Meier cumulative survival curve when patient groups are dichotomized by high or low GATA3 expression. Significantly, patients with higher GATA3 expression had a 97% 10-year disease-specific survival as compared to only 72% in patients with low GATA3 expression ($P = 0.0041$).

We further examined whether GATA3 expression was associated with any clinico-pathological variables in the patients used for outcomes analysis (see Table 1 for patient demographics and clinico-pathological variables). When we examined GATA3 as a continuous variable, we found that low protein expression was associated with high tumor grade ($P = 0.0050$), larger tumor size ($P = 0.0165$), negative estrogen receptor (ER) status ($P = 0.0001$), and negative progesterone receptor (PR) status ($P = 0.0366$). These associations were consistent when we repeated the analysis with GATA3 as a dichotomized variable. The only exception was that low GATA3 was no longer significantly associated with negative PR status (Table 1).

Because GATA3 expression was associated with tumor grade and hormone receptor status, we grouped patients by ER status or by tumor grade and examined GATA3 expression levels in these subpopulations. Notably, we found that lower levels of GATA3 was still predictive in a subset of individuals who had ER-positive tumors (Figure 4A; $P = 0.0163$) and low-grade

tumors (Figure 4B; $P = 0.0172$); this was not the case in individuals with ER negative tumors ($P = 0.2542$) or high-grade tumors ($P = 0.2173$).

To assess whether GATA3 expression was a significant predictor of disease outcome after correcting for standard prognostic variables, we used a multivariate Cox model that included clinical stage (dichotomized as stage I and II vs. III and IV), tumor grade (dichotomized as grade I and II vs. III), lymphovascular invasion, and ER, PR, and HER-2/neu status (Table 3). In this model, GATA3 as a dichotomized variable remained a significant independent predictor of survival (HR = 0.0928, 95% CI = 0.0104 – 0.83, $P = 0.033$) along with stage (HR = 6.68, 95% CI = 1.67 – 26.8, $P = 0.0073$). Of note, although the P value for tumor stage is highly significant, the confidence interval for the hazard ratio is very wide indicating decreased accuracy for this clinical parameter in this particular data set. Nevertheless, if we carry out an additional analysis of restricted the Cox analysis to patients of a given stage, we find a highly significant P -value for GATA3.

GATA3 expression neither correlated with the proliferative capacity of the tumor cells nor the apoptotic index consistent with the role of GATA3 in differentiation in these tumors [36]. In representative tumor samples, an average of 83.2 ± 15.0 cells stained positively for nuclear GATA3; there was a corresponding proliferative index (as measured by Ki67) of 14.4 ± 9.8 and apoptotic index (as measured by cleaved caspase 3) of 2.5 ± 2.2 .

DISCUSSION

GATA3 is a member of the GATA transcription regulatory family and is important in directing cell fate, development, and/or differentiation in a number of cell types including luminal epithelial cells of the mammary gland [7–13]. As such, there has been keen interest in the potential role of GATA3 dysregulation in the pathogenesis of breast cancer. In this study, we have used TMA technology to reassess the associations of GATA3 expression with breast cancer development and progression. Importantly, we observed a number of findings that were consistent with results from other independent cohorts therefore building on the strength of GATA3 as a beneficial biomarker for potential use clinically.

We observed that GATA3 protein expression levels were relatively higher in invasive ductal carcinoma and metastatic cells, compared to normal glandular or ductal breast epithelium. This pattern was similarly consistent in DCIS lesions compared to normal epithelium. Somewhat surprisingly, GATA3 expression levels in ductal hyperplasia were relatively elevated as well suggesting that GATA3 may not be an appropriate gauge of early detection; however, it should be noted that there were relatively few hyperplasia samples represented on the TMA, and all these lesions were in the context of women with breast cancer. On the other hand, the elevated GATA3 levels in our system may in part be a function of proliferation rather than necessarily malignant transformation.

The observed higher expression of GATA3 in hyperplastic, DCIS, and malignant lesions is interesting because it is seemingly counter to the prevailing data from *in vitro* and animal studies suggesting that deletion or depletion of GATA3 in normal mammary epithelium leads to dedifferentiation and increased cell proliferation [12–13,36]. However, a recent study by Pei, *et al.* suggests that the story may not be quite so linear as they found that a normal function of GATA3 is the suppression of p18^{INK4C}, an inhibitor of the cell cycle [37]. In their system, luminal A type breast malignancies which expressed higher GATA3 and lower levels of p18^{INK4C}, had a more favorable outcome [37]. While this would be consistent with our observations presented here, certainly the exact function of GATA3 during malignant development and progression awaits further clarification and refinement.

While we and others have shown that GATA3 levels were generally higher in malignant cells compared to morphologically normal epithelium, within malignant tissue, relatively lower levels of GATA3 portended a poorer outcome compared to tumors with relatively higher expression. This phenomenon is shown dramatically in Figure 2A, where 97% of individuals with tumors expressing higher levels of GATA3 survived 10 years or more. It is important to note that these findings are consistent with and validate previous observations further emphasizing the potential relevance of GATA3 as a clinically useful biomarker [14, 16–17, 19, ²¹, 23, 26–27, 38]. Consistent with the role of GATA3 in the development and differentiation of normal mammary epithelium, we further observed that lower levels of GATA3 were generally associated with a higher grade, less differentiated malignancy. Nevertheless, even in low grade tumors, GATA3 was a powerful predictive marker (Figure 4B).

GATA3 expression was correlated with expression of the estrogen receptor both in our cohort and in others [23,39–40]. The correlation of GATA3 and ER expression levels continues to be intriguing both mechanistically and because of its potential clinical application. Despite this correlation, to date there is no compelling consensus data suggesting that ER or GATA3 directly regulate one another. Rather, the interplay between GATA3 and the ER stimulated pathways may be bridged by the forkhead family transcription factor, FOXA1 [12,29,39]. GATA3 can potentially regulate FOXA1 expression, and in turn, FOXA1 appears to be required to promote expression of many if not all estrogen-responsive genes [12,39]. Of note, our results (Figure 4A) as well as those of Mehra *et al.* [23] show that there is not always a direct correlation between GATA3 and ER expression; within the patient population with ER + tumors, relatively higher GATA3 expression levels remains a powerful predictor of future survival.

It is important to point out that our results confirm the findings of other groups who have used TMA or gene expression technology to investigate the prognostic value of GATA3 [19,23, 39–41]. While our intent when we initiated this project was not necessarily to conduct a validation study, it should be noted that the relevance and importance of independent validation is critical for identifying tangible disease biomarkers [42–44]. Therefore, that GATA3 levels were predictive of outcome, as observed in multiple separate and independent patient populations, greatly strengthens its potential utility as a prognostic indicator.

Acknowledgments

We would like to thank Samson Schatz for helpful discussion.

This work was supported in part by the Early Detection Research Network NCI CA-86366 (LG and DC) and Gonda Family Foundation (HRC).

Abbreviations

TMA	tissue microarray
DCIS	ductal carcinoma <i>in situ</i>
DH	ductal hyperplasia
IDC	invasive ductal carcinoma
ER	estrogen receptor
PR	progesterone receptor
HR	hazard ratio

CI confidence interval

REFERENCES

1. Jemal A, et al. Cancer statistics, 2009. *CA Cancer J Clin* 2009;59(4):225–249.
2. Burch JB. Regulation of GATA gene expression during vertebrate development. *Semin Cell Dev Biol* 2005;16(1):71–81. [PubMed: 15659342]
3. Ting CN, et al. Transcription factor GATA-3 is required for development of the T-cell lineage. *Nature* 1996;384(6608):474–478. [PubMed: 8945476]
4. Pai SY, et al. Critical roles for transcription factor GATA-3 in thymocyte development. *Immunity* 2003;19(6):863–875. [PubMed: 14670303]
5. Hendriks RW, et al. Expression of the transcription factor GATA-3 is required for the development of the earliest T cell progenitors and correlates with stages of cellular proliferation in the thymus. *Eur J Immunol* 1999;29(6):1912–1918. [PubMed: 10382753]
6. Grogan JL, Locksley RM. T helper cell differentiation: on again, off again. *Curr Opin Immunol* 2002;14(3):366–372. [PubMed: 11973136]
7. Kaufman CK, et al. GATA-3: an unexpected regulator of cell lineage determination in skin. *Genes Dev* 2003;17(17):2108–2122. [PubMed: 12923059]
8. Grote D, et al. Pax 2/8-regulated Gata 3 expression is necessary for morphogenesis and guidance of the nephric duct in the developing kidney. *Development* 2006;133(1):53–61. [PubMed: 16319112]
9. Tong Q, et al. Function of GATA transcription factors in preadipocyte-adipocyte transition. *Science* 2000;290(5489):134–138. [PubMed: 11021798]
10. Martinez-Monedero R, et al. Differentiation of inner ear stem cells to functional sensory neurons. *Dev Neurobiol* 2008;68(5):669–684. [PubMed: 18278797]
11. Tsarovina K, et al. Essential role of Gata transcription factors in sympathetic neuron development. *Development* 2004;131(19):4775–4786. [PubMed: 15329349]
12. Kourou-Mehr H, et al. GATA-3 maintains the differentiation of the luminal cell fate in the mammary gland. *Cell* 2006;127(5):1041–1055. [PubMed: 17129787]
13. Asselin-Labat ML, et al. Gata-3 is an essential regulator of mammary-gland morphogenesis and luminal-cell differentiation. *Nat Cell Biol* 2007;9(2):201–209. [PubMed: 17187062]
14. Shen D, et al. Loss of annexin A1 expression in human breast cancer detected by multiple high-throughput analyses. *Biochem Biophys Res Commun* 2005;326(1):218–227. [PubMed: 15567174]
15. Hoch RV, et al. GATA-3 is expressed in association with estrogen receptor in breast cancer. *Int J Cancer* 1999;84(2):122–128. [PubMed: 10096242]
16. Bertucci F, et al. Gene expression profiling of primary breast carcinomas using arrays of candidate genes. *Hum Mol Genet* 2000;9(20):2981–2991. [PubMed: 11115842]
17. West M, et al. Predicting the clinical status of human breast cancer by using gene expression profiles. *Proc Natl Acad Sci U S A* 2001;98(20):11462–11467. [PubMed: 11562467]
18. Gruvberger S, et al. Estrogen receptor status in breast cancer is associated with remarkably distinct gene expression patterns. *Cancer Res* 2001;61(16):5979–5984. [PubMed: 11507038]
19. Jenssen TK, et al. Associations between gene expressions in breast cancer and patient survival. *Hum Genet* 2002;111(4–5):411–420. [PubMed: 12384785]
20. Sorlie T, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 2001;98(19):10869–10874. [PubMed: 11553815]
21. Sotiriou C, et al. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. *Proc Natl Acad Sci U S A* 2003;100(18):10393–10398. [PubMed: 12917485]
22. Usary J, et al. Mutation of GATA3 in human breast tumors. *Oncogene* 2004;23(46):7669–7678. [PubMed: 15361840]
23. Mehra R, et al. Identification of GATA3 as a breast cancer prognostic marker by global gene expression meta-analysis. *Cancer Res* 2005;65(24):11259–11264. [PubMed: 16357129]

24. Farmer P, et al. Identification of molecular apocrine breast tumours by microarray analysis. *Oncogene* 2005;24(29):4660–4671. [PubMed: 15897907]
25. Yang F, et al. Laser microdissection and microarray analysis of breast tumors reveal ER-alpha related genes and pathways. *Oncogene* 2006;25(9):1413–1419. [PubMed: 16261164]
26. Perou CM, et al. Molecular portraits of human breast tumours. *Nature* 2000;406(6797):747–752. [PubMed: 10963602]
27. van 't Veer LJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002;415(6871):530–536. [PubMed: 11823860]
28. Bouchard MF, Taniguchi H, Viger RS. Protein kinase A-dependent synergism between GATA factors and the nuclear receptor, liver receptor homolog-1, regulates human aromatase (CYP19) PII promoter activity in breast cancer cells. *Endocrinology* 2005;146(11):4905–4916. [PubMed: 16109788]
29. Eeckhoutte J, et al. Positive cross-regulatory loop ties GATA-3 to estrogen receptor alpha expression in breast cancer. *Cancer Res* 2007;67(13):6477–6483. [PubMed: 17616709]
30. Kononen J, et al. Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nat Med* 1998;4(7):844–847. [PubMed: 9662379]
31. Shen D, et al. Decreased expression of annexin A1 is correlated with breast cancer development and progression as determined by a tissue microarray analysis. *Hum Pathol* 2006;37(12):1583–1591. [PubMed: 16949910]
32. Seligson DB, et al. Expression of X-Linked Inhibitor of Apoptosis Protein Is a Strong Predictor of Human Prostate Cancer Recurrence. *Clin Cancer Res* 2007;13(20):6056–6063. [PubMed: 17947468]
33. Seligson DB, et al. Global histone modification patterns predict risk of prostate cancer recurrence. *Nature* 2005;435(7046):1262–1266. [PubMed: 15988529]
34. Yoon NK, et al. Higher expression levels of 14-3-3sigma in ductal carcinoma in situ of the breast predict poorer outcome. *Cancer Biomark* 2009;5(4):215–224. [PubMed: 19729831]
35. Gulbinas A, et al. Aberrant gata-3 expression in human pancreatic cancer. *J Histochem Cytochem* 2006;54(2):161–169. [PubMed: 16087702]
36. Chou J, Provot S, Werb Z. GATA3 in development and cancer differentiation: cells GATA have it! *J Cell Physiol* 2010;222(1):42–49. [PubMed: 19798694]
37. Pei XH, et al. CDK inhibitor p18(INK4c) is a downstream target of GATA3 and restrains mammary luminal progenitor cell proliferation and tumorigenesis. *Cancer Cell* 2009;15(5):389–401. [PubMed: 19411068]
38. Wang Y, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* 2005;365(9460):671–679. [PubMed: 15721472]
39. Albergaria A, et al. Expression of FOXA1 and GATA-3 in breast cancer: the prognostic significance in hormone receptor-negative tumours. *Breast Cancer Res* 2009;11(3):R40. [PubMed: 19549328]
40. Voduc D, Cheang M, Nielsen T. GATA-3 expression in breast cancer has a strong association with estrogen receptor but lacks independent prognostic value. *Cancer Epidemiol Biomarkers Prev* 2008;17(2):365–373. [PubMed: 18268121]
41. Ciocca V, et al. The significance of GATA3 expression in breast cancer: a 10-year follow-up study. *Hum Pathol* 2009;40(4):489–495. [PubMed: 19084267]
42. Feng Z, Prentice R, Srivastava S. Research issues and strategies for genomic and proteomic biomarker discovery and validation: a statistical perspective. *Pharmacogenomics* 2004;5(6):709–719. [PubMed: 15335291]
43. Maruvada P, Srivastava S. Joint National Cancer Institute-Food and Drug Administration workshop on research strategies, study designs, and statistical approaches to biomarker validation for cancer diagnosis and detection. *Cancer Epidemiol Biomarkers Prev* 2006;15(6):1078–1082. [PubMed: 16775163]
44. Maruvada P, et al. Biomarkers in molecular medicine: cancer detection and diagnosis. *Biotechniques* 2005 Suppl:9–15. [PubMed: 16528918]

Figure 1. Representative images of GATA3 expression in breast tissues

Staining was observed predominantly in the nuclei of glandular epithelial cells, as shown using a 20x objective in spots containing (A) normal ductal epithelium, (B) ductal hyperplasia without atypia, (C) ductal carcinoma *in situ*, and (D) invasive ductal carcinoma.

Figure 2. GATA3 expression stratified by histologic type and nuclear grade

The mean GATA3 protein expression for each category is shown using bar plots. The error bars represent the standard error of the mean; n is number of spots. (A) GATA3 expression was significantly increased in ductal hyperplasia (DH, $P < 0.0001$), ductal carcinoma *in situ* (DCIS, $P = 0.0006$), and invasive ductal carcinoma (IDC, $P < 0.0001$) compared to adjacent normal tissue or tissue from voluntary breast reductions (BR). (B) GATA3 expression decreased with increasing tumor grade ($P < 0.0001$).

Figure 3. Low GATA3 is associated with poor outcome in breast cancer patients

Kaplan-Meier analysis of patients with low GATA3 expression (median GATA3 expression < 1.8) or high GATA3 expression (median GATA3 expression ≥ 1.8) is shown; n is the number of patients in each group.

Figure 4. GATA3 expression is predictive in ER-positive and low grade patients

Kaplan-Meier curves of patients in ER-positive (A) and low grade (B) subgroups are shown; n is the number of patients in each group. Patients with low GATA3 expression in either subgroup showed a significantly increased risk of disease-related death.

Table 1
Clinico-pathologic parameters and nuclear GATA3 expression

The reported p-values were the results from the following non-parametric statistical analyses:

Variables	All patients	Number of patients Low GATA3	High GATA3	GATA3 Mean $\pm$ S.E.	Dichotomized P-value	Continuous P-Value
All invasive patients	86	49	37			
Age at diagnosis						
Median (Range)	51 (30 – 89)	52 (35 – 82)	50 (30 – 89)	$\rho = 0.005^b$	0.8683 ^a	0.9605 ^b
25 th to 75 th Quartile	45 – 68	44 – 64	45 – 68			
Clinical stage						
I	32	13	19	1.602 $\pm$ 1.042		
II	37	24	13	1.084 $\pm$ 0.983	0.1073 ^d	0.1366 ^e
III	16	11	5	1.030 $\pm$ 0.910		
IV	1	1	0	1.150		
Tumor grade						
1	24	10	14	1.466 $\pm$ 1.151		
2	24	12	12	1.583 $\pm$ 0.766	0.0343 ^d	0.0050 ^e
3	34	25	9	0.842 $\pm$ 0.929		
Unknown	4	2	2			
Lymph node metastasis						
Absent	37	18	19	1.477 $\pm$ 1.076	0.1522 ^c	0.1286 ^a
Present	33	22	11	1.064 $\pm$ 0.984		
Unknown [†]	16	9	7			
Tumor size (cm)						
Median (Range)	2.1 (0.1 – 9.0)	2.3 (0.1 – 9.0)	1.5 (0.5 – 7.3)	$\rho = -0.266^b$	0.0466 ^a	0.0165 ^b
25 th to 75 th Quartile	1.2 – 3.0	1.5 – 3.0	1.0 – 2.5			
Lymphovascular invasion						
Absent	56	31	25	1.296 $\pm$ 1.030	0.8198 ^c	0.6403 ^a
Present	30	18	12	1.214 $\pm$ 0.983		
ER status						
Positive	63	29	34	1.530 $\pm$ 0.949	0.0038 ^c	0.0001 ^a
Negative	20	17	3	0.564 $\pm$ 0.869		
Unknown	3	3	0			
PR status						
Positive	61	32	29	1.407 $\pm$ 0.974	0.2338 ^c	0.0366 ^a
Negative	25	17	8	0.926 $\pm$ 1.031		
HER-2/neu status						
Positive	20	10	10	1.365 $\pm$ 0.919	0.6062 ^c	0.6544 ^a
Negative	63	37	26	1.267 $\pm$ 1.040		
Unknown	3	2	1			

(a) Mann-Whitney,

(b) Spearman Correlation,

(c) Fisher's Exact Test,

(d) Chi Square Test, and

(e) Kruskal-Wallis Test.

[†] No lymphadenectomies were performed for these patients.

Table 2**Univariate Cox Model**

Grade was dichotomized as low (grade 1 and 2) vs. high (grade 3). Stage was dichotomized as low (stage I and II) vs. high (stage III and IV).

Variable	Hazard Ratio	95% Confidence Interval	P-Value
GATA3 continuous	0.544	0.292 – 1.01	0.055
GATA3 dichotomized	0.092	0.012 – 0.708	0.022
Age at diagnosis	0.955	0.909 – 1.00	0.072
Clinical stage	8.2	2.67 – 25.2	0.00024
Tumor grade	3.31	1.02 – 10.8	0.046
Lymph node metastasis	15.4	1.98 – 119	0.0089
Tumor size	1.23	0.943 – 1.61	0.13
Lymphovascular invasion	3.2	1.05 – 9.78	0.041
ER status	0.347	0.117 – 1.03	0.057
PR status	0.616	0.201 – 1.89	0.40
HER-2/neu status	3.59	1.21 – 10.7	0.022

Table 3
Multivariate Cox Model

Grade was dichotomized as low (grade 1 and 2) vs. high (grade 3). Stage was dichotomized as low (stage I and II) vs. high (stage III and IV).

Variable	Hazard Ratio	95% Confidence Interval	P-Value
GATA3 dichotomized	0.0928	0.0104 – 0.83	0.033
Stage	6.6798	1.6679 – 26.75	0.0073
Grade	2.1267	0.5271 – 8.58	0.29
Lymphovascular invasion	1.4671	0.3885 – 5.54	0.57
ER status	1.2381	0.2995 – 5.12	0.77
PR status	0.5155	0.1145 – 2.32	0.39
HER-2/neu status	3.4416	0.9525 – 12.44	0.059